

Datum: Montag, 27.01.2025

11:00 - 12:00	Eröffnung Ort: O 102 Chair der Sitzung: Oliver Dickhäuser Chair der Sitzung: Karina Karst Chair der Sitzung: Andreas Rausch
O 102	
12:30 - 20:30	Kinderbetreuung Agentur Farbenfroh Ort: O 226 - 228
O 226 - 228	
13:00 - 14:45	IS-01: Invited Symposium: Digitale Kompetenzen von Schüler*innen und Rahmenbedingungen des Kompetenzerwerbes im dritten internationalen Vergleich - Befunde und vertiefende Analysen der IEA-Studie ICILS 2023 Ort: O 102 Chair der Sitzung: Birgit Eickelmann Chair der Sitzung: Nadine Fröhlich Chair der Sitzung: Johannes Naumann
O 102	
13:00 - 14:45	OF-1: Workshop: Suchst du noch oder findest du schon? Strategien für komplexe Literaturrecherchen in der Bildungsforschung Ort: EO 159 Chair der Sitzung: Ingeborg Jäger-Dengler-Harles
EO 159	
13:00 - 14:45	PS-01: Qualität und Unterstützung im Lernprozess: Perspektiven auf Unterrichtsinteraktion, Kontext und Lernerfolg Ort: EO 169 Chair der Sitzung: Alexander Jonas Jung
EO 169	
13:00 - 14:45	PS-02: Digitale Medien in der frühen Bildung: Chancen und Herausforderungen Ort: EO 256 Chair der Sitzung: Dave Möwisch
EO 256	
13:00 - 14:45	PS-03: Ansätze in der Lese- & Schreibförderung Ort: EO 154 Chair der Sitzung: Joachim Grabowski
EO 154	
13:00 - 14:45	PS-04: Passung und Commitment in der beruflichen Bildung: Erfolgsfaktoren und Herausforderungen Ort: SO 418 Chair der Sitzung: Sylvia Rahn Chair der Sitzung: Sebastian Rahn
SO 418	
13:00 - 14:45	PS-05: Sprache im Unterricht: Förderung, Unterrichtspraxis und Lehrerperspektiven Ort: EO 162 Chair der Sitzung: Kristin Kersten
EO 162	
13:00 - 14:45	PS-06: Professionelles Wissen und Handeln von Lehrkräften Ort: EO 242 Chair der Sitzung: Birte Oetjen
EO 242	
13:00 - 14:45	PS-07: Professionalisierung im Lehramt Ort: SO 318 Chair der Sitzung: Christin Lotz
SO 318	
13:00 - 14:45	PS-08: Learning Mathematics: Motivation, Self-Concept, and Achievement Ort: EO 382 Chair der Sitzung: Charlott Rubach English Track
EO 382	
13:00 - 14:45	PS-09: Diagnosekompetenz in Theorie und Praxis: Perspektiven und Interventionen Ort: SO 322 Chair der Sitzung: Juliane Rutsch
SO 322	
13:00 - 14:45	PS-10: Selbstkonzept, Motivation und Feedback: Dynamiken und Einflüsse im Bildungsprozess Ort: EO 157 Chair der Sitzung: Emely Hoch
EO 157	
13:00 - 14:45	S-01: Bridging Minds: A Transdisciplinary Approach on Executive Functions, Metacognition, Self-Regulation, and Self-Regulated Learning in Education Research Ort: O 129 Chair der Sitzung: Friederike Blume Chair der Sitzung: Maria Theobald Chair der Sitzung: Charlotte Dignath Chair der Sitzung: Julia Kerner auch Körner English Track
O 129	
13:00 - 14:45	S-02: Facets of early self-regulation: Valid assessment, short- and long-term dynamics and links to social development Ort: O 133
O 133	

	Chair der Sitzung: Manja Attig Chair der Sitzung: Jessica Willard Chair der Sitzung: Hanna Mulder English Track
13:00 - 14:45	S-03: Methodische Innovationen zur Nutzung computerisierter adaptiver Tests in Bildungskontexten
O 135	Ort: O 135 Chair der Sitzung: Joachim Ulf Kröhne Chair der Sitzung: Andreas Frey Chair der Sitzung: Johannes Hartig
13:00 - 14:45	S-04: Sonderpädagogische Perspektiven auf Bullying und Viktimisierung in der Schule
O 142	Ort: O 142 Chair der Sitzung: Philipp Nicolay Chair der Sitzung: Christian Huber Chair der Sitzung: Karolina Urton
13:00 - 14:45	S-05: Wie Lehrkräfte ihre Schüler*innen einschätzen: Potentielle Verzerrungen in Leistungserwartungen, Leistungsbewertung und Umgang mit Fehlverhalten
O 145	Ort: O 145 Chair der Sitzung: Sophie Harms Chair der Sitzung: Lars Meyer-Jenßen Chair der Sitzung: Hanna Dumont
13:00 - 14:45	S-06: Chancenungleichheit als gesellschaftliche Herausforderung: Interdisziplinäre und internationale Perspektiven auf die Förderung bildungsbenachteiligten Schüler:innen in Deutschland und Schweden
SN 163	Ort: SN 163 Chair der Sitzung: Haiqin Ning Chair der Sitzung: Marianne Schüpbach
13:00 - 14:45	S-07: Domänenübergreifende und domänenspezifische Bildungsprozesse in der frühen Kindheit
SO 108	Ort: SO 108 Chair der Sitzung: Elisa Oppermann Chair der Sitzung: Hans-Günther Roßbach
13:00 - 14:45	S-08: Zur Rolle von Emotionen von Lehrer*innen und Lehramtsstudierenden in verschiedenen pädagogischen Kontexten
O 148	Ort: O 148 Chair der Sitzung: Irina Rosa Kumschick Chair der Sitzung: Gerda Hagenauer Chair der Sitzung: Tina Hascher
13:00 - 14:45	S-09: Aktuelle Mindset-Interventionen im deutschsprachigen Raum: Was können sie für wen bewirken?
O 151	Ort: O 151 Chair der Sitzung: Katharina Asbury Chair der Sitzung: Anke Heyder Chair der Sitzung: Caroline Villiger
13:00 - 14:45	S-10: Einstellungen gegenüber Diversitätsmerkmalen in schulischen Kontexten: Interventionsmöglichkeiten
EO 165	Ort: EO 165 Chair der Sitzung: Stefanie Bosse Chair der Sitzung: Ursula Kessels
13:00 - 14:45	S-11: Professionelle Wahrnehmung von Lehrkräften anforderungsbezogen betrachten
SN 169	Ort: SN 169 Chair der Sitzung: Julia Bruns Chair der Sitzung: Rebekka Stahnke Chair der Sitzung: Marita Friesen
13:00 - 14:45	S-12: Sozio-emotionale Lehr/Lernerfahrungen im Schulkontext
EO 145	Ort: EO 145 Chair der Sitzung: Diana Raufelder
13:00 - 14:45	S-13: To KI or not to KI? Potentiale maschinellen Lernens, vortrainierter Deep Learning-Modelle und natürlicher Sprachverarbeitung für formatives Assessment
M 003	Ort: M 003 Chair der Sitzung: Sonja Hahn Chair der Sitzung: Fabian Zehner Chair der Sitzung: Carolin Hahnel
13:00 - 14:45	Special session: Meet the Editors
EO 150	Ort: EO 150 Chair der Sitzung: Sabine Glock
15:15 - 17:00	IS-02: Invited Symposium: Frühe Bildung und Ungleichheit
O 102	Ort: O 102 Chair der Sitzung: Guido Neidhöfer

	Chair der Sitzung: Simon Wiederhold Chair der Sitzung: Henning Hermes
15:15 - 17:00	OF-2: Podiumsdiskussion: Transparenz im Bildungssystem – welche Chancen bietet ein Bildungsverlaufsregister?
EO 145	Ort: EO 145 Chair der Sitzung: Kerstin Schneider
15:15 - 17:00	PS-11: Akademische Selbstwirksamkeit und Bildungserfahrungen von Schüler:innen mit Migrationshintergrund
SO 322	Ort: SO 322 Chair der Sitzung: Claudia Diehl
15:15 - 17:00	PS-12: Politische Bildung und Engagement: Perspektiven aus verschiedenen Bildungskontexten
EO 242	Ort: EO 242 Chair der Sitzung: Sabine Manzel
15:15 - 17:00	PS-13: Potenziale und Herausforderungen für das Lernen mit Virtual Reality
EO 256	Ort: EO 256 Chair der Sitzung: Miriam Hansen
15:15 - 17:00	PS-14: Künstliche Intelligenz in der Lehrerbildung und -praxis:
EO 159	Ort: EO 159 Chair der Sitzung: Franziska Baier-Mosch
15:15 - 17:00	PS-15: Sportunterricht als sozialer Kontext
EO 169	Ort: EO 169 Chair der Sitzung: Tino Edelmann
15:15 - 17:00	PS-16: Inklusion und Teilhabe im Bildungssystem: Rahmenbedingungen und Perspektiven
EO 382	Ort: EO 382 Chair der Sitzung: Julia Wittig
15:15 - 17:00	PS-17: Mathematik als Herausforderung: Emotionen, Ängste und Unterstützung im Lernprozess
SO 418	Ort: SO 418 Chair der Sitzung: Lena Keller
15:15 - 17:00	PS-18: Führung und Professionalisierung im Schulwesen: Einblicke in Fortbildung und Kooperation
EO 157	Ort: EO 157 Chair der Sitzung: Stephan Gerhard Huber
15:15 - 17:00	PS-19: Technologie und selbstreguliertes Lernen: Perspektiven aus der Bildungsforschung
EO 154	Ort: EO 154 Chair der Sitzung: Yoana Omarchevska
15:15 - 17:00	PS-20: Lesekompetenz und Motivation
SO 318	Ort: SO 318 Chair der Sitzung: Julia Künstler-Sment
15:15 - 17:00	S-14: Wissen und Stereotype zum Thema Hochbegabung von (angehenden) Lehrkräften
SN 163	Ort: SN 163 Chair der Sitzung: Julia Schiefer Chair der Sitzung: Friederike Zimmermann Chair der Sitzung: Jörn Sparfeldt
15:15 - 17:00	S-15: Multidimensionale Ansätze zur Beschreibung von Interaktionsprozessen in der Schule
O 151	Ort: O 151 Chair der Sitzung: Madita Frühauf Chair der Sitzung: Britta Wenzel Chair der Sitzung: Natalie Fischer
15:15 - 17:00	S-16: Mehr als ein Augenblick – Wie Eye-Tracking die Unterrichtsforschung bereichert
SN 169	Ort: SN 169 Chair der Sitzung: Özün Keskin Chair der Sitzung: Tosca Daltoè Chair der Sitzung: Kathleen Stürmer
15:15 - 17:00	S-17: Heimatnah oder fern der Heimat? Empirische Befunde zur regionalen Mobilität von (angehenden) Lehrkräften
EO 150	Ort: EO 150 Chair der Sitzung: Svenja Mareike Schmid-Kühn Chair der Sitzung: Denise Demski
15:15 - 17:00	S-18: Selbstgesteuertes Lernen in digitalen Lernumgebungen – Erfassung und Analyse der Kompetenzen von Studierenden im Umgang mit Online-Informationen
SO 108	Ort: SO 108 Chair der Sitzung: Olga Troitschanskaia

	Chair der Sitzung: Johannes Hartig Chair der Sitzung: Samuel Greiff
15:15 - 17:00	S-19: Wirksamkeit von Interventionen in der Lehrkräftebildung – Curriculare und methodische Innovationen
M 003	Ort: M 003 Chair der Sitzung: Johannes König Chair der Sitzung: Sarah Strauß Chair der Sitzung: Dirk Richter
15:15 - 17:00	S-20: Zeitlicher Bezug als vernachlässigter Aspekt bei der Untersuchung von kognitiver Aktivierung
EO 165	Ort: EO 165 Chair der Sitzung: Benjamin Herbert Chair der Sitzung: Anna-Katharina Praetorius Chair der Sitzung: Matthias Martens
15:15 - 17:00	S-21: Gelingende Ko-Konstruktion von Bildungswissenschaften, Schulpraxis und Bildungsadministration: Empirische Ergebnisse zu ko-konstruktiver Zusammenarbeit in der Initiative Schule Macht Stark (SchuMaS)
O 148	Ort: O 148 Chair der Sitzung: Marie Baesch Chair der Sitzung: Simon Ohl Chair der Sitzung: Birgit Pikowsky
15:15 - 17:00	S-22: Wohlbefinden und psychische Gesundheit verschiedener Akteure im Bildungsbereich. Messung, Erklärungen, Entwicklungen.
O 145	Ort: O 145 Chair der Sitzung: Felix Kruse Chair der Sitzung: Arvid Nagel Chair der Sitzung: Sarah Bürger
15:15 - 17:00	S-23: Vielschichtige Perspektiven auf das Wohlbefinden von Schüler*innen: Prädiktoren, Profile und förderliche Faktoren in heterogenen Lernumwelten
O 142	Ort: O 142 Chair der Sitzung: Alyssa Laureen Grecu Chair der Sitzung: Katharina Molitor Chair der Sitzung: Gerda Hagenauer
15:15 - 17:00	S-24: Gestaltung und Wirksamkeit früher Sprachförderinterventionen
O 135	Ort: O 135 Chair der Sitzung: Leonie Dargiewicz Chair der Sitzung: Svenja Wehrhöfer Chair der Sitzung: Annika Ohle-Peters Chair der Sitzung: Gisela Kammermeyer
15:15 - 17:00	S-25: Bildung für nachhaltige Entwicklung ist mehr als nur Wissen – Einstellungen, Ängste und Kompetenzen von Lernenden als Untersuchungsgegenstände in der Bildungsforschung von Deutschland, Österreich und der Schweiz
O 133	Ort: O 133 Chair der Sitzung: Jennifer Stemmann Chair der Sitzung: Werner Rieß Chair der Sitzung: Benjamin Fauth
15:15 - 17:00	S-26: Förderbezogene Diagnostik im inklusiven Mathematikunterricht
O 129	Ort: O 129 Chair der Sitzung: Fynn Töllner Chair der Sitzung: Poldi Kuhl Chair der Sitzung: Michael Besser Chair der Sitzung: Daniel Sommerhoff
15:15 - 17:00	S-27: Erfassung und Förderung digitalisierungsbezogener Kompetenzen von (angehenden) Lehrkräften
EO 162	Ort: EO 162 Chair der Sitzung: Isabell Runge Chair der Sitzung: Jakob Thomas Fischer Chair der Sitzung: Andreas Lachner
17:15 - 18:45	Podiumsdiskussion: Wie hältst du's mit der Evidenz?
SO 108	Ort: SO 108 Chair der Sitzung: Oliver Dickhäuser
17:15 - 18:45	Podiumsdiskussion: Wie hältst du's mit der Evidenz? - Digitale Live-Übertragung
SN 169	Ort: SN 169
17:15 - 18:45	Podiumsdiskussion: Wie hältst du's mit der Evidenz? - Digitale Live-Übertragung
SN 163	Ort: SN 163
19:00 - 20:15	Posters & Wine
O 001	Ort: O 001 Chair der Sitzung: Karina Karst
19:00 - 20:15	

O 102	Posters & Wine Ort: O 102 Chair der Sitzung: Karina Karst
19:00 - 20:15 EO 154	Vorstandssitzung GEBF Ort: EO 154
20:15 - 20:30 O 102	Verleihung des Waxmann-Posterpreises Ort: O 102 Chair der Sitzung: Karina Karst
Datum: Dienstag, 28.01.2025	
8:30 - 17:30 O 226 - 228	Kinderbetreuung Agentur Farbenfroh Ort: O 226 - 228
9:00 - 10:45 SO 108	IS-03: Invited Symposium: Kompetenzdiagnostik in der beruflichen Bildung Ort: SO 108 Chair der Sitzung: Stephan Abele Chair der Sitzung: Carolin Hahnel
9:00 - 10:45 EO 159	OF-3: Special session: Die GEBF Open Science Initiative: Aktueller Stand, Arbeitsfelder und zukünftige Beteiligungsmöglichkeiten Ort: EO 159 Chair der Sitzung: Jürgen Schneider
9:00 - 10:45 SO 322	PS-21: Chemieunterricht: Herausforderungen und digitale Ansätze Ort: SO 322 Chair der Sitzung: Amina Zerouali
9:00 - 10:45 EO 169	PS-22: Hochschulbildung Ort: EO 169 Chair der Sitzung: Zouhier Kassabali
9:00 - 10:45 EO 256	PS-23: Selbstwirksamkeit angehender Lehrkräfte Ort: EO 256 Chair der Sitzung: Stefanie Rach
9:00 - 10:45 O 129	PS-24: Digitale Personal- und Organisationsentwicklung Ort: O 129 Chair der Sitzung: Nina Carolin Werth
9:00 - 10:45 EO 162	PS-25: Digitales Lesen und Schreiben: Herausforderungen und Potenziale für die schulische Bildung Ort: EO 162 Chair der Sitzung: Johannes Naumann
9:00 - 10:45 SO 418	PS-26: Förderung und Transfer Ort: SO 418 Chair der Sitzung: Svenja Lemmrich
9:00 - 10:45 EO 242	PS-27: Innovation und Qualität im MINT-Unterricht: Beiträge zur Bildungsforschung und Praxis Ort: EO 242 Chair der Sitzung: Stefan Ufer
9:00 - 10:45 SO 318	PS-28: Education and Equal Opportunities: Impacts of the COVID Pandemic Ort: SO 318 Chair der Sitzung: Markus Vogelbacher English Track
9:00 - 10:45 EO 382	PS-29: Sprachliche Bildung und Mehrsprachigkeit: Grundlagen und Förderung Ort: EO 382 Chair der Sitzung: Patrick Dahdah
9:00 - 10:45 EO 157	PS-30: Identität, Diskriminierung und interkulturelles Lernen im Schulkontext Ort: EO 157 Chair der Sitzung: Sharleen Pevec-Zimmer
9:00 - 10:45 EO 154	PS-31: Language and Multilingualism in Education Ort: EO 154 Chair der Sitzung: Kerstin Göbel English Track
9:00 - 10:45 SN 163	S-28: Diagnostik schriftsprachlicher Fähigkeiten: Effekte der Beurteilenden und der Ratingkriterien auf Urteilsqualität und -prozesse. Ort: SN 163 Chair der Sitzung: Johannes Hartig Chair der Sitzung: Johanna Fleckenstein Chair der Sitzung: Stefan Keller
9:00 - 10:45	

EO 145	S-29: Urteile von Lehrkräften: Entwicklung, Veränderbarkeit und Auswirkungen Ort: EO 145 Chair der Sitzung: Jan Beck Chair der Sitzung: Detlef Urhahne
9:00 - 10:45	S-30: Merkmale von Unterrichtsqualität für Schüler:innen mit sonderpädagogischem Förderbedarf: Welche Erkenntnisse liegen vor? Ort: EO 150 Chair der Sitzung: Elisabeth Moser Opitz Chair der Sitzung: Mireille Tabin Chair der Sitzung: Jasmin Decristan
EO 150	
9:00 - 10:45	S-31: KI im Bildungskontext: Angebote von Lehrkräften und Nutzungsverhalten von Schüler:innen Ort: M 003 Chair der Sitzung: Jana Kemmler Chair der Sitzung: Jessica Heesen
M 003	
9:00 - 10:45	S-32: Wandel von Schul- und Unterrichtskulturen über die letzten 50 Jahre – interdisziplinäre Einblicke in das Projekt „Schule im Wandel“ Ort: O 102 Chair der Sitzung: Tobias Feldhoff Chair der Sitzung: Sabine Reh Chair der Sitzung: Eckhard Klieme Chair der Sitzung: Katharina Maag Merki
O 102	
9:00 - 10:45	S-33: Selbstregulation in digitalen Lernumgebungen – Operationalisierung und Validierung verschiedener Facetten Ort: O 151 Chair der Sitzung: Lea Marie Nobbe Chair der Sitzung: Leonard Tetzlaff Chair der Sitzung: Laura Dörrenbächer-Ulrich
O 151	
9:00 - 10:45	S-34: Auslandsaufenthalte während der Schulzeit – Zugang, Entwicklungsverläufe und deren Einflussfaktoren Ort: SN 169 Chair der Sitzung: Matthias Kreuch Chair der Sitzung: Annabell Daniel Chair der Sitzung: Marita Jacob
SN 169	
9:00 - 10:45	S-35: Unterrichtstörungen aus Sicht von Lehrpersonen und Schüler*innen Ort: EO 165 Chair der Sitzung: Sabine Glock Chair der Sitzung: Claudia Schuchart
EO 165	
9:00 - 10:45	S-36: Ursachen von digitalen Ungleichheiten und Konsequenzen für den Bildungserfolg Ort: O 148 Chair der Sitzung: Markus Lörz
O 148	
9:00 - 10:45	S-37: Modeling Learners Beyond the Mean: Erhebung und Statistische Modellierung intraindividuellder Dynamiken im Hochschulkontext Ort: O 145 Chair der Sitzung: Peter Adriaan Edelsbrunner Chair der Sitzung: Julia Dietrich
O 145	
9:00 - 10:45	S-38: From Data to Performance: Understanding and Enhancing Student Learning Behaviors Ort: O 142 Chair der Sitzung: Jakob Schwerter Chair der Sitzung: Carolin Baumann Chair der Sitzung: Markus Dresel English Track
O 142	
9:00 - 10:45	S-39: Forschungssynthesen und Digitalisierung: Empirische Befunde und methodische Ansätze Ort: O 133 Chair der Sitzung: Lisa Ziernwald Chair der Sitzung: Marcel Capparozza Chair der Sitzung: Frank Fischer
O 133	
9:00 - 10:45	S-40: Application of Generative Large Language Models in Teacher-Centered Educational Research Ort: O 135 Chair der Sitzung: Yizhen Huang Chair der Sitzung: Michael Sailer English Track
O 135	
11:15 - 12:15	Heike Solga: Neue Chancen oder verstärkte Benachteiligung? Der Übergang von leistungsschwachen Schulabgänger:innen in die Ausbildung Ort: SO 108 Chair der Sitzung: Irena Kogan
SO 108	
11:15 - 12:15	

SN 169	Heike Solga: Neue Chancen oder verstärkte Benachteiligung? Der Übergang von leistungsschwachen Schulabgänger:innen in die Ausbildung - Digitale Live-Übertragung Ort: SN 169
11:15 - 12:15	Heike Solga: Neue Chancen oder verstärkte Benachteiligung? Der Übergang von leistungsschwachen Schulabgänger:innen in die Ausbildung - Digitale Live-Übertragung Ort: SN 163
SN 163	
12:20 - 12:30	Kongressfoto Ort: Ehrenhof
Ehrenhof	
13:15 - 15:00	DFG: Special session: Förderung der DFG im Bereich der empirischen Bildungsforschung Ort: EO 150 Chair der Sitzung: Annabell Koch
EO 150	
13:15 - 15:00	IS-04: Invited Symposium: Programme for the International Assessment of Adult Competencies (PIAAC 2023): Aktuelle Ergebnisse in Deutschland und im internationalen Vergleich Ort: O 102 Chair der Sitzung: Debora Maehler Chair der Sitzung: Beatrice Rammstedt
O 102	
13:15 - 15:00	OF-4: Special session: Care-Arbeit in der Wissenschaft: Herausforderungen und Perspektiven auf allen Karrierestufen Ort: EO 145 Chair der Sitzung: Meike Bonefeld Chair der Sitzung: Merle Thielmann
EO 145	
13:15 - 15:00	OF-5: World Café: GEBF-AG zur sinkenden Teilnahmebereitschaft an Erhebungen im Schulkontext – Gemeinsame Diskussion der Leitlinien des künftigen GEBF-Positionspapiers Ort: EO 159 Chair der Sitzung: Jutta von Maurice Chair der Sitzung: Jasmin Decristan Chair der Sitzung: Roman Auriga Chair der Sitzung: Bayer Sonja Chair der Sitzung: Karina Karst
EO 159	
13:15 - 15:00	PS-32: Bildungsdisparitäten Ort: EO 382 Chair der Sitzung: Katharina Fink
EO 382	
13:15 - 15:00	PS-33: Digitale Medien: Herausforderungen für die Entwicklung von Jugendlichen im digitalen Zeitalter Ort: SO 422 Chair der Sitzung: Mareike Thumel
SO 422	
13:15 - 15:00	PS-34: Inklusion und sozial-emotionale Entwicklung in der Grundschule Ort: SO 322 Chair der Sitzung: Stefan Kulakow
SO 322	
13:15 - 15:00	PS-35: Motivation und Emotionen im Studium Ort: EO 169 Chair der Sitzung: Robert Grassinger
EO 169	
13:15 - 15:00	PS-36: Soziale Ungleichheit und Lesekompetenz – Herausforderungen, Interventionen Ort: EO 256 Chair der Sitzung: Josefine Lüh
EO 256	
13:15 - 15:00	PS-37: Motivation in Hochschule und Freizeit Ort: EO 242 Chair der Sitzung: Martin Gartmeier
EO 242	
13:15 - 15:00	PS-38: Motivation und Selbstwirksamkeit im Mathematiklernen: Interventionen und ihre Wirkung Ort: EO 157 Chair der Sitzung: Janina Krawitz
EO 157	
13:15 - 15:00	PS-39: Digitale Transformation und Lehrerbildung Ort: SO 418 Chair der Sitzung: Christian Seyferth-Zapf
SO 418	
13:15 - 15:00	PS-40: Gender and Equality in the Classroom Ort: EO 154 Chair der Sitzung: Till Stefes English Track
EO 154	
13:15 - 15:00	

SO 318	PS-41: Familie, Bildung und Lebensverläufe Ort: SO 318 Chair der Sitzung: Maximilian Seitz
13:15 - 15:00	S-41: Geschlechtstypische Berufswahlprozesse: neue Befunde zu Mechanismen und Interventionsmöglichkeiten Ort: SO 108 Chair der Sitzung: Brigitte Schels Chair der Sitzung: Corinna Kleinert Chair der Sitzung: Kathrin Leuze
SO 108	
13:15 - 15:00	S-42: Gesundheitsfördernde Führung in der Schule – Drei Perspektiven Ort: EO 165 Chair der Sitzung: Anita Sandmeier
EO 165	
13:15 - 15:00	S-43: Open Science in der (quantitativen) Bildungsforschung: Herausforderungen, Werkzeuge und Perspektiven Ort: SN 163 Chair der Sitzung: Jürgen Schneider
SN 163	
13:15 - 15:00	S-44: Augmented Reality beim Lernen durch Experimentieren Ort: O 151 Chair der Sitzung: Sarah Malone Chair der Sitzung: Jochen Kuhn Chair der Sitzung: Andreas Lachner
O 151	
13:15 - 15:00	S-45: Reflexion im Fokus: Einblicke in den Reflexionsprozess von der Initiierung bis zur Förderung Ort: O 148 Chair der Sitzung: Katrin Arendt Chair der Sitzung: Anja Friedrich Chair der Sitzung: Frank Fischer
O 148	
13:15 - 15:00	S-46: KI zur Förderung und Analyse von Lehr- und Lernprozessen Ort: M 003 Chair der Sitzung: Ulrike Cress Chair der Sitzung: Marcus Specht
M 003	
13:15 - 15:00	S-47: Potenziale besser erkennen und fördern: Neue Ansätze in der Begabungsforschung Ort: O 145 Chair der Sitzung: Anne-Kathrin Stiller Chair der Sitzung: Tanja Gabriele Baudson
O 145	
13:15 - 15:00	S-48: Ein guter Start ins Studium? Eingangsvoraussetzungen und Erwartungen der Studierenden als Gelingensbedingungen eines Hochschulstudiums Ort: O 142 Chair der Sitzung: Svenja Kaiser Chair der Sitzung: Katharina Vogel Chair der Sitzung: Leif Döring Chair der Sitzung: Stefan Münzer Chair der Sitzung: Stefan Janke
O 142	
13:15 - 15:00	S-49: Meta-Meta: Die Rolle von systematischen Übersichten zu Meta-Analysen in der aktuellen Bildungsforschung Ort: SN 169 Chair der Sitzung: Maximilian Knogler Chair der Sitzung: Valentin Emslander Chair der Sitzung: Doris Holzberger
SN 169	
13:15 - 15:00	S-50: Wie Lehrkräfte Familien sehen: Erkenntnisse aus Vignetten-Studien über Stereotype von Lehrkräften auf familiäre sozioökonomische Herkunft Ort: O 135 Chair der Sitzung: Janina Dickert Chair der Sitzung: Oscar Yendell Chair der Sitzung: Melanie Olczyk
O 135	
13:15 - 15:00	S-51: Interdisziplinäre Transferstrategien im Kompetenzverbund lernen:digital Ort: O 133 Chair der Sitzung: Luisa Scherzinger Chair der Sitzung: Annika Goeze
O 133	
13:15 - 15:00	S-52: Stimmgesundheit bei Lehrkräften: Innovative Ansätze zur Messung und Prävention Ort: O 129 Chair der Sitzung: Christian Gegner Chair der Sitzung: Uta Klusmann
O 129	
13:15 - 15:00	S-53: Motivationale Wirkung von Feedback: Welche Rolle spielen Heterogenitätsmerkmale der Schüler*innen? Ort: EO 162 Chair der Sitzung: Stefanie van Ophuysen Chair der Sitzung: Gruehn Sabine Chair der Sitzung: Lintorf Katrin Chair der Sitzung: Susanne Narciss
EO 162	

15:15 - 17:00	IS-05: Invited Symposium: Üben im Unterricht: Theoretische und empirische Überlegungen zu einer zentralen Dimension von Unterrichtsqualität
SO 108	Ort: SO 108 Chair der Sitzung: Stefan Keller Chair der Sitzung: Erin Gerlach
15:15 - 17:00	IS-06: Invited Symposium: Wissen oder Annahmen? Beiträge zur Validierung von Interpretationen aus Logdaten
SN 163	Ort: SN 163 Chair der Sitzung: Carolin Hahnel
15:15 - 17:00	OF-6: Diskussion: Was bringen die Öffnungen bei der Differenzierung des Bildungssystems?
O 102	Ort: O 102 Chair der Sitzung: Steffen Schindler Chair der Sitzung: Hans-Peter Blossfeld
15:15 - 17:00	PS-42: Sprachentwicklung und -Förderung
EO 154	Ort: EO 154 Chair der Sitzung: Steffi Sachse
15:15 - 17:00	PS-43: Herausforderungen und Ressourcen im Lehrberuf
SO 318	Ort: SO 318 Chair der Sitzung: Carmela Aprea
15:15 - 17:00	PS-44: Wahrnehmung und Wirkung von digitalem Feedback im Lernprozess
SO 418	Ort: SO 418 Chair der Sitzung: Marlit Annalena Lindner
15:15 - 17:00	PS-45: Innovationen und Professionalisierung in der Lehrkräftebildung
EO 382	Ort: EO 382 Chair der Sitzung: Sara Großbruchhaus
15:15 - 17:00	PS-46: Bildung und Berufseinstieg unter den Bedingungen der COVID-19-Pandemie
SO 322	Ort: SO 322 Chair der Sitzung: Wilhelm Koschel
15:15 - 17:00	PS-47: Kooperation und Problemlösen im Unterricht
EO 256	Ort: EO 256 Chair der Sitzung: Christina Wekerle
15:15 - 17:00	PS-48: Vertrauen in Wissenschaft und die Beurteilung wissenschaftlicher Argumente
EO 242	Ort: EO 242 Chair der Sitzung: Nina Udvardi-Lakos
15:15 - 17:00	PS-49: KI und sprachliche Bildung: Innovationen in der Lernförderung und Leistungsbewertung
EO 159	Ort: EO 159 Chair der Sitzung: Fabian Zehner
15:15 - 17:00	PS-50: Psychosoziale Herausforderungen und schulisches Wohlbefinden von Jugendlichen
EO 157	Ort: EO 157 Chair der Sitzung: Tim Hamer
15:15 - 17:00	PS-51: Schulleitung und digitale Transformation: Impulse aus der Bildungsforschung
EO 169	Ort: EO 169 Chair der Sitzung: Andreas Harder
15:15 - 17:00	S-54: Demokratie lernen: Schulische Einflüsse auf die politische Entwicklung junger Menschen
EO 162	Ort: EO 162 Chair der Sitzung: Katharina Eckstein Chair der Sitzung: Miriam Schwarzenhal Chair der Sitzung: Peter Noack
15:15 - 17:00	S-55: Vielfalt in der Hochschule: Chrononormativität an Deutschlands Hochschulen – oder doch nicht?
O 129	Ort: O 129 Chair der Sitzung: Jessica Ordemann Chair der Sitzung: Laura Naegele
15:15 - 17:00	S-56: Individuelle Unterschiede im selbstregulierten Lernen
O 133	Ort: O 133 Chair der Sitzung: Charlotte Dignath Chair der Sitzung: Olena Kryshko Chair der Sitzung: Franziska Perels
15:15 - 17:00	S-57: Forschungsdaten und Open Science – Praktische Hinweise zum Erheben, Teilen und Nachnutzen von Forschungsdaten und Angebote des Verbund Forschungsdaten Bildung
O 135	Ort: O 135

	Chair der Sitzung: Aleksander Kocaj Chair der Sitzung: Bayer Sonja
15:15 - 17:00	S-58: Mehrdimensionale Kompetenzorientierung im mathematisch-naturwissenschaftlichen Unterricht – Ergebnisse aus PISA-Ceco 2022 Ort: SN 169 Chair der Sitzung: Doris Lewalter Chair der Sitzung: Anja Schiepe-Tiska Chair der Sitzung: Samuel Greiff
15:15 - 17:00	S-59: „Ein soziales Merkmal kommt selten allein“: Intersektionalitätseffekte in Lehrkrafturteilen und dem Erleben von Schüler*innen Ort: O 142 Chair der Sitzung: Ineke M. Pit-ten Cate Chair der Sitzung: Sabine Glock Chair der Sitzung: Katja Scharenberg
15:15 - 17:00	S-60: Veränderungen beruflicher Orientierungen im Kontext struktureller Herausforderungen - Die Rolle sozialer, regionaler und fluchtspezifischer Faktoren Ort: M 003 Chair der Sitzung: Alexandra Wicht Chair der Sitzung: Oliver Winkler Chair der Sitzung: Helen Hickmann
15:15 - 17:00	S-61: Educational Chatbots: Ein interdisziplinärer Blick auf aktuelle Forschungsfragen zu hybriden und generativen Chatbots in Unterricht und Lehrkräftebildung Ort: O 145 Chair der Sitzung: Steven Beyer Chair der Sitzung: Maria Klar Chair der Sitzung: Sebastian Becker-Genschow
15:15 - 17:00	S-62: Chancen und Risiken KI-gestützter Literaturüberblicke und Metaanalysen Ort: EO 145 Chair der Sitzung: Tim Fütterer Chair der Sitzung: Olga Chernikova Chair der Sitzung: Jürgen Schneider
15:15 - 17:00	S-63: Inter- und intraindividuelle Variation im Zusammenspiel von selbstreguliertem Lernen, Metakognition und exekutiven Funktionen: Erkenntnisse aus verschiedenen Altersgruppen (DFG Netzwerk SeReNe) Ort: O 148 Chair der Sitzung: Laura Dörrenbächer-Ulrich Chair der Sitzung: Maria Theobald
15:15 - 17:00	S-64: Understanding Student Motivation, Emotion, and Engagement in Higher Education: Perspectives from Longitudinal and Context-specific Studies Ort: EO 165 Chair der Sitzung: Luise von Keyserlingk Chair der Sitzung: Jakob Schwerter Chair der Sitzung: Stefan Janke English Track
15:15 - 17:00	S-65: Partizipation und Transfer in der kulturellen Bildung Ort: O 151 Chair der Sitzung: Lisa Birnbaum Chair der Sitzung: Katrin Rakoczy
15:15 - 17:00	S-66: Herausforderungen im Umgang mit Belastungen durch eine Stärkung personaler Ressourcen von Lehrkräften und Schüler:innen begegnen Ort: EO 150 Chair der Sitzung: Franziska Greiner-Döcher Chair der Sitzung: Hendrik Lohse-Bossenz Chair der Sitzung: Gerda Hagenauer
17:15 - 19:00	Mitgliederversammlung Ort: SO 108 Chair der Sitzung: Tina Seidel
19:30	Gesellschaftsabend
Datum: Mittwoch, 29.01.2025	
8:30 - 17:30	Kinderbetreuung Agentur Farbenfroh Ort: O 226 - 228
9:00 - 10:45	IS-07: Invited Symposium: Zentrale Ergebnisse aus TIMSS 2023 im Überblick und ihre Bedeutung für die Bildungsforschung, -politik und -praxis Ort: O 102 Chair der Sitzung: Christin Beese Chair der Sitzung: Luise Scholz Chair der Sitzung: Christine Sälzer
9:00 - 10:45	

O 129	PS-52: Soziale Mobilität und Bildungserfolg Ort: O 129 Chair der Sitzung: Thomas Zimmermann
9:00 - 10:45	PS-53: Diversität in der Bildung Ort: O 133 Chair der Sitzung: Diana Raufelder
O 133	
9:00 - 10:45	PS-54: Führung und Entwicklung im Schulwesen: Internationale Perspektiven und aktuelle Herausforderungen Ort: SO 318 Chair der Sitzung: Olivia Wohlfart
SO 318	
9:00 - 10:45	PS-55: Führung und Innovation in benachteiligten Schulen: Perspektiven für eine gerechte Bildungsentwicklung Ort: SO 418 Chair der Sitzung: Rukiye Ates
SO 418	
9:00 - 10:45	PS-56: Professionelle Perspektiven auf Inklusion Ort: EO 169 Chair der Sitzung: Veronika Dumbacher
EO 169	
9:00 - 10:45	PS-57: Bildungsqualität und Fachkräfte: Herausforderungen und Strategien im Schul- und Betreuungswesen Ort: O 135 Chair der Sitzung: Chantal Mose
O 135	
9:00 - 10:45	PS-58: Herausforderung digitale Lehre: Perspektiven auf Lernen und Zusammenarbeit Ort: SO 422 Chair der Sitzung: Maik Beege
SO 422	
9:00 - 10:45	PS-59: spezielle Themen der Motivation & Metakognition Ort: EO 256 Chair der Sitzung: Miriam Wunsch
EO 256	
9:00 - 10:45	PS-60: Diversität im MINT-Bereich als Herausforderung für Bildung und Chancengleichheit Ort: O 142 Chair der Sitzung: Charlotte Popp
O 142	
9:00 - 10:45	PS-61: Eltern-LehrerInnen Kooperation & elterliches Handeln im Schulkontext Ort: SO 322 Chair der Sitzung: Josef Strasser
SO 322	
9:00 - 10:45	S-67: Diagnosegeleitete Leseförderung in der und für die Schulpraxis Ort: EO 154 Chair der Sitzung: Sarah Bez Chair der Sitzung: Karina Karst Chair der Sitzung: Vanessa Völlinger
EO 154	
9:00 - 10:45	S-68: Übergänge im Jugendalter und der (unsichere) Blick auf die Zukunft: Berufsorientierung und Zukunftsangst von Jugendlichen in Deutschland und Österreich Ort: O 148 Chair der Sitzung: Lisa Hasenbein Chair der Sitzung: Stefan Hofherr
O 148	
9:00 - 10:45	S-69: Methodische Innovationen bei der Konstruktion von Test- und Fragebogenverfahren Ort: EO 159 Chair der Sitzung: Andreas Frey Chair der Sitzung: Johannes Hartig Chair der Sitzung: Joachim Ulf Kröhne
EO 159	
9:00 - 10:45	S-70: Erfassung musikbezogener Kompetenzen Ort: EO 242 Chair der Sitzung: Stefan Münzer Chair der Sitzung: Johannes Hasselhorn Chair der Sitzung: Andreas C. Lehmann
EO 242	
9:00 - 10:45	S-71: Einzelschulen im Kontext räumlicher und sozialer Ungleichheit: Die Rolle des schulischen Angebots und der elterlichen Schulwahl Ort: O 151 Chair der Sitzung: Carolin Zink Chair der Sitzung: Nicky Zunker Chair der Sitzung: Stefan Immerfall
O 151	
9:00 - 10:45	S-72: Inklusive Bildung: Die Einstellungen von Lehrkräften im Mittelpunkt Ort: EO 157 Chair der Sitzung: Thorsten Fahrbach Chair der Sitzung: Carmen Zurbruggen
EO 157	
9:00 - 10:45	S-73: Benefits of Generative Learning and Desirable Difficulties for Meaningful Learning: Combining Learning Activities and Addressing Learner Characteristics
SN 163	

	Ort: SN 163 Chair der Sitzung: Veit Kubik Chair der Sitzung: Matthias Nückles English Track
9:00 - 10:45 EO 150	S-74: Historisch denken und (mit-)fühlen lernen. Arbeit mit Zeitzeug*innen ein Beitrag für die Demokratiebildung? Ort: EO 150 Chair der Sitzung: Christiane Bertram Chair der Sitzung: Martin Nitsche
9:00 - 10:45 O 131	S-75: Museen als Orte informeller Bildung Ort: O 131 Chair der Sitzung: Alexandra W. Busch Chair der Sitzung: Olaf Köller
9:00 - 10:45 EO 162	S-76: Lernen durch Feedback von Lehrpersonen, Eltern und Schüler:innen Ort: EO 162 Chair der Sitzung: Sebastian Röhl Chair der Sitzung: Fabian Hoya
9:00 - 10:45 SO 108	S-77: Studienabbruch im Lehramt: Ursachen, Prädiktoren und Lösungsansätze Ort: SO 108 Chair der Sitzung: Stefan Sorge Chair der Sitzung: Thilo Kleickmann Chair der Sitzung: Martin Neugebauer
9:00 - 10:45 EO 145	S-78: Förderung evidenzinformierter Entscheidungen von Lehrkräften Ort: EO 145 Chair der Sitzung: Samuel Merk Chair der Sitzung: Kirstin Schmidt Chair der Sitzung: Ulrich Trautwein
9:00 - 10:45 SN 169	S-79: Challenges and opportunities for young refugees in Germany: How institutional conditions shape competencies and educational careers Ort: SN 169 Chair der Sitzung: Melanie Olczyk Chair der Sitzung: Julian Seuring Chair der Sitzung: Gisela Will Chair der Sitzung: Birgit Becker English Track
9:00 - 10:45 EO 382	S-80: Heterogenität der Schülerschaft und herkunftsbedingte Bildungsungleichheiten: Welche Rolle spielen Lehrkräfte, Lehrkrafturteile und Unterrichtsgestaltung? Ort: EO 382 Chair der Sitzung: Katja Scharenberg Chair der Sitzung: Katharina Maag Merki
9:00 - 10:45 M 003	S-81: Computational modeling and educational sciences: Synergies and advancements Ort: M 003 Chair der Sitzung: Blazej Baczkowski Chair der Sitzung: Manuel D.S. Hopp Chair der Sitzung: Kou Murayama English Track
11:15 - 12:15 SO 108	Emilia Del Bono: The Trouble (and Value) of Child Socio-Emotional skills – Co-Note by Michèle Tertilt Ort: SO 108 Chair der Sitzung: Guido Neidhöfer
11:15 - 12:15 SN 169	Emilia del Bono: The Trouble (and Value) of Child Socio-Emotional skills – Co-Note by Michèle Tertilt - Digitale Live-Übertragung Ort: SN 169
11:15 - 12:15 SN 163	Emilia del Bono: The Trouble (and Value) of Child Socio-Emotional skills – Co-Note by Michèle Tertilt - Digitale Live-Übertragung Ort: SN 163
13:15 - 15:00 O 102	IS-08: Invited Symposium: Die Integration von Geflüchteten in das deutsche Bildungssystem: Herausforderungen und Perspektiven Ort: O 102 Chair der Sitzung: Irena Kogan Chair der Sitzung: Jörg Dollmann
13:15 - 15:00 O 129	PS-62: Inklusion in der Schule aus Perspektive von Lehrenden & Lernenden Ort: O 129 Chair der Sitzung: David Scheer

13:15 - 15:00	PS-63: Motivation in der Hochschulbildung 2 Ort: O 133 Chair der Sitzung: Florian Klapproth
O 133	
13:15 - 15:00	PS-64: Education and Career Prospects Ort: O 135 Chair der Sitzung: Hanna Beißert
O 135	
13:15 - 15:00	PS-65: Digital Media and Professional Development: Importance for Lifelong Learning and Career Growth Ort: EO 256 Chair der Sitzung: Falk Scheidig
EO 256	
13:15 - 15:00	PS-66: Evidenzorientierung und Wahrnehmung in der Lehrkräftebildung Ort: EO 169 Chair der Sitzung: Linda Schirle
EO 169	
13:15 - 15:00	PS-67: Emotionale und soziale Dimensionen der (Lehrkräfte-)Professionalisierung Ort: O 142 Chair der Sitzung: Anett Wolgast
O 142	
13:15 - 15:00	PS-68: Feedback in Lern- und Leistungsprozessen: Wahrnehmung, Wirkung und Gestaltung Ort: SO 322 Chair der Sitzung: Nikolai Zinke
SO 322	
13:15 - 15:00	PS-69: Digitale Lehrmethoden und ihre Rolle im Lernprozess: Neue Ansätze für die Fachbildung Ort: SO 318 Chair der Sitzung: Hanna Meiners
SO 318	
13:15 - 15:00	PS-70: Digitale Lernumgebungen als Schlüssel zur Förderung von wissenschaftlichem Denken und Praxis Ort: O 145 Chair der Sitzung: Miriam Lechner
O 145	
13:15 - 15:00	PS-71: Pädagogische Überzeugungen und Qualität in der frühkindlichen Bildung Ort: SO 422 Chair der Sitzung: Franziska Cohen
SO 422	
13:15 - 15:00	S-103: Auf dem Weg zu adaptiven Online-Fortbildungen? Voraussetzungen von Lehrkräften und ihr Nutzungsverhalten Ort: O 148 Chair der Sitzung: Tim Fütterer Chair der Sitzung: Anika Dreher Chair der Sitzung: Daniela Rzejak
O 148	
13:15 - 15:00	S-82: Technologiebasierte Diagnostik und Förderung von Schreibkompetenzen: Herausforderungen und Potenziale Ort: O 151 Chair der Sitzung: Johanna Fleckenstein Chair der Sitzung: Johannes Hartig Chair der Sitzung: Vera Busse
O 151	
13:15 - 15:00	S-83: Digitale Förderung von Kompetenzen im selbstregulierten Lernen – Konzeptionen, Chancen und Herausforderungen Ort: M 003 Chair der Sitzung: Carmen Nadja Hirt Chair der Sitzung: Mathias Mejeh Chair der Sitzung: Dirk Ifenthaler
M 003	
13:15 - 15:00	S-84: Emerging Pathways: Gelingendes Studieren in Fernstudiengängen Ort: O 131 Chair der Sitzung: Stefanie Habermann Chair der Sitzung: Michael Hast
O 131	
13:15 - 15:00	S-85: Herausforderungen einer indikatorengestützten Bildungsberichterstattung Ort: SN 169 Chair der Sitzung: Susanne Lochner Chair der Sitzung: Bettina Arnoldt Chair der Sitzung: Josefine Lühe
SN 169	
13:15 - 15:00	S-86: Wie können soziale Bildungsungleichheiten reduziert werden? Eine interdisziplinäre Perspektive auf formale und non-formale Bildung Ort: SO 108 Chair der Sitzung: Jan Scharf Chair der Sitzung: Hannah Kleen Chair der Sitzung: Hanna Dumont
SO 108	
13:15 - 15:00	S-87: Transitionen und Karriereverläufe in der Wissenschaft Ort: EO 145 Chair der Sitzung: Martin Daumiller Chair der Sitzung: Johanna Ott Chair der Sitzung: Ronja Steinhauser
EO 145	
13:15 - 15:00	

EO 157	S-88: Non-formale berufliche Weiterbildung als Schlüssel zur Anpassung an eine sich wandelnde Arbeitswelt Ort: EO 157 Chair der Sitzung: Benjamin Schimke
13:15 - 15:00	S-89: Reproduzierbarkeit und Datenqualität bei Forschungssynthesen: Status Quo und Lösungsansätze für eine transparente und nachhaltige Methode Ort: EO 159 Chair der Sitzung: Nadeshda Jung Chair der Sitzung: Tamara Heck Chair der Sitzung: Tanja Burgard
EO 159	
13:15 - 15:00	S-90: Forschungswissen für die Bildungspraxis: Forschungsbasierte Gestaltung von Unterstützungsangeboten für (angehende) Lehrkräfte Ort: EO 162 Chair der Sitzung: Pia Sander Chair der Sitzung: Steven Beyer Chair der Sitzung: Isabell van Ackeren-Mindl
EO 162	
13:15 - 15:00	S-91: ENTFÄLLT: Persönlichkeitsmerkmale und ihre Relevanz auf Beurteilungen im Bildungsbereich Ort: EO 242 Chair der Sitzung: Valerie Gitzi Chair der Sitzung: Vanda Capon-Sieber Chair der Sitzung: Richard Göllner
EO 242	
13:15 - 15:00	S-92: Anstrengung als Schlüssel zur Bewältigung von Lernanforderungen Ort: EO 382 Chair der Sitzung: Laura Kehle Chair der Sitzung: Detlef Urhahne Chair der Sitzung: Diana Raufelder
EO 382	
13:15 - 15:00	S-93: Erwachsenenbildung und gesellschaftlicher Wandel. Methodische Herausforderungen und Innovationen in der Erfassung von (Weiter-) Bildungsinhalten. Ort: EO 150 Chair der Sitzung: Jana Costa Chair der Sitzung: Lena Nusser Chair der Sitzung: Sonja Hahn
EO 150	
13:15 - 15:00	S-94: Eine ganzheitliche Perspektive auf Bildung wagen: Lehrkräfte und Schüler*innen an Schulen in herausfordernden Lagen stärken Ort: SN 163 Chair der Sitzung: Patrick Hawlitschek Chair der Sitzung: Simone Jambor-Fahlen Chair der Sitzung: Hans Pant
SN 163	
13:15 - 15:00	S-95: Educational Neuroscience: Bridging Brain Mechanisms and Learning Practices Ort: EO 154 Chair der Sitzung: Sabine Oligschläger Chair der Sitzung: Michiko Murayama Chair der Sitzung: Kou Murayama English Track
EO 154	
15:15 - 17:00	IS-09: Invited Symposium: Digitale Souveränität als Schlüssel zur Bewältigung gesellschaftlicher Herausforderungen - Beiträge der politischen Bildung Ort: O 102 Chair der Sitzung: Monika Oberle Chair der Sitzung: Johannes Hartig
O 102	
15:15 - 17:00	PS-72: Prüfungen und Innovationen: Fairness, Emotionen und neue Ansätze im Hochschulkontext Ort: SO 418 Chair der Sitzung: Hernán González Cruz
SO 418	
15:15 - 17:00	PS-73: Mathematiklernen: Didaktik, Aufgaben und Kompetenzen Ort: EO 154 Chair der Sitzung: Ann-Katrin van den Ham
EO 154	
15:15 - 17:00	PS-74: Social Background, Educational Success, and Integration: The Impact of Education on Inequality and Aspiration Ort: EO 150 Chair der Sitzung: Ann-Marie Sommerfeld English Track
EO 150	
15:15 - 17:00	PS-75: Mitbestimmung und Interaktion im Unterricht Ort: EO 157 Chair der Sitzung: Victoria Wiederseiner
EO 157	
15:15 - 17:00	PS-76: School Development in Transition: Current Challenges and Changes Ort: EO 382 Chair der Sitzung: Marko Neumann
EO 382	

	English Track
15:15 - 17:00	PS-77: Perspektiven auf Heterogenität im Unterricht
O 129	Ort: O 129 Chair der Sitzung: Andrea Vorderobermeier
15:15 - 17:00	PS-78: schulisches Engagement und Entwicklung: Individuelle, soziale und familiäre Faktoren
EO 242	Ort: EO 242 Chair der Sitzung: Felix Bittmann
15:15 - 17:00	PS-79: Berufliche Anforderungen und Bildungswege
O 133	Ort: O 133 Chair der Sitzung: Helena Geisler
15:15 - 17:00	PS-80: Bildungsentscheidungen und ihre Auswirkungen auf Persönlichkeit und Kompetenzen
O 135	Ort: O 135 Chair der Sitzung: Kamila Cygan-Rehm
15:15 - 17:00	PS-81: Weiterbildung: Motive, Gelingensbedingungen, Effekte
SO 318	Ort: SO 318 Chair der Sitzung: Luca Farina Holtricher
15:15 - 17:00	PS-82: Bildungsverläufe: Bedeutung kontextueller Faktoren
O 142	Ort: O 142 Chair der Sitzung: Lukas Lazzara
15:15 - 17:00	PS-83: Einstellungen: Bedeutung für Bildungsentscheidungen, -verläufe und -outcomes
O 145	Ort: O 145 Chair der Sitzung: Dennis Hauk
15:15 - 17:00	S-100: Affektives Lernen mit XR-Technologien in der Umweltbildung
EO 162	Ort: EO 162 Chair der Sitzung: Jule Krüger Chair der Sitzung: Josef Buchner
15:15 - 17:00	S-101: Understanding and Fostering Giftedness: Approaches to Talent Development from Multiple Disciplines
SN 163	Ort: SN 163 Chair der Sitzung: Ulrich Trautwein Chair der Sitzung: Katrin Kunz English Track
15:15 - 17:00	S-102: Förderung der Bewertungskompetenzen von Informationen im postfaktischen Zeitalter
SN 169	Ort: SN 169 Chair der Sitzung: Philipp L. Marten Chair der Sitzung: Ingo Kollar
15:15 - 17:00	S-104: Berufsorientierung und Übergänge in berufliche Ausbildung von geflüchteten Jugendlichen und jungen Erwachsenen
SO 108	Ort: SO 108 Chair der Sitzung: Regina Becker Chair der Sitzung: Gisela Will Chair der Sitzung: Mona Granato
15:15 - 17:00	S-105: Prozessuale Perspektiven auf Lehrkraftmotivation und Unterricht
EO 145	Ort: EO 145 Chair der Sitzung: Markus Dresel Chair der Sitzung: Rebecca Lazarides
15:15 - 17:00	S-106: Unterrichtsqualität zwischen Interdisziplinarität, Intersubjektivität und interaktivem Wissensaustausch
O 151	Ort: O 151 Chair der Sitzung: Erin Gerlach Chair der Sitzung: Alfred Lindl Chair der Sitzung: Wiebke Langer Chair der Sitzung: Mirjam Steffensky
15:15 - 17:00	S-96: Adaptive Unterrichtsplanung: Konzepte, Praxis und Förderung schüler*innenbezogener Entscheidungsprozesse bei Lehrkräften
O 148	Ort: O 148 Chair der Sitzung: Daniel Scholl Chair der Sitzung: Simon Küth Chair der Sitzung: Katharina Schnitzler
15:15 - 17:00	S-97: (Post-)Sekundäre Sozialisationskontexte und ihr Einfluss auf Bildungs- und Berufsaspirationen
O 131	Ort: O 131 Chair der Sitzung: Annabell Daniel Chair der Sitzung: Frederick de Moll Chair der Sitzung: Steffen Schindler

15:15 - 17:00	S-98: Beobachtungsbasierte Erfassung von Unterrichtsqualität: Auf zu neuen Wegen? Ort: EO 159 Chair der Sitzung: Anna-Katharina Praetorius Chair der Sitzung: Thilo Kleickmann Chair der Sitzung: Richard Göllner
15:15 - 17:00	S-99: Herausforderungen in der Entwicklung von neuen Lehr-Lernformaten am Beispiel der Universitätsschule Dresden Ort: M 003 Chair der Sitzung: Anke Langner Chair der Sitzung: Birgit Lütje-Klose
Datum: Donnerstag, 30.01.2025	
7:00 - 7:45	Nachwuchstagung: Running with the Professors
8:30 - 15:30	Kinderbetreuung Agentur Farbenfroh Ort: O 226 - 228
9:00 - 9:15	Nachwuchstagung: Begrüßung Ort: O 102
9:15 - 10:45	Nachwuchstagung: Postersession "Auswertungsstrategien" Ort: O 102
9:15 - 10:45	Nachwuchstagung: Postersession "Erhebungsdesign" Ort: O 001
11:15 - 12:15	ENTFÄLLT: Nachwuchstagung: Workshop: Analyse von Log-Daten mit R (Teil 1) Ort: EO 159 PD Dr. Ulf Kröhne (DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation) TeilnehmerInnen wurden per Mail informiert.
11:15 - 12:15	Nachwuchstagung: Workshop: Academic Writing (Teil 1) Ort: EO 157 Prof. Dr. Jochen Gebauer (Universität Mannheim)
11:15 - 12:15	Nachwuchstagung: Workshop: Einführung in die Item-Response-Theorie (IRT) mit R: Theoretische Grundlagen und praktische Anwendungen (Teil 1) Ort: EO 169 Leonard Tetzlaff und Tobias Deribo (DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation)
11:15 - 12:15	Nachwuchstagung: Workshop: Erfolgreiches Forschungsdatenmanagement: Von der Datenarchivierung bis zur Forschungsethik (Teil 1) Ort: EO 162 Larissa Obst (RatSWD, Rat für Sozial- und Wirtschaftsdaten) Dr. Bernhard Miller (GESIS) Dr. Katarina Blask (ZPID) Dr. Salome Wagner (DIPF)
11:15 - 12:15	Nachwuchstagung: Workshop: Smart AI-Hacks: Mit Prompt-Engineering den Wissenschafts-Workflow optimieren (Teil 1) Ort: ExpLab (BWL-Bibliothek) Tina Helbig (Zentrum für Lehren und Lernen), Jorge Murcia Serra, Thomas Schmidt (Universitätsbibliothek Mannheim)
11:15 - 12:15	Nachwuchstagung: Workshop: Using State-of-the-Art Machine Learning Models: An Introduction to the Hugging Face Ecosystem (Teil 1) Ort: EO 154 Dr. Steffen Brandt (OpenCampus.SH)
13:30 - 14:30	ENTFÄLLT: Nachwuchstagung: Workshop: Analyse von Log-Daten mit R (Teil 2) Ort: EO 159 PD Dr. Ulf Kröhne (DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation) TeilnehmerInnen wurden per Mail informiert.
13:30 - 14:30	Nachwuchstagung: Workshop: Academic Writing (Teil 2) Ort: EO 157 Prof. Dr. Jochen Gebauer (Universität Mannheim)
13:30 - 14:30	Nachwuchstagung: Workshop: Einführung in die Item-Response-Theorie (IRT) mit R: Theoretische Grundlagen und praktische Anwendungen (Teil 2) Ort: EO 169 Leonard Tetzlaff und Tobias Deribo (DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation)

13:30 - 14:30	Nachwuchstagung: Workshop: Erfolgreiches Forschungsdatenmanagement: Von der Datenarchivierung bis zur Forschungsethik (Teil 2)
EO 162	Ort: EO 162 Larissa Obst (RatSWD, Rat für Sozial- und Wirtschaftsdaten) Dr. Bernhard Miller (GESIS) Dr. Katarina Blask (ZPID) Dr. Salome Wagner (DIPF)
13:30 - 14:30	Nachwuchstagung: Workshop: Smart AI-Hacks: Mit Prompt-Engineering den Wissenschafts-Workflow optimieren (Teil 2)
ExpLab (BWL-Bibliothek)	Ort: ExpLab (BWL-Bibliothek) Tina Helbig (Zentrum für Lehren und Lernen), Jorge Murcia Serra, Thomas Schmidt (Universitätsbibliothek Mannheim)
13:30 - 14:30	Nachwuchstagung: Workshop: Using State-of-the-Art Machine Learning Models: An Introduction to the Hugging Face Ecosystem (Teil 2)
EO 154	Ort: EO 154 Dr. Steffen Brandt (OpenCampus.SH)
14:30 - 15:00	Nachwuchstagung: informal farewell
O 102	Ort: O 102

Präsentationen

Special session: Meet the Editors

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 150

Chair der Sitzung: Sabine Glock

Das Publizieren eigener Arbeiten in wissenschaftlichen Journals ist ein wesentlicher Bestandteil des Arbeitsprozesses in der empirischen Bildungsforschung.

Unterschiedliche Journals unterscheiden sich dabei allerdings in ihrem Scope, den besonders erwünschten Beitragsformaten und den Spezifika des Review-Prozesses.

In dieser Session haben die Teilnehmenden Gelegenheit, mit Herausgeber*innen verschiedener Journals zusammenzutreffen. Die Personen auf dem Podium stellen in einer ersten Runde zunächst kurz „ihr“ Journal vor. In einer anschließenden offenen Runde besteht die Möglichkeit, Fragen zu aktuellen Themen des Publikationsprozesses an die Herausgeber*innen zu stellen.

Auch journalübergreifende Fragen zum Publizieren in der empirischen Bildungsforschung können dabei diskutiert werden.

Podium:

- Sabine Glock, Chief Editor, Social Psychology of Education
- Samuel Greiff, Chief Editor, Learning and Individual Differences, Associate Editor, Journal of Educational Psychology
 - Katrin Arens, Associate Editor, Learning and Individual Differences
 - Andreas Rausch, Associate Editor, Vocations and Learning
- Jürgen Seifried, Chief Editor, Zeitschrift für Berufs- und Wirtschaftspädagogik
 - Pia Pinger, Editorial Board Member, Review of Economic Studies
 - Stefan Ufer, Editor, Journal für Mathematikdidaktik
- Christina Felfe, Outgoing associate Editor, European Economic Review

Sonderveranstaltung: Meet the Editors

Sabine Glock

Bergische Universität Wuppertal, Deutschland

Das Publizieren eigener Arbeiten in wissenschaftlichen Journals ist ein wesentlicher Bestandteil des Arbeitsprozesses in der empirischen Bildungsforschung.

Unterschiedliche Journals unterscheiden sich dabei allerdings in ihrem Scope, den besonders erwünschten Beitragsformaten und den Spezifika des Review-Prozesses.

In dieser Session haben die Teilnehmenden Gelegenheit, mit Herausgeber*innen verschiedener Journals zusammenzutreffen. Die Personen auf dem Podium stellen in einer ersten Runde zunächst kurz „ihr“ Journal vor. In einer anschließenden offenen Runde besteht die Möglichkeit, Fragen zu aktuellen Themen des Publikationsprozesses an die Herausgeber*innen zu stellen. Auch journalübergreifende Fragen zum Publizieren in der empirischen Bildungsforschung können dabei diskutiert werden.

Podium:

- Sabine Glock, Chief Editor, Social Psychology of Education
- Samuel Greiff, Chief Editor, Learning and Individual Differences, Associate Editor, Journal of Educational Psychology
- Katrin Arens, Associate Editor, Learning and Individual Differences
- Andreas Rausch, Associate Editor, Vocations and Learning

- Jürgen Seifried, Chief Editor, Zeitschrift für Berufs- und Wirtschaftspädagogik
- Pia Pinger, Editorial Board Member, Review of Economic Studies
- Stefan Ufer, Editor, Journal für Mathematikdidaktik
- Christina Felfe, Outgoing associate Editor, European Economic Review

IS-01: Invited Symposium: Digitale Kompetenzen von Schüler*innen und Rahmenbedingungen des Kompetenzerwerbes im dritten internationalen Vergleich - Befunde und vertiefende Analysen der IEA-Studie ICILS 2023

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 102

Chair der Sitzung: Birgit Eickelmann

Chair der Sitzung: Nadine Fröhlich

Chair der Sitzung: Johannes Naumann

Digitale Kompetenzen von Schüler*innen und Rahmenbedingungen des Kompetenzerwerbes im dritten internationalen Vergleich - Befunde und vertiefende Analysen der IEA-Studie ICILS 2023

Chair(s): **Birgit Eickelmann** (Universität Paderborn, Deutschland), **Nadine Fröhlich** (Universität Paderborn, Deutschland)

Diskutant*in(nen): **Johannes Naumann** (Bergische Universität Wuppertal)

Angesichts der voranschreitenden gesellschaftlichen und technologischen Entwicklungen im Kontext der digitalen Transformation ergibt sich die Relevanz einer zeitgemäßen Gestaltung des Bildungssystems und damit einhergehend von Schule und Lehr-Lernprozessen. Die empirische Bildungsforschung hat dabei die Aufgabe, die zur Bewältigung dieser Entwicklungen notwendigen Kompetenzen und die Rahmenbedingungen des Kompetenzerwerbes zu untersuchen, um so wissenschaftlich-analytisches Wissen zu generieren und darauf aufbauend Steuerungswissen zur gezielten Entwicklung des Bildungsbereiches zu erzeugen (Gerick & Eickelmann, 2021).

Die internationale vergleichende Schulleistungsstudie International Computer and Information Literacy Study 2023 (ICILS 2023; Eickelmann et al., 2024, in Druck; Fraillon & Rožman, 2023) wird von der International Association for the Evaluation of Educational Achievement (IEA) koordiniert und untersucht zum dritten Mal nach ICILS 2018 (Eickelmann, Bos, Gerick, Goldhammer et al., 2019; Fraillon et al., 2019) und ICILS 2013 (u.a. Bos, Eickelmann, Gerick, Goldhammer et al., 2014) die computer- und informationsbezogenen Kompetenzen von Schüler*innen der achten Jahrgangsstufe sowie die Rahmenbedingungen des Kompetenzerwerbs. Die Studie wird, auch in Deutschland, mit ICILS 2023 zum zweiten Mal um das internationale Zusatzmodul ‚Computational Thinking‘ ergänzt, sodass zwei Kompetenzbereiche erfasst werden können.

Das hier eingereichte interdisziplinär angelegte Symposium bearbeitet beitragsübergreifend die Fragestellung nach dem Entwicklungsstand beider in der Studie betrachteter Kompetenzbereiche bei Achtklässler*innen und nutzt hierzu für vertiefende Zusammenhangsanalysen zudem Daten zu den Rahmenbedingungen des Kompetenzerwerbes in Deutschland im internationalen Vergleich. Die jeweiligen Ergebnisse werden vor dem Hintergrund des Tagungsschwerpunkts ‚Bildung als Schlüssel für gesellschaftliche Herausforderungen: Interdisziplinäre Beiträge aus der Bildungsforschung‘ diskutiert.

Die repräsentative Datengrundlage für die Analysen des hier eingereichten Symposiums bilden die Kompetenzdaten von Achtklässler*innen aus einer computerbasierten Testung sowie die nationalen und internationalen Fragebogendaten dieser Schüler*innen, sowie der Lehrkräfte und Schulleitungen in den 35 an der Studie ICILS 2023 beteiligten Bildungssystemen.

Im Rahmen des Symposiums werden in einem ersten Zugang (Beitrag 1) die ICILS-2023-Ergebnisse zu den computer- und informationsbezogenen Kompetenzen und den Kompetenzen im Bereich ‚Computational Thinking‘ von Achtklässler*innen im internationalen Vergleich betrachtet. Dabei werden im Hinblick auf Trendanalysen und die Untersuchung von Gruppenunterschieden Vergleiche über einen zehn- bzw. fünfjährigen Zeitraum betrachtet. Der zweite Beitrag (Beitrag 2) untersucht die Perspektive der Schüler*innen auf das Lernen mit digitalen Medien im Zusammenhang mit dem Erwerb computer- und informationsbezogener Kompetenzen. Der dritte Beitrag (Beitrag 3) fokussiert die unterrichtliche Nutzung digitaler Lernwerkzeuge und Prädiktoren der Nutzung digitaler Medien durch Lehrkräfte in Deutschland im internationalen Vergleich, während der vierte Beitrag (Beitrag 4) die Perspektive der Schulleitungen auf digitalisierungsbezogene Transformationsprozesse in den Blick nimmt. Für Deutschland werden neben vertiefenden Zusammenhangsanalysen u.a. schulförmenspezifische Analysen präsentiert.

Beiträge des Symposiums

Die Studie ICILS 2023 – computer- und informationsbezogene Kompetenzen und Kompetenzen im Bereich ‚Computational Thinking‘ von Schüler*innen der 8. Jahrgangsstufe in Deutschland im dritten bzw. zweiten internationalen Vergleich

Birgit Eickelmann¹, Jan Vahrenhold², Frank Goldhammer³, Nadine Fröhlich¹

¹Universität Paderborn, ²Universität Münster, ³DIPF

Die fortwährenden digitalisierungsbezogenen Transformationsprozesse erfordern einen kompetenten Umgang mit digitalen Medien, um selbstbestimmt, reflektiert und verantwortlich am gesellschaftlichen, politischen und beruflichen Leben teilhaben zu können (Fraillon et al., 2023; Vuorikari et al., 2022). In diesem Kontext stellen die digitalen Kompetenzen eine fächerübergreifende Schlüsselkompetenz für alle Schüler*innen dar (Council of the European Union, 2021; Fraillon et al., 2023; KMK, 2021). Aufgabe von Schule ist es, den Erwerb ebendieser Kompetenzen zu fördern (KMK, 2016).

Mit den ICIL-Studien werden computer- und informationsbezogene Kompetenzen (CIL) von Achtklässler*innen zum dritten Mal nach ICILS 2018 und ICILS 2013 und Kompetenzen im Bereich ‚Computational Thinking‘ (CT) zum zweiten Mal nach ICILS 2018 international vergleichend erfasst. Dabei zeigen bisherige empirische Befunde basierend auf den vorherigen ICILS-Zyklen ein Abschnitten Deutschlands für CIL im internationalen Mittelfeld (Bos, Eickelmann & Gerick, 2014; Eickelmann, Bos, Gerick & Labusch, 2019). Bei den Kompetenzen im Bereich CT liegt Deutschland in ICILS 2018 hingegen unter dem internationalen Mittelwert (Eickelmann, Vahrenhold & Labusch, 2019), wobei 2018 - im Gegensatz zu 2023 - nur eine kleine Anzahl von Staaten an dem Zusatzmodul teilgenommen hat. Auffällig waren durchgehend die großen Leistungsunterschiede zwischen Gymnasien und anderen Schulformen der Sekundarstufe I in Deutschland sowie der hohe Anteil der Schüler*innen auf den unteren beiden Kompetenzstufen (2013: 29.2% und 2018: 33.2%). An diese Ergebnisse anknüpfend sowie der Nutzung der erstmaligen Möglichkeit, Entwicklungen abzubilden, stellt der vorliegende Beitrag als Auftaktbeitrag für das Symposium zunächst die Anlage der Studie ICILS 2023 dar und beantwortet daran anknüpfend folgende Forschungsfragen:

1. Über welche computer- und informationsbezogenen Kompetenzen verfügen die Achtklässler*innen in Deutschland im internationalen Vergleich und wie verteilen sich deren Kompetenzen auf die unterschiedlichen Kompetenzstufen in ICILS 2023?
2. Welche Entwicklungen lassen sich seit ICILS 2013 in Bezug auf die computer- und informationsbezogenen Kompetenzen für Deutschland aufzeigen?

3. Welche Unterschiede in den computer- und informationsbezogenen Kompetenzen zeigen sich im Vergleich zwischen Achtklässler*innen in Deutschland, die ein Gymnasium besuchen und Achtklässler*innen anderer Schulformen?
4. Über welche Kompetenzen im Bereich CT verfügen die Schüler*innen in Deutschland im internationalen Vergleich und im Vergleich zu ICILS 2018?
5. Welche Unterschiede zeigen sich im Vergleich zwischen den Kompetenzen der Achtklässler*innen im Bereich CT in Deutschland, die ein Gymnasium besuchen und Achtklässler*innen anderer Schulformen?

Um die Forschungsfragen zu beantworten, werden die über computerbasierte Tests in einer Live-Softwareumgebung erhobenen Kompetenzdaten von Achtklässler*innen in repräsentativen Stichproben (CIL: Ninternational = 122.763; Nnational = 5065; CT: Ninternational = 80.570; Nnational = 5065) in allen an der Studie teilnehmenden Länder der IEA-Studie ICILS 2023 herangezogen und deskriptiv sowie interferenzstatistisch ausgewertet. Durch die Besonderheiten der Stichprobenziehung in Deutschland sind darüber hinaus Vergleiche zwischen den Kompetenzständen von Schüler*innen unterschiedlicher Schulformen möglich (Eickelmann et al., 2024, in Druck). Diese Ergebnisse geben mögliche Diskussionsimpulse für die Gestaltung von Bildung in einer Kultur der Digitalität und werden interdisziplinär diskutiert.

Nach der ICILS-2023-Veröffentlichungssperre stehen die zentralen Ergebnisse der Studie zur Verfügung, die im Rahmen des Symposiums vertieft und ergänzt werden (vgl. Hinweis Mantelabstract).

Die Perspektive der Schüler*innen auf das Lernen mit digitalen Medien und den Erwerb computer- und informationsbezogener Kompetenzen in Deutschland im internationalen Vergleich

Heike Schaumburg¹, Nadine Fröhlich², Martin Senkbeil³, Birgit Eickelmann²

¹Humboldt-Universität zu Berlin, ²Universität Paderborn, ³IPN

In diesem Beitrag zum ICILS-2023-Symposium wird die Perspektive von Schüler*innen auf das Lernen mit digitalen Medien und den Erwerb von computer- und informationsbezogenen Kompetenzen in Deutschland im internationalen Vergleich fokussiert. Der Einbezug der Perspektive der Schüler*innen ist dabei von besonderer Relevanz und bildet den Kern der Studie (Eickelmann et al., 2024, in Druck). Entsprechend theoretischer Modelle hängt der Erwerb vorgenannter Kompetenzen insbesondere von Voraussetzungs- und Prozessmerkmalen ab (Eickelmann & Drossel, 2019; Fraillon et al., 2023). Diesbezügliche empirische Befunde auf Basis der IEA-Studie ICILS 2018 zeigten, dass schulische Bildungsprozesse in Deutschland hier ein hohes Entwicklungspotenzial aufwiesen: So erwies sich mit Blick auf die Voraussetzungsebene, dass beispielsweise in keinem ICILS-2018-Teilnehmerland die Anteile der Schüler*innen, die angaben, seit mehr als fünf Jahren Erfahrungen mit digitalen Medien zu haben, geringer als in Deutschland ausfielen. Hingegen zeigte sich, dass Schüler*innen in Deutschland ähnlich zu den Anteilen im internationalen Mittel positive Einstellungen gegenüber digitalen Medien und deren gesellschaftliche Relevanz aufwiesen (Gerick et al., 2019). Darüber hinaus zeigte sich bezogen auf die Prozessebene, dass beispielsweise weniger als ein Viertel (22.8%) der Achtklässler*innen in Deutschland nach eigenen Angaben mindestens einmal in der Woche digitale Medien in der Schule für schulbezogene Zwecke nutzten. Dieser Anteil lag signifikant unter dem internationalen Mittelwert (44.0%) sowie dem Wert der Vergleichsgruppe EU (45.1%). Vertiefende Analysen haben zudem ergeben, dass anders als in den meisten ICILS-2018-Teilnehmerländern die schulische Nutzungshäufigkeit digitaler Medien für schulbezogene Zwecke in Deutschland in einem negativen Zusammenhang mit den computer- und informationsbezogenen Kompetenzen der Schüler*innen stand (Schaumburg et al., 2019).

Der vorliegende Beitrag fokussiert nun die Ergebnisse und Analysen der ICILS-2023-Studie und damit einerseits den Status quo in Bezug auf das Lernen mit digitalen Medien und den Erwerb computer- und informationsbezogener Kompetenzen aus Perspektive der Schüler*innen in Deutschland in ICILS 2023 und im internationalen Vergleich. Somit werden folgende Fragestellungen bearbeitet:

1. Wie stellt sich die aktuelle Situation in Deutschland und im internationalen Vergleich zu ausgewählten Merkmalen auf der Voraussetzungs- und Prozessebene im Kontext des Erwerbs computer- und informationsbezogener Kompetenzen aus der Perspektive der Schüler*innen in ICILS 2023 dar?
2. Welche Prädiktoren des Erwerbs computer- und informationsbezogener Kompetenzen durch Schüler*innen lassen sich in Deutschland und im internationalen Vergleich identifizieren?

Die Fragestellungen werden auf Basis der in ICILS 2023 in Deutschland und international erhobenen Daten der Schüler*innenbefragung sowie den durchgeführten computerbasierten Kompetenztests beantwortet (Eickelmann et al., 2024, in Druck).

Während Frage 1 zunächst international vergleichend mittels deskriptiver Analysen betrachtet wird, wird der Zusammenhang zwischen den Prädiktoren des Erwerbs computer- und informationsbezogener Kompetenzen und den computer- und informationsbezogenen Kompetenzen der Schüler*innen mit einem regressionsanalytischen Ansatz (Schendera, 2014) untersucht. Die Ergebnisse werden vor dem Hintergrund des empirischen Forschungsstandes und im Hinblick auf Entwicklungen über die Jahre sowie ihre Bedeutung für eine interdisziplinär angelegte empirische Bildungsforschung diskutiert.

Nach der ICILS-2023-Veröffentlichungssperre stehen die zentralen Ergebnisse der Studie zur Verfügung, die im Rahmen des Symposiums vertieft und ergänzt werden (vgl. Hinweis Mantelabstract).

Die Nutzung digitaler Lernwerkzeuge aus der Perspektive der Lehrkräfte und Prädiktoren der Nutzung in Deutschland im internationalen Vergleich

Kerstin Drossel¹, Julia Gerick², Jan Niemann¹, Birgit Eickelmann¹

¹Universität Paderborn, ²TU Braunschweig

Lehrkräfte und ihre professionellen Kompetenzen haben eine hohe Relevanz für das Lehren und Lernen mit und über digitale(n) Medien und nehmen bei der Förderung des Erwerbs digitalisierungsbezogener Kompetenzen der Schüler*innen eine Schlüsselrolle ein (Hoffmann, Richter & Weirich, 2023; KMK, 2021; Scheiter, 2021). Grundvoraussetzung für den schulischen Erwerb digitalisierungsbezogener Kompetenzen ist dabei entsprechend theoretischer Modelle die Nutzung digitaler Medien im Unterricht, die wiederum von Prädiktoren auf der Voraussetzungs- und Prozessebene abhängt (Eickelmann & Drossel, 2019; Fraillon et al., 2023).

Ein mit besonderer Aufmerksamkeit wahrgenommenes Ergebnis der Studien ICILS 2013 und ICILS 2018 war der Befund, dass der Anteil der Lehrpersonen, der regelmäßig digitale Medien im Unterricht einsetzte, in Deutschland teilweise deutlich geringer ausfiel als in den meisten anderen Teilnehmerländern (Drossel et al., 2019; Eickelmann et al., 2014). Dabei wurde, wie auch in anderen Studien, darauf hingewiesen, dass die Quantität der unterrichtlichen Nutzung digitaler Medien durch Lehrpersonen allein

keine Rückschlüsse auf die Qualität ihrer Nutzung erlaubt, die Nutzung allerdings als eine notwendige Voraussetzung für die Realisierung der viel beschriebenen Potenziale digitaler Medien für das Lernen und Lehren, auch in fachdidaktischen Kontexten, anzusehen ist. In dem Zusammenhang wurde in den bisherigen ICIL-Studien durch internationale Vergleiche ebenfalls deutlich, dass in Deutschland die Rahmenbedingungen schulischen Lehrens und damit die Voraussetzungen für Lehrpersonen, überhaupt digitale Medien in den Unterricht integrieren zu können, deutliche Entwicklungsbedarfe aufwiesen (Gerick et al., 2014, Drossel et al., 2019). Dies bezog sich u.a. auf lehrkräftebezogene Voraussetzungen, wie digitalisierungsbezogene Kompetenzen sowie auf lehrkräftebezogene Prozesse, wie die digitalisierungsbezogene Lehrkräftekooperation (Eickelmann et al., 2019; Gerick & Eickelmann, 2020; Gerick et al., 2019). Damit waren zentrale Voraussetzungen für Lehrkräfte, den pädagogischen Handlungsspielraum des Lehrens mit digitalen Medien gestalten zu können, in Deutschland nicht in der Fläche gegeben. Aktuelle empirische Daten für Deutschland, insbesondere im internationalen Vergleich, liegen nun mit ICILS 2023 vor und werden in diesem Beitrag mit folgenden Forschungsfragen adressiert:

Wie stellt sich die unterrichtliche Nutzung digitaler Lernwerkzeuge in Deutschland und im internationalen Vergleich aus der Perspektive der Lehrkräfte in ICILS 2023 dar?

Welche Prädiktoren der unterrichtlichen Nutzung digitaler Lernwerkzeuge durch Lehrkräfte lassen sich in Deutschland und im internationalen Vergleich identifizieren?

Die Forschungsfragen werden mittels der in ICILS 2023 international erhobenen Daten der Fragebogenbefragung der Lehrkräfte (n = 62.294), die in der achten Jahrgangsstufe unterrichten, beantwortet. Für die erste Forschungsfrage werden deskriptive Statistiken genutzt, während für die zweite Forschungsfrage mittels Regressionsanalysen (Schendera, 2014) die Zusammenhänge zwischen der Häufigkeit der Nutzung digitaler Lernwerkzeuge (abhängige Variable) und verschiedenen lehrkräftebezogenen Voraussetzungen (u.a. Lehrkräfteprofessionalisierung, selbsteingeschätzte digitalisierungsbezogene Kompetenzen der Lehrkräfte) sowie lehrkräftebezogenen Prozessen (u.a. digitalisierungsbezogene Lehrkräftekooperation) entlang des ICILS-2023-Rahmenmodells untersucht werden.

Nach der ICILS-2023-Veröffentlichungssperre stehen die zentralen Ergebnisse der Studie zur Verfügung, die im Rahmen des Symposiums vertieft und ergänzt werden (vgl. Hinweis Mantelabstract).

Die Ergebnisse werden vor dem Hintergrund des Forschungsstandes und der Entwicklungen seit der ersten ICILS-Studie sowie hinsichtlich weiterführender Implikationen und Entwicklungsbedarfe für die Gestaltung von Professionalisierungsprozessen im schulischen Bildungsbereich diskutiert.

Die Perspektive von Schulleitungen auf digitalisierungsbezogene schulische Transformationsprozesse in Deutschland

Julia Gerick¹, Birgit Eickelmann², Knut Schwippert³, Jan Niemann²

¹TU Braunschweig, ²Universität Paderborn, ³Universität Hamburg

Schulleitungen sind zentrale Akteur*innen bei der Initiierung und Gestaltung digitalisierungsbezogener Transformationsprozesse in der Schule, die angesichts gesellschaftlicher Entwicklungen und der Notwendigkeit des Erwerbs computer- und informationsbezogener Kompetenzen durch Schüler*innen zunehmend an Relevanz gewinnen (Dexter, 2018; Ruloff & Petko, 2021). Eine solche mit den Transformationsprozessen einhergehende digitalisierungsbezogene Schulentwicklung lässt sich entsprechend des theoretischen Modells von Eickelmann und Gerick (2017) in fünf Dimensionen ausdifferenzieren: (1) Organisationsentwicklung, (2) Personalentwicklung, (3) Unterrichtsentwicklung, (4) Kooperationsentwicklung und (5) Technologieentwicklung.

Forschungsbefunde deuten auf die Relevanz der Schulleitungen für die Initiierung und Gestaltung der Dimensionen zur digitalisierungsbezogenen Schulentwicklung hin. In Bezug auf die (1) Organisationsentwicklung können Schulleitungen Rahmenbedingungen herstellen, um Lehren und Lernen in einer Kultur der Digitalität zu fördern. So können beispielsweise Konzepte entwickelt oder zeitliche Ressourcen bereitgestellt werden (Tondeur, Coopert & Newhouse, 2010). Im Kontext der (2) Personalentwicklung können Schulleitungen durch die Unterstützung der Lehrkräfteprofessionalisierung Maßnahmen initiieren, um digitalisierungsbezogene Transformationsprozesse an der eigenen Schule voranzubringen (Kafyulilo, Fisser & Voogt, 2016). Der Bereich der (3) Unterrichtsentwicklung bezieht sich auf die Unterstützung der Entwicklung digitaler Lehr-Lernformate (Gerick, Tulowitzki & Oelkers, 2023). Mit der Förderung der (4) Kooperationsentwicklung kann ebenfalls der Einsatz digitaler Medien im Unterricht forciert werden, indem Schulleitungen Kooperationsmöglichkeiten sowohl innerschulisch als auch außerschulisch eröffnen (Tołwińska, 2021). Im Kontext der (5) Technologieentwicklung haben Schulleitungen u.a. die Aufgabe eine adäquate IT-Ausstattung und technischen Support bereitzustellen, was sich auf die Integration digitaler Medien in das Lehren und Lernen auswirken kann (Gerick, Tulowitzki & Oelkers, 2023).

Das Zusammenwirken der Dimensionen digitalisierungsbezogener Schulentwicklung bestimmt maßgeblich darüber, wie Transformationsprozesse gestaltet werden und damit, in welchem Umfang es Schüler*innen ermöglicht wird, computer- und informationsbezogene Kompetenzen zu erwerben. Dabei fehlen jedoch aktuelle empirische Befunde zum Status quo der Dimensionen, um aktuelle Entwicklungen mit Blick auf die digitale Transformation in Schule zu beschreiben. Zudem ist der Zusammenhang zwischen den Dimensionen aus Perspektive der Schulleitungen und den computer- und informationsbezogenen Kompetenzen von Schüler*innen in Deutschland unerforscht, sodass dieser Beitrag folgenden Fragestellungen nachgeht:

1. Wie nehmen Schulleitungen in Deutschland ihr Schulleitungshandeln in den fünf Dimensionen der digitalisierungsbezogenen Schulentwicklung im Jahr 2023 wahr und inwiefern hat sich die diesbezügliche Perspektive seit 2018 verändert?
2. In welchem Zusammenhang steht das Schulleitungshandeln im Kontext der fünf Dimensionen der digitalisierungsbezogenen Schulentwicklung mit den computer- und informationsbezogenen Kompetenzen von Schüler*innen der achten Jahrgangsstufe in Deutschland?

Beide Forschungsfragen werden auf Basis der repräsentativ gezogenen und gewichteten Stichprobe der ICILS-2023-Daten von über 5000 Schüler*innen und 200 Schulleitungen in Deutschland bearbeitet (Eickelmann et al., 2024, in Druck). Forschungsfrage 1 wird auf dieser Datengrundlage und unter Hinzunahme der ICILS-2018-Daten mittels deskriptiver Analysen und t-Tests basierend auf der Schulleitungsbefragung und gewichtet auf die Schüler*innen beantwortet (Frailon & Rožman, 2023). Darüber hinaus wird im Kontext der Forschungsfrage 2 mittels latenten Regressionsmodellen (Schendera, 2014) das Zusammenwirken der fünf Dimensionen mit den über Kompetenztests erfassten computer- und informationsbezogenen Schüler*innenkompetenzen betrachtet.

Nach der ICILS-2023-Veröffentlichungssperre stehen die zentralen Ergebnisse der Studie zur Verfügung, die im Rahmen des Symposiums vertieft und ergänzt werden (vgl. Hinweis Mantelabstract).

Die Ergebnisse werden vor dem Hintergrund des Forschungsstandes und in Bezug auf Implikationen für schulische digitalisierungsbezogene Transformationsprozesse und die Rolle von Schulleitungen diskutiert.

OF-1: Workshop: Suchst du noch oder findest du schon? Strategien für komplexe Literaturrecherchen in der Bildungsforschung

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · *Ort:* EO 159

Chair der Sitzung: Ingeborg Jäger-Dengler-Harles

Die Suche nach Fachliteratur stellt für Wissenschaftler*innen in der Bildungsforschung trotz vielversprechend erscheinender Ansätze in der KI immer noch eine Herausforderung dar. Spätestens bei umfangreichen Literaturrecherchen für Forschungssynthesen und Meta-Analysen sind tiefergehende Kenntnisse über die Inhalte und Suchoptionen von Online-Datenbanken, die Mechanismen der Suchmaschinen und über die Vielzahl an weiteren Informationsquellen (für eine sinnvolle Recherchestrategie) notwendig. Dieser Workshop zeigt, worauf es bei einer zielgerichteten Literatursuche in der Bildungsforschung ankommt und warum es sich in diesem transdisziplinären Forschungsfeld lohnt, einen Blick auf die inhaltliche Abdeckung der Datenbanken und die Informationsangebote auf den unterschiedlichen Plattformen zu werfen. Suchwege und Suchtechniken werden vorgestellt und diskutiert. Die Teilnehmenden können eigene Erfahrungen und Fragen beisteuern. Anhand der neu entwickelten Profisuche im Fachportal Pädagogik wird gezeigt, wie eine komplexe Suche Schritt für Schritt aufgebaut werden kann, ohne dass der Überblick dabei verloren geht.

Suchst du noch oder findest du schon? Strategien für komplexe Literaturrecherchen in der Bildungsforschung

Ingeborg Jäger-Dengler-Harles

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Die Suche nach Fachliteratur stellt für Wissenschaftler*innen in der Bildungsforschung trotz vielversprechend erscheinender Ansätze in der KI immer noch eine Herausforderung dar. Spätestens bei umfangreichen Literaturrecherchen für Forschungssynthesen und Meta-Analysen sind tiefergehende Kenntnisse über die Inhalte und Suchoptionen von Online-Datenbanken, die Mechanismen der Suchmaschinen und über die Vielzahl an weiteren Informationsquellen (für eine sinnvolle Recherchestrategie) notwendig. Dieser Workshop zeigt, worauf es bei einer zielgerichteten Literatursuche in der Bildungsforschung ankommt und warum es sich in diesem transdisziplinären Forschungsfeld lohnt, einen Blick auf die inhaltliche Abdeckung der Datenbanken und die Informationsangebote auf den unterschiedlichen Plattformen zu werfen. Suchwege und Suchtechniken werden vorgestellt und diskutiert. Die Teilnehmenden können eigene Erfahrungen und Fragen beisteuern. Anhand der neu entwickelten Profisuche im Fachportal Pädagogik wird gezeigt, wie eine komplexe Suche Schritt für Schritt aufgebaut werden kann, ohne dass der Überblick dabei verloren geht.

PS-01: Qualität und Unterstützung im Lernprozess: Perspektiven auf Unterrichtsinteraktion, Kontext und Lernerfolg

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 169

Chair der Sitzung: Alexander Jonas Jung

Effects of Student Groups on Teachers' and Students' Perceptions of Teaching Quality

Alexander Jonas Jung, Ann-Kathrin Jaekel, Richard Göllner, Kou Murayama

Universität Tübingen, Deutschland

Theoretical Background

The ability of teachers to effectively impart knowledge and educational values to students is crucial for our society's progress and success (OECD, 2014). Various frameworks have been developed to assess the quality of teachers and their teaching (e.g., Creemers & Kyriakides, 2007; Ferguson, 2012; Klieme et al., 2009; Pianta et al., 2008). These frameworks describe generic characteristics of effective teaching desirable across all subjects and age groups and are largely similar at their core.

The framework of Three Basic Dimensions (TBD; Klieme et al., 2001, 2009; Praetorius et al., 2018) with its main dimensions classroom management, support, and cognitive activation, is especially popular in Germany. While most indicators for teaching quality in the TBD describe teacher behaviors also student behaviors are relevant: High levels of student engagement, low levels of disturbances, and a positive classroom climate indicate high teaching quality (Fauth et al., 2021; Praetorius et al., 2018). It is not clear to which degree classroom characteristics independent of the teacher influence teaching quality. Initial findings suggest effects of student characteristics and classroom composition (Fauth et al., 2020; Nurmi & Kiuru, 2015).

Research Question

Does teaching quality, measured via teacher and student ratings across the three dimensions classroom management, support, and cognitive activation, vary depending on the taught student group?

Method

To address this question, we employed a quasi-experimental longitudinal design. Teachers either taught the same class (control condition) or two different classes (experimental condition) at two time points one year apart. 146 teachers German and mathematics teachers from various tracks in the German federal state of Baden-Württemberg provided ratings of eight facets of teaching quality.

We assessed whether teacher and student rated teaching quality ratings for teachers that taught the same class at both measurement points were more similar, in terms of absolute differences, than those for teachers that taught two different classes at both measurement points. Statistical analyses were conducted using mixed-effects modeling to account for the clustering of data within teachers and classes.

Results and Significance

For seven out of eight teacher-rated scales, ratings were descriptively more similar in the same-class condition than in the different-class condition. Only the scale for disturbances (a facet of classroom management) remained significant after Bonferroni-Holm correction for multiple testing. For all eight student-rated scales, ratings were descriptively more similar in the same-class condition than in the different-class condition, with generally larger effects than for teacher-ratings. Significant effects after Bonferroni-Holm correction were found for disturbances, monitoring, and rule clarity (classroom management), motivating students and feedback (support), and challenging tasks (cognitive activation).

These findings indicate that teaching quality ratings are impacted by student groups. It may be that some classes are generally more teachable than others, or that some teachers may be better suited to teach certain classes (Fauth et al., 2020). Therefore, teaching quality should not be attributed solely to the teacher but should be interpreted as an indicator of the interactions between students and teachers' behaviors.

This finding is relevant for interpreting correlations between teaching quality and student outcomes, which should not be interpreted as causal effects of teaching quality on motivation and achievement. For practitioners, it suggests that classroom composition needs careful consideration to ensure educational justice. Creating classrooms that are disbalanced according to some criteria, e.g., gender, SES, or motivation, may affect the quality of teaching that students in such classrooms experience. We call for more research to explore specific characteristics of student groups that may positively or negatively impact teaching quality.

Die Bedeutung des Vorwissens für die gemeinsame Steuerung von Lernprozessen

Svenja Rieser¹, Katja Adl-Amini²

¹Technische Universität Dortmund, Deutschland; ²Technische Universität Darmstadt, Deutschland

Theoretischer Hintergrund

Kooperative Lernformen sind in der Grundschule weit verbreitet und besonders effektiv (Kyndt et al., 2013). Schüler*innen steuern dabei ihren gemeinsamen Lernprozess, wobei die Qualität der Steuerung zum Lernerfolg beiträgt (Hadwin et al., 2018; Panadero & Järvelä, 2015). Nach der Forschung zum selbstregulierten Lernen (Pintrich, 2004) werden dabei die Steuerungsaktivitäten in Planung, Überwachung, Regulation und Evaluation unterteilt, die sich auf inhaltliche, organisatorische oder motivationale Aspekte beziehen können (Grau & Whitebread, 2012; Iskala et al., 2011; Larkin, 2009).

Es wird vermutet, dass die gemeinsame Steuerung von individuellen Schülermerkmalen beeinflusst wird, aber entsprechende Forschung fehlt bislang (Panadero & Järvelä, 2015). Aus der Forschung zum selbstregulierten Lernen könnte das Vorwissen der Lernenden als ein positiver Einflussfaktor abgeleitet werden (z.B. Whitherby & Carpenter, 2021), allerdings bringen kooperativ Lernende oft unterschiedliches Vorwissen mit (z.B. bei leistungsheterogener Gruppenzusammensetzung). Daher stellt sich die Frage, inwiefern sich die Steuerung je nach Zusammensetzung der Lerngruppe unterscheidet.

Die vorliegende Arbeit befasst sich daher mit folgenden Fragestellungen:

1. Welche Bedeutung hat das Vorwissen für die gemeinsame Steuerung von Lernprozessen?
2. Besteht ein Zusammenhang zwischen Unterschieden im Vorwissen und der gemeinsamen Steuerung?

Methode

Die Datenbasis der Studie umfasst zehn dritte Klassen, die an einer Peer-Tutoring-Intervention teilnahmen. In jeder Klasse wurde eine Doppelstunde des naturwissenschaftlichen Sachunterrichts videographiert und die Interaktionen von zwei Teams transkribiert. Insgesamt wurden Transkripte von 40 Schüler*innen in 20 Teams erstellt. Das Durchschnittsalter der Kinder betrug 8,70 Jahre ($SD = 0,31$), und 50 % von ihnen waren weiblich.

Zur Erfassung des Vorwissens wurde die naturwissenschaftliche Kompetenz der Schüler*innen vor Beginn der Unterrichtseinheit mit einem Test gemessen (EAP-Reliabilität = XX). Die gemeinsame Steuerung wurde anhand eines Kodierschemas erfasst, das Planung, Überwachung, Regulation und Evaluation in Bezug auf Inhalt, Organisation und Motivation umfasste (adaptiert nach Grau & Whitebread, 2012; $\kappa \geq .83$).

Aufgrund der inhaltlichen Abhängigkeit innerhalb der Teams wurden die Werte für alle Auswertungen auf Teamebene aggregiert. Zur Beantwortung beider Fragestellungen berechnen wir zunächst bivariate Korrelationen zwischen der kodierten gemeinsamen Steuerung und (1) dem mittleren Vorwissen der Lernenden bzw. (2) der Differenz im Vorwissen der Teampartner.

Ergebnisse und ihre Bedeutung:

Insgesamt deuten die Ergebnisse darauf hin, dass das Vorwissen sowie die diesbezügliche Zusammensetzung der Teams einen Einfluss auf die gemeinsame Steuerung des Lernprozesses haben. In Teams mit einem höheren durchschnittlichen Vorwissen fand eine intensivere Steuerung des inhaltlichen Lernprozesses statt ($r = .57$, $p = .01$), während organisatorische Aspekte weniger intensiv thematisiert wurden ($r = -.67$, $p = .00$). Den Gruppen scheint also eine stärkere Fokussierung auf den fachlichen Lernprozess gelungen zu sein.

Je höher die Differenz im Vorwissen der Teams, desto seltener wurden Planungsaktivitäten beobachtet ($r = -.40$, $p = .09$) und desto ausgeprägter war die Überwachung ($r = .55$, $p = .02$). Eine mögliche Erklärung für dieses Ergebnis könnte darin bestehen, dass der leistungsstärkere Lernende die Führung übernahm und so den schwächeren Partner kognitiv entlastete (vgl. auch Pino-Pasternak et al., 2017). Somit wäre weniger gemeinsame Planung notwendig, aber dafür eine kleinschrittigere gemeinsame Überwachung des Lernprozesses.

Um diese Unterschiede vertieft zu ergründen, laufen derzeit rekonstruktive Auswertungen von Transkripten kontrastiver Fälle mithilfe der Grounded Theory Methodologie (Strauss, 1998) orientiert am Konzept der Ko- und Socially Shared Regulation (Hadwin et al., 2018). Erste Ergebnisse werden im Vortrag berichtet.

ENTFÄLLT Qualität von Lernunterstützung in Schülerarbeitsphasen: Analyse und Einflussfaktoren

Eva Prinz

PH Karlsruhe, Deutschland

Theoretischer Hintergrund

Lernunterstützung als „eine Form der Individualisierung“ (Krammer 2009, 18) ist Teil der unterrichtlichen Angebotsseite (Vieluf et al., 2020), für deren Umsetzung vor allem Schülerarbeitsphasen Gelegenheit bieten. Im Mathematikunterricht der Sekundarstufe nicht-gymnasialer Schularten nehmen Schülerarbeitsphasen einen großen Anteil der Unterrichtszeit ein (Kleinknecht, 2010; Seidel, 2011). Dies verweist auf deren Potenzial für eine hochwertige individuelle Unterstützung von Schüler:innen gerade in heterogenen Lerngruppen, wobei insbesondere die Unterstützungsqualität von Bedeutung ist (Lemmrich, Ehmke & Reusser, 2024). Die kurzfristige Lernunterstützung in Lehrer-Schüler-Interaktionen bezieht sich vorwiegend auf konkrete fachspezifische Lernaufgaben und ist von einer längerfristigen Lernbegleitung und -beratung, die stärker fachübergreifend die individuelle Lernentwicklung und den Bildungsweg der Schüler:innen adressiert, zu unterscheiden (Schnebel, 2013; Pauli et al. 2018). Anknüpfungen lassen sich deshalb zum Konzept des ‚Feedbacks‘ finden, welches ebenso ‚on-the-fly‘ während des Unterrichts stattfinden kann und dessen Lernwirksamkeit gut belegt ist (z.B. Hattie & Timperley, 2007; Rakoczy et al., 2019). Wiederholt konnte ebenso gezeigt werden, dass ein Unterricht, in dem die Schüler:innen von der Lehrperson individuell oder kleingruppenbezogen unterstützt werden, einem Unterricht ohne Unterstützung bezüglich der Lernleistung überlegen ist (z.B. Lazonder & Harmsen 2016). Bisherige quantitative Beobachtungsstudien zielen hauptsächlich auf die Beschreibung der Lehreraktivitäten in Schülerarbeitsphasen und berücksichtigen aggregierte Schülerdaten (z.B. Bohl et al., 2012; Kobarg & Seidel, 2007; Krammer, 2009). Zusammenhänge mit proximalen Kontextvariablen, wozu die Gestaltung der Schülerarbeitsphase sowie Merkmale der Lehrenden und Lernenden zählen, wurden bislang nicht umfassend untersucht. Die vorliegende Studie nimmt deshalb eine konsequente Einbettung der Lernunterstützung in eine Angebots-Nutzungs-Modellierung von Unterricht (zsf. Vieluf et al., 2020) unter Berücksichtigung der Unterscheidung von Oberflächen- und Tiefenmerkmalen (Decristan et al., 2020) vor und geht explorativ diesen drei Forschungsfragen nach:

1. Inwiefern hängt die Qualität der Lernunterstützung mit Oberflächen- und Tiefenmerkmalen der Schülerarbeitsphase zusammen?
2. Inwiefern ist die Unterstützungsqualität mit den differenziellen Voraussetzungen der Schüler:innen assoziiert?
3. Inwiefern präzisieren Merkmale der Lehrperson die Qualität der Lernunterstützung?

Methode

In einer quasi-experimentellen Feldstudie (Döring & Bortz, 2016) wurden 60 Unterrichtsstunden bei 30 Lehrpersonen in achten Mathematikklassen an Werkreal-/Haupt- und Realschulen in Baden-Württemberg videografiert. An den Einführungs- und Übungsstunden zu dem curricular verankerten Thema ‚Vermehrter/Verminderter Grundwert‘ nahmen $N = 608$ Schüler:innen teil. Die Erfassung der Lernunterstützung und der rahmenden Schülerarbeitsphase erfolgte mit videoanalytischen Beobachtungsverfahren unterschiedlicher Inferenzniveaus (z.B. Seidel & Thiel, 2017), wobei die Einschätzung der Qualität der Lernunterstützung mittels eines mittel-inferenten Kategoriensystems von Krammer (2009) erfolgte. Fragebögen und Leistungstest wurden für die Erhebung von Schüler- und Lehrervariablen eingesetzt, womit die Studie mehrperspektivisch angelegt ist (Clausen, 2002). Die multivariaten regressionsanalytischen Schätzungen wurden auf der Ebene der einzelnen Interaktion sowie auf individueller Schüler:innenebene durchgeführt.

Ergebnisse

Die Qualität der Lernunterstützung als Tiefenmerkmal variiert in Abhängigkeit von der Gestaltung der Schülerarbeitsphase als Unterrichtsoberfläche, z.B. hinsichtlich der gewählten Sozialform oder der Funktion der Schülerarbeitsphase. Sozialformen mit Kooperationsmöglichkeit z.B. haben einen positiven Effekt auf eine stärker kognitiv aktivierende und schülerorientierte Lernunterstützung seitens der Lehrperson ($OR = 2.06$; $p = .002$) (Forschungsfrage 1). Auch unterscheidet sich die Qualität der Lernunterstützung innerhalb einer identischen Oberfläche für Schüler:innen in Abhängigkeit ihrer individuellen Merkmale (z.B. Vorwissen, Geschlecht). In Klassen mit höherem mittlerem Leistungsniveau z.B. fällt die Unterstützung der Lehrpersonen kognitiv

aktivierender aus ($OR = 1.63$; $p = .033$) (Forschungsfrage 2). Seitens der Lehrperson variiert die Qualität der Lernunterstützung ebenso in Abhängigkeit u.a. ihrer lehr-lerntheoretischen Überzeugung sowie ihrer Lehrbefähigung für das Fach Mathematik. Fachfremd unterrichtende Lehrpersonen unterstützen ihre Schüler:innen z.B. stärker instruktional und kleinschrittig während Schülerarbeitsphasen ($OR = 1.84$; $p = .004$) (Forschungsfrage 3). Aus den Ergebnissen der Studie lassen sich Impulse für Entwicklungs- und Professionalisierungsaufgaben im didaktischen Handeln von Lehrpersonen ableiten.

Scaffolding & Kontextbasiertes Lernen: Auswirkungen auf Lernleistung und Interesse von Lernenden

Anna Liskes, Helena van Vorst

Universität Duisburg-Essen, Deutschland

Theoretischer

Ein Ansatz, um der Heterogenität in Schulklassen zu begegnen, ist die Binnendifferenzierung, bei der Lernende in kriteriengeleitet in Gruppen unterteilt werden und angepasste Lernangebote erhalten (Dumont, 2019). Grundsätzlich zeigt Binnendifferenzierung positive Effekte auf kognitive Lernergebnisse (Hattie, 2009). Die Gruppierung der Lernenden erfolgt dabei meist anhand kognitiver Fähigkeiten (Vock & Gronostaj, 2017). Eine Möglichkeit der Umsetzung von Binnendifferenzierung bietet das Scaffolding, das nach einer Meta-Analyse positive Effekte auf kognitive Faktoren, wie den Lernzuwachs, aufweist (Belland et al., 2017). Ein alternativer Ansatz ist eine interessenbasierte Binnendifferenzierung, die bisher weniger Beachtung findet. Im MINT-Unterricht kann dies durch kontextbasiertes Lernen erfolgen, das den Erwerb von Fachwissen mit alltäglichen Situationen verbindet (Parchmann et al., 2008). In einer interessenbasierten Binnendifferenzierung können Lernende zwischen verschiedenen Kontexten wählen (Güth & van Vorst, 2024). Da die Forschung zur Binnendifferenzierung sich bisher hauptsächlich auf leistungsorientierte Parameter konzentriert, gibt es kaum Erkenntnisse zur Effektivität einer interessenbasierten im Vergleich zu einer leistungsorientierten Binnendifferenzierung.

Hintergrund

Fragestellung

Aus dem bisherigen Forschungsstand ergibt sich folgende Forschungsfrage: Inwiefern werden affektive und kognitive Lernendenfaktoren im Chemieunterricht durch die Anwendung von interessenbasierter Binnendifferenzierung, leistungsorientierter Binnendifferenzierung und eine Kombination beider Differenzierungsansätze beeinflusst?

Methode

Zur Beantwortung der Forschungsfrage wurde eine quasi-experimentelle Interventionsstudie im 2x2-Design mit einer Prä-/Post-Datenerhebung durchgeführt. Dazu wurde digitales Lernmaterial für den Chemieunterricht der 10. Klasse entwickelt. Die Lernenden wurden klassenweise randomisiert auf vier Interventionsgruppen verteilt:

- Gruppe LI: leistungsorientierte (Scaffolding) und interessenbasierte Binnendifferenzierung (Wahl zwischen drei Kontexten)
- Gruppe L: leistungsorientierte Binnendifferenzierung (nur Scaffolding)
- Gruppe I: interessenbasierte Binnendifferenzierung (nur Kontextwahl)
- Kontrollgruppe: identisches Lernmaterial ohne Binnendifferenzierung.

Die Intervention fand im regulären Chemieunterricht über zwölf Unterrichtsstunden (je 45 Minuten) statt und wurde von einer angeleiteten Lehrperson umgesetzt.

Vor Beginn der Intervention wurden demografische Daten, kognitive und affektive Parameter mithilfe von Fragebögen und Tests erhoben. Der eingesetzte Fachwissenstest für organische Chemie als abhängige Variable zeigte im Prä-Test eine mäßige Reliabilität von $\alpha = .59$, verbesserte sich jedoch im Post-Test auf $\alpha = .81$. Als abhängige affektive Variablen wurden das Interesse der Lernenden vor und nach der Intervention (Reliabilitäten: $\alpha_{pre} = .90$, $\alpha_{post} = .93$) (Li et al., 2023) sowie das situationale Interesse (Reliabilitäten: $\alpha_{MZP1-3} = .91$) (Engeln, 2004) während der Intervention erhoben. Nach der Intervention wurde zudem die Zufriedenheit mit der Lernumgebung und deren Benutzerfreundlichkeit erfasst. Daten von $N = 703$ Lernenden aus 13 Gymnasien konnten in die Analyse einbezogen werden.

Ergebnisse

Die Kontrollvariablen zeigen keine signifikanten Unterschiede zwischen den Interventionsgruppen. Allerdings zeigt sich, dass Gruppe LI im Prä-Test signifikant bessere Leistungen erzielte als die anderen Gruppen, $F(3, 699) = 4.44$, $p = .004$, $\eta^2 = .02$. Auch beim Interesse gab es vor der Intervention signifikante Unterschiede zwischen den Gruppen, $F(3, 658) = 2.63$, $p = .049$, $\eta^2 = .01$. Diese Variablen wurden in den anschließenden Analysen als Kovariaten einbezogen. Die Ergebnisse zeigen, dass alle Lernenden ihre Leistung über die Zeit signifikant verbessern, $F(1, 699) = 436.63$, $p < .001$, $\eta^2 p = .38$. Nach dem Einbeziehen der Kovariaten gibt es signifikante Gruppenunterschiede im Post-Test, $F(3, 656) = 3.36$, $p = 0.02$, $\eta^2 = .02$. Das Interesse der Lernenden nimmt signifikant über die Zeit zu, $F(1, 458) = 12.73$, $p < .001$, $\eta^2 p = .03$, jedoch ohne signifikante Unterschiede zwischen den Gruppen. Ähnlich verhält es sich mit dem situationalen Interesse, das ebenfalls signifikant ansteigt, $F(2, 542) = 51.77$, $p < .001$, $\eta^2 p = .06$, jedoch ohne signifikante Gruppenunterschiede. Eine Vermutung ist, dass die hohe Attraktivität der gesamten Lernumgebung dazu geführt haben könnte, dass spezifische Effekte der einzelnen Interventionen nicht deutlich wurden. Ergebnisse vertiefter Datenanalysen sollen auf der Konferenz vorgestellt werden.

PS-02: Digitale Medien in der frühen Bildung: Chancen und Herausforderungen

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 256

Chair der Sitzung: Dave Möwisch

Analyse von Lehrperson-Kind-Interaktionen mit Fokus auf den Einsatz von Tablets als visuelle Lernunterstützung im frühen naturwissenschaftlichen Lernen

Priska Jossen, Ueli Studhalter, Annette Tettenborn, Marco Seeli

PH Luzern, Schweiz

Auf das Lernen bezogene Interaktionen zwischen Lehrpersonen und Kindern sind zentrale Bildungsmomente der frühen Bildung. Die Lehrperson-Kinder-Beziehung sowie die sozialen Interaktionen stellen entscheidende Determinanten für gelingende schulische Lehr-Lernprozesse (Raufelder & Wettstein, 2020) dar. Der stetig zunehmende Einsatz von Tablet-Computern im Unterricht von jungen Kindern (miniKIM, 2024) löst teils Befürchtungen aus, dass genau diese Interaktionen unterbunden oder eingeschränkt werden könnten. Der Einsatz von digitalen Medien wird andererseits seit längerem mit einer Steigerung der Motivation der Kinder in Verbindung gebracht, wobei die lernfördernde Wirkung von zahlreichen Faktoren abhängt (Neuss, 2017). In Studien fehlen jedoch oft detaillierte Hinweise, wie Tablets Lehr- und Lernprozesse in bestimmten Lernszenarien bereichern könnten (Haßler et al., 2016).

Es gibt Modelle und Empfehlungen für den effektiven und sinnstiftenden Einsatz digitaler Geräte in der frühkindlichen Bildung. Bailey und Blagojevic (2015) weisen darauf hin, dass die gezielte Nutzung der Kamerafunktion auf einem Tablet bestehende Lernaktivitäten bereichern und transformieren kann. Die Möglichkeit der Verzögerung und Fokussierung scheint für das konzeptuelle Lernen von naturwissenschaftlichen Inhalten bedeutsam, weil kognitive Umstrukturierungen viel Zeit und variable Lernanlässe erfordern (Carey, 2000).

Der Einsatz des Tablets ermöglicht, die physikalischen Vorgänge visuell sichtbar zu machen und als Erinnerungspunkte zu verwenden. Visuell dokumentierte Erinnerungspunkte können Lehrpersonen im Gespräch mit den Kindern im Sinne des «Reminiscing» nutzen, womit die gemeinsame sprachliche Verarbeitung vergangener Erlebnisse gemeint ist (Neale & Pino-Pasternak, 2017). Das gemeinsame Erinnern auf Basis der Tablet-Videos, ergänzt mit fachspezifischem «Scaffolding» (Wood et al. 1976, zit. in Eshach, 2011), soll das Lernen der Kinder anregen und unterstützen.

Forschungsfragen:

F1: Inwieweit spiegeln in ausgewählten Video-Sequenzen die Tablet-unterstützten Episoden Qualitätsindikatoren der Lehrperson-Kind-Interaktion wider?

F2: Welche Tablet-spezifischen Muster lassen sich aus den ausgewählten Episoden ableiten?

Zwei Kindergartenklassen (zwei Lehrpersonen, 41 Kinder) haben während vier Wochen an der Spiel- und Lernumgebung «Türme bauen» gearbeitet. Die Kinder durften die Stabilität ihrer Türme testen und sie schließlich gezielt zum Einsturz bringen. Diese Situationen filmte die Lehrperson mit einem Tablet und nutzte die entstandenen Videos direkt im Anschluss als Erinnerungspunkte im Gespräch, um die Phänomene mit den Kindern zu reflektieren.

Interaktionssequenzen wurden gemäss König (2009) definiert als Situationen, in denen die Lehrperson und ein oder mehrere Kinder durch gemeinsame Handlungen oder verbalen Austausch inhaltsspezifische Beziehungen eingehen. Für die qualitative Inhaltsanalyse (Kuchartz & Rädiker, 2022) wurden pro Lehrperson 8 Sequenzen mittlerer Dauer ausgewählt, in denen das Tablet eine aktive Rolle spielt und die zeitliche Dauer eine Vertiefung in ein Gespräch zulässt. Die Analyse wurde für die erste Frage deduktiv unter Verwendung des aus sechs Hauptkategorien bestehenden Kodierleitfadens durchgeführt. Für die zweite Frage wurden die Analysekategorien induktiv am Material entwickelt.

Ergebnisse und ihre Bedeutung

Insgesamt scheint das beobachtete Unterstützungsverhalten der Lehrpersonen bei Nutzung des Tablets im Hinblick auf kognitiv aktivierende Interaktionen vorteilhaft zu sein. Es zeigte sich ein reichhaltiges Scaffolding, Reminiscing und die häufige Verwendung inhaltspezifischer Sprache. Im Weiteren haben wir vier Muster der Tablet-Nutzung identifiziert, die sich auf oft genannte Potenziale des Einsatzes digitaler Geräte beziehen. So kann die beobachtete *Verzögerung* im Handlungsablauf im Sinne einer vertieften Wahrnehmung und Auseinandersetzung mit dem Phänomen in Bezug auf das konzeptuelle Lernen (Vosniadou, 1994) von Bedeutung sein. Das Medium scheint anhand der bewegten Bilder eine Sogwirkung (*Magnet-Effekt*) auf die Kinder auszulösen und vermag so die Aufmerksamkeit der Kinder auf natürliche Weise zu bündeln. Der Aspekt der *Wiederholung*, welcher sowohl durch das Medium wie auch durch das *Erinnern* erfolgt, scheint ebenfalls konzeptuelles Lernen zu begünstigen. Die Kinder bekundeten in diesem konkreten Lernszenario Interesse sowohl am inhaltlichen Lerngegenstand wie auch am Tablet selbst.

Die Studie zeigt, dass Tablets ein beträchtliches Potenzial haben, einen gemeinsamen Denk- und Lernprozess auf der Basis von Erinnerungspunkten zu initiieren und damit das Lernen zu unterstützen.

Digitale Medien in der frühen Bildung: Der Einfluss professioneller Kompetenzen von Fachkräften auf die Nutzung von ICT in Kindertageseinrichtungen

Franziska Cohen¹, Theresia Hummel², Sarah Schroeter², Yvonne Anders²

¹PH Freiburg, Deutschland; ²Otto-Friedrich-Universität Bamberg

Mit der rasanten Weiterentwicklung digitaler Technologien und der zunehmenden Digitalisierung aller Lebensbereiche kommen junge Kinder immer früher mit Informations- und Kommunikationstechnologien (ICT) in Kontakt (Guth 2021). Kindertageseinrichtungen (Kitas) spielen eine zentrale Rolle in dieser Entwicklung, indem sie Kindern ermöglichen, einen angemessenen und reflektierten Umgang mit digitalen Medien zu erlernen. Zahlreiche internationale Studien belegen zudem, dass digitale Medien das Entwicklungspotenzial von Kindern fördern können, wenn sie pädagogisch sinnvoll eingesetzt werden (Walter-Laager, Brandenburg, Tinguely, Schwarz, Pfiffner & Moschner 2017; Schacter & Jo, 2016; Egert, Cordes & Hartig, 2022).

Basierend auf dem Struktur-Prozess-Ergebnis-Modell (SPE-Modell) pädagogischer Qualität (Oppermann & Anders 2024) spielen professionellen Kompetenzen der frühpädagogischen Fachkräfte eine zentrale handlungsleitende Rolle für die pädagogische Qualität in Bildungssituationen. Für den Einsatz von ICT in Kitas gibt es bisher keine Untersuchungen zu Überzeugungen, emotionalen und motivationalen Kompetenzaspekten und der Nutzung von ICT mit Kindern. Diese Forschungslücke greift diese Studie auf und untersucht den Zusammenhang zwischen den Überzeugungen, dem Enthusiasmus, den Ängsten sowie dem

Interesse der Fachkräfte an der pädagogischen Arbeit mit ICT und der Nutzungshäufigkeit digitalgestützter Bildungsangebote in Kitas.

Die Studie basiert auf Daten einer Querschnittserhebung mit 266 frühpädagogischen Fachkräften aus 97 Kitas in Sachsen-Anhalt, Bayern, Nordrhein-Westfalen und Bremen. 44 Fachkräfte nahmen zusätzlich an einer standardisierten Tagebuchstudie teil (N=398 Tagebücher). Mittels standardisierter Fragebögen wurden die Fachkräfte sowohl zu den professionellen Kompetenzfacetten als auch zu ihrer täglichen Nutzung von ICT befragt. Institutionelle Merkmale wie Gruppengröße, Fachkraft-Kind-Schlüssel und technische Ausstattung wurden ebenfalls erhoben.

Es nahmen an der Studie überwiegend weibliche Fachkräfte (88%), mit einem Durchschnittsalter von 37.27 ($SD=11.21$) Jahren, teil. 17 Prozent der Fachkräfte haben einen Hochschulabschluss. In den teilnehmenden Kitas sprachen 30,5% der Kinder teilweise oder überwiegend eine andere Sprache als Deutsch und 21,8% der Kitas arbeiten nach dem offenen Konzept.

Die emotional-motivationalen Kompetenzfacetten der Fachkräfte, Selbstwirksamkeitsüberzeugungen, Angst und Enthusiasmus, wurden jeweils auf einer vierstufigen Skala mit jeweils drei Items erfasst (1 = stimme nicht zu; 4 = stimme zu.) (Selbstwirksamkeit z. B. „Ich bin mir sicher, dass ich über die besten Praktiken und Strategien verfüge, um den Kindern zu helfen, mit/über digitale(n) Medien zu lernen“, Angst z.B. „Wenn mir die Kinder Fragen zu digitalen Medien stellen, werde ich nervös“, Enthusiasmus: „Mir macht es Spaß zusammen mit den Kindern digitalen Medien im Kita-Alltag anzuwenden.“). Die Überzeugungen der Fachkräfte wurden in Form eines semantischen Differentials mit sieben Gegensatzpaaren über ein sieben-stufiges Antwortformat erfragt (z.B. „Digitale Medien in der pädagogischen Arbeit in Kindertageseinrichtungen sehe ich im Allgemeinen eher ... nutzlos – ... hilfreich“).

Die abhängige Variable „pädagogische Nutzung von ICT“ wurde als Summenscore aus zwei Items mit sechsstufiger Skala (1 = nie; 6 = täglich) verwendet (z.B. „Verwendung von Medien als Werkzeuge in Bildungssituationen, in denen Medien selbst nicht im Mittelpunkt stehen“). Für die allgemeine Nutzungshäufigkeit von ICT mit Kindern wurden die Fachkräfte danach befragt, welche Geräte sie mit den Kindern wie häufig verwenden (z.B. Tablets, Smartphones, E-Books; dichotom; 0 = nein; 1 = ja). Weiterhin wurde im Tagebuch gefragt, ob und wenn ja, welche ICT-Geräte in Lehr-Lernsituationen eingesetzt wurden. Die Daten wurden mithilfe von Mehrebenenanalysen ausgewertet und fehlende Daten mit FIML berücksichtigt. Strukturmerkmale (z.B. Fachkraft-Kind-Schlüssel, ICT-Ausstattung) wurden in allen Modellen kontrolliert.

Vorläufige Ergebnisse zeigen, dass Überzeugungsmerkmale und motivationale Merkmale in mittlerer Höhe signifikant korrelieren. Es zeigen sich ebenfalls signifikant positive Zusammenhänge zwischen der Selbstwirksamkeit und der Häufigkeit der Nutzung von ICT sowie der Häufigkeit pädagogischer Nutzung von ICT. Überzeugungen zeigen sich als nur relevant im Hinblick auf die Häufigkeit der pädagogischen Nutzung von ICT.

Die Ergebnisse betonen die Bedeutung von Selbstwirksamkeitserwartung und positiver Überzeugungen der Fachkräfte gegenüber ICT für eine qualitativ hochwertige Integration von ICT in den frühkindlichen Bildungskontext. Der Beitrag wird abschließend mit Blick auf Handlungsempfehlungen diskutiert.

Mit 7 Monaten Medienexperte: Welche Nutzungsprofilen mit digitalen Medien zeigen sich bei Säuglingen und was bedeuten sie für ihre Sprachentwicklung?

Dave Möwisch, Jessica Willard, Manja Attig

Leibniz-Institut für Bildungsverläufe

Soziale Interaktion ist bedeutsam für die frühe Sprachentwicklung (Tomasello, 2003), doch die Entwicklungskontexte haben sich durch die digitale Revolution verändert. Säuglinge sind von digitalen Geräten umgeben und interagieren mit diesen. Es gibt erhebliche Unterschiede in den spezifischen Geräten, ob damit interagiert wird oder sie im Hintergrund laufen, und im Umfang der digitalen Exposition (Palaologou, 2016; Rideout & Robb, 2020): Bereits Säuglinge erfahren zwischen null und über sechs Stunden digitale Exposition pro Tag, besonders in benachteiligten Familien (Brushe et al., 2023). Bisher ist diese Heterogenität nicht untersucht worden, sodass typische Muster nicht bekannt sind. Säuglinge können zwar durch Bildschirmmedien sprachliche Fähigkeiten erlernen, sie zeigen aber ein "Video-Defizit": Im Rahmen von sozialen Interaktionen ist es möglich, sich besser auf ihre kognitiven Fähigkeiten einzustellen, als dies typische kommerzielle digitale Medien tun könnten. Mit letzterem haben Säuglinge Schwierigkeiten, visuelle Repräsentationen mit ihrem 3D-Äquivalent zu verbinden und können kaum ein schnelles reizgesteuertes Orientieren zu flackernden oder bunten Regionen steuern (Guellai et al., 2022; Linebarger & Vaala, 2010). Dementsprechend ist nicht zu erwarten, dass typische digitale Expositionsmuster frühe Sprachentwicklung stimulieren. Säuglinge erfahren neben digitaler Exposition aber viele andere nicht-soziale Aktivitäten (Needham & Nelson, 2023), was die Frage aufwirft, ob jede Form der typischen digitalen Exposition die Sprachentwicklung negativ beeinflussen muss. Häufigere Bildschirmnutzung und Hintergrundfernsehen sind in Meta-Analysen mit geringeren Sprachkenntnissen verbunden (Madigan et al., 2020; Xie et al., 2024). Solche Ergebnisse könnten jedoch von Subgruppen mit sehr starker digitaler Exposition verursacht sein. Die Auswirkungen einer moderaten Exposition sind bisher unklar. Zu den möglichen Mechanismen negativer Effekte gehört das "Displacement" von Aktivitäten und Erfahrungen, die die Sprachentwicklung stimulieren; diese Effekte könnten bei moderater Exposition allerdings vernachlässigbar sein (McArthur et al., 2021). Zudem wird in bisherigen Untersuchungen oft das „analoge“ Umfeld vernachlässigt (z. B. Home Learning Environment, Swider-Cios et al., 2023), obwohl diese teilweise gering oder unerwartet hoch zusammenhängen (Attai et al., 2020; Lehl et al., 2021; Then et al., 2024). Widersprüchliche Ergebnisse aus Interaktionen digitaler Exposition mit strukturellen Familienvariablen (z. B. sozioökonomischer Status; Slobodin et al., 2024) könnten deshalb entstehen, da familiäre Prozessvariablen selten berücksichtigt werden. Die Berücksichtigung solcher Variablen kann Aufschluss geben, ob für Säuglinge je nach elterlicher Interaktionsqualität digitale Exposition nur im Sinne von „dual risk“ oder „kompensatorischen“ Prozessen negative Effekte hat. Die vorliegende Arbeit zielt darauf ab, für Säuglinge aus benachteiligten Familien in Deutschland (1) typische Muster der digitalen Exposition zu identifizieren, (2) Zusammenhänge mit dem Co-Viewing und der elterlichen Intention bei der Medienexposition zu untersuchen, (3) die Familienstruktur und die häusliche Lernumgebung sowie (4) Verbindungen zur Sprachkompetenz der Säuglinge zu analysieren. Zudem wird untersucht, (5) ob die Unterschiede mit "Displacement", „dual risk“/kompensatorischen Effekten im Einklang stehen. Über die theoretischen Implikationen hinaus sind die Ergebnisse relevant für die Unterstützung benachteiligter Familien und für die selten befolgten Forderungen nach frühzeitiger vollständiger digitaler Abstinenz (McArthur et al., 2022). Methode und Stichprobe Die Stichprobe von N = 300 Säuglingen stammt aus der BRISE-Studie (Bremer Initiative zur Stärkung frühkindlicher Entwicklung), die in sozioökonomisch benachteiligten Stadtvierteln Bremens vom Säuglingsalter an verfolgt. Mit sieben Monaten berichteten die Eltern über die digitale Exposition (Geräte, Vorder-/Hintergrundnutzung, Häufigkeit, Dauer, Co-Viewing und elterliche Intentionen) und die häusliche Lernumwelt (HLE und strukturelle Merkmale). Im Alter von 12 Monaten berichteten die Eltern zudem über die sprachlichen Fähigkeiten der Kinder (ELFRA, Grimm & Doil, 2006). Ergebnisse Vorläufige latente Profilanalysen zeigten, dass 40 % der benachteiligten Kinder keine digitale Exposition hatten, während es Profile mit starker (10 %) und mittlerer Exposition

(50 %) gab. Weitere Analysen werden die Zusammenhänge zwischen der Profizugehörigkeit und den Sprachkompetenzen untersuchen sowie die Rolle der häuslichen Lernumwelt, sowohl strukturell als auch prozessbezogen, analysieren.

Wirksame Sprachunterstützung durch digitale Bilderbuchbetrachtungen in Kindertageseinrichtungen

Anne-Kristin Cordes¹, Franziska Egert², Fabienne Körner^{1,4}, Wynona Kühn³

¹Staatsinstitut für Frühpädagogik und Medienkompetenz, Deutschland; ²Katholische Stiftungshochschule München; ³Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau; ⁴Pädagogische Hochschule Heidelberg

Für den Bildungserfolg von Kindern sind sprachliche Kompetenzen entscheidend (Burchinal et al., 2016). Doch weisen mindestens 20% der Vorschulkinder einen Sprachförderbedarf auf (Autorengruppe Bildungsberichterstattung, 2020), was die Unterstützung kindlicher Sprachfähigkeiten zur drängenden Aufgabe frühkindlicher Bildungseinrichtungen macht. Dazu eignen sich u.a. Bilderbuchbetrachtungen (Flack, Field, & Horst, 2018; Noble et al., 2019). Sprachunterstützend können dabei wiederkehrende Elemente in der Handlungsstruktur wirken, die mit Wiederholungen sprachlicher Strukturen einhergehen. Eine enge Passung (Kongruenz) von Handlung/Sprache und Illustrationen kann ebenso zum Verstehen beitragen (Mayer, 1989). In diesem Sinne können Bilderbücher *sprachlich-inhaltliches Scaffolding* bieten. Bilderbuchbetrachtungen bieten Raum für intensive Fachkraft-Kind-Interaktionen (Noble et al., 2019; Whitehurst et al., 1994). Durch den Einsatz von Sprachunterstützungsstrategien (wie Fragen, Impulse, korrekтивem Feedback oder Erweiterungen) – also durch *adaptives, kindgerichtetes Scaffolding* – können Fachkräfte die kindlichen Sprachfähigkeiten stärken (De Rivera et al., 2005; Whitehurst et al., 1988). Auch digitale Bilderbücher können die Sprachentwicklung unterstützen (Egert, Cordes & Hartig, 2022). Dabei sind Multimedia-Elemente, wie Hotspots oder Animationen, in den Blick zu nehmen. Diese bergen das Potenzial zum *multimedialen Scaffolding*, wenn sie kongruent zur Geschichte sind, können aber auch ablenkend wirken (Takacs et al., 2015; Bus et al., 2020). Die vorliegende Studie untersucht, wie digitale Bilderbücher über unterschiedliche Formen des Scaffolding möglichst sprachunterstützend eingesetzt werden können verglichen mit traditionellen Bilderbuchbetrachtungen (gedrucktes Buch) bzw. dem regulären Einrichtungssalltag. In drei cluster-randomisierten Studien wurden der sprachliche Input (Vorlesefunktion vs. Vorlesen vs. adaptives Vorlesen), das Sprachunterstützungshandeln der Fachkraft (Dialogisches Lesen nach Whitehurst vs. selbstgesteuertes Lesen) und das Buchformat (digitales vs. traditionelles, gedrucktes Bilderbuch) variiert. Rezeptiver und expressiver Wortschatz sowie Bildungssprache (Morphologie) wurden im Prä- und Posttest erhoben. 89 4- bis 6-Jährige wurden (basierend auf dem PPVT) in homogene Kleingruppen von 2-3 Kindern eingeteilt, welche anschließend zufällig den Bedingungen zugeteilt wurden (13-18 Kinder pro Bedingung). Die Kurzzeitinterventionen umfassten 3 Einheiten mit der Geschichte „Oskar und der sehr hungrige Drache“ (Krause, 2007; Oetinger Verlag & Krause, 2015), die insgesamt 60 bis 90 Minuten dauerten. Für das Dialogische Lesen wurde ein Leitfaden entwickelt, der mit hoher Implementierungsqualität umgesetzt wurde. Im selbstgesteuerten Lesen betrachteten Kinder das digitale Bilderbuch selbstständig unter Begleitung einer Erwachsenen, die Fragen an die Kinder zurückspeiegelte, um Peer-Interaktionen anzuregen. Es ergaben sich signifikante Vorteile des Dialogischen Lesens für den expressiven Wortschatz sowie kleine Effekte für Bildungssprache (Morphologie) verglichen mit dem regulären Kindergartenalltag. Das Dialogische Lesen erwies sich als vergleichbar wirksam für digitale und traditionelle Bilderbuchbetrachtungen. Selbstgesteuerte digitale Bilderbuchbetrachtungen waren nicht sprachunterstützender als der reguläre Kindergartenalltag. Vorlesen und adaptives Vorlesen waren nahezu so hilfreich wie die Vorlesefunktion digitaler Bilderbücher für die Wortschatzerweiterung. Im Vergleich zu früheren Studien mit Dialogischem Lesen in Deutschland, die kleine bis mittelgroße Effekte zeigten (u.a. Ennemoser, Kuhl & Pepouna, 2013; Ennemoser, Lehnigk, Hohmann & Pepouna, 2015), erscheinen unsere Effektstärken als relativ groß, was möglicherweise auf die hohe Implementierungsqualität, die hohe Anwesenheit der Kinder oder den dreifach wiederholten Input der Geschichte zurückzuführen sein könnte. Unsere Befunde legen nahe, dass die Sprachfähigkeiten mit digitalen Bilderbüchern, die nach der Methode des Dialogischen Lesens betrachtet werden, ähnlich gut zu stärken sind wie mit traditionellen Bilderbuchbetrachtungen und Dialogischem Lesen. Aus den Daten ergab sich kein signifikanter Mehrwert für multimediales Scaffolding. Möglicherweise geht dies auf die Umsetzung der Hotspots zurück, die zwar überwiegend inhaltlich kongruent waren, aber gleichzeitig eine ablenkende Komponente aufwiesen (Blinken während des Lesens), was Ressourcen im Arbeitsgedächtnis beansprucht haben könnte, die somit nicht für die Verarbeitung der Geschichte zur Verfügung standen (Cordes, Egert & Hartig, 2020; Mayer 2022). Die Weiterentwicklung digitaler Bilderbücher nach kognitions- und lernpsychologischen Überlegungen könnte hier Abhilfe schaffen. Nichtsdestotrotz stellen bereits vorhandene, überwiegend lernunterstützend aufbereitete digitale Bilderbücher in Kombination mit Dialogischem Lesen eine Chance zur Stärkung der sprachlichen Fähigkeiten von Kindern in frühkindlichen Bildungseinrichtungen dar.

PS-03: Ansätze in der Lese- & Schreibförderung

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 154

Chair der Sitzung: Joachim Grabowski

Systematische Leseförderung mit dem Hamburger Leseband – Wirksamkeit und Bedeutung für das Leseverstehen in der Grundschule.

Anna Rau, Jonna Wolters, Britta Pohlmann

Institut für Bildungsmonitoring und Qualitätsentwicklung, Deutschland

Theoretischer Hintergrund

Das Thema Sprachförderung steht nicht zuletzt aufgrund der Erkenntnisse internationaler und nationaler Schulleistungsstudien (Deutsches PISA-Konsortium, 2001; Bos & Pietsch, 2006) im Fokus schulpolitischer Maßnahmen. Die dabei vielfach konstatierte mangelnde Lesekompetenz hat in Hamburg zur Implementation eines umfassenden Sprachförderkonzepts geführt, das unter anderem eine systematische Leseförderung vorsieht. Mit der Einführung einer verbindlichen regelmäßigen Lesezeit, dem sogenannten Leseband, wird das Ziel verfolgt, Schüler*innen an Schulen in herausfordernder Lage in ihrer Lesesozialisation zu fördern und langfristig die Lesekompetenz zu steigern (Wolters et al., in Druck). Anhand verschiedener Lautlesemethoden (Rosebrock & Nix, 2011) wird die Leseflüssigkeit gefördert, die als wesentliche Voraussetzung für das verstehende Leseverstehen gilt (Fuchs et al., 2001; Kim et al., 2021; Klauda & Guthrie, 2008; Sappok et al., 2020).

Der Hamburger Leseansatz wurde im Rahmen der Bund-Länder-Initiative „Bildung durch Sprache und Schrift“ (BiSS) in Hamburg von 2015 bis 2018 an sechs Pilotschule erprobt (Gailberger, 2019). Aufgrund der positiven Ergebnisse wurde die Maßnahme auf weitere Schulen mit heterogenerer Schülerschaft ausgeweitet (so genannte Transferschulen), die weniger eng begleitet wurden als die Pilotschulen. Damit das Lesetraining auch in dieser erweiterten Gruppe an Schulen erfolgreich ist, sollten der Implementationsforschung zufolge die Kriterien Akzeptanz, Machbarkeit, Kooperation und Wiedergabetreue erfüllt sein (Souvignier et al., 2021). Dies ist nur möglich, wenn die Fördermaßnahme gut in das schulische Setting integriert ist, die Notwendigkeit der Maßnahme vom Kollegium anerkannt wird und das Fortbildungskonzept passgenau ist (Meudt et al., 2020; Souvignier et al., 2021).

Fragestellungen

- 1) Wie gelingt die Implementierung des Lesetrainings an den Transferschulen?
- 2) Kann die Wirksamkeit des Lesetrainings auf die Leseflüssigkeit auch in den Transferschulen nachgewiesen werden?
- 3) Zeigt sich der Wirkzusammenhang zwischen Leseflüssigkeit und Leseverstehen?
- 4) Welche Gelingensbedingungen sind besonders relevant?

Methode

Zur Untersuchung der Implementierung (Fragestellung 1) werden Kurzskalen zur Erfassung des Transfererfolgs von Leseförderkonzepten von Meudt et al. (2020) eingesetzt. Neben einer deskriptiven Auswertung der Implementationsmerkmale sollen verschiedene Gruppen von Schulen (nach Zusammensetzung der Schülerschaft, Teilnahmedauer) vergleichend analysiert werden. Insgesamt liegen Daten aus 61 Schulen vor.

Die Wirksamkeit des Trainings auf die Leseflüssigkeit (Fragestellung 2) wird mittels latenter Wachstumskurvenmodelle (Bauer & Curran, 2021) untersucht. Die Datengrundlage liefern 3.077 Schüler*innen aus insgesamt 49 Schulen, die von Schuljahr 2019/20 bis 2021/22 am Lesetraining teilgenommen haben und mittels des Salzburger Lese-Screenings 2-9 (Mayringer & Wimmer, 2014) halbjährlich auf ihre Leseflüssigkeit getestet wurden.

In Verbindung mit weiteren Leistungsdaten aus den Lernstandserhebungen KERMIT („Kompetenzen ermitteln“, Lücken et al., 2014) wird anhand von Strukturgleichungsmodellen überprüft, wie sich die Wirkzusammenhänge zwischen Leseflüssigkeit auf Leseverstehen über mehrere Zeitpunkte (Jahrgangsstufe 2, 3, und 4) hinweg gestalten (Fragestellung 3).

Schließlich soll durch die Koppelung der Angaben zur Implementierung und der Leistungsdaten mithilfe von Mehrebenenstrukturgleichungsmodellen (Hox, 2013) analysiert werden, welchen Einfluss die Art der Umsetzung sowie die Überzeugungen und Einstellungen der Lehrkräfte auf den Erfolg des Trainings haben (Fragestellung 4).

Ergebnisse und ihre Bedeutung

Hinsichtlich der Implementation berichten die Schulen von einer hohen Akzeptanz und Machbarkeit des Trainings, wobei sich keine wesentlichen Unterschiede zwischen Teilgruppen von Schulen ausmachen lassen. Die Kompetenzdaten zeigen, dass das Lesetraining zu einer signifikanten Steigerung der Leseflüssigkeit im Verlauf der Primarstufe führt. Die Schüler*innen nähern sich in ihrem Leistungsniveau dem altersgemäßen Erwartungswert an und können damit ungünstige Ausgangslagen kompensieren. Darüber hinaus kann der vermutete Wirkzusammenhang zwischen Leseflüssigkeit und Leseverstehen nachgewiesen werden. Der Einfluss der Implementationsmerkmale auf die Lernerfolge Schüler*innen wird derzeit noch untersucht.

Die vorliegenden Ergebnisse liefern wichtige Erkenntnisse zur Wirksamkeit systematischer Leseförderung im schulischen Setting. Mit dem Lesetraining wird ein Ansatz beschrieben, der sich in der Praxis gut umsetzen lässt und zur gezielten Verbesserung basaler Kompetenzen und der gesellschaftlichen Teilhabe von Schüler*innen beiträgt. Das Vorhaben zeigt, wie Wissenschaft Erkenntnisse bereitstellt, die von Bildungsadministration und -Politik zur Steuerung genutzt werden.

Effizient Lesen lernen mit FLINK: Individuelle Förderung der Leseflüssigkeit

Fenke Kachisi^{1,2}, Jelena Marković^{1,2}, Jan-Henning Ehm^{2,3}, Telse Nagler^{1,2}

¹DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ²IdEA - Center for Research on Individual Development and Adaptive Education of Children at Risk; ³Pädagogische Hochschule Heidelberg

Ein frühzeitiges Interventionsangebot für Kinder mit Leseschwierigkeiten ist entscheidend, um Lesefähigkeiten zu verbessern und schulische Probleme zu vermeiden. Ein Forschungszweig der Leseintervention befasst sich dabei mit der experimentellen Beeinflussung der Lesegeschwindigkeit durch schrittweises Ausblenden von Text (Breznitz & Berman, 2003; Nagler et al., 2014). Hierbei wird das Lesematerial, angepasst an das individuelle Leseniveau, Buchstabe für Buchstabe in Leserichtung ausgeblendet (Text-Fading). Diese Methode wurde bereits in mehreren Sprachen als Training zur Verbesserung der Leseflüssigkeit eingesetzt (Korinith & Nagler, 2021). Auch für die deutsche Sprache konnte die Trainingswirksamkeit hinsichtlich Lesegeschwindigkeit und -verständnis bei Grundschul*innen nachgewiesen werden (Nagler et al., 2021).

Dieses Text-Fading-Training wurde zu einem browserbasierten Förderangebot weiterentwickelt, um eine leichte Zugänglichkeit für eine große Anzahl von Kindern zu ermöglichen. Das digitale Leseflüssigkeitstraining (FLINK) ist in die Lernverlaufsdiagnostik-Plattform *quop* integriert und kann von Lehrkräften für ausgewählte Kinder freigeschaltet werden. Aktuell wird die Teilnahme für Kinder im mittleren und unteren Leistungsbereich im Lesen der Klassenstufe 3 und 4 bereitgestellt. Ab nächstem Schuljahr (24/25) wird FLINK ebenfalls für die Klassen 5 und 6 erweitert. Die Kinder absolvieren im Verlauf eines Schuljahres alle drei Wochen die für die Klassenstufe vorgesehenen *quop*-Tests. Zwischen zwei *quop*-Tests können bis zu sechs FLINK-Übungseinheiten mit jeweils 30 Aufgaben bearbeitet werden, wodurch im Schuljahr bis zu 54 FLINK-Übungen durchgeführt werden können. Alle Kinder beginnen mit der gleichen Ausblendegeschwindigkeit des präsentierten Lesematerials. Im Verlauf des Trainings wird diese Ausblendegeschwindigkeit, basierend auf der Verständnisleistung, individuell angepasst.

Ziel dieses Projektes ist, die Wirksamkeit von FLINK im Schulalltag zu überprüfen. Für die Trainingsevaluation wurde untersucht, ob der Verlauf der Lesegeschwindigkeit im externen *quop*-Test über das Schuljahr 22/23 hinweg durch die FLINK-Trainingsteilnahme beeinflusst wurde. Anhand eines Latent Growth Curve Model wurde an 725 deutschsprachigen Drittklässler*innen, die sowohl an FLINK als auch an der Lernverlaufsdiagnostik *quop* teilnahmen, untersucht, ob sich die Trainingsteilnahme prädiktiv auf Lesegeschwindigkeitsverbesserungen auswirkt. Hierzu wurde die Stichprobe in eine Gruppe mit niedriger (1-6 Trainingssessions, $M = 3.06$) sowie eine Gruppe mit hoher Trainingsintensität (7-40 Trainingssessions, $M = 13.2$) eingeteilt. Die Analysen ergaben einen signifikanten linearen ($\mu_s = 1.477$, $z = 6.049$, $p < .001$) und quadratischen ($\mu_s = 0.070$, $z = 2.213$, $p < .05$) Anstieg der Lesegeschwindigkeit im Laufe des Schuljahres. Die Gruppe mit hoher Trainingsintensität wies einen signifikant höheren linearen Zuwachs der Lesegeschwindigkeit im externen Lesetest *quop* auf als die Gruppe mit niedriger Trainingsintensität ($\beta = 0.297$, $SE = 0.490$, $z = 1.973$, $p = .048$), was auf die Wirksamkeit des Trainingsansatzes bezüglich einer Steigerung der Lesegeschwindigkeit hinweist.

Da das FLINK-Training im Schuljahr 22/23 eingeführt wurde, konnten interessierte Lehrkräfte über das ganze Schuljahr hinweg Schüler*innen zum Training anmelden. Dadurch kamen im Laufe des Schuljahres etliche neue Schulen hinzu, sodass die Anzahl an teilnehmenden Kindern zwar stetig anstieg, jedoch nur diejenigen Schüler*innen das Training im erhofften Umfang durchführen konnten, die zu Beginn des Schuljahres mit dem Training gestartet haben. Im Schuljahr 23/24 hingegen wurde das Training fortgeführt, sodass die Anzahl an Kindern, die von Anfang an teilgenommen haben, deutlich höher ausfiel. Aktuell haben 18907 Kinder mit dem Training im Schuljahr 23/24 gestartet.

Im geplanten Konferenzbeitrag werden die Ergebnisse der abgeschlossenen Datenerhebung des Schuljahres 22/23 sowie Ergebnisse aus dem Schuljahr 23/24 präsentiert. Da das FLINK-Training in unterschiedlichen Bundesländern durchgeführt wurde, ist das Ende der Datenerhebung an die jeweiligen Sommerferienzeiten gebunden. Aus diesem Grund konnten die Daten aus dem Schuljahr 23/24 zu diesem Zeitpunkt noch nicht analysiert werden. Neben der Lesegeschwindigkeit im externen *quop*-Test werden dann ebenfalls weitere Variablen wie bspw. Lesegenauigkeit, Leseverständnis oder Lesemotivation in die Analysen einbezogen. Diese neu analysierten Daten sollen auf der Konferenz vorgestellt und die praktischen Implikationen des Trainings für den Einsatz an deutschen Schulen sowie Weiterentwicklungsmöglichkeiten diskutiert werden.

Der Zaubertrick mit der Karte: Lassen sich die Fähigkeiten zur Kohärenzbildung und zur Perspektivenübernahme beim Schreiben separat und erfolgreich fördern?

Joachim Grabowski, Moti Mathiebe, Christian Weinzierl

Leibniz Universität Hannover, Deutschland

Aufbau und Förderung literaler Fähigkeiten gelten als zentrale Aufgaben vor allem der frühen Angebote des Bildungssystems. Neben dem Lesen spielt auch die – im öffentlichen Diskurs weniger saliente und bildungswissenschaftlich seltener beforschte – Schreibkompetenz eine wichtige Rolle. Die sprachbiografisch heterogene Zusammensetzung von Schulklassen stellt dabei eine zusätzliche Herausforderung dar (Marx, 2017).

Konstrukte literaler Fähigkeiten umfassen in der Regel kognitive und sprachliche Facetten; ihre Erforschung profitiert von der psychologischen und sprachdidaktischen Zusammenarbeit. Hier zeigte sich, dass neben allgemeinen kognitiven und sprachlichen Fähigkeiten (wie Arbeitsgedächtniskapazität, Wortschatz und Leseflüssigkeit) vor allem die Komponenten des Kohärenzmanagement (Verstehen und sprachliche Gestaltung von Zusammenhangsstrukturen) und der Perspektivenübernahme (andere als die jeweils eigene Perspektive einnehmen und berücksichtigen zu können) die Qualität geschriebener Texte (unterschiedlicher Genres) vorhersagen und deshalb bei einer Förderung der Schreibkompetenz entsprechend berücksichtigt werden sollten (Becker-Mrotzek et al., 2015).

Die vorliegenden Interventionsstudie ist eine empirische Miniatur zur Vorbereitung einer umfassenderen Intervention auf der Basis eines ausgearbeiteten Unterrichtskonzepts. Es geht hier darum, ob sich mit Blick auf die Textqualität bei Fünftklässler*innen die Fähigkeitskomponenten des Kohärenzmanagement (Weinzierl, 2022) und die der Perspektivenübernahme (Schmidt, 2022) separat durch spezielle didaktische Einheiten fördern lassen und ob sich die jeweils fokussierte Förderung in zugehörigen Eigenschaften der geschriebenen Texte diskriminativ niederschlägt. Dazu wurden für die beiden Fähigkeitskomponenten jeweils fünf Unterrichtseinheiten konzipiert, die von Mitarbeitenden des Forschungsteams in zwei (jeweils halbierten) fünften Gymnasialklassen an zwei großstädtischen Standorten ($n = 56$) über fünf Wochen hinweg durchgeführt wurden. Die quasi-experimentelle Variation besteht also in der fünfständigen unterrichtlichen Förderung der einen bzw. der anderen Fähigkeitskomponente. Diese nutzte aufwändig ausgearbeitete Aufgaben- und Instruktionsmaterialien aus früheren Interventionen, die von den durchführenden Personen in der Rolle einer Lehrkraft instruiert und begleitet wurden.

Vor und nach der Intervention (= Prä-post-Design) wurden korrespondierende diagnostische Einheiten (Vor- und Nach-Test) durchgeführt; diese umfassten jeweils drei testartige Aufgaben zur unmittelbaren Indikation der Kohärenz- und Perspektivenfähigkeit sowie zwei Schreibaufgaben, einen Bericht sowie die Zaubertrick-Schreibaufgabe. Diese Aufgabe wurde so konzipiert und umgesetzt, dass sie in besonderer Weise geeignet ist, die zu prüfende Korrespondenz zwischen den Förderinhalten und den resultierenden Aspekten der Textqualität zu prüfen. Der Zaubertrick, bei dem eine Dose scheinbar auf einer Spielkarte stehen kann, ohne zu schwanken oder herunterzufallen, wird dabei filmisch aus zwei Perspektiven gezeigt: Zunächst sieht man den Trick aus der Perspektive der Zuschauer, erlebt also den Effekt, ohne ihn sich erklären zu können. Die hiermit korrespondierende Schreibaufgabe ist eine *Beschreibung* des Tricks. Danach wird der Trick sozusagen von „hinter der Bühne“ gezeigt; man sieht nun, wie der Trick funktioniert. Die mit dieser Perspektive korrespondierende Schreibaufgabe ist eine *Erklärung* des Tricks. Die Verteilung der beiden Aufgabenperspektiven auf die beiden Messzeitpunkte wurde in beiden Interventionsgruppen permutiert, so dass Reihenfolgeeffekte kontrolliert werden können. Die Qualität der geschriebenen Texte (pro Kind jeweils zwei Texte an beiden Messzeitpunkten) wurde mithilfe eines semi-holistischen Ratingverfahrens (sog. „naives“ Rating; vgl. Grabowski et al., 2014, 2018) gemessen.

Bei der direkten Indikation der Fähigkeitskomponenten ergab sich nur für den Test „Konjunktionen verwenden“ (als Kohärenzaspekt) eine stärkere Verbesserung der Kohärenzgruppe im Vergleich zur Perspektivenübernahmegruppe ($F(1, 56) = 4.1$; $p < .05$), keine bedeutsamen Unterschiede bei den beiden anderen direkten Indikatoren (Reihenfolge Bildergeschichten beurteilen; Zuweisung von Emotionsausdrücken). Die Schüler:innen scheinen bei der Beschreibungsaufgabe stärker von der kohärenzfokussierenden Intervention zu profitieren als von der Fokussierung der Perspektivenübernahme; bei der Erklärungsaufgabe ergab sich kein Unterschied. Die Detailanalysen der Textqualität und ihrer konstituierenden Eigenschaften sind jedoch noch nicht abgeschlossen.

Im Vortrag wird neben dem empirischen Grundbericht der Studie vor allem die Umsetzung der psychologischen Konstrukte in didaktisierte Materialien und die Entwicklung geeigneter diagnostischer Aufgaben erläutert und diskutiert. Solche komponentiellen Aufgaben halten wir (gegenüber der Anforderung von Langtexten) für die unterrichtliche Förderung der Schreibkompetenz für besonders relevant.

Transfer evidenzbasierter Schreibfördermaßnahmen in die Praxis – Analyse transferrelevanter Einflussfaktoren auf Lehrkräfteebene

Pia Sieveke¹, Steve Graham², Vera Busse¹

¹Universität Münster, Deutschland; ²Arizona State University, USA

Theoretischer Hintergrund:

Kompetentes Schreiben ist eine Schlüsselkompetenz des 21. Jahrhunderts, die für den erfolgreichen Übergang von der Grundschule zur weiterführenden Schule von zentraler Bedeutung ist (Cutler & Graham, 2008). Obwohl zahlreiche Studien die Wirksamkeit von Schreibförderungsstrategien belegen, erweist sich der Transfer in den Unterrichtsalltag häufig als herausfordernd (Broekamp & van Hout-Wolters, 2007; Koster et al., 2017). Es besteht ein dringender Forschungsbedarf hinsichtlich der erfolgreichen Übertragung validierter Förderstrategien in den Schulkontext sowie der nachhaltigen Unterstützung von Lehrkräften bei deren Anwendung.

Ziel:

Übergeordnetes Ziel des Projekts ist es, Schreibkompetenz, Schreibmotivation und soziale Partizipation an Grundschulen zu fördern. Das Ziel der vorliegenden (Teil-)Studie besteht darin, den Transfer zwischen Schreibforschung und Unterrichtspraxis zu verbessern und ein vertieftes Verständnis der Bedingungen erfolgreichen Transfers von Schreibfördermaßnahmen zu erlangen.

Methode:

Das Projekt ist als (quasi)experimentelle Interventionsstudie mit einem Pre- und Posttest angelegt. Die Stichprobe umfasst elf Grundschulen ($\gg 1.500$ Schüler*innen). Zur Evaluierung der Intervention wurden standardisierte schriftliche Befragungen ($n = 23$) und Interviews ($n = 10$) mit Lehrkräften der Experimentalgruppe durchgeführt. In dieses Gesamtprojekt eingebettet, fokussiert die vorliegende Studie den Implementationsprozess, wobei die Lehrkräfte als Übermittler der Transfermaßnahme im Mittelpunkt stehen. Es sollen folgende Forschungsfragen beantwortet werden:

FF 1: Wie schätzen Lehrkräfte vor der Fortbildung ihr Wissen zum Schreibunterricht und die Effektivität ihres Feedbacks ein?

FF 2: Wie beeinflusst die Fortbildung das Wissen der Lehrkräfte zum Schreibunterricht und die wahrgenommene Effektivität des Feedbacks?

FF 3: In welchem Umfang implementieren Lehrkräfte Schreibstrategien und Feedback?

FF 4: Welche Faktoren beeinflussen die Entscheidung der Lehrkräfte, Schreibstrategien und Feedback im Unterricht umzusetzen?

Um eine umfassende Analyse der Implementierungsprozesse zu ermöglichen, wurde ein Mixed-Methods-Ansatz gewählt, der die Triangulation verschiedener Datenquellen beinhaltet. Lehrkräfte der Experimentalgruppe nahmen an einer Fortbildungsmaßnahme teil, die darauf abzielte, die Implementierung evidenzbasierter Strategien für Schreibunterricht und Feedback im Klassenraum zu fördern. Quantitative Daten zum Wissen der Lehrkräfte über Schreibunterricht und Feedback wurden vor und nach der Fortbildung erhoben. Veränderungen des selbsteingeschätzten Wissens wurden mittels t -Tests für abhängige Stichproben untersucht. Die Implementierung der Interventionsmaßnahme im Unterricht werteten wir mittels Dokumentationsmappen deskriptiv aus. Ergänzend wurden leitfadengestützte Interviews mit Lehrkräften durchgeführt, um Faktoren zu identifizieren, die die Implementierung beeinflussen. Die transkribierten Interviews wurden unter Verwendung der MAXQDA-Software inhaltsanalytisch (Mayring, 2008) ausgewertet.

Ergebnisse:

Die Fragebogendaten zeigen, dass die Lehrkräfte ihr Wissen über Schreibunterricht ($M = 4,2$; 6-Point-Likert-Skala) und Feedback ($M = 4,9$; 6-Point-Likert-Skala) zum Zeitpunkt T1 als hoch einschätzten; das selbstberichtete Schreibwissen stieg nach der Fortbildung signifikant an ($t(20) = -5,936$, $p < .001$, $d = 1,295$). Allerdings ergab sich keine signifikante Veränderung in der wahrgenommenen Wirksamkeit des eigenen Feedbacks zwischen T1 und T2 ($t(19) = -1,469$, $p = .158$).

Die Analyse der Dokumentationsmappen ergab eine deutliche Varianz in der Intensität der Umsetzung der Intervention durch die Lehrkräfte. Zudem wurden die Dokumentationsmappen nicht von allen Lehrkräften ausgefüllt. Um dieser Herausforderung zu begegnen, wurde die Implementierung der einzelnen Interventionselemente zusätzlich über die durchgeführten Interviews rekonstruiert. Positiv hervorzuheben ist, dass die Betrachtung der Nutzungsmuster der verschiedenen Interventionselemente nahelegt, dass bestimmte Elemente trotz hoher Komplexität (z.B. das Modellieren; zur Methode des Modellierens vgl. Siekmann et al., 2022) häufig eingesetzt wurden.

Die Interviewdaten offenbaren jedoch, dass die allgemeine Akzeptanz und wahrgenommene Angemessenheit der Intervention zwischen den Lehrkräften stark variierten. Unter Berücksichtigung transferrelevanter Einstellungen (Akzeptanz und Angemessenheit) können drei Implementationsprofile herausgearbeitet werden (hohe Akzeptanz u. hohe Angemessenheit; mittlere Akzeptanz u. mittlere Angemessenheit; geringe Akzeptanz u. geringe Angemessenheit). Weiterhin zeigen die Interviews konzeptuelle Diskrepanzen zwischen Fortbildungsinhalten und deren Umsetzung im Unterricht, die die Implementationsqualität beeinflussen.

Die Ergebnisse werden im Hinblick auf Fortschritte in den Schreibleistungen der Kinder diskutiert, die zurzeit ausgewertet werden. Gelingensbedingungen für einen erfolgreichen Transfer und eine nachhaltige Implementation werden vor dem Hintergrund der Ergebnisse diskutiert.

PS-04: Passung und Commitment in der beruflichen Bildung: Erfolgsfaktoren und Herausforderungen

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: SO 418

Chair der Sitzung: Sylvia Rahn

Chair der Sitzung: Sebastian Rahn

ENTFÄLLT: Person-vocation fit and vocational educational training success

Anika Bela, Daniel Bela

Leibniz-Institut für Bildungsverläufe (LIfBi), Deutschland

The transition from school to the labor market is the foundation for young people's future labor market outcomes and overall life opportunities. However, the rate for premature vocational educational training (VET) contract termination lies at about 25% in Germany, with an upward trend since the 1990s (BIBB 2015). This interruption in an individual's CV leads to loss of income and time and can have serious negative consequences for the labor market career. A high dropout rate also poses an additional threat to the already growing shortage of skilled workers in Germany, which is a further reason why it is important to expand knowledge about the determinants of VET completion.

Determinants for successful VET can be found at regional or individual level as well as in the characteristics of the training occupation or the companies (Böhn & Deutscher 2022). At the individual level, different studies have shown, that both cognitive skills and socio-emotional abilities have considerable predictive power for VET outcomes and for what happens after a VET contract ends (Michaelis & Findeisen 2022, Michaelis & Richter 2022, Holtmann & Solga 2023).

As a theoretical basis, we rely on the assumptions of the person-environment fit approach (Kristof-Brown et al. 2005). This approach emphasizes the congruence between individual characteristics of a person and various aspects of a particular job as important factors for job satisfaction and performance. In this study, we focus on the person-vocation fit, which reflects the match between an individual's skills and abilities on one, and the general requirements of the occupation on the other side. We therefore want to investigate the relevance of the initial level of both cognitive and socio-emotional skills, measured just before leaving school, for an individual's chance to successfully complete a first VET. We complement previous research by considering both types of skills simultaneously and aim to answer the question of how important the fit between the trainees' initial skills and the requirement-related characteristics of the training occupation is for predicting the VET outcome. Additionally, we are interested in whether certain skills compensate for deficits in the fit between individual abilities and occupation-specific requirements.

For this study, we use data from the German National Educational Panel Study's (NEPS) Grade 9 cohort (NEPS Network 2021). The data includes measures regarding a person's self-perception, social skills and personality as well as in-depth assessments of cognitive competencies. Preliminary results from probit models indicate that the probability for successful completion of the VET increases with higher mathematical competences, particularly in VETs for occupations in business administration. Reading competences generally seem to play less of a role. Social skills are beneficial especially for the success in VETs within the service sector. These results can be interpreted as an initial indication that it plays a role whether a trainee is already equipped with the relevant skills for meeting the requirements of a particular vocation.

To explicitly investigate the relevance of the fit between an individual's skills and the skills required within the occupation, we plan to use two further data sets with additional useable information. We want to use variables from the Occupational Panel for Germany (Grienberger et al. 2023) that classify occupations in terms of five task types (Dengler et al. 2014). We intend to deepen these analyses by using the tasks-data conducted in NEPS-Starting Cohort 6, which we can link to the data of NEPS-Starting Cohort 4. With this data we expect to be able to examine the importance of a person-vocation fit for VET completion, but also to shed light on skills that can compensate for a lack in the person-vocation fit.

Vertragslösungsneigung in geschlechtsuntypischen Ausbildungsverhältnissen – empirische Prüfung und praktische Implikationen eines Erklärungsmodells

Sascha Hanel¹, Christoph Fuhrmann¹, Sylvia Rahn^{1,2}

¹Bergische Universität Wuppertal, Deutschland; ²Deutsches Institut für Erwachsenenbildung, Deutschland

Theoretischer Hintergrund und Fragestellung

Die Förderung geschlechtsuntypischer Ausbildungsverhältnisse gilt seit geraumer Zeit als probater Beitrag zu mehr Chancengerechtigkeit zwischen den Geschlechtern im Berufsbildungssystem und auf dem Arbeitsmarkt. Allerdings wird die Geschlechtersegregation im deutschen Berufsbildungssystem nur sehr langsam reduziert. Dies liegt auch daran, dass geschlechtsuntypische Ausbildungsverhältnisse häufiger wieder vorzeitig gelöst werden (Rohrbach-Schmidt & Uhly 2015).

Vor diesem Hintergrund wird im Vortrag der dreifachen Frage nachgegangen:

1. warum das so ist,
2. ob a) ein aus mehreren Theorieangeboten und dem Forschungsstand synthetisiertes Erklärungsmodell die Vertragslösungsneigung (bspw. Deuer 2006) in geschlechtsuntypischen Ausbildungsverhältnissen statistisch erklären kann und b) diese erwartungsgemäß besser erklärt als in geschlechtstypischen Ausbildungsverhältnissen,
3. inwieweit sich die Faktoren bzw. das relative Gewicht, mit dem sie die Vertragslösungsneigung der jeweiligen geschlechtlichen Minorität in typischen Männer- und Frauenberufen erklären, voneinander unterscheiden.

Zu diesem Zweck wird im Vortrag zunächst aus erwartungs-wert-theoretischen (Eccles & Wigfield 2002; 2020), passungs- und sozialisationstheoretischen (bspw. Stalder & Schmid 2016) Ansätzen ein Erklärungsmodell für die Vertragslösungsneigung in geschlechtsuntypischen Ausbildungsverhältnissen extrahiert. Postuliert wird, dass die Vertragslösungsneigung in geschlechtsuntypischen Ausbildungsverhältnissen – neben Individualmerkmalen der Auszubildenden und Kontextmerkmalen des Ausbildungsverhältnisses – von folgenden übergeordneten Faktoren beeinflusst wird: der Berufsorientierung im Vorfeld der Ausbildung und dem sozialen Ausbildungsklima (Beckmann 2023), der wahrgenommenen Prozessqualität der Ausbildung und etwaigen aus dem geschlechtlichen Minderheitenstatus resultierenden Benachteiligungen. Angenommen wird ferner, dass diese in eine mehr oder weniger ausgeprägte berufliche Beanspruchung während der Ausbildung münden. All dies wird durch die sozialen Normen des privaten Umfeldes der Auszubildenden beeinflusst und mündet in variierende Erwartungen der Auszubildenden an und Ziele für ihre berufliche Zukunft.

Methode

Das aufgestellte Erklärungsmodell wurde durch die Adaption eingeführter Skalen und nur wenigen Neuentwicklungen mit durchweg akzeptablen bis sehr guten internen Konsistenzen (Omega zwischen .68 und .93) operationalisiert. Anhand einer Gelegenheitsstichprobe aus insgesamt 1373 Auszubildenden (119 Frauen, 124 Männer in geschlechtsuntypischen sowie 363 Frauen, 767 Männer in geschlechtstypischen Ausbildungsverhältnissen) wird im Vortrag durch die Anpassung eines Strukturgleichungsmodells die Gültigkeit des Erklärungsmodells für die Gesamtstichprobe geprüft.

Nach der Prüfung der Invarianz des Messmodells werden sodann mittels Gruppenvergleichen die Gemeinsamkeiten und Unterschiede in der statistischen Determination der Vertragslösungsneigung junger Frauen in Männerberufen und junger Männer in Frauenberufen identifiziert. Ergänzend wurden differenzierter operationalisierte Regressionsmodelle für die Teilgruppen gerechnet, um ein genaueres Bild von den jeweiligen Gründen für die Neigung zur vorzeitigen Auflösung geschlechtsuntypischer Ausbildungsverhältnisse zu erhalten.

Ergebnisse

Nach aktuellem Stand der Analysen gelingt die Anpassung eines „schlank“ operationalisierten Erklärungsmodells mit je einer Skala pro Dimension an die Daten der Gesamtstichprobe mit einem guten absoluten, aber nur knapp akzeptablen relativen Modellfit. Die strenge Invarianz des Messmodells zwischen den Gruppen ist gegeben, d.h. Vergleiche der Mittelwerte und Koeffizienten zwischen den Teilgruppen (männliche Minoritäten in Frauenberufen und weibliche Minoritäten in Männerberufen miteinander und der jeweiligen Majorität) sind möglich. Die Testung der Koeffizienten und die Gruppenvergleich stehen noch aus.

Aktuelle Befunde aus den Regressionsmodellen lassen bei einer jeweils aufgeklärten Gesamtvarianz von mindestens 61% folgende Gründe für die Vertragslösungsneigung in geschlechtsuntypischen Ausbildungsverhältnissen bzw. Unterschiede zwischen Männern in Frauen- und Frauen in Männerberufen erkennen: Ein hohes Eingangsinteresse am Ausbildungsberuf sowie die soziale Unterstützung der Berufswahl scheinen die Vertragslösungsneigung für junge Frauen in Männerberufen, nicht aber für junge Männer in Frauenberufen zu reduzieren. Gefördert wird die Vertragslösungsneigung der jungen Frauen in Männerberufen durch eine geringe wahrgenommene Passung zum Ausbildungsberuf und ein hohes Belastungserleben, Faktoren, die in dem Modell für die jungen Männer in Frauenberufen nicht signifikant werden. Deren Vertragslösungsneigung steht wiederum mit der Größe des Ausbildungsbetriebs, der Reflektiertheit der Berufswahlentscheidung und der Absicht, zukünftig im Ausbildungsberuf zu arbeiten, in einem gegenläufigen Zusammenhang.

Der Vortrag mündet in die Diskussion unter anderem der praktischen Implikationen der signifikanten, aber auch der nicht signifikanten Ergebnisse für die Prävention vorzeitiger Vertragslösungen in geschlechtsuntypischen Ausbildungsverhältnissen.

Die Entwicklung des normativen Commitments zum Ausbildungsbetrieb bei nicht-traditionellen Studierenden im Dualen Studium

Sebastian Rahn¹, Steffen Wild²

¹htw saar, Deutschland; ²Technische Universität Dortmund, Deutschland

Theoretischer Hintergrund

Unternehmen stehen zunehmend vor der Herausforderung, dem Fachkräftemangel und dem demografischen Wandel zu begegnen. Eine zentrale Strategie hierbei ist es, Arbeitskräfte frühzeitig zu rekrutieren und langfristig an das Unternehmen zu binden (Brunello & Wruuck, 2021). Ein möglicher Weg, um mit dieser Strategie leistungsstarke Arbeitskräfte zu rekrutieren, besteht in einem Engagement als Praxispartner in dualen Studiengängen (Graf, 2014). Der Investition in die Ausbildung von Nachwuchskräften steht hier die Chance gegenüber, den eigenen Fachkräftebedarf durch Absolvent*innen eines dualen Studiums selbstständig decken zu können.

Allerdings schließen nicht alle dual Studierenden das begonnene Studium erfolgreich ab. Als besonders vulnerable Gruppe wurden nicht-traditionelle Studierende identifiziert, definiert als beruflich Qualifizierte ohne schulische Hochschulzugangsberechtigung/ohne Abitur (Kerst & Wolter, 2022). Dieser Gruppe wurde vor dem Hintergrund der Diskussion um die Defizithypothese und der Annahme einer funktionalen Äquivalenz seit 2009 ein Hochschulstudium ermöglicht, um „Bildungsreserven“ zu rekrutieren (Kerst & Wolter, 2022; Nickel & Püttmann, 2015; Otto, 2021; Wolter & Kerst, 2015). Studien zeigen diesbezüglich, dass die Zahl der nicht-traditionellen Studierenden in Deutschland zwischen 2000 und 2020 bei den Studienanfänger:innen deutlich von 0,6 Prozent auf 3,09 Prozent angestiegen ist, insbesondere an Fachhochschulen (Nickel & Thiele, 2022), und deren Studienerfolg diskutiert wird (Dahm, 2022).

Betrachtet man diese Entwicklungen verschränkt, so stellt sich die Frage, inwieweit nicht-traditionell Studierende im dualen Studium nicht nur „akademisch“ erfolgreich sind, sondern ob sie auch eine Bindung an ihr Praxisunternehmen im Sinne eines Organisationalen Commitments entwickeln. Im Rahmen des Drei-Komponenten-Modells des Organisationalen Commitments (Meyer & Allen, 1991) gilt das normative Commitment, definiert als „internalized normative pressures to act in a way that meets organizational interests“ (Wiener, 1982, S. 418), im Vergleich zum affektiven und kalkulatorischen Commitment als untererforscht (Meyer & Parfyonova, 2010). Die wenigen empirischen Forschungsergebnisse zum normativen Commitment weisen darauf hin, dass ein Zusammenhang zwischen dem normativen Commitment und der tatsächlichen Fluktuation, der Arbeitsleistung sowie dem organisationalen Sozialverhalten besteht (Meyer et al., 2002). Andere Studien zeigen, dass normatives Commitment positiv mit beruflichem Aufstieg und respektvollem Umgang im Unternehmen zusammenhängt (Donkor, 2022).

Fragestellung

Die Entwicklung des normativen Commitments zum Ausbildungsbetrieb bei nicht-traditionellen Studierenden im dualen Studium gilt als unerforscht. Vor dem Hintergrund der empirischen Zusammenhänge von normativem Commitment und Arbeitsverhalten gilt es, Entwicklungsverläufe herauszuarbeiten sowie Einflussvariablen hierauf zu identifizieren, beispielsweise um Studienabbruch entgegenzuwirken. Im Detail werden folgende Forschungsfragen untersucht: a) existieren unterschiedliche Entwicklungskurven im normativen Commitment und b) weisen demografische Faktoren, wie Geschlecht, Alter oder die soziale Herkunft, einen Einfluss auf die Verlaufskurven auf.

Methode

Wir verwenden die erhobenen Daten aus der Panelstudie „Studienverlauf – Weichenstellungen, Erfolgskriterien und Hürden im Verlauf des Studiums an der DHBW“ (Deuer & Meyer, 2020) zwischen 2016 und 2019 mit 173 nicht-traditionellen Studierenden bei jährlichen Messungen. Das normative Commitment wurde anhand einer Skala von Felfe et al. (2002) gemessen und zeigt annehmbare Reliabilitäten auf ($\omega = .74-.84$). Die demografischen Variablen wurden von der Universitätsverwaltung zur Verfügung gestellt. Anhand eines Growth Mixture Modeling werden die Daten hin auf verschiedene Entwicklungskurven analysiert (Ram, & Grimm, 2009). Eine multinomiale logistische Regression untersucht im Anschluss die demografischen Einflussfaktoren auf die Entwicklungskurven.

Ergebnisse und ihre Bedeutung

Drei Entwicklungskurven konnte die Analyse des normativen Commitments anhand des Growth Mixture Modeling herausarbeiten ($AIC = 1015.29$; $BIC = 1043.67$; $SABIC = 1015.17$; $Entropy = .55$). Gruppe 1 (13% der Teilnehmenden) zeigt unveränderliche niedrige Werte auf. Die Gruppe 2 (47 % der Teilnehmenden) weist eine stark sinkende Verlaufskurve des normativen Commitments auf und gilt aufgrund der hohen Häufigkeitsbesetzung als besonders virulente Gruppe für künftige Forschung. Gruppe 3 (60 % der Teilnehmenden) zeigt einen kontinuierlichen hohen Score auf. Die multinomiale logistische Regression zeigt, dass männliche Studierende eher in der Gruppe 3 mit kontinuierlich hohem Commitment sind ($OR = 4.24$; $p = .008$).

Längsschnittliche Effekte von Gewissenhaftigkeit und erhaltenem Feedback auf die Passungswahrnehmung bei jungen Erwerbstätigen

Lukas Ramseier, Markus P. Neuenschwander

Pädagogische Hochschule FHNW, Schweiz

Theoretischer Hintergrund

In der Schweiz entscheiden sich zwei Drittel der Jugendlichen nach der Sekundarstufe I für eine Berufsausbildung (SKBF, 2023). Im vorliegenden Beitrag werden Prädiktoren der wahrgenommenen Person-Umwelt-Passung (Passungswahrnehmung) zu Beginn der Erwerbstätigkeit analysiert. Passungswahrnehmung geht mit erhöhter Arbeitszufriedenheit und einer geringeren Dropout-Wahrscheinlichkeit einher (Findeisen et al., 2023) und wird als Aspekt von Bildungserfolg betrachtet.

Gemäss Theory of Work Adjustment wird die Passungswahrnehmung durch Persönlichkeitsmerkmale des Individuums sowie proximale Umweltaspekte geschaffen (Dawis & Lofquist, 1984). Jugendliche, die Persönlichkeitsmerkmale aufweisen, welche in der Berufswelt gefordert sind, fügen sich besser in den Arbeitskontext ein und nehmen diesen daher als passender wahr. Ein zentrales solches Merkmal ist Gewissenhaftigkeit (Etzel & Nagy, 2016; Zimmerman, 2008). Studien zeigen Effekte der Gewissenhaftigkeit auf die Passungswahrnehmung in Bildungskontexten (Etzel & Nagy, 2016). Aufgrund der Stabilität der Gewissenhaftigkeit sind diese Effekte auch längsschnittlich erkennbar (Pawlowska et al., 2014).

Einen wichtigen Umweltaspekt bildet das Feedback von Vorgesetzten, da es dem Individuum ermöglicht, Informationen zu erhalten, die das Erreichen der eigenen Ziele unterstützen und die soziale Eingebundenheit fördert (De Stobbeleir et al., 2016; Newman et al., 2021). Feedback beeinflusst somit ebenfalls die Ausprägung der Passungswahrnehmung.

Basierend auf diesen Annahmen wurde ein Arbeitsmodell getestet, in dem längsschnittliche Effekte der Gewissenhaftigkeit in der 9. Klasse sowie des Feedbacks durch Vorgesetzte am Anfang der beruflichen Grundbildung auf die Passungswahrnehmung Jugendlicher fünf Jahre nach Austritt aus der Sekundarstufe I angenommen werden.

Methode

Die Hypothesen wurden anhand von Daten aus einem längsschnittlichen Forschungsprojekt getestet. Es wurden Daten von Welle 4 (t_1 : letztes Jahr der Sekundarstufe I), Welle 5 (t_2 : erstes Jahr der beruflichen Grundbildung) und Welle 6 (t_3 : fünf Jahre nach Abschluss der Sekundarstufe I) verwendet.

Es wurden alle Schüler:innen ausgewählt, die 1) nach der Sekundarstufe I eine Berufsausbildung begonnen hatten, 2) einen Abschluss der Sekundarstufe II erworben hatten und 3) zu t_3 erwerbstätig waren. Dies resultierte in einer Stichprobe von 230 Jugendlichen (weiblich: 54,8%; Durchschnittsalter $t_1 = 15,3$, $SD = 0,56$) aus vier Schweizer Kantonen, von denen zu allen drei Messzeitpunkten Daten vorlagen. Die Konstrukte wurden durch Fragebogen mit etablierten Instrumenten gemessen (Passungswahrnehmung: fünf Items nach Neuenschwander & Frank, 2009; Gewissenhaftigkeit: fünf Items nach Borkenau & Ostendorf, 1993; Feedback: vier Items nach Ashford & Black, 1996; Kammeyer-Mueller et al., 2011; Parker & Collins, 2010).

Zur Überprüfung der Fragestellung wurde ein Strukturgleichungsmodell in Mplus berechnet, das Passungswahrnehmung sowie Gewissenhaftigkeit zu allen drei Messzeitpunkten bzw. Feedback von Vorgesetzten zu t_2 und t_3 enthält. Die Messinvarianzen zwischen den drei Erhebungszeitpunkten wurden für alle Konstrukte bestätigt.

Ergebnisse

Das finale Modell wies zufriedenstellende Fit-Werte auf ($\chi^2(631) = 795.301$; $p < .001$; CFI = .95; RMSEA = .03; SRMR = .08). Die Passungswahrnehmung zu t_2 hatte einen signifikanten direkten Effekt auf die Passungswahrnehmung zu t_3 ($\beta = .35^{***}$). Es zeigte sich ein signifikanter indirekter Effekt der Gewissenhaftigkeit zu t_1 auf die Passungswahrnehmung zu t_3 ($\beta = .22^{***}$), vermittelt über die Gewissenhaftigkeit und die Passungswahrnehmung zu t_2 . Weiterhin zeigte sich ein signifikanter indirekter Effekt des Feedbacks durch Vorgesetzte zu t_2 auf die Passungswahrnehmung zu t_3 ($\beta = .14^{**}$), vermittelt über die Passungswahrnehmung zu t_2 sowie das Feedback zu t_3 .

Bedeutung

Die Ergebnisse bestätigen Annahmen der Theory of Work Adjustment. Für Jugendliche bedeutet dies, dass Gewissenhaftigkeit in der Sekundarstufe I dem Finden einer passenden Ausbildung nach der Sekundarstufe I und langfristig nach der Sekundarstufe II zuträglich ist, da sie hilft, die verlangten Aufgaben zu bearbeiten und sich in den Arbeitskontext einzufügen. Durch ihre hohe Stabilität wirkt sich die Gewissenhaftigkeit auch nach zwei Bildungsübergängen auf die Passungswahrnehmung aus. Auszubildende im Berufsbildungskontext wiederum können durch konstruktives Feedback zu einer als passend wahrgenommenen Arbeitsumwelt in der späteren Erwerbstätigkeit beitragen.

PS-05: Sprache im Unterricht: Förderung, Unterrichtspraxis und Lehrerperspektiven

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 162

Chair der Sitzung: Kristin Kersten

Language-sensitive and cognitively activating teaching techniques predict language competence and buffer social effects

Kristin Kersten

Universität Mannheim, Deutschland

In Germany, educational attainment and language skills in both academic language ('Bildungssprache') and foreign languages are strongly affected by social differences (Klieme et al., 2006). However, research suggests that the quality of teachers' input-scaffolding techniques can reduce these differences (Helmke, 2021).

According to the Cognitive-Interactionist Framework (Ellis & Shintani, 2014; Long, 2015), 'input' is a crucial predictor of language acquisition (Gass et al., 2020). Numerous studies identified characteristics of 'input quality' that positively predict first and second/foreign language competence (overview: Kersten et al., under review). Its operationalization, however, is challenging. In their Modular Cognition Framework, Truscott & Sharwood Smith (2019), define 'input' comprehensively as all external sensory stimuli that induce intake, processing and storing, thus contributing to language comprehension and production (Kersten, 2021, 2023).

Based on this definition, the *Teacher Input Observation Scheme* (TIOS, Kersten et al., 2018) operationalizes input with 41 instructional techniques across four scales: *Cognitively Stimulating Tasks*, *Verbal Input*, *Nonverbal Input*, and *Support of Learners' Output* (interrater reliability: Krippendorff's $\alpha=.88^*$). They describe a distinct, observable communicative behavior or activity in the classroom in direct contact with the learner/s, and are compatible with approaches to language-sensitive teaching (Gogolin et al., 2020; Rossner & Bolitho, 2022).

Research questions.

1. Does the use of TIOS-techniques predict learners' language proficiency in German and English?
2. Does teachers' total TIOS-score moderate the effect of socio-economic status (SES) on learners' language proficiency in both languages?
3. Do empirically derived TIOS-factors predict learners' language proficiency in both languages differentially?

Data was elicited online (SoSciSurvey) from $N=131$ teachers, grades 4-6, in two strands: $n=37$ teachers of German-language subjects ($n=836$ learners), and $n=94$ English teachers ($n=1935$ learners). Techniques were operationalized using TIOS items, learners' language proficiency using CEFR and *Lingualevel* descriptors (Mettler et al., 2007) (A1.1-C1). Teachers also reported learners' multilingualism status and average class SES.

Results. 1. Separate regression analyses controlling for SES and multilingualism showed significant effects of the total TIOS-score, *Stimulating Tasks* (Scales 1), *Verbal Input* (Scale 2), and SES on German proficiency; multilingualism had a negative effect. For English, no effects were found for the full sample, but in the subsample of grade 4 ($n=39$ teachers, $n=690$ learners), the total score, *Stimulating Tasks* and SES predicted English proficiency.

2. Two moderation analyses yielded buffering effects of the total scores on learners' proficiency in both languages.

3. A principal-axis factor analysis (oblimin-rotation) yielded five factors in the English sample (36.36% of variance) that were labeled regarding their content. Given our cross-sectional sample from grade 4 (low English proficiency) to grade 6 (more advanced), we expected differential effects for factors hypothesized to be used with novice learners (factor: *Input Scaffolding*) and with more advanced learners (*Focus on Meaning*, *Cognitive Stimulation*, *Linguistic Input*), with no hypothesis for *Output Scaffolding*. Due to the lack of directionality in cross-sectional designs, negative effects suggest an association with low proficiency. A linear mixed effect regression with classes as a random-effects variable (level 2) yielded significant predictions of *Focus on Meaning* ($p=.032$), *Cognitive Stimulation* (marginal, $p=.053$), and *Early Scaffolding* (negative, $p=.006$) ($R^2_{\text{marginal}}=.219$, $R^2_{\text{conditional}}=.737$). No effects were found for the German sample.

These results highlight the importance of clearly operationalizing input quality to measure effects of language-sensitive techniques both for majority and foreign languages. They underline their potential to mitigate social effects on language competences.

Negative cross-sectional effects of *Early Scaffolding* suggest that teachers use differential techniques for novice vs. advanced learners, supporting models of "responsibility transfer" (Conceptual Model of Scaffolding, van de Pol et al., 2010; CAPA, Tedick & Lyster, 2020). Positive effects demonstrate the importance of techniques activating cognitive involvement and meaningful classroom interactions. The findings call for corroboration in longitudinal analyses.

Unpacking Divergences in Classroom Talk: How Teacher Demands Shape Learner Responses in Linguistically Diverse Classrooms

Simay Birce Cirit

Universität Kassel, Deutschland

Theoretical Background and Research Question

This study is rooted in the sociocultural theory (Vygotsky 1978, 1981), which highlights the social and communicative nature of learning. The theoretical framework also incorporates Brousseau's (1997, 2014) 'didactic contract', which refers to implicit expectations, rules, and norms between teachers and students. Didactic contract views learning as a modification of learners' knowledge. In any "learning situation," learners' initial answers are typically insufficient in terms of what the teacher intends to teach (Brousseau 1997: 228), which constitute so-called divergences in teacher-learner interactions. In this study, divergence is defined as a discrepancy between teachers' verbal demands, which are typically conveyed through instructions or questions, and the responses provided by learners. This insufficiency prompts learners to modify their understanding to cope with a new situation, guided by teachers. The deeper these modifications, the more valuable the learning process, as students adapt to the constraints of the environment rather than merely meeting teachers' expectations (Brousseau, 1997). As part of an ongoing PhD project, this study will focus on one aspect of a larger study, investigating the following question: How do divergences in whole-class interactions vary according to teachers' verbal demands?

Methods

This study is part of the larger video study INTERFACH (<https://interfach.de/>), which was conducted in primary school contexts and financed by the DFG. This exploratory qualitative study is based on data from 13 mathematics lessons in five 3rd grade classrooms with high proportions of emerging multilingual students (53%-87% per classroom). Data were collected through video-based classroom observations and audio recording. Qualitative content analysis was employed, beginning with the transcription and coding of whole-class interactions, to identify divergent moments. Initially, a basic coding scheme was developed inductively and deductively to distinguish divergent interactions. Subsequently, a detailed coding scheme was developed to categorize different types of divergences, which were validated through pilot testing and inter-coder agreement checks.

Preliminary Results

Preliminary findings indicate that different types of divergence occur in whole-class interaction, one of which is a discursive/linguistic mismatch. This mismatch occurs when learners fail to recognize or fulfill the linguistic activity in teacher demand. As a result, the provided response does not match the teacher's expectation, which is observed when teachers try to guide learners toward their demands by making adjustments in response to the divergence. These mismatches were often triggered by unclear or multiple simultaneous teacher demands. These findings highlight the critical importance of clear and precise teacher demands. Furthermore, teacher demands that require discursive activity (e.g., explaining and justifying) are regarded as high cognitive demands that foster higher thinking and cognitive skills (Prediger & Neugebauer, 2021; Quasthoff et al., 2022). Being precise and specific while engaging learners with these high cognitive and linguistic demands benefits all children; however, they are particularly essential for emerging multilingual learners to ensure what they are being asked (Gibbons, 2015) (referred as "comprehensible input" by Krashen, 1982). These findings support Gibbons' (2015) assertion that clarity in instructional language is essential for effective subject and language learning in linguistically diverse classrooms.

Überzeugungen angehender Lehrkräfte zur Bedeutung von Sprache im Fachunterricht

Dorothee Kohl-Dietrich¹, Axel Zinkernagel², Rebecca Zerwas¹

¹PH Karlsruhe, Deutschland; ²Bergische Universität Wuppertal

Theoretischer Hintergrund und Fragestellung

Die sprachliche Heterogenität an Schulen und damit verbunden das Postulat einer integrierten Sprachbildung und eines wertschätzenden, ressourcenorientierten Umgangs (Gogolin 2021) mit Mehrsprachigkeit im Unterricht stellen besondere Anforderungen an zukünftige Lehrkräfte, Unterricht sprach- und diversitätssensibel zu planen und zu gestalten. Ausgehend von der Annahme, dass nicht nur Kompetenzen in Bezug auf Deutsch als Zweitsprache (DaZ) und Mehrsprachigkeit (vgl. DaZKom-Modell, Ohm 2018) sondern auch Einstellungen von Lehrkräften ihre Wahrnehmung von und ihren Umgang mit Mehrsprachigkeit im Unterricht beeinflussen können, untersuchen die hier präsentierten Studien, ob Einstellungen (*beliefs*) zu Mehrsprachigkeit durch universitäre Lehrveranstaltungen verändert werden können. Dabei wird angenommen, dass *beliefs* keine stabilen Merkmale im Sinne von *traits* sind, sondern durch Kompetenzerwerb in relevanten Lerngelegenheiten verändert werden können (Fischer, Hammer & Ehmke 2018).

Methode

Das Paper stellt zwei Studien vor. Die erste Studie ($N = 75$) ist eine Studie mit Pre-Post-Test-Design, die im Rahmen von drei themengleichen Seminaren zu sprachsensibler Unterrichtsgestaltung (Erhebungszeitraum 2020/2021) in den Bildungswissenschaften der Lehrämter Realschule Plus und Gymnasium durchgeführt wurde. Als Erhebungsinstrument wurde der Fragebogen von Fischer, Hammer & Ehmke (2018) zu Überzeugungen von Lehrkräften zu Sprache im Fachunterricht eingesetzt sowie der Fragebogen zur Bedeutung von Lerngelegenheiten für den Erwerb von DaZ-Kompetenz (Lemmrich & Ehmke 2018), die im Rahmen des DaZKom Projektes entwickelt wurden. Das validierte Messinstrument zu Überzeugungen (Fischer, Hammer & Ehmke 2018) enthält drei Subskalen, die sich an den Dimensionen des DaZKom Strukturmodells (Ohm 2018) orientieren: 1) Sprachsensibilität im Fachunterricht, 2) Zuständigkeit für Sprachförderung, 3) Wertschätzung von Herkunftssprachen. Beim Messinstrument zu DaZ-bezogenen Lerngelegenheiten wurden die Skalen zu Bereichen der (Angewandten) Linguistik sowie zu DaZ-bezogenen Handlungen verwendet. Die Daten wurden im Rahmen des DaZKom-Transfer Projektes (Koch-Priewe, Köker, Daase, Ohm, Ehmke) erhoben.

Die zweite Studie ($N = 144$) ist eine querschnittlich angelegte Studie, in der die Überzeugungen von angehenden Lehrkräften in unterschiedlichen Phasen des Studiums (zu Beginn des Bachelorstudiums, am Ende des Bachelorstudiums, am Ende des Masterstudiums, Erhebungszeitraum 2023) erhoben wurden. Die Stichprobe besteht aus Studierenden des Grundschullehramtes mit dem Studienfach Sachunterricht.

Die Studie untersucht folgende Forschungsfragen:

RQ1: Verändern sich Überzeugungen von angehenden Lehrkräften zu Sprache im Fach vor und nach dem Besuch eines Seminars zu sprachsensibler Unterrichtsgestaltung? (Studie 1)

RQ2: Verändert sich die Wahrnehmung zu sprachbezogenen Lerngelegenheiten vor und nach dem Besuch eines Seminars zu sprachsensibler Unterrichtsgestaltung? (Studie 1)

RQ3: Unterscheiden sich Überzeugungen von Lehramtsstudierenden mit dem Fach Sachunterricht signifikant voneinander an unterschiedlichen Stationen ihres Studiums? (Studie 2)

RQ4: Gibt es statistische Zusammenhänge zwischen den Überzeugungen von Lehramtsstudierenden des Sachunterrichts und wahrgenommenen Lerngelegenheiten? (Studie 2)

RQ5: Unterscheiden sich Lehramtsstudierende des Sachunterrichts mit einem naturwissenschaftlichen Schwerpunkt von Lehramtsstudierenden des Sachunterrichts mit einem sozialwissenschaftlichen Schwerpunkt im Hinblick auf ihre Überzeugungen zu Sprache im Fachunterricht? (Studie 2)

Ausgewählte Ergebnisse:

Die im Rahmen von Studie 1 eingesetzten Messinstrumente zeigen gute bis akzeptable Werte für die interne Konsistenz der Skalen (Subskala 1: Cronbachs $\alpha = .51$ (Prä-Test), Cronbachs $\alpha = .68$ (Post-Test); Subskala 2: Cronbachs $\alpha = .77$ (Prä-Test), Cronbachs $\alpha = .79$ (Post-Test); Subskala 3: Cronbachs $\alpha = .64$ (Prä-Test), Cronbachs $\alpha = .62$ (Post-Test)).

Zu Forschungsfrage 1 (RQ1) kann festgehalten werden, dass in Bezug auf 1) Sprachsensibilität im Fachunterricht ($p = .01$), 2) Zuständigkeit für Sprachförderung ($p = .003$) und 3) Wertschätzung der Herkunftssprachen ein statistisch signifikanter Zuwachs stattgefunden hat ($p = .04$). Allerdings ist der Zuwachs gering und unter dem als sprachsensibel definierten Wert (vgl. Fischer, Hammer & Ehmke 2018).

Erste Auswertungen zu Forschungsfrage 4 (RQ4) deuten darauf hin, dass z.B. die Wertschätzung von Herkunftssprachen in Abhängigkeit von Lerngelegenheiten signifikante Unterschiede aufweist.

Der Vortrag versucht im Sinne evidenzbasierter Lehre auf der Grundlage der Ergebnisse beider Studien Schlussfolgerungen zu ziehen, inwiefern der Umfang und die Qualität von Lerngelegenheiten im Studium Überzeugungen angehender Lehrkräfte beeinflussen.

Die VERA-BiSS-Fortbildung zur datengestützten Unterrichtsentwicklung im Kompetenzbereich Lesen: Wie profitieren Lehrkräfte in den Bereichen Wissen, Einstellungen, Selbstwirksamkeit und professionelles Handeln?

Daria Ferencik-Lehmkuhl¹, Carola Schnitzler², Charlotte Stehr¹, Sofie Henschel², Jörg Jost¹

¹Universität zu Köln, Deutschland; ²Insitut zur Qualitätsentwicklung im Bildungswesen (IQB), Berlin, Deutschland

Aktuelle Ergebnisse aus Schulleistungsstudien, z. B. dem IQB-Bildungstrend (Stanat et al., 2023), weisen auf einen Rückgang der Lesekompetenz von Schüler:innen hin. Dies unterstreicht die Notwendigkeit einer gezielten und systematischen schulischen Leseförderung, die darauf ausgerichtet ist, dass Schüler:innen die in den bundesweiten Bildungsstandards festgelegten Anforderungen für den Kompetenzbereich Lesen (KMK, 2018) erfüllen. Diesen Herausforderungen kann mit einer auf Daten basierenden Entwicklung der Unterrichtsqualität begegnet werden (Mandinach & Schildkamp, 2021). So können Deutschlehrkräfte mit Ergebnissen aus standardisierten diagnostischen Verfahren wie den Vergleichsarbeiten (VERA) ihren bisherigen Leseunterricht reflektieren und den zukünftigen so planen, dass er an die Lernvoraussetzungen und -bedarfe der Schüler:innen ihrer Klasse möglichst optimal anschließt.

Eine Nutzung der VERA-Ergebnisse zur datengestützten Unterrichtsentwicklung findet allerdings häufig nicht oder nur wenig systematisch statt (Nachtigall & Jantowski, 2007). Das liegt neben mangelnden zeitlichen und personellen Ressourcen auch daran, dass Lehrkräfte diagnostische Daten als wenig nützlich wahrnehmen (Groß Ophoff, 2013). Sie verfügen zudem oftmals nicht über eine ausreichende Datenkompetenz im Umgang mit Testergebnissen (Koch, 2011). Als Konsequenz lässt sich ein deutlicher Unterstützungsbedarf von Lehrkräften feststellen, beispielsweise in Form von Fortbildungen (Henschel & Stanat, 2019; McElvany et al., 2023). Deshalb wurde im Rahmen des Forschungsvorhabens VERA-BiSS ein neuer Blended-Learning-Fortbildungsansatz entwickelt, der darauf abzielt, Deutschlehrkräfte der Sekundarstufe I dabei zu unterstützen, die VERA-8-Ergebnisse effektiv für die Weiterentwicklung ihres Leseunterrichts zu nutzen.

Der Fortbildungsansatz wurde mit 84 Deutschlehrkräften von 32 Schulen in einem quasi-experimentellen Design bezüglich seiner Wirksamkeit evaluiert (vgl. Lipowsky & Rzejak, 2019). Nur Lehrkräfte der Experimentalgruppe ($N=61$, 85.25% weiblich, $M=40.38$ Jahre, $SD=9.15$) nahmen an der Fortbildung teil, die übrigen Lehrkräfte ($N=23$, 61.90% weiblich, $M=37.65$ Jahre, $SD=8.77$) bildeten die Wartekontrollgruppe. Der Beitrag berichtet Effekte der VERA-BiSS-Fortbildung auf die professionelle Kompetenz von Lehrkräften (Baumert & Kunter, 2006) bezogen auf VERA und die Nutzung der VERA-Ergebnisse basierend auf Pre-Post-Befragungen in den Bereichen Wissen (Selbsteinschätzung: 6 Items, $\alpha_{\text{Pretest}}=.89$; $\alpha_{\text{Posttest}}=.90$), Überzeugungen (Nützlichkeitswahrnehmung: 6 Items, $\alpha_{\text{Pretest}}=.91$; $\alpha_{\text{Posttest}}=.88$) und motivationale Orientierungen (Selbstwirksamkeit: 10 Items, $\alpha_{\text{Pretest}}=.84$; $\alpha_{\text{Posttest}}=.93$). Außerdem wurden Auswirkungen auf das professionelle Handeln (Nutzung auf Klassenebene: 6 Items, $\alpha_{\text{Pretest}}=.89$; $\alpha_{\text{Posttest}}=.84$) untersucht. Berechnet wurden Mehrebenenmodelle (MLM) mit festen Effekten der Zeit (pre/post) und Gruppe (Experimental-/Kontrollgruppe) und deren Interaktion bei Berücksichtigung der hierarchischen Datenstruktur (Zeit genestet in Lehrkräften) und unter Kontrolle schulischer Kontextfaktoren (wahrgenommene Unterstützung durch die Schulleitung und kollegiale Kooperation).

Die Interaktion zwischen Zeit und Gruppe war für alle vier untersuchten Variablen signifikant, was auf eine positive Entwicklung in den hier untersuchten Facetten der professionellen Kompetenz bei den Lehrkräften in der Experimentalgruppe spricht. Dasselbe galt für Veränderungen bei der Nutzung der VERA-Ergebnisse auf Klassenebene. Gleichzeitig ließ sich jedoch auch eine negative Entwicklung in der untrainierten Kontrollgruppe beobachten. Die Ergebnisse werden bezugnehmend auf bereits vorliegende Ergebnisse zur Wirksamkeit von Fortbildungsmaßnahmen zur datengestützten Unterrichtsentwicklung diskutiert.

PS-06: Professionelles Wissen und Handeln von Lehrkräften

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 242

Chair der Sitzung: Birte Oetjen

ENTFÄLLT: Psychometrische Qualität des TEDS-M Kurztests zur Erfassung des pädagogischen Unterrichtswissens

Stephan Abele, Diana Vogel-Blaschka, Beck Marlen, Bärbel Fürstenau

Technische Universität Dresden, Deutschland

Das pädagogische Unterrichtswissen (PUW) gilt als wichtiges Merkmal professioneller (angehender) Lehrkräfte (Voss et al., 2015). Es beinhaltet das, was (angehende) Lehrkräfte aller Lehrämter über das Unterrichten wissen sollten, und zwar unabhängig vom Fach, das sie unterrichten (König & Blömeke, 2009). Zur Messung des PUW liegen mehrere Skalen bzw. Tests vor (z. B. Hohenstein et al., 2017; Kunina-Habenicht et al., 2020; Müser et al., 2023; Voss et al., 2011). Eine Besonderheit des TEDS-M Kurztests ist, dass er mit 20 min eine geringe Testzeit erfordert (König & Blömeke, 2010), was im Hinblick auf die Testmüdigkeit und Testteilnahmequote günstig ist. Einiges, was wir über die Effekte und Entwicklung des PUW wissen, basiert auf diesem Kurztest (z. B. Heine et al., 2023; König, Darge, et al., 2020; König et al., 2018). Im Beitrag thematisieren wir ausgewählte psychometrische Qualitätsaspekte des Kurztests.

Der Kurztest umfasst Aufgaben und Items. Bisherige Studien gingen (implizit) davon aus, dass die psychometrische Qualität des Tests auf Item- und nicht auf Aufgabenebene analysiert werden sollte (z. B. Tachtsoglou & König, 2017). Jede Aufgabe beinhaltet einen Stimulus (unterrichtsrelevante Situation) und meist mehrere Items. Zwischen den Items einer Aufgabe bestehen aufgrund des gemeinsamen Stimulus' möglicherweise lokale stochastische Abhängigkeiten, weshalb bisherige Befunde zur psychometrischen Qualität verzerrt sein könnten (Frey et al., 2016).

Keine Befunde liegen dazu vor, ob der Kurztest ein- oder mehrdimensional misst. Obwohl er mehrere Inhaltsdimensionen umfasst (Umgang mit Heterogenität, Strukturierung, Klassenführung, Motivierung, Leistungsbeurteilung), vermuten König und Blömeke (2010), dass er eindimensional misst, da er wenige Items beinhaltet. Voss et al. (2011) erfassten das PUW jedoch mit weniger Items und fanden heraus, dass ihr Test mehrdimensional misst. Schließlich unterstellen König et al. (2020) (implizit), dass der Kurztest das PUW mehrdimensional erfasst: Mit Rückgriff auf den Test untersuchten sie die Effekte des PUW auf die Unterrichtsplanung und differenzierten dabei nach PUW-Dimensionen.

Fraglich ist also, ob die psychometrische Qualität des Kurztests auf der Aufgaben- oder Itemebene analysiert werden soll und ob er ein- oder mehrdimensional misst. In Anlehnung an den Forschungsstand nehmen wir an, dass eine Analyse auf Aufgabenebene empirisch gerechtfertigt ist (H01) und er mehrdimensional misst (H02).

Geprüft wurden die Hypothesen mit einer Stichprobe von 381 Studierenden des beruflichen Lehramts der TU Dresden. 213 der Studierenden studierten 1 bis 3 Semester, 168 studierten mindestens 4 Semester. Der Kurztest war Teil einer 45-minütigen Erhebung, die zunächst in Präsenzveranstaltungen, dann aufgrund der Pandemie online stattfand. Bei der Datenanalyse wurden konfirmatorische Faktorenanalysen durchgeführt, verschiedene Modelle gebildet und unter gängigen Fitwerten miteinander verglichen. Verglichen wurde z.B. ein Modell, das über die Testaufgaben gebildet wurde, mit einem Modell, das über Testitems gebildet wurde.

Entgegen unserer Erwartung (H01) zeigte sich, dass das „Aufgaben-Modell“ besser auf die Daten passt als das „Item-Modell“. Erwartungswidrig stellten wir ferner fest (H02), dass der Test eindimensional, nicht aber mehrdimensional misst. Die Aufgabenschwierigkeiten (P) lagen zwischen .09 (Lösungsquote 9 %) und .70, die Aufgabentrennschärfen (r_a) zwischen -.04 und .40; zwei Trennschärfen waren $< .25$. Die interne Konsistenz der eindimensionalen Aufgabenskala lag bei $\alpha = .72$ und war damit geringer als in früheren Studien (Tachtsoglou & König, 2017).

Unsere Studie legt nahe, dass das PUW mit dem TEDS-M Kurztest in zufriedenstellender psychometrischer Qualität gemessen werden kann, sofern mit Aufgaben und eindimensional gemessen wird. Es ist denkbar, dass die Reliabilität des Kurztests in früheren Studien überschätzt wurde, da diese über Items ermittelt wurde. Angesichts unsere Ergebnisse raten wir eher davon ab, den Kurztest für dimensionsspezifische Effekt- oder Entwicklungsanalysen zu verwenden. Fraglich ist, ob unsere Ergebnisse verzerrt sind, weil sie auf Studierenden des Lehramts an beruflichen Schulen beruhen. Dagegen spricht, dass der Kurztest so konzipiert ist, dass er das PUW in allen Lehrämtern reliabel und valide erfassen soll.

Ein systematisches Review zur Konzeptualisierung professionellen Wissens von Grundschullehrkräften

Birte Oetjen, Saskia Liebner, Leonora Gerbeshi, Sophia Römer, Lotta Bärtlein, Victoria Wiederseiner, Rebecca

Baumann, Selma Cejvan, Petra Ackerlauer, Sabine Martschinke

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Theoretischer Hintergrund und Fragestellungen

Für erfolgreiches professionelles Handeln und um Kindern eine bestmögliche Lern- und Persönlichkeitsentwicklung in ihrer Grundschulzeit zu ermöglichen, benötigen Grundschullehrkräfte u.a. fachliche sowie vor dem Hintergrund aktueller gesellschaftlicher Herausforderungen (z.B. Digitalität, Inklusion, zunehmender Extremismus, Klimakrise) fachübergreifende professionelle Kompetenzen. Trotz verschiedener theoretischer Ansätze zur professionellen Kompetenz von Lehrkräften hat sich im deutschsprachigen Raum das Kompetenzmodell COACTIV (Baumert & Kunter, 2006) etabliert, in welchem das Professionswissen als zentral beschrieben wird. Dabei wird im COACTIV-Modell, das ursprünglich bei (angehenden) Sekundarschullehrkräften im Fach Mathematik empirisch geprüft wurde, das Professionswissen in pädagogisch-psychologisches, fachdidaktisches und fachliches Wissen sowie Beratungswissen und Organisationswissen differenziert. Es fehlen bislang allerdings systematische Überblicksarbeiten, die das Professionswissen von Grundschullehrkräften konzeptualisieren. Gerade hinsichtlich spezifischer Anforderungen der Grundschule bleibt unklar, welche Rolle das COACTIV-Modell für die Konzeptualisierung von Professionswissen spielt, wie sich das Professionswissen von Grundschullehrkräften in den Diskurs um verschiedene Fachbezüge und Wissensdimensionen einordnen lässt, inwiefern es in den unterschiedlichen Lehrer:innenbildungsphasen berücksichtigt wird und inwiefern Professionswissen für fachübergreifende Herausforderungen adressiert wird. Diesen Fragestellungen geht der vorliegende Beitrag nach.

Methodisches Vorgehen

Im Projekt "Professionelle Kompetenzen in der Grundschule" (Profi-G) wurde für die Beantwortung der Forschungsfragen ein systematisches Review (Zawacki-Richter, 2020) umgesetzt. Die Datenbanken BASE, PSYINDEX, JSOTR und FISBildung wurden im Juni 2023 mit dem Suchstring "Wissen AND Profession* OR Pädagog* OR Fach* OR Fachdidakt* AND Grundschul* OR

Primar* AND Lehr* OR Ref** durchsucht. Aufgrund des Bezugs zum COACTIV-Modell wurden im Review ausschließlich deutschsprachige Publikationen seit 2006 berücksichtigt. Das Review wurde entsprechend des PRISMA Flow Diagramms (Page et al., 2021) anhand aufeinanderfolgender Screeningschritte mit EPPI Reviewer 6 durchgeführt. Einschlusskriterien waren der Fokus auf (angehende) Grundschullehrkräfte sowie deren professionelles Wissen. Bei der Datenbankrecherche wurden insgesamt 10.471 Treffer identifiziert, nach Ausschluss von Duplikaten verblieben 7.361 Beiträge für den Screeningprozess. Im ersten Schritt konnten 3.037 Publikationen anhand der Publikationstitel eingeschlossen werden, welche in einem zweiten Schritt auf Grundlage des Abstracts gesichtet wurden. Davon verblieben 749 Publikationen für die Recherche und Sichtung der Volltexte, von denen weitere 570 aufgrund fehlenden Bezugs zum professionellen Wissen von Grundschullehrkräften exkludiert wurden. Nach Abschluss der Screeningschritte wurden 179 Publikationen für die inhaltliche Auswertung zum professionellen Wissen von Grundschullehrkräften identifiziert. Diese Publikationen wurden anhand eines deduktiv-induktiv entwickelten Kategoriensystems (Kuckartz, 2016) systematisiert: Schulart (Grundschule; schulartübergreifend), Kompetenzmodell (COACTIV; anderes Modell), Wissensdimension (z.B. pädagogisch-psychologisches Wissen; fachdidaktisches Wissen; Fachwissen), Unterrichtsfach (z.B. Mathematik; Deutsch; Sachunterricht), fachübergreifende Anforderungen (z.B. Inklusion; BNE; Demokratiebildung; Digitalität), Professionalisierungsphase (z.B. 1., 2., 3. Phase). Für ausgewählte Kategorien waren Mehrfachcodierungen möglich.

Ergebnisse und Diskussion

Von 179 Publikationen beziehen sich 28 Publikationen auch auf weiterführende Schularten. Als theoretische Rahmung wird in den Publikationen überwiegend das COACTIV-Modell genutzt ($n=122$). Vorrangig beziehen sie sich auf die erste und dritte Phase der Lehrer:innenbildung ($n=58$ bzw. $n=122$), nicht aber auf die zweite Phase ($n=10$). Bezüglich der verschiedenen Wissensdimensionen werden vor allem fachdidaktisches Wissen ($n=129$) und Fachwissen ($n=114$) adressiert, weniger pädagogisch-psychologisches Wissen ($n=75$). Dabei stehen insbesondere die Unterrichtsfächer Sachunterricht ($n=59$), Mathematik ($n=50$) und Deutsch ($n=25$) im Fokus. Professionelles Wissen zu fachübergreifenden Anforderungen wie Inklusion ($n=13$), Digitalität ($n=6$) und BNE ($n=3$) wird hingegen nur in sehr wenigen Publikationen thematisiert. Es konnte keine Publikation zum professionellen Wissen bezüglich Demokratiebildung in der Grundschule identifiziert werden. Zusätzlich werden im Vortrag die verschiedenen Kategorien miteinander in Verbindung gesetzt.

Der Beitrag zeigt insgesamt, dass die Forschung um professionelles Wissen in der Grundschule noch immer ein Desiderat darstellt. Insbesondere zur für die Professionalisierung prägenden zweiten Phase der Lehrer:innenbildung (Kosinar, 2014) finden sich kaum Publikationen zum professionellen Wissen. Auch überwiegen Publikationen zum fachdidaktischen Wissen und Fachwissen im Fächerkanon der Grundschule. Gerade Wissen zu fachübergreifenden An- und Herausforderungen, die Gesellschaft und deren Zusammenhalt nachhaltig prägen, muss im zukünftigen Diskurs um professionelle Kompetenzen von Grundschullehrkräften stärker berücksichtigt werden.

Textvignetten zu Vorbereitung auf komplexe Herausforderungen der Schulpraxis? Eine Längsschnittanalyse zur Entwicklung der Selbstwirksamkeit von Lehramtsstudierenden

Simon Wagner¹, Karsten Krauskopf², Michel Knigge¹

¹Universität Potsdam, Deutschland; ²Fachhochschule Potsdam, Deutschland

Selbstwirksamkeit im Kontext der Lehrkräftebildung beschreibt die Überzeugung, schulbezogene Anforderungssituationen aufgrund eigener Fähigkeiten erfolgreich bewältigen zu können (Täschner et al., 2023; Tschannen-Moran et al., 1998). Da sich Unterrichtspraxis durch pädagogisch herausfordernde Anforderungssituationen auszeichnet (Krauskopf et al., 2020), ist Selbstwirksamkeit äußerst bedeutsam und wird als Facette professioneller Handlungskompetenz modelliert (Baumert & Kunter, 2013). Für Menschen mit wenig Erfahrung identifizierte Bandura (1997) stellvertretende Erfahrungen als Quelle der Selbstwirksamkeitsentwicklung. Bei Übertragung des Ansatzes auf Noviz*innen im Lehrberuf wird deutlich, dass insbesondere stellvertretende Erfahrungen, durch reale (in Praxisphasen) oder symbolische Modelle (in Fallvignetten), relevant sind. D. h. die Beobachtung von Modellen, die kompetente Handlungen (masterly) zeigen, oder die bei der Überwindung von Hürden (coping) beobachtet werden, ist entscheidend (Bach, 2022).

Symbolisch Modelle in Form von Video- bzw. Textvignetten stellen eine Möglichkeit dar, Lehramtsstudierenden solche stellvertretenden Erfahrungen bereitzustellen. Während die Analyse von Videovignetten im Zusammenhang mit einer Zunahme der Selbstwirksamkeit steht (Gold et al., 2017; Kumschick et al., 2017; Schlosser & Paetsch, 2023; Thiel et al., 2020), ist der Forschungsstand für Textvignetten unklar (Schlosser & Paetsch, 2023; Wagner et al., 2024). Es liegen nur wenige Studien zur Vorhersage der Selbstwirksamkeit von Lehramtsstudierenden durch stellvertretende Erfahrungen während Schulpraktika vor (Bach, 2022; Clark & Newberry, 2019; Pfitzner-Eden, 2016), insbesondere fehlen Längsschnittstudien.

Entsprechend stellt sich die Frage, inwiefern der Einsatz von Textvignetten Studierende bei der Vorbereitung auf Praktika in Bezug auf ihre Selbstwirksamkeitsentwicklung unterstützen kann. In einem quasi-experimentellen Prä-Post-Follow-Up-Kontrollgruppendesign wird untersucht, wie sich die Selbstwirksamkeit von Lehramtsstudierenden im Semesterverlauf inklusive eines Praktikums entwickelt, wenn Textvignetten (versus akademische Fachtexte) als Vorbereitung auf ein Praktikum bereitgestellt werden.

Die Textvignetten wurden in ein Vorbereitungsseminar für ein 30-stündiges Praktikum integriert. Die Stichprobe umfasst $N = 443$ Lehramtsstudierende Sekundarstufe I/II (61% weiblich, $M = 22.23$ Jahre, $SD = 3.61$, $M = 4.60$ Fachsemester, $SD = 1.65$), die sich - ohne vorherige Kenntnis des Formats - entweder für ein vignettenbasiertes ($n = 10$) oder textbasiertes ($n = 24$) Begleitseminar angemeldet hatten. In der Gruppe *Vignette* (IG, $n = 94$) wurden drei Textvignetten sowohl als Masterly als auch als Coping Modell bereitgestellt, während in der Gruppe *Fachtext* (KG, $n = 349$) inhaltlich kongruente Fachtexte verwendet wurden. Zu drei Messzeitpunkten (Vorlesungsbeginn t_1 , Vorlesungsende t_2 , nach Praktikumsabsolvierung t_3) wurden Onlinefragebögen zur Datenerhebung verwendet. Während zu t_1 vollständige Daten vorlagen, fehlten zu t_2 15.35% und zu t_3 33.84% der Daten. Die Lehrkräfteselbstwirksamkeit wurde mithilfe der Selbsteinschätzungsskala von Schmitz und Schwarzer (2000) erfasst ($\alpha_{t_1} = .74$). Zur Analyse der Längsschnittdaten wurden Latent-Change-Score-Modelle (McArdle, 2009) unter Verwendung von R (R Core Team, 2022) und dem Paket lavaan (Rosseel, 2012) berechnet. Voranalysen deuteten auf eine starke Messinvarianz für die messwiederholten Konstrukte hin. Fehlende Werte wurden mithilfe von FIML geschätzt.

Die Ergebnisse zeigen eine kurvenförmige Veränderung der Selbstwirksamkeit bei Studierenden der Interventionsgruppe. Die Teilnahme an der Intervention war zwischen t_1 und t_2 negativ mit der Veränderung der Selbstwirksamkeit verbunden ($\beta = -.12$, $p = .045$), aber zwischen t_2 und t_3 positiv ($\beta = .19$, $p = .007$). Dies bedeutet, dass die Selbstwirksamkeit der Studierenden in der Interventionsgruppe erst nach Abschluss des Praktikums signifikant anstieg, während sie in der Kontrollgruppe abnahm.

Zusammenfassend deuten die Ergebnisse darauf hin, dass Textvignetten ein wirksames Instrument zur Förderung der Selbstwirksamkeitsentwicklung angehender Lehrkräfte sein können, wenn sie vorbereitend auf praktische Erfahrungen eingesetzt werden. Dabei dienen sie ggf. als Vorwegnahme eines Praxis-Schocks (Veenman, 1984). Die ausschließliche Verwendung von Textvignetten ohne praktische Erfahrung ist hingegen kritisch zu betrachten, ggf. müssten differenzierte Effekte von Mastery und Coping Modellen näher untersucht werden (Wagner et al., 2024). Die Ergebnisse zeigen auch, dass akademische Fachtexte keine adäquaten stellvertretenden Erfahrungen bieten. Zukünftige Forschung sollte daher das Zusammenspiel von Vignetten und praktischen Erfahrungen für die Bedeutung der Selbstwirksamkeitsentwicklung näher untersuchen.

Antinomien im professionellen Handeln von Lehrpersonen

Elisabeth Maué¹, Tobias Kärner², Michael Goller³, Hakan Peker⁴

¹Universität Konstanz, Deutschland; ²Universität Hohenheim, Deutschland; ³Universität Kassel, Deutschland; ⁴Seminar für Ausbildung und Fortbildung der Lehrkräfte Karlsruhe (Berufliche Schulen), Deutschland

Die Frage, was das professionelle Handeln einer Lehrperson ausmacht und in welche konstitutiven Elemente, Bedingungen und Wirkungsweisen dieses eingebettet ist, wird je nach theoretischer Perspektive unterschiedlich beantwortet und empirisch erforscht (Überblick: Idel et al. 2021). Im strukturtheoretischen Ansatz beschreibt Helsper (2004) elf dem Lehrberuf inhärente widersprüchliche und unvereinbare Situationen, Zielkonflikte und Spannungsverhältnisse (Antinomien), die von gesellschaftlichen Transformationsprozessen beeinflusst sind: Begründungs-, Praxis-, Subsumptions-, Ungewissheits-, Symmetrie-, Vertrauens-, Nähe-, Sach-, Differenzierungs-, Organisations- und Autonomieantinomie. Professionelles Handeln umfasst trotz aller Unsicherheit und Ungewissheit demnach einen souveränen, ausbalancierten Umgang mit den unauflösbaren Antinomien durch Reflexion der spezifischen Handlungs- und Anforderungsstruktur von Aufgaben, Arbeitsprozessen und Routinen.

Der theoretisch-konzeptionell sehr ausführlichen Beschreibung der antinomischen Bedingungen des professionellen Handelns von Lehrpersonen (z.B. Helsper 2004, 2016a, 2016b) steht eine geringe empirische Erforschung gegenüber. Bisherige Studien zu den Antinomien nutzen hauptsächlich qualitative Methoden und berücksichtigen lediglich ausgewählte Antinomien und nicht das komplette Modell. Ausgehend von diesen Desiderata verfolgt unsere Studie das Ziel der Entwicklung eines Instruments zur fragebogenbasierten Erfassung der von Helsper beschriebenen elf Antinomien im professionellen Handeln von Lehrpersonen inklusive der Identifikation und Überprüfung der Faktorenstruktur und einer externalen Validierung anhand des beruflichen Rollenerlebens der befragten Lehrpersonen.

Die Forschungsfragen lauten:

FF1) Lassen sich die elf Antinomien mit Fragebogenitems trennscharf abbilden?

FF2) Korreliert das Erleben der Antinomien im Berufsalltag negativ mit Rollenerfahrungen positiver Valenz und positiv mit Rollenerfahrungen negativer Valenz?

Basierend auf den theoretisch-konzeptionellen Beschreibungen der einzelnen Antinomien (Helsper 2004, 2016a, 2016b) erfolgte ein mehrstufiger iterativer Prozess zur Entwicklung geeigneter inhaltsvalider Items unter Einbezug externer Expert:innen wie einschlägigen Forscher:innen und praktizierende Lehrpersonen. Die entwickelten Items sollten Widersprüchlichkeiten innerhalb einer Person in Bezug auf Verhaltenstendenzen sowie affektive Zustände, d.h. das Erleben von Zielkonflikten, abbilden (Vorgehensweise analog Michels et al., 2011).

Final wurden 116 Antinomien-Items (9-12 Items je Antinomie; Skala: 1=nie, 2=selten, 3=manchmal, 4=häufig, 5=sehr häufig) in einer deutschlandweiten online-Befragung von (angehenden) Lehrpersonen eingesetzt. Zudem wurde als externes Validierungskriterium das berufliche Rollenerleben erhoben. Hierzu wurden die Lehrpersonen gefragt, wie sie sich in ihrer Rolle als Lehrperson im Allgemeinen fühlen. Die Abfrage erfolgte mit 12 selbst entwickelten Adjektiven, welche sich je häufig auf Erlebensqualitäten mit positiver und Erlebensqualitäten mit negativer Valenz beziehen (Skala: 1=gar nicht, 2=ein bisschen, 3=einigermaßen, 4=erheblich, 5=äußerst; $\chi^2(53)=85.47$, $p=.003$, CFI=.970, RMSEA=.049, SRMR=.067).

Die Stichprobe beinhaltet $n=260$ (angehende) Lehrpersonen (71% weiblich) verschiedener Schulstufen und -arten mit durchschnittlich 15.5 Jahren Berufserfahrung.

Explorative und konfirmatorische Faktorenanalysen sprechen für ein Modell mit elf empirisch trennbaren Faktoren mit je drei Items (CFA: $\chi^2_{SB}=557.15$, $df=440$, $p<.001$, CFI=.949, RMSEA=.032, SRMR=.051; Konstruktreliabilitäten: .68–.83). Gemäß FF1 lassen sich Begründungs-, Praxis-, Subsumptions-, Ungewissheits-, Symmetrie-, Vertrauens-, Nähe-, Sach-, Differenzierungs-, Organisations- und Autonomieantinomie abbilden. Die Lehrpersonen erleben alle Antinomien (Ausnahme Vertrauen) eher häufig (durchschnittliche Mittelwertscores über der Skalenmitte). Am häufigsten erleben sie die Symmetrieantinomie ($M=4.02$, $SD=0.78$) und die Näheantinomie ($M=4.00$, $SD=0.81$). Signifikante geringe lineare Zusammenhänge finden sich zwischen 21 ($rx_{y=.24-.49}$, $p\leq.01$) und signifikante mittlere Zusammenhänge zwischen 32 Antinomiepaaren ($rx_{y=.50-.77}$, $p<.001$). Der stärkste lineare Zusammenhang besteht zwischen der Organisations- und Autonomieantinomie ($rx_{y=.87}$, $p<.001$). Das Rollenerleben mit negativer Valenz korreliert signifikant positiv mit den Antinomien ($r=.18-.50$, $ps<.01$). Das Rollenerleben mit positiver Valenz korreliert signifikant negativ mit den Antinomien ($r=-.14- -.28$, $ps<.05$) (FF2).

Trotz der inhaltlichen Komplexität der Antinomien wurde ein schlankes Erhebungsinstrument mit guten psychometrischen Eigenschaften entwickelt, das den strukturtheoretischen Ansatz der empirischen Untersuchung in quantitativen Studien zugänglich machen und die empirische Forschung zu Antinomien im professionellen Handeln von Lehrpersonen bereichern kann. Gleichzeitig bietet es Möglichkeiten, die verschiedenen Perspektiven auf professionelles Handeln von Lehrpersonen zu verbinden. Der Befund, dass das häufige Erleben von Antinomien mit negativen Erlebensattributen des Lehrberufs einhergeht (und umgekehrt) verweist auf Aufgaben der Professionalisierung. Hierfür bietet sich der Fragebogen als (Selbst-)Evaluations- und Reflexionsinstrument an.

PS-07: Professionalisierung im Lehramt

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: SO 318

Chair der Sitzung: Christin Lotz

Zwischen Theorie und Praxis: Differentielle Entwicklungsverläufe des Unterrichts- und Fachenthusiasmus in theoretischen und schulpraktischen Phasen des Lehramtsstudiums

Christin Lotz¹, Patrick Hawlitschek², Anne Deiglmayr¹

¹Universität Leipzig, Deutschland; ²Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin

Theoretischer Hintergrund: Unterrichtsenthusiasmus und Fachenthusiasmus als die beiden Dimensionen von Lehrpersonenenthusiasmus sind von besonderer Relevanz im schulischen Kontext, da hoher Enthusiasmus mit besserer Unterrichtsqualität und sogar besserer Schüler:innenleistung einhergeht (Kunter et al., 2011; 2013). Aufgrund dieser Befunde ist es von besonderem Interesse, Lehramtsstudierende beim Aufbau und bei der Aufrechterhaltung von positiven intrinsischen Orientierungen – schon während früher Studienphasen – zu unterstützen. Trotzdem ist noch weitgehend unklar, wie sich frühe universitäre und schulpraktische Erfahrungen von Lehramtsstudierenden auf die Entwicklung beider Enthusiasmus-Dimensionen auswirken. Insbesondere die Erlebnisse während Praktikumsphasen, in denen Studierende erste eigene Unterrichtserfahrungen sammeln, könnten einen entscheidenden Einfluss auf die Veränderungen der beiden Enthusiasmus-Dimensionen in dieser frühen und besonders prägenden Phase des Lehramtsstudiums haben (Kunter & Holzberger, 2014.).

Fragestellung: Diese längsschnittliche Studie untersucht an Lehramtsstudierenden des zweiten Fachsemesters, wie sich Fachenthusiasmus und Unterrichtsenthusiasmus im Verlauf des Semesters und eines daran anschließenden vierwöchigen, verpflichtenden Unterrichtspraktikums entwickelten.

Methode: Die Stichprobe bestand aus N=671 Lehramtsstudierenden im zweiten Semester, die aus einem Pflichtmodul rekrutiert wurden, zu welchem ein vorbereitendes Seminar während des Semesters und ein vierwöchiges Blockpraktikum in der anschließenden vorlesungsfreien Zeit gehörten. Die Studierenden wurden dreimalig befragt: 1) zu Beginn des Semesters in der ersten Seminarsitzung, 2) nach der Hälfte des Semesters in der letzten Seminarsitzung und 3) am Ende des Semesters, nach dem vierwöchigen Schulpraktikum. Alle Befragungen erfolgten online und pseudonymisiert. An jedem Messzeitpunkt beantworteten die Studierenden, im Rahmen einer umfangreicheren Erhebung, auch jeweils 4 Items, die ihren Fach- und Unterrichtsenthusiasmus erfassten (Kunter et al., 2011).

Die beiden Enthusiasmus-Dimensionen wurden zunächst getrennt auf Messinvarianz über die drei Messzeitpunkte hinweg geprüft. Anschließend wurden die Veränderungen der Mittelwerte (latent difference scores) in einem gemeinsamen latent neighbor change model mittels Mplus berechnet. Dabei wurden für sowohl Unterrichts- als auch Fachenthusiasmus drei latent state Faktoren (pro Messzeitpunkt einen) durch ihre jeweils 4 Items indiktorisiert. Zusätzlich wurden indiktor-spezifische Faktoren modelliert. Die latenten Differenz-Faktoren repräsentieren jeweils die Veränderungen vom ersten zum zweiten und von zweiten zum dritten Messzeitpunkt. Zusätzlich wurden Korrelationen zwischen den Differenz-Faktoren von Fach- und Unterrichtsenthusiasmus zugelassen, die sich auf den gleichen Zeitraum bezogen.

Ergebnisse: Die beiden Enthusiasmus-Dimensionen erfüllten die Anforderungen der skalaren Messinvarianz über die drei Messzeitpunkte hinweg (alle $\Delta CFI < .01$). Das gemeinsame latent neighbor change Modell wies einen guten Modellfit auf. Die Ergebnisse zeigten, dass sich der Enthusiasmus für das Fach weder über die Dauer der universitären Vorlesungszeit noch über die Dauer des Schulpraktikums signifikant veränderte. Der Enthusiasmus für das Unterrichten blieb ebenfalls über die Spanne der Vorlesungszeit stabil; nach dem Absolvieren des Praktikums aber zeigte sich eine signifikante Differenz: Studierende berichteten dann statistisch bedeutsam höhere Unterrichtsenthusiasmus-Werte.

Differenz-Faktoren korrelierten innerhalb der gleichen Enthusiasmus-Dimension negativ, d.h. beispielsweise eine niedrigere Fachenthusiasmus-Differenz zwischen dem ersten und zweiten Messzeitpunkt ging mit einer höheren Fachenthusiasmus-Differenz zwischen dem zweiten und dritten Messzeitpunkt einher. Zeitgleiche Differenzfaktoren korrelierten positiv, d.h. beispielsweise eine größere Differenz zwischen dem Unterrichtsenthusiasmus vom ersten zum zweiten Messzeitpunkt ging mit einer größeren Differenz zwischen dem Fachenthusiasmus vom ersten zum zweiten Messzeitpunkt einher.

Diskussion: Zusammenfassend liefern unsere Analysen Hinweise darauf, dass der Fachenthusiasmus in der frühen Phase des Lehramtsstudiums eine hohe Stabilität aufweist. Ein differenzielles Befundmuster zeigt sich allerdings für den Unterrichtsenthusiasmus: Während sich der Unterrichtsenthusiasmus zunächst über die Zeitspanne der universitären Lehrveranstaltung nicht bedeutsam veränderte, stieg dieser substantiell nach dem Absolvieren des Schulpraktikums an. Frühe schulpraktische Erfahrungen erscheinen daher essentiell und prägend für die Entwicklung von Unterrichtsenthusiasmus. Mögliche unterstützende oder hemmende Bedingungen für diesen Entwicklungsverlauf, wie z.B. eine gute Mentor:innenbeziehung oder eine positive Feedback-Kultur während des Praktikums, werden diskutiert.

„Und ich habe mich jetzt entschieden, das dann halt einfach in den Besuchen durchzuziehen, weil ich weiß, dass ihr das gefällt“ - Bildungs- und Professionalisierungsprozesse angehender Lehrkräfte als Aushandlung externer Erwartungen und eigenen Überzeugungen in sozialen Netzwerken im Vorbereitungsdienst

Ariane Schmidt

Universität Siegen, Deutschland

Theoretischer Hintergrund

Referendar*innen in der zweiten Phase der Lehrer*innenbildung stehen mit verschiedenen Ausbildungsakteur*innen (z.B. Seminarleitungen und Ausbildungslehrkräfte) in Interaktion. Diesen werden im Rahmen der Ausbildungsordnung (MSB, 2024) unterschiedliche Aufgaben und Funktionen wie die Beurteilung der Unterrichtspraxis oder die Vermittlung von fachdidaktischen Inhalten zugeschrieben. Gleichzeitig belegen Studien die Bedeutsamkeit von weiteren Akteur*innen wie beispielsweise Mitreferendar*innen als Quelle sozialer Unterstützung (u.a. Richter et al., 2011).

Die sozialen Beziehungen zu all diesen Akteur*innen beeinflussen die Lern- und Professionalisierungsprozesse der angehenden Lehrkräfte maßgeblich. Forschungsbefunde zu diesen Relationen sind dabei überwiegend auf das Problemfeld „der formalen Struktur der personalen Ungeschiedenheit von Ausbilder- und Prüferrolle“ (Wernet, 2009, S. 47) bezogen. Denn diese birgt das Risiko einer Rollenambiguität für Lehramtsanwärter*innen, die sie sowohl mit den potentiell heterogenen Wahrnehmungen und Ansichten ihrer Ausbilder*innen als auch ihren eigenen Überzeugungen und Haltungen vereinen müssen (u.a. Haupt et al., 2021).

In der Konsequenz sind soziale Beziehungen zu vielen (Ausbildungs)akteur*innen von Konformitätsdruck und Anpassungsmechanismen geprägt (Pres, 2001; Schubarth et al., 2007; Käner et al., 2021). Nachgewiesen ist zudem, dass soziale Beziehungen insgesamt einen hohen Einfluss auf das Belastungserleben von Referendar*innen haben (u.a. Käner et al., 2022). Sozialen Beziehungen sowie deren Auswirkungen wurden jedoch erst (anteilig) in dyadischen Beziehungskonstellationen untersucht (Schubarth et al., 2009; Kosinár, 2013). Bisher unberücksichtigt ist das Zusammenspiel dieser Vorbereitungsdienst-Akteur*innen und deren Wirkung in Bezug auf die Bildungs- und Professionalisierungsprozesse und das Beanspruchungserleben. Es bedarf somit einer netzwerkanalytischen Perspektive auf sozialen Beziehungen unter Einbezug von Netzwerkmechanismen (Kläner & von der Lippe, 2020), denn diese liefern Erklärungsansätze für Netzwerkveränderungen und Professionalisierungsprozesse. Die Mechanismen soziales Lernen und sozialer Druck ergeben sich aufgrund von Analogieschlüssen aus bereits bestehenden Forschungen (u.a. Maetens, 2012) für die Untersuchung der Referendar*innennetzwerke als besonders geeignet.

Fragestellung

Der vorzustellende Beitrag bearbeitet dieses Forschungsdesiderat unter der Betrachtung folgender Fragestellungen:

- (1) Unter welchen Bedingungen entwickeln und verändern sich die sozialen Referendar*innennetzwerke im Verlauf des Vorbereitungsdienstes?
- (2) Inwiefern beeinflussen die verschiedenen Akteur*innen die Bildungs- und Professionalisierungsprozesse der Referendar*innen unter besonderer Berücksichtigung der sozialen Netzwerkmechanismen *soziales Lernen* und *sozialer Druck*?

Methode

Zur Beantwortung dieser Forschungsfragen wurde eine qualitativ-netzwerkanalytische längsschnittliche Interviewstudie (Hollstein & Straus, 2006) durchgeführt. Die Referendar*innen (N=15) der November-Kohorte 2022 wurden über drei Messzeitpunkte hinweg (Anfang, Mitte und Ende Vorbereitungsdienst) befragt. Für die Datenerhebung wurden halbstrukturierte Leitfadenterviews (Döring & Bortz, 2016) mit offenen Frageimpulsen genutzt. Egozentrierte Netzwerkkarten (angelehnt an Kahn & Antonucci, 1980), die die Referendar*innen während der Interviews ausfüllten, unterstützen die Visualisierung der berichteten Beziehungen. Die Daten wurden mithilfe der qualitativ strukturalen Analyse (Herz, Truschkat & Peters, 2015) und der qualitativ strukturierenden Inhaltsanalyse (Kuckartz & Rädiker, 2022) ausgewertet.

Ergebnisse

Die Ergebnisse, limitiert durch die anspruchsvollen Erinnerungsleistungen, lassen Muster in den heterogenen Netzwerken erkennen, die – je nach Beziehungsqualität und -richtung zu den jeweiligen Akteur*innen – die Wahrnehmung von professionalisierungsbedeutsamen Lern- und Drucksituationen beeinflussen. Es ergibt sich zudem eine Typologie an Netzwerk(veränderungen), die auf ein unterschiedliches Wichtigkeitsempfinden der Akteur*innen, ein heterogenes Niveau an wahrgenommener Netzwerkbelastung sowie Professions- und emotional-sozialer Unterstützung und spezifischen Professionalisierungsmotive der Referendar*innen zurückzuführen ist.

Die Bildungs- und Professionalisierungsprozesse werden, entgegen der Hypothese eines Professionalisierungserleben in Triaden oder Cliquen mit mehreren Akteur*innen, vordergründig in Dyaden wahrgenommen. Beide Netzwerkmechanismen sind nachweisbar und ergeben sich als bedeutsam für Netzwerkveränderungen. Je mehr soziales Lernen erfahren wird, desto wichtiger wird diese*r Akteur*in eingeschätzt. Sozialer Druck hingegen birgt das potenzielle Risiko einer Wichtigkeitsabnahme der Person, sofern der Druckgegenstand nicht oder nur schwer mit der eigenen Haltung vereint werden kann.

Die Ergebnisse unterstreichen die Bedeutsamkeit sozialer Netzwerke in der Lehrer*innenbildung, da für Bildungs- und Professionalisierungsprozesse angehender Lehrkräfte multiple Akteur*innen für die Weiterentwicklung unterschiedlicher Kompetenzbereiche bedeutsam sind und Akteur*innen daher nicht nur getrennt, sondern auch gemeinsam betrachtet werden müssen, um die ganzheitliche Professionalisierung zu fassen.

Lernen von Primarstudierenden durch kooperative Tätigkeiten im sozialen Netzwerk Schulpraktikum

Marco Galle¹, Annelies Kreis¹, Sonja Hiebler¹, Esther Brunner², Sanja Stankovic²

¹Pädagogische Hochschule Luzern, Schweiz; ²Pädagogische Hochschule Thurgau, Schweiz

In der Lehrpersonenbildung gelten Schulpraktika als bedeutsam für die Entwicklung professioneller Kompetenzen (u.a. Arnold et al., 2014). Praktika lassen sich als soziales Netzwerk beschreiben, in dem Studierende mit Akteur*innen mit und ohne formalem Ausbildungsauftrag interagieren. Ausgehend von der Activity Theory (Engeström, 1999; Kreis & Brunner, 2022) wird argumentiert, dass Studierende durch kooperative Tätigkeiten im Praktikum (z.B. Einholen relevanter Informationen, gemeinsame Unterrichtsplanung) ihre Kompetenzen weiterentwickeln. Aus einer kognitiv-konstruktivistische Perspektive haben kooperativ aktivere Studierende mehr Möglichkeiten, ihre Kompetenzen und ihre Wissensstrukturen auszubauen. Damit wird ihr Wissen beweglich (reversibel) und neu zu erwerbendes Wissen kann leicht anschließen (Reusser, 2019). Das kann dazu führen, dass die kooperativ aktiven – vergleichend zu weniger kooperativ aktiven – Studierenden weitere Lerngelegenheiten besser nutzen, wie etwa Besprechungen mit Ausbilder*innen, in denen Unterricht geplant und reflektiert wird.

Vor diesem Hintergrund wird untersucht, (1) ob sich Studierende hinsichtlich der selbsteingeschätzten Häufigkeit von kooperativen Tätigkeiten mit Akteur*innen ohne formalen Ausbildungsauftrag gruppieren lassen und (2) ob sich die zwei Gruppen (aktiv vs. weniger aktiv) in der wahrgenommenen Qualität und dem Nutzen von Unterrichtsbesprechungen mit Ausbilder*innen unterscheiden?

Im SNF-Projekt DiaMaNt (Kreis & Brunner, 2022) wurden 358 Studierende (Rücklauf=43%) nach einem drei-/vierwöchigen Praktikums gefragt, (1) wie häufig sie in ihrer Praxisschule über die Interaktionen mit Akteur*innen mit Ausbildungsauftrag hinaus mit Akteur*innen kooperierten, die keinen formalen Ausbildungsauftrag haben (andere Klassenlehrpersonen, Fachlehrpersonen, Fachperson schulische Sonderpädagogik, Lehrperson für DaZ). (2) Zusätzlich schätzten die Studierenden die Qualität der Unterrichtsbesprechungen mit Ausbilder*innen ein (Praxislehrperson [Deutschland: Schulmentor:in], Mentor:in [Hochschulmentor:in] und Dozent:in der Mathematikdidaktik):

(A) Qualität des Feedbacks und der Reflexion der Praxislehrperson (6 Items; Cronbachs $\alpha=.87$) sowie

(B) Anschlussfähigkeit von Wissens- und Erfahrungsbeständen der/des Praxislehr:in, Mentor:in, Dozent:in der Mathematikdidaktik (6 Items; Cronbachs $\alpha=.91-.93$).

(C) Zusätzlich beurteilten die Studierenden den Nutzen der Gespräche mit ihrer/ihrem Praxislehr:in, Mentor:in, Dozent:in der Mathematikdidaktik für die eigene Professionalisierung (4 Items; Cronbachs $\alpha=.71-.78$).

Das Antwortformat ist bei Feedback und Reflexion (A) fünf- und bei den anderen zwei Instrumenten (B, C) sechsstufig; der höchste Wert entspricht der höchsten Zustimmung.

(1) Eine Analyse latenter Klassen (Gollwitzer, 2012) mit Mplus (MLR-Estimator, fehlende Werte mit full information maximum likelihood [FIML] geschätzt; Muthén & Muthén, 2017) kam zu einer Zwei-Klassen-Lösung (Pearson Chi-Quadrat(20)=26.1, $p=.162$; Entropy=.62; Vong-Lo-Mendell-Rubin Likelihood=136.39, $p<.001$). Die zwei Klassen lassen sich als aktiv ($n=191$) vs. weniger aktiv ($n=167$) im sozialen Netzwerk Praktikum kooperierende Studierende beschreiben. (2) Ein t-Test (Field, 2018) in SPSS zeigte, dass die kooperativ aktiveren Studierenden die Qualität der Unterrichtsbesprechungen mit ihrer *Praxislehrperson* signifikant positiver einschätzten:

- Feedback & Reflexion (A): $M_{\text{aktiv}}=3.69$; $SD_{\text{aktiv}}=0.78$; $M_{\text{weniger aktiv}}=3.42$; $SD_{\text{weniger aktiv}}=0.82$; $t(356)=3.16$; $p=.002$; $d=.34$;

- Anschlussfähigkeit (B): $M_{\text{aktiv}}=5.29$; $SD_{\text{aktiv}}=0.93$; $M_{\text{weniger aktiv}}=5.07$; $SD_{\text{weniger aktiv}}=1.09$; $t(326)=1.97$; $p=.05$; $d=.22$.

Zudem beurteilten die kooperativ aktiveren Studierenden den Nutzen der Gespräche (C) mit *ihrer/ihrer Praxislehr:in und Mentor:in* signifikant positiver ein:

- Praxislehrperson: $M_{\text{aktiv}}=5.22$; $SD_{\text{aktiv}}=0.94$; $M_{\text{weniger aktiv}}=4.88$; $SD_{\text{weniger aktiv}}=1.09$; $t(253)=2.74$; $p=.007$; $d=.34$;

- Mentor:in: $M_{\text{aktiv}}=4.49$; $SD_{\text{aktiv}}=1.17$; $M_{\text{weniger aktiv}}=4.09$; $SD_{\text{weniger aktiv}}=1.18$; $t(303)=2.95$; $p=.004$; $d=.34$.

Bei den Dozierenden der Mathematikdidaktik ergaben sich keine signifikanten Unterschiede.

Die Ergebnisse zeigen, dass sich die Studierenden in der Intensität kooperativer Tätigkeiten und in der Einschätzung des Nutzens von Unterrichtsbesprechungen unterscheiden. Theoretisch erklärt werden kann dies damit, dass während des vermehrten Kooperierens im sozialen Netzwerk Praktikum mehr Lerngelegenheiten genutzt werden und so eigene Wissensstrukturen beweglicher werden sowie neues Wissen, etwa aus Unterrichtsbesprechungen, besser daran angeknüpft werden kann. Trotzdem sich dies bei den Dozierenden der Mathematikdidaktik nicht bestätigen lässt, weil evtl. in Unterrichtsbesprechungen anderes Wissen besprochen wird, als Studierende in ihrem sozialen Netzwerk erfahren, kann folgendes Fazit gezogen werden: aktives Kooperieren im sozialen Netzwerk Praktikum ist bedeutsam für Professionalisierungsprozesse. Bis zur Tagung werden weitere Analysen mit Erklärungsvariablen für die Unterschiede durchgeführt: z.B. bereitgestellte Lernangebote durch schulbasierte Ausbilder:innen oder Persönlichkeitsmerkmal Extraversion.

Unterrichtsbezogene Interventionen durch Mentor:innen aus Studierendensicht

Nicole Bosse, Alexander Gröschner

Friedrich-Schiller-Universität Jena, Deutschland

Theoretischer Hintergrund

Schulische Mentor:innen während des Unterrichtspraktikums werden von angehenden Lehrpersonen als prägende Wegbegleiter:innen der Professionalisierung wahrgenommen (Izadinia, 2015; Ulrich & Gröschner, 2020). In der Forschung sind dabei unterrichtliche Interventionen von Mentor:innen bei studentisch durchgeführten Unterrichtsstunden in den Blick gerückt (z. B. Jaspers et al., 2018; 2022; Marciano et al., 2019). Dabei wurde gezeigt, dass Mentor:innen vornehmlich bei Unterrichtsstörungen intervenieren (Gardiner, 2017; Jaspers et al., 2018; 2022), insbesondere bei einem stärker gegenüber den Schüler:innen wahrgenommenen Verantwortungsgefühl (Jaspers et al., 2022). Außerdem spielt das eigene Professionalisierungsverständnis als Lehrperson eine Rolle (Matsko et al., 2024). Da sich die Forschung bislang ausschließlich auf die Mentor:innenperspektive beschränkt, wird mit der vorliegenden Untersuchung die Studierendenwahrnehmung ergänzt.

Folgende Fragestellungen standen dabei im Fokus:

- In welchen Unterrichtssituationen sind Studierende während eigener Unterrichtsstunden grundsätzlich mit Interventionen durch die Mentor:innen einverstanden?
- Wie nehmen Studierende Interventionen während eigener Unterrichtsstunden nach dem Unterrichtspraktikum wahr?

Methodisches Vorgehen

Die Daten wurden mittels einer Onlinebefragung im Praxissemester ($N = 189$, $n = 99$ % Sekundarschullehramt; 54,5 % weiblich) sowie anhand semistrukturierter Interviews ($N = 14$; $n = 9$ weiblich) an einer deutschen Universität erhoben. Die Teilnehmenden der Onlinebefragung befanden sich im fünften beziehungsweise sechsten Fachsemester.

Der Fragebogen zur Erfassung der Einstellungen zu Interventionen nahm Bezug auf bisherige Studien (Gardiner, 2017; Gardiner & Weisling, 2014; Jaspers et al., 2018; Kiel et al., 2023; Post, 2007; Wang, 2010), wobei unterschiedliche Unterrichtssituationen auf einer 6-stufigen Likert-Skala hinsichtlich der Akzeptanz von Interventionen einzuschätzen waren (1 = „stimme überhaupt nicht zu“ bis 6 = „stimme voll und ganz zu“). Erhoben wurden Einstellungen zu Interventionen bei fachlichen Fehlern (1 Item), Einstellungen zu Interventionen bei optimierbaren, jedoch akzeptablen Situationen („Ich wäre damit einverstanden, dass meine Mentor:innen in meinen Unterricht eingreifen, wenn ich eine Aufgabe formuliere, welche noch besser formuliert werden könnte.“, vier Items; $\alpha_{11} = .76$), zu Interventionen bei Unterrichtsstörungen („... wenn Schüler:innen stören, während ich spreche“, drei Items; $\alpha_{11} = .86$) sowie bei weiteren unerwarteten Problemen („... wenn Schüler:innen Probleme bei der Aufgabenbearbeitung haben.“, vier Items; $\alpha_{11} = .80$). Die Interviews wurden mit Praktikumsabsolvent:innen durchgeführt und dauerten zwischen 18 bis 60 Minuten. Alle Interviews begannen mit einem offenen Eingangsstimulus, woraufhin die Studierenden zunächst darüber berichten konnten, inwiefern es zwischen ihnen und ihren Mentor:innen im eigenen Unterricht zu Interaktionen kam. Um den Erinnerungsprozess zu unterstützen, wurden anschließend die quantitativen Ergebnisse der Kohorte vorgestellt, woraufhin die explorative individuelle Ergründung folgte (aktuell 125 Codes).

Ergebnisse und Diskussion

Die Ergebnisse zeigen, dass die größte Offenheit gegenüber Interventionen bei fachlichen Fehlern ($M = 4.84$, $SD = 1.37$) vorliegt, gefolgt von anderen Problemsituationen ($M = 3.85$, $SD = 1.25$). Am wenigsten werden Eingriffe bei Unterrichtsstörungen gewünscht ($M = 3.18$, $SD = 1.45$), gefolgt von optimierbaren, jedoch akzeptablen Unterrichtssituationen ($M = 3.53$, $SD = 1.19$). Es zeigen sich schwache bis starke Korrelationen zwischen allen Einstellungen. Die Einstellung zu Interventionen bei fachlichen Fehlern korreliert insgesamt geringer mit jenen in anderen Unterrichtssituationen ($r = -.188$ bis $.427$, $p < .001$) als dies bei der Einstellung zu Interventionen bei Unterrichtsstörungen der Fall ist.

Die Ergebnisse weisen auf eine der Praxis von Mentor:innen konträre Einstellung von Studierenden zu möglichen Interventionen hin. Die vorliegenden Befunde ergänzen die vorhandenen Erkenntnisse demzufolge und weisen auf weiteren Forschungsbedarf

hin, indem beide Perspektiven (Mentor:innen und Studierende) in unterrichtlichen Interaktionen stärker fokussiert werden. Die erhobenen qualitativen Daten sollen Aufschluss geben, welche Faktoren bei der Wahrnehmung von Interventionen eine besondere Rolle spielen. Zur Tagung werden diese qualitativen Befunde ergänzend vorgestellt.

PS-08: Learning Mathematics: Motivation, Self-Concept, and Achievement

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 382
Chair der Sitzung: Charlott Rubach

English Track

Ability grouping in German secondary schools: The effect of non-academic-track schools on the development of Math competencies

Sonia Herrmann^{1,2}, Katharina Bach²

¹Deutsches Jugendinstitut e.V.; ²Department Psychologie, Ludwig-Maximilians-Universität München

Background

Vertical segmentation in secondary education, where students are placed in different schools based on academic performance, aims to cater to students' diverse needs and improve teaching efficiency (e.g., Ariga & Brunello, 2007; Hattie, 2002). However, across different subjects, students in academic-track schools in particular benefit from tracking, while it is often disadvantageous for students in non-academic track schools (e.g., Angelone, 2019; Scharenberg, 2014; Traini et al., 2021). These differences in competence gains between academic and non-academic-track schools are often attributed to selection effects and are explained by students' primary school performance and socioeconomic status (SES) (e.g., Ditton et al., 2005; Maaz, Neumann, et al., 2008; Scheeren, 2022). However, little attention has been given to how the competencies of comparable students (in terms of school performance and social background) who attend different tracks develop.

Aims

As the differences in students' performance at academic-track schools compared to non-academic-track schools are particularly pronounced for Math (Ireson, Clark, et al., 2002), we investigated whether comparable students diverge in their Math competencies due to attending different types of secondary schools. More precisely, we aimed to determine whether students suitable for academic-track schools but placed in non-academic-track schools fall behind their classmates who transferred to an academic-track school or can maintain their performance levels while attending a non-academic-track school. In doing so, we aimed to contribute to the ongoing debate on whether inaccurate stratification may lead to social disadvantages.

Methods

Using data from the National Education Panel Study (Kindergarten Cohort SC2, N = 4111), we examined students' competencies development from the fourth to the seventh grades.

We employed a quasi-experimental design (Propensity Score Weighting with Inverse Probability Treatment Weights) comparing similarly capable students at academic and non-academic school tracks to make causal inferences. Further, we employed multiple imputations with chained equations combining treatment effect estimates (Leyrat et al., 2019). The outcome variable was the students' Math competence in the seventh grade. Matching was conducted using fourth-grade competency measures in Math and reading and other variables such as sex, migration background, family SES, class composition, special educational needs, school grades and school location.

Results

Results revealed a significant average treatment effect on the treated (ATT). The ATT indicates that attending an academic track school significantly improves Math competence by 0.281 units for the students who attended these schools, compared to if they had attended non-academic track schools. Moreover, an additional analysis for students with high vs. low SES indicated that the effect was more pronounced for students with low SES, meaning they lag further behind comparable students in the academic track.

Discussion

Our quasi-experimental analysis suggests that the school types and learning environments have a causal effect on students' competence development. Students with Math competences comparable to those in academic tracks may not fully realize their potential if they attend non-academic track schools. These barriers to competence development have long-term detriments in students' educational and labor market outcomes (Ozer & Perc, 2020) and in motivational aspects, such as their self-concept (Ireson et al., 2001).

Moreover, we conclude that the effect of SES on competence development is probably lower than assumed and that greater importance should be attached to the school learning environment. It would be beneficial to facilitate transitions to other school tracks later on through programs and establish an official protocol for early detection of misattribution within the first six months of tracking. Moreover, improving teaching and learning conditions at non-academic-track schools, focusing on teaching quality, aligning the teacher training, and nurturing students' potential could reduce the differences between school types, an important step toward educational equity (Ditton et al., 2005; Solga & Wagner, 2001).

Längsschnittliche Untersuchung von wahrgenommenen Fremd- und Selbstattributionen als Erklärung von Geschlechtsunterschieden in mathematischen Selbstkonzepten

Katharina Reschke, Silke Hertel, Birgit Spinath

Universität Heidelberg, Deutschland

Theoretischer Hintergrund

Der Befund, dass Mädchen trotz gleicher Leistung ein niedrigeres Fähigkeitsselbstkonzept (FSK) in Mathe aufweisen, wirft die Frage auf, wodurch Geschlechtsunterschiede im mathematischen FSK erklärt werden können. Das Erwartungs-Wert-Modell (Eccles & Wigfield, 2020) postuliert, dass wahrgenommene Fremdeinschätzungen vermittelt über Selbsteinschätzungen das FSK beeinflussen. Die Ursachenzuschreibung (Attribution) von (Miss-)Erfolgen stellt in diesem Kontext einen möglichen Erklärungsfaktor dar. Empirische Studien legen nahe, dass Jungen Erfolge in Mathematik häufiger auf Fähigkeiten und Mädchen auf Anstrengungen zurückzuführen, während Jungen Misserfolge in Mathematik stärker auf mangelnde Anstrengung und Mädchen auf mangelnde Fähigkeiten attribuieren (Dickhäuser & Meyer, 2006; Jacobs & Eccles, 1992). Im Rahmen des UP2ME-Projekts konnten Reschke und Kolleginnen (2024) diese Geschlechtsunterschiede in schülerseitiger Attribution replizieren. Darüber hinaus zeigten sich diese Geschlechtsunterschiede auch für von den Schüler*innen wahrgenommene Fähigkeitsattributionen von wichtigen Bezugspersonen, wie den Eltern und Lehrpersonen. Die wahrgenommenen Fremdattributionen sagten vermittelt über die Attributionen der Schüler*innen die Fähigkeitsselbstkonzepte vorher. Bislang wurde

allerdings nicht untersucht, wie Selbst- und wahrgenommene Fremdattributionen längsschnittlich mit Fähigkeitsselbstkonzepten zusammenhängen.

Forschungsfragen und Hypothesen

In unserer Studie wurde untersucht, inwieweit geschlechtsspezifische Unterschiede in der Zuschreibung von Erfolg und Misserfolg in Mathematik über den Zeitraum von zwei Jahren bestehen und inwieweit diese Attributionen die Fähigkeitsselbstkonzepte der Schüler*innen im Längsschnitt vorhersagen. Folgende Hypothesen wurden abgeleitet: (1) Jungen glauben im Vergleich zu Mädchen stärker, dass sie selbst, ihre Eltern und Lehrpersonen Erfolg in Mathe auf Fähigkeiten zurückführen. (2) Mädchen glauben im Vergleich zu Jungen stärker, dass sie selbst, ihre Eltern und Lehrpersonen Misserfolge auf Fähigkeiten zurückführen. (3) Wahrgenommene Fremdattributionen von Eltern und Lehrpersonen sagen die Fähigkeitsattributionen der Schüler*innen im Längsschnitt vorher. (4) Fähigkeitsattributionen der Schüler*innen sagen Fähigkeitsselbstkonzepte im Längsschnitt vorher. (5) Effekte vorangegangener wahrgenommener Fremdattributionen auf nachfolgende Fähigkeitsselbstkonzepte werden über die eigenen Attributionen der Schüler*innen mediert.

Methode

$N = 696$ Schüler*innen wurden von der achten bis zur zehnten Klassenstufe begleitet und zu fünf Messzeitpunkten befragt. Dabei machten sie Angaben zu ihrem mathematischen FSK und eigenen Ursachenzuschreibungen. Zudem gaben sie an, inwiefern ihre Eltern und Lehrkräfte (Miss-)Erfolge in Mathematik auf ihre Fähigkeiten attribuieren. Neben der Berechnung von Varianzanalysen für die Prüfung der Geschlechtsunterschiede wurden Cross-lagged Modelle (1: Erfolgsattribution; 2: Misserfolgsattribution) für die Vorhersage der Fähigkeitsselbstkonzepte analysiert. Im Rahmen von Multigruppenmodellen wurde analysiert, inwiefern Unterschiede in den Pfadkoeffizienten zwischen Mädchen und Jungen bestehen.

Ergebnisse

Wie erwartet zeigten sich für alle Messzeitpunkte signifikante Geschlechtsunterschiede im FSK ($d = 0.57$) und in den Fähigkeitsattributionen der Schüler*innen. Jungen sahen Erfolge in Mathe stärker in ihren eigenen Fähigkeiten begründet ($d = 0.23$), während Mädchen stärker daran glaubten, dass mangelnde Fähigkeiten für einen Misserfolg in Mathe verantwortlich sind ($d = 0.32$). Für die wahrgenommenen Fremdattributionen konnten nur teilweise Geschlechtsunterschiede festgestellt werden. Die Ergebnisse der Strukturgleichungsmodelle zeigten, dass vorauslaufende wahrgenommene Fremdattributionen nachfolgende Attributionen der Schüler*innen vorhersagten ($b = .25 - .45$). Darüber hinaus stellten sowohl wahrgenommene Fremdattributionen ($b = .14$ für Erfolgsattributionen und $b = -.11$ für Misserfolgsattributionen) als auch die eigenen Attributionen der Schüler*innen ($b = .18$ für Erfolgsattributionen und $b = -.16$ für Misserfolgsattributionen) signifikante Prädiktoren der FSKe dar, wobei die längsschnittlichen Effekte für Mädchen und Jungen unterschiedlich waren. Hierbei zeigte sich beispielsweise, dass die Effekte der wahrgenommenen Lehrerattribution lediglich für Mädchen signifikant waren. Zudem zeigten sich indirekte Effekte von wahrgenommenen Fähigkeitsattributionen der Bezugspersonen vermittelt über die eigenen Attributionen der Schüler*innen auf deren FSKe.

Diskussion

Diese Ergebnisse liefern erstmalig Hinweise darauf, dass wahrgenommene Attributionen von Bezugspersonen, wie Eltern und Lehrpersonen bei der Erklärung von Geschlechtsunterschieden im FSK eine wesentliche Rolle spielen. Die Bedeutsamkeit der Bezugspersonen für die Attribution der Schüler*innen sowie Ansatzpunkte für Interventionen werden diskutiert.

Mastery vs. Performance: How Achievement Goals Relate to State Academic Self-Concept in Mathematics and German

Anke Maria Weber¹, Valentin Emslander², Christoph Niepel¹

¹University of Luxembourg, Luxembourg; ²Justus Liebig University Giessen, Germany

Theory

Academic self-concept (ASC) refers to a person's perception of their ability¹, and has been extensively investigated for relations with various domain-specific psychological components^{2,3}. An example are achievement goals which describe what a student strives for in education^{4,5}. According to the trichotomous model⁶, achievement goals consist of performance-approach (outperforming others), performance-avoidance (avoiding performing worse than others), and mastery goals (acquiring knowledge/competence). If a student has mainly performance-related goals, their ASC might vary depending on their perception of how they performed compared to their peers. However, a student who has mainly mastery goals might experience less variation, as they rely less on peer comparisons. Accordingly, Niepel et al. (2014)⁷ found that performance goals were related to changes in ASC, but mastery goals were not. However, the effect of achievement goals on between-person differences in mean levels and variability in state mathematics (MSC) and German self-concept (GSC) has never been investigated.

Therefore, we address the following research questions:

Are achievement goals associated with between-person differences in...

RQ1. mean levels of state MSC and GSC?

RQ2. within-person variability in state MSC and GSC?

Method

This study was part of a larger experience-sampling project focusing on students' ASC. Students' habitual achievement goals were assessed in week 1 (paper/pencil questionnaire). After each mathematics or German lesson in weeks 2 to 4, students completed e-diaries (via smartphones) assessing their MSC and GSC, respectively. 381 (49.87% female; age: $M = 15.33$, $SD = 0.67$) students participated in the study. Of these, 263 (for mathematics) and 217 students (for German) fulfilled the following criteria: (1) responses on all achievement goals items; (2) at least 3 observations on the MSC/GSC items (3) with variance > 0 .

Achievement goals were assessed with an adaption of the Achievement Goal Questionnaire⁸ using the subscales performance-approach, performance-avoidance and mastery goals with three items each and domain-specific wording (0 = *does not apply at all* to 5 = *completely applies*).

State MSC and state GSC were each assessed with three items adapted from the Self-Description Questionnaire⁹; 0 = *strongly disagree* to 5 = *strongly agree*.

We used mixed-effects location scale models¹⁰ to analyse the statistical effects of achievement goals on mean-level and variability in state MSC and state GSC. Mathematics and German grades, reasoning ability (Intelligence Structure Test-Screening¹¹), and gender were controlled for.

Results

For state MSC, results showed that performance-approach, $\beta = .16$, $p = .005$, and mastery goals, $\beta = .22$, $p < .001$, but not performance-avoidance goals, $\beta = -.02$, $p = .660$, were related to mean-level state MSC. Achievement goals were not related to variability in state MSC.

For state GSC, results showed that mastery goals, $\beta = .12$, $p = .045$, were related to mean-level state GSC, but performance-approach goals, $\beta = .11$, $p = .225$, and performance-avoidance goals, $\beta = .10$, $p = .277$, were not. Achievement goals were not related to variability in state GSC.

Discussion

Regarding RQ1, results suggest that students who strive to be better than their peers (performance-approach goals) in mathematics or to acquire an understanding of a topic (mastery goals) have higher ASC, while avoiding doing worse than peers (performance-avoidance goals) is not related to ASC.

Regarding RQ2, we found no relation between variability in state MSC and state GSC with achievement goals, suggesting that variability in state ASC is independent of students' achievement goals.

This study contributes to the literature on state ASC, highlighting the importance of achievement goals for self-perceptions.

A cross-culture comparison of the links between adolescents' motivational beliefs and their performance in math

Charlott Rubach¹, Glona Lee-Poon², Christy Starr³, Nayssan Safavian², Anna-Lena Dicke², Jacque Eccles², Sandra Simpkins²

¹Universität Rostock, Germany; ²University of California, Irvine; ³University of Wisconsin, Madison

Theoretical and empirical foundation. Some students do not get their chance of a STEM career even if they have the skills and interest in doing so. Educational research thus continues to explore how schools and teachers can reduce educational disparities in STEM, ensuring that all students have equal opportunities to pursue careers that align with their skills and interests. In this paper, we focus on performance as consequential for students college applications in the U.S. (Newton et al., 2011) and investigate the extent to which high school students' initial motivational beliefs and the development of those beliefs predict their performance in math. Using an intersectional lens, we extending previous results on gender and racial/ethnic differences (Abu-Hilal et al., 2014; Adler et al., 2024; Else-Quest et al., 2013; McKellar et al., 2024) and examine the extent to which these psychological processes vary across three of the largest racial/ethnic groups in the U.S also considering gender groups, i.e., male and female Asian-Americans, European-Americans and Latiné-Americans. With this, we aimed to understand the universality and specificity of educational processes (Myint et al., 2024). Such results allow to identify patterns that can inform more equitable educational policies and practices.

Methods. We used two diverse U.S. datasets with 4,211 (California Achievement Motivation Project) and 18,352 (High School Longitudinal Study) students from 9th to 11th grade. Students' utility value beliefs ($\omega \geq .77$), competence-related beliefs ($\omega \geq .77$) and standardized test scores were the instruments used. Background variables like parent education, SES, prior scores, and math courses were added as covariates. After testing measurement invariance across time*gender and time*ethnicity/race (Authors, 2022), we imputed missing data on motivational beliefs and performance indices and specified one multi-group latent change score model for each dataset. All latent-change score models had good model fit indices.

Results. Before using an intersectional lens, models for the overall sample showed that students with higher initial math competence-related beliefs scored higher in math tests, even after controlling for prior scores. In the HSLS dataset, changes in these beliefs positively predicted math test scores, while in both CAMP and HSLS, changes in utility value beliefs were related to test scores, with positive changes leading to higher scores.

Using an intersectional lens revealed group-specific effects. For **Asian-Americans**, higher initial utility value beliefs predicted higher math test scores for both gender groups. For male Asian-Americans, higher competence-related beliefs and positive changes in both beliefs predicted test scores, but only in HSLS. For **European-Americans**, higher initial competence-related beliefs and less negative changes in both beliefs predicted higher scores, mostly in HSLS. For **Latiné-Americans**, higher initial competence-related beliefs predicted higher scores in both datasets, with changes in competence-related beliefs relevant only for Latino-Americans in HSLS.

Discussion and significance. In summary, we mostly found subgroup specific effects with some hints of similarities across groups. For privileged groups in STEM, i.e., male Asian- and European-Americans, both competence-related beliefs and utility value beliefs help to understand their educational progress in math. For other groups like female Asian-Americans and Latina-Americans, however, only a few links were meaningful and explained little variance in their math test scores, indicating that researchers need to consider other factors that matter for these groups' performance in math.

In terms of the replication of effects across datasets, 13 out of 24 effects replicated across the two datasets. This highlights not only the importance of replications in psychological research (Plucker & Makel, 2021), but also to investigate context-specific factors of students' educational environments.

These findings will be used to discuss the implications of such insights for the German education sector and possible differences in the largest racial/ethnic groups.

PS-09: Diagnosekompetenz in Theorie und Praxis: Perspektiven und Interventionen

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: SO 322
Chair der Sitzung: Juliane Rutsch

Akkurat und sicher? Zur klassenbezogenen Diagnosekompetenz von Grund-schullehrkräften in Zusammenhang mit der Leistungsrückmeldepraxis

Juliane Rutsch¹, Nicolas Hübner², Torsten Eckermann³, Evelin Ruth-Herbein¹, Ann-Kathrin Schnelle¹, Benjamin Fauth¹

¹Institut für Bildungsanalysen Baden-Württemberg, Deutschland; ²Universität Tübingen; ³Pädagogische Hochschule Freiburg

Einleitung. Diagnostische Kompetenzen sind zentral für das professionelle Handeln von Lehrkräften (Souvignier, 2022). So wurde für die klassenbezogene Diagnosekompetenz - eine am Leistungsniveau einer bestimmten Klasse orientierte Auswahl an geeigneten Lernaufgaben - gezeigt, dass diese die Leistungsentwicklung von Schülerinnen und Schülern positiv vorhersagen kann (Karst et al., 2014; Anders et al., 2010). Bislang sind Faktoren, die das diagnostische Urteil beeinflussen, jedoch wenig beforscht (Südkamp et al., 2012); in der Literatur werden die Unterrichtserfahrungen der Lehrkräfte in der jeweiligen fachlichen Domäne sowie die Klassengröße diskutiert (Anders et al. 2010).

Noch nicht untersucht wurde bislang, ob und wenn ja inwiefern die Leistungsrückmeldepraxis von Lehrkräften mit der klassenbezogenen Diagnosekompetenz in Zusammenhang steht. Im Rahmen des baden-württembergischen Schulversuchs „Lernförderliche Leistungsrückmeldung in der Grundschule“ (LLr) verzichteten Lehrkräfte an Modellschulen auf die Vergabe von Noten zugunsten eines verstärkten Einsatzes von alternativen Formen der Leistungsrückmeldung, was eine differenziertere und individualisierte Leistungsrückmeldung ermöglichen soll. Im Rahmen der begleitenden längsschnittlichen Evaluationsstudie bietet sich daher die einzigartige Möglichkeit zu untersuchen, ob sich die klassenbezogene Diagnosekompetenz durch eine veränderte Leistungsrückmeldepraxis im Zeitverlauf ebenfalls verändert. Zudem wird die Entscheidungssicherheit der Lehrkräfte bei diagnostischen Urteilen in den Blick genommen. Dies erscheint insbesondere deshalb von hoher Bedeutung, da der Verzicht auf Noten den Modellschulen andere Möglichkeiten für die unterrichtliche Diagnose bietet. Gleichzeitig entsteht durch den Verzicht auf Noten für die Lehrkräfte aber auch die Notwendigkeit, die bisherige diagnostische Praxis anzupassen, um den Überblick über die Lernstände der einzelnen Schülerinnen und Schüler zu behalten.

Fragestellungen. 1) Unterscheiden sich die Grundschullehrkräfte der Modell- und der Vergleichsschulen zu Beginn des Schulversuchs im Hinblick auf ihre klassenbezogene Diagnosekompetenz? 2) Unterscheiden sich Lehrkräfte an Modell- und Vergleichsschulen hinsichtlich ihrer wahrgenommenen Entscheidungssicherheit? 3) Besteht ein Zusammenhang zwischen der Entscheidungssicherheit und der klassenbezogenen Diagnosekompetenz?

Methode. Zu Beginn von Klassenstufe 2 wurden die mathematischen Fähigkeiten von 1946 Schülerinnen und Schülern mit dem DEMAT1+ getestet (Krajewski, Küspert & Schneider, 2002). Die zugehörigen Lehrkräfte an den Modellschulen ($N = 35$) und den Vergleichsschulen ($N = 25$) schätzten dazu für fünf Aufgaben ein, wie viele Kinder in ihrer Klasse diese richtig gelöst haben. Zudem schätzten die Lehrkräfte ihre Entscheidungssicherheit für jede Aufgabe ein (1 = *sehr unsicher* bis 4 = *sehr sicher*, $\alpha = .93$). Als Maß für die klassenbezogene Diagnosekompetenz wurde der Niveaufehler nach Schrader (1989) herangezogen.

Ergebnisse. Die eingeschätzte und tatsächliche Anzahl an Schülerinnen und Schülern, die die Aufgaben korrekt lösten, korrelierten zwischen $r = .46$ und $r = .66$ ($ps < .001$). Lehrkräfte an Modellschulen und Vergleichsschulen unterschieden sich zu Beginn des Schulversuchs weder hinsichtlich ihrer klassenbezogenen Diagnosekompetenz ($t(58) = .368$, $p = .714$, $d = 0.1$) noch der Entscheidungssicherheit ($t(58) = .879$, $p = .383$, $d = .23$). Die Entscheidungssicherheit stand nicht in Zusammenhang mit der klassenbezogenen Diagnosekompetenz ($r = .084$, $p = .521$).

Diskussion und Ausblick. Die befragten Lehrkräfte konnten die Schülerleistungen akkurat einschätzen (vgl. Südkamp et al. (2012): mittlere Urteilstgenauigkeit bei $r = .56$). Die Entscheidungssicherheit war leicht überdurchschnittlich ausgeprägt ($M = 2.51$, $SD = .64$), stand aber nicht in Zusammenhang mit der klassenbezogenen Diagnosekompetenz. Dieser Befund schließt an die Expertise-Forschung an, wonach bei Lehrkräften, die innerhalb einer Klasse länger unterrichten und denen demnach mehr Informationen über die Schülerinnen und Schüler ihrer Klasse vorliegen müssten, das diagnostische Urteil jedoch nicht genauer ausfällt (Oerke et al., 2016). Es zeigten sich zudem (noch) keine bedeutsamen Unterschiede zwischen den Lehrkräften der Modell- und Vergleichsschulen. Hierbei gilt allerdings zu berücksichtigen, dass sich die Daten auf den Beginn des Schulversuchs beziehen. In nachfolgenden Analysen wird untersucht werden, ob sich die klassenbezogene Diagnosekompetenz aufgrund der Teilnahme am Schulversuch positiv entwickelt. Ergebnisse hierzu werden zum Zeitpunkt der Tagung vorliegen und im Rahmen des Vortrags näher erläutert.

Diagnosekompetenzen und ihre Entwicklung – ein theoretisches Modell und seine (wissenschafts-)praktischen Implikationen

Anika Radkowsch¹, Daniel Sommerhoff¹, Stefan Ufer², Frank Fischer²

¹Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland; ²Ludwig-Maximilians-Universität München, Deutschland

Kompetenzen zur Lösung diagnostischer Probleme sind in vielen Disziplinen wie beispielsweise der Medizin oder im Lehramt relevant. Entsprechend ist der Erwerb von Diagnosekompetenzen ein wichtiges Ziel der Hochschulbildung (Heitzmann et al., 2019). Um Diagnosekompetenzen effektiv und passend zu den Voraussetzungen der Lernenden zu fördern, ist es notwendig, zu Grunde liegende Entwicklungsmechanismen zu beschreiben. Bisherige Förderansätze fokussieren dabei insbesondere zwei Mechanismen: eine Veränderung der kognitiven Voraussetzungen (z.B. durch die Vermittlung von relevantem Wissen, z.B. Rieu et al., 2022) und eine Veränderung des Verhaltens in Diagnosesituationen (z.B. durch das Anbieten von Übungsgelegenheiten in Simulationen, z.B., Chernikova et al., 2020). Beide Ansätze wurden jedoch bisher nicht in ein kohärentes Modell integriert. Ziel des vorliegenden Beitrags ist es, basierend auf Forschung zur Kompetenzentwicklung im Bereich des (diagnostischen) Problemlösens beide Ansätze in ein theoretisches Modell zu integrieren und Konsequenzen für die Erforschung und Förderung von Diagnosekompetenzen abzuleiten.

Bisherige Forschung zeigt, dass Personen mit hohen diagnostischen Kompetenzen nicht nur quantitativ über mehr Wissen, sondern vielmehr über eine qualitativ bessere Wissensstruktur verfügen (Schmidt & Rikers, 2007). Diese ermöglicht die Unterscheidung zwischen relevanten und irrelevanten Informationen und einen effizienteren Abruf dieser Informationen aus dem Langzeitgedächtnis (Gegenfurtner et al., 2011). Basierend auf den so identifizierten relevanten Problemmerkmalen kann eine

kohärente Problemrepräsentation aufgebaut werden, die Vorhersagen bezüglich weiterer Problemmerkmale erlaubt und effizientes diagnostisches Verhalten in diagnostischen Situationen begünstigt.

Daraus ergibt sich die Frage, durch welche Mechanismen sich die Wissensstruktur mit wachsender Kompetenz verändert. Basierend auf einem systematischen Literaturreview haben Boshuizen und Kollegen (2020) insbesondere die Bearbeitung von authentischen Problemen als wichtigen Faktor für die Entwicklung einer solchen effizienten Wissensstruktur identifiziert. Genauere Wirkmechanismen wurden jedoch nicht beschrieben. Einen möglichen Ansatz dafür bietet die Theorie des Fallbasierten Lernens (Kolodner, 1992), in der davon ausgegangen wird, dass Probleme anhand eines Abgleichs zwischen den Problemmerkmalen und Lösungen zuvor gelöster Probleme und des aktuellen Problems gelöst werden. In dem von uns entwickelten Modell zur Beschreibung der Entwicklungsmechanismen von Diagnosekompetenzen nehmen wir Folgendes an: 1) Gelöste Probleme werden in einer Falldatenbank mittels Indizierung gespeichert. Das bedeutet, dass die im Problemlöseprozess als relevant identifizierten Merkmale und die dazugehörige Problemlösung abstrakt gespeichert werden. 2) Bei der Problembearbeitung wird die so ausdifferenzierte Falldatenbank mittels Mustererkennung einbezogen. Das bedeutet, dass ein Abgleich zwischen den relevanten Merkmalen des aktuellen Problems im Vergleich zu schon gelösten Problemen durchgeführt wird. Kann so keine passende Problemlösung für das aktuelle Problem identifiziert werden, wird idealerweise auf bewusste Problemlöseprozesse basierend auf dem Vorwissen zurückgegriffen. Das Lösen von neuen Problemen führt dabei zu einer Differenzierung und Generalisierung der mentalen Falldatenbank und somit einer Umstrukturierung von Wissen hin zu prototypischen abstrahierten Diagnoseproblemen.

Bezüglich der Förderung von Diagnosekompetenzen impliziert das Modell, dass das Lösen diagnostischer Probleme für die Kompetenzentwicklung hilfreich ist. Da Anfänger*innen jedoch über keine gut differenzierte Falldatenbank verfügen, sollten sie dabei insbesondere auf wissensbasierte Problemlöseprozesse zurückgreifen, wobei sie von Wissensunterstützung profitieren könnten. Weiter fortgeschrittene Lernende könnten insbesondere von der Kontrastierung von Problemmerkmalen profitieren, wodurch die Differenzierung und Generalisierung der Falldatenbank unterstützt werden sollten. Somit bietet das Modell eine theoriebasierte Grundlage für die Adaptation von Unterstützung an das Vorwissen der Lernenden. Gleichzeitig legt das Modell nahe, dass klassische *multiple choice* Tests lediglich die Quantität, nicht aber die Struktur von Wissen erfassen können und somit als Instrument zur Erfassung des Vorwissens als Grundlage zur Adaption eher ungeeignet sein dürften. Darüber hinaus zeigt das Modell Bedarf für die Erforschung besonders günstiger Sequenzen von Diagnoseproblemen (z.B. von typischen zu atypischen Problemen) und von Unterschieden in den Entwicklungsmechanismen zwischen Domänen auf.

Kurzintervention zur Förderung diagnostischer Kompetenzen von Lehrkräften in mehrdeutigen Situationen

Elif Özel^{1,2}, Timo Leuders¹, Katharina Loibl¹

¹PH Freiburg, Deutschland; ²PH Heidelberg, Deutschland

Theoretischer Hintergrund

Um unterrichtliche Aktivitäten an den Lernvoraussetzungen der Schülerinnen und Schüler zu orientieren, ist es notwendig, Einsicht in deren Denkweisen zu gewinnen (Helmke, 2010). Solche Denkweisen (z. B. Fehlvorstellungen) sind der unmittelbaren Wahrnehmung einer Lehrkraft nicht zugänglich und müssen rekonstruktiv zum Beispiel aus Aufgabenbearbeitungen erschlossen werden (Loibl et al., 2020).

In diesem Beitrag werden diagnostische Urteile über Denkweisen von Schülerinnen und Schülern als das Ergebnis kognitiver Informationsverarbeitung einer Lehrkraft konzeptualisiert (Loibl et al., 2020). Grundsätzlich entstehen diagnostische Urteile aus einer dynamischen Interaktion zwischen Informationen, die in der diagnostischen Situation bestehen (z.B. Fehlermuster in den Aufgabenbearbeitungen) und Informationen, die im Langzeitgedächtnis gespeichert sind (z.B. fachdidaktisches Wissen zu Fehlvorstellungen). Diese werden zu einem kohärenten Urteil verknüpft. Das fachdidaktische Wissen kann dabei zum einen die Aufmerksamkeit der Lehrkraft auf die relevanten Informationen in den Aufgabenlösungen lenken (Wissensgesteuerte Aufmerksamkeitslenkung; Van Es & Sherin, 2021). Zum anderen kann das fachdidaktische Wissen die Bandbreite an möglichen Fehlvorstellungen aufzeigen und so die Ursachensuche für das identifizierte Fehlermuster organisieren (Knowledge-based reasoning; Van Es & Sherin; 2002).

Wenn die Informationssuche in mehrdeutigen Aufgabenbearbeitungen nach der ersten identifizierten Fehlvorstellung vorzeitig abgebrochen und alternative Fehlvorstellungen nicht erprobt werden, resultieren unzulässige Urteile (Cognitive-Biases; Gäbler, 2017; Rieu et al., 2024). Aufgabenbearbeitungen sind mehrdeutig, wenn ein in den Aufgabenbearbeitungen identifiziertes Fehlermuster durch unterschiedliche Fehlvorstellungen erklärt werden kann.

Bisher gibt es wenig Forschung dazu, wie der Vernachlässigung relevanter Informationen beim Diagnostizieren von Fehlvorstellungen in mehrdeutigen Aufgabenbearbeitungen entgegengewirkt und das Erkennen der Mehrdeutigkeit gezielt trainiert werden kann.

Fragestellung

Die vorliegende Studie untersucht, ob eine kurze Intervention der Vernachlässigung relevanter Informationen beim Diagnostizieren mehrdeutiger Aufgabenbearbeitungen entgegenwirken kann.

Methode

Im Prä-Post-Design wurden 40 Lehrkräfte aufgefordert, Aufgabenbearbeitungen von insgesamt drei Schülerinnen bzw. Schülern zum Größenvergleich von Brüchen zu analysieren und auf der Basis ihrer Analysen Diagnosen über zugrundeliegende Fehlvorstellungen zu formulieren. Das hierfür notwendige fachdidaktische Wissen wurde zuvor unterrichtet.

Alle zu analysierenden Aufgabenbearbeitungen sind mehrdeutig, d.h. auf mehrere Fehlvorstellungen zurückzuführen.

Eine Woche nach dem Prätest wurden die Lehrkräfte zufällig in zwei Gruppen (Kontroll- und Trainingsgruppe) aufgeteilt.

Die Trainingsgruppe (N=20) nahm an einer 45-minütigen Intervention teil. In der Intervention trainierten sie gezielt, Hypothesen auf der Basis von Informationen zu testen, Alternativhypothesen aufzustellen (Debiasing-Strategien; Croskerry et al., 2013) und die Mehrdeutigkeit von Aufgabenbearbeitungen aufzulösen, indem diagnostische Aufgaben konstruiert wurden. Abweichend vom Prä-Posttest fand die Intervention zum Inhaltsbereich Wahrscheinlichkeiten statt. Die im Bereich der Wahrscheinlichkeiten erworbenen Kompetenzen mussten im Posttest eigenständig auf den Bereich der Brüche übertragen werden. Die Lehrkräfte der Kontrollbedingung bearbeiteten den Posttest ohne vorangehende Intervention.

Zur Untersuchung der Wirksamkeit der Intervention wurden die im Prä- und Posttest gebildeten Urteile der Lehrkräfte analysiert. Dabei wurde die Anzahl der von ihnen als Ursache diagnostizierten Fehlvorstellungen erfasst. Jede richtig identifizierte Fehlvorstellung erhielt einen Punkt (max. 9).

Die erzielten Punkte im Prätest wurden mit den erzielten Punkten im Posttest verglichen. Zur statistischen Auswertung wurde der Mann-Whitney U-Test verwendet.

Ergebnisse und ihre Bedeutung

Während sich in der Anzahl der diagnostizierten Fehlvorstellungen der Lehrkräfte der Kontrollgruppe im Prä-Posttest-Vergleich keine signifikanten Unterschiede zeigten, diagnostizierten Lehrkräfte der Trainingsgruppe im Posttest signifikant mehr Fehlvorstellungen ($U=349.0$, $Z=4.343$, $p=0.001$).

Die Lehrkräfte der Kontrollgruppe identifizierten im Posttest durchschnittlich 4,8, die Lehrkräfte der Trainingsgruppe durchschnittlich 8,5 Fehlvorstellungen. Das Fortbildungsmodul erzielte folglich eine große Effektstärke ($r=0,69$) auf die Anzahl der identifizierten Fehlvorstellungen.

Nur 4 von 40 Lehrkräften erkannten im Prätest die Mehrdeutigkeit der zu untersuchenden Aufgabenbearbeitungen vollständig und identifizierten alle Fehlvorstellungen als mögliche Ursache. Die Ergebnisse des Prätests zeigen, dass das Diagnostizieren von Fehlvorstellungen eine kognitiv anspruchsvolle Aufgabe ist und Lehrkräfte dazu tendieren, die Mehrdeutigkeit von Aufgabenbearbeitungen nicht zu erkennen (Rieu et al., 2024). Die Ergebnisse des Posttests zeigen, dass die Intervention eine wirksame Möglichkeit bietet, Lehrkräfte auf diese anspruchsvolle Aufgabe vorzubereiten.

Förderung von Noticing-Kompetenzen angehender Lehrkräfte von nicht-salienten Unterrichtsstörungen mit der Unterstützung von Cues

Maxie Kilbury, Anja Böhnke, Felicitas Thiel

Freie Universität Berlin, Deutschland

Die frühzeitige Wahrnehmung von störungskritischem Schüler*innen-Verhalten stellt für Lehrkräfte in der Unterrichts-Praxis eine besondere Herausforderung dar, u.a., weil störungskritisches Verhalten ganz unterschiedlich salient (auffällig) sein kann (z.B. lautes Quatschen vs. leises Tuscheln). Lehrkräfte müssen sowohl salientes, als auch nicht-salientes störungskritisches Verhalten frühzeitig wahrnehmen, um den Unterrichtsfluss zu gewährleisten. Diese Kompetenz wird in der Forschung allgemein als *Noticing* bezeichnet. Die folgende Studie lehnt sich an Barths (2017) Verständnis von *Noticing* an und fokussiert auf den Prozess der „Wahrnehmung“. Wahrnehmung wird gelenkt von der Aufmerksamkeit und kann reizinduziert oder zielgesteuert sein. Erfahrene Lehrkräfte lenken ihre Aufmerksamkeit meist zielgesteuert und verfügen über ein sogenanntes „professionelles Monitoring-Blickverhalten“ (Grub et al., 2022b; Kounin, 1970). Sie lassen ihren Blick oft schweifen und blicken zwar kurz, aber häufig auf einzelne Schüler*innen. Sie erhöhen dadurch die Chance, dass alle Schüler*innen Aufmerksamkeit erhalten (Kosel et al., 2023). Des Weiteren selektieren sie gezielt relevante Informationen (Gegenfurtner, 2020). Noviz*innen verfügen noch nicht über die Erfahrung zum Aussortieren der irrelevanten Ereignisse. Zudem wird ihre Aufmerksamkeit oft nur reizinduziert und richtet sich auf einzelne, saliente Ereignisse, wodurch andere (relevante) Ereignisse übersehen werden (Schulden et al., 2019; van den Bogert, 2016). Daher sollte *Noticing* von (nicht-salienten) störungskritischen Ereignissen bereits bei angehenden Lehrkräften gefördert werden.

Der Einsatz von Unterrichtsvideos bietet die Möglichkeit professionelle Kompetenzen möglichst praxisnah und ohne Handlungsdruck zu erlernen. Allerdings transportieren auch Unterrichtsvideos eine Menge an komplexen Informationen, die für Noviz*innen herausfordernd sind (Syring et al., 2015). Sie sollten also durch gestalterische Elemente unterstützt werden, um Informationen für Noviz*innen zu reduzieren und helfen relevante Informationen zu identifizieren. In diesem Zusammenhang haben sich *Cueing*-Techniken als eine vielversprechende Möglichkeit erwiesen (De Koning et al., 2009; Lowe & Boucheix, 2011). *Cues* sind eine Manipulation visuell-räumlicher Merkmale von Lehrmaterialien mit dem Ziel die Vermittlung komplexer Inhalte zu unterstützen, indem relevante Merkmale (z.B. störungskritische Ereignisse) hervorgehoben werden (van Gog, 2014). Des Weiteren können nacheinander im Video erscheinende *Cues* helfen, den Blick von Noviz*innen über die gesamte Dauer eines Unterrichtsvideos so zu lenken, wie in der Expertise-Forschung aufgezeigt („Monitoring-Blickverhalten“).

Bislang sind die Methoden entwickelter Video-Instrumente, die *Noticing* adressieren, technisch sehr aufwendig (z.B. Eye-Tracking, Virtual Reality und 360-Grad-Videos, z.B.: Huang et al., 2023; Theelen et al., 2019; Wolff et al., 2016) und die Auswertung der erhobenen Daten ressourcenintensiv (z.B. „laut Denken“-Protokolle, Gruppendiskussionen oder recall-Interviews). Kilbury et al. (in Vorb.) haben ein Video-Instrument entwickelt, was *Noticing*-Kompetenzen „on-the-fly“ erfasst und im Lehr-Lernkontext ressourcenschonend eingesetzt werden kann.

Diese Studie geht der Fragestellung nach, ob *Noticing*-Kompetenzen von (nicht salienten) störungskritischen Ereignissen von Lehramtsstudierenden mit der Unterstützung von *Cues* im Rahmen dieses Video-Instruments gefördert werden können. Dafür wurden Video-Clips mit salienten und nicht-salienten störungskritischen Ereignissen produziert und in ihnen grafische *Cues* implementiert, die jedes störungskritische Ereignis kurzzeitig im Verlauf des Video-Clips hervorheben.

Im Rahmen einer randomisierten, experimentellen Erhebung ($N=333$) wurden Lehramtsstudierende dazu aufgefordert, sich 15 Video-Clips (zunächst ohne *Cues*) anzusehen und dabei jedes störungskritische Ereignis im laufenden Video mit einem Kreuz (via Mausclick oder Touch) zu markieren. Nach der Bearbeitung jedes Videos sahen sie das gleiche Video noch einmal, wobei eine Gruppe das Video dann mit *Cues* auf die störungskritischen Ereignisse sah (EG) und die andere Gruppe ohne *Cues* (=KG).

Insgesamt identifizierte die EG im Verlauf der Intervention signifikant mehr nicht-saliente Ereignisse, als die KG ($t(279) = -3.01$, $p = .003$, $d = -.359$). Bei der Identifikation von salienten Ereignissen konnte kein Gruppenunterschied festgestellt werden, was sich mit bisherigen Forschungsergebnissen deckt (van Driel et al., 2021). Über den Verlauf der gesamten Intervention zeigt sich auf der deskriptiven Ebene ein deutlicher Anstieg der erkannten nicht-salienten Ereignisse durch die EG nur bei den ersten zehn Clips, bei der KG ist gar kein Anstieg zu sehen.

PS-10: Selbstkonzept, Motivation und Feedback: Dynamiken und Einflüsse im Bildungsprozess

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 157

Chair der Sitzung: Emely Hoch

Eine Untersuchung der intraindividuellen Variabilität im akademischen Selbstkonzept mittels Experience Sampling: Die Bedeutung von Erwartungen und Werten

Christoph Niepel¹, Jennifer Hausen¹, Anke Maria Weber¹, Jens Möller²

¹Universität Luxemburg, Luxembourg; ²Christian-Albrechts-Universität zu Kiel

Das akademische Selbstkonzept, definiert als die mentale Repräsentation der eigenen akademischen Fähigkeiten (Shavelson et al., 1976), ist entscheidend für den Bildungserfolg und das Wohlbefinden von Schüler:innen (Marsh et al., 2019). Forschungsarbeiten, die das akademische Selbstkonzept von Schüler:innen als States konzeptualisiert haben und damit situative mentale Repräsentationen der eigenen Fähigkeiten von einer Unterrichtsstunde zur nächsten untersucht haben, liegen nur sehr begrenzt vor (z.B. Niepel et al., 2022). Folglich wissen wir wenig darüber, welche Rolle habituelle lernbezogene Erwartungen und Werte für die mittlere Ausprägung von und tägliche Schwankungen in State-Selbstkonzepten spielen. Die vorliegende Studie untersucht daher, wie Trait-Erwartungen und -Werte von Schüler:innen im Sinne des Expectancy-Value-Modells von Eccles et al. (1983; Eccles & Wigfield, 2020) mit (a) der mittleren Ausprägung und (b) intraindividuellen Schwankung im momentanen (State) Selbstkonzept zusammenhängen. Als Indikatoren für habituelle Erwartungen wurden akademische Trait-Selbstkonzepte und für habituelle Werte schulisches Interesse und Prüfungsängstlichkeit (Komponenten „Worry“ und „Emotionality“; Liebert & Morris, 1967) untersucht. Dabei wurden die untersuchten Zusammenhänge sowohl auf allgemein-schulischer Ebene (z.B. „Jetzt gerade denke ich, dass ich in den meisten Schulfächern gut bin“) als auch mit dem Fokus auf dem Fach Mathematik (z.B. „Jetzt gerade denke ich, dass ich im Fach Mathe gut bin“) untersucht.

Es wurden mikrolängsschnittliche Daten mit vielen Messzeitpunkten genutzt, die über einen Zeitraum von drei Wochen mittels digitaler Tagebücher via Smartphones (Experience Sampling-Methode) bei deutschen Gymnasialschüler:innen erhoben wurden. Die Daten stammen aus einem größeren Projekt. Die hier genutzten Daten für die habituellen Erwartungen und Werte wurden vor der digitalen Tagebuchstudie erhoben (Paper/Pencil); die für die State-Selbstkonzepte nach jeder Schulstunde in den Fächern Mathematik, Physik, Deutsch und Englisch (für das allgemein-schulische State-Selbstkonzept; $N = 289$, $N_{obs} = 6,211$; $M_{lessons} = 21.49$) bzw. nach jeder Mathematikstunde (State-Matheselbstkonzept; $N = 243$, $N_{obs} = 2,075$; $M_{lessons} = 8.54$).

Die State-Selbstkonzepte (allgemein-schulisch und mathematik-spezifisch) wurden mit jeweils drei Items auf der Grundlage von Gogol et al. (2014) und Marsh et al. (1983) erfasst. Prüfungsängstlichkeit (allgemein-schulisch und mathematik-spezifisch) wurde mit jeweils fünf Items für Worry und Emotionality (Hodapp 1991; Hodapp et al., 2011), Interesse mit fünf (allgemein-schulisches Interesse) bzw. sechs (mathematik-spezifisch) Items (Pohmann et al., 2005) und Trait-Selbstkonzept (allgemein-schulisch und mathematik-spezifisch) mit jeweils sechs Items (Gogol et al., 2014; Marsh et al., 1983) erfasst.

Wir verwendeten mixed-effects location scale-Modelle, um die Zusammenhänge zwischen Trait-Selbstkonzept, Interesse und Prüfungsängstlichkeit einerseits und dem mittleren Niveau und der intraindividuellen Variabilität im State-Selbstkonzept andererseits zu untersuchen. Diese Analysen wurden jeweils einmal für die allgemein-schulischen und einmal für die mathematik-spezifischen Konstrukte durchgeführt. In unseren Analysen kontrollierten wir für Schulnoten (selbstberichteter Notenschnitt und Mathenote), schlussfolgerndes Denken (IST-Screening, Liepmann et al., 2012) und Geschlecht.

Wir fanden, dass höhere Ausprägungen in Interesse und Trait-Selbstkonzept jeweils mit höheren mittleren Ausprägungen im State-Selbstkonzept zusammenhängen. Die beobachteten Effekte waren durchweg stärker in den mathematik-spezifischen Analysen (im Vergleich zu den allgemein-schulischen Analysen). Höhere Ausprägungen im mathematik-spezifischen Interesse waren mit einer geringeren intraindividuellen Variabilität im State-Matheselbstkonzept verbunden.

Unsere Ergebnisse deuten insgesamt darauf hin, dass habituelle Merkmale von Schüler:innen in beiden Erwartungs-Wert-Komponenten eine bedeutsame Rolle bei der Herausbildung von State-Selbstkonzepten spielen.

Metakognitive Einschätzungen und Selbstkonzept: Eine Längsschnittuntersuchung bei Schüler*innen der Sekundarstufe

Emely Hoch¹, Birgit Brouër², Tim Fütterer³

¹Leibniz-Institut für Wissensmedien, Deutschland; ²Christian-Albrechts-Universität zu Kiel; ³Hector Institut für Empirische Bildungsforschung, Eberhard Karls Universität, Tübingen

Theoretischer Hintergrund

Metakognitive Prozesse, wie das Überwachen (*monitoring*) und Steuern (*control*) des eigenen Lernprozesses, spielen eine wichtige Rolle beim selbstregulierten Lernen (Nelson & Narens, 1990). Beim Monitoring überwachen Lernende ihre kognitiven Prozesse: sie beobachten, bewerten und reflektieren diese. Dieses Ergebnis nutzen sie, um ihr Lernverhalten gezielt zu steuern. Lernende haben jedoch oft Schwierigkeiten, ihren Wissenstand oder ihre Testleistung akkurat einzuschätzen (z.B. Prinz et al., 2020). In empirischen Studien wurden verschiedene informations- und erfahrungsbasierte Hinweise bezüglich der Lern- oder Testaufgaben untersucht, die metakognitiven Einschätzungen und deren Genauigkeit zugrunde liegen (z.B. Koriart, 1997; Koriart et al., 2008). Zudem wurden motivationale oder persönlichkeitsbezogene individuelle Merkmale als Einflussgrößen identifiziert (z.B. Händel et al., 2020; Kröner & Biermann, 2007). Händel et al. (2020) zeigten in zwei Studien mit studentischen Stichproben, dass insbesondere das domänenspezifische Selbstkonzept prädiktiv für metakognitive Einschätzungen und deren Genauigkeit ist. Inwiefern sich die Befunde für Schüler*innen replizieren lassen, bleibt zu untersuchen. Da das Selbstkonzept im Jugendalter noch volatil ist (Marsh et al., 2018), ist neben der Replikation der Befunde auch die Betrachtung über die Zeit in einer solchen Stichprobe interessant.

Fragestellung

Die vorliegende Studie adressiert die Frage, inwiefern sich ein Zusammenhang von (domänenspezifischem) Selbstkonzept mit lokalen und globalen metakognitiven Einschätzungen in einer Schüler*innenstichprobe replizieren lässt. Zudem soll explorativ untersucht werden, inwiefern sich dieser Zusammenhang über die Zeit verhält.

Methode

Daten wurden von 311 Schüler*innen (48% weiblich) in zwei Erhebungswellen zu drei Messzeitpunkten erhoben (Klasse 5, 6 und 7). Zu jedem Messzeitpunkt wurden das domänenspezifische Selbstkonzept (mathematisches Selbstkonzept, z.B. „Ich war

schon immer gut in Mathematik.“; $\alpha = .89$) und das akademische Selbstkonzept (z.B. „Ich kann gut lernen.“; $\alpha = .76$) je mit 3 Items auf einer 4-stufigen Antwortskala erhoben (Kunter et al., 2002). Anschließend bearbeiteten die Schüler*innen jeweils neun curricular valide Mathematikaufgaben (TIMSS: Baumert et al., 1998; DEMAT5+: Götz et al., 2013a, DEMAT6+: 2013b), die jeweils als korrekt gelöst (1) oder nicht gelöst (0) kodiert wurden. Nach jeder Aufgabe erfolgte eine lokale metakognitive Einschätzung (ob eine bestimmte Aufgabe sicher gelöst [1], weiß nicht [0,5], sicher nicht gelöst [0] wurde). Am Ende des Tests wurde die globale metakognitive Einschätzung erhoben (wie viele Aufgaben richtig gelöst wurden). Lokale und globale metakognitive Einschätzung sowie die metakognitive Genauigkeit (Abweichung der Einschätzung von der tatsächlichen Aufgabenleistung) wurden als abhängige Variablen untersucht.

Ergebnisse

Für beide metakognitiven Maße berechneten wir verallgemeinerte lineare gemischte Modelle mit Selbstkonzept, Messzeitpunkt und deren Interaktion als Prädiktoren (unter Kontrolle von Alter, Geschlecht, Leistung, der verschachtelten Struktur wiederholter Messzeitpunkte und Schüler*innen gruppiert in Klassen). Aufgrund hoher Multikollinearität ($VIF > 5$), wurden domänenspezifisches und akademisches Selbstkonzept in separaten Modellen analysiert. Leistung, domänenspezifisches und akademisches Selbstkonzept waren statistisch signifikante Prädiktoren für lokale und globale metakognitive Einschätzungen und deren Genauigkeit, alle $p < .05$, womit die Ergebnisse von Universitätsstudierenden repliziert wurden. Darüber hinaus gab es statistisch signifikante Interaktionen jeweils zwischen domänenspezifischem bzw. akademischem Selbstkonzept und den Messzeitpunkten für lokale metakognitive Einschätzungen und die lokale metakognitive Genauigkeit, alle $p < .05$. Die negativen Interaktionen deuten darauf hin, dass die Auswirkungen beider Selbstkonzeptmaße auf die lokalen metakognitiven Einschätzungen und die metakognitive Genauigkeit in den Klassenstufen 6 und 7 schwächer ausgeprägt waren als in Klasse 5.

Diskussion

Übereinstimmend mit Händel et al. (2020) replizieren unsere Ergebnisse einen Zusammenhang zwischen domänenspezifischem Selbstkonzept und metakognitiven Maßen in einer Schüler*innenstichprobe, der bei jüngeren Schüler*innen stärker ausgeprägt zu sein scheint. Ob der schwächere Zusammenhang darauf zurückzuführen ist, dass sich ältere Schüler*innen bei metakognitiven Einschätzungen stärker auf andere Hinweise verlassen (z.B. Zugänglichkeit: Roebers et al., 2019), sollte in zukünftigen Untersuchungen betrachtet werden.

Differenzierung der Zusammenhänge zwischen Selbstkonzept in Musik, Note im Schulfach Musik und musikpraktischen Kompetenzen

Daniel Fiedler¹, Johannes Hasselhorn¹, Florian Lill², Andreas C. Lehmann²

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland; ²Hochschule für Musik Würzburg

Theoretischer Hintergrund

Das Selbstkonzept ist eines der am häufigsten untersuchten Konstrukte in Psychologie und Pädagogik. Shavelson et al. (1976) definierten das Selbstkonzept als die Selbstwahrnehmung einer Person in verschiedenen Bereichen, die „[...] the totality of the individual's thoughts and feelings [...]“ (Rosenberg, 1979, S. 7) umfasst und multidimensional sowie hierarchisch strukturiert ist. Vispoel (2003) konzipierte basierend auf Shavelson et al. (1976) das Selbstkonzept in Musik als „self-perceptions of competence in music formed through experiences with the environment and interpretations of those experiences“ (S. 153). Dabei wird das Selbstkonzept in Musik eher durch die Selbstwahrnehmung von Fertigkeiten der auditiven Kognition und instrumentaler Leistungen als durch Fertigkeiten in Musik und Bewegung/Tanz oder Gesang geprägt (Fiedler et al., 2024; Morin et al., 2016; Vispoel, 2003).

Ziel der vorliegenden Studie

Basierend auf diesem theoretischen Hintergrund war das Ziel der vorliegenden Studie, die Zusammenhänge zwischen Selbstkonzept in Musik, Note im Schulfach Musik und musikpraktischen Kompetenzen zu untersuchen, um herauszuarbeiten, ob eher Erfahrungen mit der Umwelt (Schulnote) oder eher die Interpretation dieser Erfahrungen (Kompetenzen) das Selbstkonzept in Musik von Schüler:innen mit unterschiedlichen musikalischen Hintergrund prägen.

Methodisches Vorgehen

Unsere Stichprobe bestand aus 1324 Schüler:innen (50% weiblich) der Klassen 5 bis 10 an weiterführenden Schulen in Baden-Württemberg und Bayern, von denen 72% angaben, Instrumentalunterricht erhalten zu haben. Demografische Hintergrundvariablen wie etwa Geschlecht und Alter, das Selbstkonzept (Fiedler & Spychiger, 2017; Schwanzer., 2002) und die Schulnote in Musik wurden mittels Selbstauskunftfragebogen erfasst. Die musikpraktischen Kompetenzen wurden mit dem KOPRA-M (Lill et al., 2020) in den drei Dimensionen Musizieren, Rhythmusproduktion und Gesang gemessen. Die Daten werteten wir zum einen mit Korrelationsanalysen aus, um die Zusammenhänge zu untersuchen, zum anderen führten wir Strukturgleichungsmodellanalysen (SEM) durch, um die Zusammenhänge weiterführend für die beiden Subgruppen musikalisch aktive und nicht-aktive Schüler:innen zu differenzieren.

Ergebnisse

Die Korrelationsanalyse zeigte, dass das Selbstkonzept in Musik sowohl mit der Note ($r = .33, p < .01$) als auch mit den Kompetenzen ($r = .36, p < .01$) zusammenhängt, wobei die Zusammenhänge zwischen Selbstkonzept in Musik und der Dimension Musizieren der musikpraktischen Kompetenzen höher war ($r = .41, p < .01$) als die mit den Dimensionen Rhythmusproduktion ($r = .26, p < .01$) und Singen ($r = .33, p < .01$). Entsprechend zeigen sich diese Zusammenhänge – auch unter Kontrolle von Geschlecht und Jahrgangsstufe – auch im SEM. Eine Differenzierung der Zusammenhänge mittels der beiden Subgruppen ergab, dass der Zusammenhang zwischen Selbstkonzept in Musik und Schulnote in Musik sich zwischen musikalisch aktiven und nicht-aktiven Schüler:innen nicht unterscheidet ($r = .36$ und $r = .33, p < .01$), sich allerdings Unterschiede in den Zusammenhängen zwischen Selbstkonzept in Musik und musikpraktischen Kompetenzen ($r = .36$ und $r = .48, p < .01$) ergeben.

Diskussion und Implikation

Die Ergebnisse der vorliegenden Studie lassen den Schluss zu, dass Erfahrungen mit der Umwelt (Schulnote) und die Interpretation dieser Erfahrungen (Kompetenzen) für die Entwicklung des Selbstkonzepts in Musik von Bedeutung sind. Es scheint also wichtig zu sein, dass Schüler:innen sowohl praktisch gut musizieren und dass die Lehrer:innen ihnen das auch rückmelden. Dies hat wiederum Implikationen für die musikpädagogische Praxis, da nicht nur praktisch musiziert, sondern auch entsprechend individuell wertschätzendes Feedback – vor allem musikalisch nicht-aktiven Schüler:innen – gegeben werden soll.

Somit können die Ergebnisse wertvolle Hinweise für die Erarbeitung von Interventionsansätzen zur Stärkung des Selbstkonzepts in Musik geben.

Motiviert gutes Peer-Feedback geben: Ergebnisse zweier Feldexperimente zur Förderung von Motivation und Feedbackqualität

Melanie V. Keller, Martin Daumiller, Markus Dresel
Universität Augsburg, Deutschland

Theoretischer Hintergrund

Wenn Peer-Feedback als vielversprechende Lernmethode (Double et al., 2020) in der Hochschullehre eingesetzt wird, stehen Herausforderungen im Hinblick auf die Motivation der Studierenden und die Feedbackqualität im Raum (Nicol et al., 2014). Die Situated-Expectancy-Value-Theory (Eccles & Wigfield, 2020) bietet hierbei vielversprechende Ansatzpunkte, um die Motivation von Studierenden zu fördern und bei der Peer-Feedbackgabe zu unterstützen. Dabei sind insbesondere die Erfolgserwartungen der Studierenden, Peer-Feedback geben zu können, sowie der instrumentelle Wert, den sie der Feedbackgabe zuschreiben (d.h. der Wert für das eigene Lernen oder die Entwicklung von Feedbackkompetenzen), bedeutsam für Motivation und Leistung (hier: Feedbackqualität). Erfolgserwartungen könnten durch Strukturierungs- und Formulierungshilfen unterstützt werden (Bürgermeister et al., 2021), während Informationsmaterial in Kombination mit Elaborationsaufgaben zum instrumentellen Wert einer Aufgabe (hier: Feedbackgeben) beitragen kann (Hulleman & Harackiewicz, 2021).

Fragestellung

Um Motivation und Feedbackqualität in der Peer-Feedbackgabe zu fördern, entwickelten wir auf diesen Grundlagen eine niedrigschwellige Förderung und testeten diese in zwei präregistrierten Feldexperimenten.

Methode

In Studie 1 gaben sich 356 Studierende (Experimentalgruppe: 254, Kontrollgruppe: 101; 289 weiblich, Alter: $M=20.2$, $SD=2.60$ Jahre) in einem Hochschulkurs gegenseitig Feedback auf eine schriftliche Aufgabenbearbeitung; in Studie 2 gaben sich 278 Studierende (Experimentalgruppe: 174, Kontrollgruppe: 102; 172 weiblich, Alter: $M=21.0$, $SD=2.80$) zu sechs Aufgabenbearbeitungen im Verlauf des Semesters gegenseitig Feedback.

Die Studierenden der Experimentalgruppe erhielten in beiden Studien direkt vor der Feedbackgabe eine kurze Darstellung der Nützlichkeit von Peer-Feedback für das Lernen und das Trainieren von Feedbackkompetenzen und eine Anleitung für wirksames, motivierendes Feedback, die sie in einer kurzen Elaborationsaufgabe in eigenen Worten zusammenfassten. In der längsschnittlichen Studie 2 wurde diese Darstellung in sechs Teile geteilt und über die sechs Messzeitpunkte verteilt. Die Kontrollgruppe erhielt lediglich technische Anweisungen zur Feedbackgabe.

Zum Zeitpunkt der Feedbackgabe schätzten die Studierenden Erfolgserwartung, instrumentellen Wert sowie Kosten in Bezug auf die Feedbackgabe ein, zudem erfassten wir die Länge des Feedbacks (computerbasiert) und die Feedbackqualität. Für letzteres kodierten zwei trainierte Rater die Feedbacks entlang eines Kriterienkataloges (Übereinstimmung: $\kappa=.84$). Um die Studierendenperspektive einzubeziehen, wurde die Qualität des Feedbacks zudem durch die Studierenden, die das Feedback erhielten, eingeschätzt.

Studie 1 wurde mit Varianzanalysen (nrmv 2.4.0; Burchett et al., 2017), Studie 2 mit Wachstumskurvenmodellen (lavaan 0.6.17; Rosseel et al., 2023) ausgewertet.

Ergebnisse und Bedeutung

In Studie 1 schrieben die Studierenden der Experimentalgruppe der Feedbackgabe einen statistisch signifikant höheren instrumentellen Wert, aber auch signifikant höhere Kosten zu, $F(3, 737)=4.616$, $p=.002$, während es keine signifikanten Unterschiede in den Erfolgserwartungen gab. Die Feedbacks der Experimentalgruppe hatten eine statistisch signifikant höhere Qualität, waren allerdings nicht signifikant länger als die der Kontrollgruppe ($F(3, 500)=20.424$, $p<.001$).

In Studie 2 zeigten Wachstumskurvenmodelle einen schwach positiven zeitlichen linearen Trend für Erfolgserwartungen und einen schwach negativen Trend für Feedbacklänge und -qualität über die sechs Messzeitpunkte hinweg, jedoch keine Trends für instrumentellen Wert oder Kosten. Die Gruppenzugehörigkeit zur Experimental- oder Kontrollgruppe sagte weder den Intercept (Mittelwert) noch die Steigung (zeitliche Änderung) von Feedbacklänge oder -qualität, instrumentellem Wert oder Kosten statistisch signifikant vorher. Dagegen waren Erfolgserwartungen für die Feedbackgabe in der Experimentalgruppe statistisch signifikant höher als in der Kontrollgruppe und stiegen über die Zeit auch stärker an.

Die Ergebnisse von Studie 1 weisen darauf hin, dass niedrigschwellige Förderungen wie diese grundsätzlich den instrumentellen Wert des Feedbackgebens im Peer-Feedback sowie Feedbackqualität steigern könnten. Dagegen könnten die abweichenden Ergebnisse in Studie 2 darauf zurückzuführen sein, dass die Förderung nur dann die erwartete Wirkung entfaltet, wenn sie als Ganzes eingesetzt und nicht über sechs Messzeitpunkte verteilt wird. Studie 2 wirft zudem die Frage auf, ob mit einer Verteilung der Tipps zur Feedbackgabe über die Zeit in kleineren Portionen die Erfolgserwartungen der Studierenden besser gesteigert werden könnten als durch einen einmaligen Einsatz. Dieser Frage kann in weiteren Studien nachgegangen werden.

S-01: Bridging Minds: A Transdisciplinary Approach on Executive Functions, Metacognition, Self-Regulation, and Self-Regulated Learning in Education Research

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 129

Chair der Sitzung: Friederike Blume

Chair der Sitzung: Maria Theobald

Chair der Sitzung: Charlotte Dignath

Chair der Sitzung: Julia Kerner auch Körner

English Track

Bridging Minds: A Transdisciplinary Approach on Executive Functions, Metacognition, Self-Regulation, and Self-Regulated Learning in Education Research

Chair(s): **Friederike Blume** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland), **Maria Theobald** (Universität Trier, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland), **Charlotte Dignath** (Goethe-Universität Frankfurt, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland), **Julia Kerner auch Körner** (Helmut-Schmidt-Universität Hamburg, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland)

The ability to withhold impulsive responses (i.e., inhibition), to retain and process information (i.e., updating), and to shift attention between tasks (i.e., shifting) are collectively referred to as *executive functions* (Miyake & Friedman, 2012). It is hypothesised that executive functions form the cognitive basis of several higher-order functions, which can be ranked in ascending order of complexity: *metacognition*, which refers to an individuals' capacity to reflect on their cognitive processes (Flavell, 1979), *self-regulation*, which denotes the ongoing, domain-general process of aligning one's thoughts, emotions and actions with pre-defined objectives (Baumeister et al., 1994; Karoly, 1993), and *self-regulated learning*, the process whereby learners act as active participants in their own learning, systematically directing cognitions, affect, and behaviour towards the achievement of learning goals (Zimmerman & Schunk, 2011). All these concepts are central for lifelong learning and academic success (Dent & Koenka, 2016; Diamond, 2013; Robson et al., 2020; Veenman et al., 2004). Additionally, they provide the basis for dealing with current societal challenges, such as the increasing digitalisation of schools (e.g., Daumiller & Dresel, 2019; OECD, 2019).

Nevertheless, while the concepts themselves should be theoretically and empirically related, they have primarily been of interest only in specific research domains (e.g., cognitive psychology, developmental psychology, health psychology, education sciences), with few exceptions (cf. Greene, 2017). Consequently, interrelations among these concepts and joint contributions to explaining learning and academic success have only been insufficiently examined. Furthermore, if interrelations have been investigated, the concepts have usually been operationalised differently (e.g., executive functions measured using cognitive tasks vs. using questionnaires), thus preventing the discussion and acceptance of findings across disciplines. Additionally, research designs traditionally employed in one research domain have rarely been applied in other domains (e.g., field studies in education research vs. laboratory studies in cognitive psychology). The objective of the present symposium is therefore to facilitate interdisciplinary approaches to research on executive functions, metacognition, self-regulation, and self-regulated learning, transcending the boundaries of traditional research domains. The present symposium will include four contributions, which will be discussed based on a set of guiding questions.

The first contribution will present findings from a study examining the relationship between executive functions and metacognitive monitoring. This will contribute to elucidating the significance of executive functions for metacognition and academic success, linking the fields of cognitive and developmental psychology, as well as educational sciences. The second contribution, situated within the fields of cognitive psychology and educational sciences, will present results from an investigation comparing the power of different facets of executive functions and self-regulation in predicting school performance. The findings highlight the importance of self-regulation in academic achievement. The third contribution, which seeks to establish a link between cognitive psychology and educational sciences, aimed to examine individual differences in intra-individual effects of momentary affect, lesson-specific interest, and success experiences on working memory performance. To this end, analyses will be conducted on intensive longitudinal data, which have gained prominence in cognitive psychology and educational sciences only recently. The contribution provides insights into the short-term effects of individual classroom experiences on cognitive performance and illustrates the role of individual differences in self-regulation in these relationships. The fourth contribution will present findings of a study investigating the role of executive functions and analogical reasoning in the use of educational apps to support self-regulated learning. Situated within cognitive and motivational psychology, as well as education and computer sciences, the study provides interdisciplinary insights into the optimisation of learning processes through individualised, digital interventions. The discussion will aim at theoretical clarification by highlighting similarities and differences between executive functions, metacognition, self-regulation and self-regulated learning.

Beiträge des Symposiums

Are executive functions needed to be a metacognitive learner? Experimentally investigating why executive functions and metacognitive monitoring are associated

Donna Bryce¹, Janina Eberhart²

¹Universität Augsburg, Deutschland, ²Eberhard Karls Universität Tübingen, Deutschland

Challenging academic tasks make demands on a range of self-regulatory skills. For instance, to maintain a goal in mind, resist distractions and switch flexibly between solutions a learner uses their executive functions (EF, a subcomponent of self-regulation). To self-monitor their understanding and progress, and adjust in the face of challenges, a learner uses their metacognitive skills (a subcomponent of self-regulated learning). Fernandez-Duque and colleagues (2000) first hypothesised a theoretical connection between EFs and metacognition which has since been empirically supported by correlational studies that highlight positive relationships between children's EFs and metacognitive skills (e.g., Bryce et al., 2015; Kälin & Roebbers, 2022). It has been suggested that EFs are necessary but not sufficient precursors of metacognitive skills and there is evidence that metacognitive skills are a stronger predictor than EFs of early school success (Bryce et al. 2015; Dutemple et al., 2023). However, until recently there has been a paucity of empirical evidence regarding the nature of the relationship between EFs and metacognition. Previous

findings with adults and the EF component working memory (WM) indicated that when the WM demands of a task increased, the accuracy of monitoring judgments suffered (Bryce et al., 2023) indicating that sufficient WM resources are required to make accurate monitoring judgments. In the current project, we aimed investigate why EFs and metacognitive monitoring are associated in children using a similar approach.

Using a cross-sectional and longitudinal design we tested whether two EFs, updating of WM and inhibition, have a dependent relationship with metacognitive monitoring and whether this relationship changes with age. Children in the first ($n=32$, $M_{Age}=7.3$) and third ($n=26$, $M_{Age}=9.0$) grade in Germany completed two computerised metacognitive EF tasks. The first graders completed the tasks again after an 18-month delay. In each task, children processed a typical EF task and provided monitoring judgments via a scale about their performance after each trial (either confidence judgments or estimated reaction times). We manipulated the EF demands of each EF task to investigate whether this affected the accuracy of the children's monitoring judgments. Monitoring accuracy for each child and each level of EF demand was calculated as the Pearson correlation between performance and monitoring judgment, indicating sensitivity to trial-by-trial variations in one's performance. To assess whether monitoring accuracy is affected by the EF demands, data were analysed using mixed-design ANOVAs with the within-subject factor EF Demand (low vs. high) and the between-subject factor Age (1st vs. 3rd grade) for the cross-sectional analyses or Time (1 vs. 2) for the longitudinal analyses. We hypothesised that if EFs are required for making monitoring judgments, monitoring judgments will be less accurate when EF demands increase; no specific hypotheses regarding developmental effects were stated.

In the present study the hypothesised dependency between EF and monitoring was observed only for WM in the longitudinal data. That is, monitoring accuracy suffered when WM demands increased at Time 2, but not at Time 1, $F(1,24)=5.91$, $p=.023$. Surprisingly, in the cross-sectional analyses we observed only a main effect of WM demand on monitoring accuracy, whereby monitoring accuracy slightly improved when WM demands increased, $F(1,55)=4.57$, $p=.037$. Interestingly, increasing the inhibition demands of the EF task had no effect on monitoring accuracy in either age group or at either testing time. These findings highlight that different EFs are related to metacognition for different reasons. To understand mechanisms, we need to look at specific components of metacognition and EF. While WM may help a learner to form accurate monitoring judgments (because many different sources of information can contribute to judgments), inhibition may allow a learner to engage in monitoring behaviours (because they can better pause to monitor themselves).

Comparing cognitive and emotional facets of self-regulation in their prediction of academic achievement

Julia Kerner auch Körner¹, Henning Heinze²

¹Helmut-Schmidt-Universität Hamburg, Deutschland IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland, ²Helmut-Schmidt-Universität-Hamburg, Deutschland

Theoretical background

Self-regulation refers to the ongoing, domain-general process of directing one's thoughts, emotions and actions towards a predetermined goal (Baumeister et al., 1994; Karoly, 1993) and is an important predictor of academic achievement (Diamond, 2013; Robson et al., 2020). According to the integrated model of self-regulation for preschool (Jones et al., 2016), self-regulation includes both cognitive aspects such as executive functions (inhibition, working memory, shifting) and emotional aspects such as effortful control (e.g. delay of gratification). While meta-analyses compare different facets of executive functions (Spiegel et al., 2021) or effortful control (Nasvytienė & Lazdauskas, 2021), cognitive and emotional aspects of self-regulation are rarely compared in their prediction of academic achievement in large longitudinal studies.

Research question

The aim of this study is therefore to compare different operationalizations of self-regulation to find out which one best predicts academic achievement.

Method

In the newborn sample of the National Educational Panel Study (NEPS; $N = 3481$, 1706 girls, 1775 boys), self-regulation was assessed using various instruments: Inhibition by a flanker task at the age of 4 years, working memory by a number span at the age of 3 and at 6 years, delay of gratification by a wait task at the age of 3 and 5 years, and the temperament trait effortful control was rated by parents at the age of 3, 4 and 5 years using three items of the child behavior questionnaire (Putnam & Rothbart, 2006). Academic achievement was assessed with tasks in science and language comprehension at the end of first grade and tasks in math at end of the second grade.

Results and Implications

Inhibition and working memory correlate with later performance in math ($r = .165^{**}$ and $.314^{**}$), science ($r = .141^{**}$ and $.262^{**}$) and language ($r = .060^{*}$ and $.172^{**}$). Delay of gratification and effortful control correlate significantly less with academic achievement ($r = .015$ to $.138^{**}$). This suggests that executive functions as cognitive facets of self-regulation are better suited to predict later academic achievement than emotional facets of self-regulation. However, limitations of the measurement must also be considered, as an adult was present during the delay of gratification task and effortful control was only assessed with three of the 12 items of the original scale. The results will be discussed in the framework of the integrated model of self-regulation (Jones et al., 2016) and in the context of early identification of children at risk for academic development.

ENTFÄLLT: Exploring Determinants of Working Memory Variability in Classroom Settings: The Role of Interest and Classroom Management

Lena M. Wieland¹, Florian Schmiedek²

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland; Goethe-Universität Frankfurt, Deutschland

Among executive functions, working memory in particular has been shown to be associated with students' academic performance (e.g. Friso-Van Den Bos et al., 2013; Blume et al., 2021; Lechuga et al., 2014). In addition, there are findings that demonstrate not only individual differences in students' working memory (WM) performance, but also considerable within-person variation (e.g. Dirk & Schmiedek, 2016; Galeano-Keiner et al., 2022; Riediger et al., 2014). However, there is limited research that addresses influential determinants of systematic variance in students' working memory performance in classroom settings. The present study draws on a data set of ($N = 225$) students from grades 5 and 6 who participated in ecological momentary assessments via smartphones over the course of 20 school days, including four measurements per day (in the morning, at school, in the afternoon and in the evening). Work memory measures were available for a total of ($n = 2558$) class sessions. In this contribution, we focus

on students' perceived classroom interest and classroom management as relevant indicators for their working memory performance. Based on the theory of cognitive resources by Kanfer and Ackerman (1989), we assume that interested students use their cognitive resources more efficiently and thus achieve better WM results (see also Ainley et al., 2002; Hidi & Renninger, 2006; Schiefele, 1991). Accordingly, it was expected that the students would show better WM performance after the lessons in which they reported an above-average interest in the class content (3 items derived from Marsh et al., 2005). In addition, teachers' competence in managing students' behavior is relevant to students' motivation and engagement. Research found that primary school teachers with better classroom organization skills had more engaged students (Clunies-Ross et al., 2008; Rimm-Kaufman et al., 2015). In secondary schools, classroom management strategies have also been associated with greater student interest and emotional engagement (Hochweber et al., 2013; Kunter et al., 2007). On the basis of the available intensive longitudinal data, we therefore additionally investigate whether students' individual WM accuracy was found to be positively influenced by good perceived classroom management, reported at the end of the lesson (3 items, derived from Fauth et al., 2014). In addition, we exploratively investigate how WM accuracy, which was recorded during school hours, affects school-related experiences of success and failure reported by students in the afternoon (3 items each, derived from Neubauer & Voss, 2016). The analysis of the intensive longitudinal data takes into account fixed and random effects in a linear mixed model. Controlling for individual differences in students' self-regulation, which was assessed by parental responses using the Strengths and Difficulties Questionnaire (SDQ), but also measures of general school self-concept, will be an essential part of the analysis to gain a more comprehensive understanding of the dynamics involved.

App-basierte Lernaktivitäten im Vergleich: unter Laborbedingungen und in der (digitalen) Praxis

Daniel Biedermann¹, Jasmin Breitwieser², Lea Nobbe², Hendrik Drachsler³, Garvin Brod⁴

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland, ³DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; Goethe-Universität Frankfurt, Deutschland, ⁴DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; IDeA Research Center for Individual Development and Adaptive Education of Children at Risk, Deutschland; Goethe-Universität Frankfurt, Deutschland

Bildungs-Apps sind bereits heute ein alltäglicher Bestandteil des schulischen und außerschulischen Lernens von Kindern. Neben gängigen Inhalten wie Vokabellernen bieten sie auch die Möglichkeit mobiler Interventionen, die Lernende beim selbstregulierten Lernen (SRL) unterstützen. So nutzten Breitwieser und Kolleg*innen (2023) eine App, um Fünftklässler*innen an selbst erstellte Wenn-Dann Pläne zu erinnern. Durch diese Erinnerungen konnte verteiltes Lernen bei den Teilnehmenden erhöht werden.

Wenn-Dann Pläne verknüpfen eine auslösende Situation mit einer zielgerichteten Handlung, um die Verbindung zwischen Situation zu stärken und das zielgerichtete Handeln zu automatisieren (Gollwitzer, 1999). Während die positive Wirkung dieser Planstrategie vielfach belegt ist (Sheeran et al., 2024), birgt ihre Anwendung auch Herausforderungen. Eine davon ist, dass Lernende in der Lage sein müssen sich ihren Plan zu merken, damit er seine positive Wirkung entfaltet.

Die Gedächtnisleistung für Pläne sollte dabei auch davon abhängen, über welche Aktivitäten sie sich eingeprägt wurden. Diese Aktivitäten wurden bislang nicht systematisch verglichen. Unser Ziel war daher der Vergleich unterschiedlicher Lernaktivitäten, die sich für eine appbasierte Präsentation eignen. Dies entwickelten wir basierend auf dem ICAP-Framework (Chi et al., 2014), so dass sie verschiedene tiefe Verarbeitungen von Information beinhalteten. Dabei gingen wir davon aus, dass sich entwickelnde exekutive Funktionen und analoges Schlussfolgern wichtige Grundlagen bilden, um von den komplexeren Lernaktivitäten zu profitieren (Brod, 2021).

In einer Längsschnittstudie präsentierten wir 106 Schüler*innen (9-14 Jahre, $M = 11.80$, $SD = 1.04$) in einer selbstentwickelten App 27 Wenn-Dann Pläne über 27 Tage. Jeden Morgen wurde den Kindern ein neuer Plan präsentiert und nach mindestens sechs Stunden die Gedächtnisleistung in einem Free Recall abgefragt. Eingeprägt wurden sich die Pläne durch drei unterschiedliche Lernaktivitäten: 1) dreifaches Lesen (nach ICAP: passiv), 2) Zusammenpuzzeln des Plans aus den einzelnen Worten, nachdem er zuvor gelesen werden konnte (aktiv) und 3) Auswahl von Emojis zur Repräsentation des Plans (konstruktiv). Dabei wurden jeweils drei Pläne mit derselben Aktivität gelernt und dann die Aktivität gewechselt. Die Reihenfolge der Pläne und Aktivitäten war zwischen Kindern ausbalanciert. Nach Ende der Studie wurden die Fähigkeit zum analogen Schlussfolgern (Analogical Reasoning Test des SON-R 6-40 (Tellegen et al., 2012)) sowie Exekutive Funktionen (modifizierte Hearts and Flowers Task (Wright & Diamond, 2014)) erfasst.

Entgegen der Erwartung erinnerten Kinder den Plan 1.65 Mal eher durch die passive als aktive und 1.47 Mal eher durch die konstruktive als aktive Lernaktivität, während sich zwischen der passiven und konstruktiven Lernaktivität kein Unterschied fand. Es zeigten sich allerdings große interindividuelle Unterschiede zwischen den Kindern, die wir darauffolgend zu erklären versuchten. Die Effektivität der aktiven Lernaktivität steigerte sich, wenn Kinder mehr Zeit mit ihr verbrachten ($\chi^2(2) = 17.45$, $p < .001$); gleiches gilt für die konstruktive Lernaktivität ($\chi^2(2) = 8.19$, $p = .017$). Die Zeit, die Kinder mit dem Einprägen verbrachten, wurde wiederum durch ihre Fähigkeit zum analogen Schlussfolgern vorhergesagt. So verbrachten Kinder mit besseren Fähigkeiten im Bereich des analogen Schlussfolgerns mehr Zeit mit der konstruktiven Lernaktivität. Für die exekutiven Funktionen zeigte sich nur für Inhibition ein leicht negativer Zusammenhang mit der Gedächtnisleistung in der konstruktiven Bedingung.

Zusätzlich zur bereits durchgeführten Studie planen wir die Befunde im Labor zu replizieren. Wiederum werden Kinder im Alter von 10-13 Jahren sich 27 Pläne mithilfe der passiven, aktiven und konstruktiven Lernaktivität einprägen, dies jedoch über nur eine Session. Zusätzlich wird auch die Fähigkeit zum analogen Schlussfolgern sowie exekutive Funktionen wieder erhoben. Zur Konferenz werden die Daten ausgewertet sein, was es erlaubt, die Ergebnisse der beiden Studien zu vergleichen.

Die Ergebnisse der Längsschnittstudie deuten darauf hin, dass App-basierte Lernaktivitäten je nach Lernvoraussetzungen unterschiedlich effektiv sind, was für eine individualisierte Vorgehensweise spricht. Basierend auf den Daten der Laborstudie wird diskutiert, inwieweit diese Schlussfolgerung hält und was Implikationen für die Praxis sind.

S-02: Facets of early self-regulation: Valid assessment, short- and long-term dynamics and links to social development

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 133

Chair der Sitzung: Manja Attig

Chair der Sitzung: Jessica Willard

Chair der Sitzung: Hanna Mulder

English Track

Facets of early self-regulation: Valid assessment, short- and long-term dynamics and links to social development

Chair(s): **Manja Attig** (Leibniz Institute for Educational Trajectories, Germany), **Jessica Willard** (Leibniz Institute for Educational Trajectories, Germany)

Diskutant*in(nen): **Hanna Mulder** (Utrecht University)

Self-regulation is a central predictor of academic success, socio-emotional skills and positive life outcomes (e.g., Cortés Pascual et al., 2019; Robson et al., 2020). This broad construct encompasses the processes underlying the regulation of emotion, cognition, attention and behavior and subsumes various concepts including the primarily cognitive executive functions, attentional control as a cognitive process involving emotional involvement and the emotion-involved delay of gratification (Bailey & Jones, 2019; Jones et al., 2016). Precursors of self-regulation are evident already early in life. Accordingly, self-regulation has been the focus of diverse disciplines, research traditions, and methodologies. Within the LERN-Network, the Leibniz Center of Excellence for Early Childhood education brings together interdisciplinary perspectives on self-regulation that are highlighted in this symposium. The symposium presents psychometric, experimental and large-scale studies that are furthering the assessment of self-regulation, the understanding of its underlying processes and its links to the development of socially (mal)adaptive behaviors.

Study 1 evaluates and standardizes the freely available U.S. test battery “EF Touch” on a representative sample of German-speaking preschoolers (3 to 5 years old) and discovers favorable psychometric properties and evidence of validity, suggesting that this measure can be successfully used to collect norm-referenced data for this population. Study 2 is the first to investigate the mechanisms underlying audio-visual distraction during the preschool age. Using a novel paradigm and comparing different age-groups (4 to 10 years old), the study pinpoints a period of maturation of attentional control with direct implications for early education classrooms. Study 3 disentangles how various facets of self-regulation at 3 to 5 years of age contribute to social development from the preschool to the elementary school years over and above a rich set of child and family covariates, thus shedding light on how self-regulation can facilitate the social well-being during the challenging transition to formal education. Using a large sample, this longitudinal study is able to show that the facets of self-regulation may be differentially connected to social development. The discussant is a renowned expert in childhood self-regulation and will discuss the significance of findings for research and the role of self-regulation in children's lives.

Beiträge des Symposiums

Validation of the EF Touch test battery for the assessment of executive functions in German-speaking preschool children

Bianca Ulitzka¹, Monika Daseking¹, Susanne Enke², Anne Eppinger Ruiz de Zarate³, Caterina Gawrilow³, Cathérine Gunzenhauser⁴, Henning Heinze¹, Julia Karbach⁵, Henrik Saalbach², Julia Kerner auch Koerner¹

¹Pädagogische Psychologie, Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, Deutschland; Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Deutschland, ²Pädagogische Psychologie, Universität Leipzig, ³Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Deutschland; Schulpsychologie, FB Psychologie, Eberhard Karls Universität Tübingen, Deutschland, ⁴Institut für Psychologie, Fakultät für Erziehungs- und Gesellschaftswissenschaften, Pädagogische Hochschule Ludwigsburg, ⁵Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Deutschland; Entwicklungspsychologie und Pädagogische Psychologie, Fachbereich Psychologie, Kaiserslautern-Landau (RPTU), Landau, Deutschl

Theoretical background

The ability of self-regulation develops significantly during preschool age (Zelazo et al., 2018). Skills such as the ability to wait and switch between different tasks are categorised as executive functions (Miyake, 2000; Roebbers, 2017). Preschool children with well-developed executive functions have more favourable individual prerequisites for school readiness (Bierman et al., 2008), academic learning success (Becker et al., 2014) and better mental health (Fergusson et al., 2013; Moffitt et al. 2011). However, a comprehensive assessment of executive functions in children of preschool age is not yet possible due to the lack of a test battery for measuring executive functions in German-speaking countries (Ulitzka, 2022).

Research question

This study aimed to evaluate and standardize the computerized test battery, EF Touch (Willoughby & Blair, 2015), a test battery to measure executive functions, for German speaking preschool children.

Method

Data collection for the representative sample of 442 children (215 girls; 220 boys; mean age = 4.10 years, SD = 0.87) aged between 3;0 and 5;11 years was completed in April 2023. The children completed eight EF Touch tasks on a touchscreen over two testing sessions, alongside additional validation tasks. Parents assessed their children's executive functions and psychopathological symptoms using questionnaires.

Results and their significance

The findings from the construction sample (n = 90) have already been published (Ulitzka et al., 2023). In the representative sample of 442 children almost all EF Touch tasks demonstrated acceptable to very good internal consistencies ($\alpha = .620-.925$, $n = 336-442$) and significant inter-correlations ($r = .192-.904$, $n = 270-430$). Older children performed better across all tasks. No gender effects were found in the results of the tasks, only in the shifting task girls performed better than boys. Additionally, convergent validity was supported by significant correlation with parental assessments of executive functions as measured by

the Behavior Rating Inventory of Executive Function – Preschool Version ($n = 415-422$, $r = -.212^{**}$; BRIEF-P, Daseking & Petermann, 2013) and another behavioural measure of executive function the Head-Toes-Knees-Shoulders task ($n = 97$, $r = .448^{**}$; HTKS; McClelland et al., 2014; Ponitz et al., 2008). Furthermore, calculations showed significant correlations with a range of symptoms of psychopathology assessed with the Strengths and Difficulties Questionnaire ($n = 406$, $r = -.300^{**}$ to $-.126^{**}$; SDQ; Goodman, 1997) as well as with performance (raw scores) in the Wechsler Preschool and Primary Scale of Intelligence - Fourth Edition ($n = 295$, General IQ: $r = .523^{**}$; Vocabulary Inquisition: $r = .494^{**}$; WPPSI-IV; Wechsler, 2018). Children that did wait and did not wait in a gift task differed in their executive functions ($n = 28$, $U = 138$, $p = .021$) but children that did wait and did not wait in a Marshmallow task did not ($n = 28$, $U = 121$, $p = .294$). All correlations were controlled for age.

These results from the evaluation of the test battery's quality criteria indicate that the instrument is suitable for measuring executive functions in preschool-aged children in German-speaking regions. Assessing executive functions before school entry could facilitate the promotion and thus the improvement of these skills, as well as potentially allow for the postponement of school entry.

Can't ignore – Control of attentional distraction in children aged 4-10 years

Nicole Wetzel¹, Florian Scharf², Andreas Widmann³

¹RG Neurocognitive Development, Leibniz Institute for Neurobiology, Magdeburg; University of Applied Sciences Magdeburg-Stendal, Stendal, ²Department of Psychology, University of Kassel, Kassel, ³Wilhelm Wundt Institute for Psychology, Leipzig University, Leipzig

Adequate self-regulation is an important factor in successful learning. Kindergarten and elementary school years are of particular interest, as important self-regulation skills are undergoing significant development. A key self-regulatory process is attention control (Kaunhofen et al., 2017). Attention control develops throughout childhood, with different aspects of attention following different developmental trajectories (Amso & Scarif, 2015; Hoyer et al., 2021). Selective attention includes the selection of task-relevant information and ignoring of task-irrelevant information. Auditory information has privileged access to perception and consciousness and needs to be controlled by attention. Attention control in the context of task-irrelevant sounds is especially important in typical learning environments such as classrooms. Despite this important function, little is known about the development of auditory attention in early childhood. In a series of studies, we examined the distraction of attention by unexpected and task-irrelevant sounds (Wetzel et al., 2019; Volkmer et al., 2022).

Four groups of children aged 4 to 10 years and one adult group participated in an audio-visual distraction paradigm ($N = 106$). Participants performed a visual categorization task while a task-irrelevant sound sequence consisting of frequently repeated standard sounds and infrequent environmental novel sounds (distractors) was presented. Multilevel analyses revealed increased reaction times in the visual categorization task when distractor sounds were presented. This indicates involuntary distraction of attention. The distraction effect decreased with age and differed significantly between age groups. This distraction effect extended to the subsequent trial in the 4- and 5-year-olds. This suggests a delayed reallocation of attention back to the visual categorization task in the youngest groups. In sum, the results demonstrate a significant maturation of attention control within a few years during early childhood that continues throughout middle childhood. We then pooled data from three studies ($N=320$) using the same experimental protocol and examined the short-term dynamics of distraction control in children aged 4-10 years and in adults. Results from multilevel models again revealed slower responses in the categorization task when distractor sounds were presented. This distraction effect was more pronounced in children than in adults. Throughout the experimental session, the initially strong distraction effects declined to the level of adults in the groups of 6- to 10-year-olds. No such decline was observed in the groups of the 4- and 5-year-olds, who showed consistently high levels of distraction throughout the session. In contrast, adults showed constantly low levels of distraction independent from exposure.

The results suggest that control of attentional distraction is a highly dynamic process that differs qualitatively and quantitatively between age groups. We conclude that the analysis of short-term dynamics provides valuable insights into the development of attention control. Furthermore, models of attention control may need to be refined to include age-dependent rapid learning mechanisms. Implications for the design of age-appropriate learning environments will be discussed.

Early self-regulation and social development: Understanding the role of developmental associations in the childhood years

Aashna Doshi¹, Wei Huang², Sabine Weinert²

¹Bamberg Graduate School of Social Sciences, University of Bamberg, ²Department of Developmental Psychology, University of Bamberg

Children's self-regulatory abilities are considered beneficial for social development during early childhood (Jones et al., 2016; Liew et al., 2012; Zhou et al., 2012). Despite the suggested associations, the role of various self-regulatory facets - specifically, emotionally neutral executive functions (EF) and more emotion-related aspects - in the development of prosocial behaviour and peer relationships remains largely unexplored (O'Toole et al., 2016; Paulus et al., 2015; Rademacher et al., 2022; Traverso et al., 2020; William et al., 2017). This study aims to analyse how preschool self-regulation (ages 3 to 5) predicts social development from preschool through primary school years, considering several factors: (i) inclusion of controls, (ii) incorporation of social development at age 3, and (iii) analysis of the separate associations of emotionally neutral EF and emotion-related facets with social development. To analyse these associations, approximately 2,000 children from the National Educational Panel Study – Newborn Cohort Study (NEPS-SC1), a large-scale longitudinal study conducted in Germany, were included. Self-regulatory facets assessed between ages 3 and 5 included measures of phonological working memory, inhibitory control, cognitive flexibility, effortful control, and delay of gratification. Social development, encompassing prosocial behaviour and peer relationships, was assessed at ages 5, 7, and 9. Control variables included receptive vocabulary, negative affectivity, surgency, non-verbal cognitive functioning, home interaction language, socio-economic status, and gender. Additionally, prosocial behaviour and peer relationships at age 3 were included in one of the analyses. To analyse these associations, several conditional growth curve models were conducted. The results show cognitive flexibility, delay of gratification, and effortful control to be significant predictors of the linear and quadratic slope of prosocial behaviour. Although a significant association existed between effortful control and peer relationships, it did not remain significant after including covariates in the model. While still under development, the study also aims to further elucidate these findings in relation to other developmental associations (e.g., parental behaviours or academic achievement). This study aims to analyse the complex interplay between self-regulation and various developmental factors. Overall, this body of research offers valuable insights into the critical role of early self-regulation in shaping various aspects of childhood development. By highlighting the interconnectedness of various developmental aspects, the study paves the way for targeted interventions that support holistic child development.

S-03: Methodische Innovationen zur Nutzung computerisierter adaptiver Tests in Bildungskontexten

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 135

Chair der Sitzung: Joachim Ulf Kröhne

Chair der Sitzung: Andreas Frey

Chair der Sitzung: Johannes Hartig

Methodische Innovationen zur Nutzung computerisierter adaptiver Tests in Bildungskontexten

Chair(s): **Ulf Kröhne** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland), **Andreas Frey** (Goethe-Universität Frankfurt am Main)

Diskutant*in(nen): **Hartig Johannes** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland)

Das Messen bildungsrelevanter Merkmale ist ein zentrales Element der empirischen Bildungsforschung. Der Differenziertheitsgrad der Ergebnisse, die statistische Power und die Belastbarkeit abgeleiteter Schlüsse hängen direkt von der Qualität der Messungen ab. Eine effiziente Art psychometrischer Messung ist das computerisierte adaptive Testen (CAT). Bei CAT orientiert sich die Auswahl der zur Bearbeitung vorgelegten Items an den zuvor gegebenen Antworten der getesteten Person. Dabei können jegliche in Echtzeit bewertbare Antwortformate genutzt werden. Um Testwerte miteinander vergleichen zu können, obwohl die Testpersonen unterschiedliche Items bearbeitet haben, werden bei CAT Itempools eingesetzt, die mit einem Modell der Item Response Theory (IRT) kalibriert wurden. Die Forschung zu CAT geht bis in die 1970er Jahre zurück. Bis heute konnte ein solider und breiter Kenntnisstand erreicht werden, der einen Routineeinsatz von CAT möglich macht. Es bestehen aber noch zahlreiche Möglichkeiten zur weiteren Optimierung adaptiver Tests, die sich aus dem konkreten Anwendungskontext ergeben können. Zudem ist das Verständnis der motivational-emotionalen Effekte von CAT bislang noch lückenhaft und teilweise von Fehlannahmen geprägt.

Das vorliegende Symposium vereinigte drei Arbeiten zu aktuellen Forschungsthemen im Bereich CAT. Die ersten beiden widmen sich der Grundlagenforschung zu neuen CAT-Methoden. Die dritte ist eine empirische Studie zu motivationalen Effekten von adaptiven Testdesigns im Kontext eines operativen low-stakes Assessments.

Im ersten Beitrag mit dem Titel „Psychometrische Modelle für Answer-Until-Correct Items bei computerisierten adaptiven Tests“ wird gezeigt und mit einer Simulationsstudie untersucht, wie das spezielle Antwortformat bei CAT genutzt werden kann. Dieses Antwortformat ist für die empirische Bildungsforschung interessant, da es im Vergleich zu üblichen Antwortformaten ein besseres emotionales Testerleben und eine höhere Messpräzision erwarten lässt.

Der zweite Beitrag mit dem Titel „Ein Propensity Score Matching Ansatz für die Analyse von Differential Item Functioning (DIF) in computerisierten adaptiven Tests“ greift die bislang wenig Erforschte Thematik auf, wie bei adaptiv erhobenen Antwortdaten verzerrungsfreie DIF-Analysen realisiert werden können. Übliche DIF-Methoden sind nicht gut nutzbar, da die Itemvorgabehäufigkeiten stark zwischen den auf DIF zu untersuchenden Gruppen variieren kann. Es wird eine neue Methode vorgestellt und mit einer Simulationsstudie untersucht.

Der dritte Beitrag mit dem Titel „Ein adaptiver Test für VERA – Design, Befunde und die Frage nach der Instruktion“ untersucht die Nutzung eines Multi-Stage-Tests bei der Vorgabe der Vergleichsarbeiten. Im Rahmen einer empirischen Studie ($N = 857$) werden Machbarkeit und Effekte auf Testbearbeitungsmotivation und Testleistung untersucht.

Die Beiträge werden von einem ausgewiesenen Experten Deutschlands zur Kompetenzmodellierung in einen größeren Rahmen eingeordnet und diskutiert.

Beiträge des Symposiums

Psychometrische Modelle für Answer-Until-Correct Items bei computerisierten adaptiven Tests

Cosima Schenk, Aron Fink, Patrick Naumann, Andreas Frey

Goethe-Universität Frankfurt am Main

Theoretischer Hintergrund

Computerisiertes adaptives Testen (CAT) stellt eine spezielle Art des Testens dar, bei der sich die Auswahl der zur Bearbeitung vorgelegten Items an den bisherigen Antworten einer Person orientiert. Es wird jeweils dasjenige Item aus einem mit einem Modell der Item Response Theorie (IRT) kalibrierten Itempool vorgelegt, das ein statistisches Optimalitätskriterium am besten erfüllt. Dies führt dazu, dass die Getesteten Items vorgelegt bekommen, die gut zu ihrem Fähigkeitsniveau passen und resultiert in einer erheblichen Steigerung der Messeffizienz im Vergleich zum nicht adaptiven Testen (Frey, 2023). Empirische Befunde legen zudem nahe, dass CAT sich insbesondere dann positiv auf das emotionale Erleben der getesteten Personen während der Testung auswirkt, wenn Personen so viele Antwortversuche bekommen wie benötigt werden, um eine Aufgabe korrekt zu lösen (Schenk et al., 2024). Dieses Antwortformat wird als Answer-Until-Correct bezeichnet (AUC).

Auch wenn in der Forschungsliteratur bereits der Vorschlag gemacht wurde, AUC und CAT zu kombinieren (z.B. Muñiz & Menéndez, 2011), wurde dies bislang nicht psychometrisch ausgearbeitet. Insbesondere ist zu klären, welches IRT-Modell für AUC-Items im Rahmen von CAT am besten genutzt werden sollte. Verschiedene IRT-Modelle kommen für die Modellierung von mehreren Antwortversuchen in Frage und wurden in vorherigen Studien zu nicht adaptiven Tests miteinander verglichen. Hierzu gehören das Graded Response Model (GRM, Bergner et al., 2019), das Generalized Partial Credit Model (GPCM, Muñiz & Menéndez, 2011) sowie ein sequentielles IRT-Modell mit festem Diskriminationsparameter pro Item und veränderlichen Schwellen für Antwortversuche (SM-2p(j), Bergner et al., 2019).

Unklar ist jedoch, welches dieser Modelle bei adaptiver Itemauswahl hinsichtlich der Messpräzision am geeignetsten ist. Die vorliegende Simulationsstudie beantwortet diese Frage, indem sie die Messpräzision des GRM, des GPCM, des SM-2p(j) sowie des 2PL-Modells (stellvertretend für ein Testszenario, in dem kein AUC möglich ist) sowohl bei adaptiver als auch bei nicht-adaptiver Itemauswahl miteinander vergleicht.

Fragestellung

Die folgenden Fragen werden untersucht:

1. Welches IRT-Modell (GRM, GPCM, SM-2p(j)) erzielt bei adaptiver Itemauswahl von AUC-Items im Vergleich zum 2PL die höchste Messpräzision?

2. Wie unterscheidet sich die Messpräzision bei Nutzung von AUC-Items zwischen CAT und nicht adaptivem Testen?
3. In wie weit variieren die Ergebnisse zu 1. und 2. in Abhängigkeit der Fähigkeitsausprägung der Testteilnehmenden?

Methode

Um die genannten Fragen zu untersuchen, wird eine Simulationsstudie mit einem vollständig gekreuzten faktoriellen 4x2x13-Design durchgeführt. Als unabhängige Variablen werden dabei der Faktor IRT-Modell (GRM, GPCM, SM-2p(j), 2PL), der Faktor Adaptivität der Itemauswahl (nicht adaptiv, adaptiv) sowie die wahre Personenfähigkeit (-3 bis 3 in 0,5er Schritten) herangezogen. Als abhängige Variable werden der mittlere Standardfehler sowie der RMSE der resultierenden Personenparameterschätzungen betrachtet.

Für die Modellierung wird das R-Paket mirtCAT (Chalmers, 2016) verwendet.

Ergebnisse und ihre Bedeutung

Die Analysen laufen aktuell noch, werden jedoch zum Zeitpunkt der Konferenz abgeschlossen sein. Ausgehend von bisherigen Studien und vorläufigen Ergebnissen ist davon auszugehen, dass AUC grundsätzlich eine präzisere Messung erlaubt als ein Antwortformat, bei dem Personen nur einen Antwortversuch pro Item haben. Bei AUC in Verbindung mit CAT sind weitere Messpräzisionssteigerungen und zudem eine Angleichung der Messpräzision über das gesamte Kompetenzspektrum zu erwarten.

Die Ergebnisse werden Hinweise liefern, inwiefern sich die Kombination von CAT mit AUC hinsichtlich der Messeigenschaften als vorteilhaft erweist im Vergleich zu derzeit üblichen CAT-Varianten bzw. im Vergleich zu nicht adaptivem Testen. Die Analysen stellen die zentrale psychometrische Grundlage für die praktische Nutzung dieses für das emotionale Testerleben vorteilhaften Antwortformats bei CAT dar.

Ein Propensity Score Matching Ansatz zur Analyse von Differential Item Functioning (DIF) bei computerisierten adaptiven Tests

Lara Weiß, Aron Fink, Andreas Frey
Goethe-Universität Frankfurt am Main

Theoretischer Hintergrund

Computerisiertes adaptives Testen (CAT) ist ein spezielles Vorgehen zur computerbasierten Messung individueller Merkmalsausprägungen, bei dem sich die Auswahl der zur Bearbeitung vorgelegten Items am vorherigen Antwortverhalten der Testperson orientiert (Frey, 2023). CAT wird im Bildungsbereich insbesondere bei Bildungsmonitoringstudien, Hochschulklausuren und der Eignungs- und Zulassungsdiagnostik eingesetzt. Um die Testergebnisse zwischen Personen vergleichen zu können, die aufgrund der adaptiven Itemvorgabe unterschiedliche Items beantwortet haben, werden CAT-Itempool mit einem Modell der Item Response Theory skaliert (IRT; z.B. van der Linden, 2016). Die aus der Skalierung hervorgehenden Itemparameterschätzungen werden beim adaptiven Testen bei Itemauswahl und Merkmalsschätzung fixiert. Es ist es deshalb wichtig, dass die Itemparameter kein Differential Item Functioning (DIF; Holland & Wainer, 1993) aufweisen. DIF liegt vor, wenn Itemparameterschätzungen nach Kontrolle mittlerer Merkmalsunterschiede substantiell zwischen Gruppen (z.B. Herkunftsland) variieren.

Bei CAT sind für DIF-Analysen drei Datenlagen zu unterscheiden (Zwick, 2010):

1. Potentielle DIF-Items wurden adaptiv vorgegeben.
2. Potentielle DIF-Items wurden mit „Seeding“ in den adaptiven Test integriert.
3. Potentielle DIF-Items wurden in separatem nicht-adaptivem Teil administriert.

Die für nicht-adaptive Tests entwickelten DIF-Methoden zielen auf Datenlage 3 ab. Die wenigen bislang für CAT vorgeschlagenen DIF-Methoden basieren auf unrealistischen Annahmen, sind nicht IRT-basiert; können nicht bei allen drei Datenlagen genutzt werden und haben eine bestenfalls ausbaufähige Performanz.

Ein zentrales Problem der Analyse von Itemantwortdaten der Typen 1 und 2 ist, dass die Anzahl der Antworten pro Item in Abhängigkeit des Fähigkeitsniveaus der jeweiligen Gruppe teilweise stark variiert. Dies führt u.a. zu deutlich über Items variierender Teststärke bei inferenzstatistischen DIF-Tests.

Die Problematik kann aufgelöst werden, wenn vor der Durchführung einer DIF-Analyse mittels Propensity Score Matching (Rosenbaum & Rubin, 1983) die Anzahl an Antworten je Item zwischen den zu vergleichenden Gruppen angeglichen wird. Nach dieser Angleichung kann eine DIF-Analyse mit dem Multi-Facetten-Rasch-Modell (MFRM; Linacre, 1994) erfolgen. Diese Kombination (nachfolgend: PSM-MF-Methode) führte bei empirischen Daten zu einer substantiellen Verringerung von Items mit signifikantem DIF und somit vermeintlich zu einer Reduktion falsch detektierter DIF-Items (Frey & Weiß, 2023). Mit der vorliegenden Simulationsstudie, bei der die DIF-Items bekannt sind, wird nun die Leistungsfähigkeit der neuen PSM-MF-Methode differenziert untersucht.

Fragestellungen

1. Welchen Einfluss hat die PSM-MF-Methode auf Sensitivität und Spezifität der DIF-Detektion im Vergleich zur DIF-Analyse mit MFRM?
2. Welchen Einfluss hat die DIF-Effektgröße auf die Unterschiede bezüglich Sensitivität und Spezifität von PSM-MF und MFRM?
3. Welchen Einfluss hat die Fähigkeitsdifferenz zwischen Gruppen auf die Unterschiede bezüglich Sensitivität und Spezifität von PSM-MF und MFRM?
4. Welchen Einfluss hat die relative Gruppengröße auf die Unterschiede bezüglich Sensitivität und Spezifität von PSM-MF und MFRM?

Methode

Die Fragestellungen werden mit einer in R durchgeführten Simulationsstudie auf Basis eines vollständig gekreuzten 4-faktoriellen Versuchsplans mit den unabhängigen Variablen DIF-Methode (PSM-MF; MFRM), DIF-Effektgröße (A-kein; B-mäßig, C-extrem), relative Gruppengröße (1:1; 1:4) und Fähigkeitsdifferenz zwischen Gruppen (0.0; 0.5; 1.0) bezüglich der abhängigen Variablen Sensitivität und Spezifität der DIF-Detektion beantwortet. Die Simulation umfasst zwei Gruppen von insgesamt N = 5,000, mit Mittelwerten gemäß jeweiliger Stufe des Faktors Fähigkeitsdifferenz und SD = 1.0. Der Itempool umfasst 300 raschkonforme Items mit M(Itemschwierigkeit) = 0 und SD(Itemschwierigkeit) = 1.

Ergebnisse und ihre Bedeutung

Die Simulation läuft derzeit und wird bis zur Tagung abgeschlossen sein. Vorläufige Ergebnisse weisen auf eine höhere Spezifität bei vergleichbarer Sensitivität der neuen PSM-MF-Methode im Vergleich zu MFRM hin. Der Anteil irrtümlich detektierter Items (falsch Positive) kann gesenkt werden bei vergleichbarem Anteil irrtümlich nicht detektierter Items (falsch Negative).

Bestätigen sich diese Ergebnisse nach Abschluss der Simulation, liegt mit der neuen PSM-MF-Methode die derzeit leistungsfähigste DIF-Methode für alle drei Datenlagen bei CAT vor. Dies ist für typische CAT-Anwendungen im Bildungsbereich bei großen Studien hochinteressant. DIF-Analysen können dann im laufenden Betrieb und ohne separate Teilstudie realisiert werden.

Ein adaptiver Test für VERA – Design, Befunde und die Frage nach der Instruktion

Janine Buchholz, Pauline Wernet

Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Theoretischer Hintergrund

Neben der Teilnahme an nationalen und internationalen Schulleistungsstudien stellen die Vergleichsarbeiten („VERA“, auch „Lernstandserhebungen“, „KERMIT“ oder „Kompetenztests“) einen wichtigen Bestandteil der Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (Kultusministerkonferenz, 2016) dar. Die schriftlichen Tests werden jährlich und in vielen Ländern verpflichtend in der Primar- und Sekundarstufe I durchgeführt, um „die Unterrichts- und Schulentwicklung jeder einzelnen Schule durch eine an den Bildungsstandards orientierte Rückmeldung als Standortbestimmung mit Bezug zu den Landesergebnissen zu unterstützen“ (ebd., S. 13).

Im Rahmen der Weiterentwicklung der Vergleichsarbeiten (Kultusministerkonferenz, 2018) wurde beschlossen, den diagnostischen Erkenntnisgewinn zu erhöhen, indem die Testaufgaben so ausgewählt werden, dass sie dem „Kompetenzniveau der Schülerinnen und Schüler besonders gut entsprechen“ (ebd., S. 5). Sog. Multistage adaptive Tests (MSAT), bei denen sich die Auswahl von Aufgabenbündeln („Stages“) am Antwortverhalten auf zuvor bearbeiteten Aufgabenbündeln orientiert (vgl. Frey, 2020), werden beispielsweise seit PISA 2018 eingesetzt (OECD, 2019). Aufgrund der zunehmend technologiebasierten Durchführung von VERA erscheint der Einsatz von MSAT daher auch für VERA vielversprechend.

Da es sich bei VERA um ein sog. low-stakes Assessment handelt, besteht die Gefahr, dass Unterschiede in der Testleistung nicht nur auf Unterschiede in der gemessenen Kompetenz, sondern auch der Testteilnahmemotivation zurückzuführen sind (Finn, 2015; Wise & DeMars, 2005). Als Einflussfaktoren auf die Testteilnahmemotivation wurde u.a. die Testform (adaptiver vs. linearer Test) untersucht (Akhtar & Firdiyanti, 2023; Frey et al., 2024), wobei sich differentielle Effekte je nach Leistungsniveau aus der Expectancy-Value-Theorie (Wigfield & Eccles, 2000) ableiten lassen: beispielsweise könnten leistungsstarke Schüler*innen durch die typischerweise zu erwartende geringere Lösungshäufigkeit demotiviert werden. So kommt der Testinstruktion eine besondere Bedeutung zu, also der Aufklärung der Schüler*innen über den adaptiven Charakter des Tests. Allerdings konnte hierfür bisher kein statistisch signifikanter Effekt gefunden werden, da bislang kaum Studien hierzu vorliegen (Frey et al., 2024).

Fragestellung

Diese Studie verfolgt zwei Fragestellungen. Zum einen soll die Eignung von MSAT für VERA untersucht werden. Dafür werden mögliche adaptive Testdesigns besprochen, die technisch korrekte Funktionsweise im operativen Einsatz erprobt und das Bearbeitungsverhalten der Schüler*innen exploriert (z.B. Vor- und Zurückspringen).

Zum anderen soll untersucht werden, welchen Effekt die Instruktion über den adaptiven Testcharakter auf verschiedene Maße der Testbearbeitungsmotivation und die gemessene Testleistung hat. Dabei wird ein positives Framing vorgenommen, also eine hohe Schwierigkeit als Anzeichen guter Testleistung angeführt.

Methode

Die Fragestellungen werden anhand einer Stichprobe von $N = 857$ Schüler*innen der zweiten und vierten Jahrgangsstufe untersucht (19 bzw. 18 Klassen). Die Schüler*innen bearbeiteten u.a. jeweils einen Multistage-Test mit VERA-Aufgaben aus dem Kompetenzbereich Deutsch-Sprachgebrauch und einer maximalen Dauer von 20 Minuten. In 19 von 37 zufällig gezogenen Klassen wurde die mündlich dargebotene Instruktion um den Hinweis auf den adaptiven Charakter des Tests erweitert. Vor und nach Bearbeitung des adaptiven Tests wurde die Anstrengungsbereitschaft per Selbstbericht gemessen. Des Weiteren wurden die Antwortzeiten erfasst und auf Basis der NT10-Methode (Wise & Ma, 2012) der Response-Time-Effort (RTE) berechnet. Zudem stehen Lösungshäufigkeiten, Häufigkeiten fehlender Antworten und Fähigkeitsschätzungen bereit, um den Effekt der adaptiven Instruktion zu untersuchen.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse weisen auf die korrekte Funktionsweise des in der IQB-Testumgebung implementierten adaptiven Mechanismus. Bezüglich der Instruktion zeigt sich, entgegen der Erwartung, dass die über den adaptiven Testcharakter instruierten Schüler*innen eine niedrigere Anstrengung berichten (sowohl vor als auch nach dem Test: $t(df = 675.07) = -3.20$, $p = .001$, $d = -0.236$ bzw. $t(df = 502.43) = -2.41$, $p = .016$, $d = -0.209$). Vertiefende Analysen werden Gruppenunterschiede im RTE, Lösungshäufigkeiten, Häufigkeiten fehlender Antworten und Fähigkeitsschätzungen einbeziehen. Zusätzlich wird der Effekt der adaptiven Instruktion bedingt auf ein externes Leistungsmaß (Lesen) analysiert, um so evtl. vorhandene differentielle Effekte je nach Leistungsniveau identifizieren zu können. Die Bedeutung dieser Ergebnisse wird abschließend im Kontext von Schulleistungsuntersuchungen diskutiert.

S-04: Sonderpädagogische Perspektiven auf Bullying und Viktimisierung in der Schule

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 142

Chair der Sitzung: Philipp Nicolay

Chair der Sitzung: Christian Huber

Chair der Sitzung: Karolina Urton

Sonderpädagogische Perspektiven auf Bullying und Viktimisierung in der Schule

Chair(s): **Philipp Nicolay** (Bergische Universität Wuppertal, Deutschland), **Christian Huber** (Bergische Universität Wuppertal, Deutschland), **Karolina Urton** (Universität Münster)

Bullying im schulischen Kontext stellt ein weit verbreitetes Phänomen dar, das für betroffene Schüler:innen gravierende Folgen haben kann. So ist die Viktimisierung in Bullyingprozessen u.a. mit der Ausbildung internalisierender (Burke et al., 2017; Moore et al., 2017) und externalisierender Störungen (Sigurdson et al., 2021), einer geringeren Schulleistung (Kochenderfer-Ladd, 2021), Schulabsentismus (Schlesier et al., 2023) und einem geringeren sozioökonomischen Status im Erwachsenenalter (Takizawa et al., 2014) assoziiert. Folgt man der Definition von Olweus (1993), dann ist Bullying u.a. gekennzeichnet durch ein Machtungleichgewicht zwischen Täter:innen und Opfern. Entsprechend verdeutlicht eine Vielzahl von Studien, dass die Rollenübernahme in Bullyingprozessen auch entlang von Heterogenitätsdimensionen (u.a. Migrationshintergrund, sonderpädagogischer Förderbedarf) verläuft. So sind bspw. Schüler:innen mit sonderpädagogischem Förderbedarf, insbesondere solche mit Verhaltensproblemen, häufiger Täter:innen oder Opfer in Bullyingprozessen (Eilts & Koglin, 2022; Jenkins, et al., 2017).

Eine alleinige Betrachtung von individuellen Risikofaktoren entlang von Dichotomien (Förderbedarf, Migrationshintergrund ja/nein) erklärt diese Prävalenzen allerdings nur bedingt und vernachlässigt kontextbezogene Einflussfaktoren auf das Auftreten von Bullying (Schuster, 2017). Das vorliegende Symposium versucht daher aus einer dezidiert sonderpädagogischen Perspektive Bullying Zielgruppen differenziert und als Gruppenprozess im spezifischen sozialen Raum der Schulklasse zu betrachten und hierbei auch Wirkmechanismen zu identifizieren, die Ansatzpunkte zu Präventions- und Interventionsmaßnahmen geben.

Der erste Beitrag betrachtet vor dem Hintergrund eines traumapädagogischen Ansatzes mit der Gruppe der ukrainischen Schüler:innen mit Fluchterfahrung eine besonders vulnerable und häufig bereits psychisch belastete Gruppe und vergleicht diese mit Schüler:innen mit Migrationshintergrund und Schüler:innen ohne Migrationshintergrund hinsichtlich ihrer Bullying- und Viktimisierungserfahrungen. Die zwei darauffolgenden Beiträge nehmen insbesondere die Rolle von Lehrkräften in der Entstehung von Bullyingprozessen in Schulklassen in den Fokus. Der zweite Beitrag betrachtet hierbei die Bedeutung, die schulleistungsbezogene Wettbewerbsorientierung und das Betonen von Leistungswettbewerb durch die Lehrkraft für soziale Vergleichsprozesse entlang akademischer Leistung in Schulklassen haben und inwiefern diese Bullyingverhalten befördern können. Der dritte Beitrag untersucht aus der Perspektive der sozialen Referenzierungstheorie längsschnittlich die Beziehung zwischen Verhaltensproblemen, negativem Lehrkraftfeedback und der Viktimisierung im Rahmen von Bullyingprozessen und geht dabei der Frage nach, inwiefern öffentliches negatives Feedback von Lehrkräften ein Erklärungsansatz für die erhöhte Prävalenz von Viktimisierungserfahrungen dieser Schüler:innengruppe sein könnte. Der abschließende vierte Beitrag wirft einen kritischen Blick auf die psychometrische Messung von Bullying und die damit häufig verbundene klare Zuweisung von festen Rollen in Bullyingprozessen, indem er eine Überprüfung der faktoriellen Struktur des Bully Participant Behaviors Questionnaire an einer deutschsprachigen Stichprobe vornimmt.

Die Beiträge bilden die Grundlage für eine abschließende Diskussion, die insbesondere Handlungsmöglichkeiten von Lehrkräften in den Fokus stellen soll.

Beiträge des Symposiums

Bullying- und Viktimisierungserfahrungen von Schüler:innen mit Fluchterfahrungen

Friedrich Linderkamp

Bergische Universität Wuppertal

Hintergrund: Die Viktimisierung durch Gleichaltrige umfasst ein breites Spektrum von Handlungen wie Schlagen, Beschimpfungen und soziale Ächtung. Die Folgen der Viktimisierung durch Gleichaltrige sind gut dokumentiert. Kinder, die viktimisiert wurden, haben ein höheres Risiko für verschiedene psychische Störungen, die chronifizieren können (Chokprajakhat & Kuanliang, 2018).

Die Viktimisierung durch Gleichaltrige erleben Kinder und Jugendliche zumeist auf dem Schulgelände, wo die Aufsicht gering ist (Vaillancourt et al., 2010).

Eine diesbezüglich besonders vulnerable Schüler:innengruppe stellen die Kinder und Jugendlichen mit Fluchthintergrund dar, für die, sobald sie ihren Wohnsitz in der Zuweisungsgemeinde bezogen haben, die gesetzliche Schulpflicht greift und die wachsende Gruppe dieser Schüler:innen recht direkt in das Schulsystem eintreten.

In Deutschland wurden 2023 über 150.000 Asylanträge gestellt, davon fast die Hälfte von Kindern und Jugendlichen bis zu 18 Jahren. Hinzu kommen bis Ende des Jahres 2023 über 1 Millionen Kriegsflüchtlinge aus der Ukraine (BAMF, 2024). Laut UNICEF (2023) machen Kinder etwa 40-50% der Geflüchteten aus der Ukraine aus.

Bei diesen Kindern und Jugendlichen besteht ein hohes Risiko für Bullying- und Viktimisierungserfahrungen, bspw. aufgrund direkter rassistischer Diskriminierung und Ausgrenzung im Aufnahmeland und in Verbindung mit den Erfahrungen physischer Gewalt sowie krimineller oder sexueller Ausbeutung und kriegsrechtlicher Zwangsrekrutierung in der Heimsituation und auf der Flucht (Schröder, Zok & Faulbaum, 2018). Möhrle und Kollegen (2016) ermittelten bei 35-61% der männlichen unbegleiteten minderjährigen Geflüchteten (UMF) emotionale Störungen. In anderen Studien indizieren klinische Interviews innerhalb der Gruppe der UMF psychische Erkrankungen im Umfang von 42 % bis 56 % vor allem im Bereich internalisierender Störungen (Höltermann, Scharf, Romer & Möller-Kallista, 2022; Jakobsen, Demott & Heir, 2014; Huemer et al., 2011).

Zielsetzung: Die Studie verfolgt das Ziel, die besonders vulnerable Gruppe der Kinder mit Fluchterfahrungen aus dem Kriegsgebiet der Ukraine im Hinblick auf Viktimisierungs- aber auch auf Bullyingerfahrungen („Täter-Opfer-Umkehr“) im Vergleich mit Schüler:innen mit Migrationshintergrund aus gesellschaftlich-politischen und ökonomischen Motiven und mit in Deutschland geborenen Schüler:innen zu untersuchen.

Methode: Die Stichprobe besteht aus 60 ukrainischen Flüchtlingen, 60 Flüchtlingen mit Migrationshintergrund und 60 in Deutschland geborenen Schüler:innen beider Geschlechter im Alter von 9-11 Jahren. Zur Erfassung der Bullying- und Viktimisierungserfahrungen wurde ein standardisiertes Instrument (BVF-K, Marées & Petermann, 2010) eingesetzt. Zur ergänzenden Kontrolle etwaiger Lern- und Verhaltensstörungen wurde zudem das Lehrkrafturteil gemäß SDQ (Goodman, 1997) erhoben.

Ergebnisse: Bisher wurde die Datenerhebung zur Hälfte abgeschlossen. Die ersten Ergebnisse weisen auf deutliche Gruppenunterschiede, vor allem im Hinblick auf Viktimisierungserfahrungen hin, die in der besonders vulnerablen Gruppe der Kinder mit Fluchthintergrund im Unterschied zu den Kindern mit Migrationshintergrund und der Kontrollgruppe ohne diese Merkmale deutlich häufiger und ausgeprägter vorkommen. Die Kinder mit Fluchthintergrund weisen zudem in den SDQ-Skalenwerten mit externalisierenden und internalisierenden Verhaltensmerkmalen deutlich erhöhte Werte auf. Nach Komplettierung der Stichprobe werden die Zusammenhänge der erhobenen Maße unter Berücksichtigung weiterer Faktoren (Alter, Geschlecht etc.) regressionsanalytisch modelliert.

Diskussion: Die Ergebnisse dürften Implikationen für die Konzipierung Zielgruppen differenter Präventionsmaßnahmen gegen Bullying und Viktimisierungserfahrungen von Schüler:innen mit Flucht- aber auch mit Migrationshintergrund mit sich bringen.

Eine Untersuchung der Beziehung von Bullying, schulleistungsbezogener Wettbewerbsorientierung und leistungsorientierter Lehrkraftpräferenz

Christian Drengk, Moritz Börnert-Ringleb

Leibniz Universität Hannover

Untersuchungen zu Entstehungsprozessen und Hintergründen von Bullying machen deutlich, dass dieses als instrumentalisiertes Verhalten betrachtet werden kann, das eingesetzt wird, um sich in Konkurrenz- bzw. Wettbewerbssituationen Vorteile zu verschaffen (Olthof et al., 2011; Schäfer & Stoiber, 2013). Dabei haben vorherige Analysen gezeigt, dass Bullying mit einer entsprechenden persönlichen Wettbewerbsorientierung in Verbindung gebracht werden kann (Drengk & Börnert-Ringleb, in Begutachtung). Personen mit dieser Einstellung zeigen vermehrt aggressive und feindselige Tendenzen (Ross et al., 2003; Spence et al., 1987). Die Entwicklung einer persönlichen Wettbewerbsorientierung von Lernenden kann dabei in Zusammenhang mit Unterrichtsgestaltung und Lehrkraftverhalten stehen, die einen Wettbewerb unter den Lernenden forcieren (Butera et al., 2021). Betont eine Lehrkraft den Leistungswettbewerb durch die Präferenz von leistungsstarken Lernenden, führt dies potentiell dazu, dass Lernende eine entsprechende Werthaltung als schulleistungsbezogene Wettbewerbsorientierung internalisieren. Diese steht mit Bullyingprozessen in Verbindung, sodass der leistungsorientierten Lehrkraftpräferenz an dieser Stelle eine indirekte Beziehung zu Bullyingprozessen zugeschrieben werden kann. Neben dieser indirekten Beziehung von leistungsorientierter Lehrkraftpräferenz und Bullying weisen bestehende Ergebnisse darüber hinaus auf einen direkten Zusammenhang hin. Die leistungsorientierte Lehrkraftpräferenz verstärkt Vergleiche der Lernenden hinsichtlich ihrer Schulleistung und ihrer Position in der akademischen Hierarchie der Klasse. Durch die Lehrkraft etablierte akademische Hierarchien können sich dabei auch in sozialen Hierarchien widerspiegeln (Mikami et al., 2012). Lehrkräfte, die Lernende aufgrund ihrer Position in der akademischen Hierarchie bevorzugen, könnten ihrer Klasse vermitteln, dass eine soziale Hierarchie gleichermaßen angemessen ist, in der Lernende als sozial ungleichwertig gelten, was einen Nährboden für Bullyingprozesse darstellen kann.

Der vorliegende Beitrag greift diese Überlegungen auf und setzt die Variablen Bullyingverhalten, schulleistungsbezogene Wettbewerbsorientierung und die Wahrnehmung einer leistungsorientierten Lehrkraftpräferenz in Beziehung. Dabei wird die Hypothese aufgestellt, dass Bullyingverhalten in direkter Beziehung zur Wahrnehmung einer leistungsorientierten Lehrkraftpräferenz steht und sich Teile dieser Beziehung durch eine individuelle schulleistungsbezogene Wettbewerbsorientierung erklären lassen, die einen direkten Zusammenhang mit Bullyingverhalten und der Wahrnehmung einer leistungsorientierten Lehrkraftpräferenz aufweist.

Zur Überprüfung dieser Annahmen werden die Daten von N = 541 Lernenden aus 38 Klassen der Klassenstufen sieben bis zehn im Rahmen eines Mehrebenen-Strukturgleichungsmodells analysiert.

Die Ergebnisse deuten an, dass eine schulleistungsbezogene Wettbewerbsorientierung sowohl auf Individuumsebene (Level 1) als auch auf Klassenebene (Level 2) im Zusammenhang mit Bullyingverhalten steht. Verbindungen der Wahrnehmung einer leistungsorientierten Lehrkraftpräferenz zur schulleistungsbezogenen Wettbewerbsorientierung und zum Bullyingverhalten der Lernenden lassen sich nur auf Individuumsebene beschreiben.

Insgesamt weisen die gewonnenen Erkenntnisse darauf hin, dass eine wahrgenommene leistungsorientierte Lehrkraftpräferenz eine schulleistungsbezogene Wettbewerbsorientierung und damit auch Bullyingprozesse begünstigen kann. Lehrkräfte sollten sich ihrer Stellung als zentrale Referenzperson für ihre Lernenden bewusst sein und erkennen, dass ihr Verhalten gegenüber einzelnen Lernenden Auswirkungen auf alle weiteren Klassenmitglieder haben kann. Insbesondere der öffentlich, sozial vergleichende Umgang mit heterogenen Leistungen kann dabei persönliche Einstellungen und soziale Dynamiken hervorbringen, die Bullyingprozesse begünstigen können.

Die Rolle von negativem Lehrkraftfeedback für die Viktimisierung von Schüler:innen mit Verhaltensproblemen in Bullyingprozessen

Philipp Nicolay, Christian Huber, Plutz Raphael

Bergische Universität Wuppertal

Theoretischer Hintergrund: Zahlreiche Studien zeigen, dass Schüler:innen mit sonderpädagogischem Förderbedarf in inklusiven Schulklassen eine geringere soziale Inklusion aufweisen als ihre Mitschüler:innen ohne Förderbedarf (Huber & Willbert, 2012; Huber et al., 2022). Dies betrifft im Besonderen Schüler:innen mit Verhaltensproblemen (Krull et al., 2014; Schürer, 2020). In mehreren Studien der letzten Jahre konnte ausgehend von der sozialen Referenzierungstheorie gezeigt werden, dass Lehrkräfte und ihre Beziehung zu Schüler:innen hier eine bedeutende Rolle einnehmen (Hendrickx et al., 2017; Hendrickx et al., 2020). Die Beziehung der Lehrkraft zu einzelnen Schüler:innen, scheint für die übrigen Schüler:innen als Referenz für die Wahl zukünftiger Interaktionspartner:innen zu dienen. Diese Beziehungsinformationen werden dabei insbesondere durch Feedback von Lehrkräften transportiert (Huber et al., 2018; Nicolay & Huber, 2021), das den Zusammenhang von Verhaltensproblemen und sozialer Akzeptanz zu medieren scheint (Nicolay et al., 2024).

Konzeptuell können auch Viktimisierung und Bullying als Dimensionen sozialer Inklusion betrachtet werden (Koster, et al. 2009). Studien zeigen in diesem Kontext, dass Schüler:innen mit Verhaltensproblemen besonders häufig von Bullying betroffen sind (Eilts & Koglin, 2022; Jenkins, et al., 2017) und, dass auch hier die Lehrkraft und ihre Beziehung zu Schüler:innen bedeutend ist

(Dietrich & Cohen, 2021; Marengo et al., 2021). Dies legt nahe, dass auch hier soziale Referenzierungsprozesse wirksam sein könnten. Diesen Überlegungen folgend, würde die Lehrkraft eine soziale Referenz im Klassenraum darstellen, die Viktimisierungsprozesse durch ihr öffentliches negatives Feedbackverhalten fördert oder hemmt. Dementsprechend wären insbesondere Schüler:innen, die von Lehrkräften viel negatives Feedback erhalten stärker von Viktimisierungserfahrungen betroffen.

Fragestellung: Ausgehend von den dargestellten Befunden zur Bedeutung von Lehrkraftfeedback für die soziale Inklusion von Schüler:innen soll im Beitrag überprüft werden, inwiefern diese auch auf Bullyingprozesse übertragbar sind. Entsprechend wird die Fragestellung in den Fokus genommen, in welcher längsschnittlichen Beziehung zueinander Verhaltensprobleme, negatives Lehrkraftfeedback und Viktimisierungserfahrungen stehen.

Methode: Der vorliegende Beitrag basiert auf einer Studie in $n = 37$ Grundschulklassen der dritten und vierten Jahrgangsstufe (MAlter = 9.14, SDAlter = 0.85) mit $N = 754$ Schüler:innen, die über ein Schulhalbjahr an 3 Messzeitpunkten befragt wurden.

Die drei Untersuchungsvariablen wurden jeweils durch unterschiedliche Akteur:innen beurteilt: Verhaltensprobleme der einzelnen Schüler:innen wurden durch die Lehrkraft eingeschätzt. Da für soziale Referenzierungsprozesse die subjektive Wahrnehmung von sozialen Hinweisreizen von größerer Bedeutung ist als ihre objektive Wahrnehmung durch externe Beobachter:innen, wurde die Häufigkeit mit der Schüler:innen negatives Feedback erhalten durch ihre Peers eingeschätzt und pro Schüler:in zu einem Mittelwert aggregiert (vgl. Nicolay et al., 2024). Viktimisierungserfahrungen wurden mit der Kinderversion des BVF-K (Marées & Petermann, 2010) im Selbstbericht erhoben.

Ergebnisse: Die Datenauswertung erfolgte mit Cross- Lagged-Panel Modellen mit dem Paket lavaan (Rosseel, 2012) in R. Die Fragestellung betreffend zeigten sich zwischen dem ersten und zweiten Messzeitpunkt signifikante Effekte sowohl von Verhaltensproblemen auf negatives Lehrkraftfeedback ($\beta = .19$, $SE = 0.02$, $p < .001$) als auch von negativem Lehrfachfeedback auf Verhaltensprobleme ($\beta = .20$, $SE = 0.04$, $p < .001$). Weiterhin zeigten sich sowohl zwischen dem ersten und zweiten als auch zwischen dem zweiten und dritten Messzeitpunkt signifikante Effekte von negativem Lehrkraftfeedback auf Viktimisierung ($\beta = .08$, $SE = 0.03$, $p = .028$ bzw. $\beta = .12$, $SE = 0.04$, $p = .025$) bzw. von Viktimisierung auf negatives Lehrkraftfeedback ($\beta = .04$, $SE = 0.03$, $p = .036$ bzw. $\beta = .08$, $SE = 0.04$, $p = .003$).

Diskussion: Die gefundenen mittelstarken Effekte (Orth et al., 2024) weisen darauf hin, dass negatives Lehrkraftfeedback einen Einfluss auf die Viktimisierungserfahrungen hat, die Schüler:innen in ihrer Klasse machen. Die starken wechselseitigen Effekte zwischen Verhaltensproblemen und negativem Feedback, die zwischen dem ersten und zweiten Messzeitpunkt gefunden wurden, deuten an, dass dieses Phänomen insbesondere für Schüler:innen mit Verhaltensproblemen bedeutend ist. Implikationen, die sich hieraus für Lehrkraft Handeln im Kontext heterogener Schüler:innenschaft ergeben, werden diskutiert.

Mehr als nur Victims und Bullys - Eine faktorenanalytische Untersuchung des Bullying Participant Behaviors Questionnaire

Jule Eilts¹, Jessica Wilke²

¹Ludwig-Maximilians-Universität München, ²Carl von Ossietzky Universität Oldenburg

Bullying in der Schule ist ein weit verbreitetes Phänomen, das erhebliche negative Auswirkungen auf die betroffenen Schülerinnen und Schüler hat (Krieg et al., 2020; Li et al., 2022; Montes et al., 2022). Salmivalli et al. (1996) identifizierten erstmals verschiedene Rollen in der Bullying-Dynamik, die die gesamte Klasse einbeziehen (z. B. Bully, Victim, Defender, Assistent, Outsider). Da situative Faktoren die Rollenübernahme beeinflussen, können den Schülerinnen und Schülern jedoch keine festen Rollen zugeschrieben werden.

Der Bully Participant Behavior Questionnaire (BPBQ; Demaray et al., 2016) ist einer der ersten Selbsteinschätzungsfragebögen zur Rollenübernahme (Bully, Victim, Defender, Assistent, Outsider) in der Bullying-Dynamik. Internationale Studien zu psychometrischen Untersuchungen des Instruments zeigen ebenfalls, dass den Schülerinnen und Schülern in den meisten Fällen keine eindeutige Rolle zugeschrieben werden kann. Vielmehr diskutieren die Studien, die Involviertheit in mehrere verschiedenen Rollen (z. B. Pro-Bully oder Pro-Victim) sowie den Rollenwechsel in der Bullying-Dynamik (z.B. Jenkins & Canivez, 2021; Jenkins et al., 2022; Qiu et al., 2020). Eine psychometrische Überprüfung des BPBQ für eine nationale Stichprobe steht jedoch noch aus. Ziel der vorliegenden Untersuchung ist es daher, die in internationalen Studien identifizierten Faktorenstrukturen des BPBQ an einer nationalen Stichprobe zu überprüfen und daraus Implikationen für die Praxis und die Forschung abzuleiten. Die Überprüfung des Instruments bietet Schulen die Möglichkeit, es im Klassensetting ökonomisch anzuwenden, sodass Lehrkräfte die Bullying-Dynamik besser erkennen und entsprechend darauf reagieren können.

In einer querschnittlichen Fragebogenstudie wurden $N = 241$ Schülerinnen und Schüler der 5. bis 10. Klasse aus Niedersachsen und Bremen zu ihren Erfahrungen und Rollen im Zusammenhang mit Bullying befragt. Der BPBQ umfasst 50 Items (10 pro Rolle). Es wurden konfirmatorische Faktorenanalysen (CFA) durchgeführt, um die in den internationalen Studien identifizierten unterschiedlichen Modelle zu überprüfen und das beste Faktorenmodell zu identifizieren. Die in der Literatur identifizierten Modelle umfassen sowohl fünf faktorielle Modelle (Bully, Victim, Assistent, Defender, Outsider) als auch vier faktorielle Modelle (Bully/Assistent, Victim, Defender, Outsider) und hierarchische Modelle mit zwei übergeordneten Faktoren (pro-Bully (Bully, Assistent, Outsider) und pro-Victim (Victim und Defender)).

Die Ergebnisse dieser Untersuchung bieten einen tieferen Einblick in die Dynamiken von Bullying in Schulklassen und verdeutlichen die Komplexität der Rollenübernahme. Es wird diskutiert, dass Schülerinnen und Schüler nicht starr in eine einzige Rolle eingeordnet werden können bzw. das übergeordnete Modelle nach pro-Bully und pro-Victim sinnvoll sein können (Demaray et al., 2021). Diese Erkenntnis unterstützt die Entwicklung flexiblerer Ansätze im Umgang mit Bullying, die die situativen Einflüsse berücksichtigen. Lehrkräfte können mithilfe dieser Erkenntnisse individuellere und kontextsensitivere Unterstützungsmaßnahmen entwickeln. Konkreten Maßnahmen dazu sollen in dem vorliegenden Beitrag diskutiert werden.

Für die Forschung lässt sich ableiten, dass zukünftige Studien die Dynamik der Rollenübernahme noch detaillierter untersuchen müssen, um Einflussfaktoren zu identifizieren und besser zu verstehen. Langfristig könnte dies dazu beitragen, präzisere Messinstrumente und Präventions- sowie Interventionsmaßnahmen zu entwickeln und die Wirksamkeit dieser zu erhöhen.

S-05: Wie Lehrkräfte ihre Schüler*innen einschätzen: Potentielle Verzerrungen in Leistungserwartungen, Leistungsbewertung und Umgang mit Fehlverhalten

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 145

Chair der Sitzung: Sophie Harms

Chair der Sitzung: Lars Meyer-Jenßen

Chair der Sitzung: Hanna Dumont

Wie Lehrkräfte ihre Schüler*innen einschätzen: Potentielle Verzerrungen in Leistungserwartungen, Leistungsbewertung und Umgang mit Fehlverhalten

Chair(s): **Sophie Harms** (Humboldt-Universität zu Berlin, Berliner Institut für empirische Integrations- und Migrationsforschung),

Lars Meyer-Jenßen (Humboldt-Universität zu Berlin)

Diskutant*in(nen): **Hanna Dumont** (Universität Potsdam)

Der schulische Alltag von Schüler*innen ist maßgeblich von Begegnungen mit Lehrkräften geprägt. Dabei haben Lehrkräfte immer auch die Aufgabe, die Schüler*innen in Bezug auf ihre Kompetenzen und ihr Verhalten zu beurteilen und ihr Entwicklungspotential einzuschätzen. Diese Einschätzungen bestimmen, wie Lehrkräfte den Schüler*innen gegenüber agieren, ob sie die Lernenden herausfordern, unterstützen, sie loben oder sie zurechtweisen, was sich wiederum auf deren Entwicklung auswirkt (Wang et al., 2018). Um Schüler*innen entsprechend ihres tatsächlichen Entwicklungspotentials zu fördern, sollten die Einschätzungen der Lehrkräfte und ihr Umgang mit den Schüler*innen sich vornehmlich daran orientieren, welche veränderbaren Kompetenzen und Fertigkeiten die Schüler*innen zeigen und dabei möglichst akkurat sein (Loibl et al., 2020). Gleichzeitig haben auch askriptive Merkmale der Schüler*innen, Kontextfaktoren wie die Klassenzusammensetzung sowie Eigenschaften der Lehrkräfte selbst einen Einfluss darauf, wie Lehrkräfte den Schüler*innen begegnen. Dadurch kann es zu Verzerrungen kommen, die sich darin zeigen, dass Lehrkräfte Schüler*innen bei gleichen Kompetenzen und Fertigkeiten unterschiedlich einschätzen und behandeln (Wenz, 2020). Aus solchen Verzerrungen können sich nachteilige Effekte für die schulische Entwicklung der Betroffenen ergeben, z. B. in Form von selbsterfüllenden Prophezeiungen (Gentrup et al., 2020). Um dem entgegenzuwirken, ist es zunächst zentral, potentielle Verzerrungen aufzudecken. Die Beiträge des vorliegenden Symposiums widmen sich dieser Aufgabe und bedienen sich unterschiedlicher methodischer Zugänge von experimentellen Designs bis hin zu Felduntersuchungen unter naturalistischen Bedingungen. Aus einer intersektionalen Perspektive betrachtet der erste Beitrag Verzerrungen in Leistungserwartungen von Lehrkräften aufgrund des Geschlechts und des sozio-ökonomischen Status (socioeconomic status, SES) von Schüler*innen und analysiert zusätzlich die Bedeutung von Stereotypen für diese Verzerrungen. Der zweite Beitrag fokussiert auf die diagnostische Kompetenz von angehenden Mathematiklehrkräften und untersucht mithilfe von Vignetten die Diagnose- und Förderfertigkeiten von Lehramtsstudierenden im Bereich Mathematik in Abhängigkeit von SES, Geschlecht und Migrationshintergrund der Schüler*innen unter Berücksichtigung kognitiver und affektiv-motivationaler Faktoren aufseiten der Studierenden. Der dritte und der vierte Beitrag des Symposiums beschäftigen sich mit Verzerrungen im Disziplinarverhalten von Lehrkräften. Im dritten Beitrag wird in einem Online-Experiment untersucht, ob Lehramtsstudierende in Klassen mit einem höheren Anteil von als arabisch wahrgenommenen Schüler*innen strengere Disziplinarmaßnahmen einsetzen. Der vierte Beitrag betrachtet auf Basis von Videodaten Disziplinarmaßnahmen von Grundschullehrerinnen gegenüber Erstklässler*innen auf individueller Ebene je nach Geschlecht der Kinder. Somit nimmt das Symposium potentielle Verzerrungen in verschiedenen Momenten der Begegnung zwischen Lehrkräften und Schüler*innen und in Abhängigkeit von verschiedenen Faktoren in den Blick. Die Befunde und ihre Bedeutung werden von Hanna Dumont diskutiert.

Beiträge des Symposiums

Stereotype und Erwartungen von Lehrkräften an der Schnittstelle von Geschlecht und sozioökonomischem Status der Schüler:innen

Sarah Gentrup¹, Melanie Olczyk², Georg Lorenz³

¹Institut zur Qualitätsentwicklung im Bildungswesen, ²Martin-Luther-Universität Halle-Wittenberg, ³Universität Potsdam, Universität Leipzig

Geschlecht und sozioökonomischer Status (englisch: socioeconomic status; SES) sind zentrale Dimensionen bildungsbezogener Ungleichheit (Liu et al., 2022; Rosén et al., 2022), die u.a. den Kompetenzerwerb von Schüler:innen wechselseitig prägen (Parker et al., 2020). So hängt der SES stärker mit den schulischen Leistungen von Jungen als von Mädchen zusammen. In Mathematik resultiert daraus ein umso stärkerer Leistungsunterschied zugunsten von Jungen, je höher der SES der Familien ist (Lühe et al., 2017; Mensah & Kiernan, 2010; Penner & Paret, 2008). Im sprachlichen Bereich hingegen fällt der Leistungsunterschied zugunsten von Mädchen umso geringer aus, je höher der SES der Familien ist (Entwisle et al., 2007; Lühe et al., 2017; Mensah & Kiernan, 2010).

Leistungserwartungen von Lehrkräften können zu derartigen (intersektionalen) Bildungsdisparitäten beitragen, wenn sie für verschiedene Gruppen von Schüler:innen inakkurat unterschiedlich sind (Olczyk, Gentrup et al., 2023; Olczyk, Kwon et al., 2023). Bisherige Forschung bestätigt Verzerrungen in Lehrkräfterwartungen in Abhängigkeit vom Geschlecht und SES der Schüler:innen (Wang et al., 2018). Inwieweit das Geschlecht und der SES die Lehrkräfterwartungen allerdings wechselseitig prägen, ist bislang unklar. Um solche möglichen Wechselwirkungen zu verstehen, bedarf es zudem vertiefter Kenntnisse über Stereotype als Ursache für Verzerrungen in Leistungserwartungen (Fiske & Neuberg, 1990). Einzelne Studien belegen, dass Stereotype von Lehrkräften zu Verzerrungen in ihren Erwartungen beitragen (Glock et al., 2016; Lorenz, 2021; Muntoni & Retelsdorf, 2018; Tiedemann, 2002). Studien mit intersektionaler Perspektive auf mögliche Wechselwirkungen verschiedener Stereotype fehlen jedoch.

Der vorliegende Beitrag nimmt eine intersektionale Perspektive auf Verzerrungen von Lehrkräfterwartungen im sprachlichen und mathematischen Bereich ein. Es wird untersucht, (1) inwieweit Geschlecht und SES von Schüler:innen Lehrkräfterwartungen wechselseitig beeinflussen und (2) ob Geschlechterstereotype oder sozioökonomische Stereotype (oder beide) die Erwartungen von Lehrkräften für Mädchen und Jungen mit unterschiedlichem SES (wechselseitig) prägen.

Als Basis der Zwei-Ebenen-Regressionsanalysen dienten Daten der Studie «Kompetenzerwerb und Lernvoraussetzungen» (KuL; Kristen et al., 2018a; Lehrkräfte: N=69, 94% weiblich, Schüler:innen: N=1.049 (sprachlicher Bereich)/1.027 (Mathematik), 48% weiblich) und der Startkohorte Kindergarten des Nationalen Bildungspanels (NEPS-SC2; Blossfeld & Roßbach, 2019; NEPS-Netzwerk, 2022; Lehrkräfte: N=698 (sprachlicher Bereich)/614 (Mathematik), 94% weiblich; Schüler:innen: N=4.732 (sprachlicher Bereich)/4.117 (Mathematik), 51% weiblich).

Im ersten Schritt wurde analysiert, ob Verzerrungen in Leistungserwartungen in Abhängigkeit vom Geschlecht und SES der Schüler:innen additiv oder intersektional auftreten. Der inakurate Anteil der sprachlichen bzw. mathematischen Leistungserwartungen wurde dazu vorab anhand des Residualansatzes (Madon et al., 1997) unter Einbezug vielfältiger leistungsbezogener Merkmale bestimmt. Dieses Maß inakurater Erwartungen wurde daraufhin durch das Geschlecht und den SES der Lernenden sowie deren Interaktion regressionsanalytisch vorhergesagt. Im zweiten Schritt wurde untersucht, ob Geschlechterstereotype und sozioökonomische Stereotype von Lehrkräften etwaige Erwartungsverzerrungen erklären und ob eventuelle Effekte für Mädchen und Jungen sowie für Schüler:innen mit unterschiedlichem SES vergleichbar oder unterschiedlich ausfallen. Der Zuwanderungsstatus der Lernenden wurde kontrolliert.

Die Ergebnisse bestätigen additive Verzerrungen der sprachlichen und mathematischen Leistungserwartungen von Lehrkräften in Abhängigkeit vom Geschlecht und SES der Lernenden, jedoch keine intersektionalen Verzerrungen. Demzufolge traten geschlechtsbezogene Verzerrungen unabhängig vom SES der Schüler:innen auf und Verzerrungen in Abhängigkeit vom SES waren für Mädchen und Jungen vergleichbar. Die Richtung der Verzerrungen entsprach dabei etablierten Befunden (positive Verzerrungen für Mädchen im sprachlichen Bereich, für Jungen in Mathematik und für Lernende mit höherem SES in beiden Bereichen). Zudem zeigten Lehrkräfte mit traditionelleren Geschlechterstereotypen stärker nach dem Geschlecht verzerrte Leistungserwartungen, während sozioökonomische Stereotype nicht bedeutsam mit Verzerrungen zusammenhingen.

Wir diskutieren, inwieweit das Befundmuster mit dem breiteren Konzept von Intersektionalität vereinbar ist. Die Ergebnisse unterstützen die theoriebasierte Annahme, dass insbesondere geschlechtsbezogene Stereotype zu Verzerrungen in Leistungserwartungen beitragen. (Angehende) Lehrkräfte sollten für die Bedeutung stereotypgeprägter Erwartungen in der unterrichtlichen Praxis sensibilisiert werden.

Kompetenzen von Grundschullehramtsstudierenden im Bereich Diagnose und Förderung unter Berücksichtigung von Geschlecht und sozialer Herkunft der Schüler*innen - Effekte von fachbezogener Angst und Fachwissen

Tjorven Seibold¹, Lars Meyer-Jenßen², Christin Laschke³, Bettina Rösken-Winter¹

¹Universität Münster, ²Humboldt-Universität zu Berlin, ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Diagnosen von Lehrkräften bilden die Basis für eine adaptive Förderung in Lernprozessen, die sich an den Lernständen der Schüler*innen, also ihrem noch fehlenden Verständnis sowie ihrem bereits bestehenden Potenzial, orientiert (Hardy et al., 2019). In Bezug auf das Kompetenzmodell von Blömeke et al. (2015) kann die dafür erforderliche Diagnosekompetenz als situationsspezifische Fertigkeit betrachtet werden, in die kognitive Dispositionen sowie affektiv-motivationale Merkmale einfließen (Leuders et al., 2018).

Hinsichtlich der Dispositionen wurde für Mathematiklehrkräfte gezeigt, dass diese bei der Diagnose auf mathematisches Wissen zurückgreifen (Philipp, 2018). In Bezug auf affektiv-motivationale Merkmale wurde sowohl für Grundschullehrkräfte (Porsch, 2023) als auch Lehramtsstudierende (Porsch, 2019) die Relevanz von Mathematikangst betont. Weiterhin zeigte sich, dass die Mathematikangst von Lehrkräften die Diagnoseprozesse in heterogenen Schüler:innengruppen beeinflussen kann (Copurk-Gencturk et al., 2023). Die Beeinträchtigung von Diagnosen von Lehrkräften und Lehramtsstudierenden durch stereotypische Einstellungen wurden bereits in vielen Studien nachgewiesen, insbesondere für die Schüler:innenmerkmale Geschlecht, Migrationshintergrund und sozio-ökonomischer Status (Costa et al., 2021; Pit-ten Cate & Glock, 2019; Costa et al., 2022). Welchen Effekt Fachwissen und Mathematikangst auf die Diagnose- und Förderfertigkeiten bei Lehramtsstudierenden haben und welche Effekte Schüler:innenmerkmale haben, wurde bisher nicht in einem gemeinsamen Modell untersucht.

Es wird untersucht, welche Effekte Mathematikfachwissen und Mathematikangst von Grundschullehramtsstudierenden auf Diagnose und Förderung haben und inwiefern sich bezüglich des zugeschriebenen Geschlechts, Migrationshintergrunds und sozioökonomischen Status (SES) der Schüler:innen Effekte zeigen.

Mit einem vignettenbasierten Test wurden bei N = 181 Grundschullehramtsstudierenden Kompetenzen im Bereich Diagnose und Förderung erfasst. Als mathematischer Inhalt wurde der Bereich Größen und Messen gewählt. Die Studierenden wurden randomisiert verschiedenen Vignetten zugeordnet, die Informationen zum Geschlecht (weiblich/männlich), Migrationshintergrund (deutsch/arabisch) und sozioökonomischen Status (niedrig/hoch) der Schüler:innen enthielt. Zur Erfassung der Diagnosekompetenz wurden den Studierenden 20 Schüler:innenlösungen präsentiert, anhand derer sie den Kompetenzstand des Kindes im Rahmen eines offenen Erhebungsformats diagnostizieren sollten. Die Antworten der Studierenden wurden deduktiv hinsichtlich der Dimensionen Fehlendes Verständnis und Potenzial kategorisiert (Inter-Rater-Übereinstimmung: Kappa = .87). Anschließend sollten die Teilnehmenden zu den 20 berichteten Schüler:innenlösung eine adäquate Förderung aus vier Optionen auswählen (Cronbachs Alpha = .70). Um das Fachwissen zu erfassen, lösten die Studierenden die Aufgaben vorab selbst (Cronbachs Alpha = .75). Mathematikangst wurde mit vier Items (Cronbachs Alpha = .70) erfasst (Jenßen et al., 2021).

Die Effekte von Mathematikangst und Fachwissen sowie den in den Vignetten angesprochenen Merkmalen (SES, Geschlecht, Migrationshintergrund) auf die Diagnose und Förderung wurden in einem Pfadmodell überprüft.

Das Modell zeigte eine gute Passung ($\chi^2(6) = 9.269$, $p = .41$, RMSEA = .01 (.00/.09), CFI = .99, SRMR = .03). Hinsichtlich der Dispositionen der Studierenden ergaben sich positive Zusammenhänge zwischen Fachwissen und der Fertigkeit fehlendes Verständnis ($\beta_{\text{ind}} = 0.33$, $p < .001$) sowie Potenzial ($\beta_{\text{ind}} = 0.39$, $p < .001$) in den Schüler:innenlösungen zu diagnostizieren. Fachwissen zeigte einen signifikanten Effekt auf die Auswahl der adäquaten fachdidaktischen Förderung ($\beta_{\text{ind}} = 0.17$, $p < .001$). Erwartungsgemäß konnte ein negativer Effekt der Mathematikangst auf das Wissen der Studierenden festgestellt werden ($\beta_{\text{ind}} = -0.25$, $p < .001$). Dementsprechend zeigten sich indirekt negative Effekte (jeweils $p < .001$) von Angst auf die Diagnose des fehlenden Verständnisses ($\beta_{\text{ind}} = -0.80$) und des Potenzials ($\beta_{\text{ind}} = -0.10$) sowie auf die Auswahl der adäquaten Förderung ($\beta_{\text{ind}} = -0.11$).

Potenzial in den Lösungen wurde eher bei Schüler:innen mit hohem sozioökonomischen Status als mit niedrigem Status erkannt ($\beta_{\text{ind}} = 0.16$, $p < .001$). Darüber hinaus zeigten sich keine signifikanten Effekte der Schüler:innenmerkmale auf die Diagnose oder Förderung.

Die Ergebnisse zeigen, dass vermitteltes Fachwissen und Reduktion von Mathematikangst im Lehramtsstudium dazu beitragen können, Diagnose- und Förderkompetenzen zu stärken. Eine Sensibilisierung für Stereotype könnte verzerrte diagnostische Urteile reduzieren.

„In dieser Klasse gibt's eh' nur Stress!“ Wie die Zusammensetzung von Schulklassen den Umgang mit Störungen beeinflusst

Bisherige Forschung zeigt, dass Lehrkräfte Schüler*innen, die rassifizierten Gruppen angehören, anders behandeln als Schüler*innen, die nicht-rassifizierten Gruppen angehören. Studien aus den USA zeigen, dass bei störendem Verhalten von als Schwarz wahrgenommenen Schüler*innen eher eingegriffen wird als bei als Weiß wahrgenommenen Schüler*innen (Okonofua & Eberhardt, 2015; Skiba et al., 2002). Ähnliche Befunde zeigen sich in Reaktionen auf störendes Verhalten von als Türkisch oder Arabisch wahrgenommenen Schüler*innen in Deutschland (Glock, 2016). Psychologische Erklärungen derartiger Disparitäten fokussieren darauf, dass Lehrkräfte ihren Schüler*innen negative Stereotypen zuschreiben. Jedoch berücksichtigen psychologische Erklärungen seltener inwiefern strukturelle Eigenschaften von Bildungskontexten die soziale Kognition und das Entscheidungsverhalten von Lehrkräften beeinflussen. Um diese Forschungslücke zu füllen wurde untersucht, inwieweit kontextbezogene Faktoren sich auf soziale Kognition und Verhaltensintentionen bei Lehrkräften auswirken und somit potentiell zu Disparitäten im Bildungswesen beitragen. Eine zentrale Annahme der vorliegenden Untersuchung war, dass Reaktionen von Lehrkräften nicht nur durch Stereotype gegenüber einzelnen Schüler*innen geprägt werden, sondern auch von unmittelbaren Kontextfaktoren wie der Zusammensetzung von Schulklassen abhängen. Eine präregistrierte Studie untersuchte diese angenommenen kausalen Effekte der Klassenzusammensetzung auf die soziale Kognition und Verhaltensintentionen anhand einer Stichprobe von N = 384 zukünftigen Lehrkräften (Studierenden des Lehramts), die an einem Online-Experiment mit zwei Experimentalbedingungen teilnahmen. Die zentrale Hypothese lautete, dass Teilnehmende bei Beschreibungen von störendem Verhalten im Klassenzimmer strengere Disziplinarmaßnahmen befürworten, wenn die Störungen in Klassen mit einem höheren Anteil von als arabisch wahrgenommenen Schüler*innen auftreten. Die Befunde stützten diese Hypothese. Teilnehmende, die mit Störverhalten in einer Klasse mit überwiegend als arabisch wahrgenommenen Schüler*innen konfrontiert wurden, fühlten sich durch das Verhalten stärker gestört als diejenigen, die mit Fehlverhalten in einer Klasse mit als mit überwiegend deutsch wahrgenommenen Schüler*innen konfrontiert wurden ($d_z = 0.27$). Die Teilnehmenden befürworteten in Klassen mit als arabisch wahrgenommenen Schüler*innen zudem eher strengere Formen der Disziplinierung ($d_z = 0.21$). Teilnehmende erlebten den gesamten Klassenkontext in Klassen mit als arabisch wahrgenommenen Schüler*innen zudem als herausfordernder und negativer ($d_z = 0.36$). Zusammenfassend liefert die Studie empirische Belege für kontextbasierte Diskriminierung und deutet darauf hin, dass strukturelle Faktoren wie die Zusammensetzung von Schulklassen Auswirkungen auf die soziale Kognition und das Entscheidungsverhalten im Klassenzimmer haben können. Die untersuchten Faktoren wirken auf Kontextebene und betreffen demnach viele Schüler*innen gleichzeitig. Dies unterstreicht die Relevanz und die Notwendigkeit zukünftiger Forschung zu kontextbasierter Diskriminierung sowie zu dessen Auswirkungen auf Disparitäten im Bildungswesen.

Sei (k)ein Mann! Eine videobasierte Analyse von Unterschieden im Disziplinarverhalten von Grundschullehrkräften in Abhängigkeit vom Geschlecht der Schüler*innen

Sophie Harms¹, Sarah Gentrup², Aileen Edele¹

¹Humboldt-Universität zu Berlin, Berliner Institut für empirische Integrations- und Migrationsforschung, ²Institut zur Qualitätsentwicklung im Bildungswesen

Interaktionen zwischen Lehrkräften und Schüler*innen unterscheiden sich in vielen Aspekten je nach Geschlecht der Kinder. So beschreiben Lehrkräfte die Beziehungen zu ihren Schülern bereits zu Beginn der Grundschule im Durchschnitt als weniger gut als die Beziehungen zu ihren Schülerinnen (Horn et al., 2021). Weiterhin assoziieren Lehramtsstudierende Jungen eher mit störendem Verhalten als Mädchen (Glock & Kleen, 2017) und praktizierende Lehrkräfte ermahnen Jungen häufiger als Mädchen (Beaman et al., 2006). Dies kann wiederum die Lehrkraft-Schüler*innen-Beziehung, aber auch Lernprozesse direkt beeinträchtigen (Morgan & Sideridis, 2013; Nurmi, 2012) und sich somit nachteilig auf die schulische Entwicklung der Jungen auswirken (O'Connor et al., 2011; O'Connor & McCartney, 2007). Naheliegender ist, dass die häufigere Disziplinierung von Jungen auf das Verhalten der Schüler zurückgeht. So ist gut belegt, dass Jungen den Unterricht häufiger stören als Mädchen (Arbuckle & Little, 2004; McClowry et al., 2013; Trautwein et al., 2004). Denkbare ist aber auch, dass Lehrkräfte in ihrem Umgang mit störendem Verhalten von Geschlechterstereotypen beeinflusst werden und ihre Disziplinarmaßnahmen nicht nur vom Verhalten der Schüler*innen abhängt, sondern auch von deren Geschlecht. Aus den USA ist gut belegt, dass Lehrkräfte Schüler*innen sozialer Minoritäten, die von negativen Stereotypen betroffen sind, bei gleichem Verhalten strenger disziplinieren als Schüler*innen, die der Majorität angehören (Bradshaw et al., 2010; Sullivan et al., 2013). Das Disziplinarverhalten der Lehrkräfte gegenüber minorisierten Schüler*innen scheint somit stereotypbasierten Verzerrungen zu unterliegen. Ob die Unterschiede in der Disziplinierung je nach Geschlecht der Schüler*innen vollständig auf das Verhalten der Schüler*innen zurückführbar sind, oder ob auch hier Verzerrungen aufseiten der Lehrkräfte bestehen, ist bislang unklar. Eine aktuelle Untersuchung aus Deutschland ergab, dass Lehrkräfte auch unter Kontrolle der Leistung der Schüler*innen ihr eigenes Verhalten in Interaktionen mit Jungen als eher fordernd-streng und ermahnend beschrieben als in Interaktionen mit Mädchen (Hannover et al., 2022). Das Verhalten der Schüler*innen wurde dabei jedoch nicht erfasst.

Der vorliegende Beitrag adressiert diese Lücke und untersucht auf Basis von Videodaten (Kristen et al., 2018b) die Frage, ob Lehrkräfte Jungen auch bei gleichem Schüler*innenverhalten häufiger disziplinieren als Mädchen und ob sie dabei mehr strafende und weniger lernförderliche Disziplinarmaßnahmen einsetzen. Mithilfe selbstentwickelter niedrig-inferenter Kodierschemata wird aktuell das Verhalten von Grundschüler*innen (voraussichtliches N ca. 180) und ihren Lehrkräften (N = 17) in jeweils 10-minütigen Ausschnitten aus je einer videografierten Mathematik- und einer Deutschstunde pro Lehrkraft kodiert. Alle untersuchten Lehrkräfte identifizierten sich im Fragebogen als weiblich, wodurch Einflüsse der Geschlechtsidentität der Lehrkraft auf die Interaktion mit den Schüler*innen konstant gehalten werden. Die Videos wurden in der Mitte des ersten Schuljahres aufgenommen. Die Kodierschemata beinhalten Beschreibungen von 35 unterschiedlich stark störenden Schüler*innenverhaltensweisen und 24 Disziplinarmaßnahmen aufseiten der Lehrkräfte. Mithilfe von Mehrebenenmodellen soll zunächst überprüft werden, ob sich der Befund bestätigen lässt, dass Jungen häufiger, eher strafend und weniger lernförderlich diszipliniert werden als Mädchen (H1a und H1b). In einem zweiten Schritt soll getestet werden, ob diese Unterschiede auch unter Kontrolle des Schüler*innenverhaltens bestehen bleiben (H2a und H2b). Da die Studie als Registered Report bei einer Fachzeitschrift eingereicht werden soll, werden voraussichtlich keine Ergebnisse präsentiert. Die Bestätigung der ersten beiden Hypothesen würde jedoch darauf hinweisen, dass Jungen bereits zu Beginn ihrer Schullaufbahn häufiger in für ihre schulische Entwicklung ungünstige Interaktionen mit ihren Lehrkräften involviert sind als Mädchen. Sollte sich auch die Hypothesen 2a und 2b bestätigen, würde dies auf eine geschlechterbezogene Verzerrung im Umgang mit störenden Schüler*innen durch die Lehrkräfte hindeuten, wodurch Jungen gegenüber Mädchen zusätzlich benachteiligt wären.

S-06: Chancenungleichheit als gesellschaftliche Herausforderung: Interdisziplinäre und internationale Perspektiven auf die Förderung bildungsbenachteiligten Schüler:innen in Deutschland und Schweden

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: SN 163

Chair der Sitzung: Haiqin Ning

Chair der Sitzung: Marianne Schüpbach

Chancenungleichheit als gesellschaftliche Herausforderung: Interdisziplinäre und internationale Perspektiven auf die Förderung bildungsbenachteiligten Schüler:innen in Deutschland und Schweden

Chair(s): **Haiqin Ning** (Freie Universität Berlin), **Marianne Schüpbach** (Freie Universität Berlin)

Mit der zunehmenden Vielfalt der Schülerschaft sehen sich immer mehr Gesellschaften in der Welt mit der Herausforderung konfrontiert, gleiche Bildungschancen für alle Schüler:innen zu schaffen, einschließlich derjenigen, die sich in einer benachteiligten Situation befinden, wie es im Sustainable Development Goal 4.5 der UNESCO gefordert wird (UNESCO, o. D.). Empirische Studien zeigen, dass Kinder mit Migrationshintergrund und niedrigem sozioökonomischen Status in vielen Bildungssystemen niedrigere Leistungen erbringen und geringere Bildungschancen haben als ihre Mitschüler:innen (z.B. AB, 2022; Boman & Wiberg, 2024). Dies lässt sich nicht nur im Zusammenhang mit den individuellen Lernvoraussetzungen der Kinder, sondern auch mit dem familiären Hintergrund und den institutionellen Rahmenbedingungen betrachten. Um dieser Herausforderung zu begegnen, bedarf es deshalb interdisziplinäre Ansätze und internationale Perspektiven, die sowohl die Bildungsfinanzierungspolitik, das individuelle kulturelle Kapital als auch die Qualität von Bildungsangeboten einbeziehen.

Das Symposium widmet sich der zentralen Frage: Wie kann die Chancengerechtigkeit von bildungsbenachteiligten Schüler:innen auf verschiedenen Ebenen und in unterschiedlichen Kontexten gefördert werden? Dieser Frage gehen drei Beiträge aus verschiedenen Disziplinen – der Grundschulpädagogik, der Bildungsökonomie und -soziologie – nach. Dabei wird ein Blick sowohl auf die Makroebene des Bildungssystems als auch auf die Mikroebene des Bildungsangebots geworfen. Methodisch umfassen die Beiträge sowohl quantitative als auch qualitative Verfahren. Neben der deutschen werden auch schwedische und weitere internationale Perspektiven einbezogen, um vielfältige Möglichkeiten des Umgangs mit Chancenungleichheiten aufzuzeigen.

Der erste Beitrag beschäftigt sich mit der Bedeutung des individuellen kulturellen Kapitals für die Chancengerechtigkeit. Basierend auf quantitativen Analysen von Daten des Nationalen Bildungspanels Deutschlands wird festgestellt, dass sich das individuelle kulturelle Kapital aus dem in der Familie erworbenen sowie dem durch individuelle Neigungen entwickelten Anteil zusammensetzt. Für sozioökonomisch benachteiligten Schüler:innen kann kulturelles Kapital Aufstiegschancen bieten, insbesondere der Anteil, der durch individuelle Zuneigung erworben wird. **Der zweite Beitrag** widmet sich einer qualitativen Untersuchung der Prozessqualität außerunterrichtlicher Angebote an Ganztagsgrundschulen in Deutschland zur Förderung bildungsbenachteiligter Schüler:innen. Dabei werden konkrete Realisierungsideen der durch Interviews mit internationalen wissenschaftlichen Expert:innen gewonnenen Dimensionen der Prozessqualität mit Hilfe von Fokusgruppen im deutschen Kontext erhoben. **Der letzte Beitrag** befasst sich mit der bedarfsorientierten Schulfinanzierung im schwedischen Kontext, d.h. mit der bedarfsorientierten Zuweisung von Finanzmitteln auf der Ebene der einzelnen Schulen. Als Teil eines größeren Projekts untersucht er den Einfluss dieses Ansatzes auf die Bewegung qualifizierter Lehrkräfte unter Berücksichtigung des SES-Index der Schulgemeinde.

Abschließend werden die Beiträge unter der Moderation der Chairs mit dem Publikum anhand von der oben genannten Leitfrage diskutiert, einschließlich der Unterschiede und Gemeinsamkeiten zwischen interdisziplinären und internationalen Perspektiven.

Beiträge des Symposiums

Kulturelles Kapital in der Migrationsgesellschaft. Über individuelles kulturelles Kapital als Determinante des Bildungserfolgs

Jannis Burkhard

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Theoretischer Hintergrund

Dass kulturelles Kapital förderlich für den Erwerb von Bildungsabschlüssen ist, gilt international als robuster Befund. Dieses kann sowohl die Reproduktion sozialer Strukturen fördern, als auch soziale Mobilität ermöglichen (Bourdieu & Passeron, 1971; DiMaggio, 1982). Gleichzeitig konnte herausgearbeitet werden, dass die Unterscheidung zwischen elterlichem und individuellem kulturellem Kapital bedeutsam ist (Chin & Phillips, 2004). Unter dieser Perspektive wird das individuelle kulturelle Kapital, in Abgrenzung zum kulturellen Kapital der Eltern, zur Ressource für sozialen Aufstieg. Gleichzeitig ist Deutschland inzwischen das Land mit der weltweit zweitgrößten Nettozuwanderung (McAuliffe & Triandafyllidou, 2022). Dies wirft die Frage auf, inwiefern sich die Definition und Wirkungsweisen von kulturellem Kapital durch Zuwanderung verändern können.

Fragestellung

Der Beitrag widmet sich der Frage, welche Bedeutung dem individuellen kulturellen Kapital für die Chancengerechtigkeit in einer Migrationsgesellschaft zukommt. Dafür werden die Ergebnisse einer Dissertation vorgestellt, die sich der Frage in drei Teilstudien mit ihren jeweiligen Forschungsfragen widmet: 1) Wie stark hängt das individuelle kulturelle Kapital von dem der Eltern ab? 2) Wie können Jugendliche auch unabhängig von ihren Eltern kulturelles Kapital aufbauen? 3) Begünstigt das individuell erworbene kulturelle Kapital den Abiturerwerb? Zur Beantwortung dieser Fragen werden fünf Formen kultureller Betätigung über alle Teilstudien hinweg untersucht: [1] hochkulturelle Aktivitäten, [2] der Besuch einer Schule mit musikalischem Profil, [3] außerschulische Kurse, [4] kulturelle Vereinsaktivitäten und [5] kulturelle Aktivitäten im Jugendzentrum. Durch die drei Teilfragen wird insgesamt der Frage nachgegangen, wie das individuelle kulturelle Kapital inner- und außerfamiliär erworben wird und welchen Effekt dieses auf den Bildungserwerb hat.

Methode

Alle Analysen beruhen auf Daten der Startkohorte 3 des Nationalen Bildungspanels (Blossfeld & Roßbach, 2019). Je nach Teilstudie liegt die Stichprobengröße zwischen 5778 und 8317 Schüler:innen. In den Teilstudien wurden unterschiedliche quantitative Verfahren verwendet: Strukturgleichungsmodelle, Entropy Balancing und Regressionen mit Dekompositionsanalysen.

Ergebnisse und ihre Bedeutung

In der Gesamtschau zeigt sich, dass die kulturellen Tätigkeiten mit der sozialen Herkunft assoziiert sind (Teilstudie 1) und gleichzeitig von individuellen Neigungen abhängen (Teilstudie 2). Dies bedeutet auf konzeptueller Ebene, dass sich das individuelle kulturelle Kapital aus zwei Teilen zusammensetzt: dem in der Familie erworbenen und dem durch individuelle Neigungen erworbenen Anteil. Schließlich zeigt sich, dass letzterer die Wahrscheinlichkeit, das Abitur zu erwerben, erhöhen kann (Teilstudie 3). Insgesamt verdeutlichen die Ergebnisse, dass kulturelles Kapital sozioökonomisch benachteiligten Kindern Aufstiegschancen bieten kann. Der Beitrag schließt mit einem Ausblick auf die Bedeutung für durch Zuwanderung geprägte Gesellschaften: Wie verändert sich die Definition von "Hoch"kultur durch Zuwanderung? Und welches Potenzial bietet dafür die Idee, dass sich die Vorstellung davon, was kulturelles Kapital ist, je nach individueller sozialer Gruppe unterschiedlich sein kann (vgl. Davies & Rizk, 2017)? Diese Fragen werden abschließend diskutiert.

Prozessqualität außerunterrichtlicher Angebote zur Förderung bildungsbenachteiligter Schüler:innen. Eine qualitative Studie zur Konkretisierung internationaler Perspektiven für die Ganztagsgrundschulentwicklung in Deutschland.

Jule Schmidt¹, Haigin Ning¹, Nanine Lilla¹, Wim Willem Nieuwenboom², Marianne Schüpbach¹

¹Freie Universität Berlin, ²Fachhochschule Nordwestschweiz

An den Ausbau der Ganztags(grund)schule in Deutschland ist die Erwartung geknüpft, allen Schüler:innen unabhängig ihres sozialen Status, ihrer Migrationsgeschichte sowie anderer Diversitätsdimensionen gleiche Bildungschancen zu eröffnen (Kielblock, 2021). Während die Förderung von Chancengerechtigkeit im deutschen Ganztagsgrundschulkontext bislang nicht ausreichend gelingt (Sauerwein, Thieme, & Chiapparini, 2019), gibt es international erweiterte Bildungsangebote (in Englisch: extended education, vgl. Schüpbach, 2018), die vergleichbar sind und die intendierten Wirkungen erzielen, insbesondere wenn diese eine hohe pädagogische Qualität aufweisen (Durlak et al., 2010). Pädagogische Qualität setzt sich nach Tietze et al. (2005) aus den drei Ebenen Prozess-, Orientierungs-, und Strukturqualität zusammen. Insbesondere die Prozessqualität, das heißt die Interaktionen zwischen den Schüler:innen und dem pädagogisch tätigen Personal sowie die pädagogischen Stimulationen, sind entscheidend für die Wirksamkeit der außerunterrichtlichen Angebote (Tietze et al., 2005).

In der vorliegenden Studie ist die Frage leitend, wie eine hohe Prozessqualität zur Förderung bildungsbenachteiligter Schüler:innen in außerunterrichtlichen Angeboten im deutschen Ganztagsgrundschulkontext erzielt werden kann.

Als Datengrundlage werden Transkripte aus 14 Fokusgruppen (Zwick & Schröter, 2012) herangezogen, die mit Personen, die in unterschiedlichen Funktionen an der Gestaltung der außerunterrichtlichen Angebote an offenen Ganztagsgrundschulen in vier verschiedenen Bundesländern beteiligt sind, z.B. Schulleitungen, Ganztagskoordinator:innen, Lehrkräfte, Erzieher:innen, Sozialpädagog:innen, außerschulische Kooperationspartner:innen sowie Elternvertretungen, stattfanden. Die Teilnehmenden der Fokusgruppen diskutierten aus ihren unterschiedlichen Perspektiven die Realisierung pädagogischer Qualität mit Blick auf die Förderung bildungsbenachteiligter Schüler:innen.

Die anonymisierten Transkripte wurden mittels inhaltlich strukturierender qualitativer Inhaltsanalyse in einem deduktiv-induktiven Verfahren (Kuckartz & Rädiker, 2022) und unterstützt durch MAXQDA ausgewertet. Als deduktive Kategorien dienten drei Dimensionen der Prozessqualität (Autor:innen, eingereicht), welche sich in einer Vorstudie in Expert:inneninterviews (Witzel, 2000) mit internationalen Bildungsforscher:innen aus sechzehn Ländern weltweit hinsichtlich der Förderung bildungsbenachteiligter Schüler:innen relevant zeigten: (a) konzeptionelles Prinzip: Kindzentriertheit, (b) Formate und Methoden: Variation und (c) Interaktionen zwischen Personal und Kindern: Fürsorge und Begleitung. Ergänzend zu den deduktiven Kategorien wurden induktiv Subkategorien gebildet.

Erste Ergebnisse zeigen Konkretisierungen für die mögliche Umsetzung der Qualitätsdimensionen auf. Um Kindzentriertheit zu erzielen, werden die Definition individueller (Etappen-)Lernziele und systematisches Feedback zur Angebotskonzeption durch die Schüler:innen als wirkungsvoll eingeschätzt. Hinsichtlich der Variation sprechen die Ergebnisse vor allem für erfahrungsbasiertes Lernen im Sozialraum und die Etablierung von Pat:innensystemen, in denen leistungsschwächere Schüler:innen von leistungstärkeren Schüler:innen beispielsweise beim Lesenlernen unterstützt werden. Mit Blick auf Fürsorge und Begleitung zeigen sich konsistente Regeln, eine offene Frage-Kultur sowie Aufgabenstellungen und Materialien differenziert entlang unterschiedlicher Entwicklungsstände als mögliche Konkretisierungen für die Realisierung hoher Prozessqualität.

Die gewonnenen Ergebnisse geben wichtige Hinweise für einen möglichen Transfer der aus internationaler Perspektive erwachsenen Dimensionen der Prozessqualität und bieten für die Bildungspraxis und -politik sowie weitere Forschung Implikationen für die Förderung von Chancengerechtigkeit und die Notwendigkeit einer diversitätssensiblen Ganztags(grund)schulentwicklung.

ENTFÄLLT: Ist die bedarfsorientierte Schulfinanzierung ein Weg zur Entkopplung von sozialer Herkunft und Bildungserfolg?

Annika Elwert

Department of Sociology, Lund University, Sweden

Der Beitrag untersucht die Effekte der bedarfsorientierten Schulfinanzierung auf den Bildungserfolg von Schüler:innen in Schweden. In Schweden wie in anderen Ländern gibt es einen starken Zusammenhang zwischen sozialer Herkunft und Bildungserfolg (Eriksson et al. 2021; Kao und Thompson 2003). Kinder mit Migrationshintergrund und Kinder aus bildungsfernen Haushalten erzielen schlechtere Noten, verlassen das Bildungssystem häufiger ohne Abschluss und haben insgesamt kürzere Bildungskarrieren. Ein in der internationalen Literatur vielversprechender Ansatz, um die Kopplung zwischen sozialer Herkunft und Bildungserfolg abzuschwächen, ist die bedarfsorientierte Schulfinanzierung (Card und Payne 2002; Rothstein und Schanzenbach 2022). Dabei wird die soziale Zusammensetzung der Schülerschaft im Schulbezirk bzw. an den einzelnen Schulen explizit bei der Zuweisung von Finanzmitteln berücksichtigt, sodass Schulen mit einer sozial schwächeren Schülerschaft mehr Finanzmitteln erhalten.

In Schweden liegt die Hoheit über die Schulfinanzierung bei den Gemeinden und Kommunen. Auf kommunaler Ebene wurde die bedarfsorientierte Bildungsfinanzierung bereits Ende der 1990er Jahre von einzelnen Kommunen eingeführt und hat sich seither rasant verbreitet (SKR 2018). Eine Besonderheit des schwedischen Systems der bedarfsorientierten Schulfinanzierung im internationalen Vergleich ist, (1) dass die bedarfsorientierte Schulfinanzierung auf Ebene der einzelnen Schule (und nicht wie in anderen Ländern üblich auf Ebene des Schulbezirks) erfolgt. Dies ermöglicht eine sehr zielgenaue finanzielle Ausstattung der Schulen, die für einen höheren Personalschlüssel, Sprachförderung oder die Einstellung qualifizierter Lehrkräfte genutzt werden kann. (2) Die Lehrergehälter werden nicht nach Schablonenprinzip festgelegt, sondern sind individuell verhandelbar. Dies

ermöglicht es, an Schulen mit sozial schwächerer Schülerschaft höhere Gehälter zu zahlen, um besonders qualifizierte Lehrkräfte zu gewinnen. (3) Die bedarfsorientierte Schulfinanzierung setzt bereits im Vorschulbereich an, was besonders effektiv sein kann.

Mit Hilfe ökonometrischer Methoden der kausalen Inferenz (Difference-in-Difference-Modelle) wird untersucht, ob die bedarfsorientierte Schulfinanzierung benachteiligten Schüler:innen mit Migrationshintergrund zugutekommt. Bei der Untersuchung der Mechanismen, die möglichen Effekten zugrunde liegen, konzentriert sich der Beitrag auf die Rolle der Arbeitsmärkte für Lehrkräfte, einschließlich der Gehältern und der Fluktuation von Lehrkräften. Der Beitrag ist Teil eines interdisziplinären Projekts der Bildungsökonomie und der soziologischen Ungleichheitsforschung. Um die Effekte auf die Schülerschaft zu untersuchen, werden longitudinale schwedische Registerdaten (Mikrodaten) verwendet, die Informationen über die gesamte schwedische Schüler- und Lehrerschaft erhalten, wie z.B. Schulzugehörigkeit, Lehrergehältern, Schulnoten und standardisierte Tests. Diese Daten werden mit Makrodaten zur Schulfinanzierung kombiniert, um die Verwendung der oben genannten Kausalmethoden zu ermöglichen.

Vorläufige Analysen zeigen, dass 72% (58%) aller Gemeinden die bedarfsorientierte Finanzierung im Primarbereich (Vorschulbereich) eingeführt haben. In den letzten 25 Jahren ist der Anteil der Gemeinden, die die Richtlinie im Primarbereich eingeführt haben, kontinuierlich gestiegen und erreichte 2010 14%, 2015 36% und 2020 57%. Ein ähnlicher, wenn auch weniger steiler Anstieg ist bei der Einführung der Richtlinie im Vorschulbereich zu beobachten. Die Korrelationsanalyse zeigt, dass Gemeinden, die früh mit der Einführung begonnen haben (early adopters), im Jahr 2022 einen höheren Anteil an zertifizierten Lehrkräften aufweisen als Gemeinden, die die Einführung später umgesetzt haben (late adopters). Trotz ihres deskriptiven Charakters deuten die Ergebnisse darauf hin, dass die bedarfsorientierte Finanzierung die Möglichkeiten der Gemeinden, qualifizierte Lehrkräfte anzuziehen, erhöhen könnte. Unter Berücksichtigung eines Gemeinde-SES-Index lässt sich zudem feststellen, dass die bedarfsorientierte Finanzierung zur Förderung der Chancengerechtigkeit zugunsten der Schüler:innen aus Gemeinden mit niedrigem SES beitragen kann.

S-07: Domänenübergreifende und domänenspezifische Bildungsprozesse in der frühen Kindheit

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: SO 108

Chair der Sitzung: Elisa Oppermann

Chair der Sitzung: Hans-Günther Roßbach

Domänenübergreifende und domänenspezifische Bildungsprozesse in der frühen Kindheit

Chair(s): **Elisa Oppermann** (Universität Leipzig)

Diskutant*in(nen): **Hans-Günther Roßbach** (Otto-Friedrich-Universität Bamberg)

Der Bildungsauftrag in der frühen Kindheit in Kindertageseinrichtungen (Kitas) ist ganzheitlich und zielt auf die Förderung kindlicher Kompetenzen u.a. in den Bildungsbereichen Sprache, Mathematik und Naturwissenschaften (Hebenstreit, 2008; KMK/JFMK 2022). Forschungsbefunde zeigen, dass der kindliche Kompetenzerwerb domänenspezifisch erfolgt und abhängig ist von der Häufigkeit und (kognitiven) Anregungsqualität der Lerngelegenheiten in der frühen Kindheit (Ulferts et al., 2019; von Suchodoletz et al., 2023). Lerngelegenheiten in der Kita finden jedoch selten klar getrennt nach Domänen statt. Ausgehend von einer situationsorientierten Frühpädagogik in Deutschland, werden alltagsintegrierte und spielbasierte Lerngelegenheiten betont. Dies birgt ein hohes Potenzial für domänenübergreifende Bildungs- und Lernprozesse, z.B. wenn ein Konflikt um die faire Verteilung von Spielzeug (sozial-emotionale Kompetenzen) durch verbale Aushandlungen (Sprache) bzw. Abzählen (Mathematik) gelöst wird. Bisher ist kaum erforscht, inwieweit Bildungsprozesse in der Kita domänenspezifisch versus domänenübergreifend stattfinden und in welchem Zusammenhang diese mit domänenspezifischen Kompetenzen der Kinder stehen. Das geplante Symposium widmet sich dieser Forschungslücke, indem es Studien zusammenführt, die Bildungsprozesse und kindliche Kompetenzen in verschiedenen Domänen vergleichend betrachten (u.a. Sprache, Mathematik, Naturwissenschaften und sozial-emotionale Entwicklung). Dadurch ermöglicht das Symposium eine umfassende Analyse domänenspezifischer und domänenübergreifender Bildungsprozesse sowie eine weiterführende Diskussion über die betrachteten Domänen hinweg.

Der erste Beitrag untersucht Bildungsprozesse in den Domänen Mathematik und Naturwissenschaften mithilfe von Beobachtungen des Kita-Alltags, wobei die kognitive Anregungsqualität und die Häufigkeit von domänenspezifischen sowie domänenübergreifenden Bildungsprozessen im Vergleich betrachtet werden.

Der zweite Beitrag untersucht die domänenspezifische Anregungsqualität in den Naturwissenschaften im Zusammenhang mit der Entwicklung kindlicher Fähigkeiten in den Domänen Sprache und Naturwissenschaften.

Der dritte Beitrag fokussiert die zielkindbezogene Anregungsqualität und kindliche Fähigkeiten jeweils im sprachlichen und sozial-emotionalen Bereich.

Im Fokus des vierten Beitrags stehen die domänenspezifischen und domänenübergreifenden Kompetenzprofile von Kindern in den Bildungsbereichen körperlich-motorische Entwicklung, Sprache, Mathematik, Naturwissenschaft, digitale Medien sowie sozial-emotionale Entwicklung und Wohlbefinden.

Insgesamt tragen die Beiträge des geplanten Symposiums zu einem besseren Verständnis der Umsetzung domänenübergreifender sowie domänenspezifischer Bildungsprozesse in der frühen Kindheit sowie deren Bedeutung für die kindliche domänenspezifische Entwicklung bei.

Beiträge des Symposiums

Frühe mathematische und naturwissenschaftliche Bildungsprozesse in der Kita: Beobachtung alltagsintegrierter Fachkraft-Kind-Dialoge

Elisa Oppermann¹, Simone Dunekacke²

¹Universität Leipzig, ²RPTU Kaiserslautern-Landau

Theoretischer Hintergrund:

Domänenspezifische Lerngelegenheiten in Mathematik und Naturwissenschaften in Kindertageseinrichtungen sind Voraussetzung für den Erwerb frühkindlicher mathematischer und naturwissenschaftlicher Kompetenzen. Dazu gehören insbesondere Lerngelegenheiten, bei denen Lerninhalte in Alltagssituationen (z.B. im kindlichen Spiel) eingebettet werden (Gasteiger et al., 2020; Steffensky, 2017). Aus fachdidaktischer und kindheitspädagogischer Perspektive ist es denkbar, dass gerade Alltagssituationen Inhalte aus verschiedenen Domänen adressieren, beispielsweise kann das Eingießen von Wasser beim Tischdecken mathematische Aspekte enthalten (Ermitteln der Wassermenge) sowie auch naturwissenschaftliche Aspekte (Bedeutung von Wasser für die Umwelt). So fördern Ansätze des domänenübergreifenden Lernens die Vernetzung verschiedener Lerninhalte (Falloon et al., 2020). In der Forschungsliteratur wird mit Blick auf die Umsetzung früher Bildungsprozesse vor allem die Bedeutung kognitiv anregender Dialoge zwischen Fachkraft und Kind(ern) hervorgehoben (vgl. Hopf, 2012; Siraj-Blatchford et al., 2002). Bislang ist kaum erforscht, wie häufig Alltagssituationen für domänenübergreifende bzw. domänenspezifische Dialoge in Mathematik und Naturwissenschaften genutzt werden und wie der kognitive Anregungsgehalt in solchen Dialogen ist. Dieses Forschungsdesiderat adressiert der vorliegende Beitrag und bildet damit eine wichtige Grundlage für weiterführende Forschung zu Bedingungsfaktoren und Wirkmechanismen frühe mathematischer und naturwissenschaftlicher Bildung.

Fragestellung:

Wie fallen die Häufigkeit sowie kognitive Anregungsqualität alltagsintegrierter Fachkraft-Kind-Dialoge in Mathematik und Naturwissenschaften aus? Zeigen sich domänenspezifische Unterschiede bzw. Überschneidungen?

Methode:

Datengrundlage bilden Beobachtungen des pädagogischen Alltags über drei bis vier Stunden bei 21 Fachkräften aus 14 österreichischen Kindertageseinrichtungen. Die Fachkräfte waren zu 89% weiblich und hatten im Mittel 11.7 Jahre (SD = 11.5) Berufserfahrung. Die Fachkräfte betreuten durchschnittlich 22 Kinder (SD = 2.4) im Alter von drei bis sechs Jahren. Es wurde ein von den Autorinnen selbst entwickeltes Beobachtungsinstrument eingesetzt, mit dem die Häufigkeit von Fachkraft-Kind-Dialogen dokumentiert sowie die kognitive Anregungsqualität der Dialoge kodiert wurde. Aufgrund der Annahme, dass alltagsintegrierte Lerngelegenheiten meist von kurzer Dauer sind (kurze Dialoge < 3 min; Hopf, 2012) wurden in Anlehnung an vorhandene Ratingsysteme (u.a. Leuchter & Saalbach, 2014; Siraj-Blatchford et al., 2015; Studhalter, 2017) vier Indikatoren entwickelt und anhand konkreter Beispiele geschrieben, welche ein Spektrum von geringer (1) bis hoher (4) Anregungsqualität aufwiesen.

Ergebnisse und ihre Bedeutung:

Insgesamt wurden N = 445 Dialoge erfasst, dabei kamen mathematische (43.9%) und naturwissenschaftliche (46.9%) Inhalte vergleichbar häufig vor; 9.2 % der Dialoge waren domänenübergreifend. Im Mittel erstreckten sich die Dialoge über 3.84 Minuten (SD = 7.43), wobei domänenübergreifende Dialoge länger ausfielen (M = 8.97 Minuten, SD = 7.27) als domänenspezifische Lerngelegenheiten in Mathematik (M = 3.59 Minuten, SD = 11.62) oder Naturwissenschaften (M = 3.11 Minuten, SD = 6.18). In Mathematik dominierte der Inhaltsbereich Zahlen, Mengen und Operationen, in Naturwissenschaften waren es die Inhalte der belebten Natur. Die meisten Dialoge wurden im Freispiel in Form von kurzen, weitgehend von den Kindern initiierten, Interaktionen beobachtet. Die kognitive Anregungsqualität der Lerngelegenheiten fiel in beiden Bereichen gering aus; etwas höhere Qualitätsratings konnten in domänenübergreifenden Lerngelegenheiten erreicht werden.

Zusammengenommen unterstreichen die Befunde das Potenzial alltagsintegrierter domänenübergreifender Lerngelegenheiten sowie die Notwendigkeit der Professionalisierung frühpädagogischer Fachkräfte mit Blick auf das Erkennen und Umsetzen regelmäßiger und kognitiv anregender Lerngelegenheiten in den Bereichen Mathematik und Naturwissenschaften. Durch weitere Studien sollte zudem untersucht werden, ob und wie das Potenzial der Dialoge nachhaltiger genutzt werden könnte, beispielsweise durch spezifische Fortbildungen (Schmitt, 2023).

Sprachliches und domänenspezifisches Lernen von Kindern im Kontext früher naturwissenschaftlicher Bildungsangebote

Ada Cecil Haen¹, Anika Bürgermeister², Mirjam Steffensky¹, Ilonca Hardy³, Franziska Bednarski², Miriam Leuchter⁴, Sina-Kristin Koschick³, Laura Venitz⁵, Lukas Lazzara⁴, Julia Barentzien¹, Henrik Saalbach²

¹Universität Hamburg, ²Universität Leipzig, ³Goethe-Universität Frankfurt am Main, ⁴RPTU Kaiserslautern-Landau, ⁵Hochschule für Technik und Wirtschaft des Saarlandes

Theoretischer Hintergrund:

Der Sprachfähigkeit wird eine zentrale Bedeutung für Lern- und Entwicklungsprozesse von Kindern zugesprochen (z.B. Schneider et al., 2012). Dies gilt für domänenübergreifende, aber auch domänenspezifische Bereiche wie den Naturwissenschaften (z.B. Haug & Ødegaard, 2014). Der Kindertagesstätte (Kita) kommt dabei die wichtige Rolle zu, Kinder mit unterschiedlichen Voraussetzungen sprachlich zu fördern. Häufig findet eine solche Förderung im Kontext der Förderung anderer Bildungsbereiche statt (Kempert et al., 2019).

Entscheidend dafür, dass sprachliche und domänenspezifische Lernprozesse stattfinden, ist u.a. die Qualität der Interaktionen zwischen Fachkraft und Kind bzw. zwischen Kindern. Interaktionsqualität wird häufig generisch konzeptualisiert und umfasst dabei z.B. die drei Domänen Emotionale Unterstützung, Gruppenführung und Anregungsqualität (Classroom Assessment Scoring System – CLASS-PreK; Pianta et al., 2008). Die Interaktionsqualität wird durch die professionelle Kompetenz der pädagogischen Fachkraft (und weiteren Merkmalen, vgl. Anders & Oppermann, 2024) beeinflusst. Darüber hinaus wird die aktive (verbale) Partizipation der Kinder als wesentlich für das individuelle Lernen erachtet und wirkt sich auf die Effektivität von aktivierenden und unterstützenden Maßnahmen aus, z. B. in Bezug auf die Aktivierung von Vorwissen (Bürgermeister et al., 2019; Herrmann et al., 2021; Mannel et al., 2016; Sedova et al., 2019), was allerdings im naturwissenschaftlichen Kontext bislang wenig untersucht wurde.

Fragestellung:

Vor diesem Hintergrund zielt die vorliegende Studie darauf ab, die Anregungsqualität pädagogischer Fachkräfte (N=66) sowie das verbale Partizipationsverhalten von Kindern (N=188) in videografierten naturwissenschaftlichen Bildungsangeboten in der Kita zu untersuchen. Darüber hinaus wird durch Multilevel-Regressionsanalysen in einem längsschnittlichen Studiendesign analysiert, ob die Anregungsqualität die Entwicklung der Sprachfähigkeiten und des naturwissenschaftlichen Wissens der Kinder auf Individualebene beeinflusst. Zudem soll untersucht werden, inwieweit die individuelle verbale Beteiligung der Kinder eine moderierende Rolle bei der Entwicklung der Sprachfähigkeiten und des naturwissenschaftlichen Wissens der Kinder einnimmt.

Methode:

Die Anregungsqualität in den Videoaufnahmen (M=28.35 Minuten; SD=8.64; Min=12; Max=46) wurde mit CLASS-PreK (ICC = .89; Pianta et al., 2008) eingeschätzt. Die verbale Partizipation der Kinder wurde anhand der Anzahl der Redebeiträge pro Minute und des Wortumfangs der Redebeiträge operationalisiert und mithilfe der Software MAXQDA 2022 erhoben (VERBI Software, 2021). Die Sprachfähigkeiten der Kinder wurden mittels Wechsler Preschool and Primary Scale of Intelligence-IV (Wechsler, 2018) erfasst. Das naturwissenschaftliche Wissen im Themenbereich „Materialien“ wurde in standardisierten Einzelinterviews erhoben (12 Items, Cronbach's alpha = .80; in Anlehnung an Steffensky et al., 2018).

Ergebnisse und ihre Bedeutung:

Die Einschätzungen der Anregungsqualität in den aufgezeichneten naturwissenschaftlichen Bildungsangeboten weisen auf eine niedrige Qualität hin (M=2.89, SD = .70, Min=1.67, Max=4.67). Die Anzahl der Redebeiträge der Kinder auf Individualebene pro Minute beträgt M=1.71 (SD=1.21, Min=0.00, Max=6.76). Ein Redebeitrag umfasst im Durchschnitt 5.17 Wörter (SD=3.50, Min=0.00, Max=23.25).

Erste Ergebnisse der Mehrebenen-Regressionsanalysen zeigen sowohl in Bezug auf die Entwicklung der Sprachfähigkeiten als auch auf die Entwicklung des naturwissenschaftlichen Wissens signifikante Effekte des Vorwissens als Prädiktor. Die Modelle zeigen keine signifikanten Effekte der Anregungsqualität (Gruppenebene) auf die individuelle Entwicklung der Kinder in den Bereichen Sprache und Naturwissenschaften. Es zeigt sich aber ein positiver Effekt des Wortumfangs der Redebeiträge der Kinder auf die Entwicklung der Sprachfähigkeiten, die Anzahl der Redebeiträge der Kinder pro Minute zeigte hingegen negative Effekte. Es zeigen sich keine Effekte dieser Partizipationsvariablen in den Modellen zum naturwissenschaftlichen Wissen, jedoch haben die Sprachfähigkeiten der Kinder als Prädiktorvariable einen signifikanten Effekt auf die Entwicklung des naturwissenschaftlichen Wissens.

Der ausbleibende Effekt der Anregungsqualität auf Entwicklung im Bereich Sprache und Naturwissenschaften könnte auf einen Mangel hochwertiger Anregungen durch die Fachkräfte in dieser Stichprobe zurückgeführt werden. Des Weiteren limitiert die Kürze der naturwissenschaftlichen Bildungsangebote die Aussagekraft der vorliegenden Analysen. Der signifikante Effekt der Länge der Redebeiträge der Kinder auf ihre sprachliche Entwicklung verdeutlicht die Relevanz des Umfangs von Redebeiträgen für den sprachlichen Lernprozess, und nicht die der reinen Anzahl der Redebeiträge.

Inwieweit profitieren Kinder in ihren Fähigkeiten von einer hohen zielkindbezogenen Anregungsqualität in der Kita?

Theoretischer Hintergrund und Forschungsstand:

Ausgehend von der bioökologischen Theorie (Bronfenbrenner & Morris, 1998) ist die frühkindliche Entwicklung durch das Zusammenspiel der Anregungen und Interaktionen zwischen dem Kind und seinen Lernumgebungen (z.B. Familien, Kita) geprägt. Diese werden als "proximale Prozesse" (Bronfenbrenner & Morris, 1998, S. 996) bezeichnet, die in gängigen wirkungsorientierten Modellen zur Kita-Qualität als Teil der Prozessqualität bezeichnet werden (Kluczniok, 2018). Zahlreiche Studien verweisen auf die hohe Bedeutung der kindlichen Anregungsprozesse in der Kita und deren Qualität für die Entwicklung und den Bildungserfolg der Kinder (Ulferts & Anders, 2016; von Suchodoletz et al., 2023). Auf Ebene der Anregungsprozesse wird dabei zwischen eher globalen Aspekten wie das generelle Interaktionsklima oder die emotionale Unterstützung in der Kita-Gruppe und bereichsspezifischen Aspekten (z.B. sprachliche oder mathematische Bildungsprozesse) unterschieden (Roßbach et al., 2024) und in ihrem jeweiligen Einfluss auf die kindlichen Kompetenzen untersucht (im Überblick: Lehrl, Roßbach & Weinert, 2024). So zeigen Broekhuizen et al. (2016) für globale Qualitätsaspekte, dass Kinder in ihren sozialen Fähigkeiten und im Problemverhalten von einer hohen Qualität emotionaler Unterstützung und Gruppenorganisation im Kitaalter profitieren. Für bereichsspezifische Aspekte berichten beispielsweise Lehrl et al. (2016) positive Zusammenhänge zwischen der mathematikbezogenen Anregung in der Kita und der mathematischen Kompetenzentwicklung der Kinder ins Grundschulalter hinein (ähnlich Anders et al., 2013; Howard et al., 2018). Darüber hinaus finden sich auch domänenübergreifende Zusammenhänge, z.B. zwischen der literacy-bezogenen Anregung in der Kita und frühen mathematischen Kompetenzentwicklung (Howard et al., 2018), allerdings nicht zwischen einer mathematik-bezogenen Anregung und der Wortschatzentwicklung oder sozial-emotionalen Entwicklung (Howard et al., 2018; ähnlich Oppermann et al., 2023).

Die genannten Studien messen jedoch Kita-Qualität auf Gruppenebene und insofern ungenau, da sie häufig von den Interaktionen einzelner Kinder abstrahieren (Chien et al., 2010). Eine auf einzelne Kinder gerichtete zielkindbezogene Anregungsqualität berücksichtigt die Individualität jedes Kindes und ist damit differenzierter als eine gruppenbezogene Qualität. Dies greift der vorliegende Beitrag auf und fokussiert auf die von Kindern erfahrene pädagogische Anregungsqualität in Kitas in ihrer Bedeutung für die kindlichen Fähigkeiten.

Fragestellung:

Der Beitrag untersucht die Zusammenhänge zwischen zielkindbezogener Anregungsqualität und kindlichen Fähigkeiten unter Berücksichtigung struktureller einrichtungs- und kindspezifischer Merkmale (z.B. Fachkraft-Kind-Schlüssel, Familiensprache des Kindes). Im Vordergrund stehen dabei zielkindbezogene Anregungsprozesse und kindliche Fähigkeiten jeweils im sprachlichen und sozial-emotionalen Bereich. Dadurch wird es möglich, domänenspezifische als auch domänenübergreifende Zusammenhänge zu analysieren.

Methode:

Datengrundlage ist das Forschungsprojekt „Qualität in bilingualen Kindertagesstätten“. Die Stichprobe umfasst N = 200 Kinder aus N = 10 Kitas (Abschluss der Datenerhebungen bis 31.07.2024). Die kindbezogenen Anregungsprozesse sind über standardisierte Beobachtungen auf Basis eines neu entwickelten Zielkindverfahrens (Kluczniok, Resa & Voss, 2024) erfasst. Über einen Vormittag hinweg werden ausgewählte Zielkinder (n=200) über jeweils vier zehnminütige Zyklen bei ihren Interaktionen im Kita-Alltag von geschulten Datenerheberinnen beobachtet und anschließend in der Qualität der erfahrenen Anregung auf einer vierstufigen Skala (1=unzureichend, 2=minimal, 3=gut, 4=ausgezeichnet) eingeschätzt. Die kindlichen Fähigkeiten werden durch Eltern und Fachkräfte in den Bereichen Kommunikation (Zuhören und Verstehen, Sprechen) und Soziale Fertigkeiten (Umgang mit anderen, Spielen und Freizeit) anhand der Vineland Adaptive Behavior Scales (Sparrow et al., 2021) eingeschätzt. Über Fragebögen wurden die strukturellen Merkmale auf Einrichtungs-/Gruppen- sowie Kindebene erfasst.

Zur Analyse der Zusammenhangsmuster zwischen zielkindbezogener Anregungsqualität und kindlichen Fähigkeiten werden neben deskriptiven und korrelativen Auswertungen insbesondere multivariate Regressionsmodelle unter Einbezug von Kontrollvariablen auf Einrichtungs- und Kindebene durchgeführt.

Ergebnisse und ihre Bedeutung:

Die Ergebnisse werden vorgestellt, in den Forschungsstand eingebettet und diskutiert. Der erwartete Erkenntnisgewinn bezieht sich auf ein genaueres Verständnis der Anregungsqualität in Kitas auf Zielkindebene in zwei zentralen Bereichen (sprachlich, sozial-emotional) und deren Bedeutung für kindliche Fähigkeiten. Daraus können Erkenntnisse für gezielte Qualitätsentwicklungsmaßnahmen in den Kitas sowie Fördermöglichkeiten von Kindern abgeleitet werden.

ENTFÄLLT: Identifizierung der domänenspezifischen Profile von Kindern am Übergang zur Grundschule mit Hilfe der Übergangsdokumentation „Mika“ – Erkenntnisse zur Notwendigkeit einer individualisierten Bildungsbegleitung im Übergang

Beatrice Rupprecht
Universität Leipzig

Theoretischer Hintergrund und Forschungsstand:

Gemäß der UN-Kinderrechtskonvention hat jedes Kindes den Anspruch auf eine individuelle und bedürfnisbezogene Unterstützung seines Bildungs- und Entwicklungsprozesses (Heimlich & Ueffing, 2018). Nationale Studien zeigen, dass eine entsprechende Individualisierung des kindlichen Bildungsprozesses in der Kitapraxis noch nicht immer hinreichend realisiert wird (Hasselhorn & Kuger, 2014; Smidt, 2012). Es bedarf dabei eines besonderen Augenmerks auf die Anschlussfähigkeit der Bildungsprozesse zwischen der Kita und der Grundschule, um der Kontinuität auch im Lebensabschnitt des Schuleintritts gerecht zu werden (Liebers, 2020; Liebers & Rupprecht, 2018). Um einem Bruch im Bildungsverlauf zu begegnen und eine Verstärkung krisenhafter Erfahrungen, wie bspw. der Rückstellung und dem ausbleibenden Schulerfolg (Liebers, 2022) zu vermeiden, wurde die Relevanz der Weitergabe von Lernprozessdokumentationen an die nachfolgende Bildungseinrichtung betont, auch für den Erfolg im Anfangsunterricht (Fabian & Dunlop, 2007; Athola et al., 2011; Heimlich & Ueffing, 2018). Dafür eignen sich prinzipiell Übergabedokumentationen, die über die individuelle Lernausgangslage des Kindes Auskunft geben (Liebers, 2022), um ein Anknüpfen im Anfangsunterricht zu ermöglichen. Sie müssen sich am Anspruch der Berücksichtigung der unterschiedlichen Bildungsverständnisse und den spezifischen Kulturen in Kita und Grundschule messen lassen (Liebers, Geiling & Prengel, 2020; Viernickel, 2014) und zugleich dem Kind als wichtigstem Akteur des Übergangs Gestaltungs- und Reflexionsräume für die eigenen Lernprozesse eröffnen. An der Universität Leipzig wurde von sechs Arbeitsbereichen die Übergabedokumentation „Mika“ entwickelt (gefördert vom Ministerium für Bildung, Jugend und Sport des Landes Brandenburg). Es besteht aus einem Kinderportfolio, um u. a. die Partizipation des Kindes zu stärken, und aus einer Dokumentation zum individuellen Stand des

Kindes in den Bereichen Bio-psychosoziales Wohlbefinden, körperlich-motorische Entwicklung, Frühe Literalität/Sprache, Mathematik, Naturwissenschaft und Digitale Medien. In die inhaltliche Gestaltung flossen der aktuelle Forschungsstand und theoretische Kompetenzmodellierungen ein.

Fragestellung:

Inwieweit lassen sich mit Hilfe von Mika die individuellen domänenspezifischen Kompetenzprofile von Kindern identifizieren? Welche Subpopulationen ergeben sich im Vergleich der Gruppenprofile?

Methode:

Der Beitrag greift auf die Daten aus der Praxiserprobung von Mika im Land Brandenburg zurück, die derzeit im Land Brandenburg realisiert wird (bis August 2024, vorauss. Stichprobengröße $N = 100$). Hierbei werden die von den Kita-Fachkräften ausgefüllten Mika Teil 2 Dokumentationen (für Kinder, die im Sommer 2024 eingeschult werden), erfasst und mit Hilfe der latenten Profilanalyse (Nylud-Gibson et al., 2014; Pastor et al., 2007; Vermunt, 2014) ausgewertet. Im Fokus steht die Analyse hinsichtlich der Profilhöhe und -streuung im Zusammenhang mit den domänenspezifischen Kompetenzen der Kinder über die verschiedenen Bildungsbereiche hinweg. Darüber hinaus werden Unterschiede und Schnittmengen im Vergleich der Profile aus der Gesamtstichprobe der erfassten Kinder ermittelt.

Ergebnisse und ihre Bedeutung:

Im Vortrag wird „Mika“ zunächst konzeptionell sowie die Ergebnisse der Praxiserprobung vorgestellt (Skalengüte, Nebengütekriterien des Verfahrens) vorgestellt, bevor vertieft auf die Ergebnisse der latenten Profilanalyse eingegangen wird.

In diesem Zusammenhang wird der Beitrag von Mika zur Identifizierung individueller Profile als Ausgangspunkt einer individualisierten Bildungsbegleitung am Übergang diskutiert

S-08: Zur Rolle von Emotionen von Lehrer*innen und Lehramtsstudierenden in verschiedenen pädagogischen Kontexten

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 148

Chair der Sitzung: Irina Rosa Kumschick

Chair der Sitzung: Gerda Hagenauer

Chair der Sitzung: Tina Hascher

Zur Rolle von Emotionen von Lehrer*innen und Lehramtsstudierenden in verschiedenen pädagogischen Kontexten

Chair(s): **Irina Rosa Kumschick** (PH Luzern, Schweiz), **Gerda Hagenauer** (Paris Lodron Universität Salzburg)

Diskutant*in(nen): **Tina Hascher** (Universität Bern)

Die Lehrprofession zeichnet sich durch zahlreiche und intensive Interaktionen, etwa mit Lernenden, anderen Lehrpersonen oder mit der Schulleitung, aus (Sutton & Wheatley, 2003; Waber et al., 2021). Solche sozialen Interaktionen wiederum sind eine Quelle für Emotionen (Denzin, 1984) und beeinflussen das professionelle Handeln bedeutsam (Anttila et al., 2016; Frenzel et al., 2016; Hagenauer & Hascher, 2018; Kumschick et al., 2018). Dieses Symposium vereint drei empirische Studien, in denen die Emotionen, die durch verschiedene Interaktionen im Lehrberuf entstehen, untersucht werden.

Die erste Studie untersucht, ob und wie ein positiver Führungsstil der Schulleitung mit dem emotionalen Erleben der Lehrpersonen zusammenhängt (Pekrun et al., 2023). Basierend auf dem PERMA-Modell (Seligman & Flourish, 2011) wird erforscht, ob unterstützende und wertschätzende Führung positive Emotionen bei Lehrkräften fördert und negative reduziert. Erste Ergebnisse zeigen positive Effekte auf das Emotionserleben von Lehrpersonen und bieten eine Diskussionsgrundlage zu der Frage, wie Führung im schulischen Kontext gestaltet sein sollte, um positive Emotionen von Lehrpersonen zu unterstützen.

Die zweite Studie widmet sich den Emotionen im Teamteaching. Hierbei wird untersucht, inwiefern die Emotionen, die beim Unterrichten vom/von der Teamteachingpartner*in ausgelöst werden, mit der wahrgenommenen Unterrichtsqualität zusammenhängen. Als Rahmenmodell dient das appraisaltheoretische Modell der Lehrer*innenemotionen von Frenzel et al. (2016). Die Ergebnisse resultierend aus Tagebucheinträgen belegen den Zusammenhang zwischen Emotionen und Qualitätsindikatoren von Unterricht und geben Anlass zur Reflexion, wie ein emotionsgünstiger Teamteaching-Unterricht gestaltet werden könnte.

Die dritte Studie simuliert die Emotionen von Schüler*innen während Unterrichtsstörungen. Untersucht wird, wie Schüler*innen (in dieser Studie sind dies Lehramtsstudierende, die sich in die Rolle der Schüler*innen hineinversetzen) die Effektivität des Störungsumgangs der Lehrperson wahrnehmen und welche Emotionen sie als proximale Prädiktoren beim Lernen erleben (Pekrun, 2006). Differenziert wird untersucht, wie die Valenz dieser Emotionen zwischen Lernenden mit unterschiedlichen Verhaltensweisen variiert. Die Ergebnisse zeigen unterschiedliche Interaktionen je nach Störverhalten der Lernenden und der Effektivität des Störungsumgangs der Lehrperson.

Alle drei Studien bieten bedeutsame Einblicke in die Zusammenhänge zwischen emotionalem Erleben, dessen Bedingungen und professionellem Handeln im Lehrberuf. Diese Erkenntnisse unterstützen die Entwicklung gezielter Interventionen und Weiterbildungsmaßnahmen, die sowohl das positive emotionale Erleben der Lehrpersonen als auch das Lernumfeld für Schüler*innen verbessern. Damit legen die Studien eine fundierte Grundlage für zukünftige Forschung und praktische Anwendungen zur Schaffung einer positiven und unterstützenden Schulkultur, in der alle Beteiligten ihr volles Potenzial entfalten können.

Beiträge des Symposiums

Positive School Leadership und das emotionale Erleben von Lehrkräften

Thomas Goetz¹, Elouise Botes², Lea M. Resch¹, Sina Weiss¹, Anne C. Frenzel³, Markus Ebner¹

¹Universität Wien, Österreich, ²University of Luxembourg, Luxemburg, ³Ludwig-Maximilians-Universität München, Deutschland

Theoretischer Hintergrund

Das emotionale Erleben von Lehrkräften in der Schule ist als suboptimal zu bezeichnen – hierzu gibt es zahlreiche empirische Indikatoren (z.B. hohe Burnout-Raten, García-Carmona et al., 2019; geringes Wohlbefinden, Skaalvik & Skaalvik, 2017). Dies kann zu zahlreichen Problemen führen, wie beispielsweise zum Lehrkräftemangel mit entsprechend negativen Auswirkungen auf unsere Gesellschaft (OECD, 2021). Die Forschung zu den Emotionen von Lehrkräften wurde jedoch bisher vernachlässigt (Frenzel et al., 2021) – insbesondere was die distalen Antezedenzen emotionalen Erlebens anbelangt (z. B. das soziale Umfeld in der Schule; Pekrun, 2023). Unsere Studie sollte dazu beitragen, diese Forschungslücke zu schließen. Sie basiert auf den theoretischen Annahmen von Pekruns (2006, 2023) Kontroll-Wert-Theorie der Leistungsemotionen: Kontroll- und Wertappraisals der Lehrkräfte (proximale Emotionsantezedenzen) sollten die Wirkungen des sozialen Umfelds (distale Emotionsantezedenzen) auf ihre Emotionen medieren. Wir haben uns bezüglich der distalen Emotionsantezedenzen auf den Führungsstil der Schulleitung konzentriert. Hierbei beziehen wir uns auf das PERMA-Lead-Modell (Ebner, 2019, 2020). Das Akronym steht für fünf Faktoren von Führung, die dazu führen sollten, dass Menschen im Berufsleben ihre Potenziale entfalten können: (1) „Positive Emotions“ erlebbar machen, (2) „Engagement“ fördern, (3) „Relationships“ (also Beziehungsqualität) fördern, (4) „Meaning“, d.h. die Sinnhaftigkeit der Tätigkeiten hervorheben und (5) „Accomplishment“, d.h. Zielerreichungen explizit sichtbar machen. Entsprechend der Kontroll-Wert-Theorie sollten die PERMA-Lead-Facetten zu höheren Kontroll- und Werteinschätzungen bei den Lehrkräften und folglich zu höheren positiven und niedrigeren negativen Emotionen führen.

Fragestellung

Auf der Basis der Kontroll-Wert-Theorie zu Lern- und Leistungsemotionen (Pekrun, 2006, 2023) leiteten wir die folgende Hypothese ab: (H1) Lehrkräfte profitieren emotional von positiver Führung in der Schule. (H2) Der Effekt von positiver Führung auf die Emotionen der Lehrkräfte wird durch deren Kontroll- und Wertappraisals in Bezug auf ihre Tätigkeiten an der Schule mediert.

Methode

Wir haben Schulleitungen in Österreich, Deutschland und der Schweiz gebeten, ihre Lehrkräfte an unserer Online-Studie (SoSci Survey) teilnehmen zu lassen. Insgesamt nahmen N = 446 Lehrkräfte (80% weiblich, 19% männlich, 1% non-binär) aus Grund-, Sekundar- und beruflichen Schulen teil. Positive Leadership wurde anhand des PERMA-Lead-Profilers (Ebner et al., 2019; 15

Items) erhoben, welcher alle fünf PERMA-Facetten aus der Perspektive der Lehrkräfte erfasst. Kontrollappraisals bezüglich der Tätigkeiten an der Schule wurden anhand der Teacher Self-Efficacy Scale (Pfitzner-Eden, 2016; 12 Items) erhoben, Wertappraisals anhand des Meaning in Work Inventory (ME-Work; Schnell & Hoffmann, 2020; 12 Items). Zur Erfassung der Emotionen Freude, Angst und Ärger wurden die Teacher Emotions Scales (TES, Frenzel et al., 2016; 4 Items pro Emotion) verwendet. Zudem wurden Skalen zu Stolz und Langeweile entwickelt (jeweils 4 Items). Die Reliabilitäten aller Skalen waren zufriedenstellend (McDonald's [1999] $\omega \geq 0.80$). Zur Hypothesenprüfung wurde ein latentes Mediationsmodell anhand des lavaan packages (Rosseel, 2012) in RStudio gerechnet. Hierbei wurde jeweils ein latentes Konstrukt für PERMA-Lead, Kontrolle, Wert und positive Emotionen der Lehrkräfte (negative Emotionen invertiert) gebildet.

Ergebnisse und ihre Bedeutung

Das latente Mediationsmodell wies einen sehr guten Fit auf. Die beiden indirekten Effekte (PERMA-Lead $\rightarrow$ Kontrolle $\rightarrow$ positive Emotionen; PERMA-Lead $\rightarrow$ Wert $\rightarrow$ positive Emotionen) waren statistisch signifikant und jeweils positiv. Der verbleibende direkte Effekt (PERMA-Lead $\rightarrow$ positive Emotionen) war hingegen nicht mehr signifikant, was für eine vollständige Mediation spricht. Insgesamt bestätigen unsere Ergebnisse somit die angenommenen Zusammenhänge: PERMA-Lead zeigte positive Effekte auf das positive Emotionserleben von Lehrkräften (H1), wobei dieser Effekt vollständig durch die Kontroll- und Wertappraisals mediiert wurde (H2). Obwohl unsere Studie querschnittlich ist, liefert sie erste Hinweise darauf, dass die Emotionen der Lehrkräfte durch den Führungsstil an Schulen positiv beeinflusst werden können. Dies könnte eine sehr effektive und kostengünstige Maßnahme darstellen, die durch entsprechende Aus- und Fortbildungsprogramme für Schulleitungen initiiert werden könnte – mit potenziell großen Auswirkungen, da teilweise mehr als 100 Lehrkräfte an einer Schule tätig sind.

Emotionen von Teamteaching-Lehrpersonen und die wahrgenommene Unterrichtsqualität – Ergebnisse einer Tagebuchstudie

Franziska Muehlbacher¹, Gerda Hagenauer¹, Christoph Helm²

¹Paris Lodron Universität Salzburg, Österreich, ²Johannes-Kepler-Universität Linz, Österreich

Theoretischer Hintergrund und Fragestellung

Emotionen von Lehrpersonen beim Unterrichten stehen in einem bedeutsamen Zusammenhang mit der Unterrichtsqualität (Frenzel, 2014; Frenzel et al., 2021). Dieser Zusammenhang zeigte sich in zahlreichen empirischen Studien bezogen auf Lehrpersonen, die allein in Klassen unterrichten (z.B. Becker et al., 2014; Kalchgruber et al., 2021). Allerdings unterrichten Lehrpersonen regelmäßig auch kooperativ: eines dieser kooperativen Settings stellt das Teamteaching dar. Im Teamteaching-Unterricht sind zwei Fachlehrpersonen für die Planung, Durchführung und Evaluation des Unterrichts in einer Klasse verantwortlich (Krammer et al., 2017). Muehlbacher und Hagenauer (2023) konnten aufzeigen, dass auch die Interaktion mit dem Teampartner/mit der Teampartnerin zahlreiche positive und auch negative Emotionen bei Lehrpersonen hervorruft. Noch ist aber unklar, ob diese Emotionen ebenso wie beim alleinigen Unterrichten bedeutsam mit der Unterrichtsqualität in Verbindung stehen. Einem appraisal-theoretischen Ansatz folgend (Scherer & Moors, 2019), wie er auch im Modell der Lehrer*innenemotionen nach Frenzel (2014) postuliert wird, gehen wir daher der Frage nach, 1) inwiefern die positiven und negativen Emotionen, die Lehrpersonen in einer konkreten Teamteaching-Unterrichtsstunde erleben (gemessen über positiven und negativen Affekt), mit der wahrgenommenen Unterrichtsqualität in Verbindung stehen (auf der intraindividuellen und interindividuellen Ebene). 2) Zudem überprüfen wir spillover Effekte: Inwieweit kann das emotionale Erleben von Lehrpersonen in einer Teamteaching-Einheit die wahrgenommene Unterrichtsqualität der folgenden Unterrichtseinheit erklären und vice versa?

Methode

Um diese beiden Fragestellungen zu prüfen, wurde eine längsschnittliche Tagebuchstudie durchgeführt, an der 47 österreichische Teamteaching-Lehrpersonen partizipierten (MAlter = 40.45; 34 weiblich). Die Teilnehmer*innen wurden gebeten, 15 Tagebucheinträge jeweils nach einer Teamteachingstunde vorzunehmen (N=652 Tagebucheinträge; Responserate = 92%). Die positiven und negativen Emotionen wurden mit Hilfe der deutschsprachigen PANAS Scale erfasst (Positive and Negative Affect Schedule; Breyer & Bluemke, 2016). Aufbauend auf den drei Basisdimensionen der Unterrichtsqualität wurde diese anhand von vier Dimensionen mit jeweils einem Item erfasst: effektives Zeitmanagement (Kunter et al., 2017), Klarheit der Instruktion (Frey et al., 2009), Differenzierung (gemessen als individualisierte Unterstützungsleistung; Mang et al., 2023) und Lehrer*innen-Schüler*innen-Beziehung (Aldrup et al., 2018).

Als statistisches Analyseverfahren kam die dynamische Strukturgleichungsmodellierung (DSEM; Asparouhov et al., 2018; Hamaker et al., 2018) in Mplus Version 8 (Muthén & Muthén, 1998-2017) zum Einsatz. Bei diesem Ansatz kann eine Mehrebenenmodellierung mit einer Zeitreihenmodellierung kombiniert werden, um die Lead-Lag-Beziehung zwischen wiederholt beurteilten Variablen zu untersuchen.

Ergebnisse und ihre Bedeutung

Bezogen auf Forschungsfrage 1 zeigen die Ergebnisse, dass die durchschnittlichen Korrelationen zwischen den untersuchten Variablen auf der intraindividuellen (within) Ebene moderat und statistisch signifikant waren und von $\beta = -0.282$ [-0.364, -0.194] bis $\beta = -0.099$ [-0.183, -0.004] und $\beta = 0.112$ [0.030, 0.191] bis $\beta = 0.249$ [0.163, 0.327] in der erwarteten Richtung reichten. Auf der interindividuellen Ebene (between) wurden die folgenden Beziehungen beobachtet: Erstens waren Klarheit der Instruktion ($\beta = 0.248$ [0.069, 0.674]) und Differenzierung ($\beta = 0.512$ [0.184, 0.738]) positiv mit positiven Emotionen verbunden. Zweitens war die Differenzierung negativ mit negativen Emotionen verbunden ($\beta = -0.411$ [-0.687, -0.035]). Spillover-Effekte (Forschungsfrage 2) wurden nur in zwei Fällen bestätigt: Erstens stand die die Lehrer-Schüler-Beziehung der vorigen Teamteaching-Unterrichtsstunde positiv mit den positiven? Emotionen in der nächsten Teamteaching-Unterrichtsstunde in Verbindung ($\beta = 0.101$ [0.003, 0.200]). Zweitens war das Zeitmanagement von der vorigen Teamteaching-Unterrichtsstunde negativ mit den negativen Emotionen in der nächsten Teamteaching-Unterrichtsstunde assoziiert ($\beta = -0.151$ [-0.271, -0.026]). Die Effektgröße war jedoch in beiden Fällen schwach.

Die Ergebnisse der vorliegenden Studie untermauern, dass die von einem/einer Teamteaching- Partner*in ausgelösten Emotionen mit der Qualität des Unterrichts in der jeweiligen und der nachfolgenden Unterrichtsstunde in Verbindung stehen. Dadurch werden die bisherigen Forschungsarbeiten zu Lehrer*innenemotionen um einen bedeutsamen Kontext erweitert. In Bezug auf praktische Implikationen sprechen die Ergebnisse dafür, dass in Trainings zur Förderung sozial- emotionaler Kompetenzen auch der Umgang mit dem Teamteaching-Partner/der Teamteaching- Partnerin fokussiert werden sollte.

Die Effektivität des Störungsumgangs einer Lehrperson interagiert mit den Emotionen im Klassenraum

Irina Rosa Kumschick¹, Cécile Tschopp¹, Larissa Troesch², Annette Tettenborn¹

¹PH Luzern, Schweiz, ²PH Bern, Schweiz

Theoretischer Hintergrund und Fragestellung

Die Kontroll-Wert-Theorie von Pekrun (2006) beschreibt leistungsbezogene Emotionen als proximale Prädiktoren beim Lehren und Lernen. Obwohl die Annahmen dieser Theorie im schulischen Kontext häufig überprüft wurden, gibt es kaum Studien, die schülerbezogene Emotionen bei Unterrichtsstörungen untersuchen. Unterrichtsstörungen werden als Ereignisse definiert, die Lernprozesse beeinträchtigen oder unterbrechen (Doyle, 2013). Eine wenig untersuchte Frage ist, inwiefern die Emotionen der Schüler*innen mit dem Störungsmanagement der Lehrperson zusammenhängen. Deshalb untersucht diese Studie, wie die Effektivität des Störungsmanagements der Lehrperson mit den Emotionen von engagierten oder störenden Schüler*innen zusammenhängt. Bezogen auf das Störverhalten im Klassenraum orientiert sich die Studie an der Systematisierung von Marzano und Marzano (2003), in welcher nicht nur Lernende, die aktiv stören, sondern auch solche, die sich passiv, aggressiv und sich sozial unangemessen verhalten, als störend betrachtet werden. Diese werden von Marzano und Marzano (2003) als 'high-needs students' klassifiziert und in der vorliegenden Studie abgegrenzt von engagierten Schüler*innen, die interessiert sind und der von den Lehrpersonen intendierten Unterrichtsaktivität folgen (Kumschick et al., under Review).

Methode

In einer experimentellen In-situ-Simulation reagierten Lehramtsstudierende während einer selbst geplanten, zehnminütigen Unterrichtseinheit auf Störungen, die von anderen Lehramtsstudierenden als engagierte oder störende Schüler*innen simuliert wurden. Als Selbsteinschätzungen der Schüler*innen wurden folgende Skalen eingesetzt: Positive (McDonald's Omega = .87) und negative Emotionen (McDonald's Omega = .82) wurden mit der Kurzform des PANAS (Thompson, 2007) erhoben. Zudem wurde die Effektivität des Störungsumgangs anhand von vier Subskalen gemessen (Störung gestoppt: Omega = .78, Störung abgeschwächt: Omega = .65, Störung unvermindert: Omega = .60, Störung vergrößert: Omega = .66). Insgesamt wurden 168 Fragebögen von engagierten (n = 81) und störenden Schüler*innen (n = 87) aus 17 Experimentdurchführungen im Frühling 2023 und 2024 analysiert. Eine Mehrebenenanalyse mit einem Random-Intercept-Modell prüfte, inwieweit wahrgenommene positive bzw. negative Emotionen von störendem oder engagiertem Verhalten mit der Effektivität des Störungsumgangs interagierten.

Ergebnisse

Die Analysen ergaben einen signifikanten Haupteffekt der zugewiesenen Rolle auf die positiven Emotionen, wobei störende Individuen ein höheres Maß an positiven Emotionen erlebten als Personen in engagierten Rollen. Zudem wurde die Interaktion zwischen der Effektivität des Störungsumgangs und der Bedingung „Schüler*innenverhalten (störend vs. engagiert)“ signifikant. Die Ergebnisse weisen darauf hin, dass engagiertes Verhalten im Klassenraum stärker mit positiven Emotionen korreliert, wenn der Störungsumgang der Lehrperson als effektiv wahrgenommen wird. Andererseits berichteten störende Schüler*innen, dass sie mehr positive Emotionen empfanden (als engagierte Schüler*innen), wenn sie das Störungsmanagement der Lehrperson als ineffektiv empfanden (d.h. die Störung wurde nicht beeinflusst oder sogar vergrößert durch die Lehrperson). Gleichzeitig zeigte sich, dass engagierte Schüler*innen unter dieser Bedingung (ineffektiver Störungsumgang) mehr negative Emotionen wahrnahmen als störende Schüler*innen. Ferner zeigten sich bei störenden Individuen mehr negative Emotionen, wenn die Lehrperson in der Lage war, die Störung zu beenden.

Bedeutung und Relevanz

Die vorliegende Studie untermauert die Kontroll-Wert-Theorie (Pekrun, 2006). Die individuellen Kontroll- und Werte-Appraisals sind für die negativen und positiven Emotionen während Unterrichtsstörungen bedeutsam und hängen mit dem gezeigten Verhalten der Lernenden zusammen (störend versus engagiert). Problematisch ist der Befund, dass ein misslungener Störungsumgang positive Emotionen bei störenden Schüler*innen hervorruft. Landrum und Kauffman (2013) betonen, dass das Verhalten der Schüler*innen durch positive und negative Verstärkung beeinflusst wird. Lehrpersonen müssen daher vermeiden, störendes Verhalten durch Verstärkung positiver Emotionen zu fördern. Wissenschaftliche Untersuchungen zeigen jedoch, dass Berufsanfänger*innen ihre Aufmerksamkeit nahezu ausschliesslich auf das störende Verhalten richten und es dadurch verstärken (Glock & Kleen, 2023; Wolff et al., 2021). In-situ-Simulationen können Lehramtsstudierenden helfen, schnell und effektiv auf Störungen zu reagieren, ohne selbst zur Störquelle zu werden.

Schlüsselwörter: Störungsmanagement, engagierte Schüler*innen, störende Schüler*innen, Emotionen im Klassenraum, In-situ-Simulation

S-09: Aktuelle Mindset-Interventionen im deutschsprachigen Raum: Was können sie für wen bewirken?

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: O 151

Chair der Sitzung: Katharina Asbury

Chair der Sitzung: Anke Heyder

Chair der Sitzung: Caroline Villiger

Aktuelle Mindset-Interventionen im deutschsprachigen Raum: Was können sie für wen bewirken?

Chair(s): **Katharina Asbury** (IPN Kiel, Deutschland), **Anke Heyder** (Ruhr-Universität Bochum)

Diskutant*in(nen): **Caroline Villiger** (PH Bern)

Growth Mindsets (die Überzeugung, dass kognitive Fähigkeiten durch Anstrengung wachsen können) sind positiv assoziiert mit der Motivation und Leistung Lernender (z.B. Dweck, 2006). US-Amerikanische Studien legen nahe, dass sie sich durch Interventionen fördern lassen (z.B. Burnette et al., 2023). Dieses Symposium fokussiert vier aktuelle Interventionen zur Förderung von Growth Mindsets aus dem deutschsprachigen Raum und untersucht aus psychologischer, erziehungswissenschaftlicher und fachdidaktischer Perspektive, welche Veränderungen diese Interventionen bewirken oder nicht. Durch eine Kombination aus verschiedenen quantitativen und qualitativen Studiendesigns und Maßen soll beleuchtet werden, wie diese Interventionen Growth Mindsets im sekundären und tertiären Bildungsbereich fördern können. Dabei werden verschiedene Zielgruppen, z.B. Lehramtsstudierende und leistungsschwache Schüler*innen, wie auch Kontexte, z.B. MINT-Fächer (Beiträge 2, 3 und 4) und Nicht-MINT-Fächer (Beitrag 1), berücksichtigt. Damit berührt das Symposium auch aktuelle gesellschaftliche Herausforderungen wie den Mangel an MINT-Fachkräften und die wachsende Zahl an Schüler*innen mit unzureichenden Kompetenzen.

Der erste Beitrag untersucht die Effekte einer Growth-Mindset-Intervention bei Psychologiestudierenden und setzt dabei verschiedene Messinstrumente ein, darunter Fragebögen, fallbasierte Aufgaben und einen reaktionszeitbasierten Impliziten Assoziationstest (IAT). Die Ergebnisse zeigen, dass die Intervention das selbstberichtete Growth Mindset signifikant erhöhte, insbesondere bei Studierenden mit einem ausgeprägten Fixed Mindset zu Beginn. Dies deutet darauf hin, dass besonders Lernende mit anfänglich ungünstigen Überzeugungen von der Intervention profitieren.

Der zweite Beitrag untersucht die Rolle von Growth Mindsets in der MINT-Studieneingangsphase. Eine Online-Intervention zielte darauf ab, das Growth Mindset der MINT-Studierenden zu fördern, um Studienabbrüche zu reduzieren. Die Intervention konnte positive Veränderungen im Growth Mindset bewirken und zeigt das Potenzial, die Abbruchquoten zu senken. Diese Ergebnisse unterstreichen die Bedeutung von Mindset-Interventionen im MINT-Hochschulbereich.

Der dritte Beitrag untersucht die Effekte einer reflexionsbasierten Mindset-Kurzintervention bei Mathematik-Lehramtsstudierenden. Die Ergebnisse zeigen, dass die Interventionsgruppe im Vergleich zur Kontrollgruppe ein signifikant niedrigeres Fixed Mindset sowie weitere günstigere motivationale Merkmale aufwies. Zusätzliche qualitative Analysen der Feedbacktexte bereichern die quantitativen Ergebnisse. Diese Erkenntnisse sind besonders relevant für die Ausbildung von Lehrkräften, da sie aufzeigen, wie bereits im Studium die Überzeugungen zukünftiger Mathematiklehrkräfte beeinflusst werden können.

Der vierte Beitrag untersucht die Auswirkungen einer Mindset-Intervention auf das Growth Mindset von angehenden Lehrkräften sowie ihrer leistungsschwachen Schüler*innen. Die Intervention zeigte, dass Lehramtsstudierende eine positive Veränderung ihres Mindsets erlebten. Zudem wiesen die leistungsschwachen Schüler*innen, die von den angehenden Lehrkräften der Interventionsgruppe unterrichtet wurden, ebenfalls ein erhöhtes Growth Mindset auf. Diese Ergebnisse zeigen auf, wie auch leistungsschwache Schüler*innen von Mindset-Interventionen für Lehrende profitieren können.

Die Beiträge werden abschließend aus erziehungswissenschaftlicher Sicht diskutiert.

Beiträge des Symposiums

Wie Lernvoraussetzungen die Veränderung des Growth Mindsets beeinflussen: Eine Analyse expliziter und impliziter Maße

Helene Zeeb, Inga Glogger-Frey
Universität Erfurt

Growth-Mindset-Interventionen zielen darauf ab, Überzeugungen über die Veränderbarkeit kognitiver Fähigkeiten (Growth Mindset) zu stärken (Dweck & Yeager, 2019). Die Forschungslage ist jedoch uneinheitlich (z. B. Sisk et al., 2018): Können Interventionen solche tief verwurzelten Überzeugungen tatsächlich verändern? Und welche Messinstrumente sind zur Erfassung der Interventionseffekte geeignet? In der Regel werden Fragebögen verwendet; diese messen aber nur selbstberichtete, dem Bewusstsein zugängliche Überzeugungen und sind anfällig für Verzerrungen (Lüftenegger & Chen, 2017). In unserer Studie nutzten wir daher auch fallbasierte Aufgaben und einen reaktionszeitbasierten Impliziten Assoziationstest (IAT). Eine weitere offene Frage ist, ob Growth-Mindset-Interventionen bei allen Lernenden gleich gut wirken. Einige Studien deuten darauf hin, dass Lernende mit ungünstigen Voraussetzungen – zum Beispiel einem ausgeprägten Fixed Mindset hinsichtlich ihrer eigenen Fähigkeiten – besonders von Growth-Mindset-Interventionen profitieren (Burnette et al., 2023).

Wir adressierten folgende Forschungsfragen:

1. Wie wirkt sich eine Growth-Mindset-Intervention auf verschiedene Mindset-Maße aus? Wir erwarteten, dass die Intervention (a) das selbstberichtete Growth Mindset, (b) das fallbasierte Growth Mindset und (c) das reaktionszeitbasierte Growth Mindset erhöht.
2. Werden die Interventionseffekte durch vorab bestehende Überzeugungen über die eigenen Fähigkeiten moderiert? Wir erwarteten, dass Lernende mit einem Fixed Mindset hinsichtlich ihrer eigenen Fähigkeiten besonders von der Intervention profitieren.

An der experimentellen, präregistrierten Laborstudie nahmen 73 Psychologiestudierende teil (Alter: $M = 21.2$ Jahre, $SD = 1.7$; 88% weiblich), die entweder die Growth-Mindset-Intervention ($n = 35$) oder eine Kontrollintervention bearbeiteten ($n = 38$). Beide Interventionen waren computerbasiert, hatten die gleiche Struktur und Länge (ca. 45 min) und dieselbe Anzahl an Reflexionsaufgaben. Die Growth-Mindset-Intervention enthielt neurowissenschaftliche Informationen über das lernende Gehirn und vermittelte die Kernbotschaft, dass kognitive Fähigkeiten veränderbar sind. Die Kontrollintervention enthielt ebenfalls neurowissenschaftliche Informationen über das Gehirn, aber ohne Bezug zu Veränderungs- oder Lernprozessen. Wir erfassten

(a) das selbstberichtete Growth Mindset mittels Fragebögen (Dweck, 2000), (b) das fallbasierte Growth Mindset mit zwei Fallvignetten, zu denen die Teilnehmenden ihre Gedanken, Gefühle und Handlungsideen beschreiben sollten, und (c) das reaktionszeitbasierte Growth Mindset mit einem IAT (Mascret et al., 2015). Nach einer Woche folgte ein Follow-up mit allen drei Messinstrumenten.

Kontrastanalysen zeigten, dass das selbstberichtete Growth Mindset durch die Intervention deutlich verstärkt wurde, sowohl im unmittelbaren Posttest ($\eta^2 = 0.30$, $p < .001$) als auch im Follow-up ($\eta^2 = 0.20$, $p < .001$). Die Auswertung der fallbasierten Aufgaben steht noch aus. Hinsichtlich des reaktionszeitbasierten Growth Mindsets zeigten sich keine Interventionseffekte (η^2 s < 0.01 , p s $> .63$). Die Dreifach-Interaktion zwischen Zeit, Gruppe und Vor-Überzeugungen war jedoch für das selbstberichtete und das reaktionszeitbasierte Growth Mindset signifikant ($\eta^2 = 0.11$, $p = .020$ und $\eta^2 = 0.19$, $p = .011$). Je stärker das Fixed Mindset hinsichtlich der eigenen Fähigkeiten zu Beginn war, desto positiver entwickelte sich das Growth Mindset durch die Intervention.

Insgesamt wirkte die Intervention also vor allem auf die „klassische“ Mindset-Messung, den Fragebogen. Beim IAT konnten auf Gruppenebene keine Interventionseffekte festgestellt werden. Interessanterweise zeigten sich aber bei beiden Messinstrumenten Moderationseffekte im Sinne einer Aptitude-Treatment-Interaktion (Snow, 1991): Bei Studierenden, die zu Beginn überzeugt waren, ihre eigenen Fähigkeiten nicht verändern zu können, zeigten sich die stärksten Growth-Mindset-Veränderungen. Diese Befunde passen zu Studien, in denen Lernende mit ungünstigen Voraussetzungen besonders von der Intervention profitierten (Burnette et al., 2023). Zugleich deuten die Ergebnisse darauf hin, dass unsere Intervention auch implizite Überzeugungen veränderte – zwar nicht bei allen, wohl aber bei den Lernenden, die mit ungünstigeren Überzeugungen gestartet waren. Um weitere Einblicke zu gewinnen, wie Growth-Mindset-Interventionen auf indirekte Maße wirken, sollen zunächst die fallbasierten Aufgaben ausgewertet werden. Zudem scheint es sinnvoll, in Folgestudien längere Zeiträume zu betrachten, da es Zeit braucht, bis eine „Aufwärtsspirale“ aus positiveren Überzeugungen, höherer Motivation und größeren Lernerfolgen in Gang kommt (Yeager & Walton, 2011), in deren Verlauf sich Überzeugungen grundlegend verändern können.

Die Rolle impliziter Fähigkeitstheorien in der MINT-Studieneingangsphase: Deskriptive Analysen und Beeinflussung durch eine Online-Intervention

Malte Diederich, Verena Spatz
TU Darmstadt

Gut ausgebildete Fachkräfte im MINT-Bereich sind notwendig, um die Herausforderungen der Digitalisierung, Dekarbonisierung und des demografischen Wandels zu bewältigen (Anger et al., 2024). Gerade im MINT-Bereich sind allerdings die Abbruchquoten an den Universitäten besonders hoch (Heublein et al., 2022). Dabei spielen hier insbesondere Leistungsprobleme bei über 84 % der Studierenden eine zentrale Rolle, stärker als in anderen Fächergruppen (Heublein et al., 2017). Studienabbrecher:innen begründen diese Probleme mit zu hohen Studienanforderungen, Schwierigkeiten beim Einstieg sowie Zweifeln an ihrer persönlichen Eignung (ebd.).

Universitäten versuchen dieses Problem z.B. durch Beratungsangebote oder fachliche Vorkurse zu adressieren (Klöpping et al., 2017; Kürten, 2019). Da Studienabbrecher:innen diese Angebote allerdings im Mittel als weniger hilfreich (Heublein et al., 2017) bewerten, rücken zunehmend auch motivationale Faktoren in den Blick, wie die impliziten Fähigkeitstheorien nach Dweck und Leggett (1988). Hierbei handelt es sich um Grundüberzeugungen zur Veränderbarkeit von persönlichen Eigenschaften wie Intelligenz oder den grundlegenden Fähigkeiten in Fächern wie Physik oder Mathematik (Yeager & Dweck, 2012). Eine Unveränderbarkeitstheorie (auch „Fixed Mindset“) betrachtet diese Eigenschaften als statisch, was mit lernhinderlichen Auswirkungen auf die Motivation verbunden ist. Im Gegensatz dazu geht eine Wachstumstheorie (auch „Growth Mindset“) davon aus, dass sich diese durch Anstrengung, gute Strategien und Unterstützung durch Mentor:innen positiv verändern lassen (ebd.).

Implizite Fähigkeitstheorien haben international unter dem Begriff „Growth Mindset“ viel Aufmerksamkeit erlangt. Einige Autor:innen warben hier populärwissenschaftlich mit großen Versprechungen (Boaler, 2013), denen zum Teil mit fundamentaler Kritik an der Theorie begegnet wurde (Burgoyne et al., 2020). Dabei belegt eine umfangreiche Studienlage die lernförderliche Auswirkung einer Wachstumstheorie auf die Motivation, vor allem in Situationen, welche den eigenen Selbstwert bedrohen (Burnette et al., 2013; Yeager & Dweck, 2020). Darüber hinaus liefern Studien im Schulbereich (Rege et al., 2020; Yeager et al., 2019) und auch im Hochschulbereich (Yeager, Walton et al., 2016) überzeugende Belege dafür, dass selbst eine kurze Online-Intervention von unter einer Stunde unter bestimmten Bedingungen eine langfristige, lernförderliche Wirkung erzielen kann.

In der von uns vorgestellten Studie wird daher der Frage nachgegangen, welche Rolle implizite Fähigkeitstheorien in der MINT-Studieneingangsphase spielen können. Die Studie liefert hierbei Beiträge auf zwei Ebenen: Zum einen werden durch deskriptive Analysen zentrale Ergebnisse aus dem internationalen Bereich zum „Mindset Meaning System“ repliziert und das Zusammenspiel von impliziten Fähigkeitstheorien mit dem fachspezifischen akademischen Selbstkonzept sowie Studienwechsel- und Studienabbruchsintentionen beschrieben. Zum anderen wird untersucht, ob eine gezielte Beeinflussung der impliziten Fähigkeitstheorien durch eine Online-Intervention in der Studieneingangsphase zur Reduktion von Studienabbrüchen im MINT-Bereich beitragen kann. Hierzu wurde der Prototyp einer Online-Intervention aus der Physik (Rehberg et al., 2023) auf den MINT-Bereich ausgeweitet und ihr Einfluss auf die oben genannten Faktoren untersucht. Die Intervention verwebt dabei bewährte Elemente aus US-amerikanischen Mindset-Interventionen (Yeager, Romero et al., 2016) mit der Vermittlung effektiver Lernstrategien (Dunlosky et al., 2013) auf synergetische Weise.

Hierfür wurden in den Wintersemestern 21-22 und 22-23 in allen MINT-Grundlagenveranstaltungen an der TU Darmstadt Fragebogenerhebungen zu zwei Messzeitpunkten durchgeführt: T1 zu Beginn des Semesters und T2 zwei Monate später ($N_{\text{Max}} = 1505$, $N_{\text{Min}} = 335$). Im zweiten Jahr hatten die Studierenden der neuen Kohorte die Möglichkeit, freiwillig an der Intervention teilzunehmen (Interventionsgruppe $N = 90$). Die Ergebnisse dieser Interventionsgruppe können mit den Ergebnissen der Baseline (ohne Kontakt zur Intervention) verglichen werden.

Die erhobenen Daten sprechen dafür, dass eine Tendenz zur Unveränderbarkeitstheorie in der MINT-Studieneingangsphase mit Studienabbruchsintentionen zusammenhängt (kleine bis mittlere Effektstärke), besonders in einer Subgruppe von Studierenden, die (nach Selbstauskunft) bleibende inhaltliche Schwierigkeiten im Studium haben (mittlere bis große Effektstärke). Zusätzlich zeigen die Analysen zur Online-Intervention, dass implizite Fähigkeitstheorien in Richtung einer Wachstumstheorie positiv beeinflusst werden können. Die Intervention ist inzwischen als OER frei verfügbar unter <https://lernbar.uni-frankfurt.de/MaDi/BesserStud> und wurde mit dem OER-Preis 2023 des HessenHub ausgezeichnet.

Effekte Mathematikspezifischer Mindsets von Lehramtsstudierenden im Kontext von schulischem Feedback

Theoretischer Hintergrund

Die subjektive Überzeugung, inwiefern Personenmerkmale wie Fähigkeiten statisch oder dynamisch sind, wird als Fixed oder Growth Mindset bezeichnet (z.B. Dweck & Yeager, 2019). Bisherige Forschung zeigt, dass ein Mindset bereichsspezifisch sein kann und insbesondere Mathematik als ein Fach betrachtet wird, welches mehr angeborener, unveränderbarer Begabung bedarf (Heyder et al., 2020; Leslie et al., 2015). Für den Hochschulkontext konnte gezeigt werden, dass ein Fixed Mindset zu stärkeren Leistungsattributionen auf stabile Faktoren sowie der Anwendung tröstender Strategien führt, die das weitere Engagement im Mathematik verringern (Rattan et al., 2012). Bislang fehlt es jedoch an experimentellen Arbeiten, die dieses für den schulischen Bildungsbereich untersuchen. Dieses ist wichtig, um Aussagen über die Bedeutsamkeit mathematikspezifischer Mindsets im Schulkontext treffen und mögliche Schlüsse für die Lehramtsausbildung herleiten zu können.

Fragestellung

Die vorliegende experimentelle Studie untersucht den Effekt einer Growth Mindset fördernden Kurzintervention (Heyder et al., 2023) auf das Verhalten und antizipierte Handeln von Mathematik-Lehramtsstudierenden. Dabei wird erstens untersucht, ob die Intervention replizierbare Effekte für das Mathematikspezifische Mindset zeigt. Zweitens wird analysiert, ob sich die Intervention auf die Einschätzung der Leistung eines fiktiven Schülers in einem Mathematiktest, die antizipierte Leistungsverbesserung und die antizipierten Lehrerreaktionen gegenüber dem fiktiven Schüler auswirkt. Wir erwarteten, dass die Mathematik-Lehramtsstudierenden der Interventionsgruppe durch ihre Reflexion ein niedrigeres Fixed Mindset aufweisen. In Bezug auf den korrigierten Mathematiktest des fiktiven leistungsschwachen Schülers erwarteten wir zudem, dass die Interventionsgruppe das Testergebnis weniger auf mangelnde Begabung und mehr auf mangelnde Anstrengung zurückführt, eine höhere Wahrscheinlichkeit der Leistungsverbesserung, eine geringere Wahrscheinlichkeit von tröstenden und längerfristig nicht hilfreichen Reaktionen berichtet sowie mehr Growth-Mindset-orientiertes Feedback gibt als die Kontrollgruppe.

2

Methode

In einem Online-Experiment wurden 105 Lehramtsstudierende (80 weiblich; Alter $M = 23.83$ Jahre) mit Mathematik als Unterrichtsfach einer weiterführenden Schulform randomisiert der Interventions- ($n = 54$) oder Kontrollgruppe ($n = 51$) zugewiesen. Die Interventionsgruppe reflektierte über ihre Mission als zukünftige Mathematik-Lehrkraft, die Kontrollgruppe über das Ruhrgebiet (vgl. Heyder et al., 2023). Anschließend berichteten sie ihr Mathematikspezifisches Mindset (Manipulationscheck; Heyder et al., 2020). Dann lasen sie einen korrigierten Mathematiktest eines fiktiven Schülers und verfassten ein Feedback an den Schüler, welches dieser zusätzlich zur Note erhalten solle. Daraufhin schätzten sie mit einem Item ein, inwiefern das Testergebnis auf eine mangelnde mathematische Begabung des Schülers zurückzuführen ist, und bewerteten mit einem weiteren Item, zu wie viel Prozent ihrer Meinung nach die Testleistung auf mangelnde Anstrengung und auf mangelnde mathematische Begabung zurückzuführen ist (Rattan et al., 2012). Zudem machten sie Angaben zur zukünftigen Leistungsverbesserung und schätzten ein, wie wahrscheinlich sie verschiedene Reaktionen gegenüber dem Schüler (z.B. trösten; Rattan et al., 2012) zeigen würden.

Ergebnisse

Die Interventionsgruppe berichtete wie erwartet ein signifikant niedrigeres Fixed Mindset als die Kontrollgruppe ($d = -0.34$, $p = .044$), was die Effektivität der Intervention stützt. Zudem führte die Interventionsgruppe erwartungskonform das Testergebnis signifikant weniger auf mangelnde Begabung zurück ($d = -0.33$, $p = .047$) und erklärte die Leistung deskriptiv mehr durch mangelnde Anstrengung statt mangelnder Begabung ($d = 0.20$, $p = .15$). Entgegen den Annahmen unterschieden sich die beiden Gruppen nicht in der erwarteten Leistungsverbesserung und der Wahrscheinlichkeit der Verwendung tröstender und nicht hilfreicher Reaktionen. Zusätzlich zu den quantitativen Analysen werden die Feedbacktexte derzeit inhaltlich hinsichtlich ihrer Mindset-Orientierung ausgewertet. Orientiert wird sich dabei an der inhaltlichen Strukturierung nach Mayring (2022). Dazu werden die Feedbacktexte zunächst thematisch induktiv codiert. Anschließend werden diese Textstellen hinsichtlich der Mindset-Orientierung durch zwei unabhängige Codierer*innen anhand deduktiv abgeleiteter Kriterien beurteilt. Der anschließende Vergleich codierter Textstellen in den Feedbacktexten der Interventions- und Kontrollgruppe liefert zusätzliche Hinweise zur Rolle des Mathematikspezifischen Mindsets für das Geben von Feedback. Die Ergebnisse werden hinsichtlich ihrer Limitationen und Implikationen diskutiert.

Eine Growth-Mindset-Intervention bei angehenden Mathematik-Lehrkräften lohnt sich auch für ihre leistungsschwachen Schülerinnen und Schüler

Katharina Asbury, Bastian Carstensen
IPN Kiel

Theoretischer Hintergrund

Mindset-Interventionen basieren auf dem Konzept von Carol Dweck (2006) und zielen darauf ab, die Überzeugungen von Individuen hinsichtlich ihrer kognitiven Fähigkeiten zu beeinflussen. Bisherige Studien haben das Potenzial solcher Interventionen zur Verbesserung von Motivation und Lernerfolg von Lernenden aufgezeigt (Burnette et al., 2023; Yeager et al., 2022). In der bisherigen Forschung wurde jedoch bisher wenig Augenmerk auf Lehramtsstudierende gelegt, die für aktuelle Probleme des deutschen Bildungssystems, wie dem akuten Lehrkräftmangel, eine entscheidende Rolle spielen könnten (SWK, 2023). Diese Studie legt einen besonderen Schwerpunkt auf leistungsschwache Schülerinnen und Schüler im Fach Mathematik: Der Anteil leistungsschwacher Schülerinnen und Schüler, die nach der Grundschulzeit in basalen Grundkompetenzen in den Bereichen Lesen und Rechnen nicht einmal die unterste Kompetenzstufe erreichen, war bereits in den letzten Jahren Grund zur Sorge, wie uns der IQB-Bildungstrend im Primarbereich und wiederholte PISA-Ergebnisse vor Augen führten (Stanat et al., 2022). Der letzte PISA-Bericht weist nicht darauf hin, dass sich der Anteil dieser Schülerinnen und Schüler verringert hat; im Gegenteil (OECD, 2023). Interventionsstudien mit Growth-Mindset-Schwerpunkt zeigten dabei, dass insbesondere leistungsschwache Kinder hinsichtlich ihres eigenen Growth Mindsets, aber auch in Bezug auf Motivation und Leistung von Mindset- Interventionen profitieren können (Burnette et al., 2023; Hwang et al., 2016). Die Studie setzt an genau diesem Punkt an und schlägt zusätzlich einen ökonomischen Forschungsansatz vor, bei dem untersucht wird, ob eine Mindsetintervention, die auf Lehrkräfte als wichtigste Akteure im Bildungssystem abzielt, außerdem positive Effekte bei deren leistungsschwachen Schülerinnen und Schülern hat. Vorarbeiten deuten in die Richtung einer positiven Assoziation der Lehrkraft- und Schüler-Mindsets (Mesler et al., 2021), jedoch wurde diese Annahme bislang nicht in einem experimentellen Setting mit leistungsschwachen Kindern und angehenden Lehrkräften untersucht.

Fragestellung

Diese Studie untersucht die Auswirkungen einer Mindset-Intervention mit insgesamt 83 Lehramtsstudierenden und 173 leistungsschwachen Schülerinnen. Ziel ist es, das Growth Mindset angehender Lehrkräfte sowie ihrer leistungsschwachen Schülerinnen und Schüler zu erhöhen. Die beiden Forschungsfragen lauten daher:

1. Kann das Growth Mindset angehender Mathematik-Lehrkräfte mit einer Growth Mindset erhöht werden?
2. Zeigen leistungsschwache Kinder, die von diesen angehenden Lehrkräften 12 Wochen lang in einem Mathematik - Förderunterricht unterrichtet werden, ebenfalls ein erhöhtes Growth Mindset?

Methode

Die 83 angehenden Lehrkräfte wurden randomisiert auf drei Versuchsbedingungen Kontrollgruppe (N=54), Fachdidaktik-Interventionsgruppe (N= 16) und Fachdidaktik + Mindsetinterventionsgruppe aufgeteilt. Das Studiendesign sah für die beiden Interventionsgruppen ein 20-stündiges Training mit fachdidaktischer Unterweisung vor. Die Fachdidaktik + Mindsetinterventionsgruppe erhielt dabei zusätzlich eine zweistündige Mindsetintervention, um sowohl die Growth Mindsets der Lehramtsstudierenden als auch die der Schülerinnen und Schüler steigern. Dabei wurden die Growth Mindsets bei Lehrenden und Lernenden zu drei Messzeitpunkten (Prätest, Posttest und Follow-Up erhoben).

Ergebnisse

Die Ergebnisse zeigten, dass die Lehramtsstudierenden in der Fachdidaktik + Mindsetinterventionsgruppe signifikante Erhöhungen im Growth Mindset im Vergleich zur anderen Interventionsgruppe aufwiesen ($\beta = .38$, $p = .02$). Zudem zeigten die Kinder, die von den angehenden Lehrkräften der Fachdidaktik + Mindsetinterventionsgruppe unterrichtet wurden, ebenfalls erhöhte Growth Mindsets zum Post-test ($\beta = .29$, $p = .01$). In der Fachdidaktikgruppe zeigten die Lehramtsstudierenden hingegen eine signifikante Verringerung ihres Growth Mindsets ($\beta = .18$, $p = .03$), wohingegen die Mindsets der Kontrollgruppe stabil blieben.

Relevanz und Implikationen

Diese Ergebnisse unterstreichen das Potenzial von Mindset-Interventionen, indem sie sowohl die Mindsets der Lehramtsstudierenden als auch der Schülerinnen beeinflussen. Die dynamische Interaktion zwischen Lehrenden und Schülerinnen und Schülern und die Notwendigkeit gezielter Interventionen zur Unterstützung beider Gruppen wird betont. Diese Studie liefert wertvolle Erkenntnisse für Bildungspraktiken und trägt zur laufenden Diskussion über die Förderung von Growth Mindset, insbesondere bei leistungsschwachen Schülerinnen und Schülern bei. Die praktischen Implikationen hinsichtlich der konkreten Umsetzung der Interventionsinhalte im Förderunterricht durch Unterrichtspraktiken der Lehrenden werden diskutiert.

S-10: Einstellungen gegenüber Diversitätsmerkmalen in schulischen Kontexten: Interventionsmöglichkeiten

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 165

Chair der Sitzung: Stefanie Bosse

Chair der Sitzung: Ursula Kessels

Einstellungen gegenüber Diversitätsmerkmalen in schulischen Kontexten: Interventionsmöglichkeiten

Chair(s): **Stefanie Bosse** (Universität Potsdam, Deutschland)

Diskutant*in(nen): **Ursula Kessels** (Freie Universität Berlin)

Bildung wird als der zentrale Schlüssel für gesellschaftliche Herausforderungen gesehen. Anerkennende und wertschätzende Einstellungen werden dabei als zentrale Gelingensbedingungen für ein gleichberechtigtes, soziales Miteinander in schulischen Kontexten betrachtet (Altrichter et al., 2016; Bohl et al., 2023). Vor dem Hintergrund voranschreitender inklusionsorientierter Entwicklungen durch die Umsetzung der UN-BRK und gesellschaftlicher Migrationsbewegungen durch internationale Konflikte, rücken Einstellungen gegenüber Heterogenität und Diversität zunehmend in den Fokus. Aus der Forschung geht indes hervor, dass sowohl Schüler*innen, als auch (angehende) Lehrkräfte abhängig vom Diversitätsmerkmal positive bis negative inklusionsorientierte Einstellungen und Haltungen zeigen (z.B. Denessen et al., 2022; Hellmich & Loeper, 2018; Neuenschwander & Niederbacher, 2021) und auch im schulischen Kontext mit stereotypen Vorstellungen agieren (Glock & Baumann, 2022). Da Einstellungen geplantes Verhalten bedingen (Ajzen & Fishbein, 2005), ist der Wunsch nach Einstellungsveränderungen groß. Vor diesem Hintergrund widmen sich die drei Beiträge des Symposiums den Möglichkeiten, Einstellungen, Haltungen und damit verbundenes Verhalten unterschiedlicher Akteur*innen gegenüber Diversitätsmerkmalen im schulischen Kontext positiv zu stärken.

Die Reihenfolge der drei Beiträge orientiert sich am Alter der untersuchten Akteur*innen, beginnend mit Grundschüler*innen, über Lehramtsstudierende, bis hin zu Lehrkräften.

Der erste Beitrag widmet sich den Einstellungen von $N = 374$ Schüler*innen an inklusiven Grundschulen. In längsschnittlichen autoregressiven Modellen wird gezeigt, dass sich Einstellungen gegenüber adaptiver Unterrichtsgestaltung für Lernende mit dem Förderbedarf Lernen und jenen mit einer anderen Herkunftssprache über die Teamfähigkeit stärken lassen, die eingesetzte Intervention jedoch keinen direkten Effekt auf die Einstellungen hatte.

Der zweite Beitrag betrachtet die inklusionsbezogene Einstellungsentwicklung von $N = 124$ Lehramtsstudierenden des Fachs Sport in Abhängigkeit von Vor- bzw. Kontakterfahrungen mit Menschen mit besonderen Bedürfnissen. Über ein quasi-experimentelles Prä-Post-Design mit zwei Messzeitpunkten zeigt das eingesetzte Lehr-Lernkonzept positive Effekte auf die Entwicklung einer die Vielfalt wertschätzenden Einstellung der angehenden Sportlehrkräfte.

Im dritten Beitrag untersuchen die Autor*innen welche Wirkungen die inklusive Haltung der Lehrkraft auf die soziale Akzeptanz von Schüler*innen mit unterschiedlichen Diversitätsmerkmalen hat ($N = 632$ Schüler*innen, 39 Klassen). Aus hierarchischen Regressionsmodellen geht hervor, dass Schüler*innen mit Diversitätsmerkmalen weniger akzeptiert werden, die inklusive Haltung der Lehrkraft hingegen nicht auf die soziale Akzeptanz innerhalb der Klassen wirkt.

Aus den Beiträgen lassen sich zentrale Erkenntnisse darüber gewinnen, welche Möglichkeiten, aber auch Grenzen es in den Bemühungen, Einstellungen gegenüber Diversitätsmerkmalen zu verändern, gibt. Die Beiträge werden von einer führenden Expertin im Bereich der Einstellungs- und Haltungsforschung diskutiert.

Beiträge des Symposiums

Einstellungen und Teamfähigkeit von Schüler*innen in inklusiven Grundschulen

Stefanie Bosse, Jenny Lenkeit, Nadine Spörer
Universität Potsdam

Theoretischer und empirischer Hintergrund

Gelingende Inklusion zielt durch die Berücksichtigung der heterogenen Lernvoraussetzungen von Schüler*innen unter anderem auf die soziale Teilhabe. In der Grundschule ist neben guter Unterrichtsqualität das Sozialverhalten der Lernenden entscheidend für soziale Teilhabe im inklusiven Unterricht. Nach Modellen von Kunter und Stanat (2002) und Bosse und Spörer (2022) wird das Sozialverhalten der Lernenden wiederum durch soziale und personale Kompetenzen sowie soziale Einstellungen beeinflusst.

Aus der Forschung geht indes hervor, dass die Einstellungen von Schüler*innen gegenüber Peers mit besonderen Unterstützungsbedarfen bzw. dem gemeinsamen Unterricht mit ihnen tendenziell neutral bis positiv sind (de Boer et al., 2011; Schwab, 2015). Abhängig von der Art des Unterstützungsbedarfes aber auch negativ ausfallen (z. B. gegenüber Schüler*innen mit Verhaltensschwierigkeiten oder körperlichen Beeinträchtigungen) (Hellmich & Loeper, 2018). Weil Einstellungen geplantes Verhalten bedingen (Ajzen & Fishbein, 2005), ist der Wunsch nach Einstellungsveränderungen groß. Die wenigen Bemühungen mit dem Ziel der Einstellungsveränderung verzeichnen teils erwartungswidrige Ergebnisse (de Boer et al., 2014, Meyer et al., 2015). Darüber hinaus, werden diese möglichen Veränderungen meist nicht mit anderen sozialen oder personalen Kompetenzen in Zusammenhang gebracht.

Das Wirkmodell von Bosse und Spörer (2022) geht jedoch davon aus, dass Einstellungen nicht nur das Sozialverhalten beeinflussen, sondern auch mit sozialen und personalen Kompetenzen zusammenhängen. Vor diesem Hintergrund untersucht der Beitrag, ob sich Einstellungen gegenüber Unterrichtsadaptionen für Schüler*innen mit dem Förderbedarf Lernen und Schüler*innen mit einer nicht-deutschen Herkunftssprache in Abhängigkeit von Veränderungen in der Teamfähigkeit, als Facette der sozialen Kompetenz, positiv stärken lassen.

Forschungsdesign

Als Datengrundlage dient die unterrichtsintegrierte Intervention „Soziale Entwicklung beim gemeinsamen Lernen“ (SEGEL), in der kooperative Lernformen (Syring, 2019) eingesetzt wurden, um das Sozialverhalten zu fördern (Bosse & Spörer, 2022). Der Interventionserfolg wurde zu drei Messzeitpunkten mittels Fragebögen erfasst. Informationen zur Teamfähigkeit und Einstellungsfacetten liegen von 374 Schüler*innen der vierten und fünften Jahrgangsstufe aus 16 Klassen in inklusiv arbeitenden Grundschulen vor. Nach Überprüfung der Messinvarianz wurden folgende Skalen ausgewählt.

Teamfähigkeit (Furthmüller, 2015): Beispielim „Ich arbeite gern mit anderen zusammen, auch wenn wir nicht befreundet sind.“, Cronbach's Alpha $\alpha_0 = 0.71$, $\alpha_1 = 0.71$, $\alpha_2 = 0.80$.

Einstellungen zur Unterrichtsadaption für Schüler*innen mit Sprachschwierigkeiten (Eigenentwicklung): Beispielitem „Es ist okay, wenn Kinder mit Schwierigkeiten in der Sprache wenige Aufgaben machen müssen.“, Cronbach's Alpha $\alpha_0 = 0.86$, $\alpha_1 = 0.86$, $\alpha_2 = 0.84$.

Einstellungen zur Unterrichtsadaption für Schüler*innen mit Lernschwierigkeiten (Eigenentwicklung): Beispielitem „Es ist okay, wenn Kinder mit Schwierigkeiten im Lernen wenige Aufgaben machen müssen.“, Cronbach's Alpha $\alpha_0 = 0.71$, $\alpha_1 = 0.73$, $\alpha_2 = 0.80$.

Alle Items wurden auf einer Likertskala von 0 = „stimmt gar nicht“ bis 3 = „stimmt genau“ eingeschätzt. Berichtet werden standardisierte Ergebnisse. Um die Veränderungen der Einstellungsfacetten in Abhängigkeit der Veränderungen der Teamfähigkeit durch die Intervention zu prüfen, wurden autoregressive Modelle je Einstellungsfacetten unter Berücksichtigung der geschachtelten Datenstruktur in MPLUS berechnet.

Ergebnisse & Diskussion

Die Ergebnisse zeigen, dass die Teamfähigkeit zum jeweils vorher gemessenen Zeitpunkt die Einstellungsfacetten signifikant ($p < 0.05$) positiv vorhersagt. Das gilt sowohl für die Einstellungen zur Unterrichtsadaption für Schüler*innen mit Sprachschwierigkeiten ($\beta_0 = .09$, $\beta_1 = .11$), als auch zur Unterrichtsadaption für Schüler*innen mit Lernschwierigkeiten ($\beta_0 = .13$, $\beta_1 = .11$).

Die Ergebnisse verdeutlichen, dass die Veränderung von Einstellungen über die Stärkung sozialer Kompetenzen gedacht werden muss, insbesondere vor dem Hintergrund, dass diese leichter trainierbar zu sein scheinen (z.B. Petermann et al., 2016; Urban et al., 2018), als Einstellungen.

„Immer überraschend“ - Zur Bedeutung biographischer Vorerfahrungen auf die inklusionsbezogene Einstellungsentwicklung Sportlehramtsstudierender

Helena Sträter¹, Michael Pfitzner²

¹Bergische Universität Wuppertal, ²Universität Duisburg-Essen

Theoretischer Hintergrund

Ausgewiesenes Ziel der Lehrkräfteprofessionalisierung ist die Ausprägung einer anerkennenden und wertschätzenden Einstellung der Studierenden zur Vielfalt von Schüler*innen. Eine solche ist als Schlüssel zum gelingenden inklusiven Unterricht anzusehen. Angehende Sportlehrkräfte bewegen sich dabei in einem besonderen Spannungsfeld. Ihre (Berufs-)Biographie sowie Vor- bzw. Kontakterfahrungen sind häufig durch ein auf Homogenität ausgerichtetes, teils exkludierendes und an den Perspektiven Leistung und Wettkampf orientiertem Sportsystem geprägt.

In diesem Rahmen arbeitet der Beitrag zum einen anhand bisheriger Erkenntnisse die (berufs-)biographischen Stationen angehender Sportlehrkräfte sowie die Bedeutung der Fachpraxis im sportwissenschaftlichen Studium auf. Es wird deren potentieller Einfluss auf die didaktischen Orientierungen der Lehramtsstudierenden für (inklusionsorientierten) Sportunterricht beleuchtet.

Bevor angehende Sportlehrkräfte sich offen dafür zeigen können, Sportunterricht zu inszenieren, der die Vielfalt ihrer Schüler*innen wertschätzt, braucht es eines grundlegenden Wandels von Akteur*innen im (Leistungs-) Sport zu Arrangeur*innen von Bewegung, Spiel und Sport (Baur, 1995). In diesem Kontext von Stärke und Schwäche zugleich verorten sich die schon vor dem Beginn des Studiums zeitlich sehr umfangreichen Vorerfahrungen der Studierenden an Lernorten des Sports wie dem Sportverein (Fischer & Pfitzner, 2021). Diese sind keinesfalls als inklusiv zu bezeichnen. Ganz im Gegenteil weist gerade der Leistungssport Facetten der Exklusion auf, die man exemplarisch an der Unterscheidung in olympische und paralympische Spiele mit der Orientierung an schneller, höher, stärker erkennen kann.

Für die Sportlehrkräftebildung werden in dieser Gemengelage neben den in der Lehrkräftebildung verfolgten kompetenz- (Terhart, 2011) und strukturtheoretischen Ansätzen (u.a. Helsper, 2007) in besonderer Weise biographietheoretische Perspektiven eingenommen.

Es geht um die Frage, inwieweit sich die sportpädagogische Arbeit an einem inklusiven Lernort auf die Einstellungsentwicklung auswirkt und welchen Einfluss biographische Vorerfahrungen der Studierenden haben.

Forschungsdesign

Der Beitrag fokussiert eine Möglichkeit in einem universitären Setting einen Lernort hochschuldidaktisch zu gestalten, der auf die inklusionsbezogene Einstellungsentwicklung von Lehramtsstudierenden abzielt. Zusammenfassend zeichnet sich das Lehr-Lernkonzept durch Kontakterfahrungen von Sportstudierenden zu Schüler*innen mit Förderschwerpunkt aus. Die Studierenden planen Sportunterricht, führen diesen durch und reflektieren ihn.

Die Effekte des Lehr-Lernkonzepts wurden über vier Semester quasi-experimentell im Prä-Post-Design mit zwei Messzeitpunkten untersucht. Die Messung der Entwicklung von Einstellungen zum inklusiven Sportunterricht ist quantitativ ausgerichtet. Im Rahmen der Interventionsstudie mit Kontrollgruppe nahmen 124 Sportlehramtsstudierende des Master of Education an der Untersuchung teil ($n_{IG} = 54$; $n_{KG} = 70$). Die Messung der Entwicklung von Einstellungen zum inklusiven Sportunterricht erfolgt mithilfe des adaptierten Fragebogeninstruments von Seifried und Heyl (2016) in den Dimensionen Fachliche Förderung, Persönliche Bereitschaft und Soziale Inklusion.

Ergebnisse

Eine MANOVA zeigt eine statistisch signifikante Interaktion zwischen den Experimentalgruppen und den Messzeitpunkten für die Einstellungen zum inklusiven Sportunterricht (Wilk's $\Lambda = .795$, $F(4,116) = 7.47$, $p < .001$, $\eta^2 = .205$).

Die Proband*innen machen zudem Angaben zu Vor- bzw. Kontakterfahrungen mit Menschen mit besonderen Bedürfnissen. Zum Zeitpunkt der Prä-Testung gibt ca. jede*r Vierte an, dass bisher keine Kontakterfahrungen mit Menschen mit besonderen Bedürfnissen gemacht wurden. Eine lineare Regressionsanalyse zeigt, dass in der vorliegenden Stichprobe die Vor- bzw. Kontakterfahrung mit Menschen mit besonderen Bedürfnissen kein signifikanter Prädiktor für die Einstellung zum inklusionsorientierten Sportunterricht ist ($\beta = .094$, $p = .341$).

Weiterhin zeigt eine MANOVA eine statistisch signifikante Interaktion für die Proband*innen mit und ohne Vorerfahrungen in der Interventionsgruppe und den Messzeitpunkten (Wilk's $\Lambda = .793$, $F(4,40) = 2.61$, $p = .049$, $\eta^2 = .21$). Für die Kontrollgruppe stellt sich keine signifikante Interaktion ein (Wilk's $\Lambda = .924$, $F(3,55) = 1.51$, $p = .223$, $\eta^2 = .08$).

Diskussion

Durch die Teilnahme am Lehr-Lernkonzept können Effekte der Entwicklung einer die Vielfalt wertschätzenden Einstellung nachgewiesen werden. Besonders stark scheint das Lehr-Lernkonzept auf Studierende ohne inklusionsorientierte Vor- bzw. Kontakterfahrungen zu wirken. Diskutiert wird abschließend, inwiefern diese universitären Vor- bzw. Kontakterfahrungen Einfluss auf die (Berufs-)Biographie der Lehramtsstudierenden haben kann.

Soziale Akzeptanz in Grundschulklassen: Welchen Unterschied macht die inklusive Haltung der Lehrkraft?

Yvonne Erhardt, Josephine Gatzweiler, Sina Schürer, Stefanie van Ophuysen
Universität Münster

Theoretischer Hintergrund

Gelingende soziale Partizipation wird als eines der Hauptziele inklusiver Bildung verstanden (z.B. Grosche, 2015). In unserem Beitrag betrachten wir die soziale Akzeptanz als einen Indikator sozialer Partizipation. Forschungsbefunde zeigen, dass die Einbindung aller Kinder gleichermaßen in die Klassengemeinschaft nicht durch inklusive Beschulung allein gesichert werden kann. So hat sich in empirischen Studien wiederholt gezeigt, dass individuelle Faktoren wie Migrationshintergrund, schwache schulische Leistungen und geringe emotional-soziale Kompetenz mit niedrigeren Werten für soziale Akzeptanz einhergehen (Böttiger, 2021; Schürer, 2020). Neben Variablen auf Individualebene zeigen sich auch Zusammenhänge mit Variablen auf Klassenebene. Farmer (2011) sieht die Lehrkraft als entscheidende Kraft an, die als „invisible hand“ durch ihre Beziehungsgestaltung und ihr Verhalten Einfluss auf die Qualität der Peer-Beziehungen nimmt (Endedijk et al., 2022). Ein Lehrkraftmerkmal, das die soziale Partizipation aller Kinder befördert, ist – so unsere Vermutung – die inklusive Haltung der Lehrkraft gegenüber der Individualität der Schüler*innen. Hierunter verstehen wir die bedingungslose Akzeptanz und Wertschätzung aller Schüler*innen in ihrer Individualität, also unabhängig von ihren individuellen Stärken und Schwächen. Eine solche Haltung sollte sich in der Gestaltung der Lehrkraft-Lernenden-Beziehung manifestieren und hierüber Einfluss auf die soziale Partizipation nehmen. Unterschiede in sozialer Partizipation zwischen Schulklassen (z. B. Huber & Wilbert, 2012) könnten somit – so unsere Annahme – auf Unterschiede in der inklusiven Haltung der Lehrkräfte zurückzuführen sein.

Inwiefern eine stark ausgeprägte inklusive Haltung der Lehrkraft darüber hinaus dazu beiträgt, dass die Effekte von Merkmalen des individuellen Kindes abgeschwächt werden, sodass soziale Partizipation für alle Kinder gleichermaßen gelingt, ist unseres Wissens bislang nicht untersucht.

Fragestellungen

In unserem Beitrag gehen wir daher folgenden Fragen nach:

- 1) Stehen Migrationshintergrund, schulische Leistungen und sozial-emotionale Kompetenzen im Zusammenhang mit der sozialen Akzeptanz?
- 2) Steht die inklusive Haltung der Lehrkraft in positivem Zusammenhang mit der sozialen Akzeptanz der Lernenden?
- 3) Variiert der Einfluss der Individualmerkmale auf die soziale Akzeptanz je nach Ausprägung der inklusiven Haltung der Lehrkraft?

Methode

Aus einer standardisierten Befragung in 39 Klassen stehen Daten von rund 632 Lernenden der Jahrgänge 3 und 4 aus Nordrhein-Westfalen zur Verfügung. Die soziale Akzeptanz wurde über einen soziometrischen Rating-Fragebogen erfasst. Der Migrationshintergrund wurde über die Familiensprache erfasst. Für jedes Kind liegen weiterhin Lehrkraftangaben zu den schulischen Leistungen und den sozial-emotionalen Kompetenzen vor. Die inklusive Haltung wurde über ein selbstentwickeltes Instrument mit den drei Subskalen Vielfalt, Individuum und Gemeinschaft erhoben. Zur Beantwortung der Forschungsfragen werden Zwei-Ebenen-Modelle angepasst, und neben den Effekten auf Klassen- und Lernenden-Ebene werden Cross-Level-Interaktionen überprüft.

Ergebnisse

Die Ergebnisse zeigen signifikante Zusammenhänge zwischen den Individualmerkmalen und sozialer Akzeptanz: So werden Kinder mit Migrationshintergrund weniger akzeptiert als Kinder ohne Migrationshintergrund ($b = -0.098$, $p < 0.001$). Außerdem erfahren Kinder mit niedrigen emotional-sozialen Kompetenzen weniger Akzeptanz ($b = 0.142$, $p < 0.001$), ebenso wie Kinder mit geringe schulischen Leistungen ($b = 0.037$, $p < 0.001$). Darüber hinaus besteht ein signifikanter Zusammenhang zwischen Geschlecht und sozialer Akzeptanz: Mädchen werden weniger akzeptiert als Jungen ($b = -0.049$, $p < .05$). Hinsichtlich der inklusiven Haltung findet sich kein signifikanter Zusammenhang zur Akzeptanz. Auch in Bezug auf Cross-Level-Interaktionen konnte kein signifikanter Effekt gefunden werden.

Die Ergebnisse werden mit Blick auf nicht erwartungskonformen Befunden sowie Implikationen für die schulische Praxis diskutiert.

S-11: Professionelle Wahrnehmung von Lehrkräften anforderungsbezogen betrachten

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: SN 169

Chair der Sitzung: Julia Bruns

Chair der Sitzung: Rebekka Stahnke

Chair der Sitzung: Marita Friesen

Professionelle Wahrnehmung von Lehrkräften anforderungsbezogen betrachten

Chair(s): **Julia Bruns** (Universität Paderborn, Deutschland), **Rebekka Stahnke** (IPN Kiel)

Diskutant*in(nen): **Marita Friesen** (PH Heidelberg)

Die Professionelle Wahrnehmung von Lehrkräften wird als wissensbasierte Fähigkeit, lernrelevante Unterrichtssituationen zu erkennen (*perception*) und zu verarbeiten (*interpretation, decision-making*), verstanden (Blömeke et al., 2015; Stahnke et al., 2016; van Es & Sherin, 2002). Für die Untersuchung dieser situationsbezogenen Kompetenzfacette wird zunehmend gefordert, die spezifischen Anforderungen an die Lehrkräfte stärker in den Fokus zu rücken, indem inhaltliche Schwerpunkte und kulturelle Kontexte berücksichtigt werden (Bastian et al., 2022; Friesen & Kuntze, 2021; Weyers et al., 2024; Yang et al. 2021). Dieser Forderung folgend, wird im Symposium die Bedeutung konkreter Anforderungssituationen für die professionelle Wahrnehmung von Lehrkräften berücksichtigt und untersucht.

Der erste Beitrag untersucht, inwieweit sich kulturelle Unterschiede in der Bewertung unterschiedlicher Situationen im Mathematikunterricht der Sekundarstufe zeigen. Die Datengrundlage bilden schriftliche Analysen von taiwanesischer und deutscher Lehrkräften zu Vignetten aus dem *TaiGer Noticing*-Projekt. Als Analyseraster wurden die vier Dimensionen des *usable knowledge* nach Kersting et al. (2012) genutzt (suggestions for improvement, student thinking/understanding, mathematical content, level of interpretation). Um kulturelle Unterschiede festzustellen, wurden die psychometrischen Eigenschaften zwischen den Gruppen verglichen. Dabei zeigt sich, dass die Analysedimensionen zwar weitgehend übertragbar sind, sich aber kulturelle Unterschiede insbesondere bei der Dimension *student thinking/understanding* ergeben.

Im zweiten Beitrag steht ein Anforderungssituation von Sportlehrkräften im Fokus: Die Reflektion der Anspruchsdimensionen eines inklusiven Sportunterrichts mit Schüler:innen. Der Beitrag hat das Ziel auf Basis von 100 videographierten Unterrichtsstunden ein Kompetenzprofil zu bestimmen, das u.a. die notwendigen situationsspezifischen Fähigkeiten für diese Reflexion ermittelt. Dafür werden anhand der Daten die zentralen Anforderungen an die Lehrkräfte ermittelt: Gesprächsfördernde Rahmenbedingungen schaffen, aktiv beobachten und zuhören, Verletzungen inklusiver Ansprüche thematisieren sowie Reflexionsgespräche moderieren. Mit diesen Anforderungen gehen spezifische Bedarfe u.a. an die situationsspezifischen Fähigkeiten der Lehrkräfte einher, die auf Basis der empirischen Daten herausgestellt werden.

Im dritten Beitrag wird die professionelle Unterrichtswahrnehmung von (angehenden) Lehrkräften bezogen auf unterschiedliche Anforderungssituationen des Mathematikunterrichts zum flexiblen Rechnen in Jahrgangsstufe 1 betrachtet.

Die Studie untersucht mittels Videovignetten expertise-bezogene Unterschiede in der professionellen Wahrnehmung, indem je 10 angehende und erfahrene Mathematiklehrkräfte betrachtet werden. Die zunächst inhaltsanalytisch ausgewerteten Daten der beiden Gruppen werden mittels der epistemischen Netzwerkanalyse verglichen. Erste Ergebnisse deuten darauf hin, dass sich expertise-bezogene Unterschiede nicht nur auf die Teilfacetten der professionellen Wahrnehmung, sondern auch auf die spezifischen Anforderungssituationen beziehen.

Abschließend werden die Beiträge von einer renommierten Forscherin diskutiert und reflektiert. Die Beiträge zeigen auf, wie die Berücksichtigung inhaltlicher und kultureller Anforderungen an die professionelle Wahrnehmung gelingen und wichtig Erkenntnisse für die Lehrkräfteaus- und Weiterbildung generieren kann.

Beiträge des Symposiums

MEASURING MATHEMATICS TEACHERS' USABLE KNOWLEDGE – IS KERSTING AND COLLEAGUES' APPROACH UNIVERSALLY APPLICABLE AND CULTURALLY NEUTRAL?

Josephine Paul¹, Anika Dreher², Ting-Ying Wang³, Feng-Jui Hsieh³, Anke Lindmeier¹

¹Universität Jena, ²PH Freiburg, ³NTNU Taipei, Taiwan

Identifying reliable methods for assessing noticing or related constructs, like teachers' "usable knowledge" (Kersting et al., 2012), remains a relevant concern in mathematics education research. Typically, vignette-based assessments are used for this. Kersting and colleagues (2005, 2010) developed one widely-cited, reliable approach. It utilizes a decompositional procedure to evaluate teachers' analysis of vignettes regarding four dimensions: suggestions for improvement (SI), student thinking/understanding (ST), mathematical content (MC), level of interpretation (LI; Kersting et al., 2010). It is particularly effective because it allows conclusions on instructional quality (Kersting et al., 2012). Its decompositional nature makes it appear universally applicable to various written responses and suggests it to be culturally neutral and suitable for cross-cultural comparisons.

Despite these strengths, the approach is mainly used by Kersting's research group. This raises the question of whether it is actually universally applicable to written responses and culturally neutral.

Therefore, we conducted a secondary analysis using a dataset from the *TaiGer Noticing*-project (e.g., Dreher et al., 2024). This dataset, like Kersting's (2008), consists of written analyses; here from 107 Taiwanese and 115 German teachers regarding four vignettes (Lindmeier et al., in press; Paul et al., accepted). Contrastingly to Kersting (2008), a prompt addressing only two (teacher and content: "evaluate the teacher's use of the task") instead of three aspects (student, teacher, and content) was used. Additionally, Kersting used videos, whereas, in the *TaiGer Noticing*-project, text vignettes, including cultural triggers, were used (Dreher et al., 2021).

If Kersting's approach is universally applicable, psychometric properties should be similar despite these differences. To test this, teachers' analyses were coded with 0 (not addressed), 1, or 2 (comprehensively addressed) regarding the four dimensions (Kersting et al., 2010). Subsequently, the same calculations Kersting (2008) did were conducted: separability, internal consistency of dimensions, item difficulty using CTT and unidimensional IRT (Graded Response Model). Cultural neutrality was examined using a G-study (facets: culture, vignette, culture×vignette) and by comparing Taiwanese and German teachers' mean scores.

Analog to the original context, SI, ST, and MC showed good separability (correlated errors < .1). LI was less separable from the other (.21 - .41). Internal consistency was only acceptable for SI and LI (α =.65), but low for MC (α =.34). Regarding the mean

scores, similar to Kersting (2008), the MC-items appeared as most difficult ($M=3.59$). IRT showed that achieving MC-score 2 requires the highest knowledge level (for vignettes 1 to 3: $\theta > 2$). Contrary to Kersting, LI-items instead of ST-items are the easiest ($M=4.41$).

The G-study showed that most variance (36.8%) is explained by the participants, 0.7% by culture, 4.6% by the vignette, and 2% by culture $\times$ vignette. Moreover, the Taiwanese teachers addressed ST significantly more often.

To summarize, Kersting's approach could be transferred to teachers' written analyses of text vignettes, which is particularly evident in the largely good separability of dimensions and similar CTT results. Limitations arise from the secondary use of data and the design of the TaiGer instrument, which aims to represent situations diverse regarding teacher-student interactions and content, which could be one explanation for the low α -values.

Considering culture, the approach seems neutral regarding three dimensions (SI, MC, LI) but not regarding ST. This may be due to the high student engagement in three of the four vignettes, which is dissimilar to traditional teacher-centered instruction in Taiwan and might have caught the attention of Taiwanese teachers. Further analyses would be needed to substantiate what causes the Taiwanese teachers' higher ST-scores. Simultaneously, the approach does not provide insights into underlying cultural characteristics like cultural norms (e.g., Dreher et al., 2024). Consequently, the approach seems universally applicable, but there are indications that its applicability for intercultural comparisons is limited.

Mit Schüler*innen über Anspruchsdimensionen eines inklusiven Sportunterrichts reflektieren – Entwicklung eines anforderungssituations-bezogenen Kompetenzprofils für Sportlehrkräfte

André Meister¹, Björn Brandes², Jan Erhorn¹

¹Universität Oldenburg, ²Universität Osnabrück

Einleitung

Im Rahmen eines inklusiven Sportunterrichts treten vielfältige Anforderungssituationen für Lehrkräfte auf. Eine davon ist „Mit Schüler:innen über die Anspruchsdimensionen eines inklusiven Sportunterrichts (zu) reflektieren“ (Erhorn & Langer, 2022). Unklar ist bislang noch, welche Dispositionen, situationsspezifische Fähigkeiten und Handlungsweisen die Lehrkräfte benötigen, um diese Anforderungssituation bewältigen zu können. Als Grundlage für eine fundierte Professionalisierung angehender Sportlehrkräfte wird deshalb ein anforderungssituationsbezogenes Kompetenzprofil entwickelt und im Beitrag vorgestellt.

Theoretischer Rahmen

Im sportdidaktischen Diskurs wird kontrovers diskutiert, was unter einem inklusiven Sportunterricht verstanden und an welchen Prinzipien sich dieser orientieren sollte. Das in diesem Beitrag geteilte Verständnis beruht auf den Anspruchsdimensionen der Teilhabe, Anerkennung und individuellen Förderung (Jürgens & Neuber, 2020; Grimminger, 2012; Neuber, 2012). Zusätzlich wird die Reflexion bzw. das Reflektieren sportunterrichtlichen Handelns als ein zentrales Prinzip eines erziehenden Sportunterrichts gesehen (Beckers, 2013). Daraus ergibt sich u. a. die Herausforderung über die (z. B. Verletzung der) Ansprüche inklusiven Sportunterrichts kompetent mit den Schüler:innen zu reflektieren. Im hiesigen Beitrag wird Kompetenz als Kontinuum zwischen Performanz und Dispositionen betrachtet, zwischen denen situationsspezifischen Fähigkeiten eine vermittelnde Funktion einnehmen (Blömeke, Gustafsson & Shavelson, 2015). Professionelle Kompetenzen können vor dem Hintergrund beruflicher Anforderungen ermittelt und auf mittlerer Abstraktionsebene in Form von Kompetenzprofilen konkretisiert werden. Ein kompetentes Handeln, das sowohl theoretisch als auch empirisch fundiert werden kann, wird entsprechend dieser Profile als solches beurteilbar (Oser & Bauder, 2013).

Methodisches Vorgehen

Das methodische Vorgehen orientiert sich am integrativen Ansatz zur Ermittlung professioneller Kompetenzen von Erhorn, Setzer und Wohlers (2019). Dieser sieht vor, dass zunächst gegenstandsbezogenes Datenmaterial erhoben wird. Im Rahmen des QIS-Projekts wurden insgesamt 104 Unterrichtsstunden aus 14 Klassen an sechs Hamburger Schulen (drei Grund-, drei Stadtteilschulen) videographisch dokumentiert. Eine Kamera filmte die Sportstätte in der Totale, während eine andere der Lehrkraft, mit einem Funkmikrofon ausgestattet, folgte. Im Folgenden wurden Situationen, in denen Reflexionsphasen über Ansprüche inklusiven Sportunterrichts sichtbar wurden, zunächst fallspezifisch und anschließend fallübergreifend mit Blick auf Handlungsweisen, situationsspezifische Fähigkeiten und Dispositionen interpretativ ausgewertet. In einem weiteren Schritt wurden die beobachteten Handlungsweisen reflexiv-normativ eingeordnet und anschließend in ein Kompetenzprofil überführt.

Ergebnisse

Das Kompetenzprofil umfasst vier Teilanforderungen. Zum einen müssen Lehrkräfte wiederkehrend „Gesprächsförderliche Rahmenbedingungen herstellen“, wenn diese ungünstig sind (Handlungsweisen: „Bewegungsgeschehen unterbrechen“, „Positionsänderung organisieren“ und „Verbale und nonverbale Ruhe/ Geduld einfordern“). Zum anderen müssen Lehrkräfte „Aktiv Beobachten und Zuhören“ (Handlungsweisen: Bewusst hin- oder wegschauen bzw. hin- oder weghören), um „Verletzungen inklusiver Ansprüche thematisieren“ (Handlungsweisen: „Impulse der SuS aufgreifen“, „Problematisierende Nachfrage stellen“, „Direkte verbale Reaktion“, „Eigene Impulse/ Überlegungen einbringen“, und „Nonverbales Problematisieren“) zu können. Zudem müssen Sie in der Lage sein „Reflexionsgespräche führen und moderieren“ zu können (Handlungsweisen: „Beobachtungen erfragen“, „Normen benennen und explizieren“, „Perspektivwechsel anregen durch Hineinversetzen lassen“, „Perspektivwechsel anregen durch Konfrontation“, „Verstehen sicherstellen“, „Lösungen erfragen oder vorschlagen“, und „Anweisen“).

Diskussion und Ausblick

Einerseits erweist sich das forschungsmethodische Verfahren als gewinnbringend, da das Kompetenzprofil sowohl empirisch gestützt als auch normativ gerahmt wird. Zwar kann durch die Fallzahl lediglich ein Teil des Phänomenbereichs abgebildet werden, jedoch bietet die empirische Fundiertheit konkrete Möglichkeiten zur Fallarbeit in der Hochschullehre (Langer & Erhorn, 2023).

Zur Rolle der Anforderungssituationen für die professionelle Unterrichtswahrnehmung von Noviz:innen und Expert:innen im Bereich des flexiblen Rechnens

Federica Becker¹, Julia Bruns¹, Rebekka Stahnke²

¹Universität Paderborn, ²IPN Kiel

Professionelle Unterrichtswahrnehmung wird als wesentlicher Bestandteil professioneller Kompetenz von Lehrkräften betrachtet und beschreibt, wie professionelles Wissen in beruflichen Anforderungssituationen wirksam wird (Sherin, Jacobs & Philipp 2011). In der Literatur wird professionelle Unterrichtswahrnehmung unterschiedlich konzeptualisiert, wobei aktuelle Studien drei

Teilfacetten unterscheiden: wahrnehmen, interpretieren und entscheiden (Blömeke et al. 2015; Stahnke et al. 2016; Kaiser et al. 2017). Um Lehrkräfte in der Entwicklung ihrer professionellen Wahrnehmung in Aus- und Fortbildungsangeboten zu unterstützen, sind insbesondere Erkenntnisse zu Unterschieden in Abhängigkeit zu der Expertise von Lehrkräften relevant. Studien zeigen zwar eine Wechselbeziehung zwischen der Wahrnehmung und dem professionellen Wissen, konzentrieren sich jedoch entweder auf angehende oder erfahrene Lehrkräfte, während Vergleiche zwischen beiden Gruppen selten sind (Bastian et al. 2022; Gold & Holodynski 2017; König et al. 2022; Weyers et al. 2023). Die wenigen vorliegenden Forschungsergebnisse zeigen, dass Expert:innen höhere Fähigkeiten im Bereich der professionellen Wahrnehmung aufweisen als Noviz:innen (Gegenfurth et al. 2020; Friesen & Kuntze 2021; Yang et al. 2021b), ohne diese Unterschiede differenziert zu beschreiben. Als weiteres Desiderat können fach- bzw. inhaltspezifische Aspekte der professionellen Wahrnehmung beschrieben werden (Bastian et al. 2022; Stahnke & Friesen 2023). Die Relevanz der fach- bzw. inhaltspezifischen Untersuchung der professionellen Wahrnehmung zeigt sich u. a. bei Stahnke & Friesen (2023), die Unterschiede in der professionellen Wahrnehmung zum Classroom Management von Lehrkräften aus zwei Fächern berichten. Gegenstandsspezifische Noviz:innen-Expert:innen-Vergleiche könnten Aufschluss über spezifische Schwächen und Stärken der professionellen Wahrnehmung in konkreten Anforderungssituationen geben, die wiederum wichtige Ansatzpunkte für die Gestaltung von Unterstützungsangeboten bilden. Eine inhaltspezifische Betrachtung der professionellen Wahrnehmung von Lehrkräften könnte zudem wertvolle Kenntnisse darüber liefern, welche Faktoren eine höhere Wahrnehmungskompetenz bei Expert:innen in Bezug auf konkrete Anforderungssituationen hervorbringen.

Die vorliegende Studie zielt daher darauf ab, einen Beitrag zu diesem Forschungsdesiderat zu leisten, indem die professionelle Wahrnehmung von angehenden und erfahrenen Mathematiklehrkräften zum Unterrichtsgegenstand ‚Flexibles Rechnen im Zahlenraum bis 20‘ untersucht wird. Die Forschungsfrage lautet: Welche Unterschiede zeigen sich hinsichtlich der professionellen Unterrichtswahrnehmung zum Gegenstand ‚Flexibles Rechnen im Zahlenraum bis 20‘ bei angehenden und erfahrenen Mathematiklehrkräften?

Zur Beantwortung der Forschungsfrage wurden Lehramtsstudierende im Master, die ihr Praxissemester bereits absolviert haben ($n = 10$), und Lehrkräfte mit einer Berufserfahrung von mindestens 5 Jahren und einer zusätzlichen mathematikbezogenen Tätigkeit ($n = 10$) miteinander verglichen. Im Rahmen dieses Expert:innen-Noviz:innen-Vergleiches wurden die (angehenden) Mathematiklehrkräfte in Einzelinterviews befragt. Die Interviews basieren auf der Betrachtung von neun Videovignetten aus dem Mathematikunterricht zum flexiblen Rechnen im Zahlenraum bis 20. Zentrale Anforderungen im Mathematikunterricht zum flexiblen Rechnen im Zahlenraum bis 20 liegen darin, Lernstände und -prozesse zu diagnostizieren und beurteilen, Lernprozesse zu unterstützen und fördern sowie gemeinsam Gespräche moderieren. Für ebendiese Anforderungssituationen kann auch der professionellen Wahrnehmung eine besondere Bedeutung zugeschrieben werden. Die Videoszenen bilden daher Situationen der drei Anforderungsbereiche in Kombination mit den zentralen Lernzielen zum flexiblen Rechnen im Zahlenraum bis 20 (Rechenwege entdecken, Rechenwege kommunizieren und darstellen, Rechenwege begründen) ab. Die Videoanalyse dient als wesentliches Instrument zur Erfassung der Wahrnehmungsprozesse (van Es & Sherin 2005, 2008), wobei jeweils die drei Teilfacetten der professionellen Unterrichtswahrnehmung (wahrnehmen, interpretieren, entscheiden) in den Interviews berücksichtigt werden.

Die transkribierten Interviewdaten werden aktuell mittels qualitativer Inhaltsanalyse ausgewertet. Aufbauend wird eine Epistemic-Network-Analysis (ENA) durchgeführt, wodurch sich gemeinsam auftretende Codes identifizieren und Beziehungen zwischen Codes visualisieren lassen (Shaffer et al. 2016). So können Unterschiede in der gegenstandsspezifischen Unterrichtswahrnehmung zwischen den beiden Gruppen (Expert:innen und Noviz:innen) erfasst und in Netzwerkgraphen dargestellt werden. Die Datenanalyse wird bis Ende des Jahres abgeschlossen. Erste Ergebnisse lassen auf Unterschiede in der Wahrnehmung zwischen den beiden Gruppen gerade bezüglich der Anforderungssituation ‚Moderieren von Gesprächen‘ zugunsten der Expert:innen erkennen. Werden die Teilfacetten der professionellen Unterrichtswahrnehmung betrachtet, so können besonders im Bereich des Entscheidens Unterschiede erwartet werden.

S-12: Sozio-emotionale Lehr/Lernerfahrungen im Schulkontext

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: EO 145

Chair der Sitzung: Diana Raufelder

Sozio-emotionale Lehr/Lernerfahrungen im Schulkontext

Chair(s): Diana Raufelder (Universität Greifswald, Deutschland)

In den letzten zwei Jahrzehnten ist ein Anstieg an Forschungsergebnissen zu sozialen und emotionalen Schulerfahrungen zu verzeichnen (vgl. Schlesier & Raufelder, 2024). Diese Entwicklungen werden meist durch die Selbstbestimmungstheorie (Deci & Ryan, 1985) oder das Konzept der emotionalen Intelligenz (Mayer & Salovey, 1997) theoretisch untermauert. Allerdings werden diese sozio-emotionalen Aspekte im schulischen Unterricht häufig nur rudimentär oder isoliert voneinander behandelt. Dabei sind sozio-emotionale Erfahrungen – im Sinne positiver sozialer Beziehungen und damit verknüpften positiven emotionalen Erlebnissen – von zentraler Bedeutung für schulische Motivations- und Lernprozesse (Bakadorova & Raufelder 2018; Hattie et al. 2018; Roorda et al. 2011). Zudem tragen sie zu einem verminderten Stresserleben und einem gesteigerten schulischen Wohlbefinden bei (Hagenauer & Raufelder 2021). Diese Erfahrungen fördern die Entwicklung der Fähigkeiten im Umgang mit sozialen Situationen sowie das Verständnis und die Regulation von Emotionen, also die sozio-emotionale Kompetenz (Beaumont et al. 2023; Hoferichter et al. 2021; Saxer et al. 2024). Im schulischen Kontext sollte daher – neben den akademischen Bildungszielen – die Stärkung der sozialen Fähigkeiten, die Förderung von Empathie, die Entwicklung von Konfliktlösungsfähigkeiten sowie ein tiefgehendes Emotionsverständnis priorisiert werden.

Dieses Symposium stellt vier evidenzbasierte Ansatzpunkte zur Förderung sozio-emotionaler Schulerfahrungen vor:

Der erste Beitrag untersucht die Auswirkungen von Lehrerhumor auf die Lehrer-Schüler-Beziehung, die Emotionen der Schüler:innen und deren intrinsische Motivation anhand von Daten von 359 Neuntklässler:innen. Lerngegenstandsbezogener und selbstabwertender Humor fördern positive Beziehungen, während aggressiver Humor negative Auswirkungen hat. Angemessene Humorformen fördern positive Emotionen und die Motivation der Schüler:innen.

Der zweite Beitrag analysiert den Einfluss selbstgesteuerter Lernphasen auf soziale Eingebundenheit und Emotionen von 754 Schüler:innen. Selbstgesteuertes Lernen stärkt die Wahrnehmung sozialer Eingebundenheit und positive Emotionen, was das Schulengagement und die ganzheitliche Entwicklung der Schüler:innen fördert.

Der dritte Beitrag adressiert Emotionen beim Schriftspracherwerb und schriftsprachliche Kompetenzen von Grundschüler:innen im Kontext einer Service-Learning-Intervention. Durch ein längsschnittliches Design mit 252 Grundschüler:innen zeigt sich, dass Ärger beim Schriftspracherwerb negativ mit schriftsprachlichen Kompetenzen korreliert. Die Studie untersucht, ob Service-Learning-Interventionen positive Effekte auf Emotionen und schriftsprachliche Kompetenzen haben.

Der vierte Beitrag untersucht den Einfluss der Unterstützung durch Lehrkräfte und Peers auf die Entwicklung emotionaler und verhaltensbezogener Schwierigkeiten von Sekundarschüler:innen. Analysen von 733 Schüler:innen zeigen, dass Unterstützung durch Lehrkräfte und Peers Anpassungsprobleme und emotionale Schwierigkeiten verringert.

Insgesamt präsentiert das Symposium evidenzbasierte Forschungsbeiträge zur Bedeutung sozio-emotionaler Aspekte im schulischen Kontext. Die Studien bieten Einblicke in die Förderung positiver Lehr-Schüler-Beziehungen, die Auswirkungen selbstgesteuerter Lernphasen und die Relevanz von Lehrer- und Peer-Unterstützung für die Entwicklung von Schüler:innen und tragen zur ganzheitlichen Bildungsförderung bei.

Beiträge des Symposiums

Longitudinal Impact of Teacher Humor on Student Socio-Emotional Characteristics

Hassan Banarjee, Sonja Bieg

PH Weingarten

Theoretical background

Teacher humor is a multi-dimensional construct which can interplay with students' socio-emotional instructional characteristics. This research delves into how it strengthens teacher-student relationship, affecting student emotions and intrinsic motivation. According to the Instructional Humor Processing Theory (IHPT, Wanzer et al., 2010) a form of humor that fits the teaching content plays a facilitative role in student learning, while its irrelevant and inappropriate use does not (Bieg & Dresel, 2018). Four different types of teacher humor have been distinguished (Bieg & Dresel, 2016; Frymier et al., 2008); course-related, course-irrelevant, self-disparaging, and aggressive humor. Prior cross-sectional studies partially support the assumptions of the IHPT with empirical evidence and highlights course-related teacher humor as an effective tool in classroom which is positively related with a better teacher-student relationship, positive emotions and students' intrinsic motivation (Bieg & Dresel, 2018; Bieg et al., 2017). However, there is no longitudinal study examining the role of different teacher humor types on teacher-student relationship, students' emotions, and intrinsic motivation, hence this study aimed to fill this gap.

Research questions

We investigated whether student perceived socio-emotional instructional characteristics i.e. teacher-student relationship and students emotional experiences are affected by teacher humor forms in classrooms. We also explored whether only course-related teacher humor is potentially feasible to improve the perception of students' teacher-student relationship and their emotions and intrinsic motivation.

Method

The current sample consists of 365 German speaking 9th grade students from 18 classes at secondary schools - Gymnasium (51.4 % female; Mage: 14.28 years; SD = 0.50). We measured students' perceptions of teacher humor with HUMLAS (Bieg & Dresel, 2016; 16 items, $\alpha = .78 - .94$), students' perceptions of teacher-student relationship (Kunter et al., 2002; 5 items, $\alpha = .90$), students' emotion (Pekrun et al., 2011) enjoyment (3 items, $\alpha = .92$), anger (4 items, $\alpha = .84$), boredom (3 items, $\alpha = .85$), anxiety (4 items, $\alpha = .84$), and intrinsic motivation (Seidel et al., 2005; 4 items, $\alpha = .90$). All scales ranged from not at all true (1) to very true (5). Students completed the questionnaires at two time points, at the beginning of the school year and after the half-time. Structural equation modeling (SEM) using R software determined the predictive power of different teacher humor types and the mediating effects of socio-emotional instructional characteristics.

Results

The SEM model from different teacher humor types at time point 1 on positive teacher student relationship and students' emotions and motivation at time point 2 fitted the data very well (CFI = 0.98, TLI = 0.97, RMSEA = 0.06, SRMR = 0.04). The Chi-square test results indicated a significant model fit ($t = 44.822$ with $df = 20$ and $p = .001$). Regression estimates showed that course-related and self-disparaging teacher humor positively predicted positive teacher-student relationship ($\beta = 0.506$; $\beta = 0.183$ respectively), while aggressive teacher humor ($\beta = -0.409$) and irrelevant teacher humor ($\beta = -0.243$) negatively predicted positive teacher-student relationship. Positive teacher-student relationship, in turn, negatively predicted anxiety ($\beta = -0.21$), anger ($\beta = -0.52$), and boredom ($\beta = -0.65$), but positively predicted enjoyment ($\beta = 0.73$) and intrinsic motivation ($\beta = 0.48$) – all significant $p < .001$.

Discussion

Course-related and self-disparaging teacher humor emerged as potent enhancers of positive teacher-student relationships, leading to improved student emotional well-being and intrinsic motivation. Conversely, irrelevant and aggressive humor showed limited or negative effects. These findings underscore the significant role of teacher humor in shaping the educational climate and student well-being.

Soziale Eingebundenheit und positive Emotionen im Unterricht: Einfluss von selbstgesteuerten Lernphasen auf adoleszente Schüler:innen

Lia Grahl, Sabine Schweder
Universität Greifswald

Theoretischer Hintergrund

Die Qualität einer schulischen Lernumgebung und insbesondere in welchem Umfang diese das Bedürfnis nach sozialer Eingebundenheit erfüllt, beeinflusst das Ausmaß positiver Emotionen (Frenzel et al., 2009). Digitale Klassenzimmer fördern selbstgesteuertes Lernen (Dumont et al., 2010; Schleicher, 2015), was umgekehrt die Frage aufwirft, inwiefern fühlen sich in diesem Lernkonzept Schüler:innen sozial eingebunden, obwohl diese individualisiert lernen (Bolhuis & Voeten, 2001). Aktuell belegen empirische Befunde bei zunehmenden Schulalter einen Abwärtstrend bei positiven Emotionen (Pekrun et al., 2002). Empirische Forschung ist notwendig, um zu untersuchen, wie selbstgesteuerte Lernphasen (SL) als Erweiterung des Fachunterrichtes dem psychologischen Bedürfnis nach sozialer Eingebundenheit entspricht und positive Emotionen davon möglicherweise positiv berührt werden.

Neben der Kontroll-Wert-Theorie (Pekrun, 2006) implizieren Befunde zur Stage-Environment-Fit-Theorie (Eccles & Midgley, 1989), dass Lernumgebungen, die das Bedürfnis nach sozialer Eingebundenheit der Schüler:innen erfüllen, das Empfinden positiver Emotionen unterstützen (Reeve, 2002). Selbstgesteuerte Lernphasen fördern durch individuelle Lernplanung (Voskamp et al., 2019) und individualisierte Lernprozesse (Van Zyl & Mentz, 2019), individuelle Zuwendung der Lehrer*innen (Knowles, 1980), Peer-to-Peer Interaktionen (Bolhuis & Voeten, 2001) und individuelle Feedbackprozesse (Van Deur, 2020) Empfindungen soziale Eingebundenheit (Pekrun, 2016). Ein lehrer*innenzentrierter Unterricht zeichnet sich durch die zentrale Rolle der Lehrkraft aus, die Wissen direkt an die Schüler:innen vermittelt. Die Unterrichtsstruktur ist klar vorgegeben und folgt einem festen Ablauf. Interaktion und Kommunikation unter den Schüler:innen sind begrenzt, da der Fokus auf der Wissenskontrolle und -vermittlung durch die Lehrkraft liegt (Helmke, 2004; Helmke, 2009; Helmke, 2013).

Der untersuchte Ansatz umfasst, in den Fachunterricht integrierte, SL-Intervalle in denen die Lehrkräfte die Kontrolle an die Schüler:innen abgeben, die dann ihre Lernbedingungen selbst steuern.

Fragestellung

Hypothese H1a: SL-Phasen führen zu einer stärkeren Wahrnehmung sozialer Eingebundenheit im Vergleich zu lehrer*innengeleiteten Phasen.

Hypothese H1b: SL-Phasen tragen zu einer positiven Entwicklung von Emotionen im Vergleich zu lehrer*innengeleiteten Phasen bei.

Hypothese H2: Soziale Eingebundenheit in SL-Phasen ist positiver mit der Entwicklung von Emotionen assoziiert als in lehrer*innengeleiteten Phasen.

Methode

Die Studie untersucht positive Emotionen und soziale Eingebundenheit während eines Schuljahres basierend auf der Stage-Environment-Fit-Theorie und der Kontroll-Wert-Theorie. Analysiert wurden Fragebogendaten zu vier Messpunkten von 754 Schüler:innen (MAlter=13,56; SD=1,2; 49,7% weiblich) aus 37 Klassen der 6. bis 9. Jahrgangsstufe in sechs ländlichen Schulen in Mecklenburg-Vorpommern und Schleswig-Holstein. Diese Schulen integrieren ab der 6. Klasse SL-Intervalle in ihren Lehrplan. Während der SL-Intervalle bearbeiten die Schüler:innen selbstgewählte Fragestellungen. Zwei univariate Baseline Latent Change Score Modelle und ein bivariates Neighbor Latentes Change Score Modell wurden mit Mplus durchgeführt, um Veränderungen und Assoziationen zwischen den Variablen zu untersuchen.

Sozialen Eingebundenheit wurde mit einer validierten Subskala des Fragebogens zur Unterstützung der Grundbedürfnisse für Jugendliche (Müller & Thomas, 2011). Positive Emotionen wurden mit einer validierten Subskala von Prenzel et al. (1996).

Ergebnisse und Implikationen

Hypothese H1a wurde bestätigt: Ein positiver Trend in der Entwicklung sozialer Eingebundenheit wurde über die Zeit hinweg beobachtet, besonders während der SL-Intervalle. Zu t3 gab es keinen signifikanten Unterschied zu t1. Hypothese H1b zeigte, dass positive Emotionen im Laufe des Schuljahres zunahmen, hauptsächlich durch die SL-Interventionen. SL-Intervalle können somit einem negativen Trend bei der Entwicklung positiver Emotionen entgegenwirken, der bei Übergang in die Sekundarstufe und Adoleszenz ausgeprägt ist.

Soziale Eingebundenheit und positive Emotionen waren im Zeitverlauf miteinander verbunden, jedoch nur bei praktiziertem selbstgesteuertem Lernen. Die positive Entwicklung von Emotionen und die gesteigerte Wahrnehmung sozialer Eingebundenheit während SL-Phasen zeigen, dass SL langfristig zur Verbesserung des schulischen Engagements beitragen kann. Dies unterstreicht die Bedeutung von Unterrichtsansätzen, die soziale Eingebundenheit fördern. Schulen, die SL erfolgreich integrieren, leisten einen wichtigen Beitrag zur ganzheitlichen Entwicklung der Schüler:innen und schaffen Grundlagen für lebenslanges Lernen und persönliches Wachstum.

Emotionen beim Schriftspracherwerb und schriftsprachliche Kompetenzen von Grundschüler:innen: Entwicklung eines Fragebogens (AEQW-ES) und Einsatz einer Service-Learning-Intervention zur Förderung

Juliane Schlesier¹, Simon Meyer², Stephanie Lichtenfeld³, Lilith Rothmund¹, Daniela Eckhoff¹, Tanja Jungmann¹

¹Carl von Ossietzky Universität Oldenburg, ²FAU Erlangen-Nürnberg, ³Universität Hamburg

Theoretischer Hintergrund

Die sprachlichen Kompetenzen von Kindern haben aktuellen Schulleistungsstudien zufolge abgenommen (Stanat et al., 2022). Da diese als Kernkompetenzen für den weiteren schulischen Erfolg gelten, ist deren nachhaltige Förderung äußerst bedeutsam. Laut der Kontroll-Wert-Theorie (Pekrun, 2000, 2006; Pekrun et al., 2017) werden die schulischen Leistungen mitunter durch das emotionale Erleben von Schüler:innen im Unterricht bedingt; dies wird auch anhand empirischer Daten im Grundschulalter gestützt (Lichtenfeld et al., 2023). Ein Ansatz kann es sein, die Schreibkompetenzen über die Förderung positiver emotionaler Schulerfahrungen im Rahmen von einer Service-Learning-Aktivität zu steigern, wobei fachliches Lernen mit sozialem Engagement verbunden und auf diese Weise das Erleben positiver Emotionen begünstigt wird. Obwohl positive Effekte dieses Ansatzes bereits auf zahlreiche Kompetenzen nachgewiesen werden konnten (Celio et al., 2011; Yorio & Ye, 2012), wurde diese Methode für den Grundschulkontext noch nicht systematisch-empirisch in Hinblick auf ihren Effekt auf die Emotionen beim Schriftspracherwerb und die grundlegenden schriftsprachlichen Kompetenzen untersucht.

Ziel der Studie

Teilstudie 1 soll eruiert werden, ob die Emotionen Freude, Stolz, Ärger und Angst beim Schriftspracherwerb zu Beginn des Schuljahres (t1) die schriftsprachlichen Kompetenzen von Grundschüler:innen zu Ende des ersten Schulhalbjahres (t2) vorhersagen können (Forschungsfrage FF1). In Teilstudie 2 wird der Frage nachgegangen, ob eine Service-Learning Intervention, welche im zweiten Schulhalbjahr in den Grundschulen implementiert wurde, dazu beitragen kann, sowohl die Emotionen beim Schriftspracherwerb als auch die schriftsprachlichen Kompetenzen am Ende des zweiten Schulhalbjahres (t3) zu fördern (FF2).

Methodik

Die Daten wurden in einem längsschnittlichen Design mit drei Messzeitpunkten erhoben. Insgesamt wurden 252 Grundschüler:innen (MAlter = 7.48 Jahre; SDAAlter = 0.93; 47.7% Mädchen) der Klassen 1-4 aus sechs Grundschulen zu ihren Emotionen beim Schriftspracherwerb (AEQW - ES) befragt. Die schriftsprachlichen Kompetenzen wurden mit der Hamburger Schreibprobe (HSP; May et al., 2018) erfasst. Für die Teilstudie 1 wurde die Konstruktvalidität des AEQW-ES-Fragebogens anhand von zufälligen Split-half-Subsamples über explorative und konfirmatorische Faktorenanalysen überprüft. Anschließend wurden alle Daten von t1 und t2 genutzt, um FF1 nachzugehen. Dafür wurde ein Mehrebenen-Pfadmodell gerechnet, bei welchem die schriftsprachlichen Kompetenzen (alphabetische und orthographische Kompetenz) als Outcome-Variablen und die Emotionen Freude, Stolz, Ärger und Angst beim Schriftspracherwerb als Prädiktoren dienten; Geschlecht und die Anzahl der Bücher zuhause wurden als Kovariaten in den Analysen berücksichtigt. Für die Beantwortung von FF2 in Teilstudie 2 wurden die Daten aller Messzeitpunkte einbezogen: Anhand von mehrfaktoriellen ANOVAs mit Messwiederholung wurde die Entwicklung der einzelnen Emotionen beim Schriftspracherwerb über das gesamte Schuljahr hinweg untersucht. Dabei wurden die Schüler:innen, die an der Intervention teilnahmen (73,8% der Grundschüler:innen), unter Einbezug von als Kovariaten mit den Schüler:innen der Kontrollgruppe verglichen.

Ergebnisse und ihre Bedeutung

Die Ergebnisse von Teilstudie 1 zeigen nicht nur, dass sich die vier Emotionen beim Schriftspracherwerb angemessen und getrennt abbilden lassen anhand des AEQW-ES-Fragebogen (EFA-Modell: $\chi^2(62) = 79.04$, RMSEA = 0.046, SRMR = 0.023, CFI = 0.99; CFA-Modell: $\chi^2(98) = 191.87$, RMSEA = 0.085, SRMR = 0.023, CFI = 0.94); die Skalen zeigen gute interne Konsistenzen ($.81 \leq \omega \leq .92$). Im Mehrebenen-Pfadmodell (CFI = 0.97, TLI = 0.95, RMSEA = 0.050, $p < .001$, SRMR = 0.055) zeigt sich, dass die Emotion Ärger zu t1 alle schriftsprachlichen Kompetenzen der Kinder zu t2 negativ vorhersagt ($\beta = -0.18$). Die Anzahl der Bücher zuhause hängt positiv mit den schriftsprachlichen Kompetenzen zusammen ($\beta = 0.13$, $p < .01$). Die Ergebnisse von Teilstudie 2 werden aktuell noch ausgewertet und auf der Konferenz präsentiert (Daten von t3 liegen bereits im Paper-Pen-Format vor, müssen noch in die Datentabelle übertragen werden). Die Ergebnisse aus Teilstudie 1 weisen darauf hin, dass insbesondere die Emotion Ärger der Schüler:innen beim Schriftspracherwerb berücksichtigt werden sollte, um die schriftsprachlichen Kompetenzen im Grundschulalter zu fördern.

Inwiefern trägt die Unterstützung durch Lehrkräfte und Peers zur Entwicklung emotionaler und verhaltensbezogener Schwierigkeiten von Sekundarschüler:innen über ein Schuljahr bei? Ein latentes Veränderungsmodell

Frances Hoferichter¹, Marie-Christine Vierbuchen², Juliane Schlesier³, Diana Raufelder¹

¹Universität Greifswald, ²Europa-Universität Flensburg, ³Carl von Ossietzky Universität Oldenburg

Schule ist neben der akademischen Entwicklung ein bedeutsamer Ort, um die sozio-emotionale Entwicklung im Zusammenhang mit dem Verhalten von Schüler:innen (SuS) zu stärken (z.B. Brinckman, Kwon, Azen & Lease, 2024). Die Interaktion zwischen kognitiven, emotionalen und sozialen Kompetenzen mit der (Lern-) Umgebung findet sich im CASEL-Modell (Collaborative for Academic, Social, and Emotional Learning, 2021), welches hier als theoretische Grundlage dient. Unter anderem ist für die sozio-emotionale Entwicklung maßgeblich, wie SuS sich durch Lehrkräfte und Peers eingebunden und unterstützt fühlen. Die Beziehung zwischen Lehrkraft und SuS, die wahrgenommene Unterstützung durch Lehrkräfte oder Peers sowie Lob werden bislang häufig in der Primarstufe oder Vorschule erforscht, jedoch weniger bei Sekundarschüler:innen. Insbesondere der prädiktive Einfluss der Lehrkraft auf die Entwicklung von emotionalen und Verhaltensschwierigkeiten (Hamre & Pianta, 2001; O'Connor, 2010; Floress et al., 2024), sowie die hohe Qualität der Interaktionen im Klassenraum kann Entwicklungsrisiken wie emotionale Probleme (z.B. Buyse et al., 2008) oder Aufmerksamkeitsprobleme (Hamre & Pianta, 2005; Rudasill et al., 2016) abmildern.

Ziel der Studie

Die vorliegende Untersuchung hat das Ziel, den Forschungsstand für die Kohorte von Sekundarschüler:innen zu erweitern, indem der Einfluss der wahrgenommenen Unterstützung durch Gleichaltrige, durch die Lehrkraft und die instruktionale Unterstützung im Unterricht sowie Lob auf die Veränderung von Anpassungsproblemen, emotionalen Problemen, Hyperaktivität und Problemen mit Gleichaltrigen über den Verlauf eines Schuljahres untersucht werden. Dabei wird das Geschlecht, der sonderpädagogische Unterstützungsbedarf und der Schultyp berücksichtigt.

Methodik

Dafür wurde eine Stichprobe aus 733 SuS der siebten und achten Klassenstufe aus M-V (zu T1: MAlter = 13.78, SD = 0.67, 52% weiblich) aus 11 randomisiert gezogenen Schulen im Herbst 2018 und Sommer 2019 erhoben und ausgewertet. Ministerielle und Datenschutzgenehmigungen liegen vor. Erhebungsinstrumente sind der „Strength and Difficulties Questionnaire“ (SDQ, Goodman, 1997, Selbstauskunftsversion für Jugendliche), die Skalen „Unterstützung durch Gleichaltrige und Unterstützung durch Lehrkräfte“ (Torsheim et al., 2010) sowie die Subskalen „Instruktionale Unterstützung durch die Lehrkraft“ und „Lob“ aus dem Fragebogen „Lernunterstützung durch die Lehrkraft“ (Rakoczy et al., 2005; Hoferichter & Raufelder, 2022). Kontrolliert wurden das Geschlecht, der sonderpädagogische Unterstützungsbedarf sowie die Schulform.

Die Forschungsfragen wurden mittels Latent Change Models (LCM) in Mplus ausgewertet. Dabei ermöglicht LCM die Bewertung von Veränderungen innerhalb von Individuen durch die Nutzung von zwei Messzeitpunkten einer latenten Variablen (Steyer et al., 1997).

Ergebnisse und ihre Bedeutung

Es ergibt sich ein komplexes Bild, welches teilweise den bisherigen Forschungsergebnissen widerspricht und sie in anderen Bereichen bestärkt: Die Ergebnisse zeigen, dass die Veränderung der Probleme mit Gleichaltrigen von T1 auf T2 durch die Unterstützung der Lehrkraft vorhergesagt wird. Zudem hängt die wahrgenommene Unterstützung durch Gleichaltrige (T1) mit geringen Anpassungsproblemen, emotionalen Problemen und Problemen mit Gleichaltrigen (T1) zusammen, während die Unterstützung durch Lehrkräfte (T1) mit weniger Hyperaktivität und mehr Problemen mit Gleichaltrigen (T1) zusammenhängt. Die instruktionale Unterstützung im Unterricht (T1) hängt mit mehr Anpassungsproblemen (T1) zusammen. Lob durch die Lehrkraft (T1) hängt mit weniger emotionalen Problemen und Hyperaktivität (T1) zusammen. SuS mit sonderpädagogischem Unterstützungsbedarf berichten eher Anpassungsprobleme, emotionale Probleme und Probleme mit Gleichaltrigen. Mädchen berichten im Vergleich zu Jungen höhere emotionale Probleme, wohingegen SuS der Regionalschule im Vergleich zu Gymnasien eher Hyperaktivität berichten.

Die Bedeutung dieser Untersuchung ist hoch, da sie für die Sekundarstufe komplexe Zusammenhänge bestätigt. Aspekte der wahrgenommenen Unterstützung, Unterrichtsqualität wie instruktionale Unterstützung sowie Lob im Unterricht können von der Lehrkraft aktiv gestaltet werden und hängen mit emotionalen und Verhaltensschwierigkeiten von SuS zusammen. Auch die positive Rolle der Unterstützung durch Peers im Hinblick auf Verhaltensschwierigkeiten wurde deutlich. Konkrete Konsequenzen für den Unterricht und die Lehrkräftebildung werden im Vortrag thematisiert.

S-13: To KI or not to KI? Potentiale maschinellen Lernens, vortrainierter Deep Learning-Modelle und natürlicher Sprachverarbeitung für formatives Assessment

Zeit: Montag, 27.01.2025: 13:00 - 14:45 · Ort: M 003

Chair der Sitzung: Sonja Hahn

Chair der Sitzung: Fabian Zehner

Chair der Sitzung: Carolin Hahnel

To KI or not to KI? Potentiale maschinellen Lernens, vortrainierter Deep Learning-Modelle und natürlicher Sprachverarbeitung für formatives Assessment

Chair(s): **Sonja Hahn** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt, Deutschland), **Fabian Zehner** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt, Deutschland)

Diskutant*in(nen): **Carolin Hahnel** (Ruhr-Universität Bochum)

Technologische Innovationen sollten in der Bildungspraxis dort Einzug erhalten, wo sie mithilfe entsprechender Werkzeuge bedeutsame Aspekte des Lehr-/Lernprozesses unterstützen. Formatives Assessment etwa kann Lehren und Lernen evidenzbasiert steuern, indem es individuelle Leistungen oder Fähigkeitsprofile erfasst, rückmeldet oder zu gezielter Förderung führt (Schildkamp et al., 2020; Souvignier & Hasselhorn, 2018). Solche bedeutsamen Interaktionen im Lehr-/Lernprozess enthalten auch immer wieder sogenannte unstrukturierte Informationen (z. B. Sprache aus offenen Antwortformaten), deren effiziente Verarbeitung dem Einsatz moderner Methoden bedarf, wie etwa maschinellen Lernens, vortrainierter Deep Learning-Modelle oder natürlicher Sprachverarbeitung. Diese Methoden werden aktuell mit künstlicher Intelligenz (KI) in Verbindung gebracht werden und ermöglichen, Informationen und Daten verschiedenster Formate automatisiert zu erfassen, auszuwerten und so den Unterricht oder andere Lehr-Lern-Szenarien flächendeckend an die Lernenden anzupassen (Gardner et al., 2021).

Ziel des Symposiums ist es, konkrete Methodenentwicklungen und -anwendungen im formativen Assessment vorzustellen und zu reflektieren. Das Symposium zeigt außerdem deren Potential über Anwendungsfelder hinweg auf (u. a. in unterschiedlichen Fächern). Gemäß dem Motto der diesjährigen Tagung umfasst das Symposium Beiträge verschiedener Disziplinen (Fachdidaktik, Linguistik, Psychologie) sowie Fachinhalte (Physik, Mathematik, Englisch). Die Beiträge decken zudem eine Vielfalt in Hinblick auf Entwicklungsstadien, Anwendungs- bzw. Forschungsorientierung sowie die Art der erhobenen Daten (visuell vs. sprachgebunden) ab.

So stellt der erste Beitrag die Entwicklung eines Maschinenlernmodells vor, das Lehrkräfte darin unterstützt Concept Maps im Physikunterricht zielorientiert auszuwerten. Concept Maps stellen die mentale Repräsentation eines Lernergebnisses dar, das unstrukturierte Informationen (Begriffe) strukturiert in Beziehung setzt. Der zweite Beitrag hingegen behandelt unstrukturierte – nämlich visuelle – Informationen in Form von Lernenden erstellter Skizzen aus dem Bereich Mathematik und fokussiert auf den Vergleich verschiedener Auswertungsmethoden. Der dritte Beitrag nimmt eine fachunabhängige Perspektive ein und präsentiert ein Werkzeug zur Lerngruppenmobilisierung im Unterrichtsdiskurs, das es ermöglicht, alle Lernenden im Unterricht auf offene Fragen antworten zu lassen und der Lehrkraft semantisch gruppiert die wichtigsten Antworten anzuzeigen. Der letzte Beitrag behandelt ebenfalls natürlichsprachliche Daten, allerdings zusätzlich neben textueller auch in gesprochener Form. Er berichtet von drei Werkzeugen zum Einsatz im Fremdsprachenunterricht und illustriert eine Bandbreite von Einsatzzwecken für welche die Modelle genutzt werden können (z.B. automatische Erzeugung von Iteminhalten, dialogische Interaktionen).

Die Beiträge stellen eine Vielfalt von Potentialen moderner Technologien für formative Assessments vor und reflektieren deren Limitationen. Diese Potentiale und Limitationen werden in der abschließenden Diskussion von Prof. Dr. Carolin Hahnel aufgegriffen und darüber hinaus im Speziellen der Terminologiegebrauch in der wissenschaftlichen Kommunikation rund um KI entsprechend einer laufenden Debatte reflektiert (vgl. Zehner & Hahnel, 2023).

Beiträge des Symposiums

Machine-Learning-basierte Auswertung von Concept Maps: Chancen und Herausforderungen für formative Assessments im Schulalltag

Tom Bleckmann, Gunnar Friege

Leibniz Universität Hannover, Institut für Didaktik der Mathematik und Physik, AG Physikdidaktik

Concept Maps, im deutschsprachigen Raum oft auch Begriffsnetze genannt, wurden als grafisches Hilfsmittel zur Visualisierung von Wissen entwickelt (Novak & Cañas, 2008). Dabei besteht eine Concept Map aus verschiedenen Begriffen zu einem bestimmten Thema, die durch beschriftete Pfeile miteinander verbunden sind. Auf diese Weise können die Begriffe miteinander in Beziehung gesetzt werden und Zusammenhänge verdeutlicht werden. In einer Concept Map wird die Kombination aus erstem Begriff, beschrifteten Pfeil und zweiten Begriff auch als Proposition bezeichnet (Stracke, 2004).

Concept Maps eignen sich als Diagnoseinstrument, da sie flexibel im Unterricht eingesetzt werden können (Buldu & Buldu, 2010). Mit Concept Maps können unter anderem die Vorstellungen der Lernenden zu einem bestimmten Thema analysiert, Missverständnisse identifiziert und Einblicke in das Zusammenhangswissen der Lernenden gewonnen werden (Ley, 2015). Sie liefern somit wertvolle Informationen, die für ein formatives Assessment genutzt werden können (Ruiz-Primo & Shavelson, 1996).

Für ein formatives Assessment ist jedoch eine detaillierte Auswertung der Concept Maps notwendig. Dabei haben quantitative Auswertungsstrategien, wie z. B. das Zählen der korrekt verwendeten Begriffe, nur eine geringe Aussagekraft über den Inhalt einer Concept Map (Ryssel, 2018). Um einen detaillierten Einblick in den aktuellen Wissensstand der Lernenden zu erhalten, bieten sich eher qualitative Ansätze wie die Analyse einzelner Propositionen an (Bak Kibar et al., 2013). Im Gegensatz zu quantitativen Methoden ist dieser Evaluationsansatz jedoch sehr zeitintensiv, da jede einzelne Lernenden-Proposition analysiert werden muss (Stracke, 2004). Dieser hohe Zeitfaktor erschwert den Einsatz von Concept Maps für qualitativ hochwertiges formatives Assessment im regulären Schulalltag (Hartmeyer et al., 2018).

In diesem Projekt wird daher ein Machine-Learning-Modell vorgestellt, das sowohl einzelne Propositionen einer Concept Map automatisch auswertet, als auch Lernende und Lehrende im Lehr-Lern-Prozess mit Informationen versorgt, die im Rahmen eines formativen Assessments genutzt werden können.

Dafür wurde eine Concept Map zum Thema Mechanik entwickelt und die einzelnen Propositionen wurden mithilfe einer vierstufigen Skala analysiert. Die entwickelte Concept Map wurde in 14 Physikklassen der 11. Jahrgangsstufe in Niedersachsen eingesetzt und von 230 Lernenden bearbeitet. Die gesammelten Daten dienten als Grundlage, um mehrere Machine-Learning-

Modelle zu trainieren, die eine schnelle, detaillierte und einfache Auswertung für den regulären Physikunterricht ermöglichen sollen.

Die Ergebnisse zeigten, dass eine zufriedenstellende Übereinstimmung zwischen menschlicher und maschineller Bewertung erreicht werden konnte (Cohen's Kappa $\kappa = 0,75$; F1-Score = 0,82). Darüber hinaus konnte gezeigt werden, dass weder das Geschlecht noch die schulischen Leistungen einen Einfluss auf die automatische Bewertung des Machine-Learning-Modells hatten. Es wurde hingegen festgestellt, dass Propositionen, die Formeln oder Zahlen enthalten, die Leistung des Modells negativ beeinflussen.

In einer anschließenden Interviewstudie mit 8 Physiklehrkräften konnte gezeigt werden, dass die automatische Auswertung der Concept Map von den Lehrkräften generell als nützlich empfunden wurde. Die Lehrkräfte nutzten ausschließlich die bereitgestellten Informationen auf Klassenebene, jedoch nicht das individuelle Feedback für einzelne Lernende. Die Lehrkräfte offenbarten außerdem Unsicherheiten bei der Nachvollziehbarkeit und Transparenz der Machine-Learning-Auswertung. Da Machine-Learning-Modelle selten vollständig transparent sind, ist eine fundierte Interpretation der Ergebnisse ohne technische Kompetenz kaum möglich. Dies verdeutlicht den Bedarf an gezielten Fortbildungen für Lehrkräfte, um deren Kompetenz im Umgang mit Machine-Learning-Modellen und deren Ergebnisse zu fördern. Dadurch wird es ihnen ermöglicht, auch bei Nachfragen der Lernenden adäquate Antworten zu geben.

Insgesamt verdeutlichen die Ergebnisse das Potenzial automatisierter Auswertungen für das formative Assessment im regulären Schulalltag. Es konnten jedoch auch Schwächen aufgezeigt werden, die für die Zukunft kritisch diskutiert und reflektiert werden müssen. Gleichzeitig wurde aber auch deutlich, dass die Nachvollziehbarkeit von KI-basierten Modellen ein Kernaspekt für den Einsatz in sensiblen Bereichen wie der Bildung ist.

Constructed responses beyond NLP – Auswertungsansätze für graphische Antworten

Sonia Hahn¹, Leon Hammerla¹, Corinna Hankeln², Sebastian Gross², Christina Röper¹, Ulf Kroehne¹

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Fakultät für Mathematik, Technische Universität Dortmund

Für Constructed Responses, also von Lernenden produzierte Antworten auf offene Fragen, werden vielfältige Lern- und Denkprozesse sichtbar. Hier existieren diverse Möglichkeiten zur automatisierten Auswertung mittels Natural Language Processing Ansätzen (NLP; Magliano & Graesser, 2012). Während diese bereits vielfach Eingang in Praktiken des formativen Assessments gefunden haben (s. weitere Beiträge in diesem Symposium), fehlt dies für graphische Antworten (z.B. durch Lernende erstellte Skizzen) bisher (Gao et al., 2024). Dabei weisen viele Fachdisziplinen insbesondere im MINT-Bereich auf den Wert graphischer Repräsentationen hin (Smith et al., 2019), welche z.B. in der Mathematikdidaktik in unterschiedlichen Themenbereichen eingesetzt werden und neben einem didaktischen auch einen diagnostischen Mehrwert aufweisen (Bräuer et al., 2021; Prediger & Wessel, 2012), der unter anderem in formativen Settings lernförderliche Einsichten liefern kann.

Während in vielen Anwendungsbereichen (z.B. Medizindiagnostik) Verfahren des maschinellen Lernens zur computerbasierten Bildverarbeitung bereits erfolgreich zur Diagnostik eingesetzt werden, gibt es erst wenige entsprechende Ansätze aus dem Bildungsbereich (cf. Von Davier et al., 2023; Tyack et al., 2024; Zhai et al., 2022). Diese beruhen bisher zumeist auf dem Training spezifischer neuronaler Netze anhand großer Datenmengen (z.B. große Datensätze aus Large-Scale Assessments), die jedoch nicht immer bzw. bei neuen Items noch nicht zur Verfügung stehen. Zur Auswertung graphischer Constructed Responses steht eine Bandbreite von Ansätzen zur Verfügung: Klassische Computer-Vision nutzt verschiedene Algorithmen zur Bildverarbeitung, um einzelne Merkmale der Bilder zu erfassen. Die entsprechenden Algorithmen werden anhand der Bilder und zu ermittelnden Merkmale ausgewählt, so dass im Anschluss regelgeleitet Bilder automatisiert ausgewertet werden und kein maschinelles Lernen im engeren Sinne zum Einsatz kommt. Multimodale Modelle sind in der Lage, für Bilder entsprechende Encodings zu erstellen. Darauf können weitere Schritte aufbauen. Beispielsweise kann unter Verwendung geeigneter Prompts eine Einordnung der Bilder vorgenommen werden (Bewersdorff et al., 2024). Dieser Ansatz kann durch zusätzliche Beispielbilder ergänzt werden, um die Auswertung zu verbessern (Few-Shot Prompting; Latif & Zhai, 2024). Schließlich stellt das oben genannte Finetuning neuronaler Netze anhand vieler Beispielbilder, die für eine Aufgabe typisch sind, eine Methode dar, aufbauend auf größeren Datenmengen spezifische Modelle zu optimieren. Das Ziel dieses Beitrages ist es, die verschiedenen Methoden sowohl im prinzipiellen Ansatz als auch an Hand einer Beispielanwendung im Kontext eines formativen Assessments zu vergleichen.

In dieser Beispielanwendung kommen graphische Schülerantworten zum Einsatz. Lehrkräfte werden mit Methoden des maschinellen Lernens bei der Auswertung dieser Antworten unterstützt, um Ergebnisse sowie darauf aufbauende Hinweise zur Förderung der Lernenden schneller erhalten zu können. Im Bereich der graphischen Antworten fokussieren wir zunächst vier Aufgaben im Bereich Bruchrechnung. Anhand der vorliegenden Antworten (Anzahl der Antworten liegt zwischen 29 und 64, je nach Aufgabe) werden zunächst durch Personen mit fachdidaktischer Expertise diagnostisch relevante Aspekte der Schüler-Zeichnungen erarbeitet, wobei im Gegensatz zu einer reinen falsch/richtig-Einordnung eine detaillierte diagnostische Einordnung mit mehreren Endkategorien angestrebt wird, worauf eine spezifische Förderung aufbauen kann. Hierzu gehören u.a. die systematische Nutzung des Bruchstreifens und die Art und Weise der Einteilung (z.B. weisen unterschiedliche Einteilungen auf verschiedene mögliche Fehlvorstellungen zur Aussage der Bruchschreibweise hin) sowie der Markierung. In einem zweiten Schritt werden verschiedene Ansätze (klassische Computer-Vision Ansätze, multimodale Modelle) anhand des Expertenschemas optimiert (z.B. Anpassung von Algorithmen bzw. Prompts). Erste Ergebnisse weisen auf eine große Schwankung der Ergebnisse zwischen den Ansätzen hin. So werden zwischen 17% und 75% der Bilder den korrekten multiplen Kategorien zugeordnet. Der Einsatz weiterer Ansätze wie oben beschrieben ist in Arbeit.

Es ist geplant, auf der Konferenz aktualisierte Ergebnisse vorzustellen, welche eine größere Datengrundlage, weitere Auswertungsansätze und weitere Aufgabenstellungen umfassen. Zudem soll ein Einblick in den Entwicklungsprozess gegeben werden, indem vorgestellt wird, welche Wege sich bisher als nicht erfolgreich und welche Strategien sich als zielführend erwiesen haben.

ReCo-Live: Die gesamte Lerngruppe im Unterrichtsdiskurs mithilfe natürlicher Sprachverarbeitung mobilisieren

Nico Andersen¹, Franziska Baier-Mosch¹, Nora Ries¹, Tom Bleckmann², Gunnar Friege², Fabian Zehner¹

¹DIPF | Leibniz Institute for Research and Information in Education, ²Leibniz University Hannover

ReCo-Live (Automatic Text Response Coder) ist eine innovative App, die den Unterrichtsdiskurs auf alle Lernenden ausweiten kann, indem die offenen Antworten der Lernenden in Echtzeit und direkt im Klassenzimmer automatisch gruppiert werden. Die App ermöglicht es Lehrkräften, Dozierenden und Vortragenden, offene Fragen an ein Publikum beliebiger Größe zu stellen. Anders als im traditionellen Unterricht, bei dem nur eine Person auf die Frage antworten kann, wird so das Ausschließen der

Mehrheit vermieden und stattdessen die gesamte Lerngruppe mobilisiert. Im Gegensatz zu diesem Vorgehen sind herkömmliche Audience-Response-Systeme auf geschlossene Antwortformate beschränkt oder sind in ihrem Anwendungsspektrum eingeschränkt, da sie die Textantworten anhand exakter Übereinstimmung gruppieren (vgl. Kundisch et al., 2013). Durch die Ermöglichung und komplexerer Verarbeitung offener Antworten in ReCo-Live, werden alle Lernenden von Anfang an effektiv einbezogen (McLaughlin & Yan, 2017). Dies bietet das Potenzial, die kognitive Aktivierung (Kunter & Trautwein, 2013) in der gesamten Lerngruppe zu streuen und vielfältige Anwendungen des formativen Assessments zu ermöglichen, wie z. B. das Priming von Vorwissen zum Einstieg in ein Thema, das Schaffen eines Ausgangspunkts für Diskussionen oder One-Minute Papers gegen Ende des Unterrichts (z.B. Vonderwell & Boboc, 2013).

Technisch und methodisch funktioniert die App mithilfe einer Kombination aus webbasierten Technologien wie dem R-basierten Shiny, serverbasierter Kommunikation und Methoden der natürlichen Sprachverarbeitung. Die gesamte Lerngruppe oder das Publikum kann ihre freien Textantworten auf die Frage über ein mobiles Endgerät in einen Webbrowser eingeben. Die Antworten werden dann in Echtzeit an das Dashboard der Lehrkraft oder des Vortragenden übertragen und mithilfe natürlicher Sprachverarbeitung gemäß semantischer Ähnlichkeit gruppiert. Hierfür werden klassische Vorverarbeitungstechniken sowie moderne transformer-basierte vortrainierte Embeddings verwendet (z. B. XLM-RoBERTa; Conneau et al., 2020), um lexikalische Informationen zu extrahieren und die Semantik der Antworten darzustellen. Das Dashboard präsentiert mehrere lexikalische oder semantische Visualisierungen der Antworten mit unterschiedlichen Schwerpunkten, um verschiedenen Zwecken zu dienen. Als Herzstück zeigt es eine Übersicht über häufige Antwortgruppen und ihre Häufigkeiten (ähnlich zu Paiva et al., 2014) sowie Ausreißer an, zusammen mit interaktiven Diagrammen zur Visualisierung der Antworten im semantischen Raum. Im Falle der Anwendung im Unterricht, erhält die Lehrkraft alle Antworten, die in einige wenige Antworttypen gruppiert sind (Andersen & Zehner, 2021; Zehner, Sälzer & Goldhammer, 2016), um daraus potentiell sinnvolle pädagogische Maßnahmen ableiten zu können.

Aktuell läuft eine empirische Erprobung zur Evaluation, welche sowohl die Nutzungsweisen der Lehrkräfte als auch die Effektivität des App-Einsatzes untersucht. Dazu setzen Vortragende die App in drei Kontexten ein: im Schulunterricht, in der Hochschullehre oder in Vorträgen. Während ein Teil der Forschungsfragen direkt adressiert, welche Fragen gestellt werden und wie (hilfreich) anschließende Antworten gruppiert werden, werden die Studienteilnehmenden der ersten beiden Kontexte außerdem mit zusätzlichen etablierten Fragebogenskalen zur Effektivität des App-Einsatzes befragt (u. a. bzgl. der Gruppenmobilisierung, des Klassenmanagements und der kognitiven Aktivierung).

In der Präsentation werden wir die App demonstrieren, die zugrundeliegende Methodik beleuchten und erste empirische Befunde berichten.

KIBi: Integration von KI in die Bildung

Björn Rudzewitz¹, Sarah Löber¹, Julia Schmidt-Peterson², Marei Beukman¹

¹Hector-Institut für Empirische Bildungsforschung, LEAD Graduate School & Research Network, Universität Tübingen, ²Hector-Institut für Empirische Bildungsforschung, Gymnasium Neckartenzlingen, Universität Tübingen

Wir präsentieren drei von uns entwickelte, unterschiedliche KI-gestützte Systeme, welche KI (Künstliche Intelligenz) für formatives Assessment von Englisch als Fremdsprache integrieren. Neben der Kernfunktionalität beleuchten wir die Rolle der KI-Komponenten und berichten erste Praxiserfahrungen aus dem schulischen Kontext. Die Systeme illustrieren unterschiedliche Anwendungsmöglichkeiten und Rollen von KI in der Bildung (KIBi): als Dialogpartner, zur Vorbereitung von Leseverständnisfragen, als Korrekturhilfe, sowie als Generator für Test-Items.

ILAP (Intelligent Language Assessment Platform) ist eine Testplattform für Sprachkompetenztests mit Schwerpunkt Elicited Imitation (EI), eine Datenerhebungsmethode zur Einstufung der Sprachfertigkeit (Kostromitina & Plonsky, 2022) und formatives Assessment (Campfield, 2017). Testteilnehmer hören Sätze an und wiederholen diese originalgetreu. EI-Tests werden unterschiedlich realisiert, beispielsweise durch Manipulation der grammatikalischen Konstruktionen durch ungrammatische Varianten in einzelnen Items. Es wird dann beobachtet, ob, und gegebenenfalls wie, Testteilnehmer die ungrammatischen Varianten verbessern.

ILAP verwendet ein Large Language Model (LLM), um automatisch ungrammatische Test-Items auf Basis von wohlgeformten Sätzen zu erstellen. Dies ersetzt traditionelle, aufwendig zu entwickelnde Mal-Rule-basierte Verfahren. Testadministratoren können die generierten ungrammatischen Sätze selbst anpassen und den Output der KI überprüfen und editieren. Das System bietet zudem die Möglichkeit zur computergestützten Bewertung auf Basis automatisch transkribierter Antworten mit verschiedenen Distanzmaßen. Die vom System vorgeschlagene automatische Bewertung kann überschrieben werden.

Die Rolle der KI ist die automatische Generierung von ungrammatischen Test-Items sowie automatische Transkriptionen, was eine Entlastung und effizientes Assessment ermöglicht.

AISLA (Intelligenter Sprachassistent zum Englischlernen in Alltagssituationen) dient dem Üben von schriftlichen oder mündlichen Dialogen in Englisch als Fremdsprache. 100 Aufgaben bilden die Grammatikphänomene des Bildungsplans für Englisch 7. Klasse Gymnasium ab. In den Aufgaben übernimmt jeweils die KI die Rolle eines Dialogpartners. Lernende sollen das funktionale Ziel der Aufgabe unter Verwendung der sprachlichen Zielkonstrukte erreichen. Im Gegensatz zu generischen Sprachmodellen verfolgt AISLA einen taskbasierten Ansatz. Im Falle von fehlerhaften oder unerwarteten Antworten, d.h. Antworten, welche die vordefinierten Konversationslots des zugrunde liegenden Dialogsystems nicht ausfüllen, gibt das System bis zu drei Mal durch Lehrkräfte hinterlegtes, manuell erstelltes Feedback mit steigender Spezifität zu Inhalt und Grammatik. Ab dem vierten Versuch, der nicht zu einer Ausfüllung der Slots führt, bittet AISLA ein LLM mit der Aufgabenstellung, der Eingabe, einer Musterantwort sowie dem Prompt um Feedback, welches den Lernenden dann angezeigt wird.

Die KI wird bei AISLA für Transkriptionen und als unterstützender Dialogpartner bei wiederholt unerwarteten Eingaben eingesetzt.

ARES (Annotated Reading Enhancement System) ist eine Anwendung zur Leseunterstützung im englischen Fremdsprachenunterricht. Im System hochgeladene Texte werden automatisch linguistisch analysiert. Über 650 linguistische Konstruktionen aus dem English Grammar Profile (EGP, <https://www.englishprofile.org/english-grammar-profile/egp-online>) werden automatisch annotiert. Lehrkräfte können Aufgaben erstellen, indem sie Textverständnisfragen hinzufügen und Lernziele definieren. Letzteres dient dazu, die Anzeige der Konstruktionen auf den Lehrstoff anzupassen. Wenn z.B. Komparative behandelt werden, könnte die Lehrkraft Annotationen zu Zeitformen für diese spezielle Übung ausblenden. Beim Lesen der Texte öffnet ein Klick auf ein Konstrukt eine detaillierte Erklärung, inklusive z.B. Nutzungsbeispielen. Ein weiteres KI-Feature zur Unterstützung von Lehrkräften in ARES ist die Erstellung editierbarer Vorschläge für Textverständnisfragen. Die Fragen dienen dazu, die Lernenden dazu zu bewegen, sich mit den Texten und enthaltenen Konstruktionen auseinanderzusetzen. Lernende können die Fragen bearbeiten und abgeben und den Text hinsichtlich der Schwierigkeit und Interessanz bewerten. Lehrkräfte

können dann die KI automatisch die abgegebenen Freitextantworten vorkorrigieren lassen, wobei nur das Feedback, welches von Lehrkräften übernommen oder selbst erstellt wurde, angezeigt wird.

Die Rolle der KI im ARES-System ist eine automatische Annotation von linguistischen Konstrukten in Texten, der Vorschlag von Leseverständnisfragen unterschiedlicher Natur, sowie die Vorbereitung von Korrekturen zu Antworten auf Leseverständnisfragen ("teacher-in-the-loop").

Neben der technischen Komponente wird eine Englisch-Lehrkraft das AISLA- und ARES-System aus praktischer Sicht bewerten.

IS-02: Invited Symposium: Frühe Bildung und Ungleichheit

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 102

Chair der Sitzung: Guido Neidhöfer

Chair der Sitzung: Simon Wiederhold

Chair der Sitzung: Henning Hermes

Symposium „Frühe Bildung und Ungleichheit“

Chair(s): **Guido Neidhöfer** (ZEW Mannheim, Deutschland), **Simon Wiederhold** (IWH Halle)

Diskutant*in(nen): **Henning Hermes** (ifo Institut München)

Die Erforschung von Bildungsungleichheiten nimmt eine zentrale Rolle in der bildungsökonomischen Forschung ein. Insbesondere wurde betont, dass früh auftretende Ungleichheiten besonders starke Auswirkungen haben (vgl. u.a. Heckman und Mosso, 2014). Das Symposium „Frühe Bildung und Ungleichheit“ beleuchtet zentrale Aspekte und Mechanismen, die zur Reproduktion und Verstärkung von Ungleichheiten im Bildungssystem beitragen. Es vereint drei empirische Studien, die unterschiedliche Facetten der frühen Bildungsungleichheit untersuchen und bietet somit einen umfassenden Einblick in die Herausforderungen und Chancen im Kontext von Bildungsgerechtigkeit.

Der erste Beitrag von Christina Felfe (Universität Konstanz) untersucht, wie und wann Schüler:innen berufliche Aspirationen entwickeln, mit einem besonderen Fokus auf geschlechtsspezifische Unterschiede in den MINT-Berufen (Mathematik, Informatik, Naturwissenschaften, Technik). Die Studie zeigt, dass verstärkter elterlicher Druck die geschlechtsspezifische Kluft in den MINT-Aspirationen ab der sechsten Klasse signifikant vergrößert, wobei Mädchen ihre MINT-Aspirationen als Reaktion um 13,3 Prozentpunkte reduzieren. Diese Ergebnisse verdeutlichen den Einfluss externer Faktoren auf die Berufswahl von Jugendlichen und deren langfristige Auswirkungen auf Bildungs- und Berufslaufbahnen.

Pia Pinger (Universität zu Köln) beleuchtet im zweiten Beitrag die (Un-)Gleichheit der Bildungschancen beim Übergang von der Grundschule zur weiterführenden Schule. Anhand innovativer Datensätze aus Deutschland und Dänemark zeigt sie auf, dass Kinder aus sozial benachteiligten Familien trotz gleicher schulischer Leistungen deutlich seltener eine akademische Laufbahn einschlagen. Die Untersuchung der Entscheidungsfindung bei verschiedenen Gruppen offenbart, dass Schüler:innen die Bedeutung vergangener Leistungen für den zukünftigen Schulerfolg tendenziell unterschätzen. Ein niedrigschwelliges Mentoring-Programm zeigt, dass gezielte Unterstützung die Bildungswege von Kindern aus sozial schwächeren Familien signifikant verbessern kann.

C. Katharina Spieß (BiB und Johannes Gutenberg-Universität Mainz) analysiert im dritten Beitrag die Auswirkungen des Ausbaus der Kita-Infrastruktur auf frühe Bildungsungleichheiten. Trotz politischer Maßnahmen zur Verbesserung des Kita-Zugangs profitieren insbesondere Kinder aus sozioökonomisch besser gestellten Familien von qualitativ hochwertiger frühkindlicher Bildung und Betreuung. Die longitudinale Analyse zeigt, dass sich die Bildungsungleichheiten teilweise sogar verschärft haben, insbesondere im Hinblick auf die Nutzung ganztägiger Betreuungsangebote und die Kita-Qualität. Ein besonderer Fokus liegt auf Kindern mit Migrationshintergrund, deren Anteil an der Bevölkerung stark zunimmt und die besonders von diesen Ungleichheiten betroffen sind.

Gemeinsam bieten die Beiträge wertvolle Erkenntnisse darüber, wie strukturelle und individuelle Faktoren frühkindliche Bildungsungleichheiten beeinflussen. Die Diskussion von Henning Hermes (ifo Institut München) wird sich darauf konzentrieren, diese Ergebnisse zu synthetisieren und politische Implikationen abzuleiten, um die Bildungschancen für alle Kinder zu verbessern.

Beiträge des Symposiums

ENTFÄLLT Mintful career choices

Christina Felfe

Universität Konstanz

Wie und wann entwickeln Schüler berufliche Aspirationen? Wir untersuchen diese Frage mit einem Fokus auf die Rolle von Überzeugungen, Präferenzen (Entscheidungsgewichten) und elterlichem Druck bei 2.847 Schülern der sechsten bis neunten Klasse in Deutschlands nicht-akademischen Bildungswegen, wobei wir uns darauf konzentrieren, ob Mädchen und Jungen eine Karriere im MINT-Bereich (Mathematik, Informatik, Naturwissenschaften und Technik) anstreben oder nicht.

Unsere Ergebnisse zeigen, dass erhöhter elterlicher Druck den Gender-Gap bei MINT-Aspirationen in allen Klassenstufen erheblich vergrößert. Bereits in der sechsten Klasse verringern Mädchen ihre MINT-Aspirationen um 13,3 Prozentpunkte als Reaktion auf den erhöhten elterlichen Druck. Während die Entscheidungsgewichte zwischen Mädchen und Jungen nicht signifikant unterschiedlich sind, werden Überzeugungen zunehmend nach Geschlecht polarisiert, was zu einer Erhöhung des Gender-Gaps von weniger als 5 Prozentpunkten in der sechsten Klasse auf mehr als 15 Prozentpunkte in der neunten Klasse führt. Diese Ergebnisse deuten darauf hin, dass äußere Einflüsse wie elterlicher Druck berufliche Aspirationen formen können, noch bevor die Schüler genaue Vorstellungen über alternative Berufe entwickelt haben, und möglicherweise die anschließende Bildung von Überzeugungen und Aspirationen beeinflussen.

(Un-)Gleichheit der Chancen und der Übergang von der Grundschule zur weiterführenden Schule

Pia Pinger

Universität zu Köln

Frühe Bildungsweichenstellungen sind in vielen Bildungssystemen üblich. Sie führen oft zu Chancenungleichheit, wenn zwei Schüler mit identischen schulischen Leistungen aufgrund des unterschiedlichen sozioökonomischen Status (SES) ihrer Eltern unterschiedlich qualifizierte Schulen besuchen. Diese Präsentation nutzt zwei innovative Datensätze aus Deutschland und Dänemark, um neue Einsichten in den ungleichheitsfördernden Charakter des ersten Schulübergangs zu gewähren. Die Analyse gliedert sich in drei Schritte:

- 1) Es wird gezeigt, dass Kinder aus Familien mit niedrigem SES signifikant seltener eine akademische Laufbahn einschlagen und erfolgreich abschließen, auch wenn ihre vorherigen schulischen Leistungen berücksichtigt werden.
- 2) Unterschiede in den Entscheidungsprozessen verschiedener Gruppen werden beleuchtet. Hierbei werden Überzeugungen zu Leistungsunterschieden in weiterführenden Schulen auf Basis von Geschlecht, elterlichem Hintergrund und früheren Grundschulleistungen sowie die mentalen Modelle, die diese Überzeugungen stützen, analysiert. Die Ergebnisse zeigen, dass

Jugendliche zwar die Vorhersagekraft vergangener Leistungen für den Erfolg in weiterführenden Schulen qualitativ verstehen, deren Bedeutung jedoch quantitativ unterschätzen. Die Einschätzung der Rolle von früheren Leistungen, Geschlecht und elterlicher Bildung im Schulerfolg prognostiziert deutlich die Erfolgsaussichten Jugendlicher aus verschiedenen sozialen Schichten. Antworten aus offenen Umfragen legen die wahrgenommenen Mechanismen offen, die vergangene Leistungen, Geschlecht und SES mit zukünftigen Leistungen verknüpfen und diese Wahrnehmungen mit quantitativen Überzeugungen und Entscheidungen in Beziehung setzen.

3) Abschließend wird der Einfluss des sozioökonomischen Umfelds auf Bildungsentscheidungen bewertet. Kausale Belege aus einer Stichprobe deutscher Grundschulkinder zeigen, dass ein niedrigintensives Mentorenprogramm die langfristigen Bildungsergebnisse für Kinder aus benachteiligten Verhältnissen verbessern und die Chancengleichheit steigern kann. Kinder aus sozioökonomisch schwachen Familien, die zufällig für ein Jahr einem Mentor zugeteilt wurden, haben eine um 20% höhere Wahrscheinlichkeit, in die Oberstufe überzugehen. Die Mentorenbeziehung wirkt sich positiv auf Eltern und Kinder aus und führt zu langfristigen Vorteilen im Bildungsweg der Kinder.

Insgesamt bieten diese Ergebnisse neue Perspektiven auf die ungleichheitsverstärkende Natur des ersten Schulübergangs, dessen zugrundeliegenden Ursachen und mögliche Lösungsansätze.

Frühe Bildungsungleichheiten in Kita und Familie: Was hat sich die letzten Jahrzehnte mit dem Kita-Ausbau verändert?

C. Katharina Spieß

BiB und Johannes Gutenberg-Universität Mainz

Zahlreiche bildungsökonomische Studien belegen, dass insbesondere Kinder aus sozioökonomisch benachteiligten Familien von einer qualitativ guten Bildung und Betreuung in Kindertageseinrichtungen (Kita) profitieren. Qualitativ gute Bildung und Betreuung in einer Kita ist vor dem Hintergrund kompensatorischer Überlegungen relevant, da bereits in der Bildung und Betreuung in der Familie frühe Bildungsungleichheiten festzumachen sind. Empirische Studien belegen für Deutschland, dass Kitas jedoch insbesondere von Kindern sozioökonomisch besser gestellter Familien besucht werden. Dies trifft insbesondere für Kinder unter vier Jahren zu. Aus bildungsökonomischen Überlegungen führt dies zu Fehlallokationen. Mit dem Ausbau der Kita-Infrastruktur und der Erweiterung des Rechtsanspruchs bestand vielfach die Hoffnung, dass sich diese frühen Bildungsungleichheiten abbauen. Längsschnittliche Untersuchungen zeigen, dass dem nicht so war, sondern die Ungleichheiten bei bestimmten Gruppen sogar noch zugenommen haben. Darüber hinaus bestehen neben diesen Ungleichheiten am sogenannten „extensive margin“ auch bei der Nutzung ganztägiger Angebote, also am „intensive margin“, sowie der Kita-Qualität frühe Bildungsungleichheiten zu Lasten von Kindern aus sozioökonomisch schlechter gestellten Familien. Dabei sind zwischen soziodemografischen Gruppen weite Differenzen zu finden. All diese Befunde werden im Längsschnitt auf der Basis repräsentativer Daten für Deutschland dargestellt und diskutiert. Ein spezifischer Fokus wird dabei auf Kinder mit Migrationshintergrund gelegt, einer Gruppe deren Anteil an der Bevölkerung unter sechs Jahren stark zugenommen hat und auch weiterhin noch zunehmend wird. Darauf aufbauen werden Befunde für die Gründe dieser Ungleichheiten sowohl auf der Angebots- als auch der Nachfrageseite vorgestellt und bewertet. Ferner werden Wechselwirkungen zwischen der Kita-Qualität und der Bildungs- und Betreuungsqualität in der Familie diskutiert, die zu einer Verstärkung oder Abschwächung dieser Ungleichheiten beitragen können. Der Beitrag schließt mit Handlungsempfehlungen, die auf der Basis empirischer Forschungsergebnisse abgeleitet werden können.

OF-2: Podiumsdiskussion: Transparenz im Bildungssystem – welche Chancen bietet ein Bildungsverlaufsregister?

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 145

Chair der Sitzung: Kerstin Schneider

Podiumsdiskussion: Transparenz im Bildungssystem – welche Chancen bietet ein Bildungsverlaufsregister?

Kerstin Schneider

RatSWD, Bergische Universität Wuppertal

Aktuelle und qualitativ hochwertige Bildungsdaten sind essentiell sowohl für die Wissenschaft, um evidenzbasiert forschen und beraten zu können, aber auch für die Politik, um entsprechend zu steuern. Das Wissen über Bildungsverläufe und Bildungserträge ist in Deutschland im europäischen Vergleich jedoch gering, und dieser Mangel erschwert sowohl das Monitoring als auch die Steuerung des Bildungssystems. Bund und Länder planen nun ein Bildungsverlaufsregister, welches eine Besserung der Situation in Aussicht stellt. Wo genau stehen wir da und welche Herausforderungen zeichnen sich ab? Nur wenn die Wissenschaft Zugang zu den Daten des Bildungsverlaufsregisters bekommt, werden wichtige gesellschaftliche Fragestellungen beantwortet werden können und nur dann kann die Politik entsprechend beraten werden.

Seit 2004 setzt sich der RatSWD – ein unabhängiger Beirat der Bundesregierung – für eine verbesserte Forschungsdateninfrastruktur in Deutschland ein. Das Thema Bildungsverlaufsregister hat der RatSWD zuletzt in einem Positionspapier ausführlich diskutiert. Ziel dieser Podiumsdiskussion ist es nun, das Potenzial eines Bildungsregisters für Wissenschaft und Politik aus verschiedenen Positionen zu beleuchten. Dabei soll über mögliche Chancen, aber auch über offene Punkte und Herausforderungen.

Chair: Kerstin Schneider [Stellv. Vorsitzende des RatSWD, Bergische Universität Wuppertal]

Diskussionsteilnehmende:

- Alexandra Blanke, BMBF
- Prof. Dr. Corinna Kleinert, IIfBi
- Dr. Susanne Kollmann, IQHB

--

Der Rat für Sozial- und Wirtschaftsdaten (RatSWD) berät seit 2004 die Bundesregierung und die Regierungen der Länder in Fragen der Forschungsdateninfrastruktur für die empirischen Sozial-, Verhaltens- und Wirtschaftswissenschaften. Der RatSWD ist Teil des Konsortiums für die Sozial-, Verhaltens-, Bildungs- und Wirtschaftswissenschaften (KonsortSWD) in der Nationalen Forschungsdateninfrastruktur (NFDI).

Er versteht sich als institutionalisiertes Forum des Dialoges zwischen Wissenschaft und Datenproduzenten und erarbeitet Empfehlungen und Stellungnahmen. Dabei engagiert er sich für eine Infrastruktur, die der Wissenschaft einen breiten, flexiblen und sicheren Datenzugang ermöglicht. Diese Daten werden von staatlichen, wissenschaftsgetragenen und privatwirtschaftlichen Akteuren bereitgestellt. Derzeit hat der RatSWD 41 Forschungsdatenzentren (Stand: Dezember 2023) akkreditiert und fördert deren Kooperation.

PS-11: Akademische Selbstwirksamkeit und Bildungserfahrungen von Schüler:innen mit Migrationshintergrund

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: SO 322

Chair der Sitzung: Claudia Diehl

Individuelle, gruppenspezifische und systembezogene akademische Selbstwirksamkeit: Unterschiede nach sozialer Herkunft und Migrationshintergrund

Max Philipp Jansen, Birgit Becker, Susanne Garritzmann, Sigrid Roßteutscher

Goethe Universität Frankfurt, Deutschland

Die akademische Selbstwirksamkeit von Schüler*innen hat sich als ein zentraler Einflussfaktor für deren Lernerfolge und Bildungsergebnisse erwiesen. Vor dem Hintergrund, dass Bildungsungleichheiten nach Sozioökonomischem Status (SES) und Migrationshintergrund trotz zahlreicher Bemühungen weiter ein bekanntes Phänomen im deutschen Bildungswesen darstellen, stellt die Stärkung der akademischen Selbstwirksamkeit daher einen Schwerpunkt von Unterrichtskonzepten und gezielten Interventionen dar. Dabei wird bisher überwiegend die individuelle Selbstwirksamkeit adressiert. Für ein tiefergehendes Verständnis von Bildungsungleichheiten scheint es daher sinnvoll, weitere Dimensionen der Selbstwirksamkeit – jenseits der individuellen – zu betrachten. Um gruppenspezifische Unterschiede nach SES und Migrationshintergrund zu untersuchen, kombinieren wir eine Perspektive der soziologischen Ungleichheitsforschung mit dem sozialpsychologischen Konzept der akademischen Selbstwirksamkeit.

Selbstwirksamkeitsüberzeugungen beziehen sich auf den Glauben daran, bestimmte Ziele erreichen zu können (Bandura, 1977; Bandura et al., 2001). Die akademische Selbstwirksamkeit beschreibt dabei die Überzeugung von Schüler*innen, Bildungsaufgaben bewältigen und Bildungsziele erreichen zu können (Schunk und DiBenedetto 2022). Bisher wird akademische Selbstwirksamkeit jedoch fast ausschließlich auf individueller Ebene betrachtet („Was kann ich als Individuum akademisch erreichen?“; z.B. Pajares 1996; Honicke and Broadbent, 2016). Wir schlagen daher eine konzeptionelle Erweiterung um system- und gruppenspezifische Selbstwirksamkeitsüberzeugungen vor („Reagiert das System auf die Bedürfnisse der Schüler*innen?“ und „Wird meine Gruppe vom System fair behandelt?“) da wir davon ausgehen, dass wahrgenommene Gruppenvorteile (z. B. Solidarität) und -nachteile (z. B. Stigmatisierung, Diskriminierung) entscheidend für die Bildung gruppenspezifischer akademischer Selbstwirksamkeitsüberzeugungen und damit für einen erfolgreichen Bildungsverlauf sind (Lane, 1965; Bandura, 1982). Obwohl sich diese Dimensionen in Studien zu Diskriminierung oder relativer Deprivation als zentral erwiesen haben (z. B. Jetten et al., 2001; Tropp & Wright, 1999), wurde sie in Untersuchungen zu akademischer Selbstwirksamkeit bislang nicht ausreichend beachtet. Angesichts anhaltender Ungleichheiten im deutschen Bildungssektor gehen wir hingegen von einer hohen Relevanz gruppen- und systembezogener akademischer Selbstwirksamkeitsüberzeugungen aus und erwarten daher Unterschiede in individueller, gruppen- und systembezogener akademischer Selbstwirksamkeit je nach SES und Migrationsbiografie der Schüler*innen.

Zu diesem Zweck gehen wir zwei Forschungsfragen nach: (1) Lassen sich die vorgeschlagenen Dimensionen individueller, gruppenbezogener und systemischer akademischer Selbstwirksamkeit empirisch nachweisen? (2) Inwiefern zeigen sich dabei gruppenbezogene Disparitäten nach sozialer Herkunft und Migrationshintergrund?

Zur Beantwortung dieser Fragen nutzen wir Daten von 3.411 Jugendlichen aus ganz Deutschland, die über eine registerbasierte Stichprobe rekrutiert wurden und zwischen Dezember 2023 und Februar 2024 an einer Online-Umfrage teilgenommen haben. Die Umfrage erfasste detaillierte Maße der akademischen Selbstwirksamkeit, differenziert nach individueller, gruppen- und systembezogener Selbstwirksamkeit. Wir verwenden deskriptive Methoden und lineare Regressionsmodelle mit Interaktionseffekten (SES, Migration), um die verschiedenen Dimensionen der akademischen Selbstwirksamkeit und Disparitäten nach Migrationshintergrund und SES (operationalisiert durch die Bildung der Eltern) zu untersuchen.

Unsere Ergebnisse zeigen, dass sich die akademische Selbstwirksamkeit von Schüler*innen je nach sozialer Herkunft und Migrationshintergrund in den Dimensionen Individuum, Gruppe und System unterscheidet. Schüler*innen mit hohem SES zeigen im Vergleich zu Schüler*innen mit niedrigem SES auf allen Ebenen eine höhere Selbstwirksamkeit, unabhängig von Schulleistung und -form. Dies unterstützt unsere Annahme, dass Eltern mit hohem SES ihren Kindern wichtige Impulse zur Ausbildung hoher Selbstwirksamkeitsüberzeugungen weitergeben, während Kinder mit niedrigem SES-Hintergrund von solchen Einflüssen nicht profitieren. Darüber hinaus zeigen sich unterdurchschnittliche Werte für Jugendliche mit türkischem, ehemals jugoslawischem, nah-/mittelöstlichem und afrikanischem Migrationshintergrund bei der gruppenspezifischen Selbstwirksamkeit sowie überdurchschnittliche Werte für Jugendliche aus der Ukraine. Diese Unterschiede können teilweise durch Diskriminierungserfahrungen erklärt werden, wodurch der überdurchschnittliche Wert der Ukraine noch verstärkt wird. Im Bereich der systemischen Selbstwirksamkeit fällt besonders der große Unterschied zwischen der Türkei (sehr niedrige Werte) und der Ukraine (sehr hohe Werte) auf, der ebenfalls teilweise durch Diskriminierungserfahrungen erklärt wird. Diese Erkenntnisse legen nahe, dass differenziert ausgerichtete Interventionen zur Stärkung der akademischen Selbstwirksamkeit besonders für Schüler*innen mit niedrigem SES und verschiedenen Migrationshintergründen deren Partizipationschancen fördern und so dazu beitragen können, bestehenden Bildungsungleichheiten entgegenzuwirken.

Einfluss sozialer und dimensionaler Vergleiche auf das fachspezifische Fähigkeitsselbstkonzept von Grundschulkindern mit und ohne Migrationshintergrund

María Valcárcel Jiménez, Madita Frühauf, Bettina Hannover

Freie Universität Berlin, Deutschland

Das Internal/External-Frame-of-Reference Modell (I/E-Modell, [1]) sowie der Big-Fish-Little-Pond-Effekt (BFLPE; [2]) erklären die Genese des Fähigkeitsselbstkonzept (SK). Das I/E-Modell beschreibt den Einfluss individueller sozialer und dimensionaler Vergleiche, während der BFLPE den Einfluss der Leistung der gesamten Klasse auf das SK von Schüler:innen beschreibt. Verschiedene Autor:innen [3, 4, 5, 6] haben das I/E-Modell und das BFLPE integriert, um Kontexteffekte und die Effekte sozialer und dimensionaler Vergleiche gleichzeitig zu berücksichtigen.

Während die Zusammenhänge zwischen individuellen Merkmalen (Geschlecht, Alter) der Schüler:innen und ihrem SK vielfach untersucht wurden [7,8], blieb u.W. ein Migrationshintergrund (MH) unbeachtet, obwohl ca. 43% der Grundschul:innen in Deutschland einen MH aufweisen [9]. Die Berücksichtigung des MH ist jedoch essenziell, denn verschiedene Studien zeigen, dass sich Schüler:innen mit MH in ihren Leistungen überschätzen [10] oder ein ähnliches SK wie ihre Peers ohne MH [11] trotz niedrigerer Leistungen aufweisen. Eine mögliche Erklärung ist, dass das SK von Schüler:innen mit MH eher ein Abbild ihrer

Wünsche (Wunschdenkhypothese, [12]) als ein Abbild ihrer tatsächlichen Leistung ist. Zudem könnten sich Schüler:innen mit MH weniger stark mit ihrer Gesamtklasse identifizieren und lediglich Mitschüler:innen mit MH als Referenzgruppe heranziehen.

In einem Modell, in dem die Grundannahmen des I/E-Modells und des BFLPE integriert sind, untersuchen wir, ob sich die Zusammenhänge für Schüler:innen mit und ohne MH unterscheiden. Wir nehmen an, dass bei Schüler:innen mit MH relativ zu Schüler:innen ohne MH (a) bei gleichen Leistungen ein stärkeres mathematisches (H1) und ein stärkeres verbales SK (H2) vorliegt (b) die mathematische Leistung in Mathematik ein schwächerer Prädiktor des Mathe-SK (H3) und die verbale Leistung ein schwächerer Prädiktor des verbalen SK (H4) ist und (c) der Einfluss der Klassenleistung auf das mathematische SK (H5) und das verbale SK (H6) schwächer ausgeprägt ist.

Insgesamt nahmen 988 Schüler:innen (50.5% weiblich, 49% MH [als Kleinkind kein Deutsch gelernt]) der Jahrgänge 4-6 aus 57 Klassen teil. Die Schüler:innen machten Angaben zu ihrem mathematischen und verbalen SK [13]. Die Leistung der Schüler:innen in Deutsch und Mathematik wurde mittels standardisierter Leistungstests [14,15] erfasst. Wir berechneten Multi-Level-Pfadanalysen für Schüler:innen im Gesamtmodell und getrennt für Schüler:innen mit und ohne MH.

Die Annahmen des integrierten Modells bewährten sich fast vollständig in der Gesamtstichprobe: Die individuelle Leistung hing positiv mit dem SK in der jeweils gleichen Domäne zusammen ($\beta_{\text{mathe}} = .377^{***}$, $\beta_{\text{deutsch}} = .457^{***}$), die individuelle verbale Leistung korrelierte negativ mit dem mathematischen SK ($\beta = -.086^*$), die individuelle mathematische Leistung korrelierte nicht mit dem verbalen SK. In beiden Domänen hingen Klassenleistung und SK negativ zusammen ($\beta_{\text{mathe}} = -.578^{***}$, $\beta_{\text{deutsch}} = -.710^{***}$).

Unsere Ergebnisse zeigen, dass Schüler:innen mit MH kein positiveres mathematisches SK (H1, $\beta = .040$), aber ein positiveres verbales SK (H2, Wunschdenkhypothese, $\beta = .078^*$) als vergleichbar kompetente Schüler:innen ohne MH hatten. Die Leistung war bei Schüler:innen mit MH im mathematischen Bereich kein signifikanter Prädiktor ihres SK (H3, $\beta_{\text{ohneMH}} = .445^{***}$, $\beta_{\text{mitMH}} = .274^*$), aber signifikant für den verbalen Bereich (H4, $\beta_{\text{ohneMH}} = .438^{***}$, $\beta_{\text{mitMH}} = .430^{***}$). Der Einfluss der Klassenleistung auf das SK war im mathematischen Bereich bei Schüler:innen mit MH tendenziell schwächer ausgeprägt (H5, Referenzgruppenhypothese, $\beta_{\text{ohneMH}} = -.758^{***}$, $\beta_{\text{mitMH}} = -.450^*$), nicht aber im verbalen Bereich (H6, $\beta_{\text{ohneMH}} = -.749^+$, $\beta_{\text{mitMH}} = -.815^{***}$). Einen Effekt dimensionaler Vergleiche zeigte sich ausschließlich für die verbale Leistung auf das mathematische SK bei Schüler:innen mit MH ($\beta_{\text{ohneMH}} = -.068$, $\beta_{\text{mitMH}} = -.094^*$).

Unsere Ergebnisse replizieren teilweise die Annahmen des integrierten Modells. Wir diskutieren, dass Schüler:innen mit MH ein optimistischeres verbales SK haben als ihre Peers ohne MH (Wunschdenkhypothese) und dass BFLPE in Mathematik für Schüler:innen mit MH geringer ausfällt als bei Schüler:innen ohne MH, was darüber erklärt wird, dass sich Schüler:innen mit MH die Gesamtklasse weniger als Referenzgruppe für ihre eigenen Leistungen heranziehen.

On the wrong track? Perceived track mismatch among ethnic minority and majority students in the German educational system

Claudia Diehl, Katja Pomianowicz, Thomas Hinz
Universität Konstanz, Deutschland

This study examines the perceptions of ethnic minority and majority students that they belong to a higher track than the one they currently attend. We analyze these perceptions based on a survey of 7th graders in Germany, a country with a highly stratified education system that sorts students into different tracks based on academic ability at an early age. While previous research has repeatedly shown that minority students are more likely to be placed in lower tracks, little is known about how students in general, and minority students in particular, perceive the fairness of their track placement. This is important because previous research has shown that perceptions of unfair treatment influence students' motivation, self-esteem and future educational and occupational outcomes.

Theoretically, the study examines two main mechanisms that might explain why minority students believe more often they should be in a higher educational track. The first mechanism, "exposure to unfair treatment," hypothesizes that they are more likely to feel unfairly treated because they are actually more likely to attend a track that is too low given their abilities. The second mechanism emphasizes the role of attributional processes. It focuses on people's tendency to attribute a lack of success to unfair treatment ("ex post rationalization of failure"). Because minority students are more likely than majority students to attend the lowest educational tracks and to have highly ambitious parents, they may feel a greater need to attribute their limited educational success to unfair treatment in order to protect their self-esteem.

We use data from the "Students' Perceptions of Inequality and Fairness (PerFair)" survey, a classroom survey conducted in three German states: Baden-Wuerttemberg, Saxony, and Schleswig-Holstein. To measure perceived track mismatch, students were asked whether they believed they should be in a higher track or not. To test the exposure mechanism, we use a proxy for actual track mismatch by calculating the deviation of students' ability from the average ability of other students attending the same track. The attribution mechanism is tested by analyzing whether students who attend the lowest educational track and those whose parents have high educational aspirations are more likely to feel unfairly treated in terms of the track they attend.

Based on two-level random intercept models with multiple imputation of missing values, our results indicate that minority students are significantly more likely than majority students to perceive a mismatch in their track placement. In terms of potential explanations for these differences, we find no support that the "exposure mechanism" plays an important role. Attributional processes ("ex-post rationalization of failure") provide a more plausible explanation. Being placed in the lowest educational track and having parents with high educational aspirations not only exert a strong influence on perceived track mismatch, but also explain large parts of the perception gap between majority and minority students.

In sum, our findings align with previous literature on "objective" ethnic educational disparities. These studies not only show that immigrant students have high educational aspirations. They also suggest that this may reflect both "immigrant optimism" and a lack of knowledge. This supports our finding that minority students' perceptions of unfair treatment are less grounded in reality. In particular, their sense of belonging to a higher educational track is independent of how their cognitive abilities relate to those of other students in their track. Rather, minority students and their parents have high aspirations. Previous evidence suggests that this does not necessarily put them on a path to success. Our findings add to this research by showing that it also increases students' subjective perceptions of being treated unfairly in school.

Angepasst, Gespalten, Entfremdet. Einfluss eines sozialen Aufstiegs durch universitäre Bildung auf habituelle Praxisformen von Bildungsuaufsteiger*innen

Nune Arazyan

Im Beitrag wird das Phänomen des sozialen Aufstiegs durch universitäre Bildung untersucht, mit einem besonderen Fokus auf die kulturellen und sozialen Anpassungen und Herausforderungen der Bildungsaufsteiger*innen, die mit einem solchen Aufstieg einhergehen. Der American Dream suggeriert, dass soziale Mobilität nur eine Frage des persönlichen Einsatzes ist. Doch biografische sowie wissenschaftliche Literatur zeigt, dass sie in der Realität durch viele leistungsunabhängige Faktoren erschwert wird.¹ Ziel dieser Arbeit ist es, die Mechanismen und Strategien zu beleuchten, die Bildungsaufsteiger*innen anwenden, um sich an das neue soziale Milieu anzupassen, sowie die damit verbundenen Veränderungen in ihren Denk- und Verhaltensmustern zu analysieren.

Der theoretische Rahmen der Arbeit basiert auf den Konzepten von Pierre Bourdieu (insbesondere auf den zu unterschiedlichen Kapitalformen² und dem Habitus³), Didier Eribon (insbesondere Klassenscham⁴), Judith Butler (insbesondere verbale Gewalt⁵), auf der Milieutheorie,⁶ auf theoretischen Ansätzen zur Bildungsungleichheit und Klassismus,⁷ sowie auf aktuellen empirischen Erkenntnissen zur Bildungsungleichheit und Bildungsaufstiegen.⁸ Die Milieutheorie bildet den Ausgangspunkt der Forschung, indem sie darauf hinweist, dass jedes soziale Milieu spezifische Werte, Wahrnehmungsweisen und kulturelle Praktiken entwickelt, die die sozialen Identitäten und Habitus der Individuen prägen. Die Kombination der zuvor dargelegten theoretischen Ansätze und bestehender empirischer Erkenntnisse bildet die Grundlage, mit der die Dynamiken sozialer Mobilität und die kulturellen Anpassungsprozesse bei Bildungsaufsteiger*innen verstanden werden können.

Im Mittelpunkt der Arbeit steht die Frage, wie sich für Bildungsaufsteiger*innen durch die Erfahrung des Aufstiegs folgende fünf Bereiche verändern, die ihre kulturelle Zugehörigkeit, ihre Identität und ihr Verhalten, sowie kulturelle Praktiken bestimmen: Erfahrungen verbaler und symbolischen Gewalt, Identitätskonstruktion und Entfremdung, Klassenscham, (ggf. gespaltenen) Habitus und Veränderungen im Sprachgebrauch. Zudem wird untersucht, welche neuen Herausforderungen nach dem vollzogenen Aufstieg entstehen.

Die empirische Basis bilden teilstrukturierte qualitative Interviews mit Personen aus benachteiligten sozialen Milieus, die in den letzten zehn Jahren ein Studium an einer deutschen Hochschule vollendet haben. Dabei wird jeder Frageblock der in der Arbeit untersuchten Aspekten gewidmet. Als methodisches Vorgehen wurde die qualitative Inhaltsanalyse nach Udo Kuckartz benutzt,⁹ wobei mit MAXQDA deduktiv kodiert wurde, als Hilfsmethode wurde die korpuslinguistische Analyse mit AntConc nach Noah Bubenhofer herangezogen.¹⁰ Das Kodieren ermöglichte eine systematische Erfassung und Analyse wiederkehrender Themen und Muster in den Aussagen der Befragten, während die korpuslinguistische Analyse detaillierte Einblicke in den Sprachgebrauch und die sprachlichen Anpassungsstrategien bot. Es hat sich gezeigt, dass eine interdisziplinäre Problemstellung, die neben kulturwissenschaftlicher Forschung auch soziologische und bildungswissenschaftliche Ansätze beachtet, sowie der Einsatz der technikunterstützten Methoden in dieser Arbeit unter anderem vorher noch nicht entdeckte Ergebnisse offenbart haben.

Die Ergebnisse zeigen, dass der soziale Aufstieg nicht nur mit der Überwindung äußerer Barrieren verbunden ist, sei es verbale und symbolische Unterdrückung seitens anderer Menschen auf dem gesamten Bildungsweg oder strukturelle Hindernisse, sondern auch erhebliche innere Anpassungen im Denken, Verhalten und milieuspezifischen kulturellen Praktiken von den Betroffenen erfordert. Bildungsaufsteiger*innen müssen milieuspezifische kulturelle Praktiken erlernen und sich an die neuen sozialen Erwartungen anpassen. Ihr Habitus bleibt tief in den Personen verankert und wird auch durch den sozialen Aufstieg und die dazugehörigen Anpassungsanforderungen nicht vollständig verändert wird. Dadurch entwickeln viele Aufsteiger:innen einen gespaltenen Habitus, der sowohl Elemente des Herkunfts- als auch des Aufstiegsmilieus umfasst. Diese Spaltung führt oft zu Identitätskonflikten und einem Gefühl der Entfremdung, sowohl vom ursprünglichen sozialen Milieu als auch vom neuen. Gleiches gilt für den Klassenscham: während die Betroffenen sich anfangs dafür schämen, was sie von dem Aufstiegsmilieu unterscheidet, kommt es vor, dass sie sich nach dem Bildungsaufstieg für die Unterschiede zum Herkunftsmilieu schämen, was die Entfremdung begünstigt. Die Verschiebung des Sprachgebrauchs in die Richtung der Bildungssprache betont die Unterschiede ebenso.

Die Erkenntnisse legen die Notwendigkeit nahe, Chancen auf sozialen Aufstieg zu verbessern und produktive gesellschaftliche Koexistenz zu fördern, und bieten Anregungen für die Gestaltung der entsprechenden bildungspolitischen Maßnahmen.

PS-12: Politische Bildung und Engagement: Perspektiven aus verschiedenen Bildungskontexten

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 242
Chair der Sitzung: Sabine Manzel

Wie Bildungswege politisches Interesse prägen: Ein Vergleich zwischen Deutschland und der Schweiz

Julia Mayr¹, Sara Möser²

¹Leibniz-Institut für Bildungsverläufe, Bamberg; ²Institut für Erziehungswissenschaften, Universität Bern

Politische Partizipation und politisches Interesse fungieren als zentrale Pfeiler des demokratischen Systems und tragen wesentlich zu dessen kontinuierlicher Festigung bei (van Deth & Elff 2004; Hadjar & Becker 2006). Aktuelle Befunde deuten jedoch auf eine sozial induzierte politische Ungleichheit innerhalb der Gesellschaft hin (Wessels 2021), die sich bereits im jungen Erwachsenenalter verfestigt. So ist neben dem Phänomen der politischen Ungleichheit auch die vermeintliche Politikverdrossenheit Heranwachsender ein medial vielfach thematisierter Sachverhalt, der sich durch eine hohe gesellschaftspolitische Relevanz auszeichnet.

Die Beteiligung am politischen Prozess ist anspruchsvoll und bedarf nach dem *civic voluntarism model* (Verba et al. 1995) sowie dem *absoluten Bildungsmodell* (Persson 2013) gewisser (Partizipation-)Kompetenzen, da der Zugang durch Präsuppositionen beschränkt ist. Neben jenen Fähigkeiten zählen auch das politische Interesse sowie die politische Selbstwirksamkeit zu den Ressourcen, die ursächlich für das politische Engagement erscheinen (Verba et al. 1995). Es ist eine zentrale Aufgabe des Bildungssystems, die für die politische Beteiligung erforderlichen Fertigkeiten und Kenntnisse frühzeitig zu fördern (Van de Werfhorst 2014). So bietet die Schule als formale Bildungsinstitution gemäß dem *ability-motivation-opportunity triad model* die Gelegenheit, (erstmalig) mit dem politischen System in Berührung zu kommen, vermittelt dabei eben jene zentralen (Partizipations-)Kompetenzen und regt darüber hinaus das Interesse gegenüber dem politischen Geschehen an (Delli Carpini & Keeter 1996). Aufgrund der Stratifizierung des Bildungssystems erfahren jedoch nicht alle Bürger:innen eines Landes politische Bildung im gleichen Ausmaß. Es ist davon auszugehen, dass im Sinne der *Humankapitaltheorie* (Becker 1993; Mincer 1974; Schultz 1961) während eines längeren Bildungswegs sowie im Rahmen einer allgemeinbildenden Ausbildung vermehrt Humankapital und damit auch in einem höheren Maße politische Kompetenzen akkumuliert werden, was mitunter zu einer stärkeren Ausprägung des politischen Interesses führt. Daraus ergibt sich die erste Forschungsfrage, inwiefern politisches Interesse nach Bildungsweg stratifiziert ist. Ziel dieses Beitrags ist es, eine mögliche Schichtung des politischen Interesses nach Bildungswegen, sowie nach Art und Umfang des dabei akkumulierten Bildungskapitals zu untersuchen. Obwohl solch eine detaillierte Betrachtung der Bildungswege zunehmend erforderlich erscheint, um den Zusammenhang adäquat erfassen zu können, wurden diese bislang nur unzureichend berücksichtigt (Lischewski et al. 2020).

In der zweiten Forschungsfrage werden die institutionellen Rahmenbedingungen politischer Sozialisation thematisiert. In einem Ländervergleich wird dabei das politische Interesse – unter Berücksichtigung der politischen Selbstwirksamkeit – von jungen Erwachsenen in Deutschland und der Schweiz gegenübergestellt. Um die systemspezifische Stratifizierung zu identifizieren, werden somit zwei Länder verglichen, die sich in ihren politischen Systemen (repräsentative vs. direkte Demokratie) und Bildungssystemen (u. a. in Bezug auf das Stratifizierungsniveau) unterscheiden. Dabei gilt es, die Hypothese zu prüfen, dass die politische Selbstwirksamkeit und das politische Interesse in der Schweiz höher ausfallen und weniger stark ausdifferenziert sind als in Deutschland auf Grund der stärkeren Repräsentation von Nicht-Akademiker:innen im Parlament sowie der Ausgestaltung der direkten Demokratie, welche verglichen mit der repräsentativen Demokratie Deutschlands vermehrt Beteiligungsmöglichkeiten bietet (Hadjar & Schlapbach 2009).

Die Analysen basieren auf Daten der DAB-Panelstudie (Determinanten der Ausbildungswahl und der Berufsbildungschancen) und der Startkohorte 3 des NEPS (National Education Panel Study). Die zentralen Informationen, u. a. zum politischen Interesse, wurden in der sechsten DAB-Welle und in der zehnten NEPS-Welle erhoben. Die Befragten sind zu diesem Zeitpunkt gerade ins Erwachsenenalter eingetreten und zwischen 18 und 20 Jahre alt. Die Panelstruktur der Daten ermöglicht es, den Bildungsweg der Befragten in der Sekundarstufe I und II detailliert zu operationalisieren.

Zur Beantwortung der Forschungsfragen werden in einem ersten Schritt deskriptive Befunde zum politischen Interesse beider Ländersamples nach Bildungsweg herangezogen. In einem zweiten Schritt werden basierend auf ordinalen logistischen Regressionsmodellen Average Marginal Effects (AMEs) berechnet, um den Einfluss von Bildung auf das politische Interesse und die politische Selbstwirksamkeit zu prüfen. Erste Ergebnisse zeigen in beiden Samples eine starke Stratifikation nach Bildungsweg.

Diskurs und Konfliktleben im Politikunterricht - Pilotstudie zu agonaler Pluralität im Klassenzimmer und in der Freizeit

Sabine Manzel

Universität Duisburg Essen, Deutschland

Plurale Gesellschaften zeichnen sich durch eine Vielzahl verschiedener Akteure aus, die „mit- und gegeneinander um gesellschaftliche, wirtschaftliche und politische Macht“ konkurrieren (Woyke 2021, 717). Unterschiedliche Einstellungen, Werte, Lebensweisen können zu Konflikten führen, wobei der „Rechtsstaat den Ordnungsrahmen und die Regeln für den Konfliktaustrag zur Verfügung“ stellt (ebd., 718). Das Aushandeln unterschiedlicher Interessen und die politische Kompromissuche sind konstitutiv für plurale Demokratien. Jüngst wird die agonale Polarisierung der Gesellschaft beklagt, in der Meinungen und Werte unversöhnlich aufeinanderprallen, Kompromissbereitschaft sinkt und herkömmliche Lösungsmechanismen nicht mehr greifen. Mau et al. (2023) zeigen entlang vier Ungleichheitsarenen für die deutsche Gesellschaft eine mehrschichtige Konfliktlage mit noch vorhandener Aushandlungsbereitschaft. Zum Erlernen von Toleranz und einer sog. Streitkultur bedarf es bereits in der Schule entsprechender Lehr-Lerngelegenheiten. Insbesondere der Politikunterricht macht Lernangebote auf der fachinhaltlichen Wissens-ebene (z.B. Fachkonzept Konflikt) und der Kompetenzdimension der politischen Handlungsfähigkeit (Argumentation, vgl. Detjen et al. 2012). Schüler*innen in NRW erleben ein offenes Unterrichtsklima mit der Möglichkeit, eigene Meinungen zu äußern und zu diskutieren (Hahn-Laudenberg & Abs 2024, 297). Die Frage, wie Schüler*innen agonale Pluralität wahrnehmen und wie der Politikunterricht damit umgeht, weist jedoch noch Lücken auf.

Die Pilotstudie zu agonaler Pluralität im Klassenzimmer untersucht explorativ, wie sich gesellschaftliche Auseinandersetzungen im Politikunterricht und privaten Bereich, inklusive Social-Media-Aktivitäten bei Jugendlichen zeigen. Die Typologie der Ungleichheitsarenen von Mau (s.o.) wird an Interessen der Schüler*innen angelegt, um Polarisierungsfelder auszuloten. Angenommen wird, dass die Arena „Heute-Morgen“ (Klima-Generationenkonflikt) das höchste Interesse weckt. Das

Diskussionsverhalten wird für Unterricht und Freizeitbereich anhand folgender Unterfragen genauer untersucht: Welche Gründe nennen Schüler*innen für Meinungsverschiedenheiten zu politischen Themen? Welche Auseinandersetzungsformen erleben sie bei Konflikten? Gibt es Unterschiede bei der Positionierung in Streitfragen hinsichtlich Gender, Migrationsgeschichte und Jahrgangsstufe? Wie nehmen Schüler*innen die Lehr-Lernangebote zu Diskursivität und Multiperspektivität wahr?

Für die explorative Pilotstudie wurde im Juni 2024 eine Gelegenheitsstichprobe von N=333 Schüler*innen der Sek. I und II in NRW mittels eines Fragebogens mit rund 50 geschlossenen Items (Likert-skaliert) befragt. Skalen umfassen u.a. politisches Interesse, Diskussionsverhalten, Kontroversität, Lernangebote. Der Fragebogen wurde aufgrund unterschiedlicher digitaler schulischer Infrastruktur sowohl im Paper-pencil-Format als auch digital über LimeSurvey im Politikunterricht angeboten. Die Auswertung erfolgte quantitativ über SPSS.

Einzelne Ergebnisse werden exemplarisch herausgegriffen, umfassende Analysen zur Tagung präsentiert. Bei den Ungleichheitsarenen überwiegt das Interesse an „Oben-Unten“ (wirtschaftlich-soziale Verteilungskonflikte) mit 48,3% leicht über „Heute-Morgen“ (46,5%). Diskussionen über Politik mit Familie und Freunden führen über 65% der Jugendlichen, deutlich weniger hingegen mit Lehrkräften (36%). Obwohl Lehrkräfte Diskussionsangebote zu aktuellen Themen machen (66%), geben 63% der Schüler*innen an, nicht zu diskutieren und 36% erleben wenig Toleranz. Die Übereinstimmung mit Meinungen in der Klasse ist deutlich geringer als unter Freunden und Familie. Auf Klassenebene werden als Ursache andere politische Ansichten (18,3%), Erziehung (8,4%), Werte (9,3%) und religiöse Werte (4,2%) genannt. Ähnliche Prozentangaben zeigen sich bei Social-Media-Aktivitäten. 2/3 der Befragten stimmen nicht mit Social-Media-Positionen überein. Bei Konfliktformen sind diskursive Auseinandersetzungen mit knapp 76% die Regel, gefolgt von Ausgrenzung (25%), Mobbing (17%), Hate Speech und Beleidigungen (je 9,3%), Ausnahme bildet Gewalt (1,2%). Mädchen berücksichtigen bei Konflikten signifikant stärker beide Seiten ($p=.004$) und halten sich bei Streit mehr zurück als Jungen ($p=.093$). Der Migrationshintergrund hat keinen signifikanten Einfluss.

Die Pilotstudie mit Limitation einer Gelegenheitsstichprobe gibt erste Einblicke wie Schüler*innen konflikthafte Streitthemen erleben. Auch wenn der Diskurs noch als Hauptform der Auseinandersetzung gilt, so ist der Befund, dass an Gymnasien Multiperspektivität und Kontroversität stärker gefördert werden als an anderen Schulformen (Hahn-Laudenberg & Abs 2024, 300), gefährlich. Diese Benachteiligung in der politischen Kompetenzentwicklung (z.B. Fehlkonzeppte zu Konflikt, geringere Argumentationsfähigkeit) kann Polarisierungstendenzen verstärken. Im nächsten Schritt werden Fokusgruppen-Interviews mit Schüler*innen aller Schulformen geführt, um das Konzept agonaler Pluralität vertiefter auszuleuchten und schulpraktische Ansätze der agonistischen Demokratieerziehung kritisch zu reflektieren (Yazek 2021, 5).

Umstrittene Zukünfte? Politisches Wissen und sozialer Hintergrund als Prädiktoren von Krisenwahrnehmung

Daniel Deimel, Abs Hermann Josef
Universität Duisburg-Essen, Deutschland

Hintergrund & Fragestellung

Problemlagen, die sich krisenhaft darstellen, wie Klimawandel, soziale Ungleichheit und globale Sicherheitsrisiken werden die Zukunft maßgeblich prägen und stellen auch (politisch) bildende Institutionen vor Herausforderungen (Abs, 2021). Bildungsprozesse können dazu beitragen, ein Verständnis für gegenwärtige und zukünftige Herausforderungen auszubilden und Lernende zu unterstützen, die zur Adressierung dieser Herausforderungen erforderlichen Fähigkeiten und Kenntnisse zu entwickeln.

Zukünfte, im Plural als die Möglichkeit einer Vielzahl von Szenarien, lassen sich nach Urry (2016) als *umstritten* bezeichnen. Diese Beschreibung bezieht sich auf die Idee, dass verschiedene soziale Gruppen unterschiedliche Vorstellungen und Interessen bezüglich der Zukunft haben. Zukunft ist also eine Projektion, bei der unterschiedliche soziale, politische, ökonomische und kulturelle Konzeptionen miteinander in Konflikt stehen und aus Perspektive Urrys (2016) als sozial konstruiert zu verstehen.

Personen mit der Befähigung, potentielle Zukünfte zu antizipieren, können als „futures literate“ (Miller, 2018, S. 15) bezeichnet werden, mit einem Bewusstsein dafür, wie Zukunftsvorstellungen individuelle Entscheidungen und Handlungen in der Gegenwart beeinflussen. Die Befähigung, komplexe Phänomene wie den Klimawandel erklären zu können, wird häufiger von Schüler:innen an Gymnasien berichtet (Diedrich et al., 2022). Außerdem zeigen sich Jugendliche im Zusammenhang mit klimabezogenen Themen nicht uniform, sondern gesellschaftliche Bruchlinien strukturieren jugendliche Perspektiven auch entlang sozioökonomischer Kontexte (Albert et al., 2020).

Politisch bildende Angebote können auf dem Verständnis von jungen Menschen für gegenwärtige und zukünftige Bedrohungen aufbauen. Hierzu ist ein differenziertes Verständnis der Interessen von Jugendlichen erforderlich, um authentisch anknüpfen zu können. Dieser Beitrag fragt nach den Mustern der Krisenwahrnehmung Jugendlicher sowie dem Zusammenhang dieser Muster zu politischem Wissen und sozioökonomischem Hintergrund.

Methoden

Dazu wird eine vertiefende Analyse der repräsentativen Stichprobe Nordrhein-Westfalen der International Civic and Citizenship Education Study (ICCS 2022) vorgenommen. Die Stichprobe (Deimel & Welsandt, 2024). Die Stichprobe wurde mit Hilfe eines mehrfach gestuften Verfahrens gezogen und umfasst 3.269 Schüler:innen in Jahrgang 8 (45% weiblich; 67% mit mindestens einem Elternteil, das in Deutschland geboren worden ist) an 145 Schulen. Die Erhebung bestand aus einem 45-minütigen Test zur Erfassung des *politischen Wissens* sowie zwei Fragebogenmodulen (ca. 60 min). Zentral für diesen Beitrag sind die *Zukunftsperspektiven auf Bedrohungsszenarien*. Den Schüler:innen wurden verschiedenen Begriffe präsentiert, die sie mit einem vierstufigen Antwortformat bewerten sollten: „Große Bedrohung“ (1), „Mittlere Bedrohung“ (2), „Kleine Bedrohung“ (3), „Keine Bedrohung“ (4). Zur Analyse wurden die Bewertungen der folgenden sechs Begriffe ausgewählt: „Klimawandel“, „Umweltverschmutzung“, „Krieg / bewaffnete Konflikte“, „Terrorismus“, „Armut“ sowie „Globale Finanzkrisen“. Die Bewertungen der Bedrohungen wurden in eine latente Klassenanalyse (LCA; Collins & Lanza, 2010) überführt. Vor dem Hintergrund der umstrittenen Zukünfte (Urry, 2016), also dass die individuelle Annahme der Bedrohlichkeit bestimmter Zukunftsszenarien das Ergebnis komplexer sozialer Prozesse ist und in Populationen unterschiedlich ausfallen kann, erscheint die LCA hier als besonders geeignetes Verfahren. Ein weiterer Vorteil der LCA besteht darin, dass sie ein statistisches Modell anbietet, das die Wahrscheinlichkeit jedes Individuums schätzt, zu einer bestimmten latenten Klasse zugeordnet zu werden. Diese Zuordnungswahrscheinlichkeiten werden in einem weiteren Schritt zur Analyse der Interaktion zwischen sozioökonomischem Status und politischem Wissen genutzt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass Jugendliche sich hinsichtlich ihrer Bedrohungswahrnehmung zu einer von fünf latenten Klassen zuordnen lassen: Sicherheitsfokussierte (ca. 17% der Schüler:innen), Umwelt- und Klimabewusste (33%), gemäßigte Skeptiker:innen (9%), Alarmierte (40%) sowie Entwarnende (1%). Erste korrelative Analysen zeigen, dass insbesondere die Klassifizierung als Umwelt- und Klimabewusste in positivem Zusammenhang mit politischem Wissen steht, die Klassifizierung als Sicherheitsbewusste und gemäßigte Skeptiker:innen in negativem Zusammenhang. Ähnliche Muster zeigen sich auch für den sozioökonomischen Hintergrund. Jugendliche zeigen, abhängig von ihren Voraussetzungen, keine uniforme Perspektive auf die Zukunft. Stattdessen kann in den Wahrnehmungen der Jugendlichen das soziale Phänomen der umstrittenen Zukünfte nachgezeichnet werden. Eine (politisch-bildnerische) Pädagogik, welche hieran authentisch anknüpfen will, sollte dies berücksichtigen.

Prädiktoren der politischen Involviertheit von Lehramtsstudierenden und damit verbundene Implikationen für die Lehrer*innenbildung

Janis Wehde, Bea Bloh, Martina Homt

Universität Paderborn, Deutschland

Schule ist eine zentrale Institution der Gesellschaft, um Demokratie als Gesellschafts-, Lebens- und Regierungsform zu vermitteln (Himmelfmann, 2016). Dabei wird schulische Demokratiebildung als Aufgabe aller Lehrpersonen verstanden (Aktionsrat Bildung, 2020; Avenarius & Hanschmann, 2019; SWK, 2024). Politische Involviertheit kann dabei als immanenter Bestandteil der professionellen demokratiebildenden Kompetenz von (angehenden) Lehrpersonen angesehen werden (u.a. Bicheler, Heinrich & Klameth, 2023; Busch & Dittgen, 2024). Sie spiegelt eine grundlegende Orientierung gegenüber Politik wider, die als Merkmal der personalen Verankerung demokratischer Anforderungen (u.a. Partizipation und kritische Wachsamkeit) verstanden wird (Bromme & Rothmund, 2021) und stützt Lehrpersonen bei der Gestaltung prodemokratischer Bildungsgelegenheiten (z.B. einer offenen Diskurskultur im Unterricht) (Sant, 2019).

Über das Konstrukt der politischen Involviertheit von (angehenden) Lehrpersonen ist – im Vergleich zu den in diesem Konstrukt verankerten Einzelaspekten (politisches Interesse; politische Selbstwirksamkeit; gesellschaftspolitisches Engagement) sowie zu Daten aus berufsunabhängigen länderübergreifenden Untersuchungen – jedoch wenig bekannt. Eine aktuelle Untersuchung der politischen Involviertheit von angehenden Lehrpersonen und deren Einflussfaktoren kann nicht zuletzt deshalb dazu beitragen, Schlussfolgerungen für die Gestaltung der Lehrer*innenbildung zu ziehen.

Die daran anknüpfende Forschungsfrage des Beitrages lautet: Welche Faktoren bestimmen die politische Involviertheit von Lehramtsstudierenden und welche Implikationen können daraus für die Lehramtsausbildung abgeleitet werden?

Aufbauend auf der Forschungsfrage wurde die politische Involviertheit von Lehramtsstudierenden mit den Facetten wahrgenommene politische Selbstwirksamkeit (Bromme, Rothmund & Capara, 2020), politisches Interesse (Otto & Bacherle, 2011) und gesellschaftspolitisches Engagement (van Deth & Zorell, 2018) sowie die mit diesen Facetten verbundenen Einflussfaktoren Geschlecht, sozioökonomischer Status, Migrationshintergrund und die studierte Fächerkombination (Arens & Watermann, 2017; Fischer, Kauder & Potrafke, 2016; Gabriel, 2022; Reinders, 2016; van Deth, Abendschön & Vollmar, 2011) in einer Online-Fragebogenstudie mithilfe des Umfragetools Limesurvey an lehramtsauszubildenden Universitäten in Nordrhein-Westfalen untersucht. Die Stichprobenumfang betrug nach Bereinigung der Daten N = 750 Lehramtsstudierende aus elf Universitäten.

Die statistische Modellierung der politischen Involviertheit erfolgte in Anlehnung an Bromme und Rothmund (2021) als Faktor höherer Ordnung der drei Facetten wahrgenommene politische Selbstwirksamkeit, politisches Interesse und gesellschaftspolitisches Engagement. Zur Analyse der Einflussfaktoren wurden multiple Regressionsmodelle mit den in der Untersuchung erhobenen Einflussfaktoren als Prädiktoren und der politischen Involviertheit als Kriterium berechnet.

Die Ergebnisse stützen die prädikative Relevanz der untersuchten Einflussfaktoren für die politische Involviertheit der Lehramtsstudierenden. Zur Kompensation der Auswirkungen der differenten prädikativen Merkmale der Studierenden könnten studienfachübergreifende Austauschmöglichkeiten im Peer-Format (z.B. Demokratie-Laboratorien) geschaffen werden (zur Bedeutung von Peer-Kontexten für politische Partizipation und politisches Wissen u.a. Ekström & Östman, 2013). Hierbei geht auch um eine institutionelle Ausweitung der möglichen politisch- demokratischen Sozialisationsräume der Studierenden (zur Bedeutung von Sozialisationsinstanzen für politische Einstellungen und politisches Verhalten u.a. Rippl, Seipel & Kindervater, 2022). Darüber hinaus wird diskutiert, inwiefern – insbesondere vor dem Hintergrund der aktuell ausbaufähigen Verankerung demokratiebildender Inhalte im Lehramtsstudium (Berkemeyer, Franzmann & May, 2023) – Veranstaltungen des Lehramtsstudiums von den Lehrenden um sekundäre Lernziele zum offenen Diskurs (u.a. Oesterreich, 2003; Persson, 2015) und zur Partizipation im engeren Sinne (Kärner et al., 2023; Sant, 2019) ergänzt werden sollten.

PS-13: Potenziale und Herausforderungen für das Lernen mit Virtual Reality

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 256

Chair der Sitzung: Miriam Hansen

Auswirkungen von Virtual Reality - Lernaktivitäten auf emotionale, kognitive und verhaltensbezogene Reaktionstendenzen von Lehramtsstudenten auf herausfordernde Unterrichtssituationen

Julia Mendzheritskaya¹, Miriam Hansen¹, Sebastian Breitenbach²

¹Goethe Universität, Institut für Psychologie, Deutschland; ²Goethe Universität, Akademie für Bildungsforschung und Lehrkräftebildung, Deutschland

Der Einsatz von Virtual Reality (VR)-Technologie in der Lehrkräfteausbildung wird als vielversprechender Ansatz zur Förderung professioneller Kompetenzen von Lehrkräften angesehen (Breitenbach, 2024; Stavroulia et al., 2018). Es wird davon ausgegangen, dass die Integration von VR-Technologien in Lernaktivitäten aufgrund des durch Immersion entstehenden intensiven Gefühls der Präsenz (Hamilton et al., 2021; Radianti et al., 2020) die Theorie-Praxis-Verbindung im Rahmen der Lehramtsausbildung (Hascher & de Zordo, 2015) stärken kann. Insbesondere Lernaufgaben, die darauf abzielen, effektive Verhaltensstrategien zu trainieren, können durch den Einsatz von VR-Technologien ausgebaut werden, weil erworbene Kenntnisse und Fähigkeiten in direkter Interaktion mit dem Lernenden erkundet werden können (Radianti et al., 2020).

In der aktuellen Forschung wurden eine Reihe an Lernangebotsmerkmalen und didaktischen Aspekten von VR-Lernumgebungen identifiziert, die positiv auf Lernerfahrungen und das Studieninteresse der Studierenden auswirken (Radianti et al. 2020). Trotz der Existenz etablierter VR-Klassensimulationen für die Ausbildung von Lehramtsstudierenden (Breitenbach, 2024; Fischler, 2006, Lugin et al. 2016) ist immer noch sehr wenig darüber bekannt, wie das Agieren in VR-Lernumgebungen das Lernen der Studierenden auf emotionaler, kognitiver und verhaltensbezogener Ebene beeinflusst und wie gut die Teilnahme an VR-Klassensimulationen Lehrkompetenzen von künftigen Lehrkräften unterstützen kann (Stavroulia et al. 2018). Daher zielt die vorliegende Studie (gefördert durch BMBF 01JA2025) darauf ab, die Auswirkungen von VR-Lernaktivitäten auf emotionale, kognitive und verhaltensbezogene Reaktionstendenzen von Lehramtsstudierenden auf herausfordernde Unterrichtssituationen zu analysieren.

Im Rahmen des Vorhabens wurde eine Lehrveranstaltung für Lehramtsstudierende zum Thema "Förderung von digitalen Lehrkompetenzen durch den Einsatz von Virtual Reality" entwickelt. Der Kurs war in fünf Themen unterteilt und beinhaltete zwei VR-Lehr-Lern-Labore (Roth & Priemer, 2020) mit daran anschließenden Reflexionsphasen, die zum Anfang und gegen Ende des Semesters stattfanden.

Basierend auf den Dimensionen guten Unterrichts (Helmke 2012) wurden drei Szenarien für VR-Klassensimulationen entwickelt und pilotiert (Glocker et al., 2024). Die Szenarien orientierten sich inhaltlich auf professionelle Kompetenzen der Lehrkräfte in den Bereichen Klassenführung, kognitive Aktivierung und konstruktive Unterstützung (Baumert & Kunter, 2011) und wurden als herausfordernde Unterrichtssituationen konstruiert. Die Szenarien wurden nach dem „decision tree“-Ansatz in drei Ebenen gegliedert (Breitenbach, 2024; Fischler, 2006).

Die Teilnehmenden waren Lehramtsstudierende (N=80; N = 55 weiblich), die an der Lehrveranstaltung im Sommersemester 2023 und 2024 teilnahmen. Für die Prä-Post-Messung der emotionalen Erfahrungen vor und nach der Teilnahme am VR-Lehr-Lern-Labor wurden die Teilnehmer mit der PANAS (Krohne et al., 1996) befragt. Am Ende jedes VR-Labor wurden die Studierenden außerdem gebeten, einen kurzen Fragebogen zur Technologieakzeptanz (nach Davis, 1989) auszufüllen. Die Handlungsauswahl während des Agierens in VR-Klassensimulationen wurde dokumentiert und nach dem Passungsgrad für die Lösung der jeweiligen Unterrichtssituation kodiert. Darüber hinaus füllten die Teilnehmenden zu Beginn und am Ende der Lehrveranstaltung als Prä- bzw. Post-Messungen die Fragebögen 1) zu Einstellungen gegenüber digitalen Kompetenzen (adaptiert nach Vogelsang et al., 2018); 2) zu Emotionsregulationsstrategien (deutsche Version des ERQ, Abler & Kessler, 2011); 3) zu Klassenführungscompetenzen (LDK; Lenske & Mayr, 2015) aus.

Die Ergebnisse zeigen, dass sich die Teilnehmenden nach dem Agieren in der Rolle einer virtuellen Lehrkraft stolzer, weniger gelangweilt und weniger ängstlich fühlten. Die Emotion der Lernfreude dominierte in beiden VR-Lehr-Lern-Laboren. Darüber hinaus profitierten insbesondere diejenigen Studierenden von dem Trainingseffekt der Teilnahme an VR-Klassensimulationen, deren Handlungsauswahl in dem ersten VR-Lehr-Lern-Labor weniger passend ausgefallen ist. Des Weiteren weisen die Ergebnisse auf eine signifikante Intensitätsminderung der Motive für die Nutzung digitaler Medien im Allgemeinen und im Unterricht zwischen den Messzeitpunkten auf. Darüber hinaus gab es keine signifikanten Unterschiede zwischen den Prä- und Post-Messungen in der Ausprägung der Emotionsregulationsstrategien sowie in den selbsteingeschätzten Klassenführungscompetenzen. Die Implikationen der gewonnenen Erkenntnisse für die Optimierung von Lehrkräfteausbildungsprogrammen und für den Einsatz von VR-Technologien zur Förderung professioneller Lehrkompetenzen im universitären Kontext werden diskutiert.

ENTFÄLLT Kognitive Aktivierung durch Lernen mit Virtual Reality?

David Kablitz, Stephan Schumann, Jan Paul Baur, Matthias Conrad

Universität Konstanz, Deutschland

Theoretischer Hintergrund

Der Beitrag führt zwei Forschungsstränge zusammen: (1) die Forschung zum Einsatz von Virtual Reality (VR) in Lernumgebungen und (2) die Studienlage zur kognitiven Aktivierung von Lernenden im Unterricht. Zu (1): Übersichtsarbeiten, die die Wirksamkeit der VR-Technologie im Vergleich zu anderen Medien und Methoden hinsichtlich des Lernerfolgs untersuchen, zeigen, dass sich die Integration von VR in Lehr-Lern-Settings bei der Vermittlung prozeduralen Wissens als vorteilhaft erweist. Für den Erwerb deklarativen Wissens ist die Studienlage weniger eindeutig (Hamilton et al., 2021; Conrad et al., 2024; Thomann et al., 2024). Gemäß des „Cognitive and Affective Model of Immersive Learning“ (CAMIL; Makransky & Petersen, 2021) kann die lernförderliche Wirkung von VR mit den Faktoren „Handlungsfähigkeit“ und „Präsenzerleben“ begründet werden. Diese wirken auf kognitive und affektive Faktoren (z.B. Cognitive Load und Interesse), die wiederum den Wissenserwerb beeinflussen. Zu (2): Allgemein in der Unterrichtsqualitätsforschung, aber auch spezifisch in der Literatur zum Einsatz digitaler Medien im Unterricht wird darauf hingewiesen, dass dessen Effektivität von der kognitiven Aktivierung der Lernenden abhängt (Quast et al., 2021), welche wiederum als eine der drei Basisdimensionen guten Unterrichts als Prädiktor für gelingendes Lernen verstanden wird (Rieser & Decristan, 2023).

Fragestellungen

Inwiefern mittels VR-Technologie die kognitive Aktivierung im Unterricht gefördert werden kann und welche VR-spezifischen Eigenschaften mit dieser in Zusammenhang stehen, ist bislang in der Forschung jedoch noch wenig beleuchtet. Vor diesem Hintergrund untersucht der Beitrag erstens, inwiefern VR-integrierter Unterricht die kognitive Aktivierung der Lernenden fördert und wie diese mit dem Erwerb relevanten Fachwissens zusammenhängt. Zweitens wird untersucht, welche VR-spezifischen Faktoren gemäß dem CAMIL-Modell mit der kognitiven Aktivierung der Lernenden in Zusammenhang stehen.

Methode

Im Rahmen eines quasi-experimentellen Settings und mit Hilfe eines Pre-Post-Test-Designs wurde eine VR-integrierten Unterrichtsserie (Treatmentgruppe) bestehend aus insgesamt vier Unterrichtseinheiten im Fach *Betriebswirtschaft* mit analogem Unterricht (Kontrollgruppe) zum Thema „Verkaufsraumgestaltung“ verglichen. Hierbei wurden Daten in den Schuljahren 2022/23 und 2023/24 von $N=137$ Auszubildenden ($n_{VR}=71$, $n_{KG}=66$) im Beruf „Verkäufer/in“ (1. Ausbildungsjahr) aus insgesamt 8 Klassen erhoben (4 Klassen mit VR; 4 Klassen mit analog gestaltetem Unterricht). Der Erwerb von Fachwissen wurde via Pre-Posttest-Messung über einen eigens hierfür entwickelten Fachwissenstest abgebildet (26 Items; EAP/PV-Reliabilität = .71). Die weiteren Merkmale wurden über Selbstauskünfte der Lernenden anhand einer 5-stufigen Likert-Skala erfasst: *Präsenzerleben* (5 Items; $\alpha=.90$) und *Handlungsfähigkeit* (4 Items; $\alpha=.70$) am Ende der vierten Unterrichtseinheit (T5), *kognitive Aktivierung* (4 Items; $.76 < \alpha < .86$) hingegen zum Zeitpunkt des Pretests (T1, als Baseline) sowie direkt nach den ersten drei Unterrichtseinheiten (T2–T4).

Ergebnisse und ihre Bedeutung

Während vor der Unterrichtsreihe keine signifikanten Gruppenunterschiede in der *kognitiven Aktivierung* zwischen Experimental- und Kontroll-Gruppe festgestellt werden ($M(SD)_{VR}=3.59(0.55)$, $M(SD)_{KG}=3.57(0.59)$, $p=.86$), zeigen sich signifikante Unterschiede unmittelbar nach den einzelnen Unterrichtseinheiten (T2: $M(SD)_{VR}=4.23(0.74)$, $M(SD)_{KG}=3.35(0.42)$, $d=1.26$, $p<.001$; T3: $M(SD)_{VR}=4.09(0.84)$, $M(SD)_{KG}=3.50(0.67)$, $d=0.73$, $p=.001$; T4: $M(SD)_{VR}=4.13(0.81)$, $M(SD)_{KG}=3.60(0.72)$, $d=0.68$, $p=.003$), die auf eine Vorteilhaftigkeit des VR-integrierten Unterricht hinweisen. Der Zusammenhang zwischen der *kognitiven Aktivierung* und dem Erwerb von Fachwissen (Posttest; unter Kontrolle von Vorwissen) wurde für die Messzeitpunkte der kognitiven Aktivierung (T2–T4) jeweils durch Regressionen untersucht. Die Analysen zeigen signifikante Zusammenhänge für T3 ($\beta=.268$; $p=.022$) und T4 ($\beta=.453$; $p<.001$). Die Befunde zu diesbezüglich vertieften Analysen werden auf der Tagung vorgestellt. Hinsichtlich der Verbindungen zwischen den im CAMIL beschriebenen Konstrukten und der *kognitiven Aktivierung* (T4) kann zudem ein signifikanter Zusammenhang mit dem *Präsenzerleben* ($\beta=.588$; $p<.001$) beobachtet werden.

Die Befunde verweisen auf eine Vorteilhaftigkeit des VR-Einsatzes. Es wird deutlich, dass die beobachtbare kognitive Aktivierung im VR-integrierten Berufsschulunterricht insbesondere mit dem Präsenzerleben in virtuellen Lernumgebungen in Zusammenhang steht. Zudem unterstreichen die Befunde die Relevanz der kognitiven Aktivierung. Mit Blick auf die Unterrichtsqualität in VR-integriertem Unterricht sollten zukünftige Untersuchungen verstärkt didaktische und strukturelle Merkmale des Unterrichts berücksichtigen, um zu klären, wie VR möglichst kognitiv aktivierend in den Unterricht eingebunden werden kann.

Simulationsübelkeit in der schulischen Virtual Reality-Lehre: Gendereffekte und Wirkungen auf ARCS-Motivation

Maximilian Fink¹, Bianca Watzka², Bernhard Ertl¹

¹Universität der Bundeswehr München, Deutschland; ²Otto-von-Guericke-Universität Magdeburg

Theoretischer Hintergrund: Mehrere Studien belegen bereits, dass auch in der Virtual Reality (VR) effektiv Wissen vermittelt werden kann (Elme et al., 2022; Gloy et al., 2022). Unterstützt wird Wissensvermittlung in der VR in MINT-Fächern beispielsweise durch Multimedia-Elemente wie Erklärgrafiken mit verbalen Erläuterungen, die komplexe naturwissenschaftliche Themen verständlich präsentieren (Mayer, 2014). Die Entwicklung von ansprechenden und wirksamen VR-Lernumgebungen ist aufwendig und kann insbesondere bei der Erzeugung sogenannter Virtual Field Trips (Woerner, 1999) durch die Einbettung von 360-Grad-Videos vereinfacht werden (Eisenlauer & Sosa, 2022). Allerdings können insbesondere in VR-Lernumgebungen, die mit 360-Grad-Videos erzeugt werden, erhöhte Werte von Simulationsübelkeit auftreten (Shahid Anwar et al., 2020). Darüber hinaus weisen Studien darauf hin, dass Simulationsübelkeit in der VR bei Frauen häufiger auftritt als bei Männern (Tian et al., 2022). Mehrere Laborstudien liefern überzeugende Belege für einen negativen Zusammenhang zwischen Simulationsübelkeit und kognitiven Leistungen (Weech et al., 2019). Noch untersucht werden muss, wie Simulationsübelkeit in VR-Schulsettings den Wissenserwerb und die motivationalen Variablen beeinflusst. In Bezug auf die Motivation muss die Wirkung auf die vier Variablen des ARCS-Modells der Motivation (Aufmerksamkeit, Relevanz, Confidence und Satisfaction) erforscht werden, die ein motivierendes Lernumfeld schaffen können (Keller, 1987).

Fragestellungen: Inwieweit erwerben SchülerInnen Wissen in einer VR-Lernumgebung mit Erklärgrafiken und verbalen Erläuterungen? Wie unterscheiden sich Schülerinnen und Schüler in Bezug auf Simulationsübelkeit? Wie wirkt Simulationsübelkeit auf Wissenserwerb in einer schulischen VR-Lernumgebung? Inwieweit hängt Simulationsübelkeit mit den Variablen des ARCS-Modells zusammen?

Methode: $N = 126$ 10.-Klässler aus deutschen Gymnasien absolvierten die Studie, davon waren 58 (46%) männlich und 65 (52%) weiblich; 3 Teilnehmende (2%) machten keine Angabe zum Geschlecht. Die Studie bestand aus einem Prätest, der Lernumgebung und einem Posttest. Die Lernumgebung umfasste Physik- und Chemieinhalte zur Abwasseraufbereitung. An neun interaktiven Stationen mit 360-Grad-Videos lernten die Teilnehmenden anhand von Erklärgrafiken auf Konsolen, begleitet von Erzähler-Erläuterungen. Die Lernumgebung dauerte 30 Minuten und wurde im Sitzen dargeboten. Wissen wurde mit einem Single-Choice-Test erfasst, der aus fünf Items mit jeweils drei Antwortmöglichkeiten bestand. Für diese Wissenstests wird der Prozentsatz der korrekten Antworten von 0% (keine korrekten Antworten) bis 100% (ausschließlich korrekte Antworten) angegeben. Simulationsübelkeit wurde mit dem Fragebogen CSQ-VR (Kourtesis et al., 2023) erhoben, die Skala reichte von 1 (kein Gefühl von Simulationsübelkeit) bis 7 (extremes Gefühl von Simulationsübelkeit). Die ARCS-Gesamtskala und Subskalen konnten Werte von 1 (geringe Ausprägung) bis 5 (hohe Ausprägung) erreichen. Die Computer- und VR-Nutzung der SchülerInnen wurde ebenfalls erfasst.

Ergebnisse und ihre Bedeutung: Über alle Gruppen hinweg gewannen SchülerInnen vom Prätest ($M = 44$ % korrekte Antworten, $SD = 22\%$) zum Posttest ($M = 63\%$ korrekte Antworten, $SD = 19\%$) 19 Prozentpunkte Wissen hinzu, $t(125) = 7.91$, $d = 0.70$, $p < .001$. Dieses Ergebnis zeigt, dass VR mit Erklärgrafiken und verbalen Erläuterungen effektiv Wissen zu einem naturwissenschaftlichen Thema in einem Schulsetting vermitteln kann. Analysen von Gendereffekten zeigten, dass Schüler deutlich geringere Simulationsübelkeit ($M = 2.39$) als Schülerinnen ($M = 3.39$) verspürten, $t(119) = 3.55$, $d = 0.65$, $p < .001$.

Zusätzliche Regressionsanalysen ergaben, dass die entstandenen Symptome von Simulationsübelkeit nicht nur durch Gender, sondern auch Computer- und VR-Nutzung erklärt werden. Simulationsübelkeit hatte eine mittlere, negative Korrelation von $r(122) = -.27, p = .002$ mit dem Posttest für Wissen. Weitere Regressionsanalysen unterstreichen, dass Simulationsübelkeit, auch wenn andere erklärende Faktoren einbezogen wurden, ein bedeutsamer Prädiktor von Wissenserwerb blieb. Darüber hinaus zeigten sich mittlere bis hohe Zusammenhänge zwischen Simulationsübelkeit und den Variablen der ARCS-Motivation (Gesamtkorrelation $r(122) = -.36, p < .001$, Einzelkorrelationen der ARCS-Subskalen: $r = -.15$ bis $-.41$). Diese Ergebnisse verdeutlichen, dass Simulationsübelkeit die Motivation in schulischen VR-Lernumgebungen erheblich beeinträchtigen kann. Abschließend diskutieren wir Fragen für zukünftige Forschung und geben Empfehlungen zur Reduzierung von Simulationsübelkeit in VR-Lernumgebungen sowie zur Erhöhung der ARCS-Motivation.

Eine Untersuchung zur Validität einer auf Beobachtungen basierenden Bewertung von Unterrichtsqualität in einer VR-basierten Unterrichtssimulation

Nico Klausner-Thimm, Yizhen Huang, Eric Richter, Dirk Richter

Universität Potsdam, Deutschland

Theoretischer Hintergrund

Virtual Reality (VR) bietet die Möglichkeit, Lehr- und Lernprozesse in einer realistischen und sicheren Umgebung zu simulieren. In der Forschung wird VR daher insbesondere eingesetzt, um handlungsbezogene Kompetenzen, wie z.B. das Klassenmanagement, bei angehenden Lehrkräften durch VR-basierte Unterrichtssimulationen zu fördern (McGarr, 2020). VR kann aber auch zur videobasierten Reflexion (Huang, Richter, Kleickmann & Richter, 2023) und Evaluation von Unterricht (Daltoë et al., 2024) verwendet werden. In beiden Anwendungsszenarien hat sich der Einsatz von VR-Videos als effektiver erwiesen als herkömmliche Unterrichtsvideos. Daltoë et al. (2024) haben gezeigt, dass die Bewertung der Unterrichtsqualität mit Hilfe von 360-Grad-Videos in VR im Vergleich zu herkömmlichen Unterrichtsvideos signifikant besser ausfällt. In einem Vergleich ohne VR-Funktionalität konnte dies nicht festgestellt werden (Gold & Windscheid, 2020). VR bietet somit eine vielversprechende Möglichkeit, die Unterrichtsqualität durch externe Beobachter:innen zu bewerten. Bislang liegen jedoch keine Studien vor, die zeigen, dass eine reliable und valide Messung der Unterrichtsqualität auch im Rahmen einer VR-basierten Unterrichtssimulation möglich ist. Da sich der Unterricht bei VR-basierten Unterrichtssimulationen vollständig in eine virtuelle Realität verlagert, ist eine Überprüfung der Messbarkeit von besonderer Bedeutung. Insofern soll im Rahmen dieser Arbeit die Validität einer auf Beobachtungen basierenden Bewertung der Unterrichtsqualität in einer VR-basierten Unterrichtssimulation untersucht werden. Als Bewertungsinstrument wurde das standardisierte und bereits mehrfach validierte Classroom Assessment Scoring System für die Oberstufe (CLASS-S) ausgewählt.

Fragestellung

Die vorliegende Arbeit untersucht die folgende Fragestellung:

Ermöglicht eine VR-basierte Unterrichtssimulation eine valide Beschreibung der Unterrichtsqualität durch das Classroom Assessment Scoring System (CLASS-S)?

Methode

Es wurden 67 Bachelor-Studierende des Lehramts in einem Seminar zum Klassenmanagement während einer VR-basierten Unterrichtssimulation gefilmt. Die Studierenden waren zu 52,3% weiblich und durchschnittlich 24,9 Jahre alt ($SD = 4,7$ Jahre). Sie hatten die Aufgabe, einen Vortrag zum Thema „Notizen machen“ zu halten und dabei auf Unterrichtsstörungen unterschiedlichen Schweregrades wie z.B. ein Gespräch zwischen zwei SchülerInnen zu reagieren. Zwei zertifizierte CLASS-Rater:innen bewerteten anschließend unabhängig voneinander die Unterrichtsqualität in sieben Dimensionen des CLASS-S, welche sich wiederum drei übergeordneten Domänen zuordnen lassen. Die erste Domäne „Emotionale Unterstützung“ ($\alpha = .74$) beinhaltet die Dimensionen „Positives Klima“, „Lehrersensibilität“, und „Berücksichtigung der Perspektiven der Jugendlichen“. Die zweite Domäne „Klassenmanagement“ ($\alpha = .70$) umfasst die Dimensionen „Verhaltensmanagement“ und „Produktivität“. Die dritte Domäne „Instruktionale Unterstützung“ ($\alpha = .78$) setzt sich aus den Dimensionen „Instruktionale Lernformate“ und „Inhaltliches Verständnis“ zusammen. Dimensionen, die auf Handlungen der virtuellen SchülerInnen abzielen und aufgrund technischer Einschränkungen in der VR nicht simulierbar waren, wurden aus der Untersuchung ausgeschlossen. Um abschließend zu überprüfen, ob das CLASS-S die Unterrichtsqualität in einer VR-basierten Unterrichtssimulation valide erfasst, wurde mit der Maximum-Likelihood Schätzmethode (ML) in Jamovi 2.3 eine konfirmatorische Faktorenanalyse durchgeführt. Zur Behandlung fehlender Daten wurde die vollständige Maximum-Likelihood-Schätzung (FIML) verwendet.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse sprechen für einen akzeptablen bis guten Modellfit ($\chi^2 [17] = 23.00, p = 0.15$; RMSEA = 0.08; CFI = 0.96; TLI = 0.94; SRMR = 0.07). Auf Ebene der Domänen zeigt sich für das Modell, dass „Klassenmanagement“ stark positiv mit „Instruktionale Unterstützung“ ($\beta = .88; p < .001$) und „Emotionale Unterstützung“ ($\beta = .58; p < .001$) korreliert ist, ebenso wie „Instruktionale Unterstützung“ mit „Emotionale Unterstützung“ ($\beta = .53; p < .001$). Da die faktorielle Struktur des CLASS-S abgebildet werden konnte, ist eine valide Messung der Unterrichtsqualität in einer VR-basierten Unterrichtssimulation möglich. Dementsprechend kann davon ausgegangen werden, dass unterrichtsbezogene Merkmale wie das Klassenmanagement von Lehrkräften mit Hilfe bestehender Beobachtungsinstrumente in einer VR-Unterrichtssimulation valide erfasst werden können. Dies spricht auch für den möglichen Einsatz von VR-Unterrichtssimulationen zu Zwecken experimenteller Unterrichtsforschung. Daneben zeigt die Studie die Möglichkeit auf, kompetenzorientierte VR-Trainings im Rahmen der Lehrkräftebildung zukünftig durch Messungen der Unterrichtsqualität zu begleiten und unmittelbar hinsichtlich ihrer Wirksamkeit zu evaluieren.

PS-14: Künstliche Intelligenz in der Lehrerbildung und -praxis:

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 159

Chair der Sitzung: Franziska Baier-Mosch

Teachers' use of AI for monitoring, engagement, feedback, and assessment: A rapid review of research syntheses

Meg Farrell, Begüm Arvaneh, Fabian Reinwarth, Sabrina Reith, Tina Seidel

Technical University of Munich, Germany

Teachers' use of AI for monitoring, engagement, feedback, and assessment: A rapid review of research syntheses

Theory

Rapid reviews are an efficient, evidence synthesis method, often tailored for stakeholders, which adapt systematic review processes with some accelerated steps, providing timely, evidence-based knowledge (Garritty et al., 2021; Sutton et al., 2019). Rapid reviews are particularly beneficial for offering stakeholders up-to-date information on emerging research topics or technologies, (e.g., via knowledge mobilization platforms like clearing houses; Cain et al., 2016).

The expansion of AI into K-12 education (Martin et al., 2024) introduces new teacher challenges in navigating evolving discourse and practice applications (Chiu et al., 2023; Mintz et al., 2023). A rapid review, with practice-oriented design, offers a means for addressing these challenges. Thus, the present study aims to (1) elicit teachers' informational needs about AI in schools, (2) conduct a needs-specific rapid review, and (3) prototype a knowledge transfer report format via teacher-researcher co-design, for offer on XXClearing-House.

Research Question

In a preliminary stage, we identified K-12 teachers' informational needs about AI in schools. This led to a rapid review focus on the following research questions:

What AI-focused interventions are teachers using that support effectiveness for ...

- a) ...student motivation/engagement?
- b) ...student monitoring of learning?
- c) ...automatic feedback support?
- d) ...automated student assessment?

To accelerate the methodological process, we concentrate on reviewing existing systematic research syntheses rather than primary empirical works, as these secondary-level studies offer a robust evidence base (Garritty et al., 2024). Thus, our review investigates secondary research on AI-focused interventions (e.g., tools, software, instruction) used by K-12 teachers for various educational purposes.

Method

To narrow the 'AI in schools' scope, our first aim was addressed with a preliminary planning step involving stakeholders (Garritty et al., 2021; Tricco et al., 2016) via online needs-assessment survey (4 multiple-choice questions on population, topic, use, and format; e.g., What AI topics in educational research are you most interested in?). Teacher respondents ($N = 118$), recruited via snowball convenience sampling, determined our study's focus on teachers' use of AI for motivation, monitoring, feedback, and assessment. Further, their responses helped determine how they plan to use to the study results (i.e., personal interest and decision support for planning and teaching), which informs the final synthesis design. Finally, they indicated preferences for video-based and digitally interactive formats for knowledge transfer products on study outcomes.

Towards our second aim, we have begun the needs-specific rapid review. This methodology mostly aligns with standard systematic reviews (see PRISMA; Page et al., 2021), yet lacks an agreed upon consensus for specified acceleration processes (Tricco et al., 2015), and can depend on stakeholder involvement (e.g., Khangura et al., 2012). Our acceleration procedures include, for example, limiting the database search to synthesis sources, and extracting data with one coder and one verifier (e.g., Garritty et al., 2021).

Our database search (e.g., Web of Science), used study-focused keywords (e.g., "AI", "teachers", "K-12", "systematic review"), yielding 192 English language sources, supplemented with a hand search of 'AI in Education' journals (2020-2024), currently being screened. Data extraction is planned for limited aspects of study meta-data, study characteristics, population characteristics, AI tools and functions, and AI-outcomes related to teaching and learning effectiveness. Results will be narratively synthesized, presented at the conference, and transfer formats will be co-designed with stakeholders via focus-group (aim 3).

Results and Significance

This study leverages stakeholder involvement within rapid review design to model targeted, teacher-relevant research synthesis. It also promotes teacher agency in shaping practice-relevant research and knowledge transfer. The rapid review outcomes will provide valuable, teacher-tailored information, aiding them in navigating the rapidly expanding possibilities for AI in the classroom.

„Kannst du mir KI erklären?“ – Der Zusammenhang zwischen dem Alter von Lehrkräften und ihren Einstellungen und Wissen zum Unterrichten von KI

Franziska Baier-Mosch^{1,3}, Lukas Schulze-Vorberg^{2,3}

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ²Goethe-Universität Frankfurt am Main;

³IDeA-Zentrum Frankfurt am Main

Theoretischer Hintergrund

Künstliche Intelligenz (KI) führt zu tiefgreifenden Veränderungen des Alltags- und Berufslebens. Schüler*innen müssen von ihren Lehrkräften in den unterschiedlichen Fächern auf diese Veränderungen vorbereitet werden (Kim & Kwon, 2023). Dies erscheint schwierig, da KI ein relativ neues und komplexes Thema ist und Lehrkräfte in diesem Bereich bisher selbst kaum ausgebildet sind (Kim & Kwon, 2023). Besonders älteren Lehrkräften könnte es schwerfallen, die aus Sicht des viel rezipierten Will-Skill-Tool Modells (Petko, 2012) notwendigen positiven Einstellungen (*will*) und das Wissen (*skill*) zu entwickeln, um KI-Themen effektiv in den Fachunterricht zu integrieren. Studien zeigen, dass ältere Menschen mehr Angst vor der Nutzung von Computern haben (Czaja, 2006) und dass sich ältere Lehrkräfte weniger kompetent in Hinblick auf den Einsatz neuer Technologien wahrnehmen

(z. B. Inan & Lowther, 2010; Sabic et al., 2022). Speziell in Hinblick auf KI-Integration wurde bisher nicht untersucht, inwiefern das Alter von Lehrkräften mit dem *will* (kognitiven und affektiven Aspekte von Einstellungen zur KI-Integration) und dem *skill* (dem Wissen über KI-Integration in den Unterricht) zusammenhängt. Dies zu untersuchen erscheint allerdings sinnvoll, um das Alter als möglichen Risikofaktor bei Bemühungen, die Vermittlung von KI Inhalten im Schulunterricht voranzutreiben, zu berücksichtigen.

Fragestellung

Die Studie untersucht, ob das Alter von Lehrkräften ein Prädiktor für die Einstellungen (*will*) in Hinblick auf die Vermittlung von KI-Inhalten im Unterricht ist. Dabei wird eine kognitive Komponente (Relevanzüberzeugungen) und eine affektive Komponente (Freude an der Beschäftigung mit KI Inhalten zur Vermittlung im Unterricht) betrachtet. Des Weiteren wird untersucht, ob das Alter von Lehrkräften das Wissen über KI (*skill*), nämlich das Fachwissen (CK; Was ist KI?) und das fachdidaktische Wissen (PCK; Wie kann KI als Thema in den Unterricht integriert werden?) vorhersagt. Im Zusammenhang damit wird auch betrachtet, wie schwierig die Lehrkräfte die KI-Inhalte wahrnehmen (mental load).

Methode

Es handelt sich um eine Querschnittsstudie mit $N = 180$ Lehrkräften (63 % weiblich, 1 % divers), die online Fragebögen ausfüllten und eine kurze Lerngelegenheit absolvierten. Die Lehrkräfte beantworteten zunächst Fragen zu ihrem Alter ($M = 41$, $SD = 12$), ihrem Wissen über KI (*skill*; CK: 1 offenes Item $\kappa = .88$, PCK: 1 offenes Item $\kappa = .85$), ihren Relevanzüberzeugungen (*will*; 5 Items; $\alpha = .87$) und ihrer Freude in Hinblick auf die bevorstehende KI-Lerngelegenheit (*will*; 4 Items, $\alpha = .93$). Die Lehrkräfte absolvierten dann die Lerngelegenheit, die aus einem Text bestand, der erläuterte, wie das Thema KI (Chatbot Programmierung) verknüpft mit Fachinhalten im Unterricht vermittelt werden kann. Anschließend beurteilen die Lehrkräfte ihren mental load (*skill*; 4 Items, $\alpha = .92$) in Hinblick auf den Text (Textschwierigkeit).

Ergebnisse & ihre Bedeutung

Eine multivariate Regressionsanalyse zeigte, dass das Alter der Lehrkräfte signifikant mit den Einstellungen (*will*) der Lehrkräfte in Hinblick auf das Unterrichten/Lernen von KI zusammenhing. Je älter die Lehrkräfte waren, desto relevanter fanden sie es, dass Schüler*innen in der Schule KI-Kompetenzen vermittelt werden, $B = 0.013$, $t(171) = 2.08$, $p = .04$. Gleichzeitig sank mit dem Alter die Freude an der Beschäftigung mit der Vermittlung von KI-Themen, $B = -0.019$, $t(171) = -2.41$, $p = .02$. Keine Effekte zeigten Regressionsanalysen für den Zusammenhang zwischen Alter und Wissen in Hinblick auf KI-Themen (CK und PCK; *skill*). Allerdings zeigte sich, dass ältere Lehrkräfte über einen höheren mental load bei der Beschäftigung mit KI-Themen im Rahmen der Lerngelegenheit berichteten, $B = 0.014$, $t(167) = 2.57$, $p = .01$. Insgesamt waren die Effekte klein.

Die berichteten Ergebnisse haben Implikationen für die adaptive Gestaltung von Fortbildungen für ältere Lehrkräfte zum Thema KI. Eine besonders motivationsförderliche Gestaltung (affektive Komponente) sowie eine didaktische Reduktion, die den wahrgenommenen mental load verringert, sollten angestrebt werden.

Effekte der wahrgenommenen Nützlichkeit, subjektiven Norm und Selbstwirksamkeitsüberzeugungen auf die Nutzungsdauer von KI-Tools bei Lehramtsstudierenden

Fabian Hoya¹, Kira Weber², Christopher Neil Prilop³, Dana-Kristin Mah⁴, Lucas Jacobsen⁴

¹Pädagogische Hochschule Vorarlberg; ²Universität Hamburg; ³Aarhus University; ⁴Leuphana Universität Lüneburg

Dem Einsatz von Künstlicher Intelligenz (KI), verstanden als „Computerprogramme und Algorithmen, die menschliche Fähigkeiten wie Lernen, Problemlösungen und Entscheidungsfindungen nachahmen“ (Nepper & Ruch 2023, S. 3), wird in der Lehrer:innenbildung ein großes Potenzial zugesprochen. KI-Tools bieten eine Vielzahl an innovativen Einsatzmöglichkeiten, um Lernende stärker individuell zu fördern oder um Lehrkräfte in ihrer pädagogischen Praxis zu unterstützen und sie zu entlasten (vgl. Autor:innen, 2023; Martin et al. 2024). Dennoch zeigt sich, dass KI-Tools gegenwärtig noch vergleichsweise selten von praktizierenden Lehrkräften eingesetzt werden (vgl. z. B. Galindo-Dominguez et al., 2024). Aus diesem Grund bedarf es empirisch gesicherter Erkenntnisse über Faktoren, die das KI-Nutzungsverhalten von (angehenden) Lehrkräften tangieren, um diese entsprechend fördern zu können. Folgt man der Theorie des geplanten Verhaltens von Ajzen (1985), kann die Absicht zur Durchführung eines Verhaltens und damit verbunden auch das tatsächlich von Personen gezeigte Handeln über ihre Einstellungen sowie die von ihnen wahrgenommenen Erwartungen signifikanter Bezugspersonen (subjektive Norm) und Möglichkeiten zur Beeinflussung des Verhaltens (Verhaltenskontrolle) erklärt werden. Mit Blick auf den gegenwärtigen empirischen Forschungsstand zeigt sich, dass in verschiedenen Untersuchungen signifikante Effekte der Einstellungen, subjektiven Norm oder Variablen der Verhaltenskontrolle auf die Verhaltensintention zum KI-Einsatz der befragten Teilnehmer:innen belegt werden konnten (vgl. z. B. An et al. 2023; Gado et al. 2022; Ivanov et al. 2024; Sing et al. 2022; Zhang et al. 2023). Allerdings wurden in den vorliegenden Studien bislang kaum Indikatoren für das tatsächlich von Personen gezeigte Verhalten einbezogen. An dieses Forschungsdesiderat knüpft unsere Untersuchung an. Im Rahmen einer quantitativen Querschnittsstudie wurden $N=931$ Lehramtsstudierende (192 männlich, 729 weiblich, 7 divers, 3 fehlende Angaben) aller Schulformen in Deutschland befragt. Diese verteilten sich auf 626 Grundschullehramtsstudierende, 58 Sonderpädagogikstudierende, 50 Studierende der Sekundarstufe I, 88 Studierende für das Lehramt an Gymnasien und 105 berufsbildende Student:innen (4 fehlende Angaben). Zum Zeitpunkt der Befragung befanden sich die Teilnehmer:innen durchschnittlich etwa im 4. Fachsemester ($M=3,71$; $SD=2,85$; $Min=1,00$; $Max=17,00$). 735 Student:innen waren zwischen 18 und 24, 177 zwischen 25 und 39 sowie 18 zwischen 40 und 69 Jahren alt (1 fehlende Angabe). Die Studierenden gaben in einem Online-Fragebogen Auskunft zu der von ihnen perzipierten Nützlichkeit von KI-Tools (4 Items; $M=4,16$; $SD=1,02$; $Min=1,00$; $Max=6,00$; $\alpha=.83$; Beispiellitem: Das Erlernen von KI wird mir bei meinem (zukünftigen) Job helfen.), der wahrgenommenen subjektiven Norm (2 Items; $M=2,87$; $SD=1,18$; $Min=1,00$; $Max=6,00$; $\alpha=.84$; Beispiellitem: Menschen, die mein Verhalten beeinflussen, sind der Meinung, dass ich mich mit KI-Tools beschäftigen sollte.), ihren Selbstwirksamkeitsüberzeugungen (4 Items; $M=3,85$; $SD=0,89$; $Min=1,00$; $Max=6,00$; $\alpha=.88$; Beispiellitem: Ich kann KI-bezogene Tools gut nutzen.) sowie zu ihrer Verhaltensabsicht zur KI-Nutzung (3 Items; $M=3,99$; $SD=0,97$; $Min=1,00$; $Max=6,00$; $\alpha=.83$; Beispiellitem: Ich werde auch in Zukunft KI nutzen.). Es wurden bestehende Instrumente von Ng et al. (2023) sowie Venkatesh und Davis (2000) genutzt, die aus dem Englischen übersetzt wurden. Allen Items lag eine 6-stufige Likert-Skala zugrunde (1=stimme überhaupt nicht zu bis 6=stimme voll und ganz zu). Des Weiteren wurden die Proband:innen mittels eines selbst entwickelten 7-stufigen Erhebungsinstrumentes (1=0h; 2=<1h; 3=1-3h; 4=4-7h; 5=8-11h; 6=12-15h; 7=>15h) darum gebeten, Stellung zur durchschnittlichen wöchentlichen Nutzungsdauer von KI-Tools zu nehmen ($M=2,06$; $SD=0,77$; $Min=1,00$; $Max=7,00$). Die Ergebnisse eines in Mplus berechneten Strukturgleichungsmodells belegen bei einem zufriedenstellenden Modellfit ($CFI=.95$; $TLI=.93$; $RMSEA=.064$; $SRMR=.038$), dass die Verhaltensabsicht der Studierenden zur zukünftigen Nutzung von KI-Tools signifikant über die von ihnen perzipierte Nützlichkeit ($\beta=.54^{***}$), die von ihnen wahrgenommene subjektive Norm ($\beta=.16^{***}$) sowie ihre Selbstwirksamkeitsüberzeugungen ($\beta=.35^{***}$) erklärt werden kann. Darüber hinaus erweisen sich die Verhaltensabsicht der

Studierenden ($\beta=.38^{***}$) und ihre Selbstwirksamkeitsüberzeugungen ($\beta=.17^{***}$) als signifikante Prädiktoren zur Vorhersage ihrer durchschnittlichen wöchentlichen KI-Nutzungsdauer. Die Ergebnisse unserer Untersuchung liefern wichtige Erkenntnisse für eine stärkere Integration von KI-Tools in der Lehrer:innenbildung, die in weiterführenden Studien im Längsschnitt auf mögliche Ursache-Wirkungs-Zusammenhänge untersucht werden müssten.

ENTFÄLLT: Einstellungen und Selbstwirksamkeitserwartungen von Lehrkräften gegenüber der Nutzung digitaler Medien. Eine repräsentative Studie auf Basis des Deutschen Schulbarometers

Hermann Dzingel¹, Eric Richter¹, Uta Klusmann², Dirk Richter¹

¹Universität Potsdam, Deutschland; ²Humboldt-Universität zu Berlin

Der rasante digitale Wandel der Gesellschaft verändert den Bildungsbereich nicht erst seit Kurzem. So wurde jedoch durch aktuelle technische Entwicklungen die Transformation der Schulen nochmals beschleunigt (Richter, Lazarides & Scheiter, 2024). Schülerinnen und Schüler nutzen digitale Medien in ihrer Freizeit, werden bei dieser Nutzung jedoch nicht immer ausreichend durch die Angebote der Schule medienpädagogisch begleitet (Schenk, 2023). Lehrkräfte und ihre digitalen Kompetenzen spielen eine zentrale Rolle bei der effektiven Integration von Technologie in den Unterricht und in der Medienerziehung der Schülerinnen und Schüler (Fernández-Batanero et al. 2022; Scheiter & Gogolin 2023). Gerade wenn Lehrkräfte geringere digitale Kompetenzen zeigen, zögern sie eher, Technologie vollständig in ihre pädagogischen Methoden zu integrieren, und verwenden sie eher als Ergänzung traditioneller Methoden (Butler et al., 2018; Starkey, 2020). Somit sind Einstellungen und Selbstwirksamkeitserwartungen gegenüber digitalen Medien laut aktueller Forschung u.a. entscheidend für die Integration digitaler Medien in den Unterricht und können durch Professionalisierungsmaßnahmen verändert werden (Drossel et al., 2019; Pozas & Letzel, 2023; Runge et al., 2023). Der derzeitige Forschungsstand bietet wenige aktuelle sowie repräsentative Erkenntnisse, die Einstellungen und Selbstwirksamkeitserwartungen von Lehrkräften gegenüber digitalen Medien aufzeigen sowie Zusammenhänge zwischen Fortbildungsaktivitäten und diesen Kompetenzfacetten untersuchen (Drossel et al., 2019). Diese Forschungslücke versucht die vorliegende Studie zu füllen.

Fragestellung

Folgende spezifische Forschungsfragen stehen im Fokus der Studie:

1. Welche Einstellungen und welches Selbstwirksamkeitsempfinden haben Lehrkräfte gegenüber der professionellen Nutzung digitaler Medien?
2. Wie können sowohl personen- und schulbezogene Merkmale als auch die Teilnahme an formellen, informellen sowie Lernaktivitäten mit Bezug zu digitalen Medien Einstellungen und das Selbstwirksamkeitserwartungen von Lehrkräften vorhersagen?

Methode

Die Daten wurden durch eine repräsentative Befragung unter 1608 Lehrkräften durch das Institut forsa erfasst. Die Erhebung fand zwischen dem 16. November 2023 und dem 3. Dezember 2023 statt und zielte auf Lehrkräfte der Primar- und Sekundarstufe sowie auf alle Schularten ab. Zur Messung der Einstellungen und des Selbstwirksamkeitsempfindens wurden bestehende Skalen verwendet und zum Teil leicht abgewandelt (Carstensen et al, 2022). Die Skalen umfassen 4 bzw. 5 Items und wurden auf einer vierstufigen Likert-Skala eingeschätzt. Die Reliabilität der Skalen erwies sich zufriedenstellen (Einstellungen: $\alpha = .80$, Selbstwirksamkeit: $\alpha = .85$). Zusätzlich wurden die Teilnahme an formellen und informellen Lernaktivitäten sowie demografische Variablen wie Berufserfahrung, Geschlecht, Schulart und Fächerkombination erfasst. Die Daten wurden mittels multipler linearer Regression analysiert, wobei separate Modelle für Einstellungen und Selbstwirksamkeit gegenüber der Nutzung digitaler Medien erstellt wurden.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zu Forschungsfrage 1 zeigen, dass die Einstellungen gegenüber digitalen Medien ($M= 2.73$) und Selbstwirksamkeitserwartung ($M = 2.73$) mittleren Bereich ausgeprägt sind.

Die Befunde zu Forschungsfrage 2 machen deutlich, dass die Berufserfahrung negativ mit den Einstellungen als auch mit der Selbstwirksamkeitserwartung der Nutzung digitaler Medien zusammenhängt. Lehrkräfte mit längerer Berufserfahrung zeigen weniger positiv ausgeprägte Einstellungen und Selbstwirksamkeitserwartungen gegenüber der Anwendung digitaler Medien. Die Teilnahme an Lernaktivitäten mit Bezug zu digitalen Medien hängt weiterhin signifikant positiv mit Einstellungen und Selbstwirksamkeit zusammen. Insbesondere formelle Lernaktivitäten stehen signifikant positiv in Beziehung mit der Selbstwirksamkeitserwartung, während informelle Lernaktivitäten sowohl mit Einstellungen als auch mit Selbstwirksamkeit positiv in Beziehung stehen.

Die Studie kann zeigen, dass formelle und informelle Lernaktivitäten einen jeweils eigenen Beitrag zur Professionalisierung von Lehrpersonen leisten können, um Lehrkräften die notwendigen Kompetenzen im Umgang mit digitalen Medien zu vermitteln.

PS-15: Sportunterricht als sozialer Kontext

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 169

Chair der Sitzung: Tino Edelmann

Zielorientierungen in Peer-Netzwerken: Die Rolle multidimensionaler Leistungsziele im Sportunterricht

Annabell Schüßler, Cornelius Holler, Yannick Hill, Rüdiger Heim

Heidelberg University, Deutschland

Theoretischer Hintergrund

Das Streben nach Kompetenz, oft als Verfolgung von Zielorientierung in der Forschungsliteratur verhandelt, manifestiert sich als omnipräsenter Aspekt unseres täglichen Lebens (Weissman & Elliot, 2023). Leistungs-Zielorientierungen fungieren als Leitlinie für individuelles Handeln, sei es durch das Bestreben, eine Aufgabe besonders gut meistern zu wollen (task-approach), die vergangene eigene Leistung zu optimieren (self-approach) oder andere zu übertreffen (other-approach). Diese Annäherungs-Zielorientierungen finden ihre Ergänzung in den korrespondierenden Vermeidungs-Zielorientierungen (task-, self-, other-avoidance) (Mascret et al., 2015; Van Yperen, 2022). Die Forschungshistorie verdeutlicht, dass das Konstrukt der Leistungs-Zielorientierung bisher vornehmlich den Erfolg auf individueller Ebene in den Blick genommen hat (z.B. Elliot, 1999; Hulleman et al., 2010). Allerdings betont die jüngere Forschung auch die sozialen Kontexte, indem sich Peers aufgrund ähnlicher Zielorientierungen zusammenschließen. Insbesondere scheinen Kinder, die ähnliche Zielorientierungen verfolgen, miteinander befreundet zu sein, ein Phänomen, das als Homophilie bekannt ist (z.B. Laninga-Wijnen et al., 2017). Darüber hinaus können Heranwachsende von Peers angezogen werden, die eine besonders hohe Zielorientierung aufweisen, wodurch sie eigene individuelle Ziele erreichen könnten. Im Gegenzug könnten Schüler*innen Freunde für Kooperationen bevorzugen, um von den Vorteilen wie Sympathie, Loyalität, Vertrauen und gegenseitiger Unterstützung zu profitieren (Bukowski et al., 2011).

Fragestellung

Die vorliegende Studie untersucht daher, ob Schüler*innen im Sportunterricht eine stärkere Präferenz für Kooperationen zeigen, wenn sie Ähnlichkeiten in Zielorientierungen bzw. höhere Zielorientierungen aufweisen oder wenn sie über Freundschaften berichten.

Methode

Zur Erfassung der Zielorientierung im Sportunterricht setzten wir den 3x2 Achievement Goal Questionnaire (AGQ) für Sport (Mascret et al., 2015) ein, der Aufgaben-, Selbst- und Fremdziele mit Annäherungs- und Vermeidungswertung abbildet. Neben soziodemografischen Daten wie dem Geschlecht erfassten wir zusätzlich mit dem Peer-Nominationsverfahren (Wasserman & Faust, 1994) verschiedene Arten von Peer-Beziehungen, einschließlich Freundschaften und kollaborative Lernaktivitäten wie Teampartner*innen innerhalb eines Sportspiels. In unserer Querschnittsstudie erhoben wir Daten von 364 Schüler*innen in 16 Lerngruppen im Alter von 10 bis 16 Jahren während ihres Sportunterrichts. Zur Analyse wurden Exponential Random Graph Models (ERGMs) eingesetzt, um einen tieferen Einblick in die Rolle der Zielorientierung für die Netzwerkstruktur der Teamkamerad*innenwahl bei Sportspielen zu gewinnen (Lusher et al., 2013).

Ergebnisse und ihre Bedeutung

Unsere Befunde illustrieren primär, dass Schüler*innen im Rahmen eines Sportspiels Gleichaltrige präferieren, die ähnliche Leistungs-Zielorientierungen im Sportunterricht aufweisen. Des Weiteren bevorzugen Schüler*innen Kollaborationen mit Freund*innen und wählen häufig Gleichaltrige desselben Geschlechts, wobei Jungen häufiger als Mädchen ausgewählt werden. Wenn Schüler*innen im Rahmen eines Sportspiels nicht ihre Freund*innen wählen, dann neigen sie dazu, Gleichaltrige mit einem höheren Grad an Zielorientierung zu präferieren, insbesondere für Ziele, die auf den Sieg ausgerichtet sind. Dieses Auswahlmuster betont die Wettbewerbsnatur von Sportspielen und die Priorisierung von Schüler*innen, die Freundschaften miteinander pflegen und solchen, die starke vergleichsorientierte Zielausrichtungen zeigen, im Vergleich zu jenen mit ähnlichen oder geringen Zielorientierungen. Allerdings zeigte sich auch, wenn Schüler*innen für Sportspiele auf Basis von Freundschaft wählen, dass diese Freund*innen niedrige Leistungs-Zielorientierungen aufweisen können. Dies weist auf eine puffernde Wirkung der Freundschaft gegenüber leistungsbezogenen Aspekten hin. Die Befunde unterstreichen die Komplexität der Teambildung im Sportunterricht, wobei sowohl persönliche Beziehungen wie Freundschaft als auch der Siegeswille ausbalanciert werden müssen. Die Ergebnisse liefern erste Hinweise zum Einsatz von Sozial- und Organisationsformen für (Sport)-Lehrkräfte, um die Unterrichtsqualität nachhaltig zu optimieren.

Exponiertheit, die Bedeutung von Sozialbeziehungen und das emotionale Erleben von Sportunterricht bei Schüler:innen der Sekundarstufe I.

Tino Edelmann, Arno Müller, Susanne Jurkowski

Universität Erfurt, Deutschland

Theoretischer Hintergrund

Im Sportunterricht sind die Körper von Schüler:innen und deren Bewegungen zentraler Gegenstand des unterrichtlichen Geschehens und Bezugspunkt für Bewertungen vielfältiger Art (Böhle & Hunger, 2017). So sind Schüler:innen für Andere, je nach Art der Unterrichtsgestaltung mehr oder weniger exponiert. Diese Exponiertheit geht mit Gelegenheiten für soziale Bewertungsprozesse einher, die potenziell das emotionale Erleben im Sportunterricht beeinflussen können (Åsebø et al., 2022). Das Erleben von Sportunterricht hat nicht nur Einfluss auf das Engagement der Schüler:innen im Sportunterricht (Pollmann, 2022), sondern auch auf ihre Teilnahme an außerschulischen sportlichen Aktivitäten (Cardinal et al., 2013; Jaakkola et al., 2017). Negative emotionale Erfahrungen im Sportunterricht, die mit Exponiertheit in Verbindung stehen, können so prägend sein, dass die Freude an Bewegung langfristig beeinträchtigt ist (Ladwig et al., 2018). Positive Beziehungsqualitäten zwischen den Schüler:innen und ihrer unterrichtenden Sportlehrkraft, sowie zwischen den Schüler:innen und ihren Peers erscheinen vor diesem Hintergrund besonders wichtig (Heemsoth, 2014). Im Beitrag wird eine Studie vorgestellt, die sich der Frage nach den Zusammenhängen zwischen Exponiertheit vor den Mitschüler:innen, Beziehungsqualitäten und emotionalem Erleben im Sportunterricht widmet. Ferner wird die moderierende Rolle der Sozialbeziehungen für das emotionale Erleben in exponierten Sportunterrichtsphasen beleuchtet.

Methode

82 Schüler:innen aus vier Klassen der Jahrgangsstufen fünf bis sieben ($M = 11.65$, $SD = 0.91$, 56,1% weibl.) bearbeiteten Fragebögen zur Qualität der von ihnen wahrgenommenen Peerbeziehungen und der Beziehung zu ihrer Lehrperson (Heemsoth

& Miethling, 2012). Exponiertheit wurde messwiederholt für mehrere Unterrichtsphasen aus insgesamt 12 Sportstunden anhand eines Kodiersystems auf der Grundlage von Beobachtungen beurteilt. Darüber hinaus berichteten die Schüler:innen für jede Unterrichtsphase über ihr emotionales Erleben (Zufriedenheit, Freude, Wut/Ärger).

Ergebnisse

Mehrebenenanalysen zeigen, dass höhere Exponiertheit mit weniger Zufriedenheit ($f^2 = 0.13$) und Freude ($f^2 = 0.05$) und mehr Wut/Ärger ($f^2 = 0.10$) zusammenhängt. Günstige Sozialbeziehungen wirken sich hingegen positiv auf das emotionale Erleben aus. Schüler:innen mit einer positiveren Beziehung zur Lehrperson berichten eher von Freude ($f^2 = 0.08$) und weniger von Ärger/Wut ($f^2 = 0.08$) als Schüler:innen mit einer weniger positiven Beziehung zur Lehrkraft. Die Qualität der Peerbeziehungen steht in positivem Zusammenhang mit Zufriedenheit ($f^2 = 0.19$) und Freude ($f^2 = 0.17$) im Sportunterricht. Eine positive Beziehung zur Lehrkraft moderiert zudem die Effekte der Exponiertheit auf das emotionale Erleben in dem Sinne, dass auch bei stärkerer Exponiertheit mehr Freude ($\Delta\omega^2 = 0.28$) und weniger Wut/Ärger ($\Delta\omega^2 = 0.25$) erlebt werden.

Diskussion

Angeichts der hohen gesundheitlichen Bedeutung sportlicher Betätigung und des Anspruchs von Sportunterricht, diese bei Heranwachsenden zu fördern, gewinnen Befunde, die das emotionale Erleben einschließen, besondere Relevanz. Freude und Zufriedenheit im Sportunterricht, welche durch den Grad der Exponiertheit mitbestimmt werden, stehen in direktem Zusammenhang mit der gesundheitsrelevanten sportlichen Aktivität in der Freizeit (Sanchez-Oliva et al., 2014; Woods et al., 2012). Die Ergebnisse der hier vorgestellten Studie zeigen, dass eine höhere Exponiertheit zu weniger Zufriedenheit und Freude und mehr Ärger/Wut bei den Lernenden führt. Die Lehrperson-Schüler:innen-Beziehung erweist sich dabei als zentraler Kontextfaktor für den Erhalt positiven emotionalen Erlebens unter besonders exponierten Bedingungen im Sportunterricht. Die Erkenntnisse zur Exponiertheit und zum Pädagog:inneneffekt sollten bei der Planung und Gestaltung von Sportunterricht berücksichtigt werden.

Leistungsperspektiven und ihre Rollen für Peernetzwerke im Sportunterricht

Cornelius Holler, Annabell Schüßler, Rüdiger Heim

University of Heidelberg, Deutschland

Die sportliche Leistungsfähigkeit im Schulsport spielt eine wichtige Rolle für die soziale Eingebundenheit in fachspezifischen Peerbeziehungen sowie für Anerkennungs- und Missachtungsprozesse in der Sporthalle (Grimminger, 2012). Hierbei kommt der Bewertung der Leistungsfähigkeit der Mitschüler:innen insbesondere im Vergleich zur eigenen Leistungsfähigkeit besondere Bedeutung zu. So gründen sich Fähigkeitsselbstkonzepte (FSK) nicht nur auf temporalen und dimensional, sondern insbesondere auf sozialen Vergleichen (Marsh, 1990), wobei eine systematische Verzerrung im Sinne des *Big-Fish-Little-Pond-Effektes* auch im Sportunterricht auftritt (Gerlach, 2006). Während das FSK sowohl ein normatives Ziel von Sportunterricht darstellt und vielfach Zusammenhänge mit sportunterrichtlichen Leistungen interessiert haben, wurden Peer-Leistungseinschätzungen bisher nicht thematisiert. Hierbei stellt sich die Frage, inwieweit die Schüler:innen ihre eigene sowie die Leistungsfähigkeit ihrer Mitschüler:innen akkurat einschätzen können, und welchen Einfluss die wahrgenommene Leistungsfähigkeit auf sportunterrichtsspezifische Peerbeziehungen hat.

Der vorliegende Beitrag untersucht, inwiefern sportunterrichtliches FSK sowie Peer- und Lehrkräfteeinschätzungen der sportlichen Leistungsfähigkeit untereinander sowie mit der sozialen Eingebundenheit in fachspezifischen instrumentellen Netzwerken zusammenhängen und inwiefern diese Perspektiven die soziale Eingebundenheit in diesen Kontexten aufklären. Hierfür wurden 302 Schüler:innen aus 14 Klassen der 5.-9. Jahrgangsstufe und ihre jeweiligen Sportlehrkräfte mittels Fragebögen befragt. Neben dem sportunterrichtlichen Fähigkeitsselbstkonzept und der Lehrkraftbewertung der Leistung wurden soziale Netzwerke und die Peer-Einschätzungen über Nominierungsfragen erhoben. Die Schüler:innen gaben an, welche Mitschüler:innen sie mögen (Sympathie) sowie mit welchen Mitschüler:innen sie bei einem Spiel im Sportunterricht in ihrer Mannschaft haben möchten (Sport Spiel). Für die Peerbewertung konnten sowohl bis zu fünf besonders leistungsstarke als auch -schwache Mitschüler:innen angegeben werden.

Für die Auswertungen werden sowohl Zusammenhänge zwischen attributbasierten Daten in den Blick genommen, wobei die Peernominierungen der Leistungsfähigkeit in einen Gesamtscore umgewandelt wurden, als auch netzwerkanalytische Zusammenhänge zwischen den Peernominierungen der Leistungsfähigkeit und den Peerbeziehungen im Sport Spiel mittels exponential random graph models (ERGM, Lusher & Robins, 2013).

Die Ergebnisse zeigen erwartbar hohe Übereinstimmungen zwischen FSK und Lehrkraftbewertung und sehr große Korrelationen zwischen Lehrkraftbewertung und Peer-Score. Die Ergebnisse der ERGMs zeigen den signifikanten Einfluss der Leistungsfähigkeit auf die sportunterrichtlichen Peerbeziehungen. Während das Fähigkeitsselbstkonzept und die Lehrkraftbewertung keine signifikanten Zusammenhänge mit den Peerbeziehungen aufweisen, werden Schüler:innen mit einem hohen sowie ähnlich hohem Peer-Score häufiger für die Mannschaft im Sport Spiel gewählt. Des Weiteren standen die individuellen Nominierungen der Leistungswahrnehmung signifikant mit der Netzwerkstruktur in Zusammenhang. Diese Zusammenhänge blieben bestehen, selbst wenn für den signifikanten Einfluss der Sympathie und des Geschlechts kontrolliert wurde.

Die Ergebnisse unterstreichen die wichtige Rolle von Peernormen, wie z.B. der sportlichen Leistungsfähigkeit, für die Beziehungen zwischen Gleichaltrigen im Sportunterricht.

Professionelle Wahrnehmung pädagogischer Qualität fördern – Wirkungsanalyse verschiedener Seminarkonzepte im Sportstudium

Kathrin Kohake

Universität Münster, Deutschland

Einleitung und theoretischer Hintergrund

In verschiedenen Kompetenzmodellen wird die professionelle Wahrnehmung von Lehrkräften als medierende Variable zwischen professioneller Kompetenz und beobachtbarem Lehrverhalten angenommen (Blömeke et al., 2015). Die menschliche Wahrnehmung ordnet Sinnesindrücke anhand hierarchischer, kategorialer Systeme, auf die sie ständig zurückgreift. Professionen entwickeln spezielle Klassifizierungen („Coding Schemes“), die fachlich begründeten Relevanzstrukturen entsprechen (Goodwin, 1994). Diese Systeme müssen im Rahmen der Professionalisierung erlernt werden (Bromme, 1992). Professionelle Wahrnehmung kann verstanden werden als Ergebnis der Aneignung eines fachlich relevanten Kategoriensystems. Der Lernaufwand für Kategorien auf der Verhaltensebene ist geringer als auf tieferen Ebenen (Casale et al., 2012). Deshalb

sind Beobachtungssysteme mit diesem Fokus reliabler; für Fortschritte von Lernenden sind jedoch Tiefenmerkmale des Unterrichtens relevanter (Lotz et al., 2013). Elaborierte Beobachtungssysteme ordnen hierarchisch Oberflächenmerkmale solchen Tiefenkonzepten zu (Prätorius & Charalambous, 2018). Die Beobachtung von Unterricht und Training wird durch extremen Zeit- und Komplexitätsdruck erschwert: Relevante Ereignisse geschehen simultan, dynamisch und komplex vernetzt.

Für die Auswahl von Tiefenkonzepten sowie zugehörigen Verhaltensmerkmalen wurde im vorliegenden Projekt basierend auf dem „Classroom Assessment Scoring System“ (Pianta et al., 2008) das Beobachtungssystem „Pädagogische Qualität im Kinder- und Jugendsport (PQ-Sport)“ verwendet (Richartz et al., 2021). Um das Erlernen dieser kategorialen Ordnung zu fördern, wurde das instruktionale Design systematisch an der Dosierung der kognitiven Belastung („cognitive load“) orientiert (Sweller et al., 2019). Folgende Forschungsfrage wurde untersucht: Wie verändern sich die 1) internalisierte kategoriale Klarheit und 2) Reliabilität professioneller Wahrnehmung in Abhängigkeit unterschiedlicher Seminarformate?

Methode

In einem quasi-experimentellen Prä-Post-Test-Design lernten Studierende in drei Seminartypen sechs Dimensionen pädagogischer Qualität auf Basis des PQ-Sport-Instruments kennen. In Seminartyp A ($N = 64$) wurden ein Manual (Klärung und Schärfung von Kategorien), Beobachtungsbögen (gezielte Aufmerksamkeitslenkung) und Videoaufnahmen (Übungsaufgaben) von Sportunterricht und Training eingesetzt. Teilnehmer*innen in Seminartyp B ($N = 36$) erhielten das Manual und Beobachtungsbögen, verwendeten diese jedoch nach Klärung der Konzepte ohne videobasierte Übung und Feedback beim selbstregulierten Lernen im Praktikum. In Seminartyp C ($N = 20$) erlernten und reflektierten Studierende exakt dieselben Tiefenkonzepte an Videobeispielen, aber ohne Manual, Beobachtungsbögen und Feedback. Die Seminare unterscheiden sich damit in kategorialer Klarheit (A, B, C⁻) und kognitiv dosierter Übung (A, B⁻, C⁻). Prä- und Post-Beobachtungsurteile wurden mittels Videobeispielen und Fragebögen erhoben (Übereinstimmung mit Master-Codes als Percentage-Within-One). Internalisierte kategoriale Klarheit (in Prozent korrekter Antworten) wurde mittels eines Tests zur Zuordnung schriftlicher Beispiele zu sechs Dimensionen pädagogischer Qualität ermittelt. Die Daten wurden varianzanalytisch ausgewertet.

Ergebnisse und Diskussion

Am Ende des Semesters ordneten Studierende des Seminartyps A im Durchschnitt 70 % ($SD = 11$ %) der Beispiele der korrekten Dimension zu und erzielten damit die besten Ergebnisse in der internalisierten kategorialen Klarheit, gefolgt von Teilnehmer*innen des Seminartyps B mit 65 % ($SD = 13$ %) und C mit 63 % ($SD = 19$ %). Ein signifikanter Gruppeneffekt ($F(2,114) = 3,14$; $p < ,05$; *part. $\eta^2 = ,05$*) konnte mittels Tukey-Kramer-Post-Hoc-Tests nicht bestätigt werden.

Die Ergebnisse zur Reliabilität professioneller Wahrnehmung zeigen neben signifikanten Haupteffekten für Zeit und Gruppe einen signifikanten Interaktionseffekt mittlerer Größe für Zeit x Gruppe ($F(2,232) = 10,64$, $p < 0,001$, *part. $\eta^2 = 0,08$*). Die Tukey-Kramer-Post-Hoc-Analyse bestätigt die erwarteten Unterschiede: Die professionelle Wahrnehmung der Studierenden im Seminartyp A änderte sich signifikant von Beginn zu Ende des Semesters, nicht jedoch für die der Seminartypen B und C. Zusätzlich übertrafen im Post-Test die Studierenden des Seminartyps A ($M = 82$ %; $SD = 16$ %), ihre Kommilitonen aus den Typen B ($M = 63$ %; $SD = 19$ %) und C ($M = 55$ %; $SD = 24$ %).

Die Ergebnisse bestätigen die angenommene Bedeutung sowohl der Vermittlung klarer Beobachtungskategorien als auch kognitiv dosierter Übung. Konsequenzen für die Ausbildung von Studierenden werden vor dem theoretischen Hintergrund diskutiert.

PS-16: Inklusion und Teilhabe im Bildungssystem: Rahmenbedingungen und Perspektiven

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 382

Chair der Sitzung: Julia Wittig

Werden soziale Teilhabe und weitere inklusive Bildungsziele besser an verschiedenen Formen von Ganztagsschulen erreicht als an Halbtagschulen?

Julia Wittig, Cornelia Gresch, Lena Külker

Humboldt-Universität zu Berlin, Deutschland

Theoretischer Hintergrund

Mit der im Jahr 2009 in Deutschland in Kraft getretenen Behindertenrechtskonvention der Vereinten Nationen (UN-BRK) gehen notwendige Änderungen für ein inklusives Bildungssystem einher. Artikel 24 der UN-BRK konkretisiert soziale Teilhabe, Menschenrechtsbildung, Persönlichkeitsentfaltung und Empowerment als Ziele inklusiver Bildung (vgl. Piezunka et al., 2018). Ähnliche Ziele hat auch der Ausbau der Ganztagsschule in Deutschland (Schurig & Gebhardt, 2020): die Ganztagsschule soll alle Schüler:innen individuell fördern, soziale Teilhabe ermöglichen, Demokratiebildung unterstützen und die Persönlichkeitsentfaltung und Selbstentwicklung anregen (BMBF, 2003). Angelehnt an das Kontext-Input-Prozess-Output-Modell (Scheerens, 1990) und die Theorie der Ökologie der menschlichen Entwicklung (Bronfenbrenner, 1981) argumentieren wir, dass sowohl offene als auch teil- und vollgebundene Ganztagsschulen aufgrund des erweiterten Zeitrahmens und der Einbindung von weiterem pädagogischen Personal im Vergleich zur auf den Fachunterricht fokussierten Halbtagschule bessere Voraussetzungen haben um inklusive Bildungsziele zu fördern (Klemm & Preuss-Lausitz, 2011), etwa indem Schüler:innen außerhalb des Unterrichts in klassen- und jahrgangsübergreifenden Lern- und Freizeitangeboten weitere Kontakte knüpfen und sich interessenbasiert entfalten können, wobei das pädagogische Personal aushandlungsbereit und unterstützend unter Berücksichtigung der individuellen Voraussetzungen der Schüler:innen agiert.

Fragestellung

Bisher gibt es keine Studien, die den Zusammenhang zwischen schulischer Organisationsform und dem Erreichen inklusiver Bildungsziele gezielt adressieren. Entsprechend wird folgende Frage untersucht: Besteht ein Zusammenhang zwischen der Organisationsform der Schule und dem Erreichen inklusiver Bildungsziele? Damit verbunden wird folgende Hypothese getestet: Schüler:innen zeigen höhere Ausprägungen bei inklusiven Bildungszielen, wenn sie eine Ganztagsschule und insbesondere eine teil- oder vollgebundene Ganztagsschule besuchen und zwar hinsichtlich der sozialen Teilhabe, der Demokratiebildung, der Persönlichkeitsentfaltung und des Empowerments.

Methode

Die Analysen basieren auf Daten von 1097 Schüler:innen der 7.Klassenstufe an inklusiv arbeitenden Schulen der Studie „Inklusion in der Sekundarstufe in Deutschland“ (INSIDE) (Schmitt et al., 2020). Inklusivitätsziele wurden über die Wahrnehmung der Schüler:innen erfasst. Auf einer vierstufigen Likert-Skala (1=überhaupt nicht; 4=voll und ganz) gaben die Schüler:innen an, wie sehr sie verschiedenen Aussagen zum Erleben der Bildungsziele in ihrer Schule zustimmen. Soziale Teilhabe wurde über vier Skalen (Külker et al. 2021) erfasst (α jeweils > 0.740), Demokratiebildung ($\alpha=0.866$), Persönlichkeitsentfaltung ($\alpha=0.808$) und Empowerment ($\alpha=0.744$) mit jeweils einer Skala. Die schulische Organisationsform wurde bei den Schulleitungen erfragt. Es wurden Mehrebenenmodelle (Random-Intercept-Modelle) in der Statistiksoftware Mplus (Version 8; Muthén & Muthén, 2017) unter Verwendung robuster Maximum-Likelihood-Schätzer mit Schüler:innen als Level 1- und Schulen als Level 2-Einheiten modelliert und für den individuellen Hintergrund der Schüler:innen auf Level 1 und die Heterogenität der Schüler:innenschaft auf Level 2 kontrolliert. Fehlende Werte wurden über die Full-Information-Maximum-Likelihood-Methode ersetzt.

Ergebnisse

Die Analysen zeigen, dass mit dem Besuch einer Ganztagsschule kein genereller Vorteil für das Erreichen inklusiver Bildungsziele einhergeht, lediglich die Persönlichkeitsentfaltung wird an teilgebundenen Ganztagsschulen ($b=0.120$, $p=0.038$) höher eingeschätzt als an Halbtagschulen. Schüler:innen an vollgebundenen Ganztagsschulen berichten entgegen der Erwartung eine signifikant geringere Selbstwahrnehmung ihrer sozialen Teilhabe ($b=-0.178$, $p=0.017$), allerdings liegen die Mittelwerte der Skala für die Selbstwahrnehmung der sozialen Teilhabe in allen Organisationsformen deutlich über dem theoretischen Mittelwert von 2.5 (jeweils $M>3.2$), was bedeutet, dass Schüler:innen ihre soziale Teilhabe auch an vollgebundenen Ganztagsschulen sehr positiv einschätzen. Die Ergebnisse können in die generelle Debatte um die Wirksamkeit des schulischen Ganztags eingeordnet werden. Immer wieder wird hier festgehalten, dass weniger die Organisationsform als vielmehr die Qualität der Einzelschule sowohl im Leistungsbereich als auch bei überfachlichen Merkmalen entscheidend ist (Fischer & Kielblock, 2020). Die Befunde bestätigen diese Argumentation auch mit Blick auf inklusive Bildungsziele. Das bedeutet jedoch auch, dass die Organisationsform Halbtags trotz begrenzter Möglichkeiten in der Lage ist inklusive Bildungsziele umzusetzen.

Einfluss des lokalen Förderschulangebots auf den Anteil inklusiv beschulter Kinder im Bundesland Thüringen (Arbeitstitel)

Sebastian Steinmetz

Leibniz Institut für Bildungsverläufe, Deutschland

Regionale Disparitäten der sonderpädagogischen Förderung sind in der Forschungsliteratur gut dokumentiert. Der Anteil der Kinder mit sonderpädagogischem Förderbedarf (SPF) schwankt nicht nur zwischen den Bundesländern sehr stark, sondern auch innerhalb kleinerer Gebietskörperschaften und zwischen einzelnen Schulstandorten (Topsch 1975, Moser/Dietze 2015, Goldan/Grosche 2021). Zwar kann ein Teil dieser Varianz durch die regional unterschiedlich ausgeprägte Heterogenität der Schülerschaft erklärt werden, weil Kinder und Jugendliche aus bildungsarmen und ökonomisch deprivierten Haushalten ein höheres Risiko haben, einen sonderpädagogischen Förderbedarf attestiert zu bekommen und somit Schulen in sozial prekärer Lage überproportional viele Kinder mit Förderbedarf unterrichten (Bos et al. 2010, Möller/Bellenberg 2017, S.47). Die gesamte Varianz der regionalen Disparitäten lässt sich damit jedoch nicht erklären. In der Forschungsliteratur werden deshalb verschiedene Ansätze diskutiert, um die starken regionalen Unterschiede in der sonderpädagogischen Förderung zu plausibilisieren. Mittlerweile belegen Studien einen Zusammenhang zwischen der räumlichen Nähe von Förderschulen und Grundschulen sowie dem Anteil der als förderbedürftig attestierten Kinder und Jugendliche. Goldan und Grosche (2021) zeigen

für die Grundschulen im Bundesland NRW, „dass die Wahrscheinlichkeit einen sonderpädagogischen Förderbedarf diagnostiziert zu bekommen, regional ungleich verteilt ist und dies mitunter davon abhängig ist, ob sich eine Förderschule in der Nähe befindet.“

Dieser „Sog-Effekt“ der Förderschulen kann systemtheoretisch hergeleitet werden (Orthmann-Bless 2007). Die allgemeine Schule und die Förderschule gelten in dieser Perspektive jeweils als Subsysteme, die jeweils auf Komplexitätsreduktion und Selbsterhalt orientiert sind (Luhmann 2000, 2002). Der „Sog-Effekt“ sonderpädagogischer Unterstützungssysteme entsteht demnach einerseits aus dem Selbsterhaltungsinteresse der Förderschulen, die durch eine Ausweitung der Diagnostik an räumlich nahegelegenen Schulstandorten für ausreichend Schülerinnen und Schüler mit SPF sorgen (vgl. Helbig/Steinmetz 2021). Die Grundschulen sind hingegen bestrebt jene Kinder als „förderbedürftig“ zu deklarieren, die von den Normerwartungen der allgemeinen Schule abweichen. Mit der Vergabe eines sonderpädagogischen Förderbedarf, der in der Regel mit einem reduzierten Lehrplan und Leistungsanforderungen einhergeht, wird deren mangelnde Passung zum allgemeinen Schulsystem signalisiert und zusätzliche Ressourcen mobilisiert. Umso geringer die räumliche Distanz zu einem Förderzentrum, desto eher könnte die allgemeine Schule dazu neigen, die sonderpädagogischen Unterstützungsangebote zu beanspruchen (vgl. Goldan/Grosche 2021).

Da die Ausgestaltung der sonderpädagogischen Förderung zwischen den Bundesländern stark variiert, untersucht der vorliegende Beitrag, ob auch im Bundesland Thüringen ein solcher „Sog-Effekt“ nachgewiesen werden kann. Mit Daten der amtlichen Schuldatenbank Thüringens wird untersucht, ob die räumliche Distanz (in km) zwischen einer Förderschule und einer Grund- bzw. Gemeinschaftsschule einen signifikanten Einfluss auf den Anteil der inklusiv beschulten Kinder mit SPF ausübt. Der Datensatz umfasst die Schuljahre 2011 bis 2020, untersucht werden alle Grund- und Gemeinschaftsschulen. Berechnet werden sogenannte Tobit-Regressionsmodelle unter Berücksichtigung der Mehrebenenstruktur der Daten und unter Kontrolle des sozialen Umfeldes der Schulen. Die abhängige Variable ist der „Anteil der Kinder mit SPF“ auf Schulebene, die zentrale unabhängige Variable ist die räumliche Distanz zwischen der allgemeinen Schule und dem nächstgelegenen Förderzentrum. Zur Kontrolle des sozialen Umfeldes werden Daten auf Rasterebene des RWI-Leibniz-Instituts für Wirtschaftsforschung (Breidenbach/Eilers 2018) verwendet. Die Raster (Grids) umfassen eine Fläche von 1 km² und entsprechen nicht den gängigen administrativen Einheiten. Die Verwendung der Raster ermöglicht eine kleinräumige Betrachtung des sozio-ökonomischen Umfeldes der Schulen.

Auch unter Kontrolle des sozialen Umfeldes, des Schulträgers und der Schulform zeigt sich ein signifikanter Distanz-Effekt für die Grund- und Gemeinschaftsschulen in Thüringen. Je näher eine allgemeine Schule an einer Förderschule liegt, desto höher fällt der Anteil der Kinder mit SPF aus. Der Effekt ist besonders stark für Schulen in freier Trägerschaft. Die Befunde haben Implikationen für das meritokratische Bildungsverständnis unseres Schulsystems, da offenbar die Vergabe eines sonderpädagogischen Förderbedarfs nicht nur mit den individuellen Merkmalen der betroffenen Schülerinnen und Schüler zusammenhängt, sondern auch vom Wohnort und der lokalen Angebotsstruktur der Förderschulen beeinflusst wird.

ENTFÄLLT: Schüler:innen mit sonderpädagogischem Förderbedarf im Bereich Lernen (SPF-L) in Deutschland. Strukturelle Veränderungen der Schülerschaft an Förderschulen seit der Ratifizierung der UN-BRK (hinsichtlich verschiedener bildungsrelevanter Indikatoren)

Sandra Hubert, Lena Nusser

Leibniz-Institut für Bildungsverläufe (LifBi), Deutschland

Im Jahr 2009 unterzeichnete Deutschland die Konvention über die Rechte von Menschen mit Behinderungen (UN-BRK). Mit Artikel 24 verpflichtet sich Deutschland, ein inklusives Bildungssystem aufzubauen und Schülerinnen und Schüler (SuS) mit SPF vor Diskriminierung und Marginalisierung zu schützen. Im Vergleich zu SuS an segregierten Förderschulen erreichen SuS mit SPF an Regelschulen – auch unter Berücksichtigung der Eingangsselektivität – höhere Kompetenzwerte und haben eine bessere Lernentwicklung (Kocaj et al. 2014; Krämer et al. 2021; Stranghöner et al. 2017). Außerdem verlassen inklusiv lernende SuS mit SPF die Schule seltener ohne Schulabschluss (Autorengruppe Bildungsberichterstattung 2014, 2024, Hollenbach-Biele & Klemm 2020). Allerdings zeigt sich, nicht zuletzt aufgrund der Rüge im Jahr 2023 durch den zuständigen UN-Fachausschuss, dass einige Bundesländer den Umbau des Schulsystems in Richtung Inklusion mit Zurückhaltung angehen (Committee on the Rights of Persons with Disabilities 2023).

In den vergangenen Jahren erhöhte sich sowohl die Zahl der Diagnosen als auch der Anteil inklusiv lernender SuS. Diese Entwicklung fand an Förderschulen keine gegenläufige Entsprechung. Darüber hinaus zeigen sich neben anderen Indikatoren soziale Selektivitäten bei der Zuweisung eines SPF sowie bei der Entscheidung des Schultyps als auch bei der Bereitstellung notwendiger schulischer Maßnahmen.

Vor diesem Hintergrund möchte der Beitrag erstens quantitative Veränderungen auf der Makroebene für SuS mit SPF insgesamt als auch für SuS mit SPF-L als dem quantitativ bedeutsamsten Förderschwerpunkt in den letzten 20 Jahren und somit sowohl in der Zeit vor als auch nach der Ratifizierung der UN-BRK nachzeichnen. Die Veränderungen auf der Makroebene lassen, in Kombination mit den Selektivitäten der Zuweisung eines SPF und der Schultypentscheidung Konsequenzen auf der Mikroebene vermuten. Daher möchte der Beitrag in einem zweiten Schritt die Komposition der SuS mit und ohne SPF-L an unterschiedlichen Schultypen hinsichtlich sozialstruktureller und bildungsrelevanter Merkmale über die Zeit hinweg vergleichend untersuchen.

Datengrundlage für Fragestellung 1 stellen die Daten von KMK 2024 und destatis 2023 dar. Fragestellung 2 wird anhand der jeweils ersten Erhebungswellen der Startkohorten 3 (2010, n = 6.112 SuS) und 8 (2022, n = 5.192 SuS) des Nationalen Bildungspanels (NEPS, Blossfeld & Roßbach 2019) untersucht. Die SuS befanden sich jeweils in der 5. Klasse. Beide Studien beinhalten ein Oversampling von Förderschulen mit dem Schwerpunkt Lernen und fanden kurz nach der Ratifizierung der UN-BRK sowie etwa eine Dekade später statt und bieten somit eine optimale Datenbasis, um Veränderungen der Struktur von Förderschüler:innen untersuchen zu können. Bildungsrelevante Faktoren, die vergleichbar in beiden Startkohorten erfasst wurden, umfassen neben familialen Merkmalen das schulische Selbstkonzept, die Lesegewohnheiten, die kognitiven Grundfähigkeiten (erfasst mit zwei Subtests des KFT; Gnams & Nusser 2024) sowie die idealistische und realistische Bildungsaspiration (aus der Perspektive der Eltern).

Die Auswertungen offenbaren auf Bundesebene, dass, während sowohl der Inklusionsanteil als auch die Inklusionsquote ansteigen, die Exklusionsquote nicht im vergleichbaren Ausmaß sinkt. Für Fragestellung 2 lassen sich für das Jahr 2010 große Unterschiede zwischen den Schülergruppen festhalten, und zwar sowohl im Hinblick auf die soziale Komposition als auch auf die angeführten bildungsrelevanten Indikatoren. Förderschüler:innen stammen z. B. deutlich häufiger aus bildungs- und vermögensbenachteiligten Elternhäusern. Die Bildungsaspirationen weichen signifikant nach unten ab, überschätzen die spätere Realität aber dennoch bei Weitem.

Befunde zur Veränderung über die Zeit können zum jetzigen Zeitpunkt nicht berichtet werden, aber für die Konferenz in Aussicht gestellt werden, da der SUF der NEPS-SC8 am Ende des Jahres veröffentlicht wird.

Der Zusammenhang zwischen Bildungschancen und sozialer Herkunft scheint in Deutschland stark ausgeprägt zu bleiben, was sich auch an den hier vorgestellten Ergebnissen deutlich zeigt. Für SuS, die eine Förderschule besuchen, kann das langfristige Konsequenzen hinsichtlich ihrer Chancen auf dem Ausbildungs- und Arbeitsmarkt haben. Die Potentiale von Bildung gilt es auch für SuS mit SPF aufzuzeigen und die Anstrengungen der Bundesländer zum Ausbau eines effektiven und gerechten inklusiven Schulsystems sind auszubauen.

Erkenntnisse über die Teilhabe von Menschen mit Behinderung vom Übergang der Förderschule auf den allgemeinen Arbeitsmarkt.

Ann Christin Schulz, Daniel Krüger, Lisa Preissner

Technische Universität Dortmund, Deutschland

Theoretischer Hintergrund

Durch die UN-Behindertenrechtskonvention (UN-BRK) ist die Teilhabe aller Menschen am Arbeitsleben gesetzlich verankert. Nichtsdestotrotz spiegelt dieser Anspruch einer offenen, inklusiven Arbeitswelt nicht die Beschäftigungsrealität wider (York & Jochmaring, 2022). Menschen mit Behinderungen (MmB) sind im Vergleich zu Menschen ohne Behinderungen auf den allgemeinen Arbeitsmarkt unterrepräsentiert (vgl. Bundesagentur für Arbeit, 2024). Weiterhin liegen nicht dieselben Zugänge vor (Kardorff et al., 2013). Das zeigt sich darin, dass eine Vielzahl von MmB in Werkstätten für behinderte Menschen (WfbM) beschäftigt sind (Bundesarbeitsgemeinschaft Werkstätten für behinderte Menschen e.V. [BAG WfbM], 2021). Trotz dass WfbM MmB so fördern und unterstützen sollen, dass ihnen die Teilhabe auf dem allgemeinen Arbeitsmarkt (wieder) ermöglicht wird, wird dieses Ziel kaum erreicht. Eine Befragung von 311 WfbM wies im Jahr 2019 auf eine Übergangsquote von 0,6% aller Teilnehmer*innen von Eingangsverfahren und Berufsbildungsbereich sowie von 0,35% von Personen aus dem Arbeitsbereich hin (Engels et al. 2023). Der durch die UN-BRK politisch und im Bundesteilhabegesetz gesetzlich verankerte Anspruch auf Teilhabe am Arbeitsleben für MmB misslingt dadurch, weshalb Ansätze zur Förderung der Teilhabe notwendig werden. Das vom Bundesministerium für Arbeit und Soziales geförderte Projekt „Arbeiten – wie ich es will“ (AWIEW) adressiert diese Problemlage, indem ein transdisziplinäres Team aus Sozialdesigner*innen, Technikexpert*innen, Sozial- und Rehabilitationswissenschaftler*innen sowie einer Wohlfahrtsorganisation ein digitales Tool entwickelt, das MmB beim Übergang auf den allgemeinen Arbeitsmarkt unterstützt. Grundlage dafür sind partizipative Ansätze, wie das Co-Design (Sanders & Stappers, 2008), das MmB als Co-Forscher*innen einbezieht, oder die eingebettete Forschung (McGinity & Salokangas 2014; Vindrola-Padros et al., 2016). Der Fokus des Einzelbeitrags wird dabei auf letzterem liegen.

Fragestellung

Um eine Erhöhung der Übergangsquoten auf den allgemeinen Arbeitsmarkt zu unterstützen, steht die Frage im Vordergrund, welche förderlichen und hinderlichen Faktoren die berufliche Orientierung von MmB im Kontext einer Wohlfahrtsorganisation beeinflussen. Demzufolge wird der Beitrag einen Einblick über solche Faktoren für den Übergang von MmB auf den allgemeinen Arbeitsmarkt liefern.

Methode

Um die Fragestellung zu beantworten, bedient sich das Team der Sozial- und Rehabilitationswissenschaftler*innen dem Ansatz der eingebetteten Forschung. Dies ist ein Ansatz, in dem Wissenschaftler*innen für einen festgelegten Zeitraum Teil einer Organisation werden (McGinity & Salokangas, 2014; Vindrola-Padros et al., 2016). Dabei ist das Ziel, Einblicke in Beziehungen, Strukturen und weiteren Elementen der beforschten Organisation zu erhalten, die die Forscher*innen durch Beobachtungen nicht erhalten hätten. Aus diesem Grund haben fünf Sozial- und Rehabilitationswissenschaftler*innen jeweils eine Arbeitswoche (fünf Tage) in sechs verschiedenen Arbeitsbereichen der beforschten Organisation verbracht, um tiefgehende Einblicke zur Berufsorientierung von MmB sowie Einflussfaktoren bezüglich ihres Übergangs auf den allgemeinen Arbeitsmarkt zu erhalten.

Ergebnisse und ihre Bedeutung

Ein Arbeitsbereich der eingebettet erforscht wurde, war die Förderschule. Hierbei zeigte sich, dass die Förderschule auf die Berufswelt vorbereitet, indem Praktika angeboten werden. Weitere fördernde Faktoren für den Berufsübergang wurden durch die Berufsberatung der Arbeitsagentur, dem Unterrichtsblock mit der Themenstunde „Fit für das Vorstellungsgespräch“, der Berufsvorbereitenden Bildung (BvB) sowie der gezielten Förderung von Softskills erkennbar. Andersherum wurde deutlich, dass in der Förderschule zumeist der Übergang in die WfbM als der „Goldweg“ vermittelt wurde. Übergänge auf den allgemeinen Arbeitsmarkt wurden den Schüler*innen als zu komplex vermittelt, was durch folgende Aussagen erkennbar wird: *„Das schaffen halt nicht alle.“* oder *„Jedes Schuljahr werden 12 Schüler*innen entlassen. Davon geht aber so gut wie niemand auf den ersten Arbeitsmarkt [...] Wir dürfen halt nicht vergessen, dass wir hier nicht umsonst in einer Förderschule sind.“* MmB sind daher nach der Förderschule zumeist nicht auf dem allgemeinen Arbeitsmarkt beschäftigt. Sie werden vielmehr in WfbM eingegliedert, sodass die durch die UN-BRK postulierte Teilhabe am Arbeitsleben misslingt. Weitere Erkenntnisse, beispielsweise aus Interviews, dem Bereich der Arbeitsmarktintegration und der dort durchgeführten BvB-Maßnahmen sowie ihre Bedeutung sollen auf der GEBF-Tagung vorgestellt und mit anderen Teilnehmer*innen diskutiert werden.

PS-17: Mathematik als Herausforderung: Emotionen, Ängste und Unterstützung im Lernprozess

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: SO 418

Chair der Sitzung: Lena Keller

Mathematik als Stressor: Emotionale und Psychophysiologische Reaktionen angehender Grundschullehrkräfte in einer mathematischen Leistungssituation

Franziska Eckerskorn, Miriam Wünsch, Anne C. Frenzel

Ludwig-Maximilians-Universität, Deutschland

Für viele angehende wie praktizierende Grundschullehrkräfte ist Mathematik ein "notwendiges Übel" ihres Berufs. Studien berichten übereinstimmend, dass geringe Kompetenzüberzeugungen und erhöhtes negatives Emotionserleben in Mathematik, wie z.B. Mathematikangst, insbesondere unter weiblichen Grundschullehrkräften weit verbreitet sind (Artemenko et al., 2021, Beilock et al., 2010, Porsch, 2017, 2019). Die Untersuchung von Lehrkräfteemotionen ist besonders relevant, da die Emotionen von Lehrkräften nicht nur wichtig für ihr eigenes Wohlbefinden und ihre Verweildauer im Beruf (Chang, 2009) sind, sondern weil sie auch in Zusammenhang mit der Qualität des Unterrichts und der Emotionen der Schülerinnen und Schüler stehen (Frenzel et al., 2018; Sutton & Wheatley, 2003). Darüber hinaus kann Mathematikangst bei Lehrkräften potenziell Mathematikangst bei den Lernenden hervorrufen (Hembree, 1990), insbesondere weil Mathematik allgemein als ein Fachbereich bekannt ist, der bei Schülern tendenziell Ängste auslöst (Ashcraft, 2002; Ma, 1999; Wigfield & Meece, 1988), besonders bei Mädchen (z.B. Beilock et al., 2010; Frenzel et al., 2007; Hyde et al., 1990). Eine methodische Schwäche der meisten existierenden Befunde zu negativen Mathematik-Emotionen liegt darin, dass sie die selbstberichteten „Trait-Emotionen“ der Betroffenen erfassen (also deren Überzeugung dazu, wie sie sich „üblicherweise“ in Bezug auf Mathematik fühlen), was zu Verzerrungen ggü. den tatsächlich „in-situ“ erlebten Emotionen führen kann (Goetz et al., 2013). Die vorliegende Studie widmete sich entsprechend der Frage, wie sich angehende Grundschullehrerinnen in einer konkreten Situation fühlen, in der sie mit Mathematik konfrontiert werden. Dazu schufen wir eine mathematische Leistungssituation, um insbesondere die bisher unerforschte Frage zu klären, ob ein niedriges mathematisches Selbstkonzept tatsächlich mit erhöhter negativer Emotionalität und – ergänzend zum Selbstbericht – mit erhöhter psychophysiologischer Erregung bei der Bearbeitung mathematischer Aufgaben einhergeht. Unsere Stichprobe umfasste $N = 70$ angehende Grundschullehrerinnen (Durchschnittsalter = 20,6 Jahre). Die Teilnehmerinnen berichteten vorab ihr mathematisches Selbstkonzept und bearbeiteten dann im Labor fünf Blöcke mit relativ schwierigen mathematischen Aufgaben, während wir kontinuierlich ihre Herzfrequenz und ihren Hautleitwert sowie unmittelbar nach jedem Block ihr subjektives Erleben von Frustration, Stress, Ärger, Angst und Langeweile maßen. Zusätzlich erhoben wir die von den Teilnehmerinnen in der Situation wahrgenommene Schwierigkeit, Herausforderung und ihr Kompetenzerleben nach jedem Block. Wie erwartet zeigten sich mittels regressionsanalytischer Verfahren negative Zusammenhänge (Benjamin-Hochberg-korrigiert) zwischen dem Selbstkonzept und allen negativen selbstberichteten State Emotionen (Frustration: $\beta = -.45$, $p < .001$, Stress: $\beta = -.44$, $p < .001$, Ärger: $\beta = -.35$, $p = .004$, Angst: $\beta = -.26$, $p = .037$, Langeweile: $\beta = -.31$, $p = .013$). Ebenso fanden sich hypothesenkonforme Zusammenhänge zwischen dem Selbstkonzept und der wahrgenommenen Schwierigkeit: $\beta = -.58$, $p < .001$, der wahrgenommenen Herausforderung: $\beta = -.58$, $p < .001$ sowie dem Kompetenzerleben in der Situation: $\beta = .53$, $p < .001$. Die baseline-korrigierten physiologischen Maße während der Aufgabenausführung waren jedoch nicht signifikant mit dem mathematischen Selbstkonzept verbunden (Herzrate: $\beta = .54$, $p = .270$, Hautleitwert: $\beta = .25$, $p = .171$), was auf aufgabenbezogenes kognitives Disengagement (Solhjo et al., 2019, Vanneste et al., 2020) oder methodische Einschränkungen bei der Messung von physiologischer Erregung in Between-Person-Designs zurückzuführen sein könnte (Kiuru et al., 2022). Unsere Ergebnisse unterstreichen die Relevanz von domänenspezifischen Überzeugungen wie dem akademischen Selbstkonzept für situationsspezifische emotional-motivationale Erfahrungen bei angehenden Grundschullehrerinnen. Wenn man unsere Ergebnisse auf den beruflichen Alltag im Mathematikunterricht überträgt, kann man davon ausgehen, dass das Unterrichten von Mathematik für Lehrkräfte mit geringem mathematischen Selbstkonzept wiederholt zu Situationen mit verstärkt negativem emotionalen Erleben führen kann, mit potenziell negativen Folgen für ihr Wohlbefinden und die Lernergebnisse ihrer Schülerinnen und Schüler. Daraus folgt, dass angehende Lehrkräfte im Rahmen ihrer universitären Ausbildung beim Aufbau eines stärkeren mathematischen Selbstkonzepts unterstützt werden sollten (Holzberger et al., 2021).

Feinfühligkeit der Lehrkraft und Mathematikangst von Sekundarschüler:innen: eine Mehrebenenanalyse des Zusammenhangs unter der Berücksichtigung individueller und klassenbezogener Faktoren

Lea Grosser, Paulina Feige, Rainer Watermann

Freie Universität Berlin, Deutschland

Um dem Abwärtstrend der Mathematikleistungen der Schüler:innen in Deutschland (OECD, 2023) zu begegnen, können auch emotionale Faktoren, wie die Mathematikangst, relevant sein (Pekrun, 2000). Dieses relativ verbreitete Phänomen unter Sekundarschüler:innen (OECD, 2013) kann negative Auswirkungen auf die Mathematikleistung zeigen (z. B. Aldrup et al., 2020). Die Mathematikangst besteht aus den getrennten Komponenten Worry (kognitive Sorgenkomponente) und Emotionality (affektive Komponente, z. B. Nervosität) (Liebert & Morris, 1967). Bisherige Studien fokussieren jedoch eher auf die Emotionality-Komponente (Gunderson et al., 2018; Ma & Xu, 2004).

Nach der Kontroll-Wert-Theorie (Pekrun, 2000) kann die Unterrichtsqualität, welche auch die Feinfühligkeit der Lehrkraft hinsichtlich akademischer sowie emotionaler Bedürfnisse der Schüler:innen (Hamre & Pianta, 2007) impliziert, die Kognitionen und Leistungsempfinden der Schüler:innen und damit auch die Leistungen beeinflussen. Theoretisch können emotional feinfühligkeitsvolle Lehrkräfte Angst vermindern und durch akademische Feinfühligkeit Unterrichtsmaßnahmen initiieren (Bowlby, 1969; Pianta, 1999). Bisher wurde die Lehrkraftsensibilität nur vereinzelt explizit und mit einem stärkeren Fokus auf der akademischen Komponente untersucht, wobei sich Hinweise auf die Bedeutung der Lehrkraftsensibilität für die Mathematikangst zeigten (Aldrup et al., 2020; Westphal et al., 2018). Differenzierte Untersuchungen zur Feinfühligkeit der Lehrkraft mit einer Unterscheidung zwischen akademischer und emotionaler Feinfühligkeit fehlen jedoch.

Im Zusammenhang mit der Unterrichtsqualität ist auch der Klassenkontext bedeutsam (Gruehn, 2000); so entsteht im Unterricht ein kollektiv geteiltes Klima, das zu einer hohen Übereinstimmung der Schüler:innenwahrnehmungen führt (Pekrun, 1985). Unklar ist jedoch, inwieweit die Feinfühligkeit der Lehrkraft als geteiltes oder individuelles (psychologisches) Merkmal wahrgenommen wird.

Der vorliegende Beitrag untersucht erstens, welche Bedeutung sowohl die individuelle als auch die in der Klasse geteilte Wahrnehmung der akademischen bzw. emotionalen Feinfühligkeit der Mathematiklehrkraft in Klasse 5 für die Mathematikangst

(Worry und Emotionality) der Schüler:innen in Klasse 5 haben. Zweitens wird untersucht, wie diese Wahrnehmungen in Klasse 5 die Mathematikangst in Klasse 6 beeinflussen.

Die Datengrundlage bildet die Teilstudie BiKS-8-14 des Projekts „Bildungsprozesse, Kompetenzentwicklung, Selektionsentscheidungen im Vorschul- und Schulalter (BiKS)“ (Artelt et al., 2013). Genutzt wurden Welle 4 (Klasse 5, Sekundarstufe) und Welle 5 (Klasse 6, Sekundarstufe). Es gingen $N = 1592$ Schüler:innen aus 159 Klassen in die Analysen ein ($M_{\text{Alter}} = 11.45$ Jahre). 54% der Schüler:innen waren weiblich. Fehlende Werte im Längsschnitt (18% Prozent in Welle 5) wurden unter Beachtung der Mehrebenenstruktur der Daten mit dem R Package *pan* (Schafer & Zhao, 2014) multipel imputiert (vgl. Grund et al., 2016).

Zur Beantwortung der Fragestellungen wurden in R Mehrebenenanalysen durchgeführt (R Core Team, 2022). In allen Analysen wurde für die Mathematikleistung auf Individualebene und die Schulform auf Klassenebene kontrolliert. Es wurden separat für Emotionality und Worry jeweils zwei Modelle gerechnet (also insgesamt vier Modelle). In allen Modellen gingen auf Individualebene beide Dimensionen der Feinfühligkeit (emotional, akademisch) ein. Auf Klassenebene ging jeweils nur eine Dimension der Feinfühligkeit ein. Die Ergebnisse ergaben folgendes Bild: auf der Individualebene war in allen Modellen die emotionale Feinfühligkeit im Hinblick auf Worry und Emotionality prädiktiv, nicht jedoch die akademische Feinfühligkeit. Auf Klassenebene war die emotionale Feinfühligkeit im Hinblick auf Emotionality prädiktiv. Für die akademische Feinfühligkeit zeigten sich auch hier keine Effekte. Die vier Modelle wurden anschließend auch im Längsschnitt gerechnet, wobei jeweils für den Ausgangswert der Mathematikangst kontrolliert wurde. Hier zeigten sich keine Effekte auf die Mathematikangst in Klasse 6.

Die Ergebnisse werden unter Berücksichtigung der Limitationen aufgrund des Studiendesigns im Vergleich zu Ergebnissen bisheriger Studien (z. B. Aldrup et al., 2020; Westphal et al., 2018) diskutiert. Auf Basis bisheriger Ergebnisse deutet sich eine besondere Rolle durch die wahrgenommene emotionale Feinfühligkeit der Lehrkraft an. Methodisch könnte dies darauf hinweisen, zukünftig eine differenzierte Betrachtung dieser beiden Dimensionen vorzunehmen und spezifischere Messinstrumente hierfür weiterzuentwickeln.

The Relationship Between Math Anxiety and Math Achievement: New Perspectives From an Individual Participant Data Meta-Analysis

Lena Keller¹, Franzis Preckel², Thomas Götz³, Oliver Lüdtke^{4,5}, Martin Brunner¹

¹University of Potsdam, Germany; ²University of Trier, Germany; ³University of Vienna, Austria; ⁴Leibniz Institute for Science and Mathematics Education, Germany; ⁵Centre for International Student Assessment, Germany

Theoretical background: Math anxiety and math achievement are reciprocally related (e.g., Carey et al., 2016), a relationship that is likely to affect individuals' agency and educational and career trajectories (e.g., Daker et al., 2021). Several meta-analyses have synthesized the research on this relationship by integrating the correlations reported in primary studies (i.e., aggregated data, AD; e.g., Barroso et al., 2021; Caviola et al., 2022). However, the generalizability of these findings is limited because (a) most primary studies were based on small, nonrepresentative samples of college students or came from Western countries and (b) they largely neglected the results and individual participant data (IPD) from international large-scale assessments (ILSAs). Moreover, previous AD meta-analyses could not provide conclusive evidence of moderating effects (e.g., gender) due to several methodological issues, including low statistical power and ecological bias. These moderating effects are especially important for understanding the boundaries and applicability of theories in the field. This IPD meta-analysis advances our knowledge about the relationship between math anxiety and math achievement by integrating representative, high-quality IPD from ILSAs and examining important moderating effects, which have not yet been examined in previous AD meta-analyses.

Research questions:

- (1) What is the magnitude of the relationship between math anxiety and math achievement, considering representative, high-quality IPD from ILSAs?
- (2) Is this relationship moderated by grade/age group, math subdomains, components of math anxiety, student gender, and low and high math achievement?

Methods: We systematically searched three major repositories of ILSA data and identified 11 relevant ILSAs ($N_{\text{total}} = > 2,000,000$; 96 countries). In all ILSAs, math anxiety was measured by self-report questionnaires and math achievement by standardized achievement tests. We integrated the IPD using the two-stage approach for an IPD meta-analysis of ILSAs (Brunner et al., 2023). In Stage 1, we estimated effect sizes (i.e., Fisher z-transformed correlations, regression coefficients, and group differences in regression coefficients), which were then meta-analytically integrated in Stage 2 using random-effects models. Dependencies between effect sizes were accounted for by applying multilevel meta-analytic models with robust variance estimation. Moderator analyses were conducted with mixed-effects meta-regression models. Additionally, missing data were imputed using multiply nested imputations.

Results and discussion: We found a robust negative relationship between math anxiety and math achievement ($r = -.28$) that was moderated by several factors. Math anxiety was significantly and negatively associated with math achievement across all grade levels/age groups and in all math achievement subdomains (e.g., algebra, calculus). Furthermore, all math anxiety components (e.g., worry, emotionality) were significantly and negatively associated with math achievement. Moreover, math anxiety and math achievement were consistently negatively related for both genders. However, the relationship was somewhat stronger and more negative for male than for female students in grade 4, but slightly more negative for female than for male students in grades 8, 9, and 10. Gender differences almost disappeared for students in advanced math classes in grade 12. The relationship between math anxiety and math achievement was more negative for students with average math skills compared to those with low skills. The weakest negative relationship occurred in students below the 5th percentile. However, this relationship was more pronounced for students scoring between the 75th and 95th percentiles, but not for top performers above the 95th percentile. Overall, meta-analyzing IPD provided new insights into moderating effects of individuals' gender and the level of math achievement. This evidence is reliable, generalizable, and nuanced, making it particularly valuable for research, practice, and policy development. Theoretical implications will be discussed.

Lernverlaufsdiagnostik effektiv einsetzen: Effekte von Unterstützungsmaßnahmen bei Datenauswertung und Unterrichtsadaption

Amelie Fuchs, Anika Radkowsch, Daniel Sommerhoff

IPN Kiel, Deutschland

Hintergrund und Zielsetzung

Die zunehmende Heterogenität der Schülerschaft und der Mangel an qualifizierten Lehrkräften stellen das Bildungssystem vor große Herausforderungen, auch bei der individuellen Förderung aller Schüler*innen. Um einen evidenzbasierten und bedarfsgerechten Unterricht (Espin et al., 2017) zu gestalten, ist eine kontinuierliche Erfassung des Lernstandes und individuellen Lernfortschritts aller Schüler*innen notwendig (Jung et al., 2018). Ein vielversprechender Ansatz hierfür ist Lernverlaufsdagnostik (LVD), bei der individuelle Lernfortschritte der Schüler*innen durch wiederholte, kurze Tests desselben Inhalts gemessen und den Lehrkräften (und ggf. den Schüler*innen) rückgemeldet werden (Deno, 1985; Fuchs, 2004). Diese formative Beurteilung (Black & Wiliam, 2009) ermöglicht eine effiziente, individuelle Förderung, da positive wie negative Leistungsentwicklungen zeitnah sichtbar werden und als Orientierung für Lehrkräfte und Schüler*innen dienen können.

Zentral für die effektive Nutzung von LVD durch Lehrkräfte ist (i) die sinnvolle Interpretation der erhobenen Daten und (ii) die darauf basierende, adäquate Anpassung des Unterrichts. Beide Aspekte stellen jedoch häufig eine Herausforderung für Lehrkräfte dar (Zeuch et al., 2017). In der vorgestellten Studie wird daher untersucht, inwiefern Unterstützungsmaßnahmen (i) zur Interpretation der Leistungsdaten der Schüler*innen und (ii) zur Unterrichtsadaption die Effektivität von LVD erhöhen können. Dabei soll geklärt werden, welchen Effekt LVD mit Unterstützungsmaßnahmen im Vergleich zu LVD ohne Unterstützungsmaßnahmen längsschnittlich auf die Schülerleistung hat.

Methode

Zur Beantwortung der Forschungsfrage wird derzeit eine präregistrierte quasi-experimentelle Längsschnittstudie in einem 2x2-Design mit Prä- und Posttest über einen Zeitraum von ca. sechs Monaten durchgeführt. Teilnehmende Lehrkräfte wurden zu Beginn randomisiert einer von vier Gruppen zugeordnet: Unterstützung bei Dateninterpretation (IG), Unterstützung bei Unterrichtsadaption (AG), Unterstützung in beiden Bereichen (IAG) oder einer Kontrollgruppe, die nur LVD ohne Unterstützung erhält (KG). In allen Gruppen setzen die Lehrkräfte LVD im Mathematikunterricht ein. Die Studie umfasst 47 zweite Klassen mit ca. $N = 996$ Schüler*innen und fokussiert inhaltlich auf Addition und Subtraktion im Hunderterraum.

Vorläufige Ergebnisse

Auf der Grundlage der bisher gesammelten Daten zeigt sich deskriptiv zunächst in allen Gruppen ein Leistungszuwachs vom Prä- zum Posttest, gleichzeitig aber eine große Heterogenität zwischen den Klassen und innerhalb der Gruppen sowohl in den Ausgangswerten als auch in den Lernzuwächsen. Die Mittelwerte der verwendeten Linking-Items liegen im Prätest bei 8,54 (IG), 8,63 (AG), 8,60 (IAG) und 7,74 (KG) sowie im Posttest bei 9,62 (IG), 9,69 (AG), 9,59 (IAG) und 9,59 (KG).

Diskussion und Ausblick

Die vorläufigen Ergebnisse zeigen, dass die Schüler*innen in allen Gruppen vom Prä- zum Posttest Leistungszuwächse verzeichnen. Dieser Befund deutet auf eine allgemeine Wirksamkeit der LVD hin, welche allerdings nicht im Zentrum der Studie stand und mangels Kontrollgruppe ohne LVD nicht weiter betrachtet werden soll. Vielmehr deuten die vorläufigen deskriptiven Ergebnisse an, dass die spezifische Wirkung der experimentellen Unterstützungsmaßnahmen gering ist und sich nicht direkt in den Gruppenmittelwerten widerspiegelt. Dies könnte aber auch der hohen Heterogenität zwischen den Klassen und innerhalb der Gruppen geschuldet sein.

Aufgrund dieser Heterogenität umfassen die nächsten Analyseschritte vor der inferenzstatistischen Auswertung neben einer IRT-Skalierung der Gesamtitens aus Prä- und Posttest und der Imputation fehlender Werte eine Analyse der Vergleichbarkeit der Gruppen. Die vollständigen Ergebnisse der experimentellen Variation auf Basis von Multilevel-Analysen werden auf dem Kongress präsentiert.

Ein weiterer Faktor, der die Wirksamkeit der experimentellen Variation beeinflussen könnte, ist die tatsächliche Nutzung der LVD durch die Lehrkräfte, da die Integration der LVD mit anderen Maßnahmen in den Schulalltag herausfordernd sein kann (Hebbecke & Souvignier, 2018). Darüber hinaus könnte die bereits vor der Intervention vorhandene Kompetenz der Lehrkräfte, die LVD-Leistungsdaten zu interpretieren, den Effekt der Unterstützungsmaßnahmen moderieren, da Lehrkräfte mit höheren Kompetenzen möglicherweise weniger von der Unterstützung profitieren. Diese Aspekte werden bei der Analyse der vollständigen Daten durch weitere Kontrollvariablen wie den Umfang der Nutzung der LVD durch die Lehrkräfte und die Kompetenz der Lehrkräfte berücksichtigt.

PS-18: Führung und Professionalisierung im Schulwesen: Einblicke in Fortbildung und Kooperation

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 157

Chair der Sitzung: Stephan Gerhard Huber

Qualifizierungsangebote für schulische Führungskräfte in Österreich - Eine Analyse des Einflusses von Prozessmerkmalen auf Ergebnismerkmale

Larissa Lusnig³, Stephan Gerhard Huber^{1,2}

¹Johannes Kepler Universität Linz, Österreich; ²Pädagogische Hochschule Schwyz, Schweiz; ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland

Für die Qualität von Schule sowie für eine kontinuierliche Weiterentwicklung sind professionelle Schulleitungen unabdingbar (Sammons et al. 1995; Hallinger & Huber 2012; Huber, 2010; Huber & Mujis 2010). Für das Amt der Schulleitung sind demnach adäquate Qualifizierungsangebote und Maßnahmen der Weiterbildung nötig. Auf diese Weise können schulische Führungskräfte ihre beruflichen Kompetenzen weiterentwickeln und auch ihre berufsbezogenen Werte und Haltungen reflektieren (Huber, 2003, 2005, 2012).

Die unterschiedliche Ausgestaltung des didaktischen Rahmens derartiger Qualifizierungs- und Weiterbildungsangebote wird in der einschlägigen Fachliteratur diskutiert. Es stehen nicht mehr nur klassische Kurse oder Seminare im Fokus. Zu multiplen Lernanlässen (Huber, 2013; Huber & Schneider, 2013) gehören kognitiv-theoretische Lernformen (Vorträge und Referate), aber auch kooperative und kommunikativ-prozessorientierte Verfahrensweisen (z.B. Projektarbeit) sowie reflexive Methoden (z.B. Self-Assessment und Feedback sowie Beratungsansätze wie Supervision und Coaching).

Ziel der vorliegenden Studie ist es, wichtige Prozessmerkmale für die Qualitätssicherung und -weiterentwicklung von Qualifizierungsangeboten für Schulleitungen zu identifizieren. Die Zusammenhänge verschiedener Prozess- und Ergebnisqualitäten wurden auf der Grundlage des Angebots-Nutzungs-Wirkungs-Modells für die Aus- und Weiterbildung analysiert (vgl. Fend 1987, 1998; Helmke, 2003; Huber et al., 2008; Huber, 2011). Dieses integriert die Evaluationsstufen nach Kirkpatrick (1994).

Für die vorliegende Studie, die in einem Mixed-Methods-Design durchgeführt wurde, wurden 782 Teilnehmer zu den Prozess- und Ergebnisqualitäten eines vorbereitenden nationalen Qualifizierungsangebots für Schulleitungen befragt, was einer Rücklaufquote von 51,79 entspricht. Die Befragungsteilnehmenden kamen von 11 nationalen Institutionen, welche die nationale Qualifizierung anbieten und waren in unterschiedlichen Schularten beschäftigt. Die quantitative Auswertung wurde im Rahmen von Korrelations- und linearen Regressionsanalysen durchgeführt. Für die Auswertung der offenen Fragen wurden neben den quantitativen Methoden der Datenerhebung und -analyse auch qualitative Methoden eingesetzt.

Die Ergebnisse der quantitativen Auswertung der vorliegenden Studie zeigen, dass der Angebotsnutzen umso größer eingeschätzt wird, je besser die Befragungsteilnehmenden ihre Gesamtzufriedenheit, ihre Kompetenzentwicklung, die praktische Umsetzung und die Zufriedenheit mit der Organisation und den Dozierenden einschätzen. Reflexion, Prozessgestaltung und die Bewertung der Gruppengrößen stehen hingegen nicht im Zusammenhang mit der Bewertung des Angebotsnutzens.

Die Kompetenzentwicklung wird umso größer eingeschätzt, je besser die persönliche Entwicklung, die praktische Umsetzung, der Angebotsnutzen und die Relevanz der Lerninhalte eingeschätzt werden. Die Zufriedenheit mit der Organisation und den Dozierenden und die Nützlichkeit der Lernanlässe stehen hingegen nicht im Zusammenhang mit der Bewertung der Kompetenzentwicklung.

Die Motivationssteigerung wird umso größer eingeschätzt, je besser die persönliche Entwicklung, der Angebotsnutzen und die Nützlichkeit der Lehrveranstaltungen eingeschätzt wird. Die Praktische Umsetzung, die Gesamtzufriedenheit, Reflexion und die Nützlichkeit der Lernanlässe stehen hingegen nicht im Zusammenhang mit der Bewertung der Motivationssteigerung.

In Anlehnung an die Entwürfe des BIO-Strategiemodells (Huber, 2020) wird dargestellt, welche Prozessmerkmale des Qualifizierungsangebots aus Sicht der Teilnehmer erhalten, erneuert und optimiert werden sollten (qualitative Analyse). Dabei fällt auf, dass die Umfrageteilnehmer besonders von der Vernetzung und dem Austausch mit anderen Teilnehmern profitierten.

Kooperationsmuster und ihre Auswirkungen auf Selbstwirksamkeit, Innovationsbereitschaft und Berufszufriedenheit von Fortbildenden

Ella Dorothea Hansen¹, Rebekka Stahnke², Malte Lehmann³, Bettina Rösken-Winter³

¹Technische Universität Dortmund; ²IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik;

³Universität Münster

Über Lehrkräftefortbildungen kann der Transfer innovativer Unterrichtskonzepte in die Schulen gelingen (Borko, 2004). Werden Fortbildende qualifiziert, die wiederum selbstständig Fortbildungen ausbringen, können in einem Kaskadenmodell zahlreiche Lehrkräfte erreicht werden (Rösken-Winter et al., 2021; Thurm et al., 2024). Kooperieren die an der Fortbildung teilnehmenden Lehrkräfte zudem mit ihrem Kollegium, können die Fortbildungsinhalte weiter multipliziert und die Umsetzung innovativer Ansätze gefördert werden (Sun et al., 2013; Vigerske, 2017). Kooperation unter Lehrkräften hat weitere positive Effekte, sowohl für unerfahrene als auch erfahrene Lehrkräfte (Richter et al., 2011): Gute Kooperationsstrukturen verbessern die fachliche Zusammenarbeit an Schulen (Gräsel et al., 2006) und steigern Selbstwirksamkeit und Berufszufriedenheit (Clement & Vandenberghe, 2000; Duyar et al., 2013), was wiederum mit höherer Motivation zum Transfer sowie einer besseren Transferleistung einhergeht (Grossman & Sales, 2011; Skaalvik & Skaalvik, 2014). Zudem fördert Kooperation die Innovationsbereitschaft (Bergmann & Rollett, 2008). Auf Ebene der Lehrkräfte gibt es somit umfangreiche Evidenz für die positiven Effekte von Kooperation auf die Lehrkräfte selbst sowie den Transfer. Bisher ist jedoch wenig über die Kooperationsmuster von Fortbildenden bekannt, die zwar oft selbst Lehrkräfte sind, jedoch eine übergeordnete Rolle im Kaskadenmodell einnehmen.

Hier setzt der Beitrag an und widmet sich den Fragen, welche Muster von Kooperation bei Fortbildenden vorliegen und wie diese Kooperationsmuster mit der Erfahrung als Lehrkraft und als Fortbildende, der Selbstwirksamkeit, Innovationsbereitschaft und Berufszufriedenheit zusammenhängen.

Das Projekt *QuaMath* zielt darauf ab, die mathematische Bildung in Deutschland langfristig zu verbessern, indem Fortbildende für die Ausbringung von Lehrkräftefortbildungen im Fach Mathematik qualifiziert werden. Vor Beginn dieser Qualifizierung wurden 190 Fortbildende mit mindestens einem Jahr Fortbildungserfahrung befragt. Ein standardisierter Fragebogen erfasste die

Erfahrung als Lehrkraft und Fortbildende, die Einstellung zur Kooperation mit anderen Lehrkräften und Fortbildenden sowie zur Kooperation unter Lehrkräften in gegebenen Fortbildungen (jeweils 4 Items; adaptiert nach Meudt et al., 2020). Zusätzlich wurden die Selbstwirksamkeit (18 Items; adaptiert nach Schwarzer & Jerusalem, 1999), Innovationsbereitschaft (4 Items; adaptiert nach Meudt et al., 2020) und Berufszufriedenheit (4 Items; adaptiert nach Ditton & Merz, 2013) erfasst. Die Reliabilitätsanalyse zeigte überwiegend gute interne Konsistenzen ($\alpha = .54$ bis $\alpha = .89$). Mit einer Latenten Profilanalyse (LPA) wurden die Kooperationsmuster untersucht. Mithilfe einer Multivariaten Varianzanalyse (MANOVA) wurde der Zusammenhang dieser Muster mit den weiteren Variable überprüft.

Die Analyse identifizierte drei Kooperationsmuster (AIC=759.80, BIC=805.26, Entropie=0.82; $p < .001$): Selektive-Kollaborateure (N=26; 13.7%), Moderate-Kooperationsbefürworter (N=38; 20%) und Kooperationsenthusiasten (N=126; 66.3%). Selektive-Kollaborateure kooperieren hauptsächlich mit anderen Fortbildenden und weniger mit anderen Lehrkräften. Moderate-Kooperationsbefürworter zeigen moderate Kooperation in allen Bereichen und betrachten Kooperation unter Lehrkräften in gegebenen Fortbildungen als weniger wichtig. Kooperationsenthusiasten weisen hohe Werte in allen drei Kooperationsformen auf.

Zwischen den Kooperationsmustern zeigen sich keine signifikanten Unterschiede bezüglich der Erfahrung als Lehrkraft und als Fortbildende. Auch hinsichtlich der Berufszufriedenheit ergaben sich keine signifikanten Unterschiede. Die Selbstwirksamkeit ($F(2, 187)=7.637$, $p < .001$, partial $\eta^2 = .076$; MDiff=-0.31, 95%-CI[-0.49, -0.12], $p < .001$) und Innovationsbereitschaft ($F(2, 187)=125.8$, $p < .001$, partial $\eta^2 = .069$; MDiff=-0.23, 95%-CI[-0.34, -0.08], $p = .015$) war bei den Kooperationsenthusiasten jedoch signifikant höher als bei den Moderaten-Kooperationsbefürwortern.

Der Beitrag konnte drei Kooperationsmuster identifizieren und somit neue Einblick in die Kooperation von Fortbildenden als zentrale Figuren im Kaskadenmodell des Transfers geben. Weiterhin zeigte sich, dass ein Kooperationsmuster mit geringerer Einbindung anderer Fortbildender und Lehrkräfte mit geringerer Selbstwirksamkeit sowie Innovationsbereitschaft zusammenhängt, denn die Kooperationsenthusiasten zeigen sich selbstwirksamer und offener für innovative Fortbildungskonzepte als die Moderaten-Kooperationsbefürworter. Insgesamt legen die Ergebnisse nahe, dass Kooperation ähnlich wie bei Lehrkräften eine wichtige Rolle für den Transfer spielen könnte. Dies deutet darauf hin, dass die Einbindung in kooperative Netzwerke und Arbeit in Tandems bei der Qualifizierung von Fortbildenden positive Auswirkungen auf diese Merkmale haben könnte. Die identifizierten Kooperationsmuster zeigen jedoch auch, dass Fortbildende möglicherweise unterschiedliche Unterstützungsbedarfe hinsichtlich ihrer Kooperationsstrukturen haben. Vertiefende Einblicke werden zukünftige Forschungsarbeiten in *QuaMath* liefern.

Welche Aufgaben übernehmen Fortbildungsbeauftragte an Schulen? Eine empirische Untersuchung aus Nordrhein-Westfalen

Julia Hartung, Eric Richter, Dirk Richter
Universität Potsdam, Deutschland

Theoretischer Hintergrund

Fortbildungsveranstaltungen sind zentral für die berufliche Weiterentwicklung von Lehrkräften (Darling-Hammond et al., 2017; Desimone, 2009). Schulleitungen spielen hierbei eine wichtige Rolle, da sie im Rahmen der Personalentwicklung Fortbildungen zur Förderung der Lehrkräfte gestalten. Sie haben die Möglichkeit, die Verantwortung für die Planung von Fortbildungsmaßnahmen an Fortbildungsbeauftragte zu delegieren (Thiel et al., 2022). Die Art der übertragenen Aufgaben kann zwischen Bundesländern aber auch Schulen erheblich variieren, was zur unterschiedlichen Ausgestaltung und Erfüllung der Aufgaben durch Fortbildungsbeauftragte führt. Aus diesem Grund geht diese Arbeit der Frage nach, welche spezifischen Aufgaben Lehrkräfte in der Rolle als Fortbildungsbeauftragte übernehmen und welche individuellen Merkmale die Übernahme dieser Aufgaben bedingen.

Fragestellung

Die vorliegende Studie untersucht die folgende Forschungsfragen:

1. Welche Aufgaben übernehmen Fortbildungsbeauftragte an Schulen in Nordrhein-Westfalen?
2. Welcher Zusammenhang besteht zwischen den objektiven und subjektiven Arbeitsvoraussetzungen der Fortbildungsbeauftragten und dem Umfang der übernommenen Aufgaben?

Methode

Die Daten wurden im Rahmen des Projektes E-Format (Universität Potsdam, Laufzeit 09/2023 – 02/2025), welches die Implementierung des Fortbildungsmanagement-Tools *FORMAT* begleitet, erhoben. Insgesamt nahmen 125 Fortbildungsbeauftragte (76,8 % weiblich) an der Online-Befragung teil. Die Mehrheit (16,8 %) übt die Tätigkeit seit 12 Monaten (Range wenige Wochen bis 192 Monate) aus.

Erfragt wurden soziodemografische Angaben, wie *Geschlecht*, *Dauer der Tätigkeit* und *Qualifikation* als Fortbildungsbeauftragte:r sowie der Erhalt von *Entlastungsstunden* für die Ausübung der Tätigkeit. Darüber hinaus wurde der Erhalt von Ressourcen, bspw. *Besoldung* und *Weisungsbefugnis* erfasst. Neben diesen objektiven Arbeitsvoraussetzungen wurden sogenannte subjektive Arbeitsvoraussetzungen erhoben, um zu beschreiben, wie die Fortbildungsbeauftragten ihre Arbeit interpretieren. Dazu gehört die Wahrnehmung von *fehlenden Ressourcen* (z. B. fehlende Zeit, Informationen; $\alpha = .61$), die Klarheit über die eigene Rolle (*Aufgabenklarheit*; $\alpha = .88$), und den Umfang der Zusammenarbeit mit der Schulleitung (*Kooperation Schulleitung*, $\alpha = .83$). Alle subjektiven Arbeitsvoraussetzungen wurden über eine 4-stufige Likert-Skala erfasst. Abschließend wurden die verschiedenen Aufgaben erfasst, die von Fortbildungsbeauftragten ausgeübt werden können. Dazu wurde auf Basis von Literatur eine Liste von 18 Aufgaben abgeleitet (z. B. Erfassung des Fortbildungsbedarfs, Dokumentation der Fortbildungsplanung, Organisation schulinterner Fortbildungen oder Koordination der Fortbildungsaktivitäten der Kolleg:innen), aus der die tatsächlich ausgeübten Aufgaben ausgewählt werden konnten.

Für die Beantwortung der ersten Forschungsfrage wurde deskriptiv ausgewertet, welche Aufgaben von den Fortbildungsbeauftragten übernommen werden. Für die Analyse der zweiten Forschungsfrage wurden drei multiple lineare Regressionsmodelle spezifiziert und geschätzt (Tabachnick & Fidell, 2007).

Ergebnisse und ihre Bedeutung

Die 18 Aufgaben werden in vier Kategorien gegliedert: 1. Fortbildungsplanung/-koordination; 2. Verwaltung/Ressourcenmanagement; 3. Durchführung/Evaluation; 4. Kommunikation/Unterstützung. Von den 125 befragten Fortbildungsverantwortlichen übernehmen 94,4 Prozent mindestens eine Aufgabe aus der Kategorie 1, 74,4 Prozent aus der

Kategorie 2, 88,8 Prozent aus der Kategorie 3 und 82,4 Prozent aus der Kategorie 4. Am häufigsten ($n = 106$) erfolgen Absprachen mit der Schulleitung zur Fortbildungsplanung und die Organisation schulinterner Fortbildungsveranstaltungen. Am seltensten ($n = 8$) wird ein individueller Fortbildungsplan für Kolleg:innen erstellt.

Mittels drei multipler linearer Regressionsmodelle wurde der simultane Einfluss mehrerer unabhängiger Variablen auf die abhängige Variable *Summe an übernommenen Aufgaben* quantifiziert und prognostiziert. Im ersten Modell wurden die objektiven Voraussetzungen (Besoldung, Weisungsbefugnis, Entlastungsstunden, Qualifikation) sowie Geschlecht und Tätigkeitsdauer als unabhängige Variablen aufgenommen. Es zeigen sich signifikante positive Zusammenhänge zw. Tätigkeitsdauer sowie Qualifikation und der Summe an übernommenen Aufgaben. Im zweiten Modell wurden die subjektiven Arbeitsvoraussetzungen (fehlende Ressourcen, Aufgabenklarheit, Kooperation Schulleitung) sowie Geschlecht und Tätigkeitsdauer als unabhängige Variablen aufgenommen. Es zeigen sich signifikante positive Zusammenhänge zw. Aufgabenklarheit sowie Kooperation Schulleitung und der Summe an übernommenen Aufgaben. In einem dritten Modell wurden die Prädiktoren kombiniert. Es zeigen sich signifikante Zusammenhänge zw. Qualifizierung, Kooperation Schulleitung, fehlende Ressourcen sowie Aufgabenklarheit und der Summe an übernommenen Aufgaben. Die Ergebnisse bieten Hinweise für die Gestaltung von Unterstützungs- und Qualifizierungsmaßnahmen, um Fortbildungsbeauftragte in ihrer Rolle zu stärken.

Berufszufriedenheit und eingeschätzter Fortbildungsbedarf von Lehrkräften. Ergebnisse einer Bedarfserhebung

Pia Schäfer, Jonas Burkhart, Michael Becker, Leo van Waveren
RPTU Kaiserslautern-Landau, Deutschland

Als dritte Phase der Lehrkräftebildung erfüllen Fort- und Weiterbildungen, mehrere Funktionen, darunter, die Weiterentwicklung „professioneller Handlungskompetenzen“ (Lipowsky 2023, S. 16) und die Umsetzung aktueller Themen wie Digitalisierung in Schulen (vgl. Kleinespel & Lütgert 2008; Kultusministerkonferenz 2017). Diese Angebote zielen auf eine Sicherung der Unterrichtsqualität hin (vgl. SWK, 2023), die u.a. aufgrund des technologischen Fortschritts sowie den stetigen Wissenszuwachs erforderlich sind (vgl. Bonnes et al. 2022; Kultusministerkonferenz, 2021; Runge et al. 2024).

Fort- und Weiterbildungen können sich positiv auf die Berufszufriedenheit von Lehrkräften auswirken (vgl. OECD 2019). Wobei Dicke et al. (2020) zeigen konnten, dass die Berufszufriedenheit von Lehrkräften eine bi-faktorielle Struktur mit den Faktoren Zufriedenheit mit dem Beruf sowie Zufriedenheit mit der beruflichen Situation aufweist. Bei einer detaillierteren Betrachtung der Berufszufriedenheit ergibt sich über die Zeit ein nicht-linearer Zusammenhang, der in den ersten 15 Jahren auf eine abnehmende und anschließend eine zunehmende Berufszufriedenheit hinweist (vgl. Crossman & Harris, 2006; OECD, 2014).

Festzuhalten ist, dass die Teilnahme an Fort- und Weiterbildungen nicht automatisch zu einem Transferprozess führt (vgl. Darling-Hammond et al., 2017), da diese „lediglich ein Angebot [darstellen]. Ob und wie effizient dieses Angebot genutzt wird, hängt von einer Vielzahl dazwischenliegender Faktoren ab“ (Helmke, 2012, S. 71f.). Daher wird empfohlen Fort- und Weiterbildungen entlang der Bedarfe und Interessen der Teilnehmenden zu entwickeln (vgl. Arthur et al. 1998; Fütterer et al. 2024; Lacerenza et al. 2017; Lipowsky & Rzejak 2015).

Im Rahmen des Projektes MINT-ProNeD (Standort Kaiserslautern) wird der Frage nachgegangen, welche Fortbildungsbedarfe Lehrkräfte zum Einsatz digitaler Technologien im Unterricht haben. Die Ergebnisse der Bedarfsanalyse fließen in die Konzeption der im Projekt MINT-ProNeD geplanten Angebote für MINT-Lehrkräfte ein. Die dem Beitrag zugrunde liegende querschnittliche Bedarfsanalyse adressiert zwei Forschungsfragen:

FF1: Wie stellt sich die berufliche Zufriedenheit vor dem Hintergrund des Fortbildungsbedarfs zu digitalen Technologien von Lehrkräften dar?

FF2: In welchen Bereichen können Fortbildungsbedarfe zu digitalen Technologien identifiziert werden, die im MINT-Unterricht relevant sind?

Das Erhebungsinstrument umfasst u.a. Skalen zu Demografie, Berufszufriedenheit sowie der Einschätzung des eigenen Fortbildungsbedarfs zu digitalen Technologien (z.B. VR/AR, 3D-Druck). Die Items zur Berufszufriedenheit sowie zu digitalen Technologien, wurden über ein fünfstufiges Likert-Antwortformat erfasst. Im Rahmen der vorläufigen deskriptiven Analysen und Korrelationsanalysen, wurde die Software R eingesetzt.

An der Ende 2023 durchgeführten Erhebung haben 101 Personen (50 Männer, 48 Frauen, 3 ohne Angaben) teilgenommen, die im Mittel 47 Jahre [$SD = 9,1$ Jahre] alt (Dienstalter: 18 Jahre, $SD = 8,5$ Jahre) waren. Die bi-faktorielle Struktur der Berufszufriedenheit (Dicke et al. 2019) konnte repliziert werden ($r = .7$, $p < .001$). Vorläufige Analysen zeigen einen Zusammenhang zwischen dem Alter ($r = .4$, $p < .001$) sowie den Dienstjahren ($r = .26$, $p = .01$) und der Zufriedenheit mit dem Beruf, nicht jedoch mit der Zufriedenheit mit der beruflichen Situation. Für das Geschlecht konnte kein signifikanter Zusammenhang mit beiden Faktoren der Berufszufriedenheit nachgewiesen werden.

FF1: Während die Lehrkräfte insbesondere für Erklärvideos, AR- und VR-Anwendungen sowie digitale Lösungsbeispiele einen Fortbildungsbedarf sehen, zeigt sich kein signifikanter Zusammenhang zwischen den geäußerten Bedarfen und der Berufszufriedenheit.

FF2: MINT-Lehrkräfte sehen einen signifikant größeren Fortbildungsbedarf ($.22 < r > .33$, $p < .05$) für neuere Technologien (u.a. AR/VR, 3D-Druck) wohingegen nicht-MINT-Lehrkräfte einen größeren Bedarf ($-.35 < r > -.24$, $p < .05$) für Tablets, Smartboards und Lernvideos sehen. Unter Berücksichtigung der Dienstjahre, zeigt sich ebenfalls eine kleine bis mittlere, negative Korrelation ($p < .05$) mit dem eingeschätzten Fortbildungsbedarf für neuere Technologien (u.a. AR/VR, 3D-Druck).

Die Befunde werden für die Entwicklung von Angeboten für Lehrkräfte (u.a. Fortbildungen) herangezogen, um diese auf die aktuellen Bedarfe der Lehrkräfte zuzuschneiden.

PS-19: Technologie und selbstreguliertes Lernen: Perspektiven aus der Bildungsforschung

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 154
Chair der Sitzung: Yoana Omarchevska

From Passive to Metacognitive: The Role of Mode of Engagement for Lasting Learning with Technology

Yoana Omarchevska¹, Brendan Schuetze¹, Juliane Richter^{2,4}, Sascha Schanze³, Nina Ulrich⁵, Katharina Scheiter^{1,2}

¹University of Potsdam; ²Leibniz Institut für Wissensmedien; ³Leibniz University Hannover; ⁴Hochschule der Medien Stuttgart;

⁵Humboldtschule Hannover

Theoretical Background

Digital learning environments provide a wealth of fine-grained logfile data measuring students' activities, from which learning processes can be inferred. We adopted the ICAP framework to classify logfile data of students' engagement with an interactive digital textbook (eChemBook). The ICAP framework (Chi & Wylie, 2014) hypothesizes that interactive activities are more effective than constructive activities, active activities, and passive activities in that order ($I > C > A > P$). Due to their flexible and self-paced nature, digital learning environments enable high amounts of *learner control* (Scheiter & Gerjets, 2007). In the eChemBook, students can rewind videos, re-engage with simulations, and write summaries to understand the learning material better. Students are prompted to make metacognitive judgments for each learning unit and can review learning material if needed. The non-linear structure of the eChemBook allows for studying the material in a self-regulated manner. The present study investigated how students' engagement with a digital textbook predicts their immediate and delayed knowledge.

RQ1. To what extent does students' engagement predict their immediate and delayed knowledge?

RQ2. Which modes of cognitive engagement are most effective for learning?

RQ3. To what extent does metacognitive engagement predict immediate and delayed knowledge above-and-beyond prior knowledge and the modes of cognitive engagement?

Methods

Data from 248 secondary school students were analyzed ($M_{Age} = 14.23$, $SD = 0.67$, 49% male). Prior knowledge was assessed with 15 items ($\alpha = .43$). Immediate and delayed knowledge was measured with 15 items ($\alpha = .65/.71$), 29 multiple-choice items ($\alpha = .87/.89$), and 8 drawing items ($\alpha = .86/.84$). Students' engagement with the eChemBook was recorded as logfiles and classified as either passive (e.g., playing videos), active (e.g., working with simulations), constructive (e.g., writing summaries), and metacognitive (e.g., confidence in explaining the learning content of a unit). Data were collected during chemistry lessons, including a pretest, two 90-minute learning sessions, an immediate post-test (1-7 days), and a delayed post-test (6-8 weeks).

Results

We constructed structural equation models to test four mediators—active, passive, constructive, and metacognitive engagement—between the pre-test and immediate and delayed knowledge. The pre-test and metacognitive engagement were the only significant predictors of immediate knowledge ($R^2 = .43$, CFI = 0.94, TLI = 0.92, RMSEA = 0.06). Diagnostic checks revealed that the active and passive factors, though theoretically well-grounded, contributed very little to the prediction of students' knowledge. Acknowledging the multilevel structure of the data, we compared two models, one with varying regression coefficients and another that constrained them to be equal. The model with varying regression coefficient did not fit better, $\chi^2(5, 248) = 1.80$, $p = 0.88$; thus, we report the single-group model results. A condensed model (CFI = 0.94, TLI = 0.92, RMSEA = 0.06), including only constructive and metacognitive engagement, indicated that both constructive ($\beta = 0.21$, $p < .001$) and metacognitive engagement ($\beta = 0.37$, $p < .001$) predict immediate knowledge ($R^2 = .44$) over-and-above the pre-test ($\beta = 0.28$, $p < .001$). Despite removing two predictors, the condensed model explained the same proportion of variance as the full model. A condensed model (CFI = 0.96, TLI = 0.93, RMSEA = 0.07) showed that metacognitive engagement ($\beta = 0.37$, $p < .001$) was the only predictor of delayed knowledge ($R^2 = .34$) over-and-above the pre-test ($\beta = 0.25$, $p < .001$). Our results corroborate findings that constructive activities are more effective for learning than passive activities. Moreover, metacognitive engagement was the strongest predictor for achievement, highlighting the importance of self-regulation for learning. Our findings help to understand the efficacy of different technology-enhanced learning activities in relation to lasting learning.

Effekt eines Lehrkräfte-Trainings zur Förderung des selbstregulierten Lernens mit digitalen Medien im Unterricht

Sabrina Reith, Maria Bannert

Technical University of Munich, Deutschland

Theoretischer Hintergrund

Die Bedeutung der Fähigkeit, das eigene Lernen zu steuern, ist in den letzten Jahren, insbesondere im Hinblick auf die voranschreitende Digitalisierung immer deutlicher geworden. Lernende benötigen Selbstregulation, um ihren Lernprozess in Lernumgebungen mit digitaler Mediennutzung zu optimieren (Zimmerman, 2008). Zudem kann die Nutzung digitaler Medien das selbstregulierte Lernen (SRL) bei den Lernenden unterstützen (Winne & Nesbit, 2009).

Damit SchülerInnen zu selbstregulierten Lernenden werden können, müssen Lehrkräfte in der Lage sein, die Entwicklung von SRL bei diesen zu unterstützen (Kramarski & Heaysman, 2021). Neben der Fähigkeit das selbstgesteuerte Lernen zu fördern, ist auch die Kompetenz der Lehrkräfte, digitale Medien in den Unterricht zu integrieren, angesichts der zunehmenden Digitalisierung des Bildungswesens wie auch der Bedeutung von digitalen Medien für das SRL, zunehmend wichtig. Um beide Fähigkeiten der Lehrkräfte zu verbessern, können Trainings eine effektive Methode sein (Lachner et al., 2021; Moos & Ringdal, 2012).

In Anbetracht der zahlreichen Forschungsarbeiten, die die Auswirkungen von Trainings für Lehrkräfte zur Integration digitaler Medien sowie zur Förderung des selbstregulierten Lernens untersucht haben, konnten die AutorInnen keine Forschungsarbeiten finden, die sich bisher mit einer Kombination aus beiden Komponenten befasst haben.

Fragestellung

Ziel unserer Studie ist es, herauszufinden, inwieweit ein Training zur Förderung von SRL mit digitalen Medien im Unterricht zu einem erhöhten Wissenserwerb bei Lehrkräften hinsichtlich der Förderung von SRL (mit Integration digitaler Medien) führen kann. Dabei untersuchen wir die Auswirkungen des SRL-Trainings *MIT Input zur Integration digitaler Medien* im Vergleich zu *OHNE*.

Methode

Hierfür wurde ein experimentelles Pretest-Posttest-Design durchgeführt. Beide Gruppen erhielten ein Training zur Förderung von SRL im Unterricht. Die EG-Teilnehmenden erhielten zusätzlichen Trainingsinhalt über die Umsetzung der Förderung von SRL mit digitalen Medien, während sich die Trainingsinhalte für die KG-Teilnehmenden auf die allgemeine Umsetzung von SRL ohne explizite Nutzung digitaler Medien konzentrierten. Final nahmen 66 Lehrkräfte aus verschiedenen Sekundarschultypen an der Studie teil (48 weiblich). Die Teilnehmenden wurden zufällig der Experimental- ($N = 33$) oder Kontrollgruppe ($N = 33$) zugewiesen.

Das Training wurde als kleiner privater Online-Kurs (SPOC) auf der Plattform *edX edge* durchgeführt. Sowohl der Inhalt als auch die Struktur des Trainings basierten auf dem zyklischen Phasenmodell von Zimmerman (2008) und dem Konzept der Lernstrategien von Wild und Schiefele (1994). Neben der Erfassung des Wissens über Strategieinstruktion (Dignath-van Ewijk & van der Werf, 2012) wurden die Teilnehmenden in einer Aufgabe zum Wissenstransfer gebeten, zu beschreiben wie sie SRL (mit digitalen Medien) bei ihren SchülerInnen in einer ihrer Unterrichtsstunden fördern würden. Zur Kodierung der erstellten Unterrichtspläne wurde das Verfahren wie auch das entwickelte Kodierschema an ATES von Dignath-van Ewijk et al. (2013) angelehnt. Dieses erfasste Variablen zu leistungsstarken Lernumgebungen, Strategieinstruktion und digitalem Medieneinsatz.

Ergebnisse und ihre Bedeutung

Während für das Wissen über Strategieinstruktion, für Variablen zur leistungsstarken Lernumgebung, sowie für die Häufigkeit und die Qualität des digitalen Medieneinsatzes keine allgemeinen Trainingseffekte gefunden werden konnten, hat das Training bei den Teilnehmenden für implizite und explizite Strategieinstruktion zu signifikanten Lernzuwächsen geführt. Dabei zeigten sich für das Wissen über Strategieinstruktion, wie erwartet, keine Gruppenunterschiede. Entgegen unserer Erwartung, gab es keine Gruppenunterschiede für Variablen zur leistungsstarken Lernumgebung, zur Strategieinstruktion, sowie zur Qualität des Einsatzes digitaler Medien; jedoch zur Häufigkeit des digitalen Medieneinsatzes. Bei der Analyse der Beziehung zwischen den Variablen zur Förderung des SRL und der Integration digitaler Medien fanden wir zudem einige signifikante positive Korrelationen; u.a. zwischen der impliziten Strategieinstruktion und der Häufigkeit und Qualität des digitalen Medieneinsatzes. Die Ergebnisse und daraus ableitbare Implikationen sollen in diesem Beitrag diskutiert werden.

Die Studie leistet damit einen wichtigen ersten Schritt zur theoretischen Forschung und zu praktischen Implikationen für die Förderung des selbstregulierten Lernens im Unterricht mit Integration digitaler Medien in der Lehrkräfteaus- und -weiterbildung.

Effekte von erweiterten interaktiven Elementen und metakognitiven Prompts in Lernvideos auf Lernprozesse und Wissenserwerb

Vincent Dusanek, Ingo Kollar

Psychologie m.b.B.d. Pädagogischen Psychologie, Universität Augsburg, Deutschland

Theoretischer Hintergrund

Digitale Videos bieten gemäß des multimedialen Lernens beträchtliches Potenzial für den Wissenserwerb (Mayer, 2014; Noetel et al., 2021; Rolfe & Gray, 2011). Allerdings bergen die dort nur kurzzeitig präsentierten Informationen die Gefahr einer kognitiven Überlastung des Arbeitsgedächtnisses (Costley et al., 2021; Sweller et al., 1998), da gegebenenfalls zu viele Informationen simultan verarbeitet werden müssen (Merkel & Schwan, 2016; Scheiter, 2014). Ein weiteres Problem besteht darin, dass Zuschauer die präsentierten Inhalte meist nur passiv rezipieren (Mayer & Chandler, 2001; Spanjers et al., 2012), was gemäß des ICAP-Modells nach Chi und Wylie (2014) für den Wissenserwerb weniger effektiv ist als aktivere Lernaktivitäten (Schwan & Riempp, 2004).

Vor diesem Hintergrund erscheint es vielversprechend, Lernvideos mit *basalen interaktiven Funktionen* zur Wiedergabekontrolle (Start/ Pause, Geschwindigkeitskontrolle etc.) anzureichern. Einerseits wirken solche Funktionen einer Überlastung des Arbeitsgedächtnisses entgegen (Merkel et al., 2022; Moreno & Mayer, 2007), andererseits fördern sie die aktive Informationsverarbeitung (Chi & Wylie, 2014; Findeisen et al., 2019). Die zusätzliche Implementierung von *erweiterten interaktiven Funktionen* mithilfe von H5P-Elementen, wie Lernzielkontrollaufgaben, stößt unter Umständen konstruktive Lernprozesse und eine tiefergehende Auseinandersetzung an (Chi, 2009; Jacob & Centofanti, 2023). Derzeit existieren nur wenige Studien mit größeren Stichproben zur Anreicherung von Lernvideos mit erweiterten interaktiven Funktionen mittels H5P (Mutawa et al., 2023; Ploetzner, 2022). In diesem Zusammenhang konnten keine eindeutigen Erkenntnisse gewonnen werden, was jedoch auf die geringe Nutzung der erweiterten interaktiven Funktionen zurückgeführt wurde (Jacob & Centofanti, 2023).

Metakognitive Prompts stellen eine Möglichkeit dar, um Lernende darauf hinzuweisen, interaktive Elemente wie Lernzielkontrollaufgaben wahrzunehmen. Die indirekte Förderung metakognitiver Lernstrategien (Mandl & Friedrich, 1992) in Lernumgebungen kann anhand von heuristischen Hilfen nicht nur das Engagement (Bergdahl, 2022) und den Lernerfolg erhöhen (Bannert, 2003; Bannert & Reimann, 2012), sondern bietet Lernenden auch die Chance, ihr Nutzungsverhalten in Bezug auf Lernzielkontrollaufgaben zu reflektieren und anzupassen (Bannert & Mengelkamp, 2013). Folglich sollte gezieltes Prompting die Nutzungsrate von interaktiven Funktionen und den Wissenserwerb steigern.

Fragestellung

FF: Welche Effekte haben (a) die Integration von erweiterten interaktiven Funktionen, (b) metakognitive Prompts sowie (c) deren Kombination in Lernvideos auf den Wissenserwerb und das Nutzungsverhalten von Studierenden?

Methode

Zur Klärung der Forschungsfrage nahmen deshalb $N=149$ Lehramtsstudierende ($M_{\text{Alter}}=20.89$, $SD_{\text{Alter}}=2.80$, $M_{\text{Fachsemester}}=3.50$, $SD_{\text{Fachsemester}}=2.50$) an einer Laboruntersuchung teil, sahen sich in einer virtuellen Lernumgebung sieben Videos zu multimedialem Lernen an und beantworteten schriftlich, wie sie einen Vortrag unter Nutzung lerntheoretischer Erkenntnisse ausgestalten würden. Um den Einfluss auf den Lernerfolg zu untersuchen, wurden die Faktoren "Interaktive Funktionen" (basale vs. erweiterte interaktive Funktionen) und "Metakognitive Lernprompts" (Fehlen vs. Vorhandensein metakognitiver Lernprompts) in einem 2x2-faktoriellen Design experimentell variiert und die Teilnehmenden randomisiert einer der vier Bedingungen zugewiesen. Die Erfassung des deklarativen Wissens erfolgte vor und nach der Maßnahme mittels Multiple-Choice-Fragen. Zwei unabhängige Gutachter werteten außerdem die schriftlichen Antworten zur Ermittlung des Anwendungswissens aus, $ICC_{\text{unjust}}(2,1)=.89$, $p<0.001$, 95%-KI [0.65,0.97]. Zusätzlich lieferten Bildschirmmitschnitte der Studierenden relevante Prozessdaten, um das Nutzungsverhalten zu untersuchen.

Ergebnisse und ihre Bedeutung

Weder die Art der interaktiven Funktionen, $F(1,144)=0.20$, $n.s.$, noch die der metakognitiven Lernprompts, $F(1,144)=0.43$, $n.s.$, hatten positive Effekte auf den deklarativen Wissenserwerb. Die Interaktion zwischen beiden Faktoren lieferte ebenfalls kein

signifikantes Ergebnis, $F(1,144)=1.13$, $n.s.$. Bislang gibt es keine Hinweise darauf, dass die Integration von erweiterten interaktiven Funktionen und metakognitiven Lernhilfen die Effektivität von Lernvideos steigert. Die Ergebnisse schließen sich somit an die heterogenen Befunde zu H5P-Elementen an (Jacob & Centofanti, 2023). Weiterhin soll überprüft werden, ob sich für das Anwendungswissen ähnliche Resultate ergeben. Die Auswertung der Prozessdaten ist noch nicht abgeschlossen. Sie ist jedoch entscheidend, um festzustellen, in welchem Zusammenhang unterschiedliche Nutzungsarten von interaktiven Elementen mit dem Erwerb von deklarativem Wissen sowie Anwendungswissen stehen. Der Vortrag thematisiert mögliche Erklärungen und stellt sämtliche Ergebnisse zur gemeinsamen Diskussion.

“Do-it-yourself” und adaptiv? Eine Implementationsstudie zur Wirksamkeit von selbst-entwickelten adaptiven Systemen im Schulkontext

Luisa Wellert¹, Sara Becker¹, Leonie Sibley¹, Vincent Aleven², Andreas Lachner¹

¹Eberhard Karls Universität Tübingen, Deutschland; ²Carnegie Mellon University, USA

Theoretischer Hintergrund

Adaptive Lernsysteme bieten das Potenzial, Heterogenität in Schulklassen zu adressieren und personalisiertes Lernen zu ermöglichen (Brod et al., 2023). Allerdings ist die Entwicklung solcher Lernsysteme technisch-komplex und kostenintensiv, weshalb sie noch nicht in der Breite angekommen sind. Zwei Strategien werden vor diesem Hintergrund diskutiert: Einerseits kann der Entwicklungsaufwand minimiert werden, indem Lehrpersonen Authoring Tools nutzen, um selbst angepasste adaptive Systeme zu erstellen und anderen als freie Bildungsressource bereitzustellen (Aleven, 2016). Andererseits könnten die Entwicklungsanforderungen adaptiver Systeme reduziert werden, indem anstatt komplexer Wissensmodellierung Selbsteinschätzungen über die wahrgenommene Schwierigkeit der aktuellen Aufgabe zur adaptiven Anpassung genutzt werden (Heitmann et al., 2021). Ob sich von Lehrkräften selbst erstellte adaptive Lernsysteme jedoch in (meta-)kognitiven oder motivationalen Lernleistungen von Schüler:innen widerspiegeln, ist ungeklärt, da die Validität von Assessments sowohl von Lehrpersonen als auch von Schüler:innen bezüglich der Komplexität von Aufgaben verzerrt sein können (Hellman & Nückles, 2013; Prinz et al., 2020).

Fragestellung

In der Studie untersuchten wir zwei präregistrierte Hypothesen:

H1: Von Lehrkräften entwickelte adaptive Lernsysteme sind hinsichtlich der Metakognition (monitoring accuracy), Kognition (Lernleistungen), kognitiven Belastung (mentale Anstrengung, subjektive Schwierigkeit) und der Motivation von Schüler:innen effektiver als nicht-adaptive Lernsysteme.

H2: Adaptive Lernsysteme, die auf den Selbsteinschätzungen von Schüler:innen beruhen, sind hinsichtlich der Metakognition (monitoring accuracy), Kognition (Lernleistungen), kognitiven Belastung (mentale Anstrengung, subjektive Schwierigkeit) und der Motivation von Schüler:innen effektiver als testbasierte adaptive Lernsysteme.

Methode

Teilnehmende und Design

Einem Co-Design-Ansatz folgend setzten sechs Lehrpersonen gemeinsam mit Studierenden adaptive Übungssysteme für ihren Unterricht um (Roschelle et al., 2006). Um die Generalisierbarkeit potenzieller Effekte zu untersuchen, wurde eine Many-Classes Studie durchgeführt (Fyfe et al., 2021), in welcher gezielt adaptive Lernsysteme in verschiedenen Klassenstufen und Schultypen (Grundschule, Sekundarschule, Berufsschule) sowie Fächern (Biologie, Deutsch, Medizin, Religion, Naturwissenschaften, Technik) implementiert wurden. Insgesamt nahmen sieben Klassen ($N = 134$ Schüler:innen) an der Studie teil.

Realisiert wurde ein Within-Participants-Design mit Prä- und Posttest sowie drei Bedingungen:

1. Adaptive Anpassung durch Selbsteinschätzung: Die Aufgabenschwierigkeit variierte je nach wahrgenommener kognitiver Belastung. Bei niedriger wahrgenommener Belastung (1-5) folgten schwierigere Aufgaben, bei hoher wahrgenommener Belastung (6-10) einfachere (siehe Heitmann et al., 2021).
2. Adaptive Anpassung durch Testleistungen: Die Aufgabenschwierigkeit wurde je nach vorheriger Leistung angepasst, indem korrekten Antworten schwierigere Aufgaben folgten und falschen Antworten einfachere.
3. Keine Anpassung: Die Reihenfolge der Aufgaben wurde zufällig, ohne Anpassung, ausgewählt.

Ablauf

Die Studie wurde von Kultusministerium und Ethikkommission genehmigt. Die Datenerhebung erfolgte im regulären Unterricht. Zu Beginn jeder ersten Unterrichtsstunde wurde ein Prätest durchgeführt. Die Lehrkräfte hielten anschließend ihren Unterricht wie gewohnt ab. Jeweils im Anschluss an die drei Unterrichtsstunden arbeiteten die Schüler:innen an den durch die Lehrkräfte entwickelten Lernsystemen. Die Posttests folgten jeweils zu Beginn der nächsten Unterrichtsstunde.

Ergebnisse und ihre Bedeutung

Die Analyse beider Hypothesen erfolgte durch Kontrastanalysen. Fehlende Werte wurden mittels multipler Imputationen und die hierarchische Datenstruktur (Schüler:innen genestet in Klassen) mittels clusterrobuster Regressionen berücksichtigt.

H1: Die Ergebnisse zeigten keinen signifikanten Unterschied zwischen den adaptiven Lernsystemen gegenüber den nicht-adaptiven Lernsystemen bezüglich Lernergebnissen, monitoring accuracy, Motivation, mentaler Anstrengung und subjektiver Schwierigkeit ($p \geq .252$).

H2: Ebenfalls zeigten adaptive Anpassungen durch Selbsteinschätzung keine höheren Ergebnisse als adaptive Anpassungen durch die Testleistungen bezüglich Lernergebnissen, monitoring accuracy, Motivation, mentaler Anstrengung und subjektiver Schwierigkeit ($p \geq .056$).

Explorativ ermittelten wir die Schwierigkeit der Übungsaufgaben, die von den Lehrkräften für die Lernsysteme erstellt wurden, und untersuchten den Zusammenhang zu den eingeschätzten Aufgabenschwierigkeiten durch die Lehrkräfte. Es zeigte sich zwar ein signifikanter aber lediglich kleiner Zusammenhang ($\beta = 0.14$, $p = .016$). Dies deutet darauf hin, dass die Lehrkräfte die reale Aufgabenschwierigkeit nur bedingt einschätzen können, was die Effektivität ihrer selbstgenerierten adaptiven Lernsysteme schmälern könnte. Weitere Forschung zur Erhöhung der Einschätzungskompetenz bei Lehrkräften ist nötig, um adaptive Lernsysteme praxiswirksam zu implementieren.

PS-20: Lesekompetenz und Motivation

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: SO 318

Chair der Sitzung: Julia Künstler-Sment

Das Wechselspiel von Lesekompetenz, Lesefreude und Leseinteresse von Heranwachsenden im Verlauf der Sekundarstufe

Nora Heyne, Maximilian Pfoth

Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Nach aktuellem Kenntnisstand wird die Förderung der intrinsischen Lesemotivation als zentraler Ausgangspunkt gesehen, um Heranwachsende in ihrer Lesekompetenz zu fördern. Hinsichtlich der intrinsischen Lesemotivation als allgemeines Konstrukt zeigten sich in querschnittlichen Studien jedoch unterschiedliche Zusammenhänge mit der Lesekompetenz, insbesondere bei Personen unterschiedlichen Alters. Längsschnittliche Studien, in welchen die wechselseitigen Beziehungen beider Konstrukte in verschiedenen Jahrgangsstufen untersucht wurden, brachten ebenso uneinheitliche Ergebnisse hervor: Einige Befunde legten nahe, dass Lesekompetenz zum Lesen motiviert (McElvany et al., 2008; Miyamoto et al., 2018), andere deuten darauf hin, dass Lesemotivation die Lesekompetenz fördert (Artelt et al., 2010; McElvany et al., 2008; Miyamoto et al., 2018) und weitere zeigten reziproke Beziehungen (Miyamoto et al., 2018; Schaffner et al., 2016; Schiefele et al., 2016). Differenziert man zwischen verschiedenen Komponenten der intrinsischen Lesemotivation – insbesondere Lesefreude versus Leseinteresse – so zeigten Querschnittsanalysen unterschiedliche Korrelationen mit der Lesekompetenz (Retelsdorf et al., 2011; Schaffner & Schiefele, 2007), die vor allem für die Lesefreude höher ausfielen (Möller & Bonerad, 2007). Zudem ergaben Längsschnittstudien (Heyne et al., 2023; Retelsdorf et al., 2011) inkonsistente Effekte von Lesefreude und Leseinteresse auf die Lesekompetenz. Die Effekte früherer Lesekompetenz auf beide Komponenten intrinsischer Lesemotivation sind hingegen noch weitgehend unklar.

Fragestellungen

Vor diesem Hintergrund steht die Untersuchung der längsschnittlichen Beziehungen zwischen der Lesekompetenz und der Lesefreude versus dem Leseinteresse bei Heranwachsenden zwischen den Klassenstufen 5, 7 und 9 im Fokus der Studie: Dabei wird einerseits untersucht, inwieweit die jeweiligen motivationalen Merkmale in dem betrachteten Zeitraum in einer wechselseitigen Beziehung zu der Lesekompetenz stehen. Zudem werden die Effekte der Lesekompetenz auf die Lesefreude versus das Leseinteresse hinsichtlich ihres Ausmaßes verglichen wie auch die Effekte beider motivationalen Merkmale zur Vorhersage der späteren Lesekompetenz. Schließlich wird untersucht, ob sich die Beziehungen zwischen der Lesekompetenz und der Lesefreude versus dem Leseinteresse zwischen den Klassenstufen 5 und 9 verändern.

Methoden

Die Studie basiert auf Daten des Nationalen Bildungspanels (NEPS) in Deutschland, die bei Heranwachsenden ($N = 2.854$) in den Klassenstufen 5, 7, und 9, anhand von standardisierten Lesekompetenztests und Fragebögen, erhoben wurden. Zur Umsetzung der Analysen kamen zwei Crosslagged-Panel-Modelle zum Einsatz, in denen interindividuelle Unterschiede (CLPM-L2; einschließlich langfristiger Effekte über den gesamten Untersuchungszeitraum) versus intraindividuelle Veränderungen (Random-Intercept Modelle; RI-CLPM) geschätzt wurden (Hübner et al., 2023; Lüdtke & Robitzsch, 2022).

Ergebnisse und Diskussion

Die Ergebnisse zeigen, dass Lesefreude in einer wechselseitigen Beziehung mit Lesekompetenz steht, was hingegen nicht für das Leseinteresse zutrifft. Während die frühere Lesekompetenz das spätere Leseinteresse vorhersagt, werden keine gegenläufigen Effekte beobachtet. Der Vergleich der Effekte der unterschiedlichen motivationalen Merkmale zur Vorhersage der späteren Lesekompetenz verdeutlicht, dass Lesefreude positive und signifikant höhere Effekte auf die spätere Lesekompetenz zeigt, als das Leseinteresse. Darüber hinaus sind die Effekte früherer Lesekompetenz auf das spätere Leseinteresse geringer als die Auswirkung auf die Lesefreude, was besonders in jüngeren Klassenstufen deutlich wird. Schließlich sind über den betrachteten Zeitraum keine signifikanten Veränderungen in den Beziehungen zwischen den jeweiligen motivationalen Merkmalen und der Lesekompetenz festzustellen.

Fazit

Trotz einiger methodischer Einschränkungen im Design der Studie, legen ihre Ergebnisse nahe, in zukünftigen Untersuchungen wie auch Fördermaßnahmen zwischen der Lesefreude und dem Leseinteresse, als verschiedene Facetten intrinsischer Lesemotivation, zu differenzieren. Zudem heben die Befunde die Bedeutung der Förderung von Lesefreude, als zentralen Ansatzpunkt für die Förderung von Lesekompetenz, hervor, wohingegen Leseinteresse eine geringere Rolle zukommt. Nach den Ergebnissen ist es daher vielversprechend, die Lesefreude in den Mittelpunkt zu stellen, um dem eingangs genannten Förderbedarf von Heranwachsenden zu begegnen und diese hinsichtlich ihrer Lesekompetenz zu fördern.

Lesekompetenzen im Vergleich zwischen Halb- und Ganztagschulen - Welche Rolle spielt die kognitiv-instrumentelle Unterstützung von Mitschüler:innen?

Julia Künstler-Sment¹, Aleksander Kocaj^{1,2}, Georg Lorenz^{3,4}, Malte Jansen^{2,5}

¹Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Deutschland; ²Zentrum für internationale Vergleichsstudien (ZIB); ³Department Erziehungswissenschaft, Universität Potsdam, Potsdam; ⁴Institut für Soziologie, Universität Leipzig, Leipzig;

⁵Department of Psychology, HMU Health and Medical University Potsdam

Der Ausbau des Ganztagsbetriebs an deutschen Schulen ist mit ambitionierten Zielen verbunden, darunter die Erwartung, dass Ganztagschulen durch eine Erhöhung der Lernzeit zu höheren Lernzuwächsen der Schüler:innen beitragen (KMK, 2001). Trotz dieser Ambitionen zeigen repräsentative Schulleistungsstudien (Hertel et al., 2008; Holtappels et al., 2010; Willems et al., 2015) kaum Unterschiede in den schulischen Kompetenzen zwischen Halb- und Ganztagschulen, auch nach Kontrolle von Selektionseffekten (Strietholt et al., 2015) und unter Berücksichtigung von nachmittäglichen Bildungsangeboten (Steinmann & Strietholt, 2019).

Ein Mechanismus, der zu positiver Lernentwicklung an Ganztagschulen beitragen kann, sind die vielfältigen Möglichkeiten zum sozialen Austausch und gemeinsamen Lernen mit Peers (z. B. durch mehr gemeinsam verbrachte Zeit und mehr extracurriculare Angebote). Damit Schüler:innen an Ganztagschulen von diesen Möglichkeiten profitieren, ist es wichtig, positive soziale Beziehungen zu ihren Peers aufzubauen (Autoren, 2023). Oft suchen Jugendliche bei fachlichen Herausforderungen Hilfe bei ihren Mitschüler:innen (Puustinen & Rouet, 2009; Zimmerman & Pons, 1986). Diese kognitiv-instrumentelle Beziehungen unter

Peers fördern einen lernorientierten Wissensaustausch und kooperatives Lernen (Zander et al., 2017). Kooperatives Lernen hilft Peers, sich gegenseitig bei der Selbstregulierung unterstützen, indem sie erfolgreiche Strategien und Techniken zum Lernen miteinander teilen (Kyndt et al., 2013; Tenenbaum et al., 2020; Wentzel, 2009). Schüler:innen, die sich aktiv Hilfe suchen, erreichen unter Kontrolle der Vorleistung höhere schulische Leistungen (Webb & Mastergeorge, 2003). Leistungsschwache Schüler:innen nutzen seltener Hilfe (Karabenick & Berger, 2013), während Schüler:innen mit einem besseren Verständnis eher dazu neigen, anderen zu helfen als um Hilfe zu bitten (Davison et al., 2023). Beim Erklären von Inhalten profitieren auch leistungsstärkere Schüler:innen, in dem sie ihr Wissen festigen und ein tieferes Verständnis entwickeln (Lernen durch Erklären; vgl. Fiorella & Mayer, 2013, 2014, 2016). Insgesamt gibt es noch wenige aktuelle und repräsentative Studien, die kognitiv-instrumentelle Beziehungen und deren Zusammenhänge zu schulischen Leistungen in Halb- und Ganztagschulen betrachten.

1. Diese Studie vergleicht die Lesekompetenzen in Deutsch zwischen Halb- und Ganztagschulen, wobei Selektionseffekte und Unterschiede zwischen Schularten der Sekundarstufe statistisch kontrolliert werden.
2. Darüber hinaus wird untersucht, wie kognitiv-instrumentelle Beziehungen zwischen Peers bei Hilfesuchenden und Hilfegebenden mit den Lesekompetenzen in Deutsch zusammenhängen und ob diese Zusammenhänge in Halb- und Ganztagschulen unterschiedlich ausgeprägt sind.

Die Analysen basieren auf repräsentativen Daten des IQB-Bildungstrends 2022 (Stanat et al., 2023). Nach Ausschluss von Klassen mit weniger als 50 % Teilnahme an der Netzwerkbefragung sowie offener Ganztagschulen betrug die Analysestichprobe 7.202 Sekundarschüler:innen (Durchschnittsalter 15,5 Jahre; 50,0% Mädchen), davon 4.543 an Halbtagschulen und 2.659 an Ganztagschulen. Die Lesekompetenzen in Deutsch wurden anhand von Mehrebenenanalysen (Random-Intercept-Modelle) unter Berücksichtigung von Populationsgewichten separat für Gymnasien, Integrierte Gesamtschulen, Schulen mit mehreren Bildungsgängen und Realschulen zwischen Halb- und Ganztagschulen verglichen. Die Peerbeziehungen wurden über soziale Netzwerkindikatoren erfasst (Hilfesuchende: Outdegrees; Hilfegebende: Indegrees). Zur Kontrolle von Selektionseffekten wurden als Kovariaten soziodemografische und kognitive Merkmale der Schüler:innen, deren Teilnahme an bildungsunterstützenden extracurricularen Angeboten sowie die Klassengröße und das Bundesland einbezogen. Fehlende Werte wurden multipel imputiert.

In Einklang mit vorherigen Studien zeigen die vorliegenden Ergebnisse keine signifikanten Unterschiede in den Lesekompetenzen in Deutsch zwischen Halb- und Ganztagschulen unter Berücksichtigung von Selektionseffekten und Schulartunterschieden. Bezüglich der kognitiv-instrumentellen Beziehungen zwischen Peers zeigten sich differenzielle Zusammenhänge mit den Lesekompetenzen: Während die erhaltene Unterstützung nicht signifikant mit den Lesekompetenzen von Hilfesuchenden zusammenhing (mit Ausnahme eines negativen Zusammenhangs an Schulen mit mehreren Bildungsgängen in Ganztagsform), konnte ein signifikanter positiver Zusammenhang mit der gegebenen Unterstützung der Schüler:innen festgestellt werden. Dies legt nahe, dass Schüler:innen, die anderen bei schulischen Herausforderungen helfen, tendenziell über bessere Lesekompetenzen verfügen. Diese Zusammenhänge waren an Schulen mit mehreren Bildungsgängen in Ganztagschulen stärker ausgeprägt, während sie an Gymnasien nur in Halbtagschulen beobachtet wurden. Dagegen zeigen diejenigen, die Hilfe suchen, keine systematisch höheren oder niedrigeren Kompetenzen.

Einfach mehr lesen, um besser zu lesen? Longitudinale reziproke Effekte von Lesegewohnheiten und Leseverständnis

Simon Tiffin-Richards¹, Jan Retelsdorf², Sandra Preusler³, Jens Möller³

¹IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel, Germany; ²Universität Hamburg, Fakultät für Erziehungswissenschaft; ³Christian-Albrechts-Universität zu Kiel, Institut für Pädagogisch-Psychologische Lehr- und Lernforschung

Theoretischer Hintergrund

Texte flüssig lesen und verstehen zu können, ist eine Schlüsselkompetenz für den Bildungserfolg und bildet eine zentrale Grundlage für den lebenslangen Wissenserwerb durch den selbstgesteuerten Zugang zu Informationen und deren Integration mit dem eigenen Vorwissen (Richter & Tiffin-Richards, 2024). Da Lesen eine mühsam über viele Jahre hinweg erlernte Kulturtechnik ist, sollte wenig überraschen, dass Lesen als Freizeitbeschäftigung positiv mit Lesekompetenz zusammenhängt (z.B. Mol & Bus, 2011; Schiefele et al., 2012). Es wird davon ausgegangen, dass Leseerfahrung basale Worterkennungsprozesse trainiert (Share, 1995, *self-teaching hypothesis*) und damit kognitive Kapazitäten für das Leseverstehen freisetzt (Perfetti, 1985, *verbal efficiency hypothesis*). Umgekehrt wird angenommen, dass Leseschwächen sich negativ auf Leseaktivitäten auswirken (Becker et al., 2010). Darüber hinaus ist eine plausible Annahme, dass der wechselseitige Einfluss von Leseleistung und Leseaktivitäten einen sich verstärkenden Kreislauf bildet (Stanovich, 1986). Erstaunlicherweise gibt es jedoch relativ wenige Studien, die sich damit befassen, wie sich das Leseverhalten in Beziehung zur Lesekompetenz im Laufe der Zeit verändert. Das Befundmuster der vorhandenen Längsschnittstudien ist zudem teilweise inkonsistent: Während generell stärkere Effekte der Lesefähigkeit auf die Leseaktivität als umgekehrt gefunden werden (z.B. Leppänen et al., 2005; Torppa et al., 2020; van Bergen et al., 2020), finden andere Studien nur einen einseitigen Einfluss der Lesefähigkeit auf die Leseaktivität (z.B. Harlaar et al., 2011) oder unabhängige Entwicklungsverläufe (z.B. Aarnoutse & van Leeuwe, 1998).

Fragestellung

In diesem Beitrag werden zwei große Längsschnittdatensätze reanalysiert, um die wechselseitige Beziehung zwischen dem Leseverhalten (Dauer und Häufigkeit des Lesens in der Freizeit) und einem standardisierten Maß des Leseverständnisses in den Klassenstufen 4 bis 9 zu untersuchen. Wir testen die Hypothese eines bidirektionalen Zusammenhangs, wonach die Lesekompetenz die Dauer und Häufigkeit des Lesens erhöht und umgekehrt das Leseverhalten das Leseverständnis positiv beeinflusst.

Methode

Als Datengrundlage verwenden wir die Längsschnittstudie EUROPA (Möller et al., 2017), die drei Messzeitpunkte von Klasse 4 bis 6 ($N = 2053$ Schüler:innen) umfasst, und die Längsschnittstudie "Lesen in der Sekundarstufe" (LISA, Retelsdorf & Möller, 2013), die vier Messzeitpunkte von Klasse 5 bis 9 ($N = 1508$ Schüler:innen) umfasst. In beiden Studien wurde das Leseverhalten anhand von Selbsteinschätzungen der Lesedauer und Lesehäufigkeit in der Freizeit erhoben sowie das Leseverständnis mit standardisierten Testverfahren. Um autoregressive und gekreuzte Effekte des Leseverhaltens und -verständnisses zu testen, wurden Cross-Lagged-Panel-Modelle (CLPM, vgl. Lüdtke & Roitzsch, 2022) in Mplus gerechnet, wobei für das Geschlecht der Schüler:innen, ihren sozioökonomischen Hintergrund und ihre nonverbalen kognitiven Fähigkeiten kontrolliert wurde. Die genestete Struktur der Daten wurde mit der Klassenzugehörigkeit der Schüler:innen in den Mehrebenenmodellen berücksichtigt. Modellparameter wurden mit Full Information Maximum Likelihood (FIML) und robusten Standardfehlern geschätzt.

Ergebnisse und ihre Bedeutung

In der Stichprobe der Schüler*innen der 4. bis 6. Klassen ergab ein CLPM signifikante autoregressive Effekte, die sowohl auf eine Vorhersage der Leseleistung durch die vorherige Leistung als auch auf eine Vorhersage des Leseverhaltens durch das vorherige Leseverhalten hindeuten. Signifikante reziproke Effekte deuten darauf hin, dass sich Lesekompetenz und Leseaktivität im Laufe der Zeit gegenseitig verstärken. Hier zeigten sich stärkere Effekte der Leseleistung auf das Leseverhalten. Ein ähnliches Befundmuster ergab ein CLPM mit der Stichprobe der Schüler*innen der 5 bis 9 Klasse, wobei die Effekte insgesamt schwächer ausfielen. Unsere Ergebnisse deuten darauf hin, dass gute Leser*innen in ihrer Freizeit mehr lesen als schwache Leser*innen. Umgekehrt gab es auch Evidenz dafür, dass Freizeitlesen einen positiven Einfluss auf die Lesekompetenz hatte, zumindest in den unteren Klassenstufen. Dieses Ergebnismuster weist auf eine reziproke Wechselwirkung zwischen Übung und Kompetenz in der Leseentwicklung hin (vgl. Leppänen et al., 2005; Troppa et al., 2020), vor allem während des frühen Leseerwerbs.

Der Einfluss professioneller Kompetenzen von angehenden Lehrkräften auf die Leseleistung und die Lesemotivation von SchülerInnen und die vermittelnde Rolle der Passung des Förderangebots

Henrike Mentel¹, Mareike Ehlert¹, Elmar Souvignier¹, Natalie Förster²

¹Universität Münster, Deutschland; ²Universität Wuppertal, Deutschland

Theoretischer Hintergrund

Professionelle Kompetenzen von Lehrkräften gelten als wichtige Grundlage für effektiven Unterricht (Kunter et al., 2013). Entsprechend ist die Vermittlung solcher Kompetenzen ein zentrales Ziel der Ausbildung von Lehrkräften. Tatsächlich finden sich in aller Regel positive Effekte, wenn Zusammenhänge zwischen professionellen Kompetenzen von Lehrkräften und SchülerInnen Outcomes untersucht werden (z.B. Yang & Kaiser, 2022). Nur in wenigen Studien wird dabei jedoch der vollständige Wirkprozess unter Berücksichtigung der Gestaltung des Unterrichts in den Blick genommen (Blömeke et al., 2022; Yang & Kaiser, 2022). Während für zentrale Merkmale der Unterrichtsqualität wie Klassenführung und kognitive Aktivierung (Blömeke et al., 2022; Fauth et al., 2019) signifikante Mediationseffekte nachgewiesen wurden, liegen bislang kaum entsprechende Studien für den Aspekt der Passung zwischen Unterrichtsangebot und Lernniveau der SchülerInnen vor (Parsons et al., 2018). Dieser Beitrag untersucht den Einfluss des Wissens, der Selbstwirksamkeit und der unterrichtlichen Überzeugungen von Lehramtsstudierenden im Praxissemester auf die Leseleistung und Motivation der SchülerInnen und die vermittelnde Rolle der Passung als eine relevante adaptive Unterrichtspraktik.

Fragestellungen

1. Welchen Einfluss haben die professionellen Kompetenzen der Lehramtsstudierenden auf die Entwicklung der Leseleistung und Lesemotivation der SchülerInnen?
2. Welchen Einfluss haben die professionellen Kompetenzen der Lehramtsstudierenden auf die realisierte Passung zwischen Lernausgangslage und Förderangebot?
3. Welchen Einfluss hat die realisierte Passung von Lernausgangslage und Förderangebot auf die Entwicklung der Leseleistung und Lesemotivation der SchülerInnen?
4. Wird der Einfluss der professionellen Kompetenzen der Lehramtsstudierenden auf die Entwicklung der Leseleistung und Lesemotivation der SchülerInnen durch die realisierte Passung zwischen Lernausgangslage und Förderangebot vermittelt?

Methode

Zur Beantwortung der Fragestellungen werden Daten aus dem Projekt „Lese-Sportler: mit PS in die Schule“ ausgewertet. In diesem Projekt wurden Lehramtsstudierende ($N = 36$) im zweiten oder dritten Mastersemester in zwei universitären Blockveranstaltungen zu differenzierter Leseförderung theoretisch ausgebildet. Sie lernten das differenzierte Lesematerial „der Lese-Sportler“ (Hebbecke et al., 2020) kennen, welches drei evidenzbasierte Methoden zur Förderung von Lesegenauigkeit, Lesegeschwindigkeit und Leseverständnis mit unterschiedlichen Schwierigkeitsstufen beinhaltet. Anschließend führten sie in ihrem Praxissemester ein Lesetraining mit Kleingruppen durch, bei dem sie für jedes Kinderpaar eine nach ihrer Einschätzung passende Fördermethode auswählten, und die Trainingseinheiten dokumentierten. Das selbsteingeschätzte Wissen über Leseförderung, die Selbstwirksamkeit (angelehnt an Schwarzer & Schmitz, 2002) und unterrichtsbezogene, konstruktivistische und transmissive Einstellungen (Behrmann & Souvignier, 2013) der Studierenden wurden vor dem Praxissemester erhoben. SchülerInnen ($N = 218$) wurden sowohl vor als auch nach der Förderung zu ihrer intrinsischen Lesemotivation (Goy et al., 2017) befragt und das Leseverständnis wurde mit dem ELFE II (Lenhard et al., 2018) getestet. Basierend auf den Lesekompetenzen der Kinder beurteilten drei Experten die optimale Lesefördermethode, um die Passung zu operationalisieren.

Ergebnisse und Diskussion

Die Datenerhebung ist abgeschlossen, die Daten sind jedoch noch nicht abschließend analysiert. Erste deskriptive Auswertungen zeigen ein hoch ausgeprägtes Wissen ($M=5.38$, $SD=0.39$) und Selbstwirksamkeit ($M=4.71$, $SD=0.57$) der Lehramtsstudierenden. Zudem zeigen sich korrelativ positive Zusammenhänge zwischen der Selbstwirksamkeit ($r=.12$, $p=.092$) und den konstruktivistischen Überzeugungen ($r=.13$, $p=.058$) der Lehramtsstudierenden mit den Leistungen der SchülerInnen am Ende des Schulhalbjahres. Die Fragestellungen nach der realisierten Passung zwischen Lernausgangslage und Förderangebot und Zusammenhänge mit professionellen Kompetenzen und SchülerInnen Outcomes werden mit polynominalen Regressionsmodellen in Kombination mit der Response Surface Analysis untersucht. Mit dieser Analysestrategie kann, im Vergleich zu Differenzscores, der tatsächliche Effekt der Passung auf die Outcomes identifiziert werden (Shanock et al., 2010). Bei allen Analysen wird die hierarchische Datenstruktur (SchülerInnen auf L1, Lehramtsstudierende auf L2) berücksichtigt.

Die Befunde werden einen differenzierten Einblick dazu geben, welche Rolle die Passung von Unterrichtsangebot und Lernniveau im Rahmen einer differenzierten Förderung der Lesekompetenz spielt. Diskutiert werden die Ergebnisse vor dem Hintergrund der Relevanz professioneller Kompetenzen von (angehenden) Lehrkräften und ihrem Zusammenspiel mit adaptiven Unterrichtspraktiken sowie der Leistung und Motivation von SchülerInnen.

S-14: Wissen und Stereotype zum Thema Hochbegabung von (angehenden) Lehrkräften

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: SN 163

Chair der Sitzung: Julia Schiefer

Chair der Sitzung: Friederike Zimmermann

Chair der Sitzung: Jörn Sparfeldt

Wissen und Stereotype zum Thema Hochbegabung von (angehenden) Lehrkräften

Chair(s): **Julia Schiefer** (Pädagogische Hochschule Ludwigsburg), **Friederike Zimmermann** (Institut für Pädagogisch-Psychologische Lehr- und Lernforschung (IPL), CAU Kiel)

Diskutant*in(nen): **Jörn Sparfeldt** (Universität des Saarlandes)

Wissen und Überzeugungen von Lehrkräften zum Thema Hochbegabung und individueller Förderung sind entscheidend für das Erkennen sowie die Förderung von hochbegabten Schülerinnen und Schülern im Schulunterricht. Dennoch wird im Lehramtsstudium kaum Wissen zum Thema Hochbegabung vermittelt und (angehende) Lehrkräfte verfügen häufig über wenig Wissen, negative Stereotype und eher ungünstige implizite Vorstellungen über hochbegabte Schülerinnen und Schüler (Matheis et al., 2019; Preckel & Vock, 2021). Überzeugungen von Lehrkräften sind dabei häufig der sogenannten Disharmonie-Hypothese zuzuordnen. Diese nimmt einen Zusammenhang von hoher Intelligenz und negativen psychosozialen Merkmalen an, also beispielsweise, dass intellektuell hochbegabte Kinder anfälliger für soziale und emotionale Probleme oder Verhaltensauffälligkeiten seien als durchschnittlich begabte Kinder (Preckel & Vock, 2021). Wissenschaftliche Untersuchungen widerlegen die Disharmonie-Hypothese, dennoch prägt diese nach wie vor Alltagsvorstellungen (implizite Theorien) und Überzeugungen über Hochbegabte (Baudson, 2016), auch im Kontext Schule. Da Lehrkräfte bei der Entwicklung Hochbegabter eine bedeutende Rolle spielen, sind ihre impliziten Theorien und vorherrschenden (häufig negativen) Stereotype zu hochbegabten Schülerinnen und Schülern in der Praxis hoch relevant (Baudson & Preckel, 2016; Preckel & Vock, 2021). Diese negativen Stereotype können einen adäquaten Umgang mit hochbegabten Schülerinnen und Schülern verhindern und deren Entwicklung beeinträchtigen. Daher ist sowohl ein Verständnis für die Entstehung von solchen Stereotypen als auch die Konzeption und Evaluation von Interventionen zur Reduzierung solcher Stereotype ein wichtiges Thema für die Bildungsforschung und Praxis. Dieses wird in dem Symposium aus verschiedenen Perspektiven beleuchtet.

Der erste Beitrag beschäftigt sich mit grundsätzlichen Konzeptionen von Lehrkräften zum Thema Hochbegabung und stellt zum anderen eine experimentelle Studie zu (stereotypen) Assoziationen von Lehrkräften hinsichtlich „hochbegabten“ vs. „besonders begabten“ Kindern vor. Der zweite Beitrag präsentiert die Ergebnisse einer experimentellen Studie zur Reduzierung des Stereotyps sozialer Inkompetenz und psychischer Auffälligkeiten bei Hochbegabung mittels verschiedener textbasierter Mini-Interventionen im Vergleich zu einer Kontrollbedingung bei angehenden Lehrkräften. Darauf aufbauend stellt der dritte Beitrag eine experimentelle Interventionsstudie vor, bei der ein Refutationstext als Video im Tiktak-Stil gegenüber neutralem Video-Material auf seine Wirksamkeit zur Reduzierung des „Disharmonie-Stereotyps“ geprüft wird. Der letzte Beitrag untersucht im Rahmen einer quasi-experimentellen Feldstudie die Frage, ob sich das Wissen, die stereotypisierten Vorstellungen, die Förderüberzeugungen sowie das selbst eingeschätzte Kompetenzerleben im Unterricht Hochbegabter von angehenden Lehrkräften durch den Besuch eines Hochschulseminars signifikant verbessern lassen. Die Ergebnisse werden im Rahmen des Symposiums hinsichtlich ihrer Relevanz für die Bildungsforschung und Bildungspraxis diskutiert.

Beiträge des Symposiums

Konzeptionen von und (stereotype) Überzeugungen über Hochbegabung bei (angehenden) Lehrkräften

Hanna Beißert¹, Nina Brück¹, Patrick Lösche¹, Jessika Golle²

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Laut dem Modell der professionellen Kompetenz von Lehrkräften (Kunter et al., 2011) sind Wissen und Überzeugungen zentrale Dimensionen professionellen Handelns. Zum Wissens zählen allgemeines pädagogisch-psychologisches Wissen, Fachwissen und fachdidaktisches Wissen. Mit Überzeugungen sind epistemologische Überzeugungen sowie subjektive Theorien über Lehren und Lernen gemeint, wozu auch stereotype Auffassungen zählen. Überträgt man das Modell auf den professionellen Umgang mit hochbegabten Schüler*innen, ist Wissen über Hochbegabung zentral (z.B. Kenntnisse von theoretischen Modellen, empirischen Befunde, didaktischen Besonderheiten). Im Hinblick auf den Aspekt der Überzeugungen ist von Bedeutung, welche subjektiven Theorien und mögliche stereotype Überzeugungen Lehrkräfte in Bezug auf Hochbegabung und gegenüber hochbegabten Schüler*innen haben.

Da das Thema Hochbegabung in der Lehramtsausbildung nur eine untergeordnete Rolle spielt, ist unklar, über welches Wissen und welche Alltagstheorien Lehrkräfte in diesem Bereich verfügen, welches Verständnis von (Hoch-)Begabung sie haben und welche stereotypen Überzeugungen möglicherweise vorherrschen. Diese impliziten Theorien der Lehrkräfte sind aber relevant, da sie ihr Verhalten und ihre Entscheidungen gegenüber hochbegabten Schüler*innen beeinflussen können (Matheis et al., 2017; Megay-Nespoli, 2001) und Einfluss darauf haben, ob Schüler*innen überhaupt als hochbegabt identifiziert werden oder nicht (Matheis et al., 2017, 2020).

Deshalb untersuchen wir, welches Verständnis von Hochbegabung Lehrkräfte haben (Studie 1) und welche (möglicherweise stereotypen) Überzeugungen über hochbegabte Kinder vorherrschen (Studie 2).

In Studie 1 wurden Grundschullehrkräfte (N=155, 154 weiblich, 1 männlich) aus Baden-Württemberg anhand einer offenen Frage gefragt, was sie persönlich unter Hochbegabung verstehen. Zur Auswertung der Antworten wurde anhand eines kombinierten deduktiv-induktiven Verfahrens ein Kodierleitfaden entwickelt. Dabei wurden zunächst mögliche Konzeptionen aus der Literatur abgeleitet und diese durch weitere aus dem Material heraus gebildete Kategorien ergänzt: „Hochbegabung als überdurchschnittliche Leistung“, „Hochbegabung als hohe Fähigkeit/Potenzial“, „Hochbegabung als hohe Kreativität“, „Hochbegabung als schnelle Auffassungsgabe“, „Hochbegabung als hohe Motivation/Leistungsbereitschaft“, „Hochbegabung als hohes Interesse“. Die Inter-Rater-Reliabilität zweier unabhängiger Coder war mit einem Cohens Kappa von .88 sehr gut. Zusätzlich wurde kodiert, wenn Hochbegabung als angeboren und/oder als domänenspezifisch beschrieben wurde.

Es zeigte sich, dass bei den Lehrkräften das Alltagsverständnis von Hochbegabung verschiedenen in der Literatur auffindbaren Konzeptionen ähnelt: Häufig waren Konzeptionen von Hochbegabung als hoher Leistung (27%) oder hoher Fähigkeit (43%) sowie ein Verständnis von Hochbegabung als schneller Auffassungsgabe (34%). Konzeptionen von Hochbegabung als hohem

Interesse (13%), hoch ausgeprägter Kreativität (6%) oder hoher Motivation/Leistungsbereitschaft (6%) waren weniger vertreten. Häufig wurde Hochbegabung als etwas Domänenspezifisches beschrieben (37%) und 5% der Befragten beschrieben Hochbegabung als etwas Angeborenes.

Es zeigte sich wie erwartet auch, dass unter den Lehrkräften kein einheitliches Verständnis von Hochbegabung vorliegt. Zudem sind einige der vorliegenden Begabungskonzeptionen nicht dienlich, um adäquate schulische Begabtenförderung zu betreiben. So sind Konzeption wie „Begabung als hohe Leistung“ oder „Hochbegabung als hohe Motivation/Leistungsbereitschaft“ auf sichtbare Aspekte beschränkt, wodurch begabte Underachiever*innen übersehen werden.

Studie 2 untersucht, welche (stereotypen) Überzeugungen (angehende) Lehrkräfte gegenüber hochbegabten Schüler*innen haben. Es gibt bereits Hinweise darauf, dass die Überzeugungen von Lehrkräften bezüglich Hochbegabten oft von negativen Stereotypen geprägt sind (Baudson, 2016; Preckel & Vock, 2020). Dies wird in Studie 2 genauer erforscht. Dabei wird anhand eines experimentellen Designs untersucht, ob negative Stereotype insbesondere durch das Label „Hochbegabung“ aktiviert werden. In einer Online-Studie werden Lehrkräfte und Lehramtsstudierende gebeten, in einer offenen Antwort zu beschreiben, was sie über hochbegabte (vs. besonders begabte) Kinder wissen. Diese experimentelle Variation wird between-subjects realisiert. Geplant ist eine Stichprobe von je 80 Lehrkräften und Lehramtsstudierenden (Datenerhebung läuft aktuell). Anhand der offenen Antworten soll kodiert werden, welche Zuschreibungen die Teilnehmenden vornehmen und ob sich die Häufigkeit von stereotypen Aussagen unterscheidet, je nachdem ob der Begriff „hochbegabt“ oder „besonders begabt“ dargeboten wurde. Zudem werden mögliche Unterschiede zwischen Lehramtsstudierenden und Lehrkräften untersucht.

Die Ergebnisse der beiden Studien sollen Ansatzpunkte liefern, um die Aus- und Fortbildung von (angehenden) Lehrkräften im Bereich Hochbegabung zu verbessern.

Textbasierte Mini-Interventionen zur Reduzierung von Stereotypen gegenüber hochbegabten Schüler*innen: Ein Vergleich der Wirksamkeit

Vanessa Clauss, Steffani Sass, Friederike Zimmermann
IPL Kiel

Theoretischer Hintergrund

Hochbegabte werden im Laufe ihres Lebens mit verschiedenen Stereotypen bezüglich ihrer sozial-emotionalen Kompetenzen konfrontiert (z.B. Baudson & Preckel, 2013). Zwei Hauptannahmen dominieren: die Harmonie-Hypothese besagt, dass intellektuelle Hochbegabung mit stärker ausgeprägten sozial-emotionalen Kompetenzen einhergeht, während die Disharmonie-Hypothese Hochbegabten Defizite in ihrer Sozialkompetenz und eine höhere Auftretenswahrscheinlichkeit für psychische Auffälligkeiten unterstellt (Preckel & Vock, 2012). Obwohl die Harmonie-Hypothese durch kumulierende Forschungsergebnisse gestützt wird (z.B. Bergold et al., 2020), neigen Lehramtsstudierende und Lehrkräfte dazu, Stereotypen entsprechend der Disharmonie-Hypothese zuzustimmen (z.B. Baudson & Preckel, 2016). Stereotype über Eigenschaften und Merkmale von Gruppen dienen der Vereinfachung von Denkprozessen, erweisen sich im Kontext von Hochbegabung jedoch als besonders problematisch. Die vorherrschenden Stereotype, die der Disharmonie-Hypothese entsprechen, basieren auf im Allgemeinen nichtzutreffenden Annahmen und können die Entwicklung hochbegabter Schüler*innen erheblich beeinflussen (Baudson & Preckel, 2013). Eine vorherige Studie (Clauss et al., in prep) zeigte, dass das Lesen eines Refutationstextes die wahrgenommene soziale Inkompetenz von hochbegabten Schüler*innen gegenüber einer Kontrollbedingung verringerte. Ein Refutationstext folgt dabei einer spezifischen Struktur: Zuerst wird eine weitverbreitete Fehlvorstellung genannt, dann folgt ein Widerlegungssignal und zuletzt die aktuell anerkannte wissenschaftliche Erklärung (Tippett, 2010). Es wurde vielfach gezeigt, dass Refutationstexte im naturwissenschaftlichen Bildungsbereich zu einer stärkeren konzeptuellen Veränderung von Fehlvorstellungen führen als expositorische Texte, die nur die wissenschaftliche Erklärung enthalten (z.B. Schroeder & Kucera, 2022). Dies kann durch die Koaktivierungs-Hypothese (Kendeou & van den Broeck, 2007) erklärt werden, die die gleichzeitige Aktivierung der Fehlvorstellung und der korrekten Annahme als notwendig ansehen, um kognitive Dissonanz zu erzeugen, die idealerweise durch konzeptuelle Veränderung aufgelöst werden soll. Auch Braasch et al. (2013) betonen die Notwendigkeit des Widerlegungssignals für konzeptuelle Veränderungen.

Fragestellung

Wie effektiv ist der Refutationstext im Vergleich zu einem expositorischen Text und einem Kontrolltext bei der Reduzierung von Stereotypen gegenüber hochbegabten Schüler*innen?

Methode

An dieser Studie nahmen 109 Lehramtsstudierende teil (72% weiblich; MAlter = 24.23 Jahre (SD = 4.09); M = 7 Fachsemester (SD = 3); 85% Gymnasiallehramt), die randomisiert einer der drei Bedingungen (Refutationstext, expositorischer Text oder Kontrolltext) zugewiesen wurden.

Instrumente

Stereotype

Erfasst durch eine adaptierte Skala zur Wahrnehmung der Gruppe hochbegabter im Vergleich zu durchschnittlich begabten Schüler*innen (nach Heyder et al., 2018) hinsichtlich a) wahrgenommener sozialer Inkompetenz (WSI) und b) wahrgenommenen psychischen Auffälligkeiten (WPA).

Refutationstext

Ein speziell für die Studie entwickelter Text basierend auf der drei-stufigen Struktur von Tippet (2010). Nach der fälschlichen Aussage der Disharmonie-Hypothese, dass hochbegabte Schüler*innen im Allgemeinen in ihrer sozial-emotionalen Entwicklung Probleme aufwiesen, folgte die Entkräftung, dass diese Hypothese nach aktuellem Forschungsstand nicht gestützt werden könne. Diese wurde mit Studienergebnissen belegt.

Expositorischer Text

Dieser Text enthielt nur die Studienergebnisse entsprechend des Refutationstextes. Um die Textlänge zu kontrollieren, wurde er durch Wiederholung der Aussagen auf die gleiche Länge gebracht.

Ergebnisse und ihre Bedeutung

Da die Datenauswertung zu diesem Zeitpunkt noch nicht vollständig abgeschlossen ist, werden hier vorläufige Ergebnisse berichtet.

Das Lesen des expositorischen Textes verringerte sowohl die WSI ($\beta = -.40$, $p < .001$) als auch die WPA ($\beta = -.27$, $p < .05$) im Vergleich zur Kontrollgruppe. Das Lesen des Refutationstextes zeigte ähnliche Wirkungen für WSI ($\beta = -.34$, $p < .05$) und WPA ($\beta = -.34$, $p < .01$). Es zeigte sich kein signifikanter Unterschied in den Effekten zwischen Refutationstext und expositorischem Text.

Die vorläufigen Ergebnisse deuten darauf hin, dass sowohl Refutationstext als auch expositorischer Text wirksame Mittel zur Reduktion von Stereotypen gegenüber Hochbegabten sind, wobei der spezifische Vorteil des Refutationstextes nicht gezeigt werden konnte. Weiterführend werden Moderationsanalysen mit verschiedenen personellen Variablen für die Effekte der textbasierten Mini-Interventionen berechnet.

Eine videobasierte Intervention zur Reduzierung des Stereotyps mangelnder sozialer Kompetenz gegenüber hochbegabten Personen

Steffani Sass, Sonia Ahmling Loureiro, Friederike Zimmermann
IPL Kiel

Theoretischer Hintergrund

Hochbegabte Personen werden in den Medien häufig als weniger sozial kompetent dargestellt (Baudson, 2016; Cross, 2005a). Wissenschaftliche Untersuchungen zeigen jedoch, dass diese Darstellungen empirisch nicht gestützt sind und hochbegabte Personen im Durchschnitt genauso sozial kompetent sind wie durchschnittlich begabte (Roznowski et al., 2000). Trotz dieser Erkenntnisse bleibt das negative Stereotyp über die mangelnde soziale Kompetenz Hochbegabter, das demnach einer Fehlvorstellung entspricht, in der Allgemeinbevölkerung bestehen (Baudson, 2016). Um Vorurteilen und Stereotypen im Allgemeinen entgegenzuwirken, wurden verschiedene Interventionen entwickelt. Bisherige Ansätze basieren oft auf der Kontakthypothese (Allport, 1954), die postuliert, dass direkter persönlicher Kontakt mit vorurteilsbehafteten Gruppen zur Stereotypenreduktion beiträgt. Angesichts der heutigen gesellschaftlichen Gegebenheiten ist dieser direkte Kontakt jedoch schwer herzustellen (Ramasubramanian, 2015). Daher wird zunehmend untersucht, ob stellvertretender oder vorgestellter Kontakt durch Medien ebenfalls Stereotype reduzieren kann (Bergold et al., 2020; Ramasubramanian, 2015). Eine Methode zur Korrektur von Fehlvorstellungen ist ein Refutationstext; das Knowledge Revision Components Framework spezifiziert, wie Fehlvorstellungen in Wissensnetzwerke eingebettet sind und durch Refutationstexte korrigiert werden können (Zengilowski et al., 2021). Refutationstexte identifizieren zunächst Fehlvorstellungen, ordnen sie als falsch ein und widerlegen sie schließlich durch wissenschaftliche Belege, was zu einer Umstrukturierung des Wissensnetzwerkes führen soll (Tippett, 2010). Traditionell werden solche textbasierten Interventionen in Form von Fallvignetten eingesetzt (Schroeder & Kucera, 2022).

Die vorliegende Studie kombiniert diese beiden Ansätze, indem sie eine Videointervention entwickelt, die sowohl vorgestellten Kontakt als auch das Prinzip des Refutationstextes nutzt und die wichtigsten Informationen über eine Fallvignette präsentiert (vgl. Amsalem et al., 2021). Ziel ist es, zu prüfen, ob diese Videointervention zu geringeren Stereotypen über Hochbegabung beiträgt.

Fragestellung

Wie effektiv ist eine Videointervention im Vergleich zu einer Kontrollgruppe bei der Reduzierung von Stereotypen gegenüber hochbegabten Personen?

Methode

An dieser Studie nahmen 253 Studierende teil, von denen 114 der Kontrollgruppe und 139 der Interventionsgruppe randomisiert zugeordnet wurden. Das Durchschnittsalter der Teilnehmenden (77% weiblich und 21% männlich) betrug 23.30 Jahre ($SD = 4.01$). Die Teilnehmenden schauten sich ein Video an und beantworteten anschließend einen Fragebogen.

Instrumente

Videointervention

Das Video ist im TikTok Stil produziert (eine hochbegabte Person spricht direkt in die Kamera) und ist nach dem Prinzip eines typischen Refutationstextes nach Tippett (2010) aufgebaut. Es beginnt mit der Darstellung des Stereotyps, dass hochbegabte Personen weniger sozial kompetent sind, und widerlegt dies dann anhand aktueller wissenschaftlicher Erkenntnisse. Studien werden im Video eingeblendet, um die Glaubhaftigkeit zu erhöhen. Die persönliche Ansprache im Video soll den vorgestellten Kontakt verstärken. Für die Kontrollgruppe wurde ein ähnlich langes Natur-/Tiervideo von YouTube ausgewählt, das keine Informationen über Hochbegabung enthält.

Stereotype

Zur Erfassung des Stereotyps mangelnder sozialer Kompetenzen von hochbegabten Personen wurde die Wärme/Kompetenz-Skala, im Original von Fiske et al. (2002), mit den Items von Amsalem et al. (2021) verwendet. Zusätzlich wurde eine adaptierte und reduzierte Skala (9 Items) nach Heyder et al. (2018) zur Erfassung des Stereotyps mangelnder sozialer Kompetenzen eingesetzt.

Ergebnisse

Da die Datenauswertung zu diesem Zeitpunkt noch nicht vollständig abgeschlossen ist, werden hier vorläufige Ergebnisse berichtet. Teilnehmende der Experimentalgruppe schätzten die Gruppe der Hochbegabten als sozial kompetenter ($t(250) = -3.53$, $p < .01$) und wärmer ($t(242) = 2.39$, $p < .01$) ein, als Teilnehmende der Kontrollgruppe. Bezüglich der Einschätzung der Kompetenz gab es keinen Unterschied zwischen beiden Gruppen ($t(247) = 1.15$, $p = .25$). Die vorläufigen Ergebnisse zeigen, dass die Videointervention effektiv dazu beiträgt, das Stereotyp der mangelnden sozialen Kompetenz hochbegabter Personen zu reduzieren.

Förderung von Wissen und Überzeugungen zum Thema Hochbegabung von angehenden Lehrkräften

Julia Schiefer
Pädagogische Hochschule Ludwigsburg

Theoretischer Hintergrund

Wissen und Überzeugungen von Lehrkräften zum Thema Hochbegabung und individueller Förderung sind entscheidend für das Erkennen und die Förderung von hochbegabten Schülerinnen und Schülern im Unterricht. Diese verfügen über weit überdurchschnittliche intellektuelle Fähigkeiten und benötigen in der Schule eine besondere Förderung und angemessene hohe kognitive Anforderungen (Preckel & Vock, 2021). Dies kann beispielsweise durch individuelle und differenzierende Aufgaben umgesetzt werden und ermöglicht den hochbegabten Schülerinnen und Schülern die Entfaltung ihre Potentiale sowie die

Aufrechterhaltung ihrer Motivation und Lernfreude. Dennoch wird im Lehramtsstudium kaum Wissen zum Thema Begabtenförderung vermittelt und Lehrkräfte verfügen häufig über eher negative Stereotype zum Thema Hochbegabung (Baudson, 2016; Preckel & Vock, 2021). Diese sind häufig der sogenannten Disharmoniehypothese zuzuordnen, die von einem Zusammenhang zwischen sozial-emotionalen Problemen und hoher Intelligenz ausgeht (Preckel & Vock, 2021). Zudem fehlen kontrollierte Studien zum Einfluss von Weiterbildungen oder Interventionen auf das Wissen, die Stereotype, die intendierten Unterrichts- und Lehrpraktiken sowie die Selbstwirksamkeitserwartungen von (angehenden) Lehrkräften.

Fragestellung

Die vorliegende Interventionsstudie widmet sich der Frage, ob sich das Wissen, die Förderüberzeugungen sowie das selbst eingeschätzte Kompetenzerleben im Unterrichten Hochbegabter von angehenden Lehrkräften durch den Besuch eines Hochschulseminars zum Thema Hochbegabung signifikant steigern sowie die stereotypen Vorstellungen signifikant reduzieren lassen.

Methode

An der quasi-experimentellen Studie mit Prä-Postdesign (T1, T2) nahmen im WS 2023/24 N = 58 Lehramtsstudierende (Alter: M = 23.48, SD = 2.74; 87.9% weiblich, 69% im Bachelor) teil. N = 31 (Interventionsgruppe, IG) besuchten ein blended-learning Seminar zum Thema Hochbegabung, die Kontrollgruppe (N = 27) ein asynchrones Onlineseminar zum Thema Sozialpsychologie. Das 14-wöchige Seminar (IG) vermittelte Wissen und Grundlagen zur Identifikation und Förderung besonders begabter und hochbegabter Schülerinnen und Schüler im Regelunterricht. Dabei wurden u. a. Konzepte und Modelle der (Hoch)Begabung und Intelligenz, Grundlagen der Diagnostik, Persönlichkeitsmerkmale und Lernvoraussetzungen Hochbegabter sowie Ansätze und konkrete Methoden zur individuellen Förderung und Beratung dieser Schülerinnen und Schüler adressiert. Aufbauend auf diesen Grundlagen entwickelten die Studierenden eine eigene Unterrichtseinheit mit integrierten Maßnahmen der Begabtenförderung in ihrem Fachbereich. Zudem wurden externe Referentinnen, die Maßnahmen zur Begabtenförderung an Schulen durchführen, in das Seminar eingeladen.

Instrumente

Es wurden zu T1 und T2 das Wissen (Steinmayr et al., 2022), die Stereotype (Spyré et al., 2023), die Bereitschaft zu Individualisierung und Differenzierung im Unterricht (Baumert et al., 2008) sowie das Kompetenzerleben im Zusammenhang mit Hochbegabung (selbst generierte Skala) erfasst. Die Reliabilitäten der Skalen lagen zwischen $.64 \leq \alpha \leq .88$.

Statistische Analyse

In einem ersten Schritt wurden als Maß der Veränderung der Zielvariablen Effektstärken für die IG (Cohen's d) berechnet. Im Beitrag werden zudem – unter Hinzunahme einer weiteren Kohorte aus dem Sommersemester 2024 – die Ergebnisse multipler Regressionsanalysen in Mplus (AVs = z-standardisierte T2 Werte; UVs = Gruppe (IG vs. CG), z-standardisierte T1 Werte, Kovariaten) vorgestellt.

Ergebnisse und ihre Bedeutung

Die vorläufigen Ergebnisse (der ersten Kohorte) deuten darauf hin, dass sich das Wissen zum Thema Hochbegabung ($d = 1.14$, $p < .001$), die Bereitschaft zu Individualisierung und Differenzierung im Unterricht ($d = 0.82$, $p = .005$) sowie das selbst eingeschätzte Kompetenzerleben im Unterrichten Hochbegabter ($d = 1.01$, $p = .001$) nach dem Besuch des Hochschulseminars signifikant gesteigert haben und die Teilnehmenden weniger Stereotype über Hochbegabte berichten ($d = 0.68$, $p = .016$). Kein signifikanter Effekt zeigte sich auf die Entwicklung der Förderüberzeugungen der Studierenden ($d = .028$, $p = .278$). Die Studie zeigt die Möglichkeit der Förderung von Wissen und Überzeugungen im Zusammenhang mit Hochbegabung und kann zum Verständnis gelingender Fort- und Weiterbildungen von (angehenden) Lehrkräften beitragen. Diese stellen eine wichtige Voraussetzung für eine differenzierende und kognitiv aktivierende Lehr- und Unterrichtspraxis dar.

S-15: Multidimensionale Ansätze zur Beschreibung von Interaktionsprozessen in der Schule

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 151

Chair der Sitzung: Madita Frühauf

Chair der Sitzung: Britta Wenzel

Chair der Sitzung: Natalie Fischer

Multidimensionale Ansätze zur Beschreibung von Interaktionsprozessen in der Schule

Chair(s): **Madita Frühauf** (Freie Universität Berlin), **Britta Wenzel** (Bergische Universität Wuppertal)

Diskutant*in(nen): **Natalie Fischer** (Universität Kassel)

Der schulische Lernraum ist geprägt von sozialen Interaktionen zwischen Schüler:innen, Lehrkräften und Schulleitungen. Dabei ist die Ausgestaltung dieser Interaktionen für den Lernerfolg der Schüler:innen (Wentzel, 2023) sowie der Berufszufriedenheit der Lehrkräfte (Kremer-Hayon & Wubbels, 1993) sehr bedeutsam. Ein Modell zur Beschreibung von Interaktionen ist das interpersonale Zirkumplex (Leary, 1957), das in verschiedenen Disziplinen wie der klinischen (Wright et al., 2023), Persönlichkeits- (Gurtman, 2009) sowie der pädagogischen (Wubbels & Brekelmans, 2005) Psychologie etabliert ist. Dabei wird Interaktionsverhalten auf zwei Dimensionen beschrieben, dem Ausmaß der affektiven Zuwendung (z. B. Wärme, Herzlichkeit, Communion, sozio-emotionale Unterstützung) und dem Ausmaß der Lenkung (z. B. Steuerung, Kontrolle, Agency, fachlich-inhaltliche Unterstützung). Auf Grundlage des interpersonalen Zirkumplex vereint dieses Symposium den Blick aus verschiedenen Disziplinen auf Interaktionsprozesse in der Schule und beleuchtet dabei den Zusammenhang mit lernwirksamen Aspekten. Beitrag 1 erweitert die Unterrichtsqualitätsforschung, indem konstruktive Unterstützung in die Subdimensionen sozio-emotionale und fachlich-inhaltliche Unterstützung unterteilt wird. Auf Basis eines neu vorgestellten Beobachtungsinstrumentes wurden Interaktionen zwischen Lehrkräften und Grundschüler:innen auf den beiden Unterstützungsdimensionen kodiert und u.a. der Zusammenhang mit der Lesekompetenz der Schüler:innen ermittelt. Es zeigte sich, dass Lehrkräfte sich überwiegend positiv sozio-emotional und fachlich-inhaltlich unterstützend gegenüber Schüler:innen verhielten, wobei sie Schüler:innen mit hoher Lesekompetenz am meisten unterstützten. Beitrag 2 und 3 erweitern den Blick und betrachten, inwiefern sich das Verhalten von Lehrkräften und Schüler:innen gegenseitig bedingt (Beitrag 2) und welchen Einfluss fachspezifische Geschlechtsstereotype auf das Verhalten der Lehrkraft in den Fächern Deutsch und Mathematik haben (Beitrag 3). Im Rahmen von Beobachtungsstudien wurde das dyadische Lehrkraftverhalten gegenüber Schüler:innen auf den Dimensionen Agency (Lenkung, Kontrolle) und Communion (Wärme, Herzlichkeit) beobachtet. Beitrag 2 zeigte, dass die Lehrkräfte ihre Agency adaptiv an die Kompetenzen der Schüler:innen anpassten und dass ihre Communion davon abhing, ob die Schüler:innen ebenfalls hohe Communion zeigten und welche Kompetenz die Schüler:innen aufwiesen. Beitrag 3 demonstriert, dass Lehrkräfte sich im Fach Deutsch gegenüber den negativ stereotypisierten Jungen stärker anleitend verhielten (hohe Agency) als gegenüber Mädchen. Im Fach Mathematik zeigte sich kein Geschlechtsunterschied im Lehrkraftverhalten. Beitrag 4 fokussiert auf Interaktionen zwischen Schulleitungen und Lehrkräften und stellt ein Instrument vor, das Schulleitungsverhalten auf den Dimensionen Agency und Communion misst. Das Instrument erwies sich als valide und es zeigten sich Zusammenhänge zwischen Schulleitungsverhalten und den durch die Lehrkräfte wahrgenommenen Leitungsaspekten Vertrauen, Führungskompetenz und Klimaarbeit. Die Beiträge dieses Symposiums bieten einen umfassenden Blick auf Interaktionen in der Schule und präsentieren verschiedene Ansätze zur Betrachtung von qualitätsvollen Beziehungen zwischen den verschiedenen schulischen Akteur:innen.

Beiträge des Symposiums

Beobachtete konstruktive Unterstützung in Lehrkraft-Schüler:innen Interaktionen während der freien Lernzeit an Grundschulen

Britta Wenzel, Jasmin Decristan

Bergische Universität Wuppertal

Konstruktive Unterstützung der Lehrkraft ist eines von drei zentralen Unterrichtsqualitätsmerkmalen (Fauth et al., 2014; Klieme et al., 2009) und fördert die Schüler:innenmotivation und das Lernen (z. B. Cornelius-White, 2007). In neueren Arbeiten wurden zwei Subdimensionen konstruktiver Unterstützung diskutiert: die sozio-emotionale und die fachlich-inhaltliche Unterstützung (Decristan et al., 2022; Sliwka et al., 2019). Sozio-emotionale Unterstützung beinhaltet respektvolle und fürsorgliche Interaktionen (Fauth et al., 2014; Klieme et al., 2009) sowie Empathie und Sensitivität (Lipowsky et al., 2009). Fachlich-inhaltliche Unterstützung fokussiert konkrete Lernunterstützung durch verschiedene Hilfsangebote, z. B. Scaffolding (Van De Pol et al., 2010) oder konstruktives Feedback (Hattie & Timperley, 2007). Die Unterscheidung der Unterstützungsdimensionen wurde theoretisch hergeleitet (Sliwka et al., 2019), eine systematische empirische Untersuchung fehlt jedoch. Zudem wird in Beobachtungsstudien konstruktive Unterstützung üblicherweise auf der Klassenebene erhoben (z. B. Fauth et al., 2014). Dabei bleibt unberücksichtigt, dass Lehrkräfte den Unterricht an unterschiedliche Lernvoraussetzungen ihrer Schüler:innen anpassen (Corno, 2008), sodass es zu interindividuellen Unterstützungsunterschieden kommen kann. Diese Unterschiede sind in Lehrkraft-Schüler:innen-Interaktionen beobachtbar. In welchem Ausmaß diese auftreten und ob sie mit individuellen Schüler:innenmerkmalen zusammenhängen, wurde bisher noch nicht untersucht.

Es wurde ein Beobachtungsinstrument entwickelt, das zwischen den beiden Subdimensionen unterscheidet und Interaktionen fokussiert. Zudem wurde geprüft, (1) wie häufig, wie lange und in welcher Valenz Lehrkräfte Grundschüler:innen in Interaktionen sozio-emotional und fachlich-inhaltlich unterstützen und (2) welche Zusammenhänge sich zwischen individuellen Schüler:innenmerkmalen und sozio-emotionaler sowie fachlich-inhaltlicher Unterstützung zeigen.

Die Stichprobe setzt sich zusammen aus 202 Schüler:innen in 14 dritten und vierten Klassen an fünf Schulen, die für ihren besonderen Umgang mit Vielfalt ausgezeichnet wurden. Die individuellen Merkmale der Schüler:innen wurden zu Beginn der Studie über Fragebögen und Leistungstests erhoben. Für die Beobachtungen von konstruktiver Unterstützung wurde für jede Klasse eine 30-minütige Sequenz freier Lernzeit ausgewählt. Im ersten Kodierschritt wurden die Beginn- und Endzeiten von Interaktionen der Lehrkraft mit einer:m Schüler:in identifiziert (PÜ=74%). Im zweiten Schritt wurde jede dyadische Interaktion hinsichtlich der Valenz (positive, neutrale oder negative Ausprägung) beider Unterstützungsdimensionen eingeschätzt (PÜ=78%).

Zur Beantwortung der ersten Forschungsfrage dienen deskriptive Statistiken. Für die zweite Fragestellung wurde ein Score berechnet, der die Häufigkeit, Dauer und Ausprägung jeder Unterstützungsdimension umfasst. Damit erhält jede:r Schüler:in einen individuellen Unterstützungsscore. Schließlich wurden Korrelationen zwischen den für das Lernen zentralen individuellen

Schüler:innenmerkmalen (Lesefähigkeiten, sonderpädagogischer Förderbedarf, Selbstkonzept, Leistungsangst) und beobachteter Unterstützung berechnet.

Insgesamt wurden 873 Interaktionen kodiert, die 193 der 202 Schüler:innen adressierten (Mittlere Dauer (MD) = 28s). Bei der sozio-emotionalen Unterstützung waren 780 Interaktionen positiv (MD = 30s), 17 neutral (MD = 20s) und 28 negativ (MD = 30s), für 48 Interaktionen konnte keine sozio-emotionale Unterstützung kodiert werden. Bei der fachlich-inhaltlichen Unterstützung wurden 523 Interaktionen positiv (MD = 41s), 2 neutral (MD = 25s) und 18 negativ (MD = 18s) kodiert, für 330 Interaktionen konnte keine fachlich-inhaltliche Unterstützung kodiert werden. Bezüglich der zweiten Fragestellung zeigt sich, dass Lehrkräfte Schüler:innen mit hoher Lesegeschwindigkeit und hohem Leseverständnis am meisten sozio-emotional (rLesegeschwindigkeit = .24, $p < .05$; rLeseverständnis = .25, $p < .05$) und fachlich-inhaltlich (rLesegeschwindigkeit = .30, $p < .05$; rLeseverständnis = .28, $p < .05$.) unterstützten. Für Schüler:innen mit Förderschwerpunkten, das Selbstkonzept und die Leistungsängstlichkeit resultierten keine signifikanten Zusammenhänge mit der beobachteten Unterstützung.

Die Befunde geben erste Hinweise darauf, dass Lehrkräfte ihr Unterstützungsverhalten an individuellen Voraussetzungen der Schüler:innen anpassen. Es ergeben sich jedoch einige Anschlussfragen, die bis zur GEBF 2025 anvisiert werden: 1. Warum wurden Schüler:innen mit hohen Lesefähigkeiten mehr unterstützt? Hierzu wird analysiert, ob diese Schüler:innen die Interaktionen initiiert und damit aktiv Unterstützung eingefordert haben. 2. Warum zeigen sich keine Zusammenhänge mit sonderpädagogischen Förderbedarfen? Es wird ermittelt, ob weiteres Fachpersonal zur Unterstützung dieser Schüler:innen im Raum war. 3. Wie verhalten sich die Befunde, wenn Unterschiede zwischen den Klassen berücksichtigt werden? Hierzu werden Mehrebenenregressionen berechnet.

Der Zusammenhang von dyadischem Lehrkraftverhalten mit dem Verhalten und der Kompetenz der Schüler:innen

Madita Frühauf, Karoline Koeppen, Madeleine Kreutzmann, Bettina Hannover

Freie Universität Berlin

Lehrkraftverhalten, das sich für Schülerinnen und Schüler als lernförderlich erwiesen hat, ist durch hohe Wärme (Roorda et al., 2017) und an die Kompetenz des Lernenden angepasstes Verhalten charakterisiert (adaptives Unterrichten; siehe Bernard et al., 2019; Corno, 2008). Gemäß der Interpersonal Theory (Leary, 1957) kann interpersonelles Lehrkraftverhalten gegenüber einem/ einer Lernenden in einem Zirkumplex beschrieben werden, das auf einer Kreuzung der beiden orthogonalen Dimensionen Communion und Agency basiert. Communion umfasst dabei Qualitäten, die für den Aufbau und die Aufrechterhaltung sozialer Beziehungen erforderlich sind (z. B. Wärme, Herzlichkeit), während Agency Qualitäten beinhaltet, die für das Erreichen von Zielen erforderlich sind (z. B. Führung, Anleitung). Doch wovon hängt es ab, wie stark sich die Lehrkraft in dyadischen Interaktionen mit einem/ einer Lernenden verhält? Wir argumentieren, dass Lehrkraft-Communion umso stärker ist, je stärker der jeweilige Lernende Communion zeigt, und dass die Lehrkraft-Agency umso geringer ist, je geringer das Niveau der schulischen Kompetenz der/des Lernenden (noch) ist (sog. Komplementaritätsprinzip in der Interpersonal Theory; Sadler et al., 2009).

In einer Beobachtungsstudie schätzten 39 verschiedene Beobachtende in 39 Schulklassen der Jahrgänge 1-12 das Verhalten der Lehrkraft gegenüber jeweils zwei Schülerinnen und Schülern ($n = 78$ Dyaden, $n_{\text{Mädchen}} = 25$) auf einem Zirkumplex-Instrument ein, das acht verschiedene Verhaltensfacetten mit unterschiedlichen Ausprägungen von Communion und Agency erfasst (22 Items, z. B. Facette Anleitend (hohe Agency und moderate Communion): „Die Lehrkraft erklärt dem Schüler/ der Schülerin die Lösung eines Problems sehr häufig Schritt für Schritt.“, Hannover et al., 2022). Außerdem machten die Beobachtenden Angaben dazu, wie stark oder schwach kommunal sich der Schüler/die Schülerin gegenüber der Lehrkraft verhält (5 Items; „In meiner Wahrnehmung verhält sich der Schüler/ die Schülerin gegenüber der Lehrkraft a) herzlich, b) kooperativ, c) höflich und respektvoll, d) kühl und reserviert, e) feindselig“, Scherzinger et al., 2021) und wie stark die überfachlichen Kompetenzen (5 Items, z. B. „In meiner Wahrnehmung arbeitet und lernt der Schüler/ die Schülerin selbstständig.“, Helm et al., 2012) des Schülers/ der Schülerin ausgeprägt sind. Zudem erfassten wir die fachlichen Kompetenzen über die Zeugnisnote im beobachteten Schulfach. Als Analysestrategie wendeten wir die korrelationsbasierte Structural-Summary-Method an (SSM; Gurtman, 1992; Zimmermann & Wright, 2017). Die SSM-Analyse bestimmt den Zusammenhang eines externen Maßes (hier: Communion und Kompetenz des Schülers/ der Schülerin) mit dem interpersonellen Lehrkraftverhalten unter der Berücksichtigung der zirkulären Struktur. Die SSM-Analysen zeigten erwartungsgemäß, dass hohe Lehrkraft-Communion mit hoher Lernenden-Communion (herzlich: $\delta = 342.2^\circ$, CI [306.8, 14.7]; kooperativ: $\delta = 314.9^\circ$, CI [288.9, 340.2]; höflich und respektvoll: $\delta = 321.6^\circ$, CI [295.3, 346.3]) und geringe Lehrkraft-Communion der Lehrkraft mit niedriger Lernenden-Communion (kühl und reserviert: $\delta = 160.5^\circ$, CI [110.4, 200.8]; feindselig: $\delta = 122.7^\circ$, CI [191.0, 152.4]) zusammenhing. Außerdem zeigte sich erwartungsgemäß, dass die Lehrkraft umso höhere Agency in dyadischen Interaktionen zeigte, je geringer die fachliche und überfachliche Kompetenz des Schülers/der Schülerin ausgeprägt war (fachliche Kompetenzen: $\delta = 310.5^\circ$, CI [288.4, 333.6]; überfachliche Kompetenzen $\delta = 318.8^\circ$, CI [293.2, 340.9]), ein Beleg für adaptives Lehrkraftverhalten. Zudem korrelierte unerwarteterweise die Communion der Lehrkraft positiv mit der Kompetenz des Schülers/der Schülerin (fachliche Kompetenz: $r = .34$, CI [.16, .52]; überfachliche Kompetenz: $r = .45$, CI [.23, .65]).

Unsere Ergebnisse unterstützen das Komplementaritätsprinzip in dyadischen Lehrkraft-Lernenden-Interaktionen: Lehrkräfte verhalten sich stärker kommunal gegenüber kommunal agierenden Lernenden und stärker agentisch gegenüber Lernenden mit geringen Kompetenzen. Wir diskutieren unsere Ergebnisse, dass Lernende, die noch nicht gelernt haben, sich sozial kompetent gegenüber der Lehrkraft zu verhalten, und Lernende, die relativ geringe Kompetenzen haben, Gefahr laufen, in dyadischen Interaktionen ungünstiges, nämlich wenig kommunales Lehrkraftverhalten zu erfahren.

Qualität des Lehrkraftverhaltens in dyadischen Lehrkraft-Lernenden-Beziehungen in den Fächern Deutsch und Mathematik: Unterschiede in Abhängigkeit des Geschlechts des Lernenden und der Lehrkraft

Ninel Kuhl, Madita Frühauf, Bettina Hannover

Freie Universität Berlin

Schülerinnen und Schüler profitieren in ihrer Motivation und Lernentwicklung von qualitativem interpersonalem Verhalten ihrer Lehrkraft (Wentzel, 2023). Qualitätsvolles Lehrkraftverhalten in dyadischen Lehrkraft-Schüler:in-Beziehungen zeichnet sich dabei durch hohe Communion (Wärme, Sensitivität) und von an die Kompetenzen des Lernenden angepasste Agency (Lenkung, fachliche Unterstützung) aus (z. B. Hannover et al., 2022; Roorda et al., 2013).

Wir haben gefragt, ob Geschlechtsstereotype einen Einfluss auf interpersonales Lehrkraftverhalten nehmen. Geschlechtsstereotype können ebenfalls auf den beiden Dimensionen Communion und Agency beschrieben werden. Dem Stereotype Content Model zufolge (Cuddy et al., 2009) wird weiblichen Personen mehr Communion zugeschrieben.

Entsprechend haben wir vorhergesagt, dass (a) weibliche Lehrkräfte sich stärker kommunal verhalten als männliche und dass (b) Schülerinnen mehr kommunales Lehrkraftverhalten erfahren als Schüler.

Im Schulkontext spielen neben dem allgemeinen Geschlechtsstereotyp (Cuddy et al., 2009) aber auch fachspezifische Stereotype eine Rolle. Die Schulfächer Deutsch und Mathematik sind mit unterschiedlichen, geschlechtsbezogenen Stereotypen assoziiert. Während hohe Leistungen in Deutsch (Lesen) eher mit Schülerinnen als mit Schülern assoziiert werden, gilt das Umgekehrte für Mathematik (z. B. Steffens & Jelenec, 2011). Entsprechend haben wir erwartet, dass im Fach Deutsch Lehrkräfte beider Geschlechter sich gegenüber Schülern stärker agentisch verhalten als gegenüber Schülerinnen, da sie annehmen, Schüler bedürften einer stärkeren Lenkung und fachlichen Unterstützung. Umgekehrt sollten im Fach Mathematik Lehrkräfte beider Geschlechter sich gegenüber Schülerinnen stärker agentisch verhalten als gegenüber Schülern.

Hierzu haben wir 109 Lehrkräfte in ihrem Verhalten gegenüber jeweils zwei Schüler:innen in ihrer Klasse in drei aufeinander folgenden Unterrichtsstunden à 45 Minuten im Mathematikunterricht (42 Lehrkräfte) oder Deutschunterricht (67 Lehrkräfte) beobachtet. Dies resultierte in Angaben zu insgesamt 216 Lehrkraft-Schüler:innen-Dyaden (109 Lehrkräfte, davon 71 Lehrerinnen [11 fehlende Werte]; 216 Schüler:innen der Jahrgangsstufen 1 bis 12, davon 85 Schülerinnen [18 fehlende Werte]). Die Beobachter:innen beschrieben das interpersonelle Lehrkraftverhalten gegenüber den einzelnen Schüler:innen mithilfe eines Zirkumplex-Instrumentes (Hannover et al., 2022): Durch die Kreuzung der beiden orthogonalen Verhaltensdimensionen Agency und Communion entsteht eine zirkuläre Struktur, die sich aus vier Quadranten (hohe Agency mit hoher Communion, hohe Agency mit niedriger Communion, niedrige Agency mit hoher Communion, niedrige Agency mit niedriger Communion) zusammensetzt (Leary, 1957). Als Analysestrategie wendeten wir die mittelwertbasierte Structural-Summary-Method an (Zimmermann & Wright, 2017), die auf Grundlage der zirkulären Struktur gruppenspezifische Profile für das Lehrkraftverhalten aus der Kombination von Agency und Communion ermittelt. Entgegen den Annahmen zeigte sich kein signifikanter Unterschied in der Agency oder Communion des Lehrkraftverhaltens in Abhängigkeit des Geschlechts der Lehrkraft. Auch konnte nicht bestätigt werden, dass sich Lehrkräfte gegenüber Schülerinnen kommunal verhalten als gegenüber Schülern. Die fachspezifischen Analysen zeigten für Deutsch erwartungsgemäß einen Unterschied im Lehrkraftverhalten in Abhängigkeit des Geschlechts der Schüler:innen: Die Lehrkräfte verhielten sich gegenüber Schülerinnen im Mittel weniger agentisch als gegenüber Schülern (Schüler: $\delta = 29.3^\circ$, 95% CI [15.9, 43.6] bzw. Schülerinnen: $\delta = 1.8^\circ$, 95% CI [345.7, 17.2]). Für das Fach Mathematik konnte hingegen die Erwartung nicht bestätigt werden, dass Schüler ($\delta = 24.1^\circ$, 95% CI [9.9, 38.0]) weniger Agency erfahren als Schülerinnen ($\delta = 17.6^\circ$, 95% CI [1.0, 32.3]). Die Befunde werden hinsichtlich der Frage diskutiert, wie geschlechtergerechte qualitätsvolle Lehrkraft-Schüler:innen-Beziehungen gefördert werden können.

Das Interaktionsverhalten von Schulleitungen aus Lehrkraftperspektive: Validierung einer deutschen Version des Questionnaire on Principal Interaction

Stefan Markus¹, M. Tim Mainhard², Cornelia Gar³

¹Bergische Universität Wuppertal, ²Universiteit Leiden, ³Ludwig-Maximilians-Universität München

Das Interaktionsverhalten von Schulleitungen spielt für die Schulentwicklung sowie die Berufszufriedenheit von Lehrkräften eine wichtige Rolle (Kremer-Hayon & Wubbels, 1993). Zur Beschreibung des Interaktionsverhaltens von Schulleitungen gegenüber ihren Lehrkräften können die beiden orthogonalen Dimensionen interpersonalen Verhaltens „Communion“ (z. B. emotionale Nähe, Unterstützung) und „Agency“ (z. B. Lenkung, Kontrolle) genutzt werden. Bisherige Befunde zu interpersonalem Verhalten, die bislang insbesondere auf der Ebene der Lehrkraft-Schulkind-Interaktionen vorliegen, zeigen übereinstimmend, dass hohe Communion mit verschiedenen positiven Outcomes zusammenhängt (Mainhard et al., 2018), während die Zusammenhänge mit Agency schwächer bzw. uneindeutig ausfallen (Hannover et al., 2022). Das Interaktionsverhalten kann dabei durch eine Kombination von Ausprägungen auf den beiden Dimensionen Communion und Agency charakterisiert werden, wodurch sich acht Verhaltensfacetten innerhalb eines Zirkumplex-Modells ergeben.

Die vorliegende Studie überprüft anhand einer übersetzten und adaptierten Version des Questionnaire on Principal Interaction (QPI, Kremer-Hayon & Wubbels, 1993) die Angemessenheit eines Zirkumplex-Modells zur Beschreibung des Interaktionsverhaltens von Schulleitungen aus Sicht der Lehrkräfte (FF1) sowie die Zusammenhänge mit den externalen Kriterien Vertrauen, Führungskompetenz und Klimaarbeit, welche mit hoher Communion einhergehen sollten (FF2).

Hierfür wurden in einer Querschnittuntersuchung 323 Lehrkräfte von 25 Schulen online befragt. Vertrauen (Hoy & Tschannen-Moran, 2003), Führungskompetenz und Klimaarbeit (Ittner et al., 2019) wurden anhand etablierter Instrumente erfasst. Die deutsche Version des QPI enthält 33 Items in acht Skalen: führend/strukturierend, helfend/freundlich, verständnisvoll/kooperativ, verantwortungübertragend/tolerierend, unbestimmt/unsicher, oppositionell, streng/kontrollierend, fordernd/streng.

Alle Skalen erwiesen sich als reliabel ($\omega = .73 - .93$). Entsprechend der angenommenen Zirkumplex-Struktur zeigten benachbarte Facetten hohe positive Korrelationen, sich im Zirkumplex gegenüberliegende Facetten hohe negative Korrelationen. Die Struktur des Modells wurde anhand der Structural Summary Method (SSM; Gurtman & Pincus, 2003; Zimmermann & Wright, 2017) für Zirkumplex-Daten unter Verwendung des circumplex-Pakets (Girard, Zimmermann & Wright, 2023) in RStudio (Posit Team, 2024) überprüft. Die Passung der theoretisch angenommenen Zirkumplex-Struktur zu den Daten zeigte mit $R^2 = .92$ einen guten Modellfit. Der Generalfaktor fiel mit $e = 0.47$ [0.47, 0.48] hoch aus, ebenso die Amplitude ($a = 0.34$ [0.32, 0.36]), was auf eine hohe Differenzierung zwischen den verschiedenen Verhaltensfacetten hinweist. Das Displacement $\delta = 353.3^\circ$ [351.2°, 355.5°] deutet darauf hin, dass das prädominante Verhaltensprofil der Schulleitungen in dem Oktanten „verständnisvoll/kooperativ“ lag. Das Verhaltensprofil der Schulleitungen in unserer Stichprobe war dabei mit hoher Communion (0.34 [0.32, 0.36]) assoziiert und unkorreliert mit der Agency-Dimension (-0.04 [-0.05, -0.03]).

Die Korrelationen mit den externalen Kriterien Vertrauen, Führungskompetenz und Klimaarbeit erwiesen sich bei Facetten mit positiver Communion durchwegs stark positiv ($.61 \leq r \leq .91$), während sie bei Oktanten mit negativer Communion stark negativ ausfielen ($-.52 \leq r \leq -.72$). Die Zusammenhänge waren bei Facetten mit mittlerer Agency meist am stärksten ausgeprägt. Auch anhand der SSM konnte gezeigt werden, dass alle drei externalen Kriterien am stärksten mit verständnisvollen/kooperativen Interaktionsstil assoziiert sind und mit positiver Communion (0.89, 0.83, 0.90) sowie leicht negativer Agency (-0.36, -0.27, -0.39) einhergehen ($.88 \leq R^2 \leq .91$).

Die vorliegende Studie kann durch die Übersetzung und Validierung des QPI nicht nur ein reliables deutschsprachiges Instrument zur Erfassung des lehrkraftperzipierten Interaktionsverhaltens von Schulleitungen vorstellen, sondern auch die theoretisch konzeptualisierte Zirkumplex-Struktur mit acht Facetten der Interaktionsgestaltung bestätigen (FF1). Die psychometrischen Kennwerte des Instruments sind als gut einzuschätzen, auch wenn die Stichprobe vermutlich eine Positivauswahl mit hohen Communion- und moderaten Agency-Werten darstellt. Beziehungen zu korrespondierenden Konstrukten belegten die konvergente Validität des Instruments. Aus Sicht der Lehrkräfte gehen Vertrauen in die Schulleitung, wahrgenommene Führungskompetenz und gute Klimaarbeit mit gemäßiger Agency und hoher Communion – insbesondere mit einem

helfend/freundlichen, verständnisvoll/kooperativen und verantwortungsübertragend/tolerierenden Interaktionsstil der Schulleitung – einher (FF2). Unter Beachtung von Limitationen können somit frühere Befunde in deutschsprachigen Ländern repliziert bzw. auf die Interaktionsebene zwischen Schulleitungen und Lehrkräften erweitert und Implikationen abgeleitet werden.

S-16: Mehr als ein Augenblick – Wie Eye-Tracking die Unterrichtsforschung bereichert

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: SN 169

Chair der Sitzung: Özün Keskin

Chair der Sitzung: Tosca Daltoé

Chair der Sitzung: Kathleen Stürmer

Mehr als ein Augenblick – Wie Eye-Tracking die Unterrichtsforschung bereichert

Chair(s): **Özün Keskin** (Universität Augsburg, Deutschland), **Tosca Daltoé** (Universität Tübingen)

Diskutant*in(nen): **Kathleen Stürmer** (Universität Tübingen)

Seit vielen Jahren untersucht die empirische Bildungsforschung die unterrichtlichen Prozesse des Lehrens und Lernens in standardisierten Studien. Auf Seite der Lehrenden werden beispielsweise professionelle Kompetenzen von Lehrkräften durch die Präsentation von Unterrichtsvideos oder virtuellen Unterrichtssimulationen untersucht (z. B. Richter et al., 2022). Auf Seite der Lernenden werden beispielsweise Effekte von Feedback auf das Lernen durch die Präsentation von Lehrvideos erforscht (z. B. Fidan & Gencel, 2022). Neben der Untersuchung von Outcomes, rückten in diesen Studien in den letzten Jahren verstärkt Prozessdaten, beispielsweise erhoben durch Eye-Tracking, in den Fokus (Keskin et al., 2024; Wang et al., 2021). Blickbewegungsdaten haben das Potenzial, wertvolle Informationen über Kognitionen während Lehr- und Lernprozessen zu liefern (Just & Carpenter, 1976). Das vorliegende Symposium stellt vier aktuelle, innovative Forschungsarbeiten aus der Unterrichtsforschung vor, die Eye-Tracking-Technologie nutzen. Alle vier Beiträge vereint die Frage, wie wir den Blick von Lehrenden (Beitrag 1-3) und Lernenden (Beitrag 4) nutzen können, um ein tieferes Verständnis von Lehr- und Lernprozessen zu erhalten.

Im ersten Beitrag werden geschlechtsspezifische Einstellungen von Lehramtsstudierenden untersucht und mit deren Blickfokus auf Schülerinnen und Schüler in Zusammenhang gebracht. Es werden implizite und explizite geschlechtsspezifische Einstellungen mit Eye-Tracking-Daten aus der Betrachtung von Unterrichtsvideos in Verbindung gebracht und untersucht, ob sich die visuelle Aufmerksamkeit auf Schülerinnen und Schüler je nach Einstellung der Lehrkraft unterscheidet.

Auch der zweite Beitrag adressiert geschlechtsspezifische Einstellungen von Lehrkräften. In diesem Beitrag wird der Einfluss geschlechtsspezifischer Stereotype von Lehramtsstudierenden auf deren professionelle Wahrnehmung von Unterrichtsstörungen anhand von Screenrecordings aus einem virtuellen, simulierten Klassenzimmer untersucht. Es wird erforscht, wie sich die visuelle Aufmerksamkeit auf Unterrichtsstörungen je nach Art der Störung und Geschlecht der störenden Lernenden unterscheidet. Darüber hinaus wird im Beitrag die Rolle impliziter und expliziter Stereotype für die visuelle Wahrnehmung von Unterrichtsstörungen in den Blick genommen.

Der dritte Beitrag widmet sich der Wahrnehmung und Einschätzung von Unterrichtsqualität. Im Beitrag werden traditionelle Unterrichtsvideos und immersive 360-Grad-Unterrichtsvideos bezüglich videobasierter Beobachtung und Einschätzung von Unterrichtsqualität verglichen. In einer randomisierten Studie wird das Blickverhalten von Lehramtsstudierenden während der Unterrichtsbeobachtung erfasst und untersucht, welche Merkmale des Blickverhaltens in den beiden Videoumgebungen (traditionell vs. immersiv) mit erfolgreichen Einschätzungen der Unterrichtsqualität zusammenhängen.

Im vierten Beitrag werden die Aufmerksamkeitsprozesse Lernender während des Lernens in einer Online-Umgebung in den Blick genommen. In einer kontrollierten Laborstudie mit Studierenden wird die Blicksynchronität Lernender als Indikator für deren Aufmerksamkeit genutzt. Es wird untersucht, wie die Blicksynchronität während der Teilnahme an einer Online-Vorlesung mit selbstberichteter Aufmerksamkeit sowie Lernerfolg zusammenhängt.

Beiträge des Symposiums

Spiegelt sich die Einstellung von angehenden Lehrkräften in ihrem Blickverhalten wider?

Sylvia Gabel, Andreas Gegenfurtner

Universität Augsburg

Theoretischer Hintergrund

Einstellungen sind positiv, negativ oder neutral bewertende Assoziationen gegenüber bestimmten Gruppen (Fazio, 2007). Bezüglich einer Geschlechtergerechtigkeit im Klassenzimmer sollten Lehrkräfte eine neutrale oder positive Einstellung gegenüber sowohl Mädchen als auch Jungen haben, ohne eine Gruppe zu bevorzugen oder zu benachteiligen. Während explizite Einstellungen bewusst geäußert werden, werden implizite Assoziationen automatisch und unbewusst verarbeitet (Greenwald et al., 2003). Um den Effekt der sozialen Erwünschtheit bei der Befragung nach expliziten Einstellungen zu vermeiden, sind implizite Assoziationstests eine wertvolle Methode, um unbewusste Einstellungen zu messen (Greenwald et al., 2003). Glock und Kleen (2017) fanden heraus, dass Lehrkräfte tendenziell ungünstigere implizite Persönlichkeitsvorstellungen über das soziale Verhalten von Jungen haben und neigen dazu, Mädchen im Klassenzimmer implizit zu bevorzugen. Ob diese Einstellungen auch mit visuellen Präferenzen korrelieren, ist Fragestellung vorliegender Studie. Eye-Tracking kann Informationen darüber liefern, wie viel Aufmerksamkeit auf Jungen und Mädchen im Klassenzimmer gerichtet wird und so ungleiche Tendenzen aufdecken. Bisher bleibt jedoch offen, inwieweit Einstellungen mit dem Blickverhalten von Lehrkräften zusammenhängen.

In dieser multimethodischen Studie wurden drei Hypothesen getestet:

H1: Lehramtsstudierende haben eine positivere implizite (a) und explizite Einstellung (b) gegenüber Mädchen.

H2: Basierend auf vorangegangenen Studien über eine positivere Einstellung gegenüber Mädchen fokussieren Lehramtsstudierende ihre visuelle Aufmerksamkeit mehr auf Mädchen als auf Jungen.

H3: Die Einstellungen korrelieren positiv mit der Anzahl der Fixationen auf Mädchen.

Methode

An dieser Studie nahmen 135 Lehramtsstudierende teil. Die Mehrheit der Teilnehmenden war weiblich (79,9%) und im ersten Semester (65,7%). Eine Woche zuvor füllten die Teilnehmenden einen Fragebogen zu ihren expliziten Einstellungen aus. Wir integrierten ein Gefühlsthermometer zur Messung expliziter Einstellungen (Alwin, 2007) für beide Geschlechtergruppen.

Im Forschungslabor sahen sie ein Unterrichtsvideo einer Sekundarstufe an. Die Teilnehmenden hatten keine Hintergrundinformationen zu den Schüler*innen. Für die Analyse wählten wir eine Szene (21 Sekunden), in der drei Mädchen und drei Jungen in einer Gruppe arbeiteten. Alle Schülerinnen verhielten sich ähnlich, sodass wir aufmerksamkeitsrelevante Effekte aufgrund auffälligen Verhaltens ausschließen konnten. Diese Schüler*innen waren die AOIs für unsere Eye-Tracking-Analyse und hatten eine ähnliche Pixelgröße. Wir operationalisierten das Blickverhalten der Lehramtsstudierenden durch die Anzahl der Fixationen für die Mädchen und Jungen.

Anschließend führten die Teilnehmenden einen Impliziten Assoziationstest (IAT) durch, der angenehme und unangenehme Attribute (z.B. fleißig/faul) und männliche und weibliche Namen (z.B. Paul/Paula) beinhaltete. Ein positiver d-Wert bedeutete eine positive Einstellung zu Mädchen.

Aufgrund schlechter Datenqualität mussten 31 Teilnehmende von der Analyse ausgeschlossen werden.

Ergebnisse

Die expliziten Einstellungen waren signifikant höher für Mädchen als für Jungen ($t(104)=2.452$; $p=.016$). Sowohl männliche als auch weibliche Lehramtsstudierende hatten eine positivere explizite Einstellung gegenüber Mädchen. Die impliziten Einstellungen zeigten eine moderate Präferenz für Mädchen ($M=.57$; $SD=.41$), wobei weibliche Lehrkräfte ($M=.72$; $SD=.27$) eine stärkere Präferenz zeigten als männliche Lehrkräfte ($M=.02$; $SD=.32$).

Die Eye-Tracking-Daten zeigten, dass Lehramtsstudierende signifikant häufiger Mädchen fixierten als Jungen ($t(104) = 10.971$; $p < .001$). Es gab eine positive, signifikante Korrelation zwischen dem d-Wert und der Anzahl der Fixationen auf Mädchen ($r = .25$; $p = .012$).

Diskussion

Trotz vergleichbarem Verhalten von Mädchen und Jungen standen Mädchen im Fokus der Blicke der Lehramtsstudierenden. Positive Tendenzen zugunsten von Mädchen bei expliziten und impliziten Einstellungen bestätigte frühere Forschungsergebnisse (Glock & Kleen, 2017). Diese Ergebnisse deuten darauf hin, dass Einstellungs- und visuelle Präferenzen gleichzeitig auftreten können, wenn Lehramtsstudierende Klassenzimmersituationen beobachten, und dass sie möglicherweise aufgrund ihrer impliziten Präferenz für Schülerinnen mehr Aufmerksamkeitsressourcen auf Mädchen als auf Jungen lenken. Wir konnten keine signifikanten Korrelationen mit expliziten Einstellungsmaßen finden, da diese möglicherweise durch soziale Erwünschtheit verzerrt sind (Greenwald et al., 2003). Zukünftige Forschung sollte kausale Einflüsse weiter untersuchen sowie den Einfluss von Erwartungen und damit auch die fachspezifischen, geschlechter-spezifischen Einstellungen miteinbeziehen.

Mädchen sind konzentriert & Jungen rebellisch. Der Einfluss geschlechtsspezifischer Einstellungen auf die professionelle Wahrnehmung von Lehramtsstudierenden: Eine Blickbewegungsstudie.

Antje Biermann, Eva Mayer, Ann-Sophie Grub
Universität des Saarlandes

Theoretischer Hintergrund

Lehrkräfte treffen eine Vielzahl an Entscheidungen, um den Lehr-Lern-Prozess aufrecht zu halten, Konflikte zu lösen, Schüler*innen individuell zu fördern und zu motivieren. In dieser komplexen Gemengelage bleibt selten Zeit, um Entscheidungen auf einer ausreichenden Informationsbasis treffen zu können. Dies erhöht die Wahrscheinlichkeit für die Nutzung kognitiver Abkürzungen wie Einstellungen und Stereotype (z.B. Martiny & Froehlich, 2020). Die (affektive) Zuschreibung von Eigenschaften und Verhaltensweisen zu Mitgliedern sozialer Gruppen kann als Filter für die Wahrnehmung von Ereignissen dienen und Entscheidungen und Verhaltensweisen der Lehrkraft beeinflussen (Glock et al., 2016; Tobisch & Dresel, 2022). Gerade in komplexen Unterrichtssituationen bei hoher kognitiver Belastung werden implizite Stereotype oder Einstellungen wirksam, was insbesondere Noviz*innen anfälliger für eine davon beeinflusste Urteilsbildung macht. Untersucht wurden stereotypbasierte Prozesse der Urteilsbildung für Lehrkräfte unter anderem für den sozialen Status (Tobisch & Dresel, 2020), für Migrationshintergrund (Schuchart, 2020) sowie für das Geschlecht (z.B. Glock & Kleen, 2017). Glock und Kleen (2017) konnten z.B. zeigen, dass externalisierende Verhaltensweisen fast ausschließlich Jungen und internalisierende Verhaltensweisen fast ausschließlich Mädchen zugeschrieben wurden. Der Einfluss von Stereotypen und Einstellungen auf die professionelle Wahrnehmung von Lehrkräften ist bisher nur ansatzweise für Migrationshintergrund und Hautfarbe von Schüler*innen untersucht: Ebricht et al. (2021) zeigten, dass der Aufmerksamkeitsfokus deutlich stärker auf Schwarzen statt auf Weißen Schüler*innen liegt und dass jede Störung, die von Schwarzen Schüler*innen ausgingen, entdeckt wurde, nicht jedoch bei Weißen Schüler*innen. Keskin et al. (2024) zeigten ebenfalls einen stärkeren Aufmerksamkeitsfokus auf Schüler*innen mit Migrationshintergrund. Negative Einstellungen gegenüber Schüler*innen aus Minoritäten könnten hier zu stärkerem Monitoring deren Verhaltens geführt haben. Studien, die den Einfluss von Einstellungen und Stereotypen zum Geschlecht von Schüler*innen auf das Wahrnehmungsverhalten von (angehenden) Lehrkräften untersucht haben, sind uns nicht bekannt. In der vorliegenden Studie wird der Zusammenhang von expliziten und impliziten Einstellungen auf die Wahrnehmung von Schülerstörungen bei Mädchen und Jungen von angehenden Lehrkräften untersucht.

Methode

Zur Überprüfung der Fragestellung haben wir ein randomisiertes Mixed Design-Experiment durchgeführt ($N=67$ Lehramtsstudierende, 67% weiblich). Als Stimulusmaterial wurden vier kurze virtuelle Unterrichtsszenarien verwendet, die mit der Software Teach-R (Wiepke et al., 2019) generiert wurden. Die Szenarien zeigen eine Stillarbeit mit je einer Unterrichtsstörung. Variiert wurde die Art der Störung (keine Störung, passive und aktive Störung) und das Geschlecht der virtuellen Schüler*innen. Die Studierenden sollten die Unterrichtssituation beobachten, wobei deren Blickbewegungsdaten mittels Eyetracker aufgezeichnet wurden. Im Anschluss wurde den Probanden ein Impliziter Assoziationstest (Greenwald et al., 1998) mit den Kategorien Geschlecht und Verhaltensweisen nach Glock und Kleen (2020), sowie ein selbst entwickelter Test zu expliziten Einstellungen vorgelegt. Für die Analysen wurden ANOVAs mit Messwiederholung auf den Faktoren Schüler-Geschlecht und Art der Störung berechnet. Implizite und explizite Stereotype gingen als Zwischensubjektfaktor (Mediansplit) in die Analyse ein. Als Parameter für Aufmerksamkeit wurden die Blickbewegungsparameter Fixationsanzahl und Revisits sowie die mittlere Fixationsdauer genutzt (Biermann et al., 2023).

Ergebnisse

Die Art der Störung wirkte sich auf die Aufmerksamkeit aus: Die aktive Störung wurde mit der größten Aufmerksamkeit belegt (höhere Fixationsanzahl [$F(1,65)=153.65$, $\eta^2=.83$] und Revisits [$F(1,65)=84.56$, $\eta^2=.73$], längere Fixationsdauer [$F(1,65)=51.57$, $\eta^2=.62$], alle $p<.05$) gefolgt von der passiven Störung und keiner Störung. Auch das Geschlecht hat einen Effekt: Das Mädchen wurde prinzipiell mit mehr Aufmerksamkeit belegt als der Junge (Fixationsanzahl: $F(1,65)=77.56$, $\eta^2=.55$, Revisits: $F(1,65)=39.73$,

$\eta^2=.38$, Fixationsdauer: $F(1,65)=20.54$, $\eta^2=.24$, alle $p<.05$). Liegt allerdings ein hohes implizites Stereotyp vor (Mädchen: positives Verhalten, Jungen: negatives), liegt der Aufmerksamkeitsfokus stärker auf dem Jungen, wenn eine Störung vorliegt (längere Fixationsdauer: $F=3.45$, $\eta^2=.55$, $p<.05$, Fixationsanzahl und Revisits: n.s.). Explizite Einstellungen haben keinen Einfluss auf die Aufmerksamkeit. Die Ergebnisse liefern damit Hinweise darauf, dass implizite Einstellungen in Hinblick auf das Schülergeschlecht den Aufmerksamkeitsfokus beeinflussen können. Zukünftige Studien sollen auch konkrete Handlungsentscheidungen mit in den Blick nehmen.

Die Qualität im Blick – Zusammenhänge von Blickverhalten und Einschätzungen der Unterrichtsqualität in verschiedenen Videoumgebungen

Tosca Daltoé¹, Tobias Appel¹, Philipp Stark¹, Birgit Brucker², Anika Dreher³, Benjamin Fauth¹, Marita Friesen⁴, Peter Gerjets², Linn Hansen³, Ulrich Trautwein¹, Richard Göllner¹

¹Universität Tübingen, ²Leibniz-Institut für Wissensmedien Tübingen, ³Pädagogische Hochschule Freiburg, ⁴Pädagogische Hochschule Heidelberg

Theoretischer Hintergrund

Unterrichtsbeobachtungen durch externe Beobachtende sind eine zentrale Perspektive zur Erfassung von Unterrichtsqualität (Fauth et al., 2020). Obwohl Unterrichtsbeobachtungen in der Schulpraxis und Unterrichtsforschung vielfach eingesetzt werden, erwies es sich als herausfordernd, die Qualität von Unterricht beobachtungsbasiert zuverlässig einzuschätzen (z. B. Praetorius et al., 2012). In der Literatur werden unterschiedliche Faktoren diskutiert, die die Qualität von Unterrichtsbeobachtungsratings beeinflussen könnten, wie beispielsweise die Intensität der Schulung (Gitzi et al., 2023) oder der fachliche Hintergrund der Raterinnen und Rater (Dreher & Leuders, 2021).

Aus der Unterrichtsforschung lassen sich zwei weitere Faktoren ableiten, die potenziell zur Erklärung der Qualität von Unterrichtsbeobachtungsratings beitragen können. Zum einen könnte das Blickverhalten von Beobachtenden Rückschlüsse auf deren Wahrnehmung des Unterrichts zulassen. Es ist zu vermuten, dass Beobachtende die Unterrichtsqualität nur dann zuverlässig einschätzen können, wenn sie qualitätsrelevante Ereignisse im Unterricht überhaupt erkennen (vgl. „Noticing“; König et al., 2022). Demnach könnte die visuelle Aufmerksamkeit auf die relevanten Unterrichtseignisse während Unterrichtsbeobachtungen mit der zuverlässigen Einschätzung der Unterrichtsqualität zusammenhängen. Darüber hinaus könnte der Pupillendurchmesser potenziell ein Indikator für die Wahrnehmung qualitätsrelevanter Ereignisse während der Unterrichtsbeobachtung sein (z. B. Goldinger & Papesh, 2012). Neben dem Blickverhalten könnte zum anderen auch die Videoumgebung und das damit einhergehende Unterrichtserleben die Unterrichtsbeobachtungsratings beeinflussen. In bisheriger Forschung zeigte sich, dass sich Beobachtende in immersiven 360-Grad-Videoumgebungen stärker in den Unterricht eingebunden fühlen als in traditionellen Videoumgebungen am PC, was teilweise mit einer veränderten Unterrichtseinschätzung einhergeht (Daltoé et al., 2024; Gold & Windscheid, 2020).

Ziel der vorliegenden Studie ist es, den Einfluss des Blickverhaltens sowie der Videoumgebung (traditionell vs. immersiv) auf die Qualität von Unterrichtsbeobachtungsratings zu untersuchen. Folgende Forschungsfragen stehen im Fokus der Untersuchung:

1. Unterscheidet sich die Verteilung der visuellen Aufmerksamkeit während qualitätsrelevanter Unterrichtseignisse in traditionellen und immersiven Videoumgebungen?
2. Hängen visuelle Aufmerksamkeit und Pupillendurchmesser in traditionellen und immersiven Videoumgebungen mit der Qualität von Unterrichtsbeobachtungsratings zusammen?

Methode

In der vorliegenden Studie betrachteten $N = 75$ Lehramtsstudierende jeweils zwei Unterrichtsvideos: eines in einer traditionellen Videoumgebung am PC und eines in einer immersiven Videoumgebung mit Virtual Reality-Brillen. Während der Unterrichtsbeobachtungen wurden in beiden Videoumgebungen Blickbewegungsdaten mit Eye-Trackern erhoben. Es wurde die visuelle Aufmerksamkeit auf die beiden Areas of Interest (AOIs) „Lehrkraft“ und „Schülerinnen und Schüler“ (SuS), die Häufigkeit der Übergänge zwischen den AOIs sowie der Pupillendurchmesser erfasst. Im Anschluss an die Unterrichtsbeobachtung schätzten die Teilnehmenden die Qualität des Unterrichts mit dem Unterrichtsfeedbackbogen Tiefenstrukturen (Fauth et al., 2022) ein, einem Beobachtungsbogen, der auf den drei Basisdimensionen der Unterrichtsqualität (kognitive Aktivierung, konstruktive Unterstützung, strukturierte Klassenführung; Klieme et al., 2009) basiert. Die Unterschiede im Blickverhalten und dessen Zusammenhänge mit Unterrichtsbeobachtungsratings zwischen den Videoumgebungen wurden mit Mehrebenen-Regressionsanalysen untersucht.

Ergebnisse

Die Ergebnisse weisen darauf hin, dass sich Beobachtende in der immersiven Videoumgebung visuell stärker auf die Lehrkraft und in der traditionellen Videoumgebung stärker auf die SuS fokussieren (Forschungsfrage 1). Darüber hinaus zeigen sich signifikante Zusammenhänge zwischen Eye-Tracking-Maßen und Rating-Qualität (Forschungsfrage 2): Je größer der Pupillendurchmesser während kritischer Unterrichtsevents, desto akkurater fiel das Rating des Aspekts „Verständnisorientierung“ (kognitive Aktivierung) aus. Je stärker Beobachtende die Lehrkraft fokussierten, desto akkurater fiel das Rating des Aspekts „Herausfordernde Aufgaben und Fragen“ (kognitive Aktivierung) aus. Signifikante Interaktionseffekte der Blickbewegungsmaße mit der Videoumgebung wiesen darauf hin, dass je nach Videoumgebungen das Blickverhalten unterschiedlich mit erfolgreichen Unterrichtsbeobachtungsratings zusammenhängt.

Diskussion

Die Befunde der präsentierten Studie liefern neuartige Erkenntnisse zur Qualität videobasierter Unterrichtsbeobachtungen. Beispielsweise zeigt der Zusammenhang des Blickfokus der Beobachtenden mit erfolgreichen Unterrichtsbeobachtungsratings, dass in Rater-Schulungen thematisiert werden sollte, worauf der Fokus der Unterrichtsbeobachtung (Lehrkraft vs. SuS) liegen sollte, je nachdem, welcher Unterrichtsqualitätsaspekt eingeschätzt werden soll. Außerdem liefern die Ergebnisse Hinweise darauf, dass der Einsatz immersiver 360-Grad-Videoumgebungen im Vergleich zu traditionellen Videoumgebungen in der Lehrkräftebildung und Unterrichtsqualitätsforschung sinnvoll sein kann.

Aufmerksam und synchron: Untersuchung der Beziehung zwischen Blicksynchronität und selbstberichteter Aufmerksamkeit beim Online Unterricht

Babette Bühler¹, Efe Bozkır¹, Hannah Deininger¹, Peter Gerjets², Ulrich Trautwein¹, Enkelejda Kasneci³

¹Universität Tübingen, ²Leibniz-Institut für Wissensmedien Tübingen, ³Technische Universität München

Theoretischer Hintergrund

Erfolgreiches Lernen beruht auf der Fähigkeit der Lernenden, ihre Aufmerksamkeit aufrechtzuerhalten, was im Online-Unterricht aufgrund der begrenzten Interaktion und Unterstützung von Lehrenden eine besondere Herausforderung darstellt und erhöhte Selbstregulation erfordert (Schachter et al., 2015). Für Lehrende hingegen ist es schwieriger, die Aufmerksamkeit der Lernenden zu verfolgen und zu steuern (Sun et al., 2019). Eine der wichtigsten Erfassungsmodalitäten für die Messung der menschlichen Aufmerksamkeit ist Eye Tracking. Madsen et al. (2022) stellten die Hypothese auf, dass aufmerksame Lernende Lehrvideos in ähnlicher Weise mit ihren Augen verfolgen. Durch die Messung der Synchronität mit der Inter-Subjekt-Korrelation (ISC) der Blicke innerhalb von Versuchsgruppen mit aufmerksamen und abgelenkten Lernenden konnten sie zeigen, dass das Niveau der Blicksynchronität unter aufmerksamen Lernenden höher ist und Testergebnisse vorhersagt (Madsen et al., 2022; Liu et al., 2023).

Eine wesentliche Einschränkung der bisherigen Forschung besteht darin, dass die Aufmerksamkeit experimentell manipuliert wurde, indem eine sekundäre Ablenkungsaufgabe eingeführt wurde (Madsen et al., 2021; Liu et al., 2023). Es bleibt ungewiss, ob dies die natürlich auftretenden Ablenkungen ausreichend widerspiegelt (Liu et al., 2023). Die Mechanismen, die Unaufmerksamkeit zugrunde liegen, reichen von offener Ablenkung durch die Umgebung (Wammes et al., 2019), bis hin zu versteckten kognitiven Prozessen wie „Mind-Wandering“, dem Abschweifen hin zu aufgabenfremden Gedanken, z. B. Tagträumen (Lindqvist et al., 2011). In realistischen Lernszenarien kann davon ausgegangen werden, dass sich Aufmerksamkeit dynamisch im Laufe der Zeit entwickelt und schwankt (Farley et al., 2013). Gedankliches Abschweifen ist möglicherweise weniger inhibitorisch, und macht sich weniger stark in Augenbewegungen bemerkbar als die künstlich auferlegte Ablenkungsaufgabe, wirkt sich aber dennoch nachteilig auf das Lernen aus.

Fragestellung

In dieser Studie untersuchen wir den verborgenen Zusammenhang zwischen Blicksynchronität und natürlich auftretender, selbstberichteter Aufmerksamkeit in einem realistischen Lernszenario, unter der Verwendung neuer Synchronitätsmaße. Darüber hinaus untersuchen wir die Eignung von Blicksynchronität als Indikator für Aufmerksamkeit, indem wir Zusammenhang zwischen Blicksynchronität und Lernerfolg analysieren.

Methode

Die für diese Studie verwendeten Daten wurden von $N = 88$ (Alter 19-33 Jahre, $M = 23,44$, $SD = 2,6$) Studierenden im Labor erhoben. Blickbewegungen wurden mit dem SMI Red Remote Eye Tracker (250 Hz) aufgezeichnet. Nach einem Test zum Vorwissen sahen sich die Teilnehmer eine 60-minütige zuvor aufgezeichnete Zoom-Vorlesung an. Danach wurde ein umfassender Wissens-test mit 14 Fragen durchgeführt, die auf Faktenwissen und tieferes Verständnis abzielten. Die Aufmerksamkeit der Teilnehmenden wurde anhand von 15 Erfahrungsstichproben erhoben. Dabei wurden sie während der Vorlesung unterbrochen und gebeten, anzugeben, worauf ihre Aufmerksamkeit in diesem Moment gerichtet war (Weinstein, 2018).

Für die Berechnung der Blicksynchronität betrachteten wir die 10-Sekunden-Zeitfenster unmittelbar vor den Erfahrungsstichproben. Als Synchronitätsmaße wurden neben der ISC, Kullback-Leibler-Divergenz von Gaze Density Maps und MultiMatch-Algorithmus-Scanpfadvergleich herangezogen. In einem nächsten Schritt analysierten wir die Beziehung zwischen der Aufmerksamkeit und der Blicksynchronität. Diese Analyse konzentrierte sich auf die Beziehung zwischen je-dem Aufmerksamkeitsselfbericht und der vorhergehenden Blicksynchronität unter der Verwendung von Mehrebenenmodellen. Schließlich analysierten wir den Zusammenhang zwischen Blicksynchronität und Lernerfolg mithilfe linearer Regressionsanalysen, wobei die mittleren Werte für Synchronie und Aufmerksamkeit pro Teilnehmer sowie das Vorwissen berücksichtigt wurden.

Ergebnisse und Bedeutung

Die Ergebnisse zeigten, dass die Blicksynchronisation unter Lernenden während selbstberichteter Aufmerksamkeit in allen verwendeten Synchronitätsmaßen signifikant höher war. Zusätzlich fanden wir einen positiven Zusammenhang zwischen dem Anteil von selbstberichteter Aufmerksamkeit und dem Lernerfolg. Im Gegensatz dazu korrelierten die Synchronisationsmaße nicht mit den Lernergebnissen. Obwohl die Hypothese, dass aufmerksame Lernende ähnliche Augenbewegungen zeigen, unterstützt wird, stellt die direkte Verwendung von Synchronität als Indikator für Aufmerksamkeit eine Herausforderung dar, die weitere Forschung über das Zusammenspiel von Aufmerksamkeit, Blicksynchronität und der Art des Videoinhalts erfordert.

S-17: Heimatnah oder fern der Heimat? Empirische Befunde zur regionalen Mobilität von (angehenden) Lehrkräften

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 150

Chair der Sitzung: Svenja Mareike Schmid-Kühn

Chair der Sitzung: Denise Demski

Heimatnah oder fern der Heimat? Empirische Befunde zur regionalen Mobilität von (angehenden) Lehrkräften

Chair(s): **Svenja Mareike Schmid-Kühn** (Universität Koblenz, Deutschland)

Diskutant*in(nen): **Denise Demski** (Ruhr-Universität Bochum)

Das Thema der regionalen Mobilität über Bundesländergrenzen hinweg gewinnt durch den steigenden Lehrkräftebedarf und den bereits jetzt schon vorherrschenden Lehrkräftemangel an Bedeutung. Zur regionalen Mobilität von Studienberechtigten beim Übergang ins Studium und zum Mobilitätsverhalten von Hochschulabsolventinnen und -absolventen beim Berufseinstieg liegen verschiedene bildungsstatistische Daten (z.B. Haußen & Übelmesser 2015; KMK 2019) und empirische Studien (z.B. Jungbauer-Gans & Gottburgsen 2022; Quast, Mentgens & Föste-Eggers 2022) vor, die mitunter auch auf einzelne Fächer(-gruppen) und/oder Bundesländer konzentriert sind (z.B. Falk & Kratz 2009). Erkenntnisse zur regionalen Mobilität von (angehenden) Lehrkräften zwischen den Bundesländern gibt es hingegen bislang kaum, worauf auch das Gutachten ‚Lehrkräftegewinnung und Lehrkräftebildung für einen hochwertigen Unterricht‘ der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz hinweist (2023, S. 14ff.). Da der Ort zur Aufnahme eines Lehramtsstudiums prinzipiell frei gewählt werden kann und sowohl der Lehrerausbildungs- als auch der Lehrerarbeitsmarkt angesichts der vereinbarten gegenseitigen Anerkennung von Lehramtsausbildungen und -abschlüssen beim Zugang zum Vorbereitungsdienst und für den Berufseinstieg (vgl. KMK 2013) im föderalen System keine Ländergrenzen kennt, kann angenommen werden, dass es auch im Lehramtskontext zu Wechselbewegungen über Ländergrenzen hinweg kommt, und zwar v.a. an zentralen Übergangspunkten (1) vom Schulabschluss mit Hochschulzugangsberechtigung in das Lehramtsstudium über (2) den Übergang in den Vorbereitungsdienst bis (3) zum Eintritt in den Schuldienst. Das Symposium ‚Heimatnah oder fern der Heimat?‘ widmet sich dieser aktuellen Thematik und stellt in drei Beiträgen empirische Befunde zur regionalen Mobilität von (angehenden) Lehrkräften vor: Vortrag 1 betrachtet unter Verwendung der Daten des Nationalen Bildungspanels deutschlandweit regionale Mobilitätsmuster von (angehenden) Lehrkräften bei der Aufnahme des Lehramtsstudiums, dem Übergang in den Vorbereitungsdienst und dem Eintritt in den Schuldienst. Es zeigt sich, dass etwa zwei Drittel der (angehenden) Lehrkräfte über den gesamten Verlauf der Lehramtsausbildung in dem Bundesland verbleiben, in dem die Hochschulzugangsberechtigung erworben wurde und auch dort in den Beruf einsteigen. Ein Drittel hingegen wechselt im Verlauf der Lehramtsausbildung bzw. beim Einstieg in den Lehrkräfteberuf – teilweise mehrfach – das Bundesland, wobei die meisten Wechselbewegungen zu Studienbeginn erfolgen und die Mobilität mit den weiteren institutionellen Übergängen abnimmt. Vortrag 2 knüpft hier unmittelbar an und geht, ebenfalls unter Verwendung der Daten des Nationalen Bildungspanels, der Frage nach, welche soziodemographischen und leistungsbezogenen Faktoren einen Bundeslandwechsel determinieren. Dabei zeigt sich, dass v.a. mit Blick auf die Studienaufnahme die soziale Herkunft das Mobilitätsverhalten beeinflusst, spätere Übergänge aber auch durch die jeweilige Abschlussnote determiniert werden. Vortrag 3 fokussiert den Übergang in das Lehramtsstudium und thematisiert angesichts des allgegenwärtigen Lehrkräftemangels die spezifischen Herausforderungen von peripheren ländlichen Regionen: Vorgestellt wird eine Studie, in der Herkunftsregionen und Präferenzen für Regionen der zukünftigen Berufstätigkeit unter Studienanfänger:innen im Lehramtsstudium in den Blick genommen werden. Dabei zeigt sich ein Zusammenhang zwischen der Vertrautheit mit ländlichen Räumen und einer positiven Einstellung gegenüber einer möglichen späteren Berufstätigkeit dort. Das Symposium schließt mit einer Diskussion der Beiträge.

Beiträge des Symposiums

Regionale Mobilität von (angehenden) Lehrkräften. Bundeslandwechsel zwischen Schulabschluss, Lehramtsstudium, Referendariat und erste Erwerbstätigkeit als Lehrkraft

Sebastian Franz¹, Seyma Gülen², Svenja Mareike Schmid-Kühn³

¹Universität Bamberg, ²Universität Tübingen, Hector-Institut für Empirische Bildungsforschung, ³Universität Koblenz

Einleitung, theoretischer Hintergrund und Forschungsstand

Regionale Mobilität von Lehramtsstudierenden ist entscheidend für die Verteilung qualifizierter Lehrkräfte in den deutschen Bundesländern. Obwohl die Kultusministerkonferenz (KMK, 2013) beschlossen hat, dass Lehrkräfte ihre Qualifikationen deutschlandweit anerkennen lassen können, bleibt der Wechsel zwischen Bundesländern oft schwierig. Unterschiede in den Schulsystemen der Bundesländer erscheinen als Hemmnisse. Einige Bundesländer könnten versuchen, Lehramtsabsolvent*innen aus anderen Bundesländern anzuwerben, um ihren Lehrkräftemangel zu mindern (Füssel, 2011). Die Untersuchung der Mobilität ist daher wesentlich, um die Flexibilität und Anpassungsfähigkeit des Lehrkräftebedarfs zu verstehen.

Regionale Mobilität umfasst geographische Bewegungen zwischen Bundesländern, einschließlich Umzug zu Studien- oder Ausbildungszwecken sowie Wechsel des Beschäftigungsortes. Diese Studie konzentriert sich auf Mobilitätsverhalten auf Bundesländerebene.

Bisherige Forschung zur Bundeslandmobilität von Lehramtsstudierenden ist begrenzt und fokussiert sich meist auf individuelle Hochschulen oder Bundesländer (Eulenberger, 2018). Erste Befunde deuten auf eine begrenzte Mobilität, bedingt durch soziale und familiäre Bindungen, hin (Eulenberger et al., 2015; Foerster, 2008). Beispielsweise stammen 72,7 % der Lehramtsstudierenden an der Universität Bamberg aus Bayern und 74,2 % der sächsischen AbsolventInnen bewerben sich für einen Referendariatsplatz im selben Bundesland.

Forschungsfrage

Ziel der Studie ist die Untersuchung regionaler Mobilitätsmuster von Lehramtsstudierenden in Deutschland. Die spezifischen Fragen sind:

1. Wie viele Personen wechseln das Bundesland zwischen Hochschulzugangsberechtigung (HZB), Lehramtsstudium, Referendariat und Erwerbstätigkeit als Lehrkraft?
2. Bei welchen Übergängen wechseln Personen die Bundesländer und wohin?
3. Wie hoch ist der Anteil derer, die in ihr Heimatbundesland zurückkehren?

Daten und Methode

Zur Beantwortung der Fragen werden Daten des Nationalen Bildungspanels (NEPS) und des Lehramtsstudierenden-Panels verwendet (Blossfeld & Roßbach, 2019; NEPS, 2023; Schaeper et al., 2023). Die Daten umfassen Studierende, die erstmals im Wintersemester 2010/11 ein Studium aufgenommen haben, sowie alle Lehramtstypen und -fächer. Die Stichprobe umfasst 1628 Personen mit Informationen zu HZB, Studienabschluss, Referendariat und erster Erwerbstätigkeit.

Der methodische Ansatz umfasst ausschließlich deskriptive Auszählungen und graphische Visualisierungen mittels Sankeyplots. Da keine spezifischen Auswertungen pro Bundesland erlaubt sind, wurde lediglich der Wechsel zwischen Bundesländern analysiert.

Erste Ergebnisse

Die Analysen zeigen, dass 68 % der Stichprobe zwischen HZB, Lehramtsstudium, Referendariat und Lehrkraftberuf nicht das Bundesland gewechselt haben. Wenn ein Bundesland-Wechsel erfolgte, dann meist zum Studieren. Zwischen Lehramtsabschluss und Referendariat (67 %) sowie zwischen Referendariat und Erwerbstätigkeit (64 %) überwiegend in ein benachbartes Bundesland gewechselt wird. Ein Studium führt häufig zu einem Wechsel in ein weiter entferntes Bundesland (82 %). Nur ein geringer Anteil kehrt nach dem Studieren oder Referendariat in das Heimatbundesland zurück.

Weitere Analysen sollen Lehramtstyp und Unterrichtsfach einbeziehen, um tiefere Einblicke in die Mobilitätsbedingungen zu gewinnen. Die Ergebnisse werden in Bezug auf ihren Einfluss auf den bundeslandspezifischen Lehrkräftemangel diskutiert.

Wer bleibt, wer geht? Determinanten regionaler Mobilität von (angehenden) Lehrkräften beim Übergang in die Lehrkräfteausbildung und in den Beruf

Sevma Gülen¹, Sebastian Franz², Svenja Mareike Schmid-Kühn³

¹Universität Tübingen, ²Hector-Institut für Empirische Bildungsforschung, ³Universität Bamberg, ³Universität Koblenz

Der Mobilitätsbeschluss der Kultusministerkonferenz, welcher einen gleichberechtigten länderübergreifenden Zugang von (angehenden) Lehrkräften in den Vorbereitungsdienst sowie in den Beruf ermöglichen soll, liegt inzwischen mehr als ein Jahrzehnt zurück (KMK, 2013). Bislang liegt allerdings nach wie vor weder über das Ausmaß der Lehrkräftemobilität innerhalb der Bundesrepublik noch über deren Determinanten hinreichende empirische Evidenz vor. Vorliegende einzelne und eher lokal ausgerichtete Studien zeigen, dass Lehrkräfte im Vergleich zu anderen Berufsgruppen eine geringere Mobilitätsrate aufweisen (Alesi & Neumeyer, 2017; Haussen & Uebelmesser, 2018). In welchem Ausmaß (angehende) Lehrkräfte auf bundesweiter Ebene mobil sind und aus welchen Gründen, bleibt dabei allerdings empirisch ungeklärt.

Vor dem Hintergrund des aktuellen und seit Jahren fortwährenden Lehrkräftemangels sind diese Fragestellungen mehr denn je von wissenschaftlicher und politischer Bedeutung. Schließlich sind Erkenntnisse über das Mobilitätsverhalten von (angehenden) Lehrkräften auf unterschiedlichen Ebenen von Relevanz, wie beispielsweise auf Ebene der Steuerung von Bildungssystemen oder für die Schulen selbst (OECD, 2006). Hinzu kommt der andauernde Bildungsmisserfolg von Schüler*innen in der Bundesrepublik (Lewalter, 2023), wobei die Situation durch den Lehrkräftemangel und die schlechte Planbarkeit und Steuerung von Bildungssystemen aufgrund fehlender wissenschaftlicher Befunde weiter verschärft wird.

Basierend auf diesem Forschungsdesiderat geht der Beitrag der Forschungsfrage nach, welche Determinanten die regionale Mobilität von (angehenden) Lehrkräften beim Übergang in die Hochschule, in den Vorbereitungsdienst sowie in den Beruf erklären. Hierbei stehen sowohl individuelle Merkmale als auch Merkmale des Bildungsverlaufes, wie beispielsweise Abschlussnoten, im Fokus des Interesses. Das Ziel des Beitrages ist es, Befunde hinsichtlich spezifischer Merkmale von (angehenden) Lehrkräften zu identifizieren, welche eine Mobilität begünstigen bzw. hindern. Basierend auf theoretischen Annahmen der Rational-Choice-Theorie sowie der Lebenslauforschung werden individuelle Merkmale, vorherige Mobilitätsverfahren sowie strukturelle Rahmenbedingungen im Kontext der Mobilitätsentscheidung berücksichtigt. Die Analysen basieren auf der Datengrundlage der Studierendenkohorte (SC 5) des Nationalen Bildungspanels (NEPS), die einen längsschnittlichen Einblick in den Bildungs- und Berufsverlauf von Lehramtsstudierenden erlaubt und somit ermöglicht, Mobilitätsprozesse über die gesamte Ausbildungs- und Berufseinstiegsphase zu untersuchen (Erhebungszeitraum von 2010 bis 2021). Basierend auf Methoden der Inferenzstatistik (Hypothesentests und multinomiale logistische Regression) zeigen die Ergebnisse, dass grundsätzlich die männlichen Lehramtsstudierenden bzw. Lehrkräfte eine größere Mobilitätsbereitschaft aufweisen als ihre weibliche Vergleichsgruppe. Gleiches gilt für (angehende) Lehrkräfte aus akademisch geprägten Elternhäusern. Mit bis zu vier Prozentpunkten Unterschied sind sie im Vergleich zu denjenigen aus bildungsfernen Elternhäusern signifikant häufiger über die gesamte Bildungs- und Berufslaufbahn hinweg mobil. Der Migrationshintergrund ist nur beim Eintritt in die Hochschule von Bedeutung, wobei an diesem Übergang Studienberechtigte mit Migrationshintergrund um zwei Prozentpunkte signifikant seltener für den Studieneintritt das Bundesland der Erlangung ihrer Hochschulzugangsberechtigung wechseln. Bei den Übergängen von Studium in das Referendariat bzw. in den Beruf unterscheiden sich diejenigen (angehenden) Lehrkräfte mit und ohne Migrationshintergrund hingegen nicht mehr voneinander. Hinsichtlich der Unterschiede im Mobilitätsverhalten nach Abschlussnoten zeigen die Befunde, dass eine bessere Hochschulabschlussnote von Lehramtsabsolvent*innen die Mobilität beim Übergang in das Referendariat positiv beeinflusst. Demgegenüber sind Lehramtsabsolvent*innen mit einer schlechteren Note der Hochschulzugangsberechtigung sowie einer schlechteren Note des 2. Staatsexamens beim Übergang zwischen Referendariat und Beruf signifikant häufiger mobil. In einer Gesamtschau zeigen die Befunde darüber hinaus, dass zunächst grundsätzlich die Mobilität mit jedem nächsten institutionellen Übergang abnimmt. Darüber hinaus nimmt mit jedem Übergang die soziale Selektivität zu, was dazu führt, dass in Folge dessen individuellen Merkmalen als Erklärungsmechanismus weniger eine Bedeutung zukommt als der jeweiligen Abschlussnote.

Herkunftsregion und regionale Präferenzen für die spätere Berufstätigkeit bei Studienanfänger:innen im Lehramtsstudium

Anja Vatterrott, Yvonne Driesner
Universität Rostock

Der Lehrkräfteersatzbedarf differenziert sich nicht nur nach Lehramtstypen und Fächern, sondern innerhalb dieser auch noch einmal in Bezug auf Regionen, spezifisch in Bezug auf städtische und ländliche Räume. In peripheren ländlichen Regionen ist die Situation in Bezug auf den Lehrkräftemangel besonders herausfordernd. Es gibt verschiedene Strategien, dieser Herausforderung zu begegnen, die sich auf Anwerbung und Bindung ausrichten. Anwerbung kann neben finanziellen Anreizen (KMK 2022) darin bestehen, Lehrkräfte aus den betroffenen Regionen zu rekrutieren („grow your own“) (McClure & Reeves 2004). In Bezugnahme auf den „grow your own“-Ansatz untersucht der Beitrag Herkunftsregionen und Präferenzen für Regionen

der zukünftigen Berufstätigkeit unter Studienanfänger:innen im Lehramtsstudium. Dabei werden Bundesländer und Stadt- und Gemeindetypen sowie die allgemeine Einstellung zur Lehrtätigkeit im ländlichen Raum betrachtet.

Datengrundlage sind standardisierte Fragebogenerhebungen unter Studienanfänger:innen im Lehramt an der Universität Rostock in den zentralen Einführungsveranstaltungen in den Jahren 2019-2023. Die Lehramtsstudierenden sind bei Aufnahme des Studiums im Vergleich mit der Gesamtheit der Studienanfänger:innen weniger mobil, etwa zwei Drittel eines Jahrgangs haben die Hochschulzugangsberechtigung im Bundesland des Studienbeginns erworben. Für jeweils über 80% eines Jahrgangs kommt das Bundesland, in dem sie studieren, auch für die spätere Tätigkeit als Lehrer:in in Frage. Der Anteil ist jedoch größer bei diejenigen, die auch schon die Hochschulzugangsberechtigung im selben Bundesland erworben haben.

Die Einstellung zu einer späteren Berufstätigkeit im ländlichen Raum steht mit dem Stadt- und Gemeindetyp, in dem die Befragten bisher die meiste Zeit ihres Lebens verbracht haben, in Zusammenhang, d.h. je kleiner die Stadt bzw. Gemeinde, in der sie bisher die meiste Zeit gelebt haben, desto größer ist die Wahrscheinlichkeit, dass die Befragten angeben, sich gut vorstellen zu können, später im ländlichen Raum berufstätig zu sein. Für die Einstellung zur Tätigkeit im ländlichen Raum spielt es keine eindeutige Rolle, ob die Studienanfänger:innen im Bundesland des Erwerbs ihrer Hochschulzugangsberechtigung studieren oder nicht. Sie steht jedoch in Zusammenhang mit dem studierten Lehramt. Die Einstellung zur Berufstätigkeit im ländlichen Raum ist in der Regressionsanalyse bei Kontrolle für die Größe der Herkunftsgemeinde und das Bundesland des Erwerbs der Hochschulzugangsberechtigung mit größerer Wahrscheinlichkeit positiv bei Studierenden im Lehramt an Grundschulen im Vergleich mit den anderen Lehramtstypen.

Da die Befragten noch am Anfang des Studiums stehen und es sich lediglich um die Erfassung von Präferenzen handelt, müssen die Zusammenhänge in Hinsicht auf zukünftiges Handeln mit Vorsicht interpretiert werden. Die Ergebnisse scheinen jedoch einen Zusammenhang zwischen Vertrautheit mit ländlichen Räumen und einer positiven Einstellung gegenüber einer Tätigkeit dort zu unterstreichen, was den „grow your own“-Ansatz unterstützen würde. Diese Präferenzen könnten durch Praxisphasen während des Studiums, die im ländlichen Raum absolviert werden, gestärkt bzw. über das Studium aufrechterhalten werden, um die Wahrscheinlichkeit einer tatsächlichen späteren Berufstätigkeit als Lehrkraft im ländlichen Raum zu erhöhen.

S-18: Selbstgesteuertes Lernen in digitalen Lernumgebungen – Erfassung und Analyse der Kompetenzen von Studierenden im Umgang mit Online-Informationen

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: SO 108

Chair der Sitzung: Olga Troitschanskaia

Chair der Sitzung: Johannes Hartig

Chair der Sitzung: Samuel Greiff

Selbstgesteuertes Lernen in digitalen Lernumgebungen – Erfassung und Analyse der Kompetenzen von Studierenden im Umgang mit Online-Informationen

Chair(s): **Olga Zlatkin-Troitschanskaia** (Johannes Gutenberg-Universität), **Johannes Hartig** (DIPF, Frankfurt a.M.)

Diskutant*in(nen): **Samuel Greiff** (Goethe Universität Frankfurt)

Studierende weltweit nutzen zunehmend Online-Informationen zum Lernen in ihrem Studium. Das selbstgesteuerte Internet-basierte Lernen ist aufgrund der hohen Anzahl von algorithmisch vorselektierten, oft unverifizierten, didaktisch kaum aufbereiteten und teilweise inkorrekten Quellen voraussetzungsreich [37]. Während die Forschung zum Lernen in moderierten digitalen Settings zunimmt, ist das selbstgesteuerte Lernen von Studierenden mit Internet noch wenig erforscht [28]. In diesem Symposium werden aktuelle Forschungsarbeiten einer interdisziplinären DFG-geförderten Forschungsgruppe zur Erfassung der Fähigkeiten von Studierenden zum kritischen Umgang mit Online-Informationen - definiert als *Critical Online Reasoning (COR)* – präsentiert. Dabei werden zum einen die Informationsverarbeitungsprozesse und deren Ergebnisse bei Lösung von Szenario-basierten COR-Aufgaben zu freier Internetrecherche erfasst. Zum anderen werden die Online-Quellen und deren Inhalte analysiert, die Studierenden bei Ihrer freien Recherche zur COR-Aufgabenlösung nutzen.

Da Studierende in ihrem Studium mit domänenspezifischen sowie mit domänenübergreifenden Anforderungen konfrontiert werden, fokussiert die Forschungsgruppe beide Kontexte: generisches (GEN)-COR und domänenspezifisches (DOM)-COR. Der erste Beitrag präsentiert das neu entwickelte Ratings-Schema zum Scoring der studentischen GEN- und DOM-COR-Aufgabenlösungen (kurze Stellungnahmen) und diskutiert Ergebnisse aus einer Validierungsstudie mit mehreren unabhängigen trainierten Ratern.

Der Internet-Rechercheprozess während der COR-Aufgabenbearbeitung wird in Echtzeit erfasst und aufgezeichnet. Die Daten umfassen die von Studierenden zur Aufgabenlösung verwendeten Quellen. Der zweite Beitrag präsentiert die Ergebnisse einer systematischen Inhaltsanalyse der meistgenutzten Quellen hinsichtlich zentraler Qualitätskriterien wie Vollständigkeit, Richtigkeit etc. und liefert damit einen unikalen Einblick in die Qualität von Internet-Informationen, die von Studierenden zum Lernen tatsächlich genutzt werden.

Die Interaktion von Lernenden mit der Internet-Informationslandschaft ist ein weiterer Analysefokus der Forschungsgruppe und wird mittels quantitativer Verfahren wie Eye-tracking mit Think-Alouds sowie innovativer Verzahnung der qualitativen hermeneutisch rekonstruktiven und narrativen Methoden untersucht. Der dritte Beitrag zeigt, inwieweit Studierende narrativen Framings und latente Sinnstrukturen in den von Ihnen genutzten Internet-Quellen erkennen und in ihrer COR-Aufgabenantworten unkritisch übernehmen, umdeuten oder kritisch reflektieren.

Um einen tieferen Einblick in die COR-Aufgabenlösungsprozesse der Studierenden (z.B. mit hoher und niedriger Performanz) bei der Nutzung der Online-Quellen zu erhalten, werden Augenbewegungen und retrospektive Think-Aloud-Protokolle von Studierenden während der GEN- und DOM-COR-Aufgabenbearbeitung aufgezeichnet. Der vierte Beitrag gibt erste Einblicke in die unterschiedlichen (z.B. heuristische vs. systematische) Informationsverarbeitungsprozesse der Studierenden bei Lösung von COR-Aufgaben.

Beiträge des Symposiums

Interaktive Testverfahren zur Bewertung der kritischen Informationsnutzung von Hochschulstudierenden im digitalen Zeitalter

Philine Drake¹, Johannes Hartig¹, Jannick Illmann¹, Carmen Köhler¹, Marie-Theres Nagel², Anika Kohmer³, Alisa Stroganova⁴

¹DIPF, Frankfurt a.M., ²Johannes Gutenberg-Universität Mainz, ³Goethe-Universität, Frankfurt a.M., ⁴Ludwig-Maximilians-Universität, München

Theoretischer Hintergrund

Hochschulstudierende nutzen das Internet als wesentliche Informationsquelle für studienrelevante Themen [11;38] und stehen vor der Herausforderung, zuverlässige Informationen zu finden, zu bewerten und zu integrieren. Diese Herausforderungen erfordern spezifische Fähigkeiten, die als „Critical Online Reasoning“ (COR) bezeichnet werden und drei Facetten umfassen: Online Information Acquisition (OIA), Critical Information Evaluation (CIE) und Reasoning based on Evidence, Argumentation, and Synthesis (REAS)[25]. OIA bezieht sich auf die effektive Suche und vorläufige Bewertung der Suchergebnisse, CIE auf die tiefere Bewertung von Quellen und deren Vertrauenswürdigkeit sowie die Anwendung von Überprüfungsstrategien. REAS umfasst die Synthese von Informationen und die Bildung evidenzbasierter Argumente [25]. COR-Fähigkeiten lassen sich nach ihrem Anwendungsbereich in allgemeinen (GEN-COR) und domänenspezifischen (DOM-COR) Kontexten unterteilen [38].

Fragestellung

Die Forschung der letzten Jahre zeigt, dass eine authentische Aufgabenumgebung mit hoher Interaktivität erforderlich ist, um COR valide zu erfassen [25;26;27;33]. Daher haben wir Aufgaben entwickelt, die im offenen Informationsraum des Internets von Studierenden bearbeitet werden, der als natürliche Umgebung für COR dient. Diese Aufgaben bewerten die COR-Leistung anhand offener Antworten (z. B. schriftliche Argumentation und Begründung für die Auswahl von Quellen) sowie anhand des aufgezeichneten Verhaltens während der Aufgabenbearbeitung (z. B. Logdaten, die die verwendeten Suchbegriffe und Quellen enthalten). Der aktuelle Beitrag untersucht, wie sich COR in einem natürlichen Handlungsraum erfassen lässt und welche Indikatoren die drei COR-Facetten abbilden. Der Fokus liegt hierbei auf Aufgaben zur Erfassung von GEN-COR, die kein spezifisches Fachwissen erfordern.

Methode

Vier GEN-COR-Aufgaben wurden entwickelt und von einer Stichprobe von N=471 Hochschulstudierenden aus vier verschiedenen Disziplinen (Medizin, Physik, Soziologie, Ökonomie) zum ersten von vier Testzeitpunkten bearbeitet. Diese Aufgaben behandeln allgemeine bzw. alltägliche Fragestellungen, die von öffentlichem Interesse sind und die politisch, journalistisch und/oder wissenschaftlich behandelt wurden. Beispielsweise sollten die Studierenden Fragen zum Recycling bestimmter Abfallprodukte oder zu möglichen Gesundheitsrisiken aufgrund von Wohnortbedingungen recherchieren und ein kritisches Statement mit Pro- und Contra-Argumenten auf Basis der im Internet gefundenen Evidenzen verfassen. Zudem sollten sie die verwendeten Quellen hinsichtlich ihrer Qualität (z.B. Vertrauenswürdigkeit) beurteilen. Alle vier GEN-COR-Aufgaben enthalten Prompts, um alle drei COR-Facetten durch verschiedene Kriterien abzubilden. Während der Internetrecherche können bereits Indikatoren der OIA-Facette erfasst werden, wie z.B. die Qualität der allerersten Suchanfrage der Studierenden, mit der sie sich i.d.R. einen ersten Überblick über den Themenbereich verschaffen, in dem die Problemstellung angesiedelt ist. Die Angabe und Bewertung der verwendeten Quellen liefert Indikatoren auf die Fähigkeit zur Quellenbewertung (CIE). Im verfassten Statement wird sichtbar, inwieweit die Studierenden unterschiedliche Argumente in einer kohärenten und inhaltlich vollständigen Stellungnahme zusammenführen und belegen können (REAS).

Ergebnisse

In unserem Beitrag präsentieren wir die Entwicklung von Kriterien zur Beurteilung der GEN-COR-Leistung, die eine menschliche Bewertung erfordern, sowie erste empirische Ergebnisse zu diesen Kriterien. Hierzu zählen Ergebnisse zur Interrater-Reliabilität bei der Leistungsbeurteilung sowie zu Skalen- und Itemeigenschaften. Zudem werden Ergebnisse zur Überprüfung der theoretischen COR-Dreifacettenstruktur mittels konfirmatorischer multidimensionaler IRT-Analysen vorgestellt. Diese Analysen basieren auf den Indikatoren zur Qualität der Suchprozesse (OIA), der Bewertung von Online-Quellen (CIE) und der abschließenden argumentativen Reflexion und Integration der Evidenzen (REAS). Wir erwarten, dass sich die GEN-COR-Leistung nach diesen drei Facetten differenzieren lässt.

Diskussion

Bezogen auf die Fragestellungen des Projekts wird diskutiert, inwieweit die theoretische Trennung der drei COR-Facetten anhand der erfassten und bewerteten Indikatoren über die vier GEN-COR-Aufgaben hinweg empirisch gestützt wird. Zudem wird kritisch betrachtet, inwieweit sich die Nutzung des freien Internets in einem diagnostischen Verfahren zur Erfassung von COR bewährt hat und welche Implikationen sich aus den Ergebnissen für die Konstruktion von COR-Aufgaben in simulierten Handlungsräumen ergeben.

Unvollständig, inkorrekt und unausgewogen? Die Qualität der von Studierenden zur COR-Aufgabenlösung genutzten Online-Inhalte

Tobias Scherer, Alice Laufer, Marcus Maurer, Christian Schemer
Johannes Gutenberg-Universität Mainz

Theoretischer Hintergrund

Onlinemedien spielen eine bemerkenswerte Rolle bei der Informationsbeschaffung und beim Lernen von Studierenden [3;35]. Demnach nutzen Studierende sowohl im Alltag als auch zur Prüfungsvorbereitung Online-Nachrichtenmedien, soziale Medien und andere Webseiten mindestens in gleichem Maße wie universitäre Lernmaterialien [35;3]. Diese Onlinemedien unterliegen oft keiner wissenschaftlichen Qualitätskontrolle, sodass sie auch unvollständige, inkorrekte und unausgewogene Informationen enthalten [9;19], die den Lernerfolg gefährden könnten [26]. Mindestens ebenso bedeutsam wie die Frage, aus welchen Quellen sich Studierende informieren, ist folglich die Frage, welche Inhalte von welcher Qualität sie dort vorfinden.

Fragestellung

Bisherige Studien zur Informationsqualität der von Studierenden genutzten Webseiten haben sich i.d.R. entweder darauf beschränkt, diese exemplarisch von Nutzenden bewerten zu lassen, oder sie in ein einfaches dichotomes "gut/schlecht"-Schema einzuteilen [39;8]. Elaboriertere Studien haben die Informationsquellen anhand einer ordinalen Klassifikation in Gruppen eingeteilt [26]. Demnach weisen wissenschaftliche Publikationen beispielsweise eine höhere Informationsqualität auf als die Webseiten von Nachrichtenmedien, diese wiederum weisen eine höhere Qualität auf als Informationen aus sozialen Medien usw. [26]. Obwohl es eine hilfreiche Heuristik ist, beantwortet sie die Frage nach der Qualität der genutzten Inhalte nicht hinreichend – vor allem deshalb, weil auch innerhalb der Gruppen von Informationsquellen erhebliche Qualitätsunterschiede bestehen [21]. So macht es für die genutzte Informationsqualität einen Unterschied, ob sich ein Studierender ein wissenschaftliches YouTube-Tutorial zum Klimawandel ansieht oder einen YouTube-Beitrag der AfD zum selben Thema. Zudem zeigen viele Studien z.T. erhebliche Unterschiede in der Informationsqualität von Nachrichtenmedien [22]. Dies gilt für die politische Berichterstattung insgesamt [22;34], sowie auch für die Berichterstattung zu medizinischen [30] und wirtschaftswissenschaftlichen [9] Themen. In diesem Beitrag wird deshalb erstmals die Qualität der von Studierenden genutzten Informationsquellen bei Lösung von Aufgaben zum Critical Online Reasoning (COR) (25) detailliert untersucht.

Methode

Die Studie greift auf Daten von 473 Studierenden aus vier Disziplinen (Medizin/Physik/Ökonomie/Soziologie), welche jeweils drei Aufgaben aus einem Tool von 28 Aufgaben zum kritischen Umgang mit Online-Informationen aus den Bereichen Allgemeinwissen, Medizin oder Wirtschaftswissenschaft im WS 2023/24 bearbeitet haben. Innerhalb der einzelnen COR-Aufgaben wurden die Probanden jeweils mit einer generischen oder fachspezifischen Fragestellung konfrontiert, welche sie in einem vorgegebenen Zeitfenster mittels einer Internetrecherche beantworten sollten. Die genutzten Webseiten dienten den Studierenden als Informationsgrundlage zur Beantwortung der Fragestellungen und Verfassen von Stellungnahmen. Das Online-Suchverhalten der Studierenden wurde getrackt und alle Logdaten aufgezeichnet.

Basierend auf diesen Logdaten mit URL-links untersuchen wir die Inhalte der zur Bearbeitung dieser COR-Aufgaben genutzten Webseiten mit der in der Kommunikationswissenschaft etablierten Methode der quantitativen Inhaltsanalyse [31]. Durch diese Methode werden große Textmengen mithilfe eines standardisierten Messinstruments (Codebuch) intersubjektiv prüfbar analysiert [10]. Dabei knüpfen wir an ein etabliertes Framework zur Analyse von Informationsqualität an [31;34], aus dem wir Qualitätsdimensionen wie Vollständigkeit, Richtigkeit und Ausgewogenheit abgeleitet und für jede genutzte Webseite auf Beitragsebene gemessen haben. Die Analyse wurde von sieben geschulten Codierern manuell durchgeführt. Die Inter-coder-Reliabilitätswerte als Maß für die Übereinstimmung der Codierer untereinander betrugen nach Holsti [13] zwischen 0.802 und 1.

Ergebnisse

Die 473 Studierenden haben zur Lösung der 28 COR-Aufgaben insgesamt über 2.300 verschiedene Webseiten genutzt – Suchmaschinen-treffer und Login-Seiten wurden aus dieser Angabe bereits herausgefiltert. Darunter waren wissenschaftliche

Publikationen und Datenbanken, aber auch Beiträge von Nachrichtenmedien, Webseiten von Unternehmen, Verbänden und Interessengruppen sowie User Generated Content. Im Vortrag werden die Analyseergebnisse zur Qualität der genutzten Webseiten für die unterschiedlichen o.g. Dimensionen präsentiert. Dabei wird zum einen die Informationsqualität in unterschiedlichen Quellen (Wissenschaft, Nachrichtenmedien, User Generated Content usw.) verglichen. Zum anderen wird gezeigt, wie die Qualität der genutzten Informationen zwischen den generischen und fachspezifischen COR-Aufgaben sowie den Studienfächern der Probanden wie Medizin vs. Ökonomie variiert.

Diskussion

Im Ergebnis wird ein differenziertes Bild der von den Studierenden genutzten Informationslandschaft und ihrer Qualität präsentiert und hinsichtlich Implikationen für die Hochschulbildung kritisch diskutiert.

Sprachgebrauch und fachliche Kontextualisierungen beim Aufgabenlösen in Internet-lernumgebungen – Critical Online Reasoning in studentischen Bildungsgängen der Ökonomie und Medizin

Dominik Braunheim, Carla Schelle, Mita Banerjee, Amina Touzos, Olga Zlatkin-Troitschanskaia
Johannes Gutenberg-Universität Mainz

Theoretischer Hintergrund

Die zunehmende Nutzung digitaler Medien in Lernprozessen stellt Studierende vor Herausforderungen wie die eigenverantwortliche Suche und Evaluation von Internetquellen. In einer DFG-geförderten Forschungsgruppe analysieren wir die Navigations- und Entscheidungsprozesse von Studierenden in einer situativen Internet-Lernumgebung, d.h. deren Verarbeitung gefundener Internetquellen im Lösungsprozess von Performance Aufgaben. Die methodenintegrativen qualitativen Analysen fokussieren auf narrative Framings [2]genutzter Website-Inhalte sowie auf latente Sinngehalte und normative Setzungen in schriftlichen Stellungnahmen Studierender zu generischen und domänenspezifischen Aufgabenstellungen. Untersucht wird dabei, inwiefern Studierende Framings, latente Sinngehalte u.a. innerhalb der Internetquellen identifizieren, beim Aufgabenlösen aufgreifen und reflektieren. Wir folgen dabei der These [4], dass Studierende als soziale Akteure mit eigener soziokultureller Vorgeschichte in die sozialstrukturierte Lebenswelt [24] eingebunden sind und dass sowohl die inkorporierte Sprache als auch Vorerfahrungen deren Lernprozesse (im Internet) konstituieren. Weiterhin folgenden wir der These, dass Herkunft, Sprache und Umwelt (Enkulturation) als habitualisierte Muster Einfluss auf das studentische Lernverhalten [32] und deren individuelle Aufgabenlöseprozesse wie Rechercheverhalten, Quellenselektion und eigener argumentativer Informationsverarbeitung im Internet haben. Diese drei Facetten der Informationsverarbeitung konstituieren das (meta)kognitive Kompetenzkonstrukt Critical Online Reasoning (COR)[25], das hinsichtlich weiterer Merkmale wie Erkennen von Bias, kritische Positionierung und Urteilsbildung ausdifferenziert wird.

Fragestellung

In diesem Beitrag untersuchen wir, inwieweit Studierende der Fächer Ökonomie und Medizin narrative Rahmungen und latente Sinnstrukturen in Website-Texten erkennen und in ihrer Aufgabenlösung (generisch zum Thema Hitzestift sowie Funkmasten) verwerfen, aufgreifen, umdeuten, unkritisch und/oder kritisch reflektieren. Den o.g. Thesen folgend, dass sich in diesem Aufgabenlösungsprozess Vorerfahrungen (enkluturiert und inkorporiert) niederschlagen und als habitualisierte Muster manifestieren können, untersuchen wir:

In einem intraindividuellen Vergleich: Zeigen sich bei Studierenden identische Entscheidungs- und Aufgabelöseprozesse bzw. Strategien usf. bei verschiedenen COR-Aufgaben oder nicht?

In einem interindividuellen Vergleich: Zeigen sich bei Studierenden unterschiedlicher Fächer (Medizin, Ökonomie) verschiedene Entscheidungs- und Aufgabelöseprozesse bzw. Strategien bei gleichen COR-Aufgaben?

Methode

Da die Rahmung von Quellen aus dem Internet als Narrative potenziell zu Verzerrungen der Informationsverarbeitung führen kann(3) werden die bei der COR-Aufgabenlösung genutzten Inhalte mit der Methode der Narration auf sprachliche Mittel und stilistische Merkmale hin untersucht.

Die schriftlichen Stellungnahmen der Studierenden als COR-Aufgabenlösungen werden mit der Objektiven Hermeneutik sequentiell rekonstruiert, davon ausgehend, dass soziales Handeln (hier die Internet-Recherche, Auswahl und Verarbeitung von Quellen) sich regelgeleitet vollzieht und stets auch ein latenter Sinn innewohnt(5).

Anschließend werden sämtliche im Aufgabenlösungsprozess entstandenen multimodalen Daten (Logdaten, Selbstreflexionsfragebögen) für die wechselseitig aufeinander bezogenen narrativen und rekonstruktiven Analysen herangezogen.

Ergebnisse

Die interindividuellen Analysen zu unterschiedlichen Aufgabenbearbeitungen und den zugehörigen Quellen von Studienanfänger:innen der Ökonomie und Medizin zeigen in deren Stellungnahmen zu einer generischen COR-Aufgabe zur Verwendung von Hitzestiften bei Mückenstichen domänenspezifische Zugänge und Kontextualisierungen (Ökonomie: materiell, marketingnahe Argumentation für ein bestimmtes Produkt; Medizin: körperbezogen, diagnostisch, medizinische Begriffe und Behandlung. Die intraindividuellen Analysen zu einer generischen Aufgabe (Funkmasten) bestätigen dies mit leichten Varianten, die kritisch diskutiert werden. Gleichzeitig zeigen sich Schwächen in beiden Stellungnahmen hinsichtlich aller drei COR-Facetten und der jeweilige domänenspezifische Fokus bezieht sich nicht auf den Kern der generischen Fragestellungen.

Die Ergebnisse sprechen für einen fachspezifischen COR-relevanten Studienhabitus, der sich bezogen auf Aufgabenlösungsprozesse bei den Studierenden zeigt. Um diese These weiter zu untersuchen, ist zu rekonstruieren, inwiefern sich über den Studienvorlauf und bezogen auf unterschiedliche Studiengänge habitualisierte Muster der Suche, Nutzung von Quellen, des Sprachgebrauchs usf. erkennen und unterscheiden lassen, worin die Unterscheidungen bestehen und inwiefern sich etwa Sprache/Sprachgebrauch (Begriffe, Fachbegriffe, Subjektpositionierung), die kritische Rezeption von Framings, Normativität im Kontext von COR verändern und entwickeln. Hierzu werden entlang mehrerer Erhebungszeitpunkte COR-Aufgabenlösungsprozesse über den Studienvorlauf erhoben. Akademisches Arbeiten und Denken als Form der Enkulturation im Bildungsprozess/Bildungsgang einzelner Studierender wird dabei weiter untersucht, um Erkenntnisse zu generieren, die dazu beitragen fachspezifische Interventionen zur effektiven Förderung von COR in der universitären Lehre zu entwickeln.

Critical Online Reasoning-Prozesse bei generischen Aufgaben: Einblicke in leistungsstarkes und leistungsschwaches Verhalten mittels Eye-Tracking

Anika Kohmer¹, Maruschka Weber¹, Verena Klose¹, Andreas Maur², Ann-Kathrin Kunz², Yavuz Dinc³, Verena Ruf³, Stefan Küchemann³

¹Goethe-Universität, Frankfurt a.M., ²Johannes Gutenberg Universität, Mainz, ³Ludwig-Maximilians-Universität, München

Theoretischer Hintergrund

Studierende bevorzugen zunehmend Online-Quellen zur Informationsbeschaffung und zum Lernen, da diese einfacher zugänglich sind als traditionelle Medien (z.B. Lehrbücher [11;20]). Allerdings stellen die Fülle und die Fehlerhaftigkeit verfügbarer Online-Informationen eine Herausforderung dar. Das erfolgreiche Lernen mit Online-Quellen erfordert eine reflektierte Recherche, Selektion und Bewertung der Qualität und Glaubwürdigkeit von Online-Informationen [39;7], die als „Critical Online Reasoning“ (COR) definiert werden [25]. COR beinhaltet die effektive Auswahl und Bewertung von Online-Informationen sowie die korrekte Interpretation und argumentative Begründung der gewonnenen Evidenzen [25;27].

In Studien [33;23] wurden Unterschiede in den COR-Aufgabenlösungsprozessen zwischen leistungsstarken und leistungsschwachen Studierenden festgestellt. Leistungsstarke Studierende verbrachten z.B. mehr Zeit auf qualitativ hochwertigen Webseiten und gingen systematischer vor, während leistungsschwächere Studierende häufiger auf weniger vertrauenswürdige Quellen zurückgriffen und unstrukturierter vorgehen [18].

Die Analyse der COR-Aufgabenlösungsprozesse durch die Aufzeichnung der Blickbewegungen bei der Internetsuche ermöglicht ein differenziertes Verständnis solcher Unterschiede [14]. Durch die Beobachtung, auf welche Informationsbereiche innerhalb einer Internetseite die Studierenden ihre Aufmerksamkeit richten, wie lange sie sich in bestimmten Informationsbereichen aufhalten und wie oft sie zwischen verschiedenen Informationsquellen wechseln, können tiefere Einblicke in ihre kognitiven Such- und Bewertungsprozesse gewonnen werden [12] (zu „Eye-Mind-“ und „Immediacy-Hypothese“: 16;15).

Fragestellung

Diese Studie untersucht die COR-Fähigkeiten von Studierenden aus vier Fachrichtungen (Wirtschaftswissenschaften, Medizin, Soziologie, Physik) an drei deutschen Universitäten im ersten Studienjahr. Dabei werden Strategien (weniger) erfolgreicher COR-Aufgabenlösungsprozesse auf häufig besuchten Webseiten verschiedener Kategorien bei den Teilnehmenden der Eye-Tracking-Studie (N=88) identifiziert. Erwartet wird, dass Blickbewegungen, einschließlich der Anzahl und Dauer der Fixationen auf relevanten und irrelevanten Inhaltsbereichen („Areas of Interest“, AOIs) auf Webseiten, sowie die Anzahl der Blicksprünge zwischen AOIs, systematische Suchprozesse und Variationen in der COR-Leistung aufweisen [12;29]. Beispielsweise könnten häufigere und längere Fixationen auf lösungsrelevante Informationsbereiche sowie häufige Blicksprünge zwischen relevanten Abschnitten auf eine gründlichere und kritischere Auseinandersetzung mit dem Inhalt hindeuten [14;1]. Demnach lauten die Hypothesen, dass leistungsstarke Studierende mehr Zeit auf lösungsrelevanten AOIs verbringen als leistungsschwache Studierende (H1) und mehr lösungsrelevante AOIs innerhalb einer Webseite betrachten (H2).

Methode

Die Studierenden wurden gebeten, eine generische COR-Aufgabe zu gesundheitlichen Auswirkungen von Mobilfunkmasten mittels Online-Recherche zu bearbeiten, während ihre Blickbewegungen aufgezeichnet wurden. Zudem wurden deren soziodemografischen Merkmale (z.B. Geschlecht, Alter) sowie weitere Persönlichkeitsmerkmale (z.B. vorherige Interneterfahrung; Schulleistungen) erfasst. Die COR-Leistung wurde durch trainierte Rater anhand eines kategorienbasierten Scoring-Schemas bewertet, und die Gesamtstichprobe in zwei Gruppen „High“ und „Low“-Performers aufgeteilt.

Die fünf meistbesuchten Webseiten aller Teilnehmenden bei der COR-Aufgabenbearbeitung wurden hinsichtlich der lösungsrelevanten AOIs kategorisiert. Anschließend wurde analysiert, wie lange und in welcher Reihenfolge „High“ und „Low“-Performers die AOIs betrachteten und ob sich mittels qualitativer Inhaltsanalyse Auffälligkeiten finden.

Ergebnisse

Wir präsentieren Ergebnisse zu Studiengangs-übergreifenden COR-Prozessen, die bei Studierenden im ersten Studienjahr beim Lösen von COR-Aufgaben identifiziert wurden. Dies beinhaltet Ergebnisse zu Unterschieden in der Nutzung und Bewertung von Online-Informationen sowie in den Suchprozessen [36;17]. Unabhängig von ihrer Fachrichtung und COR-Leistung besuchten alle Teilnehmenden fünf Webseiten am häufigsten. Lediglich die Rangfolge der besuchten Webseiten unterscheidet sich geringfügig.

Die Unterschiede in der COR-Leistung, die auf unterschiedliche visuelle Informationsverarbeitungsprozesse innerhalb der Webseiten zurückzuführen sind, können z.T. durch Eye-Tracking-Metriken aufgezeigt werden. Durch die Analyse von Blickbewegungsdaten der Studierenden geben wir tiefere Einblicke in diese visuellen Lösungsstrategien und die Entscheidungsfindung während der Online-Recherche zur COR-Aufgabenbearbeitung.

Diskussion

Diese Studie zielt darauf ab, das Verständnis über die Informationsverarbeitungsprozesse von leistungsstarken und -schwachen Studierenden bei der Verwendung von Online-Informationen zur Lösung von COR-Aufgaben zu verbessern. Die Ergebnisse geben Einblicke in die Unterschiede in den COR-Aufgabenlösungsstrategien und zeigen auf, wie Blickbewegungsmuster mit der COR-Leistung und Persönlichkeitsmerkmalen der Studierenden zusammenhängen. Diese Erkenntnisse tragen dazu bei, Instruktionsansätze zu entwickeln, die Studierende besser auf die Herausforderungen der Online-Informationsbeschaffung vorbereiten und ihre kritischen Denkfähigkeiten stärken [6;5].

S-19: Wirksamkeit von Interventionen in der Lehrkräftebildung – Curriculare und methodische Innovationen

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: M 003

Chair der Sitzung: Johannes König

Chair der Sitzung: Sarah Strauß

Chair der Sitzung: Dirk Richter

Wirksamkeit von Interventionen in der Lehrkräftebildung – Curriculare und methodische Innovationen

Chair(s): **Johannes König** (Universität zu Köln), **Sarah Strauß** (Universität zu Köln, Deutschland)

Diskutant*in(nen): **Dirk Richter** (Universität Potsdam)

Zahlreiche Bildungsreformen zielen auf die Ausgestaltung der professionellen Lernprozesse von angehenden wie berufstätigen Lehrpersonen, häufig verbunden mit dem Anspruch, die Qualität ihres (zukünftigen) Unterrichts zu verbessern sowie – daraus resultierend – das Lernen ihrer Schülerinnen und Schüler zu fördern (z.B. Liebermann & Pointer Mace, 2008; Hill et al., 2021; Ellis et al., 2023). Entsprechend finden sich curriculare Innovationen in der Lehrkräfteausbildung (z.B. die flächendeckende Implementation von Langzeitpraktika im Studium oder eine Ausweitung curricularer Inhalte zur Vorbereitung angehender Lehrkräfte auf neue Anforderungen der inklusiven Bildung und Erziehung). Sie sind zugleich anzutreffen als neuartige Maßnahmen der Lehrkräftefortbildung (z.B. im Bereich der Digitalisierung an Schulen; vgl. Richter et al., 2024).

Derartigen Innovationen bzw. Maßnahmen in der Lehrkräftebildung wird mitunter eine wichtige vermittelnde Funktion zugeschrieben (vgl. Brüsemeister, 2020; van Ackeren & Klein, 2020): Die Wirkung bildungspolitischer Forderungen und Erwartungen sollen im Zuge einer spezifischen Professionalisierung angehender wie berufstätiger Lehrkräfte zum Tragen kommen, häufig sogar verbunden mit weitreichenden Erwartungen hinsichtlich relevanter Folgewirkungen für das schulische Lehren und Lernen. Vor diesem Hintergrund ist es aus Sicht der empirischen Bildungsforschung entscheidend zu verstehen, was curriculare Innovationen oder neuartige Maßnahmen in der Aus- und Fortbildung von Lehrkräften tatsächlich effektiv macht, da nur dann – gegebenenfalls zeitlich versetzt eintretende – Erfolge oder Misserfolge von Bildungsreformen angemessen bzw. vollständig verstanden werden können (Desimone, 2009). Wissenschaftliche Erkenntnisse zum Gelingen oder Misslingen curricularer Reformen in der Lehrkräftebildung sind relevant, um zukünftig Entscheidungen zur Gestaltung von Lehrkräftebildung zu treffen (Gräsel, 2020).

Vor diesem Hintergrund setzt sich das Symposium drei Ziele: Erstens soll ein Überblick zum Stand der Forschung von Interventionen in der Lehrkräftebildung gegeben werden. Zweitens sollen an ausgewählten Beispielen aktueller Interventionsstudien im Bereich der Aus- und Fortbildung von Lehrkräften Einblicke zu curricularen und forschungsmethodischen Innovationen geliefert werden. Diese sollen, drittens, zur kritischen Diskussion gestellt werden, um Hinweise für zukünftige Forschung abzuleiten. Alle Beiträge sind interdisziplinär angelegt, und im Symposium ist insgesamt eine breite Beteiligung unterschiedlicher Disziplinen (Fachdidaktiken, Erziehungswissenschaft, Psychologie) gegeben.

Im ersten Beitrag werden zentrale Befunde aus einer aktuell erstellten Synthese von 27 Literatur-Reviews zu Interventionen in der Lehrkräftebildung vorgestellt. Die Synthese basiert auf knapp 1000 Interventionsstudien und geht u.a. den Fragen nach, mit welchen theoretischen Ansätzen, Messinstrumenten, Studiendesigns die Forschung zu Interventionen in der Lehrkräftebildung arbeitet und fasst zentrale Ergebnisse der Literaturreviews zusammen (u.a. Meta-Analyse von Effektstärken). Die drei sich daran anschließenden empirischen Beiträge berichten über die Wirksamkeit aktueller, methodisch aufwändiger Interventionen zum inklusiven Mathematikunterricht, zur didaktischen Adaptivität im Biologieunterricht und zur effektiven Klassenführung.

Beiträge des Symposiums

Interventionen in der Lehrkräftebildung: Synthese von 27 Literatur-Reviews

Johannes König, Sandra Heine, Charlotte Kramer, Sarah Strauß, Jonas Weyers, Kristina Gerhard, Janine Mühle, Tim Engartner, Jörg Großschedl, Charlotte Hanisch, Petra Hanke, Thomas Hennemann, Jörg Jost, Kai Kaspar, Benjamin Rott

Universität zu Köln

Theoretischer Hintergrund

Interventionsstudien sind ein wichtiger Bestandteil der Forschung zur Wirksamkeit von Lehrkräftebildung (Hill et al. 2021; König & Blömeke 2020; König et al. 2024). Im Vergleich zu deskriptiven Evaluationsstudien erlauben Interventionsstudien, den Effekt inhaltlicher Neuerungen, zum Beispiel eines innovativen Trainings zur Förderung digitaler Kompetenzen, auf Lernzuwächse bei teilnehmenden Lehrpersonen zu prüfen. Experimentelle Studiendesigns ermöglichen zudem das Ziehen von kausalanalytischen Schlussfolgerungen für Bildungsprozesse (König 2013). Während die Wirksamkeit der Lehrkräftebildung über viele Jahre beklagt wurde, sind über die vergangenen Jahre zunehmend Literatur-Reviews publiziert worden, die zu thematischen Schwerpunkten wie Digitalisierung (z.B. Wilson et al. 2020) oder Inklusion (z.B. Dignath et al. 2022) Interventionsstudien zusammenfassen. Angesichts der Zunahme solcher Reviews einerseits, andererseits der weiterhin bestehenden Fragen zur Wirksamkeit von Lehrkräftebildung, etwa im Zuge des Lehrkräftemangels und alternativer Qualifikationswege in den Lehrkräfteberuf (SWK 2023), benötigt die empirische Bildungsforschung ein Überblickswissen, das über zentrale Befunde solcher Reviews zusammenfassend informiert.

Fragestellung

Anhand einer systematischen Literatursuche wurden im März 2024 insgesamt 27 Literatur-Reviews zu Interventionsstudien in der Lehrkräfteaus-/fortbildung identifiziert. Der Textkorpus wurde mit einem System von 43 Analysekategorien (Anzahl in Klammern) zu fünf Fragestellungen-Schwerpunkten kodiert: I. Rahmenmodelle/theoretische Grundlegung (6); II. Input-/Outcome-Variablen mit Blick auf Anbietende (1), Teilnehmende (4), Interventionsformat (4), Interventionslänge (2), Outcomes (6); III. Studiendesign (11), Messniveau (3); IV. Effektanalyse (3); V. Forschungslücken (3). Der Fokus der Präsentation liegt auf den Fragen I, IV und V, während die Ergebnisse zu II und III bei Bedarf in die Diskussion eingebracht werden:

I. Welche theoretischen Rahmungen werden grundgelegt?

IV. Welche Effekte können aus einer Meta-Meta-Analyse berichtet werden?

V. Welche Forschungslücken lassen sich benennen, die in Zukunft adressiert werden sollen?

Methode

Die systematische Literaturrecherche nutzte die Datenbanken Web of Science und Education Resources Information Center. Verwendet wurden zahlreiche Suchbegriffe, sodass nach Ausschluss von Dubletten und Artikeln ohne peer-review insgesamt 350 Publikationen in Bezug auf Titel, Abstract und Keywords von mindestens zwei Personen der Autor:innen gelesen wurden. Über eine Volltextanalyse und der Verwendung von spezifischen Einschluss-/Ausschlusskriterien wurde die Anzahl der einzubeziehenden Reviews auf 27 eingegrenzt.

Die Kodierung des Textkorpus anhand der Analysekategorien aller Fragen erfolgte durch sechs der beteiligten Autor:innen. Frage IV wird aktuell in Form einer Meta-Meta-Analyse durchgeführt für eine Untergruppe von 14 Reviews, die zusammengefasste Effektstärken ausweisen. Dazu wurden im Rahmen der Kodierung die zusammengefassten Effekte der einzelnen Metaanalysen sowie deren Varianzen extrahiert bzw. ermittelt. Im Falle mehrerer Metaanalysen pro Publikation wurden die unterschiedlichen Zielkonstrukte als separate Metaanalysen berücksichtigt. Die Schätzung des Gesamteffekts erfolgt über ein random effects model.

Ergebnisse und ihre Bedeutung

Die vollständige Auswertung lag zum Zeitpunkt der Abstract-Einreichung noch nicht vor. Erste Datenanalysen zeigen: I. Grundsätzlich legen alle Reviews die Annahme einer Wirkkette zugrunde (häufige Referenz auf Desimone 2009), nämlich dass die Intervention zu einer Förderung der Lehrkraftvariablen (meist Kompetenzfacetten) beiträgt, die sich wiederum positiv in (erhöhter) Unterrichtsqualität und (gesteigerten) Lerngewinnen von Schüler:innen niederschlagen soll. Es fehlen aber in fast allen Reviews eigene theoretische Rahmungen und nicht immer wird der Anlass für Interventionen überhaupt begründet. IV. Erste meta-meta-analytische Befunde weisen eine durchschnittliche Effektgröße von .528 (Standardfehler = .067) aus, die Streuung der Einzeleffekte ist groß, vor allem im Intervall zwischen 0 und 1. Insbesondere Interventionen im Bereich der Digitalisierung fallen durch konsistent höhere Effekte auf. Effekte auf die Lernergebnisse von Schüler:innen der fortgebildeten Lehrkräfte sind demgegenüber eher gering. Weiterführend werden die in den einzelnen Metaanalysen aufgeführten Moderatoreffekte analysiert. Beispielsweise zeigen mehrere Metaanalysen keinen signifikanten Effekt der Fortbildungsdauer (z.B. Brock & Carter 2017; Dignath et al. 2022; Ning et al. 2022), sodass insbesondere die Qualität der Interventionsmaßnahme ins Zentrum rückt. V. Forschungsdesiderata werden insbesondere erkennbar: in der weiteren Theorieentwicklung, in der Berücksichtigung forschungsmethodischer Standards der einzelnen Interventionsstudien sowie in einer methodischen Qualitätssteigerung von Literatur-Reviews (z.B. der Ausweisung von Effektstärken).

Interventionsstudie zu Kompetenzen von (angehenden) Lehrkräften im inklusiven Mathematikunterricht – erste Ergebnisse aus der Längsschnittstudie TEDS-IME

Gabriele Kaiser¹, Conny Melzer², Sarah Strauß³, Natalie Ross¹, Anton Bastian¹, Dennis Sommer¹, Isabelle Klee-Schramm³, Benjamin Rott³, Jonas Weyers³, Nils Buchholtz¹, Johannes König³

¹Universität Hamburg, ²Universität Leipzig, ³Universität zu Köln

Theoretischer Hintergrund

Inklusive Bildung, verstanden als bestmögliche akademische Förderung aller Lernenden (Piezunka et al., 2017), stellt an Lehrkräfte hohe Anforderungen insbesondere in den Bereichen der Diagnose und Förderung von Lernvoraussetzungen ihrer Schüler*innen (König et al., 2019). Dies erfordert von Lehrkräften professionelle Kompetenzen, die im Projekt Teacher Education and Development Study – Inclusive Mathematics Education (TEDS-IME) nach Blömeke et al. (2015) als ein Kontinuum konzeptualisiert werden, das Dispositionen, situationsspezifische Fähigkeiten der Unterrichtswahrnehmung und die unterrichtliche Performanz umfasst.

Interventionsmaßnahmen zur Lehrkräfteprofessionalisierung wird eine hohe Bedeutung für den Erkenntnistransfer aus Unterrichtsforschung in Praxis zugeschrieben (Ball et al., 2008; Lipowsky & Rzejak, 2017). Trotz eines insgesamt umfangreichen Angebots fehlt es im Sekundarstufenbereich jedoch an auf Wirksamkeit überprüften Aus- und Fortbildungsveranstaltungen zur Förderung von Lehrkräftekompetenzen mit Blick auf inklusiven Mathematikunterricht, insbesondere bezogen auf den curricular besonders bedeutsamen Themenbereich Algebra (Malle, 1993). Auch für den Bereich der Kompetenzmessung von (angehenden) Lehrkräften im Hinblick auf einen inklusiven Unterricht kann konstatiert werden, dass bisher nur wenige standardisierte Instrumente vorliegen, die psychometrischen Standards genügen (Weyers et al., 2023).

Die im Rahmen von TEDS-IME entwickelte Intervention richtet sich an berufstätige Lehrkräfte, Lehrkräfte im Vorbereitungsdienst und Studierende im Masterstudium, umfasst 18 Zeitstunden und wurde im Zeitraum von 2023 bis 2024 mit Teilnehmer*innen aller drei Erfahrungsgruppen durchgeführt.

Forschungsfragen

Im Projekt TEDS-IME werden u. a. die folgenden Forschungsfragen untersucht:

Führt die Teilnahme von (angehenden) Mathematiklehrkräften an einer Intervention zum inklusiven Algebraunterricht im Pre-Post-Vergleich zu statistisch signifikanten Zuwächsen im Hinblick auf

- (1) mathematikdidaktisches und pädagogisches Wissen und
- (2) mathematikdidaktische und pädagogische Facetten der professionellen Unterrichtswahrnehmung?

Methode

In einem interdisziplinären Team mit Wissenschaftler*innen aus Mathematikdidaktik, Erziehungswissenschaft und Sonderpädagogik wurden die Intervention und Messinstrumente basierend auf Prinzipien des Constructive Alignments (Biggs & Tang, 2011) entwickelt. In der Entwicklung der Intervention war die Orientierung an evidenzbasierten Wirksamkeitsprinzipien handlungsleitend (Desimone, 2009). In Anlehnung an den Design-Based-Research-Ansatz (Gravemeijer & Cobb, 2006) wurde sowohl die Professionalisierungsmaßnahme selbst als auch zugrundeliegende Unterrichtsmaterialien in mehreren Schritten iterativ angepasst.

Ein quasi-experimentelles Pre-Post-Design der Studie mit Kontrollgruppen (nur für die Unterrichtswahrnehmung) erlaubt Aussagen über Kompetenzzuwächse und ermöglicht Vergleiche zwischen (angehenden) Lehrkräften mit und ohne Intervention. Aufbauend auf bestehenden Instrumenten (Kaiser et al., 2015) wurden hierzu ein Test zu mathematikdidaktischem und pädagogischem Wissen sowie ein innovatives videobasiertes Instrument zur Messung der professionellen Unterrichtswahrnehmung für den inklusiven Algebraunterricht entwickelt und in einem ausführlichen Expert*innen-Review-Verfahren validiert.

Die Datenerhebung in elf Bundesländern und in Kooperation mit Universitäten, Schulen und Instituten für Lehrkräfteaus- und -fortbildung ist abgeschlossen. An der Intervention haben n = 540 Personen teilgenommen (hiervon n = 313 Personen im echten

Längsschnitt). Die Kontrollgruppe umfasst n = 114 Personen (n = 89 im Längsschnitt). Im Mittel waren die Teilnehmer*innen M = 32.3 Jahre alt (SD = 10.4) und der überwiegende Teil ordnete sich dem weiblichen Geschlecht zu (63 %).

Ergebnisse

Zur Analyse der Instrumentengüte wurden je Test zweidimensionale Raschmodelle gerechnet, die jeweils eine mathematikdidaktische (MPCK-IT, M_PID) und eine pädagogische Facette (GPK-IT, P_PID) von professionellem Wissen und Unterrichtswahrnehmung modellieren. Die Modelle weisen zufriedenstellende EAP-Reliabilitäten für MPCK-IT (.76), M_PID (.70), GPK-IT (.68) und P_PID (.70) auf. Die längsschnittlichen Ergebnisse für das inklusionsbezogene pädagogische und das mathematikdidaktische Wissen zeigen signifikante Verbesserungen von kleiner praktischer Bedeutsamkeit in der Interventionsgruppe (MPCK-IT $d=0.36^{***}$; GPK-IT $d=0.25^{***}$). Die längsschnittlichen Ergebnisse für die professionelle Unterrichtswahrnehmung weisen auf signifikante Verbesserungen in der Interventionsgruppe hin (M_PID $d=0.16^{**}$; P_PID $d=0.38^{***}$), die ebenfalls eine kleine praktische Relevanz haben; für die Kontrollgruppe zeigen sich keine signifikanten Veränderungen.

Im Vortrag werden des Weiteren Ergebnisse der verschiedenen Lehrkräftegruppe vergleichend dargestellt. Es werden praktische Implikationen für Lehrkräfteaus- und -fortbildung diskutiert und es wird berichtet, wie adaptierte Konzeptualisierungen und Materialien auf Basis fest etablierter Kooperationen über die Projektdauer hinaus verankert werden können.

TEDS-IME wird gefördert durch das BMBF (Förderkennzeichen 01NV2125A/B).

Theorie-Praxis-Verzahnung durch Forschendes Lernen im Lehramtsstudium: eine Lerngelegenheit zur Verbesserung der Reflexionsfähigkeit

Pauline Sommerer, Nadine Großmann, Johannes König, Jörg Großschedl
Universität zu Köln

Theoretischer Hintergrund

Reflexion bezeichnet das Nachdenken über eigene Erfahrungen, Handlungen und deren Auswirkungen, um zukünftiges Handeln zu verbessern (Lenske & Lohse-Bossenz, 2023; Yost et al., 2000). Reflexionsfähigkeit gilt als notwendige Eigenschaft professioneller Lehrkräfte. Sie ist die Voraussetzung, um reflexive Distanz zur eigenen Unterrichtspraxis aufzubauen, diese kontinuierlich zu innovieren und die eigene Lehrkraftpersönlichkeit weiterzuentwickeln (Neuber & Göbel, 2021). Zur Förderung der Reflexionsfähigkeit wurden in der Mehrheit der deutschen Bundesländer verlängerte Praxisphasen ins Lehramtsstudium integriert (Weyland & Wittmann, 2015). Diese Praxisphasen stellen am Lernort Schule einen geeigneten situativen Rahmen für Forschendes Lernen bereit, in dem theoretische Inhalte aus dem Lernort Universität empirisch erprobt werden können. Dadurch soll Forschendes Lernen zu einer Theorie-Praxis-Verzahnung in der ersten Phase der Lehrkräftebildung beitragen und die Fähigkeit das eigene Unterrichtshandeln zu reflektieren und evidenzbasiert zu innovieren verbessern (Homt & van Ophuysen, 2019; Ulrich et al., 2020). Neben der Reflexionsfähigkeit soll Forschendes Lernen die Forschungskompetenz verbessern (Gussen et al., 2023), die in den notwendigen Fähigkeiten und Fertigkeiten, um eigenständig zu forschen (kognitive Komponente) und dem Interesse an Forschung und ihrem wahrgenommenen Nutzen (affektiv-motivationale Komponente) zum Ausdruck kommt (Besa et al., 2023; Thiel & Böttcher, 2014). Studien zeigen, dass die kognitive Komponente der Forschungskompetenz beim Forschenden Lernen zunimmt (Gussen et al., 2023; König, Darge, & Kramer, 2020). Gleichzeitig erleben Studierende Forschendes Lernen als stressauslösend und frustrierend (Schröder et al., 2023) und beurteilen seine Relevanz für die eigene Kompetenzentwicklung als gering (Gussen et al., 2023; Homt & van Ophuysen, 2019; vgl. affektiv-motivationale Komponente). Die Ziele verlängerter Praxisphasen gelten zudem als weitgehend verfehlt: Studierende werden mehrheitlich nicht bestärkt eine forschend-reflexive Haltung einzunehmen (Zurstrassen, 2018), der theoriegeleiteten Reflexion eigenen Unterrichtshandelns wird nur geringe Bedeutung zugeschrieben (Ukley & Bergmann, 2020) und die angestrebte Theorie-Praxis-Verzahnung bleibt weitgehend erfolglos (Ulrich et al., 2020). Ursächlich dafür könnte sein, dass eine kohärente Verzahnung der unterrichtspraktischen und forschenden Tätigkeit häufig nicht gelingt (Mitzschke & Neuhaus, 2016).

Fragestellung

Wirkt sich eine Interventionsmaßnahme mit gezielter Verzahnung der unterrichtspraktischen und forschenden Tätigkeit positiv aus auf...

- die Reflexionskompetenz (FF1)?
- die Forschungskompetenz in der affektiv-motivationalen Komponente (FF2) und die intrinsische Motivation forschend zu lernen (FF3)?
- das Stressempfinden in der Praxisphase (FF4)?

Methode

In einer quasi-experimentellen Interventionsstudie mit Prä- und Posttest-Design entwickelten N = 54 Masterstudierende des Lehramts Biologie in einer biologiedidaktischen Lehrveranstaltung Lernaufgaben und erprobten ihr Lernangebot im Anschluss am Lernort Schule (unterrichtspraktische Tätigkeit). Die Experimentalgruppe holte unmittelbar nach der unterrichtspraktischen Tätigkeit das Feedback der Schüler*innen zur Unterrichtsqualität über ein Onlinesurvey ein (forschende Tätigkeit), das zur Reflexion der unterrichtspraktischen Tätigkeit herangezogen werden konnte (CAP-Instrument; Gerhard et al., 2023). Forschung rund um Schüler*innenfeedback zeigt, dass Schüler*innen Unterrichtsqualität hinsichtlich ausgewählter Qualitätskriterien valide beurteilen können (Ferreira González et al., 2024; Göbel et al., 2021). Die Kontrollgruppe war frei in ihrer forschenden Tätigkeit. Die beiden Komponenten der Forschungskompetenz und das anwendungsbezogene Methodenwissen wurden zu beiden Messzeitpunkten über Selbsteinschätzungsskalen (Thiel & Böttcher, 2014; Wessels et al., 2021) bzw. einen Performanztest (Cammann, 2020) erhoben. Zum Zeitpunkt des Posttests wurden zusätzlich das Stressempfinden der Studierenden (Schneider et al., 2020) und die intrinsische Motivation (Wilde et al., 2009) erfasst. Außerdem erstellten die Studierenden schriftliche und vorstrukturierte Reflexionen über die Unterrichtsqualität, die inhaltsanalytisch ausgewertet werden. Als Treatmentcheck, der sich auf die empfundene Kohärenz zwischen der unterrichtspraktischen und forschenden Tätigkeit bezieht, wurde eine selbstentwickelte Skala aus neun Items eingesetzt.

Vorläufige Ergebnisse

Baseline-Analysen zeigen, dass sich die beiden Gruppen zum Zeitpunkt des Prätests nicht signifikant hinsichtlich ihrer Forschungskompetenz und ihres forschungsbezogenen Wissens unterschieden und beide Konstrukte reliabel erfasst werden konnten. Die Posttest-Ergebnisse werden auf der Tagung berichtet, da die Datenerhebung erst im Juli 2024 abgeschlossen wurde. Theoretische und praktische Implikationen für die zukünftige Gestaltung des Praxissemesters werden auf der Tagung diskutiert.

CLEVER – Classroom Management Video Intervention - Von der Intervention zur Implementation einer videobasierten Lehrveranstaltung zur Förderung der klassenführungsspezifischen Kompetenzen bei Lehramtsstudierenden

Sarah Strauß, Charlotte Kramer, Johannes König
Universität zu Köln

Theoretischer Hintergrund

Effektive Klassenführung gilt als Basisdimension qualitätvollen Unterrichts (Klieme, 2018) und erweist sich in Metaanalysen als ein Merkmal mit großer Vorhersagekraft für die Lernleistung von Schüler*innen (Hattie, 2012; Wang et al., 1993). Es zeigt sich jedoch, dass sich Lehramtsstudierende diesbezüglich nicht hinreichend vorbereitet fühlen (Jones, 2006; Poznanski et al., 2018). Für eine effektive Klassenführung gelten neben Wissen und affektiv-motivationale Merkmalen vermehrt auch situationsspezifische Fähigkeiten der Wahrnehmung, Interpretation und Entscheidung in konkreten Unterrichtssituationen als zentrale Voraussetzungen (Blömeke et al., 2015; Kaiser & König, 2019) und als Prädiktor für tatsächlichen Unterrichtsqualität (Blömeke et al., 2022; König & Kramer, 2016; König et al., 2021). Um die professionelle Kompetenz und vor allem die situationsspezifischen Fähigkeiten von Klassenführung zu fördern, erweisen sich Unterrichtsvideos als wirksam (Gold et al., 2013; Kramer et al., 2017; Schneider et al., 2016; Seidel et al., 2011; Sunder et al., 2016).

Obwohl sich die Forschung zu den situationsspezifischen Fähigkeiten in den letzten zwei Jahrzehnten stark entwickelt hat (König et al., 2022) und das Thema Klassenführung vermehrt in videobasierten Interventionsstudien bearbeitet wird (Junker & Holodynski, 2022; Kramer et al., 2020; Weber et al. 2020), fehlen diesbezüglich empirische Nachweise zur Wirksamkeit der (universitären) Lehrkräftebildung, die über Einzelinterventionen hinaus gehen und in methodisch anspruchsvollen Designs (standardisiert) durchgeführt werden (Hill et al., 2021; König, Heine et al., 2024).

Fragestellung

Im Rahmen des an der Universität zu Köln laufenden interdisziplinären Verbundprojekts „Proving the Effectiveness of Teacher Education“ wird über drei Semester der Frage nach der Wirksamkeit der neu entwickelten videobasierten Intervention „CLEVER – CLassroom ManagEment Video IntERvention“ zur Förderung der klassenführungsspezifischen Kompetenzen bei Lehramtsstudierenden im Masterstudium nachgegangen.

Um zu untersuchen, ob die Intervention wirksam und auf verschiedene Kontexte übertragbar ist, bspw. im Hinblick auf Merkmale der Lerngruppe sowie der Lehrenden, werden drei Gruppen unterschieden: A) Modellseminare: Durchgeführt von mit dem Thema Klassenführung vertrauten Lehrenden. Die mögliche Kompetenzveränderung von Studierenden dieser Seminare wird mit einem B) Implementationsseminar (Standardisierte Durchführung durch mit dem Thema unerfahrenen Dozierenden) und einem C) Kontrollseminar (kein Situationsbezug, anderer Inhalt) verglichen.

Es werden folgende Fragestellungen fokussiert:

F1: Führt die Intervention für die Modellseminare zu einem Zuwachs an professioneller Kompetenz, operationalisiert über situationsspezifische Fähigkeiten in Form von (a) Wahrnehmen und Interpretieren und (b) Entscheiden, bzw. Handlungsalternativen generieren, (c) im spezifischem Wissen zur Klassenführung sowie (d) zu einer Verbesserung der Selbstwirksamkeit?

F2 Ist eine Übertragung auf das Implementationsseminar möglich, d.h. eine standardisierte Durchführung durch im Thema Klassenführung unerfahrene Dozierende?

Methode

Im Fokus der standardisierten Intervention steht die Analyse von Klassenführung in authentischen Unterrichtsvideos. In acht Sitzungen werden konkrete Anforderungsbereiche in Form standardisierter Materialien (wissenschaftliche Texte, Unterrichtsvideos) bearbeitet.

Mittels eines experimentellen Designs (2 Modellseminare, 1 Implementationsseminar, 1 Kontrollgruppe; aktuell n=74) wird untersucht, ob die professionelle Kompetenz von Klassenführung bei Lehramtsstudierenden im Master gefördert werden kann. Alle Seminare werden im Prä-Post-Design mit standardisierten (u.a. videobasierten) Testinstrumenten getestet. Hierzu gehören ein neu entwickelter auf Klassenführung fokussierter Wissenstest (TCM-KT, König et al., 2024), sowie videobasierte Instrumente zur Erfassung der situationsspezifischen Fähigkeiten der Wahrnehmung und Interpretation (CME, König, 2015) und Entscheidung (CME-Decide, Weyers et al., 2024).

Bei der Pilotierung im Wintersemester 2023/24 konnten im Rahmen eines quasi-experimentelles Studiendesign unter Einschluss von fünf Seminaren (n=134) erste Wirksamkeitsnachweise der Intervention gezeigt werden.

Ergebnisse

Im Vortrag werden Ergebnisse zur Lernwirksamkeit aus der Durchführung der CLEVER-Intervention im Sommersemester 2024 erstmalig in einem experimentellen Design präsentiert. Zum aktuellen Zeitpunkt werden die Daten des zweiten Messzeitpunktes im Sommersemester noch erhoben. Erwartet werden jedoch auf Basis der vorliegenden Ergebnisse des ersten Durchlaufs im Wintersemester 2024 Zuwächse im Wissen und den situationsspezifischen Fähigkeiten im mittleren bis großen Bereich in den Modellseminaren sowie Zuwächse für das Implementationsseminar, die im Vergleich etwas geringer ausfallen. Praktische Implikationen werden diskutiert.

S-20: Zeitlicher Bezug als vernachlässigter Aspekt bei der Untersuchung von kognitiver Aktivierung

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 165

Chair der Sitzung: Benjamin Herbert

Chair der Sitzung: Anna-Katharina Praetorius

Chair der Sitzung: Matthias Martens

Zeitlicher Bezug als vernachlässigter Aspekt bei der Untersuchung von kognitiver Aktivierung

Chair(s): **Benjamin Herbert** (Goethe Universität Frankfurt, Deutschland), **Anna-Katharina Praetorius** (Universität Zürich)

Diskutant*in(nen): **Matthias Martens** (Universität zu Köln)

Kognitive Aktivierung gilt aus Sicht der generischen wie auch der fachspezifischen Unterrichtsqualitätsforschung als zentrale Qualitätsdimension von Unterricht, da ihr in besonderem Maße wünschenswerte Effekte auf die Leistungs- und Motivationsentwicklung von Schüler*innen zugeschrieben werden (Klieme et al., 2006; Praetorius, Klieme, et al., 2020). Dabei zeigt sich die empirische Befundlage heterogener, als es die Prominenz des Konstrukts erwarten ließe (Herbert, 2023; Praetorius et al., 2018). In der Folge wird derzeit viel über mögliche Gründe dieser Diskrepanz diskutiert (für einen Überblick Praetorius et al., im Druck). Eine mögliche Ursache, die bisher jedoch nur wenig beleuchtet wurde, ist der zeitliche Bezugsrahmen.

Empirische Studien zeigen, dass kognitive Aktivierung zwischen Unterrichtsstunden stark variieren kann (Gabriel-Busse & Lipowsky, 2021; Praetorius et al., 2014), wobei erste Befunde darauf hindeuten, dass eine fachspezifische Operationalisierung des Konstrukts diesen Umstand fördert (Jentsch et al., 2021). Auch im zeitlichen Verlauf einzelner Stunden variiert das Ausmaß der kognitiven Aktivierung, denn Unterrichtsphasen können unterschiedliche Zielsetzungen verfolgen und nicht jeder Inhalt bietet das gleiche Potenzial dafür, Schüler*innen ein kognitiv aktivierendes Angebot zu unterbreiten (Hermkes et al., im Druck; Hugener et al., 2007; Jentsch et al., 2021; Rakoczy et al., 2022; Wemmer-Rogh et al., 2023). Auch aus lernpsychologischer Perspektive erscheint ein durchgehend hoch kognitiv aktivierender Unterricht wenig zielführend für den Lernerfolg der Schüler*innen (Renkl, im Druck). Vor diesem Hintergrund stellt sich die Frage, wie der zeitliche Bezugsrahmen die Validität von Erhebungen der kognitiven Aktivierung beeinflusst. Entsprechend scheint es sinnvoll, die Frage nach sinnvollen zeitlichen Bezügen bei der Erforschung von kognitiver Aktivierung gezielt in den Blick zu nehmen.

Dazu wird in einem ersten Beitrag der zeitliche Bezugsrahmen bisheriger quantitativer Erhebungsinstrumente aufgearbeitet und Designunterschiede hinsichtlich ihrer Auswirkungen auf die Güte und Aussagekraft der Instrumente diskutiert. Der zweite Beitrag befasst sich mit der Frage, ob kognitive Aktivierung als Komposition zeitunabhängiger Facetten oder als prozessuale Konfiguration von Dimensionen verstanden werden sollte und untersucht dazu empirisch, ob eine Operationalisierung durch Prozessmuster die Prädiktivität für motivationale Merkmale der Schüler*innen erhöht. Anschließend werden in zwei Beiträgen aus fachdidaktischer Perspektive Einflüsse des zeitlichen Bezugsrahmens auf inhaltsorientierte Operationalisierungen untersucht: Der dritte Beitrag widmet sich der Frage nach der optimalen Länge der Analyseeinheit in Videostudien. Dabei wird kognitive Aktivierung entlang verschiedener Verstehens Elemente beim Erlernen eines ausgewählten Inhalts operationalisiert. Der vierte Beitrag schließlich widmet sich der Frage nach der Notwendigkeit von Anpassungen des Verständnisses und der Operationalisierung von kognitiver Aktivierung in Abhängigkeit von der Positionierung einer Unterrichtsstunde im Verlauf einer Unterrichtsreihe.

Die basierend auf der Logik der quantitativen Unterrichtsqualitätsforschung entwickelten Fragestellungen, Analysen und Argumentationen der vier Beiträge werden im Diskussionsbeitrag basierend auf einer qualitativ-rekonstruktiven Perspektive konstruktiv-kritisch beleuchtet.

Beiträge des Symposiums

Zeitlicher Bezugsrahmen bei der Erfassung von kognitiver Aktivierung: Eine Analyse quantitativer Instrumente

Benjamin Herbert¹, Wida Wemmer-Rogh², Patrick Schreyer³, Thilo Kleickmann⁴, Anna-Katharina Praetorius²

¹Goethe-Universität Frankfurt, ²Universität Zürich, ³Universität Kassel, ⁴Christian-Albrechts-Universität zu Kiel

Theoretischer Hintergrund. Für kognitive Aktivierung wurden in den letzten 20 Jahren kontinuierlich neue quantitative Erhebungsinstrumente entwickelt (Herbert et al., im Druck; Wemmer-Rogh et al., im Druck). Der zeitliche Referenzrahmen dieser Instrumente variiert erheblich und kann sich beispielsweise auf einen kurzen Ausschnitt einer einzelnen Unterrichtsstunde (Rakoczy et al., 2022) oder aber den gesamten bisherigen Unterricht bei einer Lehrperson beziehen (Praetorius, Fischer, et al., 2020). Diese große Varianz kann die Vergleichbarkeit von Erhebungen beeinträchtigen (Herbert, 2023), zu inkonsistenten Ergebnissen bei der Analyse theoretisch erwarteter Effekte auf Schüler*innen führen (Praetorius et al., 2018) und die Kommunikation über kognitive Aktivierung erschweren (Charalambous & Praetorius, 2020). Eine systematische Aufarbeitung des zeitlichen Referenzrahmens von quantitativen Instrumenten sowie eine Diskussion der Konsequenzen verschiedener Vorgehensweisen liegen dennoch bislang nicht vor.

Fragestellung. Diesem Desiderat widmet sich der vorliegende Beitrag anhand von zwei Forschungsfragen: (1) Wie unterscheiden sich quantitative Instrumente zur Erfassung von kognitiver Aktivierung im Hinblick auf ihren zeitlichen Bezugsrahmen? (2) Wie beeinflussen Unterschiede die Güte und Aussagekraft?

Methode. Um die Gestaltungsvielfalt der Instrumente aufzuzeigen, wurde eine umfangreiche Literaturrecherche durchgeführt, in der 75 verschiedene quantitative Instrumente identifiziert wurden. Für diese wurde der zeitliche Referenzrahmen ausgewertet und Unterschiede vor dem Hintergrund theoretischer Annahmen und bisheriger empirischer Befunde eingeordnet.

Ergebnisse und ihre Bedeutung. In keiner Beschreibung der 75 Instrumente findet sich ein expliziter Verweis darauf, dass der zeitliche Referenzrahmen aufgrund des zugrundeliegenden theoretischen Verständnisses von kognitiver Aktivierung gewählt wurde. Die ausgewerteten Instrumente beziehen sich in vielen Fällen auf einen längeren Zeitraum wie das vergangene Schuljahr (12%), den bisherigen Unterricht mit der Lehrperson oder der Klasse (12%) oder allgemein auf den Unterricht in einem Fach (13%). Häufig werden auch Rückschlüsse auf eine Unterrichtseinheit (31%), eine Unterrichtsstunde (18%) oder eine Unterrichtsphase (10%) gezogen, an die sich entweder eine Befragung anschließt oder deren Unterricht mit Beobachtungsinstrumenten evaluiert wird. Findet eine Evaluation von Unterrichtsmaterialien statt, beziehen sich die Instrumente in einigen Fällen auch auf die kognitive Aktivierung einzelner Aufgaben (4%). Bei Auswertungen des Unterrichts durch externe Rater*innen, werden zudem häufig die Ratings mehrerer Unterrichtsstunden zusammengefasst und bezogen auf einen größeren

zeitlichen Bezugsrahmen, z.B. eine Unterrichtseinheit, interpretiert. Eine entsprechende Extrapolation findet bei 58% der untersuchten Beobachtungsinstrumente statt.

In Bezug auf die zweite Forschungsfrage lässt sich feststellen, dass kein Instrument für alle zeitlichen Bezugsrahmen valide Aussagen über kognitive Aktivierung erlaubt. Dies ist unter anderem darauf zurückzuführen, dass mit dem zeitlichen Bezug auch die Gütekriterien für Instrumente variieren. Sollen Aussagen über kurze Zeiträume getroffen werden, gilt es, insbesondere darauf zu achten, dass die in Items adressierten Praktiken und Handlungen überhaupt auftreten. Denn kognitiv aktivierender Unterricht kann sich je nach Unterrichtsgegenstand und Unterrichtsphase unterschiedlich darstellen (Keller et al., in diesem Band; Wemmer-Rogh et al., 2023). Bei einer Erfassung der kognitiven Aktivierung über längere Zeiträume ist die Güte von Instrumenten u.a. abhängig von Faktoren, die sich je nach gewähltem Erhebungsmodus unterscheiden. Bei Einschätzungen durch externe Beobachtende ist zu überprüfen, inwiefern die gewählten Analyseeinheiten auf längere Zeiträume extrapoliert werden können. Dies kann beispielsweise durch eine Generalisierbarkeitsstudie erfolgen (Herbert, 2023). Die Beurteilung der kognitiven Aktivierung durch Schüler*innen und Lehrpersonen ist hingegen nur dann valide, wenn die Befragten in der Lage sind, Unterschiede zwischen verschiedenen Zeiträumen überhaupt zu differenzieren (Wagner, 2008). Diesbezüglich liegen jedoch bislang keine verlässlichen Befunde oder Empfehlungen zur Förderung präziserer Einschätzungen vor.

Die Ergebnisse zum zeitlichen Bezugsrahmen von Instrumenten für die Erfassung von kognitiver Aktivierung lassen sich auf andere Konstrukte der Unterrichtsqualitätsforschung übertragen, z.B. konstruktive Unterstützung. Sie stellen damit relevante Kriterien für die Entwicklung neuer und die Interpretation bestehender Instrumente dar und tragen dazu bei, Unterrichtsqualität - insbesondere kognitive Aktivierung - genauer zu beschreiben und zu erforschen.

Muster kognitiver Aktivierung – Zeitunabhängige Facetten oder prozessuale Konfiguration?

Rico Hermkes, Benjamin Herbert, Manon Heuer-Kinscher, Gerhard Minnameier
Goethe-Universität Frankfurt

Theoretischer Hintergrund

Kognitive Aktivierung gilt als eine von drei Basisdimensionen guten Unterrichts und ist ein häufig untersuchtes Qualitätsmerkmal (Praetorius et al., im Druck). Die temporale Dimension kommt dabei aber oft zu kurz. Das gilt in zweierlei Hinsicht: Zum einen erfolgt die Erhebung meist als globales Rating und nicht als Prozessgröße (Herbert et al., im Druck; Praetorius et al., 2018), zum anderen wird die Annahme präsupponiert, dass alle Dimensionen kognitiver Aktivierung zu jedem Zeitpunkt im Lernprozess gleichermaßen Relevanz haben. Gerade Letzteres kann aber angezweifelt werden, wenn man kognitive Aktivierung aus einer Prozessperspektive betrachtet. So macht es einen Unterschied, ob es z.B. um die initiale Evozierung kognitiver Konflikte geht oder um die nachfolgende Aufrechterhaltung kognitiver Aktiviertheit (Hermkes et al., eingereicht).

Ziel und Fragestellung

Ziel der Studie ist es, Prozessmuster kognitiver Aktivierung zu identifizieren und hinsichtlich ihrer Wirksamkeit auf Lernoutcomes zu untersuchen. Dies erfolgt konzeptionell ausgehend von einer inferentiellen Perspektive, die es ermöglicht, Wissens-Konstruktionsprozesse im Hinblick auf ihre kognitive Charakteristik zu rekonstruieren. So gibt es initiale Unterrichtssequenzen, die als Probleminduktionen rekonstruiert werden können und die der Evozierung kognitiver Konflikte dienen. Daran anschließend folgen Sequenzen, in denen Lernende eigenaktiv Lösungsansätze erarbeiten, die als abduktive Inferenzen rekonstruiert werden können. Die Hypothese ist, dass im zeitlichen Zusammenwirken beider Facetten kognitiver Aktivierung positive Zusammenhänge mit Lernoutcomes resultieren, die die jeweiligen Facetten allein nicht in gleicher Weise bewirken können.

Methode und Stichprobe

Dazu wurde ein Verfahren entwickelt, das kognitive Aktivierungs- und konstruktive Unterstützungssequenzen aus videographierten Unterrichtsstunden inferentiell rekonstruiert und erfasst, ob diese Sequenzen jeweils konstruktiv oder transmissiv von den Lernenden durchlaufen werden. Untersucht wurden 20 Schulklassen mit insgesamt 315 Schüler*innen (je eine Doppelstunde zum Thema „quantitativer und qualitativer Angebotsvergleich“) im Wirtschaftsunterricht an beruflichen Schulen in Hessen. Die Reliabilitäten der Kodierungen fielen für die Inferentielle Rekonstruktion substantiell ($\kappa = .75$) und die Kodierung der Wissenskonstruktionsaktivität hoch ($\kappa = .91$) aus.

Eine Probleminduktion, die einen klassischen Indikator für kognitive Aktivierung darstellt (Lauterbach et al., 2013), fand im Unterricht von 52% der Schüler*innen statt. Jedoch hatten nur 39% der Lernenden im Anschluss daran die Möglichkeit, zur Entwicklung eines Lösungsansatzes beizutragen.

Als Zielgrößen wurde das Kompetenzerleben der Schüler*innen, deren Autonomiewahrnehmung sowie die kognitive Belastung (intrinsic load) einbezogen. Es wurden multiple Regressionsanalysen berechnet, in denen die erwartete Leistung der Schüler*innen im Vergleich zum Klassendurchschnitt, die Range der erwarteten Leistung in der Klasse als Indikator für die Leistungsheterogenität der Klassen sowie das Interesse an Wirtschaft als Kontrollvariablen berücksichtigt wurden.

Ergebnisse

Die Analysen zeigen, dass in Stunden, in denen nur Probleminduktionen ohne anschließende eigenaktive abduktive Lernschritte stattfinden, hypothesenkonform kein Zusammenhang mit dem Kompetenzerleben vorliegt. Erst wenn sich eine solche abduktive Sequenz anschließt, in der Schüler*innen Wissen aktiv konstruieren, resultiert ein marginal signifikanter Zusammenhang ($\beta = .21$, $SE = .11$). Bezogen auf die kognitive Belastung zeigt sich ein marginal signifikanter Zusammenhang in Stunden mit initialer Probleminduktion ($\beta = -.13$, $SE = .27$) sowie ein signifikanter Zusammenhang für Stunden, in denen sich eine abduktive Sequenz mit eigenaktiven Schüler*innen anschließt ($\beta = -.14$, $SE = .29$). Bezogen auf das Autonomieerleben zeigen sich keine signifikanten Zusammenhänge.

Diskussion

Die Studie liefert erste Evidenzen für lernphasenabhängige Effekte der kognitiven Aktivierung. Limitationen resultieren, da die Studie aus einer kleinen Stichprobe bestand, so dass Effekte der Schachtelung von Schüler*innen in Klassen nur bedingt durch Kontrollvariablen berücksichtigt werden konnten. Bestätigen sich diese Befunde auf einer breiteren Evidenzbasis, würde das für ein Konzept kognitiver Aktivierung sprechen, in dem sich die relevanten Facetten in spezifischer Weise temporal konfigurieren. Eine Aggregation aller Facetten ohne Rücksicht auf deren zeitabhängige Bedeutsamkeit für Lernprozesse würde dagegen zu kurz greifen.

Das Problem des zeitlichen Bezugsrahmens bei der Einschätzung der kognitiven Aktivierung aus mathematikdidaktischer Sicht

Barbara Drollinger-Vetter, Alex Buff
Pädagogische Hochschule Zürich

Theoretischer Hintergrund. Die kognitive Aktivierung im Unterricht ist weder innerhalb einzelner Unterrichtsstunden (Herbert et al., im Druck) noch über mehrere Unterrichtsstunden hinweg (Praetorius et al., 2014) stabil. Daher ist es bei deren Einschätzung wichtig, sich zu überlegen, wie der zeitliche Bezugsrahmen beschaffen sein sollte. Hierzu existieren gemäß Herbert et al. (im Druck) bislang jedoch nur wenige klare Empfehlungen. Vor diesem Hintergrund stellt sich die Frage, was eine fachdidaktische Perspektive zur Lösung dieses Problems beitragen könnte, zumal in den letzten Jahren vermehrt über eine stärker fachdidaktisch ausgerichtete Erfassung von Unterrichtsqualitätsmerkmalen generell und der kognitiven Aktivierung im Besonderen diskutiert wurde (z.B. Brunner & Star, 2024; Dreher & Leuders, 2021; Lindmeier & Heinze, 2020; Praetorius & Gräsel, 2021; Praetorius et al., 2020; Reusser & Pauli, 2021).

In der Pythagoras-Studie (Klieme et al., 2009) wurden drei fachdidaktische Unterrichtsqualitätsmerkmale entwickelt und aus Expertensicht mittel- bis hoch-inferent eingeschätzt: das Vorkommen von Verstehenselementen, die Qualität der Repräsentationen des Satzes des Pythagoras und die inhaltlich-strukturelle Klarheit des Unterrichts (Drollinger-Vetter, 2011). Dies war nur möglich, weil der Inhalt und das Lernziel der videografierten Lektionen standardisiert waren: Verstehen des Satzes des Pythagoras nach einer dreistündigen Einführung. Da diese drei fachdidaktischen Unterrichtsqualitätsmerkmale gemäß Drollinger-Vetter (2023; siehe auch Pauli et al., 2008) als spezielle lernzielbezogene Subdimensionen der kognitiven Aktivierung aufgefasst werden können, kann darauf aufbauend die Frage des geeigneten zeitlichen Bezugsrahmens zur Erfassung der kognitiven Aktivierung lernzielspezifisch und theoriegeleitet untersucht werden.

Fragestellung. Welche Folgen hat eine Veränderung des zeitlichen Bezugsrahmens hinsichtlich der Einschätzung der drei oben genannten Subdimensionen der kognitiven Aktivierung und worin könnten die Gründe für divergierende Einschätzungen bestehen?

Methode. Für die Klärung der Fragestellung wurde der Unterricht einer Lehrperson aus der Pythagoras-Studie mit hohen Werten in den drei Subdimensionen der kognitiven Aktivierung unter Verwendung eines neuen zeitlichen Bezugsrahmens erneut eingeschätzt. Im Pythagoras-Projekt war die Einschätzung über die Theoriephasen aller drei Lektionen hinweg erfolgt. Dies schien den Projektbeteiligten mit Blick auf ihr Forschungsinteresse aus theoretischen Gründen der angemessenste zeitliche Bezugsrahmen zu sein (Drollinger-Vetter, 2011; Drollinger-Vetter & Lipowsky, 2006). Bei der Reanalyse wurde für jede der drei Lektionen einzeln eine Einschätzung (über die Theoriephasen) vorgenommen. Von Interesse ist nachfolgend weniger die Frage nach den Veränderungen der Einschätzungen selbst als vielmehr die Frage nach den möglichen Gründen, die zu den Veränderungen geführt haben könnten. Hierfür wurden die Konzeptualisierungen der drei Subdimensionen sowie die jeweiligen Itemformulierungen anhand der Videodaten miteinander in Beziehung gesetzt und qualitativ analysiert.

Ergebnisse. Es zeigte sich, dass der neue zeitliche Bezugsrahmen die Einschätzung der drei Subdimensionen teilweise verändert hatte. In Einzelfällen wurde in gewissen Theoriephasen eine Einschätzung aufgrund eines nicht mehr sinnvoll anwendbaren Items gar verunmöglicht. Diese Veränderungen traten nicht bei allen drei Subdimensionen der kognitiven Aktivierung gleichermaßen auf. Die Verkürzung des zeitlichen Bezugsrahmens führte in einigen Fällen außerdem zu einer Veränderung der inhaltlichen Bedeutung einer Subdimension.

Bedeutung der Ergebnisse. Gemäß Herbert et al. (im Druck) liegt ein Grund für solche Veränderungen in den Einschätzungen darin, dass die in den Instrumenten adressierten Unterrichtsphänomene in einem kurzen zeitlichen Bezugsrahmen gegebenenfalls nicht mehr beobachtbar sein können. Da im vorliegenden Beitrag Unterricht mit einem vorgegebenen Lernziel betrachtet wurde und die Konzeptualisierung der kognitiven Aktivierung genau darauf ausgerichtet war, deuten die Ergebnisse auf einen weiteren Grund hin, der fachdidaktischer Art ist: Die Wahl des zeitlichen Bezugsrahmens, das Lernziel der videografierten Unterrichtslektionen und die Konzeptualisierung resp. die Operationalisierung der kognitiven Aktivierung müssen aus inhaltlich-theoretischer Sicht eng aufeinander abgestimmt werden. Dies steht im Einklang mit der oben erwähnten Forderung nach einem verstärkten Einbezug der fachdidaktischen Perspektive bei der Erforschung der Unterrichtsqualität. Die Bedeutung dieser Ergebnisse wird im Beitrag ausführlicher diskutiert.

**Zeitliche Bezüge als Einflussfaktor bei der Entwicklung von Ratings zur kognitiven Aktivierung in
Unterrichtsreihen zu literarischen Texten**

Maximilian Arend, Florian Hesse, Iris Winkler
Friedrich-Schiller-Universität Jena

Theoretischer Hintergrund: Die quantitative Unterrichtsqualitätsforschung der letzten Jahre hat sich vornehmlich mit dem MINT-Unterricht auseinandergesetzt (Praetorius et al., 2018). Erst kürzlich werden auch für andere Fächer Qualitätsdimensionen modelliert (z. B. Lindl et al., 2024; Praetorius et al., 2020), darunter auch für den Literaturunterricht (Hesse, 2024; Winkler, 2020). Als zeitlicher Bezugspunkt fungieren dabei meist Einzelstunden, während die Betrachtung von Unterrichtsreihen, also mehrerer thematisch zusammenhängender Stunden, noch immer ein Desiderat darstellt (Wieser, 2019). Dies dürfte u.a. an einem Mangel an Instrumenten zur Beobachtung von Literaturunterricht in Reihenzusammenhängen liegen. Entsprechend besteht ein dringender Bedarf an einer systematischen (Weiter-)Entwicklung von Instrumenten, die an bisherige Überlegungen zur Untersuchung von Einzelstunden anknüpfen.

Fragestellung: Im Vortrag wird deshalb übergeordnet gefragt, inwiefern es gelingt, ein reliables und valides Instrument zur Beobachtung von Unterrichtsreihen aus einem bestehenden Instrument zur Beobachtung von Einzelstunden (Hesse, 2024) abzuleiten. Ein besonderer Fokus liegt hierbei auf der kognitiven Aktivierung, die fachdidaktisch bedeutsam ist, weil sie die Auseinandersetzung von Lernenden und Lerngegenstand fokussiert. Teilfragestellungen richten sich auf folgende Aspekte:

- F1: Inwiefern sind inhaltliche Anpassungen des Ratingmanuals notwendig, je nachdem, ob eine Stunde zu Beginn, in der Mitte oder am Ende einer Reihe den zeitlichen Bezugspunkt der Beobachtung darstellt?
- F2: Wie bewerten unabhängige Rater:innen in Abhängigkeit vom zeitlichen Bezugspunkt die Schwierigkeit, anhand des angepassten Manuals fachdidaktisch angemessene Urteile zu fällen sowie die Items voneinander abzugrenzen?

Methode: Als Ausgangspunkt fungierte ein Kodiermanual zur Beschreibung von kognitiv-emotionaler Aktivierung im Literaturunterricht (Hesse, 2024), das sich hinsichtlich seiner Inter-Rater-Reliabilität in einer Stichprobe von N = 22 Unterrichtsvideos zu Einzelstunden größtenteils als reliabel erwies ($.722 < ICC(A,1) < .824$; N = 3 Rater). Die kognitiv-emotionale Aktivierung wurde dabei über sieben vierstufige Items erfasst, die die Qualität von Lernaufgaben und Unterrichtsinteraktion einschätzen. Im Rahmen der vorliegenden Studie sollen die Items auf die Beobachtung von N = 10 Unterrichtsreihen zum Roman Tschick von Wolfgang Herrndorf aus dem Projekt QuaLiTyA (angebotsseitige Qualität von Literaturunterricht) übertragen werden. Zu jeder Reihe wurden drei Unterrichtsstunden (je 45–90 Minuten) videografiert: (1) die erste Stunde, (2) eine mittlere Stunde zu

einem vorgegebenen Romankapitel und (3) die Abschlussstunde (N = 30 Stunden). Zur Beantwortung beider Forschungsfragen werden im Vortrag erste Einblicke in die systematische Adaption der Items gegeben, die sich an etablierten Verfahren zur Item-Entwicklung in der Unterrichtsforschung orientiert (z. B. Lotz et al., 2013, S. 90). Hierzu sollen von August bis Oktober 2024 zunächst an einer kleineren Teilstichprobe (n = 3 zufällig gezogene Reihen mit insg. 9 Stunden) in einem deduktiv-induktiven Verfahren die bisherigen Items an die Videos aus dem QuaLiTyA-Korpus herangetragen und – ggf. in Abhängigkeit von den jeweiligen zeitlichen Bezugspunkten – modifiziert werden. Das daraus resultierende Ratingmanual wird einerseits mit Expert:innen aus den Bereichen Literaturdidaktik und Unterrichtsforschung diskutiert. Andererseits erfolgt eine unabhängige Pilotierung durch zwei unabhängige Rater:innen, um sich der Validität des Instruments zu vergewissern und dessen Tauglichkeit im Vorfeld einer statistischen Überprüfung der Inter-Rater-Reliabilität qualitativ zu explorieren. Zu diesem Zweck werden die Überarbeitungshinweise aus den Expert:innendiskussionen und die Anmerkungen der Rater:innen während der Erprobung schriftlich festgehalten und anschließend mittels qualitativ-strukturierender Inhaltsanalyse ausgewertet (Kuckartz, 2018).

Ergebnisse und ihre Bedeutung: Hinsichtlich der Einschätzung von Unterrichtsqualität in Reihenzusammenhängen lässt sich theoretisch begründen, dass sich das Potenzial an kognitiver bzw. kognitiv-emotionaler Aktivierung nicht aus der bloßen Addition von Einzelstundenratings ergibt, sondern die unterschiedlichen zeitlichen Bezugspunkte mit Blick auf Validität und Reliabilität der Messung berücksichtigt werden müssen (Herbert et al., im Druck). An welchen konkreten Punkten sich am Übergang von Einzelstunde zu Unterrichtsreihe in Ratinginstrumenten aus einer fachdidaktischen Perspektive Modifikationsbedarfe ergeben, ist bislang aber kaum erforscht. Die qualitativ gewonnen Erkenntnisse dieser Studie können zur Schließung dieser Forschungslücke beitragen und exemplarisch auch für vergleichbare Bemühungen in anderen Fächern relevant sein.

S-21: Gelingende Ko-Konstruktion von Bildungswissenschaften, Schulpraxis und Bildungsadministration: Empirische Ergebnisse zu ko-konstruktiver Zusammenarbeit in der Initiative Schule Macht Stark (SchuMaS)

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 148

Chair der Sitzung: Marie Baesch

Chair der Sitzung: Simon Ohl

Chair der Sitzung: Birgit Pikowsky

Gelingende Ko-Konstruktion von Bildungswissenschaften, Schulpraxis und Bildungsadministration: Empirische Ergebnisse zu ko-konstruktiver Zusammenarbeit in der Initiative Schule Macht Stark (SchuMaS)

Chair(s): **Marie Baesch** (DIPF, Deutschland), **Simon Ohl** (Universität Potsdam)

Diskutant*in(nen): **Birgit Pikowsky** (Pädagogisches Landesinstitut Rheinland-Pfalz)

In den letzten Jahren hat die ko-konstruktive Zusammenarbeit zwischen Wissenschaft und Praxis in der Schul- und Unterrichtsentwicklung zunehmend an Bedeutung gewonnen. Dabei wird ein einseitiges Transfervverständnis zugunsten bidirektionaler Austauschprozesse unter Einbezug der jeweiligen Systemlogiken überwunden (Farley-Ripple et al., 2018). Eine ko-konstruktive Zusammenarbeit erfordert dabei neue Formate, die gemeinsame Diskussions- und Aushandlungsprozesse ermöglichen (Manitius & Bremm, 2021). Trotz der Relevanz des Themas gibt es im deutschsprachigen Kontext bisher wenige empirische Studien, die Gelingensbedingungen solcher Formate untersuchen.

In der Initiative Schule Macht Stark werden Schul- und Unterrichtsentwicklungsmaßnahmen entwickelt und erprobt, die sich an Schulen in herausfordernder Lage richten. Dazu findet eine ko-konstruktive Zusammenarbeit von Wissenschaftler:innen und Personen aus Bildungspraxis und -verwaltung in mehreren Formaten statt. Das Symposium beleuchtet diese Prozesse multiperspektivisch anhand von vier empirischen Studien.

Der erste Beitrag geht der Frage nach, wie ein Service-Learning-Seminar in Kooperation zwischen Schule, Hochschule und Studierenden unter Berücksichtigung knapper zeitlicher Ressourcen und personeller Wechsel gestaltet werden kann. Auf Grundlage der Ergebnisse einer transdisziplinären Arbeitsphase, in der Lehrkräfte (N = 6), Universitätsmitarbeitende (N = 4) und Lehramtsstudierende (N = 13) diskutierten, wie unterschiedliche Akteursgruppen zum Gelingen von Service-Learning beitragen können, werden anhand einer Inhaltsanalyse zentrale Erfolgsfaktoren identifiziert.

Der zweite Beitrag nimmt die Erwartungen der Schulleitungen zu Beginn der Initiative in den Blick. Basierend auf einer inhaltsanalytischen Auswertung von 178 Interviews wird die Sicht von Schulleitungen auf ko-konstruktive Formate der Zusammenarbeit mit Wissenschaft herausgearbeitet. Die Ergebnisse werden unter Rückgriff auf das Konzept der Grenzarbeit (Gieryn, 1983) eingeordnet, welches Praktiken der Grenzziehung zwischen verschiedenen Akteursgruppen beschreibt.

Der dritte Beitrag untersucht in einem längsschnittlichen quantitativen Design die Frage, welche Merkmale der Netzwerkarbeit schulischer Akteur:innen mit einer positiven Bewertung der Ko-Konstruktion zusammenhängen. Dazu werden die Häufigkeit der Teilnahme sowie verschiedene Aspekte der Prozessqualität der in SchuMaS stattfindenden Schulleitungsnetzwerktreffen auf ihren Zusammenhang mit der wahrgenommenen Ko-Konstruktion durch die Schulleitungen (N= 138) untersucht.

Der vierte Beitrag betrachtet die ko-konstruktive Zusammenarbeit in der Triade von Wissenschaft, schulischer Praxis und Bildungsverwaltung. Nach einer theoretischen Einordnung dieser Akteurskonstellation werden anhand einer Einzelfallstudie im Land Hessen vertiefte Einblicke in den Versuch der Gestaltung der Ko-Konstruktion in der Triade im Rahmen der Initiative gegeben.

Das Symposium leistet einen Beitrag zur Erforschung ko-konstruktiver Zusammenarbeit zwischen Bildungswissenschaften, Schulpraxis und Bildungsadministration unter Beachtung der interdisziplinären Perspektiven am Beispiel der Initiative Schule Macht Stark. Neben Fallstricken und Herausforderungen werden dabei in allen Beiträgen vor allem Gelingensbedingungen herausgearbeitet, die handlungsleitend für künftige Schulentwicklungsprogramme sein können.

Beiträge des Symposiums

Sensibilisierung zukünftiger Lehrkräfte für diversitätssensible Schulentwicklung - Gelingensbedingungen und Wirkungen von Service Learning als Kooperation zwischen Hochschule und Schulen in herausfordernden Lagen

Merle Thielmann, Oscar Yendell, Julia Derkau
Universität Mannheim

Schulen in sozialräumlich deprivierter Lage stehen vor besonderen Herausforderungen aufgrund geringerer familiärer Ressourcen und Lernvoraussetzungen, die mit der Zusammensetzung der Schülerschaft assoziiert sind (van Ackeren et al., 2021). Die Weiterentwicklung schulischer Strukturen und Routinen zur Sicherung der Prozessqualität ist in diesen Schulen besonders relevant, um bestmögliche Lernergebnisse bei den Schüler:innen zu erzielen (Bremm et al., 2016). Entsprechend sind Lehrkräfte an diesen Schulen besonders gefragt, sich an Schulentwicklungsprozessen zu beteiligen. In der Lehrkräfteausbildung wird das Thema Schulentwicklung bislang jedoch kaum adressiert (Holtappels et al., 2021). Ein Konzept, das die Möglichkeit bietet, angehenden Lehrkräften theoretisches Wissen über Schulentwicklungsprozesse zu vermitteln und sie zugleich auf die besonderen Herausforderungen an Schulen in herausfordernden Lagen vorzubereiten, ist Service Learning (Derkau & Gerholz, 2023). Dieses Lehr-Lernformat ermöglicht eine enge Verknüpfung zwischen der Vermittlung theoretischer Inhalte und praktischem gesellschaftlichem Engagement (Bringle & Clayton, 2012). Das Engagement in realen schulischen Kontexten ermöglicht Studierenden Einblicke in die Schulentwicklungspraxis, sensibilisiert sie für gesellschaftliche Verantwortungsübernahme und stärkt ihre Selbstwirksamkeit (Derkau & Gerholz, 2023).

Der vorliegende Beitrag geht der Frage nach, wie Service Learning als Kooperation zwischen Schule, Hochschule und Studierenden unter Berücksichtigung knapper zeitlicher Ressourcen und häufiger personeller Wechsel an Schulen in herausfordernden Lagen gelingend gestaltet werden kann. Die Untersuchung fand im Rahmen eines Service Learning-Seminars statt, bei dem Lehramtsstudierende Kooperationsschulen der Initiative Schule Macht Stark durch die Visualisierung diversitätsbezogener Schulentwicklungsprozesse unterstützten. Als Datengrundlage dienen die Ergebnisse einer transdisziplinären Arbeitsphase zum Abschluss des Seminars, in der Lehrkräfte (N = 6), Universitätsmitarbeitende (N = 4) und

Lehramtsstudierende (N = 13) der Frage nachgingen, welche Beiträge die unterschiedlichen Akteursgruppen zum Gelingen von Service Learning mit Schulen in herausfordernden Lagen leisten können. Die Notizen der Arbeitsphase wurden anschließend anhand eines deduktiv-induktiv gebildeten Kategoriensystem ausgewertet, das aus den Qualitätskriterien für Service Learning (Hochschulnetzwerk Bildung durch Verantwortung/Redaktionsgruppe Qualität, 2018; Kuckartz & Rädiker, 2014) sowie den Formulierungen der Teilnehmenden der Arbeitsphase gebildet wurde.

Als besonders bedeutsam für die gelingende Umsetzung von Service Learning-Konzepten wurde von den Diskussionsteilnehmer:innen die Organisation der Zusammenarbeit zwischen allen Beteiligten genannt. Dies beinhaltete das Ausloten und zur Verfügung stellen notwendiger Ressourcen auf Seiten der Schule und Hochschule und die Herstellung einer Passung zwischen unterschiedlichen Handlungslogiken, beispielsweise in der Zeitplanung und dem Informationsfluss. Zentral wurde die effektive Kommunikation zwischen den Beteiligten hervorgehoben (25% der Nennungen), insbesondere der persönliche Kontakt zu festen Ansprechpartner:innen, sowie eine Offenheit aller Beteiligten. Zudem wurde häufig der Bedarf nach einer partnerschaftlichen Zusammenarbeit genannt (23% der Nennungen). Diese sollte von Respekt, Verständnis und Anerkennung für das Engagement aller Beteiligten geprägt sein. Als wichtig für diese partnerschaftliche Zusammenarbeit wurden eine gemeinsame Zielformulierung unter Berücksichtigung der Perspektiven der Schulen, der Anforderungen der Hochschule an das Lehrformat, der Abgleich von Erwartungen an das Format und eine Verlässlichkeit aller Partner bei der Umsetzung genannt. Seltener wurden freiwilliges Engagement der Studierenden (12%) sowie eine ausreichende organisatorische und fachliche Vorbereitung der Studierenden auf die Praxisphase (11%) genannt.

Die genannten Gelingensbedingungen spiegeln bereits etablierte Qualitätskriterien von Service Learning wider (Hofer & Derkau, 2020). Aspekte der Gestaltung des Praxisengagements und des Lehrformates traten jedoch in den Hintergrund, während gelingende Kommunikation und partnerschaftlichen Zusammenarbeit als zentrale Erfolgsfaktoren von Service Learning mit Schulen in sozialräumlich deprivierter Lage identifiziert wurden. Diese Befunde weisen darauf hin, dass eine Weiterentwicklung zeitlich begrenzter Service-Learning-Kooperationen hin zu längerfristigen Partnerschaften gewinnbringend sein könnte, die durch ihre Kontinuität das Verständnis für die gegenseitigen kommunikativen und organisatorischen Bedürfnisse sowie den Aufbau einer offenen, vertrauensvollen Kommunikation auf Basis persönlicher Kontakte fördern können (Hofer & Derkau, 2020). Die Ergebnisse werden daher vor dem Hintergrund diskutiert, wie das Potenzial langfristig angelegter Schulentwicklungsprojekte wie Schule Macht Stark oder des kommenden Startchancen-Programmes zur Etablierung erfolgreicher Service Learning-Konzepte in der Lehrkräftebildung genutzt werden kann.

Begegnung auf Augenhöhe oder Herabblicken vom Elfenbeinturm: Erwartungen von Schulleitungen an die Zusammenarbeit mit Wissenschaft im Spiegel der Grenzarbeit

Jonas Ringler, Marie Baesch, Steffen Brill, Alexandra Marx

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Zunehmend werden Überzeugungen zum Verhältnis von Wissenschaft und Praxis sowie ihrer Rolle in Prozessen des Transfers hinterfragt, insbesondere vor dem Hintergrund häufig ausbleibender Transfererfolge (Schuster et. al. 2024). So wird kritisiert, dass im unidirektionalen Verständnis die "Bildungswissenschaft als Transfergeberin und die Bildungspraxis als Transferadressatin" (Schuster et al. 2024, S. 8) aufgefasst wird. Vor diesem Hintergrund ist eine ko-konstruktive und nachhaltig angelegte Zusammenarbeit zwischen Praktiker:innen und Wissenschaftler:innen in der Schul- und Unterrichtsentwicklung in den vergangenen Jahren verstärkt in den Fokus der Bildungswissenschaften gerückt: Ansätze wie Research-Practice-Partnerships (RPP) (Coburn & Penuel 2016; Coburn et al. 2021) versuchen, das klassische Transferverständnis als Einbahnstraße (Schmiedl 2022) zu überwinden.

Ko-konstruktive Formate der transdisziplinären Zusammenarbeit (Defila & Di Giulio 2019) zwischen Wissenschaft und Praxis, die sich am bidirektionalen Transferverständnis (Farley-Ripple et al., 2018) orientieren, sind voraussetzungsreich. Sie erfordern u.a. die Integration unterschiedlicher Logiken der Akteursgruppen und Wissenssysteme als eine wesentliche Voraussetzung, wofür bspw. gemeinsame Kommunikationssettings und spezifische Diskussions- und Aushandlungsprozesse erforderlich sind (Manitius & Bremm, 2021). Bisherige Forschungsarbeiten zu Gestaltungsdynamiken der Zusammenarbeit zwischen Forschung und Praxis liegen vorwiegend aus Sicht von beteiligten Wissenschaftler:innen und nur selten aus der praktischen Perspektive vor (Coburn & Penuel, 2016). Es besteht daher eine Notwendigkeit der Analyse von Sichtweisen der Schulpraxis auf ko-konstruktive Prozesse der Schul- und Unterrichtsentwicklung.

Im Rahmen der Initiative Schule Macht Stark (SchuMaS) werden wirksame und praxisbewährte Maßnahmen in den Bereichen der Schul- und Unterrichtsentwicklung, der Professionalisierung und der Sozialraumorientierung gemeinsam mit der Schulpraxis identifiziert und vom Forschungsverbund adressat:innengerecht aufbereitet, um Schulen in sozialräumlich benachteiligter Lage in ihrer Entwicklungsarbeit zu unterstützen und somit zu einer Stärkung der Bildungsgerechtigkeit beizutragen (Maaz & Marx, im Erscheinen). Die kontinuierliche und kohärente Zusammenarbeit des Forschungsverbunds mit beteiligten Schulen erfolgt in insgesamt vier regionalen SchuMaS-Zentren. Hierbei handelt es sich um eine Schulentwicklungsbegleitung, die sich durch die Moderation der interschulischen Netzwerkarbeit, die netzwerkbasierte Schulentwicklung, die Peerberatung im Rahmen von Schulleitungsnetzwerken und darüber hinaus durch eine Schulentwicklungsberatung mit einem spezifischen Fokus auf datengestützte Schulqualitätsentwicklung auszeichnet. Die Zentren arbeiten damit angelehnt an den Ansatz der RPP ko-konstruktiv an der Schnittstelle zwischen Wissenschaft und Praxis.

Der Vortrag greift das Desiderat, die Erforschung der Sichtweise von Praxisakteur:innen zur Zusammenarbeit von Wissenschaft und Praxis, auf. Dafür werden Erwartungen von Schulleitungen an die Zusammenarbeit innerhalb von RPP mit Wissenschaftler:innen aus dem SchuMaS-Forschungsverbund analysiert. Datengrundlage sind 178 transkribierte leitfadengestützte qualitative Interviews, die zu Beginn des Projekts bundesweit an SchuMaS partizipierenden Grund-, Sekundar-, und Förderschulen durchgeführt wurden. 39% der Schulen nehmen eigeninitiativ an SchuMaS teil und 60 % der Schulen wurden entweder zur Bewerbung aufgefordert oder zur Teilnahme nominiert. Der mittels Qualitativer Inhaltsanalyse nach inhaltlich strukturierendem Typ (Kuckartz 2014) induktiv herausgearbeitete Erwartungshorizont wird zunächst deskriptiv dargestellt. Es zeigt sich beispielsweise bezogen auf die Erwartungen an die Zusammenarbeit mit Wissenschaft, dass das Format „Fortbildungen“ (6,7%) als deutlich weniger relevant wahrgenommen wird als das Format „Begleitung bei Schulentwicklungsprozessen“ (55,6%).

Anschließend an die Darstellung der empirischen Untersuchung wird das Konzept der Grenzarbeit (Gieryn, 1983), welches Praktiken der Grenzziehung zwischen verschiedenen Akteursgruppen beschreibt, zur Einordnung und Diskussion der induktiv herausgearbeiteten Erwartungen von Schulleitungen an Zusammenarbeit mit Wissenschaftler:innen genutzt. Ziel ist es dabei, die Sicht von Schulleitungen auf die Etablierung von RPP und die damit verbundene Notwendigkeit des Hinterfragens von Grenzen der eigenen Profession bezogen auf ko-konstruktive Formate der Zusammenarbeit herauszuarbeiten. Auf der einen Seite zeigt sich eine vorsichtige Erwartungshaltung der Schulleitungen an die Zusammenarbeit mit Wissenschaftler:innen, die

auf die (Re)produktion der Grenze zwischen Wissenschaft und Praxis zurückgeführt werden kann. Auf der anderen Seite kann gleichzeitig festgestellt werden, dass die Erwartungen der Schulleitungen an die Zusammenarbeit mit Wissenschaftler:innen über die Grenzen von Wissenschaft und Praxis hinweg im Sinne einer Grenztransformation überwiegen.

Prozessqualität von Schulleitungsnetzwerktreffen und ihr Einfluss auf die Bewertung von Ko-Konstruktion im SchuMaS-Projekt

Simon Ohl¹, Karina Karst², Hanna Dumont¹

¹Universität Potsdam, ²Universität Mannheim

Eine ko-konstruktive Zusammenarbeit von Bildungsforschung und Bildungspraxis hat in den letzten Jahren international an Bedeutung gewonnen und wird als zentrale Voraussetzung zur Bewältigung komplexer Herausforderungen in Bildungssystemen gesehen (Sliwka et al., 2024). Auch das Positionspapier der Landesinstitute und Qualitätseinrichtungen der Länder betont ein bidirektionales Transfervverständnis, dass eine ko-konstruktive Zusammenarbeit von Wissenschaft und Bildungspraxis und die Wertschätzung der Expertisen aller Beteiligten erfordert (Bieber et al. 2018). Schulleitungsnetzwerke, die durch Wissenschaftler:innen begleitet werden, bieten eine Plattform, solche ko-konstruktiven Beziehungen umzusetzen, indem Wissenschaft und Praxis gemeinsam und langfristig praxisrelevante Probleme bearbeiten und dabei Strategien zur Förderung der produktiven Zusammenarbeit nutzen (Berkemeyer, 2008).

Auch im SchuMaS-Projekt spielt die Zusammenarbeit in Schulleitungsnetzwerken eine zentrale Rolle (Marx, Karst & van Ackeren-Mindl, 2024). Die Schulleitungsnetzwerke bieten jeweils 5-10 Schulleitungen der insgesamt 200 teilnehmenden Schulen die Möglichkeit zum vertiefenden Austausch, basierend auf Prozessbegleitung durch Mitarbeitende des SchuMaS-Forschungsverbunds, kollegialem Austausch und inhaltlicher Arbeit mit Inputs aus dem SchuMaS-Forschungsverbund (Ringler et al, 2024). Ziel ist es, durch regelmäßige Treffen neue Praktiken und Konzepte zu entwickeln sowie Erfahrungen zu teilen. Die Treffen werden von Mitarbeitenden des SchuMaS-Forschungsverbunds geleitet und dienen auch der ko-konstruktiven Zusammenarbeit zwischen Wissenschaft und Praxis. Da Schulleitungsnetzwerktreffen in diesem Umfang erstmalig in Deutschland durch den SchuMaS-Forschungsverbund umgesetzt wurden, ist ein Ziel des SchuMaS-Forschungsverbunds Gelingensbedingungen einer ko-konstruktiven Zusammenarbeit zu identifizieren.

In dem Beitrag untersuchen wir den Einfluss der Häufigkeit der Teilnahme der Schulen sowie den Einfluss verschiedener Aspekte der Prozessqualität der Schulleitungsnetzwerktreffen auf die Bewertung der Ko-Konstruktion im SchuMaS-Projekt durch die Schulleitungen. Im Fokus steht dabei die Frage, welche Merkmale der Netzwerkarbeit mit einer positiven Bewertung der Ko-Konstruktion im SchuMaS-Projekt zusammenhängen.

Im Zeitraum von März 2022 bis Mai 2023 wurden die RZ-Mitarbeitenden nach jedem Schulleitungsnetzwerktreffen mit einem Onlinefragebogen zur Prozessqualität der Treffen befragt (persönliche Beziehung und Arbeitsbeziehung zu den Teilnehmenden, Austausch der Teilnehmenden untereinander, Transfer aus dem Netzwerk in die Schulen, Transfer aus dem Netzwerk in den Forschungsverbund). Insgesamt wurden Daten von 90 Treffen erfasst an denen Vertreter:innen von N=141 Schulen teilgenommen haben. Von Juni bis August 2023 fand eine Befragung der am SchuMaS-Projekt teilnehmenden Schulleitungen (N= 138) statt, in der unter anderem Fragen zur wahrgenommenen Ko-Konstruktion von Bildungsforschung und Bildungspraxis im Projekt gestellt wurden. Im Beitrag verknüpfen wir diese beiden Perspektiven, indem wir die wahrgenommene Ko-Konstruktion durch die Prozessqualität der Netzwerktreffen vorhersagen. In gewichteten Multiple Membership Models berücksichtigen wir die Teilnahme der Schulleitungen an mehreren Schulleitungsnetzwerktreffen, sodass wir untersuchen, inwiefern die Teilnahme an Netzwerktreffen, die durch eine bestimmte Prozessqualität gekennzeichnet waren, einen Einfluss auf die Bewertung der Ko-Konstruktion der Schulleitungen hatte. Dabei berücksichtigen wir, dass die Schulleitungen nicht an allen Treffen ihrer Netzwerke anwesend waren und untersuchen in diesem Zusammenhang, ob die relative Häufigkeit der Teilnahme einen Einfluss auf die Bewertung der Ko-Konstruktion hatte.

Die Analysen zeigen, dass verschiedene Merkmale der Prozessqualität einen Einfluss auf die Bewertung der Ko-Konstruktion der Schulleitungen hatten. Insbesondere die persönliche Beziehung ($\beta = .36, p < .01$) und die Arbeitsbeziehung ($\beta = .37, p < .01$) zu den Teilnehmenden sowie der Transfer aus dem Netzwerk in die Schulen ($\beta = .37, p < .01$) und in den SchuMaS-Forschungsverbund ($\beta = .35, p < .01$) zeigten einen positiven Zusammenhang mit Ko-Konstruktion. Der Austausch der Teilnehmenden untereinander, die Kooperation mit der Bildungsverwaltung und die relative Häufigkeit der Teilnahme hingen hingegen nicht mit der Bewertung der Ko-Konstruktion zusammen. Die Ergebnisse deuten darauf hin, dass weniger die Quantität als die Qualität (der sozialen Interaktionen) der Schulleitungsnetzwerke die Wahrnehmung der ko-konstruktiven Zusammenarbeit von Praxis und Wissenschaft beeinflussen. Darüber hinaus unterstreichen die Ergebnisse die Bedeutung persönlicher Beziehungen zwischen den beteiligten Personen (vgl. auch Ringler et al., 2024). Damit liefert der Beitrag erste Erkenntnisse zur Schulleitungsnetzwerkarbeit aus längsschnittlicher quantitativer Perspektive.

Ko-konstruktiv orientierte Zusammenarbeit von Wissenschaft, schulischer Praxis und Bildungsverwaltung: Theoretische Ansätze und Gelingensbedingungen am Beispiel der Initiative Schule Macht Stark

Anne Klein¹, Julia Michalla¹, Isabell van Ackeren-Mindl²

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Universität Duisburg-Essen

Komplexe gesellschaftliche Herausforderungen, wie z.B. soziale Ungleichheiten, erfordern meist ebenso komplexe Lösungsstrategien, die – schaut man etwa auf das Schulsystem – nicht ausschließlich auf einer Ebene von einer Akteursgruppe, wie z.B. Lehrkräften (vgl. Klopsch & Sliwka, 2021), problemlösend entwickelt und wirksam implementiert werden können. Auch die Zusammenarbeit zwischen Wissenschaft und Praxis (vgl. Kerres et al., 2022) greift zu kurz, da Veränderungspotentiale in den systemischen Strukturen begrenzt sind. Hingegen kann eine ko-konstruktiv orientierte Zusammenarbeit in der Triade von Bildungswissenschaft, schulischer Praxis und Bildungsverwaltung zu einer nachhaltigen Schul- und insbesondere Bildungssystementwicklung beitragen (vgl. Jungermann, 2021), sofern es den Beteiligten gelingt, trotz unterschiedlicher Organisationslogik, Sprache und zeitlichem Handlungsdruck, einen Weg zur zielbezogenen Kooperation zu finden. Die Förderung einer solchen ko-konstruktiv angelegten Zusammenarbeit bildet eine zentrale Zielperspektive in der Bund-Länder-Initiative ‚Schule macht stark‘. Neben dem kohärenten Handeln der Beteiligten und dem systemischen Aufbau von Entwicklungskapazitäten ist die Ko-Konstruktion als intensive wechselseitige Zusammenarbeit ein handlungsleitendes Konzept. Sie zielt darauf ab, die triadische Vernetzung zwischen den Beteiligten sowie einen nachhaltigen Wissenstransfer zu fördern (vgl. Manitiuss & Bremm, 2021).

Die Prozesse einer solchen ko-konstruktiven Zusammenarbeit werden im Beitrag einerseits theoretisch fundiert und andererseits exemplarisch mit Bezug auf die intendierte und teilweise realisierte Praxis der Kooperation in Schule Macht Stark vorgestellt und

reflektiert. Dabei werden folgende Forschungsfragen aufgegriffen: (1) Welches theoretische Verständnis von Ko-Konstruktion liegt der Initiative Schule Macht Stark zugrunde? (2) Welche Ziele werden im Hinblick auf die triadische Zusammenarbeit in der Initiative Schule Macht Stark verfolgt und welche Strukturen wurden daraufhin etabliert? (3) Wie wird die Zusammenarbeit mit der Bildungsverwaltung des beispielhaft ausgewählten Bundeslands Hessen im Programmkontext ko-konstruktiv gestaltet? (4) Welche Gelingensbedingungen für eine triadische Ko-Konstruktion lassen sich auf Basis der Erkenntnisse aus der bisherigen Programmpraxis in Schule Macht Stark ableiten?

Für die theoretische Rahmung werden Ansätze zu Zusammenarbeitsformen (vgl. Gräsel et al., 2006), kooperativen professionellen Lerngemeinschaften (vgl. Werner, 2021), lernenden Organisationen (vgl. Klein, 2018) und kollektiven Lern- und Entwicklungsprozessen (vgl. Hemmings, 2012) herangezogen. Weiterhin liegen den Überlegungen Konzepte wie das Sense-Making (Coburn, 2005), das kooperative Wissensmanagement (vgl. Manitiutius & Bremm, 2021), die evidenzbasierte Entwicklungsprozessgestaltung (vgl. Eiden et al., 2018) und die schnittstellenfokussierte sowie netzwerkbasierte Schul- und Schulsystementwicklung (vgl. Manitiutius et al., 2021) zugrunde.

Als Praxisbeispiel werden vertiefte Einblicke in die Gestaltung der ko-konstruktiv orientierten Zusammenarbeit zwischen Bildungswissenschaft, schulischer Praxis und Bildungsverwaltung im Rahmen der Initiative Schule Macht Stark im Land Hessen gegeben. Zum einen werden die landesspezifischen Rahmenbedingungen vorgestellt, die die Arbeitsgrundlage für die triadische Ko-Konstruktion bilden. Zum anderen werden Arbeits- und Kommunikationsstrukturen nachgezeichnet.

Basierend auf der theoretischen Einordnung und der mithilfe der Case-Study-Methode (Yin, 2009) durchgeführten Analyse der unter der Beteiligung der Akteur:innen der Initiative Schule Macht Stark gewonnenen Daten werden Gelingensbedingungen für ihre ko-konstruktiv angelegte Zusammenarbeit abgeleitet. Dabei wird einer abduktiven Forschungslogik gefolgt (Reichert, 2003), bei der einerseits die theoretischen Konzepte zur Ko-Konstruktion die Grundlage für eine deduktive Analyse der ko-konstruktiven Praktiken in der Initiative Schule Macht Stark bilden und bei der andererseits die Erkenntnisse aus der Fallanalyse zur triadischen Ko-Konstruktion in Hessen, einem induktiven Vorgehen folgend, die theoretische Rahmung des Vorhabens konzeptionell anreichern.

Die Analyse zeigt, dass der Erfolg, eine triadische Ko-Konstruktion zu etablieren, insbesondere von der Offenheit und Bereitschaft aller Beteiligten für Arbeitsprozesse über ihre institutionellen Logiken und Grenzen hinaus abhängig ist. Zudem ist die ko-konstruktive Zusammenarbeit auf allen Ebenen, sowohl innerhalb der Systeme als auch systemübergreifend, aktuell in SchuMaS noch nicht durchgehend als Kooperationsform gemäß der Idealvorstellung etabliert. Die Ko-Konstruktion innerhalb der Triade, und wünschenswerterweise auch über die Triade hinaus, beispielsweise mit Akteur:innen aus den Schulaufsichten und den landesseitigen Unterstützungssystemen, könnte intensiviert werden, indem sich alle Beteiligten in Schule Macht Stark gezielt der Aufgabe der Schnittstellenkommunikation widmen.

S-22: Wohlbefinden und psychische Gesundheit verschiedener Akteure im Bildungsbereich. Messung, Erklärungen, Entwicklungen.

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 145

Chair der Sitzung: Felix Kruse

Chair der Sitzung: Arvid Nagel

Chair der Sitzung: Sarah Bürger

Wohlbefinden und psychische Gesundheit verschiedener Akteure im Bildungsbereich. Messung, Erklärungen, Entwicklungen.

Chair(s): **Felix Kruse** (PH NMS Bern, PHSG), **Arvid Nagel** (PH NMS Bern)

Diskutant*in(nen): **Sarah Bürger** (Goethe Universität Frankfurt)

Wohlbefinden und psychische Gesundheit sind bedeutsame Faktoren der unterschiedlichen Akteur:innen im Bildungsbereich, sowohl als eigenes Bildungsziel und Voraussetzung für den Lernerfolg auf Seiten der Schüler:innen, als auch der beruflichen und persönlichen Entwicklung auf Seiten der Lehrpersonen (z.B. Hascher & Waber, 2021; Morinaj & Hascher, 2022; Waters, 2011). Empirische Studien konnten weiterhin zeigen, dass ein hohes Wohlbefinden bei Schüler:innen etwa mit einer positiv wahrgenommenen Beziehung zu Lehrpersonen sowie einem geringeren Internalisierungsverhalten zusammenhängt (z.B. Kern et al., 2016; Smith et al., 2020). Weiterhin konnte die Bedeutsamkeit des Wohlbefindens von (angehenden) Lehrpersonen für die Unterrichtsqualität sowie dem Verbleib im Beruf oder Studium herausgestellt werden (z.B. Hascher & Waber, 2021; Zhou et al., 2024).

Die Konzeptualisierung und entsprechend auch die Messung des Wohlbefindens ist vor diesem Hintergrund durch eine hohe Heterogenität geprägt, welche die Vergleichbarkeit von Befunden erschwert und die Relevanz einer theoretisch fundierten Messung verdeutlicht (z.B. Marsh et al., 2020).

Das Symposium thematisiert das Wohlbefinden und die psychische Gesundheit verschiedener Akteur:innen im Bildungsbereich, darunter Kinder, Jugendliche und angehende Lehrpersonen. Es wird untersucht, wie das Wohlbefinden und die psychische Gesundheit gemessen und erklärt werden können. Weiterhin wird untersucht, inwiefern Aspekte des schulischen Wohlbefindens von Schüler:innen Prädiktoren für weitere schulbezogene Variablen darstellen. Die vorliegenden Beiträge verdeutlichen die Relevanz einer ganzheitlichen und multidimensionalen Herangehensweise zur Messung und Förderung des Wohlbefindens und der psychischen Gesundheit unterschiedlicher Akteur:innen im Bildungsbereich.

Der erste Beitrag beschäftigt sich mit der Entwicklung und Validierung eines mehrdimensionalen Instruments zur Messung des psychischen Wohlbefindens von Kindern und Jugendlichen. Basierend auf internationalen Klassifikationssystemen psychischer Störungen (ICD-10; DSM-5) wurde ein Instrument entwickelt, das die komplexe Struktur des Wohlbefindens erfasst. In der Studie wurden Daten von 906 Kindern und Jugendlichen analysiert, wobei zusätzlich zur konventionellen Strukturgleichungsmodellierung komplexere Modellierungstechniken angewendet wurden. Die Ergebnisse bestätigen die mehrdimensionale Struktur des Instruments und zeigen eine adäquate Modellanpassung. Korrelationen mit weiteren Konstrukten sind erwartungskonform und werden im Rahmen des Beitrags präsentiert.

In dem zweiten Beitrag werden die reziproken Zusammenhänge zwischen schulischem Wohlbefinden, Engagement und Schulleistungen bei Schweizer Sekundarschüler:innen untersucht. Dabei wurde schulisches Wohlbefinden als ein multidimensionales Konstrukt, bestehend aus positiven und negativen Kognitionen und Emotionen, sowie schulisches Engagement bestehend aus einer kognitiven, emotionalen und verhaltensbezogenen Komponente konzeptualisiert. Zur Analyse des längsschnittlichen Datensatzes wurde ein random intercept cross-lagged panel model spezifiziert, um die Daten von 757 Schüler:innen aus drei deutschsprachigen Kantonen zu analysieren. Die Ergebnisse zeigen entgegen der Erwartung keine Effekte des schulischen Wohlbefindens auf das Engagement sowie des Engagements auf die schulische Leistung. Diese unerwarteten sowie weitere Befunde werden im Symposium diskutiert.

Der dritte Beitrag präsentiert die Entwicklung und Validierung eines Inventars zur Erfassung von Stress und Stressbewältigung im unteren Grundschulalter. Das SSKJ-mini-Inventar, eine Anpassung des bestehenden SSKJ 3-8 R Inventars, wurde speziell für jüngere Kinder entwickelt. In einer Studie mit 895 Grundschüler:innen wurden die psychometrischen Eigenschaften des neuen Instruments untersucht. Die Ergebnisse zeigen eine gute Modellpassung und interne Konsistenzen, was darauf hinweist, dass das SSKJ-mini ein zuverlässiges Instrument zur Messung von Stress und Wohlbefinden bei jüngeren Kindern ist.

Ein vierter Beitrag untersucht die Bedeutung kontextueller und persönlicher Ressourcen für das Wohlbefinden und die Abbruchsintentionen angehender Lehrpersonen. Die Analyse der Daten von 2867 Lehramtsstudierenden zeigt, dass kontextuelle Ressourcen wie die Qualität des Praktikums und die Kohärenz zwischen Praktikum und Hochschule sowie persönliche Ressourcen signifikante Zusammenhänge mit dem Wohlbefinden und den Abbruchsintentionen aufweisen. Die Ergebnisse werden hinsichtlich ihrer Komplementarität zu personenzentrierten Befunden diskutiert und sollen Impulse zur Förderung des Wohlbefindens während des Lehramtsstudiums liefern.

Beiträge des Symposiums

Psychisches Wohlbefinden von Kindern und Jugendlichen im Kontext einer positiven psychischen Gesundheit: Entwicklung und Validierung eines mehrdimensionalen Messinstruments auf Basis internationaler Klassifikationssysteme psychischer Störungen

Felix Kruse¹, Arvid Nagel²

¹PH NMS Bern, PHSG, ²PH NMS Bern

Theoretischer Hintergrund und Fragestellung Der vorliegende Beitrag beschäftigt sich mit der theoretischen Fundierung und empirischen Überprüfung eines mehrdimensionalen Instruments zur Messung des psychischen Wohlbefindens bei Kindern und Jugendlichen. Die Konzeptualisierungen und entsprechend auch die Messungen des Wohlbefindens unterscheiden sich vor diesem Hintergrund stark. So kann etwa im Kindes- und Jugendalter zwischen allgemeinen oder schulbezogenen Aspekten, einer hedonischen, eudaimonischen oder hybriden Perspektive sowie zwischen unidimensionalen und mehrdimensionalen Ansätzen differenziert werden (Hascher, 2004; Hughes & Kendall, 2009; Kern et al., 2016; Scales et al., 2000; Tian et al., 2013). Während kein Konsens über die Definition des subjektiv wahrgenommenen Wohlbefindens existiert, haben Huppert & So (2013) sowie in einer weiteren Ausdifferenzierung Marsh et al. (2020) eine systematische Definition des Wohlbefindens von Erwachsenen als positive psychische Gesundheit auf Basis der internationalen Klassifikationssysteme für psychische Störungen

(ICD-10; DSM-5) vorgeschlagen. Ein bedeutsamer Vorteil dieser Vorgehensweise liegt etwa in dem fehlenden Konsensus der Konzeptualisierung des Wohlbefindens, wohingegen die Symptome in den internationalen Klassifikationssystemen einen weitgehenden Konsens über die Zeit aufweisen. Huppert und So (2013) konnten weiterhin zeigen, dass verschiedene Gruppen bei einer Messung der Lebenszufriedenheit ähnlich abschnitten, während die Profile der Dimensionen des Wohlbefindens sehr unterschiedlich ausfielen, während Marsh et al., (2020) zusätzlich auf Probleme hinsichtlich unidimensionaler Modellierungen des Wohlbefindens aus psychometrischer Perspektive hinweisen (z.B. Tennant et al., 2007) – insgesamt die mehrdimensionale Struktur des Wohlbefindens indiziert.

Der vorliegende Beitrag untersucht in Anlehnung an die Arbeiten von Huppert & So (2013) und Marsh et al. (2020), ein adaptiertes Instrument zur Messung des psychischen Wohlbefindens von Kindern und Jugendlichen auf Basis einer positiven psychischen Gesundheit. Das Ziel des Beitrags besteht in der Überprüfung der mehrdimensionalen faktoriellen Struktur des Messinstruments mittels unterschiedlicher Modellierungstechniken sowie der Assoziationen mit verwandten Konstrukten.

Methodik Der Beitrag kann auf eine Teilstichprobe des ersten Messzeitpunkts der längsschnittlichen DiStressLS-CH Studie zurückgreifen, welche 906 Kinder und Jugendliche im Alter von 10 bis 14 Jahren der Deutschschweiz umfasst. Die Erhebung der Daten erfolgte zwischen März und Mai 2024 mittels Online-Befragung. Auf Basis einer eingehenden Literaturrecherche wurden 16 Dimensionen der positiven psychischen Gesundheit differenziert und empirisch mittels konfirmatorischer Faktorenanalyse (CFA), exploratorischer Strukturgleichungsmodellierung (ESEM) sowie einer Kombination ebendieser mit einer Bifaktor-Modellierung (B-ESEM) überprüft. Die Analysen erfolgten mittels Mplus und R, wobei der MLR-Schätzer sowie das FIML-Verfahren angewandt wurden.

Ergebnisse Die Prüfung der 16-faktoriellen Struktur mittels CFA zeigte eine adäquate Modellanpassung (z.B. RMSEA = 0.03; CFI = 0.95; TLI = 0.94). In einem nächsten Schritt wurden für einen differenzierteren Einblick in die Binnenstruktur der Daten sowie der Realisierung einer realistischeren Modellierung Exploratorische Strukturgleichungsmodelle (ESEM) sowie eine Kombination ebendieser mit Bifaktor-Modellen (B-ESEM) spezifiziert, welche eine exzellente Modellanpassung aufwiesen (z.B. B-ESEM: RMSEA = .02; CFI = .99; TLI = .98). Die Inter-Faktorkorrelationen der einzelnen Subdimensionen liegen allesamt im erwartungskonformen Bereich, gleiches gilt für die relativen und absoluten Zusammenhänge mit den Subdimensionen Depression, Angst und Stress des DASS21-YG (Kruse & Nagel, in preparation). Die Untersuchung der Verbindung mit dem schulbezogenen Wohlbefinden betont weiterhin die Bedeutung desselben für die positive psychische Gesundheit ($|\beta| = .46 - .91$).

Diskussion Das vorliegende Instrument bietet einen holistischen, theoretisch fundierten Ansatz zur Messung der positiven psychischen Gesundheit bzw. des psychischen Wohlbefindens von Kindern- und Jugendlichen. Starke Argumente für die Konstruktvalidität des betreffenden Instruments können demonstriert werden. Künftige Studien sollten auf weitere Validitätsargumente etwa hinsichtlich Messinvarianztestung oder der Untersuchung komplexerer Wirkungszusammenhänge abzielen. Die deskriptiven Befunde, Zusammenhangsanalysen sowie die Prüfung der faktoriellen Struktur sollen detailliert präsentiert, kritisch reflektiert sowie im Kontext schulbezogener Variablen diskutiert werden.

Reziproke Zusammenhänge zwischen schulischem Wohlbefinden, Engagement und Schulleistungen von Schweizer Sekundarschüler:innen

Jakob Schnell, Katja Saxer, Julia Mori, Tina Hascher
Universität Bern

Theoretischer Hintergrund

Die Förderung des Wohlbefindens ist nicht nur von zentraler Bedeutung für die psychische Gesundheit von Schüler:innen, sondern hängt auch mit ihren schulischen Leistungen zusammen, da es zum schulischen Engagement beitragen kann (Datu & King, 2018). In unserer Studie verstehen wir das schulische Wohlbefinden als multidimensionales Konstrukt, welches als Dominanz positiver Kognitionen und Emotionen gegenüber negativen Kognitionen und Emotionen in Bezug auf verschiedene Aspekte der Schule konzeptualisiert wird (Hascher, 2007). Ebenso verstehen wir schulisches Engagement als ein multidimensionales Konstrukt, das als Interesse und Beteiligung an der Schule definiert ist und kognitive, emotionale und verhaltensbezogene Komponenten umfasst (Fredricks et al., 2005). Nach dem Selbstsystemmodell der Motivationsentwicklung (Skinner et al., 2008) stehen das Wohlbefinden und das Engagement in einem wechselseitigen Zusammenhang mit psychologischen Grundbedürfnissen. Die Befriedigung dieser Bedürfnisse führt zu höherem Wohlbefinden und Engagement, was wiederum die Leistung fördert, und umgekehrt. Im schulischen Kontext können positive Erfahrungen das Engagement und die Leistung fördern (Bücker et al., 2018). Gleichzeitig kann schulischer Erfolg zu einem höheren Wohlbefinden führen (Steinmayr et al., 2016). Während die Längsschnittbeziehung zwischen den Dimensionen des Wohlbefindens und der schulischen Leistung bereits untersucht wurde (Morinaj & Hascher, 2022), wurde die Rolle des schulischen Engagements in dieser Beziehung bisher noch nicht ergründet.

Ziele

In der vorliegenden Studie wurde untersucht, wie die Dimensionen des Wohlbefindens mit den Komponenten des Engagements und der Leistung zusammenhängen. Wir erwarteten einen positiven Einfluss des Wohlbefindens auf das Engagement im Zeitverlauf (Hypothese 1). Ausserdem erwarteten wir einen positiven Einfluss des Engagements auf die Leistung (Hypothese 2). Gleichzeitig erwarteten wir einen positiven Einfluss der Leistung auf das Engagement (Hypothese 3) und das Wohlbefinden (Hypothese 3)

Methode

Die Stichprobe besteht aus N = 757 Schüler:innen der Sekundarstufe I (Klassen 7-9) aus drei deutschsprachigen Kantonen in der Schweiz, die an einem Längsschnittprojekt zum Wohlbefinden in der Schule (WESIR; 2021-2025) teilnahmen. Die Teilnehmer füllten jährlich einen Online-Fragebogen aus, der unter anderem ihr Wohlbefinden und ihr Engagement in der Schule erhob. Das schulische Wohlbefinden wurde mit der Skala zum schulischen Wohlbefinden (Hascher, 2007) gemessen, die sechs Dimensionen umfasst: (1) positive Einstellung zur Schule, (2) Freude an der Schule, (3) schulischer Selbstwert, (4) Abwesenheit von Sorgen in der Schule, (5) Abwesenheit von körperlichen Beschwerden in der Schule, (6) Abwesenheit von sozialen Problemen in der Schule. Das schulische Engagement wurde anhand der School Engagement Scale (Fredericks et al., 2005) gemessen welche drei Subskalen umfasst: (1) kognitives Engagement, (2) behaviorales Engagement, (3) emotionales Engagement. Die akademische Leistung wurde anhand des Notendurchschnitts (GPA) der Abschlussnoten am Ende der Schuljahre gemessen.

Um die reziproken Zusammenhänge zwischen dem Wohlbefinden, dem Engagement und der Leistungen zu untersuchen, wurde ein random intercept cross-lagged panel model (RI-CLPM) mit MPlus 8.10 (Muthén & Muthén, 1998-2017) berechnet.

Ergebnisse und Implikationen

Das Modell wies einen akzeptablen fit auf ($\chi^2(6542) = 13588.398$, $p < .001$; CFI = 0.825; RMSEA = 0.034; SRMR = 0.069). Die Analyse zeigte differenzielle Effekte:

Keine Dimension des Wohlbefindens hatte einen signifikanten Einfluss auf das schulische Engagement und keine Komponente des Engagements hatte einen signifikanten Einfluss auf die Schulleistung. Der schulische Selbstwert zu t1 hatte einen negativen Einfluss auf die Schulleistung zu t2 ($\beta = -0.16$, $p = .016$). Die Schulleistung zu t1 hatte negative Einflüsse auf kognitives ($\beta = -0.23$, $p < .001$) und behaviorales Engagement ($\beta = -0.19$, $p = .012$) zu t2. Der schulische Selbstwert zu t1 hatte einen positiven Einfluss auf das emotionale Engagement zu t2 ($\beta = 0.11$, $p = .047$).

Entgegen unseren Erwartungen scheinen Wohlbefinden, Engagement und Leistung kaum zusammenzuhängen. Bessere Schulnoten in der 7. Klasse scheinen sich sogar negativ auf das spätere schulische Engagement auszuwirken. Mögliche Gründe für diese unerwarteten Befunde werden im Symposium diskutiert.

Entwicklung und Validierung eines Inventars zu Stress und Stressbewältigung im unteren Grundschulalter (SSKJ-mini)

Diana Raufelder¹, Juliane Schlesier², Heike Eschenbeck³, Carl-Walter Kohlmann³, Arnold Lohaus⁴

¹Universität Greifswald, ²Universität Oldenburg, ³PH Schwäbisch Gmünd, ⁴Universität Bielefeld

Theoretischer Hintergrund

Das Stresserleben und die Stressbewältigung von Kindern sind eng mit verschiedenen, relevanten Outcomes verknüpft, wie etwa der Resilienz oder dem schulischen Erfolg (Flores-Buils & Andrés-Roqueta, 2023). Bereits im Grundschulalter berichten Kinder von Stress in schulischen und freizeithlichen Settings (Sotardi, 2016), welchen sie über verschiedene Strategien bewältigen (Vierhaus et al., 2016). Doch obwohl ein wesentlicher Teil der Entwicklung angemessener Bewältigungsstrategien insbesondere im unteren Grundschulalter stattfindet, beschränken sich die meisten Studien in dem Themenfeld bislang auf die oberen Jahrgänge der Grundschule bzw. das Sekundarschulalter (Flores-Buils & Andrés-Roqueta, 2023). Dies liegt vermutlich darin begründet, dass es bis dato noch kein Inventar gibt, welches Stressanfälligkeit, Stressbewältigung und Stresssymptomatik/Wohlbefinden im unteren Grundschulalter (SSKJ-mini) umfassend und zielgerichtet erfasst.

Fragestellung

Auf Grundlage des bereits vorhandenen Inventars SSKJ 3-8 R (Lohaus et al., 2018) für Schüler:innen der oberen Grundschuljahrgänge sowie des unteren Sekundarschulalters soll ein Inventar zur Erfassung von Stresserleben und Stressbewältigung im Vor- und unteren Grundschulalter (SSKJ-mini) konstruiert und validiert werden. Im vorliegenden Beitrag wird der Frage nachgegangen, ob und inwiefern die drei Teile des adaptierten SSKJ-mini-Fragebogens zu Stressvulnerabilität, Stressbewältigung und Stresssymptomatik/Wohlbefinden angemessene psychometrische Eigenschaften aufweisen.

Methode

Stichprobe und Ablauf. Insgesamt 895 Grundschüler:innen der ersten und zweiten Klassen (MAlter = 7.13 Jahre, SDAlter = 0.74, 52.1 % Mädchen) von 34 Schulen und 70 Klassen wurden in einem querschnittlichen Studiendesign Anfang 2024 über eine Paper-Pencil-Fragebogen befragt. Da die Grundschüler:innen in den meisten Fällen noch nicht bzw. noch nicht sicher lesen konnten, wurden die Fragen durch die Testleitungen klassen- bzw. kleingruppenweise vorgelesen.

Instrumententwicklung. Da Grundschulkinder der ersten und zweiten Klassen sich in ihren kognitiven Fähigkeiten und ihrer Konzentrationsfähigkeit von Kindern ab Klasse 3 unterscheiden, wurden die Fragen grundlegend überarbeitet, altersgerecht angepasst und der SSKJ-mini-Fragebogen insgesamt deutlich kürzer gehalten als der SSKJ. Der SSKJ-mini-Fragebogen besteht aus drei Teilen: Teil 1 Stressvulnerabilität (acht Items), Teil 2 Stressbewältigung (30 Items zu den fünf Copingstrategien soziale Unterstützung, Problemorientierte Bewältigung, Vermeidung, Palliative Emotionsregulation, Ärgerbezogene Emotionsregulation) und Teil 3 Stresssymptomatik/Wohlbefinden (physisches Wohlbefinden 6 Items, psychisches Wohlbefinden 16 Items). Alle Items des SSKJ-mini wurden jeweils in einem 3- bis 5-stufigen Likert-Häufigkeitsformat (z. B. nie, selten, manchmal, oft, immer) beantwortet.

Datenanalyse. Die Daten wurden mittels MPlus analysiert. Auf Basis der bereits vorhandenen Struktur des SSKJ 3-8 R wurde für alle drei Teile jeweils eine konfirmatorische Faktorenanalyse durchgeführt: Für Teil 1 (Stressvulnerabilität) in einem 1-Faktorenmodell, Teil 2 in einem 5-Faktorenmodell (mit den fünf Strategien als Faktoren) und Teil 3 in einem 2-Faktoren-Modell (psychisch und physisches Wohlbefinden) sowie vergleichend in einem 5-Faktoren-Modell (physisches Wohlbefinden, Freude, Traurigkeit, Ärger, Angst). Items mit Faktorladungen $< .30$ wurden exkludiert. Interne Konsistenzen wurden über Cronbachs α ermittelt. Die Modellgüte wurde über CFI, RMSEA, SRMR und einen χ^2 -Test überprüft. Anschließend wurden die gebildeten Variablen deskriptiv untersucht und Pearson-r-Korrelationskoeffizienten kalkuliert.

Ergebnisse und ihre Bedeutung

Für Teil 1 und 3 zeigte sich jeweils eine gute Modellpassung: 1-Faktor-Modell Teil 1:

$\chi^2(20) = 55.59$; $p < .001$; CFI = 0.95; RMSEA = 0.05; SRMR = .03. Für das 5-Faktoren-Modell Teil 3: $\chi^2(199) = 419.71$; $p < .001$; CFI = 0.95; RMSEA = 0.04; SRMR = .04. Das 5-Faktoren-Modell von Teil 2 zeigte mit jeweils fünf Items in je fünf Faktoren bei Ausschluss von je einem Item mit Faktorladung $< .30$ eine sehr gute Modellpassung: $\chi^2(261) = 555.70$; $p < .001$; CFI = 0.92; RMSEA = 0.04; SRMR = .05. Interne Konsistenzen sind mit einer Ausnahme (Vermeidung: $\alpha = .57$) für alle Skalen aller drei Teile gut mit $.70 \leq \alpha \leq .78$. Unsere Ergebnisse zeigen, dass der adaptierte SSKJ-mini gute psychometrische Eigenschaften aufweist und folglich als bislang einziges Testinstrument zur Ermittlung von Stressvulnerabilität, Stressbewältigung und Stresssymptomatik/Wohlbefinden bei Schüler:innen der ersten und zweiten Klassen eingesetzt werden kann.

Die Bedeutung kontextueller und persönlicher Ressourcen für das Wohlbefinden und die Abbruchsintentionen angehender Lehrpersonen

Manuela Haldimann¹, Rebecca Collie², Tina Hascher¹, Doreen Flick-Holtsch³

¹Universität Bern, ²University of New South Wales, Sydney, ³Universität Zürich

Theoretischer Hintergrund und Fragestellung

In den vergangenen Jahren wurde intensiv zum Wohlbefinden von Lehrpersonen geforscht und dessen Bedeutung für schulische Lernprozesse sowie den Verbleib von Lehrpersonen im Lehrberuf hervorgehoben (Zhou et al., 2024). Bisher wenig Beachtung fand hingegen das Wohlbefinden von angehenden Lehrpersonen, insbesondere unter Berücksichtigung eines kontextspezifischen und mehrdimensionalen Verständnisses. Mit Bezug zum Modell des schulischen Wohlbefindens von Hascher (2004, 2023) und dessen Anwendung auf die Lehrerinnen- und Lehrerbildung (Haldimann et al., 2024) wird in der

vorliegenden Studie das Wohlbefinden von Lehramtsstudierenden mit drei positiven Dimensionen (1) Positive Einstellungen zum Studium, (2) Freude im Studium, (3) Akademischer Selbstwert, sowie zwei negativen Dimensionen (4) Sorgen wegen des Studiums, (5) und Physische Probleme wegen des Studiums konzeptualisiert.

Theoretische Modelle wie das Job Demands-Resources-Modell (Bakker & Demerouti, 2017) und das Academic Demands-Resources-Modell (Martin & Collie, 2022) weisen auf die Bedeutung kontextueller und persönlicher Ressourcen für das Wohlbefinden sowie für mögliche Abbruchsintentionen hin. So zeigte eine personenzentrierte Studie mit angehenden Lehrpersonen, dass deren Wahrnehmung von kontextuellen und persönlichen Ressourcen mit der Zugehörigkeit in adaptiveren Wohlbefindensprofilen einherging. Ebenfalls wiesen angehende Lehrpersonen adaptiver Wohlbefindensprofile geringere Abbruchsintentionen als angehende Lehrpersonen maladaptiver Wohlbefindensprofile auf (Haldimann et al., submitted).

Ziel der vorliegenden Studie ist es, diese vorliegenden Befunde mit variablenzentrierten Ergebnissen zu ergänzen (z.B. durch die Prüfung indirekter Effekte). So wird der Fragestellung nachgegangen, wie zwei kontextuelle Ressourcen (Praktikumsqualität, Kohärenz zwischen Praktikum und Hochschule) und eine persönliche Ressource (Lehrer:innenselbstwirksamkeit) mit dem Wohlbefinden von Lehramtsstudierenden und deren Abbruchsintentionen bezüglich Lehramtsstudium und bezüglich Lehrberuf zusammenhängen. Die Kombination von personen- und variablenzentrierten Ergebnissen soll zu einem tieferen Verständnis und der Entwicklung zielgerichteter Interventionen beitragen (Morin et al., 2017).

Methode

Reanalysiert wurden Daten von 2867 österreichischen Lehramtsstudierenden, die im Frühjahr 2021 an einer Onlinebefragung teilgenommen haben. Die fünf Wohlbefindensdimensionen, die kontextuellen und persönlichen Ressourcen sowie die Abbruchsintention bezüglich Lehramtsstudium wurden je mit einer Skala erfasst und wiesen zufriedenstellende Reliabilitäten auf ($\alpha = .76-.94$). Die Abbruchsintention bezüglich des Lehrberufs wurde mit einem Einzelitem erfasst («Ich habe nicht vor, als Lehrperson zu arbeiten»). Als Kovariaten wurden das Geschlecht, das Alter, das Studienprogramm (Primarstufe vs. Sekundarstufe Allgemeinbildung), das Studienjahr, eine mögliche Teilzeitbeschäftigung als Lehrperson und Betreuungspflichten berücksichtigt. Die Analysen erfolgten mit Mplus (Version 8.11). Bis auf die Prüfung indirekter Effekte wurde der robuste Maximum-Likelihood (MLR) Schätzer verwendet. Fehlende Werte (<3%) wurden mithilfe des Full Information Maximum Likelihood-Verfahrens geschätzt. In einem ersten Schritt wurde eine konfirmatorische Faktorenanalyse (CFA) durchgeführt. Das Einzelitem zur Erfassung der Abbruchsintention bezüglich des Lehrberufs sowie die sechs Kovariaten wurden als manifeste Variablen (mit Ladung auf 1 und Residuen auf 0 fixiert) integriert. In einem zweiten Schritt wurde ein Strukturgleichungsmodell spezifiziert, wobei für die geteilte Varianz sowie die Kovariaten kontrolliert wurde. Zur Prüfung der indirekten Zusammenhänge der Ressourcen mit den Abbruchsintentionen via Wohlbefindensdimensionen wurde ein Bootstrapping-Verfahren mit bias-korrigierten Bootstrap-Konfidenzintervallen angewandt (1000 Ziehungen; MacKinnon et al., 2004).

Ergebnisse und deren Signifikanz

Das Strukturgleichungsmodell wies einen adäquaten Modellfit auf (SEM: $\chi^2(964) = 5062.33$, $p = .001$, RMSEA = .04, CFI = .94). Es zeigten sich direkte Zusammenhänge der kontextuellen und persönlichen Ressourcen mit den fünf Wohlbefindensdimensionen ($|\beta| = .08$ bis $.44$) sowie den beiden Abbruchsintentionen ($|\beta| = .09$ bis $.15$). Positive Einstellungen zum Studium waren direkt und negativ mit den Abbruchsintentionen bezüglich Lehramtsstudium ($\beta = -.31$) und Lehrberuf ($\beta = -.26$) assoziiert. Der akademische Selbstwert wies einen direkten negativen Zusammenhang mit der Abbruchsintention bezüglich Lehramtsstudium auf ($\beta = -.13$). Es zeigten sich auch indirekte Zusammenhänge. Die Ergebnisse werden hinsichtlich ihrer Komplementarität zu den personenzentrierten Befunden diskutiert und sollen Impulse zur Förderung des Wohlbefindens während des Lehramtsstudiums liefern.

S-23: Vielschichtige Perspektiven auf das Wohlbefinden von Schüler*innen: Prädiktoren, Profile und förderliche Faktoren in heterogenen Lernumwelten

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 142

Chair der Sitzung: Alyssa Laureen Grecu

Chair der Sitzung: Katharina Molitor

Chair der Sitzung: Gerda Hagenauer

Vielschichtige Perspektiven auf das Wohlbefinden von Schüler*innen: Prädiktoren, Profile und förderliche Faktoren in heterogenen Lernumwelten

Chair(s): **Alyssa Grecu** (TU Dortmund, Institut für Schulentwicklungsforschung), **Katharina Molitor** (TU Dortmund, Institut für Schulentwicklungsforschung)

Diskutant*in(nen): **Gerda Hagenauer** (Paris Lodron Universität Salzburg, Abteilung Bildungswissenschaft, Schulforschung und Schulpraxis)

Ein großer Teil des Alltags von Kindern und Jugendlichen findet in der Schule statt. Neben der Leistung als Bildungsergebnis hat in den vergangenen Jahren auch das Wohlbefinden von Schüler*innen an Bedeutung gewonnen. Wohlbefinden ist ein mehrdimensionales Konstrukt, welches nach Hascher (2004) eine kognitive, eine affektive, eine physische und eine soziale Komponente enthält. Der Zusammenhang mit der Leistung ist bereits gut erforscht – hier zeigen Metaanalysen reziproke Zusammenhänge (Kaya & Erdem, 2021). Bisherige Studien weisen darauf hin, dass neben der Leistung (Stang-Rabrig et al., 2023) unter anderem die soziale Herkunft (Navarro-Carrillo et al., 2020), das Geschlecht (Titzmann & Jugert 2022) sowie Klassenzusammensetzung und Lehrkraft-Schüler*innenbeziehungen (Hagenauer & Raufelder, 2021) prädiktiv wirken. Häufig wurden dabei einzelne Wohlbefindensdimensionen fokussiert. Daran anschließend nimmt sich dieses Symposium zum Ziel unter Berücksichtigung vielfältiger quantitativer Methoden unterschiedliche Dimensionen des Wohlbefindens zu beleuchten, heterogene Lernumwelten differenziert zu betrachten und Prädiktoren des Wohlbefindens von Lernenden zu identifizieren. Ein weiterer Fokus liegt auf der Identifikation förderlicher Faktoren für die Entwicklung von Wohlbefinden.

Beitrag I "Schulisches Wohlbefinden, Lehrpersonen-Schüler:innen- und Schüler:innen-Schüler:innen-Beziehungen in der Schweiz: Ein multidimensionaler, personenzentrierter Ansatz" fokussiert die emotional-affektive Komponente des Wohlbefindens und identifiziert vier Wohlbefindensprofile (engagiert, desinteressiert, ängstlich und sozial gestresst) von Sekundarschüler*innen (8. Klasse) in deutschsprachigen Schweizer Kantonen. Im zweiten Schritt werden mittels logistischer Regressionsanalysen Lehrkraft-Schülerbeziehungen sowie Schülerbeziehungen untereinander als Prädiktoren für die Zugehörigkeit zu den Profilen untersucht. Positive Lehrkraft-Schülerbeziehungen bedingen überwiegend ein positives Wohlbefindensprofil, wohingegen negative Lehrkraft-Schülerbeziehungen und Schülerbeziehungen untereinander prädiktiv für die Zugehörigkeit zu negativen Wohlbefindensprofilen sind.

Beitrag II "Mittendrin statt nur dabei – Entwicklung der sozialen Zugehörigkeit während der Jugendphase" beleuchtet die soziale Wohlbefindensdimension und untersucht die Entwicklung der sozialen Zugehörigkeit von 12- bis 16-Jährigen in Deutschland mit fixed-effects-Modellen. Die Ergebnisse weisen auf eine Abnahme des sozialen Zugehörigkeitsgefühls im untersuchten Zeitraum hin. Für männliche Schüler und Jugendliche mit sozial hohem Hintergrund fallen die Abnahmen geringer aus als für die entsprechenden Vergleichsgruppen.

Beitrag III "Wohlbefinden von Schüler*innen - eine Längsschnittstudie während der COVID-19 Pandemie" untersucht das schulische Wohlbefinden von Schüler*innen vor und während der COVID-19-Pandemie. Lineare Mixed Modelle ergaben, dass Schüler*innen während der Pandemie weniger körperliche Beschwerden und soziale Probleme, jedoch mehr Sorgen hatten. Der schulische Selbstwert sowie die Freude und Anerkennung stiegen an. Signifikante Unterschiede bezüglich Geschlecht und Migrationshintergrund wurden festgestellt.

Beitrag IV "Das Wohlbefinden von Schüler*innen in differenziellen Lernumwelten: Zur Bedeutung von familiären, individuellen und kontextuellen Faktoren" nimmt vier Wohlbefindensoutcomes (Lebenszufriedenheit, Zufriedenheit mit Freund*innen, akademisches Selbstkonzept und Hilfslosigkeit) in den Blick und analysiert Unterschiede in Abhängigkeit von schulischen Leistungssettings. Mit einer MANOVA werden Schüler*innen ohne Gymnasialempfehlung, die der Übergangsempfehlung gefolgt bzw. nicht gefolgt sind, verglichen. In Klasse 9 zeigen sich im Mittel signifikante Unterschiede im Wohlbefinden. Auch hinsichtlich der Prädiktoren des Wohlbefindens zeigten sich in den Mehrgruppenstrukturgeleichungsmodellen heterogene Befundmuster.

Insgesamt vermittelt das Symposium einen umfassenden Einblick in die Entwicklung des multidimensionalen Wohlbefindens im Jugendalter. Die Beiträge werden anschließend diskutiert und eingeordnet.

Beiträge des Symposiums

Schulisches Wohlbefinden, Lehrpersonen-Schüler:innen- und Schüler:innen-Schüler:innen-Beziehungen in der Schweiz: Ein multidimensionaler, personenzentrierter Ansatz.

Katja Saxer¹, Heta Tuominen², Markku Niemivirta², Tina Hascher¹

¹Universität Bern, Institut für Erziehungswissenschaft, ²University of Eastern Finland, School of Applied Educational Science and Teacher Education

Dem Wohlbefinden von Schüler:innen (StudWB) kommt angesichts des Rückgangs der psychischen Gesundheit weltweit (OECD, 2023), gepaart mit den ökonomischen, ökologischen und gesellschaftlichen Herausforderungen des 21. Jahrhunderts, eine Schlüsselrolle zu.

StudWB, als multidimensionales Konstrukt, kann durch positive Emotionen/Kognitionen (z.B. positive Einstellung zur Schule, Freude in der Schule, schulischer Selbstwert) sowie durch das Fehlen negativer Emotionen/Kognitionen (z.B. Sorgen wegen der Schule, körperliche Beschwerden wegen der Schule, soziale Probleme in der Schule) charakterisiert werden (Hascher, 2004).

Positive soziale Beziehungen im Klassenzimmer können als bedeutende Determinanten erfolgreicher Bildungsprozesse angesehen werden (Hagenauer & Raufelder, 2021) und korrelieren mit hohem StudWB (Lopes Moreira et al., 2021; Roffey, 2015). Die Qualität der Lehrpersonen-Schüler:innen-Beziehungen (LSB) zeichnet sich durch Nähe und geringes Konfliktniveau aus (Koomen & Jellesma, 2015), während Schüler:innen-Schüler:innen-Beziehungen (SSB) durch den berichteten Kohäsionsgrad beschrieben werden können (Diel & Nieder, 2010).

Bisherige Untersuchungen zeigen, dass das StudWB in der Sekundarstufe I aufgrund gestiegener schulischer Anforderungen und altersbedingter Veränderungen abnimmt (Green et al., 2012; Virtanen et al., 2019), wobei negative LSB und SSB diese

Tendenz verstärken können (Baumeister et al., 2001). Personenzentrierte Ansätze ermöglichen die Berücksichtigung der Heterogenität innerhalb einer Population und die Identifikation latenter Untergruppen (Bergman & Andersson, 2010). Erste Untersuchungen zeigen, dass sich Schüler:innen hinsichtlich ihrer Wohlbefindensdimensionen unterscheiden lassen (Busseri et al., 2009; Morinaj & Held, 2023). Längsschnittanalysen zur Stabilität und Veränderung der Wohlbefindensprofile sowie zu den Einflüssen der LSB und SSB bleiben rar.

Dieser Beitrag hat zum Ziel, (1) anhand latenter Profilanalysen (LPA) schulbezogene Wohlbefindensprofile zu identifizieren sowie (2) den Zusammenhang der schulbezogenen Wohlbefindensprofile mit den LSB und SSB zu analysieren.

Die Stichprobe setzt sich aus N = 757 Schüler:innen (47.8% weiblich; Alter = 13.12, SDalter = 0.60) in der 7. Klasse (w1) und N = 720 Schüler:innen (47.6% weiblich; Alter = 13.92, SDalter = 0.81) in der 8. Klasse (w2), aus drei Deutschschweizer Kantonen zusammen. Als Messinstrumente wurden das Student Well-being Questionnaire (Hascher, 2007), die Teacher–Student Relationship Scale von Koomen & Jellesma (2015) sowie die Student–Student Relationship Scale basierend auf dem Hessischen Referenzrahmen Schulqualität (HRS, 2012) verwendet.

Zur Beantwortung der Forschungsfragen wurden latente Profil-Modelle für beide Messzeitpunkte in MPlus version 8.9 (Muthén & Muthén, 1998–2017) geschätzt (Morin et al., 2016). Logistische Regressionsanalysen (R3STEP) wurden verwendet, um den Zusammenhang der Wohlbefindensprofile mit den LSB und SSB zu analysieren (Vermunt, 2010).

Basierend auf den sechs Dimensionen des StudWB liessen sich zu beiden Messzeitpunkten vier Profile des schulischen Wohlbefindens identifizieren: Schüler:innen im engagierten Profil (w1=51%; w2=52%) zeigten das höchste Level an positiven und das tiefste Level an negativen Emotionen/Kognitionen in der Schule. Im Gegensatz dazu zeigten Schüler:innen im desinteressierten Profil (w1=22%; w2=20%) die tiefsten Werte an positiven Emotionen/Kognitionen sowie hohe Sorgen wegen der Schule. Am meisten Sorgen sowie körperliche Beschwerden wegen der Schule zeigten Schüler:innen im ängstlichen Profil (w1=16%; w2=20%), wobei sie dennoch mittlere Werte in den positiven Dimensionen aufwiesen. Schüler:innen im sozial gestressten Profil (w1=11%; w2=8%) zeichneten sich durch hohe negative Emotionen/Kognitionen, insbesondere durch hohe soziale Probleme in der Schule, aus.

Die vorläufigen Resultate der logistischen Regressionsanalysen zeigten, dass nahe LSB positiv mit den positiven Wohlbefindensdimensionen korrelieren, während konfliktreiche LSB die negativen Wohlbefindensdimensionen verstärken. Darüber hinaus können Schüler:innen trotz konfliktreicher LSB hohe Werte in den positiven Wohlbefindensdimensionen aufweisen. Eine negative SSB verstärkt besonders die sozialen Probleme in der Schule. Der Verlauf der Wohlbefindensprofile wird zu einem späteren Zeitpunkt mittels einer Transitionsanalyse (LTA) untersucht.

Die bisherigen Ergebnisse weisen auf eine heterogene Wahrnehmung des schulischen Umfelds durch Schüler:innen hin und verdeutlichen damit die Relevanz eines multidimensionalen, personenzentrierten Ansatzes. Die Wohlbefindensprofile bilden eine Grundlage für die Entwicklung massgeschneiderter Interventionen und unterstützen Lehrpersonen dabei, ihre Beziehungsarbeit gezielt auf die Bedürfnisse verschiedener Schüler:innengruppen auszurichten.

Mittendrin statt nur dabei – Entwicklung der sozialen Zugehörigkeit während der Jugendphase

Jacqueline Kroh, Monja Schmitt, Lena Nusser
Leibniz-Institut für Bildungsverläufe (LIfBi)

Soziale Zugehörigkeit wird als die Integration in „ein Mindestmaß an dauerhaften, positiven und bedeutsamen zwischenmenschlichen Beziehungen“ (Baumeister & Leary, 1995, S. 497 [eigene Übersetzung]) verstanden. Sie ist ein grundlegendes, menschliches Bedürfnis und eine entscheidende Determinante für das individuelle Wohlbefinden (Tomova et al. 2021). Das Gefühl der sozialen Zugehörigkeit führt unter anderem zu mehr gesellschaftlichem Engagement (Albanesi et al. 2007), höheren akademischen Leistungen (Sirin & Rogers-Sirin 2004; Tomova et al. 2021) und besserer mentaler Gesundheit (Van Harmelen et al., 2017). In der Bildungsforschung wurde bisher ein besonderer Fokus auf soziale Zugehörigkeit in der Schule als zentralem Ort der Jugendphase gelegt (Allen et al. 2022). Hierbei zeigen sich sowohl Zusammenhänge mit individuellen Faktoren (z.B. Geschlecht, Titzmann & Jugert 2022; Migrationshintergrund, Booker, 2006; sozioökonomischer Status, Chiu et al. 2015) und kontextuellen Faktoren (Allen et al. 2022). Viele Studien deuten auf einen Rückgang der sozialen Zugehörigkeit in der Schule hin (z.B., Anderman 2002; Ding & Hall 2007). Im Prozess der Adoleszenz und den damit verbundenen Entwicklungsaufgaben der Identitätsbildung spielt soziale Zugehörigkeit auch über den schulischen Kontext hinaus eine zunehmende Rolle (Tomova et al. 2021). Trotz ihrer Bedeutung gibt es weiterhin offene Fragen, wie sich die soziale Zugehörigkeit während der Adoleszenz entwickelt und wie groß soziale Ungleichheiten in dieser entscheidenden Lebensphase sind.

Theoretisch werden in Anlehnung an den Ansatz von Allen et al. (2021) vier zentrale Bedingungsfaktoren der sozialen Zugehörigkeit angenommen: die individuellen Kompetenzen, Möglichkeiten der sozialen Interaktion, die individuelle Motivation sowie die Wahrnehmung. Diese können im Laufe der Adoleszenz Veränderungen unterliegen und somit auch das Gefühl der sozialen Zugehörigkeit beeinflussen. Es ist anzunehmen, dass individuelle, sozial stratifizierende Merkmale zu systematischen Unterschieden in den Ausgangsniveaus und der altersspezifischen Entwicklung führen sollten. Empirische Evidenz hierzu ist jedoch lückenhaft.

Diese Studie adressiert diese Forschungslücke, indem sie die Entwicklung der sozialen Zugehörigkeit im Alter zwischen 12 und 16 Jahren anhand von Daten der Startkohorte 2 (Kindergarten) des Nationalen Bildungspanels (NEPS, Blossfeld & Roßbach, 2019) untersucht. Die Analysestichprobe umfasst 2731 Schüler:innen in der Sekundarstufe (ohne zwischenzeitliche Schulwechsel) zu drei Messzeitpunkten. Die soziale Zugehörigkeit wurde anhand der Definition von Baumeister & Leary (2017) jeweils mit fünf Items erfasst (Beispiel: Mir fällt es leicht, neue Freundschaften zu schließen.; Cronbach's α = 0.68, CFI = 0.952, RMSEA = 0.128). Mit Hilfe von fixed effects-Modellen wird untersucht, wie sich die soziale Zugehörigkeit mit dem Alter verändert und wie individuelle Merkmale – Geschlecht (binär erfasst), sozioökonomischer Hintergrund, Migrationshintergrund und schulische Leistungen – diese Entwicklung moderieren. Erste Ergebnisse zeigen eine leichte, aber signifikante Abnahme des Gefühls der sozialen Zugehörigkeit zwischen 12 und 16 Jahren (b = -0.05). Darüber hinaus weisen Mädchen ein geringeres Maß an sozialer Zugehörigkeit auf, wobei der Rückgang im Vergleich zu ihren männlichen Peers ausgeprägter ist. Ebenso scheinen Jugendliche mit höherem sozialem Hintergrund einen Vorteil bei der sozialen Zugehörigkeit zu haben, wobei die Alterseffekte im Sinne differentieller Entwicklungsverläufe in dieser Gruppe geringer zu sein scheinen. Bemerkenswert darüber hinaus ist, dass zwar Niveauunterschiede in der sozialen Zugehörigkeit je nach akademischer Leistung zu beobachten sind, die Alterseffekte in diesen Gruppen jedoch konsistent bleiben. Zusammenfassend lässt sich sagen, dass soziale Ungleichheiten in Bezug auf die soziale Zugehörigkeit bereits in früheren Lebensphasen bestehen und sich teilweise sogar im Jugendalter verstärken.

Diese Ergebnisse verdeutlichen die Relevanz der frühzeitigen und kontinuierlichen Förderung sozialer Zugehörigkeit als grundlegendes menschliches Bedürfnis mit Bedeutung für das individuelle Wohlbefinden. Sozialräumliche Strukturen zur Integration von Kindern aus benachteiligten Familien im frühen Kindes- und Schulalter sollten verstärkt gefördert werden, um

positive Begegnungen unterschiedlicher Personengruppen zu ermöglichen. Pädagogisches Personal sollte in allen Bildungsbereichen entsprechend geschult werden, um solche Begegnungen zu initiieren und zu begleiten. Dadurch kann sich bei allen Kindern und Heranwachsenden langfristig ein Gefühl der sozialen Zugehörigkeit entwickeln.

Wohlbefinden von Schüler*innen - eine Längsschnittstudie während der COVID-19 Pandemie

Isabell Martin¹, Michaela Gläser-Zikuda¹, Mareike Nowak¹, Lisa Pösse², Ramona Obermeier³, Chengming Zhang¹

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut für Erziehungswissenschaft, ²Justus-Liebig-Universität Gießen, Institut für Kindheits- und Schulpädagogik, ³Universität Vechta, Institut für Erziehungswissenschaft

Theoretischer Hintergrund: Schulisches Wohlbefinden hängt positiv mit Gesundheit, der Wahrnehmung von Schule und Unterricht sowie den Lernprozessen und Leistungen von Schüler*innen zusammen (Hascher, 2004; OECD, 2019), wobei heterogene Befunde für unterschiedliche Schüler:innengruppen vorliegen. Dem Sechs-Komponenten-Modell (Hascher, 2004) folgend umfasst schulisches Wohlbefinden positive Schuleinstellungen, hohen schulischen Selbstwert, Abwesenheit von körperlichen und sozialen Problemen und Sorgen wegen der Schule sowie Freude und Anerkennung in der Schule. Mischformen aus Präsenz- und Distanzunterricht (Goldan et al., 2021; Voss & Wittwer, 2020) und Schulschließungen während der COVID-19 Pandemie stellten eine psychosoziale Herausforderung für Schüler*innen dar, die das Belastungserleben (z.B. vor allem von bildungsbenachteiligten) Schüler:innen erhöhten (Obermeier et al., 2022) und das Wohlbefinden der Schüler:innen reduzierte (OECD, 2023). Allerdings fehlt es bislang an Längsschnittstudien, die sich mit allen einzelnen Komponenten des schulischen Wohlbefindens befassen. Es lässt sich bspw. annehmen, dass die reduzierte Schulpräsenz z.B. soziale Probleme bei Schüler*innen temporär verringerte, andere Dimensionen des schulischen Wohlbefindens (z.B. Sorgen) aufgrund der unsicheren Gesamtsituation jedoch vermehrt auftraten.

Fragestellung: Daher werden folgende Fragestellungen in der vorliegenden Längsschnittstudie, die im Rahmen eines Schulprojekts (Gläser-Zikuda et al., 2024) durchgeführt wurde, fokussiert:

F1: Wie ist das schulische Wohlbefinden von Schüler*innen vor und am Ende der COVID-19 Pandemie ausgeprägt?

F2: Welche Unterschiede des schulischen Wohlbefindens lassen sich hinsichtlich individueller Schüler*innenmerkmale feststellen?

Methode: Insgesamt nahmen N = 185 Schüler*innen der Jahrgangsstufen fünf bis neun an zehn Sekundarschulen (Haupt-/Realschulen) mit relativ hohen Quoten an Schulabbrecher*innen (> 10%) aus einem deutschen Bundesland im Frühjahr 2019 und im Herbst 2021 an einer paper-pencil Befragung teil. Diese 185 Schüler:innen (41.9 % männlich; 9.2 % Migrationshintergrund; MZP1: Alter: M = 11.96 (SD = 1.13), MZP 2: Alter: M = 13.76 (SD = 1.17)) verteilten sich relativ ausgewogen auf zehn Schulen. Das schulische Wohlbefinden der Schüler:innen wurde mit allen sechs Subskalen nach Hascher (2004) erhoben. Die Reliabilitäten der Subskalen liegen zwischen $\alpha = .62 \leq .88$. Angaben zu Migrationshintergrund, Jahrgangsstufe, Geschlecht und Leistung (Mathematik- und Deutschnote im letzten Schuljahr) der Schüler:innen wurden als Kovariaten einbezogen. Um mögliche Veränderungen im schulischen Wohlbefinden zu prüfen, wurden lineare mixed models in R unter Einbezug der zwei Messzeitpunkte berechnet.

Ergebnisse: Für die Dimension Positive Einstellungen zur Schule konnte kein signifikanter Unterschied zwischen den Messzeitpunkten identifiziert werden ($B = 0.76$, $SE = .39$, 95% CI [-0.00, -1.53]). Hinsichtlich der Abwesenheit körperlicher Beschwerden zeigt sich, dass die Schüler*innen während der Pandemie weniger körperliche Beschwerden wahrnehmen ($B = -0.28$, $SE = .08$, 95% CI [-0.44, -0.12]); für Geschlecht und Migrationshintergrund konnten zudem signifikante Effekte bezogen auf die Abwesenheit körperlicher Beschwerden ermittelt werden. Der schulische Selbstwert ist bei den Schüler*innen während der COVID-19 Pandemie höher als zuvor ($B = 0.18$, $SE = .59$, 95% CI [0.07, -0.30]), wobei besonders höhere Leistungen der Schüler:innen zu einem höheren Selbstwert beitragen. Bezüglich der Abwesenheit sozialer Probleme zeigt sich, dass die Schüler*innen während der COVID-19 Pandemie weniger soziale Probleme haben ($B = 0.21$, $SE = .07$, 95% CI [0.07, -0.36]). Auch ein männliches Geschlecht ist mit einer geringeren Ausprägung sozialer Probleme assoziiert. Die Schüler*innen machen sich während der COVID-19 Pandemie mehr Sorgen als vorher ($B = -0.25$, $SE = .08$, 95% CI [-0.40, -0.09]), für das Geschlecht, den Migrationshintergrund und die Leistung konnten signifikante Effekte auf die Ausprägung der Sorgen identifiziert werden. Freude und Anerkennung werden im Vergleich zum ersten Messzeitpunkt während der COVID-19 Pandemie ebenfalls höher eingeschätzt ($B = 0.15$, $SE = .06$, 95% CI [0.03, -0.27]).

Die Studie verdeutlicht, dass der Einfluss von Schulschließungen und Distanzunterricht auf die verschiedenen Dimensionen des schulischen Wohlbefindens sowie individuelle Schüler:innenmerkmale differenziert betrachtet werden sollte. Weitere zentrale Ergebnisse der Studie werden präsentiert sowie Limitationen und praktische Implikationen abschließend diskutiert.

Das Wohlbefinden von Schüler*innen in differenziellen Lernumwelten: Zur Bedeutung von familiären, individuellen und kontextuellen Faktoren

Katharina Molitor, Justine Stang-Rabrig, Nele McElvany
TU Dortmund, Institut für Schulentwicklungsforschung

Wohlbefinden ist neben leistungsbezogenen Outcomes ein weiteres, zentrales Bildungsziel, welches mit der Leistung zusammenhängt (Kaya & Erdem, 2021). Das multidimensionale Konstrukt konstituiert sich aus einer psychischen, physischen und sozialen Dimension, wobei die psychische in eine kognitive und eine affektive Facette differenziert werden kann (z. B. Diener, 1984; Hascher, 2004). Neben familiären (z. B. soziale Herkunft [SES]) (Navarro-Carrillo et al., 2020) und personalen Faktoren (z. B. Leistung) (Bücker et al., 2018) sind auch kontextuelle Faktoren (z. B. Lernumgebung) (Praharso et al., 2017) bedeutsame Determinanten des Wohlbefindens. Diese werden bei Sekundarstufenschüler*innen, die sich in einer wichtigen Entwicklungsphase befinden, wesentlich durch die von Lehrkräften am Ende der Grundschulzeit ausgesprochene Übergangsempfehlung mitbestimmt: Die Übergangsempfehlung, die durch primäre und sekundäre Herkunftseffekte (Boudon, 1974) sozial verzerrt sein kann (Dumont et al., 2019; Neugebauer, 2010), bietet einen wichtigen Indikator, welche Sekundarschulform Schüler*innen besuchen sollten – Abweichungen von der Übergangsempfehlung sind jedoch möglich. Erste Studien deuten darauf hin, dass nach oben von der Empfehlung abweichende Schüler*innen (bspw. Gymnasiast*innen mit Haupt- oder Realschulempfehlung) im Mittel im Vergleich zu ihren Mitschüler*innen mit passender Übergangsempfehlung schlechtere Noten und Kompetenzen und ein höheres Wohlbefinden aufweisen (Bittmann, 2021; Pfost et al., 2018). Dies kann beispielsweise auf den Big-Fish-Little-Pond Effekt zurückgeführt werden kann (Marsh, 1987). Unklar ist allerdings aus der Perspektive der Wirkung differenzieller Lernumwelten betrachtet, wie sich Schüler*innen ohne Gymnasialempfehlung, die 1) der Übergangsempfehlung auf eine nicht-gymnasiale Schulform gefolgt sind, und 2) die von der Übergangsempfehlung nach oben mit einem Gymnasialbesuch abgewichen sind, in ihrem Wohlbefinden (allgemeine Lebenszufriedenheit, Zufriedenheit mit Freund*innen, akademisches Selbstkonzept und Hilfflosigkeit im Fach Deutsch) in Klasse 9 voneinander unterscheiden

(Forschungsfrage 1) und welche Prädiktoren (soziale Herkunft, Wert von Bildung, wahrgenommene elterliche Aspirationen, Erschöpfung) für die Förderung des Wohlbefindens relevant sind (Forschungsfrage 2) und ob es hier Mediationseffekte gibt (Forschungsfrage 3).

Die Forschungsfragen wurden mit den Daten des Nationalen Bildungspanels (NEPS, NEPS Network, 2023) der Startkohorte 3 beantwortet. Ein 1:1 Propensity Score Matching mit Zurücklegen wurde durchgeführt, um zu gewährleisten, dass beide Schüler*innengruppen (nNicht-Gymnasiast*innen: 726, MAlter=10,5 Jahre, SD=0,84; 47% weiblich; nGymnasiast*innen: 92, MAlter=10,4 Jahre, SD=0,86; 45% weiblich) hinsichtlich ihrer Leistung vergleichbar sind. Reliabilitäten der Skalen waren gut.

Die Ergebnisse der MANOVA zu Forschungsfrage 1 zeigen ein höheres akademisches Selbstkonzept bei Nicht-Gymnasiast*innen – der Besuch einer höheren Schulform als empfohlen ist also bei Übergängen ohne Gymnasialempfehlung nicht positiv mit dem akademischen Selbstkonzept korreliert. Hinsichtlich der weiteren Wohlbefindensmaße wurden keine Gruppenunterschiede gefunden ($F(6, 518)=53.7, p<.05, \text{Wilks's } \Lambda=.97$).

Die Mehrgruppenstrukturgleichungsmodelle zur Beantwortung der weiteren Forschungsfragen wurden schrittweise spezifiziert. Modellfitwerte waren gut ($\text{RMSEA}\leq.05; \text{CFI}\geq.91$). Bei der Betrachtung von SES und Wert von Bildung als Prädiktoren zeigten sich Gruppenunterschiede der Prädiktionskraft für den Wert von Bildung und die Lebenszufriedenheit ($\Delta\chi^2=4.0, \Delta\text{df}=1; p<.05$) bzw. das akademische Selbstkonzept ($\Delta\chi^2=36.8, \Delta\text{df}=1; p<.00$) – beide Outcomes wurden für Gymnasiast*innen negativ vom Wert der Bildung vorhergesagt, für Nicht-Gymnasiast*innen fand sich kein signifikanter Zusammenhang. Für die wahrgenommenen elterlichen Aspirationen und die Erschöpfung wurden unterschiedliche Zusammenhänge mit dem akademischen Selbstkonzept und der Hilflosigkeit in Deutsch in beiden Gruppen festgestellt. Dabei sagten die Aspirationen das akademische Selbstkonzept ($\Delta\chi^2=9.2, \Delta\text{df}=2; p<.05$) für Nicht-Gymnasiast*innen positiv und die Hilflosigkeit ($\Delta\chi^2=7.1, \Delta\text{df}=2; p<.05$) für Gymnasiast*innen negativ und für Nicht-Gymnasiast*innen positiv vorher. Unter Berücksichtigung der elterlichen Aspirationen und der Erschöpfung wurde der oben berichtete Zusammenhang zwischen dem Wert von Bildung und dem akademischen Selbstkonzept bei Gymnasiast*innen nicht mehr gefunden. Dies ist auf eine Mediation durch die Erschöpfung zurückzuführen.

Für die Praxis bedeuten die Ergebnisse, dass für die Stärkung des Wohlbefindens differenziert verschiedene Schüler*innengruppen in den Blick genommen werden müssen, da sich für unterschiedliche Gruppen verschiedene Prädiktoren als relevant erwiesen haben.

S-24: Gestaltung und Wirksamkeit früher Sprachförderinterventionen

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 135

Chair der Sitzung: Leonie Dargiewicz

Chair der Sitzung: Svenja Wehrhöfer

Chair der Sitzung: Annika Ohle-Peters

Chair der Sitzung: Gisela Kammermeyer

Gestaltung und Wirksamkeit früher Sprachförderinterventionen

Chair(s): **Leonie Dargiewicz** (TU Dortmund, Institut für Schulentwicklungsforschung, Deutschland), **Svenja Wehrhöfer** (TU Dortmund, Institut für Schulentwicklungsforschung), **Annika Ohle-Peters** (TU Dortmund, Institut für Schulentwicklungsforschung)

Diskutant*in(nen): **Gisela Kammermeyer** (Universität Koblenz-Landau)

Bildungserfolg und gesellschaftliche Teilhabe sind eng mit Sprachkompetenzen verknüpft (Dale et al., 2023). Schon früh lassen sich Unterschiede hinsichtlich verschiedener Sprachkompetenzbereiche, wie z. B. des Wortschatzes und der Grammatikkompetenzen, zwischen Kindern feststellen. Diese können beispielsweise durch einen geringen sozioökonomischen Status (SES), eine andere Muttersprache als die Mehrheitssprache oder eine Sprachentwicklungsstörung bedingt werden (z. B. Chan et al., 2023; Krok & Leonard, 2015; Levine et al., 2020). Kinder, auf die diese Faktoren zutreffen, können als *at-risk* beschrieben werden (z. B. Marulis & Neuman, 2013). Um sich vergrößernden Sprachkompetenzunterschieden (Passaretta et al., 2022) entgegenzuwirken bedarf es einer frühen systematischen Sprachförderung. Im Rahmen dieses Symposiums wird der Frage nach der Effektivität früher Sprachfördermaßnahmen für Kinder *at-risk* im Kontext institutioneller Bildung nachgegangen (Beiträge 1-3) sowie Herausforderungen der Wirksamkeitsmessung untersucht (Beitrag 4).

Die Studie StrateGe im ersten Beitrag fokussiert eine in der Praxis eingesetzte, strategiegeleitete Genusförderung von Zweitklässler*innen mit sonderpädagogischen Förderbedarf im Bereich Sprache. In einem Kontrollgruppendesign mit Prä-Post-Follow-up-Messung wurde die Wirksamkeit der Kleingruppenförderung untersucht.

Einen übergreifenden Blick über die Wirksamkeit von Interventionsstudien liefert die Studie EffectLi im zweiten Beitrag. Der Fokus liegt auf der Sprachförderung mehrsprachiger Kinder im Elementarbereich hinsichtlich der Umgebungssprache und der Familiensprache der Kinder. Die Ergebnisse zeigen u. a. einen mittleren Effekt der vorschulischen Sprachförderung auf die Sprachkompetenzen der Kinder in der Umgebungssprache.

Der dritte Beitrag stellt ebenfalls Ergebnisse einer Metaanalyse vor, die in einem engeren Rahmen die Wirksamkeit von Wortschatzinterventionen im Elementar- und frühen Primarbereich untersucht. Dabei wird ein besonderer Fokus auf Kinder mit verschiedenen Risikofaktoren (geringer SES, Zweitsprachenlernende, Migrationshintergrund, sprachliche Beeinträchtigung, geringe schulische Leistung) gelegt. Erste deskriptive Ergebnisse zeigen, dass verschiedene Risikofaktoren in den Studien berücksichtigt wurden und die Umsetzung von Sprachförderinterventionen sehr heterogen ist.

Nachdem die Wirksamkeit von Sprachinterventionen in den ersten drei Beiträgen beispielhaft und übergreifend beschrieben wurde, wird im vierten Beitrag anhand des Forschungsprojektes TRIO u. a. untersucht, warum der Effektnachweis von Sprachfördermaßnahmen in verschiedenen sprachlichen Bereichen unterschiedlich gut gelingt. In einem Prä-Post-Kontrollgruppendesign wurde die Wirksamkeit einer Fortbildungsmaßnahme für Fachkräfte hinsichtlich der Entwicklung ihrer Sprachförderkompetenzen sowie der Sprachentwicklung der von ihnen geförderten mehrsprachiger Kinder evaluiert. Auf Grundlage der Ergebnisse werden die Herausforderungen, die mit der Wirksamkeitsmessung von Sprachförderungen einhergehen, in diesem Beitrag diskutiert.

Die Beiträge dieses Symposiums liefern relevante Erkenntnisse im Hinblick auf die Wirksamkeit von systematischen Sprachfördermaßnahmen für Kinder *at-risk* vor und zu Beginn der Grundschulzeit. Implikationen für die Praxis und die empirische Bildungsforschung werden im Rahmen des Symposiums diskutiert.

Beiträge des Symposiums

Strategiebasierte Genusförderung für Grundschüler:innen mit sprachlichem Unterstützungsbedarf

Tanja Ulrich¹, Maria Lenzen²

¹Universität Duisburg-Essen, Institut für Sonderpädagogik, Pädagogik und Didaktik im Förderschwerpunkt Sprache, ²Universität zu Köln, Department Heilpädagogik und Rehabilitation

Theoretischer Hintergrund

Schüler:innen mit sonderpädagogischem Unterstützungsbedarf im Bereich Sprache haben typischerweise Schwierigkeiten im Erwerb des Genusystems, die sich unter anderem durch die Verwendung eines unpassenden Artikels vor Nomen (z.B. „Die Ball ist weg.“), unbestimmte Platzhalterformen („De Ball ist weg.“) oder das Auslassen des Artikels („_Ball ist weg.“) zeigen (Clahsen, 1988; Ruberg, 2021; Lenzen, Riehemann & Ulrich, i.V.). Auch für viele Kinder mit einem mehrsprachigen Erwerbsprofil scheint der Erwerb des deutschen Genusystems eine besondere Herausforderung darzustellen (Turgay, 2010; Montanari, 2014; Ruberg, 2015). Zwar sehen Lehrkräfte die unterrichtliche Unterstützung des korrekten Artikelgebrauchs ihrer Schüler:innen einerseits als zentrale Aufgabe an, andererseits existieren keine ausreichend evaluierten Förderkonzepte, um diese Unterstützung des Genuserwerbs evidenzbasiert leisten zu können (Lenzen, Riehemann & Ulrich, 2021). Das Förderkonzept StrateGe - Strategien zum Genuslernen (Riehemann, 2021) zielt darauf ab, bei den Schüler:innen selbstregulierte Lernprozesse zu aktivieren und sie mithilfe von Sprachlernstrategien langfristig im Genuserwerb zu unterstützen. Erste Pilotstudien konnten die Durchführbarkeit des Konzepts im sprachtherapeutischen ebenso wie im unterrichtlichen Setting an inklusiven Schulen sowie Förderschulen zeigen (Lenzen, Riehemann & Ulrich, 2021). Daran anschließend wurde die Effektivität des Förderkonzepts im Rahmen einer kontrollierten Interventionsstudie an Förderschulen mit dem Förderschwerpunkt Sprache untersucht.

Fragestellung

Die Evaluationsstudie untersucht, inwiefern sich durch eine Kleingruppenförderung mit dem Förderkonzept StrateGe die Genusfähigkeiten von Schüler:innen mit sprachlichem Unterstützungsbedarf im Vergleich zur Kontrollgruppe verbessern.

Methode

Es wird ein kontrolliertes Gruppendesign im Prä-Posttest-Design mit Follow-up-Messung angewendet. Die Stichprobe umfasst N = 80 Schüler:innen an Schulen mit dem Förderschwerpunkt Sprache im zweiten Schulbesuchsjahr. Die Hälfte der Schüler:innen sind der Experimentalgruppe zugeteilt und werden über vier Wochen hinweg zweimal pro Woche mit der StrateGe-Intervention

im Kleingruppensetting (zu je 4-6 Schüler:innen) gefördert. Zeitgleich werden die Schüler:innen der Kontrollgruppe regulär von ihren Lehrkräften gefördert (treatment as usual). Alle Proband:innen weisen im Subtest „Genus“ des standardisierten und normierten Grammatiktests ESGRAF 4-8 (Motsch & Rietz, 2019) unterdurchschnittliche Ergebnisse auf. Ausschlusskriterien für eine Teilnahme sind das Vorliegen einer Autismus-Spektrum-Störung oder eines selektiven Mutismus. Die Genusfähigkeiten der Teilnehmenden werden direkt vor und nach der Intervention (Prä- und Posttest) sowie circa fünf Monate nach der Intervention (im Follow-up Test) gemessen. Das hierzu selbst entwickelte Verfahren erhebt differenziert unterschiedliche Aspekte von Genuskompetenz über die Markierung des bestimmten Artikels an Nomina (Lenzen et al., i.V.). Um zwischen Übungs- und Transfereffekten zu differenzieren, werden einige Items während der Intervention geübt, während andere ungeübt bleiben. Langfristige Generalisierungseffekte auf den Genuserwerb werden in der Follow-up-Messung mithilfe des Genussubtests der standardisierten und normierten ESGRAF 4-8 erhoben (Motsch & Rietz, 2019).

Ergebnisse

Die Genusfähigkeiten zu den verschiedenen Messzeitpunkten werden in Hinblick auf Artikelauslassungen, das Artikelforminventar, die Benennkonsistenz und die Korrektheit der Artikel deskriptiv ausgewertet und interferenzstatistisch analysiert. Die Ergebnisse zeigen die Veränderungen der Genusfähigkeiten innerhalb der Gruppen zu den verschiedenen Messzeitpunkten sowie die Unterschiede zwischen der Experimental- und der Kontrollgruppe.

Fazit

Schwierigkeiten im Genuserwerb können sich negativ auf den Lern- und Bildungserfolg von Schüler:innen auswirken. Um Bildungschancen für Schüler:innen mit sprachlichem Unterstützungsbedarf in dieser Hinsicht zu erhöhen, bedarf es einer individualisierten und nachhaltigen Genusförderung. Die Evaluation des Förderkonzeptes StrateGe bietet Erkenntnisse darüber, inwiefern eine selbstregulierte, strategiebasierte Unterstützung des Genuserwerbs bei Grundschüler:innen effektiv sein kann.

Wie wirksam ist Sprachförderung in der Umgebungssprache sowie zur Stärkung der Familiensprache im Elementarbereich - Eine internationale Metaanalyse

Franziska Egert¹, Katarina Groth², Steffi Sachse³

¹Katholische Stiftungshochschule München, ²Staatsinstitut für Frühpädagogik und Medienkompetenz, ³Pädagogische Hochschule Heidelberg, Institut für Psychologie

In den meisten OECD-Ländern besuchen 86% der Kinder im Alter von 3 bis 5 Jahren eine Kindertageseinrichtung (OECD, 2024). Innerhalb des Angebots der frühen Bildung befinden sich eine große und wachsende Gruppe von Kindern, die mit einer anderen Familiensprache aufwächst als die, die in der Gesellschaft und in der Schule gesprochen wird (Passel, 2011). Laut dem aktuellen Bildungsbericht lag in Deutschland 2023 der Anteil an Kindern mit nichtdeutscher Familiensprache in Kindertageseinrichtungen bei 25% (Autor:innengruppe Bildungsberichterstattung, 2024). Diese mehrsprachig aufwachsenden Kinder liegen in ihrer Sprachentwicklung in der Umgebungssprache oftmals hinter ihren monolingualen Peers (Ballantyne, Sanderman, & McLaughlin, 2008). Dies steht im Zusammenhang mit einem qualitativ und quantitativ unterschiedlichen und meist nicht optimalen Input sowie sozioökonomischen Merkmalen (Dubowy, Ebert, von Maurice, & Weinert, 2008; Reardon & Galindo, 2006). Jedoch zeigen viele Studien, dass die Sprachkompetenzen in der Umgebungssprache zu den bedeutendsten Prädiktoren für schulischen Erfolg gehören (Weinert, 2008; Viesel-Nordmeyer et al. 2019). Die Förderung der sprachlichen Fähigkeiten im Vorschulalter ist somit eine zentrale Aufgabe von Kindertageseinrichtungen. Aus dem Grund werden unterschiedliche Sprachförderprogramme und -ansätze (u.a. alltagsintegriert, additive Kleingruppenförderung, Immersionsprogramme) im Elementarbereich mit mehr oder weniger Erfolg eingesetzt bzw. praktiziert (Buysse, Peisner-Feinberg, Paez, & Scheffner, 2014; Duran, Hartzheim, Lund et al., 2016; Blatter, Groth, & Hasselhorn, 2020). Um diese globale Herausforderung zu meistern, ist es wichtig die Wirksamkeit dieser Sprachfördermaßnahmen zu evaluieren und evidenzbasierte Entscheidungen zu treffen.

Die vorliegende Metaanalyse (EffectLi) untersucht den international verfügbaren Forschungsstand zur Sprachförderung für mehrsprachige Kinder im Elementarbereich. Dabei sind folgende Forschungsfragen handlungsweisend: Wie wirksam ist vorschulische Sprachförderung für mehrsprachige Kinder (1) in der Umgebungssprache sowie (2) in der Familiensprache. Wird die Effektivität von (a) der implementierenden Person, (b) dem Sprachförderansatz oder (c) methodologischen Faktoren moderiert.

Alle (quasi-)experimentellen Studien mit Prä- und Posttestmessung sowie mindestens 10 Probanden pro Interventions- oder Vergleichskondition wurden in das Review eingeschlossen. Die elektronische Suche der englisch und deutschsprachigen Publikationen in den Datenbanken PsynDEX, PsycINFO, ERIC, ProQUEST-D&T, WISO und FIS Bildung sowie eine manuelle Suche für graue Literatur aus dem Zeitraum 1970-2020 ergaben 2603 Treffer. Nach dem ersten Screening anhand vordefinierter Selektionskriterien, wurden zunächst 259 Treffer für die Volltextanalyse in Betracht gezogen. Nach dem erweiterten Screening zur Studienqualität blieben noch 44 Studien mit 52 verschiedenen Sprachfördermaßnahmen für die Metaanalyse übrig.

Der Codierprozess erfolgte durch zwei unabhängige Reviewer*innen. Nichtübereinstimmungen wurden durch eine erneute gemeinsame Codierung gelöst und ein finaler Code festgelegt. Die Interrater-Reliabilität für das Qualitätsscreening sowie die inhaltlichen Codes zu den Sprachförderprogrammen waren sehr zufriedenstellend. Der aggregierte Effekt wurde über ein random-effects-model, bei dem die Effektstärken unter die Sprachfördermaßnahmen geschichtet wurden, mit der Software Comprehensive Meta-Analysis V3 (Borenstein, Hedges, Higgins & Rothstein, 2014) berechnet.

Eine Publikationsverzerrung kann aus den Ergebnissen der Egger-Regression (Intercept=0.34; SE=0.76; p=.662) ausgeschlossen werden. Die aggregierte Wirksamkeit für die Förderung der Familiensprache der Kinder ergab $g=0.45$ (SE=0.10; $k=19$). Effekte auf produktive und rezeptive Fähigkeiten waren annähernd gleich hoch ($g=0.38$ vs. 0.45). Die größte Wirkung im mittleren Effektniveau erzielten bilinguale Immersionsprogramme ($g=0.63$) und ein Programm, bei dem frühpädagogische Fachkräfte und externe Fachkräfte zusammenwirkten. Die vorschulische Förderung der Umgebungssprache erzielte einen mittleren Effekt von $g=0.34$ (SE=0.07; $k=50$). Es fanden sich signifikante Unterschiede zwischen den Ländern (wobei Deutschland mit $g=0.05$ am niedrigsten abschnitt) und den Sprachbereichen (Produktion $g=0.16$ vs. Verständnis $g=0.40$). Es gab keine Überlegenheit eines spezifischen Sprachförderansatzes bzw. -programms oder Unterschiede zwischen den implementierenden Personen.

Der Blick auf die Wirksamkeitsbefunde macht es möglich Schlüsselemente der Sprachförderprogramme zu identifizieren, damit diese in der pädagogischen Praxis mehr beachtet und umgesetzt werden. Aufgrund der ernüchternden Befundlage muss anhand struktureller und pädagogischer Elemente kritisch diskutiert werden, warum es Ansätze aus Deutschland nur bedingt schaffen, ihr Wirkungspotential zu entfalten.

Wirksamkeit von Wortschatzinterventionen bei Kindern at-risk: Eine Meta-Analyse

Sprachkompetenzen stellen eine zentrale Voraussetzung für den Bildungserfolg und die gesellschaftliche Teilhabe dar (z. B. Dale et al., 2023). Wortschatz ist allen Teilbereichen von Sprache (Lesen, Schreiben, Sprechen, Hören) inhärent und folglich eine zentrale Dimension der Sprachkompetenz (Jude, 2008). Er ist ein wichtiger Prädiktor für den Lernerfolg in allen Schulfächern (Dale et al., 2023) und für weitere Sprachkompetenzen, insbesondere die Lesekompetenz (z. B. Li & Kirby, 2015). Bereits zu Beginn der Grundschule zeigen sich Unterschiede im Wortschatz zwischen so genannten Kindern at-risk, die z. B. einen geringen sozioökonomischen Status (SES; z. B. Levine et al., 2020), eine andere Muttersprache als die Mehrheitssprache (Chan et al., 2023) oder eine sprachliche Beeinträchtigung (Bergelson et al., 2023) haben, im Vergleich zu ihren Peers. Diese Unterschiede bleiben auch während der Grundschulzeit bestehen (z. B. Heppt & Stanat, 2020) bzw. nehmen sogar zu (Passaretta et al., 2022). Unter Berücksichtigung empirischer Evidenz, dass explizite Wortschatzförderung bisher nur selten in den Regelunterricht implementiert wird (Dwyer & Harbaugh, 2020), ergibt sich die Relevanz systematischer Wortschatzinterventionen.

Wortschatzförderung kann sowohl implizit als auch explizit erfolgen. Bei impliziten Methoden wird die Wortbedeutung aus dem Kontext erschlossen (Sternberg & Powell, 1983), wohingegen bei expliziten Methoden Wortbedeutungen konkret benannt werden (National Institute of Child Health and Human Development, 2000). Als implizite Methode wird häufig das storybook reading eingesetzt (Lowman et al., 2018), während bei expliziten Methoden neben der Nennung der Wortdefinition auch weitere Aktivitäten, wie das Zeigen von Bildkarten, genutzt werden (Heller et al., 2019). Meta-Analysen konstatierten, dass die Kombination beider Methoden besonders wirksam ist (Marulis & Neuman, 2010, 2013).

So zeigte die im Jahr 2013 erschienene Meta-Analyse zur Wirksamkeit von Wortschatzinterventionen bei Kindern at-risk über alle Studien hinweg eine große positive Effektstärke (Marulis & Neuman, 2013). Seitdem erschien eine Reihe weiterer Interventionsstudien zur Wortschatzförderung. Zusätzlich hat der gesellschaftliche Fortschritt, wie die voranschreitende Digitalisierung (Baker & Gowda, 2018), neue Möglichkeiten (z. B. individuelles Lernen; Hammerschmidt-Snidarich et al., 2019) gebracht, Wortschatzinterventionen zu gestalten. Der vorliegende Beitrag knüpft an die Meta-Analyse von Marulis und Neuman (2013) an und betrachtet das Forschungsanliegen unter Berücksichtigung aktueller Studien. Vor diesem Hintergrund sollen folgende Forschungsfragen beantwortet werden:

1. Wie effektiv sind Wortschatzinterventionen für Schüler*innen at risk zu Beginn der formalen Bildung?
2. Moderiert die Art und die Anzahl der Risikofaktoren (geringer SES, andere Muttersprache als Mehrheitssprache, sprachliche Beeinträchtigung) die Effektivität von Wortschatzinterventionen?
3. Gibt es Moderationseffekte hinsichtlich (a) methodischer, (b) pädagogischer oder (c) didaktischer Merkmale auf die Effektivität von Wortschatzinterventionen?

Unter Berücksichtigung der PRISMA-Richtlinien (Page et al., 2021) wurde eine systematische Literatursuche in einschlägigen Datenbanken (Academic Search Ultimate, APA PsychArticles, APA PsychInfo, Education Source, ERIC, PsychIndex, Web of Science) durchgeführt. Zusätzlich erfolgte eine manuelle Literatursuche. Es wurden Interventionsstudien mit einem (quasi-)experimentellen Design berücksichtigt, die Kinder zwischen 4;0 bis 7;11 Jahren in die Untersuchung einschlossen, von denen mindestens 90% der jeweiligen Stichprobe einen oder mehrere der Risikofaktoren aufwiesen. Von den anfangs $k = 2386$ identifizierten Studien konnten nach den Selektionsschritten $k = 36$ für die Analysen inkludiert werden ($N = 6701$ Kinder).

Erste deskriptive Ergebnisse zeigen, dass unterschiedliche Risikofaktoren in den Studien berücksichtigt wurden, wie z. B. Schüler*innen mit einem geringen SES (47%), einer anderen Muttersprache als der Mehrheitssprache (33%) sowie einer Zugehörigkeit zu einer ethnischen Minorität (31%). In 31% der Studien wurden Lernende mit mehr als einem Risikofaktor gefördert. Besonders häufig verwendete Aktivitäten waren storybook reading (86%) mit anschließender Diskussion (58%), der Einsatz von Bildkarten (61%) und die Benennung von Wortdefinitionen (61%). Analysen zu der Wirksamkeit der Wortschatzinterventionen, Unterschieden nach Risikofaktoren und Moderationseffekten werden aktuell durchgeführt und die Ergebnisse im Rahmen des Vortrags präsentiert. Abschließend werden Implikationen für die Forschung und Praxis vorgestellt und diskutiert.

Heißt „nicht nachweisbar“ auch „nicht wirksam“? Über die Herausforderungen, die Wirksamkeit früher Sprachförderung zu messen

Barbara Voet Cornelli, Rabea Lemmer, Johanna Campean, Petra Schulz

Goethe-Universität, Institut für Psycholinguistik und Didaktik der deutschen Sprache

Sprachfördermaßnahmen gelten als wirksam, wenn die geförderten Kinder nachweislich größere Erwerbsfortschritte machen als Kinder in einem ähnlichen Spracherwerbskontext, die lediglich unspezifische sprachliche Anregungen erhalten; diese Effekte sollten noch nach einem längeren Zeitraum nachweisbar sein (Hager, 2000). Wie die Wirksamkeit von Sprachförderung erfasst werden sollte, ist noch ungeklärt (Egert, 2024). Untersuchungen setzen häufig auf einer von mehreren Ebenen an:

- 1) Kindliche Sprachfähigkeiten. Untersucht wird, welche Effekte Sprachfördermaßnahmen auf die Sprachentwicklung von geförderten Kindern im Vergleich zu einer Kontrollgruppe zeigen, bei Verwendung verschiedener Sprachförderprogramme (Groth et al., 2027; Voltmer et al., 2021; Polotzek et al., 2009; Lemmer et al., 2019), bei alltagsintegrierter vs. additiver Förderung (Gagarina et al., 2018) und in unterschiedlichen Förderbereichen (Grammatik, Wortschatz).
- 2) Sprachförderliches Handeln in der Praxis. Es wird angenommen, dass Sprachförderkompetenzen eine Voraussetzung für fundiertes Handeln in konkreten Sprachfördersituationen sind. Untersucht wird daher das Handeln von Sprachförderkräften in konkreten Sprachfördersituationen vor und nach einer Fortbildung (Böse & Scherger, 2023; Geyer & Müller, 2021).
- 3) Professionalisierung der Fachkräfte. Es wird angenommen, dass Kompetenzen von Sprachförderkräften die Wirksamkeit von Sprachförderung beeinflussen. Gemessen wird daher das Wissen vor und nach einer Fortbildung (z.B. Ruberg & Rothweiler, 2018).

Nur wenige Studien stellen bisher einen Zusammenhang zwischen den verschiedenen Untersuchungsebenen her, also etwa der Teilnahme an Fortbildungen und der Sprachentwicklung der Kinder (Lemmer et al., 2019) bzw. der Verwendung von Sprachfördermethoden (Böse & Scherger, 2023). Auf der Basis voriger Ergebnisse (TRIO, s. Schierbaum et al., 2020) werden folgende Fragen diskutiert:

(F1) Welches Studiendesign ist notwendig, um die Wirksamkeit einer linguistisch fundierten Sprachförderung auf den verschiedenen Ebenen nachzuweisen?

(F2) Warum gelingt der Nachweis von Effekten in verschiedenen sprachlichen Bereichen unterschiedlich gut?

In TRIO wurde in einem Prä-Post-Kontrollgruppen-Design die Effektivität einer Fortbildungsmaßnahme zu linguistisch fundierter Sprachförderung untersucht. Die Sprachförderkompetenz von 33 Fachkräften (Ebene 3) sowie die Entwicklung der Sprachfähigkeiten (Ebene 1) der von ihnen geförderten mehrsprachigen Kinder (n=78) untersucht. Die Umsetzung der Förderung in die Praxis (Ebene 2) wurde nicht untersucht. Eingesetzt wurden die Instrumente SprachKoPF (Thoma & Tracy, 2014) für die Fachkräfte und LiSe-DaZ (Schulz & Tracy, 2011) für die Sprachfähigkeiten der Kinder (hier: regelbasierte Erwerbsbereiche ‚Satzstruktur‘ und ‚Kasus‘).

Die Ergebnisse zeigen, dass eine Fortbildung zu linguistisch fundierter Sprachförderung nachweislich positive Effekte sowohl auf die Kompetenzentwicklung der Sprachförderkräfte als auch auf die produktiven Sprachfähigkeiten der Kinder hat. Die Fähigkeiten der geförderten Kinder im Bereich Satzstruktur verbesserten sich signifikant gegenüber Kindern, deren Förderkräfte keine Fortbildung besucht hatten. Im Bereich Kasus fanden sich keine Effekte.

Ein Prä-Post-Kontrollgruppen-Design erlaubt folglich, die Wirksamkeit von Fortbildungen für Sprachförderkräfte und deren Sprachfördermaßnahmen (Ebene 1) und 3)) nachzuweisen (F1). Jedoch sind die Herausforderungen der Messung unter praxisnahen Bedingungen vielfältig. Zeigen sich Effekte, können diese je nach Förderbereich auch auf kurzfristige Trainingsergebnisse zurückzuführen sein. Lassen sich keine Effekte nachweisen, kann dies verschiedene Gründe haben; Rückschlüsse auf die Nicht-Wirksamkeit einer Maßnahme greifen oft zu kurz (F2). Insbesondere Studien zu den Ebenen 1) und 2) werden meist unter praxisnahen Bedingungen durchgeführt, um konkrete Handlungsempfehlungen für die Praxis abzuleiten. Zugleich werden die komplexen Rahmenbedingungen nach Möglichkeit so kontrolliert, dass Vergleiche zwischen Untersuchungs- und Kontrollgruppen zulässig sind. Dennoch bleiben Unterschiede bestehen, die einen Wirksamkeitsnachweis erschweren, z.B. hinsichtlich Umfang, Inhalten und linguistischer Fundiertheit des Sprachförderansatzes, konzepttreuer Umsetzung, Passung zwischen Förderbereichen und Messinstrumenten, Qualifikation der Förderkräfte, Sprachständen der Kinder. Dazu sind weitere kontrollierte Studien zum Zusammenhang der verschiedenen Ebenen dringend erforderlich.

Offen ist bis dato auch, welche Effekte auf die Sprachentwicklung von Kindern zu erwarten sind, wenn eine wirklich durchgängige Sprachförderung realisiert wird, d.h. nicht nur einzelne Fachkräfte einer Einrichtung, sondern das gesamte Team, ggf. mit kooperierenden Einrichtungen, im Sinne einer durchgängigen Sprachförderung gemeinsam fortgebildet (Müller et al., 2015) und beim Praxistransfer des Gelernten begleitet werden.

S-25: Bildung für nachhaltige Entwicklung ist mehr als nur Wissen – Einstellungen, Ängste und Kompetenzen von Lernenden als Untersuchungsgegenstände in der Bildungsforschung von Deutschland, Österreich und der Schweiz

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 133

Chair der Sitzung: Jennifer Stemmann

Chair der Sitzung: Werner Rieß

Chair der Sitzung: Benjamin Fauth

Bildung für nachhaltige Entwicklung ist mehr als nur Wissen – Einstellungen, Ängste und Kompetenzen von Lernenden als Untersuchungsgegenstände in der Bildungsforschung von Deutschland, Österreich und der Schweiz

Chair(s): **Jennifer Stemmann** (Pädagogische Hochschule Freiburg, Deutschland), **Werner Rieß** (Pädagogische Hochschule Freiburg, Deutschland)

Diskutant*in(nen): **Benjamin Fauth** (IBBW Stuttgart)

Problemstellung/Ausgangslage

Bildung ist ein wesentlicher Faktor, um Menschen auf das Lösen nachhaltigkeitsbezogener Probleme sowie auf Herausforderungen im Umgang mit dem Klimawandel vorzubereiten (Lohmann et al., 2021; UNESCO, 2021). Die Ungewissheit und Unvollständigkeit des Wissens über die komplexen Zusammenhänge sowie die Vielfalt an Zielkonflikten innerhalb der Nachhaltigkeitsdimensionen führen zu widersprüchlichen Gedanken und Emotionen (Grunwald & Kopfmüller, 2022), die von Lernenden erkannt, reflektiert und ein Stückweit auch als unauflösbar akzeptiert werden sollten (Weselek, 2022). Hierfür benötigen die Lernenden neben kognitiven Fähigkeiten auch ethische, soziale, emotionale, motivationale und verhaltensbezogene Fähigkeiten (Rychen, 2008). Diese Aspekte treten in bildungspolitischen und wissenschaftlichen Diskussion vermehrt in den Vordergrund (Azevedo & Marques, 2017; Mochizuki & Bryan, 2015), weil Wissen in diesem Kontext nicht automatisch zum Handeln führt (Braun & Dierkes, 2019). Neben der Vermittlung von Grundlagenwissen sollten demnach auch die Einstellungen und Gefühle der Schüler*innen zu nachhaltigkeitsbezogenen Themen im Unterricht in den Blick genommen werden. Wie Ängste mit dem Wissen interagieren, bzw. welchen Einfluss Einstellungen auf das Verhalten der Lernende haben, ist bislang aber weniger Gegenstand der Forschung. Inwieweit nicht-kognitive Komponenten von Nachhaltigkeitskompetenzen in schulischen und außerschulischen Bildungszielen (z.B. Siegmund et al., 2020) implementiert sind und wie sich der Stand empirischer Bildungsforschung grundsätzlich darstellt, ist ein weiteres Desiderat.

Zielsetzung/Zuschnitt des Symposiums

Die Beiträge des Symposiums greifen diese Desiderata auf und untersuchen zum einen auf individueller Ebene die Einstellungen und Ängste von Lernenden und zum anderen auf systemischer Ebene die Implementierung von Nachhaltigkeitskompetenzen in Bildungszielen und in der Bildungsforschung.

Der erste Beitrag widmet sich den Einstellungen von Schüler*innen der 9. Klassestufe in Deutschland zum Klimawandel. Basierend auf bereits vorhandenen Instrumenten zu Umwelteinstellungen auf Basis selbstberichteten Verhaltens entwickeln die Autor*innen ein Instrument, das klimarelevante Aspekte integriert und validieren die daraus resultierenden Einstellungsmaße mit traditionellen Einstellungsmaßen zum Klimawandel.

Der zweite Beitrag geht der Frage nach, welche Themen im Kontext des Klimawandels, die 13-14-jährigen Schüler*innen in Österreich am meisten beunruhigen, wie diese Ängste mit ihrem Wissen zusammenhängen und wie optimistisch sie in die Zukunft blicken.

Der dritte Beitrag nimmt die NE-Zertifikatsprogramme in Deutschland in den Blick. Die Autor*innen analysieren diese qualitativ und quantitativ in Hinblick auf die mit ihnen intendierten Teil- und Schlüsselkompetenzen sowie deren wissenschaftliche Fundierung.

In dem letzten Beitrag wird ein systematisches Literaturreview vorgestellt, das die bisherigen BNE-Forschungstätigkeiten in der Schweiz anhand eines Kategoriensystems hinsichtlich ihrer Inhalte, Methoden, Bildungskontexte und Zielgruppen analysiert.

Beiträge des Symposiums

Unterrichtsentwicklung in der Klimabildung: Validierung eines Online-Fragebogens zur verhaltensbasierten Erfassung von Einstellungen zum Klimawandel bei Schüler*innen

Magdalena Stadler, Monika Martin, Werner Rieß, Josef Künsting

Pädagogische Hochschule Freiburg, ReCCE

Schulische Klimabildung ist ein bedeutsamer Baustein, um den Klimawandel zu begrenzen und in seinen Folgen erfolgreich zu bewältigen (Conninck et al., 2018; Otto et al., 2020). Vor dem Hintergrund, dass Wissen in diesem Kontext nicht automatisch zum Handeln führt (Braun & Dierkes, 2019; Kollmuss & Agyeman, 2002), treten motivational-affektive Aspekte in der Diskussion vermehrt in den Vordergrund (Azevedo & Marques, 2017; Mochizuki & Bryan, 2015): Neben der Vermittlung von Grundlagenwissen sollten auch die Einstellungen der Schüler*innen zum Klimawandel im Unterricht in den Blick genommen werden. Unter Einstellungen werden nach Rosenberg et al. (1960) kognitive, affektive und behaviorale Reaktionen auf ein Objekt verstanden. In der schulischen Klimabildung ist vor allem die Verhaltenskomponente von Bedeutung (Dunlop et al., 2022; Howard-Jones et al., 2021), mit dem Ziel, die Bereitschaft der Schüler*innen für klimabewusstes Verhalten zu fördern.

Im Sinne einer evidenzbasierten Steuerung wird von Schulen wie auch von Lehrkräften gefordert, dass diese die Qualität ihres Unterrichts datengestützt bewerten und darauf aufbauend weiterentwickeln (Sekretariat der Kultusministerkonferenz, 2016; Thiel et al., 2019). Diese Forderung im Kontext der Klimabildung aufgreifend, widmet sich der vorliegende Beitrag der Validierung eines Instruments zur Messung von Einstellungen zum Klimawandel bei Schüler*innen der Sekundarstufe, das Lehrkräfte zur Unterrichtsentwicklung nutzen können. Grundlage für das entwickelte Instrument stellt der GEB-40 dar (Kaiser et al., 2007), der Umwelteinstellungen von Jugendlichen auf Basis von selbstberichtetem Verhalten im Rahmen des Campbell-Paradigmas (Kaiser et al., 2010) misst. Dieser wurde für den Kontext der Einstellungen zum Klimawandel adaptiert und enthält 40 Items zu Verhaltensweisen aus den Bereichen Mobilität, Energiesparen, Ernährung und Konsum und gesellschaftliches Engagement. Die Bearbeitung erfolgt online (am PC, Tablet oder Smartphone) und dauert zwischen fünf und zehn Minuten.

Zur Überprüfung entsprechend dem Konzept der argumentativen Validität (Kane, 2013) wurde der Test in einer Stichprobe von 810 Schüler*innen (MAlter = 15,6 Jahre) unterschiedlicher Schularten in [Bundesland] eingesetzt. Zusätzlich wurden ein umfassender Kompetenztest zu kognitiven Kompetenzen zum Klimawandel und Selbsteinschätzungsskalen zu Konstrukten, die mit der Einstellung zum Klimawandel assoziiert sind (z. B. Überzeugungen, Besorgnis), eingesetzt sowie relevante Hintergrundvariablen (z. B. sozioökonomischer Status) erfragt. Strukturiert entlang verschiedener Validitätskriterien nach Messick (1995) wurden Annahmen zur intendierten Testwertinterpretation und -verwendung geprüft:

Die Passung eines eindimensionalen Rasch-Modells (EAP-Rel. = .795, WLE-Rel. = .806) zeigt, dass sich die erfragten Verhaltensweisen auf ein gemeinsames zugrundeliegendes Verhaltensziel, den Klimawandel einzudämmen, zurückführen lassen (strukturelle Validität) und spricht damit – neben weiteren inhaltlich-theoretischen Argumenten (inhaltliche Validität) – für eine erfolgreiche Übersetzung des Konstrukts in das Testinstrument. Als Argumente zur externen Validität werden mittlere bis hohe signifikante Zusammenhänge herangezogen, die zwischen den Testwerten der verhaltensbasierten Einstellung zum Klimawandel und Selbsteinschätzungsmaßen zu relevanten Einstellungsfacetten ($r = .20$ bis $r = .51$) sowie Testwerten zu kognitiven Kompetenzen ($r = .46$) bestehen. Weiterhin ist die Reihenfolge der Itemschwierigkeiten für verschiedene Teilstichproben (Gymnasium vs. andere Schularten) vergleichbar, was auf Unabhängigkeit der Testwerte von der Validierungsstichprobe hindeutet (Generalisierbarkeit). Zusammengefasst kann also von einer validen Übersetzung des Testverhaltens in Testwerte ausgegangen werden.

Ausgehend von den Daten einer aktuell durchgeführten Erhebung zur Evaluation des Online-Instruments durch Lehrkräfte (angezielt: $n = 15$) sollen im Beitrag zudem Argumente zur konsequentialen Validität aufgeführt sowie allgemeinere Anforderungen aus der Praxis an Instrumente zur Unterrichtsentwicklung (im Kontext Klimabildung) diskutiert werden.

„Welche Zukunft erwartet dich?“ – Zukunftsvisionen, Umweltängste und Betroffenheit von Jugendlichen in Zeiten des Klimawandels

Freya Hutter, Alexandra Huber, Lena v. Kotzebue
Paris Lodron Universität Salzburg

Theoretischer Hintergrund

Die Klimaangst bei Jugendlichen nimmt weltweit zu, und es zeigen sich negative Konsequenzen bezüglich des umweltfreundlichen Verhaltens, Überzeugungen, Lernprozesse, sowie für das Engagement für Umweltthemen (Clayton & Karazsia, 2020; Hickman et al., 2021; Pearson, 2024; Russell, 2024). Auch ein höheres Bewusstsein über die Auswirkungen des Klimawandels steigert die empfundene Betroffenheit, welche ebenfalls Sorgen und Ängste verursachen (u.a. Léger-Goodes et al., 2022; Malka et al., 2009). Obgleich hohes faktisches Wissen vorhanden ist, zeigt sich vielfach ein Mangel an tiefgreifendem Verständnis (Feldbacher et al., 2024) sowie die Unfähigkeit, mit Klimaangst adäquat umzugehen (Pearson, 2024). Fehlende Coping-Strategien können einer erfolgreichen Umweltbildung hinderlich sein, insbesondere, wenn Jugendliche mit ihren Gefühlen im Unterricht allein gelassen werden (Léger-Goodes et al., 2022).

Im Fokus des Beitrages stehen die Zukunftsvisionen und Umweltängste von Jugendlichen im Jahr 2023, welche die Einschätzungen zur eigenen Betroffenheit von den Folgen des Klimawandels, der Umweltverschmutzung und des Ressourcenmangels einbeziehen. Zudem werden die Ergebnisse mit denen aus den Studien von Unterbruner (1991, 2011) aus den Jahren 1988 und 2008 verglichen und diskutiert.

Fragestellung

Im Rahmen der Untersuchung wurden folgende Forschungsfragen nachgegangen:

- Welche Themen in Bezug auf den Klimawandel beunruhigen Schüler:innen am meisten? Was sind die größten Ängste und Wünsche Schüler:innen für die Zukunft?
- Welchen Zusammenhang gibt es zwischen der eigenen Betroffenheit der Schüler:innen bezüglich der Folgen des Klimawandels, der Umweltverschmutzung und des Ressourcenmangels und den Ängsten der Schüler:innen und ihrem Wissen zu diesen Themen?
- Zeichnen Schüler:innen ihr Heimatland (Österreich) in 20 Jahren pessimistisch, neutral oder optimistisch? Unterscheiden sich die Zukunftsvisionen je nach Geschlecht? Und wie haben sich die Zukunftsvision im Verlauf der letzten 35 Jahren bei den Schüler:innen verändert?

Methode

Im Rahmen der Erhebung bei einer Stichprobe aus $N = 280$ Schüler:innen (Alter: $M=13,9$ Jahre; $SD=1,04$ Jahre; $w=127$; $m=153$) wurde ein quantitativer Fragebogen mit Likert-Skalen zur eigenen Betroffenheit bezüglich Klimawandel, Umweltverschmutzung und Ressourcen-/Rohstoffmangel; Bewertung von 20 Ängsten; Angabe der größten Wünsche und Ängste und einem Single-Choice Klimawandel-Wissenstest eingesetzt. Zudem sollten die Schüler:innen ihr Heimatland (Österreich) in 20 Jahren zeichnen. Große Teile des Fragebogens und die Zeichnung wurden in den Studien von Unterbruner (1991, 2011) bereits bei einer vergleichbaren Stichprobe ($NU1=302$, $NU2=677$) eingesetzt und ermöglichen einen Vergleich über 35 Jahre. Zur Sicherstellung der Reliabilität wurde die interne Konsistenz (Cronbachs α) der Skalen und die objektive Übereinstimmung der Codierung der Zeichnungen durch Cohens Kappa ($k>.9$) geprüft.

Ergebnisse und ihre Bedeutung

Es konnte festgestellt werden, dass das Wissen der Schüler:innen signifikant mit der eigenen Betroffenheit bezüglich Klimawandel, Umweltverschmutzung und Ressourcen-/Rohstoffmangel zusammenhängt ($r=.309$, $p<.001$). Desweiteren zeigte sich ein signifikanter Zusammenhang zwischen eigener Betroffenheit und den Ängsten der Schüler:innen ($r=.450$, $p<.001$), betreffend u.a. die „Angst vor Umweltzerstörung“, welche mehr als jede/r Zweite als beängstigend bewertete (59,3%).

Dies steht im Einklang mit den Ergebnissen der Bertelsmann Stiftung aus den Jahren 2022/ 2023 (Liz Mohn Center gGmbH, 2023), in welcher mehr als ein Drittel der Befragten (37% bzw. 38%) kritisch auf die Zukunft Deutschlands blickten (16% bzw. 20% positiv), und vergleichbaren Studien aus den Jahren 1989 und 2008/2009 (Unterbruner 1991, 2011). Mädchen sehen die Zukunft pessimistischer (61%) als Jungen (43%). Es zeigt sich auch, dass Mädchen in ihren Zukunftsvorstellungen häufiger die Natur und die Umwelt miteinbeziehen und diese zerstört dargestellt wird. Ältere Schüler:innen stellen tendenziell die Umwelt eher als zerstört dar als jüngere.

Trotz dystopischer Visionen und Klimaangst lässt sich ein positiver Ausblick einnehmen: Die Natur spielt in der Mehrheit (60%) der gezeichneten Zukunftsvisionen der Jugendlichen immer noch eine große Rolle. Jedoch ist ein Rückgang von

Naturdarstellungen (z.B. Umweltidylle oder Naturkatastrophen) in den Zeichnungen bzw. Zukunftsvisionen der Jugendlichen erkennbar: 80% (1988), 62% (2008) und 60% (2023) (Unterbruner, 1991, 2011).

Mehr als „nur“ Wissen? Kompetenzziele in Nachhaltigkeitszertifikatsprogrammen deutscher Hochschulen aus der Perspektive eines integrierten Rahmenmodells für NE-Kompetenz

Maximilian Irion¹, Christiane Bertram², Benjamin Nagengast¹

¹Hector-Institut für Empirische Bildungsforschung, Eberhard Karls Universität Tübingen, ²Kompetenzzentrum für Nachhaltige Entwicklung, Eberhard Karls Universität Tübingen

Theoretischer Rahmen

In Nachhaltigkeitszertifikatsprogrammen (im Folgenden: NE-Zertifikatsprogramme) setzen sich Studierende verschiedener Studiengänge und Fachrichtungen mit nachhaltigkeitspezifischen Themen und Fragestellungen auseinander. In der Regel umfassen die Programme einen Grundlagenbereich, in dem Studierende eine oder mehrere Einführungsveranstaltungen zu Themen Nachhaltiger Entwicklung (NE) besuchen, sowie einen Vertiefungsbereich, in dem sie weiterführende und/oder praxisorientierte Lehrveranstaltungen belegen (vgl. Herth et al., 2018). Durch die Teilnahme an den Lehrveranstaltungen und den abschließenden Erwerb eines Zertifikats soll sich die Nachhaltigkeitskompetenz von Studierenden verbessern. Darunter versteht man ein Set an Wissen, Fähigkeiten, Werten, Einstellungen und Bereitschaften, das Studierende bzw. Lernende dazu befähigt, in ihrem Leben, sozialen Umfeld und (späteren) Beruf zur Lösung von NE-bezogenen Problemen beitragen und eine NE auf privater, sozialer und institutioneller Ebene gestalten zu können (vgl. Brundiers et al., 2021; Rieß et al., 2018; Wiek et al., 2011).

Ein besonderes Desiderat stellt die Evaluation von NE-Zertifikatsprogrammen dar (Potthast et al., 2024). Hierfür muss jedoch zuerst definiert werden, aus welchen konkreten Teil- und Schlüsselkompetenzen sich die zu fördernde Nachhaltigkeitskompetenz zusammensetzt. Ausgehend von dem Rahmenmodell für Nachhaltigkeitskompetenz von Rieß et al. (2018), in dem zwischen kognitiver, affektiv-motivationaler, verhaltensbezogener und generischer Zieldimension unterschieden wird, werden diesen Zieldimensionen in dem hier vorgelegten integrierten Rahmenmodell die Teil- und Schlüsselkompetenzen von sechs wissenschaftlichen Kompetenzmodellen zugeordnet und ggf. verdichtet.

Forschungsfragen

Damit Studierende zur nachhaltigen Gestaltung der Gesellschaft befähigt werden können, ist in NE-Zertifikatsprogrammen Klarheit über die jeweils zugrunde liegenden Kompetenzziele erforderlich (Anders et al., 2021). Mithilfe des integrierten Rahmenmodells soll deshalb untersucht werden,

- welche Kompetenzen durch NE-Zertifikatsprogramme an deutschen Hochschulen gefördert werden sollen und
- ob die Konzeption dieser Programme auf wissenschaftlich fundierten Kompetenzmodellen beruht.

Methode

Zur Untersuchung der ersten Forschungsfrage werden nach Döring und Bortz (2016) sowie Hsieh und Shannon (2005) eine quantitative und qualitative Dokumentenanalyse durchgeführt, die Zertifikatsbeschreibungen von 14 NE-Zertifikatsprogrammen sowie Kursbeschreibungen von 14 dazugehörigen Einführungsveranstaltungen umfassen. Das integrierte Rahmenmodell dient dabei als deduktiv entwickeltes, nominales Kategoriensystem, mit welchem analysiert werden kann, ob sich in den einzelnen Dokumenten Verweise auf die Teil- und Schlüsselkompetenzen finden oder nicht. Darüber hinaus werden ausgehend vom Material weitere Teil- und Schlüsselkompetenzen induktiv identifiziert. Zur Untersuchung der zweiten Forschungsfrage wird eine quantitative Dokumentenanalyse durchgeführt und analysiert, inwieweit die Zertifikatsbeschreibungen der 14 NE-Zertifikatsprogramme auf gängige wissenschaftliche Kompetenzmodelle verweisen. Ziel der Dokumentenanalysen ist eine frequenzanalytische Auswertung, durch die herausgearbeitet werden kann, welche Teil- und Schlüsselkompetenzen bzw. Zieldimensionen in NE-Zertifikatsprogrammen vorwiegend gefördert werden sollen und inwieweit wissenschaftliche Kompetenzmodelle bei der Programmkonzeption berücksichtigt werden.

Ergebnisse und Ausblick

Eine Pilotierung des zur Untersuchung der ersten Forschungsfrage entwickelten Kategoriensystems anhand von Beispielmateriale deutet auf eine Dominanz verhaltensbezogener Kompetenzziele hin. Darüber hinaus gibt es Hinweise auf weitere Teilkompetenzen, die nicht in den wissenschaftlichen Kompetenzmodellen berücksichtigt werden, beispielsweise die Kompetenz zur Vermittlung von nachhaltigkeitspezifischen wissenschaftlichen Erkenntnissen. Die finale Untersuchung, die bis zur Durchführung des Symposiums abgeschlossen sein wird, verspricht schließlich Hinweise darauf, inwieweit die Praxis der Hochschul-BNE, in diesem Fall in Form von NE-Zertifikatsprogrammen, den Erwartungen und Anforderungen wissenschaftlicher Kompetenzmodelle gerecht wird und inwieweit darüber hinausgehende, eigene Kompetenzziele adressiert werden (Singer-Brodowski & Kminek, 2023). Zudem sollen perspektivisch auf Grundlage der Ergebnisse passgenaue Messinstrumente zur Untersuchung der Wirksamkeit der hier fokussierten NE-Zertifikatsprogramme entwickelt werden.

Forschungsstand zu Klimabildung und BNE in der Schweiz – Über- und Ausblick

Matthias Probst¹, Noé Balsiger², Moritz Gubler¹, Sebastian Tempelmann²

¹Pädagogische Hochschule Bern, Oeschger-Zentrum für Klimaforschung, Universität Bern, ²Pädagogische Hochschule Bern

Bildung wird als wesentlicher Hebel zur Bewältigung drängender sozial-ökologischer Herausforderungen wie dem menschengemachten Klimawandel angesehen (Otto et al., 2020; UNESCO, 2021). Für die Entwicklung und Umsetzung von wirkungsvollen Bildungsprogrammen und Unterrichtszugängen im Bereich der Klimabildung (KB) und der Bildung für nachhaltige Entwicklung (BNE), sind einerseits entsprechende bildungspolitische Rahmenbedingungen entscheidend. Andererseits braucht es aber auch umfassenden Überblick über wissenschaftlich gesicherte Erkenntnisse zu deren Anwendbarkeit in unterschiedlichen Unterrichtskontexten, zu ihrer Wirkung auf verschiedene Zielgruppen oder zu ihrer Verbreitung entlang unterschiedlicher Umsetzungsebenen der Bildungslandschaft. Solche systematischen Analysen des aktuellen Forschungsstandes liegen auf internationaler Ebene sowohl im Kontext der KB (z.B. Monroe et al. 2019) wie auch hinsichtlich BNE bereits vor (z.B. Ssossé et al. 2021; Yang & Xiu, 2023). Jedoch besteht auf nationaler Ebene oft Nachholbedarf, was insbesondere auf kleinere Nationen wie die Schweiz zutrifft (Adamina et al., 2018; Pache et al. 2023).

Mit dem Ziel, diese Lücke für den schweizerischen Forschungskontext zu schliessen, wird im Rahmen der Zusammenarbeit im «International Consortium for Climate Change Education and Education for Sustainable Development» deshalb aktuell eine umfangreiche Analyse zum Stand der empirischen Forschung zu KB und BNE in allen Sprachregionen der Schweiz durchgeführt. Dabei steht die Beantwortung folgender Fragestellungen im Vordergrund:

- 1) Welche Bildungskontexte und Zielgruppen wurden durch bisherige Forschungstätigkeiten im Bereich KB und BNE in der Schweiz abgedeckt (z.B. schulisch vs. ausserschulisch, Schulstufen)?
- 2) Welche forschungsmethodischen Rahmenbedingungen und Zugänge kamen in bisherigen Forschungstätigkeiten im Bereich KB und BNE in der Schweiz zur Anwendung (z.B. Stichprobengrössen, Analyseverfahren)?
- 3) Welche inhaltlichen Bezugsfelder wurden in bisherigen Forschungstätigkeiten im Bereich KB und BNE in der Schweiz tangiert (z.B. Fachkonzepte, Themen)?

Um den aktuellen Stand eines Forschungsfeldes fundiert aufzuarbeiten, eignet sich die Methode der systematischen Literaturanalyse (Fink, 2019). In dieser Studie werden hierzu die Einträge mehrerer Literaturdatenbanken (Web of Science, Scopus, ERIC, Pedocs, FIS Bildung, érudit) entlang von unterschiedlichen Suchbegriffen und deren Übersetzungen in Deutsch, Englisch, Französisch und Italienisch gesammelt. Die ausgewiesenen Literatureinträge werden mittels eines auf die Forschungsfragen abgestimmten Kategorienrasters analysiert, um darauf aufbauend Antworten auf die oben aufgeführten Fragestellungen zu finden.

Erste Einblicke in die bisher untersuchte Forschungsliteratur zu KB und BNE in der Schweiz zeigen, dass vor allem im Bereich der KB noch erhebliche Lücken klaffen und die Forschungstätigkeit in der Schweiz grundsätzlich relativ tief ist. Die meisten Publikationen im Bereich der KB beziehen sich auf die Sekundarstufe 1 und beschränken sich geographisch meist auf die Deutschschweiz. In der französischsprachigen Schweiz wurde bisher kaum nennenswerte Aktivitäten in der empirischen Bildungsforschung zu KB festgestellt. Im inhaltlich und konzeptionell umfassenderen Feld der BNE wurden erwartungsgemäss markant mehr Literatureinträge gefunden. Dabei fällt auf, dass empirische Untersuchungen zu BNE häufig auf die tertiäre Bildungsstufe fokussieren, während die publizierten Forschungsaktivitäten auf der Primarstufe klar unterrepräsentiert sind. Geographisch lassen sich die empirischen Untersuchungen im Bereich BNE in sämtlichen Sprachregionen verorten.

Im vorgeschlagenen Beitrag an der GEBF 2025 werden einerseits die finalen Resultate der systematischen Literaturanalyse zu aktuellen Forschungstätigkeiten im Bereich KB und BNE in der Schweiz präsentiert und Gründe für die gefundenen Muster und Lücken multiperspektivisch diskutiert. Auf Basis dieser Auslegeordnung werden prioritär anzugehende Umsetzungsebenen, Zugänge und Felder der empirischen Bildungsforschung zu KB und BNE im schweizerischen Kontext benannt, um den Beitrag der empirischen Bildungsforschung zur Bewältigung drängender sozial-ökologischer Herausforderungen zu stärken.

S-26: Förderbezogene Diagnostik im inklusiven Mathematikunterricht

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: O 129

Chair der Sitzung: Fynn Töllner

Chair der Sitzung: Poldi Kuhl

Chair der Sitzung: Michael Besser

Chair der Sitzung: Daniel Sommerhoff

Förderbezogene Diagnostik im inklusiven Mathematikunterricht

Chair(s): **Fynn Töllner** (Leuphana Universität Lüneburg, Deutschland), **Poldi Kuhl** (Leuphana Universität Lüneburg, Deutschland), **Michael Besser** (Leuphana Universität Lüneburg, Deutschland)

Diskutant*in(nen): **Daniel Sommerhoff** (IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik)

Der Übergang zu einem inklusiven Schulsystem stellt alle beteiligten Akteur*innen vor Herausforderungen (Borsch, 2018). Insbesondere die Frage der Unterstützung aller Lernenden unabhängig von deren individuellen Lern- und Leistungsvoraussetzungen ist dabei aufgrund der gestiegenen (Leistungs-)Heterogenität in inklusiven Schulklassen von zentraler Bedeutung. Eine Möglichkeit, diese Heterogenität zu adressieren, stellt die Individualisierung von Lernprozessen dar (Posch, 2019). Eine solche Individualisierung, welche sich an den Bedürfnissen der Schüler*innen orientiert, setzt jedoch eine genaue Diagnostik der individuellen Leistungsstände der Lernenden voraus (Langner & Jugel, 2019). Dies kann beispielsweise im Rahmen eines formativen Assessments erfolgen. Es ist jedoch bislang unzureichend untersucht, wie eine solche Diagnostik im inklusiven Mathematikunterricht erfolgreich umgesetzt und auch zur individuellen Förderung der Lernenden produktiv genutzt werden kann. Diese Forschungslücke aufgreifend, werden in diesem Symposium Vorträge gebündelt, die sich empirisch im inklusiven Mathematikunterricht mit verschiedenen Facetten erfolgreichen Diagnostizierens und darauf aufbauenden Feedback- und Fördermaßnahmen auseinandersetzen. Dabei werden alle Projekte dieses Symposiums im Rahmen der Förderlinie „Förderbezogene Diagnostik in der inklusiven Bildung“ durch das BMBF gefördert.

Hierfür werden zunächst mit der Ebene der Lehrkräfte, die Akteure in den Blick genommen, die im Unterricht die Diagnostik durchführen. Entsprechend fokussiert der erste Beitrag auf Lehrkräfteebene Gelingensbedingungen einer erfolgreichen Diagnostik in inklusiven Klassen. Dazu nimmt dieser Vortrag die Auswirkung der Selbstwirksamkeit und des fachdidaktischen Wissens von Lehrkräften bei der Durchführung formativer Diagnostik auf die mathematische Kompetenzentwicklung von Schüler*innen in den Blick. Auch wird untersucht, ob sich die Auswirkungen dieser Merkmale von Lehrkräften in Abhängigkeit vom Strukturierungsgrades des formativen Assessments unterscheiden, d. h. ob Unterschiede zwischen den beiden Bedingungen „planned for interaction“ und „curriculum-embedded assessment“ auftreten.

Im zweiten Beitrag wird ein Förderkonzept für die Mathematik- und Lesekompetenzen von Schüler*innen in inklusiven Schulen am Übergang von der Grundschule zur Sekundarstufe 1 vorgestellt und damit die Materialebene in den Blick genommen. Die Präsentation und empirische Diskussion einer diagnosegeleiteten Förderkartei bilden den Kern dieses Beitrags.

Im dritten Vortrag werden die Schüler*innen als Rezipienten von diagnosegeleiteten Handlungen betrachtet. Hierzu werden Wirkmechanismen formativen Assessments auf der Ebene der Lernenden analysiert. Es wird untersucht, wie Schüler*innen inklusiv arbeitender Schulen Feedback verarbeiten. In diesem Kontext erhielten Schüler*innen Feedback mit unterschiedlichem Informationsgehalt, um zu analysieren, welches Feedback wie verarbeitet wird und schließlich am lernwirksamsten ist.

Abschließend erfolgt eine Einordnung der Vorträge aus mathematikdidaktischer Perspektive. Dabei werden die Beiträge im Hinblick auf Implikationen für eine evidenzbasierte Entwicklung einer förderbezogenen Diagnostik für den inklusiven Mathematikunterricht kritisch zusammengefasst und Schlussfolgerungen für die Bildungsforschung und alltägliche Bildungspraxis abgeleitet.

Beiträge des Symposiums

Effekte von Lehrkräftemerkmalen auf die Wirksamkeit zweier Ansätze formativen Assessments im Mathematikunterricht

Larissa Aust¹, Jeanne-Celine Linker², Luise Eichholz², Janina Lenhart¹, Marcus Nührenböcker¹, Jana Schiffer¹, Christoph Selter², Benjamin Weiß², Elmar Souvignier¹

¹Universität Münster, ²Technische Universität Dortmund

Theoretischer Hintergrund

Bei Formativem Assessment werden lernprozessbegleitend diagnostische Informationen über den Lernstand von Schüler:innen erfasst, um diese für eine adaptive Unterrichtsgestaltung zu nutzen (Black und Wiliam, 1998). Positive Effekte formativen Assessments auf die Leistungsentwicklung von Schüler:innen sind belegt; die Wirkung ist jedoch insbesondere abhängig von der spezifischen Implementation im Unterricht (z. B. Lane et al., 2019). In der vorliegenden Studie werden daher zum einen zwei unterschiedlich stark strukturierte Ansätze formativen Assessments (Shavelson et al., 2008) verglichen und zum anderen werden Effekte von Merkmalen der Lehrkräfte untersucht. Während Lehrkräfte in der stärker strukturierten Variante „curriculum-embedded assessment“ (CE) vorbereitete, schriftliche Diagnoseaufgaben zur systematischen Erfassung des Lernstandes aller Schüler:innen einsetzen, lernen Lehrkräfte in der weniger stark strukturierten Variante „planned-for-interaction assessment“ (PI), Diagnosegespräche auf Basis reichhaltiger Aufgaben zu führen und entsprechende diagnostische Impulse zu gestalten. Im Hinblick auf die Rolle von Lehrkräftemerkmalen werden Selbstwirksamkeit und fachdidaktisches Wissen untersucht, da beide Merkmale mit einer intensiveren und erfolgreichereren Umsetzung von formativem Assessment im Unterricht assoziiert sind (Heitink et al., 2016; Lane et al., 2019; Schildkamp et al., 2020; Yan et al., 2021). Inwieweit die Effekte dieser Merkmale abhängig vom Strukturierungsgrad der Ansätze sind, ist jedoch eine offene Frage. Es ist anzunehmen, dass der Einfluss der Lehrkräftemerkmale mit zunehmendem Strukturierungsgrad sinkt.

Fragestellungen

1. Welchen Effekt haben (a) die Selbstwirksamkeit und (b) das fachdidaktische Wissen der Lehrkräfte auf die Entwicklung der mathematischen Kompetenzen von Schüler:innen bei der Implementation von formativem Assessment?
2. Unterscheiden sich die Effekte der Lehrkräftemerkmale in Abhängigkeit von dem Strukturierungsgrad der Ansätze formativen Assessments?

Methode

An der Studie nahmen N = 923 Schüler:innen (MAlter = 7.09; SDAAlter = 0.48; 51.6% weiblich) aus 41 Klassen des zweiten Schulbesuchsjahres in Nordrhein-Westfalen teil. Die beiden Varianten formativen Assessments wurden in einem Zweigruppen-Prä-Post-Design realisiert. Die Lehrkräfte beider Gruppen (MAlter = 39.93; SDAAlter = 10.21, 95.1% weiblich) besuchten über ein Schuljahr hinweg jeweils sechs Fortbildungsveranstaltungen zu diagnosebasierter differenzierter Mathematikförderung in inklusiven Klassen. Die jeweiligen Konzepte (CE vs. PI) wurden jeweils im Anschluss an die Fortbildungen mit hoher Wiedergabetreue (Selbstauskunft der Lehrkräfte) in den Klassen umgesetzt. Die mathematischen Basiskompetenzen der Schüler:innen wurden zu Beginn des Schuljahres mithilfe des Basis-Math-G 1+ (Moser-Opitz et al., i. V.) und am Ende des Schuljahres mit dem Basis-Math-G 2+ (Moser-Opitz et al., 2020) erfasst. Auf Ebene der Lehrkräfte wurden die Selbstwirksamkeit bezüglich des Umgangs mit Heterogenität (Junker et al., 2020) sowie das fachdidaktische Wissen im Bereich Stellenwertverständnis und Operationsverständnis mithilfe von Fallvignetten erhoben (angelehnt an Brandt 2022). Es wurde ein hierarchisches Regressionsmodell geschätzt, in welches die Lehrkräftemerkmale als Level-2-Prädiktoren eingingen (Fragestellung 1). Zur Prüfung möglicher Interaktionseffekte (Fragestellung 2) wurden die Cross-Level-Interaktionen zwischen Lehrkräftemerkmalen und Gruppenzugehörigkeit spezifiziert.

Ergebnisse und Implikationen

Es zeigte sich weder für die Selbstwirksamkeit ($\beta = -.059$, $p = .175$) noch für das fachdidaktische Wissen ($\beta = -.000$, $p = .488$) der Lehrkräfte ein Effekt auf die mathematischen Kompetenzen der Schüler:innen. Für beide Merkmale wurden zudem keine Interaktionseffekte gefunden (Selbstwirksamkeit: $\beta = -.012$, $p = .469$; fachdidaktisches Wissen: $\beta = -.005$, $p = .403$).

Entgegen unseren Erwartungen erwies sich die Entwicklung mathematischer Kompetenzen der Schüler:innen als unabhängig von der Selbstwirksamkeit und dem fachdidaktischen Wissen der Lehrkräfte. Auch unterschied sich der Einfluss der Lehrkräftemerkmale nicht zwischen den beiden Gruppen. Ein Interpretationsansatz für die unerwarteten Befunde ist, dass bei sehr umfassenden und eng begleiteten Interventionen potenzielle Moderatoren auf Lehrkräfteebene einen eher geringen Einfluss auf die Schüler:innenleistungen haben. Zudem können Merkmale von Lehrkräften nur indirekt und vermittelt über zahlreiche unterrichtspraktische Prozesse Einfluss auf die Schüler:innenleistungen nehmen. Diese große Transferdistanz (Yan & Cheng, 2015) könnte folglich zu den nicht-signifikanten Effekten beigetragen haben.

Förderbezogene Diagnostik in Mathematik inklusive Lesen (FORMEL) – Ergebnisse zur diagnosegeleiteten Förderung

Lina Kortüm¹, Maike Mentrop², Anke Hußmann², Stephan Hußmann², Marcus Nührenbörger¹

¹Universität Münster, ²Technische Universität Dortmund

Befunde aus zahlreichen Studien bestätigen seit Jahren den Zusammenhang von sprachlichen und mathematischen Fähigkeiten von Schüler*innen (z. B. Heinze et al., 2011; Mann Koepke & Miller, 2013; Martin & Mullis, 2013; Peng et al., 2020; Schröder & Ritterfeld, 2014; Ufer et al., 2013; Walzebug, 2014). Für die Entwicklung des mathematischen Verständnisses sind sprachliche Kompetenzen der Schüler*innen elementar (Prediger et al., 2015) und gelten als Prädiktor für den Lernzuwachs (Paetsch et al., 2016). Unter den sprachlichen Fähigkeiten gilt das Leseverstehen als Schlüsselkompetenz (Grimm, 2008). Schwierigkeiten beim Lesen können zu Problemen bei der Erschließung von mathematischen Aufgabentexten führen (Prediger et al., 2019; Stephany, 2018). So stellt das Arbeiten in einem sachlichen Kontext für viele Kinder eine Herausforderung dar (Greefrath, 2018). Unklar ist jedoch, wie genau sich Lesekompetenzen auf die Entwicklung mathematischer Kompetenzen auswirken (Moser Opitz et al., 2018; Rosebrock & Nix, 2017; Thomas et al., 2015). Zugleich bestätigen Befunde aktueller Schulleistungsstudien den dringenden Bedarf an validen Diagnose- und Fördermaterialien, um dem steigenden Anteil an Kindern mit nur rudimentären Lese- und Mathematikkompetenzen am Ende der Grundschulzeit angemessen im Unterricht zu begegnen (McElvany et al., 2023; Selter et al., 2020; Stanat et al., 2022).

Das Projekt FORMEL setzt an diesem Bedarf an. Entwickelt, erprobt und evaluiert wurde ein kombiniertes Diagnostik- und Förderkonzept von Mathematik- und Lesekompetenzen in inklusiven Schulen am Übergang von der Grundschule zur Sekundarstufe I. Das FORMEL-Konzept wurde im Rahmen der fachdidaktischen Entwicklungsforschung (Hußmann et al., 2013; Nührenbörger et al., 2019) umgesetzt. Die entwickelten Diagnostik- sowie Fördermaterialien wurden in zyklisch aufeinander aufbauenden Designexperimenten (in Klassen- und Einzelsettings) erprobt und (weiter-)entwickelt.

Der Vortrag stellt sowohl deskriptive Ergebnisse der Haupterhebung (N = 177) zum Zusammenhang mathematischer sowie leseverstehensbezogener Kompetenzen vor als auch die daraus abgeleiteten Fördertypen, die mittels Clusteranalysen ermittelt wurden. Je nach Fördertyp ist das Fördermaterial auszuwählen, dass im Rahmen von Einzelfallanalysen (N = 15) in analoger Form erprobt wurde. Aktuell wird es in digitaler Form weiterentwickelt. Im Beitrag werden Einblicke in den Aufbau und Einsatz der diagnosegeleiteten, digitalen Förderkartei gegeben: Die Förderkartei enthält verschiedene Themenkomplexe, welche jeweils eine Textaufgabe sowie zugehörige Förderaufgaben beinhalten. Die Gestaltung der Textaufgaben orientiert sich an den Gestaltungsmerkmalen der Textaufgaben im Diagnostiktest. Die Förderaufgaben sind nach Aufgaben zur Erarbeitung, zur Erklärung (So geht es) sowie zur vertiefenden Übung differenziert. Sie fokussieren die diagnostizierten Kompetenzbereiche aus dem Diagnostiktest. Die Förderkartei ist nach dem im Test zugeschriebenen Fördertypen strukturiert. Die Lernenden können selbstständig damit arbeiten und jederzeit auf ihren individuellen Lernpfaden zu einfacheren oder herausfordernden Aufgaben wechseln.

Wie verarbeiten Lernende Feedback im inklusiven Mathematikunterricht?

Fynn Töllner, Poldi Kuhl, Michael Besser

Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Die Adaption von Lerngelegenheiten an die unterschiedlichen Bedürfnisse von Lernenden ist ein Merkmal lernerfektiver Settings (Parsons et al., 2018) und wird entsprechend im Kontext der Gestaltung eines inklusiven Mathematikunterrichtes diskutiert (z.B. Schöttler, 2019). Die erhöhte Heterogenität der Lernenden im inklusiven Unterricht stellt dabei für Lehrkräfte eine besondere Herausforderung dar (Borsch, 2018).

Formatives Assessment gilt als eine Möglichkeit, dieser Herausforderung zu begegnen (Mitchell, 2014), Diagnostik und Feedback bilden dabei die Kernaspekte formativen Assessments (Heritage, 2007). Allerdings ist unzureichend untersucht, unter welchen Umständen formatives Assessment effektiv wirkt (Lui & Andrade, 2022). Empirische Studien deuten darauf hin, dass der Feedbackkomponente eine zentrale Rolle zukommt (Decristan et al., 2015). Diese kann u.a. anhand der zur Verfügung gestellten Informationen kategorisiert werden. Hierbei wird häufig zwischen nicht elaborierten und elaborierten Feedbackarten unterschieden. Nicht elaboriertes Feedback besteht allein aus der Information über die Korrektheit einer Antwort (knowledge of

result (KR)) oder der richtigen Antwort (knowledge of correct result (KCR)). Elaborierte Feedbackarten enthalten zusätzliche Informationen (Narciss, 2006; Shute, 2008) und gelten als besonders lernförderlich (Mertens et al., 2022). Hinsichtlich des Unterrichtsfach Mathematik kann elaboriertes Feedback sowohl durch konzeptuelle als auch prozedurale Wissensfacetten angereichert werden.

Neben der Auseinandersetzung mit Feedbackarten wird diskutiert, inwiefern die Wahrnehmung und Nutzung von Feedback durch Lernende Unterschiede in der Effektivität formativen Assessments erklären könnten (Lui & Andrade, 2022). Hierzu haben Lui und Andrade (2022) ein theoretisches Modell entwickelt, welches die internen Prozesse der Feedbackverarbeitung durch Lernende beschreibt. Demgemäß befinden sich Lernende in „Initial States“, welche (a) Überzeugungen über Lern- und Leistungskontrollen, (b) motivationale Ausgangszustände und (c) das Vorwissen beinhalten. Diese Initial States wiederum beeinflussen die „Responses to Feedback“, welche sich aus (1) ausgelösten Emotionen, (2) der Entscheidungsfindung mit Plänen für die nächsten Schritte und (3) der Feedbackinterpretation zusammensetzen. Insgesamt sollen sich diese schließlich auf Leistung und gezeigtes Verhalten auswirken.

Fragestellungen

Um die Wirkung von Feedback als Momentum formativen Assessments zu untersuchen, werden folgende Forschungsfragen adressiert:

(I) Welche Feedbackart ist im inklusiven Mathematikunterricht am lernwirksamsten?

(II) Welche „Responses to Feedback“ wirken auf die Leistung von Schüler*innen?

Methode

Im LERN-IF-Projekt wurde eine Interventionsstudie mit randomisierten Prä-Postdesign durchgeführt. An der Studie haben N=229 Siebtklässler*innen (12.8 Jahre, SD=0.66; 42% m/41% w/1% d/15% missing) von inklusiven nicht-gymnasialen Sekundarschulen teilgenommen, von n=155 liegen Daten vom Pre- und Posttest vor. Diese wurden vorab zufällig einer von vier Versuchsbedingungen zugeteilt: konzeptuelles Feedback (n=41), prozedurales Feedback (n=37), KR-Feedback (n=39) oder kein Feedback (Kontrollgruppe; n=38). Die Erhebung umfasste zwei Messzeitpunkte (MZP) im Abstand von einer Woche. An MZP1 wurde ein Bruchrechentest zur Vorwissenserfassung durchgeführt. An MZP2 beantworteten die Schüler*innen einen Fragebogen zu ihren Initial States, bearbeiteten daran anschließend Bruchrechenaufgaben auf Tablets und erhielten entsprechend der Zugehörigkeit zur Versuchsgruppe Feedback. Danach wurden die „Responses to Feedback“ mithilfe eines Fragebogens erfasst und abschließend ein weiterer Bruchrechentest als Posttest administriert.

Ergebnisse und Diskussion

Bezüglich der ersten Forschungsfrage nach Unterschieden zwischen Feedbackarten deuten erste Analyseergebnisse einer ANOVA auf signifikante Unterschiede zwischen den vier Versuchsbedingungen bezüglich des Posttestscores hin ($F(3, 151)=12.43$, $p<.001$). Tukey's HSD Test belegt signifikante Unterschiede zwischen der Kontrollgruppe, die kein Feedback erhalten hat, und den drei Feedbackgruppen, die sich untereinander nicht signifikant voneinander unterscheiden.

Für die Beantwortung der zweiten Forschungsfrage nach der Wirkung von den „Responses to Feedback“ auf die Leistung der Schüler*innen wurden Regressionen berechnet. In der Gruppe, die konzeptuelles Feedback erhalten hatte, sagten positive ausgelöste Emotionen ($\beta=.38^*$), informative Bedeutungsgebung ($\beta=.35^*$) und eine positive Beurteilung des Feedbacks ($\beta=.39^*$) den Posttestscore vorher. In den anderen Gruppen wurden keine bedeutsamen Zusammenhänge zwischen den „Responses to Feedback“ auf die Leistung ersichtlich.

Diese Ergebnisse geben erste Hinweise darauf, dass Lernende unterschiedliche Feedbackarten differentiell verarbeiten, und sollen abschließend vor dem Hintergrund ihrer Verwendung im inklusiven Mathematikunterricht diskutiert werden.

S-27: Erfassung und Förderung digitalisierungsbezogener Kompetenzen von (angehenden) Lehrkräften

Zeit: Montag, 27.01.2025: 15:15 - 17:00 · Ort: EO 162

Chair der Sitzung: Isabell Runge

Chair der Sitzung: Jakob Thomas Fischer

Chair der Sitzung: Andreas Lachner

Erfassung und Förderung digitalisierungsbezogener Kompetenzen von (angehenden) Lehrkräften

Chair(s): **Isabell Runge** (Universität Potsdam), **Jakob Thomas Fischer** (Universität Potsdam)

Diskutant*in(nen): **Andreas Lachner** (Universität Tübingen)

Digitalisierungsbezogene Kompetenzen von Lehrkräften sind eine grundlegende Voraussetzung für die Gestaltung eines qualitativ hochwertigen digitalgestützten Unterrichts (KMK, 2021; Lachner et al., 2020). Für die Aus- und Fortbildung digitalisierungsbezogener Kompetenzen sind eine zuverlässige Erfassung zur Identifikation von Bedarfen sowie gezielte Förderungsmaßnahmen von entscheidender Bedeutung (KMK, 2021). Orientierung dafür bieten Kompetenzmodelle wie das „European Framework for the Digital Competence of Educators“ (DigCompEdu), das pädagogische und didaktische Kompetenzen zur unterrichtlichen Integration digitaler Medien beschreibt (Redecker & Punie, 2017). Die ICAP-Taxonomie (interactive, constructive, active, passive) ergänzt diese kompetenzbasierte Perspektive um eine differenzierte Betrachtung (digitalgestützter) kognitiver Aktivierungsprozesse, die das DigCompEdu im Bereich der aktiven Einbindung Lernender beschreibt (Chi & Wylie, 2014).

Für die Erfassung (digitalisierungsbezogener) Kompetenzen haben sich Vignetten als geeignet erwiesen (Brovelli et al., 2014; Lachner et al., 2021). Im Gegensatz zu Selbsteinschätzungen erfordern unterrichtsbezogene Vignetten die Beurteilung der Angemessenheit von Unterrichtsstrategien und die Auseinandersetzung mit berufsspezifischen Problemsituationen (Neuweg, 2015; Rutsch & Dörfler, 2018).

Durch eine gezielte Erfassung digitalisierungsbezogener Kompetenzen können Defizite identifiziert und bedarfsorientierte Förderungsmaßnahmen gestaltet werden. Eine Möglichkeit für die Förderung von Kompetenzen bieten videobasierte Simulationen. Diese bilden reale (unterrichtsspezifische) Situationen reduziert nach und ermöglichen ein praxisnahes Handeln (Cook, 2014). Trotz ihrer Potentiale wurde videobasierte Simulationen bisher jedoch kaum zur Förderung digitalisierungsbezogener Kompetenzen von Lehrkräften eingesetzt.

Vor diesem Hintergrund bringt das vorliegende Symposium vier Beiträge zu aktuellen empirischen Befunden zur kompetenzorientierten Erfassung und Förderung digitalisierungsbezogener Kompetenzen von Lehrkräften zusammen.

Die ersten beiden Beiträge basieren auf dem theoretischen Hintergrund des DigCompEdu-Modells. Der *erste Beitrag* beschreibt Teilstudien zur Validierung eines entwickelten vignettenbasierten Instruments zur Erfassung digitalisierungsbezogener Kompetenzen von Lehrkräften im Bereich Lehren und Differenzierung. Der *zweite Beitrag* thematisiert die Förderung digitalisierungsbezogener Kompetenzen angehender Lehrkräfte mithilfe eines videobasierten Beobachtungstools, in welchem der digitale Medieneinsatz analysiert wird.

Die beiden weiteren Beiträge basieren auf dem ICAP-Modell. Der *dritte Beitrag* stellt den aktuellen Entwicklungsstand eines objektiven Messinstruments vor, mit dem die Fähigkeiten von Lehrkräften erfasst werden, das Potential der Einbindung digitaler Medien für die Förderung der kognitiven Aktivierung von Lernenden einzuschätzen. Der *vierte Beitrag* thematisiert die Validierung einer simulationsbasierten Lernumgebung zur Förderung der Diagnosekompetenz in Bezug auf die kognitive Aktivierung im digitalgestützten Unterricht.

Zusammenfassend bietet das Symposium eine Synthese von Erkenntnissen der situativen Erfassung und Förderung digitalisierungsbezogener Kompetenzen auf Grundlage des DigCompEdu- und ICAP-Modells. Das Symposium leistet somit einen Beitrag zum Austausch über Forschungsansätze vor dem Hintergrund etablierter Kompetenzmodelle. In der anschließenden Diskussion erfolgt eine kritische Auseinandersetzung mit den dargelegten Ansätzen und es werden zukünftige Forschungsrichtungen definiert.

Beiträge des Symposiums

Erfassung von Kompetenzen von Lehrkräften zum Unterrichten mit digitalen Medien: Validierung eines textbasierten Vignetteninstruments basierend auf dem DigCompEdu

Isabell Runge¹, Jakob Thomas Fischer¹, Florian Hebibi¹, Frauke Voitle², Moritz Krell², Rebecca Lazarides³, Katharina Scheiter¹

¹Universität Potsdam, ²Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, ³Universität Potsdam, Technische Universität Berlin

Theoretischer Hintergrund

Lehrkräfte benötigen digitalisierungsbezogene Kompetenzen, um digitale Medien effektiv und didaktisch sinnvoll einzusetzen (Krumsvik, 2014; Vilppola et al., 2022). Ein etabliertes Modell zur Beschreibung digitalisierungsbezogener Kompetenzen von Lehrkräften ist das European Framework for the Digital Competence of Educators (DigCompEdu-Modell; Redecker & Punie, 2017). Das Zentrum des Kompetenzrahmens bilden vier pädagogische und didaktische Kompetenzdimensionen zur fachübergreifenden Integration digitaler Medien. Der vorliegende Beitrag fokussiert die Subdimension Lehren, die als grundlegende Kompetenz das Gestalten, Planen und Einsetzen von digitalen Medien beschreibt. Zusätzlich wird die Subdimension Differenzierung einbezogen, da der Einsatz digitaler Medien insbesondere Potentiale zur Förderung unterschiedlicher Lernbedürfnisse und Lernwege aufweist (Aleven et al., 2017; Schaumburg, 2021). Die bisher wenigen empirischen Forschungsarbeiten zur Operationalisierung des DigCompEdu beziehen sich vorrangig auf Selbsteinschätzungen von Lehrkräften (Antonietti et al., 2022; Lucas et al., 2021). Studien diskutieren jedoch deren Eignung zur Erfassung digitalisierungsbezogener Kompetenzen (Scherer et al., 2017). Alternativ hat sich der Einsatz textbasierter Vignetten zur Erfassung digitalisierungsbezogener Kompetenzen bei (angehenden) Lehrkräften als geeignet erwiesen (Lachner et al., 2021; Schulze-Vorberg & Horz, 2023). Dabei wurden jedoch bisher keine spezifischen digitalisierungsbezogenen Kompetenzdimensionen wie im DigCompEdu berücksichtigt.

Zielstellungen

Die entlang des DigCompEdu entwickelten Vignetten werden im Rahmen von drei Teilstudien überprüft. Ziel der ersten Teilstudie mit Lehrkräften ist es, Unklarheiten hinsichtlich der Verständlichkeit zu ermitteln sowie die Authentizität der Unterrichtssituationen zu prüfen. Ziel der zweiten Teilstudie mit Expert*innen ist es, die Passung der Vignetten zum DigCompEdu zu ermitteln. In der dritten Teilstudie werden die überarbeiteten Vignetten mit einer größeren Stichprobe an Lehrkräften erprobt sowie psychometrisch analysiert. Zum Zeitpunkt der GEBF-Tagung liegen die vollständigen Daten aller Studien vor.

Methode

Es wurden 24 Vignetten entwickelt, die Unterrichtssituationen zum Einsatz digitaler Medien in den Bereichen Lehren und Differenzierung beschreiben. Aus vier vorgegebenen Antwortmöglichkeiten soll diejenige ausgewählt werden, die zur Förderung des formulierten Stundenziels am besten beiträgt. In den ersten zwei Teilstudien werden schriftliche Rückmeldungen zur Konstruktpassung, Verständlichkeit sowie Authentizität der Vignetten zur Einschätzung der Inhaltsvalidität qualitativ ausgewertet. Für die Itemanalyse in der dritten Teilstudie werden die psychometrischen Itemeigenschaften hinsichtlich ihrer Itemschwierigkeit und Trennschärfe mithilfe der Item-Response-Theorie evaluiert, indem ein 2-parametrisches logistisches Modell für dichotome Items durchgeführt wurde. Nach Befragungsabschluss im September 2024 werden zudem bivariate und multivariate Verfahren zur Ermittlung von Zusammenhängen mit digitalisierungsbezogenen Kompetenzselbsteinschätzungen zur Überprüfung der Konstruktvalidität durchgeführt.

Perspektivische Ergebnisse und ihre Bedeutung

Die Vignetten wurden inhaltlich von sechs Lehrkräften sowie neun Expert*innen überarbeitet. Die Einschätzung der Lehrkräfte zur Verständlichkeit lag bei $M = 3.40$ ($SD = 0.48$; 1 = unverständlich; 6 = verständlich) und die Einschätzung der Authentizität lag bei $M = 4.69$ ($SD = 1.01$; 1 = unauthentisch; 6 = authentisch). Die Expert*innen haben die Passung der Vignetten zur DigCompEdu-Subdimension Lehren durchschnittlich mit $M = 3.42$ ($SD = 0.94$) eingeschätzt und die Passung zur Differenzierung mit $M = 2.81$ ($SD = 0.90$; 1 = unpassend; 4 = passend). Die Passung der Vignetten zur Erfassung digitalisierungsbezogener Kompetenzen wurde mit $M = 2.67$ ($SD = 0.87$) eingeschätzt. Im Zuge der Überarbeitungen aus der Lehrkräfte- und Expert*innen-Befragung wurden vier Vignetten aufgrund unzureichender Einschätzungen hinsichtlich ihrer Passungen ausgeschlossen. Daraus resultieren je zehn Vignetten zu beiden DigCompEdu-Subdimensionen, die in der dritten Teilstudie getestet werden. Im Folgenden werden vorläufige Ergebnisse der Angaben von 418 Lehrkräften aus Brandenburg vorgestellt. Ergebnisse der Itemanalyse zeigen, dass die Itemschwierigkeit der Vignetten zu Lehren von -0.85 bis 13.08 und zu Differenzierung von -2.41 bis 7.56 reichte. Die Trennschärfe lag für Lehren zwischen -0.02 und 4.67 und für Differenzierung zwischen -0.86 und 6.74. Die vorläufigen Ergebnisse deuten darauf hin, dass vier Vignetten möglicherweise zu schwierig sind und nicht zuverlässig zwischen Personen mit unterschiedlichen Fähigkeitsniveaus differenzieren. Finale Ergebnisse zur Itemanalyse und zur Konstruktvalidität werden auf der GEBF berichtet.

Stichworte: digitalisierungsbezogene Kompetenzen, digitale Medien, Vignetten, DigCompEdu, Instrumentenentwicklung

Entwicklung und Validierung eines Beobachtungstools für digitale Kompetenzen von Lehrkräften

Melissa Oezsoy¹, Julia Murböck¹, Meral Roeben², Elisabeth Bauer³, Michelle Stannard⁴, Florian Schultz-Pernice¹, Michael Sailer³, Frank Fischer¹

¹Ludwig-Maximilians-Universität München, Department für Psychologie, ²Ludwig-Maximilians-Universität München, Department für Psychologie, Ludwig-Maximilians-Universität München, Department für Anglistik und Amerikanistik, ³Universität Augsburg, Learning Analytics and Educational Data Mining, ⁴Ludwig-Maximilians-Universität München, Department für Anglistik und Amerikanistik

Eine effektive Nutzung digitaler Medien im Unterricht erfordert digitale Fähigkeiten und Wissen von Lehrpersonen (Mishra & Koehler, 2006; Sailer et al., 2021). Zu deren Erwerb ist eine kontinuierliche Förderung essenziell (Fernández-Batanero et al., 2020). Im Kompetenzrahmen DigCompEdu (Redecker et al., 2017) werden diese digitalen Kompetenzen in sechs Bereiche unterteilt, darunter Lehren und Lernen mit digitalen Medien. Der DigCompEdu bietet eine Grundlage zur systematischen Erfassung (Runge et al., 2023) und Förderung digitaler Kompetenzen. Um den Kompetenzerwerb von Lehrpersonen zu fördern, können Annäherungen an die Praxis (Grossmann et al., 2009), z.B. in Form von Unterrichtsvideos (Farell et al., 2024), mit Instrumenten zur systematischen Wahrnehmung kombiniert werden (Seidel & Stürmer, 2014), da Wissen in realitätsnahen Situationen angewendet wird.

Der DigCompEdu Observe ist ein digitales Tool, das auf dem DigCompEdu basiert und zur kriteriengeleiteten Beobachtung und Analyse digitaler Kompetenzen bei Lehrpersonen in Unterrichtsvideos dient. Bei der Konzeption des DigCompEdu Observe wurde der DigCompEdu durch Ergänzung durch pädagogisch-psychologischer Evidenz operationalisiert, um verschiedene Ausprägungen digitaler Kompetenzen zu definieren. Der DigCompEdu Observe soll die professionelle Wahrnehmung und somit wichtige digitale Lehrkompetenzen durch das Identifizieren und Analysieren des digitalen Medieneinsatzes in Videos fördern. Die inhaltliche Validität wurde durch die Orientierung an pädagogisch-psychologischer Evidenz sichergestellt. Das Tool sowie die Videos von digitalen Englischunterrichtsszenarien wurden von sechs Expert*innen evaluiert, überarbeitet und in eine Lernplattform integriert.

Zur Überprüfung der Konstruktvalidität und Kriteriumsvalidität werden folgende Forschungsfragen und Hypothesen untersucht: Inwiefern lässt sich die faktorielle Struktur des Kompetenzbereichs Lehren und Lernen mit digitalen Medien des DigCompEdu Observe empirisch bestätigen? Es wird vermutet, dass sich die sechs entwickelten Subskalen des Kompetenzbereichs Lehren und Lernen mit digitalen Medien des DigCompEdu Observe abbilden lassen und auf einen gemeinsamen Faktor laden. Inwieweit hängt technologisch-pädagogisches Wissen mit der Performanz im DigCompEdu Observe zusammen? Es wird angenommen, dass ein signifikanter Zusammenhang zwischen technologisch-pädagogischem Wissen und der Performanz besteht.

Die Konstruktvalidität sowie die Kriteriumsvalidität wurden anhand einer Studie mit 77 angehenden Lehrpersonen (67 % weiblich, 30 % männlich) mit dem Unterrichtsfach Englisch überprüft. Das Durchschnittsalter der Teilnehmenden betrug 23,47 Jahre ($SD = 4,30$). Die Studie beinhaltete die Identifizierung und Analyse des digitalen Medieneinsatzes in vier Videos, die durch den DigCompEdu Observe unterstützt wurden.

Die Konstruktvalidität wurde anhand einer konfirmatorischen Faktorenanalyse für ein Video überprüft. Jede Subskala wurde durch einen latent modellierten Faktor repräsentiert. Um die Konstruktvalidität des DigCompEdu Observe zu sichern, wurden Kriterien mit negativen und Faktorladungen $\beta < 0.300$ ausgeschlossen. Infolgedessen wurden ein Item sowie eine Subskala exkludiert. Die Ergebnisse ergaben, dass alle Faktoren auf einen übergeordneten Faktor laden. Das reduzierte Modell wies einen guten Modellfit auf ($\chi^2(83) = 97.285$; $p = .135$; $RMSEA = .051$; $CFI = .957$; $TLI = .945$; $SRMR = .068$; Hu & Bentler, 1999).

Die Kriteriumsvalidität wurde durch die Korrelation der Performanz im DigCompEdu Observe mit dem technologisch-pädagogischen Wissen überprüft. Das technologisch-pädagogische Wissen konnte nicht reliabel gemessen werden und es zeigte sich keine signifikante Korrelation mit der Performanz im DigCompEdu Observe.

Die Struktur des DigCompEdu konnte in das Beobachtungs- und Analysetool DigCompEdu Observe überführt und etliche Validitätskriterien wie Inhaltsvalidität und Strukturvalidität konnten abgedeckt werden. Auf der Grundlage der aktuellen Ergebnisse können weitere Analysen durchgeführt werden, um festzustellen, welche Facetten von DigCompEdu bereits erfolgreich umgesetzt werden konnten – diese Ergebnisse werden auf der Konferenz vorgestellt. Die Kriteriumsvalidität muss in weiteren Studien noch etabliert werden, da sich die erhobene Skala zur Überprüfung der Kriteriumsvalidität als nicht reliabel erwiesen hat. Limitationen liegen in der Überprüfung des Tools mit Studierenden und einer kleinen Stichprobe. Die Schwierigkeit der Praxisrepräsentationen wurde von Expert:innen eingeschätzt, allerdings bleibt offen, wie schwierig Studierende die Praxisrepräsentation wahrnahmen. In weiteren Studien soll der DigCompEdu Observe in weiteren Fachdidaktiken wie Biologie und Mathematik, eingesetzt werden.

Stichworte: Digitale Lehrkompetenz, DigCompEdu, Beobachtungstool, Professionelle Wahrnehmung, Validierung

Ein Test zur Erfassung der ICAP-basierten Analysefähigkeit von Lehrkräften bezüglich des kognitiv aktivierenden Einsatzes digitaler Medien im Unterricht

Annika Goeze¹, Emely Hoch², Michael Sailer³, Tim Fütterer¹

¹Universität Tübingen, ²Leibniz-Institut für Wissensmedien, Tübingen, ³Universität Augsburg

Theoretischer Hintergrund

Digitale Medien haben großes Potenzial dafür, Lehr-Lern-Settings lernwirksam auszugestalten (Scheiter, 2017). Die Potenziale entfalten sich jedoch nicht automatisch, sondern müssen durch eine didaktisch sinnvolle Orchestrierung in einem qualitativ hochwertigen Unterricht von Lehrkräften ausgeschöpft werden (Fütterer et al., 2022; 2023a). Eine wichtige Voraussetzung dafür ist, dass Lehrkräfte über technologisch-pädagogisches Wissen (TPK) verfügen – also beispielsweise das Wissen darüber, wie digitale Medien eingesetzt werden können, um kognitive Prozesse beim Lernen zu fördern (Mishra & Koehler, 2006; Koehler et al., 2014). Ein Modell, das darauf fokussiert, ist das Interactive-Constructive-Active-Passive-(ICAP)-Modell (Chi & Wylie, 2014). Es geht davon aus, dass bestimmte beobachtbare Lernaktivitäten mit höherwertigerer kognitiver Aktivierung einhergehen, welche es als Lehrkraft zu analysieren und zu initiieren gilt. Um dies erfassen und entsprechende Interventionsmaßnahmen wie beispielsweise Lehrkräftefortbildungen passend zu Vorkenntnissen anbieten zu können, braucht es eine gute Diagnostik. Die Erfassung von Wissensfacetten wie TPK erfolgte bisher überwiegend mittels Selbstauskünften (Koehler et al., 2012; Willermark, 2018), die Nachteile haben, z.B. die Anfälligkeit für Verzerrungen bei der eigenen Fähigkeits-Bewertung (Fütterer et al., 2023). Objektive Tests zur TPK-Erfassung, die psychometrisch überzeugen, sind jedoch rar.

Situational Judgment Tests (SJT) sind besonders geeignet, Analysefähigkeiten und insbesondere das Verhalten und die Entscheidungsfindung von Personen beispielsweise in beruflichen Situationen zu erfassen (Christian et al., 2010).

Ziele

In diesem Beitrag stellen wir den aktuellen Stand der Entwicklung eines Rasch-skalierten, vignettenbasierten SJT vor, der darauf abzielt, objektiv erfassen zu können, wie gut Lehrkräfte (basierend auf ihrem ICAP-Modell-Wissen) für unterschiedliche Unterrichtsszenarien in der Lage sind, zu analysieren, welche Art der Einbindung digitaler Medien mit höherer Wahrscheinlichkeit Schüler:innen kognitiv aktiviert. Im Zentrum stehen die Forschungsfragen, wie zuverlässig und valide der Test ist.

Methode

Zunächst wurde ein Itempool aus 40 single-choice Items mit jeweils vier Antwortoptionen entwickelt, die sich auf unterschiedliche Unterrichtssituationen (Vignetten) beziehen, bei denen (digitale) Medien eingesetzt werden. Jede ICAP-Dimension wurde mit jeweils zehn Items abgedeckt. Danach wurde sowohl systematisches Feedback von drei ICAP-Modell-Expert:innen und vier Lehrer:innen eingeholt, um die Itemformulierungen anzupassen. Die überarbeiteten Items wurden in einer Pilotierungsstudie mit N = 135 Lehrkräften (79 angehende und 56 Lehrkräfte im Beruf) eingesetzt, um eine Itemauswahl basierend auf Infit- und Outfit-Werten vornehmen zu können. Anschließend haben wir eine Validierungsstudie mit N = 219 Lehrkräften im Beruf (70% weiblich, M = 11,45 Jahre Berufserfahrung) durchgeführt und dabei u.a. einen Test zur technologischen Bedienfertigkeit (T-TK; Fütterer et al., 2023b) und eine ICAP-Skala (Antonietti et al., 2023) als Validierungsinstrumente eingesetzt. Schließlich haben wir eine Think-Aloud-Studie mit 12 Erwachsenen (davon zehn Lehrkräfte im Beruf und zwei Professionsfremde) durchgeführt, um tiefere Einblicke in die kognitiven Prozesse bei der Itembearbeitung zu erhalten.

Ergebnisse

Um jede ICAP-Dimension mit je drei Items abzudecken, haben wir basierend auf Infit und Outfit-Werten der Pilotierungsstudie und inhaltlicher Abwägungen 12 aus den 40 Items ausgewählt, bei denen die Wright-Map eine relativ gleichmäßige Abdeckung der Aufgabenschwierigkeit- / Personenfähigkeitenintervalle gezeigt hat und zugleich (insbesondere für Lehrkräfte im Beruf, für die der Test entwickelt wird) eine angemessene Reliabilität vorliegt (KR20 = .66, RELAndrich = .64, nur für Lehrkräfte RELAndrich = .72). Die Ergebnisse der Validierungsstudie haben die gute Verteilung der Items Aufgabenschwierigkeit- / Personenfähigkeitenintervalle bestätigt. Während sich hier erwartungsgemäß keine Zusammenhänge zu T-TK-Testwerten ($r = .04$) zeigten, haben sich entgegen unseren Erwartungen allerdings nur eine geringe Reliabilität und geringe Zusammenhänge zur ICAP-Skala gezeigt (bspw. $r = .17$). Basierend auf den Erkenntnissen der Think-Aloud-Studie (z.B. Schwächen in Formulierungen der Items, unzureichende Konkretion von Unterrichtssituationen, Zwiespalt von ICAP-basiertem Analysieren vs. eigenem Vorgehen in der Praxis) steht aktuell eine gründliche Überarbeitung der 12 Items an. Ab November 2024 planen wir eine weitere Validierungsstudie für die überarbeiteten Items durchzuführen.

Stichworte: ICAP, Vignetten, Digitale Medien, kognitive Aktivierung, Situational Judgment Test

Förderung medienbasierter Diagnosekompetenzen hinsichtlich kognitiver Aktivierung: Validierung einer Simulation und Analyse ihrer Fälle

Meral Roeben¹, Johanna Vejvoda¹, Julia Murböck¹, Nicole Heitzmann¹, Florian Schultz-Pernice¹, Michael Sailer², Matthias Stadler¹, Frank Fischer¹

¹Ludwig-Maximilians-Universität München, ²Universität Augsburg

Kognitive Aktivierung spielt eine wichtige Rolle bei der effizienten Gestaltung des Lernprozesses (Chi & Wylie, 2014; Schrader, 2013), insbesondere im Kontext digitaler Medien (Stegmann, 2020). Um digitale Medien kognitiv aktivierend einzusetzen, benötigen Lehrpersonen entsprechende digitale Kompetenzen - unter anderem Diagnosekompetenz (Schultz-Pernice et al.,

2017). Wir konzeptualisieren Kognitive Aktivierung mithilfe des ICAP-Modells mit den vier Stufen passiv, aktiv, konstruktiv, interaktiv (Chi & Wylie, 2014). Beim Planen von Unterricht sollte die angestrebte Stufe der kognitiven Aktivierung von Lernziel mit dem Arbeitsauftrag übereinstimmen. Bei der Unterrichtsdurchführung kann die Stufe entweder auf Basis der beobachtbaren Lernendenaktivität bestimmt oder muss zusätzlich mithilfe der Lernprodukte inferiert werden (Chi & Wylie, 2014). Aktuell fehlt systematische Forschung dazu, welche Stufen schwerer zu diagnostizieren sind. Zur Förderung der komplexen Kompetenz des Diagnostizierens kognitiver Aktivierung, eignen sich Simulationen. Simulationen bieten eine gewisse Annäherung an die Unterrichtspraxis bei gleichzeitiger Reduktion der Komplexität (Grossman, 2009; Chernikova, 2020). Ziel unserer Studie war die simulationsbasierte Förderung von Diagnosekompetenz hinsichtlich kognitiver Aktivierung im Kontext eines digitalgestützten Unterrichts und in diesem Zuge auch die systematische Erfassung der Schwierigkeit, verschiedene Stufen zu diagnostizieren. Bevor Simulationen als Lernumgebungen eingesetzt werden können, müssen sie validiert werden. Ein gängiges Validierungskriterium ist, ob das Vorwissen prädiktiv für die Performanz in der Simulation ist (Radkowsitch et al., 2020).

Daraus ergaben sich folgende Forschungsfragen:

FF1: Inwieweit ist das Vorwissen von Lehramtsstudierenden prädiktiv für ihre Performanz in der Simulation?

FF2: Inwiefern wird die Diagnoseschwierigkeit der Stufen kognitiver Aktivierung davon beeinflusst, ob die Stufen kognitiver Aktivierung inferiert werden müssen und welche Phasen des Unterrichtens (Planen und Durchführen) vorliegen?

In einer Onlinestudie bearbeiteten $N = 213$ Lehramtsstudierende die Simulation Digivate. Ihr Vorwissen erfassten wir mit einem Single-Choice-Test des konzeptuellen Wissens zum ICAP-Modell (Cronbachs Alpha = 0.67). Die Proband:innen schlüpften in die Rolle von Referendar:innen und bearbeiteten Fälle, die ihnen randomisiert zugewiesen wurden. Die Simulation ist in zwei Phasen unterteilt: Die Planungs- (sechs Fälle) und die Durchführungsphase (fünf Fälle). In der Planungsphase erhielten die Teilnehmer:innen fertige Unterrichtspläne und bestimmten die potenziell erreichbare Stufe kognitiver Aktivierung des Lernziels und des Arbeitsauftrags. In der Durchführungsphase sahen sie videobasierte Bildschirmaufnahmen von Schüler:innenaktivitäten und diagnostizierten die Stufe kognitiver Aktivierung der Lernenden.

Wir untersuchten FF1 mit einer linearen Regression, die unser Validierungskriterium unterstützt, da Vorwissen prädiktiv ($\beta = .454$) für die Performanz in der Simulation ist ($(F(1, 211) = 54.65, p < .05)$), wobei Vorwissen 21% der Varianz in der Performanz erklärt ($R^2 = .21$).

Für FF2 nutzen wir Konfusionsmatrizen. So stellten wir die tatsächlich diagnostizierten Stufen kognitiver Aktivierung der akkuraten Stufe gegenüber. Grundsätzlich ist die Übereinstimmung der tatsächlich diagnostizierten Stufe mit der akkuraten immer am höchsten. Jedoch zeigen sich Schwierigkeitsunterschiede zwischen den Stufen. Über beide Phasen hinweg ($\chi^2(1) = 177.36, p < .05$) und in der Planungsphase (Lernziel: $\chi^2(1) = 34.09, p < .05$, Arbeitsauftrag: $\chi^2(1) = 183.05, p < .05$) werden aktiv und konstruktiv am häufigsten verwechselt, in der Durchführungsphase konstruktiv und interaktiv ($\chi^2(1) = 177.36, p < .05$). Das impliziert, dass Stufen, die inferiert werden müssen, schwerer zu diagnostizieren sind. Eine leicht höhere Sensitivität in der Durchführungsphase sowie das Ergebnis, dass unterschiedliche Stufen verwechselt wurden, unterstützt eine unserer Annahmen, dass sich die Phasen in ihrer Schwierigkeit unterscheiden.

Hieraus kann die Implikation abgeleitet werden, dass für das Verständnis von Inferenzprozessen (konstruktiv, interaktiv) Unterstützung benötigt wird, wofür das Potenzial von Simulationen weiter genutzt werden kann, indem gezielt und datengestützt Aspekte der Fälle manipuliert werden.

Unsere Ergebnisse geben wichtige Hinweise auf die Validität der Simulation sowie Implikationen für eine Weiterentwicklung von Digivate. Durch das Untersuchen der Schwierigkeiten der Stufen kognitiver Aktivierung erhielten wir zudem ein tieferes Verständnis des ICAP-Modells, aus dem wir relevante theoretische Annahmen ableiten können.

Stichworte: Digitale Lehrkompetenz, Diagnosekompetenz, Simulationen, Validierung, Diagnoseschwierigkeit

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 29: Digitale Souveränität im Schreibunterricht: Ein simulationsbasiertes Fortbildungsprogramm für Lehrkräfte zur Schreibdiagnostik und Schreibförderung

Daria Podwika, Anita Schilcher, Johannes Wild

Universität Regensburg, Deutschland

Theoretischer Hintergrund

Angeichts der rasanten Entwicklung einer „Kultur der Digitalität“ (Stalder, 2016) ermöglichen digitale Medien die Entwicklung hochwertiger und interaktiver Fortbildungsprogramme, die adaptiv auf das Vorwissen der Lehrkräfte abgestimmt werden können (vgl. Blossfeld et al., 2019) und somit einen Beitrag zum Transfer innovativer Konzepte leisten können. Dies ist gerade auch bei komplexen Kompetenzen wie der Schreibkompetenz eine Möglichkeit, wo nach wie vor traditionelle Vorgehensweisen und Methoden verwendet werden. Innovative Formate werden nur langsam eingeführt (vgl. Sieberer, 2016). Dies verdeutlicht die dringende Notwendigkeit, neue didaktische Ansätze zügig in den Unterricht zu integrieren.

Die Regensburger Arbeitsgruppe im Verbundprojekt „Digitale Souveränität als Ziel wegweisender Lehrer:innenbildung für Sprachen, Gesellschafts- und Wirtschaftswissenschaften in der digitalen Welt“ (DiSo-SGW) hat das Ziel, eine simulationsbasierte Fortbildung zu entwickeln, zu evaluieren und umzusetzen, die die schreibdidaktischen Kompetenzen und Überzeugungen von Lehrkräften stärkt. In der Projektgruppe wird eine digitale Fortbildungsplattform zur Schreibdidaktik entwickelt, die auf Forschungsergebnissen der Schreibpsychologie und der empirischen Schreibdidaktik fußt. Der Fokus liegt dabei auf der Entwicklung von Diagnosekompetenz und dem Wissen zur Förderung der Schreibkompetenz von Schülerinnen und Schülern. Gleichzeitig soll die digitale Fortbildungs- und Lernumgebung nicht nur fachliches Wissen vermitteln, sondern Verständnis dafür aufbauen, wie eine effektive digitale Lernumgebung gestaltet sein sollte.

Die digitale Fortbildung, die den Schreibunterricht aus der Perspektive der Lehrkraft simuliert (Fischer et al., 2022), umfasst vier Module, die sich an zentralen Core Practices orientieren und diese auf die Schreibdidaktik anwenden: (1) Diagnose von Schülertexten, (2) Sequenz- und Unterrichtsplanung, (3) Entwicklung von Schreiblernumgebungen und (4) konstruktives Feedback. Ziel ist es, die Plattform so zu gestalten, dass Lehrkräfte ihren eigenen Lernpfad wählen können. Je nach Vorwissen und Interesse können sie entweder praxisnahe oder vertiefte theoretische Verarbeitungsmodi, um sich das nötige Hintergrundwissen anzueignen. Im Modul zur Diagnostik beispielsweise betreten Lehrkräfte ein simuliertes Klassenzimmer, in dem Schülerinnen und Schüler an einer argumentativen Schreibaufgabe arbeiten. Lehrkräfte können einzelne Schülerinnen und Schüler auswählen, deren Gedanken (Müller & Trautwein, 2015) und Texte analysieren sowie Videos zur Schreibflüssigkeit (Odersky, 2018) betrachten. Für die effektive Bewertung argumentativer Texte wurden spezifische Kriterien wie Funktion, Formulierung, Textstrukturierung und Rechtschreibung definiert, die in vier Kompetenzstufen unterteilt sind. Ergänzend werden in Erklärvideos Fallbeispiele für Bewertungsverfahren modelliert. Das zweite Modul vermittelt Lehrkräften, wie sie erfolgreich Unterrichtssequenzen planen können. Es stellt anhand des Textmusters Erzählen das SRSD-Modell nach Graham & Harris (2005) vor und konzentriert sich auf Erzählstrategien wie Ereignis, Figur und Situation (Wild, Schilcher & Pissarek, 2018) und deren Vermittlung im Unterricht. Lehrkräfte lernen, Unterrichtseinheiten zu strukturieren und verschiedene didaktische Maßnahmen, wie etwa die Vermittlung von Schreibstrategien, zu implementieren, um den Schreibunterricht effektiver zu gestalten. Auch dieses Modul wird durch Erklärvideos, Informationstexte und interaktive Elemente unterstützt.

Fragestellung

Daraus lässt sich folgende Fragestellung ableiten: Wie wirksam und nachhaltig ist eine simulationsbasierte, digitale Fortbildung in der Verbesserung der schreibdidaktischen Kompetenzen und Überzeugungen von Lehrkräften?

Methode

Die Entwicklung aller vier Module ist bis Ende Juli 2024 abgeschlossen, sodass sie anschließend in eine Lernplattform implementiert werden können. Die Erprobung der Fortbildung beginnt im Herbst 2024 und wird wissenschaftlich begleitet. Die Evaluationsstudie wird im Prä-Post-Follow-Up-Design durchgeführt, wobei allgemeine Fragen zur Wirksamkeit von Lehrkräftefortbildungen (vgl. Lipowsky & Rzejak 2019) mit fachspezifischen und themenspezifischen Fragestellungen kombiniert werden. Dabei kommt ein Fragebogen mit verschiedenen Items zum Einsatz. Basierend auf dem DPACK-Modell (Döbeli-Honegger, 2021) werden insbesondere die Inhaltskompetenz (IK), die digitale Inhaltskompetenz (DIK) sowie die pädagogische Inhaltskompetenz (PIK) evaluiert. Zusätzlich sind qualitative Datenerhebungen in Form von Interviews oder Gruppendiskussionen mit Lehrkräften geplant.

Ergebnisse und ihre Bedeutung

Obwohl zum gegenwärtigen Zeitpunkt noch keine Ergebnisse vorliegen, werden bis Januar 2025 erste Einblicke in die Wirksamkeit der Fortbildung erwartet. Diese vorläufigen Ergebnisse können helfen, erste Schlüsse zur Effektivität der entwickelten Module zu ziehen und mögliche Anpassungen und Verbesserungen der Fortbildungsprogramme zu diskutieren.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 40: Partizipative Interventionsentwicklung: Fokusgruppendifkussionen mit Jugendlichen zur Stärkung des Freizeit-Engagements in FabLabs und Makerspaces

Kathrin Smolarczyk¹, Marios Mouratidis², Sophie Uhing³, Rolf Becker¹, Stephan Kröner³

¹Hochschule Rhein-Waal, Deutschland; ²Universität Siegen; ³Friedrich-Alexander-Universität Erlangen-Nürnberg

Theoretischer Hintergrund

Die Anzahl von Makerspaces und FabLabs nimmt weltweit immer weiter zu. Dabei handelt es sich um Räume in denen Menschen aller Altersgruppen digitale und physische Technologien kombinieren und kreativ arbeiten, um Ideen zu entwickeln, technische Fertigkeiten zu erlernen und neue Produkte zu kreieren (Sheridan et al., 2014) sowie um ihre eigenen Probleme zu lösen (Gershenfeld, 2005). Solche Making-Aktivitäten, die als kreativer Prozess der Herstellung eines Produkts verstanden werden, auch in Kombination mit Programmierung (Sefton-Green, 2013), haben das Potenzial, Kinder zu empowern und ihre Möglichkeiten zu verbessern, an der Gesellschaft und der Digitalisierung teilzuhaben und diese mitzugestalten (Hartikainen, Ventä-Olkkonen, Kinnula & Iivari, 2021). Trotz des großen Potentials besteht jedoch noch großer Forschungsbedarf zu Making-Aktivitäten, um Ungleichheiten zu verringern und den Zugang zu Making-Aktivitäten für Kinder mit unterschiedlichem Hintergrund zu erleichtern (Bar-El & Worsley, 2021; Smolarczyk & Kröner, 2023).

Fragestellung

Um Erkenntnisse darüber zu gewinnen, wie junge Menschen, die sich zuvor noch nicht mit Making-Aktivitäten in Makerspaces oder FabLabs beschäftigt haben, motiviert werden können, insbesondere auch Mädchen und Jugendliche mit niedrigem sozioökonomischen Status (Dreessen, Schepers & Leen, 2016; Melo, March, Hirsh & Arnsberg, 2023; Tomko, Alemán, Newstetter, Nagel & Linsey, 2021), untersuchte die vorliegende Studie ko-konstruktiv die wahrgenommenen Hürden und Hilfen für Making-Aktivitäten von Kindern und Jugendlichen und band sie in einen partizipativen Prozess ein.

Methode

Wir führten acht Fokusgruppen mit insgesamt $N = 61$ Jugendlichen (52% weiblich, 48% männlich; Alter: $M = 12,65$, Spanne = 11-16) durch. Um sowohl Mädchen als auch Jungen und Jugendliche mit unterschiedlichem sozioökonomischem und Migrationshintergrund in die Studie einzubeziehen, wurde bewusst kein außerschulischer Workshop angeboten, zu dem sich die Jugendlichen selbst anmeldeten, sondern eine Kooperation mit einer nahe gelegenen Schule, die den Besuch des FabLabs in einen regulären Schultag integrierte.

Ergebnisse und ihre Bedeutung

Die identifizierten Hürden für Jugendliche, sich an Making-Aktivitäten zu beteiligen, waren ein unklarer Zweck, irreführende Annahmen bezüglich der Kosten, frustrierende frühere Erfahrungen mit dem Thema ökologische Nachhaltigkeit, Zeitmangel, die Vermeidung elterlicher Beteiligung und ein Mangel an Selbstwirksamkeit. Zu den förderlichen Faktoren für Jugendliche, sich an Making-Aktivitäten zu beteiligen, gehörten ein wahrgenommener intrinsischer Wert, die Möglichkeit, Dinge für andere herzustellen, die FabLab-Umgebung, die Verfügbarkeit zusätzlicher Optionen zu Hause, die Möglichkeit, Gegenstände individuell zu gestalten und zu personalisieren, die Möglichkeit, an Spielen und Videos zu arbeiten, Empowerment-Prozesse im Hinblick auf ökologische Nachhaltigkeit sowie Gefühle von Autonomie, sozialer Eingebundenheit und die Nutzung von Schulen für Werbung.

Weitere neue Erkenntnisse wurden im Hinblick auf Empowermentprozesse gewonnen. Dazu gehören die persönliche Sinnhaftigkeit (Iivari, 2020; Kinnula, Sánchez Milara, Norouzi, Sharma & Iivari, 2021), das Schaffen neuer, individuell relevanter Produkte (Sheridan et al., 2014), das Lösen persönlich relevanter Probleme (Katterfeldt, Dittert & Schelhowe, 2015), das "Machen" für andere oder das Inspirieren anderer sowie die Bedeutung von Autonomie und Selbstwirksamkeit. Die Tatsache, dass die Jugendlichen selbst diese Aspekte als wichtig für Making-Aktivitäten bezeichneten, ist eine neue Erkenntnis und sehr vielversprechend für die Förderung von Empowerment-Prozessen bei Making-Aktivitäten. Die Studie verfolgte einen innovativen Ansatz, indem sie die Jugendlichen in die Gestaltung der Aktivitäten einbezog. Dieser Ansatz lieferte Erkenntnisse darüber, wie sehr man Jugendlichen zutrauen kann, zum Gestaltungsprozess beizutragen, und wie gut sie die Funktionen und Grenzen verschiedener Gestaltungsaspekte verstehen, zum Beispiel ihre Fähigkeit, die Wirksamkeit von Werbung in sozialen Medien kritisch zu beurteilen. Wir präsentieren somit ein Beispiel für einen explorativen Ansatz zur Einbeziehung von Jugendlichen in den Gestaltungsprozess von Making-Aktivitäten, der die Kluft zwischen dem Erreichen aller Kinder in formalen Settings in der Schule und der Untersuchung nur derjenigen, die bereits in außerschulischen Settings involviert sind, überwindet. Weitere Implikationen für die Forschung und Praxis werden diskutiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 33: Modeling and Measuring Critical Online Reasoning in Higher Education – First-year Students' Results in Four Disciplines

Olga Troitschanskaia¹, Dimitri Molerov¹, Jennifer Fischer¹, Lukas Trierweiler¹, Paul Hodes¹, Lisa Martin de los Santos Klein¹, Johannes Hartig²

¹Johannes Gutenberg Universität Mainz, Deutschland; ²DIPF Leibniz Institut für Bildungsforschung und Bildungsinformation Frankfurt, Deutschland

Theoretical background. Students in higher education (HE) use the Internet daily to obtain information, complete assignments, or prepare for exams [9]; university students indicate using the Internet more frequently than learning materials supplied by their lecturers (e.g., (online) textbooks; [10]). Nevertheless, they frequently lack the skills required for successful Internet-based learning. They employ poor search, selection, and evaluation strategies, often using deficient learning information [8;12]. Current research associates skill deficits with the specific challenges students face when learning online. They must process and synthesize information from multiple sources, and determine accuracy from algorithmically preselected, predominantly unverified, didactically ill-suited, and partly biased sources [7;11;12;13].

Thus, effective self-directed learning with web sources requires a skillset among students to search for, and assess the quality (e.g., credibility) of online information they found and employ for their learning, defined as critical online reasoning (COR) [3], comprising three overlapping cognitive facets: online information acquisition, critical information evaluation, and reasoning based on evidence, arguments, and synthesis.

Funded by the German Research Foundation (DFG) since 2023, the interdisciplinary research unit “Anonymized”, consisting of nine projects, developed COR assessments to validly measure the required skillset in the real Internet environment. [5] The first of four measurements of a longitudinal panel took place in winter 2023-24.

Question

The overarching question of the poster presentation is: How did first-semester students in the four disciplines (economics, sociology, medicine, physics) perform on the generic (GEN) and domain-specific (DOM-COR) Critical Online Reasoning assessment tasks at the beginning of their studies?

Method. For the representative sampling, an initial survey was administered (N=2,458) to obtain the population of all first-semester students in these four disciplines at the three main participating institutions (and supplemental sampling at over four more institutions). A purposeful sample for a longitudinal panel across disciplines was drawn, with N=471 participants completing the GEN- and DOM-COR assessment. The sample size, combined with the Bayesian model estimation methods, allowed for the inclusion of all theoretically expected person and institutional-context control variables (e.g. socio-demographics, educational backgrounds, lectures and courses attended, grades, general and domain-specific information use) as well as test-induced and course-learning effects, including study design effects (e.g., multiple measurements, missing-by-design distribution).

Every participant received two DOM-COR tasks and one GEN-COR task or vice versa in a virtual assessment platform, self-administered at home over two months. Breaks and logouts were allowed between tasks, for a maximum of three online sessions, and a maximum test-taking time of 3h (30 min of additional covariates including reading, intelligence, domain-specific knowledge tests, and socio-demographics). The distribution of the COR tasks across students and measurements is based on the booklet design with linked behavioral indicators to allow for Item Response Theory equating.

After assessment completion, a rating of student responses by trained scorers has been performed for the GEN- and DOM-COR tasks, ensuring the interrater reliability of Krippendorff's alpha over 0.8.

Results/Discussion. We present results on students' GEN- and DOM-COR skills from the first measurement across involved projects, and discuss key lessons learned, e.g., regarding students' use of AI-chatbots like GhatGPT to solve the COR tasks and implications for valid measuring of COR skills in the real Internet environment.

The examination of COR skills within and across disciplines based on the performance data and method-integrative analysis of response process data, tracked while students solved tasks (including log data, visited websites) proves to be innovative and impactful, given the growing importance of online learning. In the methods-integrative framework, the COR skills and the online sources students use are analyzed in conjunction and contribute to establishing a technologically and methodologically innovative research of key importance for the next decades.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 41: Gefährdungsbeurteilung psychischer Belastung an Schulen – Ein Konzept des Arbeitsmedizinischen Instituts für Schulen in Bayern

Daniel Ossenschmidt, Simone Sailer, Meike Sons, Victoria Heumann, Stöhr Sarah, Tina Tischer

Landesamt für Gesundheit und Lebensmittelsicherheit, Deutschland

Lehrkräfte sind an Schulen verschiedenen psychischen Belastungen ausgesetzt, die sich ungünstig auf die Gesundheit auswirken können. Darunter fallen zum Beispiel schwieriges Schülerverhalten, schwieriges Verhalten der Eltern oder Zeitmangel (Robert Bosch Stiftung, 2024).

Gemäß § 5 Arbeitsschutzgesetz (ArbSchG) sind Schulleitungen zur Beurteilung der für die Beschäftigten mit ihrer Arbeit verbundenen Gefährdung verpflichtet. Dabei ist der Frage nachzugehen, inwiefern vorhandene psychische Arbeitsbelastungen die Gesundheit der Lehrkräfte gefährden und wie die Belastungssituation reduziert werden kann. Um staatliche Schulen in Bayern bei der Durchführung der GBU psychischer Belastung zu unterstützen, hat das Arbeitsmedizinische Institut für Schulen in Bayern (AMIS-Bayern) als beauftragte Fachbehörde ein leitlinienkonformes Verfahrenskonzept in Anlehnung an der Gemeinsamen Deutschen Arbeitsschutzstrategie (GDA, Arbeitsprogramm Psyche, 2022) und der Bundesanstalt für Arbeits- und Betriebsmedizin (BAuA, 2021) entwickelt.

Im Rahmen des Konzeptes werden Schulen in den einzelnen GBU-Phasen - bestehend aus der Prozessvorbereitung, der Gefährdungsanalyse und dem Maßnahmenmanagement - begleitend unterstützt. Hierzu werden spezielle Handlungshilfen sowie ein selbst entwickeltes Erhebungsinstrument zur Belastungsermittlung in Form eines Online-Fragebogens mit 127 Items bereitgestellt. Zur Maßnahmenableitung und -management unterstützt AMIS-Bayern mit Handlungsempfehlungen und stellt zudem Planungs- und Umsetzungshilfen zur Verfügung.

Im Rahmen einer seit 2022 laufenden Evaluationsstudie wird das Konzept und das Ermittlungsinstrument in Bezug auf Praxistauglichkeit untersucht. Grundlage sind die Daten von bisher 27 Schulen verschiedener Schulformen mit insgesamt 851 Lehrkräften, die das Verfahren zur Gefährdungsbeurteilung psychischer Belastung des AMIS-Bayerns genutzt haben. Als Teil der jeweiligen Befragung zur Gefährdungsermittlung und zur Maßnahmenwirksamkeit beantworteten die Lehrkräfte zusätzlich Evaluationsfragen zum Verfahren und zum Ermittlungsinstrument. Gefragt wird unter anderem, ob das Ermittlungsinstrument die Belastungssituation vollständig erfasst und die Items gut verständlich sind. So fiel die Evaluation des Fragebogens beispielsweise mit Zustimmungsraten zwischen 81 und 91 Prozent bei den Items positiv aus.

Im Rahmen des Posterbeitrages werden sowohl das entwickelte Verfahrenskonzept zur GBU psychischer Belastung als auch die damit verbundenen Evaluationsergebnisse vorgestellt und diskutiert.

Pinnwand 42: Erfolgreicher Schulstart

Bianca Strohmaier

Pädagogische Hochschule Schwäbisch Gmünd, Deutschland

Fragestellung und Zielsetzung

Ziel dieses Projektes ist es zu zeigen, dass gezielte, auf die Einschulungsuntersuchung abgestimmte, additive Förderung im Vorschulalter auch längerfristig wirksam ist und so Bildungsungleichheiten im schulischen Verlauf entgegenwirken kann.

In Baden-Württemberg ist die Teilnahme an der Einschulungsuntersuchung verpflichtend, um gesundheitliche Risiken und eventuelle Entwicklungsverzögerungen frühzeitig zu erkennen, so dass sie bei Bedarf rechtzeitig gefördert bzw. gezielt behandelt werden können (Ministerium für Soziales, Gesundheit und Integration Baden-Württemberg 2023). Je nach Termin der Untersuchung ist 12-24 Monate vor der Einschulung der Förderbedarf der Kinder bereits erkannt und bekannt. Entsprechend der Ergebnisse wird häuslicher Förderbedarf oder gezielte Sprachförderung im Anschluss an die Einschulungsuntersuchung empfohlen. Ob es zu einer Sprachförderung tatsächlich kommt hängt allerdings von den Kapazitäten des Kindergartens ab und ist daher nicht zwingend der Fall. Der häusliche Förderbedarf wird den Eltern zwar empfohlen, jedoch wird die Umsetzung weder begleitet noch unterstützt oder kontrolliert.

Innerhalb dieses Forschungsprojektes werden Kinder, welche bereits bei der Einschulungsuntersuchung Auffälligkeiten in den Fächern Mathematik und Deutsch aufweisen bis zu ihrer Einschulung additiv in den Kindertagesstätten bis zu zweimal wöchentlich in Kleinstgruppen mit 2-3 Kindern über einen Zeitraum von ca. 15 Monaten in sprachlichen und mathematischen Kompetenzen gefördert.

Die geförderten Kinder werden bis zum Übergang in die Sekundarstufe I begleitet. Sie, ihre Lehrkräfte und die Eltern werden in regelmäßigen Abständen befragt.

Ihre schulischen Leistungen am Übergang zur Sekundarstufe werden anhand des IQB-Bildungstrends mit Kindern aus ähnlichem sozio-ökonomischen Hintergrund verglichen.

Hinweise auf die Effektivität der additiven Förderung während des letzten Kindergartenjahrs sollen in dieser Längsschnittstudie die Übergangsquoten auf die einzelnen Schulformen an der Schwelle zur Sekundarstufe sowie die Leistungsvergleiche mit entsprechenden IQB-Daten in Klassenstufe vier liefern.

Hintergrund und Forschungsstand

Wie der IQB-Bildungstrend 2021 gezeigt hat ist der soziale Gradient in Baden-Württemberg in den letzten 10 Jahren eher gestiegen und erreicht beispielsweise in der Orthographie den höchsten Wert im Ländervergleich (Stanat et al. 2022). Die PISA-Studie und weitere Forschungen haben gezeigt, dass Kinder aus sozial schwachen Familien an allen Gelenkstellen im Bildungssystem benachteiligt sind (Baumert 2011; Anders und Maaz 2010). Der erste Übergang vom Elementarbereich in den Primarbereich stellt bereits wichtige Weichen. Selbst der Zeitpunkt der Einschulung ist vom sozio-ökonomischen Hintergrund der Eltern abhängig (Kratzmann und Schneider 2009). Die leistungsbezogene Heterogenität beim Schuleintritt ist sowohl im Fach Deutsch als auch in der Mathematik sehr groß (Dollinger 2014) und sowohl vom Migrationsstatus als auch vom sozioökonomischen Hintergrund des Elternhauses abhängig. Kenntnisse der Unterrichtssprache sind beispielsweise ein guter Prädiktor für den späteren Schulerfolg (Dollmann 2019). Die Dauer der besuchten Kindertageseinrichtung vor Schuleintritt kann

die Abhängigkeit etwas reduzieren (Kratzmann und Schneider 2009). In Hamburg ist das letzte Schuljahr vor geplanter Einschulung für Kinder mit Förderbedarf im Bereich Sprache beispielsweise obligatorisch. Entweder besuchen die Kinder im Jahr vor der Einschulung eine Vorschule oder sie erhalten eine additive Förderung in der Kindertageseinrichtung. Dies erscheint in Anbetracht der sich weitenden Schere mit Beginn der Einschulung als lohnenswert. Die leistungsbedingte Heterogenität, welche bei der Einschulung festzustellen ist, vergrößert sich nach Schuleintritt (Dollinger 2014). So ist anzunehmen, dass sich die herkunftsbedingten Unterschiede im Laufe der Schulzeit vergrößern was an der Schwelle zur Sekundarstufe deutlich wird (Stanat et al. 2022; Baumert 2011; Schulze und Unger 2009; Paulus und Blossfeld 2007). Um eine Gymnasialempfehlung zu bekommen müssen Kinder aus bildungsfernen Milieus sogar bessere Leistungen erbringen als ihre MitschülerInnen aus bildungsnahen Elternhäuser (Bos 2004).

Eine gezielte Förderung in frühen Jahren sorgt somit nicht nur für mehr Bildungsgerechtigkeit sondern erweist sich auch als langfristige Investition in die Zukunft, wie bereits in mehreren Studien gezeigt werden konnte (Kahl und Barenboim 2006; Belfield et al. 2006).

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 7: Transferbegriff im Wandel? Analyse von Veränderungen des Transferverständnisses in deutschsprachigen Fachzeitschriften

Caroline J. Scherer^{1,2}, Ulrike Hartmann^{1,2}, Hannah Kleen^{1,2}, Katharina Vogel^{1,3}, Mareike Kunter^{1,2,4}

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; ²IDeA-Center for Research on Individual Development and Adaptive Education of Children at Risk; ³Humboldt-Universität zu Berlin; ⁴Goethe-Universität Frankfurt am Main

Theoretischer Hintergrund

Fragen des Transfers von Forschungswissen in die Bildungspraxis haben in den vergangenen Jahren zunehmend an Relevanz gewonnen (Bieber et al., 2018). Mit der wachsenden Forderung nach Evidenzbasierung im Bildungsbereich (Slavin, 2002) einhergehend sind auch Herausforderungen für gelingenden Transfer zwischen Bildungsforschung und -praxis in den Fokus gerückt (Coburn, 2003). Auf nationaler Ebene macht sich der anhaltende Bedeutungszuwachs der Transferthematik in der Bildungsforschung z.B. durch vermehrte Diskussion in wissenschaftlichen Publikationen bemerkbar, u.a. in entsprechenden Themenheften deutschsprachiger Zeitschriften (Thiel & Rott, 2022). Wissenschaft-Praxis-Transfer wird zudem von gegenwärtigen Projektverbünden wie *lernen:digital* als zentrale Zielsetzung definiert (BMBF, 2023). In der Auseinandersetzung mit der Thematik wird allerdings schnell deutlich, dass der Transferbegriff keineswegs einheitlich definiert und verwendet wird, sondern das Begriffsverständnis je nach Zusammenhang (implizit) variiert (Gräsel, 2010). So spiegelt das Transferverständnis die jeweiligen Annahmen zur Beziehung von Bildungsforschung und Schulpraxis wider – ein Spannungsverhältnis, das Gegenstand vielfältiger Debatten ist (Hartmann & Kunter, 2022).

In bildungswissenschaftlichen Diskursen ist ein Verständnis von Transfer als lineare Übertragung von Forschungswissen in die Praxis weit verbreitet (Mohajezad & Schrader, 2022) – im Sinne eines Sender-Empfänger Modells (Schmiedl, 2022) oder einer ‚Einbahnstraße‘ (Gräsel, 2019). Dieser lineare Transferansatz wird jedoch zunehmend kritisiert (Schmiedl, 2022), während beziehungsorientierte und systemische Ansätze verstärkt Beachtung finden (Blatter & Schelle, 2022), welche Interaktion und Kommunikation zwischen Forschung und Praxis betonen (Hartmann & Kunter, 2022). Dieser Annahme eines Bedeutungswandels stehen allerdings Hinweise auf ein fortbestehendes lineares Transferverständnis gegenüber, das beispielsweise in der Forschungsliteratur häufig (implizit) zugrunde gelegt wird (Blatter & Schelle, 2022).

Ziel der Studie

Vor diesem Hintergrund verfolgt unsere Studie das Ziel, Veränderungen des Transferbegriffs bzw. des Transferverständnisses in der Bildungsforschung empirisch zu erfassen. Das Poster stellt die Ergebnisse einer automatisierten Textanalyse vor, mit der die Verwendung des Begriffs „Transfer“ in wissenschaftlichen Artikeln deutschsprachiger Fachzeitschriften explorativ untersucht wurde. Dabei liegt der Fokus einerseits auf der Erfassung von Begriffsvariationen (z.B. Wissenschaft-Praxis-Transfer), andererseits werden begriffliche und thematische Zusammenhänge identifiziert und im Zeitverlauf betrachtet, um so mögliche Veränderungen des Transferverständnisses nachzuvollziehen.

Methode

Die Datenbasis der automatisierten Textanalysen bilden theoretische und empirische Beiträge deutschsprachiger Fachzeitschriften, die der Bildungsforschung zugeordnet werden können (voraussichtlich *Zeitschrift für Erziehungswissenschaft*, *Zeitschrift für pädagogische Psychologie*, *Die Deutsche Schule*, *Unterrichtswissenschaft* und *Zeitschrift für Bildungsforschung*) in den Jahren 2000-2024. Als Textkorpus für die nachfolgenden Analysen werden aus den Volltexten zunächst sämtliche Passagen extrahiert, die die Zeichenkette „transfer“ enthalten. Auf dieser Grundlage wird im ersten Schritt die Verwendung des Transferbegriffs bezogen auf den Gesamtzeitraum untersucht, wofür unterschiedliche Variationen des Begriffs und deren jeweilige Häufigkeit ermittelt werden. Im nächsten Schritt werden Veränderungen im Zeitverlauf fokussiert: Zunächst wird durch eine Wortfrequenzanalyse bestimmt, wie sich im Laufe der Jahre die Anzahl der Nennungen des Transferbegriffs entwickelt hat. Anschließend werden für Zeitabschnitte von jeweils fünf Jahren durch die Bildung von Bigrammen charakteristische Wortkombinationen mit dem Begriff „Transfer“ identifiziert, um so wesentliche Begriffe und deren Veränderungen wahrzunehmen. Im letzten Analyseschritt wird schließlich *Topic Modeling* angewendet mit dem Ziel, Themenfelder im Textkorpus zu identifizieren. Dabei wird das Erscheinungsjahr der Artikel als Kovariate einbezogen, um auch zeitliche Entwicklungen der Themen zu ermitteln. Die Ergebnisse der Analysen werden zur Tagung vorliegen.

Ausblick

Die Studie kann zeigen, wie in der deutschsprachigen Bildungsforschung der Transferbegriff verwendet wird und Bedeutungsveränderungen des Begriffs – sowohl quantitativer als auch inhaltlicher Art – sichtbar machen. Die Ergebnisse ermöglichen so einen Einblick in die (impliziten) Annahmen und Vorstellungen Forschender im Zusammenhang mit Transfer.

Pinnwand 13: Inklusives Experimentieren – Erhebung potenzieller Barrieren und didaktische Antworten mittels Vignetten

Nathalie Beck, Mathias Ropohl, Helena van Vorst

Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Die UN-Behindertenrechtskonvention fordert seit 2006 neben der gesellschaftlichen Teilhabe für alle Menschen die Reform der Lehrkräftebildung, um Lehrkräfte auf die Anforderungen einer inklusiven Bildung vorzubereiten (UNESCO, 2005). Entsprechend hat das Land NRW Lehrinhalte bzgl. des konstruktiven Umgangs mit Inklusion in den Inhalten der Lehramtsstudiengänge verankert (MSW NRW, 2023). Im Bereich der Naturwissenschaften ist eine Vorbereitung auf inklusiven Unterricht in der Lehrkräftebildung noch immer selten und nicht zufriedenstellend entwickelt (Bosse & Spörer 2014; Schlüter, 2018).

Die Implementierung von Experimenten im inklusiven Unterricht als eine zentrale Arbeitsweise ist ein wichtiges Ziel (Gierl & Kauertz, 2009; Windt, 2011; Windt, 2017), die konkrete Umsetzung jedoch noch nicht ausreichend untersucht (Menthe & Hoffmann 2015; Mester, 2019). Erste theoretische Grundlagen bzgl. der Umsetzung von Experimenten im inklusiven Unterricht zeigen, dass das Erkennen von Barrieren und das Ermöglichen von Partizipation wichtige Gelingensbedingungen darstellen (Gonzales et al., 2021; Stinken Rösner et al., 2021).

Vor diesem Hintergrund stellt die Identifizierung der für die Arbeitsweise des Experimentierens möglichen Barrieren sowie die Untersuchung didaktischer Handlungsweisen zum Abbau der Barrieren ein Forschungsdesiderat dar.

Ziel

Das Ziel des Promotionsvorhabens ist die Identifizierung möglicher Barrieren sowie die Erhebung methodischer Handlungsweisen zur Ermöglichung der Partizipation von Schüler:innen mit den Förderschwerpunkten „Lernen“ sowie „emotionale und soziale Entwicklung“ beim Experimentieren im inklusiven Sach- und Chemieunterricht. Das Promotionsvorhaben fokussiert auf die sonderpädagogischen Förderschwerpunkte „Lernen“ sowie „emotionale und soziale Entwicklung“, da Schüler:innen mit diesen Förderschwerpunkten die beiden größten Gruppen von Schüler:innen mit sonderpädagogischem Förderschwerpunkt bilden (KMK, 2024).

Methodik und Design

Zunächst wurden die Förderschwerpunkte „Lernen“ sowie „emotionale und soziale Entwicklung“ mittels der vom Land NRW vorgegebenen Diagnoseaspekte für die Feststellung von Förderbedarfen voneinander abgegrenzt. Anschließend wurde ausgehend von den für das Land NRW vorgegebenen Diagnoseaspekten und mit Hilfe eines Literaturreviews eine Übersicht der jeweiligen Förderschwerpunkte und die dazugehörigen Verhaltensweisen und möglichen Diagnosen in ihren verschiedenen Ausprägungen ausgearbeitet. Zusätzlich wurden die für Unterricht im Allgemeinen bekannten Barrieren abgebildet. Die erstellte Übersicht wurde den fachlichen Anforderungen des Experimentierens gegenübergestellt, um das Potenzial möglicher Barrieren in verschiedenen Kombinationen aus Voraussetzungen und Anforderungen abzubilden.

Die Erhebung möglicher Barrieren und partizipationsermöglichender Handlungsweisen wird durch den Einsatz von Vignetten realisiert. Vignetten sind „kurze, authentische Darstellungen von Situationen aus dem schulischen Fachunterricht“ (Rutsch & Seidenfuß, 2017, S. 490) und besitzen Diagnosepotenzial, das Fachwissen sowie professionelle Kompetenzen von (angehenden) Lehrkräften zu erheben (von Aufschnaiter et al., 2017). Die genutzten Vignetten enthalten Szenen typischer Unterrichtssituationen des Experimentierens im Sach- und Chemieunterricht und wurden auf Basis realer Unterrichtsvideos erstellt. Die dargestellten Szenen wurden durch Schüler:innenprofile mit den jeweiligen Förderschwerpunkten ergänzt. Die Auswahl der Szenen und Schüler:innenprofile basiert auf dem zuvor ausgearbeitetem Potenzial für Barrieren innerhalb der jeweiligen Förderschwerpunkte.

Die erarbeiteten Vignetten werden mit Hilfe eines Onlinebefragungstools erfahrenen sonderpädagogischen Lehrkräften aus NRW ($N \geq 60$) zur Bearbeitung vorgelegt. Durch den Einsatz eines offenen Antwortformats werden die Lehrkräfte darum gebeten für die dargestellten Schüler:innen die in den dargestellten Szenen bekannten Barrieren sowie ihnen bekannte Handlungsweisen zum Abbau der dieser Barrieren zu benennen. Die Antworten werden nach der strukturierten qualitativen Inhaltsanalyse (Mayring, 2022) ausgewertet. Als Ergebnis werden den Experimentierphasen zuzuordnende, förderschwerpunktspezifische Barrieren und didaktische Lösungen erwartet.

Ergebnisse und Ausblick

Im Rahmen des Projekts wurde eine Übersicht über die Förderschwerpunktspezifischen schülerseitigen Voraussetzungen mit Bezug auf das Experimentieren als naturwissenschaftliche Arbeitsweise erstellt. Diese Übersicht ermöglicht die Identifizierung exkludierender Aspekte des Experimentierens im Unterricht. Die auf dieser Basis ausgearbeiteten Vignetten ermöglichen das Erheben des Wissens um mögliche Barrieren beim Experimentieren sowie Handlungsweisen, mit denen Barrieren abgebaut werden können. Die Erhebung findet aktuell statt, so dass Ergebnisse im Rahmen des Posterbeitrags präsentiert werden können. Aufbauend auf den gewonnenen Erkenntnissen erfolgt eine Erarbeitung und Evaluation von Vignetten zur Lehrkräftebildung an der Universität Duisburg-Essen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 43: Zahlendreher und Zehnerübergang - Eine Analyse schwierigkeitsgenerierender Merkmale der Items zur Messung mathematischer Kompetenzen in der ersten Klasse

Nils Goldboom¹, Jasmin Walter², Jörg-Tobias Kuhn³, Daniel Sommerhoff², Natalie Förster¹

¹Bergische Universität Wuppertal; ²Leipniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik; ³Technische Universität Dortmund

Theoretischer Hintergrund

Diagnostische Informationen über die individuelle Lernentwicklung von Schüler*innen stellen für Lehrkräfte eine wichtige Entscheidungsgrundlage für Unterrichts Anpassungen dar. So zeigen Klassen in denen eine lernverlaufs begleitende Diagnostik im Sinne des Formativen Assessments (Black & Wiliam, 1998, 2009) durchgeführt wurde, stärkere Leistungszuwächse als Klassen, in denen lediglich statusdiagnostische Informationen zum Lernstand vorliegen (Förster & Souvignier, 2014, 2015). Die Lernentwicklungen werden dabei klassischerweise mit Paralleltests erfasst. Insofern stellt die Äquivalenz der einzelnen Tests einer Testreihe eine notwendige Voraussetzung für die sinnvolle Interpretation der Leistungsveränderungen dar.

Um äquivalent schwierige Testitems für Paralleltests einer Testreihe zu konstruieren, können Itemmerkmale gezielt bei der Konstruktion von Aufgaben berücksichtigt werden. Wenn bekannt ist, welche Merkmale die Beantwortung der Aufgaben beeinflussen, können diese systematisch kombiniert werden, so dass mehr oder weniger schwierige Aufgaben entstehen. Gleichzeitig sollte die Anwendung derselben Konstruktionsregeln auch zu gleich schwierigen Aufgaben führen. Diese Form der regelgeleiteten Konstruktion von Testitems hat sich für verbale, figurale und numerische Aufgaben als vielversprechend gezeigt (Anderson et al., 2022; Förster & Kuhn, 2023; Freund et al., 2008; Holling et al., 2009; Varnauer et al., 2022). Zudem lässt sich die Konstruktvalidität der Items im Hinblick auf kognitionspsychologische Theorien überprüfen (Embretson, 1998), wobei vertiefende Einblicke möglich sind, wenn für die einzelnen Aufgaben nicht nur Informationen über die Akkuratheit der jeweiligen Antwort, sondern zudem die benötigten Beantwortungszeiten vorliegen.

Basierend auf entwicklungspsychologischen Modellen (Dehaene, 1997, 2001, 2011; Krajewski, 2008) sowie fachdidaktischen Untersuchungen (Benz, 2007) mathematischer Kompetenzen sollten Merkmale wie z.B. der Zehnerübergang, die Größe des Zahlenraums oder die geforderte mathematische Operation die Beantwortung der Aufgaben beeinflussen. Aufgaben sollten also z. B. schwieriger sein und mehr Antwortzeit benötigen, wenn sie einen Zehnerübergang bei der Rechenoperation erfordern (Anderson et al., 2022).

Fragestellung

Die vorliegende Studie untersucht anhand der Daten von $N = 4377$ Erstklässlern mit der Konstruktvalidität sowie der Äquivalenz zwei wesentliche psychometrische Gütekriterien von Verfahren zur Lernverlaufsdiagnostik im Fach Mathematik. Konkret untersuchen wir,

- a) ob die bei der Testkonstruktion der Items berücksichtigten schwierigkeitsgenerierenden Merkmale wie angenommen die Akkuratheit und/oder die Antwortzeit beeinflussen und
- b) ob die Berücksichtigung der Konstruktionsregeln wie angenommen in äquivalent schwierigen Paralleltests resultiert.

Methode

Die Daten wurden mit dem web-basierten System zur Lernverlaufsdiagnostik *quop* erhoben (Souvignier et al., 2021). In *quop* werden im Laufe eines Schuljahres Paralleltests im dreiwöchigen Abstand durchgeführt. Ausgewertet werden die integrierten Daten der Mathematiktestreihe für die ersten Klassen (*quop*-M1) aus den letzten drei Prä-Corona-Kohorten aus den Jahren 2016/17 – 2018/19. Insgesamt gehen die Daten von $N = 4377$ Kindern (47.29% weiblich; 6.99 Jahre alt) in die Analyse ein. Die Reliabilität, Sensitivität und prädiktive Validität der Testreihe wurde bereits in einer früheren Arbeit untersucht (Salaschek & Souvignier, 2013).

Alle Aufgaben der Testreihe wurden regelgeleitet konstruiert (Salaschek & Souvignier, 2013). Für alle Items liegt neben der Information über die Akkuratheit der Antwort auch die benötigte Antwortzeit in Millisekunden vor. Die Fähigkeit und Antwortgeschwindigkeit auf Personenseite sowie die Itemschwierigkeit und Antwortlatenz auf Aufgabenebene, werden simultan mittels hierarchischer IRT-Modelle geschätzt (Förster & Kuhn, 2023; Klein Entink et al., 2009). Der Effekt der postulierten Itemmerkmale auf die Itemschwierigkeiten und Antwortlatenzen wird dann mithilfe von linearen Regressionsmodellen analysiert. Die Überprüfung der Äquivalenz der Paralleltests erfolgt über den Vergleich der Testinformationsfunktionen, Korrelationen zwischen den Paralleltests sowie Varianzanalysen.

Ergebnisse und ihre Bedeutung

Aktuell wird die Studie im Sinne von Open Science präregistriert, weswegen noch keine Ergebnisse vorliegen. Die Ergebnisse der Studie werden Erkenntnisse zur Konstruktvalidität und Äquivalenz der Testreihe liefern. Darüber hinaus werden die Einblicke in die Relevanz verschiedener schwierigkeitsgenerierender Itemmerkmale zu einem vertieften Verständnis mathematischer Prozesse bei Kindern in der Schuleingangsphase beitragen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001

Chair der Sitzung: Karina Karst

Pinnwand 2: Untersuchung relevanter Kompetenzen von Schüler:innen zum Erkennen von und Handeln in komplexen Systemen im Kontext von Nachhaltigkeit

Silja Herholz, Mathias Ropohl

Universität Duisburg-Essen, Deutschland

Bildung für nachhaltige Entwicklung (BNE) betrachtet Phänomene und Prozesse aus der Perspektive von Nachhaltigkeit, um nachfolgende Generationen zu zukunftsorientiertem Handeln zu befähigen (UNESCO, 2017). Das Fach Chemie leistet dazu einen Beitrag, da Aspekte von Nachhaltigkeit auf chemischen Konzepten beruhen. Allerdings werden im Chemieunterricht Inhalte auf Basiskonzepte reduziert, die ein kumulatives Lernen ermöglichen, jedoch für die Betrachtung komplexer Sachverhalte zu stark vereinfacht sind (MSB NRW, 2022). Nachhaltigkeitsthemen werden daher nicht vollumfänglich im Sinne von BNE betrachtet und Kompetenzen in Verbindung mit Systemen werden im Chemieunterricht aktuell nur indirekt adressiert (MSB NRW, 2022). Ein konzeptübergreifender Ansatz, der klassische chemische Themen mit Zielen von Nachhaltigkeit verknüpft, ist notwendig (Orgill et al., 2019; Whalen et al., 2022). Systemkompetenz ist ein solcher Ansatz, der verschiedene Fähigkeiten zum Umgang mit komplexen Systemen vereint und bereits im Zuge von PISA 2025 und von BNE als eine Schlüsselkompetenz für Nachhaltigkeit identifiziert wurde (Arnold & Wade, 2015; UNESCO, 2017; White et al., 2023). Die Forschung dazu im Fach Chemie konzentriert sich bisher auf den Hochschulbereich und die wenigen Studien im Schulkontext fokussieren auf die Umsetzung im Lehrplan. Erste Untersuchungen weisen darauf hin, dass Systemkompetenz bereits in der Schule erworben werden kann (Budak & Ceyhan, 2024) und dass Fachwissen und kognitive Fähigkeiten Prädiktoren für den Erwerb von Systemkompetenz sein könnten (Assaraf & Orion, 2005; Mambrey et al., 2020). Zur Überprüfung dieser Vermutungen, fehlt aktuell ein Testinstrument.

Vor diesem Hintergrund zielt das Promotionsvorhaben darauf ab, den Kompetenzstand von Lernenden der Sekundarstufe II (Sek. II) zum Erkennen von und Handeln in komplexen Systemen im Fach Chemie mit einem schriftlichen Testinstrument zu erheben. Es wird die Forschungsfrage untersucht, inwieweit Lernende der Sek. II Fähigkeiten der Systemkompetenz im Kontext Wasser erworben haben. Wasser als Kontext vereint Nachhaltigkeitsthemen mit Chemiebezug, zum Beispiel Wasserverschmutzung, Trinkwassergewinnung oder Wasserstoffproduktion. Eine zweite Forschungsfrage untersucht den Einfluss von Fachwissen und kognitiven Fähigkeiten auf den Erwerb von Systemkompetenz in dieser Population.

Zur Beantwortung der Fragen umfasst das Projekt drei Schritte. Im ersten Schritt erfolgte eine Literaturrecherche zur Synthese von Merkmalen chemischer Systeme und die Entwicklung eines Kompetenzmodells von Systemkompetenz für das Fach Chemie. Das an bisherigen Modellierungen (Rempfler & Uphues, 2011) anknüpfende Kompetenzmodell umfasst die Dimensionen *Systemorganisation* und *Systemverhalten*. Auf Grundlage der Strukturkomplexität von Systemen werden drei Kompetenzstufen für beide Dimensionen unterschieden. Für jede Kompetenzstufe wurden dimensionsbezogen spezifische Fähigkeiten adaptiert (vgl. York & Orgill, 2020). Zusätzlich wurde ein Flussdiagramm mit den Merkmalen chemischer Systeme erstellt, das durch eine logische Verknüpfung der Merkmale das Erstellen oder Auswählen passender Systeme im nächsten Projektschritt erleichtert. Dieser erste Projektschritt ist abgeschlossen und die Ergebnisse dieses Schritts liegen vor. Der zweite Schritt, der aktuell bearbeitet wird, umfasst die Konzeptionierung von chemischen Systemen, die anschließend für die Entwicklung von Testaufgaben zur Messung der Fähigkeiten aus dem Kompetenzmodell genutzt werden. Die Systeme werden als *System Maps* konzipiert, da diese einzelne Systemelemente und deren Verbindungen darstellen, und anschließend durch eine Expertengruppe auf Konstrukt- und inhaltliche Validität geprüft. Als Format für die Testaufgaben wird eine Variante von Multiple Choice Aufgaben, die auf System Maps basieren, adaptiert (Mambrey et al., 2020). So werden insgesamt 30 Testitems mit dem Kontext Wasser entstehen. Die Testitems werden auf Testhefte verteilt. Jedes Testheft wird 18 Items enthalten; zu jeder Kompetenz drei Items. Zusätzlich werden Fachwissen und kognitive Fähigkeiten als Kontrollvariablen erhoben, um die zweite Forschungsfrage beantworten zu können. Die Testaufgaben werden im dritten Schritt des Projekts validiert und zur Messung von Systemkompetenz von Lernenden der Sek. II bearbeitet. Es wird eine Gesamtstichprobe von $N = 500$ bis 600 Lernenden angestrebt, da so im Sinne der probabilistischen Testtheorie eine ausreichende Anzahl von Antworten pro Testitem sichergestellt werden kann (Rost, 2004).

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 44: Behalten von Texten: Propositionsanalyse vereinfacht durch ChatGPT

Helmut Niegemann¹, Gesine Seeber²

¹Universität des Saarlandes/Goethe-Universität Frankfurt, Deutschland; ²Technische Universität Berlin

Die Erfassung des Behaltens und Verstehens von Textinformationen erfolgt häufig durch geschlossene Testformate, die jedoch in mehrfacher Hinsicht problembehaftet sind: Zum einen kann meist ein Wiedererkennungseffekt nicht ausgeschlossen werden, zum anderen können entsprechende Tests sehr lang und damit ermüdend werden, wenn es auf Nuancen ankommt, zum Beispiel inwieweit Details der Textinformationen (u.a. Adjektive, Adverbien) aus unterschiedlichen Passagen eines Texts unterschiedlich reproduziert werden.

Studien zum Behalten und Verstehen von Texten sind dann darauf angewiesen die freie Wiedergabe durch Lernende exakt mit dem Ausgangstext zu vergleichen. Da unterschiedliche syntaktische Formen gleiche semantische Bedeutungen repräsentieren können, bietet sich für den Vergleich die Zerlegung sowohl des jeweiligen Ausgangstexts als auch aller von Lernenden reproduzierten Texte nach identischen Regeln an. Seit ca. 50 Jahren wird in der Textpsychologie dazu die Analyse der Texte in Propositionen verwendet. Propositionen sind kleinste Bedeutungseinheiten von Aussagen.

Ein Problem der Verwendung solcher Propositionenanalysen ist, dass die Auswertung nicht wie bei geschlossenen Tests durch Schablonen erfolgen kann, sondern sehr zeitaufwendig „händisch“ von geschulten Personen durchgeführt werden muss, zur Gewährleistung der Objektivität möglichst durch jeweils zwei Auswertende. Dieser Aufwand kann dazu führen, dass zulasten der Gütekriterien eher geschlossene Testformen gewählt werden.

Mit der zunehmenden Leistungsfähigkeit generativer Künstlicher Intelligenz (z.B. ChatGPT) stellt sich die Frage, ob diese Zerlegung von Texten in Propositionen nicht mit erheblichem Effizienzgewinn automatisiert erfolgen kann. Entscheidend ist bei solchen Anwendungen generativer KI das geeignete Prompting, d.h. die Entwicklung entsprechender Anweisungen an das System, den gewünschten Output zu liefern.

Hierzu sind Dokumente, die sowohl theoretische Überlegungen zum Konzept von Propositionen als auch Beispiele zur Zerlegung kurzer Texte enthalten geeignet. Ein solches Dokument kann in ChatGPT mit dem Befehl hochgeladen werden, sich auf die gleiche Art und Weise die Propositionserstellung anzueignen.

Chat GPT liest das Dokument in wenigen Sekunden ein und der zu analysierende Text kann mit der Aufforderung (Prompt), ihn entsprechend dem Beispieldokument in Propositionen zu zerlegen, hochgeladen werden. Dieses Vorgehen wurde mit Lehrtexten erprobt.

Das KI-System erstellt eine Liste mit den kleinsten Informationseinheiten gemäß der Darstellung des anleitenden Dokumentes. Eine Kontrolle der Propositionen auf Vollständigkeit und Schlüssigkeit sollte durch methodisch geschulte Personen erfolgen. Chat GPT erwies sich mit dem von uns entwickelten Prompting als geeignet: Trotz erforderlicher Nachkontrolle wurde die Arbeit, die sonst hohe personelle und zeitliche Ressourcen erfordert, in wenigen Minuten, also einem Bruchteil der bei „händischer“ Analyse benötigten Zeit und entsprechenden Kosten durchgeführt.

Es kann erwartet werden, dass bei Verfügbarkeit dieses Verfahrens die freie Textproduktion als Verfahren zur Erfassung differenzierter Behaltens- und Reproduktionsfähigkeiten häufiger eingesetzt werden wird.

Eine automatisierte Weiterverarbeitung der geprüften Propositionslisten im Sinne eines Vergleichs mit den Listen der jeweiligen Ausgangstexte ist als nächster Schritt vorgesehen. Dies ist relativ einfach zu programmieren. Angestrebt werden neben dem numerischen Vergleich der Propositionsanzahlen auch differenzierte Kennwerte für ausgezeichnete Propositionen (z.B. aus bestimmten Teilen des Ausgangstexts).

Unser Poster enthält einen QR-Code mit dem ein Download des verwendeten Prompts möglich ist.

Literatur

Ballstaedt, Steffen-Peter (1997): Wissensvermittlung. Die Gestaltung von Lernmaterial. Beltz. S. 21-41.

Klieme, E.; Jude, N.; Rauch, D.; Ehlers, H.; Helmke, A.; Eichler, W.; Thomé, G. & Willenberg, H. (2008). Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts. In E. Klieme (Hrsg.), Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie (S. 319-344). Beltz.

Niegemann, H. M. (1982). Influences of titles on the recall of instructional texts. In A. Flammer & W. Kintsch (eds.), Discourse processing (pp. 392-399). North-Holland Publishing Company.

Redder, A. & Weinert, S. (Hrsg.) (2013). Sprachförderung und Sprachdiagnostik. Interdisziplinäre Perspektiven. Waxmann.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 23: Beratung als Belastung? Herausforderungen in der Kommunikation mit Eltern begegnen.

Alexandra Merkert, Jerome Mitchell, Josef Strasser

Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU), Deutschland

Theorie

Die Gestaltung von Gesprächssituationen mit Eltern ist nicht nur fester Bestandteil des Aufgabenspektrums von Lehrkräften (Baumert & Kunter, 2006, 2011; Hertel, 2017; Schnebel, 2017), sondern auch als Voraussetzung für die Gestaltung passgenauer Bildungsangebote zu verstehen, deren Ausgestaltung Schüler:innen sowie deren Lebenssituation ernst nimmt. Gute Kommunikation kann damit für alle Seiten sowie insbesondere für die (schulische) Entwicklung von Kindern und Jugendlichen gewinnbringend sein (Aich & Behr, 2019; Ankrum, 2016; Đurišić, & Bunijevac, 2017; Rubach & Lazarides, 2020; Sacher, 2014; Santiago et al., 2016) und einen Beitrag zu mehr Bildungsgerechtigkeit leisten (Jeynes, 2024; Lenske et al., 2022). Gleichzeitig werden Elterngespräche von Lehrkräften häufig als Belastung erlebt (Nuding et al., 2018; Robert Bosch, 2019, 2024), wobei Schwierigkeiten auf mehreren Ebenen konzeptualisiert werden können (u. a. Gartmeier, 2018; Hertel & Schmitz, 2010; Nuding et al., 2018). Genannt werden bspw. Probleme wie geringes elterliches Kooperationsinteresse, aber auch eigene Unsicherheit und Schwierigkeiten, in emotional aufgeladenen Situationen Professionalität zu wahren (Nuding et al., 2018). Mangelnde Vorbereitung auf sensible Gesprächssituationen mit Eltern im Kontext der ersten Phasen der Lehrer:innenbildung (Aich & Behr, 2019; Gartmeier, 2018; Hertel, 2017; Schnebel, 2017; Nuding et al., 2018) sowie fehlende Copingstrategien können dazu führen, dass solche Belastungen als Überlastung empfunden werden (Seiferling, 2021). In Rückbezug auf Forschung zum Belastungs- und Stresserleben von Lehrkräften (u. a. Klusmann et al. 2016, Schaarschmidt et al., 2010) ist davon auszugehen, dass überlastete Lehrkräfte ihren beruflichen Aufgaben kaum noch angemessen nachkommen können, was sowohl die Gestaltung der Kommunikation mit Eltern als auch des Unterrichts betrifft. Dies kann sich wiederum auf die Schüler:innenleistungen auswirken (Klusmann et al., 2016), was besonders den am meisten auf schulische Unterstützung Angewiesenen schadet. Es gilt daher ein tieferes Verständnis der Belastungsfaktoren und ihrer Wirkungsweise im Kontext von Lehrer:innen-Eltern-Gesprächen zu erhalten, um Aus- und Fortbildungsangebote (weiter-)entwickeln zu können. Dabei geht es einerseits, um belastende Themenbereiche, Situationen und strukturelle Rahmenbedingungen sowie andererseits, um Strategien, die von Lehrkräften als hilfreich erlebt werden.

Fragestellung

Um einen Beitrag zum Verständnis der Belastungsfaktoren und deren Bewältigung zu leisten, sollen folgende Fragestellungen bearbeitet werden:

- Welche Faktoren erleben Lehrkräfte unterschiedlicher Schularten in Elterngesprächen als (besonders) beanspruchend und wie können diese unter Berücksichtigung bisheriger Konzeptualisierungen klassifiziert sowie empirisch ausdifferenziert werden?
- Welche Strategien wenden Lehrkräfte in der Praxis an, um mit den geschilderten Belastungen umzugehen?

Methode

Zur Beantwortung der Forschungsfragen wurden mit 34 Lehrkräften (1,5–40 Jahre Berufserfahrung) an Primar-, Sekundar- sowie Förderschulen leitfadengestützte problemzentrierte Interviews nach Witzel (2000) durchgeführt. Demografische Daten wurden mittels Fragebogen erhoben. Die standardnah-orthographisch transkribierten Interviews wurden qualitativ inhaltsanalytisch nach Mayring (2022) mit dem Programm MAXQDA ausgewertet. Das Kodiersystem wurde einem Interraterverfahren unterzogen. Als Übereinstimmungsmaß wurde Cohens Kappa bestimmt.

Ergebnisse und Diskussion

Bei der Auswertung wurden die genannten Belastungen in Anlehnung an Hertel und Schmitz (2010) klassifiziert und weitere Subkategorien ausdifferenziert. Das Interraterverfahren führt für alle Subkategorien zu Cohens-Kappa-Werten zwischen .89 und 1.00., was für die Eignung des Kategoriensystems spricht (siehe Landis & Koch, 1977). Die Ergebnisse bestätigen z. T. bisherige Befunde, bspw. zu Schwierigkeiten im Umgang mit eigener und elterlicher Emotionalität (Gartmeier, 2018), sprachlichen Barrieren oder mangelnder Kooperationsbereitschaft (Nuding et al., 2018). Der Umgang mit eigener Emotionalität fällt insbesondere bei Themen wie des Verdachts auf Kindeswohlgefährdung schwer. Eine zentrale berufsbezogene Ressource im Umgang mit Belastungen scheint der Austausch mit Kolleg:innen zu sein (siehe Cramer et al., 2018). Insgesamt ist aus den Antworten der Lehrkräfte abzuleiten, dass Copingstrategien inklusive der Entwicklung einer gesunden Distanzierungsfähigkeit sowie Strategien zum Aufbau unterstützender professioneller Netzwerke als feste Bestandteile von Aus- und Fortbildungsprogrammen zu integrieren sind, damit Lehrkräfte langfristig (emotional) belastbar bleiben, um empathisch und engagiert agieren zu können. Eine systematische Vermittlung dieser Fähigkeiten und Kenntnisse scheint Lehrkräfte bislang noch (zu) wenig zu erreichen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 45: Brokering 4.0 - Der Handel mit Wissen zum digital gestützten Unterricht

Aysegül Serpemen¹, Luisa Scherzinger², Till Schmäing³

¹Universität Duisburg-Essen, Deutschland; ²Eberhard Karls Universität Tübingen, Deutschland; ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Deutschland

Wie lässt sich der Dialog zwischen (Bildungs-)Forschung und Schulpraxis gestalten? Wie kann man die Kommunikation und Kooperation zwischen verschiedenen Handlungs- und Systemebenen für die digitale Transformation von Schulen verstärken? Unter anderem diesen Fragen widmen sich die Knowledge-Broker:innen (im Deutschen auch als Wissensvermittler:innen bezeichnet) der Transferstelle des Kompetenzverbund lernen:digital, in welchem aktuell in rund 200 länderübergreifenden Forschungs- und Entwicklungsprojekten evidenzbasierte Fort- und Weiterbildungen für die Schul- und Unterrichtsentwicklung in einer Kultur der Digitalität entstehen. Knowledge Broker:innen nehmen die Funktion von Zwischenhändler:innen ein (van Ackeren-Mindl & Endberg, 2023), indem sie gezielt Wissen mobilisieren sowie transformieren und die Partnerschaften zwischen verschiedenen Interessengruppen in der Bildungsforschung und -praxis vertiefen (Cooper & Shewchuk, 2015).

Die Relevanz der Broker:innen im Kontext der Bildungsforschung lässt sich durch die oft zitierte Lücke bzw. schwache Verknüpfung von Forschung und Praxis (Cooper & Shewchuk, 2015; Rechsteiner et al., 2024) und durch das wachsende Interesse an einer Verbindung dieser beiden Akteur:innen ableiten (Rycroft-Smith, 2022). Damit nehmen Broker:innen verschiedene Praktiken der Wissensmobilisierung und -vermittlung ein. Diese sind nach Meyer (2010) die Identifizierung und Lokalisierung von Wissen, die Umverteilung und Verbreitung von Wissen sowie die Transformation von Wissen. Besonders relevant ist die dritte Praktik der Wissensvermittlung, wonach es nicht um die Übertragung eines statischen Konstruktes geht (Cooper et al., 2020). Vielmehr schließt die Transformation von Wissen Aushandlungsprozesse, Anpassung- und Übersetzungsleistungen durch die Broker:innen ein. Auf verschiedenen Forschungsergebnissen beruhende Implikationen für die Praxis sind grundlegend nicht als „fertig“ zu interpretieren, sondern können je nach den Bedarfen der Zielgruppen eine Adaptation verlangen. Eine Adaptation von Forschungsergebnissen für Lehrkräfte, Schulleitungen, aber auch Verantwortliche der Lehrkräftefort- und Weiterbildung kann beispielsweise durch den systematischen Einsatz von Disseminationsplattformen wie ComPLeTT (Kiesler & Schiffner, 2022) gelingen. Weitere Adaptationsmöglichkeiten durch die Broker:innen bieten Formen der Komplexitätsreduzierung oder Fokussierung wissenschaftlicher Erkenntnisse durch Präsentations- und Darstellungsformate wie Tagungen. Handlungsleitend für die Ausdifferenzierung der Broker:innen-Rolle ist die Frage: Wie können Forschungsergebnisse so aufgegriffen und transformiert werden, dass sie in der Praxis wahrgenommen und genutzt werden? Übergreifende Ziele des *knowledge brokerings* sind beispielsweise die Stärkung der Nutzung von Forschungsergebnissen in der Praxis und die Verbesserung der Zugänglichkeit für die Wissenschaft (Cooper & Shewchuk, 2015). Außerdem sollen die sozialen Interaktionen zwischen den Akteur:innen erleichtert werden (Stovel & Shaw, 2012), was wiederum den Aufbau vertrauensvoller Beziehungen und Kooperationen sowie Partnerschaften ausgehend von den Broker:innen voraussetzt.

Broker:innen haben im Kompetenzverbund lernen:digital eine zentrale Aufgaben. Dabei gibt es eine Erweiterung der ausgeführten Rolle hinsichtlich der Zielgruppe, mit denen die Broker:innen der vier Kompetenzzentren agieren. Die Broker:innen-Teams übernehmen eine Schnittstellenfunktion zwischen den Projektverbünden innerhalb sowie außerhalb ihres fachlichen Kompetenzzentrums und vernetzen damit Akteur:innen aus der Wissenschaft miteinander. Demnach stehen die Broker:innen im engen Kontakt zu den Projektverbünden ihres Kompetenzzentrums und identifizieren in einem fortlaufenden Dialog mit den Projekten, Schnittmengen und Gemeinsamkeiten im Hinblick auf die thematischen Schwerpunkte und Transfermaßnahmen der rund 200 Forschungs- und Entwicklungsvorhaben. Aufbauend auf den Synergien in den Kompetenzzentren und darüber hinaus übernehmen die Broker:innen vielfältige Vernetzungs- und Vermittlungsaufgaben, indem sie beispielsweise Impulse setzen, Dialogformate etablieren und Transferprodukte (Fortbildungsmaterialien für die Schulpraxis) mitgestalten. Das zentrale Ziel dieser Brokerage-Ansätze ist es, dass vorhandene Konzepte und Kenntnisse zwischen unterschiedlichen Akteursgruppen geteilt, gebündelt und im Sinne einer erhöhten Sichtbarkeit für Wissenschaft und Praxis möglichst breit disseminiert werden (van Ackeren-Mindl & Endberg, 2023).

Das Poster gibt einen detaillierten Überblick über das Konzept und die Aufgaben der Knowledge-Broker:innen im Kompetenzverbund lernen:digital und verdeutlicht die Relevanz dieser innovativen Rolle in der deutschen Bildungsforschung.

Pinnwand 46: Argumentieren über gesellschaftliche Kontroversen – Warum ist es (nicht) in Ordnung, Tiere zu essen? Argumentationsmuster von Kindern, Jugendlichen und Erwachsenen im Vergleich.

Tina Bagus¹, Luke McGuire², Hanna Beißert¹

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ²University of Exeter, UK

Bildung als Selbstbildung beinhaltet die Entwicklung und Reflexion eigener Werthaltungen in einer ausdifferenzierten, sozialen Welt. Teil dieses Entwicklungsprozesses ist die Auseinandersetzung mit gesellschaftlichen Kontroversen, die Menschen herausfordert und gleichzeitig darin fördert, ihren eigenen Standpunkt zu finden und zu vertreten. Die Beurteilung sozialer Situationen und das Begründen der eigenen Position kann mithilfe der Social Domain Theory (SDT) beschrieben werden. Nach der SDT integrieren und koordinieren Menschen in solchen Prozessen Aspekte, die sich verschiedenen sozialen Bereichen zuordnen lassen: Die moralische Domäne umfasst u.a. Gerechtigkeitskonzepte, während in der sozial-konventionellen Domäne u.a. Normen oder gesellschaftliche Erwartungen enthalten sind (Turiel, 2023; Smetana et al., 2014). Daneben steht die persönliche Domäne, in der beispielsweise Motive persönlicher Entscheidungsfreiheit subsummiert sind (Nucci, 1981). Das Wissen innerhalb dieser Bereiche differenziert sich durch soziale Interaktionen und Erfahrungen mit zunehmendem Alter aus; gleichzeitig weisen die Konzepte in den einzelnen Bereichen unterschiedliche Entwicklungsverläufe auf (Smetana et al., 2014).

Ausgehend von dieser theoretischen Basis widmet sich unsere Forschung altersbezogenen Unterschieden in den Argumentationsstrukturen bei kontroversen sozio-moralischen Fragestellungen. Als Beispiel für eine solche gesellschaftliche Kontroverse wurde das Thema *Vertretbarkeit von Fleischkonsum* gewählt. Vor dem Hintergrund des Klimawandels als gesamtgesellschaftliche Herausforderung stellt dieses Thema eine aktuelle und relevante Diskussion dar, da der Verzicht auf tierische Produkte eine der wirkungsvollsten Maßnahmen ist, um auf persönlicher Ebene dem Klimawandel entgegenzuwirken (Dixon et al., 2023). Eine fleischliche Ernährung ist mit deutlich höheren Treibhausgasemissionen und einem größeren individuellen CO₂-Fußabdruck verbunden (Dixon et al., 2023; Chai et al., 2019).

Wir untersuchten die Beurteilung der Vertretbarkeit von Fleischkonsum bei 115 Kindern ($M=9.68$, $SD=1.03$), 125 Jugendlichen ($M=14.24$, $SD=2.05$) und 266 Erwachsenen ($M=33.57$, $SD=13.02$). Dazu fragten wir die Teilnehmenden anhand einer sechsstufigen Likert-Skala, wie sehr es in Ordnung ist, Tiere zu essen („überhaupt nicht in Ordnung“ bis „vollkommen in Ordnung“). Anschließend begründeten die Befragten ihre Einschätzung anhand einer offenen Antwort.

Zur Auswertung der offenen Antworten verwendeten wir ein bereits etabliertes Kategoriensystem (McGuire et al., 2023), das auf der Social Domain Theory basiert und durch Kategorien ergänzt wurde, die induktiv aus unserem Material abgeleitet wurden. Die Interrater-Reliabilität von zwei unabhängigen Kodierenden betrug $\kappa > .80$.

Eine univariate ANOVA zeigte altersbezogene Unterschiede in der Einschätzung der Vertretbarkeit von Fleischkonsum ($F(2, 482) = 34.25$, $p < .001$, $\eta_p^2 = .12$). Kinder ($M=2.56$, $SD=.11$) fanden es im Vergleich zu Erwachsenen ($M=3.64$, $SD=.08$) und Jugendlichen ($M=3.56$, $SD=.11$) weniger in Ordnung, Tiere zu essen.

Weiterhin zeigte eine ANOVA mit Messwiederholung altersbezogene Unterschiede im Argumentationsverhalten bei der Begründung, warum es (nicht) in Ordnung ist, Tiere zu essen ($F(5.52, 1341.06) = 15.77$, $p < .001$, $\eta_p^2 = .06$). *Hierbei verwendeten Kinder häufiger als Erwachsene und Jugendliche moralische Argumente, während letztere häufiger sozial-konventionelle Argumente aufführten ($ps \leq .001$). Argumente aus der persönlichen Domäne wurden von allen Altersgruppen gleichermaßen selten genannt ($ps > .05$).*

Die Ergebnisse deuten darauf hin, dass Kinder die Frage nach der Vertretbarkeit von Fleischkonsum eher aus einer moralischen Perspektive beurteilen, während Jugendliche und Erwachsene diese eher aus sozial-konventioneller Sicht betrachten.

Logistische Regressionsanalysen deuteten außerdem auf Zusammenhangsmuster zwischen der Einschätzung der Vertretbarkeit von Fleischkonsum und der Argumentation in den Begründungen hin: Je eher eine Person Fleischkonsum ablehnt, desto wahrscheinlicher begründeten sie dies unter Einbezug moralischer Argumente ($B = -1.21$, $p < .001$). Je eher eine Person Fleischkonsum befürwortet, desto wahrscheinlicher begründet sie diese Einschätzung mit sozial-konventionellen Argumenten ($B = .43$, $p < .001$). Diese Muster zeigten sich über alle untersuchten Altersgruppen hinweg.

Diese Befunde stärken bisherige Forschungsergebnisse, in denen Kinder größere moralische Bedenken hinsichtlich der Vertretbarkeit von Fleischkonsum äußern und stärken die These, dass das Befürworten einer fleischlichen Ernährung mit Sozialisationsprozessen in Zusammenhang steht (McGuire et al., 2023). Die Ergebnisse bieten daher einen Anhaltspunkt für eine wertgeleitete und reflexive Auseinandersetzung mit der eigenen Ernährung im Rahmen von Bildung für nachhaltige Entwicklung und hinsichtlich der Verantwortung in der Gestaltung einer zukunftsfähigen Gesellschaft.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 8: Entwicklung eines Schuleingangsscreenings zur frühzeitigen Erkennung von Entwicklungsrisiken im Lesen, Rechtschreiben und Mathematik

Kristina Clausen-Suhr

Universität Vechta, Deutschland

Viele sich auf dem Markt befindliche einschlägige Diagnose-Verfahren für den Einsatz um den Einschulungszeitraum herum sind als Einzelverfahren konzipiert (ausgenommen z.B. das PB-LRS, Barth & Gomm, 2014). Diese erfordern (abgesehen von ihrer unterschiedlichen Zuverlässigkeit) dadurch einen sehr hohen Aufwand an Zeit und Personal, was häufig ein fast unüberwindliches Hindernis für den flächendeckenden Einsatz in Grundschulen darstellt (Walter & Clausen-Suhr, 2022; Clausen-Suhr & Walter, 2022). Mit der Entwicklung des *FleSch* (Walter & Clausen-Suhr, i.V.) wurde ein *Gruppenverfahren* im Sinne eines Filter-Screenings konzipiert, das zeitökonomisch zum Einsatz gebracht werden kann und gleichzeitig den Anspruch verfolgt, gute klassifikatorische Resultate zu liefern.

Des Weiteren sollte das *FleSch* nicht ausschließlich in der Lage sein, lediglich Kinder frühzeitig zu identifizieren, bei denen in umschriebenen Bereichen wie z.B. des Schriftspracherwerbs (Jansen et al., 2002; Barth & Gomm, 2014; Endlich et al., 2016) oder der Rechenkompetenz (von Aster, Bzufka & Horn, 2009; Walter, 2020) mit späteren Minderleistungen zu rechnen ist. Vielmehr ist es aus sonderpädagogischer Sicht notwendig, Kinder im Blick zu haben, bei denen das Risiko einer breiten Lernschwäche, nämlich beim Lesen, Rechtschreiben und Rechnen im Sinne einer kombinierten Lernschwäche besteht. (Kanter, 1977; Ricken, Fritz & Hofmann, 2003).

Auch wenn mit dem Response-to-Intervention-Ansatz (RTI) ein während der laufenden Beschulung realisierbarer präventiver Umgang mit Lernschwierigkeiten auf der Basis curriculumbasierter Messens (CBM) zur Verfügung steht (Walter, 2008; Huber & Grosche, 2012), würde ein bereits zum oder kurz vor dem *Einschulungszeitpunkt* durchgeführtes prognostisch valides Filter-Screening (Tröster, 2009) Kinder mit erhöhtem Risiko im Bereich des schulischen Lernens gezielt und frühzeitig identifizieren können. Das *FleSch* soll mit minimalem Ressourcen-Einsatz als Gruppenverfahren eine individuelle Risiko-Aussage bezüglich eines Kindes ermöglichen, auf dessen Grundlage sich dann gegebenenfalls eine weiterführende Diagnostik bzw. gezielte Maßnahmen z.B. der Primären Prävention anschließen (Fuchs & Fuchs, 1986; Kretschmann, 2007; Huber & Grosche, 2012; Tröster, 2009; Walter, 2008, 2020).

Folgende Fragestellungen können formuliert werden:

1. Lässt sich auf der Basis eines breit angelegten Satzes von 22 Prädiktorvariablen per Datenreduktionsverfahren ein *reduziertes Set* von Prognosevariablen (prognoserelevantes Variablen-Set) identifizieren?
2. Die Stabilität eines Prognose-Ansatzes ist u.a. auch von der Beantwortung der Frage abhängig, ob sich anhand des identifizierten prognoserelevanten Variablensets dessen Prognosekraft auch auf der Basis von unabhängigen Teil-Stichproben (Kreuzvalidierung) replizieren lässt.
3. Befunde von Seeboth und Möttus (2018), Möttus und Rozgonjuk (2019) legen nahe, dass ein formatives (auf Item-Ebene) im Vergleich zu einem summativen Messmodell (der übliche Latent-Trait-Ansatz) im Zusammenhang mit prognostischen Fragestellungen u.U. bessere Ergebnisse liefert.

Auf der Basis von insgesamt 3 längsschnittlich angelegten Untersuchungswellen und N= 510 vollständigen Fällen (N=510) wurde, ausgehend von einem breiten Prädiktoren-Set von 22 möglichen prognostisch relevanten Vorläuferfertigkeiten, ein valides Prognosemodell mit insgesamt 7 unspezifischen und spezifischen Prädiktoren nach dem klassifikatorischen Verfahren der logistischen Regression (stepwise) ermittelt. Der auf der Basis einer ROC-Analyse ermittelte AUC-Wert von .89 (95 % KI [.86, .92]) steht für eine hervorragende Vorhersagegüte des Modells. Vor dem Hintergrund des recht befriedigenden gegenseitigen Modell-Transfers von jeweils einer Teilstichprobe auf die andere und umgekehrt (Kreuzvalidierung) sollte man davon ausgehen dürfen, dass die Gesamt-Stichprobe insgesamt eine stabile Basis für das Prognose-Modell darstellt. Das formative Messmodell, also die logistische Modellbildung auf Item-Ebene, liefert im Vergleich zum summativen Messmodell zwischen .03 und .04 höhere AUC-Werte und damit für alle Kriteriumsvariablen (LRS+RS; LRS; RS) sehr gute AUC-Werte zwischen .88 und .93. Die Differenzen sind höchst signifikant von Null verschieden.

Die Ergebnisse machen insgesamt deutlich, dass bei Erfüllung hoher Standards an angelegte Güteindizes einer Entwicklungsprognose ein diagnostisches Verfahren entwickelt wurde, das auch hinsichtlich der Praktikabilität und ökologischen Validität vielversprechend erscheint.

Pinnwand 34: Schulleitungshandeln und Kooperationsprozesse an inklusiven Schulen

Stefanie van Helden¹, Kathrin Fussangel¹, Karolina Urton², Philipp Krämer¹

¹Bergische Universität Wuppertal, Deutschland; ²Universität Münster, Deutschland

Theoretischer Hintergrund und Fragestellung

Die Kooperation von Regelschullehrkräften und Sonderpädagog*innen gilt als zentrale Gelingensbedingung und Grundvoraussetzung für die erfolgreiche Umsetzung schulischer Inklusion, wie sie in der UN-Behindertenrechtskonvention (UN-BRK) gefordert wird (Lütje-Klose & Urban, 2014).

Trotz des hohen Stellenwerts gestaltet sich die praktische Umsetzung oft schwierig, mit einem deutlichen Unterstützungsbedarf seitens Schulen und Lehrkräften (Bartling et al., 2021). Gemeinsame Ziele der Lehrkräfteteams sind entscheidend für eine gelingende Kooperation, doch unterschiedliche Rollen, Erwartungen sowie divergierende Ansichten zur Inklusion erschweren häufig deren Verfolgung (Lütje-Klose et al., 2024; Urton, 2022). Ob übereinstimmende Ziele ausgehandelt werden, hängt von unterschiedlichen Faktoren ab. Kooperationsförderliche Rahmenbedingungen sowie ein inklusives Schulklima können Kooperationsprozesse begünstigen (Richter & Pant, 2016; Urton, 2022). In diesem Zusammenhang spielt die Schulleitung als „Change Agent“ eine zentrale Rolle, indem sie durch indirekte Führung (vgl. Wunderer, 2003) kooperationsförderliche Rahmenbedingungen schaffen und durch direkte Interaktion auf die Kooperationsprozesse des Kollegiums einwirken kann (Harzad & Drossel, 2011). Um Schulleitungshandeln zu analysieren und konzeptualisieren, werden in der Schulleitungsforschung verschiedene Führungsmodelle diskutiert. Das Modell des „Instructional Leader“, das die Unterrichtsqualität direkt beeinflusst,

wird durch neuere Ansätze ergänzt. Dazu gehören transaktionale Führung, die auf Austausch und Management basiert, transformationale Führung, die Innovationen vorantreibt sowie Distributed Leadership, das Führungsaufgaben auf mehrere Personen innerhalb der Organisation verteilt (Dubs, 2019; Scheer, 2020).

Obwohl das Interesse an der Schulleitungsforschung im deutschsprachigen Raum in den letzten Jahrzehnten gewachsen ist und eine weitgehende Einigkeit über die Schlüsselfunktion der Schulleitung besteht (Bonsen, 2010), ist deren Rolle im Kontext von Inklusion noch nicht ausreichend untersucht. Insbesondere liegen kaum Befunde darüber vor, wie Schulleitungen die für Inklusion notwendigen Kooperationsprozesse beeinflussen.

Das vorliegende Forschungsvorhaben geht daher der Frage nach, welche Rolle das Führungshandeln von Schulleitungen aus Sicht der Lehrkräfte im Zusammenhang mit der Umsetzung von Inklusion und der im Rahmen dessen praktizierten Kooperation innerhalb des Kollegiums spielt.

Methode

Die vorliegende Studie ist die erste von vier Teilstudien des Forschungsprojekts KoLeGe (Kooperationsziele von Lehrkräften beim gemeinsamen Unterrichten), wobei Lehrkräfte über einen Fragebogen zu ihrer (a) Kooperationsaktivität und ihrer (b) Einschätzung des Schulleitungshandelns befragt werden. Für die Validierung des vorliegenden Fragebogens wurde im Rahmen einer Pilotstudie von März bis Mai 2024 eine Stichprobe von N=68 Lehrkräften (davon 56 Regelschullehrkräfte und 13 Sonderpädagog*innen) an zwei inklusiven Gesamtschulen in NRW erhoben. Zur Messung der (a) Kooperationsaktivität der Kollegien wurden etablierte Skalen zur Kooperation an Schulen verwendet (Stommel, 2014; Fussangel et al., 2023). Die Einschätzung des Schulleitungshandelns (b) erfolgte anhand von vier Skalen, die die im Theorieteil beschriebenen Führungsmodelle abbilden (vgl. Urton & Wilbert, 2017).

Die querschnittliche Datenerhebung der Hauptstudie beginnt im Sommer 2024. Die geplante Stichprobe soll ca. N=2.100 Lehrkräfte an insgesamt 30 inklusiven Gesamt- und Sekundarschulen sowie an vier inklusiven Grundschulen umfassen. Der Fokus der Hauptstudie wird auf der Intensität der Kooperation und deren Zusammenhang mit dem Führungshandeln der Schulleitung liegen. Der Einfluss des schulischen Führungshandelns auf die Kooperationsprozesse der Lehrkräfte soll anhand einer Regressionsanalyse untersucht werden.

Ergebnisse

Mithilfe der faktorenanalytischen Überprüfung der Skalen im Anschluss an die Pilotstudie konnten für die (a) Kooperationsaktivität drei zuverlässige Skalen identifiziert werden: Entwicklung von Material ($\alpha = 0.83$), Austausch und Reflexion ($\alpha = 0.78$) sowie Entwicklung von Verfahren/Konzepten ($\alpha = 0.73$). Auch die vorläufigen Ergebnisse der Faktorenanalyse der (b) Einschätzung des Schulleitungshandelns zeigen, dass sich zwei der Führungsmodelle, Instructional Leadership und Transformational Leadership, abbilden lassen und gute Reliabilitäten aufweisen (Instructional Leadership $\alpha = 0.95$, Transformational Leadership $\alpha = 0.93$). Die weiteren Ergebnisse der Hauptstudie werden ebenfalls auf dem Poster präsentiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 47: Mehr als nur Daddeln? – Eine Untersuchung zu Bildungserfolgen von kompetitiven Spieler*innen und E-Sportler*innen

Christel Eva Piersig

Georg-August-Universität Göttingen, Deutschland

Das Thema Videospiele führt immer wieder zu hitzigen Debatten. Dieses Medium wird häufig als sinnlose Zeitverschwendung betitelt, welcher unter starken Kritiker*innen auch negative Attribute wie Aggressionen und Suchtpotenzial zugesprochen wird (bspw. (dpa, 2020)). Dennoch hat das sogenannte *Gaming* über die letzten Jahrzehnte immer mehr an Beliebtheit gewonnen und es hat sich ein professioneller Bereich des Wettkampfs in Form des E-Sports herausgebildet, wo Spieler*innen in nationalen und internationalen Ligen und Turnieren gegeneinander antreten. Ähnlich wie beim traditionellen Sport können zwei Arten von Spieler*innen unterschieden werden: die Personen, die Videospiele zum lockeren Zeitvertreib spielen und die Personen, welche eher den Fokus auf den Wettkampf legen. Erstere werden häufig als *casual* Spieler*innen bezeichnet und sind vergleichbar mit dem Breitensport, während auf der anderen Seite die *competitiven* Spieler*innen bzw. die E-Sportler*innen versuchen, ihre Fähigkeiten stetig zu verbessern und sich mit anderen messen (vgl. ESBD, 2024). In sogenannten E-Sport Titeln – dazu gehören Spiele wie *League of Legends*, *Counter Strike*, *Rocket League* – gibt es Rangsysteme, welche ähnlich funktionieren wie das ELO-System im Schach. Hierdurch ist ein direkter Vergleich mit anderen Spieler*innen innerhalb eines Titels möglich.

Auch in der Wissenschaft wird der Konsum von Videospiele des Öfteren zum Forschungsgegenstand: von problematischem Spielverhalten von Jugendlichen (vgl. Rehbein *et al.*, 2009) bis hin zu körperlichen und gesundheitlichen Aspekten von E-Sportler*innen (vgl. Rudolf *et al.*, 2020) variieren die Perspektiven auf den Konsum dieses Mediums. Es kommt jedoch die Distinktion zwischen *casual* und *competitive* Spieler*innen sowie sozioökonomischen Hintergründe zu kurz. Welche Faktoren können bei der Wettkampforientierung in Videospiele eine Rolle spielen, und steht das Spielen von Videospiele dem Bildungserfolg im Weg? Eine theoretische Überlegung bietet Bourdieus Klassentheorie, insbesondere sein Begriff des „Habitus“ (vgl. Bourdieu, 2021): anders als es die Vorurteile vermuten lassen würden, kann mit dem Begriff des Habitus argumentiert werden, dass die intrinsische Motivation, einen hohen Rang zu erhalten und sich mit anderen zu messen, ähnlich begründet werden wie bei einer hohen Bildungsaspiration. Es stellt sich also die Frage, ob der Bildungserfolg und der Erfolg in Videospiele korrelieren.

Die vorgestellte Untersuchung ist ein erster explorativer Schritt, der sich mit den Spieler*innen von Videospiele auseinandersetzt. Es wurde versucht, einen Einblick in die Gruppe von sogenannten *Gamer*innen* zu erhalten: wer spielt Videospiele eher *casual* oder *competitive*, und besteht ein Zusammenhang zwischen dem Bildungserfolg und dem Erfolg in Videospiele?

Die Daten wurden von Sep. bis Okt. 2022 mit Hilfe eines Online-Fragebogens erhoben, welcher primär über Netzwerke deutscher E-Sport Organisationen sowie Social Media verteilt wurde. Die Befragung umfasst, neben den demografischen Daten und Informationen zum Bildungsverlauf (Bildungsabschluss, aktuelle Tätigkeit), Items zur Intensität des Spielens (bspw. Spielzeit in der Woche, Rang). Darüber hinaus wurden auch Informationen über die soziale Herkunft der Person, wie z.B. der Bildungsabschluss der Eltern abgefragt. Der Rang soll den Erfolg im Gaming messen. Da jeder Spieltitel ein eigenes Rangsystem hat, wurden die unterschiedlichen Ränge in einer 1-5 Skala zusammengefasst.

Mit einer Fallzahl von n= 212 stellt dieser Beitrag die deskriptiven sowie bivariaten Ergebnisse vor. Die Stichprobe ist vergleichsweise jung mit einem Durchschnittsalter von 23 Jahren, männlich (80%) und studiert (> 70%). 2/3 der Befragten verstehen sich als *competitive* Spieler*innen.

Die Daten zeigen keine signifikanten Zusammenhänge zwischen dem Bildungserfolg und einem hohen Rang in den jeweiligen Spielen. Trotzdem bietet diese Untersuchung einen ersten Einblick in den kompetitiven Bereich des *Gamings* in Deutschland. Die kontaktierten E-Sport Organisationen scheinen überproportional aus Studierenden zu bestehen, was damit begründet werden kann, dass viele eingetragenen Vereine aus Hochschulgruppen entstanden sind oder die Organisationen selbst Hochschulgruppen sind. Hier könnten zukünftige Untersuchungen anknüpfen inwieweit die Infrastrukturen von Hochschulen solche Netzwerke begünstigen. Des Weiteren könnte ein Vergleich zwischen Spieler*innen und Nicht-Spieler*innen herangezogen werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 15: „Es geht wirklich um nichts? Dann spiele ich lieber!“ – Motivation von Grundschulkindern in low-stakes Assessments

Pauline Wernet, Janine Buchholz
IQB Berlin, Deutschland

Theoretischer

Groß angelegte Schulleistungsuntersuchungen wie PISA, PIRLS/IGLU oder der IQB-Bildungstrend stellen wichtige bildungspolitische Instrumente dar, welche die Kompetenzen von Schüler*innen in verschiedenen Domänen messen und vergleichen und deren Ergebnisse häufig als Grundlage für bildungspolitische Maßnahmen dienen. Für die Schüler*innen selbst handelt es sich dabei jedoch um sogenannte „low-stakes“ Assessments, das heißt, die erzielten Ergebnisse haben für sie keine direkten persönlichen Konsequenzen (wie z.B. Noten). Es ist gut belegt, dass Schüler*innen solche Tests oft mit weniger Motivation bearbeiten und sich dies negativ auf ihre Testleistung auswirkt (Finn, 2015; Wise & DeMars, 2005). Dabei spielen verschiedene Faktoren auf Personenebene, wie Geschlecht (z. B. Soland, 2018) und situationsspezifisches Interesse (z. B. Penk & Schipolowski, 2015), sowie auf Itemebene, wie Itemtyp (z. B. Michaelides & Ivanova, 2022) und Itemposition (z. B. Penk & Richter, 2017), eine Rolle. Bisherige Studien konzentrierten sich vorwiegend auf Schüler*innen der Sekundarstufe und Studierende, sodass nur wenige Erkenntnisse über die Testbearbeitungsmotivation von Grundschulkindern vorliegen. Es stellt sich daher die Frage, ob die Testergebnisse von Grundschul*innen bei low-stakes Tests in ähnlicher Weise durch ihre Testbearbeitungsmotivation beeinflusst werden wie bei älteren Schüler*innen und Studierenden, was eine Gefahr für die Validität der Ergebnisinterpretation darstellen könnte.

Der vorliegende Beitrag möchte die aufgezeigte Lücke schließen und typische Indikatoren der Testbearbeitungsmotivation wie die selbstberichtete Anstrengungsbereitschaft und schnelles Raten (*Rapid Guessing*) bei Grundschul*innen untersuchen und Zusammenhänge mit Personen- und Aufgabenmerkmalen analysieren.

Fragestellung

Der Beitrag geht der Fragestellung nach, ob interindividuelle Unterschiede zwischen Grundschul*innen in ihrer Testbearbeitungsmotivation bestehen und ob diese sich durch bestimmte, bei älteren Schüler*innen und Studierenden identifizierte, personen- und itembezogene Merkmale erklären lassen. Außerdem behandelt der Beitrag die Frage, ob Unterschiede in der Testbearbeitungsmotivation auch bei Grundschulkindern mit der Testleistung zusammenhängen.

Methode

Die Fragestellungen werden anhand einer Stichprobe von N = 357 Schüler*innen (53 % männlich) der dritten Jahrgangsstufe untersucht. Die Schüler*innen bearbeiteten jeweils zwei Lesetexte mit mehreren Aufgaben zum Leseverständnis am Tablet bzw. Computer. Zu drei Zeitpunkten (vor, zwischen und nach Bearbeitung der Texte) wurde ihre Anstrengungsbereitschaft als Selbstbericht gemessen. Zudem wurde nach jedem Text das textbezogene Interesse (Lepper et al., 2022) der Schüler*innen erhoben. Des Weiteren wurden die Antwortzeiten erfasst und auf Basis der NT10-Methode (Wise & Ma, 2012) der *Response-Time-Effort (RTE)* berechnet. Im Anschluss an die deskriptiven Analysen sind vertiefende Pfadanalysen vorgesehen.

Ergebnisse	und	ihre	Bedeutung
Erste deskriptive Analysen zeigen, dass viele Schüler*innen (ca. 25%) einen kritischen RTE aufweisen, also mehr als 10 % ihrer Antworten als Rapid Guesses zu klassifizieren sind. Dabei sind vor allem Aufgaben mit offenem Antwortformat betroffen. Dagegen schätzten die Schüler*innen zu allen drei Messzeitpunkten ihre Anstrengungsbereitschaft als hoch ein. Es konnte kein signifikanter Zusammenhang zwischen den beiden Indikatoren festgestellt werden. Es zeigen sich keine geschlechtsspezifischen Differenzen: Jungen und Mädchen unterscheiden sich weder hinsichtlich ihres RTE noch hinsichtlich ihrer selbstberichteten Anstrengungsbereitschaft signifikant voneinander. Des Weiteren zeigt sich keine systematische Korrelation zwischen dem textbezogenen Interesse und dem im jeweiligen Text gezeigten RTE. Diese Ergebnisse sind insofern überraschend, als in Studien mit größtenteils älteren Schüler*innen meist geringere Anteile von Teilnehmenden berichtet werden, die einen niedrigen RTE unter dem Schwellenwert von 90 % aufweisen (z. B. Wise & Kingsbury, 2022). Zudem werfen die von der Erwartung abweichenden Ergebnisse die Fragen auf, ob Rapid Guessing ein passender Indikator für die Testbearbeitungsmotivation in der Grundschule ist und welche Faktoren bei Grundschulkindern dazu führen, dass sie besonders schnell durch die Aufgaben klicken. Vertiefende Analysen, die auch die Testleistung mit einbeziehen, sollen sich diesen Fragen widmen und mitsamt ihren Ergebnissen im Rahmen des Beitrags vorgestellt werden.			

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 48: Extraordinary Success against All Odds. A Comprehensive Study of Professors from Non-Academic Backgrounds in Germany.

Claudia Begemann¹, Sandra Buchholz^{1,2}

¹DZHW, Deutschland; ²Leibniz Universität Hannover

Educational inequalities, beyond income differences, remain a key concern in sociological research. Despite significant educational expansion post-World War II (Bar Haim & Shavit, 2013; Vaira, 2009) and policy efforts, disparities keep persisting (Möller, 2017). Social background continues to impact educational success across all stages, from early childhood to adulthood (Dräger, 2022; Jackson, 2013; Vergolini & Vlach, 2017).

However, some individuals from disadvantaged backgrounds achieve professorial status against all odds. The amount, distribution, and qualities of these professors remain largely unknown due to data scarcity. Our study uses the prof*panel dataset from "Scientific Careers in Transition: DZHW Professors Panel" to examine 1) How many individuals from non-academic backgrounds attain professorship and whether they are segregated by academic disciplines, career positions, compensation, and gender? And 2) What personality traits distinguish professors from non-academic backgrounds from those with academic backgrounds? Are there further gender differences?

According to Merton (1973), recognition in the academic labour market should be tied to one's scientific accomplishments and achievements. The best performing academics should hence be rewarded with success. Following Becker's (1964) human capital theory it is also productivity that is pivotal to success in the labour market. In this meritocratic ideal, non-meritocratic aspects like social background should not have a bearing on success. Sociological arguments, however, suggest that individuals from non-academic family backgrounds face formidable barriers, seemingly making the attainment of a professorship nearly impossible. Boudon's (1974) suggests that primary effects manifest as the direct impact of social background on academic success, with factors like financial means determining access to educational resources, thereby shaping performance and outcomes. Additionally, secondary effects encompass how individuals interpret and respond to their educational experiences, despite performance. Boudon accentuates the pivotal role of decision making, aspirations, and perceptions in shaping one's educational trajectory, extending beyond the initial impact of social background.

In Germany's stratified educational system, with multiple transition points, prestige and overrepresentation of higher social strata increases in every step (Buchmann & Park, 2009; Dollmann, 2021; Möller, 2017). Individuals from non-academic backgrounds are less likely to enrol in and complete university degrees and transition to postgraduate studies (Auspurg & Hinz, 2011; Finger, 2022). Reaching professorship requires exceptional performance and nuanced decisions influenced by personality traits (Coenen et al., 2021) hence our study incorporates the Big Five personality traits.

Using the 2022 prof*panel dataset with 2,586 participants, our study focuses on 2,101 professors appointed post-2010. Preliminary analysis shows 35.2% (N=740) come from non-academic families, while 64.8% (N=1,361) have at least one parent holding a higher education degree. Notably, 22% of professor have a parent with a doctorate, underscoring the selectivity of the German professoriate, compared to 1.2% of the German population with doctorates in 2019 (Schade, 2023). We categorise professors into six fields: Humanities, social sciences, life sciences, natural sciences, engineering sciences, and others. Using the Big Five traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism), we analyse those who succeeded against the odds.

Our initial analysis reveals the highly selective nature of the German professoriate. We found significant differences in the representation of non-academic background professors across various disciplines and positions. Multivariate regression analyses will further assess the relationship between personality traits and academic background, considering gender and field interactions. Our findings highlight the persistence of educational inequalities in academia, despite efforts to create alternative pathways like Tenure Track and Junior Professorships (GWK, 2020; Kreckel, 2015). The overrepresentation of professors from academic backgrounds suggests that non-meritocratic factors still play a significant role. Understanding the personality traits that help individuals overcome these barriers can inform policies to support more inclusive academic advancement.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 22: Analyse von Lernprozessdaten aus der Prüfungsvorbereitung mittels Lernsoftware: Der Einfluss von Zielsetzung auf die tägliche Lerndauer von Studierenden

Julia Hilpert, Marc Philipp Janson
Universität Mannheim, Deutschland

Selbstregulation spielt eine entscheidende Rolle im Lernprozess und ist besonders während der Prüfungsvorbereitung von großer Bedeutung (Zimmerman & Schunk, 2011). Vor allem in digitalen Lernumgebungen ist selbstreguliertes Lernen essenziell (Winters et al., 2008). Trotzdem haben viele Lernende Schwierigkeiten, Lernaktivitäten zu initiieren und aufrechtzuerhalten. Diese Herausforderungen führen oft zu ineffizientem Lernen, erhöhter Prokrastination und schlechteren Prüfungsergebnissen (Steel, 2007).

Ursächlich ist hierbei oftmals kein Mangel an Motivation, d.h. der Intention zu lernen, sondern vielmehr ein Mangel an Volition, d.h. der konsequenten Umsetzung der gebildeten Lernintentionen (Achtziger & Gollwitzer, 2018).

Daher ist es wichtig, mithilfe gezielter Maßnahmen die Volition von Lernenden im selbstregulierten Lernprozess zu stärken. Eine vielversprechende Möglichkeit sind Zielsetzungsinterventionen (Epton et al., 2017; Locke & Latham, 2013), die auf dem Rubikonmodell der Handlungsphasen (Heckhausen & Gollwitzer, 1987) aufbauen: die konkrete Festlegung auf ein Ziel führt zu einer stärkeren Zielbindung, die auch bei auftretenden Hindernissen nicht direkt aufgegeben wird. Stattdessen werden im Angesicht von Hindernissen Anstrengungen erhöht, mehr Ressourcen eingesetzt oder alternative Wege zur Zielerreichung gesucht (Gollwitzer, 1990).

Die vorliegende präregistrierte Feldstudie (https://aspredicted.org/S6P_3TK) untersuchte eine Zielsetzungsintervention zur Erhöhung der täglichen Lerndauer von Studierenden innerhalb einer Lernsoftware. Mit einem einfaktoriellen within-subjects-Design wurde untersucht, ob die Studierenden an Tagen, an denen sie aufgefordert wurden, ein konkretes Lernziel zu formulieren (Experimentalbedingung), länger lernten als an Tagen ohne diese Aufforderung (Kontrollbedingung). Als mögliche Moderatorvariablen untersuchten wir die Prokrastinationstendenz der Studierenden sowie den zeitlichen Abstand der jeweiligen Lerneinheit zur Prüfung. Erwartet wurden höhere Effekte der Zielsetzungsintervention auf die tägliche Lerndauer bei Personen mit hoher Prokrastinationstendenz (Gröpel & Steel, 2008) und bei großem zeitlichem Abstand der Lerneinheit zur Prüfung (Capelle et al., 2022, 2023).

Insgesamt werteten wir 1513 Lerneinheiten von $N = 102$ Studierenden aus, die ein intelligentes tutorielles Lernsystems im Mittel 21.87 Stunden ($SD = 15.26$) zur selbstregulierten Klausurvorbereitung nutzten. 753 der Lerneinheiten beinhalteten die Angabe eines Lernziels (Experimentalbedingung), 760 der Lerneinheiten beinhalteten diese Angabe nicht (Kontrollbedingung). Die tägliche Lernzeit (in Minuten) der Studierenden wurde über Logfiles der Lernsoftware ermittelt. Die Prokrastinationstendenz wurde in einem Prätest zum Beginn der Lernphase mittels einer deutschen Übersetzung der Tuckman Procrastination Scale (Tuckman, 1991) erfasst.

Mehrebenenanalysen (L1: Lerntag, L2: Person) zeigten, dass Studierende mit jedem Tag näher an der Prüfung mehr lernten, $b = 1.59$, $p < .001$. Die tägliche Lernzeit war jedoch unabhängig von der Interventionsbedingung, $b = 11.88$, $p = .415$, sowie von der Interaktion zwischen Interventionsbedingung und zeitlicher Distanz zur Prüfung, $b = -0.10$, $p = .685$. Auch der erwartete interindividuelle Moderationseffekt zwischen der Interventionsbedingung und der Prokrastinationstendenz der Studierenden zeigte sich nicht, $b = -5.21$, $p = .237$.

Explorative between subject Analysen des ersten Lerntags zeigten jedoch, dass Studierende mit der Experimentalbedingung an ihrem ersten Lerntag signifikant mehr lernten als Studierende mit der Kontrollbedingung, $t(89) = 1.89$, $p < 0.05$.

Überraschend war das Ausbleiben des erwarteten Effekts der Zielsetzungsintervention im within-Design. Der signifikante Befund der explorativen Analyse legt jedoch einen starken Generalisierungseffekt nahe: Studierende, die zu Beginn einer Lerneinheit zur Festlegung eines Lernziels aufgefordert wurden, könnten diese Maßnahme als hilfreich empfunden haben und sie deshalb bewusst oder unbewusst auch ohne Aufforderung durch die Lernsoftware angewendet haben. Dieses Verhalten könnte sich sowohl in der Kontrollbedingung innerhalb der Lernsoftware als auch auf andere Lernaktivitäten übertragen haben (vgl. Janson & Dickhäuser, 2019). Am ersten Lerntag war noch keine Generalisierung oder Treatmentdiffusion möglich, was hier den Effekt der Zielsetzung auf die Lerndauer erklären kann.

Die Ergebnisse dieser Studie illustrieren eindrucksvoll die Risiken von Treatmentdiffusion bei within-subject-Designs sowie das bemerkenswerte Potential solcher minimalen Interventionen (vgl. Walton & Wilson, 2018): Allein die Angabe eines Lernziels am ersten Lerntag führte anscheinend zu einer unmittelbaren Generalisierung.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 49: Sozioökonomischer Status und digitale Kompetenzen – Eine meta-analytische Untersuchung der PISA 2022 Studie

Maren Müller, Tamara Kastorff

Zentrum für Internationale Vergleichsstudien (e.V.), TU München, Deutschland

Theoretischer Hintergrund

Digitale Kompetenzen sind spätestens seit der Pandemie von hoher Relevanz. Sie umfassen den kompetenten Umgang mit digitalen Geräten sowie die Bewertung der Vertrauenswürdigkeit von Informationsquellen (Vourikari et al., 2022). Für die Zeit vor der Pandemie ist ein positiver Zusammenhang zwischen sozioökonomischem Status (SES) und digitalen Kompetenzen gut dokumentiert (Siddiq & Scherer, 2019). Der Zusammenhang von SES und digitalen Kompetenzen nach der Pandemie muss jedoch noch eingehender untersucht werden. Da Lernende mit einem niedrigen SES möglicherweise weniger Zugang zu digitalen Ressourcen haben und somit in ihren digitalen Kompetenzen eingeschränkt sind (van Dijk, 2020), haben viele Wissenschaftler*innen Bedenken geäußert, dass die durch die Pandemie bedingten Veränderungen in den Unterrichtspraktiken die sogenannte „digitale Kluft“ vergrößert haben könnten (Beaunoyer et al., 2020).

Gleichzeitig zeigen weitere Studien, dass diese Zusammenhänge durch die Verfügbarkeit von und das Lehren und Lernen mit digitalen Ressourcen abgemildert werden können: González-Betancor et al. (2021) zeigen, dass die Verfügbarkeit und qualitativ hochwertige Nutzung digitaler Ressourcen in Schulen während der Pandemie kompensatorische Effekte für Schüler*innen mit niedrigem SES haben können. Im vorliegenden Beitrag werden diese Befunde anhand der PISA 2022 Studie mit Daten aus 80 Ländern überprüft und sowohl der Zusammenhang zwischen SES und digitalen Kompetenzen als auch kompensatorische Potenziale auf Individual-, Schul- und Staatenebene analysiert.

Fragestellungen

Inwiefern lassen sich die digitalen Kompetenzen von Schüler*innen anhand des SES erklären?

Inwiefern wird der Zusammenhang zwischen dem SES und den digitalen Kompetenzen von Schüler*innen durch die Verfügbarkeit und Nutzung digitaler Ressourcen auf Individual-, Schul- und Staatenebene moderiert?

Methode

Die Studie basiert auf Daten der PISA 2022 Studie, welche $N = 613,773$ 15-Jährige aus 80 Ländern umfasst. Digitale Kompetenzen wurden durch zwei vierstufige Likert-Skalen gemessen: Selbstwirksamkeit bezüglich digitaler Kompetenzen (ICTEFFIC, $\alpha = .92$) und Umgang mit digitalen Informationen (ICTINFO $\alpha = .82$). Die Skalen basieren auf der Selbsteinschätzung der Schüler*innen. Der SES wurde durch einen Index basierend auf dem Bildungsstatus der Eltern, häuslichen Besitztümern und dem beruflichen Status der Eltern gemessen. Moderatorenanalysen berücksichtigen Schüler*innen-Angaben zur Verfügbarkeit und Nutzung digitaler Ressourcen auf Individual- und Schulebene sowie den Digital Readiness Index (DRI) auf Staatenebene (Cisco, 2024), um Varianz in Bezug auf den Fortschritt der Digitalisierung auf Staatenebene zu berücksichtigen.

Zur Beantwortung der Forschungsfragen wurde für alle Staaten mit zwei getrennten linearen Modellen geschätzt, inwiefern der SES digitale Kompetenzen beeinflusst. Anschließend wurden die Ergebnisse meta-analytisch betrachtet und Moderatorenanalysen durchgeführt (vgl. Brunner et al., 2022).

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen einen signifikanten positiven Zusammenhang zwischen SES und digitalen Kompetenzen (ICTEFFIC: $\beta_{summary\ effect} = .16$, 95% CI [.15, .17], $I^2 = 92.05\%$, $r^2 = 0.0382$, CI[0.011, 0.0021], ICTINFO: $\beta_{summary\ effect} = .13$, 95% CI [.12, .13], $I^2 = 82.64\%$, $r^2 = 0.0256$, CI[0.0004, 0.0009]). Die Stärke der Zusammenhänge variiert jedoch signifikant zwischen den Ländern (ICTEFFIC: ($Q(79) = 932.49$, $p < .0001$), ICTINFO: ($Q(33) = 130.03$, $p < .0001$)).

Die Moderatorenanalysen zeigen, dass der Zusammenhang des SES mit den digitalen Kompetenzen für beide untersuchte Skalen bei höherer Verfügbarkeit von digitalen Medien zu Hause (-0.02^{***} , [CI -0.04, -0.01]) und höherem Digitalisierungs-Index (-0.03^{***} , [CI -0.04, -0.02]) schwächer ist. Die Nutzung digitaler Ressourcen zum forschungsbasierten Lernen zeigt einen abschwächenden Effekt auf den Zusammenhang zwischen SES und den Umgang mit digitalen Informationen (-0.03^* [CI -0.05, -0.001]). Dagegen hat die Verfügbarkeit digitaler Ressourcen in der Schule lediglich einen abschwächenden Einfluss auf den Zusammenhang zwischen SES und die Selbstwirksamkeit im Umgang mit digitalen Ressourcen (-0.06^* , [CI -0.11, -0.01]).

Zusammenfassend zeigen die Ergebnisse, dass die Verfügbarkeit und die qualitativ hochwertige Nutzung digitaler Ressourcen sowie der Digitalisierungsgrad von Staaten den positiven Zusammenhang zwischen SES und digitalen Kompetenzen kompensieren können und eröffnen hiermit konkrete Handlungsoptionen zur Verringerung der digitalen Kluft.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 18: Diagnostik heterogener Lernvoraussetzungen im Religionsunterricht

Jonas Bonke, Annalena Sieveke, Claudia Gärtner, Alexander Unser

Technische Universität Dortmund, Deutschland

Diagnostische Kompetenz gilt als ein zentraler Bestandteil des Professionswissens von Lehrkräften. Obwohl in den letzten Jahren sowohl in der allgemeinen empirischen Bildungsforschung (Südkamp & Preatorius, 2017; Herppich et al. 2018) als auch in einzelnen Fachdidaktiken (Brunner et al., 2011; Leuders et al., 2018) eine Vielzahl von Studien zur Diagnosekompetenz vorgelegt wurden, liegen in der Religionsdidaktik bislang nur einzelne Beiträge vor, die sich dezidiert mit Diagnosekompetenz beschäftigen (Klose, 2014; Sieveke et al. 2023) oder diese zumindest als eine Teilkompetenz des fachspezifischen Professionswissens untersuchen (Hofmann, 2008; Fricke, 2017; Leven, 2019).

Diese Forschungslücke will das von der Deutschen Forschungsgemeinschaft (DFG) geförderte Projekt schließen. Mit dem eingereichten Poster möchten wir das methodische Design der Studie auf der Grundlage unserer Pretestergebnisse vorstellen und diskutieren.

Das Projekt besteht aus einer qualitativen (TS 1) und einer quantitativen Teilstudie (TS 2), deren Sample miteinander verbunden sind. Es wird untersucht, (1) ob, wie und mit welchem Ziel Lehrkräfte im Religionsunterricht heterogene Lernvoraussetzungen (Einstellungen, Präkonzepte, kulturelles Kapital und Religiosität) diagnostizieren, (2) ob ihre diagnostischen Urteile korrekt sind, (3) welche Aspekte ihres Professionswissens (Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen) dabei relevant sind, und (4) ob eine korrekte Diagnostik mit einer höheren Unterrichtsqualität (Aktivierung von Schüler*innen, lernförderliches Unterrichtsklima, vielfältige Motivierung, Umgang mit heterogenen Lernvoraussetzungen) im Religionsunterricht einhergeht.

Mittels problemzentrierter Interviews, die inhaltsanalytisch ausgewertet werden, werden in TS 1 25 Religionslehrkräfte zu (a) ihrem Wissen und ihren Überzeugungen bezüglich Diagnose, (b) ihrem für die Diagnose relevanten Fachwissen und fachdidaktischen Wissen sowie ihrem (c) pädagogischen Wissen befragt. Dabei orientieren wir uns an den durch von Aufschnaiter et al. (2015, 750) vorgeschlagenen „Facetten diagnostischer Kompetenz und exemplarische Standards“. Weiter werden die Religionslehrkräfte in TS 1 befragt, wie sie diagnostische Situationen in ihrem Unterricht interpretieren, welche diagnostischen Entscheidungen getroffen werden (sowohl zum Zweck der Diagnose als auch zur Adaption des Unterrichts) und es wird in TS 2 untersucht, wie genau sie Schüler*innenmerkmale wahrnehmen (Urteilsgenauigkeit). Mit dieser Dreiteilung orientieren wir uns an den im Modell von Blömeke et al. (2015) ausgewiesenen situationsspezifischen Fähigkeiten von Lehrer*innenkompetenz: Wahrnehmung, Interpretation und Entscheidung.

Entsprechend ist TS 1 über das Konstrukt der Urteilsgenauigkeit mit TS 2 verbunden. Die 25 in TS 1 interviewten Lehrkräfte füllen in TS 2 für ihre Schüler*innen jeweils einen Fragebogen aus, um deren Lernvoraussetzungen einzuschätzen. Die Schüler*innen (Ziel: N=500-625) füllen einen parallelen Fragebogen aus, der darüber hinaus auch Skalen zur Einschätzung der Unterrichtsqualität beinhaltet. Dabei wird auf etablierte Items und Skalen zurückgegriffen, um die Vergleichbarkeit mit früheren Studien zu gewährleisten. Fachliche und fachbezogene Einstellungen werden über zwei Einstellungsskalen zu Religion in der Gesellschaft und zum Religionsunterricht erhoben (Ziebertz et al. 2003); Präkonzepte mittels Skalen aus Klose (2014); das kulturelle Kapital der Schüler*innen wird mittels einer Skala zu deren Lesegewohnheiten erfasst (Unser 2019); ihre Religiosität über die Kurzversion der Centrality of Religiosity Scale (Huber & Huber 2012) sowie über eine Skala zum Gottesbild (Ziebertz et al. 2003).

Zur Bestimmung der Urteilsgenauigkeit werden zwei Kennwerte berechnet: die Urteilstendenz und die diagnostische Sensitivität (vgl. McElvany et al. 2009; Schrader & Helmke 1987). Die Urteilstendenz berechnet sich über die Differenz von Schüler*innenselbstausskunft und Lehrer*inneneinschätzung. Mittels der Urteilstendenz lassen sich Über- bzw. Unterschätzungen von Schüler*innenmerkmalen auf Seiten der Lehrer*innen beschreiben. Die diagnostische Sensitivität wird über Produkt-Moment-Korrelationen zwischen Schüler*innenselbstausskunft und Lehrer*inneneinschätzung berechnet. Positive Werte sind ein Indikator für eine gute Einschätzung der einzelnen Schüler*innen, wohingegen Werte um 0 für eine nicht-systematische und negative Werte sogar für eine fehlerhafte Erfassung der Schüler*innenmerkmale stehen.

Der Einfluss der Diagnose auf ausgewählte Merkmale der Unterrichtsqualität wird über Mehrebenen-Modelle (Gelman & Hill 2007; Hox et al. 2017) berechnet. Relevante Aspekte der Unterrichtsqualität werden ebenfalls mittels validierter Skalen erhoben (Trautwein et al. 2015), die für den Religionsunterricht sprachlich adaptiert werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 50: Bourdieus kulturelles Kapital im 21. Jahrhundert. Über die Passung der häuslichen Lernumwelt privilegierter Kinder zu den formalen Anforderungen der Grundschulbildung

Ronja Lämmchen

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Soziale Ungleichheiten des Bildungserwerbs sind in ihren Erscheinungsformen und ihrem Ausmaß gut dokumentiert (Bachleitner et al., 2022). Bourdieus Ansatz des kulturellen Kapitals (1982, 1983; Bourdieu & Passeron, 1971) betrachtet zur Erklärung bestehender Ungleichheiten die Rolle (früher) familiärer Lernprozesse für den Bildungserfolg. Demzufolge profitieren Kinder aus privilegierten Familien von kulturellen Praktiken (Aktivitäten), objektiviertem kulturellem Kapital (Lernressourcen) und klassenspezifischen Bildungswerten, welche im institutionellen Bildungssystem gewürdigt werden. Kulturelles Kapital ist somit sozial ungleich verteilt, wirkt sich positiv auf den Bildungserfolg aus (Jaeger, 2022) und entspricht den (formalen und informellen) Anforderungen des Bildungssystems (Lareau & Weiniger, 2003). Welche Ausprägungen kulturelles Kapital dabei annehmen kann und ob sich eher ein *Fokus auf Hochkultur* (z.B. De Graaf et al., 2000) oder die *Passung an Erwartungen im Bildungssystem* (Lareau, 2003) anbietet, ist weiterhin Teil der wissenschaftlichen Debatte (Davis & Rizk, 2018; Tan, 2017; Tan et al., 2019). An dieser Stelle setzt die vorliegende Studie an. Für die Grundschulbildung in Deutschland soll analysiert werden, welche Vorerfahrungen privilegierter Schüler*innen zu den formalen Bildungsanforderungen passen und somit als *Ausprägungen kulturellen Kapitals* gelten.

In einem sequenziell explorativen Mixed-Methods-Design (Creswell et al., 2008) wurde eine qualitative Inhaltsanalyse (Kuckartz & Rädiker, 2022) der Schulgesetze der Bundesländer (N= 16), der Bildungsstandards für die Primarstufe für Mathematik und Deutsch (N= 2) sowie der Rahmenlehr- und Bildungspläne der Bundesländer für die Grundschulschule in Mathematik und Deutsch (N= 32) durchgeführt. Codiert wurden Bildungswerte (z.B. Lernen für den Schulerfolg), Lernaktivitäten (z.B. (Vor-)Lesen), Lernressourcen (z.B. Nachschlagewerke) und außerschulische Lernorte (z.B. Bibliotheken), die in den Dokumenten genannt wurden. Im anschließenden quantitativen Schritt wurden die in allen Bundesländern sowie den Bildungsstandards genannten potenziellen Ausprägungen kulturellen Kapitals auf ihre soziale Stratifikation vor Schuleintritt mit Daten des NEPS (Startkohorte 2, Welle 1; NEPS-Netzwerk, 2024) getestet. Die Analysetichprobe umfasst 2116 Elternteile mit 4-5-jährigen Kindern aus dem Jahr 2011. Es wurden binominale logistische Regressionen mit dem höchsten elterlichen Bildungsabschluss (institutionalisiertes kulturelles Kapital) durchgeführt.

Folgende Ergebnisse können berichtet werden: In der *qualitativen Analyse* wurden für das Fach Deutsch flächendeckend die Lernaktivitäten Gespräche führen und Rollenspiele sowie die Lernressourcen Sach- und Gebrauchstexte, Literatur und Lyrik, Kinderbücher, Nachschlagewerke und Software in den Dokumenten genannt. Im Fach Mathematik finden sich die Lernaktivitäten Gespräche führen, Bauen/ Basteln sowie Malen/ Zeichnen in allen Lehrplänen. Als Lernressourcen werden nur nicht-lineare Textsorten, z.B. Tabellen, in allen Bundesländern erwähnt. Bei den Bildungswerten wurde nur Lernfreude/ Interessensentwicklung einheitlich für das Fach Deutsch berichtet. Keiner der codierten außerschulischen Lernorte (z.B. Bibliotheken) wurde in allen Bundesländern genannt. In die *quantitativen Analysen* konnte (aufgrund der Verfügbarkeit in den Daten) die Häufigkeit folgender häuslicher Aktivitäten einbezogen werden: Bauspiele, Basteln und Rollenspiele, wobei sich nur bei Rollenspielen ein signifikant positiver Zusammenhang mit der elterlichen Bildung zeigt. Folgende Lernressourcen konnten einbezogen werden: der Besitz von klassischer Literatur, Gedichtbänden sowie eines Wörterbuchs – hierbei zeigt sich bei allen ein signifikant positiver Zusammenhang mit der elterlichen Bildung.

Folgende Schlüsse können aus der vorliegenden Studie gezogen werden: *Erstens* kann die von Bourdieu postulierte Passung zwischen der Primärsozialisation privilegierter Kinder und den (formalen) Anforderungen des Bildungssystems für Deutschland im 21. Jahrhundert bestätigt werden. Kritik zur Gültigkeit und Relevanz dieser Theorie (Kingston, 2001; Goldtrope, 2007) kann folglich entkräftet werden. *Zweitens* kann gezeigt werden, dass traditionelle Indikatoren der Hochkultur (Gedichtbände, klassische Literatur; Jaeger, 2009) bereits in der Grundschule als Lernmaterial verwendet werden und damit eine Zusammenführung der o.g. Diskurse vollzogen wird. *Drittens* zeigt die vorliegende Arbeit den Bedarf einer Weiterentwicklung bestehender Instrumente in quantitativen Befragungsstudien, etwa hinsichtlich der Nutzung von Software in frühen Bildungskontexten (Flewitt & Clark, 2020) oder die Bedeutung familiär vermittelter Bildungswerte in der frühen Bildung. *Viertens* dient die Studie zur genaueren Erfassung bestehender Bildungsungleichheiten und kann dazu beitragen bildungspolitische Maßnahmen und Interventionen zu spezifizieren.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 5: Untersuchung von Schüler:innenwissen zu Green Chemistry im Kontext von Bildung für nachhaltige Entwicklung

Laura Freude, Mathias Ropohl

Universität Duisburg-Essen, Deutschland

Der Bildungsauftrag von Schule besteht darin, die Lernenden zu einer erfolgreichen Teilhabe und einer selbstbestimmten Gestaltung ihrer Zukunft zu befähigen (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2019). Dieses Ziel, der Vermittlung von Wissen, Fähigkeiten, Werten und Einstellungen zur Gestaltung einer sozial gerechten, wirtschaftlich erfolgreichen und ökologisch verträglichen gesellschaftlichen Entwicklung, wird im Allgemeinen durch die Bildung für nachhaltige Entwicklung (BNE) beschrieben (UNESCO, 2017). Auch der naturwissenschaftliche Unterricht soll einen Teil zu BNE beitragen, dies ist u.a. im Kernlehrplan (KLP) für Chemie verankert. Einen Ansatz BNE in den Chemieunterricht zu implementieren, stellt Green Chemistry dar (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2019).

Green Chemistry beschreibt zwölf Prinzipien (Anastas & Warner, 2000), die in der Industrie Anwendung finden, um die Herstellung chemischer Erzeugnisse nachhaltig zu gestalten und so zu Klimazielen beizutragen (Armstrong et al., 2018). Aus Perspektive von BNE kann Green Chemistry dazu beitragen, Lernenden zu vermitteln, chemische und soziale Entscheidungen unter Berücksichtigung der zwölf Prinzipien von Green Chemistry zu treffen (Armstrong et al., 2018). Erste Studien zum Thema Nachhaltigkeit zeigen, dass die nachhaltigkeitsbezogenen Einstellungen der Lernenden mit zunehmender Jahrgangsstufe sinken, das anwendungsbezogene Wissen über das Gesamtkonstrukt Nachhaltigkeit allerdings ansteigt (Waltner et al., 2021). Es stellen sich die Fragen, inwieweit Lernende über Wissen zu Green Chemistry verfügen und inwieweit sie dieses Wissen anwenden können.

Im Rahmen des Vorhabens werden drei Erhebungsinstrumente entwickelt und erprobt, um die beiden Forschungsfragen beantworten zu können. Für die Erhebung des Wissens über Green Chemistry wird dieses als Strategiewissen aufgefasst, welches deklaratives und prozedurales Wissen umfasst. Unter dem deklarativen Wissen über Green Chemistry wird das Fachwissen, welches potentiell für die Anwendung der Prinzipien benötigt wird, verstanden. Das prozedurale Wissen bezieht sich hingegen auf die Anwendung der Prinzipien. Für beide Formen des Wissens wird ein jeweils eigenes Testinstrument entwickelt.

Das Testinstrument zum deklarativen Wissen fragt das chemische Fachwissen im Sinne von Dispositionen ab, welches potenziell für die Anwendung der Prinzipien benötigt wird. Der Test wird im Multiple-Choice-Format konstruiert. Der Test zum prozeduralen Wissen wird so konstruiert, dass die Lernenden zunächst einen Informationstext mit dem für die anschließende Aufgabe benötigten deklarativen Fachwissen erhalten, um Vorwissenseinflüsse zu minimieren. Die Aufgabe besteht dann darin einen vorgegebenen Teil einer Versuchsdurchführung, welche ein Prinzip der Green Chemistry missachtet, möglichst nachhaltig und sicher zu gestalten. Das Antwortformat ist ein offenes Format.

Die Inhaltsvalidität wird durch eine Lehrplananalyse der Lehrpläne für die Sekundarstufe I in NRW, Baden-Württemberg und Berlin-Brandenburg geprüft.

Das dritte Erhebungsinstrument, welches entwickelt und erprobt wird, stellt einen Fragebogen zu nachhaltigkeitsbezogenen Einstellungen dar. Der Fragebogen soll als Kontrollinstrument zu den zuvor beschriebenen Testinstrumenten eingesetzt werden. Inhaltlich bezieht sich der Fragebogen neben nachhaltigkeitsbezogenen Einstellungen auch auf selbstberichtetes Verhalten in Bezug auf Nachhaltigkeit und Green Chemistry. Der Fragebogen liegt in Form von Aussagen vor, die anhand einer sechs-stufigen Likert-Skala beantwortet werden.

Eine Pilotierung der Erhebungsinstrumente ist für Herbst 2024 geplant, sodass im Winter 2024 erste Ergebnisse vorliegen. Die Pilotierung findet an einer Gesamtschule in NRW in der achten Jahrgangsstufe und in der Einführungsphase in jeweils zwei Klassen bzw. Kursen mit je ca. 30 Lernenden statt. Auf diese Weise können 50 bis 75 Antworten je Aufgabe von Proband:innen für eine Analyse mit probabilistischer Testtheorie erhoben werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 51: Die Entwicklung der sozialen Kompetenzen und Partizipation von Grundschüler:innen mit besonderen Förderbedarfen mittels temporärer Lerngruppen

Anne Hartmann, Jenny Lenkeit, Antje Ehlert, Michel Knigge, Nadine Spörer

Universität Potsdam, Deutschland

Die soziale Partizipation von Heranwachsenden mit sonderpädagogischem Förderbedarf (SPF) ist ein zentrales Ziel inklusiver Beschulung (Schürer, 2020). Die Einbindung in Peerbeziehungen ist im Schulkontext von besonderer Bedeutung und erfüllt zahlreiche Entwicklungsaufgaben (Hannover & Zander, 2016). Jedoch stehen besonders Schüler:innen mit einem SPF im Bereich emotional-soziale Entwicklung (ESE) in inklusiven Settings neben akademischen vor allem sozialen Herausforderungen gegenüber (Crede et al., 2019; Lütje-Klose, 2018). Ihre häufig ungünstig ausgeprägten sozialen Kompetenzen erschweren ihnen eine positive soziale Partizipation in der Klassengemeinschaft (Blumenthal et al., 2020; Garrote et al., 2017). Schüler:innen mit einem SPF ESE benötigen somit ein erhöhtes Maß an pädagogischer Zuwendung (Crede et al., 2019).

Eine Möglichkeit, diese Schüler:innengruppe in inklusiven Lehr-Lern-Settings differenziert zu fördern, stellen temporäre Lerngruppen (LG) dar. In einer LG werden 5-8 Schüler:innen täglich in 2 Unterrichtsstunden parallel zum herkömmlichen Unterricht hinsichtlich fach- und fachübergreifender Kompetenzen gefördert. Das Lernen in diesen typischerweise jahrgangsübergreifenden Kleingruppen ist durch intensive sonder- und sozialpädagogische Betreuung gekennzeichnet. Die Umsetzung und Wirksamkeit dieses neuartigen Förderkonzepts wurde bislang nicht systematisch untersucht. In Hinblick auf das langfristige Ziel, die soziale Partizipation zu fördern, ist derzeit offen, ob positive Effekte einer intensiven individualisierten Förderung mögliche negative Effekte der temporären Separierung von der Klassengemeinschaft ausgleichen können. Somit stellen sich folgende Fragen:

Wie entwickeln sich die Sozialkompetenzen sowie die peer- und selbsteingeschätzte soziale Partizipation von Schüler:innen der LG im Verlauf eines Schuljahres?

Welche individuellen und strukturellen Merkmale lassen sich für eine positive Entwicklung der sozialen Kompetenzen und Partizipation identifizieren?

Die Daten stammen aus einem längsschnittlichen Evaluationsprojekt zur Einführung der LG in inklusiven Grundschulen in Brandenburg (2023-2024). Zu T1 (Frühjahr 2023) umfasste die Stichprobe 728 Schüler:innen der Jahrgangsstufen 2-5. Davon lernten 41 Schüler:innen in insgesamt 7 LG. Zu T2 (Frühjahr 2024) wurden 1541 Schüler:innen der Jahrgangsstufen 1-6, darunter 71 Schüler:innen aus 11 LG befragt.

Klassenleitungen der Schüler:innen, die temporär die LG besuchen, schätzten die Sozialkompetenzen auf einer vierstufigen Likertskala 0–3 ein ($\alpha_{T1} = .75$, $\alpha_{T2} = .80$; Petermann & Petermann, 2006). Die selbsteingeschätzte soziale Partizipation wurde durch die Skala Soziale Integration mit 6 Items auf einer vierstufigen Likertskala (0–3) erfasst ($\alpha_{T1} = .68$, $\alpha_{T2} = .68$; Rauer & Schuck, 2003). Die Peereinschätzungen zur sozialen Akzeptanz wurden über ein soziometrisches Forced-Choice-Verfahren erfasst (angelehnt an Frederickson & Furnham, 2001). Zur Bildung der Akzeptanzvariable wurden erhaltene Freundschaftsnennungen durch mögliche Nennungen geteilt (Range 0-1). Darüber hinaus wurden Schüler:innen-Informationen (z.B. Zeitraum bzw. Dauer des Besuchs in LG) sowie strukturelle Merkmale zur LG (z.B. personell Ressourcen) erfasst.

Erste längsschnittliche Ergebnisse zeigten, dass Klassenleitungen die Sozialkompetenzen der Schüler:innen der LG zu T2 positiver einschätzten, z.B. im Sozialkontakt: $M_{T1} = 1.53$, $SD = 0.52$; $M_{T2} = 1.81$, $SD = 0.48$ ($p = .005$, $d = 0.55$). In Bezug auf die selbsteingeschätzte soziale Integration in die Klassengemeinschaft zeigte sich hingegen keine signifikante Veränderung ($M_{T1} = 1.73$, $SD = 0.64$; $M_{T2} = 1.88$, $SD = 0.54$; $p = .127$). Gleiches gilt für die peereingeschätzte soziale Akzeptanz der Schüler:innen der LG ($M_{T1} = 0.28$, $SD = 0.19$; $M_{T2} = 0.26$, $SD = 0.19$; $p = .156$).

Die vorläufigen Befunde deuten somit an, dass die intensive pädagogische Arbeit in der LG einen möglichen Ansatzpunkt darstellt, um prosoziales Verhalten und soziale Kompetenzen der Schüler:innen in inklusiven Schulen zu fördern. Möglicherweise benötigt es einen längeren Zeitraum, um positive Entwicklungen auch in der sozialen Partizipation erfassen zu können.

In weiteren Auswertungen werden Informationen der Schüler:innen sowie zur LG berücksichtigt, um mögliche individuelle und strukturelle Merkmale zu identifizieren, die mit der Entwicklung der Sozialkompetenzen sowie der sozialen Partizipation der Schüler:innen assoziiert sind. Diese Zusammenhänge können zum Zeitpunkt der Tagung präsentiert werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 24: Interdisziplinäre Interventionsstudie zu formativem Peer- und Selfassessment im Mathematikunterricht der Primarschule

Karolin Maskos¹, Andreas Schulz¹, Ladina Feucht¹, Selin S. Oeksuez², Sandra Zulliger², Alois Buholzer²

¹Pädagogische Hochschule Zürich, Schweiz; ²Pädagogische Hochschule Luzern, Schweiz

Das prozessbegleitende, formative Assessment (FA) trägt eine entscheidende Rolle zum Lernen in der Schule bei (Black & Wiliam, 2009; Black et al., 2011; Hattie, 2009). Es dient im deutschschweizerischen Lehrplan als eine von drei Beurteilungssäulen kompetenzorientierten Unterrichts (Bildungsdirektion des Kantons Zürich, 2017). Von transparenten Kriterien mit engem fachlichen Bezug ausgehend, werden im FA Lernaufgaben eingesetzt, die dem Kind und der Lehrperson den aktuellen Lernstand aufzeigen. Dieser wird zur Erreichung neuer Lernziele in lernförderlichem Feedback rückgemeldet und der Unterricht entsprechend adaptiert. Sowohl im Peerassessment (PA), dem kooperativen Beurteilen, als auch im Selfassessment (SA), der Selbstreflexion und -beurteilung, werden die Schüler:innen als Ressource füreinander und als Agenten ihres eigenen Lernens aktiviert (Black & Wiliam, 2009; Heritage, 2007). Eine wichtige Rolle bei der Umsetzung von PA und SA kommt dem fachdidaktischen und pädagogischen Professionswissen der Lehrpersonen zu (Ball et al., 2008; Harris & Brown, 2013; Heritage, 2007; Schmidt, 2020).

Der Unterrichtseinsatz von PA und SA im Mathematikunterricht wird als deutlich entwicklungsfähig beschrieben (Buholzer et al., 2020). Als erfolgsversprechend gelten zielgerichtete Weiterbildungen, die das Professionswissen von Lehrpersonen zu PA und SA kombiniert mit einem engen fachlichen Fokus vermitteln und konkrete unterrichtspraktische Anregungen geben (Oeksuez et al., 2024). Es fehlen jedoch entsprechend evaluierte Weiterbildungskonzepte insbesondere im deutschsprachigen Raum (Hebbecker & Souvignier, 2018). Hier setzt die folgend vorgestellte, vom Schweizerischen Nationalfond finanzierte, interdisziplinäre Interventionsstudie an.

Im Rahmen der Studie wurde eine dreiteilige, hybride Weiterbildung (Intervention) von insgesamt 8,5 Stunden (inklusive Vor-/Nachbereitung) für Primarlehrpersonen der Klassenstufen 4 und 5 durchgeführt. Die Intervention wurde auf Grundlage von pädagogisch-psychologischen und mathematikdidaktischen Theorien konzipiert (Oeksuez et al., 2024). Sie befasst sich mit der lernförderlichen Umsetzung von PA bzw. SA am Beispiel des Rechnens auf eigenen Wegen in der Multiplikation und Division (Greiler-Zauchner, 2022; Schulz, 2018). Das variable Anwenden von Rechenstrategien und die Freiräume in deren Notation (Schütte, 2004) eignen sich für die Beurteilung im Rahmen von PA oder SA in Verbindung mit dem Vergleichen von Rechenwegen (Rittle-Johnson & Star, 2011; Smith & Stein, 2018). Vorangegangene Studien (Schulz & Leuders, 2018) zeigten grosse Effekte des Vergleichens und Beurteilens von Rechenwegen, das eine vertiefte, lernförderliche Durchdringung von Rechenstrategien ermöglicht.

Das Poster fokussiert die mathematikdidaktischen Forschungsziele der Studie: (1) Hat eine inhaltspezifisch ausgerichtete Weiterbildung zu PA/SA eine Wirkung auf das fachliche Lernen der Schüler:innen?, (2) Hat diese Weiterbildung auch eine Wirkung auf das fachdidaktische Professionswissen der Lehrperson?, (3) Beeinflusst das Professionswissen der Lehrperson das Lernen der Schüler:innen? und (4) Vermittelt das Professionswissen der Lehrperson die Wirksamkeit der Weiterbildung auf das Lernen der Schüler:innen? (Mediatoranalyse).

Die Stichprobe ($N_{\text{Lehrpersonen}} = 96$, $N_{\text{Schüler:innen}} = 1731$, Stand 27.06.24) unterteilt sich in 50 vierte, 43 fünfte und drei jahrgangsgemischte Klassen der deutschsprachigen Schweiz. Die randomisierte Zuteilung erfolgte in Interventionsgruppen PA ($n=28$ Klassen) bzw. SA ($n=28$) und Wartekontrollgruppe ($n=39$).

Die Daten wurden in einem randomisierten Prä-Post-Follow-up-Wartekontrollgruppendesign im Schuljahr 2023/24 erhoben. Die Lehrpersonen beantworteten vor der Weiterbildung und nach der Unterrichtsdurchführung Fragebögen. Diese umfassten im mathematikdidaktischen Teil das Professionswissen zu Multiplikations- und Divisionsstrategien sowie möglichen Fehlern. Ausserdem gaben die Lehrpersonen Prognosen an, wie oft sie die Strategien bzw. Fehler in der eigenen Klasse erwarten. Die Schüler:innen beantworteten zu drei Messzeitpunkten Textaufgaben (nur t1) sowie Multiplikations- und Divisionsaufgaben. Die Daten sollen mittels mehrbenenanalytischen Regressionsmodellen ausgewertet werden.

Zum Zeitpunkt der 12. Tagung der GEBF ist die Datenerhebung und Kodierung der mathematikdidaktischen Lehrpersonen und Schüler:innen-Fragebögen abgeschlossen. Das Poster berichtet erste Befunde zur Wirkung der Weiterbildung auf das mathematikdidaktische Professionswissen der Lehrpersonen sowie auf das mathematische Lernen der Schüler:innen. Angedacht sind hierfür mehrbenenanalytische Auswertungen. Zudem wird auch der Abgleich der Prognosen der Lehrpersonen zusammen mit den tatsächlichen Leistungsdaten ihrer Klassen thematisiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinwand 52: Kann eine ganzheitliche Lehrkräftefortbildung mehr Lernzuwachs erreichen als ein Intelligentes Tutorsystem allein? Ein Feldexperiment im Englischunterricht der 7. Klasse

Florian Berens¹, Katharina Wendebourg¹, Mareike Kholin², Julia Schmidt-Peterson¹, Leona Colling³, Manuel Hopp¹, Tanja Heck³, Stephen Bodnar³, Walid El Hefny³, Florian Nuxoll³, Christoph Deeg⁴, Katja Krey⁴, Josef Schrader², Hannes Schröter², Benjamin Nagengast¹, Detmar Meurers³, Ulrich Trautwein¹

¹Universität Tübingen, Deutschland; ²Deutsches Institut für Erwachsenenbildung, Bonn; ³IWM Leibniz-Institut für Wissensmedien, Tübingen; ⁴Zentrum für Schulqualität und Lehrerbildung Baden-Württemberg

„Ignoriert die Erkenntnisse aus Schweden!“ (Cress & Trautwein, 2024) rufen prominente Bildungsforscher:innen am 10. April 2024 in Zeit Online auf und fordern nachhaltige Digitalisierung an deutschen Schulen. Hintergrund war eine Stellungnahme des Karolinska-Instituts, in der das Institut die schwedische Digitalisierungsstrategie kritisiert und vor Gefahren der Digitalisierung warnt (Karolinska Institute, 2023). Cress und Trautwein kritisieren Einseitigkeit in der Warnung vor psychischen Risiken unreflektierter Nutzung digitaler Medien, während Vorteile gut orchestrierten digitalen Lernens unterschlagen werden, die beispielsweise Chauhan (2017) oder Hilmayr und Kollegen (2020) belegen. Außerdem werde zu sehr Hardware fokussiert, obwohl digitaler Unterricht ebenso gute Software und Fortbildung der Lehrkräfte benötigt. Letztere Kritik wird auch auf die Digitalisierung in Deutschland übertragen, wo die Digitalpakete ebenfalls vor allem Hardware beschaffen (Deutscher Bundestag, 2023).

Dabei gibt es auch aus der deutschen Bildungsforschung Belege dafür, dass gute Software lernförderlich ist (Scheiter, 2021). Intelligente Tutorsysteme (ITS) beispielsweise können Lernen unterstützen, indem sie Feedback geben, adaptiv auf den Lernprozess reagieren und die Lernenden motivieren (Mousavinasab et al., 2021). In früheren Arbeiten wurde daher ein ITS entwickelt, das für den Einsatz im Englischunterricht in 7. Klassen an Gymnasien konzipiert ist und auf direkte Unterstützung der Schüler:innen durch Feedback einerseits und adaptive Aufgabenzuweisung andererseits abzielt (Autoren, 2017). Eine Studie zum Einsatz des ITS in der Praxis hat gezeigt, dass das Feedback eines ITS die Lernleistung tatsächlich erhöht (Autoren, 2019).

Um das in diesem ITS liegende Potential voll auszuschöpfen, ist Fortbildung für Lehrkräfte zur angemessenen Orchestrierung des Unterrichts mit ITS nötig. Unter Verwendung eines designbasierten Forschungsansatzes (Anderson & Shattuck, 2012) entwickeln wir in einem ko-konstruktiven Prozess, mit Wissenschaftler:innen, Fortbildner:innen und Lehrkräften über mehrere Evaluierungsrunden hinweg ein Lehrkräftefortbildungsprogramm, dass die von Lipowski und Rzejak (2021) angeregten Kriterien für effektive Lehrkräftefortbildungen erfüllt. Insbesondere erstreckt sich unsere Fortbildung über einen längeren Zeitraum (fünf Treffen verteilt über 4,5 Monate), verknüpft Input-, Erprobungs- und Reflexionsphasen und wissenschaftliche Inhalte mit praktischen Erfahrungen, stärkt kollegiale Kooperation und Coaching. Dabei orientiert sie sich inhaltlich am aktuellen Stand der Unterrichtsforschung, bezieht sich auf zentrale unterrichtliche Anforderungen und vertieft dazu ausgewählte Inhalte.

In einem Feldexperiment testen wir die Hypothese, dass Schüler:innen, die mit unserem ITS arbeiten, bessere Lernleistungen zeigen, wenn ihre Lehrkraft an unserer Lehrkräftefortbildung teilnimmt. Zu diesem Zweck haben 45 Lehrkräfte Anfang Februar 2024 unser ITS in ihren Englischklassen der Jahrgangsstufe 7 eingeführt und damit bis Mitte Juli gearbeitet. Eine zufällige Auswahl von 23 dieser Lehrkräfte hat außerdem an unserer Lehrkräftefortbildung teilgenommen, die anderen 22 Lehrkräfte sind nur technisch instruiert worden.

Lehrkräfte und Schüler:innen haben vorher und im Anschluss an ausführlichen Befragungen und Tests teilgenommen. Beide wurden über ihre Wahrnehmung des Unterrichts (Jaekel et al., 2021), ihre Motivation für den Englischunterricht (u.a. Gaspard et al., 2017) und ihre Persönlichkeit (u.a. Soto et al., 2017) befragt, zudem wurden soziodemographische Daten erhoben. Die Schüler:innen wurden bezüglich ihrer Englischkenntnisse getestet, die Lehrkräfte bezüglich ihres technologisch-pädagogischen Wissens (TPACK) (Lachner et al., 2021).

Die primäre Zielvariable unseres Feldexperiments ist die Englischleistung der Schüler:innen in drei selbst entwickelten Tests, die sich auf drei Grammatikthemen beziehen, die während des Schulhalbjahres behandelt wurden. Als wichtige sekundäre Zielvariablen betrachten wir einen C-Test als ergänzendes Maß der allgemeinen Englischkompetenz, das TPACK der Lehrkräfte, die von den Schüler:innen wahrgenommene Unterrichtsqualität mit Schwerpunkt auf der Subdimension Binnendifferenzierung, die Motivation der Schüler:innen und die Nutzung des ITS, gemessen an den entstandenen digitalen Verhaltensdaten.

Erste Ergebnisse deuten darauf hin, dass sich tatsächlich Unterschiede zwischen den Gruppen zeigen. Insbesondere in Nutzungsumfang und Art der Nutzung unseres ITS unterscheiden sich die Gruppen erheblich. Dies zeigt, dass eine Digitalisierung, die geeignete Software und insbesondere auch die Lehrkräfte mitdenkt, fruchtbar für den Lernerfolg ist.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 30: ADA MINToring - Förderung der Lernwirksamkeit und der Geschlechtergerechtigkeit im MINT-Unterricht durch einen weiblich konnotierten KI-Chatbot

Cilia Rücker, Prof. Dr. Sebastian Becker-Genschow

Universität zu Köln, Physikdidaktik, Digitale Bildung mit Schwerpunkt KI, Deutschland

Die **Integration von KI-Chatbots in die schulische Bildung** stellt einen vielversprechenden Ansatz dar, um Herausforderungen wie dem Mangel an individueller Betreuung sowie unterschiedlichen Lerngeschwindigkeiten und Lernstilen des traditionellen Schulsystems wirksam zu begegnen (Tahiru, 2021). Als **virtuelle Lernassistent:innen** bieten KI-Chatbots die Möglichkeit, Schüler:innen sofortige, interaktive, personalisierte Lernunterstützung bieten zu können (Crompton & Burke, 2023). Erste Studien geben Hinweise darauf, dass auf diesem Weg die Lernmotivation, das Lerninteresse, das Lernengagement und damit einhergehend die Lernleistung gefördert werden können (Chen et al., 2020, Deng & Yu, 2023, Okonkwo & Ade-Ibijola, 2021, Wollny et al., 2021). Außerdem können negativ-deaktivierende Emotionen wie Ängste und Stress während des Lernprozesses reduziert werden (Pérez et al., 2020).

Kaum erforscht ist bisher der Ansatz, KI-basierten Lernassistent:innen eine virtuelle Persönlichkeit zu verleihen, um simultan zur **Förderung der Lernwirksamkeit** einen möglichen **Abbau von Stereotypen** zu untersuchen. Dabei ist hinreichend belegt, dass ein Kontakt zu weiblichen Vorbildern geschlechterstereotype Vorstellungen besonders im MINT-Bereich abbauen kann (Gioannis et al., 2023). (Amando et al., 2023, Plant et al., 2009) liefern Hinweise darauf, dass dies auch mit computerbasierten Chatbots möglich ist.

Ziel dieser Studie ist es deshalb, die Auswirkungen des Einsatzes eines weiblich konnotierten, personalisierten KI-Chatbots als Lernassistentin im Mathematikunterricht neunter Klassen an weiterführenden Schulen auf die Lernwirksamkeit und die Einstellungen gegenüber dem Mathe-Gender-Stereotyp zu untersuchen. Letzteres entspricht der vereinfachten, generalisierten Überzeugung, dass weiblich gelesene Personen aufgrund ihres Geschlechts über geringere mathematische Fähigkeiten verfügen als männlich gelesene Personen (Nosek et al., 2010).

Die **zentrale Forschungsfrage** lautet daher: *Führt der Einsatz der KI-gestützten Lernassistentin Ada im Vergleich zu traditionellen Differenzierungsmaßnahmen zu einer Verbesserung des Lernerfolgs und einem Abbau des Mathe-Gender-Stereotyps?*

Im Rahmen einer **präregistrierten, quasi-experimentellen, cluster-randomisierten Kontrollstudie** werden Daten von bis zu 400 Schüler:innen neunter Klassen an weiterführenden Schulen erhoben. Die Intervention umfasst in beiden Gruppen drei Unterrichtseinheiten, deren didaktischer Inhalt das Heron-Verfahren ist, ein iteratives Verfahren zur näherungsweisen Berechnung von Quadratwurzeln. Basierend auf einem **Mixed Methods Design** werden quantitative Daten mittels Leistungstests und Fragebögen sowie qualitative Daten durch leitfadengestützte Interviews mit einer repräsentativen Stichprobe von Schüler:innen erhoben. Untersucht werden die unabhängige Variable Lernunterstützungsstrategie und die abhängigen Variablen wahrgenommene Männlichkeit der Mathematik, kognitive Belastung, situatives Interesse, Emotionen, Bedienbarkeit und Nützlichkeit des KI-Chatbots, Einstellungen gegenüber dem KI-Chatbot im Unterricht, wahrgenommene Kompetenzen des KI-Chatbots sowie soziodemographische Daten.

Die Studie befindet sich aktuell in der Phase der Datenerhebung, welche noch nicht vollständig abgeschlossen ist. Die bereits erhobenen Daten geben Hinweise darauf, dass der Einsatz des KI-Chatbots Ada positive Effekte sowohl auf den Lernerfolg als auch auf die Reduktion geschlechtsspezifischer Stereotype im Mathematikunterricht haben könnte. Die vorläufigen Ergebnisse deuten darauf hin, dass ADA MINToring einen vielversprechenden Ansatz zur Schließung der aufgezeigten Forschungslücke darstellt.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 53: Individuelle kognitive Aktivierung – Potenziale ausgehend von Lehrperson, Schüler:in und Material

Julia Warmdt, Marvin Bauersfeld, Elisabeth Zißler, Sanna Pohlmann-Rother, Christoph Ratz, Henrik Frisch

Julius-Maximilians-Universität Würzburg, Deutschland

Theoretischer Hintergrund

Spätestens seit der Ratifizierung der UN-BRK im Jahr 2009 wird ein gemeinsames Lernen von Schüler:innen mit und ohne geistige Behinderung im Regelunterricht als Ziel deklariert. Zur Gestaltung solcher inklusiven Unterrichtssituationen wird digitalen Medien ein Potenzial zugeschrieben, da durch digitale Medien verschiedene Teilhabechancen für eine heterogene Schüler:innenschaft ermöglicht werden können (Böttiger, 2023). Ein exemplarisches Konzept zur inklusiv-digitalen Unterrichtsgestaltung stellt das *Digital Storytelling* dar. Dabei kann eine literarische Geschichte neben herkömmlichen Ausdrucksmöglichkeiten (Bild, Text) auch mit erweiterten (Foto, Video, Audio) gelesen und auf verschiedenen Leistungsniveaus von Kindern mit heterogenen Lernvoraussetzungen weitererzählt werden (Warmdt et al., 2023).

Normativ wird die Forderung gestellt, einen solchen digital-inklusive Unterricht, wie beispielsweise zum *Digital Storytelling*, qualitativ hochwertig zu gestalten. Aktuelle Veröffentlichungen zur Qualität im inklusiven (Bohl, 2017) und digitalgestützten Unterricht (Quast et al., 2021) nehmen Bezug auf die klassische Unterrichtsqualitätsforschung und einhergehend auf die kognitive Aktivierung. In einem kognitiv aktivierenden Unterricht ist es das Ziel, alle Lernenden entsprechend ihrer individuellen Lernvoraussetzungen zur Auseinandersetzung mit anspruchsvollen Aufgaben anzuregen, welche Verstehen und Schlussfolgern fördern (Groß-Mlynek et al., 2022).

Fragestellung

Wird ein Blick auf die (fachspezifische) Operationalisierung sowie die Wirkmechanismen zur kognitiven Aktivierung gerichtet, lässt sich in den Forschungsarbeiten ein divergentes Bild verzeichnen (Groß-Mlynek et al., 2022). Ein Grund hierfür kann sein, dass die Effekte der kognitiven Aktivierung in den Studien über die gesamte Klasse und nicht mit Blick auf die einzelnen Kinder in den Fokus genommen werden (Begrich et al., 2023). Gerade in inklusiven Unterrichtssituationen, welche von Schüler:innen mit heterogenen Lernausgangslagen besucht werden, kann davon ausgegangen werden, dass unterschiedliche Potenziale des Lehrpersonen- oder Schüler:innenhandelns bzw. der Materialgestaltung zu individuell kognitiv aktivierenden Lehr- und Lernsituationen für das einzelne Kind führen können (Rieser & Decristan, 2023). Diesem differenziellen Blick auf die kognitive Aktivierung wird im Rahmen des Beitrags gefolgt, indem die Potenziale ausgehend von dem Lehrpersonen- oder Schüler:innenhandeln bzw. der Materialgestaltung für individuell kognitiv aktivierende Lehr- und Lernsituationen im Fokus stehen. Es stellt sich dabei folgende übergeordnete Forschungsfrage: Von wem oder was geht das Potenzial zur individuellen kognitiven Aktivierung beim *Digital Storytelling* im inklusiven Unterricht aus?

Methode

Um das Potenzial zur individuellen kognitiven Aktivierung auf der Prozessebene des Unterrichts zu analysieren, wurde eine Projektwoche zum *Digital Storytelling* in inklusiven Kleingruppen konzipiert, die in acht inklusiven, ersten bzw. zweiten Klassen durchgeführt und videografiert wurde. Die erhobenen Videodaten von 38 Schüler:innen in 18 Kleingruppen wurden aufbereitet, transkribiert und ausgewertet. Den Fokus der Auswertung bildet eine Unterrichtsphase, in der die Schüler:innen in Kleingruppen eine Geschichte digital weitererzählen. Ausgangspunkt der Auswertung stellen Kodierungen zur individuellen kognitiven Aktivierung dar ($N = 1927$) (z. B. Aktivieren von Vorwissen, Selbstständige Suche nach Problemlösung). Zu diesen individuell kognitiv aktivierenden Lehr- und Lernsituationen wurde das jeweilige Potenzial ausgehend von dem Lehrpersonen- oder Schüler:innenhandeln bzw. der Materialgestaltung nachkodiert. Hierzu wurde ein induktiv-deduktives Kategoriensystem mittels inhaltlich-strukturierender, qualitativer Inhaltsanalyse erstellt (Kuckartz & Rädiker, 2022).

Ergebnisse und ihre Bedeutung

Es wurden Oberkategorien definiert, die bis zur vierten Unterebene ausdifferenziert sind: Potenzial ausgehend von Material, Lehrperson und Schüler:in. Zu knapp dreiviertel können Potenziale zu individuell kognitiv aktivierenden Lehr- und Lernsituationen von dem Material beschrieben werden ($n = 1413$). In circa jeder fünften individuell kognitiv aktivierenden Lehr- und Lernsituation geht das Potenzial von dem:der Schüler:in aus ($n = 353$) und selten ist die Lehrperson potenzialgebende Person ($n = 114$). Entgegen der häufig vorgenommenen Fokussierung der Lehrperson als potenzialgebende:n Akteur:in (z. B. Quast et al., 2021) betonen die Ergebnisse dieser Studie die Bedeutung des Materials für individuell kognitiv aktivierende Lehr-Lernsituationen und damit die qualitätsvolle Materialgestaltung durch Lehrpersonen. Außerdem zeigen sich von Mitschüler:innen ausgehende Potenziale, die genutzt werden können, um gemeinsame Lernsituationen in inklusiven Kleingruppen zu beschreiben, die für eine qualitativ hochwertige inklusive Unterrichtsgestaltung bedeutsam sind.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 11: Kompetenz, Kontext und kooperatives Lernen: Die Vorhersage von qualitativ hochwertiger Gruppenarbeit durch Merkmale der Lehrkraft und des Schulumfelds

Nora Ries^{1,3}, Franziska Baier-Mosch^{1,3}, Mareike Kunter^{1,2,3}

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt a. M., Deutschland; ²Goethe-Universität Frankfurt, Frankfurt a. M., Deutschland; ³DeA Center for Research on Individual Development and Adaptive Education for Children at Risk, Frankfurt a. M., Deutschland

Mit kooperativem Lernen können Schüler*innen in vielerlei Hinsicht gefördert werden (Hattie, 2023), dabei scheinen gerade Schüler*innen mit bildungsrelevanten Risiken vom gemeinsamen Lernen zu profitieren (z. B. Ginsburg-Block et al., 2006). Dennoch nutzen Lehrkräfte kooperatives Lernen eher selten und wenn, häufig mit geringer Umsetzungsqualität (Adl-Amini et al., 2023). Dies wirft die Frage auf, welche Faktoren für die Umsetzung relevant sind. Dabei könnten nicht nur Merkmale auf Ebene der Lehrkräfte (wie Aspekte professioneller Kompetenz; Abramczyk & Jurkowski, 2020), sondern auch des Schulkontextes (wie schulische Ressourcen; Sarkio et al., 2023) eine wichtige Rolle spielen. Dabei weisen erste Studien darauf hin, dass Aspekte professioneller Kompetenz von Lehrkräften nur genutzt werden (können), wenn sie einen günstigen Kontext vorfinden (Filippou et al., 2022; Ries & Wolf et al., 2024).

Daher stellen sich folgende Forschungsfragen:

1. Häufigkeit kooperativer Lernangebote

- Welche individuellen und kontextuellen Merkmale von Lehrkräften sagen die Einsatzhäufigkeit voraus?
- Wird ein möglicher Zusammenhang von schulischem Anteil an Risikoschüler*innen und Einsatzhäufigkeit durch individuelle und kontextuelle Ressourcen moderiert?

2. Qualität kooperativer Lernangebote:

- Welche individuellen und kontextuellen Merkmale von Lehrkräften sagen die Einsatzqualität voraus?
- Wird ein möglicher Zusammenhang von schulischem Anteil an Risikoschüler*innen und Einsatzqualität durch individuelle und kontextuelle Ressourcen moderiert?

An unserer Fragebogenstudie nahmen 238 Lehrkräfteallgemeinbildender Schulen in Deutschland teil. Dabei wurden Qualität und Häufigkeit kooperativer Lerneinheiten (Kunter et al., 2016) im Selbstbericht erfasst. Als Lehrkraftmerkmale wurden Aspekte der professionellen Kompetenz nach Kunter et al. (2013) erfasst: Wissen über kooperatives Lernen (*Thought Listing Task* aus Kunter et al., 2016), Überzeugungen zum kooperativen Lernen (Kunter et al., 2016), Enthusiasmus (adapt. n. Kunter et al., 2016) und selbstregulative Fähigkeiten beim Einsatz kooperativen Lernens (adapt. n. Martin & Marsh, 2008). Als Kontrollvariable wurde die emotionale Erschöpfung (Enzmann & Kleiber, 1989) einbezogen. Als kontextuelle Merkmale wurden zeitliche, materielle, kollegiale und räumliche Ressourcen sowie der geschätzte Anteil an Risikoschüler*innen auf Schulebene im Selbstbericht der Lehrkräfte erhoben. Zur Beantwortung der Fragestellungen wurden multiple Regressionen gerechnet mit Einsatzqualität und Einsatzhäufigkeit als Outcome-Variablen und den zuvor aufgeführten Variablen als Prädiktoren. Dabei wurden Interaktionsterme zwischen Anteil an Risikoschüler*innen und den lehrkraftbezogenen sowie kontextuellen Ressourcen modelliert.

Ergebnisse:

Insgesamt können 37% der Varianz der Einsatzhäufigkeit und 27 % der Varianz der Einsatzqualität durch unsere Prädiktoren erklärt werden.

1. Die Umsetzungshäufigkeit wird von Enthusiasmus, Überzeugungen und zeitlichen Ressourcen, einer Lehrkraft vorausgesagt (Haupteffekte). Es zeigt sich ein signifikanter Interaktionseffekte zwischen Risikoanteil und materiellen Ressourcen.
2. Die Umsetzungsqualität wird von selbstregulativen Fähigkeiten vorausgesagt (Haupteffekt). Es wurde keiner der Interaktionsterme signifikant.

Ziel unserer Studie war es, Merkmale von Lehrkräften und ihrem schulischen Kontext zu identifizieren, die mit der Implementierung von kooperativem Lernen in Schulen zusammenhängen. Es ergaben sich motivationale, überzeugungsbezogene und selbstregulative Merkmale der Lehrkräfte sowie deren zeitliche Ressourcen als wichtige Faktoren. Bezogen auf das Unterrichten von Risikoschüler*innen betonen unsere Studienergebnisse die hohe Bedeutung einfach zugänglicher Unterrichtsmaterialien.

Auch wenn unsere Ergebnisse hinsichtlich der Kontextmerkmale aufgrund des fehlenden Mehrebenenendesigns mit Vorsicht zu interpretieren sind, liefern sie dennoch erste Ansatzpunkte beispielsweise für die Schulentwicklung. Es stellt sich etwa die Frage, wie ein mit dem kooperativen Lernen verbundene Zeitaufwand reduziert werden kann. Kompetenzaspekte (wie Motivation, Überzeugungen und Selbstregulation beim kooperativen Lernen) und zeiteffiziente Einsatzmöglichkeiten kooperativen Lernens könnten sinnvolle Weiterbildungsziele für Lehrkräfte darstellen. Auch einfach zugängliche und adaptierbare Unterrichtsmaterialien zum kooperativen Lernen könnten Hilfestellung bieten, damit gerade die Schüler*innen vom kooperativen Lernen profitieren können, die wohl am meisten davon hätten.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 54: Urteilen über religiöse Konflikte – Messung eines Strukturmodells zur Urteilskompetenz im Religionsunterricht

Frederike Gabelt

Technische Universität Dortmund, Deutschland

Theoretischer Hintergrund

Entgegen der Säkularisierungsthese und der Zunahme von Nichtreligiosität in modernen Gesellschaften werden religiöse Fragen in der Öffentlichkeit weiterhin breit und oft kontrovers diskutiert. Aus diesem Grund gewinnt die Fähigkeit, Fragen und Konflikte im Zusammenhang mit Religion zu beurteilen, zunehmend an Bedeutung. Eine der Schlüsselkompetenzen, die Schüler:innen im Religionsunterricht entwickeln sollen, ist daher die Fähigkeit zur reflektierten Urteilsbildung. (vgl. Kernlehrplan NRW 2019) Bisher ist in den offiziellen Vorgaben für das Fach allerdings unzureichend formuliert, was im Sinne der Urteilskompetenz gelernt und geprüft werden soll. (vgl. Muth 2021, 249f.) Zudem fehlt es Lehrkräften an fachdidaktischem Wissen, wie kontroverse Themen im Religionsunterricht thematisiert werden können. (vgl. von der Lippe 2021) Es bedarf daher einer Klärung, wie Urteilskompetenz von Schüler:innen im Religionsunterricht gefördert und geprüft werden kann.

Um diese Forschungslücke zu schließen, wird empirisch validiertes Wissen über die Struktur des Urteilsprozesses und der Faktoren, die die Kompetenz beeinflussen, benötigt. Die Entwicklung eines solchen Kompetenzstrukturmodells und eines Instruments zur Kompetenzmessung stellt nach Klieme et al. (2013, 6) ein zentrales Ziel jeder Fachdidaktik dar. Bisher wurden für verschiedene Schulfächer Modelle entwickelt, die die Urteilskompetenz von Schüler:innen abbilden, z.B. in der Biologiedidaktik (Reitschert et al. 2007) und der Politikdidaktik (Manzel/Weißenö 2017). Um passende Lernangebote für religiöse Lernprozesse zu generieren, bedarf es jedoch Erkenntnisse über die fachspezifische Struktur religiöser Urteilskompetenz.

Fragestellung

In der Studie wird die Frage untersucht, ob sich ein solches theoretisch formulierte Modell religiöser Urteilskompetenz anhand empirischer Daten validieren lässt. Da bisher nur wenig Kenntnisse vorliegen, wie Urteilskompetenz empirisch gemessen werden kann (vgl. Weißenö/Weißenö 2021), soll mit dieser Untersuchung ein Beitrag zur empirischen Erforschung von Urteilskompetenz geleistet werden. Aus den Ergebnissen sollen zudem fachdidaktische Schlussfolgerungen abgeleitet werden, um religionsdidaktische Fragen zur Förderung von Urteilskompetenz zu klären.

Methode

Zur Modellmodellierung wurde sich an dem vierstufigem Schema von Mayer und Wellnitz (2014) zur Entwicklung von theoretischen Kompetenzstrukturmodellen orientiert. Der konzeptionelle Rahmen des Kompetenzmodells ist an das Modell der Scientific Literacy aus der PISA-Studie angelehnt. (vgl. Reiss et al. 2016) Die Konzeptualisierung der Teilkompetenzen wurde von Strukturmodellen aus anderen Fachdidaktiken (vgl. Reitschert et al. 2007; Manzel/Weißenö 2017) beeinflusst.

Das Modell wird mit einem quantitativen Ansatz mittels eines Fragebogens getestet. Zur Messung der Urteilskompetenz absolvieren Schüler:innen einen vignettenbasierten Test, wobei die Antworten auf einer Niveauskala von richtig, teilweise richtig und falsch codiert werden. Die Validität des Modells wird mittels der konfirmatorischen Faktorenanalyse überprüft. (vgl. Backhaus/Erichson/Weiber 2015) Die externen Faktoren Wissen, Motivation und Einstellungen werden durch einen Begleitbogen erhoben. Zur Messung des Wissens lösen die Schüler:innen vierzehn Items, die Faktenwissen zu den religiösen Konflikten aus den Vignetten abfragen. Darüber hinaus werden Items aus internationalen Bildungsstudien verwendet, um Motivation und Einstellungen zu erfassen. (u.a. ICCS-Studie 2016, PISA-Studie 2015) Um zu messen, wie diese externen Faktoren Urteilskompetenz beeinflussen, werden Strukturgleichungsmodelle berechnet. (vgl. Backhaus/Erichson/Weiber 2015)

Ergebnisse und ihre Bedeutung

Im Rahmen der Studie wurde ein Kompetenzmodell theoretisch entwickelt und in ein quantitativ-empirisches Testinstrument zur Messung der Urteilskompetenz überführt. Dieses Testinstrument wurde im Rahmen eines Pretests Anfang 2023 evaluiert und verbessert. Die Hauptstudie fand von Mitte 2023 bis Anfang 2024 an nordrhein-westfälischen Schulen mit Schüler:innen der Sekundarstufe II (N=519) statt. Die Daten wurden anschließend bis Ende Juni 2024 mittels eines Manuals codiert, wobei eine sehr gute Inter-coder-Reliabilität ($k=.816$) erreicht wurde. Derzeit werden die Daten mittels der konfirmatorischen Faktorenanalyse und Strukturgleichungsmodellen ausgewertet.

Erste Eindrücke aus der Berechnung der Faktorenanalyse weisen darauf hin, dass sich die Grundstruktur des Kompetenzmodells validieren lässt ($CFI=.957$, $TFI=.950$, $RMSEA=.042$). Ein paar Items weisen allerdings eine geringere Faktorladung auf das latente Konstrukt der entsprechenden Teilkompetenz auf und müssen daher näher untersucht werden. Auf dem Poster sollen die final ausgewerteten Ergebnisse zur Validität des Modells und seiner Einflussfaktoren dargestellt und diskutiert werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 17: Wie komplex rezipieren Lehrkräfte Lernstandsergebnisse? Ergebnisse einer Think-Aloud-Studie

Fabian Burkart¹, Sarah Bez², Samuel Merk¹

¹Pädagogische Hochschule Karlsruhe; ²Universität Tübingen

Theoretischer Hintergrund

Die Nutzung unterrichtsbezogener Daten, wie z.B. Ergebnisse formativer oder summativer Assessments, als Basis für pädagogisch-didaktische Entscheidungen (sog. *data-based decision making*, DBDM) wird von Lehrkräften zunehmend gefordert (Lai & Schildkamp, 2013). Obwohl zu DBDM Studien vorliegen, welche keine positiven Effekte fanden (bspw. Henderson et al., 2007), existieren Quasi-Experimente und Metaanalysen, die vielversprechende Effekte auf Schülerleistungen zeigen (Carlson et al., 2011; Faber et al., 2023; Visscher, 2021).

Für erfolgreiches DBDM gelten Fähigkeiten im Generieren, Analysieren und Interpretieren von Daten, sog. *data literacy*, als notwendige Voraussetzungen (Kippers et al., 2018; Wurster et al., 2023). Nach Mandinach & Gummer (2016) ist eine zentrale Komponente dieser *data literacy* ein hinreichendes Verständnis von "*statistics and psychometrics*" (S. 370), wie etwa das Interpretieren von Maßen der zentralen Tendenz oder Streuung. Weiterhin wird angeführt, dass Lehrkräfte grafische Abbildungen von Daten verstehen und nutzen können sollten, um Daten in Informationen umzuwandeln (S. 371 f.).

Eine solche *graph literacy* (Friel et al., 2001) wird oftmals in drei hierarchische Level gegliedert: *Reading the data* bezeichnet die Fähigkeit des Ablesens direkt gegebener Informationen aus Grafiken (Chick & Pierce, 2013) während *reading between the data* das Erkennen und Beschreiben von Beziehungen zwischen Datenpunkten meint. *Reading beyond the data* bezieht sich z.B. auf zusammenfassende Aussagen (Bez et al., 2021) wie etwa das Erkennen schiefer Verteilungen.

Fragestellungen

Da für funktionales DBDM Rezeptionsprozesse idealerweise auf allen drei Leveln vorkommen (Espin et al., 2018), wird vorliegend untersucht, wie häufig Lehrkräfte bei der Rezeption von Lernstandsergebnissen unterschiedlich komplexe statistische Entitäten, wie Mittelwertsunterschiede, Bimodalität und Schiefe, durchschnittlich adressieren und inwieweit eine höhere allgemeine *graph literacy* (Kompetenz) diese Häufigkeit (Performanz) prädiziert.

Methode

Zur Beantwortung der Forschungsfragen wurden Bildschirminterviews mit 47 baden-württembergischen Lehrkräften ($M_{\text{Alter}} = 37.5$, $SD_{\text{Alter}} = 9.4$; 64% weiblich) der Primar- und Sekundarstufe geführt, welche in einer interaktiven App artifizielle Halbjahresnoten rezipierten. Zur Erfassung kognitiver Elaborationen während der Datenrezeption wurde die nicht-reaktive simultane Think-Aloud-Methode angewendet (Leighton, 2017).

Die Lehrkräfte konnten selbstständig unterschiedliche Visualisierungen für die Fächer Deutsch, Mathe und Englisch zu jeweils zwei Klassen einblenden. In den Daten waren, in Anlehnung an Mandinach & Gummer (2016), sieben statistische Entitäten erkennbar: Ausreißer (1), Bimodalität (2), Mittelwertsdifferenz (3), Schiefe (4), Streuung im Sinne des Wertebereichs (5a), der Mean Average Deviation (5b) und nicht näher spezifizierter Heterogenität (5c). Mithilfe eines induktiv-deduktiv generierten Kategorienschemas wurden die alltagssprachlichen Aussagen der Lehrkräfte (z.B. "Die Noten der Klasse sind weit gestreut.") von zwei unabhängigen Ratern in Eventkodierungen diesen Kategorien zugeordnet (Krippendorfs $\alpha > .63$). Separat wurde die allgemeine *graph literacy* der Lehrkräfte durch einen schriftlichen Test (Merk et al., 2020) erfasst.

Ergebnisse

Durchschnittlich adressierte jede Lehrkraft 3.1 mal (95% CI[2.6, 3.8]) die Kategorie 5c (nicht näher spezifizierte Heterogenität) und 1.9 mal (95% CI[1.5, 2.4]) den Wertebereich (5a). Die am dritt- und vierthäufigst adressierten Kategorien waren die Mittelwertsdifferenz (1.6 mal, 95% CI[1.3, 2.1]) und die Bimodalität (0.9 mal, 95% CI[0.6, 1.2]), während die Streuung im Sinne der Mean Average Deviation (5b) am seltensten erwähnt wurde (0.1 mal, 95% CI[0.0, 0.2]).

Eine Prädiktion dieser Häufigkeiten mithilfe einer bayesianischen Mehrebenen-Poissonregression zeigte einen kleinen Effekt (Incidence Ratio = 1.95), der inferenzstatistisch abgesichert werden konnte (Posterior Probability $P[\text{Incidence Ratio} > 1] = .96$). Obwohl sich Limitationen u.a. durch die kleine Stichprobe und die Tatsache, dass die Think-Aloud-Methode nicht alle kognitiven Prozesse erfassen kann (Fox et al., 2011), ergeben, zeigt sich, dass die allgemeine *graph literacy* von Lehrkräften das Adressieren von komplexeren statistischen Entitäten prädizieren kann.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 55: Code in the head: Cognitive activation and individualized support through digital learning with informative performance feedback

Sabrina Navratil, Samuel Wissel, Christian Müller, Jochen Schäfer, Frederik Armknecht, Marina Friedrich-Schieback, Stefan Münzer

Universität Mannheim, Deutschland

Background:

Universities still face the challenge of adapting the dynamic digital age's demands in teaching and learning, particularly in understanding how students engage with digital learning materials. This project focuses on the digitalization of previously paper-based learning task sheets in an informatics course with active and generative learning assignments such as writing code, developing formulas and algorithms, or describing data structures. Previously, the paper-based learning materials limited the depth of insight that could be gained into student engagement and lacked the opportunity to provide individual feedback to students. By using an interdisciplinary approach that combines expertise in higher education didactics and psychology, this study analyzes digital log data and relates indicators derived from the log data with exam success.

Research question:

Does the use of the provided digital learning materials over the course of the semester predict final course performance?

Method:

Students from the School of Business Informatics and Mathematics of a German university used a digital learning management system to complete learning tasks on programming, algorithms, and data structures. The platform allowed repeated task-solving and informative performance feedback during self-study (Roediger et al., 2011; Roediger & Pyc, 2012). Students thus could repeatedly solve the tasks and received direct automated feedback ("correct" or "incorrect"). In addition to individual digital learning, exercise sheets were completed in tutorials and submitting weekly learning tasks was required for exam admission. Thus, engagement with the digital learning materials was one learning opportunity among others in this course.

Ninety-eight students provided consent for analyzing their log data and exam performance. We derived indicators for the attempts in solving the exercises, such as the average time spend solving, the number of task repetitions, accuracy of the solutions, or (in addition) the points achieved in the mandatory weekly assignments.

Results and Discussion

The results of a multiple regression ($F(5,90) = 4.18$, $SE = 15.09$, $adj. R^2 = .1434$, $p = .0018$) show that the digital learning indicators were related to exam performance with the following differential pattern: A higher overall number of answered items was negatively related to exam performance ($b = -.70$, $SE = .29$, $t = -.28$, $p = .018$). The opposite was true for the number of correct solutions of the learning tasks ($b = .89$, $SE = .26$, $t = 3.44$, $p = .001$). This number was positively related to exam performance.

In the present context, engaging with digital self-study exercises was a relevant component of students' preparation for the exam. Both indicators were related to exam performance. The negative correlation of the number of repetitions with exam success suggests that this could be an indicator of low-quality learning with regard to constructive insight. Repeating the tasks until the feedback indicates that the solution was correct might not be a good learning strategy to prepare for the exam. This result might be taken into account when informing students about learning strategies with the digital system, and eventually as an online-measure that could be provided to students during learning. Such insights from log-data can help to improve actual learning processes.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 4: Systemdenken als Schlüsselkompetenz für Biologielehrkräfte: Wie ausgeprägt ist das Systemdenken bei Lehramtsstudierenden?

Till Schmäing, Moritz Krell

IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Das Systemdenken ist essentiell für das Erschließen von komplexen Zusammenhängen in der Ökologie und im Kontext von Bildung für nachhaltige Entwicklung (Fanta et al., 2020; Fanta et al., 2017; Schuler et al., 2018; Streiling et al., 2019). Systemdenken wird als eine von acht Schlüsselkompetenzen der Bildung für nachhaltige Entwicklung herausgestellt (UNESCO, 2017). Die Fähigkeit, Zusammenhänge in lebenden Systemen zu betrachten ist zentraler Bestandteil der biologischen Sachkompetenz in den Bildungsstandards des Faches Biologie für die Allgemeine Hochschulreife (KMK, 2020). Auch international wird der Stellenwert des Systemdenkens für das Verständnis biologischer Inhalte bekräftigt (Woodin et al., 2010). Für das Verständnis ökologischer Konzepte ist es notwendig, einzelne Merkmale eines Systems sowie ihre Zusammenhänge erkennen zu können (Mambrey et al., 2020). In der Forschungsliteratur wird das Systemdenken teilweise als *ecological thinking* bezeichnet (Riess & Mischo, 2010). Damit wird die Nähe zwischen dem theoretischen Ansatz und seiner Berücksichtigung in der biologiendidaktischen Forschung im Kontext Ökologie deutlich. Dementsprechend liegen viele Studien vor, welche zu ökologischen Inhalten Lehr- und Lernprozesse in der biologischen Bildung zum Systemdenken untersucht haben (z. B. Lankers et al., 2023; Mambrey et al., 2022; Nesimyan–Agadi & Ben-Zvi Assaraf, 2023).

Ein aktuelles Review belegt, dass die Anzahl an Studien zum Systemdenken in der MINT-Bildung seit 2016 stark gestiegen ist (Bielik, Delen et al., 2023). Trotz dieser Entwicklungen ist in der fachdidaktischen Forschung zum Systemdenken eine geringe Anzahl an Studien mit (angehenden) Lehrkräften vorhanden (Budak & Ceyhan, 2023; Yoon et al., 2018). Da (angehenden) Lehrkräften eine zentrale Verantwortung für die Förderung des Systemdenkens ihrer (zukünftigen) Lernenden zukommt, sind mehr Studien zum Systemdenken von Lehrkräften notwendig (Bielik, Jagemann et al., 2023).

Der vorliegende Beitrag stellt ein Projekt vor, welches sich diesem Forschungsdesiderat mit einer explorativen Querschnittsstudie annimmt. Dabei wird unter anderem der Fragestellung nachgegangen, wie ausgeprägt das Systemdenken von Lehramtsstudierenden des Unterrichtsfaches Biologie im Kontext Ökosystem Wald zu verschiedenen Zeitpunkten im Studium ist. Die Gelegenheitsstichprobe besteht aus Studierenden, welche biologiendidaktische Lehrveranstaltungen am an einer deutschen Universität belegen. Die Datenerhebungen werden sowohl in zwei Seminaren des Bachelors als auch in drei Seminaren des Masters erfolgen. Bei der Erhebung sollen die Befragten eine Concept-Map erstellen, die ihr Verständnis vom Ökosystem Wald beschreibt. Diese Aufgabenstellung wird in Anlehnung an Bielik, Jagemann et al. (2023) formuliert, welche dieses Vorgehen bereits erfolgreich mit Lehramtsstudierenden umgesetzt haben. Concept-Maps stellen ein visuelles Mittel zur Externalisierung von internen mentalen Modellen dar (Kinchin, 2020a, 2020b). Das Concept-Mapping ist eine etablierte Methode zur Analyse des Systemdenkens (Brandstädter et al., 2012). Die Erhebung wird mit der Software CoMapEd (Mühling, 2016) erfolgen.

Die Auswertung der Daten erfolgt vor dem Hintergrund des Struktur-Beziehungs-Funktions-Modells (Hmelo-Silver et al, 2007). Nach diesem sind Strukturen die Komponenten, aus denen das System besteht; Beziehungen sind die Mechanismen, die erklären, wie Strukturen in Beziehung stehen. Strukturen und Mechanismen führen zusammengenommen zu einer bestimmten Systemfunktion (Momsen et al., 2022). Die Datenanalyse wird sowohl qualitativ als auch quantitativ in Anlehnung an Nesimyan–Agadi & Ben-Zvi Assaraf (2023) durchgeführt. So werden in einem ersten Schritt mit einer induktiven Analyse die einzelnen Concept-Maps analysiert, um die jeweilige Systemkomplexität identifizieren und charakterisieren zu können. Dabei wird auf Basis der Beschriftung der Linien und Knotenpunkte ermittelt, inwiefern es sich jeweils um Strukturen, Beziehungen und Funktionen handelt. Nach dieser qualitativen Analyse werden die Concept-Maps hinsichtlich der Anzahl an Strukturen, Beziehungen und Funktionen betrachtet. Um einen Vergleich zwischen den Concept-Maps der Studierenden aus den unterschiedlichen Kursen vornehmen zu können, wird jeweils eine ANOVA durchgeführt.

Auf der GEBF soll mit einer Posterpräsentation ein Einblick in dieses Projekt aus der Biologiendidaktik gegeben werden. Ein Schwerpunkt wird neben einer Darstellung des Vorhabens besonders die Vorgehensweise bei der empirischen Untersuchung darstellen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 56: Können Nachbarschaften die sozialen Herkunftsunterschiede auf die Studienintention verstärken oder kompensieren?

Jacqueline Niemietz¹, Christoph Jindra¹, Melinda Erdmann², Marcel Helbig³

¹Institut zur Qualitätsentwicklung im Bildungswesen; ²Wissenschaftszentrum Berlin für Sozialforschung; ³Leibniz-Institut für Bildungsverläufe

Theorie & Fragestellung

Chancengleichheit setzt voraus, dass Bildungsentscheidungen unabhängig vom sozialen Hintergrund sind. Dennoch nehmen Studienberechtigte aus privilegierten Elternhäusern auch bei gleichen Noten häufiger ein Studium auf als solche aus weniger privilegierten Familien (Blossfeld et al., 2020; Lörz & Schindler, 2016; Quast et al., 2023; Schneider et al., 2021). Dieses Phänomen wird als sekundärer Herkunftseffekt bezeichnet (Boudon, 1974; Scharf et al., 2020; Schindler & Reimer, 2010; Watermann et al., 2014). Bisher wurde im deutschsprachigen Raum jedoch kaum untersucht, wie sozioökonomische Charakteristiken einer Nachbarschaft diese Herkunftseffekte beeinflussen, und das, obwohl aktuelle Forschungsergebnisse zeigen, dass sich die demographische und sozioökonomische Zusammensetzung von Nachbarschaften innerhalb deutscher Städte in den letzten Jahren zunehmend segregiert hat (Helbig, 2023; Helbig & Jähnen, 2018, 2019).

Dieser Beitrag untersucht erstmals mögliche Nachbarschaftseffekte auf die Studienintention von Oberstufenschüler:innen in Nordrhein-Westfalen. Auf Grundlage von bisherigen, vorwiegend internationalen Befunden, die darauf hindeuten, dass das Aufwachsen in einer privilegierten Nachbarschaft mit tendenziell besseren schulischen Kompetenzen (z. B. Gresch et al., 2023; Helbig, 2010; Wodtke et al., 2011) und einer höheren Rate an College-Besuchen (z. B. Chetty & Hendren, 2018; Chetty et al., 2016) verbunden ist, werden folgende Hypothesen aufgestellt:

1. Die Nachbarschaft beeinflusst die Studienintention der Oberstufenschüler:innen dahingehend, dass sie in ökonomisch schwächeren Nachbarschaften sinkt.
2. Die ohnehin geringere Studienintention von Oberstufenschüler:innen aus nichtakademischen Elternhäusern verringert sich zusätzlich in ökonomisch schwächeren Nachbarschaften.

Methode

Zur Untersuchung der Fragestellung wurden Daten der Studie „Zukunfts- und Berufspläne vor dem Abitur“ (ZuBAb, N=1766; Pietrzyk et al., 2019) genutzt. Die Nachbarschaft der Oberstufenschüler:innen wurden sozioökonomisch in Quintilen anhand der Kaufkraft kategorisiert (1. Quintil: geringe Kaufkraft bis 5. Quintil: hohe Kaufkraft). Die abhängige Variable (AV) ist die realistische Studienintention der Schüler:innen etwa 1,5 Jahre vor dem Abitur. Die Bildungsherkunft wird anhand des beruflichen Abschlusses operationalisiert und in zwei Kategorien unterteilt: Nichtakademiker:innenkinder (NAK) und Akademiker:innenkinder (AK). Als Kontrollvariablen fließen das Geschlecht, der Migrationshintergrund und die Schulform ein. Zudem wird die Deutschnote in Leistungspunkten berücksichtigt. Es werden Ergebnisse für 5 Mehrebenenmodelle, mit Schulebene als Level 2, präsentiert. Modell 1 (M1) berechnet den Einfluss der Bildungsherkunft, Modell 2 (M2) den Einfluss der Nachbarschaft auf die Studienintention. Modell 3 (M3) enthält beide Variablen, Modell 4 (M4) die Interaktion, und Modell 5 (M5) ergänzt die Deutschnote. In drei weiteren Modellen (M6-M8) werden separate Analysen basierend auf den Terzilen der Deutschleistungen durchgeführt.

Ergebnisse

Die Ergebnisse zeigen keinen Zusammenhang zwischen Studienintention und Nachbarschaft (M2). Die Bildungsherkunft ist ein signifikanter Prädiktor zu Ungunsten der NAK (M3). Mit der Interaktion (M4) sinkt die Studienintention der NAK im 5. Quintil der Kaufkraft im Vergleich zum 1. Quintil. NAK im 1. Quintil haben eine niedrigere Studienintention als AK im 1. Quintil. Diese Effekte verschwinden mit der Deutschnote (M5), die signifikant die Studienintention beeinflusst. In getrennten Modellen zeigen NAK mit guten Deutschleistungen eine höhere Studienintention im 3. Quintil (M6). NAK mit schlechten Leistungen haben eine niedrigere Studienintention in reichen Ortsteilen (M8).

Zusammengefasst konnten die Hypothesen nicht bestätigt werden. Die Studienintention sinkt nicht in ökonomisch schwächeren Nachbarschaften. Sie sinkt tendenziell bei NAK in Nachbarschaften mit hoher Kaufkraft, dieser Effekt verschwindet jedoch unter Kontrolle der Leistung und verstärkt sich für schlechte Schüler:innen. Die Studienintention der AK wird nicht beeinflusst.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 16: Berufliche Orientierung in bildungsbereichsübergreifender Verantwortung — Gelingensbedingungen aus der Perspektive institutioneller Akteure

Ariane Neu

FernUniversität in Hagen, Deutschland

Das Modellvorhaben „VerOnika“ zielt auf die Entwicklung und Erprobung verzahnter Orientierungsangebote zur beruflichen und akademischen Ausbildung. Diese Orientierungsangebote wenden sich an hochschulzugangsberechtigte Personen und sollen den Teilnehmenden authentische Einblicke in ein Hochschulstudium sowie in eine berufliche Ausbildung bieten, um so eine erfahrungsbasierte Bildungsentscheidung zu ermöglichen. Entwickelt und erprobt werden diese verzahnten Orientierungsangebote an den drei Standorten Berlin, Darmstadt und Karlsruhe jeweils in Zusammenarbeit von Hochschulen und Partnern der beruflichen Bildung. Die FernUniversität in Hagen hat im Rahmen dieses BMBF-geförderten Verbundvorhabens seit Dezember 2019 die wissenschaftliche Begleitung übernommen. Ziel der wissenschaftlichen Begleitung ist es, prozessbezogen die erfolgreiche Entwicklung und Erprobung der verzahnten Orientierungsangebote durch empirische Untersuchungen zu unterstützen und abschließend weitergehende Handlungsempfehlungen für die Entwicklung und Durchführung verzahnter Orientierungsangebote an der Schnittstelle zwischen beruflicher und akademischer Bildung abzuleiten. Als zentrale empirische Erhebungsmethoden kommen dazu u. a. Dokumentenanalysen, leitfadengestützte offene Expert*inneninterviews, problemzentrierte Gruppeninterviews mit Teilnehmenden der Orientierungsangebote sowie quantitative Online-Befragungen der Teilnehmenden zum Einsatz.

Der Beitrag stellt Befunde dieser wissenschaftlichen Begleitung zur Diskussion, die sich insbesondere auf die Perspektive der institutionellen Akteur*innen dieser verzahnten Orientierungsprogramme beziehen. So werden Ergebnisse aus neun leitfadengestützten offenen Expert*inneninterviews präsentiert, die in den Jahren 2021 und 2022 mit institutionellen Akteur*innen der verzahnten Orientierungsangebote des VerOnika-Verbundes geführt wurden bzw. mit auf strategischer und operativer Ebene an Entwicklung und Durchführung beteiligten Projektinitiator*innen und Programmkoordinator*innen. Im Mittelpunkt dieser Interviews steht die Frage, was aus Sicht der beteiligten Akteur*innen Gelingensbedingungen für die erfolgreiche gemeinsame Entwicklung und Durchführung verzahnter Orientierungsangebote sind. Denn vor dem Hintergrund der traditionell stark ausgeprägten institutionellen Trennung von höherer Allgemeinbildung und Berufsbildung (von Baethge (2006) als deutsches Bildungs-Schisma bezeichnet) stehen diese Akteur*innen vor der Herausforderung, auf ganz unterschiedlichen Ebenen die Logiken und Mechanismen der höheren Allgemeinbildung mit denen der beruflichen Bildung in Einklang zu bringen (vgl. Elsholz/Neu 2020). Für die Durchführung der Expert*inneninterviews wurden in methodischer Hinsicht die Ausführungen von Meuser/Nagel (1991) zugrunde gelegt. Inhaltlich war das Curriculum-Modell der Erziehungswissenschaftler Merckens/Strittmatter (1975) leitend, welches eine systematische Analyse sowohl des Entwicklungsprozesses als auch des entwickelten Produktes selbst, also des Bildungsformates ermöglicht. So gehen die Autoren in ihrem Modell auf die Wechselbeziehungen zwischen den Elementen eines Curriculums sowie zwischen diesem und externen Variablen ein, die Einfluss auf die Ausgestaltung von Lernzielen und -inhalten, die Lernorganisation sowie die Lernevaluation haben (vgl. Merckens/Strittmatter 1975). Ausgewertet wurden die geführten Interviews mittels qualitativer Inhaltsanalyse nach Mayring (2015); wobei eine Kombination aus deduktiven und induktiven Verfahren zum Einsatz kam. Leitend für die Auswertung war dabei insbesondere das Ziel, Handlungsempfehlungen abzuleiten, die sich nicht zuletzt an bildungspolitische Akteur*innen richten und förderlich für einen Transfer solcher verzahnten Orientierungsangebote auf andere Standorte sein können.

So ergibt sich aus den geführten Expert*inneninterviews u. a. die Handlungsempfehlung, bildungspolitisch Rahmenbedingungen zu schaffen, die es verzahnten Orientierungsprogrammen ermöglichen, den Teilnehmenden einen von der klassischen Dichotomie zwischen beruflicher und hochschulischer Bildung entkoppelten „neutralen“ Status zu verleihen (also weder Studierenden- noch Auszubildenden-Status). Dies würde die Gleichwertigkeit der beiden Bildungsbereiche im Rahmen der Kooperation noch mehr stärken und einen offeneren (d. h. nicht durch einen eventuell bereits vorhandenen Status beeinflussten) Orientierungsprozess der Teilnehmenden ermöglichen. Aufgrund der bislang vorliegenden rechtlichen Rahmenbedingungen ist ein solcher neutraler Status mit entsprechender sozialversicherungsrechtlicher Absicherung der Teilnehmenden noch nicht möglich. Stattdessen wurden die Orientierungsprogramme teilweise als vorbereitendes Studium oder als verlängerte Studieneingangsphase eines hochschulischen Studienganges verortet, um so für die Teilnehmenden einen Studierendenstatus und damit einen sozialversicherungsrechtlich abgesicherten Status sowie gegebenenfalls die Beantragung von BAföG zu ermöglichen. Neben diesem zentralen Punkt, stellt der Beitrag weitere Gelingensbedingungen sowie daraus abgeleitete Handlungsempfehlungen zur Diskussion.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 57: Interindividuelle Unterschiede in der Wahrnehmung schulischer Demokratiebildung von Schüler:innen

Annika Francke, Cornelia Gresch
Humboldt-Universität zu Berlin, Deutschland

Theoretischer Hintergrund

Studien weisen auf unterschiedliche Wahrnehmungen von Schüler:innen bezüglich des Zugangs von schulischem Demokratielernen, abhängig von familiären Hintergrundmerkmalen (Deimel, 2020; Hoskins & Janmaat, 2019). Angelehnt an die Sozialisationstheorie Bronfenbrenners (1981) und die Kapitaltheorie Bourdieus (1987) argumentieren wir, dass diese Unterschiede mit variierenden demokratischen Vorerfahrungen aufgrund der familiären ökonomischen, kulturellen und sozialen Lage zusammenhängen können. Empirisch stützt sich diese Annahme auf Studien aus Deutschland, die überwiegend positive Zusammenhänge zwischen dem kulturellen Kapital einer Familie und dem politischen Wissen von Schüler:innen konstatieren (z. B. Baykara-Krumme & Deimel, 2017; Goll et al., 2010; Greßer, 2016). Weiterhin zeigen Familien mit höherer Bildung häufiger demokratische Praktiken wie Diskussionen und die Einbeziehung der Kindermeinung (Alt et al., 2005; Oesterreich, 2002). Auch gibt es Hinweise, dass demokratiebildungsrelevante Merkmale je nach familiärem Zuwanderungsstatus variieren. So wurden Disparitäten im politischen Wissen und der Partizipationsbereitschaft zwischen Schüler:innen mit und ohne Zuwanderungshintergrund (ZHG) nachgewiesen, unabhängig vom familiären kulturellen Kapital (z. B. Baykara-Krumme, 2017; Goll et al., 2010; Greßer, 2016). Zudem ist denkbar, dass die Regierungsform des elterlichen Herkunftslandes die in der Familie gelebten demokratischen Werte und Praktiken beeinflusst. Müssig (2020) zeigte, dass Zugewanderte aus Demokratien seltener im Zielland politisch partizipieren als Zugewanderte aus Autokratien. Es bleibt zu untersuchen, ob der Demokratisierungsgrad des elterlichen Herkunftslandes darüber hinaus familieninterne Praktiken beeinflusst und so die Wahrnehmung Jugendlicher außerhalb der Familie prägt. Der vorliegende Beitrag intendiert, systematische Disparitäten im Zugang zur schulischen Demokratiebildung aufzudecken, damit diesen entgegengetreten werden kann, um Ungleichheiten in demokratischen Kompetenzen zu verringern.

Fragestellung

Wurden Schüler:innenwahrnehmungen zur schulischen Demokratiebildung für Deutschland bisher ausschließlich vor dem Hintergrund ihres sozioökonomischen Status fokussiert, erweitert der vorliegende Beitrag Analysen um weitere Merkmale. Es wird untersucht, wie die von Schüler:innen wahrgenommene schulische Demokratiebildung mit

- a) dem elterlichen Bildungshintergrund
- b) ihrem eigenen ZHG sowie
- c) dem Demokratisierungsgrad des elterlichen Herkunftslandes zusammenhängt.

Methode

Die Datengrundlage umfasst Angaben von 1.870 Siebtklässler:innen und deren Eltern aus der Studie „INSIDE – Inklusion in der Sekundarstufe I in Deutschland“ (Schmitt et al., 2020). Untersucht wurde die von Schüler:innen wahrgenommene schulische Demokratiebildung, erhoben mittels einer eigens entwickelten Skala, die angelehnt wurde an Abs et al. (2007) sowie eine IQB-Bildungstrend-Pilotierung. Schüler:innen bewerteten neun Aussagen auf einer Skala von 1 (trifft überhaupt nicht zu) bis 4 (trifft voll und ganz zu). Unabhängige Variablen waren der höchste Schulabschluss der Eltern und der Zuwanderungshintergrund des Kindes. Weiterhin wurde erhoben, ob ein Elternteil, beide oder keines aus einem demokratischen Herkunftsland stammt. Hierzu wurden Daten zum elterlichen Herkunftsland mit Daten des Democracy Index¹ der Economist Intelligence Unit verknüpft. Kontrollvariablen waren das Geschlecht der Schüler:innen und die Schulform. Es wurden multiple Regressionsanalysen unter Berücksichtigung der Mehrebenenstruktur durchgeführt, wobei fehlende Werte mit FIML imputiert wurden.

Ergebnisse

Die Analysen zeigen keine signifikanten Zusammenhänge zwischen der wahrgenommenen schulischen Demokratiebildung und den familiären Hintergrundmerkmalen von Schüler:innen unter Kontrolle von Geschlecht und Schulform (höchster elterlicher Bildungsabschluss Abitur: $b = .01$, $p = .80$; Zuwanderungshintergrund: $b = -.06$, $p = .29$; ein Elternteil nicht aus Demokratie: $b = .01$; $p = .87$; kein Elternteil aus Demokratie: $b = .06$, $p = .40$ (Referenzgruppe: beide Eltern aus Demokratie)). Bei den Kontrollvariablen zeigt sich, dass Schüler:innen an Gesamtschulen die schulische Demokratiebildung signifikant marginal ausgeprägter wahrnehmen als Schüler:innen anderer Schulformen ($b = .06$, $p = .03$).

Diskussion

Die Ergebnisse legen nahe, dass familiäre Vorerfahrungen wenig Einfluss auf die Wahrnehmung der Demokratiebildung in der Schule haben. Möglicherweise werden unterschiedliche Vorerfahrungen im Laufe der Zeit durch schulische Erfahrungen ausgeglichen. Zudem könnten Unterschiede in den Wahrnehmungen schulischer Demokratiebildung eher auf Unterrichtsfaktoren als auf familiär bedingte demokratische Vorerfahrungen zurückzuführen sein. Hinweise darauf, Unterschiede eher auf Schul- und Klassenebene zu untersuchen, gibt auch der Befund der signifikanten Schulformunterschiede.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 37: Lehrkräfteprofessionalisierung für die Förderung digitaler Souveränität im Bereich der Lesediagnostik und -förderung: eine Blended-Learning-Fortbildung zu LeOn (Leseraum online)

Janna Gutenberg¹, Charlotte Stehr², Daria Ferencik-Lehmkuhl², Michael Krelle¹, Jörg Jost²

¹Technische Universität Chemnitz; ²Universität zu Köln

Lesen spielt als Schlüsselkompetenz für die Persönlichkeitsentwicklung und die gesellschaftliche Teilhabe sowie als Voraussetzung für das Lernen in allen Fächern eine zentrale Rolle in der schulischen Bildung. Aktuelle nationale und internationale Schulleistungstudien weisen allerdings auf erhebliche Defizite im Bereich der Lesekompetenz von Schüler:innen der Primar- und Sekundarstufe hin (Stanat et al., 2022; McElvany, 2023; OECD, 2023). Mit Blick auf eine zunehmende Digitalisierung bergen digitale Werkzeuge vielversprechende Potenziale zur Leseförderung: Mit der webbasierten Anwendung LeOn (Leseraum Online) beispielsweise wird die Förderung der Leseflüssigkeit sowie des Leseverstehens auf hierarchiehöherer Ebene in den Blick genommen (Krelle et al., 2023).

Allerdings zeigen Befunde, dass weniger als ein Viertel der Lehrkräfte davon überzeugt ist, dass mithilfe digitaler Werkzeuge der Lernerfolg der Schüler:innen verbessert werden kann (Schmid et al., 2017, S. 6). Die internationale Studie ICILS zeigt etwa, dass digitale Werkzeuge von Lehrkräften in Deutschland vor allem zur frontalen Präsentation von Informationen (44,1 %) eingesetzt werden. Nur selten werden sie für individuelle Fördermaßnahmen genutzt (14,1 %) (Eickelmann et al., 2019, S. 18). Weiterhin äußert nur ein kleiner Anteil von Lehrkräften, an digitalisierungsbezogenen Fortbildungen teilgenommen zu haben (ebd., S. 16).

Damit der Einsatz digitaler Werkzeuge für die Leseförderung in der Schule etabliert werden kann, ist demnach nicht nur Wissen im Bereich der Lesedidaktik erforderlich, sondern auch eine ‚digitale Souveränität‘ aufseiten der Lehrkräfte. Hierunter „wird die Möglichkeit verstanden, digitale Medien selbstbestimmt und unter eigener Kontrolle zu nutzen und sich an die ständig wechselnden Anforderungen in einer digitalisierten Welt anzupassen“ (Blossfeld et al., 2018, S. 12).

An dieser Stelle setzt das Verbundprojekt „Digitale Souveränität als Ziel wegweisender Lehrer:innenbildung für Sprachen, Gesellschafts- und Wirtschaftswissenschaften in der digitalen Welt“ (DiSo-SGW) an. Es zielt auf die (Weiter-)Entwicklung, Evaluation und Implementation von Fortbildungsmodulen zur Vertiefung digitaler Kompetenzen und Überzeugungen von Lehrkräften. In der Projektgruppe ‚Sprachförderung und ihre Diagnostik‘ wird im Teilprojekt Lesen eine Blended-Learning-Fortbildung zur digitalen Souveränität von Lehrkräften der Primar- und Sekundarstufe I im Bereich der Lesediagnostik und -förderung entwickelt. In der Fortbildung steht der kompetente Einsatz des in Nordrhein-Westfalen bereits etablierten digitalen Werkzeugs LeOn im Vordergrund. Dabei handelt es sich um eine webbasierte Anwendung zur Leseförderung, die von der zweiten bis zur sechsten Klasse eingesetzt werden kann. Die Fortbildung setzt sich aus drei E-Learning-Modulen und Online-Workshops zusammen: Während in Modul A gängige Ansätze zur Lesediagnostik und -förderung thematisiert werden, widmet sich Modul B der didaktischen Umsetzung dieser Ansätze innerhalb von LeOn. Die Fortbildungsinhalte in Modul A und B werden in Form des simulationsbasierten Lernens (Chernikova et al., 2019) eingeübt, bevor Modul C die Lehrkräfte schließlich bei der eigenen Arbeit mit LeOn begleitet.

Die Erprobung der Fortbildung startet im Herbst 2024 mit rund 100 Lehrkräften aus Nordrhein-Westfalen und wird wissenschaftlich begleitet. Ziel der empirischen Studie ist, die Wirksamkeit der Fortbildung im dargelegten Themenspektrum zu evaluieren. Die Evaluationsstudie findet im Prä-Post-Follow-Up-Design statt, wobei in einem Fragebogen allgemeine Fragen zur Wirksamkeit der Fortbildung (Lipowsky & Rzejak, 2019) mit fachspezifischen und -didaktischen Fragestellungen verbunden werden.

Das Poster präsentiert das Fortbildungskonzept sowie das Design zur Begleitforschung im Rahmen der Erprobung der Fortbildung. Zudem werden erste Erkenntnisse aus der Erprobung referiert, u. a. zum selbsteingeschätzten Wissen in der Lesedidaktik und zu digitalitätsbezogenen Kompetenzen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 58: Einfluss sozialer Unterstützungsstrukturen auf Abbruchgedanken im Referendariat in Bayern: Eine quantitative Analyse der Rolle von Mentoren

Bianca Burkert, Maïke Reimer

Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (IHF), Deutschland

Relevanz

Die Kultusministerkonferenz Deutschland prognostiziert bis 2035 ein Defizit von 68.000 Lehrkräften (Kultusministerkonferenz, 2023), insbesondere in den MINT-Bereichen des Sekundarbereichs II oder für das Gymnasium (Anger, C., Betz, J., Geis-Thöne, W. & Plünnecke, 2023; Kultusministerkonferenz, 2023). Ein bedeutsamer Schwund entsteht während des Vorbereitungsdienstes/Referendariats, während dessen viele angehende Lehrkräfte Abbruchgedanken mit sich tragen (Maué, Goller, Bonnes & Käner, 2024) und letztendlich 30 % tatsächlich auch abbrechen, wie sich aus Gülen, Müller und Schmid-Kühn (2023) ergibt. Daher bietet die Verringerung von Abbruchintentionen und tatsächlichen Abbrüchen einen wichtigen Ansatzpunkt, um dem Lehrkräftemangel entgegenzuwirken (Klemm & Zorn, 2018).

Theoretischer Hintergrund

Ein erheblicher Faktor für Abbruchgedanken und Abbrüche ist die hohe Belastung während des Referendariats (Dicke & Waldeyer, 2020; Peitz, 2022). Anhaltende hohe Belastung erhöht das Risiko für emotionale Erschöpfung (Dicke & Waldeyer, 2020), was die erste Phase des Burnouts darstellt (Maslach & Leiter, 1999). Dies geht einher mit geringerer Arbeitszufriedenheit und geringerem beruflichen Engagement (Admiraal, 2022; Hakanen, Bakker & Schaufeli, 2006) und kann zu erhöhten Abbruchintentionen beitragen (Skaalvik & Skaalvik, 2011, 2016; Struyven & Vanthournout, 2014), die wiederum ein Indikator für tatsächlichen Abbruch sind (Bean, 1982; Gold, 1988).

Dies ist nicht nur auf Realitätsschocks und herausfordernde Arbeitsbedingungen zurückzuführen, sondern auch auf unzureichende Unterstützungsstrukturen während des Referendariats (Huber & Lusnig, 2022). Hingegen erhöht eine professionelle Ausbildung im Referendariat die Wahrscheinlichkeit, dem Lehrerberuf treu zu bleiben (Goldhaber, Krieg, Theobald & Goggins, 2022). Insbesondere geht ein qualitativ hochwertiges und engagiertes Mentorat, das emotionale und informationelle Unterstützung bietet, mit geringerer Erschöpfung und höherer Berufszufriedenheit einher (Richter, Kunter, Lüdtkke, Klusmann & Baumert, 2011; Richter et al., 2013) und hat daher ein hohes Potential, die aktuellen hohen Abbruchraten zu senken (Berliner, 2004). Allerdings ist noch nicht ausreichend erforscht, ob bzw. unter welchen Bedingungen soziale Unterstützungsstrukturen, insbesondere das Mentorat, bei Abbruchgedanken unterstützend wirken – auch weil vorliegende Studien nicht auf die vielfältigen Unterschiede in der Ausgestaltung des Referendariats/Vorbereitungsdienstes zwischen den Bundesländern eingehen können.

Fragestellung

Die Studie konzentriert sich auf Referendarinnen und Referendare eines einzigen Bundeslandes (Bayern) und untersucht folgende spezifische Fragestellungen:

1. Wie häufig treten Abbruchgedanken auf?
2. Werden Mentorierende bei Abbruchgedanken als Unterstützungspersonen herangezogen, und welche Faktoren beeinflussen die Häufigkeit der Nutzung?
3. Welche anderen sozialen Unterstützungsstrukturen stehen den angehenden Lehrkräften zur Verfügung?
4. Können soziale Unterstützungsstrukturen, insbesondere durch das Mentorat, der Realisierung von Abbruchgedanken entgegenwirken?

Methode

Informationen zu Abbruchgedanken und sozialer Unterstützung, inklusive Mentorat, werden im Rahmen der neunten und zehnten Bayerischen Absolventenstudie (BAS) zwischen 2023 und 2025 erhoben, einer Befragung von Absolventinnen und Absolventen von Lehramtsstudiengängen an vier Universitäten ca. 1 Jahr nach Abschluss, und mittels multivariater Regressionsanalysen ausgewertet. Eine Follow-Up-Untersuchung im Abstand von ca. einem Jahr gibt Anhaltspunkte darüber, inwieweit Abbruchgedanken persistieren, sich abschwächen oder umgesetzt werden, und wie dies mit der Unterstützung im Mentorat zusammenhängt.

Ergebnisse und ihre Bedeutung

Die vorläufigen[1] Ergebnisse zeigen, dass etwa 64% der Referendarinnen und Referendare Abbruchgedanken hegen, wobei Belastungsfaktoren wie Eigenanspruch und Fremdbewertung häufig genannte Gründe sind. Die Verfügbarkeit von unterschiedlichen Formen der Unterstützung (informationell, instrumentell, emotional) üben keinen signifikanten Einfluss auf die Häufigkeit der Abbruchgedanken aus. Etwa 25% derjenigen mit Abbruchgedanken haben mit Mentorierenden darüber gesprochen, und die Häufigkeit der Abbruchgedanken ist hierfür ein signifikanter Prädiktor.

Dies bestätigt die hohe Belastung während des Referendariats und die starke Verbreitung von Abbruchgedanken sowie die Bedeutung eines engagierten Mentorats und einer hochwertigen Mentoratsbeziehung.

[1]Aktuell sind nicht alle Erhebungen beendet; insbesondere liegen die Analysen der Follow-Up-Studie noch nicht vor. Dies wird bis Februar 2025 voraussichtlich der Fall sein.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 38: Geschlechterstereotype bei Kindergartenkindern

Johanna Landvoigt, Jan Lenhart

Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Geschlechterstereotypes Wissen kann schon bei Kindern ab zwei Jahren beobachtet werden (Martin und Ruble, 2010). Es zeigen sich Auswirkungen auf unterschiedliche Lebens- bzw. Entwicklungsbereiche wie das Aussehen, Intelligenz/ Brillanz, Berufe oder auch die Vorliebe für Spielzeuge (u.a. Bian, Leslie & Cimpian, 2017; Halim et al., 2024; Solbes-Canales, Valverde-Montesino & Herranz-Hernández, 2020, Todd et al., 2018). Hinsichtlich der Entwicklung wird im Sinne kognitiver Theorien ein sequentieller Verlauf angenommen, wobei es im frühen Kindesalter zu einer zunehmenden Rigidität kommt, mit höchster Ausprägung im Alter von 5 bis 6 Jahren, bis diese erneut flexibilisiert wird (Trautner et al., 2005). Es zeigt sich eine große Heterogenität der bisherigen Forschung bezüglich der Frage, ob Mädchen und Jungen sich hinsichtlich der Entwicklung der Stereotype unterscheiden. Während einige Untersuchungen Unterschiede in der Anzahl von Geschlechterstereotypen finden (z.B. Baker, Tisak & Tisak, 2016; Kanka, Spiel, Schober & Spiel, 2014; Signorella et al., 1993), können diese in anderen Untersuchungen nicht gezeigt werden (z.B. Leinbach, Hort & Fagot, 1997; Solbes-Canales et al., 2020). Neben quantitativen Unterschieden müssen jedoch auch qualitative Unterschiede berücksichtigt werden (Miller, Wheeler & Woods, 2024).

Fragestellung

In der vorliegenden Studie wurde untersucht, ob Kindergartenkinder stereotype Zuordnungen kongruent zu denen Erwachsener vornehmen. Weiterhin wurde geprüft, ob mit zunehmendem Alter diese stereotypen Zuordnungen der Kinder zunehmen. Schließlich wurde untersucht, ob sich die Geschlechterstereotype sowie deren Entwicklung zwischen Mädchen und Jungen unterscheiden.

Methode

Die Stichprobe umfasste 66 Kindergartenkinder (38 Mädchen und 28 Jungen) im Alter von 3 bis 6 Jahren ($M = 63$ Monate, $SD = 11.58$) aus Deutschland und Österreich. Es wurden insgesamt 56 Wörter präsentiert, welche zuvor in einer Erwachsenenstichprobe als „typisch männlich“, „typisch weiblich“ oder „neutral“ eingeschätzt worden waren. Die Kinder sollten diese Items auf einer 5-stufigen Skala (von „nur Männer“ bis „nur Frauen“) einschätzen.

Ergebnisse und ihre Bedeutung

Zunächst wurde die Übereinstimmung der Einschätzungen der Kinder bezüglich der Stereotype mit den Erwachsenenratings untersucht. Es zeigte sich, dass bei 84% der Wörter die Ratings der Kinder signifikant mit denen der Erwachsenen übereinstimmten (davon 90% „richtig“ weiblich, 88 % „richtig“ männlich und 60% „richtig“ neutral). Es ließen sich Geschlechtsunterschiede auf Itemebene für einzelne Wörter finden, teilweise zeigte sich eine stärker dem Stereotyp entsprechende Zuschreibung einzelner Begriffe durch Jungen (umbringen, kräftig, marschieren), teilweise durch Mädchen (kochen, brav, wunderschön).

Es ergab sich weder für „typisch weibliche“ noch „typische männliche“ Begriffe eine statistisch signifikante Zunahme an stereotyper Bewertung der Begriffe mit zunehmendem Alter. Bei separater Analyse von Jungen und Mädchen zeigte sich ein signifikanter Einfluss des Alters auf kongruente stereotype Zuordnungen bei Jungen (Jungen: $\beta = .03$, $t = 2.4$, $p = .024$; Mädchen: $\beta = .001$, $t = 0.21$, $p = .834$).

Während in Übereinstimmung mit den Hypothesen gezeigt werden konnte, dass Kindergartenkinder geschlechterstereotype Zuordnungen analog zu denen Erwachsener vornahmen, konnte in dieser Stichprobe kein allgemeiner Alterstrend bezüglich einer zunehmend verstärkten stereotypen Bewertung beobachtet werden. Während bei Jungen eine Zunahme stereotyper Bewertungen mit höherem Alter bestand, konnte dies für Mädchen nicht gezeigt werden. Auf Itemebene ergaben sich Unterschiede zwischen Jungen und Mädchen, die darauf hinweisen, dass Stereotype bezüglich des eigenen Geschlechts stärker verinnerlicht werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 59: „So ne Malochermentalität, das hilft schon“: Potenziale des Aufwachsens in nicht akademischen Elternhäusern für die Bewältigung der Anforderungen der Promotion aus Sicht von Bildungsaufsteiger:innen

Stephan Otto

IU International University, Deutschland

Theoretischer Hintergrund & Forschungsstand

Der Zusammenhang von sozialer Herkunft und Bildungserfolg ist trotz zahlreicher Reformmaßnahmen weiterhin ein zentraler Problembereich des deutschen Bildungssystems (Autor:innengruppe Bildungsberichterstattung 2022). Betrachtet man hierzu Bildungsinstitutionen im Verlauf, zeigt sich eine sogenannte *Leaky Pipeline*: Der Anteil von Menschen aus niedrigeren sozialen Schichten nimmt auf höheren Qualifizierungsstufen immer weiter ab. Besonders augenscheinlich wird dies bei der Promotion: Während von 100 Akademiker:innenkinder 10 eine Promotion erfolgreich abschließen, ist dies bei 100 Nicht-Akademikerkindern lediglich bei einer Person der Fall (Kracke, Middendorff & Buck 2018).

Eine Begründung dafür, dass soziale Herkunft und Bildungserfolg eng miteinander verbunden sind und dass Bildungssysteme nicht gleiche Chancen für alle bieten, sondern soziale Ungleichheiten sogar verstärken, liefert das Habituskonzept Pierre Bourdieus. Laut diesem prägt die Herkunftsfamilie den Habitus einer Person, also die jeweiligen Denk- und Handlungsmuster sowie das Auftreten (Bourdieu 1987). Die familiären Grundmuster des Habitus gelten als relativ stabil, wobei zwar eine Modifizierung grundsätzlich möglich ist, allerdings werden elterliche Grundmuster des Habitus üblicherweise nicht komplett aufgegeben (Vester 2015). Diese habituellen Prägungen sind insbesondere im Kontext der Promotion besonders bedeutsam, da die individuelle Leistung hier nur ein notwendiger Erfolgsfaktor ist. Vielmehr zeigt sich, dass als bildungsbürgerlich definierte Kapitalien wie Sprachästhetik, Auslandsaufenthalte oder ehrenamtliches Engagement den erfolgreichen Abschluss der Promotion erleichtern (Möller et al. 2020). Bildungsaufsteiger:innen dagegen können im System Hochschule Fremdheitserfahrungen machen, da ebensolche Kapitalien und Umgangsformen des Systems zunächst erlernt werden müssen (Alheit 2005). Sie müssen also ihren ursprünglichen Habitus zunächst modifizieren, um in privilegierte Positionen zu gelangen. Laut El-Mafaalani (2012) gibt es dafür zwei Ansätze: Einerseits können sie sich bewusst von ihrem Herkunftsmilieu distanzieren (reflexive Opposition), andererseits können sie kreative Wege und Strategien entwickeln, um ihre Herkunft und die neuen Lebensbedingungen zu verbinden, wobei sie zu beiden Milieus eine gewisse Distanz wahren (empraktische Synthese).

In diesem Zusammenhang sind Studien zu Bildungsaufsteiger:innen aufschlussreich, die vom „Mythos der Fremdheit der Bildungsaufsteiger_innen“ (Miethe 2017, 686) sprechen und Beispiele anführen, dass gerade die Nutzung von als unüblich im akademischen Kontext bewerteten Verhaltensweisen („auf den Tisch hauen“ ebd.) vorteilhaft für die eigene Karriere genutzt werden können.

Bislang vorliegende Forschungsarbeiten befassen sich allerdings dennoch zumeist mit der Frage, wie diese Personen es *trotz* ihrer Herkunft bis zur Promotion geschafft haben.

In diesem Beitrag soll allerdings ein anderer Weg gewählt werden, um zu ermitteln, warum aus Akteurssicht *gerade* die Herkunft aus nicht akademischen Elternhäusern dazu beigetragen hat, eine Promotion erfolgreich abzuschließen.

Fragestellung

Für den Beitrag sind die folgende Oberfrage und die daran anschließenden Unterfragen zentral:

Inwieweit erleben Promovierte aus nichtakademischen Elternhäusern ihre familiäre Sozialisation als hilfreich für die Anforderungen der Promotion?

- Inwieweit erlebten die Promovierten Unterstützung durch ihre Familie während der Promotion?
- Inwieweit erleben sich die Promovierten als beeinflusst durch das vom Elternhaus geprägte Verhältnis gegenüber Arbeit?
- Welche Unterschiede nehmen die Promovierten gegenüber Doktorand:innen aus akademischen Elternhäusern wahr?

Methode & Ergebnisse

In dieser qualitativen Studie wurden acht Leitfadeninterviews mit promovierten Personen aus einem nicht akademischen Elternhaus mittels der Habitushermeneutik (Bremer & Teiwes-Kügler 2014) ausgewertet.

Die ersten Befunde dieser Studie verweisen darauf, dass aus der Perspektive der Untersuchungsteilnehmer:innen ihre spezifischen familiären Prägungen (Verhältnis zu Arbeit, ihnen bei der Bewältigung der Anforderung an eine Promotion geholfen haben. Des Weiteren leiten sie spezifische Fähigkeiten und Kompetenzen explizit aus der eigenen, nicht akademisch geprägten Herkunft ab (Resilienz, bessere Fokussierung), die sogar als Vorteile gegenüber Promovierenden aus akademischen Elternhäusern wahrgenommen werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 1: Inklusiver naturwissenschaftlicher Unterricht durch forschendes Lernen: Erkenntnisse aus einem systematischen Literaturreview

Leonie Willmes, Helena van Vorst, Mathias Ropohl

Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Seit der Ratifizierung der UN Behindertenrechtskonvention im Jahr 2006 steht das inklusive Lernen und damit die Individualisierung von Lehr-/Lernprozessen in heterogenen Lerngruppen im Fokus der Forschung. Für eine ganzheitliche Umsetzung dieser Anforderung im Bereich des experimentellen naturwissenschaftlichen Unterrichts bietet sich das Konzept des forschenden Lernens an, das als schülerzentriertes, hypothesengeleitetes Vorgehen bei der Bearbeitung einer naturwissenschaftlichen Problemstellung (Fischer et al., 2018) beschrieben wird. Durch eine Variation von Strukturierung und Offenheit kann dieser Ansatz zur Individualisierung von Lehr-/Lernprozessen in heterogenen Lerngruppen beitragen (z.B. Blanchard, 2010). Dabei durchlaufen die Schüler:innen verschiedene Aktivitäten forschenden Lernens, z.B. das Planen und Durchführen von Untersuchungen (u.a. Rönnebeck et al., 2016). Im inklusiven naturwissenschaftlichen Unterricht unterliegt das forschende Lernen jedoch bestimmten Herausforderungen (z.B. Wahrnehmungseinschränkungen berücksichtigen). Die Entwicklung von Kriterien, wie forschendes Lernen für Schüler:innen mit bestimmten Bedarfen im Rahmen eines inklusiven naturwissenschaftlichen Unterrichts umgesetzt werden kann, ist bisher im deutschsprachigen Raum wenig systematisch erfolgt (Menthe & Hoffmann, 2015).

Vor diesem Hintergrund ist das Ziel des Vorhabens, den gegenwärtigen internationalen Forschungsstand zum forschenden Lernen im inklusiven naturwissenschaftlichen Unterricht zu synthetisieren und hinsichtlich der Umsetzung in Bezug auf einzelne Förderbedarfe zu analysieren.

Fragestellung

FF: Welche methodischen und inhaltlichen Kriterien zur Gestaltung von (experimentellen) Lerngelegenheiten können Barrieren im Unterricht abschwächen?

Methode

Zur Beantwortung der Forschungsfrage wurde ein systematisches Literaturreview durchgeführt. Die Vorgehensweise orientiert sich an den von Torgerson et al. (2017) definierten Schritten und dem PRISMA Schema (Page et al., 2021). Neben dem interessierenden Konstrukt und sich daraus ergebenden Suchstrategien, die in den Datenbanken Web of Science, ERIC und FIS Bildung sowie relevanten Journals Anwendung fanden, wurden Einschlusskriterien definiert, nach denen die Titel, Abstracts und Volltexte der Publikationen sukzessive ausgewertet wurden. Insgesamt wurden $N = 34$ Publikationen in das Review eingeschlossen. Da im Zuge der Ratifizierung der UN-Behindertenrechtskonvention im Jahr 2006 neue Forschung initiiert wurde, wurden die Datenbanken und Journals für den Zeitraum von 2009 bis 2022 durchsucht. Nur englischsprachige und peer-reviewte Publikationen mit einem Bezug zu inklusivem naturwissenschaftlichem Unterricht und forschendem Lernen wurden berücksichtigt.

Ergebnisse

Ein zentrales Ergebnis des Reviews ist die Gegenüberstellung der Aktivitäten forschenden Lernens nach Rönnebeck et al. (2016) mit den in den Publikationen angesprochenen Förderbedarfen in einer Heat Map, wodurch Forschungslücken identifiziert werden können. Auf deskriptiver Ebene zeigt sich, dass die Publikationen mehrheitlich aus den USA stammen und sich auf den Förderbedarf „Sehen“ beziehen. Es zeigt sich außerdem, dass forschendes Lernen in verschiedenen Publikationen als lernförderlich für Schüler:innen unterschiedlicher Förderbedarfe beschrieben wird. Insbesondere ein strukturiertes forschendes Lernen erweist sich als gewinnbringend. Vor allem die kleinschrittige Strukturierung der Lerngelegenheiten (z.B. durch Graphic Organisers) sowie Visualisierungen (z.B. Icons und Bilder) und sprachliche Anpassungen (z.B. Glossare) werden als relevante Gelingensbedingungen für inklusive Lerngelegenheiten im forschenden Lernen beschrieben. Es zeigt sich aber auch, dass die in den Publikationen beschriebenen Kriterien oft unspezifisch sind und als erweiterte Kriterien guten Unterrichts betrachtet werden können. Zum Beispiel fällt insbesondere im Bereich des Hypothesenaufstellens auf, dass für Schüler:innen mit Hörbeeinträchtigungen bisher nahezu keine Forschung existiert.

Ausblick

Um die praktische Relevanz der im Review identifizierten Kriterien zu beleuchten, wurden die Ergebnisse des Reviews durch $N = 7$ Expert:inneninterviews ergänzt und mittels qualitativer Inhaltsanalyse (Mayring, 2022) ausgewertet. Einige der im Review identifizierten Kriterien konnten bestätigt werden (z.B. kleinschrittige Strukturierung oder Visualisierung mittels digitaler Medien), aber auch einige neue Kriterien herausgearbeitet werden (z.B. die Arbeit in Kleingruppen nach systematisch zugeordneten Rollen je nach Fähigkeiten der Schüler:innen). Im Anschluss werden die besonders empfohlenen Herangehensweisen im Bereich der Visualisierung (Fotos vom Material und Schritt-für-Schritt-Videos) aus Review und Interviews genutzt, um experimentelle Lerngelegenheiten des strukturierenden forschenden Lernens für Schüler:innen der Förderbedarfe Lernen und emotionale-soziale Entwicklung zu entwickeln und ihre Wirksamkeit in einer videogestützten Querschnittstudie unterrichtspraktisch zu evaluieren.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 60: Erfassung professioneller Wahrnehmung zu naturwissenschaftlichen Denk- und Arbeitsweisen im Bereich Früher Bildung

Eva Bühler¹, Markus Rehm¹, Markus Wilhelm², Hendrik Lohse-Bossenz³

¹PH Heidelberg, Deutschland; ²PH Luzern, Schweiz; ³Uni Greifswald, Deutschland

Theoretischer Hintergrund: Die Ergebnisse der aktuellen PISA-Studie zeigen wiederholt, dass Schüler_innen im Vergleich zu anderen Ländern in den deutschen Schulen in den naturwissenschaftlichen Fächern niedrigere Punktwerte erzielen (Lewalter, Diedrich, Goldhammer, Köller & Reiss, 2023). Darüber hinaus wird zunehmend über einen Mangel von Fachkräften in naturwissenschaftlichen und technischen Bereichen berichtet (acatech MINT Nachwuchsbarometer, 2023). Als Antwort darauf wurden in den letzten Jahren Bildungsreformen eingeleitet, so wurden beispielsweise in allen Bundesländern Bildungspläne für den Elementarbereich eingeführt (Steffensky, 2017; Fölling-Albers, 2013). Diese beinhalten neben der Förderung sprachlicher, sozialer und emotionaler Kompetenzen auch die Bereiche der frühen mathematischen und der frühen naturwissenschaftlichen Bildung (OP BW, 2011).

Kinder zeigen früh Interesse an ihrer Umgebung und beschäftigen sich mit Phänomenen der belebten und unbelebten Natur (Koerber, Sodian & Osterhaus, 2022). Neugieriges Fragen und Staunen sind jedoch noch nicht hinreichend für naturwissenschaftliche Bildungsprozesse. Die Ausrichtung einer kindgerechten Beschäftigung an (natur-)wissenschaftlichen Denk- und Arbeitsweisen kann naturwissenschaftliche Bildungsprozesse adressieren (Leuchter, 2017). Ausgangspunkte sind dabei die epistemische Neugier (van Schijndel, Jansen & Raijmakers, 2018; Jirout, 2020) und Prinzipien des forschenden Lernens (Aufenanger, 2019).

Lernarrangements in der frühen naturwissenschaftlichen Bildung basieren oft auf Alltagssituationen, in denen Kinder zwar Denk- und Arbeitsweisen erleben (Steffensky et al., 2018), entscheidend für die Förderung naturwissenschaftlichen Bildungsprozesse sind jedoch professionell gestaltete Lerngelegenheiten, die Kinder zur Nutzung einzelner Denk- und Arbeitsweisen wie dem Beobachten, dem Ordnen und dem Fragenstellen auffordern und ermutigen (Sodian & Mayer, 2013; Kuhn, 2010; Eshach & Fried, 2005). Fachkräfte benötigen für die Bereitstellung qualitativ hochwertiger Lernangebote ein breites Handlungsrepertoire (Steffensky, 2017) und Fähigkeiten der professionellen Wahrnehmung.

Fragestellung und Methode: Ziel dieser Studie ist die Entwicklung und Validierung eines Tests zur Erfassung der professionellen Wahrnehmung von pädagogischen Fachkräften im Bereich der naturwissenschaftlichen Denk- und Arbeitsweisen. Zur kontextualisierten und situativen Erfassung ebendieses haben sich in der Lehrkräftebildung sogenannte Vignettentests – in Text- und Comicform – bereits bewährt (vgl. Rehm & Bölsterli, 2014; Rutsch, 2016; Friesen, 2017; Syring et al., 2015). Der Test umfasst acht Vignetten zu vier verschiedenen Denk- und Arbeitsweisen mit jeweils parallelen Text- und Comic-Versionen (=16 Vignetten). Eine Musterlösung wurde theoriegestützt erstellt und durch Expert_innen früher naturwissenschaftlicher Bildung validiert. Die Testwertberechnung wurde gemäß Rutsch, Vogel, Rehm & Dörfler (2018) und Billion-Kramer, Lohse-Bossenz, Dörfler & Rehm (2020) auf der Itemebene durchgeführt (partial-credit-model). Die Testwerte lagen zwischen 0 und 1, höhere Werte zeigten größere Übereinstimmung mit der theoretischen Norm. Die Vignetten wurden zwischen Herbst 2022 und Frühjahr 2024 an mehreren Stichproben (N1 = 59, N2 = 55 und N3 = 36; Alter: M = 39.05 Jahre, Berufserfahrung: M = 12.10 Jahre, belegte Fortbildungsreihen: M = 1.94) auf Praxisnähe und Authentizität geprüft und in drei Durchgängen pilotiert.

Ergebnisse: Die Fachkräfte erzielten über alle Stichproben hinweg vergleichbare Mittelwerte bei ausreichender Streuung der Testwerte. Den höchsten Mittelwert erzielten die pädagogischen Fachkräfte bei der ersten Vignette (Comic 1 = .61, SD = .18; Text 1 = .62, SD = .18), die kleinsten Lösungshäufigkeiten werden bei Vignette 7 (Comic 7 = .28, SD = .16; Text 7 = .20, SD = .14) gemessen. Zur Anpassung der Aufgabenschwierigkeit und zur Vermeidung von Bodeneffekten wurden für die Haupterhebung einzelne Formulierungen der Items überarbeitet.

Ausblick: In der Hauptstudie (Erhebungszeitraum Juli-September) soll die psychometrische Struktur des Tests validiert und erste Zusammenhänge zu externen Merkmalen identifiziert werden. Wir erwarten, dass die explorative Faktorenanalyse (EFA) eine klar strukturierte Faktorstruktur aufdeckt, die den theoretischen Dimensionen entspricht. Strukturgleichungsmodelle (SEM) sollen gute Fit-Indizes (z.B. CFI, RMSEA) zeigen und die theoretischen Beziehungen bestätigen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 12: Untersuchung der Qualität selbstberichteten Umweltverhaltens von Münsteraner Jugendlichen zur Evaluation einer BNE-Zooführung

Gesche Barg, Elmar Souvignier
Universität Münster, Deutschland

Theoretischer Hintergrund

Bildung für nachhaltige Entwicklung (BNE) zielt darauf ab Menschen zu befähigen, im Sinne einer nachhaltigen Entwicklung zu handeln. Für verschiedene BNE-Bildungsangebote konnte ein positiver Einfluss auf die Wissen, Fähigkeiten, Einstellungen und moralischen Überzeugungen in Bezug auf Umweltschutz beobachtet werden (O'Flaherty & Liddy, 2018). Die Wirksamkeit in Bezug auf Verhaltensänderungen wurde deutlich weniger erforscht (Algurén, 2021). Einzelne Studien weisen jedoch einen kleinen positiven Einfluss von Umweltbildungsprogrammen auf umweltfreundliches Verhalten nach (van de Wetering et al., 2022). Vor diesem Hintergrund zielt eine BNE-Zooführung des Allwetterzoos Münster darauf ab, bei den Teilnehmenden durch direkte Anschauung von Tieren und Ökosystemen, sowie durch Diskussionen über die ökologischen, ökonomischen und sozialen Auswirkungen menschlichen Handelns Verhaltensänderungen im Sinne umweltfreundlicherer Verhaltensweisen zu fördern. Die entsprechende Wirksamkeit dieser BNE-Zooführung wurde in der vorliegenden Studie mit einer Stichprobe von Schüler:innen der 7. bis 9. Klasse als Teil eines größeren Evaluationsprojektes untersucht. Im Unterschied zu den üblicherweise eingesetzten geschlossenen Antwortformaten wurden in dieser Untersuchung offen formulierte Aussagen der Jugendlichen zu ihrem umweltbezogenen Verhalten herangezogen, um authentische Selbstauskünfte zu analysieren. Auf diese Weise kann zudem ausgewertet werden, ob die Qualität der genannten Verhaltensweisen steigt, denn Jugendliche schätzen auch Verhaltensweisen als umweltschützend ein, die geringe ökologische Auswirkungen haben und im eigentlichen Sinne nicht als „umweltschützend“ gelten (Cologna et al., 2022; Nielsen et al., 2021; Wynes & Nicholas, 2017).

Angenommen wird, dass die Vermittlung von Umweltwissen in einer BNE-Zooführung dazu führt, dass Jugendliche sich intensiver mit den ökologischen Auswirkungen ihrer Verhaltensweisen auseinandersetzen und sich dadurch umweltschützender verhalten (Díaz-Sieffer et al., 2015; Otto & Pensini, 2017; Roczen et al., 2014). Vorliegende Studien (Barth & Schmitz, 2020; Meitinger et al., 2021) weisen dabei darauf hin, dass eine höhere Qualität und Ausdifferenzierung offener Antworten in Selbstauskünften (Antwortlänge, eindeutige Interpretierbarkeit, Anzahl genannter Themen) als Indikatoren umweltschätzenden Verhaltens gelten können.

Fragestellung

Diese Studie untersucht, ob sich das umweltfreundliche Verhalten von Jugendlichen durch die Teilnahme an einer BNE-Zooführung des Allwetterzoos Münster verändert. Es wird erwartet, dass Jugendliche nach der Teilnahme an einer BNE-Zooführung kurz- und langfristig eine Zunahme an umweltschützendem Verhalten, operationalisiert über die Qualität offener Antworten (hinsichtlich Wortanzahl, Interpretierbarkeit und Themenanzahl) angeben. Zudem gehen wir von einem Anstieg des Anteils an Antworten mit umweltfreundlichem Potential aus.

Methode

Die Stichprobe umfasste $N = 225$ Schüler:innen der Jahrgangsstufen 7 bis 9. Die Studie folgte einem randomisierten Kontrollgruppendesign mit Prä-Post-Messungen. Das Umweltverhalten wurde zu drei Testzeitpunkten durch jeweils eine offene Frage zum Naturschutz und zum Klimaschutz erfasst und anschließend inhaltsanalytisch ausgewertet. Zur Überprüfung der Wirksamkeit der BNE-Zooführung wurden die Antworten in Hinblick auf die Antwortcharakteristika „Wortanzahl“, „Interpretierbarkeit“ und „Themenanzahl“ durch Welch-Tests verglichen. Die Antworten wurden außerdem in ein Kategoriensystem eingeordnet, in dem eine Einschätzung des umweltfreundlichen Potentials der Verhaltensweisen getroffen wurde. Auch diesbezüglich wurde ein Gruppenvergleich vorgenommen.

Ergebnisse

Die Ergebnisse zeigen, dass die Experimentalgruppe weder kurz- noch langfristig nach der BNE-Zooführung einen signifikanten Unterschied hinsichtlich der Wortanzahl ($p = [.166; .655]$), der Interpretierbarkeit ($p = [.124; .842]$) und der Themenanzahl ($p = [.104; .820]$) beider offenen Fragen im Vergleich zur Kontrollgruppe aufwies. Obwohl Jugendliche spontan eine Vielzahl umweltschützender Verhaltensweisen nannten (47 Verhaltensweisen aus 10 Oberkategorien), konzentrierten sie sich vorwiegend auf solche mit lokal sichtbaren Auswirkungen (bspw. „Ich trete auf keine Pflanzen“). Inhaltlich wurden die Antworten durch die Teilnahme an der Zooführung nicht differenzierter. Die Antworten der Gruppen unterschieden sich hinsichtlich des umweltfreundlichen Potentials der Verhaltensweisen zu keinem Testzeitpunkt. Die Studie legt somit nahe, dass die gegenwärtige Ausgestaltung der BNE-Zooführung im Allwetterzoo Münster nicht ausreichend wirksam ist, um Verhaltensänderungen bei Jugendlichen herbeizuführen. Die inhaltsanalytische Auswertung der offenen Antworten verdeutlicht, dass eine differenziertere Wissensgrundlage im Hinblick auf umweltfreundliches Verhalten anzustreben ist.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 61: Wann verhalten sich Studierenden nachhaltig? Zusammenhänge zwischen motivationalen Einstellungen und alltäglichen Entscheidungen.

Annika Koch, Fani Lauermann

Universität Bonn, Deutschland

Ziele/Hintergrund

Es besteht ein starker Konsens über die Verantwortung der Menschheit, sich nachhaltiger zu verhalten. Eine wesentliche Voraussetzung dafür ist eine nachhaltige Einstellung (Ajzen, 1991; Kollmuss & Agyeman, 2002). Studien zeigen jedoch, dass sich positive Einstellungen und Absichten zum Umweltschutz nicht immer und vollständig im Verhalten der Menschen widerspiegeln, selbst wenn kontextbezogene Faktoren grundsätzlich umweltfreundliches Verhalten ermöglichen (Carrington et al., 2014; Juwan & Dolnicar, 2014; Kennedy et al., 2009; Kollmuss & Agyeman, 2002). Diese Annahme spiegelt sich auch in gemischten empirischen Befunden zum Zusammenhang zwischen Einstellungen und nachhaltigem Verhalten wider (siehe Review Zhuo et al., 2022; Miller et al., 2022). Nach der Theorie des geplanten Verhaltens (Ajzen, 1991) müssen Einstellungen, subjektive Normen, wahrgenommene Verhaltenskontrolle und Hintergrundmerkmale der Person berücksichtigt werden, um ein bestimmtes Verhalten vorherzusagen. Das bedeutet, dass Menschen sich nicht umweltbewusst verhalten können, wenn sie nicht das Gefühl haben, dass sie die Fähigkeiten, Möglichkeiten oder Ressourcen dazu haben (Ajzen, 1991; Kollmuss & Agyeman, 2002). Vor allem motivationale Faktoren scheinen also eine entscheidende Rolle bei der Einstellungs-Verhaltens-Lücke zu spielen. So konnte beispielsweise empirisch gezeigt werden, dass persönliche Kosten und Selbstkontrolle (Wyss et al., 2022) sowie umweltbezogene Selbstwirksamkeit (Miller et al., 2022) in einem Zusammenhang mit nachhaltigen Verhalten stehen.

Viele Studien beziehen sich jedoch häufig auf Intentionen anstatt auf direktes Verhalten oder Wahlentscheidungen oder definieren Nachhaltigkeit eindimensional über Umweltthemen.

Dementsprechend fokussieren wir uns in der folgenden Studie auf den Zusammenhang zwischen direkten Verhaltensindikatoren wie das Spendenverhalten und Wahlentscheidungen in alltäglichen Einkaufsszenarien von Lebensmitteln und Kleidung und dem selbstberichteten nachhaltigen Verhalten sowie Einstellungen und Motivationsfacetten von Studierenden an deutschen Hochschulen. Dabei untersuchen wir:

Fragestellung: Welche Zusammenhänge bestehen zwischen Einstellungen, Motivation, selbstberichteten Verhaltensweisen im Alltag und direkten Wahlentscheidungen (Spenden- und Kaufentscheidungen) von Studierenden?

Methodik und Analysen

Wir nutzen erste Daten eines selbst entwickelten Tests zu Nachhaltigkeitskompetenzen von Studierenden (OSA|NK), in welchem Fragen zu Wissen, Einstellungen und Verhalten von Studierenden integriert werden. Bisher konnten 55 Studierende (58,2% weiblich) an deutschen Hochschulen aus den Fachbereichen Agrarwissenschaften (30,9%), Rechtswissenschaften (41,8%) und Ingenieurwissenschaften (27,3%) befragt werden. Die Haupterhebung ist für das Wintersemester geplant, für die eine größere Stichprobe erwartet wird.

Die zuvor validierten Skalen wurden zielgruppenspezifisch angepasst. Die Verhaltensweisen der Studierenden wurden über selbstberichtete Alltagssituationen sowie traditionelle Conjoint-Items und eine Spendenfrage erhoben.

Mittels bivariater Korrelationen wurden erste Zusammenhänge zwischen den Teilnutzenwerten der Wahlaufgabe, der Spendenhöhe an eine Nachhaltigkeitsorganisation, den selbstberichteten Verhaltensweisen, den allgemeinen Einstellungen zur Nachhaltigkeit und den motivationalen Einstellungen untersucht.

Ergebnisse/Diskussion

Es zeigt sich (vgl. FF 1), dass Studierende, die positive Einstellungen zum Thema Nachhaltigkeit angeben, mehr Geld an Nachhaltigkeitsorganisationen spenden ($r = .32^*$ bis $.53^*$) und sich für Müllvermeidung, Recycling, nachhaltigen Konsum und Umweltorganisationen engagieren ($r = .32^*$ bis $.58^*$). Darüber hinaus spenden Studierende mehr Geld, wenn sie höhere motivationale Einstellungen (Selbstkonzept, persönliche Wichtigkeit, intrinsische Motivation und Nutzen) berichten ($r = .33^*$ bis $.53^*$) und entscheiden sich bei höherer Motivation eher für Lebensmittel und Kleidung mit geringer Umweltbelastung ($r = .34^*$ bis $.45^*$), wenn der Preis einen geringen Teilnutzen darstellt ($r = -.35^*$ bis $-.45^*$) (siehe Tabelle 2).

Wir erweitern die bisherige Forschung, indem wir das Spendenverhalten als direkten Verhaltensindikator und alltägliche Kaufsituationen als Wahlentscheidungen betrachten. Der Zusammenhang zwischen Einstellungen und Verhaltens- bzw. Wahlentscheidungen ist auch in dieser Untersuchung nicht ganz eindeutig. Allerdings scheinen die angegebenen Spenden einen hilfreichen direkten Verhaltensindikator darzustellen, über den die Mechanismen der Einstellungs-Verhaltens-Lücke besser erklärt werden könnten. Vor allem motivationale Einstellungen scheinen hier entscheidende Zusammenhänge widerzuspiegeln.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 20: Core Practices als boundary object – ein Kodierschema zu interinstitutionellen Aushandlungen bezüglich der Beschreibung von Kernpraktiken für die Lehrkräftebildung im Fach Latein

Judith Schellenbach-Zell, Roswitha Ritter

Bergische Universität Wuppertal, Deutschland

Einleitung und theoretischer Hintergrund

Die aktuelle Forschung weist aus verschiedenen Perspektiven auf Inkohärenzen in der Lehrkräftebildung hin, Stellen also, an denen es für Studierende und angehende Lehrkräfte schwierig ist, das wissenschaftliche Wissen aus der Universität für das Handeln in der unterrichtlichen Praxis zu nutzen.

Ein Grund dafür lässt sich auf institutionell-formeller Ebene lokalisieren, wenn Inhalte und Lerngelegenheiten verschiedener Ausbildungsphasen kaum aufeinander bezogen werden (Cramer, 2020). Dabei spielt eine Rolle, dass die verschiedenen lehrkräftebildenden Institutionen abgeschlossene Systeme bilden und unterschiedliche Kulturen, Logiken und Perspektiven aufweisen (Farley-Ripple et al., 2018). Die Forschung zum boundary crossing (Akkermann & Bakker, 2011) beschreibt diese Systemgrenzen als „socio-cultural difference leading to discontinuity in action or interaction“ (Akkermann & Bakker, 2011, S. 133). Arbeiten Akteur*innen dieser Institutionen zusammen, so treten sie aus ihren professionellen Systemen heraus: Die Schnittstelle ergibt einen Raum, in dem gemeinsames Lernen und somit das Erarbeiten einer gemeinsamen Wissensbasis und einer gemeinsamen Sprache als Grundlage der Kooperation möglich ist (Hartmann & Decristan, 2018).

Im Beitrag gehen wir davon aus, dass die gemeinsame Arbeit von Akteur*innen aus verschiedenen Ausbildungsfacetten (Fachwissenschaft, Fachdidaktik, Ausbildungsinstitutionen für die zweite Phase, Lehrkräfte) an einem phasenübergreifenden, praxisorientierten Ausbildungskonzept (die sogenannten "Core Practices", vgl. Grossman et al., 2009) dazu führt, dass angehende Lehrkräfte die Lehrangebote als zueinander stimmig und kohärent erleben (Hellmann, 2019). Das praxisorientierte Konzept der Core Practices lässt sich vor diesem Hintergrund verstehen als boundary object, also als Artefakt, das in allen Systemen erkannt wird und somit einen Ausgangspunkt für die Entwicklung einer gemeinsamen Sprache darstellen kann (z.B. Star, 2010). Auf diese Weise können die Akteur*innen in ihrem Lernen und der gemeinsamen Wissenskonstruktion unterstützt werden (Schellenbach-Zell & Hartmann, 2022).

Fragestellung

Vor diesem Hintergrund gehen wir der Frage nach, wie Akteur*innen aus verschiedenen professionellen Kontexten einzelne Core Practices für das Fach Latein identifizieren, entwickeln und die inhaltliche Ausgestaltung aushandeln. Wir wollen dabei nachzeichnen, welche Themen in die Aushandlungen eingebracht werden und welche Muster sich in den Gesprächsbeiträgen der Akteur*innen aus den verschiedenen professionellen Kontexten identifizieren lassen.

Methode

Im Rahmen eines zweitägigen Workshops mit dem Thema „Core Practices für die Lehrkräftebildung im Fach Latein?!“ wurden Fokusgruppendifkussionen audiografiert. Insgesamt fünf Teilnehmer*innen aus der Fachdidaktik, der Fachwissenschaft und der zweiten Phase der Lehrkräftebildung diskutierten die Ausrichtung und Beschreibung der Core Practice „Textarbeit anleiten“.

Das Poster wird das Kodierschema präsentieren, anhand dessen die Gruppendiskussion analysiert werden soll. Wir orientieren uns dabei am Vorgehen von Popp und Goldman (2016) und kodieren das sprachliche Material zweischrittig: Im ersten induktiven Schritt erfolgt eine Kodierung der eingebrachten Themen mit dem Ziel, inhaltliche Schwerpunkte der Aushandlungen zu identifizieren (z.B. Verweise auf institutionelle Usancen); die Kodierung erfolgt hier konsensuell. Der zweite Schritt gestaltet sich stärker deduktiv-induktiv, indem wir jeden Sprecherwechsel bestimmten Diskursbewegungen zuordnen. Dabei orientieren wir uns u.a. an der Übersicht von Diskursbewegungen von Popp und Goldman (2016), die verschiedene dieser Bewegungen als dem gemeinsamen Lernen zuträglich oder weniger zuträglich charakterisieren. Hier werden wir uns bei Vorliegen einer ausreichenden Rater-Übereinstimmung das sprachliche Material aufteilen und getrennt kodieren.

Ausblick auf Ergebnisse und ihre Bedeutung

Auf dieser Datengrundlage erhoffen wir uns Erkenntnisse dazu, wie eingebrachte Themen aufgenommen und verhandelt werden, welche Argumentationsmuster sich zeigen und inwiefern dabei ko-konstruktive, dem Aufbau neuen kollektiven Wissens zuträglich Interaktionsmuster erkennbar werden. Dies kann ein erster Hinweis darauf sein, welche Möglichkeiten und Grenzen eine Zusammenarbeit auf der institutionell-formellen Ebene aufweist und wie diese ausgestaltet werden kann.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 62: Individuelle Unterschiede in der korrelierten Veränderung von situativer Langeweile und intrinsischem Wert von Studierenden in einer spezifischen Lernsituation

Lina Wirth¹, Saskia Samland¹, Timo Ehmke¹, Jan Retelsdorf², Poldi Kuhl¹

¹Leuphana Universität Lüneburg, Deutschland; ²Universität Hamburg

Theoretischer Hintergrund

Langeweile ist eine häufig erlebte Emotion in Leistungssituationen. Sie gilt als unangenehm und aversiv und ist u. a. gekennzeichnet durch eine verlängerte Zeitwahrnehmung oder dem Wunsch nach Veränderung (Goetz & Hall, 2014; Pekrun et al., 2010). Durch die Einschränkung kognitiver Prozesse wie Aufmerksamkeit, Selbstregulierungs- oder Lernstrategien wirkt sich Langeweile nachteilig auf das Lernen und die individuellen Leistungen aus (Camacho-Morles et al., 2021; Pekrun et al., 2010; Tze et al., 2016). Um die Entstehung von Langeweile und anderen Leistungsemotionen zu verstehen, gilt die Kontroll-Wert-Theorie von Pekrun (2006) als eines der etabliertesten Erklärungsmodelle. Gemäß dieser bedingen das Gefühl der Kontrolle über leistungsbezogene Aktivitäten und der subjektive Wert, der diesen Aktivitäten zugeschrieben wird, die Leistungsemotionen (Pekrun, 2006). In Bezug auf den subjektiven Wert postulieren Pekrun et al. (2010), dass ein Mangel an intrinsischem Wert von besonderer Bedeutung für die Entwicklung der Langeweile ist.

Daran anknüpfend untersucht diese Studie die Entwicklung der situativen Langeweile von Studierenden in Korrelation zum intrinsischen Wert in einer spezifischen Leistungssituation. Während der Forschungsfokus im Kontext der Leistungsemotionen zunehmend auf Untersuchungen der Ausprägung von Langeweile und anderen Leistungsemotionen in einer Situation oder der Entwicklung von Situation zu Situation liegt (z.B. Pekrun & Marsh, 2022; Krannich et al., 2022), ist über Entwicklungen über mehrere Messzeitpunkte *innerhalb* einer Leistungssituation bislang wenig bekannt.

Fragestellungen

- (1) Inwieweit entwickeln sich die situative Langeweile und der intrinsische Wert bei Studierenden in einer spezifischen Lernsituation in Korrelation zueinander?
- (2) Lassen sich Gruppen von Studierenden mit unterschiedlichen korrelierten Entwicklungsverläufen identifizieren?
- (3) Inwieweit ergeben sich aus den Ergebnissen der vorangegangenen Fragen Unterschiede in den Leistungen der Studierenden?

Methode

An der Datenerhebung nahmen $N=123$ Lehramtsstudierende teil (87,0% weiblich, 73,1% 21–25 Jahre). Diese schauten sich ein fünfminütiges Erklärvideo zu einem statistischen Sachverhalt an und absolvierten anschließend einen Leistungstest bestehend aus neun Multiple-Choice-Fragen. Die situative Langeweile der Studierenden wurde mit dem Single-Item „Ich bin gelangweilt.“ und der intrinsische Wert mit drei Items (u.a. „Ich bin daran interessiert, etwas über das Thema zu lernen.“) zu drei Zeitpunkten erfasst: vor (t1) und während (t2) des Erklärvideos sowie nach dem Leistungstest (t3). Sowohl die Langeweile als auch der intrinsische Wert wurden auf einer vierstufigen Skala von 1 („trifft nicht zu“) bis 4 („trifft zu“) bewertet. Zur Beantwortung der Forschungsfragen wurde eine multivariate latente Klassenwachstumsanalyse (Wickrama et al., 2022) durchgeführt.

Ergebnisse

Die situative Langeweile und der intrinsische Wert korrelierten zu allen drei Messzeitpunkten signifikant negativ, verblieben im Verlauf der Lernsituation aber auf einem mittleren Niveau, sodass eine korrelierte Veränderung nicht untersucht werden konnte (Forschungsfrage 1). Außerdem ließen sich zwei Gruppen von Studierenden mit unterschiedlicher Ausprägung der Langeweile und des intrinsischen Werts identifizieren (Forschungsfrage 2). Während Gruppe 1 ($n=62$) mehr Langeweile und einen geringeren intrinsischen Wert berichtete, berichtete Gruppe 2 ($n=61$) weniger Langeweile und einen höheren intrinsischen Wert. Beide Gruppen unterschieden sich signifikant in ihren Mittelwerten zu allen drei Zeitpunkten und beide Variablen korrelierten signifikant innerhalb jeder Gruppe über alle drei Zeitpunkte hinweg. Eine signifikante Veränderung über die Zeitpunkte wurde jedoch nur für den intrinsischen Wert in beiden Gruppen festgestellt. Infolgedessen konnte auch hier keine korrelierte Veränderung mit der Langeweile untersucht werden. Es zeigte sich jedoch ein signifikanter Unterschied in den Leistungstestergebnissen zwischen den beiden Gruppen (Forschungsfrage 3). Gruppe 1 schnitt signifikant schlechter ab als Gruppe 2.

In weiteren Analyseschritten sollen weitere Emotionen, Erfolgserwartungen und das statistikbezogene Selbstkonzept der Studierenden, die ebenfalls im Rahmen dieser Studie erhoben wurden, in die Auswertungen einbezogen werden. Damit trägt die Studie zu einem differenzierten Verständnis darüber bei, wie sich Leistungsemotionen in spezifischen Leistungssituationen entwickeln. Die Ergebnisse werden vor dem Hintergrund der Kontroll-Wert-Theorie sowie im Hinblick auf theoretische und praktische Implikationen für die weitere Forschung und die Bildungspraxis diskutiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 32: Governance und Erhebung der Schulischen Gesundheitsförderung in der Schweiz

Simona Betschart, Anita Sandmeier, Guri Skedsmo

Pädagogische Hochschule Schwyz, Schweiz

Seit der Veröffentlichung der Ottawa-Charta durch die Weltgesundheitsorganisation werden Schulen als zentrale Knotenpunkte für Gesundheitsförderung anerkannt (WHO, 1986). Die WHO befürwortet einen ganzheitlichen Ansatz, bei dem sämtliche Personen und Ebenen einer Schule einbezogen werden, allgemein bekannt als "Health Promoting Schools Ansatz" (HPS) (WHO Expert Committee, 1997). Während HPS verschiedene Gesundheitsaspekte beinhalten kann (SHE Network, 2013), geht diese Arbeit speziell auf die Förderung der psychischen Gesundheit ein, da die globale Pandemie Probleme im Zusammenhang mit der psychischen Gesundheit von Kindern und Jugendlichen in ganz Europa verschärft hat (z. B. Barbieri et al., 2022; Weare & Nind). Innerhalb der Implementierung von HPS werden Schulleiter/innen als Schlüsselfiguren betrachtet. Sie wirken als Gatekeeper/innen und beeinflussen die Durchführung und Aufrechterhaltung von Gesundheitsförderungsaktivitäten in ihren Schulen massgeblich (z. B. Dadaczynski et al., 2022). Bis heute gibt es keinen umfassenden Überblick über den Grad der Umsetzung der HPS-Massnahmen an Schweizer Schulen, weder allgemein noch mit Fokus auf die Förderung der psychischen Gesundheit. Bestehende Studien zu Schulgesundheitsförderungsaktivitäten fehlen häufig entweder ein Fokus auf HPS oder auf die psychische Gesundheit (Kunz Heim & Mittag, 2016; Zumbrunn et al., 2016). Zwei Hauptgründe tragen zu dieser Lücke bei: Erstens führt das dezentrale Bildungssystem der Schweiz zu einer Vielzahl von Gesundheitsförderungsprogrammen, Anbietenden und Angeboten, was es herausfordernd macht, einen umfassenden Überblick über alle Aktivitäten und beteiligten Akteure/innen auf nationaler, subnationaler und lokaler Ebene zu erhalten. Zweitens, aufgrund der komplexen Natur des HPS-Ansatzes gestaltet sich die Bewertung des Grades der HPS-Implementierung als herausfordernd. Um dieser Herausforderung entgegenzutreten, entwickelten Dix et al. (2019) eine Skala für den australischen Kontext (SSPESH), welche darauf abzielt, den Grad der HPS-Implementierung umfassend und ökonomisch zu bewerten. Dieses Projekt die Governance-Strukturen und beteiligten Akteure/innen identifizieren, die an der HPS-Implementierung in der Schweiz beteiligt sind, und das Instrument von Dix et al. (2019) im schweizerischen Kontext validieren.

Die folgenden Forschungsfragen waren von Interesse:

- Welche Governance-Strukturen auf nationaler, subnationaler und lokaler Ebene beeinflussen die Umsetzung von HPS-Massnahmen in der Schweiz?
- Ist die Skala zur Förderung emotionaler und sozialer Gesundheit in der Schule (SSPESH) für den schweizerischen Kontext geeignet?
- Wie hoch ist der Implementierungsgrad der Förderung emotionaler und sozialer Gesundheit in Schweizer Schulen?

Es wurde eine Dokumentenanalyse auf nationaler und subnationaler Ebene mit MAXQDA durchgeführt. Zusätzlich wurden zwei Online-Umfragen durchgeführt: eine im Juni 2021 mit N / 385 Schulleiter/innen und eine im Juni 2022 mit N / 172 teilnehmenden Schulleiter/innen. Beide Stichproben wurden verwendet, um die Konstruktvalidität der SSPESH-Skala (Dix et al., 2019), die auf dem HPS-Rahmenwerk basiert (Langford et al., 2015), zu testen. Es wurde eine explorative Faktorenanalyse mit SPSS 26.0 und eine konfirmatorische Faktorenanalyse mit R durchgeführt. Darüber hinaus wurden die Konstruktzuverlässigkeiten für die latenten Faktoren berechnet und schliesslich die Beziehungen zwischen SSPESH und anderen Faktoren (z.B. Ressourcen für HPS, Schulrichtlinien für HPS) untersucht.

Erste Ergebnisse deuten auf eine schwache Verankerung von HPS auf nationaler Ebene in der Schweiz hin. Während das Bundesamt für Gesundheit (BAG) und die Konferenz der kantonalen Erziehungsdirektionen (EDK) jeweils unverbindliche Hinweise zur schulischen Gesundheitsförderung machen, ist die systematische Zusammenarbeit zwischen Gesundheits- und Bildungssektoren minimal. Die konkrete Umsetzung von HPS erfolgt meist über diese NGOs, wobei diese nur sporadisch miteinander kooperieren. Die kantonalen Gesundheits- und Bildungsbehörden zeigen sehr unterschiedliche Strukturen, was die Organisation von schulischer Gesundheitsförderung betrifft. Eine umfassende Analyse über alle 26 Kantone war aufgrund des Umfangs der Dokumente nicht praktikabel.

Des Weiteren werden Ergebnisse zur Validität und Reliabilität der SSPESH-Skala bei Anwendung an Schweizer Schulen präsentiert, was zur Anpassung und Verfeinerung von Bewertungsinstrumenten für Gesundheitsförderung in unterschiedlichen Kontexten beiträgt. Die Ergebnisse werden den aktuellen Implementierungsstand der Förderung emotionaler und sozialer Gesundheit in Schweizer Schulen aufzeigen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 63: Automatisierte Kodierung von Kurzantworten im Primarbereich anhand von VERA-3-Aufgaben

Sarah Paul, Thomas Canz

Institut zur Qualitätsentwicklung im Bildungswesen, Deutschland

Theoretischer Hintergrund und Fragestellung

Die automatisierte Kodierung kann dazu beitragen, die Effizienz von Bewertungsprozessen im Bildungskontext zu steigern und die Validität von Testergebnissen zu verbessern (Hahn et al., 2021). Jede Aufgabe erfordert spezialisierte Einstellungen, beispielsweise durch geeignete Wahl eines semantischen Raums oder ausgewählter Preprocessing-Schritte, um eine präzise Kodierung zu gewährleisten (Andersen & Zehner, 2021). Der Erfolg von automatisierten Kodierungen hängt auch von äußeren Faktoren wie der Rechtschreibung von Schüler*innenantworten, der Länge der Antworten und dem Bewertungsschema ab (Zhai et al., 2021). Horbach und Zesch (2019) wiesen ebenfalls auf die Relevanz der Antwortlänge für den Kodierprozess hin. Bisherige Forschungen zum automatisierten Kodieren fokussierten Deutschaufgaben. Für Mathematikaufgaben, die stark in ihrer Länge als auch in der Schreibweise von Formeln und Gleichungen variieren (Baral et al., 2021) liegen bisher nur wenig Befunde vor. Des Weiteren konnte eine Forschungslücke bei der automatisierten Kodierung im Primarbereich festgestellt werden. Ziel des Beitrags ist es zu prüfen, wie gut bereits in anderen Kontexten erfolgreich evaluierte automatische Kodiertools - im vorliegenden Fall shinyReCoR (Zesch et al., 2023, evaluiert anhand von PISA-Daten) - Freitextantworten von Grundschüler*innen beurteilen können.

Methode

Für die Analyse wurden Kurztextantworten zu VERA-3-Aufgaben in den Fächern Deutsch und Mathematik mit der Software shinyReCoR kodiert. Hierbei wurde die Beurteilung von Schüler*innenantworten auf Grundlage von Mensch-Maschine-Übereinstimmungsmaßen geprüft. Als Übereinstimmungsmaß diente das quadrierte Cohens Kappa (Zhai et al., 2021). Als differentielle Variablen wurden die Domänen der Aufgaben und die durchschnittliche Antwortlänge betrachtet. Grundlage des Datensatzes waren Aufgaben aus den Pilotierungen für die VERA-3-Durchgänge 2013 bis 2024 (inklusive Normierungsaufgaben) für die Fächer Deutsch und Mathematik. Hierbei wurden aufgrund der benötigten Fallzahl insgesamt zwölf Ankeritems aus den Pilotierungsstudien gewählt, welche alle Kompetenzen der Fachbereiche abdecken. Bei der Auswahl wurde die Diversität der sprachlichen Anforderungen berücksichtigt. Diese variierten von eindeutigen Ein-Wort-Antworten bis zu Begründungsformaten (für Mathematik inklusive Gleichungen). Die Aufgaben wurden handschriftlich erhoben und für diese Arbeit digital transkribiert und manuell kodiert.

shinyReCoR ist eine R-basiertes automatisches Kodiertool für Kurztextantworten und stellt mehrere semantische Räume zur Verfügung (Andersen & Zehner, 2021). Aufgrund des geringen semantischen Anspruchs der Aufgaben wurde der vorproduzierte semantische Raum „Allgemein“ verwendet. Verschiedene Preprocessing-Schritte (Andersen & Zehner, 2021) wie das Entfernen von Zeichen und Symbolen oder die Korrektur der Rechtschreibung wurde je nach Kompetenzbereich und Aufgabenanforderung angewendet. Aus den so entstandenen Antwortvektoren werden in der Antwortanalyse Cluster mittels eines varianzbasierten hierarchischen agglomerativen Clusterverfahrens (Ward Jr., 1963) bestimmt. Die Clustergröße wurde experimentell bestimmt.

Ergebnisse und Diskussion

Die beobachteten absoluten Übereinstimmungen zeigten geringe Unterschiede zwischen Deutsch und Mathematikaufgaben ($M(\text{ges.}) = 0,84$; $M(\text{Mathe}) = 0,79$; $M(\text{Deutsch}) = 0,88$) wohingegen die quadrierten Kappa zwischen den Fachbereichen stark variierten ($M(\text{ges.}) = 0,55$; $M(\text{Mathe}) = 0,39$; $M(\text{Deutsch}) = 0,67$). Ebenfalls konnte der negative Zusammenhang zwischen Wortlänge und Cohens Kappa nachgewiesen werden ($\beta = -0,046$). Die Ergebnisse weisen auf eine gute Beurteilung von Freitextantworten von Grundschüler*innen hin. Für diejenigen Aufgaben mit geringen quadrierten Kappa ($< 0,80$, vgl. Zhai et al., 2021) könnten möglicherweise unter Anwendung aufgabenspezifisch definierter semantischer Räume Performance-Zugewinne erzielt werden. Eine potentielle Erklärung für die in der vorliegenden Studie ausgemachten Defizite bezüglich der Übereinstimmungsmaße von Mathematik-Aufgaben könnte die Formatierung von Gleichungen in den Freitextantworten sein. Zukünftige Analysen mit einem größeren Aufgabenpool sollen die unterschiedlichen Kompetenzbereiche innerhalb der Fächer als weitere mögliche Varianzquelle betrachten.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 26: Kritische Nutzung von Online-Informationen von Studierenden der Wirtschaftswissenschaften – Ein Literature Review der internationalen Forschung

Olga Zlatkin-Troitschanskaia, Paul Frederik Hodes, Jennifer Fischer

Johannes Gutenberg-Universität Mainz, Deutschland

Theoretischer Hintergrund/Fragestellung

Im 21. Jahrhundert ist das Internet eine primäre Informationsquelle für Studierende geworden [1,2]. Digitale Technologien ermöglichen es Studierenden, jederzeit und überall auf eine Vielzahl von Informationen zuzugreifen, was dazu führt, dass sie Online-Ressourcen den traditionellen Lehr-Lern-Medien, wie Lehrbüchern, vorziehen [3,4,5]. Diese Entwicklungen bringen neue Herausforderungen, wie Verbreitung von fehlerhaften Informationen, weshalb ein kritisch-reflektierter Umgang mit Online-Informationen eine entscheidende Kompetenz für den Studienerfolg darstellt [6].

Im Vergleich zu moderierten digitalen Lehr-Lernumgebungen erfordert selbstgesteuertes Lernen im Internet von Studierenden die Fähigkeit, zuverlässige Quellen zu finden, die Qualität deren Inhalte kritisch zu bewerten sowie zentrale Informationen aus mehreren, oft widersprüchlichen Quellen zusammenzuführen [7]. Diese Fähigkeiten werden im Konstrukt Critical Online Reasoning (COR) subsumiert, das aus drei Facetten besteht: Online-Informationsbeschaffung, Kritische Informationsbewertung und Schlussfolgern mit Evidenzen, Argumenten und Synthese [8]. Während COR-Facetten als zentral für den Studienerfolg angesehen werden [8], zeigen Studien erhebliche Defizite bei Studierenden, insbesondere im Bereich der Wirtschaftswissenschaften, vor allem bei der effizienten Suche und der kritischen Bewertung von Online-Informationen.

Trotz umfangreicher Forschung zum Online-Lernen und Konzepten wie Information Problem-Solving [12] und Multiple Sources Use/Comprehension [13] zur Analyse von COR-bezogenen Fähigkeiten, ist die Forschung zum selbstgesteuerten Lernen im Internet und Strategien der Studierenden fragmentiert und begrenzt [14,15,16,17,18,19]. Für ein systematisches Verständnis effektiven Online-Lernens im Hochschulbereich ist eine umfassende Synthese bestehender Forschungsansätze/-ergebnisse erforderlich. Daher haben wir eine systematische Übersicht zur internationalen Forschung zu COR-bezogenen Fähigkeiten von Studierenden, einschließlich ihres Online-Verhaltens, Bewertungsstrategien und Einsatzes von Online-Quellen für akademisches Lernen, mit Fokus auf Studierenden der Wirtschaftswissenschaften, erstellt. Die Posterpräsentation zielt darauf ab, Entwicklungen in der Nutzung von Online-Medien im Studium und den erforderlichen Fähigkeiten der Studierenden seit 2020 zu beschreiben, um diagnostische und instruktionale Herausforderungen und Chancen im Hochschulbereich zu identifizieren.

Methode

In 2023/2024 wurde basierend auf einem bestehenden Forschungs-Review bis inkl. 2020 [19] eine strukturierte Literaturübersicht erstellt. Die systematische Literatursuche erfolgte nach dem „Mixed Studies Review“-Ansatz [20,21] und umfasste die Datenbanken ERIC, Web of Science und Google Scholar. Durch die Einbeziehung der etablierten wissenschaftlichen Datenbanken sowie Suchmaschinen, wurde ein breites Spektrum an relevanter Forschungspublikationen wie Zeitschriftenartikeln, Dissertationen und Konferenzbeiträge einbezogen. Im Rahmen der vorliegenden Untersuchung wurden über 500 Publikationen aus den Jahren 2021 bis 2024 identifiziert. Basierend auf den zuvor festgelegten Selektionskriterien wurden anschließend mehr als 25 empirische Studien aus 18 Ländern für eine systematische, kriterienbasierte Inhaltsanalyse ausgewählt. Der Fokus dieser Studien liegt auf Studierenden der Wirtschaftswissenschaften und ihren Strategien bei der Nutzung von Online-Informationen.

Ergebnisse

Die ausgewählten Studien wurden u.a. mit Blick auf die Befunde zu den COR-Facetten betrachtet. Dazu zählen Strategien bei der Nutzung von Suchmaschinen/Datenbanken, bei der Bewertung von Qualität und Glaubwürdigkeit der Quellen/des Inhalts sowie bei der Nutzung von Online-Quellen für domänenspezifisches Lernen. Die Zusammenstellung beinhaltet zudem individuelle und kontextuelle Faktoren, die mit der Online-Informationsnutzung zusammenhängen, wie epistemische Überzeugungen, Vorwissen, Studienjahr oder soziodemografische Merkmale (wie Geschlecht). Ein weiterer Analyseschwerpunkt ist die Entwicklung des Nutzungsverhaltens von Online-Informationen im Studienverlauf und dessen Einfluss auf den Studienerfolg.

Die Analyse zeigt, dass die Nutzung von Online-Medien im Hochschulbereich zunehmend an Bedeutung gewinnt (im Vergleich zur Überprüfung aus 2021 [19]). Gleichzeitig werden erhebliche Kompetenzdefizite bei den Studierenden sichtbar, was auf die Notwendigkeit gezielter Schulungen im Umgang mit Online-Medien und validierter Diagnostik hinweist.

Die systematische Analyse konzentriert sich auf empirische Erkenntnisse und trägt wesentlich zur Konsolidierung der aktuellen internationalen Forschung in diesem interdisziplinären Bereich bei. Die Ergebnisse betonen die Bedeutung von COR-Kompetenzen bei Studierenden der Wirtschaftswissenschaften (und anderer Disziplinen) und offenbaren gleichzeitig erhebliche Forschungslücken bezüglich der domänenspezifischen Nutzung von Online-Informationen und deren Einfluss auf den Studienerfolg. Diese und weitere Ergebnisse des Literatur-Reviews werden auch im Hinblick auf die Limitationen der Studie, wie etwa den ausschließlichen Einbezug englischsprachiger Publikationen, kritisch diskutiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 64: Nur eine Dimension? Geographiedidaktisches Basiswissen angehender Lehrpersonen zu inklusivem Geographieunterricht in einem Partial Credit Modell

Dietmar Göllitz, Ann-Sophie Winklmaier, Christiane Tellesch-Bülow, Jan Christoph Schubert

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Theoretischer Hintergrund

Geographieunterricht inklusiv zu gestalten und dabei geographisch-fachliche Lernprozesse zu ermöglichen, stellt hohe Anforderungen an angehende Lehrpersonen und deren professionelle Kompetenzen (Winklmaier & Schubert, im Druck). Neben Einstellungen (Gorges, Grumbach, Micheel & Neumann, 2020; Schwab, Resch & Alnahdi, 2021) (Song, Sharma & Choi, 2019) und motivationalen Grundhaltungen kommt dem professionellen Wissen (Heyder, Vaskova, Hußmann & Steinmayr, 2018; Steinmayr, Heyder & Tometten, 2022) und insbesondere dem fachdidaktischen Wissen (Shulman, 1982; Smit, Tuithof & Béneker, 2023; Smit, Tuithof, Savelsbergh & Béneker, 2023) eine hohe Bedeutung zu. Im Gegensatz zu anderen Fachdidaktiken (Krauss et al., 2017) existieren im deutschsprachigen Raum bislang keine Instrumente, die das geographiedidaktische Wissen bezüglich inklusivem Geographieunterricht messen, so dass keine Aussagen über die Wissensbasis von (angehenden) Lehrpersonen sowie zu den Wirkungen von entsprechenden universitären Veranstaltungen (Tellesch-Bülow et al., 2021) getroffen werden können. Ziel der Testentwicklung war daher, die inhaltliche Breite inklusiven geographiedidaktischen Wissens (FDW) zu operationalisieren. Im Rahmen eines Projekts zur Qualifizierung von Geographielehrpersonen für Inklusive Bildung (Winklmaier et al., 2022) erhoben wir paper-pencil-basiert mit Studierenden unter anderem ihr fachdidaktisches Wissen als Teil der Kompetenzen (Baumert & Kunter, 2006) von (angehenden) Lehrpersonen.

Nach Analyse von unterschiedlichen Klassifizierungen von fachdidaktischem Wissen wurden die Bereiche „Wissen über geographiespezifische Lernvoraussetzungen von Schüler*innen“, „Wissen über die Diagnose von geographiebezogenen Lernständen und entsprechenden Rückmeldungen“ und „Wissen über geographiespezifische Lehrstrategien“ als relevant identifiziert. Das Messinstrument sollte neben diesen Bereichen die sonderpädagogischen Förderbedarfe berücksichtigen sowie fachlich-geographische Bezüge in der Breite herstellen.

Fragestellung

Inwieweit lassen sich FDW Items in einer Stichprobe von Studierenden der Geographiedidaktik als eindimensional beschreiben?

Methode

Die Items wurden mithilfe einer strukturierenden Matrix (FDW-Bereiche, sonderpädagogische Förderbedarfe, geographische Themen) entwickelt, nutzten ein offenes Antwortformat und teilweise Multiple-Choice-Format und wurden mehrfach im Sinne der Inhaltsvalidität mit Expert*innen diskutiert sowie nach einem Pret-Test (Göllitz, Tellesch-Bülow et al., 2021) überarbeitet. Es wurden letztendlich 15 offene und 6 geschlossene Items konstruiert (Schubert, Winklmaier & Göllitz, 2021).

Diese überarbeitete Fassung wurde anhand einer virtuellen Stichprobe von $N = 512$ Studierenden, für welche erhobene Daten zu t1 und zu t2 zusammengefasst wurden, auf Basis eines Item-Response-Modells (Boone & Noltemeyer, 2017; Boone, Staver & Yale, 2014; Linacre, 2024; Rasch, 1980) analysiert. Der Rohdatensatz wurde im Rahmen der Projektberichterstattung veröffentlicht (Göllitz, Thieroff et al., 2021). Die Itemschwierigkeiten wurden mit einer Wright Map für ein Partial Credit Modell dargestellt und auf ihre Dimensionalität (Linacre, 2024) geprüft.

Resultat

Der Itemfit (MNSQ) ist akzeptabel, wobei einzig der Outfit des schlechtesten Items einen relevanten Grenzwert (1.50) erreicht, während alle weiteren Items unter diesem Grenzwert liegen. Der Testscore erreicht eine ausreichende Differenzierungskraft für ungefähr zwei bis drei Gruppen (mit einem standardisierten Cronbachs Alpha = .89, Personenreliabilität = .75, Testseparation = 1.72, latente Retestreliabilität für t1 auf t2 der FDW Scores = .66). Das Ergebnis einer Hauptkomponentenanalyse der Residuen (PCAR) weist auf eine im Verhältnis zum Testziel ausreichende Eindimensionalität hin (Wert im ersten Kontrast = 1.53).

Diskussion

In dieser Studie wurde weder das Ziel verfolgt empirisch zwischen den einzelnen FDW-Bereichen, noch zwischen den einzelnen Förderbedarfen oder geographischen Themenbereichen zu unterscheiden, sondern es sollten Hinweise auf die relative Eindimensionalität des Tests gewonnen werden. Die Outfit- und Infitwerte der Items liegen innerhalb üblicher Grenzwerte, der Eigenwert im ersten Kontrast der PCAR liegt unter dem Grenzwert 2 (Linacre, 2024). Im Verhältnis zum Ziel der inhaltlichen Validität sind diese Werte Hinweise auf die Eindimensionalität des Testscores. Wir diskutieren die Potenziale (Vorher-Nachher-Messung für Ermittlungen von Gruppenwerten) und Grenzen der Verwendung des Messinstruments für die geographiedidaktische Evaluations- und Transferforschung aus Sicht der internalen und externalen Validität. Der Test ist aufgrund des Konstruktionsprinzips und der Werte zur Testgüte für Individualanalysen nicht einsetzbar, als Maß bei Assessments des Gesamtlernerfolgs von Qualifizierungsmaßnahmen jedoch geeignet für gruppenbezogene Analysen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 31: Wie verändert sich die elterliche Einschätzung sprachlicher Kompetenzen im Verlauf der Grundschule?

Anna Ratka-Pauler¹, Jan Lenhart²

¹Leibniz-Institut für Bildungsverläufe, Wilhelmsplatz 3, 96047 Bamberg, Deutschland; ²Otto-Friedrich-Universität Bamberg, Markusplatz 3, 96047 Bamberg, Germany

Theoretischer Hintergrund

Eltern nehmen entscheidenden Einfluss auf die Entwicklung ihrer Kinder, indem sie indirekt (beispielsweise durch die Gestaltung von Rahmenbedingungen) und direkt durch konkretes Verhalten einwirken (vgl. Sénéchal, 2006; Stephenson et al., 2008). Neben anderen Aspekten ist die Fähigkeit der Eltern die Entwicklung ihrer Kinder adäquat einzuschätzen Voraussetzung für die Gestaltung förderlicher Lernangebote und einer angereicherten Lernumwelt (vgl. Gniewosz, 2010; Schmitt et al., 2022). Darüber hinaus ist die elterliche Einschätzung für die Übertrittsentscheidung in die Sekundarstufe I relevant, da diese Entscheidung nach Empfehlungen der Lehrkräfte zumeist den Eltern obliegt (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK), 2015). Die vorliegende Studie untersucht inwiefern sich die elterliche Fähigkeit die Sprachkompetenz ihrer Kinder einzuschätzen im Verlauf der Grundschule verändert. Der Fokus auf die sprachliche Kompetenz wurde gewählt, da sich diese als prädikativ für den weiteren Bildungsverlauf der Kinder erweist (Weinert et al., 2019) und darüber hinaus als ein essenziellen Bestandteil der Kompetenzentwicklung in anderen Domänen wie Mathematik zu verstehen ist (Paetsch et al., 2016).

Bisherige Ergebnisse zeigten, dass die elterliche Einschätzung kognitiver Kompetenzen einen stabilen mittleren Zusammenhang mit den tatsächlichen Kompetenzen der Kinder aufweisen (Deimann et al., 2005; Tabeling et al., 2022). Des Weiteren legen diverse Befunde zur Urteilsgenauigkeit nahe, dass die elterliche Einschätzung mit zunehmendem Alter des Kindes an Genauigkeit gewinnt (vgl. Schmitt et al., 2022). Unter anderem durch die besseren Einblicke in die kindliche Kompetenz anhand der Beurteilung durch Lehrkräfte und der besseren Vergleichbarkeit mit Mitschülerinnen und Mitschülern (vgl. Loibl et al., 2020). Aufgrund dessen ist anzunehmen, dass es ab Klasse drei, in welcher in den meisten deutschen Grundschulen erstmals Noten vergeben werden, zu einer weiteren Annäherung der elterlichen Einschätzung und der kindlichen Sprachkompetenz kommt (Gniewosz, 2010).

Fragestellung

Ziel der vorliegenden Studie ist es die Veränderung der elterlichen Einschätzung längsschnittlich zu analysieren. Es stellt sich die Frage, inwiefern sich die relative Urteilsgenauigkeit der elterlichen Einschätzung sprachlicher Kompetenz im Verlauf der Grundschule verbessert.

Methode

Hierfür wurde auf Daten der Startkohorte 1 des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019) zurückgegriffen. Um die Veränderung der elterlichen Kompetenzeinschätzung über die Primarstufe hinweg bestmöglich abzubilden, werden Daten aus Welle 6 (MZIP1 Vorschule; Alter Zielkind $M = 5.11$), Welle 8 (MZIP2 zweite Klasse; Alter Zielkind $M = 7.09$) und Welle 10 (MZIP3 vierte Klasse; Alter Zielkind $M = 9.07$) verwendet. Zur Erfassung der elterlichen Einschätzung wurde ein Einzelitem verwendet, in welchem die Eltern die sprachlichen Fähigkeiten ihres Kindes im Vergleich zu Gleichaltrigen einordnen. Die Antwortskala reicht von (1) „viel schlechter als andere gleichaltrige Kinder“ bis (5) „viel besser als andere gleichaltrige Kinder“. Die Sprachkompetenz der Kinder wurde mittels der deutschsprachigen Version des „Peabody Picture Vocabulary Test“ (PPVT-4; Lenhart et al., 2005) gemessen. Zu MZIP3 wurden erstmalig die Schulnoten im Fach Deutsch (1 „sehr gut“ bis 6 „ungenügend“) miterfasst, durch welche die Eltern möglicherweise einen Einblick in die tatsächlichen Kompetenzen ihrer Kinder bekommen.

Ergebnisse und ihre Bedeutung

Erste Analysen geben Hinweise auf einen stabilen positiven Zusammenhang (MZIP1 $r = .40$; MZIP2 $r = .38$; MZIP3 $r = .37$) von Elterneinschätzung und kindlicher Sprachkompetenz über den Verlauf der Grundschule hinweg. Eine Verstärkung des Zusammenhangs ist in den vorläufigen Analysen nicht zu beobachten. Auch bei der aufgeklärten Varianz kommt es widererwartend über die drei MZIP hinweg zu einer leichten Abnahme (MZIP1 $R^2 = .16$; MZIP2 $R^2 = .15$; MZIP3 $R^2 = .14$). Betrachtet man allerdings den Zusammenhang der elterlichen Kompetenzeinschätzung zu MZIP3 mit der Schulnote im Fach Deutsch kann auch hier ein signifikanter mittlerer Effekt festgestellt werden ($r = -.32$). Um einen tieferen Einblick in die Veränderung der elterlichen Kompetenzeinschätzung zu erlangen ist es geplant, die Effekte über die verschiedenen MZIP hinweg zu analysieren.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 65: Validierung einer prompt-gestützten Intervention zur differenzierten Förderung selbstregulierten Lernens Lehramtsstudierender in der universitären Prüfungsvorbereitung

Janina Schel, Barbara Drechsel

Otto-Friedrich-Universität Bamberg, Deutschland

Vor allem für Lehramtsstudierende sind Kompetenzen im selbstregulierten Lernen von hoher Bedeutung für das eigene Lernen im Studium wie für die Diagnostik und Förderung ihrer zukünftigen Schüler:innen (Michalsky & Schechter, 2013; Kramarski & Michalsky, 2009). Lerntypologische Ansätze weisen jedoch auf die Heterogenität Lehramtsstudierender bezüglich dieser Kompetenzen hin (z.B. Muwonge et al., 2020), sodass es hier besonders relevant erscheint, mit einer differenzierten und frühzeitigen Förderung anzusetzen (z.B. Theobald, 2021). Typologien können einen systematischen Überblick über die Kompetenzen sowie Entwicklungspotentiale unterschiedlicher Lernprofile bieten, an denen Förderansätze differenziert werden können (Dörrenbächer & Perels, 2016; Schel & Drechsel, *under review*). Interventionsansätze, die Prompts verwenden, können Lernende im selbstregulierten Lernprozess effizient sowie vielfältig fördern, etwa motivational (z.B. Daumiller & Dresel, 2019), meta-kognitiv (z.B. Hübner et al., 2010) sowie kognitiv (z.B. Glogger et al., 2012) und scheinen damit auch für eine differenzierte Förderung geeignet zu sein (Theobald, 2021). Besonders im Lehramtsstudium erscheinen zudem web-basierte Interventionsansätze vielversprechend, da diese ermöglichen, auch große Kohorten zu erreichen (Bellhäuser et al., 2016). Bisher gibt es, trotz Befunden zu unterschiedlichen Lernprofilen in verschiedenen Studiengängen (Parpala et al., 2022), jedoch nur wenige Studien mit personenzentriertem Ansatz, die gezielt Kompetenzen und Entwicklungspotentiale von Lehramtsstudierenden untersuchen (Muwonge et al., 2020), die als Ansatzpunkt für eine differenzierte Intervention dienen könnten. Zudem zeigen generell wenige Interventionsstudien zum selbstregulierten Lernen Möglichkeiten zur Differenzierung der Förderansätze zwischen unterschiedlichen Lernprofilen (Theobald, 2021).

Der vorliegende Beitrag setzt an der genannten Forschungslücke an und geht der Frage nach, inwieweit eine nach Lernprofilen differenzierte, prompt-gestützte Intervention im Rahmen der universitären Prüfungsvorbereitung den Lernerfolg von Lehramtsstudierenden erhöhen kann.

Für die vorliegende Studie wurde eine bereits validierte Lerntypologie zum selbstregulierten Lernen Lehramtsstudierender (Schel & Drechsel, *under review*), die auf empirisch ausgewählten Parametern aus allen drei Phasen des Prozessmodells nach Schmitz & Schmidt (2007) basiert, angewendet. Hierzu füllten die Proband:innen ($N=363$ Lehramtsstudierende) zu Semesterbeginn einen Online-Fragebogen aus, der mittels latenter Profilanalyse ausgewertet wurde (Ferguson et al., 2020) und erhielten eine individuelle Übersicht zu ihren Kompetenzen und Entwicklungspotentialen entlang der untersuchten Parameter sowie eine Empfehlung für eine Prompt-Bedingung. So haben Studierende an der Universität Bamberg die Möglichkeit, wöchentlich zwei DIN A5-Zusammenfassungen (insgesamt maximal 22) semesterbegleitend zur jeweiligen Vorlesungssitzung digital einzureichen, die als zulässiges Hilfsmittel in der Klausur genutzt werden dürfen (Pridöhl, 2022). Die Prompt-Vorlagen für die Erstellung der Vorlesungszusammenfassungen bestehen jeweils aus zwei Seiten mit je einem validierten Prompt (z.B. Glogger et al., 2012), davon je ein kognitiver (drei Bedingungen: Organisieren, Elaborieren und Kritisches Prüfen) auf der ersten Seite und ein einheitlicher meta-kognitiver Prompt zur individuellen Reflexion auf der zweiten Seite. Unterstützend standen den Proband:innen ausführliche Anleitungen und Anwendungsbeispiele zur Umsetzung der Prompts zur Verfügung. Der Lernerfolg wird mittels des Ergebnisses der Klausur am Semesterende ermittelt, die zu gleichen Teilen aus der Prüfung zweier Vorlesungen besteht, wobei die beschriebene Intervention nur zu einer der beiden Vorlesungen durchgeführt wurde.

Die Ergebnisse der Studie stehen aktuell noch aus. Es soll u.a. analysiert werden, ob sich ein signifikanter Unterschied in den durchschnittlichen Klausurpunkten pro Vorlesung und gegenüber der Klausurergebnisse der Vorsemester ohne Intervention zeigt, hypothesenkonform zugunsten der aktuellen Vorlesung mit Intervention.

Das Anreizsystem der vorliegenden Studie könnte volitionale und motivationale Parameter, in denen alle Lernprofile relative Einschränkungen zeigen, im selbstregulierten Lernprozess unterstützen, während differenzierte Prompts, die Nutzung unterschiedlich komplexer kognitiver Lernstrategien für eine tiefere Verarbeitung der Prüfungsinhalte in Abhängigkeit der Lernvoraussetzungen anregen (siehe auch Neroni et al., 2019). Potentiale (z.B. Ökonomie der Intervention) und Limitationen (u.a. das quasi-experimentelle Versuchsdesign) des Ansatzes werden diskutiert. Weiterhin wird angestrebt, die Vorlesungszusammenfassungen auch mittels Lernproduktanalyse (Schel et al., 2023) zu untersuchen, um Entwicklungen im selbstregulierten Lernen auch in konkreten Lernaktivitäten validieren zu können (Theobald, 2021).

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 14: Besser Lehren und Lernen mit Tandemlehre? Kooperative, interdisziplinäre und kollaborative Seminare in der Lehrerbildung im Fokus.

Sibylle Schneider

Universität Augsburg, Deutschland

In Anbetracht der gewachsenen Heterogenität und Pluralität im Unterricht von heute, vor allem aber im Hinblick auf die Lernvoraussetzungen und damit verbundenen Bedürfnisse der Schüler*innen, die höhere Flexibilität im Unterrichten sowie die daraus gestiegenen (fach-)didaktischen Anforderungen und dafür notwendigen neuen Formen des Lehrens und Lernens macht es aus konstruktivistischer Perspektive unausweichlich, bereits in der universitären Ausbildung der Lehrkräfte von morgen diese Zusammenhänge in der Hochschuldidaktik strukturell zu etablieren, um so Lehramtsstudierenden besser zu qualifizieren und auf ihre zukünftigen Aufgaben vorzubereiten.

Vor diesem Hintergrund werden im Projekt koko „kompetent kooperieren“ kooperative und interdisziplinäre Seminare mit Tandemlehre und kollaborativen Lernarrangements unter Studierenden (u.a. Zusammenarbeit in der Gruppe, im Team oder Tandem, Hadwin et al., 2011) in der Lehrerbildung an Universitäten unter besonderer Berücksichtigung folgender lernrelevanter Merkmale der Lehramtsstudierenden und Outcomes evaluiert: Lernmotivationsregulationsstrategien (Boekaerts, 1997), Motivationsziele, epistemologische Überzeugungen (Zinn, 2013), Einstellungen zu Kollaboration, Kompetenzorientierungen, Fachinteresse sowie personale, soziodemographische und studienbezogene struktureller Faktoren.

Auch aus Sicht der Lehrerprofessionsforschung sind diese hochschuldidaktischen Lehr- und Lernmethoden aufschlussreich, da Schulen in ihrer traditionellen Struktur eher als kooperationshemmend wahrgenommen werden (Rothland, 2016; Altrichter & Eder, 2004). Im Zuge des Trends zum fächerübergreifenden Unterrichten und nicht zuletzt inklusions- und integrationsbedingt wird sich der Lehrberuf jedoch weg vom klassischen Einzelkämpfer hin zur Zusammenarbeit in (multiprofessionellen) Teams (Bertels, 2018) entwickeln.

Im Projekt koko wurden solche Seminare bislang in zwei Teilstudien untersucht, davon Teilstudie I mit Schwerpunkt in der Religionsfachdidaktik und -pädagogik hinsichtlich der Förderung religionssensibler und interkultureller Dialog- und Kooperationskompetenzen und Teilstudie II mit Erweiterung auf alle Fachdidaktiken mit Fokus auf die Förderung allgemeiner professioneller Kompetenzen von Lehrkräften (Baumert & Kunter, 2006). Jeweils wurde sowohl ein quantitativer als auch ein qualitativer Forschungszugang gewählt (Mixed-Methods-Ansatz), bestehend aus einer schriftlichen Befragung der Studierenden mit einem standardisierten Untersuchungsinstrument und einer qualitativen Interviewstudie. Es handelt sich hier um eine quasiexperimentelle Interventionsstudie, mit Prä- und Postmessung zu Beginn bzw. am Ende des Seminars (Längsschnitt) sowie Experimental- und Kontrollgruppen mit den Experimentalbedingungen Tandemlehre (Kooperation der Dozierenden), kollaborative Arbeitsformen unter Studierenden und Interdisziplinarität des Seminars (Forschungsdesign). Die Untersuchungen fanden in einem Sommersemester im Zeitraum kurz vor der Corona-Pandemie unter Beteiligung (fast) aller Universitäten in Bayern statt.

Im Rahmen der Posterpräsentation werden ausgewählte quantitative Befunde aus multivariaten Varianzanalysen mit Post-hoc-Analysen je Messzeitpunkt von Teilstudie II vorgestellt, worin 23 Seminare aus verschiedenen an der Lehrerbildung beteiligten Disziplinen an sechs Universitäten in Bayern beforscht und evaluiert wurden (nicht-randomisierte Klumpenstichprobe). Insgesamt haben sich 279 Lehramtsstudierende an den Untersuchungen beteiligt, davon 255 am ersten, 164 am zweiten und 140 am ersten und zweiten Messzeitpunkt (Vollerhebung der Seminarteilnehmenden als Ziel wurde aufgrund von Selbstselektion je nach Seminar und Rahmenbedingungen mehr oder weniger erreicht). Folgende Fragestellung wurde formuliert: Inwieweit können die fokussierten hochschuldidaktischen Konzepte als Formen einer kurzzeitigen Intervention motivationsregulative Fähigkeiten, epistemologische Überzeugungen und Einstellungen zum gruppenbasierten Lernen von Lehramtsstudierenden fördern, die als professionelle Kompetenzen in ihrer späteren Berufspraxis als Lehrkräfte nützlich und relevant sind bzw. darauf vorbereiten. Die Ergebnisse aus Post-hoc-Analysen zeigen im Vergleich, dass sich einzelne Kompetenzbereiche wie selbstregulative Motivationsstrategien, Überzeugungen und Einstellungen zur Gruppenarbeit je nach realisiertem hochschuldidaktischem Konzept im Zuge der Seminarteilnahme signifikant verbessert haben.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 66: Professionalisierung von Grundschullehramtsstudierenden im Umgang mit Rassismus

Zebiba Teklay, Daniel Then, Sanna Pohlmann-Rother

Universität Würzburg, Deutschland

Theoretischer Hintergrund

Rassismus ist ein globales Phänomen, mit dem professionell umzugehen das Tätigkeitsprofil von Lehrkräften bedeutend prägt (Rau und Scheunpflug 2022; Taube 2022). Gleichwohl zeigen Forschungsbefunde, dass Lehrkräfte professionelle Defizite im Umgang mit Rassismus aufweisen (Bernstein 2020; Bernstein und Diddens 2022). Folglich besteht Bedarf, Lehrkräfte für rassismuskritisches Handeln im schulischen Kontext zu professionalisieren. Besondere Bedeutung wird in diesem Zusammenhang den Reflexionskompetenzen der Lehrkräfte zugeschrieben, d.h. den Kompetenzen zur systematische Reflexion von Situationen im Berufsalltag, die eine kritische Distanzierung zu Alltagsroutinen zulassen (Wehner und Weber 2018). Reflexionskompetenzen ermöglichen Lehrkräften grundsätzlich selbstgesteuerte Professionalisierungsprozesse (Aufschnaiter et al. 2019) und gelten im Kontext von Rassismus als „Anker“ rassismuskritischer Professionalisierung (Mecheril und Rangger 2022; Ohm und Fashtomi 2024), die sich aus den Facetten rassismuselevanter Kompetenzen – Fach- bzw. wissenschaftliches Wissen, selbstbezüglich-biographisches Wissen, kasuistisches, reflexives Fallwissen sowie Handlungs- und Erfahrungswissen (Weis 2017) – ableiten lassen. Da in der Grundschule durch die unausgelesene Schülerschaft eine Vielzahl ethnischer Hintergründe zu finden sind und bereits Grundschulkinder Rassismuserfahrungen berichten (Nguyen 2013), stehen schon Grundschullehrkräfte vor der Aufgabe, adäquat mit Rassismus in der Schule umzugehen. Obgleich Studien die Reflexionskompetenz von angehenden Lehrpersonen als ausbaufähig beschreiben (Wyss und Mahler 2021), existieren bislang kaum Professionalisierungsangebote, welche die Reflexionskompetenzen angehender Grundschullehrkräfte im Kontext von Rassismus fördern und wissenschaftlich begleitet werden. Mit der vorliegenden Studie wird auf das Desiderat reagiert.

Fragestellung

Mit dem Anspruch, einen Beitrag zur rassismuskritischen Professionalisierung von Studierenden im Lehramt an Grundschulen zu leisten, wurde an der Universität Würzburg ein Seminarkonzept zur Professionalisierung angehender Grundschullehrkräfte für den Umgang mit Rassismus entwickelt. Die wissenschaftliche Begleitung im Sommersemester 2024 und Wintersemester 2024/2025 zielt auf die Frage, welche Reflexionskompetenzen angehende Grundschullehrkräfte im Umgang mit Rassismus nach dem Besuch dieses Lehrangebots zeigen. Ziel des Vortrags ist es, das Seminarkonzept sowie erste Ergebnisse der Seminarevaluation darzustellen.

Methode

Die wissenschaftliche Begleitung umfasst das Sommersemester 2024 und das Wintersemester 2024/2025. Im Sommersemester 2024 wurden auf Basis einer vorhergehenden Studie mit Lehrkräften zu rassismuskritischem Handeln (Teklay i.V.) Textvignetten konzipiert (Benz 2021), die von den Studierenden im Anschluss an das Seminar schriftlich bearbeitet wurden. Die Vignette umfasst eine Unterrichtssituation, in der rassismuskritisches Handeln nötig wird, und dient als Gesprächsimpuls für das folgende leitfadengestützte Interview (Reinders 2022).

Im Wintersemester 2024/2025 ist geplant, ein Lerntagebuch einzusetzen, in welchem Studierende seminarbegleitend die eigenen Lernfortschritte aus eigener Perspektive entlang von offenen Impulsfragen dokumentieren. Am Ende des Semesters kommen abermals Textvignetten zum Einsatz, die von Studierenden schriftlich bearbeitet werden.

Die Datenauswertung erfolgt mit der inhaltlich-strukturierenden qualitativen Inhaltsanalyse (Kuckartz und Rädiker 2022) und bezieht sich dezidiert auf die rassismusrelevanten Kompetenzbereiche, die sich in Anlehnung an Weis (2017) und Aufschnaiter et al. (2019) sowie als Teil der rassismusbezogenen Reflexionskompetenz von Grundschullehrkräften beschreiben lassen (u.a. Handlungs- und Erfahrungswissen zum Umgang mit Rassismus). Ausgangspunkt ist das Wissen um die bewusste analytische Selbstreflexion, die es ermöglicht, eigene Konstruktionen aufzudecken und zu verstehen.

Ergebnisse und Diskussion

Erste Evaluationsergebnisse zeigen, dass eine rassismuskritische Lehrkräftebildung unterschiedliche Wissensformen anspricht. So äußern die Studierenden sowohl in Bezug auf Selbstreflexionen (reflexiv-kasuistisches Wissen) als auch in Bezug auf die Wissensaneignung zum Themenkomplex Rassismus (Fach- bzw. wissenschaftliches Wissen) Bedarfe für weitere und ausdifferenzierte Angebote. Insbesondere der Umgang mit Rassismus in Interaktionen mit unterschiedlichen Akteur:innen der Schule wird als sinnvolle Lerngelegenheit beschrieben, die Eingang in weiterführende Professionalisierungsangebote finden sollte. Der Besuch des Seminars scheint bei den Studierenden somit ein Bewusstsein dafür geschaffen zu haben, sich vertieft mit Rassismus in der Schule auseinanderzusetzen. Gleichzeitig verweisen die Evaluationsergebnisse auf die Notwendigkeit der weiterführenden Professionalisierung und Weiterentwicklung des Seminars, um angehende Lehrkräfte auf den Umgang mit Rassismus in der Schule vorzubereiten. Das Seminar bildet somit einen ersten, doch nicht hinreichenden Baustein für die rassismuskritische Professionalisierung von Grundschullehrkräften in der ersten Phase der Lehrkräftebildung.

ENTFÄLLT Pinnwand 67: Das Orientierungspraktikum als kritische Phase: Einflüsse auf Selbstwirksamkeit und Studienabbruchneigungen bei Lehramtsstudierenden

Denise Reisch, Robert Grassinger

PH Weingarten, Deutschland

Studien zeigen, dass Studienabbrüche im Lehramtsstudium vorwiegend zu Studienbeginn auftreten (Heublein et al., 2017). Das Orientierungspraktikum stellt per Definition eine vulnerable Phase für Studienabbruch dar, die mitunter durch einen Ist-Soll-Vergleich vormaliger Erwartungen mit der Berufsrealität vermehrt Zweifel an ihrer Berufseignung hervorrufen kann (Grassinger & Schnebel, 2024). Diese Zweifel wiederum sind Indikatoren einer Wahrnehmung geringer Passung eigener Interessen, Erwartungen oder Fähigkeiten zu den Inhalten und Anforderungen des Studiums, was der ersten Phase des Studienabbruchprozessmodells nach Bälke et al. (2021) entspricht. Zugleich ist wenig darüber bekannt, inwiefern das Orientierungspraktikum mit einer aufkommenden Studienabbruchintention korrespondiert oder die Studienentscheidung bestärkt und damit ihre Funktion erfüllt (Kaufer, 2018; Rother et al., 2022; Rothland & Straub, 2018; Weyland, 2012). Hingegen

bestehen Studien, die von einer Studienabbruchintention bereits nach dem ersten Semester und vor dem Orientierungspraktikum berichten (Dresel & Grassinger, 2013).

Zudem liegen mit Blick auf Praktika wenig Kenntnisse darüber vor, was eine aufkommende Studienabbruchintention begünstigt. Theoretisch erscheint plausibel, dass Erfahrungen im Praktikum, die an der eigenen Eignung für die beruflichen Anforderungen zweifeln lassen, eine Studienabbruchintention begünstigen. Korrespondierend damit verweisen erste Befunde von Pfitzner-Eden (2014) darauf, dass geringes Selbstwirksamkeitserleben in Praxisphasen mit einer erhöhten Studienabbruchintention einhergeht.

Es wird argumentiert, die Wahrnehmung einer geringen Selbstwirksamkeit auch als Hinweis auf individuelle Entwicklungspotenziale spezifischer Fähigkeiten verstanden werden, die es im Studium zu entwickeln gilt. Folglich erscheint theoretisch plausibel, dass eine geringe Selbstwirksamkeit nicht per se mit aufkommender Studienabbruchintention einhergeht, sondern im Besonderen dann, wenn Praktikant:innen diese vorwiegend als Hinweis auf mangelnde stabile Fähigkeiten statt auf zu entwickelnde Fähigkeiten verstehen. Dies tun sie per Definition dann, wenn sie im Orientierungspraktikum vorwiegend Leistungsziele statt Lernziele verfolgen (Elliot & Harackiewicz, 1996; Nicholls, 1984).

Die Forschungsdesiderata adressierend und die Argumentation aufgreifend werden folgende Hypothesen geprüft: Die Wahrnehmung geringer Passung zum Studium als erste Phase des Studienabbruchprozesses nach Bäumke et al. (2021) ist im Orientierungspraktikum ausgeprägter als gegen Ende des ersten Semesters (H1). Je geringer die Lehrerselbstwirksamkeit im Orientierungspraktikum, desto stärker nehmen Lehramtsstudierende eine geringe Passung zum Studium wahr (H2). Der Zusammenhang zwischen Lehrerselbstwirksamkeit im Orientierungspraktikum und der Wahrnehmung geringer Passung zum Studium ist stärker ausgeprägt, wenn im Orientierungspraktikum Leistungsziele verfolgt werden (H3a) und schwächer ausgeprägt, wenn im Orientierungspraktikum Lernziele verfolgt werden (H3b).

Zur Überprüfung dieser Hypothesen wurden $N = 203$ Lehramtsstudierende vor und nach dem Orientierungspraktikum befragt. Im Schnitt waren die Studierenden 21.23 Jahre alt ($SD = 4.22$). 86% der Studierenden waren weiblich. Die Hypothesen wurden mittels t-Test für abhängige Stichproben (H1) sowie Regressions- (H2) und Moderationsanalysen (H3) geprüft.

Die Ergebnisse bekräftigen die angenommenen Zusammenhänge zwischen dem Selbstwirksamkeitserleben und der Wahrnehmung geringer Passung im Orientierungspraktikum und der Moderation durch Lern- und Leistungsziele.

Korrespondierend mit den Befunden scheint es bedeutsam zu sein, das Orientierungspraktikum über eine Lernzielstruktur zu gestalten, so dass selbstwirksamkeitsmindernde Erfahrungen verstärkt als Hinweis auf weitere Entwicklungspotenziale erlebt werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 35: Anbahnung informatischer Kompetenzen von Grundschullehramtsstudierenden – Gestaltungsmerkmale eines universitären Theorie-Praxis-Seminars

Holger Wilhelm, Tina Jocham
Universität Würzburg, Deutschland

Theoretischer Hintergrund/Fragestellung

Digitalisierungsbezogene Kompetenzen sind zum Verständnis der „Infosphäre“ (Floridi, 2015) unumgänglich. Ein wesentlicher Teilbereich davon sind informatische Kompetenzen, zu denen „algorithmisches Denken, Data Literacy, Computational Thinking und Datensicherheit“ zählen (van Ackeren et al., 2019, S. 108). Vor dem Hintergrund einer grundlegenden Bildung sind informatische Kompetenzen bereits in der Grundschule anzubahnen, wofür gut ausgebildete Lehrkräfte benötigt werden (Kultusministerkonferenz (KMK), 2016, 2021). Jedoch fehlt bisher eine flächendeckende, verpflichtende Integration digitalisierungsbezogener Kompetenzen in der Lehrkräftebildung (SWK, 2021).

Das integrative Modell von Beißwanger und Kolleg:innen (2020) vereint verschiedene theoretische Überlegungen im Hinblick auf digitalisierungsbezogene Kompetenzen in der Lehrkräftebildung und bildet u.a. eine „konzeptionelle Grundlage für ein fächerübergreifendes Lehrveranstaltungsangebot“ (ebd., S. 12). Da die Vermittlung informatischer Kompetenzen in der Grundschule eine querliegende Aufgabe darstellt, eignet sich das Modell als theoretische Grundlage für unser Projekt. Die Seminarumsetzung erfolgt in digitalen Lehr-Lernlaboren, die den Erwerb elementarer Informatikkenntnisse bei Studierenden unterstützen, da diese primär auf die Vermittlung digitalisierungsbezogener Kompetenzen abzielen (Grafe et al., 2023). Zudem bieten Theorie-Praxis-Seminare eine Möglichkeit, digitalisierungsbezogene Kompetenzen zu fördern (Arndt et al., 2021; Kuckuck et al., 2021; Tondeur et al., 2017; Tondeur et al., 2011). Das konzipierte Seminar lässt sich deshalb im Diskurs der Theorie-Praxis-Verzahnung verorten (Bulizek et al., 2020) und fokussiert folgende Fragestellung:

Was sind Gelingensbedingungen für die Gestaltung universitärer Seminare, um Grundschullehramtsstudierende zu befähigen, informatische Projekte zu planen, umzusetzen und kritisch zu reflektieren?

Methodisches Vorgehen

Seminarkonzept/Weiterentwicklung

Das Seminarkonzept wurde vor dem Hintergrund des integrativen Modells (Beißwanger et al., 2020) interdisziplinär als Kooperation zwischen Sonderpädagogik und Grundschulpädagogik konzipiert. In drei Durchgängen wurde das Seminar auf Basis von schriftlichen Reflexionsfragen der Studierenden mithilfe von Design-Based Research (Reeves, 2000) weiterentwickelt (n=48). Ziel der Kooperation von Sonderpädagogik und Grundschulpädagogik ist es, Studierenden spezifisches didaktisches Wissen (z.B. Differenzierung/Individualisierung) an die Hand zu geben, um Kindern mit unterschiedlichen Lernbedarfen gerecht zu werden. Im Mittelpunkt des Seminars stand das Thema „Algorithmen und Programmieren im Kontext des Sachunterrichts“ (siehe Kompetenzbereiche in Baberowski et al., 2023, S. 44; GDSU, 2021; van Ackeren et al., 2019). Auf dieser Grundlage planten die Studierenden eine Unterrichtseinheit und erprobten diese mit Grundschulkindern in den digitalen Lehr-Lernlaboren (DigiLLabs) der Universität Würzburg.

Auswertung

Um konkrete Gestaltungsmerkmale zu identifizieren, wurden die Reflexionsberichte der Studierenden aller Durchgänge mit Blick auf kritische und gelungene Seminarpraktiken (z.B. inhaltlich, methodisch, didaktisch) mittels der strukturierenden qualitativen Inhaltsanalyse ausgewertet (Kuckartz & Rädiker, 2022). Die induktiv-deduktiv entwickelten Kategorien wurden konsensuell codiert und ausdifferenziert.

Ergebnisse und ihre Bedeutung

Die qualitative Auswertung ergab verschiedene Gestaltungskriterien, die von den Studierenden unterschiedlich gewichtet wurden. Positiv hoben sie insbesondere die Möglichkeit hervor, eigene Projekte zu planen und mit Grundschulkindern umzusetzen. Intensiv betreute Theorie-Praxis-Seminare bieten dabei die Chance von unbenoteten Erprobungsräumen, in denen Selbstwirksamkeitserfahrungen gesammelt und Befürchtungen von Studierenden abgebaut werden können (Brämer et al., 2020).

In der Vermittlung des informatischen Fachwissens zeigte sich der Einbezug fachwissenschaftlich fundierter Plattformen (z.B. KI Campus von Schmid et al., 2022) und eine Konkretisierung der Inhalte anhand von analogen und digitalen Lehr-Lernmedien als bedeutsam. Insgesamt betonten die Studierenden v.a. die Relevanz allgemeindidaktischer Kenntnisse (z.B. im Hinblick auf Differenzierung, Strukturierung) in Verbindung mit fachdidaktischem Wissen (z.B. sachunterrichtliche Lerngegenstände). Beide Bereiche stellten die Studierenden vor Herausforderungen.

Aus hochschuldidaktischer Sicht lässt sich schlussfolgern, dass nicht nur die Vermittlung von informatischem Fachwissen zentral ist, sondern auch eine Einbettung des Wissens in methodisch-didaktische Überlegungen (Beißwanger et al., 2020), sodass Studierende befähigt werden, unterschiedlichen Lernbedürfnissen von Kindern gerecht zu werden (Capovilla, 2019). Die Vermittlung von informatischem Fachwissen anhand von konkreten fachbezogenen Gegenständen, ist im Sinne der Querschnittsaufgabe der Grundschule zentral (z.B. Gesellschaft für Informatik e.V. [GI], 2019).

Inwiefern die konzipierten Lernangebote der Studierenden auf die individuellen Lernbedürfnisse der Kinder zugeschnitten sind, soll zukünftig mittels teilnehmender Beobachtungen erhoben werden.

Pinnwand 39: ViCoach – Ein videobasiertes Coaching-Angebot zur Vorbereitung von Studierenden auf mündliche Prüfungssituationen

Miriam Hess
Universität Bamberg, Deutschland

Neben fachlichen, fachdidaktischen und pädagogischen Kompetenzen benötigen Lehrpersonen gerade für Gespräche mit Eltern, Kolleg:innen und weiteren Personen im beruflichen Umfeld auch Argumentationskompetenzen, beispielsweise, um eigene pädagogische Ansätze fachlich sowie verständlich zu begründen (Perschak, 2022). Eine Möglichkeit, dies bereits im Studium zu adressieren, stellen mündliche Prüfungen dar, die von Studierenden allerdings häufig als sehr herausfordernd wahrgenommen

werden. Dies ist ggf. auch darauf zurückzuführen, dass eine direkte Förderung argumentativer Gesprächsführungskompetenz (Gartmeier, 2018, Grundler, 2011) im Rahmen der universitären Lehrkräftebildung bisher nur selten explizit stattfindet. Interviews mit Lehrenden im Bereich der Grundschulpädagogik und -didaktik zeigen zudem, dass die Performanz der Studierenden in mündlichen Prüfungen in den letzten Jahren abnimmt (Dotzel, 2023; Hänisch, 2023).

Videobasierte Coachings bringen grundlegend eine Vielzahl von Chancen zur Entwicklung eigener Kompetenzen mit sich (z.B. Hess, 2019) und dürften sich gerade zur Förderung von Kompetenzen für mündliche Prüfungen besonders eignen, da es sich hier um direkte Interaktionssituationen handelt, deren prozessualer Charakter in Videos besonders deutlich werden dürfte. Mit Videos können zentrale Gelingensbedingungen mündlicher Prüfungen direkt sichtbar gemacht, Reflexionsprozesse ausgelöst und begleitet (Hagenauer & Hascher, 2018) sowie zum Abbau von Prüfungsangst beigetragen werden (Lotz & Sparfeldt, 2017). Zur Wirksamkeit videobasierter Coachings auf den Erfolg in mündlichen Prüfungssituationen existieren bisher aber noch keine Studien.

Daher zielt das Projekt ViCoach darauf ab, ein videobasiertes Coaching für mündliche Prüfungssituationen sowie Instrumente zur systematischen Analyse der Wirksamkeit dieses Coachings zu entwickeln. Das Coaching-Angebot besteht aus zwei Teilen: Zunächst analysieren die Studierenden in einer asynchronen Online-Lernumgebung verschiedene, systematisch aufbereitete Videos von gescripteten mündlichen Prüfungssituationen und erhalten dazu auch direktes, vorbereitetes Feedback der Dozierenden. Im Anschluss werden in einer Präsenzveranstaltung offene Fragen besprochen und mündliche Prüfungssituationen jeweils in Dreiergruppen (Prüfer:in, Student:in, Beobachter:in) geübt und reflektiert.

Die Erprobung des Coachings findet seit mehreren Semestern statt, indem das Coaching allen Studierenden des Grundschullehramts an der Universität Bamberg zur Vorbereitung auf die Staatsexamensprüfung angeboten wird. Zur Evaluation des Angebots schätzen die Studierenden sowohl vor als auch nach dem Coaching ihre eigenen Kompetenzen sowie deren Weiterentwicklung durch das Coaching ein. Nach der Prüfung wird die Prüfungsnote erfasst und die Prüfer:innen schätzen die Performanz der Studierenden mittels eines Kurzfragebogens in verschiedenen Bereichen ein. Dabei werden alle Prüflinge befragt, auch diejenigen, die nicht am Coaching teilgenommen haben.

Folgende Fragen werden auf dem Poster beantwortet: Wie wird das Coaching-Angebot von den Studierenden angenommen? Wie beurteilen die Studierenden 1) die videobasierte Online-Lernumgebung, 2) die Übungen in Präsenz und 3) die Kombination aus beiden Teilen? Zeigen sich Zusammenhänge zwischen der Teilnahme am Coaching und dem Ergebnis bei der mündlichen Prüfung in Form der erhaltenen Note sowie der Einschätzung der Performanz in der Prüfung durch die Studierenden selbst sowie die Prüfer:innen?

Die Ergebnisse zeigen, dass etwa die Hälfte der Studierenden das Coaching-Angebot in Anspruch nimmt. Es wird mehrheitlich sehr positiv durch die teilnehmenden Studierenden beurteilt und es zeigen sich positive Zusammenhänge zwischen der Teilnahme und den Prüfungsnoten.

Diskutiert werden kann, inwiefern ein Einsatz der Videos und des Coaching-Konzepts auch für weitere Universitäten in der Lehrkräftebildung interessant wären.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 68: Qualität der Berufsorientierung – individuelle Lernprozesse und institutionelle Verankerung in beruflichen Schulen (QUIS)

Susan Seeber¹, Stephan Schumann², Robin Busse³, Elisabeth Maué¹

¹Georg-August-Universität Göttingen, Deutschland; ²Universität Konstanz, Deutschland; ³Technische Universität Darmstadt, Deutschland

Trotz Fachkräftemangels mündet jährlich ca. eine Viertel Million Jugendliche in berufsvorbereitende Bildungsgänge des Übergangssektors (BIBB 2023). Die Berufsorientierung ist dabei in berufsvorbereitenden Bildungsgängen ein zentraler Baustein zur Förderung der Berufswahlkompetenz (Nickolaus et al. 2018). Zugleich werden die Übergangs- und Passungsprobleme beim Ausbildungszugang (Seeber et al. 2019) durch berufsvorbereitende Bildungsgänge nur teilweise „gelöst“: Etwa der Hälfte der Teilnehmenden gelingt der Übergang in Ausbildung, die anderen besuchen erneut einen berufsvorbereitenden Bildungsgang, arbeiten ungelernt oder sind arbeitslos (Menze/Holtmann 2019; Michaelis et al. 2022). Daher sind Fragen der Effektivität berufsvorbereitender Bildungsgänge und damit auch der Qualität von Berufsorientierung und deren Umsetzung durch die berufsbildenden Schulen in Zusammenarbeit mit regionalen Akteuren bedeutsam. Trotz hoher Bedeutung der Berufsorientierung für die individuellen Entwicklungswege junger Menschen (ebd.) liegt kaum empirische Evidenz über deren Gestaltung und Ergebnisse vor. Die berufsbildenden Schulen sind bei der Umsetzung einer qualitativ hochwertigen Berufsorientierung angesichts anspruchsvoller Ziele, unzureichender und inkohärenter curriculärer Einbettung, unklarer Verankerung in der Schulentwicklung (inkl. Personalentwicklung) sowie schwieriger Koordinierung vielfältiger Akteure mit diversen Interessen (u.a. Betriebe, Jobcenter, private Träger, regionale Netzwerke) mit multiplen Herausforderungen konfrontiert (Ohlemann 2021).

Vor diesem Hintergrund besteht das Ziel des Projekts „QUIS“ darin, a) die Qualität von Berufsorientierung in berufsvorbereitenden Bildungsgängen unter Einbezug der Mehrebenenstruktur des Bildungssystems (Fend 2008) und multipler Akteure zu analysieren sowie b) Implikationen für deren Weiterentwicklung für Bildungsadministration und -praxis abzuleiten.

Forschungsfragen:

- (1) Welche Qualitäts- und Zielvorstellungen von Berufsorientierung haben schulische und externe Akteure?
- (2) Welche Indikatoren sind für die Qualitätssteuerung von Berufsorientierung aus Sicht von Wissenschaft und Praxis heranzuziehen?
- (3) Wie ist die Berufsorientierung in den berufsbildenden Schulen und der Schulentwicklung verankert, wie wird sie umgesetzt (Durchführungs- und Ergebnisqualität), wie erfolgt die Qualitätssicherung?
- (4) Welche Transferkonzepte, Leitlinien und Handlungsempfehlungen lassen sich zur systematischen Qualitätsentwicklung der Berufsorientierung in berufsvorbereitenden Bildungsgängen ableiten?

Der Einbezug der Bundesländer Niedersachsen, Baden-Württemberg und Hessen mit deren teils unterschiedlicher Wirtschaftsstruktur und konjunktureller Lage sowie einer differenten Ausgestaltung der Übergangsmaßnahmen (Seeber et al. 2019) ermöglicht die Untersuchung des Einflusses bildungssystemischer, institutioneller, sozial-räumlicher und wirtschaftlicher Rahmenbedingungen auf Berufsorientierungsprozesse (Butz 2008; Ratschinski 2012). In Workshops mit zentralen Akteuren der beruflichen Bildung (z.B. berufsbildende Schulen, Jugendberufsagenturen) werden Umsetzungen, Problemlagen und Qualitätsindikatoren der Berufsorientierung sowie Transferdimensionen aus Sicht von Praxis und Wissenschaft erarbeitet. Unter Einbezug dieser Expertise wird eine mixed-methods Längsschnittstudie zur Qualität der Berufsorientierung in beruflichen Schulen entwickelt. Sie beinhaltet u.a. Qualitätsindikatoren zu Ausgangs- und Rahmenbedingungen, Angebots- und Nutzungsqualität sowie Resultate der Berufsorientierung (z.B. Berufswahlkompetenz, realisierte Übergänge in Ausbildung). Die Längsschnittstudie umfasst mehrere Messzeitpunkte (Schuljahresbeginn 2025/2026, Prozessbefragungen im Schuljahresverlauf, Schuljahresende, Verbleibsbefragungen nach Ende der Bildungsgänge). Die mehrmaligen Befragungen der Schüler:innen und Lehrpersonen ergänzt eine Kurzbefragung der Schul-/Abteilungsleitung zu Schuljahresbeginn. Nach der Durchführung der Längsschnittstudie sowie der Evaluation der Qualität der Berufsorientierung und ihrer Wirkungen folgen die Rückbindung, der Transfer und die Dissemination der Projektergebnisse (z.B. OER).

Anknüpfend an bestehende Qualitätsmodelle und Qualitätsindikatoren für die Prozessgestaltung und Instrumente der Berufsorientierung werden diese kritisch reflektiert. Gegenstand sind u.a. die zugrundeliegenden Perspektiven auf junge Menschen im Kontext von Berufsorientierungsprozessen (Subjektorientierung, „Agency“ junger Menschen im Berufsorientierungskontext; Eteläpelto et al. 2013) sowie die Prüfung der Validität und Evidenz bestehender Qualitätsindikatoren. Ziel ist die Erarbeitung theoretisch und empirisch begründeter Qualitätsindikatoren der Berufsorientierung für berufsvorbereitende Bildungsgänge zur Qualitätsentwicklung für Bildungsadministration und -praxis. In Anlehnung an Modelle von Schul- und Unterrichtsqualität (Ditton 2000; Fend 2008; Helmke 2004) werden mittels einer durchgängig engen Verzahnung von Wissenschaft und Praxis die Input-, Prozess-, Output- und Outcome-Qualität der Berufsorientierung aus Sicht zentraler Akteure mehrbenenanalytisch herausgearbeitet sowie in Rückkopplung mit der Praxis Strategien, Konzepte, Instrumente und Materialien für die Qualitätsentwicklung bereitgestellt.

An der Tagung sollen neben dem Projektdesign erste Erkenntnisse der bundeslandspezifischen Workshops mit Praxispartner:innen zu Umsetzungen, Problemlagen und Qualitätsindikatoren der Berufsorientierung präsentiert werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

(Vertretung) Pinnwand 25: Evaluation kritischen Denkens – Performanz und Denkprozesse von Studierenden bei der Nutzung multipler Online-Quellen

Katharina Frank¹, Olga Zlatkin-Troitschanskaia¹, Dominik Braunheim¹, Marie-Theres Nagel¹, Richard J. Shavelson²

¹Johannes Gutenberg-Universität Mainz, Deutschland; ²Stanford Graduate School of Education, United States

Theoretischer Hintergrund. Beim Lernen mit frei zugänglichen digitalen Ressourcen werden Studierende mit der Herausforderung konfrontiert, indem sie oft auf unzuverlässige, ungenaue oder falsche Informationen stoßen[1,2]. Die Kompetenz mit teilweise widersprüchlichen Quellen kritisch umzugehen, sie zu bewerten und anhand dieser zu einem begründeten Schluss zu kommen (*Critical Thinking*, CT), ist für Erfolg im Studium im Internet-Zeitalter unabdingbar[3]. Eine universitäre Förderung von CT ist daher unerlässlich, insbesondere da Studien zu Folge Studierende diese Kompetenz zu Beginn ihres Studiums oft nicht hinreichend aufweisen[4,5,6]. Trotz der hohen Relevanz erfolgt eine universitäre Förderung von CT oft nur begrenzt und ist wenig explizit curricular verankert[7,8]. Eine Voraussetzung der effektiven Förderung von CT im Studium ist die valide Erfassung der Kompetenz, ihrer zugrundeliegenden Facetten und Denkprozesse, wobei u.a. die Komplexität der Entwicklung eines validen CT-Assessments zu berücksichtigen ist[9,10,11,12,13].

Fragestellung. Um CT von Studierenden valide zu messen und Einblicke in Performanz und Denkprozesse zu erhalten, wurde in einem internationalen Konsortium vier Ländern (Deutschland, Kolumbien, Schweiz, USA) ein holistisches Szenario-basiertes Performance Assessment (PA) mit einer realitätsnahen Simulation einer Entscheidungssituation (staatliche Regulation von Migration) entwickelt[8,10,11]. Das PA enthält vorselektierte (un-)relevante, (un-)zuverlässige und teilweise widersprüchliche Quellen/Informationen, die den Studierenden bei ihrer Aufgabenbearbeitung und Entscheidung zur Verfügung stehen. In vier Ländern wurden *Cognitive Labs* (CogLab) mit je zehn Studierenden nach gemeinsamem Protokoll durchgeführt. Daten wurden durch Antworten der Studierenden auf das PA als CT-Performanz (erstellte Stellungnahmen) sowie Antwortprozesse durch *Think Alouds* (TA) als CT-Denkprozesse erfasst[14]. Dieser Beitrag fokussiert auf die Forschungsfrage, wie sich das CT in der Performanz der Studierenden (resultat-orientiert) im Vergleich zu den Denkprozessen beim Lösen des PA (prozess-orientiert) manifestiert - u.a. im Hinblick auf die Auswahl und Nutzung von Quellen unter Berücksichtigung deren Relevanz und Glaubwürdigkeit?

Methode. Nach einer Validierungsstudie im SS 2023 umfasste die CogLab-Studie im Wintersemester 2023/2024 Erhebungen an den vier Standorten mit jeweils $N=10$ Studierenden, am Ende ihres Bachelor- bzw. zu Beginn ihres Masterstudiums. Um die CT-Performanz zu bestimmen, wurden die schriftlichen Antworten zum PA anhand eines harmonisierten Kriterien-gestützten Scoring-Schema[15,16] mit 5-Punkte-Likertskala zu vier CT-Facetten (Analysieren und Bewerten der Informationen; Erkennen und Abwägen der Konsequenzen; Formulieren und Kommunizieren eines kohärenten Arguments; Perspektivenübernahme:[10]) mit 13 Subdimensionen von zwei trainierten Ratern (starke ICC[17]; $\text{Mean}=75$) ausgewertet. Um Einblicke in die Denkprozesse zu erhalten, wurden die transkribierten TA aus den CogLabs qualitativ kriterienbasiert analysiert.

Ergebnisse. Durchschnittlich bearbeiteten die Studierenden das PA in 114.28 Minuten, schrieben 555.40 Wörter in ihren Antworten zum PA und die transkribierten TA umfassen 8555.30 Wörter.

Die Ergebnisse der CT-Performanz aus den CogLabs deuten auf unzureichende CT-Kompetenzen der Studierenden hin ($\text{Mean}=2.827, \text{SD}=.773$) mit einer starken Varianz in allen CT-Facetten, insbesondere in der Facette der Analyse und Bewertung der Informationen, mit der Anzahl der verwendeten Quellen ($\text{Mean}=3.750, \text{SD}=.920$) und der korrekten Bewertung der Qualität von Informationen ($\text{Mean}=3.750, \text{SD}=.950$).

Die ersten Analysen von transkribierten TA zeigen Unterschiede zu den Performanzergebnissen im Hinblick auf die Selektion und Nutzung von Quellen, u.a. Anzahl der genutzten Quellen in Performanz (schriftliche Stellungnahmen) ($\text{Mean}=4.500, \text{SD}=.970$) und Prozessen (TA) ($\text{Mean}=8.000, \text{SD}=.000$).

Diskussion. Diese Ergebnisse stehen im Einklang mit früherer Forschung, die auf unzureichende CT-Kompetenz im Studium hinweist[7,18]. Um weitere Einblicke in Unterschiede zwischen den Denkprozessen bei der PA-Bearbeitung und der PA-Performanz (schriftlichen Stellungnahme) zu erhalten, werden im Beitrag weitere zentrale Ergebnisse der qualitativen Inhaltsanalysen der durch die TA erfassten Denkprozesse (z.B. im Hinblick auf die Arbeit mit Quellen unter Berücksichtigung von Relevanz und Glaubwürdigkeit) bei der PA-Bearbeitung vorgestellt, die wichtige Hinweise darauf liefern können, welche Faktoren das Verhalten der Studierenden bei der Selektion, Verarbeitung und Nutzung von Informationen beeinflussen. Diese Erkenntnisse tragen zur Weiterentwicklung der Forschung zur validen CT-Erfassung in internationalen und interdisziplinären Studien bei[19,20] sowie liefern erste wichtige Hinweise für die universitäre Förderung der Schlüsselkompetenz CT.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 69: Unloved and Depraved: Welchen Einfluss haben das Familienselbstkonzept und die individuelle Relative Deprivation auf radikale Überzeugungen von Schüler/innen?

Gülru Horozoglu¹, Ricarda Steinmayr², Ursula Kessels¹

¹Freie Universität Berlin, Germany; ²Technische Universität Dortmund, Deutschland

In der aktuellen Stellungnahme der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz (Köller et al., 2024) wird Demokratiebildung als zentraler, auch fachübergreifender Auftrag der Schulen thematisiert. Hierbei werden antidemokratische Einstellungen von Kindern und Jugendlichen, die sich politisch oder religiös radikalisierten, als besondere Herausforderung für die Entwicklung einer demokratischen Schulkultur betont (Köller et al., 2024, S. 24). Im vorliegenden Beitrag werden Faktoren untersucht, die mit einer solchen Radikalisierung im Zusammenhang stehen könnten.

Theorien zur Radikalisierung nennen verschiedene Risikofaktoren, die als Vorstufen zu gewalttätigem Extremismus gelten. Dazu gehören bestimmte Überzeugungen wie die Überlegenheit der eigenen In-group, die Illegitimität von staatlichen Autoritäten und die fehlende soziale Eingebundenheit in die Mehrheitsgesellschaft (im Folgenden insgesamt als „radikale Überzeugungen“ bezeichnet). Es hat sich gezeigt, dass diese Faktoren zur Akzeptanz gewalttätiger Handlungen anderer und zur eigenen Absicht, Gewalt auszuüben, beitragen (Doosje et al., 2013). Gleichzeitig kann die Überzeugung, benachteiligt und/oder unbedeutend zu sein, Jugendliche für radikale Überzeugungen anfällig machen und dazu führen, radikalisierte Gruppen aufzusuchen (Jasko et al., 2017). Darüber hinaus kann eine Migrationsgeschichte zu einem internen kulturellen Konflikt führen, weshalb Jugendliche mit Migrationshintergrund als stärker gefährdet gelten (Ohls et al., 2023) und häufiger Gegenstand der Radikalisierungsforschung sind.

Es gibt mehrere vorgelagerte Faktoren, die diese radikalen Überzeugungen beeinflussen können. Das Ausmaß, in dem sich eine Person im Vergleich zu anderen als ressourcenarm empfindet, kann diese Überzeugungen beeinflussen (Doosje et al., 2013). Im Gegensatz dazu kann eine gute Beziehung zu den Eltern als Schutzfaktor gegen Radikalisierung wirken (Lösel et al., 2018). Während positive familiäre Faktoren nachweislich vor extremistischen Einstellungen und extremistischer Gewalt schützen, gibt es keine Untersuchungen darüber, wie familiäre Faktoren mit den oben genannten radikalen Überzeugungen kovariieren, die möglicherweise gewalttätigen Extremismus begünstigen. Daher wurde in der folgenden Studie der Zusammenhang zwischen dem familiären Selbstkonzept (FSK), d. h. dem Gefühl, von Familienmitgliedern respektiert und geschätzt zu werden, und der wahrgenommenen relativen individuellen Deprivation (IRD) mit radikalen Überzeugungen bei Jugendlichen mit Migrationshintergrund (JMM) und ohne Migrationshintergrund (JOM) untersucht.

Die Stichprobe bestand aus 221 Zehntklässler*innen, davon 82 JMM und 139 JOM. Für die aktuelle Studie wurden die Items für das FSK (Schwanzer et al., 2005), IRD (Doosje et al., 2013), radikale Überzeugungen (Doosje et al., 2013) analysiert. Wir führten ein Multigruppen-Strukturgleichungsmodell (SEM) durch, um die Zusammenhänge zwischen FSK, IRD und radikalen Überzeugungen getrennt für JMM und JOM vergleichend zu analysieren. Alle Analysen wurden mit dem lavaan-Paket in R durchgeführt.

Mit Ausnahme der wahrgenommenen Illegitimität der Autoritäten gab es bei allen Variablen signifikante Mittelwertunterschiede zwischen Jugendlichen mit und ohne Migrationshintergrund, wobei die Mittelwerte der Jugendlichen mit Migrationshintergrund höher (bzw. niedriger für das familiäre Selbstkonzept) ausfielen. Allerdings waren die Mittelwerte für radikale Überzeugungen und individuelle relative Deprivation in beiden Gruppen insgesamt niedrig (bzw. hoch für das Familienselbstkonzept). Das SEM wurde geschätzt ($\chi^2(310) = 506.69$, $p < .001$, CFI = .96, TLI = .95, RMSEA = .076, SRMR = .08). Ein höheres FSK war mit niedrigerer wahrgenommener Illegitimität von Autoritäten verbunden, und dieser Zusammenhang war bei Jugendlichen mit Migrationshintergrund stärker ausgeprägt. Höhere FSK-Werte waren auch mit einem geringeren Gefühl der sozialen Uneingebundenheit verbunden, allerdings stärker bei Jugendlichen ohne Migrationshintergrund. Es fanden sich keine signifikanten Zusammenhänge zwischen dem FSK und der empfundenen Überlegenheit der eigenen religiösen Gruppe. Im Gegensatz zu den theoretischen Annahmen über die Bedeutung des Migrationshintergrunds war ein hohes Maß an IRD sowohl bei Jugendlichen mit als auch ohne Migrationshintergrund mit höheren Werten für die wahrgenommene Illegitimität von Autoritäten, sozialer Uneingebundenheit und Überlegenheit der eigenen religiösen Gruppe assoziiert.

Die Ergebnisse werden in Bezug auf frühere Forschungsarbeiten verortet, die mit "gefährdeten" Bevölkerungsgruppen durchgeführt wurden und im Hinblick auf Präventions-/Deradikalisierungsbemühungen und Demokratiebildung diskutiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 6: Förderung der Unterrichtsplanungskompetenz mit Fokus auf inhaltlicher Kohärenz

Marcus Schiolko, Mathias Ropohl
Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund und Fragestellung

Bisherige Studien zur professionellen Kompetenz von angehenden Lehrkräften haben sich primär auf die Unterrichtsdurchführung fokussiert, während die Unterrichtsplanungskompetenz von angehenden Lehrkräften in der empirischen Bildungsforschung wenig betrachtet wurde (König & Rothland, 2022). Allerdings wird der Unterrichtsplanung durch die gedankliche Antizipation des durchzuführenden Unterrichts eine große Wirkung auf die Unterrichtsqualität und den -erfolg zugesprochen (Tebrügge, 2001). Bisher fehlt es an spezifischen Lerngelegenheiten, welche die Unterrichtsplanungskompetenz fördern (König & Rothland, 2022).

Angehenden Lehrkräften gelingt es selten im Planungsprozess Verknüpfungen zwischen verschiedenen Planungsentscheidungen und damit inhaltliche Kohärenz hinsichtlich der Instruktionen herzustellen (Gassmann, 2013). Der Planungsprozess zeichnet sich durch lineare nicht zusammenhängende Entscheidungen aus (Westerman, 1991). In Bezug auf den Fachinhalt zeigt sich in den Unterrichtsplanungen, dass die zentralen Ideen bzw. Konzepte eines Themas selten mit anderen zentralen Ideen verknüpft werden (Park & Chen, 2012). In geplanten Sicherungsphasen liegt der Fokus auf den in der Lerngelegenheit erarbeiteten Ideen und Konzepten, wodurch Lernende kaum vernetztes Wissen aufbauen können (Weitzel & Blank, 2020).

Im amerikanischen Raum wurden visuelle Repräsentationen fachinhaltlicher Konzepte entwickelt, sogenannte Wissenslandkarten. Diese ermöglichen eine visuelle Darstellung zum sachlogischen Aufbau von Konzepten anhand von Wissensbausteinen sowie deren Weiterentwicklung und Fortführung über Schulstufen hinweg (AAAS, 2001, 2007). Die Wissenslandkarten sind zwar inkonsistent zu den nationalen Bildungsstandards, lassen sich aber für naturwissenschaftliche Fächer adaptieren, da diese einerseits anhand von Basiskonzepten organisiert werden und andererseits als Spiralcurriculum aufgebaut sind (Demuth et al., 2005; KMK, 2005, 2020).

Aus der Ausgangslage wird die Forschungsfrage abgeleitet, inwiefern eine spezifische Förderung von Unterrichtsplanungskompetenz die inhaltliche Kohärenz von durch angehende Chemielehrkräfte geplanten Lerngelegenheiten erhöht.

Methode

Um die Forschungsfrage zu beantworten, wurde eine digitale Plattform mit interaktiven Lehr-/Lernmaterialien zu einer Wissenslandkarte erstellt. Diese digitale Plattform bietet für die Lehrkräftebildung das noch ungenutzte Potential die Unterrichtsplanung von angehenden Lehrkräften zu entlasten und inhaltliche Kohärenz zu fördern. In der hier vorgestellten Voruntersuchung wurde analysiert, wie sich die Planungsergebnisse bei Nutzung bzw. Nicht-Nutzung der Plattform unterscheiden. Die $N = 10$ teilnehmenden angehenden Lehrkräfte erstellten eine Unterrichtsplanung mit der digitalen Plattform und anhand selbstgewählter Materialien. Dabei wurden sie videografiert. Die beiden Planungen wurden unmittelbar hintereinander durchgeführt. Dieselben Testpersonen arbeiteten sowohl in der Experimental- als auch in der Kontrollbedingung (quasi-experimentelles Within-Subjects-Design). Darüber hinaus wurde die wahrgenommene Usability der Plattform mit einem Fragebogen erhoben.

Ergebnisse

Um Kohärenz messbar zu machen, wurde ein Kodierleitfaden erstellt. Das resultierende Kohärenzmaß der Unterrichtsplanungen kann maximal 32 Punkte je Planung erreichen. Im Mittel liegt das Maß für Planungen mit der Plattform bei $M = 15,9$ und damit 2,5 Punkte niedriger als für die Planungen ohne die Plattform. Die Unterrichtsplanungen mit dem höchsten Kohärenzmaß sind mithilfe der Plattform erstellt worden. Über alle Unterrichtsplanungen hinweg zeigten sich besonders Schwierigkeiten bei der Erreichung von Kohärenz zwischen den einzelnen Phasen. Als Prädiktoren hatten die fachdidaktischen Noten keine Auswirkungen auf das Kohärenzmaß der Unterrichtsplanung, das absolvierte Praxissemester ($n = 5$) dagegen schon. Mit absolviertem Praxissemester erzielten angehende Lehrkräfte ohne die Plattform signifikant kohärentere Planungsergebnisse ($r = 0.775^{**}$).

Bezüglich der Usability zeigten sich hohe bis sehr hohe Zustimmungswerte zur digitalen Plattform. Der System Usability Score liegt bei einem Wert von 87,0 und kann als „best imaginable“ eingestuft werden (vgl. Lewis & Sauro, 2018). Bedeutsam aus der Rückmeldung zur Usability ist, dass die angehenden Lehrkräfte mehr Eingewöhnungszeit gebraucht hätten, um mit der Plattform besser arbeiten zu können.

Die inhaltliche Kohärenz von angehenden Lehrkräften konnte mit dieser Voruntersuchung nicht wie intendiert gefördert werden. Mögliche Ursachen können Anpassungsvorgänge bei der eigenen Unterrichtsplanung an die Logik der Plattform sein (Marchand, 2017). Aktuell wird sie angepasst und anschließend in einer Interventionsstudie zur Förderung der Kohärenz von Unterrichtsplanungen eingesetzt.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 70: How to foster professional digital competence beliefs in teacher education? A longitudinal investigation

Anne-Kathrin Hirsch, Charlott Rubach

Universität Rostock, Deutschland

Theoretical Background. As educators of future teachers, who need to prepare pre-service teachers for the complexities of modern classrooms, understanding the factors that shape pre-service teachers' professional digital competencies becomes paramount. To enhance pre-service teachers' professional digital competence, Tondeur et al. (2012) proposed the Synthesis of Qualitative Evidence (SQD) model, where six of the twelve key strategies are directly associated with the training of preservice teachers in university classes (role models, instructional design, authentic experiences, collaboration, reflection, feedback). There are several studies investigating the efficacy of programs that aim to develop pre-service teachers' professional digital competencies (Kay, 2006; Knezek et al., 2023), but still empirical knowledge on specific strategies that help to foster such competencies in university classes are lacking (Backfisch et al., 2020). Addressing this lack, the paper investigates the impact of these strategies on pre-service teachers' professional digital competence. For the aim of this paper, we depicted the Technological Pedagogical Content Knowledge (TPACK) model by Mishra and Koehler (2006) to define teachers professional digital competence. The aim of the TPACK model is to predict and explain teachers' integration of technology into their teaching in a way that enhances the learning of specific content.

Objectives. We investigate in a longitudinal research design to what extend the self-perceived TPK and TPACK changes (RQ 1) and the longitudinal relationships between strategies conceptualized in the SQD model (Tondeur et al., 2012), and the self-perceived TPK and TPACK of pre-service teachers (RQ 2) (Mishra & Koehler, 2006).

Methods. Pre-service teachers (N = 313, 30 % male, 34 % gymnasium) were asked in the beginning and at the end of the semester with eight items about their self-perceived TPK and TPACK ($\Omega \geq 0.88$). Furthermore, they were questioned about the seminar according to the six SQD strategies with 19 items ($\Omega \geq 0.70$). We adapted some SQD items to address the importance of various stakeholders, including instructors, students and teachers (Tinto, 1993).

Results. Utilizing latent-change score models, we indicated no significant change in TPK ($\beta\Delta = .146$, S.E. = .116, $p = .211$, $\sigma^2\Delta = 1.329$, S.E. = .172, $p \leq .001$) and TPACK ($\beta\Delta = .073$, S.E. = .097, $p = .454$, $\sigma^2\Delta = 1.139$, S.E. = .159, $p \leq .001$). For research question 2, we indicated through an exploratory factor analysis a different structure (5-factors) compared to the proposed SQD model (6-factors). Here, the factors Role Model and Feedback were retained, but three new factors were indicated: Interaction with students, Interaction with teachers and Empowerment. Lastly, we used a bifactor exploratory structural equation model to investigated the effects of the general factor representing teaching strategies and the specific factors of each strategy on the change of self-perceived TPK and TPACK. Results demonstrated that the G-factor significantly, positively influenced the development of self-perceived TPK and TPACK.

Discussion. In line with Knezek et al. (2023), our study demonstrated TPK and TPACK are dependent on strategies employed within courses. We expanded their study and support these finding in a longitudinal design. To facilitate TPK and TPACK in teacher education, we found that the orchestration of the five newly indicated SQD strategies had a substantial impact on the changes of self-perceived TPK and TPACK (see Knezek et al., 2023). It can be specified that no single strategy is paramount; rather, the orchestration of all five SDQ strategies is crucial, with each playing a specific role in the overall learning process. Therefore, teacher training programs with a focus on digital competence development now have a starting point to design their teaching.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 10: Förderung der professionellen Wahrnehmung: ein systematisches Review

Julia Lumen¹, Ann-Sophie Grub², Antje Biermann¹

¹Universität des Saarlandes, Empirische Bildungsforschung, Deutschland; ²Universität des Saarlandes, Zentrum für Lehrerbildung, Deutschland

Die professionelle Wahrnehmung im Unterrichtskontext beschreibt die Fähigkeit von Lehrkräften, aus dem komplexen Unterrichtsgeschehen relevante Aspekte herauszufiltern, theoriegeleitet zu interpretieren und zu beurteilen, um daraus Handlungsentscheidungen für die Unterrichtsführung abzuleiten (Mason, 2011; Schoenfeld, 2010; Sherin, 2001). Diese Fähigkeit ist ein bedeutender Faktor für die Unterrichtsqualität (Blömeke et al., 2016; Grub et al., 2022) und beeinflusst den Lernerfolg der Schüler:innen (Blömeke et al., 2016; Kersting et al., 2012). Studien zeigen Unterschiede in der professionellen Wahrnehmung zwischen Expert:innen und Noviz:innen sowie unter Berufsanfänger:innen (Gegenfurtner et al., 2011; van Driel et al., 2021; Wyss et al., 2021; Eßling et al., 2023). Expert:innen können in der komplexen Unterrichtsdynamik relevante Ereignisse sicher von irrelevanten unterscheiden (van Driel et al., 2021), Noviz:innen dagegen nehmen das Geschehen eher diffus wahr und fokussieren auf auffällige, aber oft unbedeutende Merkmale (Thiel et al., 2012). Eßling et al. (2023) zeigten in einer Längsschnittstudie, dass angehende Lehrkräfte bereits zu Beginn ihrer Ausbildung unterschiedliche Ausgangsniveaus in der professionellen Wahrnehmung haben, insbesondere in der Förderung der kognitiven Aktivierung der Schüler:innen und der Strukturierung der Unterrichtsinhalte. Es ist daher wichtig, diese Unterschiede durch gezielte Fördermaßnahmen anzugehen, um ein homogen hohes Leistungsniveau der Lehrkräfte zu erreichen. Durch die Förderung der professionellen Wahrnehmung lassen sich die professionelle Entwicklung von Lehrkräften unterstützen, die Unterrichtsqualität verbessern und somit der Lernerfolg der Schüler:innen steigern (Barth, 2017; Meschede et al., 2015; Stürmer et al., 2013, 2017).

In der Literatur gibt es verschiedene Möglichkeiten zur Förderung der professionellen Wahrnehmung (Grub et al., angenommen; Keskin et al., 2024). Diese reichen von spezifischen Interventionen in Form von Seminaren (Gold et al., 2021) bis hin zur kontinuierlichen Wissenserweiterung im Rahmen von Ausbildung und Studium (Eßling et al., 2023). Ein systematischer Überblick über diese Fördermöglichkeiten und ihre Effektivität ist entscheidend – aber bis zum aktuellen Zeitpunkt unseres Wissens nach noch in der Forschungslandschaft fehlend - um fundierte Trainingsprogramme und pädagogische Maßnahmen zu entwickeln.

Durch die Identifizierung und Analyse verschiedener Ansätze können Stärken und Schwächen einzelner Interventionen herausgearbeitet und effektive Maßnahmen bestimmt werden. Eine solche Zusammenstellung schont zudem Forschungsressourcen, indem sie zeigt, welche Informationen bereits etabliert sind und wo weiterer Forschungsbedarf besteht. Insgesamt bietet ein Review nicht nur einen umfassenden Überblick über den aktuellen Stand der Forschung zur Förderung der professionellen Wahrnehmung von Lehrkräften, sondern trägt auch zur Entwicklung evidenzbasierter Interventionen und Fortbildungsprogramme bei, die die Unterrichtsqualität verbessern und den Lernerfolg der Schüler:innen steigern.

Die vorliegende, präregistrierte Arbeit hat zum Ziel, einen Überblick über die bisher verwendeten Maßnahmen zur Förderung und Unterstützung der professionellen Wahrnehmung von angehenden und erfahrenen Lehrkräften zu schaffen, zusammenzufassen und geeignete Maßnahmen zu identifizieren. Hierfür wurde im April 2024 eine systematische Literaturrecherche durchgeführt, indem die drei Datenbanken EbscoHost, Scopus und Web of Science für den Zeitraum von 2001 bis 2024 durchsucht wurden (insgesamt: 16.323 Artikel). Als Suchstrategie wurde eine boolesche Verknüpfung der drei Kategorien „professional vision“, „support“ und „teacher“ verwendet, die jeweils verwandte Wörter und Synonyme inkludierten. Es konnten nach Duplikatsentfernung 7.771 Artikel in das aktuell noch laufende Abstractscreening mit dem KI-gestützten Tool ASReview eingehen. Zum aktuellen Zeitpunkt wurden bereits mehr als 200 Abstracts durch zwei Begutachterinnen gesichtet. Erste Interraterreliabilitäten zur Qualitätssicherung liegen im guten Bereich ($K = .74$; Altman, 1991). Bis zur Tagung wird das Volltextscreening abgeschlossen sein und erste Ergebnisse können präsentiert werden. Dabei werden insbesondere die Zielgruppe der Fördermaßnahmen, die konkrete Messmethodik und Operationalisierung der professionellen Wahrnehmung sowie die systematische Analyse der im Kontext der Lehrtätigkeit nutzbaren Fördermaßnahmen adressiert, sowie die Eignung der verschiedenen Maßnahmen und deren verschiedene fachspezifische und didaktische Schwerpunkte zur Unterstützung und Förderung der professionellen Wahrnehmung von angehenden und erfahrenen Lehrkräften diskutiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 71: Adaptive Lernbegleitung im Praxissemester: Wahrnehmungen und Einstellungen von Studierenden

Maria Hauswald, Alexander Gröschner

Friedrich-Schiller Universität Jena, Deutschland

Lernbegleitung gilt als zentrales Element der Wirksamkeit von schulpraktischen Studien (Gröschner & Hascher, 2022; Ronfeldt, 2015). Dabei spielen insbesondere schulische Begleitansätze eine Rolle. Studien zeigen, dass *adaptive Lernbegleitung*, als zielgerichtete Anpassung der Unterstützungsansätze von Praxislehrpersonen an die Bedürfnisse von Studierenden, als lernförderlich betrachtet wird (van Ginkel et al., 2015). Im deutschsprachigen Raum fehlen jedoch empirische Untersuchungen, die Ansätze der adaptiven Lernbegleitung aus Studierendensicht beschreiben.

Im Rahmen einer Untersuchung der Präferenzen und Einstellungen während der Lernbegleitung wurden $N=69$ Studierende vor und nach dem Praxissemester schriftlich befragt. Das an van Ginkel et al. (2015) angelehnte Erhebungsinstrument wurde zunächst an einer Stichprobe von $N=174$ Studierenden validiert. Die Studie zeigt, dass 84% der Studierenden sich zu Beginn eine partizipative Begleitung durch die Praxislehrpersonen wünschen und diese Präferenz während des Praktikums signifikant zunimmt ($F=6.630$; $p=.012$; $\eta^2=.089$). Ergänzende qualitative Analysen einer Fallvignette zeigen darüber hinaus intersubjektive Differenzen zwischen der wahrgenommenen Lernbegleitung und den Präferenzen der Studierenden.

Die Ergebnisse verweisen darauf, dass Praxislehrpersonen zukünftig spezifischer als bislang auf heterogene Voraussetzungen und Lernerfahrungen von Studierenden eingehen müssen, damit eine von Kollegialität und wechselseitiger Abstimmung geprägte Begleitung stattfinden kann.

Pinnwand 72: Professionelle Handlungskompetenzen (angehender) Lehrkräfte für die sprachliche Bildung: Ein Systematic Review

Saskia Samland, Lina Wirth, Poldi Kuhl

Leuphana Universität Lüneburg, Deutschland

Theorie:

Das deutsche Schulsystem steht vor der großen Herausforderung, die Entwicklung einer sprachlich inklusiven Bildungspraxis zu gewährleisten, die jegliche Formen sprachlicher Heterogenität als grundlegende Voraussetzung anerkennt und die sprachlichen Kompetenzen der Schüler*innen in jedem Unterrichtsfach zu unterstützen (Fürstenau, 2012; Koch-Priewe & Krüger-Potratz, 2016). Da Sprache der Schlüssel für Bildungserfolg und gesellschaftliche Teilhabe ist, müssen alle (angehenden) Lehrkräfte spezifische professionelle Handlungskompetenzen für den Umgang mit sprachlicher Diversität erwerben (KMK & HRK, 2015; Stangen et al., 2020). Hierfür wird in internationalen Publikationen der Einbezug des *Pedagogical Language Knowledge* als Facette professioneller Handlungskompetenz von Lehrkräften gefordert (Bunch, 2013). Auch im deutschsprachigen Raum wird die Integration u. a. linguistischen Grundlagenwissens als eine Facette sprachbezogenem Wissens neben dem pädagogischen, fachlichen und fachdidaktischen Wissen in die Lehrkräftebildung aller Unterrichtsfächer diskutiert (Krüger-Potratz, 2020). Zuletzt wurden vermehrt Lerngelegenheiten bezüglich des Umgangs mit sprachlicher Diversität für Lehramtsstudierende bereitgestellt (Becker-Mrotzek & Woerfel, 2020). Empirische Untersuchungen zeigen gleichzeitig, dass sich (angehende) Lehrkräfte darauf häufig noch nicht ausreichend vorbereitet fühlen (Brandt, 2021) und auch noch keine hinreichenden Kompetenzen dafür mitbringen (Brandt, Ehmke et al., 2024; Prediger, 2019).

Effektive Lerngelegenheiten zu schaffen, ist unabdingbar, um professionelle Handlungskompetenzen aufzubauen. Dafür ist es jedoch wichtig, theoretisch zu definieren und empirisch zu prüfen, welche Facetten die professionellen sprachbezogenen Handlungskompetenzen umfassen (Gültekin-Karakoç et al., 2016). Modelle, wie z. B. DaZKom (Köker et al., 2015) oder das Sprachförderkompetenzmodell (Hopp et al., 2010) definieren entsprechende Kompetenzen und bieten Ansätze zu deren Messung. Das *European Core Curriculum for Inclusive Academic Language Teaching* (IALT-Curriculum) wurde ebenfalls für den deutschsprachigen Raum adaptiert (Brandenburger et al., 2011). Diese und weitere Untersuchungen betrachten aus u. a. sprach- sowie bildungswissenschaftlicher Perspektive verschiedene Facetten sprachbezogener Handlungskompetenzen oder spezifischer DaZ-Kompetenzen (Brandt, Menzel et al., 2024; Göbel et al., 2023; Schröter et al., 2024). Es liegt bislang jedoch keine systematische Betrachtung der verschiedenen Konzeptualisierungen professioneller sprachbezogener Handlungskompetenzen vor, die (angehende) Lehrkräfte fachunabhängig an deutschen Schulen benötigen, um ALLEN Schüler*innen zum Bildungserfolg durch eine sprachliche Bildung zu verhelfen.

Fragestellung:

An diesem Desiderat setzt das vorliegende Systematic Review an und geht folgender Fragestellung nach: Mit welchen Konstrukten werden im deutschsprachigen Raum professionelle sprachbezogene Handlungskompetenzen bei (angehenden) Lehrkräften konzeptualisiert und operationalisiert?

Methode:

Der Systematic Review-Prozess folgt den PRISMA-Richtlinien (Page et al., 2021). Es wurden fünf Datenbanken (FIS-Bildung, ERIC, PsycInfo, PubPsych, Scopus und Web of Science) durchsucht. Der Suchstring setzte sich aus Schlagwörtern aus drei Teilbereichen zusammen: (1) Professionelle Handlungskompetenzen, (2) sprachbildendes Unterrichten sowie (3) (angehende) Lehrkräfte. Die initiale Suche ergab insgesamt 4740 Treffer. Nach dem Entfernen der Duplikate blieben 3435 Treffer übrig. Die Treffer wurden anhand vorab definierter Ein- und Ausschlusskriterien von zwei unabhängigen Wissenschaftlerinnen gesichtet. Eingeschlossen wurden deutsch- sowie englischsprachige empirische Arbeiten, die professionelle sprachbezogene Handlungskompetenzen (angehender) Lehrkräfte im deutschsprachigen Raum untersuchen.

Ergebnisse:

Durch das erste Titel- & Abstract-Screening wurden 3027 Arbeiten exkludiert. Die verbliebenen Publikationen werden derzeit durch drei unabhängige Wissenschaftler*innen in einem zusätzlichen Abstract-Screening bewertet. Erste Einblicke zeigen, dass sich die Untersuchungen der inkludierten Publikationen vor allem mit DaZ-spezifischen oder Sprachförderkompetenzen, aber auch mit Kompetenzen sowie Überzeugungen für die Sprachbildung z. B. im Fachunterricht beschäftigen. Zudem werden

Lerngelegenheiten in den Blick genommen, die sich auf den Kompetenzerwerb dieser beziehen. Betrachtet werden Lehramtsstudierende und Lehrkräfte sowohl fachübergreifend als auch fachspezifisch.

Nach dem Abschluss des Abstract-Screening-Prozesses, folgt das Screening der Volltexte. In weiteren Analysen der finalen Treffer wird der Schwerpunkt daraufgelegt, die verschiedenen Konstrukte professioneller sprachbezogener Handlungskompetenzen vergleichend gegenüberzustellen. Darüber hinaus werden Implikationen für weitere Forschung sowie die Weiterentwicklung der Lehrkräftebildung diskutiert. Diese Übersichtsarbeit leistet somit einen Beitrag dazu, den aktuellen interdisziplinären Forschungsstand zu den Konstrukten professioneller sprachbezogener Handlungskompetenzen (angehender) Lehrkräfte im deutschen Bildungssystem zu systematisieren, um im weiteren Verlauf vorhandene Lerngelegenheiten an Universitäten auf Grundlage dessen zu untersuchen.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 9: Procrastination or strategic delay? Predictors of delayed learning behavior in an e-learning context.

Sarah Nicole Wörner, Marc Philipp Janson

Universität Mannheim, Deutschland

Among university students, delaying learning behavior until close to exams can be frequently observed, leading to massed learning behavior which is stated as nonbeneficial for lasting learning (Son, 2004). Predictors of delayed learning behavior have been investigated in past research (Wessel et al., 2019) and procrastination is a prominent one. Procrastination is sometimes distinguished into passive procrastination and (active) strategic delay (Chun Chu & Choi, 2005; Wessel et al., 2019). Passive procrastination is seen as unintentional and irrational, which might also evoke negative feelings (Baumeister, 1997; Silver & Sabini, 1981). Contrary, strategic delay is seen as intentional (Chun Chu & Choi, 2005) reflecting possible (mis)beliefs of massed learning as an efficient strategy (McIntyre & Munson, 2008). While both can relate to similar observed patterns, we consider passive procrastination as a form of self-regulatory failure and strategic delay as a form of (maladaptive) self-regulation (Chowdhury & Pychyl, 2018). This differentiation might help to understand the psychological processes of observed objective learning behavior, like log-file data from e-learning. Recent research underlines the importance of differentiating different processes leading to similar patterns in e-learning behavior (Baker et al., 2020).

Hence, we aim to compare procrastination and strategic delay as predictors for delayed learning behavior in an e-learning context. As empiric evidence for both processes exists, we preregistered the hypotheses (https://aspredicted.org/YRK_8R3) that higher procrastination tendencies predict delayed learning behavior (H1) and higher tendencies towards strategic delay predict delayed learning behavior (H2).

We conducted a field experiment on university $N = 97$ students using an intelligent tutoring system for self-regulated exam preparation. We assessed passive procrastination (Tuckman, 1991, $\alpha = .93$) and strategic delay (Choi & Moran, 2009, $\alpha = .79$) at first login. As indicator of delayed learning behavior, we derived a delay parameter from the software's log files. Individual delay parameters were defined as the number of days left until the exam, where the cumulative relative learning time of a person was last below specific percentage thresholds. For clarity, cumulative relative learning time refers to the accumulated proportion of time spent learning up to each day compared to the total learning time of a learner. The percentage thresholds under consideration were 25%, 50% and 75%. $N = 97$ students using the intelligent tutoring system had an average total learning time of 23.17 hours ($SD = 15.65$), passing the 25%/50%/75% threshold on average 26.85/17.41/9.58 days ($SD = 13.23/11.88/9.58$) before the exam. As assumed, higher procrastination tendencies significantly predicted delayed learning behavior for the delay parameter referring to 25% with $R^2 = 0.075$, $F(1, 95) = 7.69$, $p = .007$. Contrary, we found no significant results whether higher tendencies towards strategic delay predict delayed learning behavior, $R^2 = 0.002$, $F(1, 95) = 0.20$, $p = .657$. For the 50% and 75% thresholds, we did not reveal any significant predictions, potentially due to reduced variance in these dependent variables. First, our findings of higher variations in the 25% threshold underline the stronger impact of interindividual differences in learning with higher temporal distance to the exam, supporting the strong situational factor of upcoming exams on students' motivation (Capelle et al., 2022). For relatively earlier chosen points in the course of learning, procrastination tendency served as a predictor of delayed learning behavior, while tendency towards strategic delay did not. Existing literature already questioned the validity of strategic delay (Hensley, 2014). The present study provides evidence of whether differences in observed e-learning behavior can be used to investigate self-regulatory successes and failures as delayed learning behavior is not associated with any reported strategic delay. This is promising for educational researchers using objective trace data from e-learning for self-regulation research.

Pinnwand 21: ALIFE: Strategies for Designing Effective Adaptive Online Professional Development Courses for Teachers

Fitore Morina, Nina Udvardi-Lakos, Armin Fabian, Ulrich Trautwein, Tim Hagen Fütterer

Eberhard-Karls Universität Tübingen, Germany

Background

Professional development (PD) is a key approach to qualify and professionalize in-service teachers (Fütterer et al., 2023, 2024a; Morina et al., 2024; Richter et al., 2024). However, attendance rates of PD courses are relatively low, most teachers who do participate in PD have high prior knowledge already, leading to a Matthew effect, and many PD courses show high drop-out rates (Fütterer et al., 2024b; Richter & Richter, 2020). A reason for these patterns could be the lack of PD courses that consider teachers' individual needs. Adaptive PD could address this issue: By adapting course content to teachers' prior knowledge and interests, teachers may be more likely to participate in and engage with PD offers. Asynchronous online PD can also reduce barriers, such as time constraints and travelling required for in-person courses, that could deter teachers from participating in PD.

Methods

In the project ALIFE (**Adaptive Learning Integrated Framework for Educators**; <https://alife.academy>), online PD will be developed that combines current knowledge about adaptivity and well-established theories concerning cognitive, metacognitive and motivational constructs (e.g., Sweller et al., 2019; Wigfield & Eccles, 2000; Zimmerman, 2002). These evidence-based PD courses should support teachers' professionalization. Adaptivity in this context means we will diagnose teachers' needs and preferences and accommodate these by making appropriate adjustments to the PD to enhance learning outcomes (Plass & Pawar, 2020). For example, adaptive PD could assess prior knowledge and then provide module contents at the right level of difficulty for each individual teacher. Within the modules, adaptivity can assess individuals' needs during a learning task, and then provide support (e.g., individualized feedback; Luckin et al., 2016).

The development of the online, adaptive PD in ALIFE follows a six-step strategy (Sandoval & Bell, 2004; Trautwein et al., 2023): First, in-service teachers' needs for the planned contents were assessed using an online survey ($N = 83$, $M_{\text{age}} = 39.7$). Second, based on the findings, the objectives and success criteria for the PDs were set, before developing the PD content and implementing adaptivity elements as third step. Fourth, the PDs will be piloted using a small sample from August to December

2024, using both questionnaires and behavioral trace data to investigate the effect of adaptivity elements on participants' motivation and learning. Fifth, a systematic implementation and evaluation will be executed. Sixth, after findings from the systematic evaluation have been analyzed and the PD reviewed accordingly, the PD should be scaled-up so in-service teachers throughout Germany can participate. The development of these adaptive PD courses should then inform the development of future PD. Throughout the entire process, a close collaboration with teachers enables the consideration of their perspectives and needs from the beginning of the development. We will present the development strategy, needs-analysis results, and first results of pilot studies here.

Results

The needs analysis conducted as part of Step 1 showed that in-service teachers prefer PD that is evidence-based and adaptive to their interests and prior knowledge ($\geq 79\%$ see these aspects as important or very important). Current PD is perceived as offering low adaptivity to teachers' prior knowledge, as only 53% of teachers said that PD offers match their prior knowledge. Many teachers reported a lack of fit of existing PD to their interests, and a lack of relevance for their teaching practice, with 69% of teachers perceiving these aspects as barriers to participating in PD. Therefore, the focus on adaptivity within ALIFE seems promising for fostering teachers' participation in PD. Currently, we are working on Step 4 of the strategy cycle, piloting two PD courses. At the conference, we will also present first results of these pilot studies.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 73: Lehrkräftebildung und Lehramtsstudierende? – Eine Analyse der Verwendung von gendermovierten Komposita am Beispiel der deutschen Bildungsforschung.

Anna-Lena Wolf

Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU), Deutschland

Das Thema der Genderfrage polarisiert in der deutschen Medienlandschaft mehr denn je. Dabei handelt sich um keine neue Diskussion, sondern um eine Debatte, die bereits in den 1990er Jahren aus der feministischen Sprachkritik heraus zum zentralen Thema sowohl des wissenschaftlichen als auch des öffentlichen Diskurses herangewachsen ist. Gendern ist politisch und das nicht nur, weil Parteien in Deutschland das Thema für ihre Wahlprogramme und den Gewinn von Wählerstimmen nutzen (1), vielmehr ist die gesamte Debatte sozial-semiotisch aufgeladen (10). Einhergehend mit der dadurch vermittelten unterliegenden Einstellung kann die Verwendung oder Nicht-Verwendung von Gendersprache nicht nur die eigene politische Ideologie zum Ausdruck bringen, sondern sie kann dadurch auch entscheidend für beruflichen oder privaten Erfolg sein (11). Zu beobachten ist außerdem, dass der Widerstand gegen gendergerechte Anpassungen zunehmend lauter wird und hohe mediale Aufmerksamkeit erfährt. Ablehnende Haltungen sind dabei nicht mehr nur in privaten Räumen vorzufinden. Durch politische Verbote von genderangepasster Sprache in Bundesländern wie Bayern dringen diese zunehmend in öffentliche Räume (2).

Sogar in der Linguistik werden Gendermovierungen vielfach diskutiert: Studien für eine grammatikalische Korrektheit des generischen Maskulinums als geschlechtsneutrale Form (13) treten hierbei in den Widerspruch zu Studien, die aufzeigen, dass eine Verwendung dieser dazu führt, dass tendenziell männliche Personen assoziiert werden (6, 8, 12), während genderangepasste Sprache eine höhere Inklusion aller Geschlechtsidentitäten befördert (5). Bei der diskutierten Genderfrage referiert der Ausdruck *Gender* nicht auf das biologische Geschlecht, sondern auf das soziale Geschlecht, zu welchem sich Menschen eigenständig bekennen und welches durch Normen eingeübt und dargestellt wird (8). Eine Auflösung der Dichotomie in der Genderzuweisung wurde bereits durch die Einführung der Kategorie ‚Divers‘ im Jahr 2018 gesetzlich verankert (3). Ein Gegenstand der linguistischen Debatte ist ein Ersatz des generischen Maskulinum durch inklusive Formen und eine dadurch erzielte Inklusion des kompletten Genderspektrums (4, 9). Das Deutsche bietet mit Sonderzeichen, Doppelnennungen oder Ersatzformen unterschiedliche Möglichkeiten, um diese Anpassung zu vollziehen. Eine solche kann ebenfalls im Erst- oder Zweitglied von Komposita auftreten, wobei eine Anpassung im Zweitglied durch den konkreten Personenbezug einfacher umsetzbar und dadurch weiter verbreitet ist (7).

Ein Bereich, der eine hohe Verwendung geschlechtsspezifischer Sprache aufweist, ist die deutsche Bildungsforschung inklusive an diese angegliederte Disziplinen aus den lehramtsspezifischen Fachwissenschaften oder Fachdidaktik. Besonders geschlechtsspezifische Komposita (*Lehrerbildung*, *Lehramtsstudierende* etc.) finden hier hohe Verwendung. Teilweise werden diese moviert, während bei anderen die Verwendung der generischen Form stattfindet. Im geplanten Poster wird ein Dissertationsprojekt vorgestellt, das sich der Frage widmet, inwiefern Gendermovierungen in Komposita in der Sprache der Veröffentlichungen der deutschen Bildungsforschung zum jetzigen Zeitpunkt umgesetzt sind.

Zur Beantwortung dieser Frage wird ein Korpus mit 274 Artikeln, veröffentlicht im Zeitraum von 2000 bis 2023, systematisch nach gendermovierten Komposita durchsucht. Erste Zwischenergebnisse der Analyse ergeben 12.503 gefundenen Komposita mit mindestens einem genderspezifischen Wortglied. Diese unterteilen sich in 7.111 Erstglieder, 4.797 Letzt-Glieder, 169 Mittelglieder und 355 Token, bei denen mehr als ein Wortglied genderspezifisch ist. Insgesamt werden bei 5.244 Token angepasste Genderschreibweisen verwendet, sowie bei 174 Token eine explizit feminine Form. Die restlichen 7.259 Token verwenden weiterhin das generische Maskulinum. Innerhalb der gendermovierten Token dominiert die Ersatzform (73%), welche für weitere Analysen nochmals in Partizipform (bspw. *Studierende*) und Neutralform (bspw. *Lehrkräfte*) zu unterteilen ist, gemeinsam mit den Doppelnennungen (8%). Die im gesellschaftlichen Diskurs teilweise hoch kritisch gesehenen Genderzeichen (Asterisk, Binnenmajuskel, Doppelpunkt etc.) finden sich nur in 19% aller angepassten Formen wieder. Weitere Ergebnisse aus den Daten der Dissertation hinsichtlich Verteilungen gefundener Lexeme sowie diachronen Entwicklungen werden auf dem Poster präsentiert. Basierend auf diesen vorläufigen Beobachtungen können dann Vorschläge diskutiert werden, wie der Gebrauch von gendersensibler Sprache im Bereich der Bildungsforschung wegweisend für eine höhere gesellschaftliche Akzeptanz von Gendersprache dienen und Maßnahmen zum gendersensiblen Sprachgebrauch unterstützen und fördern kann.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 28: Der Geschlechterunterschied in der Wirtschaftskompetenz von Jugendlichen – Eine empirische Analyse

Lucy Haag¹, Luis Oberrauch², Taiga Brahm¹, Martin Biewen¹

¹Universität Tübingen; ²Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau

Theoretischer Hintergrund

Ausgeprägte Wirtschaftskompetenzen sind essenziell für fundierte Entscheidungen im komplexen Wirtschaftsgeschehen. Studien zeigen einen positiven Zusammenhang dieser Kompetenzen mit den persönlichen Finanzen und der Jobsicherheit in Krisensituationen (Grimes et al., 2021) sowie positive Auswirkungen von Finanzkompetenz auf Investitionen und Altersvorsorge (Bucher-Koenen et al., 2021). Gleichzeitig belegen zahlreiche Untersuchungen die geringe Wirtschaftskompetenz in der Bevölkerung. Frauen schneiden dabei oft schlechter ab als Männer (Oberrauch & Brahm, 2023). Es gibt bereits einige Erklärungsansätze für diesen Geschlechterunterschied, die insbesondere aus der Literatur zur finanziellen Kompetenz stammen (z.B. Preston et al., 2023; Yao et al., 2022). Ein wiederkehrendes Ergebnis ist jedoch, dass bisher nur ein kleiner Teil des Unterschieds erklärt werden kann.

Vor diesem Hintergrund wird in der vorliegenden Studie die Wirtschaftskompetenz von Jugendlichen in Baden-Württemberg mit folgenden Forschungsfragen untersucht:

- Wie ist der Geschlechterunterschied in der Wirtschaftskompetenz von Jugendlichen ausgeprägt?
- Welche Faktoren können den (potenziellen) Geschlechterunterschied erklären?

Methode

Es wurde eine Onlineumfrage mit N=2104 Zehntklässlern/-innen an allgemeinbildenden Schulen in Baden-Württemberg durchgeführt. In diesem Alter stehen viele Jugendliche bereits vor wichtigen Entscheidungen für die Zukunft, wofür eine fundierte Wirtschaftskompetenz von großer Bedeutung ist.

Der Fragebogen umfasst zwei Tests zur Erfassung ökonomischer Kompetenzen (Test of Economic Competence (Oberrauch et al., 2023) und Test of Economic Literacy (Walstad et al., 2013)). Weiterhin enthält der Fragebogen mehrere potenzielle Erklärerfaktoren für den Geschlechterunterschied, die in vorherigen Studien untersucht wurden. Darunter sind beispielsweise mathematische Fähigkeiten (Brückner et al., 2015), Selbstvertrauen (Bucher-Koenen et al., 2021), Interesse an Wirtschaftsthemen (Förster & Happ, 2019) und Stereotype (Driva et al., 2016). Anhand einer multiplen Regressionsanalyse sowie einer Oaxaca-Blinder Decomposition wird untersucht, wie stark der Geschlechterunterschied ausgeprägt ist und welche Faktoren den Unterschied (teilweise) erklären können.

Ergebnisse und Bedeutung

Die deskriptive Analyse zeigt zunächst signifikante Unterschiede zwischen Mädchen und Jungen in der Wirtschaftskompetenz. Jungen schneiden in beiden Tests um knapp 1/3 Standardabweichung besser ab als Mädchen. Weiterhin wird ein Geschlechterunterschied in vielen (Kontroll-)Variablen deutlich. So geben Jungen an, selbstbewusster zu sein und sich mehr für Wirtschaft zu interessieren als Mädchen. Außerdem schneiden Jungen im Mathetest besser ab, während sich Mädchen in den Wirtschaftstests mehr angestrengt haben.

Die Regressionsanalyse zeigt, dass der Zusammenhang des Geschlechts mit der Kompetenz durch einige Erklärerfaktoren aufgeklärt werden kann. Die größte Rolle spielt die mathematische Kompetenz. Durch zusätzliche Kontrolle von Selbstwirksamkeit, Zeitpräferenzen (present bias), Interesse und Sozialisationsfaktoren verringert sich der Einfluss des Geschlechts auf ein Drittel des ursprünglichen Werts. Da sich die Mädchen im Test tendenziell mehr anstrengen, vergrößert sich der Gap, wenn die Anstrengung konstant gehalten wird.

Die Decomposition zerlegt den Geschlechterunterschied in einen erklärten Teil (Effekt der unabhängigen Variablen/Endowments), der ca. ein Drittel des Unterschieds erklärt, sowie einen unerklärten Teil (Coefficients), der ca. zwei Drittel des Gaps ausmacht. Wie durch die Regression bereits angedeutet, kann ein bedeutender Teil des Gaps durch die höhere mathematische Kompetenz der Jungen erklärt werden. Weitere signifikante Teile gehen auf die Selbstwirksamkeit und Zeitpräferenz zurück. Die größere Anstrengung der Mädchen trägt hingegen zur Reduzierung des Gaps bei. Andere Variablen wie Interesse und Stereotype haben keinen signifikanten Einfluss.

Insgesamt ermöglicht der Beitrag ein tieferes Verständnis für die Geschlechterunterschiede von Jugendlichen und zeigt einen bedeutsamen Unterschied in der Wirtschaftskompetenz bereits bei 15-jährigen. Mehrere Faktoren tragen zu diesem Unterschied bei, wobei das bessere Abschneiden der Jungen in der mathematischen Kompetenz den größten Anteil des Unterschieds erklärt. Ein beträchtlicher Teil des Geschlechterunterschieds bleibt jedoch durch die erhobenen Variablen unerklärt. Eine mögliche Erklärung könnten Sozialisationsfaktoren sein, welche zwar teilweise berücksichtigt werden (z.B. durch Stereotype), aber schwer quantifizierbar sind.

Aufgrund der großen Bedeutung von Wirtschaftskompetenz gilt es, Mädchen künftig stärker zu unterstützen, um das wirtschaftliche Wohlergehen aller Bevölkerungsgruppen zu stärken.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 74: Übersetzung und Validierung eines Fragebogens zur kontextbezogenen Erfassung sozial-emotionaler Kompetenzen und Schwierigkeiten – Kids in Contexts Scale (KICS)

Mia Schrage¹, Gino Casale¹, Andres De Los Reyes², Elizabeth Talbott³

¹Bergische Universität Wuppertal, Deutschland; ²University of Maryland, USA; ³William & Mary School of Education, USA

Theoretischer Hintergrund

Zunehmend mehr Kinder und Jugendliche in der Schule zeigen bedeutsame Probleme in der sozial-emotionalen Entwicklung, die sich insbesondere in Verhaltensauffälligkeiten und steigenden Förderquoten im Schwerpunkt Emotionale und soziale Entwicklung (FS ESE) äußern. Um die Bedürfnisse der Schüler:innen angemessen adressieren zu können, bedarf es einer präzisen, umfassenden und schulkompatiblen Diagnostik sozial-emotionaler Kompetenzen (von der Embse & De Los Reyes, 2024; KMK, 2022). Systematische Analysen der Diagnose- und Förderplanpraxis bei Schüler:innen mit sozial-emotionalen Entwicklungsproblemen legen jedoch nahe, dass die aktuellen Diagnosepraktiken zu Benachteiligungen und der Aufrechterhaltung von Bildungsungerechtigkeiten (u. a. Über-/ Unterrepräsentation spezifischer Schüler:innengruppen, ungleiche Zuweisung zu Fördermaßnahmen, fehlerhafte Diagnostik) beitragen (Casale, 2024; Casale & De Los Reyes, 2023; Niederöst et al., 2024). Identifizierte Problematiken betreffen unter anderem (1) die diagnostischen Informationen, die häufig aus einer einzigen Informationsquelle stammen (i. d. R. die Lehrkraft), (2) die unzureichende Berücksichtigung der Kontextrelevanz sozial-emotionaler Entwicklungsmerkmale, (3) den überwiegenden Fokus auf Defizite der Schüler:innen und (4) den Einsatz größtenteils nicht standardisierter Verfahren durch Lehrkräfte (Casale, 2024; Niederöst et al., 2024). Um diese Mängel zu überwinden, sind geeignete Messinstrumente erforderlich, die (1) diagnostischen Informationen aus Perspektive mehrerer Personen einbeziehen, um ein ganzheitliches Bild der Bedürfnisse zu erhalten, (2) berücksichtigen, dass die Förderbedarfe der Schüler:innen in der Regel kontextgebunden sind und sich in verschiedenen Kontexten unterschiedlich manifestieren können, (3) auf einem umfassenden Verständnis von psychischer Gesundheit basieren und sowohl psychische Bedürfnisse als auch das Wohlbefinden der Schüler:innen berücksichtigen sowie (4) inhaltlich und methodisch fundiert sind (Casale & De Los Reyes, 2023; von der Embse & De Los Reyes, 2024; Talbott et al., 2023). Zudem sollten die Verfahren in Schulen gut umsetzbar sein (Casale & De Los Reyes, 2023).

Im vorliegenden Beitrag wird das Design der deutschsprachigen Entwicklungs- und Validierungsstudie zur Kids in Contexts Scale (KICS) vorgestellt. Die KICS ermöglicht eine detaillierte Erfassung sowohl sozialer, emotionaler und verhaltensbezogener Entwicklungsbedarfe als auch sozial-emotionaler Stärken von Schüler:innen aus mehreren Perspektiven (Eltern, Lehrkräfte, Schüler:innen). Des Weiteren bietet die KICS die Möglichkeit, das beurteilte Verhalten sowohl in schulischen als auch in außerschulischen Situationen einzuordnen, um eine präzisere Erfassung der Entwicklungsbedarfe der Schüler:innen in den jeweiligen Kontexten zu gewährleisten. Ziel des vorliegenden Projektes ist die Übersetzung und Validierung der KICS für den deutschsprachigen Raum.

Methode

Die methodische Vorgehensweise des Projektes umfasst zwei Phasen: In der ersten Phase werden die Originalfragebögen gemäß den Vorgaben des TRAPD-Modells (Translation, Review, Adjudication, Pre-Testing, Documentation; Behr et al., 2016) ins Deutsche übersetzt. Die englischsprachigen Originalfragebögen werden von zwei unabhängigen Übersetzer:innen ins Deutsche übertragen und anschließend zu einer einheitlichen Übersetzungsversion zusammengeführt. Anschließend erfolgt eine Rückübersetzung ins Englische, um die konzeptionelle Äquivalenz der deutschen Version mit dem Original zu überprüfen. Ab Oktober 2024 werden im Rahmen von Pre-Tests kognitive Interviews mit Schüler:innen, Eltern sowie Lehrkräften durchgeführt, um die Verständlichkeit und Akzeptanz der Items zu evaluieren. Auf der Grundlage der Ergebnisse wird eine finale deutsche Version der KICS erstellt. In der zweiten Phase des Projektes wird die deutsche Version der KICS hinsichtlich der psychometrischen Güte, faktoriellen Validität und Konstruktvalidität mittels einer längsschnittlichen Validierungsstudie mit $n = 1.200$ Schüler:innen der sechsten bis achten Jahrgangsstufe zu Beginn des Schuljahres 2025/2026 überprüft.

Ausblick

Im Rahmen unseres Projektes entstehen insbesondere Erkenntnisse für eine fundierte und ganzheitliche Erfassung von sozialen, emotionalen und verhaltensbezogenen Entwicklungsbedarfen und sozial-emotionalen Stärken von Schüler:innen in verschiedenen Kontexten durch eine Integration verschiedener Perspektiven. Darüber hinaus sollen zentrale Fragestellungen zur Überprüfung der Messinvarianz zwischen unterschiedlichen kulturellen Kontexten (USA, Deutschland) als auch zwischen verschiedenen Subpopulationen (z. B. Schüler:innen mit sonderpädagogischen Förderbedarf) adressiert werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 3: Berufsorientierung in Schüler:innenpraktika – eine empirische Studie im Übergangssystem

Alexander Brodsky¹, Eveline Wuttke¹, Jürgen Seifried²

¹Goethe Universität Frankfurt, Deutschland; ²Universität Mannheim

Problemstellung und theoretischer Hintergrund

Viele Jugendliche finden nach Verlassen der allgemeinbildenden Schule keinen Ausbildungsplatz (Heisler et al., 2021) und münden – da sie noch schulpflichtig sind – in das Übergangssystem ein. Ein zentrales Ziel verschiedener Schulformen des Übergangssystems ist die berufliche Orientierung der Jugendlichen. Gemäß Ertl (2021) verfolgt Berufsorientierung das Ziel, Jugendlichen durch Informations-, Erfahrungs- und Beratungsangebote an der Schwelle von der Schule ins Berufsleben Orientierung zu geben (S. 11). Dieser länger andauernde, komplexe Prozess erfordert Lernangebote an mehreren Lernorten (Hauptausschuss des Bundesinstituts für Berufsbildung, 2005); neben dem schulischen Lernort spielt hierbei auch der betriebliche Kontext eine zentrale Rolle (Ahrens, 2007; Brüggemann, 2015).

Richter (2016, ähnlich auch Gottfredson, 1981 und 2002) verweist auf zwei Ebenen der Entwicklung einer Berufsorientierung, nämlich einerseits das Ausloten eigener beruflicher Interessen und Fähigkeiten sowie andererseits das Sammeln von Informationen über Anforderungen, Bedarfe, sowie Möglichkeiten verschiedener Berufsfelder. Schüler:innenpraktika erscheinen geeignet, um beide dieser Ebenen zu adressieren. So sollen Praktika unter anderem – insbesondere durch den Zugewinn individueller Erkenntnisse und Eignungen – eine berufsorientierende Funktion ausüben (vgl. Driesel-Lange et al., 2013 und Schudy, 2002). Auch die Förderung der zweiten Ebene (Sammeln von Informationen über Berufsfelder) erscheint durch Einblicke in Arbeitsprozesse und domänenspezifische Tätigkeiten durchaus plausibel, weshalb die Integration von Betriebspraktika als eine geeignete Maßnahme im Übergangssystem gesehen wird. Forschungsseitig ist allerdings zu konstatieren, dass wenig über die Effekte von Praktika hinsichtlich der beruflichen Orientierung von Schüler*innen im Übergangssystem bekannt ist (vgl. Heisler et al., 2021; Nickolaus & Mokhonko, 2020).

Fragestellung

An diesem Desiderat setzt der vorliegende Beitrag an und adressiert die Frage, ob und inwiefern die Erfahrungen im Betriebspraktikum die berufliche Orientierung der Schüler*innen beeinflussen.

Methode

In Hessen und Baden-Württemberg wurden Schüler*innen des Übergangssystems – jeweils vor und nach einem Betriebspraktikum – befragt, wobei neben leitfadengestützten (Gruppen-)Interviews (Eingangserhebung: 79 Schüler:innen; Ausgangserhebung: 65 Schüler:innen) auch Fragebögen eingesetzt wurden. Neben soziodemografischen Daten wurde in den Fragebögen der Berufsorientierungsindex (BOX) (Driesel-Lange et al., 2023) eingesetzt. Ergänzend zu den Befragungen der Schüler:innen wurden betreuende Lehrpersonen (derzeitige Stichprobe: $n = 8$) interviewt.

(Erwartete) Ergebnisse

Die Untersuchung der BOX-Skalen in den Eingangs- und Ausgangerhebungen zeigt, dass sich die Skalenmittelwerte zwischen Eingangs- und Ausgangserhebung nicht signifikant unterscheiden. Teilweise liegen die Mittelwerte nach dem Praktikum sogar auf einem niedrigeren Niveau als zuvor (allerdings ebenfalls keine signifikanten Unterschiede). Die Interviews befinden sich derzeit noch in Auswertung, erste Tendenzen zeigen sich allerdings schon: Ein Teil der Schüler*innen berichtet von einer Konkretisierung ihrer beruflichen Orientierung; dies ist etwa dann der Fall, wenn Schüler:innen in einem Berufsfeld Praktikum machten, für das sie bereits zuvor einen Berufswunsch hatten. In anderen Fällen hingegen geben Proband*innen an, dass ihnen das Praktikum wenig oder gar nicht bei der beruflichen Orientierung geholfen habe. Dies ist beispielsweise dann der Fall, wenn die Schüler*innen Praktika in Berufsfeldern und/oder Unternehmen aufgrund eines Mangels an Alternativen machten. In zahlreichen Fällen berichten die Schüler*innen außerdem, dass sie insbesondere mit repetitiven (Routine-)Tätigkeiten betraut wurden und dadurch nur unzureichende Einblicke in die Berufsfelder erhielten. Die Erkenntnisse zu den betrieblichen Praktika im Übergangssystem sollen u. a. dabei helfen, Erfolgspotenziale der Praktika zu identifizieren – etwa hinsichtlich besonders lernförderlicher Tätigkeiten, um so Empfehlungen für die Gestaltung von Praxisphasen weiter zu konkretisieren. Auch für schulische Entscheidungsträger:innen erscheinen die Befunde von hoher Bedeutung, da die Studie Ansatzpunkte dahingehend liefern soll, in welchen Unternehmen die Praktika aus Schüler:innenperspektive als besonders hilfreich erachtet werden; derartige Informationen können etwa in Empfehlungen bzw. Richtlinien der Praktikumssuche berücksichtigt werden.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 75: Konzeption eines fächerübergreifenden Evaluationsrahmens im Kontext videobasierter Lehrkräftefortbildungen zum digital gestützten Unterrichten

Benjamin Heinitz¹, Jennifer Janeczko¹, Habiba Mohtadi¹, Sylvia Gabel², Stephanie Moser³, Saskia Schley⁴, Andreas Gegenfurtner²

¹Universität Münster; ²Universität Augsburg; ³Technische Universität Münden; ⁴Freie Universität Berlin

Theoretischer Hintergrund

Die Förderung digitalisierungsbezogener Kompetenzen stellt vor dem Hintergrund aktueller Entwicklungen in der Lehrkräftebildung ein wichtiges Anliegen dar (Redecker, 2017). Im Rahmen des Kompetenzverbunds lernen:digital (<https://lernen.digital/>) werden deshalb Fortbildungen für Lehrkräfte mit verschiedenen Schwerpunkten entwickelt und in der Praxis eingesetzt. Dabei sind Gelingensbedingungen für erfolgreiche Fortbildungen von zentraler Bedeutung, jedoch müssen die Besonderheiten digitaler Anwendungen und die damit verbundenen Anforderungen an Schulen und Lehrkräfte berücksichtigt werden (Eickelmann & Drossel, 2020). Digitalisierungsbezogene Kompetenzen decken ein breites Spektrum an möglichen Fähigkeiten ab, wobei laut dem European Framework for the Digital Competence of Educators (DigCompEdu) sechs verschiedene Bereiche mit insgesamt 22 Kompetenzen unterschieden werden können (Redecker, 2017), sodass Fortbildungen innerhalb der Kompetenzbereiche fokussieren müssen. Diese sind häufig mit der Anwendung individueller Werkzeuge und Zugänge verbunden (z. B. im Umgang mit digitalen Tools zum außerschulischen Lernen oder im Kontext künstlicher Intelligenz), was eine fortbildungsübergreifende Evaluation erschweren kann.

Die einzelnen Fortbildungen der 12 Teilprojekte des Verbundprojekts „Videobasierte Fortbildungsmodule zum digital gestützten Unterrichten im Netzwerk bundesdeutscher Videoportale“ (ViFoNet) arbeiten mit unterschiedlichen digitalen Medien/Tools und beinhalten jeweils videobasierte Aufgabenformate mit anschaulichen Erklär- und Unterrichtsvideos zum prototypischen Einsatz dieser digitalen Ressourcen. Zusätzlich sind die Teilprojekte in unterschiedlichen Fächern verortet: Deutsch, Deutsch als Zweitsprache, Englisch, Französisch, Spanisch, Geographie, Wirtschaft und Chemie (mit Schwerpunkt Fachsprache) sowie Bildung für nachhaltige Entwicklung als übergreifendes Thema. Dies stellt für eine gemeinsame Evaluation der Fortbildungen eine Herausforderung dar.

Fragestellungen

Die unterschiedlichen Ansätze der Fortbildungen müssen für eine gemeinsame Evaluation berücksichtigt und gleichzeitig eine Vergleichbarkeit sichergestellt werden. Daraus haben sich folgende Forschungsfragen für die gemeinsame Evaluation ergeben:

1. Welche Rahmenbedingungen sollten für eine fächerübergreifende Evaluation videobasierter Lehrkräftefortbildungen berücksichtigt werden?
2. Inwiefern lassen sich Zusammenhänge zwischen den Voraussetzungen der Teilnehmenden und ihren digitalisierungsbezogenen Kompetenzen projektübergreifend abbilden?

Methoden

Im Rahmen des Verbundprojekts ViFoNet wurde ein fortbildungsübergreifender Evaluationsrahmen für ein Prä-Post-Follow-up-Design entwickelt. Dieser Rahmen beruht auf dem Angebots-Nutzungs-Modell für Lehrkräftefortbildungen (Lipowsky & Rzejak, 2019) und wurde spezifisch für die Bedarfe videobasierter Fortbildungen zum digital gestützten Unterricht adaptiert. Dabei werden die folgenden Bereiche adressiert: (1) *Voraussetzungen der Teilnehmer:innen* (Selbstwirksamkeitseinschätzungen beim Einsatz digitaler Medien), (2) *Schulkontext* (Transferklima an der Schule), (3) die wahrgenommene *Qualität und Quantität der Lerngelegenheiten* (mit Fokus auf der Nützlichkeit der eingesetzten Unterrichtsvideos als besonderem Fortbildungsmerkmal des ViFoNet-Verbundprojekts), (4) die *Wahrnehmung, Verarbeitung und Nutzung* der Fortbildungsinhalte, (5) die *Akzeptanz und Zufriedenheit der teilnehmenden Lehrpersonen*, (6) die *digitalisierungsbezogenen Kompetenzen der Lehrkräfte* und (7) die *Weiterentwicklung des unterrichtlichen Handelns* (Praktische Anwendung der Fortbildungsinhalte im eigenen Unterricht). Zusätzlich werden die bisherige Nutzung digitaler Medien sowie weitere soziodemographische Daten erhoben.

Für die Operationalisierung wurden vorrangig bestehende Skalen einbezogen, um die inhaltliche Validität sicherzustellen. Die digitalisierungsbezogenen Kompetenzen werden in Anlehnung an DigCompEdu (Redecker, 2017) durch eine Selbsteinschätzung der Lehrkräfte erfasst und umfassen vier Bereiche des 7-Faktor-Strukturmodells nach Quast et al. (2023). Dazu gehört die Planung von Unterricht (*Lesson Planning*) das Lehren und Unterstützen der Schüler:innen (*Teaching and Supporting Students*), die Förderung aller Schüler:innen (*Empowering All Students*) und die Förderung digitaler Kompetenzen der Schüler:innen (*Facilitating Students' Digital Competence*). Der Evaluationsrahmen wird in allen Fortbildungen des Projekts ViFoNet eingebunden und die Ergebnisse werden zentralisiert ausgewertet.

Ergebnisse

Die erste Forschungsfrage konnte bereits anteilig bei der theoriebasierten Erstellung des Evaluationsrahmens beantwortet werden. Um die Unterschiede der Fortbildungen der 12 Teilprojekte zu berücksichtigen, wurde ein spezifisches Framing für einzelne Skalen des Evaluationsrahmens entwickelt. Dieses Framing erlaubt es den Teilprojekten, ihre Besonderheiten hervorzuheben und dennoch projektübergreifende Vergleiche zu ermöglichen. Der Evaluationsrahmen inklusive des spezifischen Framings wird entsprechend im Rahmen des Beitrags vorgestellt.

Zur Beantwortung der zweiten Forschungsfrage werden erste Ergebnisse der fächerübergreifenden Evaluation zum Prätest vorgestellt. Dabei werden auch Zusammenhänge zwischen den Voraussetzungen der Teilnehmer:innen, dem Transferklima an ihren Schulen und den soziodemographischen Daten adressiert.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 001
Chair der Sitzung: Karina Karst

Pinnwand 19: Kooperatives Lernen in der Primarstufe – eine Videostudie zur Analyse der Qualität gruppenbezogener Lernprozesse

Anabel Bach, Felicitas Thiel

Freie Universität Berlin, Deutschland

Kooperatives Lernen ist im Unterricht mittlerweile weit verbreitet (Bleck & Lipowsky, 2022). Analog zum internationalen Raum zeigen auch Lehrkräftebefragungen in Deutschland, dass kooperative Formate von einer Mehrheit der Lehrkräfte regelmäßig im Unterricht durchgeführt werden (z.B. Stanat et al., 2023). Dass kooperative Settings im Mittel lernwirksam sein können, haben zahlreiche Metaanalysen und Reviews gezeigt (z.B. Kyndt et al., 2013). Bisherige Forschungsbefunde zeigen jedoch auch, dass lernförderliche Kooperationsprozesse in der Praxis eher selten stattfinden und das kognitive Niveau von Fragen oder Erklärungen häufig niedrig ist (z.B. Kronenberger & Souvignier, 2005; Renkl & Mandl, 1995; Van Boxtel et al., 2000). Den beteiligten Akteuren fehlen häufig die Kompetenzen, die Prozesse lernwirksam zu gestalten (z.B. Abramczyk & Jurkowski, 2020; Adl-Amini et al., 2023; Völlinger et al., 2018). Insbesondere für die Primarstufe liegen sehr wenige Studienbefunde vor (Borsch et al., 2007).

Um die Qualität kooperativer Lernprozesse zu untersuchen, ist es notwendig, die Tiefenstrukturmerkmale in den Blick zu nehmen. Für instruktionsbasierten, lehrerzentrierten Unterricht berücksichtigt die empirische Unterrichtsforschung diese schon seit einiger Zeit und es sind zahlreiche Instrumente entwickelt worden, die zur Evaluation von Unterrichtsqualität eingesetzt werden (Praetorius et al., 2018). Den spezifischen Anforderungen kooperativer wird dabei nicht ausreichend Rechnung getragen. Das Lernen erfolgt im Idealfall selbstgesteuert von den Schüler:innen; die Lehrpersonen müssen erst dann steuernd eingreifen, wenn die Gruppe allein nicht weiterkommt. Daraus folgt, dass insbesondere die Interaktionsqualität innerhalb einer Gruppe in den Blick genommen werden muss. Auf Prozessebene kann im Hinblick auf die Qualität der Gruppeninteraktion in Anlehnung an Kaendler et al. (2015) zwischen kognitiven, metakognitiven und sozial-relationalen Gruppenaktivitäten unterschieden werden.

Vor diesem Hintergrund wurde ein standardisierter Fragebogen entwickelt, der die Interaktionsqualität entlang der genannten Gruppenaktivitäten erfasst und bislang ausschließlich bei älteren Lerner:innen eingesetzt wurde. In der vorliegenden Studie soll das Instrument nun, mit sprachlich adaptiven Items, in der Primarstufe im Anschluss an eine Gruppenarbeit eingesetzt werden. Außerdem wird die kooperative Lernphase videografiert, so dass die Gruppenprozesse auch aus der Außenperspektive analysiert werden können.

Unter anderem sollen die folgenden Fragestellungen untersucht werden: Wie schätzen die Schüler:innen die Qualität der Gruppenprozesse im Hinblick auf die Transaktivität, das Gruppenklima, die aufgabenbezogene Kommunikation, den Lerngewinn und die Zufriedenheit ein? Inwieweit stimmen die Einschätzungen innerhalb der Gruppen überein? Inwieweit stimmen die Selbsteinschätzungen mit der Beobachterperspektive überein? Inwieweit zeigen sich Zusammenhänge der Prozessmerkmale mit den selbsteingeschätzten Outcomes? Untersucht werden soll außerdem, ob geschlechtsspezifische Unterschiede im kooperativen Verhalten vorliegen. Bisherige Befunde deuten darauf hin, dass im Hinblick auf kooperative Kompetenzen geschlechtsspezifische Unterschiede existieren. Eine PISA-Nebenauswertung (OECD, 2017) ergab etwa, dass Mädchen beim kooperativen Problemlösen bessere Ergebnisse erzielen als Jungen. Inwieweit sich dies bereits in der Primarstufe zeigt, ist nicht ausreichend untersucht.

Die Datenerhebung erfolgt an einer Grundschule in einer 3. Klasse mit insgesamt N = 21 Schüler:innen (7 Gruppen à drei Lernenden). Die Fragebogendaten werden unter Berücksichtigung der Mehrebenenstruktur quantitativ analysiert. Die transkribierten Gruppengespräche werden mittels qualitativer Inhaltsanalyse nach Mayring (2015) ausgewertet. Die Erhebung findet erst Anfang September 2024 statt, so dass zum jetzigen Zeitpunkt noch keine Ergebnisse präsentiert werden können.

Aus wissenschaftlicher Sicht können durch die Studie Erkenntnisse in Bezug auf die Qualität kooperativen Lernens in der Primarstufe und auf die Relevanz einzelner Prozessmerkmale gewonnen werden. Darüber hinaus ermöglicht die Studie durch den Perspektivenvergleich die Untersuchung der Validität der Selbsteinschätzungen der Grundschüler:innen. Ein schulpraktischer Nutzen ergibt sich aus den Ergebnissen insoweit, dass Lehrkräfte durch einen standardisierten Fragebogen eine evidenzbasierte Rückmeldung zu der Qualität der im Unterricht durchgeführten kooperativen Lernphase erhalten können. Durch den Einsatz eines solchen Instruments in der Praxis können Defizite in der Qualität der Interaktionsprozesse ermittelt werden und Optimierungsmöglichkeiten abgeleitet werden. Limitationen der Studie ergeben sich aus der geringen Stichprobengröße und des querschnittlichen Designs mit einem Messzeitpunkt.

Posters & Wine

Zeit: Montag, 27.01.2025: 19:00 - 20:15 · Ort: O 102
Chair der Sitzung: Karina Karst

Pinnwand 76: Expert Feedback on Responsive Teaching Behavior of Student Teachers in a Virtual Classroom Setting

Sinah Auctor¹, Nils Lauinger², Roland Brünken¹, Antje Biermann¹

¹Universität des Saarlandes, Deutschland; ²Forschungsinstitut Bildung Digital

Responsiveness is “a process in which interaction partners attend to and respond supportively to each other's needs, wishes, and circumstances, thereby promoting each other's well-being” (American Psychological Association, 2018). Responsive teaching behavior is one crucial factor in building of student-teacher-relationships, and has detrimental effects on disruptive behavior, learning and motivation (Cornelius-White, 2007, Hagenauer et al., 2015, Teistler, 2024, Tausch, 2017). Studies in teacher education show that expert feedback has a significant effect on the development of important competences (Prilop et al., 2020, 2021; Weber et al., 2018). While current studies emphasize the significant influence of expert guidance on competence development, a notable gap persists in the specific area of responsive teaching behavior. Addressing this gap, our study aims to provide a detailed assessment where student teachers receive expert feedback with video replay on their responsive teaching behavior in a virtual classroom situation, addressing the following research question: Does expert feedback improve student teachers' responsive teaching behavior in a virtual classroom?

This study follows a randomized two-group design with one intervention group (expert feedback, IG) and one control group (no feedback, CG) that teach two short self-developed introductory sequences in a virtual classroom scenario. One student disrupts the sequence. The student teachers' reaction towards this behavior is assessed by a trained student assistant. Depending on the student teacher's behavior, the classroom conflict could either escalate or fade. During the classroom sequence, one expert observes the student teacher's behavior and assesses selected aspects of responsive teaching behavior with a questionnaire (Protzel, 2015) as a basis for the following expert feedback with video replay for the IG. The CG reads a text about classroom management. After completing questionnaires, both groups teach a sequence in a second virtual classroom scenario, which is similar in the conflict addressed, but with different level of difficulty concerning aspects such as student language and behavior. Finally, there is a debriefing.

The dependent variable responsive teaching behavior is operationalized through the self-assessment and the assessment of the video data. Both groups self-assess their responsive teaching performance in the classroom after each sequence (Akram & Zepeda, 2015, King, 2017). After the experiment, the video-data is assessed by two independent raters with the responsiveness questionnaire (Protzel, 2015). Further, the perceived authenticity of the classroom and conflict situation is assessed (Seufert et al., 2022, Gomes et al., 2013, Poeschl & Doering, 2013). All questionnaires are on a 5-point Likert-scale.

The hypothesis will be tested using a MANOVA with the between-subjects factor Feedback (Feedback vs. no Feedback). We will use this to analyze the dependent variables responsive teaching behavior and self-assessment of responsive teaching behavior.

From the 49 participants we are determined to acquire until the end of the survey period, we have reached $N = 20$ ($f = 15$, $m = 5$; $M_{\text{age}} = 24.95$ years, $SD_{\text{age}} = 6.05$ years) participants for now. The analysis of the expert rating is still in progress. Therefore, we only have preliminary results of one of the two dependent variables. The participants are evenly distributed into both groups (11 IG, 9 CG) and at an intermediate level of experience ($M_{\text{semester}} = 5.90$, $SD_{\text{semester}} = 2.94$). The classroom is perceived as partly realistic and the virtual conflict was recognized ($M = 3.00$, $SD = 1.08$) and comparable to a real-life conflict ($M = 4.15$, $SD = .93$).

In the current sample, we found a slightly stronger positive trend in the intervention group in their self-assessment between the first and second virtual setting ($M_{\text{IG 1st}} = 3.39$, $M_{\text{IG 2nd}} = 4.10$, $M_{\text{CG 1st}} = 3.33$, $M_{\text{CG 2nd}} = 3.60$). We expect to finish data collection and analysis before the conference.

Pinnwand 77: Demographische Faktoren als Prädiktoren von Leistungsangst bei Grundschulkindern

Ann-Kathrin Schnelle¹, Juliane Rutsch¹, Nicolas Hübner², Eveline Ruth-Herbein¹, Torsten Eckermann³, Benjamin Fauth¹

¹Institut für Bildungsanalysen Baden-Württemberg; ²Universität Tübingen; ³Pädagogische Hochschule Freiburg

Demographische Faktoren als Prädiktoren von Leistungsangst bei Grundschulkindern

Ann-Kathrin Schnelle, Juliane Rutsch, Nicolas Hübner, Evelin Ruth-Herbein, Torsten Eckermann & Benjamin Fauth

Theorie

Im Schulkontext spielt Leistungsangst eine wesentliche Rolle für das Wohlbefinden und die schulischen Erfolge von Schülerinnen und Schülern. Leistungsangst bei Schulkindern ist eine spezifische Form der Angst, die sich in Befürchtungen und Sorgen um die eigenen schulischen Leistungen oder die Reaktionen darauf manifestiert (Martschinke & Kammermeyer, 2006). Während beispielsweise zur Prüfungsangst bereits umfangreiche Studien vorgelegt wurden (vgl. Pekrun & Linnenbrink-Garcia, 2012) und obwohl Angst die am häufigsten rückgemeldete Leistungsemotion ist (Pekrun et al., 2007) und den gesamten Bildungsverlauf erheblich beeinflussen kann (Martschinke & Kammermeyer, 2006), fehlen derzeit längsschnittliche Studien zu Wirkmechanismen von Leistungsangst (Lorhmann, 2023), besonders im Primärbereich (Martschinke & Kammermeyer, 2006).

Bisherige Studien konnten Zusammenhänge zwischen Leistungsangst und dem Geschlecht feststellen (Schröder-Lenzen, 2007). Dabei zeigte sich, dass Mädchen ein höheres Leistungsangstniveau aufweisen (Faber, 2010).

Im Zusammenhang mit Leistungsangst konnte zudem nachgewiesen werden, dass Kinder aus benachteiligten Familien häufig zusätzlichen Belastungen und höheren Erwartungen ausgesetzt sind, was die Leistungsangst verstärken kann (Bornstein & Bradley, 2014). Allerdings gibt es bisher kaum mehrdimensionale Betrachtungen, die diese beiden Heterogenitätsmerkmale im Zusammenhang mit der Leistungsangst untersuchen.

Forschungsfragen

Im vorliegenden Beitrag soll folgenden Forschungsfragen nachgegangen werden:

1) In welchem Zusammenhang stehen der sozioökonomische Status und das Geschlecht mit Leistungsangst und der Entwicklung dieser?

2) Findet sich eine zusätzliche Benachteiligung von Schülerinnen und Schülern bei einer verschränkten Betrachtung der beiden Heterogenitätsmerkmale?

Methodik

Die vorliegenden Daten wurden im Rahmen einer längsschnittlichen Befragung bei Zweitklässlerinnen und Zweitklässlern ($N=866$) zu Beginn (t1, Oktober 2022) und am Ende des Schuljahres (t2, Juli 2023) erhoben. Erfasst wurden neben der Leistungsangst zu t1 ($\alpha = .79$) und t2 ($\alpha = .80$) (Wendt et al., 2006), das Geschlecht (0 = männlich, 1 = weiblich) sowie der sozioökonomische Hintergrund (Bücherfrage, Wendt et al., 2016, Heppt et al., 2022). Es wurde eine multiple Regressionsanalyse in Mplus (Muthén & Muthén, 2017) mit robusten Standardfehlern durchgeführt (Type = complex), um die genestete Struktur der Daten (Schulkinder in Klassen) zu berücksichtigen.

Ergebnisse

Bei der Vorhersage der Leistungsangst der Kinder am Ende der zweiten Klasse erwies sich ihre Leistungsangst am Anfang der zweiten Klasse als stärkster Prädiktor ($d = 0.394$, $p < .001$). Auch für das Geschlecht fand sich ein statistisch signifikanter Zusammenhang mit der Leistungsangst am Ende der zweiten Klasse, nach Kontrolle der Leistungsangst am Anfang der zweiten Klasse ($d = 0.226$, $p = .017$), wobei Mädchen von einer substantiell höheren Leistungsangst berichteten als Jungen. Für das soziale Kapital ($d = -0.009$; $p = .841$) sowie die Interaktion aus Geschlecht und sozialem Kapital ($d = -0.109$; $p = 0.259$) fanden sich keine statistisch signifikanten Zusammenhänge mit der Leistungsangst.

Diskussion

Die Ergebnisse dieser Studie deuten darauf hin, dass das Geschlecht ein wichtiger Prädiktor für die Entwicklung von Leistungsangst in der Grundschule ist. Diese Bedeutung als Prädiktor für Leistungsangst wird durch Studien bestätigt, die geschlechtsspezifischen Unterschiede in der schulischen Angst und in den Leistungsanforderungen identifizieren (Else-Quest et al., 2010). Obwohl die Ergebnisse früherer Forschungen auf einen Zusammenhang zwischen sozioökonomischem Status und psychischer Belastung hinweisen (Bradley & Corwyn, 2002, 1998; Hattie & Timperley, 2007), konnte dieser Effekt in der vorliegenden Untersuchung nicht gefunden werden. Im Vortrag sollen weitere Maße berücksichtigt werden, zum Beispiel den über die Eltern erhobenen International Socio-Economic Index of Occupational Status (ISEI) und u. a. dessen Zusammenhang mit der Leistungsentwicklung der Schülerinnen und Schülern. Darüber hinaus wäre noch zu prüfen, ob sich diese höhere Leistungsangst bei den Grundschülerinnen und Grundschülern auf die Leistungsentwicklung auswirkt (Rost & Haferkamp 1984) und inwieweit sich diese fachspezifisch unterscheidet.

Pinnwand 78: Connected learning on social media: algorithmic practices of youth

Santiago Hurtado, Anna Keune

Technical University of Munich, Germany

Computer science (CS) still faces lopsided representation. In the past decade, women accounted for 20% of information and communications technology (ICT) graduates, which represents 2% of women graduates in total (European Commission, 2024). CS competencies include algorithmic thinking, which refers to an understanding of steps executed in specific orders to achieve a result, and is foundational for developing further CS competencies (Computer Science Teachers Association, 2017). To counter lopsided representation in CS, it is necessary to support young people's different approaches to learning by honoring youth culture and interests in learning environments (Ito et al., 2020). One approach that supports youth interests in personally meaningful and long-lasting learning experiences is the connected learning framework, which highlights youth-driven environments, such as social media, as crucial for learning (Ito et al., 2020). Young people use social media daily to share, create, and view content while connected with others through evolving social media trends (Aichner et al., 2021). We investigate social media as a powerful youth-driven context through the connected learning principles of shared practices and shared purpose, as well as youth social media practices similar to algorithms (Authors, 2023; Author, 2024) that could impact young people in understanding foundational CS concepts.

Research question

How does engaging in connected youth practices with social media involve algorithmic practices and connected learning?

Method

This qualitative study involved observing the online practices of 13 girls from Latin America and Europe through social media walkthrough interviews that lasted 70 minutes on average (15 hours in total). The walkthrough interviews focused on four main topics: (1) experiences with social media, (2) connected learning on social media, (3) algorithmic learning, and (4) risks on social media. We analyzed the interviews by thematically and iteratively coding connected learning principles (i.e., shared practices, shared purposes) and algorithmic practices (i.e., flow control structures, which include conditional and linear statements; loops, which include iterative conditions; and debugging, which include probing the inner workings of the algorithms).

Results and their significance

All participants showcased approaches to social media through steps to create, share, and curate content. In the interest of space, we present one representative case: Sarah (pseudonym), a 15-year-old student in Germany who used social media apps since she was nine years old.

Sarah showcased shared practices in her quick response to real-life events through the collaborative production of information toward a shared goal of saving lives. Sarah used social media to share information about the 2023 earthquake in Turkey, concerned about people's lives to assist first responders in finding people. Sarah also showcased a shared purpose in making meaningful contributions to her community and finding a clear pathway. She shared real-time information to make meaningful contributions to a personally meaningful cause with culturally relevant values (i.e., assisting first responders).

The shared practices toward a shared purpose were linked to algorithmic practices. Sarah's posting practices were similar to flow control structures by following steps that relied on conditional statements of how, when, where, and what to post. She retrieved necessary information and decided when, and how to share it while including links and hashtags to reach desired audiences. Moreover, Sarah's editing practices were similar to loops, as she edited or deleted content if the information shared was inaccurate. Finally, Sarah shared debugging practices as she customized her feed through liking and searching and noticed that her feed changed depending on her actions on social media.

This study points toward social media as a personally meaningful context for young people where their practices, developed through years-long experiences on social media, relate to CS concepts to support entry points into the field.

Pinnwand 79: Unterrichtsgesprächsführung im Praxissemester lernen: Die Rolle der Mentor*innen für die Entwicklung der eigenen Einstellungen

Florentine Hickethier, Alexander Gröschner
Friedrich-Schiller-Universität Jena, Deutschland

Theoretischer Hintergrund

In den Praxisphasen im Lehramtsstudium erhalten Studierende die Möglichkeit, fremden Unterricht zu beobachten und eigene Unterrichtsversuche durchzuführen. Auf diese Weise können die Studierenden ihre Unterrichtswahrnehmung verbessern und eigene Einstellungen zum Unterricht entwickeln (Ulrich et al., 2020). Damit das Praktikum einen nachhaltigen Effekt für das Lernen der Studierenden erzielt und diese sich weiterentwickeln können, wird der Qualität der Betreuung eine große Bedeutung zugeschrieben (Kreis & Staub, 2011). Die betreuenden Lehrpersonen an den Schulen stellen häufig Vorbilder für das Handeln im Unterricht dar und zählen in dieser Phase zu den wichtigen Bezugspersonen (Hascher & Moser, 2001; Rothland & Straub, 2018). Oftmals zeigt die Realität an den Praxissemesterschulen jedoch deutliche Probleme hinsichtlich der Betreuung auf, da beispielsweise die Betreuung häufig nicht aus eigenen Stücken heraus von den Lehrpersonen übernommen wird, sondern von der Schulleitung vergeben wird (Gröschner & Häusler, 2014). Diese und weitere Faktoren können sich auf die Qualität der Betreuung auswirken, was die Notwendigkeit zur Forschung am Einfluss der Mentorenrolle auf die Herausbildung eigener Überzeugungen der angehenden Lehrpersonen herausstellt. Unterrichtsgespräche als ein wesentlicher Bestandteil der unterrichtlichen Praxis sollen dabei gesondert ins Zentrum der Untersuchung rücken (Gröschner et al., 2022).

Fragestellung

In der aktuellen Studie soll daher die Rolle des/der Mentors/Mentorin für die Einstellungen von Praxissemesterstudierenden zur dialogischen Gesprächsführung in den Fokus rücken. Generell soll der Frage nachgegangen werden, inwiefern das Mentor*innenhandeln die Einstellungen von Studierenden prägt. Spezieller soll auch darauf eingegangen werden, inwiefern die wahrgenommene Unterstützung der Mentor*innen einen Einfluss auf die Einstellungen von Studierenden bezüglich dialogischer Gesprächsführung hat.

Methode

Insgesamt haben $N = 387$ Lehramtsstudierende (Alter: $M = 23.05$ Jahre, $SD = 2.58$ Jahre; Geschlecht: 68.7% weiblich, 29.5% männlich, 1.8% keine Angabe) der Universität Erfurt ($n = 137$ Studierende) und der FSU Jena ($n = 250$ Studierende) an zwei Online-Fragebögen teilgenommen, die jeweils zum Beginn und zum Ende des Praxissemesters durchgeführt wurden. Die Stichprobe wurde in zwei Kohorten erfasst (Sommersemester 2021 und Wintersemester 2021/2022). Zur Bearbeitung der Forschungsfragen wurde in den Fragebögen unter anderem eine Skala zur Wahrnehmung dialogischer Gesprächsführung im eigenen Unterricht eingesetzt. Der *Dialogic Teaching Questionnaire* (DTQ; Gröschner et al., 2020) umfasst 15 Items, die sich drei Subskalen zuordnen lassen: (1) Rahmung für dialogische Gesprächsanlässe (*Rahmung*), (2) Initiierung von Schüler*innenbeiträgen (*Initiierung*) und (3) Förderung dialogischer Beteiligung (*Förderung*). Die einzelnen Items wurden von den Studierenden auf einer sechs-stufigen Likert-Skala (1 = „stimme überhaupt nicht zu“; 6 = „stimme voll und ganz zu“) beantwortet. Zudem kamen weitere Skalen zum Einsatz, die die Wahrnehmung des Unterrichts der Mentor*innen bezüglich dialogischer Gesprächsführung sowie die wahrgenommene Unterstützung der Mentor*innen erfassten (Gröschner et al., 2020, Gröschner, Schmitt & Seidel, 2013).

Nach Abschluss des Praxissemesters wurden zudem Gruppeninterviews mit $N = 29$ Studierende (69.0% weiblich, 31.0% männlich) durchgeführt, bei denen Fragen zur generellen Wahrnehmung des Praxissemesters, der Wahrnehmung der Unterstützung der Betreuungslehrpersonen und den Möglichkeiten zur Initiierung dialogischer Unterrichtsgespräche thematisiert wurden.

Ergebnisse

Eine erste Regressionsanalyse zeigte, dass die Wahrnehmung des Unterrichts der Mentor*innen bezüglich dialogischer Gesprächsführung die Einstellungen zum eigenen Unterricht signifikant auf allen Subskalen vorhersagt, *Rahmung* ($\beta = .49, p < .001$), *Initiierung* ($\beta = .48, p < .001$), *Förderung* ($\beta = .54, p < .001$).

Für den Posterbeitrag sollen weitere Erkenntnisse zum Einfluss der wahrgenommenen Unterstützung der Mentor*innen geleistet werden. Zudem sollen deskriptive Daten sowie Auszüge aus den Interviews illustrierend herangezogen werden, um Aufschlüsse über die Beziehung und den Wert der Beziehung zu den Mentor*innen für die persönliche Professionalisierung der angehenden Lehrpersonen zu liefern. Erste Ergebnisse aus den Interviewauswertungen zeigen, dass besonders das Erhalten von Feedback bezüglich des Verhaltens der Studierenden in eigenen Unterrichtsversuchen oder zu fachlichen bzw. fachdidaktischen Inhalten für die angehenden Lehrpersonen von besonderer Relevanz ist.

Pinnwand 80: Effekte eines VR-Trainings zur Kompetenzentwicklung und Reflexion im Rahmen von Elterngesprächen

Friederike Knabbe, Yannik A. Escher, Hannes M. Petrowsky, David D. Loschelder, Poldi Kuhl
Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Die Aufgaben von angehenden Lehrkräften sind mit zahlreichen Herausforderungen verknüpft, wovon die Kommunikation mit Eltern ein wichtiger Aspekt ist. Das Führen von Elterngesprächen fordert spezifische Beratungskompetenzen von Lehrkräften. Obwohl die Beratung im Rahmen von Elterngesprächen zu den alltäglichen Kernaufgaben von Lehrkräften gehört (Hertel, 2009; KMK, 2022), findet im universitären Kontext über die theoretische Auseinandersetzung hinaus kaum eine Vorbereitung darauf statt. Neuere technologische Entwicklungen, insbesondere die virtuelle Realität (VR), bieten Lehramtsstudierenden eine innovative, realistische und aktive Übungsmöglichkeit für die Erprobung von Elterngesprächen. Bislang wurde jedoch noch nicht untersucht, inwieweit VR für das Üben solcher Situationen im Rahmen der Lehrerbildung geeignet ist. Auch mögliche vermittelnde Faktoren wie Ängste, Selbstwirksamkeitserwartungen oder motivationale Aspekte sind bisher nicht untersucht worden.

In einer kürzlich durchgeführten Pilotstudie mit $N = 15$ Lehramtsstudenten wurde herausgefunden, dass die Teilnehmenden hoch motiviert waren, VR als Übungsmöglichkeit zu nutzen, an Beratungskompetenz gewannen und eine Abnahme der aufgabenbezogenen Angst im Verlauf des VR-Trainings stattfand (Knabbe et al., 2024).

Um die bisherige Forschung zu erweitern und ein Verständnis für die Prozesse zu entwickeln, die VR-basierten Trainings zugrunde liegen, wurde die Hauptstudie durchgeführt. In der das „Cognitive Affective Model of Immersive Learning“ (CAMIL; Makransky & Petersen, 2021) empirisch getestet und erweitert wird. Dieses theoretische Modell besagt, dass Präsenz und Handlungsfähigkeit die allgemeinen psychologischen Voraussetzungen des Lernens in VR sind (Makransky & Petersen, 2021). Das Ziel der vorliegenden Untersuchung ist es, die Effektivität eines hoch immersiven VR-Trainings für Elterngespräche mit einem chat-basierten Trainingsprogramm zu vergleichen und zu evaluieren. Erstens durch die Konstruktion eines theoretischen Modells der Entwicklung sozialer Fähigkeiten in der immersiven Umgebungen, d.h. durch die Erweiterung des CAMIL-Modells, zweitens aber auch durch die Prüfung des Modells und den Vergleich der Lerneffekte mit einer starken Kontrollbedingung. Nach vier Monaten wird eine Nachbefragung durchgeführt, um die Langzeitwirkung des Trainings zu untersuchen.

Die Studie folgt einem Prä-Post-Interventionsdesign und untersucht die selbst eingeschätzten Effekte eines neu entwickelten VR-Trainings für Elterngespräche auf die Selbstwirksamkeitserwartung bezüglich schulischer Beratungsgesprächen und die Angst vor Elterngesprächen.

Methode

Derzeit wird die Hauptstudie durchgeführt und eine Gesamtstichprobe von $N = 100$ angestrebt. Bei den Teilnehmenden handelt es sich um Lehramtsstudierende der Bachelor- und Masterstudiengänge der Leuphana Universität Lüneburg. Die meisten Teilnehmenden haben noch keine Erfahrung mit VR, und nur wenige haben bereits Erfahrungen mit Elterngesprächen sammeln können.

Zur Vorbereitung erhielten die Studierenden eine Fallvignette mit Informationen über das zu führende Elterngespräch. Sie wurden randomisiert entweder (1) einem virtuellen 3D-Klassenraum mit 180°-Blickwinkel und VR-Brille oder (2) der Kontrollbedingung mit Chat zugewiesen. In beiden Bedingungen mussten die Teilnehmenden ein problemorientiertes Gespräch mit einem Elternteil führen. Die Versuchsbedingung in VR war videobasiert, und die virtuelle Elternfigur, die von einer professionellen Schauspielerin dargestellt wurde, wurde von der Trainingsleitung durch Auswahl des entsprechenden Videos gesteuert. Die Kontrollbedingung basierte auf einer Chat-Nutzeroberfläche für die Teilnehmenden, die auf einem Laptop vor ihnen platziert wurde. Die Teilnehmenden erhielten Chat-Nachrichten von dem Elternteil und antworteten schriftlich. Die Fragen des Elternteils waren identisch mit denen der VR-Bedingung formuliert. Zur Beantwortung der Forschungsfrage nach den Auswirkungen des Trainings wurde eine Pre- und Postbefragung der Studierenden durchgeführt. In beiden Fragebögen wurden die Selbsteinschätzungen erhoben, unter anderem zu Selbstwirksamkeitsüberzeugungen (De Coninck et al., 2020) und Ängsten (McCarthy und Goffin, 2004).

Ergebnisse und Diskussion

Die Studie wird seit dem Wintersemester 2023/24 mit einer angestrebten Anzahl von $N = 100$ Teilnehmern durchgeführt. In der Präsentation werden die Ergebnisse der Hauptstudie vorgestellt und die Auswirkungen des VR-Trainings auf die Selbstwirksamkeitsüberzeugungen und die Ängstlichkeit der Teilnehmenden berichtet. Auf der Grundlage der vorangegangenen Pilotstudie erwarten wir, dass die Selbstwirksamkeit zunimmt und die Ängstlichkeit abnimmt. Abschließend werden die Implikationen für den Einsatz von VR im Rahmen der Lehrkräfteausbildung diskutiert.

Pinwand 81: Wahrnehmung des digitalen Medieneinsatzes im Unterricht durch die Lehrkraft aus Sicht der SchülerInnen

Begüm Arvaneh¹, Sabrina Reith¹, Karsten Stegmann², Maria Bannert¹

¹Technische Universität München, Deutschland; ²Ludwig-Maximilians-Universität München, Deutschland

Theoretischer Hintergrund

Digitale Medien bieten leistungsstarke Möglichkeiten zur Unterstützung des Lernens (Tamim et al., 2011). Daher ist die Fähigkeit von Lehrkräften geeignete digitale Medien im Unterricht einzusetzen essenziell (Lachner et al., 2019). Studien wie ICILS (*International Computer and Information Literacy Study*) und Bitkom haben jedoch gezeigt, dass Lehrkräfte Schwierigkeiten haben, digitale Medien differenziert im Unterricht einzusetzen und auch eine Verbesserung der Lehrerbildung in Bezug auf digitale Medien gewünscht wird (Eickelmann et al., 2019; Bitkom, 2019). Zur Einschätzung, wie bestimmte digitale Informations- und Kommunikationstechnologien genutzt werden können, um wirksame Lehr- und Lernprozesse zu initiieren, wurde von Lachner et al. (2011) das TPK-Instrument (*Technological Pedagogical Knowledge*) entwickelt. Dieses erfasst unter anderem das situative Wissen der Lehrkräfte über den Einsatz digitaler Medien im Unterricht. Jedoch gibt es kaum Studien, welche betrachten, ob das Wissen der Lehrkräfte auch zum Einsatz geeigneter digitaler Medien in verschiedenen Unterrichtsphasen führt. Daher ist es notwendig zu untersuchen, ob Lehrkräfte geeignete digitale Medien auswählen und in verschiedenen Unterrichtsphasen einsetzen und, ob die ausgewählten digitalen Medien eher zu einem rezeptiven oder produktiven Einsatz bei den SchülerInnen führen. Hierfür wurde im Rahmen der Studie basierend auf dem TPK-Instrument ein neues Instrument entwickelt, welches die Wahrnehmung des Einsatzes digitaler Medien im Unterricht aus Sicht der SchülerInnen erfasst.

Fragestellung

Mit diesem Hintergrund sollen dieser Studie folgende Forschungsfragen beantwortet werden:

1. *Inwieweit zeigt sich ein Zusammenhang zwischen dem Technologisch-pädagogischen Wissen der Lehrkräfte und dem wahrgenommenen Einsatz digitaler Medien im Unterricht?*
2. *Inwieweit besteht ein Zusammenhang zwischen der Wahrnehmung des Einsatzes digitaler Medien im Unterricht und der Computernutzung von SchülerInnen?*
3. *Inwieweit zeigt sich ein Zusammenhang zwischen dem Technologisch-pädagogischen Wissen der Lehrkräfte und dem Einsatz produktiver bzw. rezeptiver wirkender digitaler Medien?*

Methode

Zur Beantwortung der Forschungsfragen wurden im Rahmen einer längsschnittlichen Studie insgesamt in Pre- und Posttest sowohl Lehrkräfte ($N = 328$) als auch SchülerInnen ($N = 3851$) der 8. Klasse an Mittelschulen, Realschulen und Gymnasien in Mathematik und Biologie zu TPK, Wahrnehmung des Einsatzes digitaler Medien im Unterricht und der Computernutzung von SchülerInnen befragt. Zur Erfassung der Computernutzung der SchülerInnen wurden Items der NEPS-Instrumente für SchülerInnen ausgewählt. Das TPK-Instrument von Lachner et al. (2019) wurde genutzt, um das Wissen der Lehrkräfte zum geeigneten Medieneinsatz im Unterricht zu messen. Zudem wurde zum Abgleich das TPK-Instrument adaptiert, um ein neues Instrument zu entwickeln. Während ein Teil der ursprünglichen TPK-Skala das situative technologisch-pädagogische Wissen der Lehrkräfte bewertet und sie auffordert, konkrete Anwendungssituationen der Integration digitaler Medien im Unterricht zu beurteilen, wurden die Vignetten und die entsprechenden Items für das neue Instrument vertauscht und aus der Perspektive der

SchülerInnen umformuliert. Somit misst das neue Instrument die von den SchülerInnen wahrgenommene Nutzung digitaler Medien oder Tools im Unterricht durch die Lehrkraft.

Für die Rasch-skalierte Berechnung des wahrgenommenen rezeptiven Einsatzes digitaler Medien im Unterricht aus Sicht der SchülerInnen liegt final eine Stichprobengröße von $N = 3152$ vor (WLE-Reliabilität = 0.62, Varianz = 2.35). Die Berechnung des produktiven Einsatzes weist ebenfalls eine ausreichende Stichprobe von $N = 3138$ auf (WLE-Reliabilität = 0.60, Varianz = 2.12).

Ergebnisse und ihre Bedeutung

Erste Ergebnisse deuten darauf hin, dass Lehrkräfte zwar ein hohes Wissen über technologisch-pädagogische Kompetenzen aufweisen, dieses Wissen jedoch nicht im Unterricht nutzen. Die Daten werden derzeit weiter ausgewertet und sollen auf der GEBF 2025 präsentiert werden. Die Ergebnisse können einen Einblick in die tatsächliche Nutzung geeigneter digitaler Medien im Unterricht geben und auch, wie die eingesetzten Medien von SchülerInnen genutzt werden können. Zudem soll aufgezeigt werden, wo noch Handlungsbedarf besteht und welche Möglichkeiten es gibt, um die Nutzung digitaler Medien in verschiedenen Unterrichtsphasen zu verbessern.

Pinnwand 82: Konstruktion eines Rechtschreibfehleridentifikationstests für die Klassenstufen 7 und 8

Verena Keimerl, Jan Lenhart

Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund

Der Erwerb von Rechtschreibkompetenz spielt nicht nur in der Grundschule, sondern auch in der Sekundarstufe trotz zunehmender Verbreitung von digitalen Hilfs- und Korrekturprogrammen eine zentrale Rolle. Angesichts der von Lehrkräften häufig kritisierten klassischen Diktate (z.B. Birkel, 2012) stellen Fehleridentifikationstests eine besonders ökonomische Alternative zur Erfassung der rezeptiven Rechtschreibkompetenz dar (vgl. Lenhart, Segerer, Marx & Schneider, 2020). Daher sollen die im deutschsprachigen Raum etablierten Fehleridentifikationstests (R-FIT 5–6+, Schneider, Martínez Méndez & Hasselhorn, 2014; R-FIT 9–10, Lenhart et al., 2020) durch die Entwicklung des Fehleridentifikationstests für die Klassenstufen 7 und 8 komplettiert werden.

Empirisch lässt sich auch in der Sekundarstufe eine hohe Zunahme von Rechtschreibleistungen in (Lückentext-)Diktaten finden (Prosch 2016), wobei Mädchen im Durchschnitt eine bessere Leistung in produktiven Rechtschreibtests erzielen (May & Ivanov, 2011). Schließlich lassen sich deutliche Leistungsunterschiede zwischen den Schulformen nachweisen (Lenhart et al., 2020).

Fragestellungen

- (1) Lassen sich Unterschiede in der Rechtschreibkompetenz mit einem Fehleridentifikationstest abbilden?
- (2) Misst der Fehleridentifikationstest die rezeptive Rechtschreibkompetenz reliabel?
- (3) Liegen Klassenstufen-, Schulform- und Geschlechtsunterschiede im Fehleridentifikationstest vor?

Methode

Die Itemkonstruktion des R-Fit 7-8 basiert auf *Large-Scale-Assessments* zur Rechtschreibleistung (z.B. KESS 7 bis 11, May & Ivanov, 2011) und Rechtschreibfehlerkategorien zu gängigen Rechtschreibfehlern in der Sekundarstufe (z.B. Oldenburger Fehleranalyse; Thomé & Thomé, 2014). Es wurden Fehler aus 15 Rechtschreibfehlerkategorien (z.B. Dehnungs-h) in 34 Wörter eines Fließtextes integriert. Die Aufgabe der Schülerinnen und Schüler bestand darin, die Fehler zu markieren. In der Pilotierung bearbeiteten 202 Schülerinnen und Schüler aus 12 Klassen zweier Gymnasien und zweier Mittelschulen der Klassenstufen 6 und 8 beide Parallelversionen des *R-Fit 7-8* in randomisierter Reihenfolge. Ein Großteil der der Schülerinnen und Schüler (67%) sprachen zuhause ausschließlich deutsch und 66% der Teilnehmenden waren weiblich.

Ergebnisse und ihre Bedeutung

Von den 34 Rechtschreibfehlern wurden im Durchschnitt etwa 12 bis 13 Fehler korrekt markiert (Version A: $M = 12.30$, $SD = 8.73$; Version B: $M = 13.18$, $SD = 8.82$), wobei die Leistungsspanne in beiden Versionen zwischen 0 und 33 korrekt markierten Fehlern lag. Wie intendiert, traten hingegen nur etwa 2 Falschmarkierungen in korrekt geschriebenen Wörtern (Version A: $M = 2.27$, $SD = 2.87$; Version B: $M = 1.95$, $SD = 2.47$) und kaum Bearbeitungsfehler (z.B. Doppelmarkierungen) (Version A: $M = 0.20$, $SD = 0.51$; Version B: $M = 0.19$, $SD = 0.56$) auf.

Die interne Konsistenz der beiden Parallelversionen war als exzellent zu bezeichnen (Version A: $\alpha = .93$; Version B: $\alpha = .94$). Die Paralleltestreliabilität bzw. die Korrelation beider Testversionen war beinahe perfekt ($r = .95$, $p < .001$).

Es zeigten sich große Leistungsunterschiede zwischen Gymnasien und Mittelschulen zugunsten der Gymnasien (Version A: $p < .001$, $\eta_p^2 = .379$; Version B: $p < .001$, $\eta_p^2 = .436$). Ebenso ergaben sich bessere Leistungen für Schülerinnen und Schüler der 8. Klassenstufe verglichen mit der 6. Klassenstufe (Version A: $p = .003$, $\eta_p^2 = .044$; Version B: $p < .001$, $\eta_p^2 = .055$). Zudem ergab sich eine signifikante Interaktion zwischen Schulart und Klassenstufe (Version A: $p = .001$, $\eta_p^2 = .052$; Version B: $p < .001$, $\eta_p^2 = .078$), die darauf zurückzuführen war, dass die Verbesserung nur im Gymnasium aber nicht in der Mittelschule zu beobachten war. Die Mädchen wiesen bei Kontrolle der Klassenstufe nur in Version B signifikant höhere Kennwerte als die Jungen auf (Version A: $p = .208$, $\eta_p^2 = .009$, Version B: $p = .022$, $\eta_p^2 = .029$).

Der Fehleridentifikationstest bildet gut und mit hoher Reliabilität Unterschiede zwischen den Schülerinnen und Schülern ab und es lassen sich die erwarteten Unterschiede zwischen den Schulformen und Klassenstufen sowie teilweise für das Geschlecht finden. Die insgesamt niedrigen Mittelwerte legen empfehlen eine Hinzunahme von leichteren Items für die weitere Testkonstruktion.

Pinnwand 83: Professionswissen und diagnostische Urteile von Chemielehrkräften

Mathias Lutz¹, Markus Rehm¹, Hendrik Lohse-Bossenz²

¹Pädagogische Hochschule Heidelberg, Deutschland; ²Universität Greifswald, Deutschland

Theoretischer Hintergrund und Fragestellung

Nach Praetorius und Südkamp (2017) lässt sich diagnostische Kompetenz zusammenfassend als „die Fähigkeit, Schüler- und Aufgabenmerkmale korrekt zu beurteilen“ (S.14) beschreiben und wird als Grundstein für professionelles Handeln angesehen. Situationen, in welchen im Kontext des Unterrichts diagnostische Urteile gefällt werden, sind vielfältig: Die Beurteilung eines/einer Schüler:in, das Einschätzen und die Auswahl des Lehrmaterials oder die allgemeine Bewertung einer Lehr-Lern-

Situation können solche Diagnosesituationen darstellen (Bolte et al., 2024; Loibl et al., 2020). Das Diagnostizieren wird hierbei als Informationsverarbeitungsprozess verstanden, um bildungsbezogene diagnostische Urteile zu fällen. Aktuelle Forschungsbemühungen fokussieren auf interindividuelle Unterschiede im Prozess der Urteilsbildung und nehmen darauf wirkende Einflüsse in den Blick (Becker et al., 2020; Heeg et al., 2021; Hoppe et al., 2020; Loibl et al., 2020; Rieu et al., 2024; Rieu et al., 2020; Volkmer, 2024). Professionswissen stellt eine mögliche Erklärungsvariable im diagnostischen Prozess dar. Im Hinblick auf die Güte diagnostischer Urteile scheinen das fachdidaktische Wissen und das Fachwissen einer Lehrkraft einen wichtigen Einfluss zu haben (Blömeke et al., 2014; Dreher & Leuders, 2021; Schreiter et al., 2021, 2022). Für das Diagnostizieren von Lernendenvorstellungen konnten empirische Studien bereits einen Zusammenhang zwischen den diagnostischen Urteilen und Wissensfacetten des fachdidaktischen Wissens und des Fachwissens aufzeigen (Hoppe et al., 2020; Lembens et al., 2024; Rath, 2017).

Da die bisherigen Befunden das Ergebnis diagnostischer Prozesse betrachtet haben, knüpft das vorliegende Projekt hier an und untersucht die Informationsverarbeitungsprozesse angehender Chemielehrkräfte bei der Diagnose von Lernendenvorstellungen (Schüler:innen) für das Schulfach Chemie: Welchen Einfluss haben das fachdidaktische Wissen und das chemiespezifische Fachwissen auf das Diagnostizieren, den zu Grunde liegenden Informationsverarbeitungsprozess und das resultierende diagnostische Urteil?

Methode

Mithilfe von quantitativen Analysen wird der Einfluss von Fachwissen und fachdidaktischem Wissen auf diagnostische Urteile beim Umgang mit Lernendenvorstellungen in unterrichtsnahen Situationen anhand von Vignetten untersucht ($n=140$ angehende Chemielehrkräfte aus der ersten und zweiten Phase der Lehrer:innenbildung). Es wird eine einmalige Online-Fragebogenerhebung mit drei Erhebungsinstrumenten durchgeführt. Die Wissensbereiche chemisches Fachwissen und fachdidaktisches Wissen werden über adaptierte bestehende Wissenstests (Dollny, 2011; Tempel, 2017; Witner & Tepner, 2011) erhoben, anschließend werden die Probanden gebeten, aufgrund von in Vignetten dargestellten Situationen Handlungsalternativen auszuwählen (adaptierter Vignettentest nach Feige (2021)). Aufgrund der Genauigkeit der Auswahl werden Rückschlüsse auf das diagnostische Handeln in Unterrichtssituationen gezogen.

Ergebnisse / Ausblick

Das auf der Tagung vorzustellende Poster zeigt erste Befunde der Studie, die darauf hindeuten, dass fachdidaktisches Wissen einen stärkeren Einfluss auf die Akkuratessse bei der Auswahl von Handlungsalternativen beim Umgang mit Lernendenvorstellungen hat, als das Fachwissen. Aufgrund dieser ersten Befunde wird eine Implementierung der Ergebnisse in die Lehrerbildung angestrebt und in fachdidaktischen Seminaren zur Lehrerbildung umgesetzt. Fundamentale Teile der diagnostischen Kompetenz von Lehrkräften wie das Erkennen von Lernendenvorstellungen und die Auswahl von Handlungsalternativen scheinen bereits in der Lehrerbildung vermittelbar und einzuüben zu sein, was ähnliche Befunde von Wirth et al. (2023) stützen könnte. Im Rahmen einer anschließenden quasiexperimentellen Prä-Post-Interventionsstudie werden angehende Chemielehrkräfte ($n=170$) in einzelnen Bereichen ihres Professionswissens geschult und die Einflüsse der Schulung auf deren diagnostische Urteile beim Umgang mit Lernendenvorstellungen betrachtet. Hierzu werden drei Interventionsgruppen I1, I2 und I3 mit unterschiedlichen Interventionsinhalten gebildet (I1: Training CK, I2: Training PCK und I3: Training CK+PCK) und mit einer Wartekontrollgruppe verglichen. Abschließend wird der Fokus auf den diagnostischen Prozess gelegt. Daher werden nach der Bearbeitung der Vignetten stimulierte Erinnerungsinterviews ($n=20$) durchgeführt, in denen die Teilnehmenden ihre diagnostischen Urteile begründen sowie ihre Wissensanwendung erläutern sollen (Messmer, 2014). Die Auswertung der Interviews erfolgt mittels qualitativer Inhaltsanalyse (Mayring & Fenzl, 2014).

Pinnwand 84: Entwicklung und Evaluation eines Planspiels zur Förderung von professionellen Kompetenzen in Bezug auf Inklusion von Schüler*innen mit dem Förderschwerpunkt esE

Madeleine Müller, Susanne Jurkowski

Universität Erfurt, Deutschland

Theoretischer Hintergrund

Schüler*innen mit dem Förderschwerpunkt emotionale-soziale Entwicklung (FSP esE) werden von (angehenden) Lehrkräften als herausfordernd im Umgang wahrgenommen (Bosse & Spörer, 2014). Zudem stehen Lehrkräfte der Inklusion von Schüler*innen mit FSP esE tendenziell eher negativ gegenüber (de Boer et al., 2011; Ruberg & Porsch, 2017; Øen & Krumsvik, 2022). Durch Auffälligkeiten im emotionalen Erleben und der Verhaltensregulation bringt die Schüler*innengruppe eine Reihe an pädagogischen Herausforderungen im sozialen Kontext wie geringe Integration in die Peergruppe und Involviertheit in Mobbingdynamiken mit sich (z.B. Huber & Wilbert, 2012; Reicher & Mattischek-Jauk, 2018; Stein & Müller, 2015; Weiß et al., 2021).

Um diesen Herausforderungen als Lehrkraft angemessen begegnen zu können, gilt es den Erwerb professioneller Kompetenzen, beispielsweise durch Trainings- und Fortbildungsmaßnahmen, zu unterstützen (z.B. Forlin & Chambers, 2011; Kim, 2011; Kopp, 2009; Steinert & Jurkowski, 2023). Eine innovative Möglichkeit für den Erwerb professioneller Kompetenzen kann die Simulation in Form eines Planspiels sein (z.B. Ade-Ojo et al., 2022; Carl & Albers, 2023), welche als Aushandlungsprozesses in Kooperation mit anderen Teilnehmenden das Erarbeiten, Anwenden von Wissen sowie das Handeln erfordern (Capaul & Ulrich, 2003).

Planspiele zu bildungswissenschaftlichen Themen zeichnen sich unter anderem durch eine spannungsgeladene Situation des Handelns zwischen institutionellen und ministerialen Vorgaben einerseits und pädagogischen Anforderungen andererseits aus (Kadel et al., 2023). Neben diesem Aushandlungsprozess der verschiedenen Perspektiven, haben Planspiele ebenso das Potential, an den aktuellen Diskurs zur evidenzorientierten Praxis anzuknüpfen (u.a. Bauer & Kollar, 2023). Dabei können ganz im Sinne einer „mindful integration“ nach Rousseau und Gunia (2016) unterschiedliche Wissensbestände zum Thema Inklusion von Schüler*innen mit FSP esE aus wissenschaftlichem Wissen, Wissen über lokale Rahmenbedingungen und professionellem Erfahrungswissen zur Lösung des im Planspiel aufbereiteten Szenarios den Planspieltteilnehmenden angeboten, abgewogen, integriert und ausdiskutiert werden.

Zielsetzung und Fragestellung

Ziel des eingereichten Projektes ist es empirisch fundiert ein Planspiel zu entwickeln und dessen Einsatz zu validieren. Im ersten Schritt wird, basierend auf Expert*inneninterviews, ein authentisches und praxisnahes Szenario für den Einsatz in einem Planspiel entwickelt und dieses konzeptionell mit entsprechendem Material zum Einsatz aufbereitet. Im zweiten Schritt erfolgt die Validierung als Interventionsmöglichkeit in der Hochschullehre mittels des Einsatzes des Planspiels im Prä-Post-Design mit einer

systematischen Erfassung der Lerneffekte bezüglich professioneller Kompetenzen. Dabei sollen folgende Fragestellung beantwortet werden:

Inwiefern können durch das Planspiel

(1) professionelles Wissen, Selbstwirksamkeit und Einstellungen hinsichtlich des Unterrichtens mit Schüler*innen mit FSP esE gefördert werden?

(2) die Integration verschiedener Wissensbestände zur Lösung eines praxisnahen Problems gefördert werden?

Methode

Für die Entwicklung des Planspiels werden erfahrene Lehrkräfte im Rahmen von Expert*inneninterviews hinsichtlich besonders herausfordernder Situationen im Umgang mit Schüler*innen mit dem FSP esE befragt. In Anlehnung an charakteristische Merkmale von komplexen Problemlösesituationen (z.B. Unsicherheit, Komplexität; nach Dörner, 1987), werden die Interviews mittels qualitativer Inhaltsanalyse ausgewertet, um daraus ein authentisches Planspielszenario zu entwerfen. Darauf aufbauend soll die strukturelle Ausgestaltung des Planspiels in Anlehnung an Konstruktionsvorschläge von De Coninck et al. (2019) und Imhof & Starker (2020) erfolgen.

Für die Validierung des Planspiels sollen mittels eines Fragebogenverfahrens vor und nach dem eingesetzten Planspiel professionelles Wissen, Selbstwirksamkeit und Einstellungen hinsichtlich des Unterrichtens mit Schüler*innen mit FSP esE sowie die Nutzung und Integration verschiedener Wissensbestände erfasst werden.

Ergebnisse und Ausblick

Das Planspiel befindet sich aktuell in der Konzeptionsphase. Zur Tagung werden basierend auf der qualitativen Inhaltsanalyse der Expert*inneninterviews erste Ergebnisse der Entwicklung des Planspiels vorgestellt und ein ausführlicher Ausblick auf die bevorstehende Validierung der Intervention präsentiert. Basierend auf vorhergehenden evaluierten Planspielen (Meßner et al., 2021; Bastian et al., 2022) wird erwartet, dass das Planspiel dazu geeignet ist, mit seinem authentischen Setting professionelle Kompetenzen zu fördern.

Im Sinne eines Beitrages zum Wissenstransfer ist geplant, das Planspiel im Nachgang universitätsübergreifend zur Open Access Nutzung zu veröffentlichen.

ENTFÄLLT: Pinnwand 85: Soziale Ungleichheitswahrnehmungen von Eltern im schulischen Bildungskontext

Manuela Ulrich

Universität Konstanz, Deutschland

Angesichts der wachsenden Anforderungen in der heutigen Gesellschaft und der Botschaft zahlreicher Bildungs- und Erziehungsratgeber, dass die Familie für den Bildungserfolg der Kinder verantwortlich ist, fühlen sich viele Eltern/Erziehungsberechtigten unter Druck gesetzt, die normativen Rollen- sowie Leistungserwartungen zu erfüllen. Insbesondere Mütter haben das Gefühl, alles tun zu müssen, um die Zukunftschancen ihres Kindes zu optimieren; dazu zählt auch, die Bildungsbiographie zu planen und zu kontrollieren (Gartmeier, 2018; Jergus et al., 2018; Schneider et al., 2015). Die internationale Forschung belegt eine Intensivierung von Elternschaft entlang mehrerer Dimensionen wie Zeitinvestitionen, finanzielle Investitionen und Investitionen in die Bildung (z. B. Craig et al., 2014; Doepke & Zilibotti, 2019; Ishizuka, 2019; Vincent & Maxwell, 2016). Zeitliche und finanzielle Investitionen haben vermutlich im Zuge höherer gesellschaftlicher Bildungserwartungen zugenommen (Sayer et al., 2004). In Deutschland wurden kaum Unterschiede im Erleben von Druck nach sozioökonomischem Hintergrund festgestellt, nur im Hinblick auf finanziellen Druck (Walper & Kreyenfeld, 2022).

Der vorliegende Beitrag stellt eine qualitative Studie vor. Im Zentrum stehen mütterliche Rollen- und Ungleichheitswahrnehmungen nach dem Übergang ihres Kindes auf die weiterführende Schule untersucht, um Erkenntnisse über Druckempfinden und die Diskrepanzen zwischen normativen und individuellen Rollenerwartungen aus der Perspektive von Müttern zu gewinnen. Die zentralen Forschungsfragen bezogen sich darauf, 1) wie Mütter verinnerlichte normative Erwartungen an ihre Rolle als Bildungsbegleiterinnen erklären und 2) wie Mütter mit wahrgenommenen Rollenkonflikten umgehen.

Methodik: Um das Ziel zu erreichen, wurden fünf Fokusgruppeninterviews (je 3-6 Teilnehmerinnen) mit 25 Müttern in Baden-Württemberg durchgeführt und vollständig transkribiert. Die qualitativen Analysen erfolgten in einem ersten Schritt mithilfe der Grounded-Theory-Methodologie (Corbin & Strauss, 1990). Auf der Grundlage dieser fokusgruppenübergreifenden, offenen Datenanalyse wurde eine Fokusgruppe (die kleinste, drei Mütter mit sehr unterschiedlichen Lebenssituationen) unter Einbindung der theoretischen Konzepte des symbolischen Interaktionismus (Mead, 1934/1968) näher untersucht.

Ergebnisse und Diskussion: Die Befunde zeigen, dass Mütter normative Rollenerwartungen im allgemeinen Erziehungskontext sowie im Hinblick auf die Unterstützung der Schulbildung in gleicher Weise internalisiert haben und ihre Rolle im Kontext sozialer Ungleichheiten (elterliche Bildung, Beruf, soziales Umfeld) reflektieren. Die Ergebnisse zeigen auch, dass das mütterliche Selbstvertrauen ihr Wohl- oder Druckempfinden als Bildungsbegleiterinnen beeinflusst. Die Mutter-Lehrpersonen-Kommunikation kann als soziale Ressource gesehen werden, wird aber auch als Mangel an mütterlicher Kompetenz und als Abweichung von der Norm interpretiert. Erlebte Rollenkonflikte können zu Belastungen der Mütter und Nachteilen der Kinder hinsichtlich einer effektiven Unterstützung führen; Vertrauen in die Lehrpersonen und eine offene Kommunikation mit der Schule könnten Nachteile mindern. Im Hinblick auf die Methodik dieser Studie birgt der Ansatz des symbolischen Interaktionismus Vorteile für das Verständnis von Rollenübernahmeprozessen im Kontext sozialer Ungleichheit.

Pinnwand 86: Arbeitsgedächtniskapazität beim Fremdsprachenlernen mit einem intelligenten Tutorsystem

Katharina Wendebourg¹, Diana Pili-Moss³, Torben Schmidt³, Stephen Bodnar¹, Heiko Holz⁴, Benjamin Nagengast¹, Ulrich Trautwein¹, Detmar Meurers²

¹Universität Tübingen; ²Leibniz-Institut für Wissensmedien, Tübingen; ³Universität Lüneburg; ⁴Pädagogische Hochschule Ludwigsburg

Intelligente Tutorsysteme (ITS) bieten großes Potenzial, den individuellen Lernprozess zu optimieren. Sie können durch individuelles, adaptives Feedback gezielt auf die Stärken und Schwächen der Lernenden eingehen und so maßgeschneidert unterstützen. Gerade beim Spracherwerb sind individuelle Unterschiede der Lernenden von direkter Bedeutung (Ellis, 2004; Skehan, 1991), was automatisiertes individuelles Feedback für eine adaptive Förderung besonders relevant macht. Eines der

wenigen wissenschaftlich evaluierten Tutorsysteme ist das [ITS], das für Englischlernende der siebten Klasse an Gymnasien entwickelt wurde. In einem ersten RCT wurde gezeigt, dass Schüler:innen, die individualisiertes, adaptives Feedback zu grammatikalischen Konstrukten erhalten hatten, bessere Lernleistungen aufwiesen, als solche, die lediglich Feedback über die Korrektheit ihrer Antwort erhalten hatten ([Autoren], 2019). Daran anknüpfend wurde das [ITS] durch zusätzliche Elemente erweitert, die das Lernen noch motivierender gestalten sollen: einen persönlichen Agenten, der die Leistung am Ende einer Aufgabe spielerisch zusammenfasst sowie ein Dashboard in dem die Lernleistungen über die verschiedenen Aufgaben hinweg aggregiert visuell präsentiert werden. Der Agent und das Dashboard zeigten sich unter bestimmten Umständen als motivationsförderlich ([Autoren], in Revision). Ziel der hier vorgestellten Studie ist die Frage, ob die Schüler:innen gleichermaßen von der Arbeit mit dem [ITS] profitierten oder ob individuelle Unterschiede den Zusammenhang und die Interaktion mit dem System beeinflussten. Der Fokus liegt dabei auf der Arbeitsgedächtniskapazität der Lernenden als Maß kognitiver Leistungsfähigkeit (Oberauer et al., 2000), genauer der Fähigkeit zum Updating, d.h. Gedächtnisinhalte zu aktualisieren und neue Inhalte kurzfristig ins Gedächtnis zu integrieren (Ecker et al., 2010). Diese Fähigkeit spielt besonders bei kognitiv komplexen Aufgaben wie dem Grammatiklernen eine wichtige Rolle (Chen et al., 2022). Die Fähigkeit zum Updating korreliert mit fluiden Intelligenz (Pelegri et al., 2015).

Im Schuljahr 2021/2022 wurde eine groß angelegte Schulstudie an Gymnasien in Baden-Württemberg, Nordrhein-Westfalen und Hamburg durchgeführt. Im Rahmen eines multi-methods-Ansatzes fanden dabei qualitative Interviews und videogestützte Erhebungen mit zwei Klassen sowie eine randomisierte kontrollierte Feldstudie mit mehr als 600 Schüler:innen statt. Neben der Lernleistung, Motivation und Persönlichkeitsvariablen wurde dabei auch die Arbeitsgedächtniskapazität der Schüler:innen mittels eines n-back-task ermittelt. Als abhängige Variablen wurden die mittlere Reaktionszeit sowie die Anzahl der richtigen Antworten gewählt. Da die Erhebung nur in den Interventionsbedingungen durchgeführt wurde, und aufgrund von Studienabbrüchen, stehen diese Daten von $N = 123$ Schüler:innen zur Verfügung. Vorläufige Ergebnisse legen nahe, dass Lernende mit höherer Arbeitsgedächtniskapazität einen größeren Lernzuwachs aufwiesen, als Lernende mit niedrigerer Arbeitsgedächtniskapazität. Das war jedoch nur in derjenigen Interventionsbedingung der Fall, in der die Lernenden ausschließlich sprachliches Feedback erhalten hatten, jedoch keinen pädagogischen Agenten und kein Dashboard. Aktuell untersuchen wir, ob und wie Lernende in Abhängigkeit ihrer Arbeitsgedächtniskapazität unterschiedlich mit dem [ITS] interagierten und wie sich dies auf die Lernleistung auswirkte. Dafür werten wir behaviorale Log-Daten der Lernenden aus. Diese Ergebnisse sollen ebenfalls im Rahmen der GEBF-Tagung im Januar vorgestellt werden.

Die Ergebnisse dieser Studie ermöglichen einen tieferen Einblick in die individuellen Lernprozesse von Schüler:innen in einem intelligenten adaptiven Tutorsystemen. Durch die Analyse unterschiedlicher kognitiver Profile können wir gezielte Maßnahmen zur Optimierung des Fremdspracherwerbs entwickeln und somit einen wichtigen Beitrag zur Chancengleichheit im Bildungswesen leisten.

Pinnwand 87: Sensibilisierung für den Umgang mit sexueller und geschlechtlicher Vielfalt im schulischen Kontext durch die theoretische und praktische Auseinandersetzung mit Inklusion im Lehramtsstudium

Alissa Müller, Michael Becker

Institut für Schulentwicklungsforschung / TU Dortmund, Deutschland

Lesbische, schwule, bisexuelle, trans- und intergeschlechtliche (LSBTI) Schüler*innen erleben Viktimisierung, Marginalisierung und Diskriminierung in ihrem Schulalltag (Birkett et al., 2014; Espelage, 2015; Gegenfurtner & Gebhardt, 2017). Empirische Befunde zeigen, dass Lehrkräfte eine wesentliche Rolle in der Intervention und Prävention von queerfeindlichem Verhalten und der Förderung von Akzeptanz sexueller und geschlechtlicher Vielfalt unter Schüler*innen haben (Gegenfurtner & Gebhardt, 2017; Klocke, 2012; Klocke et al., 2019). Dennoch weisen Befragungen von Schüler*innen und Lehrkräften darauf hin, dass die Prävention und Intervention von queerfeindlicher Diskriminierung und die Berücksichtigung von queeren Lebensrealitäten im Unterricht oftmals nicht konsequent und lediglich eigenverantwortlich stattfindet (Antidiskriminierungsstelle des Bundes, 2017; Klocke, 2012; Krell & Oldemeier, 2017).

Um Lehrkräfte für einen diskriminierungsfreien und wertschätzenden Umgang mit geschlechtlicher und sexueller Vielfalt in der Schule zu sensibilisieren, werden unter anderem die Vermittlung von Wissen und Interventionsstrategien im Rahmen von Fortbildungsmaßnahmen sowie die Förderung von positiven Einstellungen gegenüber queeren Schüler*innen diskutiert (Klocke, 2012; Klocke et al., 2019; McCabe et al., 2013). Eine solche Sensibilisierung kann etwa durch die Auseinandersetzung mit queeren Perspektiven bereits im Lehramtsstudium gefördert werden (Galano et al., 2024; Gegenfurtner, 2021; Gegenfurtner et al., 2023). Ein Blick auf lehrramtsbildende Hochschulen offenbart jedoch, dass sexuelle und geschlechtliche Vielfalt nur selten explizit behandelt wird und Dozierende der Thematisierung häufig keine hohe Bedeutung zumessen (Nietzschmann et al., 2019; Vierniesel, 2018). Obwohl der Umgang mit Heterogenität und Inklusion im Unterricht verstärkt im Lehramtsstudium behandelt wird, ist dieses nicht auf die umfassende Auseinandersetzung mit allen Differenzkategorien ausgelegt, die im schulischen Kontext unter einem weit gefassten Inklusionsverständnis relevant werden könnten. Ein pragmatischer Ansatz besteht in der Formulierung weit gefasster Lernziele, die durch eine Sensibilisierung für Vielfalt sowie die Förderung einer forschenden Haltung und der (Selbst-)Reflexion den reflektierten und anpassungsfähigen Umgang mit heterogenen Schulklassen langfristig ermöglichen sollen (Jennek et al., 2022; Knigge et al., 2020).

Bislang ist es unklar, inwiefern eine solche allgemeine Ausrichtung angehende Lehrkräfte für sexuelle und geschlechtliche Vielfalt sensibilisiert, ohne diese explizit zu thematisieren. Das folgende Forschungsvorhaben soll daher untersuchen, inwiefern der Besuch von inklusionsbezogenen Lehrveranstaltungen und Praxisphasen im Lehramtsstudium eine solche Sensibilisierung ermöglicht. Dazu werden vorhandene Daten von Teilnehmenden einer universitären Lehrveranstaltung zum Thema Inklusion aus fünf konsekutiven Semesterkohorten analysiert. Die Studierenden wurden jeweils zu vier Messzeitpunkten zu verschiedenen Inklusionsaspekten befragt: Zunächst zu Semesterbeginn, dann semesterbegleitend nach der theoretischen Vermittlung und nach einem daran gekoppelten Praktikum in außerunterrichtlichen Einsatzfeldern sowie nach einer anschließenden Reflektionseinheit. Um die Sensibilisierungsaspekte darzustellen, wird auf eine Skala aus dieser Befragung zurückgegriffen, die die Einstellungen und selbsteingeschätzte Kompetenz im Umgang mit LSBTI-Schüler*innen, die wahrgenommene Rolle von Lehrkräften im Kontext von sexueller und geschlechtlicher Vielfalt sowie ihre wahrgenommene Relevanz im Kontext schulischer Inklusion erfasst hat. Im Fokus des weiteren analytischen Vorgehens steht zum einen die Untersuchung der Stabilität des Konstruktes über die Messzeitpunkte hinweg und zum anderen die Untersuchung von Mittelwertsveränderungen mit Hilfe von latenten Wachstumskurvenmodellen.

Ableitend aus den Lernzielen der Lehrveranstaltungen ist zu erwarten, dass die theoretische und praktische Auseinandersetzung mit Inklusion im weiteren Sinne zu einer wachsenden Sensibilisierung für sexuelle und geschlechtliche Vielfalt über den zeitlichen Verlauf führt. Es ist allerdings auch denkbar, dass mit der verstärkten Auseinandersetzung mit Inklusion auch eine steigende

Messinvarianz für das zugrundeliegende Konstrukt sichtbar wird, da dieses womöglich besser verstanden wird. Das vorliegende Forschungsvorhaben zielt auf wertvolle Erkenntnisse im Kontext der Lehrkräfteausbildung ab, da nach bisherigem Kenntnisstand keine vergleichbare Untersuchung vorliegt, die eine Sensibilisierung für geschlechtliche und sexuelle Vielfalt im Zusammenhang einer weit gefassten Auseinandersetzung mit Inklusion im Lehramtsstudium ergründet.

Pinnwand 88: Personalisierungseffekt und Social Agency – Führt das Gefühl direkt angesprochen zu werden zu einer höheren Verständnisleistung?

Maïke Jana Lindhaus, Alina Nöbler, Stephan Dutke

Universität Münster, Deutschland

Die Personalisierung von schriftlichen Lernmaterialien durch den Austausch von Personal- und Possessivpronomen, kann laut Mayer et al. (2014) Lernende zu einer tieferen Verarbeitung motivieren und in einer höheren Lernleistung resultieren. Theoretisch wird dies mit der *Social Agency Theorie* begründet (Mayer et al., 2014). Demnach fühlen sich Lernende durch den Austausch von Possessiv- und Personalpronomen direkt angesprochen. Das baut eine quasi-soziale Beziehung mit dem Autor des Lernmaterials auf. Durch diese gelten die Konventionen der Mensch-zu-Mensch-Kommunikation, die die tiefere Verarbeitung bei den Lernenden anregen. Der Personalisierungseffekt wurde in der Vergangenheit mehrfach repliziert (vgl. Meta-Analyse von Ginns et al., 2013). Allerdings bleiben empirische Befunde zum Einfluss der Social Agency (d.h., eine quasi-soziale Beziehung zum Autor durch die Bearbeitung des Textes aufzubauen) auf den Personalisierungseffekt bislang aus.

Daher geht die vorliegende Studie der Fragestellung nach, welchen direkten Einfluss die sprachliche Personalisierung auf die Social Agency und auf die Lernleistung hat. Dazu wurden insgesamt 17 nicht veröffentlichte Primärstudien der Autoren ($N = 1.146$ Vpn) metaanalytisch untersucht. Alle 17 Primärstudien untersuchten in einem Between-Subject-Design den direkten Einfluss der schriftlichen Personalisierung auf die Lernleistung (erfasst als Verständnisleistung und Transferleistung). Die Social Agency wurde mithilfe der Social Agency Skala (Lindhaus et al., under review) anhand von vier Facetten (Conversational Character (CC), Sympathy and Emotional Connection (SE), Explanatory Effect (EE), Task Involvement (TI)) erfasst. Die untersuchten Primärstudien wurden innerhalb von Schulen in Nordrhein-Westfalen anhand von Paper-Pen-Befragungen durchgeführt, alternativ wurden heterogener Stichproben durch anonyme Online-Befragungen erhoben.

Die Ergebnisse zeigen einen signifikanten Effekt $d = 0.19$, $p < .001$ der Personalisierung schriftlichen Lernmaterials auf die Verständnisleistung. Zusätzlich kann ein signifikanter Effekt $d = 0.27$, $p = .01$ der Personalisierung auf die Social Agency Facette Conversational Character verzeichnet werden. Betrachtet man den Forrestplot zeigt sich der Personalisierungseffekt auf die Verständnisleistung auf der Ebene der Einzelstudien besonders bei jüngeren Schülerstichproben und bei heterogenen Stichproben die via Online-Befragung erhoben wurden. Ein Grund dafür könnte sein, dass SchülerInnen aufgrund des Schulkontextes in denen die Materialien präsentiert wurden, davon ausgingen, dass der Autor auch gleichzeitig die Lehrperson ist. Während eine Ansprache per „du“ in unteren Klassenstufen gewöhnlich ist, wird in oberen Klassenstufen eine formelle Ansprache der Lehrpersonen gegenüber der SchülerInnen gewählt. Dieses inkongruente Verhalten bei der Personalisierung, könnte bei den SchülerInnen höherer Klassenstufen dazu führen, dass die Personalisierung keinen positiven Effekt hat. Der Effekt zeigt sich bei den heterogenen Stichproben unabhängig vom Alter, da diese Studien in Form einer anonymen Online-Befragung aufgezeichnet wurden und somit keine Assoziation mit einer direkten Person als Autor stattfindet, von der das Verhalten als inkongruent wahrgenommen werden kann.

Die Ergebnisse werden auf der Tagung präsentiert, diskutiert und Aufschluss darüber geben, ob sich SchülerInnen und Schüler durch die Personalisierung von Lernmaterialien direkter angesprochen fühlen und dies zu einer höheren Verständnisleistung führt.

Pinnwand 89: Wie hängen generische und fachspezifische Unterrichtsqualitätsmerkmale zusammen? Befunde aus dem Fach Wirtschaft

Stefanie Findeisen¹, Patricia Köpfer², Liane Platz¹

¹Universität Konstanz, Deutschland; ²Universität Hohenheim, Deutschland

Theoretischer Hintergrund/Fragestellung

In der Unterrichtsqualitätsforschung gelten generische, fachübergreifende Merkmale der Unterrichtsqualität (*Basisdimensionen*: kognitive Aktivierung, konstruktive Unterstützung und Klassenführung; z.B. Klieme & Rakoczy, 2003) inzwischen als recht gut untersucht. Aktuell wird vermehrt diskutiert, ob diese fächerübergreifend anwendbar sind bzw. inwiefern fachspezifische Unterrichtsqualitätsaspekte ergänzt werden sollten. Da Unterricht fachspezifisch erfolgt, ist die Forderung nach einer fachspezifischen Untersuchung von Unterrichtsprozessen keineswegs neu (z.B. Stodolsky, 1988) und durchaus naheliegend. Dennoch liegen nur wenige Erkenntnisse darüber vor (z.B. Jentsch et al., 2020 für Mathematik). Der Beitrag widmet sich erstmals der Unterrichtsqualität im Wirtschaftsunterricht aus generischer und fachspezifischer Perspektive. Die drei Basisdimensionen werden gezielt mit drei fachspezifischen Aspekten (Einsatz handlungsorientierter Methoden, Aktualitäts- und Lebensweltbezug, Urteilsbildung; vgl. [anonymisiert]) kombiniert. Wir untersuchen, wie *Qualitätsbeurteilungen für generische und fachspezifische Unterrichtsmerkmale zusammenhängen*. Hierbei werden Zusammenhangsanalysen sowohl für Schüler*innen- als auch für Beobachtendenbeurteilungen der insgesamt sechs Qualitätsaspekte (s.u.) durchgeführt. Perspektivisch ist dabei von Interesse, welche Qualitätsaspekte den größten Erklärungswert für die Lernergebnisse der Schüler*innen haben.

Methode

Zur Beantwortung der Fragestellungen führen wir eine Videostudie durch. Aktuell umfasst der Datensatz Informationen von fünf Klassen ($N = 107$ Schüler*innen) an kaufmännischen Vollzeitschulen, die über acht zusammenhängende Unterrichtsstunden zum gleichen Thema (Volkswirtschaftslehre: Angebot/Nachfrage) hinweg gefilmt wurden. In einer Post-Befragung wurde das volkswirtschaftliche Wissen der Schüler*innen getestet (Test of Economic Literacy [TEL4]; Walstad et al., 2013). Darüber hinaus wurde deren Qualitätseinschätzung zu den drei Basisdimensionen der Unterrichtsqualität (IBBW, 2021) sowie zu den drei fachspezifischen Qualitätsmerkmalen ([anonymisiert]) erfasst (jeweils Skala 1: trifft nicht zu bis 4: trifft voll und ganz zu). Auf Basis der Videoaufnahmen (unterteilt in 20-Minuten-Abschnitte) beurteilen zwei unabhängige Beobachtende die drei generischen ($.87 \leq ICC \leq .88$) und die drei fachspezifischen Qualitätsmerkmale (Coderschritt wird im Herbst 2024 abgeschlossen sein; vorläufig: $.83 \leq ICC \leq .86$). Wir berichten Korrelationen zwischen den sechs Qualitätsmaßen aus zwei Perspektiven sowie Korrelationen mit dem volkswirtschaftlichen Wissen nach der Unterrichtseinheit.

Ergebnisse/Diskussion

Aus der Korrelationsanalyse ist zunächst ersichtlich, dass für allgemeine und fachspezifische Merkmale lediglich geringe und teilweise negative Zusammenhänge zwischen den Beurteilungen der Beobachtenden sowie der Schüler*innen bestehen (z.B. kognitive Aktivierung: $r = -.312^{**}$; Handlungsorientierung: $r = .332^{**}$). Dies passt zu vorliegenden empirischen Befunden (z.B. Clausen, 2002). Darüber hinaus korrelieren generische und fachspezifische Qualitätsaspekte nicht signifikant bis maximal moderat positiv miteinander (mit Ausnahme der konstruktiven Unterstützung und der Handlungsorientierung: $r = .737^{**}$). Vereinzelt liegen auch signifikant negative Korrelationen vor (z.B. konstruktive Unterstützung und Urteilsbildung: $r = -.775^{**}$). Um diese besser erklären zu können, sind detailliertere Analysen einzelner Unterrichtsabschnitte (Codierung erfolgte jeweils für 20-Minuten-Abschnitte) notwendig. Insgesamt sprechen diese vorläufigen Befunde aber für eine ergänzende Berücksichtigung fachspezifischer Merkmale.

Hinsichtlich des volkswirtschaftlichen Wissens zeigen die beobachtete Klassenführung ($r = .305^{**}$) und Urteilsbildung ($r = .358^{**}$) sowie die von Schüler*innen beurteilte kognitive Aktivierung ($r = .287^{**}$) die größten positiven Zusammenhänge. Hingegen korrelieren konstruktive Unterstützung ($r = -.406^{**}$) und Handlungsorientierung ($r = -.267^{**}$) negativ mit der Testleistung der Schüler*innen. Der negative Zusammenhang zwischen einem stark handlungsorientierten Unterricht und der Wissenstestleistung lässt sich vermutlich über das Testinstrument begründen, das Faktenwissen abfragt und nicht auf Handlungskompetenz der Schüler*innen ausgerichtet ist. Folglich ist die Passung von Messinstrumenten zur Erfassung der Schüler*innenleistung ebenfalls in die Diskussion über eine fachspezifisch adäquate Erfassung von Unterrichtsqualität und ihrer Wirkung einzubeziehen.

Pinnwand 90: Selektive Aufmerksamkeitsmechanismen in der Unterrichtswahrnehmung bei Lehramtsstudierenden

Aldin Alijagic¹, Ricardo Böheim², Martin Daumiller¹, Markus Dresel¹, Andreas Gegenfurtner¹, Doris Lewalter², Tina Seidel²

¹Universität Augsburg, Deutschland; ²Technische Universität München, Deutschland

Hintergrund

Lehrkräfte müssen im Unterricht versuchen, aus einer Vielzahl verschiedener Informationen sinnvolle Unterrichtshandlungen abzuleiten, da sie ihre visuelle Aufmerksamkeit darauf lenken (van Es & Sherin, 2021). Diese Lenkung, die ausschließlich auf Wissen basiert, wird auch als *Noticing* bezeichnet und ist neben dem *Selecting* und *Ignoring* ein bedeutender Bestandteil der professionellen Unterrichtswahrnehmung (Gegenfurtner et al., 2022). Bisher ist allerdings noch nicht klar, ob affektive Komponenten neben Wissen auch eine Rolle spielen. Gemäß dem PID-Modell von Blömeke wirken sich auch attitudinale und stressbedingte Determinanten auf die selektive Wahrnehmung aus (Blömeke et al., 2015). Untersuchungen, die den kausalen Zusammenhang von Einstellung (Keskin et al., 2023) und Stress (Chaudhuri et al., 2022; Muhonen et al., 2022) auf die selektive Aufmerksamkeit untersucht haben, sind rar und dementsprechend würde diese Studie neue Erkenntnisse in diesem Feld liefern. Es wird auch angenommen, dass gemäß der Selbstkategorisierungstheorie Lernende häufiger betrachtet von Lehramtsstudierende werden, die das gleiche Geschlecht wie sie haben (Turner et al., 1989). Die Erforschung von Scanpath-Mustern ist ebenfalls sehr explorativ. Scanpaths werden durch zugrunde liegende kognitive Schemata gesteuert (Noton & Stark, 1971a, 1971b). Wenn man davon ausgeht, dass Stereotypen mentale Prozesse des Verstehens und Interpretierens von Informationen über eine Person beinhalten (Eagly & Chaiken, 1998), kann man davon ausgehen, dass sich Unterschiede in den Scanpaths bei Personen mit positiver und negativer Einstellung identifizieren lassen. Aus der Expertiseforschung geht zum Beispiel hervor, dass es Unterschiede zwischen Lehramtsstudierenden und erfahrenen Lehrkräften in den Scanpath-Mustern gibt (Kosel et al., 2021; McIntyre et al., 2017).

Fragestellung

Dementsprechend stellt sich die Frage, wie sich das Geschlecht der Lernenden und der Unterrichtsstress auf die visuelle Aufmerksamkeit der Lehramtsstudierenden auswirken.

F1.1: Inwieweit hängt der von den Lehramtsstudierenden impliziten und expliziten Einstellungen mit ihrer selektiven Aufmerksamkeit zusammen?

F1.2: Gibt es Unterschiede zwischen weiblichen und männlichen Lehramtsstudierende in der selektiven Wahrnehmung?

F1.3: Inwieweit unterscheiden sich einstellungsbezogene Scanpaths, operationalisiert als implizit positiv und negativ, bei Lehramtsstudierenden voneinander?

F2: Inwieweit hängt der von den Lehramtsstudierenden subjektiven Stress mit ihrer selektiven Aufmerksamkeit zusammen?

Methode

Die Studie strebte einen Multi-Method-Ansatz an. 51 Lehramtsstudierenden (73% weiblich; $M_{Alter} = 22.0$, $SD_{Alter} = 2.0$) haben sich ein Unterrichtsvideo angesehen und währenddessen wurden ihre Augenbewegungen mit einem monitorbasiertem Tobii Pro Fusion Eye-Tracker aufgenommen. Bezogen auf Forschungsfrage 1 wurde die selektive Aufmerksamkeit der Lehramtsstudierenden als prozentualer Anteil der Fixationsanzahl und -dauer auf das Geschlecht der Schüler*innen operationalisiert (McIntyre et al., 2017). Implizite Einstellungen bezüglich des Geschlechts wurden mit einem impliziten Assoziationstest erfasst (Greenwald et al., 1998). Angepasst an dieser Studie wurden explizite Einstellungen für jede einzelne Geschlechtsdimension mit fünf siebenstufigen Likert-Items gemessen (Hachfeld et al., 2012). Zur Beantwortung der zweiten Forschungsfrage wurde hier die selektive Aufmerksamkeit als Maß der Ungleichverteilung operationalisiert (Gini-Koeffizient; Cortina et al., 2015). Subjektiver Stress wurde jeweils mit Skalen zur kognitiven Belastung (Szulewski et al., 2017), wahrgenommen Stress (Schäfer et al., 2023) und Well-Being (Lamers et al., 2011) gemessen.

Ergebnisse und ihre Bedeutung

Es wird berichtet, dass implizite und explizite Einstellungen nicht mit der selektiven Wahrnehmung zusammenhängen. Ergebnisse zeigen, dass weibliche Lehramtsstudierende prozentual eher auf weibliche Schüler*innen blicken, $t(49) = 3.42$, $p < .001$. Die Befunde sind analog zum männlichen Geschlecht. Die Analyse der Scanpath-Muster erfolgt noch und würde in der kommenden Tagung berichtet werden. Die Annahme, dass der Gini-Koeffizient höher ausfällt, wenn mehr Stress empfunden wird, konnte in den Analysen nicht bestätigt werden. Ein Grund könnte in der Art der Stresserfassung liegen und es wird postuliert, dass zukünftige Studien neben Fragebogenskalen zusätzlich physiologische Indikatoren wie die Herzrate erheben (Junker et al., 2021). Zukünftig könnte sich die Frage stellen, ob visuelle Gaze Displays als Reflektionsinstrument genutzt werden können, um Einstellungen und Stress positiv zu verändern.

ENTFÄLLT Pinnwand 91: Lernausgangslagen von Schüler:innen der gymnasialen Oberstufe in Baden-Württemberg – Erste Ergebnisse der TOSCAneo-Studie

Evelyn Schnauer¹, Wolfgang Wagner¹, Benjamin Nagengast¹, Walther Paravicini¹, Guido Pinkernell², Nicolas Hübner¹, Lisa Bäumke¹, Ulrich Trautwein¹

¹Universität Tübingen, Deutschland; ²Pädagogische Hochschule Heidelberg, Deutschland

Theoretischer Hintergrund

Mit der Öffnung des Bildungssystems kann die Qualifikation zur tertiären Bildung in Baden-Württemberg am allgemeinbildenden Gymnasium (AG), beruflichen Gymnasium (BG) und an Gemeinschaftsschulen mit gymnasialer Oberstufe (GS) erworben werden. In früheren Kohorten (Köller et al., 2004; Trautwein et al., 2010) der TOSCA-Studie zeigten sich Leistungsunterschiede in Hinblick auf Englisch und Mathematik zwischen AG und BG zum Abitur, wofür unterschiedliche Lernstände vor Oberstufeneintritt verantwortlich sein könnten. Trautwein et al. (2011) untersuchten anhand der TOSCA-10-Studie den geplanten Übertritt von Schüler:innen am AG und an der Realschule mit Berechtigung zur Oberstufe am Ende der Sekundarstufe I unter Berücksichtigung des sozialen Hintergrundes. Sie konnten zeigen, dass sich die beiden Gruppen insbesondere hinsichtlich des Bildungshintergrundes unterscheiden. Es ist daher plausibel, dass sich die Schüler:innenpopulationen an den drei Schulformen systematisch hinsichtlich ihrer Lernausgangslagen zu Beginn der Oberstufe und ihrer sozialen Zusammensetzung unterscheiden.

Fragestellung

Ziel der vorliegenden Untersuchung war ein Vergleich der Lernausgangslagen in den Domänen Deutsch, Englisch, Mathematik und Naturwissenschaften sowie des sozialen Hintergrundes der Eingangskohorte der gymnasialen Oberstufe der verschiedenen Schularten aus dem Schuljahr 2023/24. Im Falle des BG wurde weiter zwischen verschiedenen Profilen (lebens- und humanwissenschaftliches, technisches und wirtschaftswissenschaftliches Profil) differenziert.

Methode

Die TOSCAneo-Studie ist eine vom Kultusministerium Baden-Württemberg beauftragte repräsentative Längsschnittstudie. Seitens der Schüler:innen wurden, angelehnt an Trautwein et al. (2011), Angaben zum Geschlecht, Bildungshintergrund, Migrationshintergrund, Highest International Socio-Economic Index of Occupational Status (HISEI; Ganzeboom et al., 1992) und zur Art des Versetzungszeugnisses analysiert. Die Analysen wurden mit R Version 4.4.1 (R Core Team, 2024) und SAS/STAT Version 15.3 (SAS Institute Inc., 2023) durchgeführt. Für die Beantwortung der Fragestellungen wurden cluster-robuste Standardfehler mit den Schulen als Cluster ($N = 136$) geschätzt, wobei zusätzlich Populationsgewichte auf Schüler:innenebene (aktive Teilnahme: $N = 4064$) berücksichtigt wurden. Als abhängige Variablen (Fachleistungen) dienten Weighted Likelihood Estimates (WLE; Warm, 1989) auf Basis von jeweils eindimensionalen zweiparametrischen Item-Response-Theorie-Modellen (IRT; de Ayala, 2009). Unterschiede zwischen den Schultypen wurden als Effektstärken in Form von Glass' Δ standardisiert (Referenzgruppe AG). Für die Häufigkeitsanalysen von kategorialen Variablen des sozialen Hintergrundes wurden logistische Regressionen berechnet.

Ergebnisse

Hinsichtlich der Deutschkompetenz erzielten Schüler:innen an der GS sowie am BG jeweils signifikant niedrigere Ergebnisse als am AG (GS: $p < .001$; BG: $p < .001$). In Englisch ergaben sich am BG verglichen mit dem AG für die Schüler:innen geringere Kompetenzen ($p = .003$), wobei dies ausschließlich das lebens- und humanwissenschaftliche ($p = .011$) sowie das wirtschaftswissenschaftliche Profil ($p < .001$) betraf. Hinsichtlich der mathematischen und der naturwissenschaftlichen Kompetenz ließen sich Leistungsunterschiede zwischen Schüler:innen an allen Schularten zugunsten des AG nachweisen (Range von $p < .001$ bis $p = .002$). Einzige Ausnahme war der Vergleich der Schüler:innen am technischen BG und am AG, bei dem sich für beide Kompetenzen keine statistisch signifikanten Unterschiede ergaben (Mathematik: $p = .697$; Naturwissenschaften: $p = .065$).

Bezüglich des Bildungshintergrundes (Anteil der Erziehungsberechtigten, die mindestens die Fachhochschulreife erworben haben) unterschieden sich lediglich AG und BG voneinander (AG: 76%; GS: 69%; BG: 61%). Der HISEI unterschied sich signifikant zwischen AG und GS ($p = .001$) sowie zwischen AG und BG ($p < .001$). Der Anteil von Schüler:innen mit Migrationshintergrund war auf dem BG größer (AG: 32%; GS: 44%; BG: 39%). Am AG und der GS erlangte ein Großteil (97% bzw. 77%) der Schüler:innen ihr Versetzungszeugnis an derselben Schulart, während 58% der Schüler:innen am BG zuvor die Realschule besuchten.

Die Ergebnisse weisen auf einen systematischen Unterschied in den Lernausgangslagen und der sozialen Komposition zugunsten des AG hin.

Pinnwand 92: Wenn Fragen Fragen aufwerfen: Wie Primarschüler:innen Items zu ihrer Nutzung von unterrichtlichen Lernangelegenheiten interpretieren

Leonie Gossner¹, Sebastian Röhl², Wida Wemmer-Rogh¹, Urs Grob¹, Gerlinde Lenske³, Anna-Katharina Praetorius¹

¹Universität Zürich, Schweiz; ²Pädagogische Hochschule Freiburg; ³Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau

Theoretischer Hintergrund

Seit vielen Jahren gelten Schüler:inneneinschätzungen zum Unterricht als ein "ökonomisches Standarderhebungsverfahren zur Messung von Unterrichtsqualität" (Lenske & Praetorius, 2020, S. 12). Im Zusammenspiel von Angebot und Nutzung im Unterricht mediert das Urteil der Lernenden zwischen dem Unterrichtshandeln der Lehrperson und dem Lernertrag (Lenske, 2016), daher kommt den Schüler:innen eine Schlüsselposition im Unterricht zu. Während Schüler:innenratings zahlreiche Vorteile mit sich bringen (z.B. vergleichsweise hohe Reliabilität durch Aggregation auf Klassenebene; Lenske, 2016), haben sich auch Herausforderungen gezeigt (z.B. fehlende pädagogisch-didaktische Expertise; Clausen, 2002). Will man die Schüler:innenperspektive auf Unterricht erfassen, stellt sich entsprechend die Frage, welche Aspekte von Unterrichtsqualität Schüler:innen überhaupt valide beurteilen können. Wird Unterricht als von Lehrpersonen absichtsvoll geplante und initiierte Lernumgebung verstanden (Terhart, 1994), setzt die Einschätzung der Angebotsseite unumgänglich pädagogisch-didaktisches Wissen voraus, worüber Schüler:innen meist nicht verfügen. Eine Alternative bildet eine Fokussierung auf die Nutzungsprozesse, also kognitive, emotionale und motivationale Prozesse der Lernenden (Vieluf et al., 2020; Vieluf, 2022). Mit Blick auf bestehende Schüler:innenfragebogen zur Unterrichtsqualität wird ersichtlich, dass die meisten Instrumente nur das Unterrichtsangebot erheben (z.B. Bertram, 2019; Ferguson, 2012) oder Items für das Angebot und die Nutzung ohne explizite Systematik kombinieren (z.B. Gärtner et al., 2022; Rakoczy et al., 2022), was eine detaillierte Analyse der Nutzungsprozesse erschwert. Nur

einzelne Befragungsinstrumente legen den Fokus ausschließlich auf individuelle Nutzungsprozesse (z.B. Chandra Handa, 2020; Jansen et al., 2022), jedoch werden mehrere Stunden oder der Unterricht allgemein beurteilt, was für Lernende einen unklaren Bezugspunkt darstellen kann. Daher ist die Entwicklung eines Nutzungsfragebogens, der sich auf einen unmittelbaren Bezugspunkt konzentriert, von großem Potenzial.

Allerdings ist bislang ungeklärt, wie Schüler:innen Items zu ihrem Nutzungsverhalten interpretieren. Bisherige Studien, die die Inhaltsvalidierung von auf die Angebotsseite fokussierte Schüler:innenfragebogenitems untersucht haben (Lenske, 2016; Lenske & Praetorius, 2020; Guttke & Porsch, 2024), zeigen, dass viele eingesetzte Items nicht wie intendiert verstanden werden. Daher fokussiert diese Studie folgende Fragestellungen:

F1: Verstehen Schüler:innen Items zur eigenen Unterrichtsnutzung im intendierten Sinn?

F2: Lässt sich das Itemverständnis durch ein Erklärvideo optimieren?

F3: Auf welche emotionalen, motivationalen und kognitiven Erfahrungen im Unterricht greifen Schüler:innen zurück, wenn sie Items zur Unterrichtsnutzung beantworten?

Methode

Auf der Grundlage eines theoretischen Modells zu Nutzungsprozessen im Unterricht, die fragebogenbasiert zugänglich sind (Wohlbefinden, Motivation, Aufmerksamkeit, Verarbeitung, Passung; Gossner et al., in Vorb.), wurde ein Fragebogen mit fünf uni- und bipolaren Items entwickelt. Aufgrund der kurzen Bearbeitungsdauer eignet sich das Instrument perspektivisch für den Einsatz im Ambulatory Assessment.

Zur Rekonstruktion der individuellen Itemverarbeitung, die sich nach Tourangeau et al. (2000) viergestuft darstellt (Iteminterpretation, Abruf relevanter Informationen, Urteilsbildung, Selektion einer Antwortkategorie), wurden insgesamt $N=46$ Schüler:innen der 3.-6. Primarklasse in kognitiven Interviews mittels *Probing*- und *Paraphrasing*-Techniken (Lenzner et al., 2015; Prüfer & Rexroth, 2005) befragt. Das gestufte Befragungskonzept umfasste sich teilweise überlappende Teilstichproben: $N=33$ Lernende wurden zu ihrem Itemverständnis interviewt (F1), davon erhielten $n=26$ ergänzende Instruktionen zu den Items mittels Erklärvideo, das auf Mayers (2014) Theorie des multimedialen Lernens basiert (F2). Lernende aus zwei Schulklassen ($n=28$) wurden nach einer Unterrichtsstunde, in der sie den Fragebogen ausgefüllt haben, zu ihren Antworten im Hinblick auf konkrete Unterrichtserlebnisse befragt (F3).

In einer induktiven Auswertungslogik werden in den Transkripten zunächst mittels zusammenfassender qualitativer Inhaltsanalyse nach Mayring (2022) wesentliche Bereiche identifiziert, die Schüler:innen mit den Items verbinden. Sodann folgt in Anlehnung an Lenske (2016) sowie Guttke und Porsch (2024) eine strukturierende Inhaltsanalyse mit Fokus auf intendiertem und nichtintendiertem Itemverständnis.

Ergebnisse

Zum Zeitpunkt der Einreichung laufen die Auswertungen noch. Erste Ergebnisse zeigen, dass selbst korrekt paraphrasierte Items nicht immer dazu führen, dass Schüler:innen diese anhand passender Erfahrungen im Unterricht beantworten können. Zudem gibt es allererste Hinweise, dass ein Erklärvideo für eine adäquate Unterscheidung der unterschiedlichen Nutzungsaspekte zuträglich ist. Weitere Ergebnisse werden auf dem Poster präsentiert.

IS-03: Invited Symposium: Kompetenzdiagnostik in der beruflichen Bildung

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: SO 108

Chair der Sitzung: Stephan Abele

Chair der Sitzung: Carolin Hahnel

Kompetenzdiagnostik in der beruflichen Bildung

Chair(s): **Stephan Abele** (Technische Universität Dresden)

Diskutant*in(nen): **Carolin Hahnel** (Ruhr-Universität Bochum)

Das Thema des Symposiums ist die computerbasierte Diagnostik beruflicher Kompetenzen von Lernenden in der beruflichen Ausbildung. Sowohl aus wissenschaftlicher als auch aus fachpraktischer Sicht ist dieses Thema relevant, weil es sich auf die Erfassung komplexer domänenspezifischer Bildungsergebnisse bezieht und dabei mehrere gesellschaftliche Herausforderungen aus dem Bereich der beruflichen Bildung adressiert.

Berufliche Kompetenzen befähigen zur zielgerichteten und koordinierten Ausführung umfangreicher Handlungssequenzen (z.B. Steuerung technischer Systeme, Reparatur eines Autos, Auswahl geeigneter Lieferant:innen). Fraglich ist, mit welchen Verfahren und mit welcher psychometrischen Güte sich solche Kompetenzen erfassen lassen. Eine Antwort auf diese Frage erfordert Expertise in unterschiedlichen Disziplinen. Es sind berufliche Kompetenzen inhaltlich zu modellieren (Fachexpertise) sowie geeignete diagnostische Verfahren und aussagekräftige Evidenzen zur psychometrischen Güte zu generieren (diagnostische bzw. psychometrische Expertise). Die vorgestellten Studien sind das Ergebnis der Zusammenarbeit zwischen Wissenschaftler:innen unterschiedlicher Disziplinen wie Fachdidaktik und pädagogischer Psychologie sowie Akteur:innen und Entscheidungsträger:innen der Berufsbildungspraxis.

Das Symposium beinhaltet drei Beiträge und eine Diskussion. Präsentiert werden drei empirische Studien, in denen unterschiedliche berufliche Kompetenzen mit einem breiten Spektrum innovativer diagnostischer Verfahren fokussiert werden. Vorgestellt wird jeweils, welche berufliche Kompetenz wie diagnostiziert, welche Befunde zur psychometrischen Güte ermittelt und welche Schlüsse daraus gezogen werden. In der Diskussion werden die Ergebnisse der Studien kritisch gewürdigt, integrativ diskutiert und u.a. vertieft, welche Lehren aus den Studien für die Diagnostik beruflicher bzw. domänenspezifischer komplexer Kompetenzen (nicht) gezogen werden können.

Der erste Beitrag fokussiert die berufliche Kompetenz von Auszubildenden der Fachinformatik. Erfasst werden soll, inwiefern diese Auszubildenden in der Lage sind, IT-Systeme bzw. Soft- und Hardware zu nutzen, um technische Systeme zu steuern (z.B. Temperaturführung in einem Gewächshaus). Zur Diagnostik wurden Concept Maps eingesetzt, die von Auszubildenden in einer Computerumgebung erstellt wurden. Im Zentrum des Beitrags steht die Validität, d.h. die Frage, inwiefern aus den Concept-Map-Ergebnissen auf die Ausprägung der Fokalkompetenz geschlossen werden darf. Dafür werden u.a. Evidenzen zum empirischen Verhältnis der Fokalkompetenz zu verwandten Merkmalen (z.B. Fachnote oder Fachwissen) und weniger verwandten Merkmalen (z.B. kognitive Grundfähigkeit) generiert.

Der zweite Beitrag fokussiert die berufliche Kompetenz von Auszubildenden der Kfz-Mechatronik. Erfasst werden soll, inwiefern diese Auszubildenden Autos reparieren können. Als diagnostisches Verfahren wurde ein computerbasierter Test eingesetzt, in dem Auszubildende Aussagen zu Videovignetten beurteilen müssen. Die Videos zeigen Reparaturhandlungen eines Schauspielers. Im Zentrum des Beitrags stehen Fragen zur Trennschärfe und internen Konsistenz der Testergebnisse sowie zur Validität beim Einsatz des Tests als Prüfungsinstrument in der praktischen Gesellenprüfung. Mit Blick auf die Validität wird dargestellt, inwiefern die Testergebnisse im Sinne der Kfz-Reparaturkompetenz interpretiert werden können und einen Beitrag zur Beurteilung prüfungsrelevanter Fähigkeiten der Auszubildenden leisten. Dafür werden die Testergebnisse u.a. mit den Ergebnissen klassischer Prüfungsstationen mit Arbeitsproben verglichen.

Der dritte Beitrag fokussiert die Problemlösekompetenz kaufmännischer Auszubildender. Erfasst werden soll, inwiefern diese Auszubildenden in der Lage sind, geeignete Lieferanten auszuwählen. Erfasst wurde diese Kompetenz mit LUCA, einer computerbasierten Bürosimulation, die reale berufliche Anforderungen abbildet und Handlungssequenzen ermöglicht. Im Fokus dieses Beitrags steht das Problem der lokalen stochastischen Abhängigkeit von Items, die sich auf einzelne Schritte inhaltlich zusammenhängender Handlungssequenzen beziehen. Präsentiert werden Erkenntnisse dazu, inwiefern Partial-Credit-Items sowie eine zweidimensionale Modellierung der betreffenden Problemlösekompetenz dieses Problem abbildern können.

Beiträge des Symposiums

Diagnostik berufsfachlicher Kompetenz bei Fachinformatiker:innen mit Concept Maps

Benjamin Ullrich¹, Nico Link²

¹Technische Universität Dresden, ²Pädagogische Hochschule Karlsruhe

Eine Besonderheit stark vernetzter technischer Systeme ist es, dass viele Prozesse „im Verborgenen“ ablaufen und nicht direkt sichtbar sind. Dies erfordert im Ausbildungsberuf Fachinformatiker:in „die Kompetenz, die physische Welt und IT-Systeme funktional zu einem cyber-physischen System zusammenzuführen“ (Sekretariat KMK, 2019). Dazu müssen Auszubildende sowohl über analytische (wie die Funktionsweise einzelner Betriebsmittel) als auch über konstruktive Fähigkeiten (z.B. bei der Synthese von Komponenten) verfügen (Walker, Link & Nickolaus, 2016). Bei der dafür notwendigen Informationsgewinnung und -verarbeitung spielt die Bildung von mentalen Modellen (Gentner & Stevens, 1983) eine zentrale Rolle.

Eine Möglichkeit um mentale Modelle zu erfassen, stellen Concept Maps dar (Chang, 2007; Hamdiyati, Sudargo, Redjeki, & Fitriani, 2018; Khajeloo & Siegel, 2022). Während es in der Wirtschaftspädagogik bereits empirische Arbeiten zum Einsatz von Concept Maps gibt (z. B. Ryssel, 2019), liegen zu informationstechnischen Ausbildungsberufen bislang noch keine systematischen empirischen Arbeiten vor. Ungeklärt sind hierbei u.a. folgende Validitätsfragen:

1. Wie hoch korrelieren die Qualität von Concept Maps mit Leistungsbewertungen im Lernfeld (Fachwissenstest und Fachkompetenznote), kognitiven Fähigkeiten sowie Ergebnissen eines Selbsteinschätzungsbogens?
2. Welche Unterschiede in der Kompetenz, die physische Welt und IT-Systeme funktional zu einem cyber-physischen System zusammenzuführen, zeigen sich zwischen von Experten erstellten Referenz-Maps und den von Auszubildenden angefertigten Novizen-Maps?

An der Studie nahmen ≈ 130 angehende Fachinformatiker:innen im zweiten Ausbildungsjahr teil. Die Datenerhebungen erfolgten im Lernfeld 7 „Cyber-physische Systeme ergänzen“ in Einzelarbeit computerbasiert mit dem Programm CmapTools. Den Auszubildenden wurden dabei zehn Konzepte vorgegeben (Novak & Cañas, 2006), welche anschließend zu möglichst

umfangreichen, vollständigen und fachlich korrekten Concept Maps verbunden werden sollten. Eingebettet waren die Concept Maps in ein lernfeldübergreifendes berufliches Projekt. Die Bearbeitung des Projektes erfolgte in Zweiergruppen und die Abschlussbewertung ergab sich aus einem Fachwissenstest zu elektrischen Grundgrößen und dem Stand der technischen Dokumentation zu drei Messzeitpunkten. Die Datenerhebung der drei Concept Maps erfolgte, nach einer Einführung in Anlehnung an Sumfleth (2010), in unmittelbarer zeitlicher Nähe der Messzeitpunkte. Jeweils nach der Erstellung einer Concept Map wurde ein Selbsteinschätzungsfragebogen mit 10 Items (Ryssel, 2019) durchgeführt. Außerdem wurden die kognitiven Fähigkeiten (fluide Intelligenz, Wortschatz und Zahlenfolge) mit dem standardisierten CFT 20-R WS/ZF-R erfasst (Weiß, 2006).

Für den Einsatz als Diagnoseinstrument müssen Concept Maps nach zuvor festgelegten Kriterien ausgewertet werden. Graf (2014) unterscheidet dabei die Auswertungsverfahren: globale Auswertungsstrategie nach Novak, Differentialdiagnose einzelner Begriffe, graphentheoretische Aspekte und einen Vergleich mit einer Referenz-Map. Um die Nachteile einzelner Verfahren zu minimieren, wird in den anstehenden Analysen ein kombiniertes Auswertungsverfahren umgesetzt.

Die Bewertung der insgesamt 361 Concept Maps erfolgt anhand von Vergleichen mit Referenz-Maps nach den Qualitätskriterien „Umfang/Vernetztheit“, „Übereinstimmungsgrad“ und „inhaltliche Übereinstimmung“ (Ullrich, Hellmich, & Link, 2023). Die drei Referenz-Maps, als Vorlage für die Auswertung, wurden wie folgt erstellt: (a) Jeder der drei Experten skizzierte eine individuelle Concept Map mit den vorgegebenen Konzepten. (b) Diese wurden den anderen Experten vorgestellt und erläutert. (c) Eine der drei Concept Maps wurde im Team ausgewählt und im Anschluss zu einer gemeinsamen Referenz-Map erweitert.

Die Auswertung der Novizen-Maps wird von drei Personen durchgeführt, wobei die Interrater-Reliabilität ermittelt wird. Dabei werden mindestens 10 % der Daten mehrfach codiert und das Überschneidungsmaß des Holsti-Koeffizienten berechnet (Döring & Bortz, 2016; Kolb, 2014).

Um die Validität zu beurteilen, werden Korrelationen berichtet und diskutiert. Die zweite Forschungsfrage wird mittels qualitativer Analyse der einzelnen Concept Maps beantwortet.

Die Datenerhebungen wurden bereits abgeschlossen, die Datensätze sind aber erst bei einer Teilstichprobe ausgewertet, wodurch sich bisher erste Tendenzen ergeben. Bis zur Tagung im Januar 2025 werden die Daten vollständig ausgewertet sein.

Digitale Kompetenzmessung im Handwerk: Testgüte und Praxistransfer videobasierter Prüfungsstationen für Kfz-Mechatroniker:innen

Stefan Hartmann¹, Tobias Gschwendtner²

¹Technische Universität Dresden; ²Pädagogische Hochschule Ludwigsburg, ²Pädagogische Hochschule Ludwigsburg

Bei der Erfassung berufspraktischer Kompetenzen birgt der Einsatz innovativer digitaler Messverfahren große Potentiale hinsichtlich der Testgüte, Testfairness und Testökonomie (Kirschner, Park, Malone, & Jarodzka, 2017). Digitale Testformate gelten als aussichtsreiche Lösungsansätze für die zunehmenden technischen und personellen Herausforderungen im Prüfungswesen in der beruflichen Bildung (Deutscher & Winther, 2019; Zentralverband Deutsches Kraftfahrzeuggewerbe [ZDK], 2024).

Im interdisziplinären Forschungs- und Transferprojekt DigiDIn-Kfz entwickelten Akteur:innen aus Fachdidaktik, pädagogischer Psychologie und beruflicher Bildung auf Videovignetten basierende Prüfungsstationen zur Erfassung von Kompetenzen im Bereich der Kfz-Reparatur. Ziel ist der Einsatz der Stationen in der praktischen Gesellenprüfung am Ende der Ausbildung von Kfz-Mechatroniker:innen. In den vollständig computerbasierten Stationen zu fünf Themenblöcken werden kurze Videosequenzen gezeigt, die ganzheitliche Reparaturhandlungen in einzelne Arbeitsschritte unterteilt darstellen. In den Videos sind fachlich adäquate sowie inadäquate Handlungsweisen zu sehen. Auf jedes Video folgen geschlossene Fragen zu den gezeigten Arbeiten. Korrekt markierte Antwortoptionen werden mit jeweils einem Punkt gewertet; die Gesamtpunktzahl einer solchen digitalen Prüfungsstation kann als Indikator für Kfz-Reparaturkompetenz interpretiert werden (Hartmann, Güzel, & Gschwendtner, 2023).

Inwieweit diese Interpretation angemessen ist, wurde im Rahmen mehrerer Validierungsstudien evaluiert. In zwei Laborstudien wurden verschiedene Standards zur Beurteilung von Validität (AERA, APA, & NCME, 2014) zunächst experimentell untersucht; die Ergebnisse fielen gemischt aus (Hartmann, Güzel, & Gschwendtner, 2023). Basierend auf den Befunden wurden zwei Themenblöcke (Lenkung und Motorsteuerriemen) für die Gesellenprüfung ausgewählt und in Zusammenarbeit mit einem Kfz-Prüfungsausschuss optimiert. Um sie als reliable, valide, aber auch akzeptierte und justiziable Prüfungsstationen einsetzen zu können, muss ihre Testgüte evaluiert werden. Zum einen sind Reliabilitäten zu berechnen und zum anderen ist zu prüfen, welchen Beitrag die Punktzahlen der digitalen Stationen zur Kompetenzerfassung in der ansonsten auf klassischen Arbeitsproben basierten Gesellenprüfung leisten.

Hierzu wurde eine Feldstudie mit 232 Prüfungskandidat:innen in einer praktischen Gesellenprüfung durchgeführt, bei der die zwei digitalen Stationen neben sechs Prüfungsstationen mit Arbeitsproben eingesetzt wurden. An zwei dieser sechs klassischen Stationen mussten an realen Fahrzeugen Reparaturaufgaben zu denselben Themenblöcken wie an den beiden digitalen Stationen gelöst werden; die restlichen vier Stationen adressierten andere Fahrzeugsysteme und somit auch andere Aspekte von Reparaturkompetenz.

Die beiden digitalen Prüfungsstationen weisen trotz ihrer inhaltlichen Heterogenität zufriedenstellende interne Konsistenzen auf ($.68 \leq \alpha \leq .70$; $.73 \leq \lambda \leq .74$). Um zu beurteilen, wie stark die Ergebnisse der digitalen Stationen innerhalb der Prüfung zur Differenzierung zwischen Auszubildenden beitragen, wurden für alle Stationen klassische Trennschärfeparameter berechnet. Hierbei betrachtet man die Gesellenprüfung als Test und die acht Prüfungsstationen als einzelne Items. Die Itemtrennschärfen streuen zwischen .23 und .56; die Trennschärfen der computerbasierten Stationen betragen .43 und .46. Neben Trennschärfen wurden die Korrelationen aller Stationen untereinander berechnet. Sie reichen von .10 bis .59 und liegen im Mittel bei .28. Die computerbasierten Prüfungsstationen korrelieren mit den inhaltlich ähnlichen Arbeitsproben zu .38 ($p < .001$) und .59 ($p < .001$).

Die Befunde zeigen, dass die inhaltlich breit streuenden Stationen einer praktischen Kfz-Gesellenprüfung nur moderat miteinander korrelieren. Mutmaßliche Ursachen sind neben der großen inhaltlichen Breite im eingeschränkt standardisierten Bewertungsverfahren der klassischen Stationen zu suchen, welches mit geringer Objektivität und Reliabilität assoziiert wird (Bejar, Williamson, & Mislevy, 2006) und empirisch bisher wenig untersucht ist. Betrachtet man die Prüfungsstationen als in sich homogene Items eines Tests, so weisen sie zufriedenstellende Trennschärfen auf. Die beiden computerbasierten Stationen integrieren sich aus messtheoretischer Sicht gut in die Prüfung und leisten zur Kompetenzmessung einen mindestens ebenso guten Beitrag wie die Prüfungsstationen mit Arbeitsproben.

Nachholbedarf gibt es bei der theoretischen Modellierung der in den Prüfungen erfassten Fähigkeiten, die bisher weitgehend als Kriterienkataloge definiert sind, aber keinem wissenschaftlichen Kompetenzkonstrukt zugrunde liegen. Im Prüfungswesen besteht großes Interesse an Praxistransfers wissenschaftlich evaluierter Instrumente; die Weiterverwendung dieser als auch künftig zu entwickelnder digitaler Prüfungsstationen wird als notwendig erachtet.

Authentische Lehr-Lernsituationen der kaufmännischen Domäne als Werkzeuge zur Beschreibung des beruflichen Kompetenzerwerbs

Esther Winther¹, Viola Deutscher², Andreas Rausch³, Jürgen Seifried³, Fabio Fortunati¹

¹Universität Duisburg-Essen, ²Universität Göttingen, ³Universität Mannheim

Der Kompetenzerwerb kaufmännischer Auszubildender ist an Lerngelegenheiten in Schule und Betrieb gebunden und abhängig davon, wie diese Gelegenheiten instruktional begleitet werden (Seifried et al., 2021; Deutscher & Winther, 2018). Während mittlerweile berufliche Kompetenzstruktur und -niveaumodelle vorliegen (u.a. Abele et al., 2016), fehlen nach wie vor Befunde zu Effekten des Unterrichts auf den Kompetenzerwerb (z.B. Naumann et al., 2019). Eine Möglichkeit zur Generierung solcher Befunde ist es, Lerngelegenheiten möglichst exakt zu beschreiben. Mit der Bürosimulation LUCA (Rausch et al., 2021) liegt ein didaktisches Arrangement vor, das gleichsam – mittels komplexer Lernaufgaben – als Performance Assessment genutzt werden kann.

Das Assessment ist so konstruiert, dass es die gestiegene Komplexität der Arbeitswelt berücksichtigt, auf die Fähigkeit zum Problemlösen (Facette der beruflichen Handlungskompetenz) abstellt und die Erhebung des Problemlösens als komplexes Phänomen in einer authentischen und inhaltsreichen Testumgebung ermöglicht (Seifried et al., 2021). Allerdings ist die Kompetenzdiagnostik auf Basis komplexer Performanzen in authentischen Arbeitsszenarien mit psychometrischen Herausforderungen verbunden.

Eine Herausforderung bei der Konstruktion (beruflicher) Assessments besteht darin, das Validitätsparadoxon zwischen hoher Inhaltsvalidität einerseits und der Messung konstrukt-irrelevanter Merkmale andererseits (bspw. Sprachkompetenz, Komplexität der Testumgebung, Kontextualisierung, Cognitive Load etc.) zu reduzieren (Mislevy, 2019; Minkley et al., 2021). Vorliegende Befunde deuten darauf hin, dass die gezeigte Performanz abhängig vom Aufgabendesign ist. Zudem kann die Darbietung derselben Aufgabe zu unterschiedlichen Zeitpunkten zu Performanzunterschieden führen (Ruiz-Primo et al., 1993) – in der Folge besteht die Schwierigkeit herauszufiltern, welche Leistungsfacetten das zu erfassende Konstrukt abbilden. Diese Schwierigkeit wird dadurch verstärkt, dass mittels Performance oft nur inhaltlich eng definierte Konstruktausschnitte erhoben werden (Seifried et al., 2021).

Im Beitrag wird Folgendes untersucht: (1) Welche Methoden sind geeignet, die psychometrische Güte des Performance Assessments zu erhöhen, um den Prozess des Kompetenzerwerbs schrittweise besser sichtbar zu machen? (2) Welche Befunde lassen sich mit dem Assessment generieren und wie lässt damit der Kompetenzerwerb kaufmännischer Auszubildender beschreiben?

Es werden die Befunde des LUCA-Inhaltsbereiches „Lieferantenauswahl“ vorgestellt. Für die Förderung wäre ein Assessment entlang folgender Lösungsschritte günstig (Wuttke et al., 2015): (a) Handlungsbedarfe & Informationsquellen identifizieren, (b) Informationen verarbeiten, (c) begründete Entscheidungen treffen, (d) Entscheidungen angemessen kommunizieren. Ein Item-Scoring entlang dieser Schritte weist allerdings Verletzungen der lokalen Unabhängigkeit auf (bei 151 Korrelationspaaren nach Yen's Q3-Statistik; Yen, 1984). Im Beitrag werden daher (1) Scoring-Algorithmen vorgestellt, mit denen durch fachlich-inhaltslogische Aggregation Partial-Credit-Items erzeugt werden, die Folgeabhängigkeiten reduzieren, so dass die Bedingungen einer psychometrischen Modellprüfung erfüllt werden. Es wird (2) gezeigt, dass sich – statt Problemlöseschritte zu differenzieren – modellspezifische Vorteile bei einer Trennung des Problemlöseprozesses in sprachlich-argumentative und mathematisch-analytische Lösungsanteile ergeben. Hierbei ist von besonderem Interesse, dass die Lage der Personenfähigkeiten zwischen den fachlichen Lösungsanteilen signifikant differenziert, wobei übergreifend mathematisch-analytische Fähigkeiten geringer ausgeprägt sind ($t=-8,430$, $df=2486$, $p<.001$, $r=.215$).

Die zweidimensionale Modellstruktur hat den Vorteil, dass sich verschiedene Inhaltsbereiche analytisch in die jeweiligen Lösungsanteile zerlegen und damit über verschiedene Problemszenarien hinweg empirisch vergleichen lassen. Die Lösung zeigt zudem, dass es im kaufmännischen Unterricht zwei Bereiche gibt, die im Besonderen Förderbedarfe aufweisen, aber auch Fördermöglichkeiten bieten. Verschiedene Performance Assessments (im Sinne weiterer komplexer Lernaufgaben) lassen sich über berufliche Lernprozesse zusammenbinden und als diagnostische Werkzeuge zur (formativen) Beschreibung beruflichen Kompetenzerwerbs und dessen Hürden nutzen. Der Nachteil dieser Lösung liegt in einem relativ hohen Abstraktionsniveau. Die Befunde zeigen, dass authentische Performance Assessments – die in ihren Konstruktionsaufwänden weit weniger komplex sind, als kaufmännische Kompetenztests und vor diesem Hintergrund auch von Lehrpersonen erstellt und ausgewertet werden können – geeignet sind, Hinweise auf den beruflichen Kompetenzerwerb zu geben. Der Beitrag lädt zur Diskussion darüber ein, welche Anforderungen an ein Performance Assessment zu stellen sind und wie sich diese von der Kompetenzerfassung in beruflichen Kontexten unterscheiden.

OF-3: Special session: Die GEBF Open Science Initiative: Aktueller Stand, Arbeitsfelder und zukünftige Beteiligungsmöglichkeiten

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 159

Chair der Sitzung: Jürgen Schneider

Die "GEBF Open Science Initiative" (GEBF-OSI) lädt ein, ihre Arbeit und ersten Ergebnisse kennenzulernen und sich über die Entwicklung von Open Science in der Empirischen Bildungsforschung auszutauschen. Diese Initiative wurde nach einem offenen Aufruf auf der letzten GEBF-Mitgliederversammlung als „Graswurzel-Arbeitsgruppe“ ins Leben gerufen, um Mitglieder aus den verschiedenen Disziplinen der Empirischen Bildungsforschung für das Thema Open Science zusammenzubringen. Die AG hat sich mittlerweile mit einigen engagierten Personen und in Zusammenarbeit mit Kolleg:innen der Arbeitsgruppe Empirische Pädagogische Forschung (AEPF) konstituiert.

Erstes Projekt der AG wird die Ausarbeitung einer Stellungnahme zu Open Science aus der Perspektive der empirischen Bildungsforschung sein. Dieses Dokument zielt darauf ab, bestehende Positionspapiere zu Open Science (z.B. der Europäischen Kommission, der DFG und des BMBF, aber auch disziplinärer Fachverbände) auf das Forschungsfeld der Empirischen Bildungsforschung zu beziehen und einzuordnen. Es ergänzt zudem die gemeinsame Stellungnahme der DGfE, GEBF und GFD zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten auf. Dabei ist einerseits die Berücksichtigung des interdisziplinären Charakters der Empirischen Bildungsforschung leitend. Andererseits stellt sich die Frage, wie Open Science-Prinzipien in den unterschiedlichen Forschungszugängen der Empirischen Bildungsforschung jeweils sinnvoll realisiert werden können.

Ziel des Konferenzformats ist es, den aktuellen Arbeitsstand der Initiative vorzustellen und zu den genannten Fragen mit den Anwesenden in Austausch zu treten, um die Vielfalt der Perspektiven unserer Community einzubeziehen und die Anschlussfähigkeit des Dokuments zu gewährleisten. Gezielte Diskussionen zu verschiedenen Aspekten von Open Science sollen wertvolles Feedback ermöglichen.

Zukünftig plant die AG die Erstellung weiterer Papiere und Ressourcen zu spezifischen Aspekten von Open Science für die GEBF. Diese sollen praxisnahe und konkrete Handlungsschritte sowie Optionen bieten. Ein besonderer Fokus liegt darauf, die Dokumente für Personen mit unterschiedlichen Kenntnisstufen zugänglich zu machen, um eine breite Nutzbarkeit zu gewährleisten. Die Bereitstellung von Good-Practice-Beispielen wurde dabei als wichtige Orientierungshilfe hervorgehoben.

Wir laden alle Interessierten ein, sich aktiv an den Diskussionen zu beteiligen und ihre Perspektiven einzubringen. Unsere Initiative ist offen für neue Mitglieder, auch wenn sich diese lediglich für spezifische Projekte einbringen möchten. Gemeinsam wollen wir die Prinzipien von Open Science in der empirischen Bildungsforschung verankern und weiterentwickeln.

Die GEBF Open Science Initiative: Aktueller Stand, Arbeitsfelder und zukünftige Beteiligungsmöglichkeiten

Jürgen Schneider¹, GEBF Open Science Initiative²

¹DIPF, Deutschland; ²GEBF

Die "GEBF Open Science Initiative" (GEBF-OSI) lädt ein, ihre Arbeit und ersten Ergebnisse kennenzulernen und sich über die Entwicklung von Open Science in der Empirischen Bildungsforschung auszutauschen. Diese Initiative wurde nach einem offenen Aufruf auf der letzten GEBF-Mitgliederversammlung als „Graswurzel-Arbeitsgruppe“ ins Leben gerufen, um Mitglieder aus den verschiedenen Disziplinen der Empirischen Bildungsforschung für das Thema Open Science zusammenzubringen. Die AG hat sich mittlerweile mit einigen engagierten Personen und in Zusammenarbeit mit Kolleg:innen der Arbeitsgruppe Empirische Pädagogische Forschung (AEPF) konstituiert.

Erstes Projekt der AG wird die Ausarbeitung einer Stellungnahme zu Open Science aus der Perspektive der empirischen Bildungsforschung sein. Dieses Dokument zielt darauf ab, bestehende Positionspapiere zu Open Science (z.B. der Europäischen Kommission, der DFG und des BMBF, aber auch disziplinärer Fachverbände) auf das Forschungsfeld der Empirischen Bildungsforschung zu beziehen und einzuordnen. Es ergänzt zudem die gemeinsame Stellungnahme der DGfE, GEBF und GFD zur Archivierung, Bereitstellung und Nachnutzung von Forschungsdaten auf. Dabei ist einerseits die Berücksichtigung des interdisziplinären Charakters der Empirischen Bildungsforschung leitend. Andererseits stellt sich die Frage, wie Open Science-Prinzipien in den unterschiedlichen Forschungszugängen der Empirischen Bildungsforschung jeweils sinnvoll realisiert werden können.

Ziel des Konferenzformats ist es, den aktuellen Arbeitsstand der Initiative vorzustellen und zu den genannten Fragen mit den Anwesenden in Austausch zu treten, um die Vielfalt der Perspektiven unserer Community einzubeziehen und die Anschlussfähigkeit des Dokuments zu gewährleisten. Gezielte Diskussionen zu verschiedenen Aspekten von Open Science sollen wertvolles Feedback ermöglichen.

Zukünftig plant die AG die Erstellung weiterer Papiere und Ressourcen zu spezifischen Aspekten von Open Science für die GEBF. Diese sollen praxisnahe und konkrete Handlungsschritte sowie Optionen bieten. Ein besonderer Fokus liegt darauf, die Dokumente für Personen mit unterschiedlichen Kenntnisstufen zugänglich zu machen, um eine breite Nutzbarkeit zu gewährleisten. Die Bereitstellung von Good-Practice-Beispielen wurde dabei als wichtige Orientierungshilfe hervorgehoben.

Wir laden alle Interessierten ein, sich aktiv an den Diskussionen zu beteiligen und ihre Perspektiven einzubringen. Unsere Initiative ist offen für neue Mitglieder, auch wenn sich diese lediglich für spezifische Projekte einbringen möchten. Gemeinsam wollen wir die Prinzipien von Open Science in der empirischen Bildungsforschung verankern und weiterentwickeln.

PS-21: Chemieunterricht: Herausforderungen und digitale Ansätze

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: SO 322

Chair der Sitzung: Amina Zerouali

DigiProCH: Eine Fortbildung zur Gestaltung von problembasiertem, digitalen Chemieunterricht

Amina Zerouali, Doris Lewalter, Stephanie Moser, Miriam Lechner, Jenna Koenen

Technische Universität München, Deutschland

Theoretischer Hintergrund

In der naturwissenschaftsdidaktischen Literatur wird der Prozess der Erkenntnisgewinnung als komplexes, wissensbasiertes Problemlösen konzeptualisiert (Kauertz et al., 2010). Problemorientierte Lehr- und Lernansätze nutzen Experimente und Modelle als fachspezifische Methoden zur Lösung naturwissenschaftlicher Probleme (Pahl & Berchtold, 2019). Diese Ansätze fördern den Kompetenzerwerb im Bereich der Erkenntnisgewinnung und tragen wesentlich zur Entwicklung einer naturwissenschaftlichen Grundbildung (Scientific Literacy) bei (Wellnitz et al., 2017). Digitale interaktive Visualisierungen oder Simulationen können als Teil des Problemlöseprozesses eingesetzt werden, um ein tieferes Verständnis von Ursache-Wirkungs-Zusammenhängen zu ermöglichen (Nerdel, 2002).

Trotz zahlreicher Vorteile (Wiemann et al., 2010) erfordert eine didaktisch sinnvolle Implementierung digitaler, interaktiver Visualisierungen in problemorientierte Unterrichtsprozesse sorgfältige Planung, um den Lernzuwachs zu maximieren (Schnotz & Bannert, 2003). Die Umsetzung digitalgestützter, problembasierter Lehr- und Lernsituationen stellt insbesondere hohe Anforderungen an die Lehrkräfte, vor allem in Bezug auf die Herausforderungen beim Experimentieren und Modellieren. Daher müssen Lehrkräfte über Kompetenzen im Einsatz und in der Nutzung solcher digitalen Tools verfügen.

Jedoch geben lediglich ein Viertel der Lehrpersonen an, den Umgang mit digitalen Medien in ihrer universitären Ausbildung erlernt zu haben (ICLIS, 2018). Außerdem nutzen deutsche Lehrkräfte im internationalen Vergleich selten digitale Medien (Eickelmann & Drossel, 2020). Hier können Fortbildungen ansetzen, um Lehrkräfte in die Lage zu versetzen, digitale Medien und problemorientierte Ansätze effektiv in ihrem Unterricht zu integrieren und miteinander zu kombinieren.

Zielsetzung und Fragestellung

Die hier gestaltete Fortbildung zielt daher darauf ab, Lehrkräften die Gestaltung eines problemorientierten Chemieunterrichts mit digitalen Medien zu vermitteln. Der Fokus liegt auf der Nutzung digitaler interaktiver Visualisierungen und Simulationen in einem problemorientierten Setting. Zur Evaluation ergeben sich daher folgende Forschungsfragen:

1. Wie schätzen Lehrkräfte ihre Kompetenzen im Umgang mit digitalen Medien und die Fähigkeit diese im Unterricht zu implementieren vor und nach der Teilnahme an der DigiProCH-Fortbildung ein?
2. Wie bewerten Lehrkräfte ihre Fähigkeit problemorientierte Lehransätze effektiv zu implementieren vor und nach der DigiProCH-Fortbildung?

Methodik der Fortbildung

Die Konzeption der Fortbildung orientiert sich an empirisch bestätigten Richtlinien und Prinzipien (z.B. DVLfB, 2018; Lipowski & Rejzack, 2019). Die Fortbildung ist in drei Phasen unterteilt. In der ersten Präsenzphase werden die Lehrkräfte in das Thema des digitalgestützten problemorientierten Lernens eingeführt. In der zweiten, asynchronen digitalen Phase erarbeiten die Teilnehmenden Inhalte eigenständig mithilfe der Materialien der Selbstlernplattform *Toolbox Lehrerbildung*. Der Schwerpunkt liegt auf der Implementation interaktiver Visualisierungen im eigenen Unterricht. Die FortbildnerInnen stehen dabei als Coaches unterstützend zur Verfügung (Göb, 2017). In der dritten Phase wird der Einsatz der digitalen Tools im eigenen Unterricht reflektiert. Im Fokus steht dabei der Erfahrungsaustausch der praktischen Umsetzungen im Unterricht und die evidenzbasierte Reflexion der eigenen Erfahrungen.

Unterrichtsvideos dienen während aller Phasen als unterstützendes Instruktions- und Reflexionsmedium, um ein situiertes Lernen durch die Verzahnung von Theorie und Praxis zu fördern (Kleinknecht et al., 2014). Die gescripteten Unterrichtsvideos sind auf die thematischen Schwerpunkte der Fortbildung ausgerichtet und bieten aufgrund ihrer kompakten und prägnanten Darstellung wesentlicher Unterrichtsprozesse einen erheblichen didaktischen Mehrwert (Blomberg et al., 2013).

Evaluation

Zur Evaluation der Fortbildung werden drei Messzeitpunkte im Pre-Post-Follow-Up-Test-Design genutzt: vor Beginn, nach Abschluss der Fortbildung und ca. 4–6 Wochen nach dem Besuch der Fortbildung. Dabei werden Aspekte wie soziodemographische Daten, Selbstwirksamkeit beim Einsatz digitaler Medien, Transferklima, digitale Kompetenzen, Nutzung digitaler Tools, problemorientiertes Lernen und Fachsprache mittels Fragebogen erfasst. Die Instrumente beruhen auf in der Forschung bereits etablierten, standardisierten Skalen.

Die geplanten Analysen umfassen deskriptive Statistiken, *t*-Tests für verbundene Stichproben, Berechnungen der Effektstärken, Varianzanalysen (ANOVA), Korrelationen zur Untersuchung von Zusammenhängen und Regressionsanalysen zur Untersuchung der prädiktiven Kraft bestimmter Variablen.

Ausblick

Die ersten Fortbildungen sind für den Herbst 2024 angesetzt. Im Rahmen eines Vortrages sollen die ersten Ergebnisse der Evaluation der ersten Fortbildungsdurchläufe vorgestellt werden.

DigiProMIN Chemie: Evaluationsergebnisse einer Fortbildungsreihe zu digitalen Medien für den Chemieunterricht

Dominik Diermann¹, Amitabh Banerji², Sascha Bernholt³, Constantin Egerer², Carolin Flerlage³, Stefanie Lenzer³, Ilka Parchmann³, Jenna Koenen¹

¹Technische Universität München, Deutschland; ²Universität Potsdam, Deutschland; ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland

Digitale Medien können Lerninhalte aufzeigen und (be-)greifbarer machen. Sie sind deshalb auch aus der Lehrkräftebildung nicht mehr wegzudenken (Lachner, Scheiter & Stürmer, 2020; Bonnes, Wahl & Lachner, 2022) – spätestens seit dem letzten Aufschwung bezüglich der Sichtbarkeit und wahrgenommenen Relevanz für den Unterrichtsalltag durch die COVID-19 Pandemie (Fütterer et al., 2021; Huber et al., 2020). Speziell der Chemieunterricht bietet mit seinen fachspezifischen Methoden und Arbeitsweisen (vgl. Nerdel, 2017; Sommer, Wambach-Laicher & Pfeifer, 2018) vielfältige Potentiale für eine effektive Nutzung

digitaler Medien, wie etwa den Wechsel zwischen Stoff- und Teilchenebene, der Unterstützung beim Experimentieren oder das Aufzeigen von Alltagsrelevanz chemischer Inhalte. Häufig verfügen Lehrkräfte jedoch noch über unausgeprägtes digitalisierungsbezogenes themen- und fachspezifisches Wissen (vgl. ICILS-Studie; Eickelmann, Bos & Labusch, 2019) und auch die Akzeptanz und Wirksamkeitseinschätzung von innovativen, digitalgestützten Ansätzen steht oftmals zur Diskussion. Diese bilden jedoch die Grundlage für einen gewinnbringenden Einsatz. An dieser Stelle setzt das *DigiProMIN* (Digitalisierungsbezogene und digital gestützte Professionalisierung von MIN-Lehrkräften) Chemie Projekt an, welches als Teilprojekt des vom BMBF geförderten Kompetenzverbundes *lernen:digital* insgesamt fünf Fortbildungen zum Thema „digitale Medien für den Chemieunterricht“ entwickelt, anbietet und evaluiert. Die Fortbildungen wurden durch die Technische Universität München, das Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik in Kiel und die Universität Potsdam anhand empirisch validierter Prinzipien und Richtlinien konzipiert (vgl. Emden & Baur, 2017; Lipowsky & Rzejak, 2012, 2017, 2019) und werden während der Projektlaufzeit auf Grundlage der *Unified Theory of Acceptance and Use of Technology* (UTAUT; Šumak & Šorgo, 2016) sowie dem *Technologie-Akzeptanz Modell* (TAM; Davis, Bagozzi & Warshaw, 1989) evaluiert und im Sinne eines Design-Based-Research Ansatzes optimiert. Im Fokus stehen bei der Evaluation Fragen nach den Gelingensbedingungen und der Qualität der angebotenen Fortbildungen sowie den Einschätzungen von motivationalen bzw. volitionalen Hintergründen der teilnehmenden Lehrkräfte.

Das *DigiProMIN* Fortbildungsangebot teilt sich in ein einführendes Orientierungsmodul und vier Vertiefungsmodulen, die verschiedene Facetten des Chemieunterrichts fokussieren: Das online Orientierungsmodul „Chemieunterricht mit digitalen Medien innovieren“ diskutiert Möglichkeiten zur theoriegeleiteten Auswahl und Gestaltung digital gestützter Unterrichtsbausteine mit Hilfe des *DiKoLAN* Kompetenzrahmens (vgl. Becker et al., 2020; Henne, Möhrke, Thoms & Huwer, 2022) und des *SAMR*-Modells (Puentedura, 2013). Die vier Vertiefungsmodulen können unabhängig davon besucht werden und werden ab 2024 in Präsenz angeboten. Die Fortbildung mit Namen „Chemie im Kontext 2.0 - authentisch, motivierend und kollaborativ“ zeigt, wie relevante und motivierende Kontexte digital gestützt für den Chemieunterricht aufgearbeitet werden können und arbeitet mit verschiedenen digitalen Medien auf, wie die Unterrichtskonzeption Chemie im Kontext digital umgesetzt werden kann. Das Modul „Individuelle Lernverläufe aufzeigen“ thematisiert am Beispiel einer Unterrichtssequenz in Moodle Möglichkeiten zur digital gestützten Diagnose, formativen Assessment und dem Umgang mit Schülerschwierigkeiten. Die Fortbildung „Digital gestütztes Experimentieren“ beinhaltet exemplarische digitale Medien zur Unterstützung des Experimentierprozesses sowie dem Ansprechen neuartiger Lernziele. Sie setzt auf automatisierte Titrationsroboter und die Erstellung von interaktiven eBooks als digitale Experimentier-Assistenten. Im vierten Vertiefungsmodul mit dem Titel „CHAMP: chemische Animationen mit PowerPoint – Modelle zum Leben erwecken“ lernen Lehrkräfte, wie sie theoriegeleitet Animationen selbst erstellen und mit Hilfe dieser digitalen dynamischen Modelle ein Verständnis chemischer Prozesse auf der Teilchenebene ermöglichen können. Der Ansatz des *DigiProMIN*-Chemie Fortbildungsangebots beruht auf der Vermittlung von transferfähigem Wissen zum Einsatz digitaler Medien im Chemieunterricht sowie auf der Entwicklung eigener digital-gestützter Materialien.

Der Beitrag stellt dabei das Fortbildungsangebot mit Fokus auf den Evaluationsergebnissen der ersten Durchführungen im Sommer 2024 vor. Dabei können Einblicke in die Technologie-Akzeptanz von Lehrkräften und in die Wirksamkeit bzw. Qualitätseinschätzung der Fortbildungen und Materialien gewonnen und weitere Empfehlungen für alternative Angebote abgeleitet werden. Diese Qualitätssicherung dient nicht nur dem Angebot selbst, sondern ist auch von grundlegendem Interesse für eine zukunftsfähige Ausbildung von Lehrkräften in der dritten Phase, da allgemeine Gestaltungsprinzipien übertragen werden können.

Dem Scheitern auf der Spur: Risikofaktoren für den erfolgreichen Erwerb von Mindeststandards im Chemieunterricht

Dennis Kirstein

Universität Duisburg-Essen, Deutschland

Ein grundlegendes Verständnis in den Naturwissenschaften ist für eine aktive und selbstbestimmte Teilhabe in einer von naturwissenschaftlichen Erkenntnissen und Technologien geprägten Gesellschaft von zentraler Bedeutung (Gebhard, Hötter & Rehm, 2017; OECD, 2006). Die in diesem Zusammenhang notwendigen Kompetenzen erwerben Schülerinnen und Schüler unter anderem im Chemieunterricht. Der Anteil von Schülerinnen und Schülern, die am Ende der Pflichtschulzeit jedoch die Mindeststandards im Fach Chemie verfehlen, liegt gegenwärtig allerdings bei etwa 20 % (Schiepe-Tiska, Rönnebeck & Neumann, 2019; Weirich, Becker, & Holtmann, 2019). Für eine in diesem Zusammenhang effiziente und gezielte Förderung ist es unabdingbar zu wissen, welche Umstände dafür sorgen, dass Schülerinnen und Schüler an Mindeststandards scheitern. Zwar werden in diesem Zusammenhang immer wieder vor allem individuelle Merkmale wie Verständnisschwierigkeiten und unzureichende Lernressourcen diskutiert (u. a. Stalder et al., 2008; Bangel & Müller, 2018; Freesemann, 2013), Studien mit einem direkten Zusammenhang zwischen dem Scheitern an Mindeststandards in naturwissenschaftlichen Fächern und der Chemie im Speziellen liegen bislang jedoch noch nicht vor. Das gilt auch für Merkmale auf der Ebene unterrichtlichen Handelns wie Bemühungen um konstruktive Unterstützung. Hier konnte bereits gezeigt werden, dass bestimmte Maßnahmen besonders in eher leistungsschwachen Lerngruppen differenziell wirken (Merk et al., 2021; Kirstein, 2023).

Vor diesem Hintergrund wurde untersucht, welche unterrichtsbezogenen und individuellen Merkmale in einem direkten Zusammenhang zum Erreichen von Mindeststandards im Fach Chemie stehen. Dazu wurden auf der Grundlage eines bestehenden Leistungstest zum Fachwissen zum Thema Chemische Reaktionen (Celik, 2021) in einer Projektgruppe aus Chemielehrkräften an Hauptschulen und Chemiedidaktiker:innen Items im Forced-Choice-Format (vgl. Behrendt, 2022) konstruiert, die sich auf Kompetenzbeschreibungen zu Mindeststandards zu diesem Thema (Körbs, 2014; Walpuski, Sumfleth & Pant, 2019) beziehen. Als Schwellenwert für das Erreichen der Mindeststandards wurde die Hälfte der erreichbaren Punktzahl festgelegt, was bei den 24 Items zu den Mindeststandards mit jeweils 5 Antworten 60 Punkten entspricht. In einer Erhebung an unterschiedlichen Schulformen wurde mit Hilfe dieses Tests das Fachwissen zu chemischen Reaktionen von 221 Lernenden am Ende der Jahrgangsstufe 10 erfasst. Außerdem wurden die wahrgenommene sozio-emotionale (Mihaly et al., 2021) und fachlich-inhaltliche Unterstützung (Mang et al., 2021) sowie die wahrgenommene Differenzierung (Mang et al., 2021) über kurze Fragebögen erfasst. Zusätzlich wurden zu Beginn des Schuljahres von 111 der Lernenden Merkfähigkeit, Informationsverarbeitung und logisches Schlussfolgern als Facetten kognitiver Grundfähigkeiten (Kubinger & Hagenmüller, 2019), die bildungssprachlichen Kompetenzen (Österlein, Ropohl, Habig & Morek, 2023), das Fachinteresse Chemie (Hauerstein, 2016), das chemiebezogene Selbstkonzept (Hauerstein, 2016) sowie die allgemeine schulbezogene Anstrengungsvermeidung (Weber, 2017) als individuelle Personenmerkmale erfasst. Sämtliche Erhebungsinstrumente weisen eine hinreichend hohe Messgenauigkeit auf ($Rel_{WLE} \geq 0,718$). Der Test zum Fachwissen wurde zur Bestimmung der Testgüte IRT-skaliert. Der Test weist ebenfalls eine zufriedenstellende Messgenauigkeit auf Skalen- und Itemebene auf ($Rel_{WLE} = 0,725$; $0,865 \leq wMNSQ \leq 1,184$).

Von den 221 Lernenden erreichen 42,99 % nicht die Mindeststandards im Fachwissen zu chemischen Reaktionen. Am Gymnasium ist dieser Anteil mit 14,29 % signifikant niedriger als an der Hauptschule mit 87,23 % ($X^2(1)=51,064$, $p<0,001$). Die Analysen zeigen, dass Lernende, die an den Mindeststandards scheitern, über signifikant niedrige bildungssprachliche Kompetenzen verfügen ($t(108)=-5,214$, $p<0,001$, $d=-1,082$) und schlechter Informationen verarbeiten können ($t(108)=-2,313$, $p=0,011$, $d=-0,456$). Darüber hinaus bewerten Lernende, die an den Mindeststandards zu chemischen Reaktionen scheitern, die sozio-emotionale ($t(197)=-3,716$, $p<0,001$, $d=-0,527$) und fachlich-inhaltliche Unterstützung ($t(197)=-2,757$, $p=0,003$, $d=-0,391$) schlechter. Gleiches gilt für die wahrgenommene Differenzierung im Chemieunterricht ($t(197)=-4,847$, $p<0,001$, $d=-0,688$). Für die anderen Merkmale zeigen sich keine signifikanten Unterschiede. Binär logistische Regressionen mit dem dichotomen Erreichen der Mindeststandards als abhängige Variable legen nahe, dass ein Mangel an bildungssprachlichen Kompetenzen ($X^2(6)=29,903$, $p<0,001$, $R^2=0,325$, $OR=2,894^{**}$) sowie nicht an das Leistungsniveau der Lernenden ausgerichtete Lernangebote ($X^2(3)=26,554$, $p<0,001$, $R^2=0,167$, $OR=1,855^{***}$) als zentrale Risikofaktoren für das Erreichen der Mindeststandards im Fachwissen zu chemischen Reaktionen zu diskutieren sind. Im Beitrag werden die Ergebnisse fachdidaktisch eingeordnet, kritisch diskutiert und weiterführend Konsequenzen für eine gezielte Förderung erörtert.

Räumliche Fähigkeiten beim Arbeiten mit chemischen Repräsentationen - Mehr als nur mentale Rotation?

Sebastian Nickel, Steffen Brockmüller, Sebastian Habig

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Theoretischer Rahmen

Strukturformeln, Summenformeln, Diagramme und Graphen – Alle diese Repräsentationen sind typisch für die Chemie, um innerhalb der fachlichen Community zu kommunizieren (Talanquer, 2022), Fachkonzepte zu lernen (Rau, 2017) und Probleme zu lösen (Gilbert, 2008). Um mit externen Repräsentationen zu arbeiten, benötigen Lernende neben chemischem Fachwissen ausgeprägte Repräsentationskompetenzen, wobei dieser Studie die drei grundlegenden Kompetenzen von Kozma und Russell (1997, 2005) zugrunde gelegt werden: *Interpretation*, *Translation* und *Konstruktion*. Da chemische Repräsentation häufig räumlich abstrakt sind, benötigen Lernende zusätzlich elaborierte räumliche Fähigkeiten, d. h. die Fähigkeiten mental visuelle Bilder zu erzeugen, zu manipulieren und abzuspeichern (Lohman, 1979). Diese haben sich bereits als wichtiger Prädiktor für die Leistung von Lernenden in der Chemie herausgestellt (z. B. Buckley et al., 2018), besonders die Fähigkeit Probleme durch dreidimensionale mentale Rotation zu lösen (Harle & Towns, 2011). Auch wenn bereits eine Vielzahl räumlicher Faktoren als relevant für das Lernen mit Repräsentationen in der Chemie identifiziert wurde, liegt aktuell keine Studie vor, die den Zusammenhang zwischen Repräsentationskompetenzen (*Interpretation*, *Translation*, *Konstruktion*), chemischem Fachwissen und verschiedenen räumlichen Faktoren im Detail untersucht.

Forschungsanliegen und -frage

Aus der beschriebenen Forschungslücke lässt sich folgende Forschungsfrage ableiten: Welcher korrelative Zusammenhang besteht zwischen Repräsentationskompetenzen (*Interpretation*, *Translation*, *Konstruktion*), chemischem Fachwissen und chemiespezifischen Faktoren räumlicher Fähigkeiten?

Durch die Identifikation zentraler räumlicher Faktoren können Möglichkeiten abgeleitet werden, wie die Entwicklung von Repräsentationskompetenzen in der Chemie gezielt gefördert werden kann und wie der Umgang mit und das Verständnis von Repräsentationen verbessert werden kann. Beispielsweise könnte die gezielte Förderung spezifischer räumlicher Fähigkeiten einen Ansatzpunkt dafür sein.

Methoden und Design

Zur Beantwortung der Forschungsfrage wurde ein quantitatives Vorgehen gewählt, welches die Entwicklung eines Instruments zur Erhebung der Repräsentationskompetenzen notwendig machte: das *Chemical Representation Inventory: Translation, Interpretation, Construction* (CRI:TIC). Das CRI:TIC beinhaltet ausschließlich repräsentationsbasierte Items und wurde mittels (multidimensionaler) Raschanalysen evaluiert (Autoren, eingereicht). Das chemische Fachwissen der Studierenden wurde mit ausschließlich verbal-textlichen Items erhoben und ebenfalls raschskaliert ausgewertet. Zur Messung der räumlichen Fähigkeiten wurden insgesamt acht verschiedene psychometrische Instrumente eingesetzt (u. a. der Hidden Patterns Test (Ekstrom et al., 1976)) und entsprechend den Manualen auf Basis der klassischen Testtheorie ausgewertet. Der Zusammenhang zwischen den Konstrukten wird durch Korrelationsanalysen untersucht.

An der Studie haben $N = 341$ Studieneinsteiger:innen ($n_m = 71$, $n_w = 186$, $n_d = 2$, $M_{\text{alter}} = 19.24$ Jahre, $SD_{\text{alter}} = 1.90$ Jahre) verschiedener Studiengänge (Chemie, Biologie, Medizin und weitere) im Rahmen eines Chemievorkurses im Wintersemester 2023/24 teilgenommen. Die Datenerhebung hat sich über acht Tage erstreckt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen eine mittlere bis hohe Korrelation zwischen den Repräsentationskompetenzen (*Interpretation*, *Translation*, *Konstruktion*) und dem chemischen Fachwissen ($.36 \leq r \leq .47$). Mit Blick auf die räumlichen Faktoren lassen sich mittlere Korrelationen zwischen vier der acht Faktoren und den Repräsentationskompetenzen feststellen ($.30 \leq r \leq .52$), wobei die dreidimensionale mentale Rotation bei allen drei Kompetenzen die stärkste Korrelation aufweist ($.43 \leq r \leq .52$). Dieses Ergebnis deckt sich mit den Ergebnissen anderer Studien (z. B. Harle & Towns, 2011). Auffällig ist, dass bei der Translation zwei Faktoren weniger bedeutend erscheinen, die dafür bei der Interpretation und der Konstruktion bedeutend sind. Des Weiteren fällt auf, dass zwischen dem chemischen Fachwissen und den räumlichen Faktoren keine signifikante Korrelation festgestellt werden kann. Dies bestärkt die Annahmen zur Trennbarkeit der Repräsentationskompetenzen und des chemischen Fachwissens (Autoren, eingereicht; Sim & Daniell, 2014), zeigt aber gleichzeitig die Relevanz der räumlichen Fähigkeiten im Kontext des Lernens mit chemischen Repräsentationen. Diese Studie legt nahe, dass die dreidimensionale mentale Rotation ein guter Ansatzpunkt zur gezielten Förderung sein könnte, um die Repräsentationskompetenzen und generell den Umgang mit Repräsentationen zu unterstützen und so die Leistung von Lernenden zu verbessern. Diese Hypothese bedarf jedoch weiterer Forschung.

PS-22: Hochschulbildung

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 169

Chair der Sitzung: Zouhier Kassaballi

Bunching in Higher Education: Unintended Effects of the ECTS?

Zouhier Kassaballi¹, Oliver Himmler¹, Robert Jäcckle², Raphael Brade³

¹University of Erfurt, Germany; ²Technische Hochschule Nürnberg; ³ifo Institut - Leibniz-Institut fuer Wirtschaftsforschung an der Universität München e.V.

Students have strong incentives to pursue a high credit load per semester, given its association with favorable employment prospects, the signaling of high ability to the job market, and reduced opportunity costs (Aina & Casalone, 2020). However, we rarely find students ending their degrees before the regular study duration. In Germany, only a mere 0.03% of students graduate prior to their official study plan (Statistisches Bundesamt, 2021). The institutional structure of study programs and official recommendations on how students should organize their studies likely influence student credit-earning behavior. We examine the European Credit Transfer and Accumulation System (ECTS), which is used by virtually all European institutions of higher education and sets a salient reference point of 30 credits per semester. Specifically, we investigate whether the ECTS leads to bunching at earning 30 credits per semester, as its design is focused on a "typical student" obtaining 30 credits. This may promote bunching, as it creates (i) a kink in the cost of obtaining credits beyond 30, and (ii) a reference point at 30 credits. Thereby discouraging some students from earning more than 30 credits per semester and bunching at 30 credits instead.

We obtained observational data, from several sources, on credits per semester choice that would be representable on a wide-scale level. The first type of data source comes from collected administrative student-level data from university examination offices, this includes data on the credits per semester from a wide range of study programs and several cohorts. The second type of data source comes from nationally representative surveys, including The National Educational Panel Study (NEPS) and the Beginning Postsecondary Students Longitudinal Study (BPS:12/17) datasets.

We find significant bunching at 30 attempted and obtained credits per semester. Furthermore, we find missing mass in students earning between 31 and 36 credits per semester, which amounts to around 11.5 % of the student population. We show that not all students lack the ability to exceed RP. By reducing the missing mass among students earning above 30 credits through curriculum adjustments and incentive schemes, we could potentially decrease the time it takes marginal bunchers to complete their degrees. We estimate that if the 11.5% of students who currently constitute the missing mass were to earn an average of 32.5 credits per semester (one additional course beyond the reference point), they would graduate one semester earlier. The missing mass above 30 credits suggests that a substantial share of students is held back by the rigid features of ECTS, increasing time-to-degree and causing potentially high individual and societal costs.

In a companion field experiment, we tested whether an informational intervention encouraging students to go beyond 30 credits could change behavior and decrease the magnitude of bunching. We find no significant effects of this simple intervention on behavior, suggesting that changing the actual institutional setup may be necessary to sufficiently alter the ECTS cost structure and reference point influence.

Die universitäre Herkunft der ProfessorInnen deutscher Universitäten: Erhebliche Konzentration in Psychologie und Politikwissenschaften.

Tobias Grabosch, Mark Lutter, Thomas Heinze

Bergische Universität Wuppertal, Deutschland

Einleitung

Die universitäre Herkunft von ProfessorInnen konzentriert sich auf wenige einflussreiche Departments. Zahlreiche Netzwerkstudien haben dies für das US-amerikanische Universitätssystem bestätigt. Mit Gini-Koeffizienten zwischen .58 und .77 [1-9] zeigt sich eine erhebliche institutionelle Stratifizierung. Für das deutsche Universitätssystem fehlen derartige Studien. Dieser Beitrag analysiert die Verteilung der universitären Herkunft im deutschen Universitätssystem und untersucht die Ausbildungs- und Einstellungsnetzwerke von ProfessorInnen in den Fächern Psychologie und Politikwissenschaften. Diese Netzwerke bestehen aus den Promotionsorten und aktuellen Beschäftigungsstellen der ProfessorInnen.

Theorie

Eine einflussreiche Netzwerkstudie von Burris [2] untersucht die Rekrutierungsstrukturen von Universitätsdepartments und entwickelt eine Prestigetheorie basierend auf Max Weber und Pierre Bourdieu. Übertragen auf das akademische System besagt sie, dass prestigeträchtige Departments Eigenschaften und Fähigkeiten besitzen, die ihnen den Zugang zu erstrebenswerten Ressourcen ermöglichen. Diese streben nach sozialer Geschlossenheit, um ihren Status zu erhalten:

H1: Prestigeträchtige Departments an deutschen Universitäten neigen dazu, ihre ProfessorInnen aus anderen prestigeträchtigen Departments zu rekrutieren.

Prestigelose Departments versuchen, Prestige zu erlangen, scheitern aber daran, ihre DoktorandInnen mit dem notwendigen Habitus auszustatten:

H2: Prestigelose Departments neigen dazu, ihre Professoren aus prestigeträchtigen Departments zu rekrutieren.

Daten und Methode

Die Daten stammen aus einer Vollerhebung, die von Mark Lutter und seinem Team 2018/2019 durchgeführt wurde [10]. Die Daten wurden von den Webseiten aller deutschen Universitätsdepartments für Psychologie und Politikwissenschaften erhoben und umfassen den Promotionsort sowie den aktuellen Standort aller ProfessorInnen. In der Psychologie umfasst der Datensatz 495 ProfessorInnen in 74 Departments, in der Politikwissenschaft 279 ProfessorInnen in 73 Departments.

Methodisch handelt es sich um eine Netzwerkanalyse, bei der die Departments die Akteure und die Rekrutierungen der ProfessorInnen die Verbindungen im Netzwerk darstellen. Eine Verbindung zwischen Departments A und B existiert, wenn Department A eine/n DoktorIn ausgebildet hat, der/die später von Department B als ProfessorIn angestellt wird. Die Summe der ausgebildeten ProfessorInnen wird als out-degree, die der eingestellten ProfessorInnen als in-degree festgehalten. Die Departments werden dann absteigend nach ihrem out-degree sortiert, woraus sich ihre Position im hierarchischen Netzwerk ergibt.

Ergebnisse und ihre Bedeutung

Der Gini-Koeffizient für out-degree liegt für Psychologie bei .61 und für Politikwissenschaften bei .66. Damit ist die Fähigkeit der Departments, zukünftige ProfessorInnen auszubilden, in Deutschland ähnlich ungleich verteilt wie in den USA. Lorenzkurven zeigen, dass in der Psychologie 10% der Departments 53% aller ProfessorInnen ausgebildet haben. In der Politikwissenschaft haben 10% der Departments 60% der ProfessorInnen ausgebildet.

In der Psychologie wurden 77% und in den Politikwissenschaften 79% der ProfessorInnen in Quartil 4 auch dort ausgebildet, was H1 unterstützt. In der Psychologie wurden nur 1% der ProfessorInnen in Quartil 1 auch dort ausgebildet, in der Politikwissenschaft sind es 0%, was H2 unterstützt.

Das Ausbildungs- und Einstellungsnetzwerk von ProfessorInnen kann als wichtiger Mechanismus zur Entstehung und Verbreitung neuer Ideen gesehen werden. Die Ideendiversität hängt vor allem von der Ausbildungsdiversität der DoktorandInnen ab [1, 5]. Eine zu starke Konzentration auf wenige Ausbildungsstätten könnte somit ein Hindernis für die Entstehung und Verbreitung neuer Ideen darstellen.

Hausgemachte Verlierer: Wie sich manipulierte Leaderboards auf kognitive Leistung und intrinsische Motivation auswirken können

Amadeus J. Pickal¹, Michael Sailer¹, Matthias Stadler², Shurui Bai³, Manuel Ninaus^{4,5}, Samuel Greiff⁶, Nicolas Becker⁷, Marco Koch⁷

¹Learning Analytics and Educational Data Mining, Universität Augsburg; ²Institut für Didaktik und Ausbildungsforschung in der Medizin (DAM), LMU Klinikum, LMU München; ³Department of Mathematics and Information Technology, The Education University of Hong Kong; ⁴Institut für Psychologie, Universität Graz; ⁵LEAD Graduate School and Research Network, Universität Tübingen; ⁶Pädagogische Psychologie, Goethe-Universität Frankfurt; ⁷Differentielle Psychologie und Psychodiagnostik, Universität des Saarlandes

Theoretischer Hintergrund

Leaderboards (d. h. Visualisierungen, die die Leistung verschiedener Lernender im Vergleich zu anderen anzeigen) sind die am häufigsten verwendeten Spielelemente im Rahmen von Gamification-Anwendungen (Zichermann & Cunningham, 2011). Studienergebnisse dazu erwiesen sich bisher jedoch häufig als widersprüchlich. Ein Grund dafür könnte die Adaptivität von Leaderboards sein: Leaderboards zeigen im Laufe des Lernprozesses u. U. unterschiedliches Feedback zu Ausgangspositionen und Positionsverschiebungen an. So ist leaderboard-basiertes Feedback normalerweise differenziert (z. B. hinsichtlich „positivem“ oder „negativem“ Feedback) und vom Verhalten der Lernenden abhängig (z. B. hinsichtlich Engagement oder kognitiven Leistung; Bai et al., 2021; Nebel et al., 2016). Obwohl die Adaptivität von Feedback bei der Verwendung von Leaderboards zentral erscheint, wurde sie in der bisherigen Forschung oft vernachlässigt (siehe z. B. Ninaus & Nebel, 2021; Zainuddin et al., 2020). Trotzdem ist davon auszugehen, dass Adaptivität Auswirkungen auf Motivation und kognitive Leistung von Lernenden haben kann (siehe z. B. Bai et al., 2021; Ninaus et al., 2020).

Forschungsfragen

In dieser Studie wurde untersucht, wie unterschiedliche Konstellationen von manipuliertem leaderboard-basiertem Feedback die kognitive Leistung und intrinsische Motivation von Lernenden bei einer kognitiven Lernaufgabe beeinflussen können. Spezifisch wurde dabei untersucht, ob manipuliertes leaderboard-basiertes Feedback die Selbstwahrnehmung von Lernenden beeinflusst (Treatmentcheck; FF1), ob sich die kognitive Leistung von Lernenden bei der Lernaufgabe allgemein verändert (FF2) und, ob diese Veränderung und die intrinsische Motivation der Lernenden abhängig davon ist, wie positiv oder negativ das gegebene Feedback ausfällt (FF3).

Methode

Im durchgeführten Experiment wurden 427 teilnehmende Studierende aus zwei deutschen Universitäten einer von fünf Leaderboard-Bedingungen zugewiesen, die sich hinsichtlich des fiktiven Feedbacks zur Ausgangsposition (höher/niedriger) und zur Positionsverschiebung (aufwärts/abwärts) unterschieden. Zusätzlich gab es eine Kontrollbedingung ohne jegliches Feedback. Das leaderboard-basierte Feedback wurde während mehrerer Runden einer kognitiven Lernaufgabe systematisch manipuliert (d. h. das Feedback entsprach nicht der tatsächlichen Leistung der Lernenden) und dessen Auswirkungen auf die kognitive Leistung in der Aufgabe selbst (bezogen auf *Genauigkeit* und *Geschwindigkeit*) und die selbstberichtete intrinsische Motivation (bezogen auf *Interesse/Freude* an der Aufgabe sowie *wahrgenommene Kompetenz*) untersucht.

Ergebnisse

Der Treatmentcheck zeigte, dass die Manipulation des leaderboard-basierten Feedbacks wie vorgesehen funktionierte: Lernende schätzten abhängig von der Bedingung die eigene Leistung ($F_{4,339} = 11.22, p < .001, \eta^2 = .12$) sowie die Entwicklung der eigenen Leistung ($F_{4,339} = 22.73, p < .001, \eta^2 = .21$) unterschiedlich und passend zur zugewiesenen Experimentalbedingung ein (FF1). Des Weiteren konnten allgemeine positive Trends bei der *Genauigkeit* ($F_{3,1278} = 256.13, p < .001, \eta^2 = .08$) und *Geschwindigkeit* ($F_{3,1278} = 342.95, p < .001, \eta^2 = .08$) der kognitiven Leistung identifiziert werden (FF2). Hinsichtlich der Unterschiede zwischen den fünf Leaderboard-Bedingungen wurden kleine, aber nicht bedeutungsvolle, Gruppenunterschiede in der Leistung über die Zeit (*Genauigkeit*: $F_{12,1266} = 2.23, p = .009, \eta^2 < .01$; *Geschwindigkeit*: $F_{12,1266} = 1.90, p = .030, \eta^2 < .01$) gefunden. Bei der intrinsischen Motivation zeigten sich größere Unterschiede zwischen den Bedingungen (*Interesse/Freude*: $F_{4,340} = 3.18, p = .014, \eta^2 = .04$; *wahrgenommene Kompetenz*: $F_{4,340} = 17.12, p < .001, \eta^2 = .17$; FF3). Das höchste Motivationslevel hatte dabei die Bedingung mit höherer Ausgangsposition und aufwärts Positionsverschiebung. Das Motivationslevel der Kontrollbedingung lag über dem der Bedingungen mit negativem Feedback (zu Ausgangsposition und/oder Positionsverschiebung).

Diskussion

Die Ergebnisse der Studie zeigen, dass leaderboard-basiertes Feedback die intrinsische Motivation von Lernenden beeinflussen kann. Für diese scheint es besonders motivierend zu sein, wenn sie sich selbst auf einer höheren Ausgangsposition sehen und/oder sich ihre Position aufwärts verschiebt. Dies ist sogar dann der Fall, wenn das Feedback fiktiv ist und nicht auf tatsächlichen Leistungen basiert. Die Ergebnisse unterstreichen auch, wie wichtig es ist, leaderboard-basiertes Feedback für Lernende in Abhängigkeit von Leistung zu gestalten und, dass leaderboard-basiertes Feedback stets mit Bedacht eingesetzt werden sollte.

Digitale Fallstricke: Der negative Effekt vom Online-Testformat auf die Leistung

Nawael Niazy, Daria Benden, Fani Lauermann
Universität Bonn, Deutschland

Theorie/Fragestellung

Mit fortschreitender Digitalisierung und der damit verbundenen Verlagerung von Lehre und Leistungsprüfungen ins Digitale hat die Frage nach dem Einfluss der Lehr-/Lern- und Prüfungsmodalitäten immer mehr an Bedeutung gewonnen.

Die digitale Teilnahme an Hochschulveranstaltungen bringt einige Vorteile mit sich, wie bspw. eine höhere Flexibilität und Zugänglichkeit und im Falle von Leistungstests eine automatisierte Auswertbarkeit (Erdfelder et al., 2021). Insbesondere in MINT-Fächern finden regelmäßige teils wöchentliche Leistungstests während des Semesters Anwendung, weshalb Onlineformate in diesem Kontext besonders attraktiv wirken (Bego et al., 2024). Aktuelle Studien zeigen jedoch einen negativen Zusammenhang zwischen der Teilnahme im Onlineformat und der Prüfungsleistung im Vergleich zur Präsenzteilnahme (Atwater et al., 2023) sowie einen Rückgang der Teilnahmefrequenz und eine erhöhte Testabbruchquote im Laufe des Semesters (Sander et al., 2024). Geringeres Engagement wirkt sich zusätzlich negativ auf die Prüfungsleistung aus (Chevalier et al., 2018; Daniel & Broida, 2004; Schwerter et al., 2022).

Bisher liegt kaum Evidenz aus Studien vor, die verschiedene Testmodalitäten innerhalb derselben Kurse in authentischen Kontexten untersuchen. In dieser Studie wurden daher Daten in authentischen Hochschulveranstaltungen für MINT-Studierende gesammelt, in denen Studierende einen landesweiten Mathematiktest bearbeiteten, der die Voraussetzungen für das Studium eines MINT-Faches abbildete (WINT-Check; verfügbar im ORCA.nrw). Hervorzuheben ist neben der Authentizität der Testsituation – der WINT-Check findet regelmäßig in MINT-Studiengängen Anwendung – die besondere Vergleichbarkeit der Online- und in Präsenzapplikation des Tests, da es sich um hybride Hochschulveranstaltungen handelte. Darüber hinaus wurde eine hohe Vergleichbarkeit der Lehr-Lernsituation gewährleistet, da alle Studierenden den gleichen Mathematiktest bearbeiteten.

Die zentralen Forschungsfragen sind:

- 1) Wie unterscheiden sich die Eingangsmerkmale von MINT-Studierenden (z.B. HZB-Note, Fachinteresse, Lerngelegenheiten in Mathematik), die online versus in Präsenz an der gleichen hybriden Hochschulveranstaltung teilnehmen?
- 2) Unterscheiden sich Studierende, die Mathematikaufgaben in der Veranstaltung online vs. in Präsenz bearbeiteten in ihren testspezifischen motivationalen Einstellungen, der Anzahl bearbeiteter Testaufgaben und in ihrer Testleistung?

Methode

Die vorliegende Stichprobe setzt sich aus 1242 Studierenden zusammen, die hybride Kern- oder Vorbereitungskurse in Mathematik für Studienanfänger*innen in mathematikintensiven Studiengängen besuchten und am WINT-Check teilnahmen (Online: $n = 635$; in Präsenz: $n = 607$). Vor der Teilnahme am Test machten die Studierenden Angaben zu ihrer Demografie wie Sprachkenntnissen, HZB- und Mathenoten, studienbezogenen Variablen wie Zufriedenheit, Interesse und Erfolgserwartungen sowie ihrem mathematikbezogenen Fähigkeitsselbstkonzept und Interesse.

Der Mathematiktest bestand aus vier Teiltests. Nach jedem Teilttest bewerteten Studierende aufgabenbezogene Motivations- und Selbstkonzeptvariablen anhand validierter Skalen (z.B. Interesse am Inhaltsbereich, kognitives Engagement, aufgabenbezogenes Fähigkeitsselbstkonzept).

Studierende in Präsenz absolvierten den Mathematiktest in Papierform, während Online-Teilnehmende den Test zeitgleich in einer Videokonferenz durchführten. Forschungsassistent*innen waren bei beiden Modi anwesend, um die Datenerhebung zu überwachen und Fragen zu beantworten. Die Teilnahme war freiwillig.

Ergebnisse/Diskussion

Analysen zur ersten Forschungsfrage ergaben, dass Online-Teilnehmende im Mittel schlechtere HZB- und Mathematiknoten haben, weniger zufrieden mit ihrem Studium sind, ein geringeres Interesse am Studiengang berichten, geringere studienbezogene Erfolgserwartungen haben und ihr Interesse an und Fähigkeiten in Mathematik geringer einschätzen als Präsenz-Teilnehmende ($ps < .05$).

Analysen zur zweiten Forschungsfrage zeigten zudem, dass Online-Teilnehmende während der Testbearbeitung ein geringeres kognitives Engagement berichten und ihr Anteil bearbeiteter Aufgaben kleiner ausfällt, sie die Testinhalte weniger interessant finden und am Ende des Semesters schlechter abschneiden ($ps < .01$). Unter Kontrolle dieser Gruppenunterschiede zeigten hierarchische Regressionsanalysen, dass die Teilnahmemodalität über die Kontrollvariablen hinaus die Testmotivation und -leistung vorhersagt (u.a. Anteil bearbeiteter und korrekt gelöster Mathematikaufgaben).

Diese Ergebnisse deuten auf eine systematische Benachteiligung von Online-Teilnehmenden hin und unterstreichen den Bedarf, Methoden zu entwickeln, die eine qualitätsvolle Umsetzung von Online-Lehr-Lernangeboten ermöglichen. Darüber hinaus sollten Kontexteffekte bei der Interpretation von Testergebnissen berücksichtigt werden.

PS-23: Selbstwirksamkeit angehender Lehrkräfte

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 256

Chair der Sitzung: Stefanie Rach

ENTFÄLLT: Das Zusammenspiel von fachbezogener und unterrichtsbezogener Selbstwirksamkeitserwartung von Grundschullehramtsstudierenden im intensiven Längsschnittdesign

Lars Meyer-Jenßen¹, Martin Guljamow¹, Mathias Meje², Katja Eilerts¹, Frederik Grave-Gierlinger¹

¹Humboldt-Universität zu Berlin, Deutschland; ²Pädagogische Hochschule, Zürich

Theoretischer Hintergrund

Selbstwirksamkeitserwartung (SWE) wird als Überzeugung verstanden, zukünftigen Anforderungen kompetent begegnen zu können (Bandura, 1997) und stellt eine affektiv-motivationale Disposition professioneller Kompetenz von (angehenden) Lehrkräften dar (Blömeke et al., 2015). Die SWE stellt eine der Zielvariablen der universitären Lehrkräftebildung dar und ist umfassend hinsichtlich ihrer positiven Effekte auf das professionelle Wissen (Bates et al., 2011; Swars et al., 2007) als auch auf das spätere unterrichtliche Handeln untersucht worden (Scarpaloro & Subban, 2021). Die Bewältigung herausfordernder Anforderungen im Studium stellt für Grundschullehramtsstudierende die wesentliche Quelle ihrer SWE dar (Ma et al., 2022). Mathematik wird dabei als besonders herausforderndes Fach während der Grundschullehramtsausbildung wahrgenommen (Porsch, 2017). Studien unterscheiden oftmals zwischen der fachbezogenen SWE, also in Mathematik der Überzeugung mathematische Aufgaben lösen zu können, und der unterrichtsbezogenen SWE, also entsprechend der Überzeugung die mathematischen Inhalte später in der Schule vermitteln zu können, wobei sich empirisch positive Zusammenhänge zwischen beiden Formen zeigen (Bates et al., 2011; Lau, 2022). Studien zeigen, dass Überzeugungen von pädagogischen Fachkräften in Mathematik über einen längeren Zeitraum als stabil erachtet werden können (Jenßen et al., 2022), allerdings muss insgesamt ein Mangel an längsschnittlichen Studien zur fachbezogenen und unterrichtsbezogenen SWE bei Lehramtsstudierenden festgestellt werden (Ma et al., 2022).

Fragestellung

Der vorliegende Beitrag geht der Frage nach wie die fachbezogene SWE und die unterrichtsbezogene SWE im Verlauf einer Stochastiklehrveranstaltung bei Grundschullehramtsstudierenden miteinander zusammenhängen. Dabei wird zwischen Zusammenhängen auf between- und within-Ebene unterschieden.

Methode

Um die Forschungsfrage zu beantworten, wurden $n = 201$ Grundschullehramtsstudierende (Alter: $M = 23,57$ Jahre, $SD = 5,68$; Geschlecht: 73,5% weiblich, 24,5% männlich, 2,0% divers) über einen Zeitraum von acht Wochen jede Woche gefragt, inwieweit sie davon überzeugt sind a) die aktuellen Aufgaben in der Stochastiklehrveranstaltung lösen zu können (fachbezogene SWE) und b) die aktuellen Inhalte der Stochastiklehrveranstaltung später didaktisch vermitteln zu können. Um die Beziehungen zwischen den beiden SWE-Formen im Zeitverlauf und die autoregressiven Effekte innerhalb der beiden Formen untersuchen zu können, wurden die Daten mittels Random Intercept Cross-Lagged Panel Model (Mulder & Hamaker, 2021) modelliert. Dabei erlaubt das Modell die Trennung von between- und within-Varianz.

Ergebnisse und deren Bedeutung

Das Modell zeigte eine sehr gute Passung zu den Daten ($\chi^2(81) = 85.19$, $p = .35$, RMSEA = .02 [.00/.04], CFI = .99, SRMR = .07). Während sich auf der between-Ebene (zustandsunabhängig) ein starker positiver Zusammenhang zwischen beiden Formen der SWE ergab ($r = .63$, $p < .001$), zeigten sich auf der within-Ebene (zustandsabhängig) nur zu sechs der acht Sitzungen signifikante Zusammenhänge, die zumeist klein ausfielen (durchschnittliches $r = 0,32$). Es lagen nur sehr wenige autoregressive Effekte vor und nur zwischen der 5. und 6. Sitzung zeigte sich ein kreuzverzögerter Effekt der fachbezogenen SWE auf die unterrichtsbezogene SWE ($r = .26$). Auffallend ist jedoch, dass einerseits die Zusammenhänge auf within-Ebene im Zeitverlauf bezüglich der Stärke zunehmen (von $r = .22$ bis $r = .46$) und autoregressive Effekte erst ab der etwa vierten Sitzung bestehen. Die Intercepts unterschieden sich über die Zeit nicht signifikant voneinander, wobei die fachbezogene SWE höhere Werte aufwies als die unterrichtsbezogene SWE. Die Ergebnisse werden vor dem Hintergrund des Designs der Lehrveranstaltung (kumulativer Wissensaufbau eher in der zweiten Hälfte) und der Bedeutung für die Lehre in der Lehramtsausbildung (starke Zusammenhänge nur auf der zustandsunabhängigen Ebene) diskutiert. Das vorliegende hochfrequente Design erlaubt erstmals differenzierte Aussagen zu dem Zusammenhang von fachbezogener und unterrichtsbezogener SWE.

Development of teacher self-efficacy in different phases of teacher education over a period of 7 years

Jennifer Paetsch, Sebastian Franz

Universität Bamberg, Deutschland

Teacher self-efficacy (TSE) refers to teachers' beliefs about their ability to effectively manage tasks and challenges in their profession. Positive TSE is associated with job satisfaction and teaching commitment, highlighting the importance of fostering TSE throughout teacher education (Skaalvik & Skaalvik, 2007; Tschannen-Moran & Woolfolk Hoy, 2011). A recent synthesis by Täschner et al. (2024) suggests that TSE can be improved at all stages of teacher education by targeting Bandura's (1997) four sources of self-efficacy: mastery experience, vicarious experience, social persuasion, and physiological and emotional states. However, different stages of teacher education may address these sources to different extents (Pfitzner-Eden, 2016).

Empirical evidence suggests that novice teachers often overestimate their TSE, which may decrease as they gain experience. While some studies have found that TSE remains stable for in-service teachers, other findings indicate that it may increase in the initial years and then decrease after many years (Klassen & Chiu, 2010; Woolfolk Hoy & Spero, 2005). However, longitudinal studies examining the development of TSE across phases of teacher education are scarce (Klassen et al., 2011; Zee & Koomen, 2016).

The research questions of this study are as follows:

1. How does TSE develop throughout the different phases of teacher education, including the teacher education program, the preparatory phase, and in-service teaching?
2. How do other variables such as migration status, year of birth, and gender influence the development of TSE during different phases of teacher education?
3. Can different developmental profiles of TSE be identified during and across various stages of teacher education using a person-centered approach?

Method

Data were drawn from the national panel education study, which oversampled teacher education students starting in the winter term of 2010/11 and continuing with semi-annual or annual surveys until 2022. TSE was measured annually from 2015 using Schwarzer and Schmitz's (1999) instrument. Missing values were imputed using the Multivariate Imputation by Chained Equations (MICE) method, including covariates such as migration status, year of birth, and gender. In total, our analytical sample comprised 2698 teachers.

The factorial structure was tested using confirmatory factor analysis and measurement invariance across measurement points and stages of teacher education. The measurement invariance test implies that the model is at least metric-invariant. Changes in TSE across career stages were examined using structural equation modeling (SEM) regression analysis. Growth curve analyses were used to test different models, including linear, quadratic, and piecewise slopes, distinguishing between the theoretical teacher education program phase, the on-the-job training phase in teacher preparation, and the teacher working phase. A Person-centered latent growth mixture model was used to identify different developmental profiles.

Preliminary Results

The latent regression model reveals that the TSE exhibits high temporal stability, with each subsequent measurement significantly predicted by the previous one, as indicated by strong standardized coefficients ranging from 0.814 to 0.936. Although the model fit was reasonable (RMSEA = .025; SRMR = .045), some indices suggested room for improvement (CFI = .767).

We tested linear, quadratic, and piecewise growth curve models. A comparison of fit indices indicates that the piecewise model, with two separate estimated intercepts and slopes for the theoretical and practical phases of teacher education, fits the data best. However, the non-significant ($p < .05$) latent slope estimates in all models suggest high stability of the TSE over time.

Adding covariates to the model and estimating a person-centered latent growth mixture model should provide insights into individual differences regarding changes in TSE during teacher education. The results of these analyses are presented at conferences. Possible practical implications for teacher education are discussed in this presentation.

Wechselseitige Zusammenhänge zwischen Selbstwirksamkeitserwartungen und Lehrverhalten: Eine video-basierte Analyse von Lehramtsstudierenden

Isabell Hußner¹, Rebecca Lazarides², Wendy Symes²

¹IPN (Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik), Deutschland; ²Universität Potsdam, Deutschland

Theoretischer Hintergrund

Selbstwirksamkeitserwartungen gelten als zentraler Prädiktor für die Unterrichtsqualität von Lehrkräften (Klassen & Tze, 2014). Nach der sozial-kognitiven Theorie von Bandura (1989) sind Selbstwirksamkeitserwartungen individuelle Überzeugungen, eigenes Handeln erfolgreich initiieren und aufrechterhalten zu können, um herausfordernde Situationen zu bewältigen. Erfolgreiches Unterrichtshandeln kann als Bedingung für die Selbstwirksamkeit betrachtet werden (Bandura, 1997). Während die wechselseitigen Beziehungen zwischen Selbstwirksamkeit und Unterrichtsqualität bei berufstätigen Lehrkräften bisher weitreichend untersucht worden sind (Holzberger et al., 2013), ist wenig über diese Zusammenhänge bei Lehramtsstudierenden bekannt. Unsere Studie geht somit über frühere Arbeiten hinaus, die Zusammenhänge zwischen der Selbstwirksamkeit von Lehrkräften und ihrem Unterrichtsverhalten untersucht haben (Holzberger et al., 2013; Praetorius et al., 2017), indem mehrere Facetten der Selbstwirksamkeit von Lehramtsstudierenden mit mehreren theoretisch ausgerichteten Dimensionen der Unterrichtsqualität abgeglichen werden, um die Richtung der Motivation von Lehrkräften zum Unterricht zu verfolgen (Bardach & Klassen, 2021; Lazarides et al., 2022).

Vor diesem Hintergrund untersucht die Studie, wie sich die Selbstwirksamkeitserwartungen von Lehramtsstudierenden während eines Semesters verändern und wie diese Veränderungen mit der videografierten Qualität ihres Unterrichts zusammenhängen. Darüber hinaus werden die Effekte der selbsteingeschätzten Feedbackqualitäten von Peers, Dozierenden und betreuenden Lehrkräften auf die Selbstwirksamkeitserwartungen sowie auf das videografierte Unterrichtsverhalten analysiert.

Fragestellungen

Die Studie untersucht die wechselseitigen Zusammenhänge zwischen den Selbstwirksamkeitserwartungen und der videografierten und mit dem CLASS-S-Instrument bewerteten Unterrichtsqualität von Lehramtsstudierenden während eines Semesters.

Methode

Es wurden Daten von $N = 110$ Lehramtsstudierenden (48,2 % weiblich; $M_{\text{Alter}} = 23,69$; $SD = 4,02$; $M_{\text{Fachsemester}} = 4,83$; $SD = 1,92$) analysiert, um die Selbstwirksamkeitserwartungen für Instruktionsstrategien, Klassenmanagement und Schüler:innen-Engagement zu Semesterbeginn und am Ende eines Semesters mit der deutschsprachigen Version der Selbstwirksamkeitsskalen von Tschannen-Moran und Hoy (2001) zu untersuchen. Die Unterrichtsqualität wurde anhand von Unterrichtsvideografien ausgewertet und mithilfe des CLASS-Scoring-Systems (Pianta et al., 2012) von 5 Rater:innen kodiert. Insgesamt wurden $n = 290$ Unterrichtssegmente von $N = 110$ Lehramtsstudierenden untersucht. Zur Analyse der wechselseitigen Zusammenhänge der Selbstwirksamkeit und dem Unterrichtsverhalten von Lehramtsstudierenden wurden konfirmatorische Faktorenanalysen (CFA) sowie Pfadmodelle in Mplus (Version 8.6) berechnet.

Ergebnisse

Die CFA zeigte drei latente Faktoren der Unterrichtsqualität: Emotionale Unterstützung, Instruktionsunterstützung und Klassenmanagement ($\chi^2 (df = 41) = 119,94$; CFI = 0.93; TLI = 0.91; RMSEA = .13; SRMR = .06). Die Selbstwirksamkeit für Instruktionsstrategien ($M_{T1} = 4,56$, $SD = 0,34$; $M_{T4} = 4,94$, $SD = 0,39$), Klassenmanagement ($M_{T1} = 4,37$, $SD = 0,48$; $M_{T4} = 4,74$, $SD = 0,54$) und Schüler:innen-Engagement ($M_{T1} = 4,47$, $SD = 0,47$; $M_{T4} = 4,83$, $SD = 0,50$) stiegen signifikant im Semesterverlauf an. Die Ergebnisse der Pfadmodelle zeigten, dass das videografierte und fremdbewertete Klassenmanagement der Lehramtsstudierenden zu Semestermitte die Selbstwirksamkeit für Klassenmanagement am Ende des Semesters vorhersagt ($\beta = .24$, $SD = .08$, $p = .005$). Zudem sagte die Selbstwirksamkeit für Schüler:innen-Engagement zu Semesterbeginn die videografierte und fremdbewertete emotionale Unterstützung zu Semestermitte positiv vorher ($\beta = .23$, $SD = .09$, $p = .011$). Die von den Studierenden eingeschätzte Feedbackqualität der betreuenden Lehrkräfte begünstigte die fremdbewertete emotionale ($\beta = .25$, $SD = .09$, $p = .008$) sowie instruktionale Unterstützung ($\beta = .33$, $SD = .09$, $p < .001$) zu Semestermitte. Im Pfadmodell Selbstwirksamkeit für Instruktionsstrategien sagte das pädagogische Wissen (Dimension Unterrichten) der

Lehramtsstudierenden zu Semesterbeginn die videografierte und fremdbewertete instruktionale Unterstützung der Lehramtsstudierenden zu Semestermitte ($\beta = .25$, $SD = .09$, $p = .007$) positiv vorher.

Relevanz

Diese Studie liefert erste Hinweise darauf, dass qualitativ hochwertiger Unterricht, insbesondere im Klassenmanagement, die Selbstwirksamkeit von Lehramtsstudierenden verbessert. Praktische Erfahrungen in authentischen Unterrichtssituationen verbunden mit Feedback insbesondere von betreuenden Lehrkräften sind daher entscheidend für die professionelle Entwicklung von angehenden Lehrkräften. Diese Ergebnisse unterstreichen die Bedeutung von Praktika und Mikro-Teaching-Erfahrungen im Lehramtsstudium zur Förderung der professionellen Kompetenz.

Selbstwirksamkeitserwartungen im Lehramtsstudium: Förderung durch Aufgaben mit Berufsfeldbezug?

Stefanie Rach

OVGU Magdeburg, Deutschland

Die Bedeutung von Selbstwirksamkeitserwartungen für erfolgreiche Lernprozesse wird in der Literatur herausgestellt (z. B. in Erwartungs-Wert-Modellen von Eccles & Wigfield, 2020 oder in den Arbeiten von Bandura, 1977), jedoch gibt es eine große Varianz in Selbstwirksamkeitserwartungen zwischen Lehramtsstudierende im Fach Mathematik (Rach et al., 2024). Selbstwirksamkeitserwartungen als Überzeugungen, eine Aufgabe erfolgreich zu bewältigen, speisen sich sowohl aus Personenmerkmalen, z. B. dem Selbstkonzept, als auch aus situationalen Merkmalen, z. B. der Aufgabenpraktik (Rach et al., 2024; Street et al., 2024). Ein Mathematiklehramtsstudium zeichnet sich in der Studieneingangsphase dadurch aus, dass Aufgaben zur akademischen Mathematik und nicht zur Schulmathematik fokussiert werden (Geisler & Rolka, 2021).

Interventionen im Lehramtsstudium verwenden häufig Aufgaben mit Berufsfeldbezug, um z. B. Wertüberzeugungen zu steigern (z. B. Rach & Schukajlow, 2024; Weber et al., 2023). Ob diese Aufgaben, die eine Verbindung zwischen akademischer und Schulmathematik schaffen, auch Selbstwirksamkeitserwartungen steigern können, ist bisher unklar. In der vorgestellten mixed-methods-Studie wird zudem untersucht, inwiefern Studierende Aufgabenmerkmale nutzen, um die Genese von Selbstwirksamkeitserwartungen zu beschreiben.

Die Forschungsfragen lauten:

1. Inwiefern steigern Aufgaben mit Berufsfeldbezug die Selbstwirksamkeitserwartungen der Studierenden?
2. Welche Faktoren nutzen Studierende, um die Genese von Selbstwirksamkeitserwartungen zu beschreiben?

Aus den Bereichen Zahlentheorie und Analysis wurden sieben Mathematikaufgaben ausgewählt, die in einem Mathematikstudium vorkommen könnten. Diese Aufgaben haben in der Regel keinen expliziten Bezug zum späteren Berufsleben („Aufgabe ohne Berufsfeldbezug“). Jede dieser Aufgaben wurde so verändert, dass der Bezug zum späteren Beruf expliziert wird („Aufgabe mit Berufsfeldbezug“) – sowohl inhaltlich durch die Verknüpfung von Schulmathematik und akademischer Mathematik als auch durch Wörter mit Berufsfeldbezug wie „Schule“, „Mathematikunterricht“ etc. (siehe Abb. 1). Die Aufgaben eines Aufgabenpaares unterscheiden sich nicht hinsichtlich der mathematischen Anforderung. An der experimentellen, quantitativen Studie haben 131 Studierende (52% weiblich) in einem Lehramtsstudiengang für weiterführende Schulen freiwillig teilgenommen. Den Studierenden wurden sieben Mathematikaufgaben vorgelegt, wobei jeweils für jeden Teilnehmenden und jede Aufgabe der Aufgabentyp (mit oder ohne Berufsfeldbezug) zufällig zugewiesen wurde. Die Studierenden beurteilten jede Aufgabe bezüglich der Selbstwirksamkeitserwartungen mit zwei Items (Beispielitem „Ich traue mir zu, diese Aufgabe zu lösen“, $.59 < r < .91$) auf einer vierstufigen Likertskala von trifft zu (4) bis trifft nicht zu (1). Zur Analyse von Fragestellung 1 wurde ein lineares Mischmodell in R (Version 4.3.0) mit dem Package „lme 4“ (Bates et al., 2015) mit den zufälligen Faktoren „Aufgabe“ und „Person“ und dem festen Faktor „Aufgabentyp“ spezifiziert. An der qualitativen Studie haben 10 Lehramtsstudierende für weiterführende Schulen freiwillig teilgenommen. Diese wurden aufgefordert, zu vier Aufgaben, jeweils zufällig mit oder ohne Berufsfeldbezug, ihre Selbstwirksamkeitserwartungen anzugeben und dabei für ihre Entscheidungsprozesse laut zu denken. Durch deduktive und induktive Kategorienbildung nach Mayring und Fenzl (2019) wurden Faktoren zu Fragestellung 2 identifiziert, die die Studierenden als relevant für die Einschätzung der Selbstwirksamkeitserwartung ansahen.

Anhand der Ergebnisse der experimentellen Studie kann nicht gezeigt werden, dass durch einen expliziten Bezug zum späteren Berufsfeld in den Aufgaben die Selbstwirksamkeitserwartungen gesteigert werden können (Estimate=0.03, $SE=0.16$, $t=0.22$, $p=.83$, ohne Berufsfeldbezug: $M=2.82$, $SD=0.77$, $N=436$; mit Berufsfeldbezug: $M=2.79$, $SD=0.83$, $N=425$). Zudem unterscheiden sich die Studierenden bezüglich der Selbstwirksamkeitserwartungen stark voneinander ($Var=0.23$) und weniger Unterschiede liegen in den Aufgaben ($Var=0.11$). Auffällig an den von den Studierenden genannten Faktoren ist, dass sich diese weniger auf individuelle Merkmale beziehen, z. B. aufs Selbstkonzept, sondern auf Aufgabenmerkmale („weil ich mit den Summenformeln nicht so gut klarkomme“). Jedoch wurde die Verbindung zum Berufsfeld in den Aufgaben nicht explizit benannt.

Im Kontrast zu Wertüberzeugungen (Rach & Schukajlow, 2024) gibt es keine Hinweise, dass die Veränderung eines Aufgabenmerkmals, der Berufsfeldbezug, die Selbstwirksamkeitserwartungen steigern. Die Studierendenäußerungen deuten jedoch darauf hin, dass Studierende sich auf konkrete Aufgabenmerkmale bei der Einschätzung beziehen. In weiterführenden Studien sollte verstärkt der Prozess, wie Lehramtsstudierende ihre Selbstwirksamkeitserwartungen generieren, analysiert werden.

PS-24: Digitale Personal- und Organisationsentwicklung

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 129

Chair der Sitzung: Nina Carolin Werth

Organisationsentwicklung in Schulen: Gesundheitsförderung effektiv gestalten

Lars König^{1,2}

¹Stiftung Gesundheitswissen, Berlin, Deutschland; ²Charité – Universitätsmedizin Berlin, Institut für Medizinische Soziologie und Rehabilitationswissenschaft, Berlin, Deutschland

Theoretischer Hintergrund

Um die Gesundheitskompetenz von Lehrkräften sowie von Schülerinnen und Schülern langfristig zu fördern, müssen Schulen aktiv in diesen Prozess miteingebunden werden. Hierbei besteht eine zentrale Herausforderung darin, dass viele Schulen in Deutschland keine ausreichende organisationale Gesundheitskompetenz aufweisen. Zudem fehlen Schulen oftmals die zeitlichen, personellen sowie monetären Ressourcen, um zielgruppenspezifische Angebote zur Förderung der Gesundheitskompetenz von Lehrkräften sowie von Schülerinnen und Schülern zu entwickeln und dauerhaft in aktueller Form bereitzustellen.

Fragestellung

Die vorliegende Studie geht der Frage nach, inwieweit die Bereitstellung von evidenzbasierten und frei verfügbaren YouTube Health Videos zur Förderung der Gesundheitskompetenz von Lehrkräften sowie von Schülerinnen und Schülern eine Möglichkeit darstellt, um die organisationale Gesundheitskompetenz von Schulen zu stärken.

Methoden

Im Vortrag werden die Ergebnisse einer Evaluationsstudie vorgestellt, welche die Attraktivität sowie Wirksamkeit von evidenzbasierten YouTube Health Videos untersucht. In Zusammenarbeit mit einem Markt- und Meinungsforschungsinstitut wurden hierzu Lehrkräfte, Schülerinnen und Schüler sowie Personen der Allgemeinbevölkerung rekrutiert. Je nach Gruppenzugehörigkeit wurden den Teilnehmenden innerhalb eines Experimentaldesigns unterschiedliche YouTube Health Videos präsentiert. Erfasst wurden unter anderem die Gesundheitskompetenz, der Wissenszuwachs sowie die Attraktivität der eingesetzten YouTube Health Videos. Verschiedene bayesianische sowie klassische inferenzstatistische Methoden wurden zur Datenauswertung eingesetzt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass die eingesetzten YouTube Health Videos zu einem subjektiven Kompetenzzuwachs sowie zu einem Wissenszuwachs geführt haben. Zudem wurden die YouTube Health Videos sowohl von den Lehrkräften als auch von den Schülerinnen und Schülern als attraktiv wahrgenommen. Darüber hinaus konnten verschiedene schulische Einsatzkontexte für die YouTube Health Videos identifiziert werden. Eine detaillierte Ergebnisdarstellung erfolgt im Vortrag. Übergreifend lässt sich festhalten, dass YouTube Health Videos zur organisationalen Gesundheitskompetenzförderung in Schulen eingesetzt werden können. Hierzu sollte allerdings der Bekanntheitsgrad dieser frei verfügbaren und evidenzbasierten Angebote gesteigert werden. Limitationen des gewählten Studiendesigns sowie zukünftige Forschungsfragen werden im Vortrag diskutiert.

Die Rolle von Promotor*innen beim digitalisierungsbezogenen Organisationalen Lernen von Schulen

Nina Carolin Werth, Amelie Sprenger, Prof. Dr. Kathrin Fussangel, Prof. Dr. Cornelia Gräsel

Bergische Universität Wuppertal, Deutschland

Die digitalisierungsbezogene Transformation stellt Schulen vor erhebliche Herausforderungen und erfordert die Kooperation aller beteiligten Akteur*innen (Altrichter 2015; Endberg et al. 2022; Labusch et al. 2020; Eickelmann und Gerick 2018; Gerick et al. 2024). Schule muss als „lernende Organisation“ (Senge 1990) im 21. Jahrhundert mehr denn je die Fähigkeit besitzen, flexibel auf Transformationen zu reagieren und aufgrund der ständig wechselnden Anforderungen agile Strukturen aufweisen, um der Vielzahl und Vielfalt der Entwicklungen gerecht zu werden (Stricker 2022; Holtappels 2007; Feldhoff 2011). Dabei spielt das **Organisationale Lernen** eine zentrale Rolle: Eine lernende Organisation passt sich agil an Veränderungsprozesse an, erkennt Probleme und verbessert sich dadurch kontinuierlich (Argyris 1974; Marks und Louis 1999; Seashore Louis und Lee 2016; Argyris und Schön 1978-1996). Organisationales Lernen findet statt, wenn Individuen in der Auseinandersetzung mit Problemen eigene Handlungsmuster und Aktivitäten verändern (Argyris und Schön 1996). Laut Feldhoff (Feldhoff 2017) sind für eine lernende Organisation vor allem die Förderung von Kooperation, die Etablierung gemeinsamer Ziel- und Wertvorstellungen sowie die systematische Nutzung individueller und externer Wissensquellen relevant. Zur Klärung der Frage, welche Personen innerhalb eines Kollegiums diese Tätigkeiten aufgreifen, kann das **Promotor*innenmodell** herangezogen werden. Das Promotor*innenmodell stellt die These auf, dass Innovationsprozesse nicht eigenständig voranschreiten, sondern durch Schlüsselpersonen, den Promotor*innen, aktiv gefördert werden (Prasse 2012; Gerick et al. 2024; Wagner und Gerholz 2022). Dabei nehmen die Promotor*innen unterschiedliche Aktivitäten wahr und so können laut Modell verschiedene Rollen unterschieden werden: Macht-, Fach-, Prozess- sowie Beziehungspromotor*innen (Schultz et al. 2014; Witte 1973; Hauschildt und Chakrabarti 1988; Gemünden und Walter 1996; Wagner und Gerholz 2022; Gerick et al. 2024).

Ziel der vorliegenden Studie ist es, Promotor*innen im Rahmen digitalisierungsbezogener Schulentwicklung zu identifizieren und zu untersuchen, welche Rolle die unterschiedlichen Promotor*innen beim digitalisierungsbezogenen Organisationalen Lernen spielen. Die Kollegien von insgesamt 24 weiterführenden Schulen aus Nordrhein-Westfalen, Bayern und Hamburg erhalten eine Online-Fragebogen, der die Interaktionshäufigkeit von Kooperation sowie die Netzwerkmerkmale und Positionen von Promotor*innen **digitalisierungsbezogener Schulentwicklung** im Schulnetzwerk fokussiert (Lin und Lee 2018; Louis et al. 2010; Louis K. S. 2010). Das **soziale Netzwerk** jeder Schule wird auf die Fragestellung hin analysiert und visuell abgebildet (Lin und Lee 2018; Daly und Finnigan 2010; Moolenaar und Daly 2012; Wullschlegel et al. 2023). So wird beispielsweise untersucht, inwiefern bestimmte Schlüsselpositionen der Promotor*innen im Netzwerk Zugang zu geeigneten Ressourcen (z.B. Vermittlung von digitalisierungsbezogenen Wissen an andere Lehrkräfte) ermöglichen (Coburn 2003; Penuel et al. 2009; Moolenaar und Daly 2012; Moolenaar et al. 2014). Netzwerkmerkmale der Promotor*innen, die unter anderem vergleichend zum Gesamtnetzwerk analysiert werden, sind z.B. die Dichte, die Zentralität (Closeness und Betweenness) und der Grad (Jansen 2006; Carrington et al. 2005; Wullschlegel et al. 2023). Für die Kapazität des Organisationalen Lernens einer Schule sind validierte Skalen (Seashore Louis und Lee 2016; Lin und Lee 2018) verwendet worden. Darüber hinaus werden leitfadengestützte Online-Gruppeninterviews mit je drei Personen pro Schule, u.a. mit den im Fragebogen identifizierten Promotor*innen, geführt (Misoch 2019; Niederberger und Zwick 2023) und auf Basis der Theorien zum Organisationalen Lernen und den Promotor*innen analysiert.

Die Datenerhebung läuft derzeit, erste Analysen sind bereits durchgeführt. Bislang haben 506 Lehrkräfte von acht Gymnasien und Gesamtschulen in NRW an der Online-Umfrage teilgenommen – neun Interviews sind an diesen Schulen bereits durchgeführt worden. Die Kollegien der anderen Bundesländer, Bayern und Hamburg, starten nach den Sommerferien 2024 mit den Umfragen und Interviews. Insgesamt liegen bis zum Winter 2024 über 70 Interviews für die Studie vor. Auf der Tagung werden die schulischen Netzwerke einzelner Schulen mit einem besonderen Fokus auf die Rolle von Promotor*innen vorgestellt. Vor diesem Hintergrund werden Ergebnisse zum Organisationalen Lernen sowohl aus der Fragebogenerhebung als auch korrespondierende Ergebnisse aus den Interviews präsentiert.

Digitale häusliche Lernumwelt im Vorschulalter: Profile gemeinsamer digitaler und analoger Aktivitäten von Eltern mit Kindern

Anja Linberg¹, Inga Simm¹, Hanna Maly-Motta¹, Thorsten Naab¹, Susanne Kuger^{1,2}

¹Deutsches Jugendinstitut, Deutschland; ²Ludwig-Maximilians-Universität München

Hintergrund

Digitale Medien sind in der frühen Kindheit vieler Familien fester Bestandteil des Alltags und der häuslichen Lernumgebung (HLE) geworden (Kieninger et al., 2024; Lehl et al., 2021). Während gemeinsame Aktivitäten der HLE traditionell analoge Tätigkeiten wie Vorlesen, Musizieren oder Bibliotheksbesuche umfassten, sind digitale Medien mittlerweile allgegenwärtig. Diese digitalen Aktivitäten schließen die gemeinsame Nutzung von audiovisuellen Medien (TV, Sendungen auf dem Tablet), Audiomedien (Lesestifte, Hörbücher, Musik) und digitale Spiele ein. Oftmals wird in diesem Zusammenhang die Sorge geäußert, dass analoge Aktivitäten verdrängt werden oder die Nutzung digitaler Medien Gefahren birgt (Kieninger et al., 2024; Istenič et al., 2023) wenngleich Studienergebnisse dieses einseitig negative Bild nicht stützen (u.a. Hirsh-Pasek et al. 2015; Soyoof, 2023).

Klassische HLE-Modelle betrachten gemeinsame Aktivitäten von Eltern und Kindern als zentrale Mechanismen, die die kognitive und sozio-emotionale Entwicklung der Kinder fördern und von den strukturellen Bedingungen der Familie abhängen (Niklas, 2015; Klucznik et al., 2013). Differenzierungen der HLE basieren dabei meist auf der Häufigkeit bildungsförderlicher Aktivitäten (z.B. im Längsschnitt niedrige vs. hohe Typen; Hayes & Berthelsen, 2019) oder von Typen in bestimmten Domänen (z.B. Typen von home literacy behaviors: Phillips & Lonigan, 2009). Bisher berücksichtigen Typisierungen jedoch keine digitalen Aktivitäten, wodurch das Verhältnis von analogen und digitalen Aktivitäten in der HLE offenbleibt. Dieser Beitrag untersucht daher, (1) welche Typen analoger und digitaler HLE sich in Familien mit Kindern im Alter von 0 Jahren bis zum Schuleintritt identifizieren lassen und (2) von welchen Strukturmerkmalen der Familie diese Typen abhängen.

Methode

Zur Beantwortung der Fragestellung werden Daten der dritten Welle der Studie Aufwachsen in Deutschland: Alltagswelten (AID:A 2023; Kuger et al., 2024) genutzt, einer repräsentativen Stichprobe, die rund 8.750 Personen im Alter von 0 bis 32 Jahren längsschnittlich verfolgt.

Für die Erfassung der HLE wird die Häufigkeit verschiedener gemeinsamer Aktivitäten von Eltern und Kindern von beiden Elternteilen auf einer 6-stufigen Skala von nie bis mehrmals täglich eingeschätzt. Diese Aktivitäten umfassen klassische analoge Aktivitäten (z. B. Vorlesen, Malen) sowie digitale Aktivitäten (z. B. Nutzung audiovisueller Medien, digitale Spiele). Der sozioökonomische Hintergrund wird unter anderem durch den Bildungshintergrund der Eltern (CASMIN; Müller et al., 1989) und den Deprivationsindex des Haushalts (Christoph, 2016) erfasst. Die Analysen berücksichtigen Angaben zur Häufigkeit gemeinsamer Aktivitäten der Hauptauskunftsperson (63 % weiblich) zu 2.894 Kindern im Alter von 0 bis zum Schuleintritt. Zur Klassifizierung unterschiedlicher HLE-Aktivitäten wurde eine Latente Profilanalyse mittels Stata durchgeführt, wobei Kennwerte des AIC und BIC zur Bestimmung der Profilanzahl genutzt wurden.

Ergebnisse

Die Analysen identifizieren verschiedene Profile von HLE-Aktivitäten in Familien mit Kindern zwischen 0 Jahren und Schuleintritt. Vorläufige Ergebnisse deuten auf vier Hauptgruppen hin: 17 % der Kinder erleben eine geringe HLE mit wenigen gemeinsamen Aktivitäten jenseits von Ausflügen, 62 % eine vielfältig anregungsreiche HLE mit vorrangig analogen, aber auch digitalen Aktivitäten wie der Nutzung audiovisueller Medien, 15 % eine vorrangig audiovisuelle HLE und 7 % eine medienaffine HLE mit digitalen Spielen, gemeinsamer Internetnutzung, audiovisuellen Medien und analogen Aktivitäten. Ein multinomiales Logit-Modell verknüpft die Profile mit wesentlichen sozioökonomischen Merkmalen.

Diskussion

Trotz verschiedener Limitationen, wie der Altersrange der Stichprobe, zeigen die Ergebnisse, dass digitale Aktivitäten bereits in der frühen Kindheit in der HLE eine Rolle spielen, jedoch bei der großen Mehrheit der Familien analoge Aktivitäten ergänzen, anstatt sie zu verdrängen. Studienergebnisse deuten darauf hin, dass die gemeinsame Nutzung digitaler Medien die familiäre Anregung erhöht, die Medienkompetenz fördert und positiv mit anderen Entwicklungsbereichen der Kinder assoziiert ist (Soyoof, 2023; Lehl et al., 2021). Der Beitrag diskutiert, inwieweit andere Profile, insbesondere die vorrangig audiovisuelle HLE, ein anderes Bild zeichnen und gibt Ausblick auf weitere Analysen, die die Fähigkeiten der Kinder einbeziehen.

ENTFÄLLT Dynamics in adults' digital skills over the life course: On the impact of further training and skill use on ICT skills

Martin Ehler^{1,2}

¹Wissenschaftszentrum Berlin für Sozialforschung, Deutschland; ²FU Berlin

Against the background of technological innovations, digital skills are becoming increasingly important for workers' productivity and performance in labor markets and their overall inclusion in society (Falck et al., 2021; OECD, 2013). Yet, many adults did not learn the skills to master information and communication technology (ICT) sufficiently during initial education. To keep up with changes, adults need to learn and update these skills later in life. The aim of this study is to assess the dynamics of ICT skills over the life course.

Previous research based on cross-sectional data found that age is significantly associated with ICT skills with lower skills among older adults (Desjardins & Ederer, 2015). Furthermore, learning at work and ICT use are associated with higher ICT skills (Hämäläinen et al., 2019; Wicht et al., 2021). Recent research on literacy and numeracy using longitudinal data showed that these skills change over time within adults. These changes seem to be mainly influenced by skill use and not by further training (Gauly and Lechner 2019; Gauly et al 2020; Reder et al, 2020). However, compared to literacy and numeracy it seems to be more likely that ICT skills are developed through further education later in life.

This study focusses on two research questions:

1. How do adults' ICT skills develop over the life course?
2. What are the effects of ICT skill use and further education on ICT skill acquisition during adulthood?

I examine two explanations for ICT skill formation during adulthood. First, participation in further education and training on ICT topics may increase ICT skills. This is based on human capital theory which posits that profit-maximizing individuals invest in education and training to gain marketable skills and then use them to increase their incomes (Becker, 1975). Second, practice engagement theory by Reder (1994) on the other hand claims that ICT skill growth is not caused by participating in ICT training but due to ICT practices. The basic idea is that engagement with activities that make use of skills in different domains leads to higher skills. Higher skills in turn lead to more engagement with associated practices. The opposite of the coin is the "use-it-or-lose-it" hypothesis which states that workers lose skills if they do not apply them at work (de Grip et al., 2008).

I use data from the German National Educational Panel Study (NEPS), Starting Cohort (SC) 6 (adults) (Blossfeld & Roßbach, 2019; NEPS-Netzwerk, 2023). The NEPS SC6 is a representative panel study of adults born between 1944 and 1986 in Germany. ICT skills ("computer literacy") were measured in wave 5 and 14 using competence tests and a longitudinally comparable (linked) competence score was provided (Gnams & Senkbeil, 2023). In addition to ICT skills, the data provides information about participation in a wide range of adult learning activities and the topics covered. This allows for the identification of learning activities focused on ICT topics. Also, ICT use on the job as well as a wide range of socio-demographic variables are surveyed.

Preliminary analyses reveal that ICT skills decreased on average among individuals over the nine years covered by the data. This decline in skills increases with age, with particularly high losses among people aged 65 and above. Among those who took at least one ICT-related non-formal course between the measurement points, the decline is less pronounced. This finding is corroborated by multivariate models indicating a sizeable positive association between ICT training and ICT skills, even after controlling for initial ICT skills and a set of socio-demographic controls. ICT use at work also proved to be significantly correlated with increases in ICT skills.

PS-25: Digitales Lesen und Schreiben: Herausforderungen und Potenziale für die schulische Bildung

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 162
Chair der Sitzung: Johannes Naumann

Vorerfahrungen zum digitalen Schreiben in der Primarstufe

Kathrin Heller

Pädagogische Hochschule Karlsruhe, Deutschland

Die digitale Produktion von Texten gilt als Schlüsselkompetenz in der heutigen Informations- und Kommunikationsgesellschaft (Frederking & Krommer, 2019). Der Vertrautheit mit dem Schreibmedium kommt dabei eine wichtige Bedeutung zu, da diese die Textqualität systematisch beeinflusst (Philipp, 2015, S. 17). Dennoch fehlt es bislang an umfassenden Studien zur Entwicklung einer digitalen Schreibkompetenz, insbesondere im deutschsprachigen Raum und für die Primarstufe. An diesem Desiderat anknüpfend wird im Forschungsprojekt untersucht, wie sich das digitale Schreiben wirksam fördern lässt und welche Effekte dies auf die Textproduktion von Lernenden der dritten Klasse hat.

Die Entwicklung wirksamer Interventionen setzt die Anknüpfung an die Eingangsvoraussetzungen der Lernenden voraus. Abgesehen von der standardisierten Copy Task (van Waes et al., 2021), die für versierte Schreibende konzipiert ist, fehlt es jedoch bislang an Diagnoseinstrumenten. Weiterhin beziehen sich die wenigen umfangreicheren diagnostischen Untersuchungen auf Erwachsene oder ältere Lernende (z.B. Dhakal et al., 2018; Grabowski, 2008). Studien mit Lernenden aus der Primarstufe fokussieren die Tastaturkompetenz im engen Sinne (z.B. Tippgeschwindigkeit, -korrektheit) (z.B. van Weerdenburg et al., 2019). Erste Untersuchungen zeigen aber, dass neben dem Tastaturschreiben (ebd.) weitere Fähigkeiten und Fertigkeiten (z.B. Navigation, Generierung von Großbuchstaben) elementare Voraussetzungen für die digitale Textproduktion sind (Anskeit, 2022). Den Desideraten folgend wurde auf Grundlage empirischer Vorarbeiten (v.a. Grabowski, 2008, 2009; van Waes et al., 2021) ein diagnostisches Verfahren entwickelt und in einer Laborsituation mit Drittklässler*innen (n=17) durchgeführt.

Die Diagnostik der Tastaturkompetenz im weiteren Sinne (Grabowski, 2009) erfolgt anhand der Bearbeitung von fünf Aufgaben unterschiedlicher kognitiver und technischer Anforderungen (digitale Textproduktion, Abschreiben aus dem Gedächtnis, Aufgaben zu digitalen Textrevisionen, „dk“-Aufgabe, Abschreibetext). Die Eingaben und zentrale Kennwerte der Tastatur- und Mausnutzung wurden mittels des Keystroke-Logging-Programms *Inputlog* (Leijten & van Waes, 2013) erfasst. Darüber hinaus wurden unterschiedliche Kovariate erhoben (Fragebogen, analoge Textproduktion, Beobachtungsbogen etc.). Die Daten werden sowohl quantitativ als auch qualitativ ausgewertet, um die Fähigkeiten und Fertigkeiten der digitalen Textproduktion von Lernenden der Primarstufe zu beschreiben, die Varianz zu ermitteln und zentrale Herausforderungen zu identifizieren. Weiterhin werden Kennwerte der Schreibflüssigkeit zwischen den verschiedenen Aufgaben verglichen und darauf basierend das diagnostische Verfahren weiterentwickelt und kondensiert.

Erste Ergebnisse zeigen, dass insbesondere die Nutzung von Funktionstasten (z.B. Umschalttaste) sowie die technische Umsetzung digitaler Textproduktionsstrategien Herausforderungen darstellen. Weiterhin ist anzunehmen, dass sich zentrale Kennwerte der Schreibflüssigkeit je nach Aufgabe unterscheiden (Grabowski, 2008), erhebliche Unterschiede zwischen den Lernenden bestehen und sich bestimmte Kennwerte gegenseitig beeinflussen, was mittels Varianz-, Korrelationsanalysen überprüft wird.

Die Ergebnisse der Vorstudie erlauben eine erste umfassende Diagnostik der schreibbezogenen digitalen Kompetenzen von Lernenden der Primarstufe. Dies ist grundlegend für die die Förderung einer digitalen Schreibkompetenz, wie bildungspolitisch erwünscht (KMK, 2017, 2022). Vor diesem Hintergrund werden im Beitrag das konzipierte Verfahren sowie die Ergebnisse der Vorstudie vorgestellt, Konsequenzen für die Entwicklung von Maßnahmen zur Förderung digitaler Textproduktionskompetenzen diskutiert und erste Einblicke in die entwickelten Fördermaterialien präsentiert. Die Ergebnisse liefern vielfältige Erkenntnisse und Anknüpfungspunkte sowohl für das Fach Deutsch als auch für andere Fächer, in denen mit Tastaturen geschrieben wird.

Sagt die Nutzung digitaler Texte die Entwicklung von Lesekompetenz vorher?

Johannes Naumann¹, Lidia Altamura², Cristina Vargas², Amelia Mana², Ladislao Salmerón²

¹Bergische Universität Wuppertal, Deutschland; ²Universitat de Valencia, Spanien

Theoretischer Hintergrund

Querschnittlich finden sich häufig negative Zusammenhänge zwischen der Nutzung digitaler Geräte und der Frequenz von Leseaktivität am Bildschirm einerseits und Lesekompetenz andererseits (Jacobsen & Forste, 2011; Naumann & Sälzer, 2017). Für die Frequenz der Lektüre gedruckter Texte gilt i.d.R. das Umgekehrte (OECD, 2023). Dabei ist unklar, wie diese Korrelationen entstehen, ob es tatsächlich negative ursächliche Effekte der Nutzung von Digitalen Geräten oder Texten auf die Kompetenz bzw. Kompetenzentwicklung gibt – wie in der öffentlichen Debatte häufig unterstellt wird (Naughton, 2024). Ähnliches gilt für Lesemotivation. Insbesondere eine hohe intrinsische Lesemotivation kann als Bildungsziel in eigenem Recht gesehen werden, und ist darüber hinaus möglicherweise ein Einflussfaktor für die Entwicklung von Lesekompetenz (Retelsdorf et al., 2011). Vor diesem Hintergrund geht die vorliegende Untersuchung der Frage nach, ob sich jenseits querschnittlicher Zusammenhänge Effekte der Nutzung digitaler Texte auf die Entwicklung von Lesekompetenz und Lesemotivation finden lassen.

Methode

462 Grundschulkindern aus der Region Valencia (Spanien) nahmen zu 3 Zeitpunkten (4., 5., und 6. Klasse) an einer Längsschnittuntersuchung teil, bei der zu jedem MZP Lesekompetenz, intrinsische und extrinsische Lesemotivation sowie die Häufigkeit der Nutzung gedruckter und digitaler Texte erhoben wurde. Für intrinsische und extrinsische Lesemotivation sowie für Lesekompetenz wurden jeweils 2 latente Wachstumskurvenmodelle geschätzt. Im ersten Modell wurden jeweils nur die Entwicklungsverläufe geschätzt. Im zweiten Modell wurden die Frequenzen der Nutzung digitaler und gedruckter Texte in Klasse 4 als Prädiktoren für Intercept und Slope hinzugefügt. In allen Modellen mit Prädiktoren wurde das Geschlecht kontrolliert, in den Modellen mit Motivation als AV außerdem Lesekompetenz.

Ergebnisse

Die Modelle ohne Prädiktoren für intrinsische ($X^2(1)=9.72$, SRMR=0.02, CFI=0.98) und extrinsische ($X^2(1)=0.20$, SRMR=0.00, CFI=1.00) Lesemotivation zeigten einen Rückgang beider Motivationsdimensionen zwischen Klasse 4 und 6 ($\alpha_{\text{slope-intrinsisch}}=-0.17$, $p<.001$; $\alpha_{\text{slope-extrinsisch}}=-0.14$, $p<.001$), während sich für Lesekompetenz ($X^2(1)=6.44$, SRMR=0.02, CFI=0.99) ein deutlicher Zuwachs fand ($\alpha_{\text{slope-Kompetenz}}=0.67$, $p<.001$). Die slopes von sowohl intrinsischer als auch extrinsischer Lesekompetenz variierten

über Personen ($\sigma^2_{\text{Slope-intrinsisch}}=0.19$, $p<.001$; $\sigma^2_{\text{Slope-extrinsisch}}=0.16$, $p<.001$). Für die Lesekompetenz konnte die Slope-Varianz nicht geschätzt werden (negative Varianzkomponente). Ein Modell mit zu Null restringierter Slope-Varianz zeigte einen akzeptablen Fit ($X^2(3)=13.70$, SRMR=0.05, CFI=0.98).

Die Modelle mit Prädiktoren zeigten einen negativen Effekt der Nutzung digitaler Texte auf den Slope intrinsischer Lesemotivation ($\beta=-0.11$, $p<.01$), was bedeutet, dass die intrinsische Lesemotivation für Schüler*innen mit in Klasse 4 überdurchschnittlicher Nutzung digitaler Texte verstärkt abnahm. Für extrinsische Lesemotivation ergab sich kein Zusammenhang zwischen Nutzung digitaler Texte in Klasse 4 und der Entwicklung zwischen den Klassen 4 und 6. Aufgrund der geringen Varianz in den Slopes für Lesekompetenz wurde kein Modell mit Prädiktoren für den Slope geschätzt. Eine getrennte Analyse für Schüler*innen mit hoher (4.Quartil) und niedriger (1.Quartil) Nutzung digitaler Texte in Klasse 4 wies vergleichbare slopes aus ($\alpha_{\text{Slope-1Quartil}}=0.63$, $p<.001$; $\alpha_{\text{Slope-4Quartil}}=0.67$, $p<.001$). Das Lesekompetenz-Intercept war positiv mit der Nutzung gedruckter Texte ($\beta=0.29$, $p<.001$) und negativ mit der Nutzung digitaler Texte ($\beta=-0.18$, $p<.001$) assoziiert.

Diskussion

Die generelle Abnahme von Lesemotivation zwischen den Klassen 4 und 6 ist konsistent mit in der Literatur berichteten Befunden (z.B. Unrau & Schlackmann, 2006). Die beschleunigte Abnahme intrinsischer (aber nicht extrinsischer) Motivation für Schüler*innen, die überdurchschnittlich digitale Texte ist dabei ebenfalls konsistent mit Befunden, die querschnittlich negative Korrelationen zwischen intrinsischer Lesemotivation und der Nutzung digitaler Texte zeigen (McGeown et al., 2016). Im Gegensatz hierzu konnten wir keine Effekte der Nutzung digitaler Texte auf die Entwicklung von Lesekompetenz finden. Im Gegenteil entwickelte sich die Lesekompetenz von Schüler*innen, die in starkem Maße von digitalen Texten Gebrauch machen, praktisch vollständig parallel zur Entwicklung von Lesekompetenz bei Schüler*innen die unterdurchschnittlich häufig digitale Texte nutzen. Dieser Befund legt nahe, dass negativen Querschnittskorrelationen zwischen Lesekompetenz und digitalem Textgebrauch andere Mechanismen zugrunde liegen als ein negativer kausaler Effekt der Nutzung digitaler Texte auf die Entwicklung von Lesekompetenz.

Digitale Leseförderung – Einfach umzusetzen, aber auch wirksam?

Samuel Bellinghausen, Mareike Ehlert, Elmar Souvignier

Universität Münster, Deutschland

Theorie

Ein beträchtlicher Anteil an Schüler*innen in Deutschland verfügt nur über geringe Lesekompetenzen (Lorenz et al., 2023). Weiter zeigt sich, dass in deutschen Grundschulen im internationalen Vergleich wenig Unterrichtszeit für Leseaktivitäten aufgewendet wird (Schilcher et al., 2023), diese Leseaktivitäten wiederum kaum evidenzbasierte Fördermethoden beinhalten (Peters et al., 2022) und die Heterogenität in den Leseleistungen von Grundschulkindern in den vergangenen Jahren zugenommen hat (Frey et al., 2023). Nicht zuletzt angesichts dieser gestiegenen Heterogenität bedarf es differenzierter Leseförderansätze, mit denen eine Passung zwischen den Lernvoraussetzungen der Kinder und dem Lernangebot hergestellt wird (Connor, 2019).

Digitale Leseförderprogramme können aufgrund leichter Zugänglichkeit die Implementation nachweislich wirksamer, differenzierter Konzepte für Lehrkräfte erleichtern (Gold, 2023). Während sich in bisherigen Metaanalysen Befunde einer grundsätzlichen Wirksamkeit digitaler Leseförderansätze für kognitive und motivationale Konstrukte zeigen, verzeichnen andere Studien Nachteile digitalen Lesens gegenüber papierbasiertem Lesen im Hinblick auf Textverständnis und Metakognition (Delgado et al., 2018; Vanbecelaere et al., 2023). In der PISA-Studie 2018 zeigte sich zudem ein negativer Zusammenhang zwischen der schulischen Nutzung digitaler Geräte und den Leseleistungen (Sälzer, 2021). Eine Erklärung für diese Befunde bietet die *shallowing hypothesis* (z.B. Salmerón et al., 2023), wonach bei der Nutzung digitaler Geräte im Alltag eine oberflächliche Verarbeitung von Textinformationen erfolgt, die sich mitunter auf den Umgang mit digitalen Medien im schulischen Kontext überträgt (Delgado & Salmerón, 2021).

Fragestellung

Ziel der vorliegenden Untersuchung ist es, die Wirksamkeit einer digitalen Umsetzung des evidenzbasierten Leseförderprogramms „Der Lese-Sportler“ (Hebbecker et al., 2020), auf die Leseleistung (Leseflüssigkeit, Wort-, Satz- und Textverständnis) und die Lesemotivation im Vergleich zur papierbasierten Version dieses Programms sowie zu regulärem Leseunterricht zu untersuchen. Erwartet werden Vorteile der beiden Lese-Sportler-Bedingungen gegenüber der Kontrollgruppe sowie angesichts bisheriger vergleichender Studien Vorteile des analogen Lese-Sportler-Programms gegenüber der digitalen Umsetzung.

Methode

Mit dem „Lese-Sportler“-Programm arbeiten Kinder differenziert nach Förderzielen mit evidenzbasierten Methoden an den drei Leseteilkompetenzen der Lesegenauigkeit („Lese-Slalom“: silbenbasiertes lautes Lesen), der Lesegeschwindigkeit („Lese-Sprint“: wiederholendes lautes Lesen) und des Leseverständnisses („Lese-Kanu“: Anwenden von Lesestrategien durch Reziprokes Lehren). Die Methoden sind als kooperative Lernformen realisiert und ermöglichen so eine hohe Intensität der Leseförderung.

Insgesamt nahmen 31 zweite und dritte Klassen aus Nordrhein-Westfalen in der App-, Heft- bzw. Kontrollbedingung teil. Im April 2024 wurde ein Prätest mit $N=632$ Grundschüler*innen zur Erfassung der Leseleistungen (SLS 2-9; Wimmer & Mayringer, 2014; ELFE II; Lenhard et al., 2022) und der Lesemotivation (McElvany et al., 2008) durchgeführt. Anschließend erhielten alle teilnehmenden Lehrkräfte eine Rückmeldung zu den Testergebnissen auf Klassen- und Individualebene sowie eine Empfehlung über die jeweils zur Förderung anstehende Leseteilkompetenz (Lesegenauigkeit; Lesegeschwindigkeit; Leseverständnis) auf Grundlage des Prätests. Die Lehrkräfte der Lese-Sportler-Bedingungen nahmen darüber hinaus an einer Kurzfortbildung zum Lese-Sportler-Programm teil und erhielten die Förderempfehlung für jedes Kind konkretisiert in Form der jeweils passenden Lese-Sportler-Methode (Slalom; Sprint; Kanu).

Die Lehrkräfte der Experimentalgruppen setzten den Lese-Sportler über einen Zeitraum von ca. 9 Wochen mit einer empfohlenen Dosierung von 3x20 Minuten entsprechend der in NRW verbindlich eingeführten Lesezeit ein. In allen Klassen der Experimentalbedingungen wurden während des Interventionszeitraums Unterrichtsbeobachtungen zur Erfassung der Wiedergabetreue durchgeführt. Im Juni 2024 wurden die Leseleistungen und die Lesemotivation in einem Posttest erneut erhoben, dessen Daten aktuell eingegeben und aufbereitet werden.

Ergebnisse

Im Prätest zeigten sich Gruppenunterschiede, weswegen ein Propensity Score Matching durchgeführt wurde, um konfundierende Effekte durch Klassenzugehörigkeit und soziodemographische Merkmale (Geschlecht, Familiensprache) zu reduzieren. Zur Beantwortung der Fragestellung nach der Wirksamkeit werden jeweils für die Leseleistungsvariablen und die Lesemotivation Mehrebenen-Wachstumsmodelle gerechnet. Auf der Tagung werden die Ergebnisse zur Entwicklung der Lesekompetenzen in den verschiedenen Bedingungen präsentiert. Die Ergebnisse werden Hinweise auf die Wirksamkeit digitaler Leseförderansätze und möglicherweise zu treffende Abwägungen zwischen praktischer Umsetzbarkeit und potentiell geringerer Wirksamkeit digitaler Leseförderung geben.

Literarisches Textverständnis in Print- und digitalem Modus: Eine interdisziplinäre Studie zum Einfluss des Modus auf die strategische Verarbeitung und das Textverständnis

Anke Schmitz¹, Nora Kernen¹, Christina Holzwarth¹, Wiebke Dannecker²

¹Pädagogische Hochschule FHNW, Schweiz; ²Universität zu Köln

Theoretischer Hintergrund

Die fortschreitende Digitalisierung ist eine der zentralen Herausforderungen unserer Gesellschaft. Im Deutschunterricht werden entsprechend vermehrt digitale Texte eingesetzt, darunter literarische Texte, die zur Persönlichkeitsbildung beitragen und motivationsförderlich sein können (Pfost & Heyne, 2023). Klare Handlungsanweisungen für den Einsatz im Unterricht fehlen aber, da die Anzahl an Studien mit literarischen Texten gering ist ($N < 10$).

Meta-Analysen zeigen, dass papierbasierte längere (Sach-)Texte vorteilhafter sind und der digitale Lesemodus ein oberflächliches Lesen hervorruft (Clinton, 2019; Delgado et al., 2018). Ein Moduseffekt auf das Verständnis literarischer Texte konnte bei Primarstufenschüler:innen und Studierenden nicht nachgewiesen werden (Delgado et al., 2018). Offen ist bis dato, ob die Befunde auf ältere Schüler:innen und anspruchsvollere literarische Texte übertragen werden können. Zudem wurden strategische Verarbeitungsprozesse kaum beleuchtet (Delgado-Osorio et al., 2023; Goodwin et al., 2020). Goodwin et al. (2020) zeigen zwar, dass in gedruckten Sachtexten mehr markiert wird. Oberflächenbezogene Markierungen führen aber nicht zu besserem Textverständnis. Prozessnahe elaborierende Textspuren, mitunter Markierspezifika und Annotationen (Bråten & Samuelstuen, 2007), könnten näheren Aufschluss geben (Goodwin et al., 2020; Schmitz & Dannecker, 2023).

Vorgestellt wird eine interdisziplinäre Studie aus Leseforschung und Fachdidaktik. Kontrastiert wird am Beispiel von anspruchsvollen Kurzprosatexten, wie sie im Print- und digitalen Modus *verarbeitet* und *verstanden* werden und was erfolgreiche strategische Leser:innen kennzeichnet (Rogiers et al., 2020).

Fragestellungen

(FF1) Wie wirkt der Modus auf das Textverständnis und die prozessnahen Textspuren?

(FF2) Liegen Interaktionen zwischen Modus und prozessnahen Textspuren auf das Textverständnis vor? Haben Markierspezifika, Markierqualität und Annotationen einen positiven Effekt auf das Textverständnis?

(FF3) Lassen sich strategische Verarbeitungsmuster abbilden und welche strategischen Verarbeitungsmuster begünstigen das Verständnis?

Methode

Die Studie wurde mit 51 Oberstufenlernenden ($M = 16.91$, $SD = 1.38$, $m = 36\%$, $w = 64\%$) aus drei Schulen durchgeführt. Jede:r Schüler:in las zwei Prosatexte im Umfang von jeweils einer Seite (within-persons) als Printtext oder auf dem Tablet mit digitalem Stift (between-persons). Das Textverständnis wurde durch 10 Items (2 MC-Items, 8 halb-offene Items) erhoben und als Gesamtscore berechnet ($\alpha = .65$) ($M = 8.06$, $SD = 3.15$). Die prozessnahen Textspuren wurden in Form von Markierungsumfang (durchschnittliche Anzahl Wörter), Markierspezifika (Nutzung mehrerer Farben), Markierqualität (Markierungen von literaturwissenschaftlich relevanten Textstellen) und Annotationen (Notizen) erhoben. Sie wurden ebenfalls über die beiden Texte hinweg als Gesamtscores summiert, da sich die Vorgehensweisen within-persons nicht unterschieden. Weiter wurden Geschlecht und außerschulisches Leseverhalten erfasst (9 Items, Skalierung 1-4, $\alpha = .72$) (Mang et al., 2021).

Ergebnisse und ihre Bedeutung

(FF1) *t*-Tests zeigen, dass der Modus das literarische Textverständnis nicht beeinflusst ($t(1, 50) 0.22$, n.s.): Printmodus ($M = 8.83$, $SD = 2.91$), digitaler Modus ($M = 8.63$, $SD = 3.23$). Überdies beeinflusst der Modus nicht die jeweiligen prozessnahen Textspuren.

(FF2) Regressionsanalysen belegen ferner keine Interaktionen des Modus mit den einzelnen Textspuren mit Blick auf das Textverständnis. Es liegen Haupteffekte der qualitativen Markierungen ($\beta = .31$, $p = .027$), der Markierspezifika ($\beta = .34$, $p = .015$) und der Annotationen ($\beta = .33$, $p = .017$) auf das Textverständnis vor.

(FF3) Eine Clusteranalyse ergibt drei Verarbeitungsmuster: (1) Markieren relevanter Textstellen, mehrfarbiges Markieren, Annotieren. (2) Lediglich einfarbiges Markieren. (3) Keine Textspuren. Eine ANOVA mit den Verarbeitungsmustern und dem Modus ergab einen starken Haupteffekt der Verarbeitungsmuster ($F(1, 51) 5.23$, $p = .019$, $\eta^2 = .09$). Lernende, die nach dem ersten Muster vorgehen, erzielten ein signifikant besseres Textverständnis ($M = 10.05$, $SD = 2.61$) als Lernende, die nichts markierten ($M = 6.40$, $SD = 2.51$).

Die Studie verdeutlicht, dass das literarische Verständnis nicht vom Modus beeinflusst wird, sondern die qualitative, elaborierende Textbearbeitung. Ob schwächere Schüler:innen von einer expliziten Vermittlung profitieren, sollte mithilfe einer Intervention geprüft werden.

PS-26: Förderung und Transfer

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: SO 418

Chair der Sitzung: Svenja Lemmrich

Förderung adaptiver Lehrkompetenz im mathematischen Problemlösen: Evaluation eines innovativen Seminarkonzepts

Svenja Lemmrich, Dominik Leiss, Timo Ehmke

Leuphana Universität Lüneburg, Deutschland

Theoretischer Hintergrund

Die pädagogische Debatte über den Umgang mit Heterogenität im Unterricht ist in einer diversen, pluralistischen Gesellschaft aktueller denn je. Während sich der Begriff der Heterogenität früher hauptsächlich auf leistungsbezogene Unterschiede zwischen den Lernenden bezog, umfasst er heute alle Dimensionen. Zudem sollen Maßnahmen der äußeren Differenzierung im Schulsystem abgebaut werden, um eine defizitorientierte Eingruppierung der Lerner*innen zu vermeiden (Trautmann & Wischer, 2009). Lehrkräfte stehen also vor der Herausforderung, den individuellen Voraussetzungen aller Schüler*innen gerecht zu werden (Beck et al., 2008). Dem adaptiven Unterricht kommt dabei eine zentrale Rolle zu (Dumont, 2019). Anhand eines entwickelten Handlungsmodells für adaptive Lernunterstützung (Lemmrich et al., 2024), welches sowohl die Ebene der adaptive Unterrichtsplanung als auch die des adaptiven situativen Handelns im laufenden Unterricht berücksichtigt, können generische Facetten adaptiver Lernunterstützung fachspezifisch ausdifferenziert werden. Für das Fach Mathematik, welches ein primär aufgabenbasierter Unterricht ist, ergibt sich der besondere Bedarf einer Fokussierung auf Bearbeitungs- und Lösungsprozesse (Leiß, 2007), welche durch Lehrkräfte adaptiv unterstützt werden müssen. Ein ebenfalls entwickeltes, innovatives Seminarkonzept für angehende Lehrkräfte vermittelt Kompetenz der adaptiven Lernunterstützung bei mathematischen Problemlöseprozessen. Der Aufbau des Seminars zielt darauf ab, traditionelle Theorie-Praxis-Barrieren überwinden (Schilling & Leiss, 2022). Dazu berücksichtigt es zentrale Entwicklungsstufen von Kompetenz (Motivation/Wissen/Überzeugen – situationsspezifische Fähigkeiten – Performanz – Reflexion, Blömeke et al., 2015; Klock, 2020) und nutzt multiperspektivische Videos von authentischen Unterrichtssituationen. Das Seminar wird parallel zur Implementation in die Lehrkräftebildung von Mathematiklehrkräften evaluiert.

Fragestellung

Die handlungsleitende Forschungsfrage ist: Inwiefern wird durch die Teilnahme am Seminar die Kompetenz vermittelt, Schüler*innen in ihrem Problemlöseprozess adaptiv zu unterstützen?

Methode

Die Evaluation des Seminarkonzepts erfolgte in einem Prä-Post-Design. Die Experimentalgruppe (n = 51) und die Kontrollgruppe (n = 147) haben zu Beginn und am Ende eines Semesters einen 90-minütigen Online-Test bearbeitet mit Items aus den Bereichen Fachwissen, fachdidaktisches Wissen, Diagnose und Intervention. Die befragten angehenden Lehrkräfte wurden im Sommersemester 2021 und im Sommersemester 2023 jeweils in der Kohorte der Mathematiklehramtsstudierenden im vierten Bachelorsemester der Leuphana Universität Lüneburg rekrutiert. Für potenzielle Unterschiede zwischen den Untersuchungsbedingungen wurden t-Tests bei unabhängigen Stichproben und für die Veränderung der Konstrukte im Verlauf der Intervention t-Tests bei gepaarten Stichproben berechnet. Für die Berechnung eines globalen Fähigkeitswertes adaptiver Lernunterstützungskompetenz wurden die Testdaten eindimensional raschskaliert. Zur Überprüfung von Leistungsunterschieden zwischen Experimental- und Kontrollgruppe bezüglich des eindimensionalen Konstrukts adaptiver Lernunterstützungskompetenz beim mathematischen Problemlösen wurden Regressionsmodelle berechnet.

Ergebnisse und ihre Bedeutung

Es zeigt sich, dass die Kompetenz der Lehramtsstudierenden durch ein Seminarkonzept mit multiperspektivischen Videos und einer Einbettung in eine community of practice verbessert werden konnte, insbesondere im deklarativen fachdidaktischen Wissen. Bezüglich situationsspezifischer Fähigkeiten im Bereich der Diagnose/Auswahl adaptiver Intervention konnte trotz konstruktiv-reflexiver Auseinandersetzung mit den Praxiselementen, keine messbaren Veränderungen erzielt werden. Die Ergebnisse verdeutlichen die Komplexität einer „guten“ adaptiven Lernunterstützung im Fachunterricht. Diese ist lediglich ein Teil adaptiver Lehrkompetenz und konnte nicht in allen Facetten im Seminar erreicht werden. Adaptive Lernunterstützung muss auf die jeweilige fachliche Situation bezogen werden: Eine Verschränkung von Fachdidaktik und Bildungswissenschaft im Studium ist notwendig, um adaptive Lernunterstützung sowohl als allgemein-pädagogisches als auch fachspezifisches Konzept zu thematisieren. Zukünftig sollte Rolle der einzelnen Kompetenzfacetten tiefergehend in diesem fachspezifischen Kontext untersucht werden (bspw. Motivation, Einstellungen aber auch Reflexion von Praxiserfahrungen). So kann das Seminarkonzept dahingehend innoviert werden, dass die ganzheitliche Entwicklung professioneller Kompetenz unterstützt wird.

Implementation und Wirksamkeit einer wissenschaftsbasierten Lese- und Schreibförderung

Nora Fröhlich^{1,2}, Lisa Birnbaum^{1,3}, Alexandra Dehme¹, Simone Jambor-Fahlen⁴, Seda Yilmaz Wörfel⁴, Elmar Souvignier⁵, Hanna Dumont⁶, Benjamin Fauth^{1,2}

¹Institut für Bildungsanalysen Baden-Württemberg (IBBW); ²Eberhard Karls Universität Tübingen; ³Friedrich-Alexander-Universität Erlangen-Nürnberg; ⁴Mercator Institut für Sprachförderung und Deutsch als Zweitsprache, Universität zu Köln;

⁵Westfälische Wilhelms-Universität Münster; ⁶Universität Potsdam

Implementationsforschung zielt häufig darauf ab, die Wirksamkeit einer Innovation in einer Interventionsgruppe im Vergleich zu einer Kontrollgruppe zu untersuchen (z.B. Dhillon et al., 2015; Song & Herman, 2010). Wenn eine Intervention im Feld implementiert wird, wird meistens die Implementationstreue untersucht und Aspekte, die den Implementierungsprozess und die Ergebnisse beeinflussen könnten, vernachlässigt (Century & Cassata, 2016; Schrader et al. 2020). Bei diesen Studien bleibt offen, weshalb eine Innovation, für wen und unter welchen Bedingungen funktioniert (Century & Cassata, 2016). Deshalb ist es notwendig, auch die Rahmenbedingungen einer Intervention zu betrachten.

In dieser Studie werden die Wirksamkeit und die Rahmenbedingungen eines umfangreichen und wissenschaftsbasierten Pilotprojekts zur Förderung von vier Teilkompetenzen untersucht – Leseflüssigkeit, Lesestrategien, Schreibflüssigkeit und Schreibstrategien:

1. Wie wird die Förderung in den vier Modulen umgesetzt?
1. Wie wirksam sind die einzelnen Förder-Module?

1. Unter welchen Rahmenbedingungen werden die einzelnen Module umgesetzt?

Die Intervention wurde von Lehrkräften durchgeführt, die Fortbildungen und Beratungsangebote erhielten. Das Förderkonzept umfasst pro Modul 24-32 fachübergreifende, kurze Trainingseinheiten von je 20 Minuten über 6-8 Wochen für Fünftklässlerinnen und Fünftklässler an Haupt- und Werkrealschulen in Baden-Württemberg. Insgesamt nahmen $n = 996$ Schülerinnen und Schüler und $n = 102$ Lehrkräfte in der Interventionsgruppe sowie $n = 282$ Schülerinnen und Schüler und $n = 26$ Lehrkräfte in der Vergleichsgruppe an der Studie teil.

Zur Erfassung der Umsetzung des Pilotprojekts wurden nach Abschluss jedes Moduls Fragebogendaten von Lernenden und Lehrkräften z.B. zu Häufigkeiten der Einheit pro Woche und zur Dauer der Einheiten erhoben.

Zur Erfassung der Lernvoraussetzung wurden das Leseverständnis (ELFE II; Lenhard et al., 2020), die Lesegeschwindigkeit (Lernstand 5; IBBW, 2020) und die kognitiven Fähigkeiten (Matrizentest CFT 20-R; Weiß, 2019) der Lernenden erfasst. Vor und nach jedem Modul wurden außerdem Prä- und Posttests zur jeweils geförderten Teilkompetenz durchgeführt (Lese- und Schreibstrategien nach Knips et al., 2021 und 2022; beide angelehnt an Souvignier & Mokhelesgerami (2006); Schreibflüssigkeit nach Stephany et al., 2020). Eine Ausnahme ist die Leseflüssigkeit, welche über acht Online-Tests einer Lernverlaufsdiagnostik erfasst wurde (quop; Souvignier et al., 2021). Die Wirksamkeit wurde bisher anhand von Intention-to-Treat-Analysen untersucht. In einem weiteren Schritt werden Complier-Average-Causal-Effect-Modelle berechnet, um die Daten zur tatsächlichen Umsetzung zu berücksichtigen.

Um die Rahmenbedingungen zu erfassen, beantworteten Lehrkräfte z.B. Fragen zur Fortbildung und Betreuung durch die Fachberaterinnen und Fachberater. Regelmäßige Diskussions- und Feedbackrunden mit den Beteiligten (ausg. Lernende) sowie Fokusgruppeninterviews boten zusätzliche Einblicke in den Umsetzungsprozess. Außerdem wurden Interviews mit den beteiligten Personen auf administrativer Ebene durchgeführt (Institut für Bildungsanalysen Baden-Württemberg, Ministerium für Kultus, Jugend und Sport, Zentrum für Schulqualität und Lehrerbildung Baden-Württemberg, Mercator Institut für Sprachförderung und Deutsch als Zweitsprache).

Pro Modul gaben 78.3% bis 100% der Lehrkräfte an, die Materialien zwei- bis viermal pro Woche einzusetzen.

Bezüglich der Wirksamkeit wurden signifikante Gruppeneffekte zugunsten der Interventionsgruppe für die Module Lesestrategien ($\beta = .46$, $p < .001$) und Schreibflüssigkeit ($\beta = .28$, $p = .005$) gefunden. Für die Module Leseflüssigkeit ($\beta = .08$, $p = .465$) und Schreibstrategien ($\beta = .08$, $p = .411$) konnten keine Effekte gefunden werden.

Die qualitativen Daten werden bis zur Vorstellung des Beitrags auf der GEBF 2025 ausgewertet. Basierend auf ersten Datenanalysen deutet sich an, dass das Material für einige Lernende zu schwer war und die Umsetzung der Strategietrainings für die Lehrkräfte eine Herausforderung darstellte.

Durch die Untersuchung der Rahmenbedingungen kann ein vertieftes Verständnis erlangt werden, weshalb, für wen und unter welchen Umständen die Intervention (nicht) wirksam ist. Es wird diskutiert, wie daraus konkrete Hinweise zu notwendigen Anpassungen und Gelingensbedingungen der spezifischen Intervention abgeleitet werden können und inwiefern der gewählte methodische Ansatz auch in anderen Implementationsstudien zum Erkenntnisgewinn beitragen kann.

Evaluating the Impact of the Identity Project on Critical Consciousness among Adolescents

Oktaý Balci^{1,2}, Linda P. Juang¹, Maja K. Schachner², Miriam Schwarzenenthal³, Zhihao Zhong¹

¹Universität Potsdam, Deutschland; ²Martin-Luther-Universität Halle-Wittenberg; ³Bergische Universität Wuppertal

Institutional barriers, language challenges, discriminatory teaching practices, and socioeconomic disadvantages systematically disadvantage individuals with migration histories in the German education system (El-Mafaalani, 2021). These structural inequalities impede educational opportunities, exacerbating existing inequities (El-Mafaalani, 2021), and negatively impacting the development of diverse youth (Garcia Coll et al., 1996). In this context, critical consciousness (CC) becomes an essential developmental asset for youth (Schwarzenenthal et al., 2022). CC refers to recognizing and analyzing systemic inequalities and social injustices (critical reflection) and acting against these oppressive structures (critical action) (Freire, 1970; Watts, 2011). This study aims to examine predictors of greater CC among diverse youth in Germany. Specifically, we investigate how the Identity Project (IP) intervention may contribute to CC development through ethnic-racial identity (ERI) exploration, critical reflection, and a critical consciousness classroom climate (CCCC), a classroom climate which encourages discussions and awareness about societal inequities. The IP is an eight-week school-based program allowing students to explore their ERI and develop a deeper identity resolution (Umaña-Taylor et al., 2018). A German study has shown that the IP intervention leads to greater ERI exploration among diverse youth (Juang et al., 2020). ERI exploration fosters more critical reflection, which could promote more critical action (Mathews et al., 2020). Additionally, another German study shows that CCCC is related to critical actions, highlighting the importance of supportive environments in developing CC (Schwarzenenthal et al., 2022).

Our research questions are:
1. Does the IP intervention promote critical action among adolescents?
2. If so, is relationship mediated by ERI exploration and critical reflection?
3. Are these relationships moderated by CCCC?

We hypothesize that:
1. Participation in the IP intervention will positively predict critical action among youth.
2. The relationship between the IP intervention and critical action will be mediated by ERI exploration and critical reflection.
3. The CCCC will moderate these relationships, enhancing the effects of ERI exploration and critical reflection on critical action.

Our sample consists of 275 7th and 8th graders from four schools in former East Germany (Mage = 13.23 years, 54.58% female, 31.07% with migration history) who filled out self-report questionnaires. We have T1 (pre-intervention) data and report preliminary analyses based on this dataset. We use bivariate correlations and t-tests with cross-sectional T1 data to examine the relationships among ERI exploration, critical reflection, CCCC and critical action. Data for T2 (post-intervention) have been recently collected and are undergoing data cleaning. The final presentation will test the moderated mediation model using path analysis, incorporating both T1 and T2 data to test the proposed hypotheses and include the effect of the IP intervention. The study hypotheses and planned analyses are pre-registered.

ERI exploration positively correlates with critical reflection ($r = 0.13$, $p = 0.03$), CCCC ($r = 0.33$, $p < 0.01$), and critical action ($r = 0.34$, $p < 0.01$). Critical reflection also significantly correlates with CCCC ($r = 0.26$, $p < 0.01$) and critical action ($r = 0.36$, $p < 0.01$). Furthermore, CCCC correlates significantly with critical action ($r = 0.31$, $p < 0.01$). T-test analyses showed that individuals with

migration history reported higher scores in ERI exploration ($M = 7.97$, $p < 0.01$) and critical action ($M = 13.37$, $p = 0.04$) compared to those without ($M = 5.78$ and $M = 11.70$).

These preliminary findings indicate that all key variables are correlated positively, in line with previous studies. Planned longitudinal analyses with T1 and T2 data will test whether the IP intervention promotes CC among diverse youth in Germany, potentially strengthening these interrelationships. In societies marked by persistent inequities, classrooms that promote greater awareness of and motivation to act against inequities is one way that education can meet social challenges.

Der Transfer innovativer Unterrichtskonzepte: Die Rolle des Schulkontexts im Kaskadenmodell

Rebekka Stahnke¹, Dilan Şahin-Gür², Corinna Hankeln², Susanne Prediger^{1,2}

¹IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Berlin, Deutschland; ²Technische Universität Dortmund

Im Kaskadenmodell des Transfers werden Multiplikator:innen qualifiziert, um eigenständig Lehrkräftefortbildungen zu innovativen Unterrichtskonzepten zu geben. Dadurch können deutlich mehr Lehrkräfte erreicht werden als bei der Fortbildung durch die Programmentwickler:innen selbst (Borko, 2004; Maass et al., 2019; Roesken-Winter et al., 2021). Jedoch kann es im Kaskadenmodell zu Wirkungsverlusten über die verschiedenen Ebenen kommen, sodass die Innovation letztlich kaum bei den Schüler:innen ankommt (Krainer, 2015). Es wird angenommen, dass die Berücksichtigung der Adaptierbarkeit des Unterrichtskonzepts an verschiedene Rahmenbedingungen die Wirksamkeit verbessern kann (Bruns et al., 2023; Roesken-Winter et al., 2021). Weiterhin wird vermutet, dass ein günstiger Schulkontext die Implementation neuer Unterrichtskonzepte an den jeweiligen Schulen mit ihren individuellen Rahmenbedingungen unterstützen kann. Ein innovationsförderlicher Schulkontext ist gekennzeichnet durch etwa die Unterstützung durch die Schulleitung, ein positives Transferklima an der Schule sowie die Kooperation im Kollegium (Grossman & Sales, 2011; Lipowsky & Rzejak, 2019). Jedoch gibt es wenig empirische Evidenz für die vermuteten Zusammenhänge.

Hier setzt der Beitrag an und untersucht, ob (1) der Transfer eines evidenzbasierten Unterrichtskonzepts zum sprachbildenden Mathematikunterricht in einem Kaskadenmodell, das die Adaptierbarkeit des Konzepts berücksichtigt, zu Kompetenzzuwächsen auf der Ebene der Schüler:innen führt. Zudem wird überprüft, ob (2) Merkmale des Schulkontexts diesen Transfer beeinflussen.

Im Projekt *Fach-BiSS* wurde ein Kaskadenmodell für den Transfer eines evidenzbasierten Unterrichtskonzepts zum sprachbildenden Mathematikunterricht genutzt (Prediger, 2019; Prediger & Neugebauer 2023). Zunächst wurden in vier Bundesländern Multiplikator:innen qualifiziert, um Fortbildungen zum Unterrichtskonzept zu geben. In einem Prä-Post-Wartekontrollgruppendesign wurden die Daten von 21 Lehrkräften mit 277 Schüler:innen in der Interventionsgruppe und 13 Lehrkräften mit 202 Schüler:innen in der Wartekontrollgruppe erhoben. Die Lehrkräfte der Interventionsgruppe nahmen an einer 15-monatigen Fortbildung teil, während die Wartekontrollgruppe ihren Unterricht wie üblich gestaltete. Auf Ebene der Schüler:innen wurden kognitiven und sprachlichen Voraussetzungen sowie die Leistung im Bereich Prozentrechnung (prä/post) erhoben. Die Lehrkräfte wurden zu drei Schulkontextvariablen befragt: die Unterstützung durch die Schulleitung, das Schultransferklima und die Kooperation im Kollegium. Alle Instrumente wiesen mindestens zufriedenstellende Reliabilitäten auf. Zur Beantwortung der Fragestellung wurden Mehrebenenmodelle berechnet, die die Clusterung der Schüler:innen innerhalb von Klassen berücksichtigen ($ICC=.39$).

Hinsichtlich der ersten Forschungsfrage sagt die Zugehörigkeit zur Interventionsgruppe eine höhere Leistung im Bereich Prozente vorher ($b = 5.06$; $p < .001$; unter Kontrolle der individuellen Voraussetzung sowie des Prätestwerts). Bezüglich der zweiten Forschungsfrage erklären Aspekte des Schulkontextes zusätzliche Varianz in der Schüler:innenleistung: Die Unterstützung durch die Schulleitung ist dabei ein positiver Prädiktor ($b = 2.49$; $p = .02$), die Kooperation im Kollegium jedoch ein negativer Prädiktor ($b = -2.45$; $p = .02$). Das Transferklima hat keine Bedeutung für die Schüler:innenleistung ($b = 1.65$; $p = .19$). Es zeigen sich jedoch keine signifikanten Interaktionen zwischen der Gruppenzugehörigkeit und den Schulkontextvariablen, d.h. die Bedeutung dieser Variablen unterscheidet sich nicht signifikant in beiden Gruppen.

Die Ergebnisse des Beitrags sprechen für einen erfolgreichen Transfer des Unterrichtskonzepts zum sprachbildenden Mathematikunterricht im Kaskadenmodell. So ist das Konzept über die Qualifizierung von Multiplikator:innen, die dann Lehrkräfte fortbilden, welche dann ihren Unterricht verändern bis hin zu den Schüler:innenleistungen wirksam. Die Berücksichtigung der Adaptierbarkeit von Konzepten und Materialien könnte ein wichtiger Faktor für diesen Transfererfolg gewesen sein (vgl. Bruns et al., 2023). Der Schulkontext zeigt sich als allgemein bedeutsam für die Schüler:innenleistungen: Erwartungsgemäß wirkt die Unterstützung der Schulleitung positiv. Überraschend ist jedoch die scheinbar negative Wirkung der Kooperation auf die Schüler:innenleistung. Hier sind weitere Untersuchungen nötig, ob z.B. andere Eigenschaften der Schüler:innen oder der Schule sowohl einer stärkeren Kooperation als auch niedrigeren Leistungen zugrunde liegen. Zuletzt deuten die Ergebnisse darauf hin, dass die Berücksichtigung der Adaptierbarkeit im Kaskadenmodell die vermutete Rolle des Schulkontexts für den Transfer abgeschwächt haben könnte, sodass dieser für alle Schulen der Studie eine ähnliche Rolle spielt.

PS-27: Innovation und Qualität im MINT-Unterricht: Beiträge zur Bildungsforschung und Praxis

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 242

Chair der Sitzung: Stefan Ufer

ENTFÄLLT Formatives Assessment im (Mathematik-)Unterricht: Berücksichtigung von Effekten generischer und fachspezifischer Unterrichtsqualität

Michael Besser¹, Maike Hagena², Thilo Kleickmann³

¹Leuphana Universität Lüneburg, Deutschland; ²Leibniz Universität Hannover, Deutschland; ³Christian-Albrechts-Universität zu Kiel, Deutschland

Theorie. Formatives Assessment gilt als vielversprechende Möglichkeit der Unterstützung schulischer Lernprozesse (Andrade, 2010; Black & William, 2009), in der oft zitierten Studie von Hattie (2008) stellt die Implementation von Formativem Assessment einen der erklärungsstärksten Faktoren guten Unterrichts dar. Meta-Studien zeigen aber ebenso auf, dass eine Implementation von Formativem Assessment in Unterricht für Lehrkräfte herausfordernd ist und dass Effekte der Implementation in Abhängigkeit verschiedener Kontext- und Personenfaktoren stark variieren (Heitink et al., 2016; Kingston & Nash, 2011; Yan et al., 2021). Zum besseren Verstehen von Effekten von Formativem Assessment schlagen Decristan et al. (2015) daher vor, bei der Analyse der Wirkung von Formativem Assessment auch Effekte generischer und fachspezifischer Unterrichtsqualität zu berücksichtigen. In einer ersten Studie für den Mathematikunterricht berichten Pinger et al. (2018) in diesem Kontext Interaktionseffekte von generischer Unterrichtsqualität und der Implementation von Formativem Assessment – Studien, die sowohl generische als auch fachspezifische Unterrichtsqualität zum Verstehen von Wirkweisen von Formativem Assessment (im Mathematikunterricht) berücksichtigen, liegen jedoch trotz aktueller Diskussionen zur Bedeutung dieser Unterrichtsdimensionen (Charalambous & Praetorius, 2018; Praetorius et al., 2020a) nicht vor.

Forschungsfrage. Ausgehend von aufgezeigtem Desideratum adressiert der Vortrag folgende Forschungsfrage: Welchen Mehrwert für das Verstehen von Effekten der Implementation Formativen Assessments in den (Mathematik-)Unterricht bietet die Berücksichtigung sowohl generischer als auch fachspezifischer Unterrichtsqualität?

Methode. Im DFG-Forschungsprojekt COCA haben von Oktober 2010 bis März 2011 insgesamt 856 Schüler*innen aus 39 Mathematikklassen (Jahrgang 9) an einer 13 Unterrichtsstunden umfassenden Interventionsstudie zum „Satz des Pythagoras“ teilgenommen. Die Klassen wurden einer Kontrollbedingung (N=15) oder einer Untersuchungsbedingung (N=24) zugeteilt, in letzterer wurde Formatives Assessment in regulären Mathematikunterricht implementiert. Die Mathematikleistung der Schüler*innen wurde durch einen Leistungstest im Pre-Post-Design erfasst. Zur Beschreibung der Unterrichtsqualität während der Interventionsstudie wurden 72 Unterrichtsstunden videografiert. Im Rahmen einer Re-Analyse der Videos wurden als Subdimensionen generischer Unterrichtsqualität in Anlehnung an das „Modell der drei Basisdimensionen“ (Praetorius et al., 2020b) die Klassenführung, die kognitive Aktivierung und die kognitive Unterstützung mittels hoch-inferenter Ratings erfasst (vierstufige Ratingskala; drei Kodierer; ICC(3,3)>.59). Als Subfacetten fachspezifischer Unterrichtsqualität wurden, basierend auf dem Modell von Schlesinger et al. (2018) zur Erfassung fachspezifischer Unterrichtsqualität, die Unterstützung beim Verstehen des Satz des Pythagoras (Drollinger-Vetter, 2011) sowie die Unterstützung beim Verstehen prozessbezogenen Arbeitens in Mathematik (Kaiser, 2020) mittels niedrig-inferenter Ratings kodiert (Intensität des Auftretens von Inhalten bzw. Prozessen; drei Kodierer; ICC(3,3)>.68). Analysen von Effekten generischer und fachspezifischer Unterrichtsqualitätsdimensionen auf Schüler*innenleistungen wurden mittels Mehrebenen-Regressionsmodellen unter Berücksichtigung der Vortestleistung auf Individual- und Klassenebene sowie unter Berücksichtigung der Zugehörigkeit zu Kontroll-/Untersuchungsbedingung auf Klassenebene durchgeführt.

Ergebnisse. Benannte Mehrebenen-Regressionsanalysen belegen verschiedene (Interaktions-)Effekte, mit Blick auf die Forschungsfrage erscheint zentral: (a) Die Leistung der Schüler*innen im Pretest ist in allen Modellen der stärkste Prädiktor für die Leistung im Posttest ($b > .41$), direkte Effekte der Implementation von Formativem Assessment auf Schüler*innenleistung liegen hingegen nicht vor. (b) Es zeigen sich aber positive Effekte der generischen Subfacette Klassenführung sowie der beiden fachspezifischen Subfacetten von Unterrichtsqualität ($b > .60$) auf Schüler*innenleistungen, deren Stärke wiederum teilweise durch negative Interaktionseffekte mit der Implementation von Formativem Assessment reduziert wird ($b < -.71$).

Diskussion. Die Ergebnisse der vorliegenden Studie verdeutlichen den Mehrwert der Berücksichtigung generischer und fachspezifischer Unterrichtsqualität beim Verstehen der Wirkmechanismen von Formativem Assessment: Während sich kein direkter Effekt der Implementation auf Schüler*innenleistungen zeigt, liegen positive Effekte einzelner Facetten generischer und fachspezifischer Unterrichtsqualität vor. Die Ausprägung dieser positiven Effekte wird jedoch teils durch negative Interaktionseffekte reduziert, im Falle der Klassenführung gar (nahezu) vollständig kompensiert. Oder anders: Die Ergebnisse belegen die besondere Bedeutung und Notwendigkeit, im Kontext einer Auseinandersetzung mit Fragen nach Effekten von Formativem Assessment im (Mathematik-)Unterricht im Speziellen bzw. mit Fragen nach effektiver Unterrichtsgestaltung im Allgemeinen explizit sowohl generische als auch fachspezifische Dimensionen von Unterrichtsqualität zu berücksichtigen (Lindmeier & Heinze, 2020).

How to flip the right way: Fail, flip, fix and feed in practice

Richard François Conrardy^{1,2}, Christian Spannagel¹

¹Heidelberg University of Education; ²University of Teacher Education Bern

The Fail, Flip, Fix and Feed model (4F) is a four-phase instructional framework (Kapur et al., 2022) that extends the traditional two-phased flipped classroom model (Abeysekera & Dawson, 2015; Bergmann & Sams, 2012) by integrating the two-phased productive failure design (Kapur, 2008; Kapur & Bielaczyc, 2012).

The flipped classroom model (FCM) is designed to encourage in-class engagement (Abeysekera & Dawson, 2015) and has empirically shown to do so (Clark, 2015). FCM increases performance in mathematics as well as in secondary schools (Strelan et al., 2020). Furthermore FCM "may allow better management of mental load" (Abeysekera & Dawson, 2015, p. 9). A successful implementation of FCM would thus lead to high in-class engagement and high performance.

Productive failure (PF) has been extensively tested in mathematics secondary schools on the topic of dispersion (Brand et al., 2023; Kapur, 2010, 2014, 2016; Loibl & Leuders, 2018; Loibl & Rummel, 2014). PF has been found to increase performance despite also increasing mental load while not having any significant effect on engagement (Kapur, 2014). A successful implementation of PF would thus entail a high performance regardless of mental load.

While both models, FCM and PF, are well studied separately, combined models have not been sufficiently studied yet. This research compares the 4F-model with an alternative model (ALT) in which the problem solving (PS) gets flipped (i.e. is out-of-class) instead of the instruction. This is in line with the findings about the importance of contrasting students' solutions during instruction (Loibl & Rummel, 2014) and the unimportance of collaboration during problem-solving (Brand et al., 2023).

Given the indicators of a successful implementation of FCM and PF above, we have the following hypotheses: (1) Engagement is highest during in-class activities, irrespective of variant. (2) The 4F model does not outperform the ALT model. (3) Engagement predicts performance, and (4) Mental load does not predict performance.

To test these hypotheses we conducted an experimental study with 234 students aged 12-16 from 12 classes in Swiss secondary schools. After each of the four phases, spread over two days, the students filled in an engagement questionnaire. The random variation was introduced only during the first day, where the classes were separated into 4F and ALT groups. Knowledge was measured with a pretest and a further test at the end of every day.

There were no indications found that engagement was higher during in-class activities than during out-of-class activities. An equivalence-test confirmed that both variations performed equally in the knowledge tests after the first and second day, thus confirming the second hypothesis. Furthermore, linear regression analyses demonstrated that engagement is a good predictor for performance in the following knowledge test. Our analysis also confirmed that mental load is not a significant predictor for knowledge gain.

These results support the findings from Brand et al. (2023) and Loibl & Rummel (2014) around PF theory. Engagement being consistently high regardless of variation and of phase shows that neither variant has an edge over the other in FCM implementation. These results suggest that it may be possible to have the first two phases out of class without negatively impacting performance.

Separate und gemeinsame Effekte von Unterrichtsqualität und der Nutzung digitaler Tools auf die Mathematikleistung in der Sekundarstufe I

Christian Lindermayer¹, Timo Kosiol¹, Maria Bannert², Begüm Arvaneh², Karsten Stegmann³, Stefan Ufer¹

¹Chair for Mathematics Education, Department of Mathematics, Faculty for Mathematics, Computer Science and Statistics, LMU Munich, Munich, Germany; ²Professorship for Teaching and Learning with Digital Media, Department of Educational Sciences, TUM School of Social Sciences and Technology, TU Munich, Munich, Germany; ³Department of Psychology, Faculty for Psychology and Educational Sciences, LMU Munich, Munich, Germany

Hintergrund

Meta-Analysen zeigen, meist anhand experimenteller Studien, dass die Nutzung digitaler Tools die Mathematikleistungen der Schüler:innen positiv beeinflussen kann, besonders wenn Lehrkräfte fortgebildet wurden (Hillmayr et al., 2020). Auch zur Unterrichtsqualität und ihren Effekten auf Schülerleistungen liegen vielfältige Arbeiten vor, die sich im deutschsprachigen Raum oft am Modell der drei Basisdimensionen Klassenführung, Schülerunterstützung und kognitive Aktivierung orientieren (Praetorius et al., 2018). Theoretische und empirische Verknüpfungen beider Forschungsbereiche sind jedoch rar. Als einfachstes Modell kann angenommen werden, dass die häufigere Nutzung digitaler Tools und eine höhere Unterrichtsqualität nicht nur isolierte Effekte auf Schülerleistungen haben, sondern dass eine höhere Unterrichtsqualität die Effekte der Nutzung digitaler Tools verstärken könnte. Es ist jedoch weitgehend unbekannt, ob und welche Dimensionen der Unterrichtsqualität zu dieser Interaktion beitragen.

Fragestellung und Design

In einer längsschnittlichen Studie mit $N = 1752$ Schüler:innen aus $N = 83$ Klassen an Mittelschulen, Realschulen und Gymnasien wurde untersucht, inwiefern schülerberichtete Dimensionen der Unterrichtsqualität sowie die wahrgenommene Nutzungshäufigkeit digitaler Tools die Schülerleistung am Ende der Jahrgangsstufe 8 unter Kontrolle der Leistungen zum Anfang des Schuljahres vorhersagen. Weiter wurden Interaktionen der Unterrichtsqualitätsdimensionen mit der Nutzungshäufigkeit digitaler Tools auf ihre Prädiktivität untersucht.

Instrumente und Analysen

Durch Schülerratings wurde die Unterrichtsqualität entlang der drei Basisdimensionen (Kunter, 2017, $\alpha = .63-.77$) und die wahrgenommene Nutzungshäufigkeit digitaler Tools (Eigenkonstruktion, $\alpha = .76$) zu Beginn und zum Ende eines Schuljahrs erhoben. Für die Analysen wurden die Werte beider Messzeitpunkte gemittelt. Die Mathematikleistung wurde mit selbstentwickelten, für die jeweilige Schulart curricular validen Tests erhoben, die durch gemeinsame Items verknüpft wurden (Prä: $\alpha = .67$; Post: $\alpha = .60$).

Die Analyse erfolgte mit Mehrebenenmodellen mit dem R-Package lavaan, kontrolliert nach Schulart.

Ergebnisse

Die Nutzungshäufigkeit digitaler Tools zeigte einen signifikant negativen Effekt auf die Schülerleistung am Schuljahresende (unter Kontrolle der Anfangsleistung, $\beta = -.11, p < .01$). Dieser Effekt blieb bestehen, wenn jeweils eine Unterrichtsqualitätsdimension als Prädiktor ergänzt wurde. Klassenführung ($\beta = .10, p < .05$), Schülerunterstützung ($\beta = .12, p < .01$) und kognitive Aktivierung ($\beta = .11, p < .01$) zeigten positive Effekte auf die Schülerleistung. Signifikante Interaktionen mit digitaler-Tool-Nutzung wurden für Klassenführung ($\beta = .07, p < .05$) sowie Schülerunterstützung ($\beta = .06, p < .05$) gefunden, nicht jedoch für kognitive Aktivierung ($\beta = .03, p = .35$). Interaktionsplots zeigen, dass der negative Effekt digitaler Tools vor allem bei niedriger Unterrichtsqualität auftrat und nur bei sehr hoher Unterrichtsqualität ($>2SD$ über dem Mittelwert auf Klassenebene) vernachlässigbar wurde.

Diskussion

Die Studie verwendet ein einfaches Modell, um die gemeinsamen Effekte von Unterrichtsqualität und Nutzungshäufigkeit zu untersuchen, wobei die Qualität der digitalen Tools und ihrer Nutzung nicht berücksichtigt wurde. Die Ergebnisse deuten an, dass der in experimentellen Studien berichtete positive Effekt digitaler Tools im alltäglich beobachtbaren Unterricht nicht repliziert werden kann. Klassen mit häufigerer Nutzung digitaler Tools scheinen weniger vom Unterricht zu profitieren als Klassen mit seltenerer Nutzung. Dies deutet darauf hin, dass der reale Mathematikunterricht das in Meta-Analysen berichtete Potenzial digitaler Tools nicht ausschöpft. Die positiven Effekte der Unterrichtsqualität decken sich mit früheren Studien (Praetorius et al., 2018). Bezüglich der erwarteten Interaktion von Unterrichtsqualität und digitaler-Tool-Nutzung zeigt sich, dass vor allem gute Klassenführung und Schülerunterstützung die negativen Effekte der digitalen-Tool-Nutzung zu einem gewissen Maße abfedern können. Dass keine vergleichbare Interaktion für kognitive Aktivierung nachgewiesen werden konnte, könnte durch die eingeschränkte Stichprobengröße auf Klassenebene erklärt werden. Denkbar ist auch, dass kognitive Aktivierung bei der

konkreten Gestaltung und Implementation digitaler Tools für den Mathematikunterricht spezifisch berücksichtigt werden muss, um deren Potenzial für die kognitive Aktivierung voll auszuschöpfen. Unsere Studie basiert auf einem einfachen Modell der Interaktion von Unterrichtsqualität und Nutzungshäufigkeit digitaler Tools. Wünschenswert wären Arbeiten, die die Qualität dieser Nutzung genauer konzeptualisieren, beschreiben und untersuchen.

ENTFÄLLT Formatives Assessment in der universitären Ausbildung angehender MINT-Lehrkräfte – Eine Möglichkeit aus der Krise?

Maike Hagena¹, Marcus Schütte²

¹Leibniz Universität Hannover, Deutschland; ²Universität Hamburg, Deutschland

Theorie. Der eklatante Lehrkräftemangel stellt das Bildungssystem vor enorme Herausforderungen (Gehrmann, 2023; Klemm, 2020; Klemm & Zorn, 2018). Besonders betroffen sind die MINT-Fächer (Klemm, 2020). Ein Grund für den Mangel an MINT-Lehrkräften ist auch die hohe Anzahl an Studierenden, die aufgrund ungünstiger motivational-affektiver Einstellungen oder aufgrund von Leistungsdefiziten ihr Studium abbrechen (HRK, 2014). Entsprechend bedarf es im Sinne einer erfolgreichen Lehrkräftebildung geeigneter Maßnahmen, um MINT-Studierende bestmöglich in ihren individuellen Entwicklungsprozessen (motivational-affektiv und Leistung) zu unterstützen (Güldener, 2023). Ein vielversprechendes Moment einer solchen Unterstützung stellt die Idee formativen Assessments dar (Black & William, 2009), welches sich durch eine individuelle Diagnose und lernprozessbegleitende Rückmeldung von Lernständen auszeichnet (Maier, 2010). Während ein positiver Einfluss formativen Assessments auf die individuellen Entwicklungsprozesse von Schüler:innen wiederholt aufgezeigt werden konnte (Hattie & Timperley, 2007), gibt es gegenwärtig kaum empirische Evidenz für die Wirksamkeit formativen Assessments in der Hochschullehre im Allgemeinen (Gikandi et al., 2011; Yorke, 2003) und der universitären Lehrkräfteausbildung im Besonderen. Im Folgenden wird eine Studie vorgestellt, in der untersucht wird, inwieweit sich formatives Assessment wirksam in die universitäre Ausbildung angehender Grundschullehrkräfte mit Unterrichtsfach Mathematik integrieren lässt.

Forschungsfragen. Die Studie greift aufgezeigtes Forschungsdesiderat auf und adressiert folgende Forschungsfragen: Inwieweit wirkt sich der Einsatz formativen Assessments bei Grundschullehramtsstudierenden mit dem Unterrichtsfach Mathematik auf (FF1) die Veränderung der mathematikbezogenen Lernmotivation und des mathematikbezogenen Selbstkonzepts (als motivational-affektive Einstellungen) und (FF2) die Entwicklung mathematikdidaktischen Wissens (als Leistungsindikator) aus?

Methode. Im Rahmen des Forschungsprojekts DialogInMathe haben im Wintersemester 2022/2023 N=279 Grundschullehramtsstudierende (Geschlecht: m=43, w=233, d=2; Alter: MW=22.05, SD=4.46) mit dem Unterrichtsfach Mathematik an einer sich über 12 Wochen erstreckenden Interventionsstudie zur Wirkung formativen Assessments auf die Veränderung affektiv-motivationaler Einstellungen (Lernmotivation; Selbstkonzept) und die Entwicklung von Leistung (mathematikdidaktischen Wissens) teilgenommen. Über den Verlauf eines Semesters haben die Studierenden alle zwei Wochen umfangreiche Lernaufgaben bearbeitet und schriftlich ein Feedback zu ihren individuellen Aufgabenbearbeitungen erhalten. Während dem Großteil der Studierenden (EB-A; N=250) ein summatives Feedback (erreichte Punktzahl bei der Aufgabenbearbeitung) zurückgemeldet worden ist, haben Studierende, die die Vorlesung aufgrund nicht bestandener Prüfungsleistung erneut belegen mussten, neben dem summativen Feedback auch ein formatives Feedback erhalten (EB-B; N=29). Im Zuge dieses formativen Feedbacks wurden den Studierenden sowohl Stärken und Schwächen als auch Hilfestellungen für die weitere Bearbeitung zurückgemeldet. Die Leistungsmotivation (4 Items; $\alpha=.90$), das Selbstkonzept (4 Items; $\alpha=.90$) und das mathematikdidaktische Wissen (13 Items; $\alpha=.72$) der Studierenden wurden im Pre-Post-Design erfasst. Zur Kontrolle der Eingangsvoraussetzungen wurden zudem die kognitiven Grundfähigkeiten (25 Items; $\alpha=.77$) der Studierenden erhoben.

Ergebnisse. (FF1) Im Bereich der affektiv-motivationalen Einstellungen zeigen Varianzanalysen mit Messwiederholung (Faktor Zeit) und den beiden Experimentalbedingungen (Faktor Gruppe) für die Lernmotivation der Studierenden weder einen Effekt der Zeit ($p=.254$) oder der Gruppe ($p=.851$) noch einen Interaktionseffekt von Zeit und Gruppe ($p=.333$). Auch im Bereich des Selbstkonzepts zeigt sich weder ein Effekt der Zeit ($p=.624$) noch ein Interaktionseffekt von Zeit und Gruppe ($p=.197$), jedoch ein Effekt der Gruppe zugunsten der Studierenden aus EB-B ($p<.05$; Eta-Quadrat=.02). (FF2) Mit Blick auf die Entwicklung des mathematikdidaktischen Wissens belegen Varianzanalysen mit Messwiederholung (Faktor Zeit) und den Experimentalbedingungen (Faktor Gruppe) (auch unter Kontrolle der kognitiven Grundfähigkeiten) einen Effekt der Zeit ($p<.01$; Eta-Quadrat=.08), einen Interaktionseffekt aus Zeit und Gruppe ($p<.05$; Eta-Quadrat=.02) sowie einen Effekt der Gruppe zugunsten der Studierenden aus EB-B ($p<.01$; Eta-Quadrat=.06).

Diskussion. Die Ergebnisse der Studie zeigen, dass mittels formativen Assessments das Selbstkonzept und das mathematikdidaktische Wissen von Grundschullehramtsstudierenden mit Unterrichtsfach Mathematik erfolgreich gefördert werden konnte. Diese Ergebnisse ermutigen, über Möglichkeiten der Implementation formativen Assessments in die fachliche/fachdidaktische Lehre als hochschuldidaktisches Gestaltungsmoment nachzudenken (Kordts-Freundinger et al., 2021). Eine solche kritische Diskussion erscheint insbesondere vor dem Hintergrund des eklatanten Lehrkräftemangels in den MINT-Fächern einerseits bzw. den hohen Abbrecherquoten andererseits unausweichlich (Güldener, 2023; Klemm, 2020).

PS-28: Education and Equal Opportunities: Impacts of the COVID Pandemic

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: SO 318
Chair der Sitzung: Markus Vogelbacher

English Track

Hat eine COVID-19-Infektion Effekte auf die kognitive Leistungsfähigkeit von Schüler:innen der Primarstufe? Ergebnisse einer repräsentativen Studie im österreichischen Burgenland.

Thomas Leitgeb¹, Wolfram Rollett², Katja Scharenberg³

¹Pädagogische Hochschule Burgenland, Österreich; ²Universität Oldenburg, Deutschland; ³LMU München, Deutschland

Auswirkungen der COVID-19-Pandemie auf die Entwicklung von Kindern und Jugendlichen werden in der aktuellen Bildungsforschung intensiv untersucht. Der Fokus liegt dabei meist auf den Folgen von Schulschließungen und Unterrichtsausfällen (z.B. Betthäuser, et al., 2023; Patrinos, et al., 2022). Die direkten Auswirkungen von COVID-19-Infektionen werden dabei bisher ausgeblendet. Obwohl jüngere Menschen seltener erkranken, können nach einer Infektion anhaltende Symptome auftreten (Behnood et al., 2022; Lopez-Leon et al., 2022). Klinische Studien zeigen Zusammenhänge zwischen der Schwere der anhaltenden Symptome und der Schwere der COVID-19-Erkrankung (z.B. Radtke, et al., 2021). Für Erwachsene wurde dieser Zusammenhang auch für anhaltende kognitive Beeinträchtigungen empirisch nachgewiesen (Hampshire et al., 2021).

Viele klinische Studien haben sich mit den Folgen von COVID-19-Infektionen bei Kindern befasst, allerdings fehlen repräsentative Studien und Studien, die einen Vergleich von genesenen und nicht infizierten Kindern erlauben. Die vorliegende Studie greift diese Forschungslücken auf und untersucht anhand einer repräsentativen Stichprobe, ob Grundschulkindern, die von COVID-19 genesen sind, Nachteile in Bezug auf ihre kognitiven Fähigkeiten erlitten haben. Aufgrund der Ergebnisse klinischer Studien wird erwartet, dass sich entsprechende Effekte nachweisen lassen.

Im Juni 2022 wurden 1.761 Schüler:innen der zweiten und dritten Klassen (49,9% Mädchen, 50,1% Jungen) im österreichischen Burgenland mit Zustimmung ihrer Eltern untersucht (32,3% der Schüler:innen im Schuljahr 2021/22, aus 106 (63,1%) der 168 Volksschulen im Burgenland). Zusätzlich wurden die Eltern bzw. Erziehungsberechtigten der Schüler:innen befragt (n=1.438). Für die Analyse wurden vier Gruppen unterschieden: eine Gruppe von zum Untersuchungszeitpunkt noch nicht infizierten Kindern (Kontrollgruppe, n=502) und drei Genesenenengruppen (GG). Letztere bestanden aus asymptomatisch infizierten Kindern (GG1, n=251), symptomatisch erkrankten Kindern (GG2, n=850) und symptomatisch erkrankten Kindern, die ärztlich behandelt wurden (GG3, n=131). Der Arztbesuch wurde als Indikator für eine gesundheitliche Situation gewertet, die Eltern Anlass zur Sorge gab (z.B. aufgrund der Symptomatik oder Vorerkrankungen). Die zentrale unabhängige Variable war, ob die Kinder zum Zeitpunkt der Erhebung bereits mit COVID-19 infiziert waren. 78 Prozent der bereits infizierten Kinder haben sich innerhalb der fünf Monate vor der Datenerhebung infiziert.

Abhängige Variablen waren die Leistungen in den vier Subtests des Kognitiven Fähigkeitstests für Kinder (KFT1-3; Heller & Geisler, 1983): Sprachverständnis, Beziehungserkennen, Schlussfolgerndes Denken und Rechnerisches Denken.

Die Daten der drei Genesenenengruppen wurden paarweise mit denen der Kontrollgruppe verglichen. Da signifikante Unterschiede hinsichtlich des Migrationshintergrundes, der Muttersprache und Familiensprache bestanden, wurden diese Variablen mittels Kovarianzanalyse kontrolliert. Fehlende Werte wurden als zufällig fehlend behandelt und mehrfach imputiert (MICE, m=20; Buuren & Groothuis-Oudshoorn, 2011; CART, Breiman, et al., 1984). Die Daten wurden mittels iterativer proportionaler Anpassung (IPF; Deming & Stephan, 1940; Lomax & Norman, 2019) basierend auf repräsentativen Statistiken des Bundeslandes Burgenland gewichtet. Alle statistischen Tests wurden mit einer Irrtumswahrscheinlichkeit von $p < .05$ durchgeführt.

Hinsichtlich der kognitiven Fähigkeiten zeigten GG1 und GG3 eine signifikant niedrigere Testleistung im numerischen Denken als die Kontrollgruppe (GG1: $F(1,748)=7.42^{**}$, $p=.007$, part. $\eta^2=.010$; GG3: $F(1,627)=9.18^{**}$, $p=.003$, part. $\eta^2=.014$). Zudem schnitt GG3 im Sprachverständnis signifikant schwächer ab als die Kontrollgruppe ($F(1,627)=11.26^{***}$, $p<.001$, part. $\eta^2=.018$). Im Beziehungserkennen und im induktiven Denken zeigte GG3 tendenziell niedrigere Testleistungen als die Kontrollgruppe ($F(1,627)=3.57$, $p=.059$, part. $\eta^2=.006$; $F(1,637)=3.19$, $p=.075$, part. $\eta^2=.005$).

Unsere Ergebnisse deuten für zwei der drei Genesenenengruppen auf negative kognitive Effekte von COVID-19-Infektionen hin. Für symptomatisch infizierte Kinder, die medizinisch behandelt wurden, weisen die Ergebnisse stärker in diese Richtung. Die nachgewiesenen Effekte sind klein, könnten jedoch angesichts der geringen Prävalenz länger anhaltender Symptome bei Kindern (Lopez-Leon et al., 2022) bedeutsame Auswirkungen auf die betroffenen Kinder haben. Es ist zu berücksichtigen, dass die Infektion zum Zeitpunkt der Erhebung bei den meisten Kindern weniger als fünf Monate zurücklag. Um zu klären, inwieweit die nachgewiesenen Effekte zeitlich stabil sind, wurde die Stichprobe im Juni 2023 erneut mit dem Kognitiven Fähigkeitstest für Kinder getestet. Die Ergebnisse werden zum Zeitpunkt der Konferenz vorliegen.

Differential effects of the COVID-19 pandemic on social-emotional development of primary school students from different social backgrounds

Markus Vogelbacher, Christoph Homuth

LfBi – Leibniz-Institut für Bildungsverläufe, Deutschland

While many studies have been conducted on the effects of the COVID-19 pandemic on students' learning development (Betthäuser et al., 2023), and internalizing behavioural problems (Wolf & Schmitz, 2023; Rajmil et al., 2021; Kauhanen et al., 2023; Ahmed et al., 2023; Santomaro et al., 2021), fewer studies have analysed the effects on externalizing behaviour (Wolf & Schmitz, 2023; Chung et al., 2023). We extend the literature by analysing externalizing behaviour in the form of conduct problems and hyperactivity of primary school students in Germany.

Recent international research on the social-emotional development of primary school students paints a bleak picture of rising social and behavioural problems during the COVID-19 pandemic (Mistry, Benner & Kimura, 2022). However, the social-emotional development during the pandemic is rarely compared with empirical evidence on pre-pandemic development. Furthermore, the origin of variance in social-emotional development during the COVID-19 shutdown seems underexplored.

Previous studies found correlations between social-emotional problems during the pandemic-induced lockdowns and age, gender, and parental education (Chistner et al., 2021; Jarvers et al., 2023; Ravens-Sieberer et al., 2021).

We draw on the Family Investment (Conger & Donellan, 2007) and Family Stress Models (Conger et al., 1992) and expect a surge in externalizing behaviour due to the increased financial and social burdens the pandemic had on families, especially those with fewer resources from lower social strata.

We investigate pandemic effects by comparing social-emotional development in primary school students during the pandemic to the development in children of the same age group that did not experience the pandemic. We use a multicohort comparison design to identify the effects of the pandemic, for which we use data from two German longitudinal studies. Children of the Starting Cohort Newborns (SC1) of the National Educational Panel Study (NEPS-Netzwerk, 2023) are our treatment group. They experienced first school shutdown in 2020 when they were in Grade 2. We draw on children from Pairfam (Brüderl et al., 2023) for the construction of our control group. They were, depending on their birth years, in primary school at different times between 2008 and 2019. In both samples, data on children's behaviour was collected via the Strengths and Difficulties Questionnaire (Goodman, 1997, 2001), rated by their parents when the children were around six, eight, and ten years old.

Preliminary results indicate a general aggravation in hyperactivity in the treatment group but not for conduct problems. While we found no general widening of the gap in the comparison of different socioeconomic groups, the preliminary results indicate an intersectional increase in hyperactivity for girls from lower social strata.

In the next step, we utilize difference-in-differences and value-added regression models to estimate the effect of the pandemic, and further investigate if structural factors and processual characteristics in the family explain differences in social-emotional development during the pandemic.

Einfluss der Covid-19-Pandemie auf den Lernerfolg Schweizer Schüler:innen und Identifikation von Prädiktoren in der Schulumgebung

Valentin Unger¹, Fabian Grünig¹, Jan Hochweber¹, Julian Brauchle¹, Nicolas Hübner², Michael Kickmeier-Rust¹

¹Pädagogische Hochschule St.Gallen, Schweiz; ²Eberhard Karls Universität Tübingen, Deutschland

Empirischer Hintergrund:

Die Schulschließungen zur Eindämmung der Covid-19-Pandemie stellten alle Beteiligten des Schulsystems vor große Herausforderungen. Empirische Studien zeigen während dieser Zeit einen geringeren Lernzuwachs der Schüler:innen als unter Normalbedingungen^{1,2,3}, allerdings mit hoher Varianz zwischen verschiedenen Studien.¹

Die Herausforderungen variierten dabei zwischen den Schulen aufgrund unterschiedlicher kontextueller Bedingungen. So waren die Schulen bei der Etablierung des Fernunterrichts besonders von der Ausstattung des Haushalts der Schüler:innen abhängig.⁴ Das zeigt sich unter anderem in empirisch nachgewiesenen Zusammenhängen zwischen dem sozioökonomischen Status der Schüler:innen und ihrem Lernerfolg im Fernunterricht.^{1,5} Auch elterliche Unterstützung hatte einen großen Einfluss auf den Lernerfolg im Fernunterricht.⁶ Bildungsnähere Eltern konnten ihre Kinder dabei besser unterstützen als bildungsfernere.^{7,8} Nicht-deutschsprachige Elternhäuser hatten zusätzlich größere Schwierigkeiten bei der Lernunterstützung.^{9,2} Zu Beginn der Schulschließung mussten Schulen neue Wege zur Distribution der Aufgaben finden¹⁰, wobei Unterschiede zwischen ländlichen und städtischen Regionen ebenfalls den Lernerfolg beeinflusst haben könnten.¹¹

Forschungsfragen:

Die Befunde geben Hinweise auf (differenzielle) Auswirkungen der Covid-19-Pandemie auf den Lernerfolg der Schüler:innen. Die genannten Studien beruhen dabei vorrangig auf querschnittlichen Untersuchungen und/oder kleineren Datensätzen und weisen dabei teils hohe Varianz auf. Um verlässliche Ergebnisse zu generieren, werden Längsschnittstudien mit größeren Datensätzen benötigt.¹ Diese Studie untersucht hierfür folgende Forschungsfragen in einer Längsschnittanalyse mit insgesamt N=43'441 Schweizer Schüler:innen:

RQ1: Inwiefern können über den Verlauf der Covid-19-Pandemie Lerneinbußen festgestellt werden?

RQ2: Inwiefern beeinflussen die folgenden theoriebasierten Faktoren in der Umgebung der Schule den Lernerfolg während der Covid-19-Pandemie?

- Durchschnittliches Einkommen
- Durchschnittlicher höchster Bildungsabschluss
- Anteil vorrangig nicht-deutschsprechender Bevölkerung
- Anteil sozialhilfebeziehender Schüler:innen
- Urbanisierungsgrad

Methodisches Vorgehen:

Die Daten dieser Studie entstammen der digitalen Lernplattform Lernpass plus, welche teilverpflichtende Tests in den Hauptfächern Deutsch, Englisch, Französisch und Mathematik von Schüler:innen der 8. und 9. Klasse umfasst. Diese Daten ermöglichen eine Längsschnittanalyse der Testergebnisse in den Schuljahren 2018-2020 (Kohorte 1, Test in Kl.8 vor den Schulschließungen und in Kl.9 direkt nach den Schulschließungen = "Coronakohorte"), 2019-2021 (Kohorte 2, Test in Kl.8 direkt und in Kl.9 ein Jahr nach den Schulschließungen) und 2020-2022 (Kohorte 3, Test in Kl.8 ein Jahr und in Kl.9 zwei Jahre nach den Schulschließungen = "Nach-Corona-Kohorte").

Zur Beantwortung von RQ1 wurde der Lernzuwachs der einzelnen Schüler:innen von der 8.-9. Klasse pro Fach berechnet und zwischen den Kohorten verglichen. Dafür wurden längsschnittliche Mehrgruppen-Item-Response-Modelle^{12,13} berechnet. Die Testergebnisse wurden mithilfe dichotomer Rasch-Modelle geschätzt, wobei die Schwierigkeitsparameter aus Kalibrierungsstudien berücksichtigt wurden. Der Lernzuwachs wurde mit Hilfe latenter Veränderungsmodelle abgeschätzt.^{14,15} Zur Beantwortung von RQ2 wurden die Veränderungsmodelle durch Regressionsmodelle erweitert, um den Einfluss verschiedener Prädiktoren auf den Lernzuwachs zu analysieren. Dabei wurde die Differenz der Regressionskoeffizienten zwischen den Kohorten berechnet und auf statistische Signifikanz getestet.

Ergebnisse:

Die durchschnittlichen Testergebnisse sowie die Unterschiede zwischen den Kohorten geben Hinweise auf die schlechtesten Testergebnisse unmittelbar nach den und die geringsten Lernzuwächse über die Schulschließungen (Kohorte 1) und den größten Lernzuwachs in Kohorte 2 nach den Schulschließungen. Die Ergebnisse aus Kohorte 3 rangierten in Bezug auf die Leistungsentwicklung zwischen Kohorte 1 und 2.

Die Ergebnisse der berechneten Regressionsmodelle kommen zu keinen konsistenten Ergebnissen. Während sich für Deutsch und Französisch keine Unterschiede zwischen den Kohorten zeigen, finden sich in Englisch statistisch signifikante Unterschiede für das Einkommen, die elterliche Bildung und den Anteil Deutschsprechender sowie in Mathematik für den Anteil sozialhilfebeziehender Schüler:innen.

Diskussion:

Die Hinweise auf Lerneinbußen widersprechen aktuellen PISA-Ergebnissen, die in der Schweiz keine Lerneinbußen sehen.¹⁶ Unsere Daten weisen zudem auf Catch-up-Effekte hin. Unsere Daten zeigen kein einheitliches Bild bezüglich des Einflusses der Prädiktoren zwischen den Kohorten.

Auch wenn die Covid-19-Pandemie überwunden scheint, ist die Wahrscheinlichkeit für ähnliche zukünftige Schulschließungssituationen, beispielsweise aufgrund von Krankheitswellen oder Extremwetterereignissen, hoch.^{17,18} Erfahrungen der Covid-19-Pandemie sollten demnach genutzt werden, um mit Hilfe von Notfallplänen, geförderter Selbstregulation der Schüler:innen oder remedialen Maßnahmen besser reagieren zu können.

Physical fitness of primary school children – Associations with the Covid-19 pandemic and social structure

Paula Teich¹, Fabian Arntz¹, Kathleen Golle¹, Florian Bähr², Toni Wöhl², Reinhold Kliegl¹

¹Universität Potsdam, Deutschland; ²Universität Erfurt, Deutschland

Theoretical background. Children's physical fitness is an important indicator of their physical health and cognitive function. The EMOTIKON project annually tests the physical fitness of all third-graders in the Federal State of Brandenburg, Germany. The project examines temporal trends, Covid-19 pandemic effects as well as associations of physical fitness with social structure. As social structure is associated with children's health and development (Stojan et al., 2023; Wessely et al., 2023), we expected (1) better physical fitness in schools with lower social burden (higher SES). Based on earlier research (Bähr et al., 2024; Stojan et al., 2023; Teich et al., 2023) showing more pronounced Covid pandemic effects on children's fitness in relatively "fitter" schools that had "more fitness to lose", we expected (2) more negative pandemic effects in schools with lower social burden. Further, since access to sports clubs was restored after the pandemic and as children in more affluent areas likely have more sports opportunities, (3) potential rebound effects were hypothesized to be more pronounced in schools with lower social burden. *Methods.* Tests assessed cardiorespiratory endurance (6-min-run), coordination (star-run), speed (20-m linear sprint), lower (powerLOW, standing long jump), and upper (powerUP, ball-push) limbs muscle power, and static balance (one-legged stance). Social burden was indicated by a school-specific social index provided by the Brandenburg Ministry of Education, Youth and Sport (MBJS, 2024) based on social welfare rate, proportion of students with non-German primary language, and proportion of students with special educational needs. 108,308 third-graders between 8 and 9.2 years from 444 schools were tested in the falls between 2016–2023. Linear mixed models, specified for a regression discontinuity design (Lee & Lemieux, 2010; Thistlethwaite & Campbell, 1960) with random factors child and school, tested pandemic effects at the first day of school in the school year 2020/21 (i.e., critical date) and differences in linear and quadratic cohort trends before and after the pandemic onset. School social burden interacted with Covid pandemic effects and cohort trends. *Results.* Higher school social burden was associated with lower performance in tests of cardiorespiratory endurance, coordination, speed, and powerLOW. At the critical date, there were small negative pandemic effects in cardiorespiratory endurance, coordination, speed, and powerUP. Pandemic effects in coordination and speed were more pronounced in schools with lower social burden. Coordination and powerUP rebounded during 2020–2023, with more pronounced coordination rebounds in schools with lower social burden. There was no evidence for rebounds of cardiorespiratory endurance and speed. *Conclusion.* Lack of evidence of post-pandemic rebounds may indicate long-term consequences of pandemic-related movement restrictions. Results show the need to enhance children's physical fitness in disadvantaged regions.

PS-29: Sprachliche Bildung und Mehrsprachigkeit: Grundlagen und Förderung

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 382

Chair der Sitzung: Patrick Dahdah

Vergleich der Entwicklung des Zugriffs auf die Bedeutung geschriebener und gesprochener Wörter bei Kindern mit Deutsch als Erst- und Zweitsprache

Patrick Dahdah¹, Johannes Naumann¹, Maj-Britt Haffmans², Julia Schindler³, Tobias Richter⁴

¹Bergische Universität Wuppertal; ²Universität Kassel; ³Freie Universität Berlin; ⁴Julius-Maximilians-Universität Würzburg

Erfolgreiches Leseverstehen hängt laut der *lexical quality hypothesis* von der Qualität lexikalischer Repräsentationen ab (Perfetti & Hart, 2001). Vor allem der effiziente Zugriff auf Wortbedeutungen ist dabei zentral (Nouwens et al., 2018; Richter et al., 2013). Bisher wurde der Zugriff auf Wortbedeutungen ausschließlich anhand schriftlich dargebotener Wörter erfasst, weshalb eine Konfundierung mit Dekodierfähigkeiten zu vermuten ist und differenzielle Entwicklungsverläufe des Zugriffs auf Wortbedeutungen bei gesprochener und geschriebener Sprache bisher nicht untersucht sind. Ebenso ist unbekannt, in welchem Verhältnis das Dekodieren von Worten und der Zugriff auf Wortbedeutungen bei der Vorhersage von Lesekompetenz spielen.

Fragestellung

Die aktuelle Studie vergleicht den Entwicklungsverlauf des Zugriffs auf Wortbedeutungen bei geschriebenen und gesprochenen Wörtern. Da bei Kindern mit und ohne Deutsch als Erstsprache sowohl im Lesen als auch im Hören unterschiedliche Entwicklungsverläufe zu vermuten sind, differenzieren wir zwischen Kindern mit Deutsch als Erst- (L1) und Zweitsprache (L2). Wir prüfen, inwieweit diese Entwicklungsverläufe das Leseverstehen am Ende der Grundschulzeit vorhersagen.

Methode

Die Stichprobe bestand aus insgesamt 1120 Grundschulkindern, die zu fünf Zeitpunkten teilnahmen (Beginn 1. Klasse-Ende 4. Klasse). Zur Erfassung des Zugriffs auf Wortbedeutungen wurden semantische Kategorisierungsaufgaben aus ProDi-L (Richter et al., 2017) und dem auditiven Äquivalent ProDi-H verwendet. Dabei sollten die Kinder entscheiden, ob ein Wort (z.B. Banane) einer zuvor genannten Kategorie (z.B. Obst) angehört. Die auditiven Items wurden zu allen fünf Zeitpunkten, die visuellen Items erst ab dem zweiten Zeitpunkt (Ende der ersten Klasse) vorgelegt. Geschwindigkeit und Akkuratheit wurden durch das Aufsummieren der reziproken Reaktionszeiten für alle richtig beantworteten Items integriert. Das Leseverstehen wurde mit dem Untertest *Textverständnis* aus ELFE 1-6 (Lenhard & Schneider, 2006) gemessen. Als Kontrollvariablen zu t_1 wurden nonverbale Intelligenz, phonologische Bewusstheit sowie visuelle und auditive Worterkennung erhoben.

Ergebnisse

Pro Modalität wurden zwei latente Mehrgruppen-Wachstumskurvenmodelle geschätzt. Im ersten Schritt wurden jeweils nur die Entwicklungsverläufe geschätzt. Im zweiten Schritt wurde das Leseverstehen am Ende von Klasse 4 als abhängige Variable, sowie die zu t_1 erhobenen Kontrollvariablen hinzugefügt. Effekte von Intercept und Slope auf die Lesefähigkeit in Klasse 4 wurden unter Kontrolle der Lesefähigkeit in Klasse 2 geschätzt. Das Modell ohne Kovariaten hatte für die geschriebenen Wörter einen adäquaten (CFI=.935, RMSEA=.074) und für die gesprochenen Wörter einen guten Fit (CFI=.952, RMSEA=.043). Die L1 Gruppe zeigte bei der jeweils ersten Messung sowohl für die geschriebenen Wörter ($\alpha_{\text{Intercept-L1}}=367.17$; $\alpha_{\text{Intercept-L2}}=335.11$) als auch für die gesprochenen Wörter ($\alpha_{\text{Intercept-L1}}=302.88$; $\alpha_{\text{Intercept-L2}}=274.45$) einen effizienteren Zugriff auf Wortbedeutungen als die L2-Gruppe. Die L1-Kinder zeigten zudem in beiden Modalitäten eine nichtlineare Entwicklung, mit stärkeren Zuwächsen zu Beginn der Messungen, die im Verlauf abflachten ($\alpha_{\text{Slope-visuell-L1}}=45.16$; $\alpha_{\text{Quadratic-visuell-L1}}=-6.58$; $\alpha_{\text{Slope-auditiv-L1}}=57.58$; $\alpha_{\text{Quadratic-auditiv-L1}}=-5.11$). Die L2 Kinder zeigten dagegen einen langsameren, aber linearen Zuwachs bei gesprochenen Wörtern ($\alpha_{\text{Slope-auditiv-L2}}=42.02$) und eine beschleunigte Entwicklung bei geschriebenen Wörtern ($\alpha_{\text{Slope-visuell-L2}}=13.36$; $\alpha_{\text{Quadratic-visuell-L2}}=5.30$). Das erweiterte Modell für die geschriebenen (CFI=.970, RMSEA=.040) und für die gesprochenen Wörter (CFI=.956, RMSEA=.039) zeigten jeweils einen guten Fit. Das anfängliche Niveau des effizienten Zugriffs auf Wortbedeutungen sagte das Leseverstehen am Ende der vierten Klasse bei geschriebenen Wörtern ($\beta_{L1}=.75$; $\beta_{L2}=.75$) und bei gesprochenen Wörtern ($\beta_{L1}=.60$; $\beta_{L2}=.54$) vorher. Auch die Wachstumsrate sagte sowohl bei geschriebenen ($\beta_{L1}=.61$; $\beta_{L2}=.69$) als auch bei gesprochenen Wörtern ($\beta_{L1}=.33$; $\beta_{L2}=.56$) das Leseverstehen am Ende der Grundschulzeit vorher.

Diskussion

Vor allem bei Kindern mit Deutsch als Zweitsprache werden die Unterschiede in der Erfassung semantischer Repräsentationen anhand geschriebener und gesprochener Wörter deutlich. Bei geschriebenen Wörtern zeigte die L2-Gruppe eine verzögerte Entwicklung, die erst ab dem Ende der zweiten Klasse stark zunahm. Die Ergebnisse belegen deutlich differenzierte Entwicklungsverläufe des Zugriffs auf Wortbedeutungen bei gesprochener Sprache und Schriftsprache. Die Ergebnisse zeigen weiterhin, dass der effiziente Zugriff auf Wortbedeutungen sowohl hinsichtlich Niveau als auch hinsichtlich der Wachstumsrate unabhängig von der Dekodierfähigkeit einen zentralen Prädiktor für die Lesekompetenz am Ende der Grundschulzeit darstellt.

Evaluation einer sprachintegrierten Emotionsintervention hessischer Vorklassen: Ein Cross-Over Design

Miriam Naomi Menz¹, Christin Tekaath², Arianne Lydia Andreas¹, Miriam Hansen¹, Irene Corvacho del Toro²

¹Goethe-Universität Frankfurt am Main; ²Universität Siegen

Das BMBF geförderte interdisziplinäre Projekt SEM hat zum Ziel, Kompetenzen von Schüler*innen in Vorklassen in den Bereichen Sprache, Emotionen und Mathematik zu fördern. Vorklassen weisen häufig einen hohen Anteil an Kindern mit Migrationshintergrund auf. Die Sprach- und Emotionsförderung spielt dabei eine entscheidende Rolle, da Kinder mit elaboriertem Emotionswissen häufig höhere soziale Kompetenzen und einen höheren schulischen Erfolg zeigen (von Salisch & Wübker, 2021). Sprachliche Fähigkeiten nehmen auf verschiedene Art und Weise Einfluss auf die emotionale Entwicklung von Kindern, zum Beispiel auf das Emotionsverständnis (Hornback, 2010; Köckeritz et al., 2010; von Salisch & Wübker, 2021), den Ausdruck von eigenen Emotionen (Roben et al., 2013) sowie auf Emotionsregulationsstrategien (Gunst et al., 2014). In der Vorstudie zeigte sich, dass Kinder am Ende der Vorklasse im Emotionswissen unterdurchschnittlich Leistungen zeigen (Vorklasse: $M = 39,832$, $SD = 10,793$; Erste Klasse $M = 49,832$, $SD = 10,633$). Im Rahmen des Projektes wurden sprachintegrierte Fördermaterialien zum Themenbereich Emotion entwickelt, die insgesamt 12 Einheiten (à zwei Schulstunden) umfassen. Hierbei werden die Basisemotionen (Freude, Trauer, Angst, Wut, Überraschung, Ekel) sowie Emotionsausdruck und Emotionsregulationsstrategien vermittelt. Es fanden pro Kohorte zwei halbtägige Professionalisierungsmaßnahmen sowie Selbstlernphasen zu den entwickelten Materialien mit den Vorklassenleiter*innen statt. Die Intervention wurde als Cross-Over Design in 18 hessischen Vorklassen durchgeführt. Es fand eine Prä- sowie eine Posterhebung durch den Adaptiven Test des Emotionswissens (ATEM 3-9) in digitaler Form statt. Zudem wurden sprachliche Fähigkeiten erhoben (SET 5-10). Die Anwendung eines Cross-Over Designs ermöglicht

einen Vergleich der Prä- und Postdaten, zwischen Kindern mit und ohne Emotionsintervention und berücksichtigt potentielle Reihenfolgeeffekte. Diese Daten und Ergebnisse sollen in Form eines Vortrags präsentiert werden. Die vorliegende Studie bietet neue Erkenntnisse zur Wirksamkeit gezielter Emotionsförderung bei Vorklassen in Hessen.

Der Übergang von bilingualen Programmen in der Grundschule zu regulärem Fremdsprachenunterricht in der Sekundarschule mit Fokus auf mehrsprachigen Kindern

Anja Steinlen, Thorsten Piske

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

In Deutschland wird der Übergang von der Grundschule zur Sekundarstufe als besonders herausfordernd für Fremdsprachenlernende angesehen. Dies liegt an den veränderten Unterrichts- und Bewertungspraktiken sowie an der Verlagerung des Fokus von der Mündlichkeit zur Schriftlichkeit, die die langfristige Entwicklung der Fremdsprachenkompetenz von Schülerinnen und Schülern (SuS) beeinträchtigen können (z.B. Jaekel et al., 2021). In dieser Studie werden die sprachliche Entwicklung und die Einstellungen einer bestimmten Schülergruppe untersucht, nämlich SuS, die in der Grundschule an deutsch-englischen bilingualen Programmen teilgenommen haben und ab der fünften Klasse im Sekundarbereich regulären Englischunterricht erhielten. Das Hauptaugenmerk liegt auf mehrsprachigen SuS, die in der Regel Deutsch als L2 und Englisch als die erste in der Schule unterrichtete Fremdsprache (L3) lernen. Diese Kinder verdienen besondere Aufmerksamkeit, da SuS aus einem „ethnisch und sprachlich vielfältigen Hintergrund [...]“ oftmals als Risikogruppe betrachtet werden (Genesee & Fortune, 2014, S. 196).

Gibt es Unterschiede in den Einstellungen der SuS zum bilingualen Lernen in der Grundschule und zum Englischunterricht in der Sekundarschule, abhängig davon, ob ihre Erstsprache Deutsch ist oder nicht? Gibt es Unterschiede zwischen SuS mit Deutsch als L1 und L2 in Bezug auf ihre Leistungen in Englischtests in der Sekundarschule?

Die Stichprobe umfasst 500 Fünftklässler und Fünftklässlerinnen aus mehr als 150 Sekundarschulen (14% Mittelschulen, 23% Realschulen, 63% Gymnasien) in Bayern, von denen 30% einen mehrsprachigen Hintergrund haben. Die SuS füllten einen Fragebogen zu ihren Einstellungen sowohl zum bilingualen Lernen in der Grundschule als auch zum Englischunterricht in der Sekundarschule aus. Zusätzlich wurde eine Stichprobe von 100 SuS (von denen ebenfalls 30% einen mehrsprachigen Hintergrund hatten) in Bezug auf das englische Leseverstehen sowie englische das Schreiben mithilfe des Primary School Assessment Kit (Little et al., 2003) untersucht.

Vorläufige Ergebnisse zeigen, dass die SuS, unabhängig von ihrem sprachlichen Hintergrund, sowohl das bilinguale Programm in der Grundschule als auch den Englischunterricht in der Sekundarschule sehr positiv bewerteten, und zwar unabhängig von der derzeit besuchten Schulart. Zudem erzielten ein- und mehrsprachige SuS in den drei Schularten Mittelschule, Realschule und Gymnasium keine signifikant unterschiedliche Ergebnisse in den Englischtests.

Diese Befunde stützen die Annahme, dass SuS aus sprachlich vielfältigen Hintergründen im Fremdsprachenunterricht nicht generell eine Risikogruppe darstellen, trotzdem Englisch ihre L3 ist (z.B. Steinlen, 2021). Diese Ergebnisse werden im Kontext verschiedener Modelle zum Transfer in der L3-Erwerbsforschung diskutiert und die praktischen Implikationen der Ergebnisse für den Fremdsprachenunterricht in der weiterführenden Schule erörtert.

Zielgruppen, Merkmale und Effektivität früher Grammatikinterventionen: Ergebnisse eines systematischen Literatur-Reviews

Svenja Wehrhöfer¹, Leonie Dargiewicz¹, Annika Ohle-Peters¹, Fani Lauermann², Nele McElvany¹

¹TU Dortmund; ²TU Dortmund, Rheinische Friedrich-Wilhelms-Universität Bonn

Bildungserfolg und gesellschaftliche Teilhabe hängen eng mit den Sprachkompetenzen Lernender zusammen (u. a. Dale et al., 2023). Die Grammatik stellt in diesem Zusammenhang eine zentrale Sprachkompetenz dar, da sie allen vier Basiskomponenten von Sprache – Hören, Sprechen, Lesen, Schreiben – inhärent ist (Jude et al., 2008). Insbesondere frühe grammatische Kompetenzen gelten dabei als wichtige Prädiktoren für das spätere Leseverstehen von Lernenden (u. a. Hjetland et al., 2020). Es bestehen bereits früh Diskrepanzen in Bezug auf grammatische Kompetenzen, wobei vor allem Zweitsprachenlernende (u. a. Bratlie et al., 2022), Lernende mit einem niedrigen sozioökonomischen Status (u. a. Levine et al., 2020) und Lernende mit Sprachentwicklungsstörungen (u. a. Delage & Frauenfelder, 2020) über geringere grammatische Kompetenzen verfügen als ihre Peers. Dennoch finden in vorschulischen Institutionen wenige sprachanregende Lehrkraft-Kind-Interaktionen sowie spezifische Sprachlernaktivitäten statt (u. a. Bratsch-Hines et al., 2019), die den Erwerb grammatischer Kompetenzen vor Beginn der Grundschulzeit unterstützen könnten. Es besteht daher die Notwendigkeit systematischer Interventionen im Bereich der Grammatik sowohl in vorschulischen Institutionen als auch in der frühen Grundschulzeit. Bisher ist jedoch wenig über die Zielgruppen, Gestaltungsmöglichkeiten und die Wirksamkeit institutioneller Grammatikinterventionen für junge Lernende im Übergang zur formalen Bildung bekannt. Verschiedene Meta-Analysen und systematische Literatur-Reviews belegen zwar die Wirksamkeit von Grammatikinterventionen (u. a. Spada & Tomita, 2010), aber sie berücksichtigen oft nur spezifische Zielgruppen wie z. B. Kinder mit Sprachentwicklungsstörungen (u. a. Cleave et al., 2015) oder umfassen große Altersspannen (Kang et al., 2019). Zudem werden in vielen Fällen Grammatikinterventionen nicht isoliert betrachtet (u. a. Silverman et al., 2020) oder die inkludierten Primärstudien finden nicht (ausschließlich) im Bildungskontext statt (u. a. Norris & Ortega, 2000). Folglich liegen bisher kaum Erkenntnisse im Hinblick auf die Zielgruppen, die Interventionsgestaltung und die Wirksamkeit von institutionellen Grammatikinterventionen für junge Lernende im Übergang zur formalen Bildung vor. Demnach ist es das Ziel dieses systematischen Literatur-Reviews, vorliegende Interventionsstudien zur frühen institutionellen Grammatikförderung zu synthetisieren und zielgerichtet zu beschreiben. Dabei sollen folgende Forschungsfragen beantwortet werden: 1. Welche sind die häufigsten Zielgruppen von Grammatikinterventionen in vorschulischen Institutionen und in der frühen Grundschulzeit? 2. Was sind die Charakteristika der Grammatikinterventionen in Bezug auf die grammatischen Zielformen und die Interventionsgestaltung? 3. Verbessern die Grammatikinterventionen effektiv die grammatischen Kompetenzen der Kinder? Unter Berücksichtigung der PRISMA-Richtlinie (Page et al., 2021) wurde eine systematische Literatursuche in einschlägigen pädagogisch-psychologischen Datenbanken durchgeführt (Academic Search Ultimate, APA PsycArticles, APA PsycInfo, Education Source, ERIC, FIS Bildung, PSYNDEX, Web of Science). Darüber hinaus erfolgte eine manuelle Suche. Interventionsstudien mit einem (quasi-)experimentellen Design wurden berücksichtigt, wenn sie u.a. eine systematische Grammatikintervention für Kinder zwischen 4;0 und 7;11 Jahren in vorschulischen oder schulischen Institutionen beinhalteten, wobei mindestens eine abhängige Variable eine grammatische Teilkompetenz erfassen musste. Von den k = 2999 identifizierten Publikationen erfüllten k = 22 Studien die Inklusionskriterien (N = 2920 Kinder). Das systematische Review zeigt, dass überwiegend Zweitsprachenlernende, Kinder mit einem niedrigen sozioökonomischen Status und Lernende mit schwachen

sprachlichen Leistungen (keine Sprachentwicklungsstörungen) die Zielgruppen früher institutioneller Grammatikinterventionen darstellen. In Bezug auf die grammatischen Zielformen kann festgestellt werden, dass syntaktische Zielstrukturen häufiger gefördert werden als Zielformen im Bereich der Flexion, wobei ein Zusammenhang mit der jeweiligen Sprachtypologie erkennbar ist. Erste Ergebnisse zur Effektivität der Interventionsmaßnahmen zeigen, dass in 82% der Studien positive Effekte für mindestens eine grammatikbezogene abhängige Variable festgestellt wurden. Gleichzeitig verfügen 41% der Studien über mindestens eine Grammatikvariable, die einen nichtsignifikanten oder negativen Effekt der Intervention aufzeigt. Mögliche Gründe für die heterogenen Befunde sowie Ergebnisse zu den Charakteristika der Interventionen in Bezug auf die Interventionsgestaltung werden aktuell analysiert und im Rahmen des Vortrags präsentiert. Ferner werden Implikationen für die empirische pädagogische Forschung sowie die Praxis vorgestellt.

PS-30: Identität, Diskriminierung und interkulturelles Lernen im Schulkontext

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 157

Chair der Sitzung: Sharleen Pevec-Zimmer

Interethnic Contact and Perceived Ethnic Discrimination

Florian Weber¹, Heike Krüger^{2,3}

¹Leibniz-Institut für Bildungsverläufe, Deutschland; ²Institut für Soziologie und Sozialpsychologie, Universität zu Köln, Deutschland; ³Institut für Soziologie, RWTH Aachen, Deutschland

The discrimination paradox describes the empirical finding that reports of discrimination are more prevalent among structurally better-integrated migrants, for example, those who have obtained a high-level education qualification (Tuppat & Gerhards, 2021). Theoretical approaches suspect two general underlying mechanisms of the integration paradox: *sensitivity* and *exposure*. Diehl et al. (2021) test these mechanisms against each other and conclude that sensitivity, rather than exposure, is the central mechanism at play. However, their measure of exposure is the share of Germans among friends, which only covers *voluntary* contact to the ethnic majority. Furthermore, their model (even though a longitudinal hybrid model is used) does not capture the direction of the relationship: social contact to the ethnic majority might increase contact to possible discriminators and thus lead to higher levels of perceived discrimination. On the other hand, migrants might avoid contact to the ethnic majority if they feel discriminated by them (Kornienko et al., 2023), which should be especially true for voluntary contact. In this case, a negative relationship between perceived discrimination and contact to the majority population is to be expected. However, such a negative relationship can also be expected when positive interethnic contact – e.g. friendships – reduce feelings of discrimination. *Involuntary* contact should be less sensitive to feelings of discrimination, while it increases exposure and should therefore influence reports of discrimination (although this could not be confirmed by Steinmann [2019]). To analyse the causal direction and the mechanism at play for both of these concepts, longitudinal data are needed. Furthermore, to measure involuntary contact to the ethnic majority, context measures are needed. School and the residential neighborhood are two of the most important social spaces for adolescents, and the share of non-migrants in these contexts offer valuable insights regarding involuntary contact to the majority population. Thus, we want to answer the following research questions: 1. How do voluntary and involuntary contact to the ethnic majority relate to perceived discrimination, and 2. What is the causal direction of these relationships?

In this contribution, we will use data from the project “Friendship and Violence in Adolescence” (German: “Freundschaft und Gewalt im Jugendalter”, FuGJ, Kroneberg et al. [2016]). This panel dataset contains survey data as well as four waves of complete social networks on the grade level in multiple schools in the Ruhr area, assuring a high share of interethnic contacts. More recently, the subsequent project JuKo has enriched this dataset with geospatial data for pupils’ place of residence and the school environment. As *Gymnasiums* were only surveyed in the last two panel waves, we will focus on these waves to ensure a balanced panel dataset.

In order to respond to our research question with the data in question, we will employ stochastic actor-oriented models. By using RSiena co-evolution models (Snijders et al., 2010), we will be able to estimate social selection mechanisms resulting from discriminatory experiences within peer networks, as well as the social influence of inter-ethnic friendships on the perception of discrimination, while controlling for the respective other process. This will enable us to analyse the causal relationship, taking contextual variations in the respondents’ place of residence and school neighbourhood into account. Results will be presented on the conference.

Jüdische und antisemitismuskritische Bildung im Religionsunterricht!?

Anna Körs

Universität Hamburg, Deutschland

Angeichts eines wachsenden Antisemitismus werden an den Religionsunterricht in staatlichen Schulen in Deutschland – als das einzige grundgesetzlich abgesicherte Lehrfach und mit expliziter Ausrichtung auf Werthaltungen und moralische Einstellungen – hohe bildungs- und gesellschaftspolitische Erwartungen gestellt, durch Wissensvermittlung über das Judentum und jüdisches Leben auch zur Prävention von Antisemitismus beizutragen. In der in 2022 beschlossenen „Nationalen Strategie gegen Antisemitismus und für jüdisches Leben“ (NASAS) als erste Strategie der Bundesregierung, die explizit die Bekämpfung von Antisemitismus und die Förderung jüdischen Lebens benennt, heißt es: „Insbesondere im Bildungsbereich, vor allem in der Schule als zentralem Ort des kognitiven und sozialen Lernens, sollten die jüdische Gegenwart und vor allem die jüngere jüdische Geschichte stärker als bisher thematisiert werden. (...) Fächer und Projekte der historisch politischen Bildung und des Religionsunterrichts bieten vielfältige Anknüpfungspunkte“ (Beauftragter der Bundesregierung für jüdisches Leben in Deutschland 2022, S.45). Dies gilt seit dem 7. Oktober verstärkt – was vorher kaum im Blick war – auch für die Bildung bereits im Kindesalter und den Religionsunterricht in der Grundschule

Dem Religionsunterricht wird damit eine wichtige Rolle zugeschrieben, nicht nur für den Umgang mit religiöser Vielfalt allgemein und für die Vermittlung des Judentums im Besonderen, sondern auch für die Prävention und Bearbeitung von Antisemitismus. Während die jüngere Antisemitismusforschung einige Erkenntnisse über ablehnende Haltungen von Jugendlichen gegenüber anderen Religionen und speziell gegenüber Juden erbracht hat, gibt es bisher keine Forschung dazu, ob und wie bereits im Kindesalter Stereotypisierungen und Zuschreibungen mit antisemitischem Hintergrund entwickelt werden und wie diesen entgegengewirkt werden kann. Ebenso unerforscht ist die Frage, wie Vorstellungen von religiöser Differenz im didaktisch angeleiteten Rahmen des Religionsunterrichts konstruiert und verhandelt werden.

Vor diesem Hintergrund beschäftigt sich der vorliegende Beitrag mit der grundlegenden Frage: Wie wird das Judentum im Religionsunterricht thematisiert und in der Unterrichtspraxis behandelt und welche Vorstellungen über Judentum und jüdische Menschen entstehen dadurch bereits im Kindesalter? Insgesamt wurde der Religionsunterricht in sieben Klassen an vier verschiedenen Grundschulen beobachtet, es wurden Interviews mit den dazugehörigen fünf Lehrkräften geführt und mit den beobachteten Klassen fanden insgesamt fünf Gruppendiskussionen statt. Zusätzlich wurden drei Expert:inneninterviews mit Religionslehrkräften geführt, deren Unterricht aus organisatorischen Gründen nicht beobachtet werden konnte. Die gesammelten Daten aus Beobachtungen, Interviews und Gruppendiskussionen wurden in Anlehnung an die Grounded Theory in einem mehrstufigen Kodierverfahren ausgewertet.

Im Beitrag werden drei sehr unterschiedliche Fälle der Behandlung des Judentums im Religionsunterricht in der Grundschule vorgestellt, die einerseits eine erhebliche Varianz und Bandbreite in der Unterrichtspraxis zum Thema Judentum aufzeigen und andererseits auf die Gefahr der Reproduktion von antisemitischen Stereotypen und einseitigen Geschichtsbildern verweisen. Diese Ergebnisse an der Schnittstelle von Antisemitismus- und Religionsunterrichtsforschung machen damit gleichsam die

Notwendigkeit deutlich, die Behandlung des Judentums im Religionsunterricht auf die Grundlage einer antisemitismuskritischen Bildung und entsprechender Forschung zu stellen. Während die vorliegende Studie die Unterrichtspraxis fokussiert und exemplarisch zeigt, dass Judentum, jüdisches Leben und Antisemitismus im Religionsunterricht in der Grundschule sehr unterschiedlich thematisiert werden und die Lehrkraft hierbei eine zentrale Rolle spielt, verweist sie zugleich auf das institutionelle Setting als ein wichtiges Desiderat, um das Potenzial des Religionsunterrichts für antisemitismuskritische Bildung erlassen und auch strukturell entwickeln zu können.

Identität(en) als Ressource für Jugendliche im Schulkontext – die Rolle von ethnischer und nationaler Identitätsentwicklung für die Förderung von Identitätskohärenz in kulturell vielfältigen Schulkontexten

Sharleen Pevec-Zimmer¹, Linda Juang¹, Maja Schachner², Tuğçe Aral¹

¹Universität Potsdam; ²Martin-Luther-Universität Halle-Wittenberg

Kulturelle Identitäten aller Schüler*innen einzubeziehen ist ein Kernaspekt pädagogischer Ansätze wie dem kultursensiblen Unterrichten. Mit dem Ziel, Bildungsgerechtigkeit zu fördern, werden gruppenbezogene Zugehörigkeiten und Ungerechtigkeiten reflektiert und Anknüpfungs- und Bezugspunkte für die Identitäten und Lebensrealitäten aller Schüler*innen im Unterricht geschaffen (Gay, 2018). US-basierte Studien zeigen, dass ethnische Identitäten (die familiären Herkunftskulturen betreffend) insbesondere für Schüler*innen mit Migrationsgeschichte relevante Aspekte ihrer Identität darstellen und ihr Wohlbefinden und schulische Leistungen begünstigen (Rivas-Drake et al., 2014). Studien in Deutschland unterstützen die Bedeutsamkeit sozialer Identitäten für Jugendliche: Die ethnische Identität kann Selbstbewusstsein und psychisches Wohlbefinden (Schotte et al., 2018) und die kulturelle Identifikation mit Deutschland die schulische Anpassung (Hannover et al., 2013) fördern. Für die Erforschung dieser sozialen Identitätsentwicklungen als potenzielle Ressourcen werden in Deutschland jedoch vorrangig Schüler*innen mit Migrationsgeschichte fokussiert (z.B. Spiegler et al., 2018). Schüler*innen ohne Migrationsgeschichte sowie nationale und ethnische Identitätsentwicklung werden bislang kaum einbezogen (Ausnahmen sind z.B. Leszczensky & Gräbs-Santiago, 2014; Spiegler et al., 2022). Identitätskohärenz, das Entwickeln eines zusammenhängenden Selbstverständnisses, ist ebenfalls relevant für die psychosoziale Anpassungsfähigkeit von Schüler*innen (Umaña-Taylor et al., 2014) und eine zentrale Entwicklungsaufgabe im Jugendalter (Erikson, 1968). Soziale Identitäten, wie die ethnische und nationale Identität, können für die Identitätskohärenz salient sein (Safa et al., 2024; Syed & Juang, 2014).

Als wichtiger Entwicklungskontext sollten Schulen Schüler*innen dabei unterstützen, sich mit der eigenen Identität auseinanderzusetzen (Verhoeven et al., 2019). Das *Identity Project*, eine forschungsbasierte US-Schulintervention (Umaña-Taylor & Douglass, 2017), wurde mit dem Ziel der Auseinandersetzung und mehr Klarheit Jugendlicher bezüglich ihrer ethnischen Identitäten konzipiert. Die Anpassung für europäische Kontexte (Juang et al., 2022) betrifft u.a. die Reflexion der ethnischen Identität sowie der nationalen Identität (kulturelle Zugehörigkeit zu Deutschland). Dies ist in Migrationsgesellschaften bedeutend, da beide Zugehörigkeitsfragen aufwerfen und prinzipiell zur Identitätsfindung von Schüler*innen beitragen können.

Wir untersuchen, ob (H1) das für deutsche Schulkontexte angepasste *Identitätsprojekt* die Identitätskohärenz zu T3 fördert, (H2) die ethnische und nationale Identität zu T2 die Effekte auf die Identitätskohärenz vermitteln und (H3) sowohl die direkten als auch die vermittelten Effekte des Identitätsprojekts moderiert werden durch den Status von Schüler*innen als kulturell majorisiert (deutsche familiäre Herkunftskultur) oder minorisiert (ausschließlich/zusätzlich andere familiäre Herkunftskultur).

Längsschnittdaten 574 Jugendlicher ($N_{\text{Intervention}}=289$, $M_{\text{age}}=12.86$, 55% mit Migrationsgeschichte; $N_{\text{Control}}=285$, $M_{\text{age}}=12.78$, 59% mit Migrationsgeschichte) basieren auf drei Messzeitpunkten: Pre-Test (T1), Post-Test (T2), Follow-up (T3). Die Entwicklung ethnischer und nationaler Identitäten wird jeweils mittels Exploration (Erkundung, 3 Items) und Resolution (Klarheit, 3 Items) erfasst (Leszczensky & Gräbs-Santiago, 2015; Umaña-Taylor et al., 2004). Identitätskohärenz wird mit der EPSI-Subskala zur Identitätsfindung erfasst (Rosenthal et al., 1981; 5 Items).

Strukturgleichungsmodelle wurden mit Mplus gerechnet. Aufgrund unterschiedlicher Faktorstrukturen nach explorativen Faktorenanalysen wurden separate Modelle für minorisierte und majorisierte Schüler*innen spezifiziert und vorige Messzeitpunkte der Skalen als Kovariaten integriert. Beide Modelle weisen gute Modellfits auf (Hu & Bentler, 1999). Die Ergebnisse zeigen keine signifikanten direkten Effekte der Intervention auf die Identitätskohärenz von Jugendlichen. Für majorisierte Schüler*innen vermitteln ethnische und nationale Identitätsentwicklung nicht den Zusammenhang zwischen Intervention und Identitätskohärenz zu T3. Für minorisierte Schüler*innen vermitteln die Klarheit bezüglich der ethnischen Identität zu T1 ($\beta=0.15$, $p=.02$) und der nationalen Identität zu T2 ($\beta=0.18$, $p=.01$) die Identitätskohärenz zu T3.

Die Studie vereint Theorien aus Sozial-/Entwicklungspsychologie und praktische (migrations)pädagogische Ansätze für die Schule. Die Ergebnisse tragen zu einem differenzierteren Verständnis der Entwicklung ethnischer und nationaler Identitäten sowie der Förderung von Identitätskohärenz in vielfältigen Schulkontexten in Deutschland bei. So sollten Schulen die Auseinandersetzung mit beiden Identitäten von Schüler*innen mit Migrationsgeschichte unterstützen, um positiv auf ihr zusammenhängendes Selbstverständnis einzuwirken. Für Schüler*innen ohne Migrationsgeschichte sind diese Identitäten weniger salient für das Selbstverständnis und ein pluralistisches Verständnis von Zugehörigkeit zu Deutschland kann möglicherweise herausfordernd sein. Es gilt zu diskutieren, wie unterschiedliche Schüler*innengruppen von Projekten zu kultureller Identität profitieren können.

Die Bedeutung sozialpsychologischer Einflussfaktoren auf die Einstellung von Schüler:innen gegenüber Muslim:innen im Kontext des interreligiösen Lernens

Janine Kuropka

TU Dortmund, Deutschland

Theoretischer Hintergrund

In einer zunehmend religiös-pluralen Gesellschaft nehmen Schule und insbesondere das interreligiöse Lernen eine wichtige Funktion ein, um Vorurteile und Stereotype gegenüber Mitgliedern unterschiedlicher Religionsgemeinschaften entgegenzuwirken und wechselseitiges Verständnis und Toleranz zu fördern (Baur 2007; Rothgangel 2013). Ein besonderer Fokus bei dem vorgestellten Projekt liegt auf Muslim:innen, eine Gruppe, die laut der alle zwei Jahre durchgeführten Mitte-Studie im Durchschnitt der Jahre 2014-2020/21 von etwa 17 % der nicht-muslimischen Bevölkerung in Deutschland offen angefeindet wurde (Zick 2021).

Anknüpfungspunkte zur Entstehung, Funktion und Messung von Einstellungen lassen sich in der Sozialpsychologie finden. Hier werden Einstellungen über drei Komponenten definiert: Sie haben (1.) ein Thema (das Objekt), (2.) sind urteilend oder bewertend (positiv oder negativ) und (3.) sind relativ lang anhaltend (Gergen und Gergen 1986). Einstellungen betreffen somit die „Bewertungen von Sachverhalten, Menschen, Gruppen und anderen Arten von Objekten unserer sozialen Welt“. Sie haben

insofern eine besondere Relevanz, als dass sie die Wahrnehmung und das Verhalten gegenüber der Welt beeinflussen (Haddock und Maio 2014).

Weiter wird in der Sozialpsychologie zwischen expliziten und impliziten Einstellungsmaßen unterschieden. Explizite Einstellungsmaße beruhen auf Selbstauskünften der befragten zu ihren Einstellungen gegenüber einem Objekt, während implizite Einstellungsmaße indirekt aus den Reaktionen auf das Einstellungsobjekt erschlossen werden (Haddock und Maio 2023). Durch die zusätzliche Erfassung impliziter Einstellungsmaße wird die Wahrscheinlichkeit verringert, dass die Ergebnisse durch sozial erwünschtes Antwortverhalten verzerrt werden (Fazio und Olson 2003).

Einstellungstheoretische Erkenntnisse aus der Sozialpsychologie werden in der Theorie des interreligiösen Lernens bislang jedoch kaum wahrgenommen oder systematisch rezipiert (Betz 2018; Unser 2022) stattdessen werden Theorien zur Einstellungsänderung relativ unabhängig von den Theorien zur Entstehung und Funktion von Vorurteilen entwickelt (Rothgangel 2013). Diese Forschungslücke ist bedeutsam, da grundlegende Erkenntnisse zur Entwicklung von Theorien und Maßnahmen zur Einstellungsänderung von hoher Relevanz sind.

Fragestellung

Um einen Beitrag zur Schließung der dargestellten Forschungslücke zu leisten, untersucht die vorliegende Studie die Einstellung von Schüler:innen gegenüber Muslim:innen. Zudem wird analysiert, welche sozialpsychologischen Faktoren diese Einstellungen beeinflussen und welche Implikationen sich daraus für das interreligiöse Lernen ergeben.

Methode

Im Rahmen des vorgestellten Projekts wird eine quantitativ-empirische Befragung unter Schüler:innen der 9. und 10. Klasse an Gesamtschulen in Nordrhein-Westfalen durchgeführt. Ziel ist es, ihre Einstellung gegenüber Muslim:innen zu erfassen und die sozialpsychologischen Einflussfaktoren zu identifizieren. Dies geschieht sowohl durch die Messung expliziter Einstellungsmaße und Stereotype als auch durch einen Impliziten Assoziationstest (IAT). Die bereits vorliegenden Pretest-Daten wurden im Zeitraum vom 11. bis 28. Juni 2024 von einer Stichprobe (n=138) erhoben. Auf der Tagung im Januar sollen die mit Hilfe einer stratifizierten Zufallsstichprobe gewonnenen Ergebnisse der Hauptuntersuchung vorgestellt werden.

Ergebnisse und ihre Bedeutung

Die Ergebnisse dieser Studie sollen zeigen, welche Einstellungen Jugendliche gegenüber Muslim:innen haben und wie verschiedene sozialpsychologische Einflussfaktoren diese Einstellungen prägen. Bereits in den Pretest-Daten ist deutlich zu erkennen, dass es Unterschiede zwischen expliziten und impliziten Einstellungsmaßen gibt. Eine differenzierte Analyse der Daten zeigt, dass 20,4 % der nicht-muslimischen Schüler:innen selbst angeben, eine negative Einstellung gegenüber Muslim:innen zu haben, während 44,7 % sich neutral positionieren und 34,9 % eine positive Einstellung bekunden. Bei der Analyse des IATs unter Verwendung des verbesserten D-Scores nach Greenwald et al. (2003) liegt der Mittelwert bei 0,21, was auf eine tendenziell negativere implizite Einstellung gegenüber Muslim:innen hinweist.

Durch die fundierte Analyse dieser und weiterer Aspekte können wertvolle Erkenntnisse gewonnen werden, die zur Verbesserung und gezielten Förderung des interreligiösen Lernens beitragen. Die Ergebnisse dieser Studie liefern nicht nur wesentliche theoretische Implikationen für die Sozialpsychologie und Religionspädagogik, sondern auch konkrete praktische Empfehlungen für den schulischen Unterricht, um Toleranz und Verständnis gegenüber religiösen Minderheiten zu stärken. Auf der Tagung werden die detaillierten Ergebnisse der multivariaten Analysen sowie ihre Bedeutung ausführlich präsentiert und diskutiert.

PS-31: Language and Multilingualism in Education

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 154

Chair der Sitzung: Kerstin Göbel

English Track

How a more flexible language offer contributes to reducing educational inequalities in Luxembourg's highly diverse and multilingual student population: First results from the evaluation of two recent policy reforms in the field of education

Joanne Colling, Caroline Hornung, Axel Grund, Pascale Esch, Ulrich Keller, Antoine Fischbach, Sonja Ugen

Luxembourg Centre for Educational Testing, University of Luxembourg, Luxembourg

With 49.3% of foreign-born inhabitants, Luxembourg is a highly diverse country and a forerunner regarding the sociocultural and linguistic composition of its population. This diversity is reflected in the education system, where approximately two thirds of students speak another language than Luxembourgish at home (SCRIPT & MENJE, 2023). These students are particularly at risk of struggling academically in Luxembourg's multilingual education system (Boehm et al., 2016; Hornung et al., 2021), where Luxembourgish functions as integration language in preschool, before German becomes the language of literacy acquisition in Grade 1, with French introduced as additional language early on.

To deal more adequately with the country's language diversity and counter educational inequalities that are assumed to result (at least partially) from a curriculum characterized by high language expectations (Hornung et al., 2023), the Luxembourgish government has recently introduced European Public Schools (EPS) and a project offering literacy acquisition in French within Luxembourgish Public Schools (LPS) to broaden the language offer within the education system.

Whereas EPS allow students to choose a main instruction language in which they mainly pursue their education (German, French, English), the literacy acquisition project reverses the traditional order of language use in primary school (literacy acquisition in German, French as additional language) by giving 1st graders the choice to start literacy acquisition in French with German as additional language.

By including EPS and the literacy project classes in Luxembourg's well-established School Monitoring Programme, the full-cohort data of approximately $N = 29.100$ students from five grades (comparison between EPS and LPS) and of $N = 5937$ 1st graders (comparison between the literacy project and students following the regular curriculum) from 2023 allows to analyze whether the diversification of the school offer contributes to reducing the observed educational inequalities in Luxembourg.

In mathematics, EPS students performed better than their peers in LPS across all assessed primary school grades. At secondary school level, EPS students performed better than students attending the intermediary and the lowest school track, while staying below the performance of those in the highest track. In addition, at-risk groups (e.g., Portuguese-speaking students) performed better when attending EPS compared to students with similar background characteristics in LPS.

In the literacy project, results indicate that students learning to read and write in French performed better in achievement tests assessing their language of literacy acquisition (e.g., skills in French listening comprehension) than students from their reference group created via propensity score matching and consisting of students with comparable background characteristics pursuing literacy acquisition in German (e.g., skills in German listening comprehension). In addition, they expressed a higher domain-specific academic motivation (e.g., enjoyment of reading in French) than the reference group and their parents perceived themselves more able to support their child academically due to their own language skills in French than parents from the reference group learning to read and write in German.

Despite methodological limitations (e.g., small sample sizes of specific student groups, test comparability), our results provide a first indication that a language offer allowing students to learn in a language closer to their home language might contribute to reducing educational inequalities in diverse education settings. By discussing advantages and disadvantages of the two approaches (e.g., broader language offer via different language sections at all grade levels in EPS with restricted spots available vs. possibility of a French literacy acquisition in Grade 1 which might be extended to all LPS at the national level), the study is highlighting how a flexible educational offer tailored to the needs of a diverse student population could play a crucial role in tackling important societal challenges (e.g., school dropouts, unemployment, health issues).

Multilingual Information Intervention in School Choice - Evidence from Immediate Acceptance in Bremen

Kevin Jan Zuchanek

Wuppertaler Institut für bildungsökonomische Forschung (WIB) & Bergische Universität Wuppertal (BUW)

Abstract

School matching is a much-discussed topic in school policy and the public debate. Parents would like to enrol children in their preferred schools, but excess demand for some schools requires a mechanism to allocate students to schools. One possible solution is to employ a matching algorithm, like the deferred or immediate acceptance algorithm (IA). One well-known characteristic of IA is that stating the true preferences for schools is not a dominant strategy (Abdulkadiroğlu & Sönmez, 2003). Thus, if parental decisions are based on insufficient information about the working of the mechanism, this might lead to costly mistakes.

Previous matching literature highlights the impact of strategic mistakes, the value of information, and information interventions. Kapor et al. (2023) show that subjective beliefs about admission chances often differ from actual probabilities, with demographic characteristics influencing this disparity. De Haan et al. (2023) find that disadvantaged students benefit most from switching from immediate acceptance to the strategy-proof deferred acceptance. Chen & He (2021) conclude that these two assignments provide different degrees of incentives to acquire information. Studies by Ajayi et al. (2020), Corcoran et al. (2018), and Cohodes et al. (2022) indicate that information interventions lead to increased selectivity and have a significant impact on migrant students particularly. However, previous studies have not evaluated the impact of providing information in multiple languages as a mean to close information deficits.

This study assesses a multilingual information intervention in the city of Bremen within the context of IA, where understanding the assignment mechanism is crucial for strategic school choice (Dur et al., 2018). In Bremen, the IA has been in place since 2011. The strategic feature of IA is explicitly communicated to parents in written information material, initially only in German. However, due to the high share of migrant students, not all parents may fully understand the school choice mechanism if

information is only available in German. Consequently, German students might outperform migrant students regarding admission to preferred schools.

To address this potential shortcoming, the Bremen school board translated the school choice brochure into six languages—Arabic, English, French, Russian, Persian, and Turkish—before the 2017/18 school year started, covering 21% of all students in terms of language spoken at home (treated). However, 13% of all students who speak other languages at home remain without a translation (control). This paper examines how providing multilingual information affects parental school choice, utilizing administrative school statistics and data on parental school choice at the student level. This allows for an analysis of each student's trajectory in public and private schools in Bremen, alongside the usage of individual information such as gender, citizenship, language, performance, and eligibility for social benefits.

Using a difference-in-differences approach, the study finds that treated students are more likely to choose and attend academic schools. Thus, the multilingual information supported parents' understanding of the school allocation mechanism. Treated students are 2.9 percentage points (p.p.) more likely to rank an academic school first and have a 3.1 p.p. higher probability of attending an academic school. Notably, the benefits are not uniformly distributed among treated students. For academic school attendance, the results are driven by high-performing students (7.7 p.p.), female students (3.9 p.p.), and low-income students (5.7 p.p.). Treated students also exhibit increased selectivity, opting for schools with more favourable social compositions, primarily ranking and attending schools with a lower share of students from low-income families and a lower social index. This suggests that the translation intervention reduces information deficits, corrects subjective admission probabilities, and leads to more successful school choices by parents.

Determinants of Language of Instruction Policies: A Global Analysis of Non-Dominant, Imperial, and English Languages

Helen Seitzer

Universität Bremen, Deutschland

Sprache ist ein wesentlicher Bestandteil der menschlichen Kultur, wobei Bildung eine entscheidende Rolle bei der Weitergabe von ebendieser Kultur und Tradition spielt. Laut bisheriger Forschung werden nur etwa 40% der weltweit gesprochenen Sprachen als Unterrichtssprache eingesetzt (Benson, 2016), was dazu führt, dass etwa die Hälfte aller SchülerInnen nicht in einer Sprache unterrichtet wird, die Sie verstehen und fließend Sprechen (UNESCO, 2016). Unterrichtssprachen sind damit ein kritischer Ansatzpunkt zur Verbesserung von Gleichberechtigung in der Bildung.

Diese Studie untersucht die vielfältigen Faktoren, die die Einführung verschiedener Sprachen als Unterrichtsmittel in Bildungssystemen weltweit beeinflussen. Durch die Untersuchung soziopolitischer und historischer Variablen wird die komplexe Dynamik sprachpolitischer Entscheidungen beleuchtet. Anhand eines Datensatzes, der 171 Länder von 1804 bis 2023 umfasst, werden die Wechselwirkungen zwischen diesen Variablen und der Einführung nicht-dominanter Sprachen, Kolonialsprachen sowie Englisch in Schulen erfasst.

Die Studie adressiert vier Forschungsfragen: Erstens, welche soziopolitischen und historischen Faktoren die Einführung nicht-dominanter Sprachen als Unterrichtsmittel beeinflussen. Zweitens, wie sich sprachliche Vielfalt, demokratische Regierungsführung und koloniales Erbe auf die Einführung von Kolonialsprachen nach der Dekolonialisierung auswirken. Drittens, welche Länder Kolonialsprachen wieder aus dem Unterricht verbannen. Viertens, welche Aspekte der Globalisierung die Einführung von Englisch als Unterrichtssprache beeinflussen.

Die theoretische Grundlage dieser Studie liegt im Schnittpunkt von Soziolinguistik, Politikwissenschaft und postkolonialen Studien. Sie stützt sich auf die Dekolonisierungstheorie nach Fanon (Brutt-Griffler, 2002; Fanon, 2002), um die Einführung sowie die Abschaffung von Kolonialsprachen zu erklären. Historischer Institutionalismus reicht nicht aus, um zu erklären, warum Länder nach der Dekolonialisierung weiterhin Kolonialsprachen im Unterricht verwenden.

Globalisierungstheorien bieten hingegen Aufschluss darüber, warum Englisch in Ländern ohne historische Verbindung zum Britischen Weltreich als Unterrichtssprache verwendet wird. Unterrichtssprache ist zudem in die Debatte um Minderheitenrechte eingebunden. Multikulturalismus (Taylor et al., 1994) bietet Erklärungen dafür, warum Länder, in denen Gleichberechtigung eine Rolle spielt, eher Unterricht in Minderheitssprachen zulassen.

Bisherige Forschung hat sich ausgiebig mit der Implementierung (Makoni et al., 2012), Durchführung (Sah & Li, 2022) und Effektivität (Collier & Thomas, 2017) verschiedener Unterrichtssprachen weltweit beschäftigt, jedoch meist mit Fokus auf einzelne Länder oder Weltregionen. Im Gegensatz dazu betrachtet diese Studie die globale Dynamik dieser Thematik, eine Perspektive, die bisher kaum beleuchtet wurde.

Unter Verwendung von Ereignisanalysemodellen (Cox proportional hazard models) und Modellen für binäre Daten (Complementary Log-log Regressionen) testet die Studie politische, historische und ökonomische Faktoren sowie die Auswirkung verschiedener kolonialer Administrationen auf die aktuelle Einführung von Unterrichtssprachen. Diese Methoden erfassen die komplexen Beziehungen zwischen den Variablen und dem Zeitverlauf und gewährleisten die Robustheit der Ergebnisse. Der Datensatz umfasst Variablen zur sprachlichen Vielfalt, politischen Regimen, Gleichberechtigung, Kolonialgeschichte und bewaffneten Konflikten, was einen umfassenden Überblick über die Faktoren bietet, die die Sprachenpolitik beeinflussen. Die Analyse zeigt, dass sprachliche, politische, wirtschaftliche und historische Faktoren die Einführung der Unterrichtssprache beeinflussen. Größere Sprachenvielfalt führt nicht durchgängig zur Einführung nicht-dominanter Sprachen, während demokratische Länder mit größerer Wahrscheinlichkeit nicht-dominante Sprachen einführen. Die Einführung von Kolonialsprachen wird hauptsächlich von der jeweiligen Kolonialadministration beeinflusst; ehemalige britische, französische und portugiesische Kolonien haben ein höheres Risiko, diese Sprachen beizubehalten. Globalisierungsdynamiken, besonders die politische und zwischenmenschliche Globalisierung, wirken sich positiv auf die Einführung von Englisch im Unterricht aus.

Diese Untersuchung verdeutlicht das komplizierte Geflecht von Faktoren, die die Einführung nicht-dominanter Sprachen im Bildungswesen beeinflussen. Das Zusammenspiel von sprachlicher Vielfalt, demokratischer Regierungsführung, rechtlichem Schutz, historischem Erbe und geopolitischen Kontexten schafft ein komplexes Umfeld für sprachpolitische Entscheidungen. Die Ergebnisse unterstreichen die Bedeutung eines ganzheitlichen Ansatzes in der Sprachenpolitik, der die vielfältigen soziopolitischen und historischen Einflüsse berücksichtigt. Diese Studie ist ein wichtiges Hilfsmittel für politische Entscheidungsträger, Pädagogen und Wissenschaftler, die die Komplexität der Sprachenpolitik in unterschiedlichen und sich entwickelnden globalen Kontexten verstehen und angehen wollen.

Chancen der Mehrsprachigkeitsorientierung im Französischunterricht

Anerkennung und Wertschätzung der sprachlichen Vielfalt sowie der unterschiedlichen kulturellen Identitäten ist ein wichtiges Bildungsziel für die Entwicklung unserer Gesellschaften (z.B. Europarat, 2020). Die kulturellen und sprachlichen Erfahrungen der Schülerinnen und Schüler sollten daher als wertvolle Ressourcen für das Lernen betrachtet werden. Didaktische Überlegungen zum Fremdsprachenunterricht fordern eine stärkere Anerkennung von gesellschaftlicher Mehrsprachigkeit und individueller Mehrsprachigkeit (Cenoz & Gorter, 2021; Gay, 2015). Da die Familiensprachen von Schülerinnen und Schülern mit Migrationshintergrund oft mit geringem Prestige verbunden sind, könnte die Aufnahme dieser Sprachen in den Unterricht dazu beitragen, dass der Wert dieser Sprachen deutlicher wahrgenommen wird, was das Selbstwertgefühl, das Sprachbewusstsein und das Sprachenlernen der Lernenden steigern könnte (Gibson & Hufeisen, 2007). Lernende, die das Gefühl haben, dass ihre sprachlichen Kompetenzen (ob Familien- oder Schulsprachen) in mehrsprachigen Unterrichtsaktivitäten angesprochen werden, können sich kompetenter fühlen, diese Aktivitäten genießen und positiv bewerten (Deci & Ryan, 2000). Mehrsprachig aufgewachsene Schülerinnen und Schüler mit Migrationshintergrund sind in deutschen Schulen jedoch häufig mit negativen Einstellungen gegenüber ihrer Muttersprache konfrontiert (Gogolin, 2019), weshalb die Integration ihrer Herkunftssprachen in Unterrichtsaktivitäten zu Bedenken hinsichtlich einer potenziellen Beurteilung aufgrund ihrer ethnischen Zugehörigkeit führen könnte (z. B. Stereotype Threat oder Othering). Der Beitrag untersucht die möglichen Unterschiede in ihrer Bewertung mehrsprachiger Unterrichtsaktivitäten zwischen plurilingual und monolingual aufgewachsenen Schüler:innen

Die vorliegende quasi-experimentelle Prä-Post-Kontrollgruppen-Studie untersucht mehrsprachige Unterrichtsaktivitäten im Französischunterricht als dritter Fremdsprache (zweite Fremdsprache) an Gymnasien in Deutschland aus der Perspektive der Lernenden. Die Stichprobe umfasst 221 Schüler:innen aus 20 Klassen der 7. Jahrgangsstufe. Insgesamt sprachen die Schülerinnen und Schüler der Stichprobe neben Deutsch 31 verschiedene Familiensprachen; die häufigsten Familiensprachen waren Türkisch, Englisch, Arabisch, Russisch und Polnisch. Etwa 40 % der Stichprobe wurden mehrsprachig erzogen. Um mehrsprachige Wertschätzung, interkulturelles Lernen und Sprachtransferstrategien zu fördern, wurden 12 mehrsprachige Unterrichtseinheiten für den Französischunterricht (zeitl. Umfang ca. 10-20 Minuten) entwickelt (Durchführungszeitraum: 6-12 Wochen, einmal pro Woche). Die Umsetzungsintensität durch die jeweilige Lehrkraft war unterschiedlich, daher wurde die Interventionsgruppe in zwei Gruppen aufgeteilt: IG I absolvierte weniger als 10 der 12 Einheiten (n = 41), IG II absolvierte mindestens 10 der 12 Einheiten (n = 73). Die Schüler wurden vor und nach der Intervention mit einem Fragebogen zu ihrer Wahrnehmung und Wertschätzung mehrsprachiger Unterrichtsaktivitäten (Cronbachs α lag zwischen .73 und .86) und weiteren Items zur Bewertung der durchgeführten Aufgaben befragt. Eine Teilstichprobe von 37 Schülerinnen und Schülern der Interventionsgruppe nahm an Posttest-Interviews über ihre Wahrnehmung der Intervention teil. Die Umfragedaten wurden mittels t-Tests analysiert, die qualitativen Daten wurden mittels qualitativer Inhaltsanalyse mit Doppelkodierung analysiert, die Interrater-Reliabilität war zufriedenstellend.

Die Ergebnisse des Prä-Post-Tests zeigen eine stärkere Wahrnehmung der „Umsetzung mehrsprachiger Unterrichtsaktivitäten“ in der IGII-Gruppe im Vergleich zur Kontrollgruppe, aber keine relevanten Unterschiede zwischen einsprachigen und mehrsprachigen Schüler:innen. In Bezug auf die „Wertschätzung mehrsprachiger Unterrichtsaktivitäten“ äußern mehrsprachige Schüler:innen bereits beim Vortest eine positivere Einstellung zu mehrsprachigen Unterrichtsaktivitäten als ihre einsprachigen Mitschüler ($t(71) = 3,578$, $p = .001$, Bonferroni-korrigiert $p = .004$). Diese positive Einstellung spiegelt sich auch in den Interviews mit mehrsprachigen Schüler:innen wider, aus denen hervorgeht, dass die Intervention in dieser Gruppe weder „Othering“ noch „Stereotype Threat“ ausgelöst hat. Im Gegensatz dazu äußerten sich einsprachige Schüler in den Interviews kritischer über mehrsprachige Unterrichtsaktivitäten, was darauf zurückzuführen sein könnte, dass sie weniger sprachliche Ressourcen zur Verfügung hatten als ihre mehrsprachigen Mitschüler (Deci & Ryan, 2000). Die Ergebnisse zeigen jedoch auch einen Anstieg der „Wertschätzung mehrsprachiger Unterrichtsaktivitäten“ für einsprachige Lernende in IG II ($t(38) = -2,199$, $p < .001$, Bonferroni-korrigiert $p = .004$). Dieser Befund markiert das Potenzial der Aufgaben für das interkulturelle Lernen speziell für monolingual aufgewachsene Schüler:innen (Beacco et al., 2016). Die Bedeutung mehrsprachiger Unterrichtsaktivitäten sowie Implikationen für die Lehrkräfteausbildung und -forschung werden diskutiert.

S-28: Diagnostik schriftsprachlicher Fähigkeiten: Effekte der Beurteilenden und der Ratingkriterien auf Urteilsqualität und -prozesse.

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: SN 163

Chair der Sitzung: Johannes Hartig

Chair der Sitzung: Johanna Fleckenstein

Chair der Sitzung: Stefan Keller

Diagnostik schriftsprachlicher Fähigkeiten: Effekte der Beurteilenden und der Ratingkriterien auf Urteilsqualität und -prozesse.

Chair(s): **Johannes Hartig** (DIPF, Deutschland), **Johanna Fleckenstein** (Universität Hildesheim)

Diskutant*in(nen): **Stefan Keller** (Diskussion 0000-0003-0115-5314 Stefan Keller stefandaniel.keller@phzh.ch Pädagogische Hochschule Zürich)

Schriftsprachliche Fähigkeiten sind von zentraler Bedeutung für den Bildungserfolg und spielen eine Schlüsselrolle in zahlreichen gesellschaftlichen Bereichen. Die Messung dieser Fähigkeiten ist daher für viele Felder der empirischen Bildungsforschung bedeutsam und wird von verschiedenen Fachgebieten beforscht. Bei der Messung schriftsprachlicher Fähigkeiten werden in der Regel umfangreichere offene Antworten erhoben. Diese werden durch geschulte Rater hinsichtlich ihrer Qualität bewertet, um aus diesen Urteilen diagnostische Indikatoren für die zu erfassenden Fähigkeiten zu bilden. Auch wenn die Bewertung schriftlicher Antworten durch aktuelle Entwicklungen in der Computerlinguistik zunehmend automatisiert erfolgen kann, bleiben menschliche Urteile als Referenz für diese Methoden unverzichtbar. Aufgrund der großen Bedeutung dieser Urteile werden die Entwicklung von Ratingverfahren, die reliable und valide Bewertungen schriftlicher Leistungen ermöglichen, intensiv beforscht. Fragestellungen hierbei beziehen sich auf die Formulierung von Beurteilungskriterien, ihre Anwendungsbedingungen, die Prozesse bei der Abgabe der Urteile und nicht zuletzt die Personen, die die Urteile vornehmen. Das Symposium vereint Beiträge, in denen die Beurteilung schriftlicher Leistungen in verschiedenen Kontexten untersucht wird. Hierbei werden die Perspektiven verschiedener am Forschungsfeld beteiligter Disziplinen kombiniert (Erziehungswissenschaft, Linguistik, Sprachdidaktik und Psychologie).

Die ersten beiden Beiträge kommen aus einem Projekt, in dem die Leistungen in integrierten Aufgabenformaten („reading into writing“) für Englisch als Fremdsprache bei Studierenden untersucht wird. Der Fokus liegt hier auf dem Einfluss der professionellen Expertise der Rater*innen, verglichen werden studentische Hilfskräfte mit erfahrenen Sprachlehrkräften. Im ersten Beitrag wird quantitativ der Effekt der Expertise auf die Beurteiler-Übereinstimmung und Strenge untersucht. Der zweite Beitrag untersucht qualitativ die bei der Beurteilung ablaufenden Prozesse, um ein vertiefendes Verständnis der quantitativen Ergebnisse zu ermöglichen. Auch im dritten Beitrag werden verschiedene Gruppen von Rater*innen verglichen. Es wird untersucht, wie sich Studierende sowie aktive und pensionierte Lehrkräfte bei der Beurteilung von Texten aus einem nationalen Assessment in der Primarstufe hinsichtlich Milde/Strenge unterscheiden. Im vierten Beitrag liegt der Schwerpunkt auf Unterschieden zwischen verschiedenen Formaten von Beurteilungskriterien. Untersucht wird ein neu entwickeltes Ratingverfahren, mit dem die Qualität von Texten bewertet werden, die von Hochschulstudierenden bei der Bearbeitung von Aufgaben zur Erfassung von „Critical Online Reasoning“ geschrieben wurden.

Zusammenfassend zeigt das Symposium die Vielfalt der Herausforderungen bei der Beurteilung schriftsprachlicher Leistungen in verschiedenen Kontexten. Aus den Ergebnissen lassen sich praktische Implikationen für die Entwicklung von Ratingverfahren und die Durchführung der Beurteilungen ableiten. Ein abschließender Diskussionsbeitrag adressiert unter anderem Limitationen und künftigen Forschungsbedarf sowie die Entwicklungsperspektiven einer immer stärker technologiebasierten Messung sprachlicher Fähigkeiten.

Beiträge des Symposiums

Effekte von professioneller Expertise auf die Konsistenz und Strenge bei der Bewertung integrierter Schreibprodukte

Ximena Delgado-Osorio¹, Johannes Hartig¹, Valeriia Koval², Claudia Harsch²

¹DIPF Frankfurt, ²Universität Bremen

Theoretischer Hintergrund

Eine der zentralen Anforderungen an Studierende, die ein Hochschulstudium aufnehmen, sind akademische Sprachkompetenzen. Diese werden mit zunehmender Internationalisierung der Hochschulbildung auch in englischsprachigen Studiengängen erwartet. Integrierten Aufgabenformaten wie Zusammenfassungen und textbasierten Stellungnahmen kommt dabei besondere Aufmerksamkeit zu, da diese sowohl in der Vorbereitung auf das Abitur (Ringwald & Harsch, 2018) als auch im ersten Studienjahr eine wichtige Rolle spielen (vgl. Arras, 2015; Ehlich & Steets, 2003; Motz, 2005). Solche integrierten Aufgaben decken unterschiedliche Kompetenzbereiche ab, insbesondere Leseverstehen und Schreiben, und erfordern daher die Integration von rezeptiven und produktiven Fertigkeiten (Plakans, 2020).

Bei der Erfassung von Leistungen von Studierenden in integrierten Aufgaben sind menschliche Bewertungen der Schreibprodukte erforderlich. Das Zustandekommen dieser Urteile und ihre Qualität sind von erheblicher Bedeutung für die Validität der Bewertungen. Auch die Eigenschaften der Bewerter*innen, die diese Urteile vornehmen, sind daher von wissenschaftlichem und praktischen Interesse. Eine prominente Eigenschaft der Bewerter*innen ist ihre professionelle Expertise (z. B. studentische Hilfskräfte vs. Sprachlehrkräfte). Über die Rolle der Expertise für die Qualität besteht in der Literatur kein Konsens. Einige Studien legen nahe, dass erfahrene Bewerter*innen auf andere Aspekte in den Schreibprodukten achten als Bewerter*innen ohne Erfahrung (vgl. Barkaoui, 2010). Andere Studien zeigen jedoch, dass erfahrene und unerfahrene Bewerter*innen durch gezieltes Training eine ähnliche Qualität der Bewertung erzielen können (Attali, 2015; Lim, 2011). Hinsichtlich der Effekte der Expertise von Bewerter*innen besteht demnach noch Forschungsbedarf.

Fragestellung

Diese Studie konzentriert sich auf den quantitativen Vergleich zwischen den Urteilen von Bewerter*innen mit unterschiedlicher Expertise. Konkret wird untersucht, wie sich die Bewertungen von integrierten Schreibprodukten durch Noviz*innen und Expert*innen in Bezug auf Inter-Rater-Reliabilität und Strenge der Urteile unterscheiden?

Methode

Die zu Beurteilenden Antworten stammen von einer Stichprobe von Abiturient*innen und Studierenden in englischsprachigen Studiengängen, die eine oder zwei integrierte Schreibaufgaben (Zusammenfassung oder Stellungnahme) bearbeitet hatten. Von den insgesamt 758 verfassten integrierten Schreibprodukten wurden nach Rater-Schulung 106 Schreibprodukte von 9 Bewerter*innen beurteilt, davon 6 Noviz*innen und 3 Expert*innen. Die Noviz*innen waren fortgeschrittene Studierende des Studiengangs "English-Speaking Cultures", während die Expert*innen erfahrene Englischlehrkräfte mit mehr als 10 Jahren Unterrichtserfahrung waren. Jede Gruppe erhielt eine Schulung, um sich mit der Bewertungsskala und den Materialien vertraut zu machen. Die Bewertungsskala wurde auf der Grundlage von Forschungsergebnissen zu Qualitätsdimensionen bei integrierten Aufgaben entwickelt (Plakans, 2009; Plakans & Gebriel, 2013, 2017; Gebriel & Plakans, 2014) und umfasst neue Kriterien: (1) Identifikation relevanter Ideen des Ausgangstexts, (2) Korrektheit und Präzision der Ideen des Ausgangstexts, (3) Sprachliche Verarbeitung, (4) Ausgangstext-Zuordnung, (5) Argumentation, (6) Textstruktur, (7) Kohäsion, (8) Wortschatz und (9) Grammatik, die jeweils mit Deskriptoren für eine fünfstufige Bewertung entwickelt wurden.

Ergebnisse

Analysen zur Interrater-Reliabilität zeigen, dass die Intraklassenkorrelation nach Koo & Li (2016) sich teilweise zwischen den Bewertungen von Noviz*innen und Expert*innen unterscheiden. Für drei der Kriterien (Ausgangstext-Zuordnung, Argumentation, Wortschatz und Grammatik) sind die Übereinstimmungen bei den Expert*innen höher als bei den Noviz*innen. Für das Kriterium „Textstruktur“ hingegen sind die Urteile der Noviz*innen konsistenter. Deskriptive Analysen zur Strenge der Bewertungen zeigen, dass die Novizen insbesondere bei den sprachlichen Qualitätskriterien (Wortschatz und Grammatik) tendenziell strenger bewerten als die Experten.

Diskussion

Die Ergebnisse zeigen leichte Qualitätsunterschiede in den Urteilen zugunsten der Expert*innen, aber nicht in allen bewerteten Kriterien. Aus einer praktischen Perspektive wird diskutiert, inwieweit die Ergebnisse den höheren Aufwand bei Rekrutierung und Vergütung professioneller Bewerter*innen rechtfertigen. Mögliche Ursachen für die Unterschiede im Urteilsverhalten werden vertiefend in einem weiteren Beitrag in diesem Symposium betrachtet, indem zusätzlich zu den quantitativen Analysen auch qualitative Methoden zur Anwendung kommen.

Effekte von professioneller Expertise auf die Bewertungsprozesse und Strategien bei der Bewertung integrierter Schreibprodukte

Valeriia Koval¹, Claudia Harsch¹, Ximena Delgado-Osorio², Johannes Hartig²

¹Universität Bremen, ²DIPF Frankfurt

Theoretischer Hintergrund

Integrierte Aufgabenformate, bei denen schriftliche Antworten auf Basis von vorgegebenen Texten produziert werden müssen, spielen bei der Erfassung fremdsprachlicher Kompetenzen eine wichtige Rolle. Bei der Auswertung dieser Aufgaben ist die Bewertung der Schreibprodukte durch menschliche Bewerter*innen erforderlich. Das Zustandekommen dieser Urteile und ihre Qualität sind von großer praktischer und wissenschaftlicher Relevanz, und in diesem Zusammenhang auch Eigenschaften der Beurteiler*innen, die die Bewertungen beeinflussen können. Ein wichtiges Merkmal der Beurteiler*innen ist ihre professionelle Expertise. Obwohl es bereits zahlreiche Studien zum Vergleich von Expert*innen und Noviz*innen gibt (Barkaoui, 2010a, 2010b; Cumming, 1990; Lim 2011), besteht keine Einigkeit über den Einfluss von Expertise auf die Bewertungsqualität. Frühere Untersuchungen (Cumming, 1990) haben gezeigt, dass Expert*innen mehr selbstreflexive Strategien anwenden und eine umfassendere mentale Repräsentation des Problems der Bewertung von Texten haben. Sie nutzen verschiedene Kriterien, effektive Selbstkontrollstrategien und verschiedene Wissensquellen beim Lesen und Beurteilen. Neulinge hingegen konzentrieren sich eher auf ein kleineres Spektrum von Merkmalen und tendieren dazu, sprachliche Merkmale eher zur Korrektur oder Bearbeitung zu nutzen, anstatt sie als Information für eine Bewertung zu verwenden. In einer anderen Studie (Barkaoui, 2010a) wurden bei genauerer Betrachtung des Entscheidungsverhaltens und der Strategien jedoch keine signifikanten Unterschiede zwischen erfahrenen und unerfahrenen Personengruppen festgestellt. Hinsichtlich der Effekte der Expertise von Bewerter*innen besteht demnach noch Forschungsbedarf besonders in der Hinsicht auf das analytische Bewertungsverfahren.

Fragestellung

Diese Studie konzentriert sich auf den qualitativen Vergleich zwischen den Urteilsprozessen durch Bewerter*innen mit unterschiedlicher Expertise. Konkret lauten die Forschungsfragen:

1. Welche Unterschiede bestehen in Bezug auf die Bewertungsprozesse und Strategien zwischen erfahrenen und nicht-erfahrenen Bewerter*innen?
2. Inwieweit lassen sich die quantitativen Ergebnisse der Beurteilungsqualität durch die festgestellten Bewertungsstrategien und Prozesse der erfahrenen und der nicht erfahrenen Bewerter*innen erklären?

Methode

Die beiden Gruppen von Bewerter*innen in dieser Studie sind 6 Noviz*innen und 3 Expert*innen. Die Novizen*innen waren fortgeschrittene Studierende des Studiengangs "English-Speaking Cultures", während die Experten erfahrene Englischlehrkräfte mit mehr als 10 Jahren Unterrichtserfahrung waren. Jede Gruppe erhielt eine Bewerter*innen-Schulung, um sich mit der Bewertungsskala und den Materialien vertraut zu machen. Die Bewertungsskala wurde auf der Grundlage von Forschungsergebnissen zu Qualitätsdimensionen in integrierten Aufgaben entwickelt (Plakans, 2009; Plakans & Gebriel, 2013, 2017; Gebriel & Plakans, 2014): (1) Qualität der Ausgangstextnutzung, (2) Diskursynthese und (3) Sprachqualität, für die jeweils Kriterien mit Deskriptoren für eine fünfstufige Bewertung entwickelt wurden. Um Unterschiede in den kognitiven Bewertungsprozessen von erfahrenen und unerfahrenen Bewerter*innen zu analysieren, wurde eine qualitative Datenerhebung in Form von Lautes-Denken-Protokollen durchgeführt (vgl. Cumming, 1990; Barkaoui, 2010a, 2010b; Lim 2011). Während der Sitzung bewerteten die Bewerter*innen jeweils vier Textprodukte aus zwei integrierten Aufgaben (Zusammenfassung und Stellungnahme). Die entstehenden Transkripte wurden in MAXQDA codiert und analysiert.

Ergebnisse

Die Ergebnisse der Noviz*innen zeigen mehrere problematische Aspekte und hoch-individuelle Bewertungsprofile. Besonders herausfordernd für die Bewerter*innen war die Beurteilung von Stellungnahmen sowie die Bewertung der sprachlichen Kriterien, wenn die Texte viele wörtliche Übernahmen aus dem Ausgangstext enthielten. Außerdem war die Reihenfolge der Bewertung von Kriterien nicht immer dieselbe, was ein Einfluss auf die Qualität der Bewertung haben konnte. Dennoch folgten die Noviz*innen insgesamt den vorgegebenen Phasen der Bewertung (Vorbereitendes Lesen, Bewerten, Score vergeben). Die Auswertung der Lautes-Denken-Protokolle der Expert*innen ist zum Zeitpunkt der Einreichung dieses Beitrags noch nicht

abgeschlossen. Es wird erwartet, dass Expert*innen weniger interindividuelle Unterschiede aufweisen und mehr Sicherheit bei der Bewertung der Sprachlicher Kriterien demonstrieren. Die Ergebnisse der qualitativen Analysen werden mit den quantitativen Ergebnissen aus dem vorangehenden Symposiums-Beitrag in Bezug gesetzt, um ein tieferes Verständnis der quantitativen Unterschiede zwischen Expert*innen und Noviz*innen zu erreichen.

Diskussion

Die Ergebnisse werden unter anderem dahingehend diskutiert, inwieweit die Urteile durch Bewerter*innen mit unterschiedlicher Expertise vergleichbar sind und ob sie in der Praxis als gleichwertig betrachtet werden können.

Der Milde/Strenge-Bias beim kategorienbasierten Rating der Textqualität von Schreibprodukten aus der Primarstufe

Lea Hoffmann¹, Ilka Fladung², Verena Raith³, Elke Mielach³, Veronika Österbauer³, Jörg Jost¹, Marcel Illetschko³

¹Universität zu Köln, ²Universität Duisburg-Essen, ³Institut des Bundes für Qualitätssicherung im österreichischen Schulwesen (IQS) Salzburg

Theoretischer Hintergrund

Zur Erfassung von Textqualität haben sich Verfahren, bei denen geschulte Rater kategorienbasiert vorgehen, als geeignet erwiesen (Weigle, 2002; Neumann, 2017). Die Schreib- und die Assessmentforschung haben erhebliche Anstrengungen in die Entwicklung von Methoden zur objektiven, reliablen und validen Bewertung der Textqualität investiert: Verschiedene Ratingverfahren wurden weiterentwickelt und deren Gelingensbedingungen untersucht (Eckes, 2008; Skar & Jølle, 2017).

Unabhängig davon, welches Ratingverfahren im Detail eingesetzt wird, ist es für die Erfassung von Textqualität zentral, dass die Rater mithilfe der Kategorien zu übereinstimmenden Urteilen gelangen. Neben der Qualität der Ratingkategorien selbst spielt dabei insbesondere das Training der Rater eine zentrale Rolle. Dafür wurden Verfahren wie Raterschulungen und Benchmark-Konferenzen etabliert und ebenfalls beforscht (van Gasse et al., 2019; Woehr & Huffcutt, 1994). Es zeigt sich jedoch immer wieder, dass selbst sehr aufwändige Verfahren zur Kategorienentwicklung und Schulung von Ratern nicht immer den gewünschten Effekt haben (Birkel & Birkel, 2002; Cooksey, Freebody & Wyatt-Smith, 2007), sogar Expertenratergruppen betreffend (Hennes 2020). Da man jedoch für Urteile über Textqualität auf verlässliche und übereinstimmende Raterurteile angewiesen ist, besteht hier weiterer Forschungsbedarf.

Fragestellung

In diesem Beitrag soll ein typischer Raterbias, der in Ratergruppen vorkommen kann, näher beleuchtet werden: Der Milde/Strenge-Effekt. Diesem Bias werden, ähnlich wie dem Anker-, Reihenfolge- und Halo-Effekt, beständig negative Auswirkungen auf die Reliabilität von Ratings attestiert (Fahim & Birjani, 2011; Eckes, 2004; Congdon & McQueen, 2000; Honko et al., 2023). Ratertrainings sollen diesen Effekten entgegenwirken; ihre Effektivität wird diskutiert (Knoch & Read, 2007; Lamprianou et al, 2023; Woehr & Huffcutt, 1994). Dieser Beitrag fokussiert verschiedene Personengruppen beim Rating, indem er danach fragt, ob verschiedene Berufsgruppen, konkret Studierende und Lehrpersonen zu verschiedenen Stadien der Professionalisierung, dem Bias Milde/Strenge besonders unterliegen.

Stichprobe und Methode

Grundlage der Analyse sind Ratingentscheidungen aus einer Ratingphase mit 186 Ratern während einer mehrmonatigen Trainingsphase für die Auswertung der iKMPLUS Deutsch, der nationalen Kompetenzerhebung in Österreich (Verfassen von Texten, 4. Schulstufe, Frühjahr 2023). Bewertet wurden insgesamt 46 handschriftliche Schreibprodukte in bis zu vier verschiedenen Dimensionen (Inhalt, Aufbau, sprachliche Angemessenheit, sprachliche Richtigkeit). Die Aufgaben, zu denen die Texte verfasst wurden, sind profilierte, textmusterspezifische Aufgaben (je eine Aufgabe zum Argumentieren, Beschreiben, Erzählen und Instruieren). Zu allen Schreibprodukten liegen Benchmarkentscheidungen vor, die von einem Panel aus sieben Expertenratern gegeben wurden.

Die Stichprobe der Rater setzt sich aus drei verschiedenen Subgruppen zusammen: aktive Lehrkräfte (N = 128), Studierende (N = 41) und pensionierte Lehrkräfte (N = 17). Von diesen Personen liegen insgesamt N = 84.672 Ratingentscheidungen auf Kriterienebene bzw. N = 13.170 auf Dimensionsebene vor. Mithilfe von linearen und logistischen Regressionsanalysen wird untersucht, ob sich die drei Gruppen dahingehend unterscheiden, inwiefern sie milder oder strenger als die Expertengruppe urteilen (Raterbias) und wie gut sie mit den Benchmarkurteilen der Expertengruppe übereinstimmen (Akkuratheit der abgegebenen Urteile).

Ergebnisse und ihre Bedeutung

Zum Raterbias: Eine Analyse der Bewertungen auf Dimensionsebene mithilfe einer linearen Regression zeigt, dass die pensionierten Lehrkräfte im Durchschnitt signifikant milder als die Lehrkräfte sind ($p < 0,001$). Die Effektstärke ist jedoch sehr klein: Die Scores der pensionierten Lehrkräfte sind lediglich 0,17 Standardabweichungen höher (Adjustiertes $R^2 = 0,0038$, $F(12, 13157) = 5,195$, $p < 0,01$).

Zur Akkuratheit der Urteile: Eine Regressionsanalyse zeigt, dass pensionierte Lehrkräfte weniger akkurat bewerten als die Lehrkräfte ($p < 0,001$). Eine logistische Regression zeigt, dass die Wahrscheinlichkeit, dass eine pensionierte Lehrkraft mit dem Benchmarkurteil der Expertengruppe übereinstimmt, nur 0,84 mal so groß ist, wie die einer Lehrkraft (Pseudo- $R^2 = 0,0021$).

Diskussion

Die Ergebnisse werden vor allem im Hinblick auf die Herausforderungen diskutiert, die begleitende Forschung von Ratingprozessen mit sich bringen. Zudem soll dargestellt werden, in welcher Form deskriptive Auswertungen zum Milde/Strenge-Bias innerhalb der Ratertrainings der iKMPLUS auch als Feedback für die Rater genutzt werden können.

Bewertungsverfahren zur Beurteilung offener Antworttexte zum kritisch-reflektierten Umgang mit Online-Medien

Johannes Hartig¹, Anika Kohmer², Philine Drake¹, Jannick Illmann¹, Carmen Köhler¹, Stroganova Alisa³, Jan Zottmann³

¹DIPF Frankfurt, ²Johannes Gutenberg-Universität Mainz, ³LMU München

Theoretischer Hintergrund

Studierende beziehen zunehmend Informationen und Lerninhalte aus Online-Quellen, da diese leichter zugänglich sind als traditionelle Medien (z. B. Lehrbücher) (Ugur et al., 2020; Maurer et al., 2020). Dabei stellen die Vielzahl und Fehlerhaftigkeit der verfügbaren Online-Informationen eine große Herausforderung dar. Erfolgreiches Lernen mit Online-Quellen erfordert die

Anwendung kritisch-reflektierter Online-Recherche Strategien (Schoor et al., 2020), die im Folgenden als „Critical Online Reasoning“ (COR) definiert werden (Molero et al., 2020). COR beinhaltet die effektive Auswahl („Online Information Acquisition“; OIA) und Bewertung („Critical Information Evaluation“; CIE) von Online-Informationen hinsichtlich ihrer Glaubwürdigkeit und Relevanz, sowie die korrekte Interpretation und argumentative Begründung („Reasoning based on Evidence, Argumentation and Synthesis“; REAS) der gewonnenen Ergebnisse (ebd.). Zudem werden COR-Fähigkeiten nach ihrem Bezug zu allgemeinen (GEN-COR) und domänenspezifischen (DOM-COR) Kontexten unterteilt. Für die Erfassung von COR-Fähigkeiten sind realitätsnahe Aufgaben erforderlich, idealerweise mit Zugriff auf das offene Internet. Zur Erfassung der REAS-Facette sind schriftliche Argumentationen und Begründung der Quellenauswahl geeignet (Nagel et al., 2022). Indikatoren für individuelle COR-Fähigkeiten können aus der Beurteilung dieser schriftlichen Antworten gebildet werden (Zlatkin-Troitschanskaia et al., 2021).

Fragestellung

Für die Beurteilung der schriftlichen Antworten zur Erfassung der REAS-Facette müssen geeignete Methoden identifiziert werden. Für komplexe Konstrukte wie COR erscheinen analytische Ratingschemas als geeignet, bei denen Charakteristika der Antwort und die damit verbundenen Scores für mehrere spezifische Kriterien separat beschrieben werden (Chen et al., 2022; Andrich, 2011). Der vorliegende Beitrag befasst sich mit der Frage, wie COR-Fähigkeiten durch ein solches kriteriengestütztes Bewertungsschema erfasst werden können und welche Indikatoren konkret die REAS-Facette abbilden. Der Fokus liegt dabei auf allgemeinen COR-Fähigkeiten (GEN-COR).

Methode

Es wurde eine GEN-COR-Aufgabe entwickelt und einer Stichprobe von N = 424 Erstsemesterstudierenden aus vier Fachrichtungen (Wirtschaftswissenschaften, Medizin, Soziologie, Physik) an drei deutschen Universitäten vorgelegt. Die Teilnehmenden sollten im offenen Internet mögliche Gesundheitsrisiken aufgrund der Wohnbedingungen in der Nähe eines Mobilfunkmastes recherchieren und eine offene Antwort mit Pro- und Contra-Argumenten verfassen, die sie mit Online-Quellen belegen. Die Aufgabenstellung enthält Prompts, die die drei COR-Facetten abbilden. Die offenen Antworttexte zeigen, inwieweit die Teilnehmenden in der Lage sind, unterschiedliche Argumente zu einer kohärenten und inhaltlich vollständigen Stellungnahme (REAS) zusammenzuführen und zu begründen. Zur Bewertung der REAS-Facette wurde eine Musterlösung entwickelt und ein kriteriengestütztes Bewertungsschema abgeleitet. Das Schema besteht einerseits aus allgemeinen Qualitätskriterien mit abgestuften Verhaltensankern (z. B. Verwendung von Quellen, Kohärenz der Argumentation) und andererseits aus spezifischen aufgabenrelevanten Inhalten, die gefunden und einbezogen werden können (z. B. Strahlung durch Mobiltelefone ist stärker als durch Mobilfunkmasten). Anschließend wurden ausgewählte Rater*innen über mehrere Wochen systematisch geschult, um die Bewertungen eigenständig und unabhängig voneinander durchführen zu können.

Ergebnisse

Wir präsentieren die Entwicklung eines kriteriengestützten Bewertungsschemas zur Beurteilung der schriftliche GEN-COR-Leistung von Studierenden. Es werden erste empirische Ergebnisse zur Beurteilerübereinstimmung und Zusammenhangsstruktur dieser Kriterien präsentiert. Es wird untersucht, ob allgemeine und spezifische Kriterien im Hinblick auf die Beurteilerübereinstimmung gleichermaßen gut funktionieren und ob sich beide Gruppen zu einer gemeinsamen Qualitätsdimension zusammenfassen lassen.

Diskussion

Die vorliegende Studie zielt darauf ab, ein systematisches Bewertungsschema für schriftliche Antworten in Aufgaben zur Erfassung von GEN-COR zu entwickeln. Die Ergebnisse zeigen, welche Kriterien mehr oder weniger gut funktionieren. Es wird diskutiert, inwieweit die Ergebnisse für die Entwicklung von Kriterien für weitere Aufgaben übertragbar sind. Weiterhin wird allgemein diskutiert, ob sich die Erfassung von COR-Fähigkeiten mit schriftlichen Antworten und unter Nutzung des offenen Internets bewähren kann.

S-29: Urteile von Lehrkräften: Entwicklung, Veränderbarkeit und Auswirkungen

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 145

Chair der Sitzung: Jan Beck

Chair der Sitzung: Detlef Urhahne

Urteile von Lehrkräften: Entwicklung, Veränderbarkeit und Auswirkungen

Chair(s): **Jan Beck** (Universität Münster, Deutschland)

Diskutant*in(nen): **Detlef Urhahne** (Universität Passau)

Lehrkräfteurteile gelten als wesentlich für die Gestaltung von Unterricht, zum Beispiel beim adaptiven Unterrichten. Daneben wirken sich Lehrkräfteurteile nachweislich auf die Lernentwicklung von Schüler*innen aus (z. B. Brandmiller et al., 2024). So kann es zum Beispiel zu negativen Effekten der Lernentwicklung kommen, wenn Lehrkräfte niedrige Erwartungen an eben diese Lernentwicklung haben (s. Pygmalion-Effekt, Rosenthal & Jacobsen, 1968; Rubie-Davis et al., 2006). Während es hinreichende empirische Evidenz zum Befund der (Un-)Genauigkeit von Lehrkräfteurteilen über lernrelevante Variablen der Schüler*innen gibt (Urhahne & Wijnia, 2021), existieren deutlich weniger empirische Studien dazu, wie Lehrkräfte ihre Urteile über die Zeit ändern, wie sich ihre Urteile beeinflussen lassen und durch welche Mechanismen Urteile Lernprozesse beeinflussen. Zudem fokussierten viele Studien bislang auf die Beurteilung der akademischen Leistung der Schüler*innen. Weniger ist über Urteile zu motivationalen Variablen bekannt. Hier setzt das Symposium an, indem die jeweiligen Studien mit längsschnittlichen Studiendesigns einen Blick auf die Entwicklung, Beeinflussbarkeit und Auswirkungen von Lehrkräfteurteilen werfen und dabei neben der akademischen Leistung auch motivationale Variablen berücksichtigen.

Der erste Beitrag untersucht anhand einer Stichprobe von $N = 58$ Praxissemesterstudierenden und $N = 58$ Lehrkräften sowie $N \approx 1.400$ Schüler*innen, wie sich die Genauigkeit lesekompetenz-, lesemotivations- und leseelbstkonzeptbezogener Urteile von (angehenden) Lehrkräften im Laufe eines Schulhalbjahres verändert. Zudem wird geprüft, ob sich diese Veränderungen für die verschiedenen Merkmale und für Studierende und Lehrkräfte unterscheiden.

Im zweiten Beitrag wird untersucht, ob Lehrkräften die Verfügbarkeit motivationaler Informationen hilft, ihre Motivationsurteile mit denen ihrer Schüler*innen anzugleichen. Dazu wurde minimal invasiv im Sinne des formativen Assessments die Motivation von 154 Schüler*innen zu 5 Zeitpunkten über 4 Wochen erfasst und Lehrkräfte schätzten diese ebenso häufig ein. Ein leicht positiver Trend in der Lehrkraft-Schüler*innen-Übereinstimmung konnte festgestellt werden.

Der dritte Beitrag untersucht die Bedeutung von leistungsbezogenen Lehrkrafturteilen für die sprachliche, mathematische und naturwissenschaftliche Leistungsentwicklung und prüft, ob Urteilsverzerrungen für als *English Learner* gelabelte Schüler*innen zu Benachteiligung führen. Daten einer US-Längsschnittstudie zeigten nach Matching von 789 gelabelten Schüler*innen zu Schüler*innen ohne Label eine bessere Leistungsentwicklung durch höhere Lehrkrafturteile, im sprachlichen Bereich allerdings eine Benachteiligung von gelabelten Schüler*innen durch Urteilsverzerrung.

Das Symposium liefert Erkenntnisse zur Veränderung und Veränderbarkeit sowie zu Auswirkungen von Lehrkrafturteilen zu verschiedenen lernrelevanten Merkmalen von Schüler*innen. Die Studien können dazu beitragen, die Bildungsprozesse der Schüler*innen durch kompetente Lehrkräfte zu optimieren.

Beiträge des Symposiums

Entwicklung der Urteilsgenauigkeit von (angehenden) Lehrkräften

Jacqueline Zielke¹, Henrike Mentel², Mareike Ehlert², Natalie Förster¹

¹Bergische Universität Wuppertal, ²Universität Münster

Theoretischer Hintergrund

Urteile von Lehrkräften über lernrelevante Schüler*innenmerkmale gelten als wesentlich für die Unterrichtsgestaltung (Herppich et al., 2018) und beeinflussen nachweislich die Lernentwicklung (Förster et al., 2022). Während vergleichsweise viele Studien dazu vorliegen, wie (un)genau Lehrkräfte lernrelevante Merkmale ihrer Schüler*innen einschätzen (Südkamp et al., 2012), ist über die Entwicklung und Veränderung derartiger Urteile sowie über Faktoren, die die Genauigkeit beeinflussen, vergleichsweise wenig bekannt.

Die Berufserfahrung von Lehrkräften wird als möglicher Einflussfaktor auf die Urteilsgenauigkeit diskutiert, hat sich jedoch als nicht relevant gezeigt (vgl. Urhahne & Wijnia, 2021). Allerdings ist basierend auf dem Linsenmodell (Back & Nestler, 2016) davon auszugehen, dass sich insbesondere die Quantität und Qualität valider Hinweisreize auf die Urteilsgenauigkeit auswirken. Demnach sollten die Urteile von Personen, denen weniger Hinweisreize vorliegen (z. B. aufgrund geringer Kontaktzeit) ungenauer ausfallen als von Personen die über mehr Hinweisreize verfügen und Urteile damit über die Zeit genauer werden. Dieser Unterschied sollte sich allerdings insbesondere bei Merkmalen zeigen, für die viele valide Hinweisreize vorliegen (Leistung) und weniger bei Merkmalen, für die weniger Hinweisreize vorliegen (Motivation und Selbstkonzept; Urhahne & Wijnia, 2021).

Vorliegende Studien zum Zusammenhang von Kontaktzeit mit der Klasse und der Urteilsgenauigkeit bei schulleistungsbezogenen/kognitiven Merkmalen kommen zu uneindeutigen Befunden (Urhahne & Wijnia, 2021). Dabei wurden bislang jedoch Lehrkräfte untersucht, die bereits mehrere Monate in den Klassen unterrichten (Oerke et al., 2015). Möglicherweise zeigt sich der Einfluss der Kontaktzeit bei Leistungsurteilen bereits nach einigen Monaten nicht mehr. Bezüglich der Urteilsgenauigkeit motivationaler Merkmale gibt es Hinweise, dass eine längere Kontaktzeit nicht zu genaueren Urteilen führt (Praetorius et al., 2015).

Fragestellung und Hypothesen

In unserer Studie untersuchen wir die Entwicklung der Urteilsgenauigkeit von (angehenden) Lehrkräften über ein Schulhalbjahr. Darüber hinaus betrachten wir Unterschiede in der Urteilsgenauigkeit sowie in der Entwicklung der Urteilsgenauigkeit in Abhängigkeit der Kontaktzeit mit den Schüler*innen sowie in Abhängigkeit des einzuschätzenden Merkmals. Dazu vergleichen wir die Urteilsgenauigkeit von Praxissemesterstudierenden mit nur wenig Kontaktzeit zu einer Klasse mit der Urteilsgenauigkeit der in dieser Klasse unterrichtenden Lehrkraft zu Beginn und am Ende eines Schulhalbjahres bezüglich der Lesekompetenz, der Lesemotivation und dem Leseselbstkonzept der Kinder. Wir erwarten, dass a) die Leistungsurteile, nicht aber die Lesemotivations- und Leseselbstkonzepturteile zum Ende des Schuljahres genauer ausfallen als zu Beginn des Schuljahres, b) die Urteilsgenauigkeit der Lehrkräfte zu Beginn des Schuljahres für die Leistung, nicht aber für die motivationalen Merkmale signifikant höher ausfällt als bei den Praxissemesterstudierenden und c), sich die Urteilsgenauigkeit der Praxissemesterstudierenden bezogen auf die Lesekompetenz zum Ende des Schulhalbjahres nicht mehr signifikant von der

Urteilsgenauigkeit der Lehrkräfte unterscheidet, da die Praxissemesterstudierenden die Schüler*innen während des Schulhalbjahres beobachtet haben.

Methode

Zur Beantwortung der Fragestellungen werden Daten aus dem Projekt „Lese-Sportler: mit PS in die Schule“ ausgewertet. In diesem Projekt haben jeweils $N = 58$ Praxissemesterstudierende und die Lehrkräfte der untersuchten Klassen zu Beginn und Ende eines Schulhalbjahres die Leseleistung, Lesemotivation und das Leseselbstkonzept einer Klasse eingeschätzt. Die Lesekompetenz wurde mit dem standardisierten Lesetest ELFE II (Lenhard et al., 2017) und die Lesemotivation und das Leseselbstkonzept der Zweit- bis Viertklässler*innen ($N \approx 1.000$) mit entsprechenden IGLU-Skalen (Goy et al., 2017) erhoben.

Die Datenerhebung ist abgeschlossen. Die Auswertung der Daten erfolgt anhand von multilevel mixed-effects Modellen in R, wobei die gemessenen Merkmale durch die Urteile, die Messzeitpunkte, die Urteilenden, die Art des Merkmals und entsprechende Interaktionen präzisiert werden.

Ergebnisse und Diskussion

Im Sinne von Open Science wird aktuell die Präregistrierung verfasst, weshalb noch keine Ergebnisse vorliegen. Die Befunde werden einen vertiefenden Einblick dazu geben, wie sich Urteile von (angehenden) Lehrkräften innerhalb eines Schulhalbjahres entwickeln und Rückschlüsse auf die Veränderbarkeit von Urteilsprozessen erlauben. Die Ergebnisse werden im Kontext der Diagnostischen Kompetenz von Lehrkräften sowie der Rolle von schulischer Diagnostik diskutiert.

Verbesserung der Motivationsdiagnostik von Lehrkräften

Jan Beck, Stephan Dutke

Universität Münster

Motivation ist zentral für erfolgreiche Lernprozesse. Schüler*innen unterscheiden sich jedoch anhand motivationaler Beweggründe, Ziele, und Anreize. Um demnach individualisiert Motivation zu wecken und gegebenenfalls zu fördern, ist es für Lehrkräfte wichtig, Kenntnis über die Motivation ihrer Schüler*innen zu haben. In vergangenen Untersuchungen wurde allerdings gezeigt, dass die Einschätzungen zur Motivation der Schüler*innen nur gering zwischen Lehrkräften und den schülerseitigen Selbstberichten übereinstimmen (Urhahne & Wijnia, 2021), unabhängig von den motivationalen Aspekten, die sie einschätzen sollten (z.B. Zielorientierungen, Dicke et al., 2012; selbstbestimmte Motivation, Dutke & Hinnersmann, 2015). Entsprechend des Realistic Accuracy Models (Funder, 2012) könnte die mangelnde Übereinstimmung an der Verfügbarkeit relevanter Informationen liegen, die benötigt werden, um übereinstimmende Urteile zu fällen. Insbesondere im Schulalltag könnte die Motivation von Schüler*innen eine untergeordnete Rolle für Lehrkräfte spielen, zugunsten eines starken Fokus auf Kompetenzziele der Schüler*innen und deren Erreichung. Für Lehrkräfte, die mehr Kenntnis motivational relevanter Aspekte ihrer Schüler*innen haben und die sich zum Beispiel Rückmeldungen einholen, dürfte die Motivation der Schüler*innen demgegenüber stärker im Fokus stehen und die Übereinstimmung zwischen Lehrkräften und Schüler*innen größer sein (vgl. Wentzel et al., 2010).

Das Ziel der vorliegenden Untersuchung ist es deshalb, (a) Lehrkräften mehrmals Informationen zur Selbsteinschätzung der Motivation ihrer Schüler*innen zu geben, damit sie ihre Einschätzungen abgleichen können, und (b) zwischen diesen Einschätzungen mit einzelnen Schüler*innen über deren Motivation zu kommunizieren. Beides soll letztlich die Übereinstimmung in den Motivationsurteilen zwischen Lehrkräften und Schüler*innen verbessern.

Im Rahmen einer Pilotstudie nahmen 8 Lehrkräfte ($M = 45.9$ Jahre, $SD = 10.0$, 6 davon weiblich) und deren 154 Schüler*innen ($M = 14.5$ Jahre, $SD = 2.0$, 41.8% weiblich, 29.6% Migrationshintergrund) teil. Auf der Tagung werden die Daten von mindestens 13 Lehrkräften präsentiert, was eine Power-Simulation für den gewünschten Interaktionseffekt ($\alpha = .05$, power = .80, 1.000 Simulationen je Schritt) ergab. Zu 5 Zeitpunkten schätzten sowohl Lehrkräfte als auch Schüler*innen gleichzeitig die Motivation der Schüler*innen im aktuellen Fachunterricht ein. Die Schüler*innen gaben dabei ihre Annäherungs-Aufgabenziele anhand von 4 Items an (eine wesentliche Facette der Masterziele im Sinne der hexagonalen Zielstruktur nach Daumiller et al., 2019; Cronbach's $\alpha = .91$). Lehrkräfte schätzten die Motivation aus ökonomischen Gründen anhand eines Single-Items ein, indem sie berichteten, wie motiviert sie die einzelnen Schüler*innen im aktuellen Fachunterricht hielten (9-P Likert Skala, „wenig motiviert“ bis „sehr motiviert“). Lehrkräfte erhielten ein bis zwei Tage nach ihrer Einschätzung Feedback über die Annäherungs-Aufgabenziele ihrer Schüler*innen, transformiert auf die gleiche 9-P Likert Skala wie ihre eigenen Einschätzungen. Danach hatten die Lehrkräfte im Median 6 Tage Zeit, um mit den Schüler*innen individuell über die Ursachen der Einschätzungsdiskrepanz zu sprechen (in Summe 33 Gespräche), bevor beide erneut die Motivation einschätzten. Die Daten wurden mit linearen gemischten Modellen auf 3 Ebenen (713 Einschätzungen genestet in 154 Schüler*innen und in 8 Lehrkräften bzw. Klassen) analysiert, die für die vorliegende Stichprobe präzise Modellierungen erlaubten.

Das Nullmodell zeigte, dass Lehrkräfteeinschätzungen hoch zwischen Schüler*innen ($ICC = .58$) und leicht zwischen Lehrkräften ($ICC = .10$) variierten. Das Hinzufügen von Annäherungs-Aufgabenzielen zeigte, dass Lehrkräfteeinschätzungen gering mit Annäherungs-Aufgabenzielen zusammenhingen ($\beta = .20$, $p < .001$). In Interaktion mit Zeit (Zeitpunkt 1 bis 5) veränderte sich die Lehrkraft-Schüler*innen-Übereinstimmung leicht positiv ($\beta = .06$, $p = .018$).

Die vorläufigen Ergebnisse bestätigen zum einen die empirische Evidenz, dass Lehrkräfte gering mit Schüler*innen in motivationalen Aspekten übereinstimmen. Zum anderen zeigt bereits die Stichprobe der vorliegenden Pilotierung, dass die eingesetzte Intervention, regelmäßig Feedback zur Übereinstimmung der eigenen Einschätzung mit der der Schüler*innen zu erhalten, ein effektives, niedrigschwelliges Hilfsmittel sein kann, um Lehrkräften zu erleichtern, motivationale Aspekte ihrer Schüler*innen zu erkennen und gewinnbringend für die Gestaltung von Unterricht und Lernen zu nutzen.

Das Label English Learner: Führen Lehrkräfteurteile zu Leistungsunterschieden zwischen Schüler*innen?

Katharina Schnitzler¹, Ilana Umansky², Hanna Dumont¹

¹Universität Potsdam, ²University of Oregon

Theoretischer Hintergrund

Die Forschung zu Erwartungseffekten belegt, dass Lehrkräfteurteile wie selbsterfüllende Prophezeiungen wirken können* Aufgrund ihrer Einschätzungen ändern Lehrkräfte ihr Verhalten gegenüber den Schüler*innen; diese nehmen die Unterschiede wahr und ändern ihr Verhalten und ihre Motivation. Folglich fallen die Leistungen von schwächer eingeschätzten Schüler*innen geringer aus als sie könnten, während stärker eingeschätzte Schüler*innen von höheren Zuwächsen profitieren (Good, 2024). Dieses Phänomen ist problematisch, da Lehrkräfteurteile anfällig für systematische Verzerrungen sind. So konnte beispielsweise gezeigt werden, dass Leistungen von Schüler*innen mit diagnostischem Label (z.B. ADHS-Diagnose) von Lehrkräften unterschätzt werden (Franz, 2023, Wang et al., 2018). Bisher ist unklar, inwiefern solche Verzerrungen zu unterschiedlichen

Leistungsentwicklungen führen. In den USA tragen 10.6% der Schüler*innen das Label English Learner (EL; National Center für Education Statistics, 2024), welches mehrsprachige Schüler*innen mit mangelnden Englischkenntnissen identifiziert. Neben bildungspolitischen Arbeiten, die den defizitorientierten Fokus des Labels kritisieren (Umansky & Avelar, 2023), zeigen erste empirische Befunde, dass Lehrkräfteurteile für diese Schüler*innen ebenfalls negativ verzerrt sind und die Leistung unterschätzt wird (Umansky & Dumont, 2021). Vor diesem Hintergrund stellt sich die Frage, inwiefern die Einschätzungen von Lehrkräften zu systematischen Leistungsunterschieden zwischen Schüler*innen mit und ohne EL-Label beitragen.

Fragestellungen

- 1) Beeinflussen Lehrkräfteurteile die Schüler*innenleistung?
- 2) Beeinflusst das EL-Label Lehrkräfteurteile?
- 3) Mediiert Lehrkräfteurteile den Effekt des EL-Labels auf die Schüler*innenleistung?

Methode

Die Fragestellungen wurde mit dem Datensatz einer US-Längsschnittstudie (ECLS-K:2011) beantwortet. Schulen in den USA sind dazu verpflichtet, mehrsprachige Schüler*innen aufgrund ihrer Englischkenntnisse in EL und nicht-EL zu klassifizieren. Da sich die Schwellenwerte zur Klassifizierung zwischen Bundesstaaten unterscheiden, werden Schüler*innen mit den gleichen Englischkenntnissen in einigen Staaten als EL gelabelt, in anderen jedoch nicht (Abedi, 2008). Das erlaubt einen Vergleich der beiden Gruppen in einem natürlichen Experiment ($N[EL] = 2048$; $N[Nicht-EL] = 922$).

Als Baselinevariablen wurden im Kindergarten die Englisch- ($\alpha = .95$) und Mathematikleistung ($\alpha = .92$) mit standardisierten Leistungstest erfasst, exekutive Funktionen gemessen sowie das Geschlecht, ethnische Zugehörigkeit und SES als soziodemographische Merkmale erhoben. Die Lehrkräfteurteile zur Leistung wurden am Ende der ersten Klasse mit Fragebogenskalen für Englisch ($\alpha = 0.96$), Mathematik ($\alpha = .89$) und Naturwissenschaften ($\alpha = 0.96$) erfasst. Die Leistung der Schüler*innen wurde am Ende der zweiten Klasse nochmals über standardisierte Tests für Englisch ($\alpha = .93$), Mathematik ($\alpha = .93$) und Naturwissenschaften ($\alpha = .83$) ermittelt.

Nach der Imputation von 20 vollständigen Datensätzen wurden zu den Schüler*innen mit EL-Label mittels Propensity Score Matching Schüler*innen ohne Label hinsichtlich aller Baselinevariablen gematcht ($N[EL] = 789$; $N[Nicht-EL] = 789$). Strukturgleichungsmodelle wurden jeweils für Englisch, Mathematik und Naturwissenschaften geschätzt, indem die Baselinevariablen als Kontrollvariablen, das EL-Label als Prädiktor, die Einschätzung der Lehrkräfte als Mediator und die Leistung als Outcome modelliert wurden.

Ergebnisse und Bedeutung

Die Befunde zeigten für die erste Fragestellung, dass Lehrkräfteurteile die Leistung in Englisch ($\beta = 0.42$), Mathematik ($\beta = 0.31$) und Naturwissenschaften ($\beta = 0.17$) signifikant positiv vorhersagten (alle $p < .001$). Hinsichtlich der zweiten Fragestellung waren die Einschätzungen der Lehrkräfte für die Leistung in Englisch signifikant verzerrt. Schüler*innen mit EL-Label wurden unterschätzt ($\beta = -0.12$, $p = 0.03$). Für die dritte Fragestellung ergab sich, dass die Lehrkräfteurteile den Effekte des Labels auf die Leistung in Englisch mediierten und so zu einer ungünstigeren Leistungsentwicklung der gelabelten Schüler*innen führten ($\beta = -0.05$, $p = 0.02$). In Mathematik und Naturwissenschaften zeigten sich weder signifikante Effekte des EL-Labels auf die Lehrkräfteurteile noch signifikante indirekte Effekte.

Die Ergebnisse verdeutlichen die Bedeutung von Lehrkräfteurteilen für die Leistungsentwicklung von Schüler*innen und zeigen, dass neben diagnostischen auch andere Label Lehrkräfteurteile verzerren können. Der Effekt des EL-Labels scheint allerdings fachspezifisch zu sein. Lehrkräfte tragen durch ihre Einschätzungen zur systematischen Benachteiligung der Schüler*innen mit EL-Label in Englisch, nicht jedoch in Mathematik und Naturwissenschaften bei.

S-30: Merkmale von Unterrichtsqualität für Schüler:innen mit sonderpädagogischem Förderbedarf: Welche Erkenntnisse liegen vor?

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 150

Chair der Sitzung: Elisabeth Moser Opitz

Chair der Sitzung: Mireille Tabin

Chair der Sitzung: Jasmin Decristan

Merkmale von Unterrichtsqualität für Schüler:innen mit sonderpädagogischem Förderbedarf: Welche Erkenntnisse liegen vor?

Chair(s): **Elisabeth Moser Opitz** (Universität Zürich, Schweiz), **Mireille Tabin** (Universität Zürich, Schweiz)

Diskutant*in(nen): **Jasmin Decristan** (Universität Wuppertal)

Ein wichtiges Ziel der Unterrichtsforschung ist die Analyse von Faktoren, die sich positiv auf den Kompetenzerwerb und die Entwicklung der Schüler:innen auswirken. Bisherige Studien beziehen sich schwerpunktmäßig auf Schüler:innen der Sekundarstufe I und die Fächer Mathematik und Naturwissenschaften (Praetorius & Gräsel, 2021). Zur Erfassung der Unterrichtsqualität wurden häufig generische Dimensionen genutzt (Praetorius et al., 2018; Helmke, 2022). Bezüglich der Dimensionen zur Erfassung der Unterrichtsqualität und der berücksichtigten Fächer gibt es aktuell Erweiterungen. Neuentwicklungen wie das MAIN-TEACH Modell (Praetorius et al., 2023) beziehen sich neben dem Kompetenzerwerb auch auf Lernprozesse und ermöglichen die Erforschung von Unterrichtsqualität in verschiedensten Schulfächern.

Noch wenig diskutiert wurde im Kontext der Unterrichtsforschung die Frage nach qualitativ hochwertigem Unterricht für Schüler:innen mit sonderpädagogischem Förderbedarf (SPF). Dabei stellt sich erstens die Frage, ob die bislang entwickelten und genutzten Dimensionen und Subdimensionen von Unterrichtsqualität für *alle* Schüler:innen, d.h. auch für Lernende mit SPF, gelten bzw. in Studien genutzt werden. Zweitens interessiert, welche Erkenntnisse zu qualitativ hochwertigem Unterricht für diese Schüler:innen vorliegen.

Die Beantwortung und Untersuchung dieser Fragen ist aus mehreren Gründen herausfordernd. Die Gruppe der Schüler:innen mit SPF ist sehr heterogen. Zudem beeinflusst das Unterrichtsetting – inklusiv oder separativ – die Entwicklung von Lernenden mit SPF (z.B. Kocaj et al., 2014; Schnepel et al., 2022). Die Wirkung von Unterrichtsqualität kann somit nicht ohne Berücksichtigung des Förderortes analysiert werden. Auch erschweren die unterschiedlichen Voraussetzungen dieser Schüler:innen die standardisierte Erfassung von Outcomes (Dumont et al., 2024).

Vor diesem Hintergrund verfolgt das Symposium zwei Zielsetzungen:

Erstens wird betrachtet, inwiefern Merkmale und Dimensionen von Unterrichtsqualitätsmodellen, die für den Regelunterricht entwickelt wurden, genutzt werden, um die Unterrichtsqualität bei Schüler:innen mit SPF zu untersuchen. Zweitens wird exemplarisch anhand der Unterrichtsqualitätsdimension Adaptivität bzw. der Subdimension Differenzierung in zwei Studien untersucht, wie sich Differenzierungsmaßnahmen auf die Schüler:innen auswirken und welche Rolle Förderbedarfe bzw. die Kompetenzen der Lehrkräfte im Umgang mit SPF spielen.

Im ersten Beitrag wird entlang der Dimensionen des MAIN-TEACH-Modells (Praetorius et al., 2023) mittels eines Literaturreviews ein Überblick über internationale Studien zur Wirkung verschiedener Unterrichtsqualitätsdimensionen auf die Outcomes der Schüler:innen mit SPF im inklusiven und separativen Unterricht gegeben.

Der zweite Beitrag betrachtet mit einem Literaturreview anhand des Merkmalkatalogs von Helmke (2022), welche Unterrichtsqualitätsmerkmale in inklusiven Grundschulen in Deutschland untersucht werden.

Im dritten Beitrag wird längsschnittlich analysiert, ob binnendifferenzierende Maßnahmen die Entwicklung der fachspezifischen Leistung und des fachspezifischen Interesses aller Schüler*innen stärken und ob binnendifferenzierende Maßnahmen den Zusammenhang des SPF mit der Leistung und dem Interesse moderieren.

Im vierten Beitrag wird der Einfluss der Kompetenz von Regellehrkräften zur Leseförderung bei SPF auf die Differenzierungspraxis sowie der Effekt der Differenzierung auf den Einfluss, den die Motivation auf den Leistungszuwachs der Schüler:innen im Lesen hat, untersucht.

Die Ergebnisse der beiden Literaturreviews zeigen, dass die unterschiedlichen Zugänge (MAIN-TEACH Modell, Merkmalkatalog von Helmke) zu unterschiedlichen Ergebnissen führen. Die beiden Studien zur Differenzierung weisen auf komplexe Wechselwirkungen zwischen Variablen auf der Ebene der Lernenden (mit SPF) und Unterrichtsmerkmalen hin.

Ausgehend von diesen Ergebnissen werden Forschungsdesiderate und Möglichkeiten für zukünftige Studien diskutiert.

Beiträge des Symposiums

Welche Dimensionen werden zur Erforschung der Unterrichtsqualität für Schüler:innen mit sonderpädagogischem Förderbedarf genutzt?

Mireille Tabin¹, Elisabeth Moser Opitz¹, Karolina Urton², Anna Praetorius¹

¹Universität Zürich, CH, ²Universität Münster

Theoretischer Hintergrund

In den letzten Jahrzehnten hat sich die Forschung zur Unterrichtsqualität intensiv mit der Frage beschäftigt, was qualitativ hochwertigen Unterricht ausmacht. Die Vielzahl an dazu durchgeführten empirischen Studien bezieht sich häufig auf die Sekundarstufe I in Regelschulen und dort auf die Fächer Mathematik und Naturwissenschaften. Ein aktuelles Modell aus der Unterrichtsforschung, das die im Rahmen der bisherigen Forschung identifizierten Aspekte von Unterrichtsqualität zusammenfasst und explizit verschiedene Fächer berücksichtigt, ist das MAIN-TEACH-Modell (Praetorius et al., 2023). Es umfasst die Dimensionen Auswahl und Thematisierung der Inhalte, Lernziele und (Fach-)Methoden, kognitive Aktivierung, Unterstützung des Konsolidierens, Beurteilung und Feedback, Klassenführung, soziale Unterstützung sowie Unterstützung der Eigenverantwortung für das Lernen.

Unklar ist bislang, wie Unterrichtsqualität für Schüler:innen mit Förderbedarf (SPF) untersucht wird bzw. welche Dimensionen dafür verwendet werden. Inwiefern Modelle wie das MAIN-TEACH Modell genutzt werden zur Erforschung der Unterrichtsqualität bei Schüler:innen mit SPF in inklusiven, aber auch separativen Settings, wurde bislang nicht untersucht. Erkenntnisse zur dieser

Frage können als Ausgangspunkt dienen, um die Frage nach qualitativ hochwertigem Unterricht für Schüler:innen mit SPF weiter zu analysieren.

Fragestellung

Inwiefern wurden die Dimensionen und Subdimensionen des MAIN-TEACH-Modells genutzt, um die Wirkung von Unterrichtsqualität auf den Outcome bei Schüler:innen mit SPF zu untersuchen?

Methode

Es wurde ein Literaturreview durchgeführt. Die Suche bezog sich auf englisch- und deutschsprachige Publikationen in den elektronischen Datenbanken Web of Science und ERIC. Berücksichtigt wurden Artikel mit Peer-Review im Zeitraum bis Dezember 2023.

Die Analyse bezog sich auf die englischen Übersetzung der Dimensionen des MAIN-TEACH-Modells (z.B. adaptive teaching; siehe Praetorius et al., 2023) und umfasste die Keywords (1) Adaptivität, (2) Klassenführung, (3) soziale Unterstützung, (4) Unterstützung der Eigenverantwortung für das Lernen, (5) Beurteilung und Feedback, (6) Auswahl und Thematisierung der Lernziele, Inhalte und (Fach-)Methoden, (7) kognitive Aktivierung, (8) Unterstützung des Konsolidierens (Praetorius et al., 2023), kombiniert mit den Keywords, die sich auf Inklusion und sonderpädagogische Förderung (inclusion OR "inclusive education" OR "special education* needs" OR "special needs" OR SEN) sowie Outcomes (outcome* OR effect* OR impact*) beziehen.

Ergebnisse und ihre Bedeutung

Es wurden insgesamt 1105 Datensätze gefunden. 55 Studien wurden nach dem Titel-Abstract Screening in die engere Auswahl aufgenommen. Insgesamt fanden wir 14 Studien, in denen die Wirkungen der Unterrichtsqualität auf Schüler:innen mit SPF entsprechend der Einschlusskriterien untersucht wurden. Vier der 14 einbezogenen Studien nutzten qualitative Erhebungs- und Analysemethoden, die anderen Studien quantitative. Mit Ausnahme der kognitiven Aktivierung, zu der keine Studien gefunden wurden, wurden für jede Dimension des MAIN-TEACH Modells eine bis vier Studien identifiziert. Die meisten Studien wurden für die Dimension der Klassenführung gefunden (n = 4). Die untersuchten Zieldimensionen bzgl. der Wirkung von Unterricht beziehen sich auf soziale Akzeptanz, die sozialen Kompetenzen oder/und die exekutiven Funktionen (Konzentrationsprobleme, störendes Verhalten), jedoch nicht auf die Leistung.

Da insgesamt nur wenige Studien gefunden wurden, soll im Rahmen des Vortrages auch die Frage diskutiert werden, welche anderen Kriterien oder Dimensionen zur Analyse der Unterrichtsqualität bei Schüler:innen mit SPF verwendet werden bzw. verwendet werden könnten und welche theoretischen Modelle und empirischen Erkenntnisse als Grundlage für weitere Recherchen genutzt werden könnten.

Wie ist die Unterrichtsqualität an inklusiven Grundschulen in Deutschland ausgeprägt? Ein systematischer Literaturreview

Katja Martin, Stefanie Bosse, Nadine Spörer
Universität Potsdam

Theoretischer Hintergrund

In der letzten Dekade sind Grundschulen als Lernorte für Kinder mit sonderpädagogischem Förderbedarf (SPF) zunehmend bedeutender geworden (Autor:innengruppe Bildungsberichterstattung, 2024). Der Forschungsstand zeigt, dass Kinder mit SPF, die inklusiv beschult werden, eine günstigere kognitive Kompetenzentwicklung aufweisen als solche, die in Förderschulen beschult werden (Kocaj et al., 2014; Ruijs & Peetsma, 2009; Schnepel et al., 2022). Ein wichtiger Prädiktor für die Kompetenzen von Schulkindern ist die Unterrichtsqualität (Hattie, 2009). Im umfassend rezipierten Angebots-Nutzungs-Modell von Helmke (2017, 2022) werden zehn generisch konzipierte Prozessqualitätsmerkmale guten Unterrichts differenziert (z. B. Klassenführung, lernförderliches Klima), auf denen jüngere Unterrichtsqualitätsmodelle aufbauen (Praetorius et al., 2018, Praetorius & Gräsel, 2021). Ein aktuelles Literaturreview zur Unterrichtsqualität an Förderschulen ergab, dass die generisch angelegten Dimensionen des Angebots-Nutzungs-Modell eine geeignete Struktur für eine differenzierte Analyse der Qualitätsdimensionen darstellen (Autor:innen, 2024). Zugleich mangelt es bislang an Forschungsarbeiten, die die Unterrichtsqualität an inklusiven Grundschulen systematisch zusammenfassen.

Forschungsfragen

Das Ziel dieses Beitrags ist es, bisherige Forschungsergebnisse zur Unterrichtsqualität inklusiver Grundschulen in Deutschland systematisch zu analysieren und synthetisieren. Dabei sollen folgende Forschungsfragen mit einem Literaturreview beantwortet werden: Welche Merkmale von Unterrichtsqualität inklusiver Grundschulen werden in der Forschung adressiert? Wie hoch ist die Unterrichtsqualität an inklusiven Grundschulen?

Methodik

Die Literatursuche erfolgte in acht Datenbanken (z. B. PsycINFO, FIS Bildung) mit einer Suchsyntax aus 100 logischen Kombinationen deutscher Schlüsselwörter, welche u.a. die zehn Unterrichtsqualitätsmerkmale adressierten (z. B. Inklusion UND Schule UND Motivierung). Anschließend wurden die N = 8586 ermittelten Publikationen in zwei Schritten von zwei trainierten unabhängigen Rater:innen hinsichtlich ihrer Relevanz beurteilt. Dazu wurden Einschlusskriterien (empirische Studie, Bezug zum deutschen Schulsystem, Bezug zu Inklusion, zur Grundschule und zur Unterrichtsqualität) und Ausschlusskriterien (ausschließlich kein Bezug zu Deutschland, zur Inklusion, zur Grundschule und zur Unterrichtsqualität) aufgestellt.

Im ersten Selektionsschritt wurden auf Basis des Titels Publikationen ausgeschlossen, wenn mindestens ein Ausschlusskriterium erfüllt war. Es verblieben N = 2103 Publikationen im Datensatz ($\kappa = 0.78$). Auf Basis des Abstracts wurde im zweiten Selektionsschritt die Eignung anhand der Ein- und Ausschlusskriterien beurteilt. Für die Volltextanalyse wurden Publikationen eingeschlossen, wenn alle Einschlusskriterien, jedoch kein Ausschlusskriterium, erfüllt waren. Dies traf auf N = 25 Publikationen zu ($\kappa = 0.90$).

Die Analyse der Publikationen erfolgt durch zwei unabhängige Rater:innen. Um den Ausprägungsgrad der Unterrichtsqualität einzuschätzen, wurden drei Stufen unterschieden: eher gering, mittel und eher hoch. Die Kategorisierung erfolgte in Abhängigkeit von den Studiendaten. Wenn eine Studie bereits eine Kategorisierung (z. B. basierend auf Normwerten) aufwies, wurde diese zur Einordnung genutzt. Waren in einer Studie Maße der zentralen Tendenz (z. B. Mittelwert) vorhanden, wurden diese Werte zur Klassifizierung herangezogen. Dabei wurde die verwendete Messskala gedrittelt.

Ergebnisse und ihre Bedeutung

Bislang wurden N = 11 Publikationen kodiert. Die Majorität bildeten quantitative (N = 9) Querschnittsstudien (N = 8). Die Stichprobengröße variierte zwischen N = 33 bis 350 Personen. Es wurden zwei der zehn Unterrichtsqualitätsmerkmale nicht adressiert: Konsolidierung und Kompetenzorientierung. Die Anzahl der Studien pro Unterrichtsqualitätsmerkmal schwankte zwischen N = 1 und 7. Am häufigsten wurden der Umgang mit Heterogenität (N = 7 Studien), die Methodenvielfalt (N = 4 Studien) sowie die Schüler:innenorientierung (N = 4) untersucht. Insgesamt lag eine mittlere Prozessqualität des Unterrichts vor. Die bislang einbezogenen Studien zeigten an, dass das lernförderliche Klima und die Klassenführung eher eine hohe Qualität aufwiesen. Die Qualität der Strukturierung und Motivierung war mittel bis hoch ausgeprägt. Eine mittlere Qualität wiesen die Dimensionen Schüler:innenorientierung und Methodenvielfalt auf. Eher gering bis mittel fiel die Differenzierungsqualität aus. Die Aktivierungsqualität war als eher gering einzustufen. Anzumerken ist, dass die Vergleichbarkeit der Studien aufgrund von stark variierenden Methoden und Stichproben limitiert ist. Bis zur Tagung wird eine vollständige Analyse aller Studien erfolgen

Binnendifferenzierung als Unterrichtsqualitätsmerkmal und Effekte auf die leistungs- und interessenbezogene Entwicklung von Grundschulkindern

Jenny Lenkeit, Nadine Spörer, Anne Hartmann, Michel Knigge, Antje Ehlert
Universität Potsdam

Theoretischer Hintergrund

Um heterogene Lernausgangslagen produktiv zu nutzen und positive Entwicklungen aller Schüler*innen im inklusiven Unterricht zu ermöglichen, wird insbesondere die Binnendifferenzierung als bedeutsames Merkmal von Unterrichtsqualität betrachtet (Gehrer & Nusser, 2020; Lindner et al., 2021). Diese ordnet sich jüngst auch in das MAIN-TEACH-MODELL zur Unterrichtsqualität in der Dimension „Adaptivität“ ein (Charalambous & Praetorius, 2020; Praetorius & Gräsel, 2021).

Zentrale Zielgrößen eines qualitativ vollen Unterrichts sind die Leistungsentwicklung und das fachspezifische Interesse, die wiederum von einer Vielzahl von individuellen Merkmalen der Schüler*innen wie ihren kognitiven, motivationalen und sozial-emotionalen Lernvoraussetzungen beeinflusst werden (Maaz et al., 2014). Insbesondere in inklusiven Lehr- und Lernumgebungen nimmt auch der sonderpädagogische Förderbedarf (SPF) von Schüler*innen eine zentrale Rolle für die leistungs- und interessenbezogene Entwicklung ein (Kocaj et al., 2016; Kocaj et al., 2020; Lütje-Klose et al., 2018). Werden Schüler*innen mit einem SPF nicht angemessen unterstützt, können Schereneffekte für die leistungs- und interessenbezogene Entwicklung angenommen werden (Gebhardt, 2023).

Mit dem Unterrichtsqualitätsmerkmal Binnendifferenzierung verbunden ist die Annahme, dass ein differenziertes, individuell förderndes Lernangebot an den unterschiedlichen Lernausgangslagen der Schüler*innen anknüpft und Schereneffekte mindern kann (Helmke, 2017). Aus den bisherigen Studien geht jedoch hervor, dass es für die postulierten positiven Effekte von Binnendifferenzierung auf die fachspezifische Leistungs- und Interessenentwicklung im inklusiven Grundschulunterricht bislang, insbesondere für Deutschland, nur wenige und zugleich divergierende Befunde gibt (Deunk et al., 2018; Pozas et al., 2021).

Fragestellungen

Vor dem dargestellten Hintergrund untersucht der Beitrag zwei zentrale Fragestellungen.

- (1) Stärken binnendifferenzierende Maßnahmen die Entwicklung der fachspezifischen Leistung und des Interesses für alle Schüler*innen?
- (2) Moderieren binnendifferenzierende Maßnahmen den Zusammenhang des SPF mit der Leistung und dem Interesse?

Methode

Die Daten entstammen einer längsschnittlichen Untersuchung zu den Herausforderungen und Gelingensbedingungen inklusiv arbeitender Schulen aus den Jahren 2018-2020 und umfassen 635 Schüler*innen aus 33 Klassen der 2. und 3. Jahrgangsstufe. Folgende Variablen und Skalen wurden operationalisiert:

Leistungsentwicklung in Mathematik (Heidelberger Rechentest (HRT 1-4; Haffner et al., 2005), erfasst über zwei Messzeitpunkte.

Mathematikinteresse (Bos et al., 2005), erfasst für vier Items auf einer vierstufigen Likertskala von 0 = „stimmt gar nicht“ bis 3 = „stimmt genau“ (Beispielitem „Ich rechne gern.“), $\alpha_1 = 0.77$, $\alpha_2 = 0.81$.

SPF: Angaben der Klassenleitungen zu jedem bzw. jeder Lernenden, ob ein festgestellter oder vermuteter SPF vorliegt.

Binnendifferenzierung: Informationen von 33 Mathematiklehrkräften zur differenzierenden Unterrichtsgestaltung in 281 Mathematikstunden, erfasst über ein standardisiertes Logbuch (Eigenentwicklung) im ersten Schulhalbjahr. Lehrkräfte machten über einen Zeitraum von vier Wochen Angaben zu je 2 Unterrichtsstunden pro Woche zu quantitativen (3 Items, Bsp. „Schüler*innen erhielten unterschiedliche Zeitvorgaben.“) und qualitativen (4 Items, Bsp.: „Schüler*innen bearbeiteten Aufgaben mit unterschiedlichem Schwierigkeitsgrad.“) Differenzierungen je in Einzel- (EA) oder Gruppenarbeitsphasen (GA). Die Angaben wurden über vier Summenindizes (EA quantitativ, EA qualitativ, GA quantitativ, GA qualitativ) operationalisiert und pro Lehrkraft über die Unterrichtsstunden gemittelt.

Zusammenhänge wurden mithilfe von hierarchisch linearen Regressionsmodellen in MPlus geprüft.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass qualitative und quantitative Differenzierungsmaßnahmen in Gruppenarbeitsphasen die Leistungsentwicklung innerhalb eines Schuljahres positiv stärken ($p < .05$), jedoch gleichzeitig negativ mit der Entwicklung des Mathematikinteresses assoziiert sind ($p < .05$).

Darüber hinaus moderiert die Häufigkeit quantitativer Differenzierungsmaßnahmen in Einzelarbeitsphasen den negativen Zusammenhang zwischen einem SPF und der Leistung zum Schuljahresende ($p < .05$). Dieser fällt schwächer aus, je häufiger quantitativ differenziert wird (Interaktionseffekt).

Die Ergebnisse deuten insgesamt darauf hin, dass verschiedene Formen von Binnendifferenzierung unterschiedlich auf Lernindikatoren wirken können und SPF-bedingte differenzielle Leistungsentwicklungen mindern können.

Zu diskutieren bleibt die Frage, inwiefern mit dem Logbuch ein innovatives Instrument zur Erfassung binnendifferenzierender Maßnahmen entwickelt werden konnte und für welche zukünftigen Forschungsfragen dieses eingesetzt werden kann.

Die Bedeutung der Differenzierung und der Motivation für die Leseleistung von Lernenden mit Förderbedarf

Theoretischer Hintergrund

Die Forschung zu Wirkzusammenhängen im Unterricht zeigt, dass sowohl die Kompetenz der Lehrkräfte als auch die Unterrichtsqualität einen Einfluss auf den Leistungszuwachs der Lernenden haben (Baumert et al., 2010; Hill, Ball & Schilling, 2008; Kunter et al., 2013). Im inklusiven Unterricht kommt insbesondere der Differenzierung, als ein Merkmal von Unterrichtsqualität, eine hohe Bedeutung zu (Lindner et al., 2021; Praetorius & Gräsel, 2021). Jedoch sind Effekte der Differenzierung auf den Leistungszuwachs von Lernenden mit Förderbedarf (FB) nur schwer nachzuweisen, da die Leistungsentwicklung von vielen weiteren Faktoren abhängt, z.B. Interessen, Erfahrungen, sozioökonomischer Status (Valiandes, 2015) oder Merkmale, die mit einem spezifischen sonderpädagogischen FB einhergehen. Ein weiterer dieser Faktoren ist die Motivation der Lernenden. Insbesondere im Lesen hat sich gezeigt, dass eine hohe Motivation einen positiven Effekt und eine niedrige Motivation einen negativen Effekt auf den Leistungszuwachs haben kann (Schiefele et al., 2012). Der Zusammenhang von Motivation und Leseleistung ist bei Lernenden mit FB im Lesen sogar besonders stark (Logan et al., 2011). Da Lernende mit FB jedoch häufig eine niedrigere Lernmotivation aufweisen, ist bei ihnen die Gefahr eines geringen Leistungszuwachses besonders hoch (Gorges et al., 2022).

Ob die Differenzierung des Unterrichts den Effekt der Motivation auf die Leistungsentwicklung verringern kann, ist bislang jedoch nicht untersucht worden. Auf diesem Hintergrund interessiert der Zusammenhang zwischen der professionellen Kompetenz der Lehrkräfte zur Leseförderung von Lernenden mit FB, von Differenzierungsmaßnahmen, der Motivation der Lernenden sowie deren Leistungszuwachs.

Fragestellungen

Untersucht werden folgende Fragen:

Wie hängt die professionelle Kompetenz der Lehrkräfte zur Leseförderung von Lernenden mit FB mit der Umsetzung von Differenzierungsmaßnahmen im Unterricht zusammen?

Wird der Effekt der Motivation auf den Leistungszuwachs im Lesen durch die Differenzierung moderiert?

Methode

Stichprobe

Durchgeführt wurde eine Längsschnittstudie mit N = 76 Regellehrkräften von inklusiven Klassen mit N = 1269 Schüler:innen im 2. bis 4. Schuljahr. Offizielle sonderpädagogische FB wurden in der untersuchten Stichprobe nicht vergeben. Der FB wird daher durch unterdurchschnittliche Leistungen im Lesen definiert (T-Wert < 35, n = 176).

Erhebungen Schüler:innen

Leseverständnis ELFE-Test (Lenhard et al., 2017) zu Beginn (t1) und am Ende des Schuljahres (t2).

Motivation schulisches Lernen t1: selbstentwickelter Fragebogen, 5 Items, z.B. „Findest du es spannend, was du in der Schule machst?“, vierstufige Likert-Skala (nein, eher nein, eher ja, ja), Cronbachs alpha = .74

Allgemeine Denkfähigkeit: CFT 1-R (Weiß & Osterland, 2013) bzw. CFT 20-R (Weiß, 2006).

Erhebungen Lehrkräfte

Professionelle Kompetenz Leseförderung: Rating von zwei Interviews pro Lehrkraft zum Thema Leseförderung mit Bezug zu ausgewählten Lernenden mit FB Lesen (vierstufiges Item, Übereinstimmung Rater:innen g = .87)

Differenzierung: Rating von zwei beobachteten Unterrichtsstunden pro Lehrkraft: Angebot zieldifferenter Aufgaben, Hilfestellungen und Arbeitsmittel (vierstufiges Item, Übereinstimmung Rater:innen g = .78)

Analysen

Mehrebenen-Pfadmodell mit einem Cross-Level-Interaktionsterm mit dem Programm Mplus

Ebene 1: Motivation zu t1, Leseleistung zu t1, Familiensprache, Geschlecht, allgemeine Denkfähigkeit als Prädiktoren von Leseleistung zu t2

Ebene 2: professionelle Kompetenz zur Leseförderung als Prädiktor von Differenzierung; Differenzierung als Prädiktor von Leseleistung zu t2

Cross-Level-Interaktionseffekt: Differenzierung (Ebene 2) und Motivation (Ebene 1) auf Leseleistung t2

Ergebnisse und ihre Bedeutung

Der Effekt der professionellen Kompetenz zur Leseförderung von Lernenden mit FB wirkt sich tendenziell positiv auf die Umsetzung von Differenzierungsmaßnahmen im Unterricht aus ($\beta = .21$, $p = .075$). Die Differenzierung wiederum wirkt sich signifikant negativ auf den positiven Zusammenhang von Motivation und Leseleistung t2 aus ($\beta = -.24$, $p < .01$), unter Kontrolle von Leseleistung zu t1, Familiensprache, Geschlecht und allgemeiner Denkfähigkeit.

Die Ergebnisse weisen erstens auf die Bedeutung der Kompetenz der Lehrkräfte zur Leseförderung von Schüler:innen mit FB für die Umsetzung von Differenzierungsmaßnahmen im Lesen hin. Zweitens scheint der Lesezuwachs durch Differenzierungsmaßnahmen trotz niedrigerer Motivation positiv beeinflusst werden zu können.

S-31: KI im Bildungskontext: Angebote von Lehrkräften und Nutzungsverhalten von Schüler:innen

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: M 003

Chair der Sitzung: Jana Kemmler

Chair der Sitzung: Jessica Heesen

KI im Bildungskontext: Angebote von Lehrkräften und Nutzungsverhalten von Schüler:innen

Chair(s): **Jana Kemmler** (Universität Tübingen, Deutschland)

Diskutant*in(nen): **Jessica Heesen** (Universität Tübingen)

Die zunehmende Integration von Künstlicher Intelligenz (KI) im Bildungswesen erfordert, dass angehende Lehrkräfte über die notwendigen Kompetenzen verfügen, um diese Technologien didaktisch fundiert und ethisch reflektiert einzusetzen. Die Kompetenzentwicklung muss hierbei unterstützt sowie erfasst werden. Abgesehen von dem Angebot, das Lehrkräfte in ihrem Unterricht erschaffen, bleibt die Frage offen, ob und wie sich diese Angebote auf die Nutzungsart der Schüler:innen auswirken. Die Beiträge dieses Symposiums untersuchen im Rahmen des Angebot-Nutzungsmodells die Aspekte von KI- und Bildungsdatenkompetenz als Angebot von Lehrkräften sowie die Nutzungsperspektive von KI durch Schüler:innen.

Beitrag 1 beleuchtet die Wirksamkeit instruktionaler Unterstützung auf motivationaler (Utility-Value Intervention) und kognitiver Ebene (Worked Example) zur Unterstützung der Wissensintegration angehender Lehrkräfte in Bezug auf kritisch-reflexive KI-Kompetenz. Durch ein 2x2-Design werden Haupt- und Interaktionseffekte dieser Interventionen analysiert. Die Ergebnisse zeigen, wie motivationale und kognitive Unterstützung zur Förderung umfassender KI-Kompetenzen beiträgt, die angehenden Lehrkräften helfen, fundierte Entscheidungen im Bildungswesen zu treffen.

Beitrag 2 untersucht die Eignung eines validierten AI Literacy Tests zur Messung der Veränderungen von AI Literacy im Rahmen von Lehrveranstaltungen. Eine Studie zur Wirksamkeit eines Seminars zur Vermittlung grundlegender KI-Kompetenzen bei Lehramtsstudierenden wird beschrieben, wobei kognitive und affektive Variablen wie Selbstwirksamkeit und Einstellungen gegenüber KI erfasst werden. Die Ergebnisse zeigen signifikante Verbesserungen hinsichtlich AI Literacy und Selbstwirksamkeit sowie eine Verringerung negativer Einstellungen gegenüber KI, was die Eignung des Tests für Lehrevaluationen unterstreicht.

Beitrag 3 adressiert die Notwendigkeit und Erforschung der Bildungsdatenkompetenz bei angehenden Lehrkräften. Es wird untersucht, wie diese Kompetenzen von Lehramtsstudierenden eingeschätzt werden und welche Arten von Daten sie in ihrer Praxis nutzen. Die Ergebnisse zeigen eine mittelmäßige Selbsteinschätzung der Kompetenz und eine Diskrepanz zwischen verfügbaren und gewünschten Daten, was auf einen erheblichen Verbesserungsbedarf in der Lehrkräfteausbildung hinweist.

Beitrag 4 befasst sich tiefergehend mit dem Nutzungsverhalten von KI-Tools durch Schülerinnen und Schüler am Ende der Sekundarstufe I und fokussiert hierzu insbesondere unterschiedliche Bereiche der schulbezogenen instrumentellen Nutzung. Die Ergebnisse legen nahe, dass Schülerinnen und Schüler KI-Tools qualitativ und quantitativ höchst unterschiedlich einsetzen. Sie unterstreichen die Notwendigkeit einer tiefergehenden Untersuchung der unterschiedlichen Nutzungsprofile und einer möglicherweise damit einhergehenden differenziellen Wirksamkeit.

Zusammengefasst bieten diese Beiträge wertvolle Einblicke in die Förderung und Erfassung von KI-Kompetenzen und verdeutlichen die Notwendigkeit einer umfassenden Bildung, die technisches Wissen um ethisch-gesellschaftliche Implikationen erweitert, um den Herausforderungen der modernen Bildungslandschaft gerecht zu werden.

Beiträge des Symposiums

Künstliche Intelligenz und Ethik: Förderung kritisch-reflexiver KI-Kompetenz von Lehramtsstudierenden

Jana Kemmler¹, Salome Wagner², Iris Backfisch¹, Andreas Lachner¹

¹Universität Tübingen, ²Leibniz-Institut für Bildungsforschung und Bildungsinformation

Theoretischer Hintergrund und Fragestellung

Der Einsatz von Schlüsseltechnologien wie Künstlicher Intelligenz (KI) im Unterricht erfordert von Lehrkräften kritisch-reflexive KI-Kompetenz, um KI didaktisch begründet und ethisch reflektiert zu verwenden. Dafür ist integratives Wissen notwendig, das sowohl technische als auch ethisch-gesellschaftliche Kenntnisse umfasst (Long & Magerko, 2020). Die Integration von Wissen ist jedoch kognitiv anspruchsvoll und erfordert hohe motivationale Voraussetzungen, insbesondere für Novizen wie angehende Lehrkräfte (Barzilai et al., 2018; Lehmann, 2024).

Ziel der Studie war es daher, zu untersuchen, wie auf kognitiver und motivationaler Ebene die Wissensintegration im Kontext von KI und Ethik gefördert werden kann. Spezifisch überprüfen wir die Wirksamkeit motivationaler (Utility-Value Intervention) und kognitiver Scaffolds (Worked Example) zur Unterstützung der Wissensintegration von angehenden Lehrkräften in Bezug auf kritisch-reflexive KI-Kompetenz. Auf der Grundlage bisheriger Forschung stellten wir die Hypothesen auf, dass angehende Lehrkräfte, welche die Utility-Value Intervention (Backfisch et al., 2024) und/oder ein Worked Example erhielten (Barbieri et al., 2023), ein höheres Maß an Wissensintegration aufzeigen und höhere Testergebnisse erzielen als diejenigen, die keine Unterstützung erhielten.

Methode

Die experimentelle Feldstudie war als vorbereitende Selbstlernaufgabe in einer Universitätsvorlesung in der Lehrkräfteausbildung eingebettet und umfasste N = 274 Studierende (69% weiblich, Mage = 24.58, SDage = 2.2, Msemester = 9.13, SDsemester = 2.84). Die Studie wurde als 2x2-Between Participants Design (Utility-Value Intervention [ja/nein], Worked Example [ja/nein]) umgesetzt.

Nach einem Prätest erhielten die Teilnehmenden entweder eine videobasierte Utility-Value-Intervention als Prämaßnahme mit einer konsolidierenden Schreibaufgabe oder nicht (siehe Canning et al., 2018). Die anschließende Hypertextlernumgebung enthielt verlinkte Hypertexte zu technologischen und ethisch-gesellschaftlichen Aspekten von KI, welche die Teilnehmenden selbstreguliert bearbeiteten. Abschließend erhielten die Teilnehmenden entweder ein Worked Example als Schritt-für-Schritt Lösungsbeispiel zur Wissensintegration der verschiedenen Inhalte oder nicht. Zum Abschluss absolvierten die Studierenden direkt den Posttest anhand von acht Multiple-Choice Fragen (naher Transfer). Zusätzlich enthielt der Posttest vier offene Fragen (von zwei unabhängigen Ratern kodiert, ICC > .72) für die Prüfung von fernem Transfer, die lebensweltnahe Herausforderungen adressierten und technisches Wissen mit ethischen Überlegungen in gesellschaftlichen Kontexten verknüpften.

Ergebnisse und ihre Bedeutung

Die Teilnehmenden unterschieden sich nicht signifikant in ihren Ausgangsvoraussetzungen. Zur Überprüfung der Hypothesen wurden ANCOVAS jeweils mit der nahen bzw. fernen Posttestperformance als abhängige Variable, der Utility-Value Intervention und dem Worked Example und deren Interaktion als Faktoren und der Prätestperformance als Kontrollvariable berechnet.

Die Analysen zeigten einen signifikanten Haupteffekt von Worked Examples sowohl auf den nahen, $F(1, 269) = 6.093, p = .014, \eta^2 = .02$, als auch auf den fernen Transfer, $F(1, 269) = 4.165, p = .042, \eta^2 = < .01$, was zeigt, dass Lehramtsstudierende, die ein Worked Example erhielten, ein höheres Maß an Wissensintegration in Nah- und Ferntransferaufgaben aufwiesen als Lehramtsstudierende, die das Worked Example nicht erhielten. Die Utility-Value Intervention zeigte keinen signifikanten Effekt auf den Nahtransfer und Ferntransfer ($p > .203$). Es ergaben sich auch keine Interaktionseffekte. Weiterführende Analysen deuteten darauf hin, dass eine tiefere Verarbeitung während der Utility-Value Intervention zu einer höheren Leistung im Ferntransfer führte, $F(1, 131) = 18.3, p < .0001, \eta^2 = .12$.

Die Studie zeigt geeignete Unterstützungsmaßnahmen für Novizen in der Wissensintegration innerhalb des Lernprozesses zu kritisch-reflexiver KI-Kompetenz auf kognitiver und motivationaler Ebene auf. In einer weiteren Studie mit demselben Design und denselben Materialien ist geplant, die bisherige Stichprobe zur Untersuchung der Generalisierbarkeit, um den Vergleich mit Nicht-Lehramtsstudierenden zu erweitern. Auf der GEBF-Tagung werden die Ergebnisse beider Studien vorgestellt.

Validierung eines AI Literacy Tests zur Erfassung des Lernerfolgs von angehenden Lehrkräften

Marie Hornberger, Arne Bewersdorff, Claudia Nerdel
Technische Universität München

Theoretischer Hintergrund und Fragestellung

Um KI sinnvoll und ethisch in der Schule einzusetzen, benötigen Lehrkräfte ein grundlegendes Verständnis von KI und die Fähigkeit, KI-Tools zu nutzen und kritisch zu bewerten (sog. AI Literacy, Celik, 2023; Long & Magerko, 2020; Sperling et al., 2024). Neben AI Literacy sind auch affektive Variablen wie Akzeptanz relevant für die Nutzung von KI-Tools im Schulkontext (Mesenhöller & Böhme, 2024; Zhang et al., 2023). Um AI Literacy zu vermitteln, werden neue Lernangebote in der Aus- und Weiterbildung von Lehrkräften entwickelt. Bisher gibt es aber nur wenig Forschung zur Überprüfung der Wirksamkeit dieser Angebote (Mesenhöller et al., 2024; Nazaretsky et al., 2022). Eingesetzt werden dabei häufig Tests, die auf Selbsteinschätzungen beruhen (Ding et al., 2024; Laupichler et al., 2023; Mesenhöller et al., 2024) und seltener Wissenstests (Jetzinger et al., 2024; Korte et al., 2024).

In der berichteten Studie sollte untersucht werden, ob sich ein bereits validierter AI Literacy Test (Hornberger et al., 2023) für eine objektive Messung der Veränderung von AI Literacy im Rahmen von Lehrveranstaltungen eignet. In dieser Validierungsstudie mit 1286 Studierenden verschiedener Fächer wurde gezeigt, dass der Test ein breites Fähigkeitsspektrum abdecken kann. Nun soll untersucht werden, ob die AI Literacy der Studierenden im Verlauf der Lehrveranstaltung zunimmt und inwiefern sich Selbstwirksamkeit bzgl. KI, Interesse an KI, Einstellung gegenüber KI und soziale Akzeptanz von KI in der Schule positiv verändern (Hornberger et al., 2023; Mesenhöller & Böhme, 2024).

Methode

Die Stichprobe umfasst 12 Studierende aus dem WiSe 2023/24 (6 weiblich, 6 männlich, Alter: $M = 27.1, SD = 4.1$) und 10 Studierende aus dem SoSe 2024 (7 weiblich, 3 männlich, Alter: $M = 26.2, SD = 2.3$). Untersucht wurde ein Seminar, das sich an Lehramtsstudierende aller Fachrichtungen richtet und AI Literacy im Bildungskontext vermitteln soll. Im ersten Teil des Seminars werden grundlegende Konzepte der Künstlichen Intelligenz erarbeitet und in Programmieraufgaben vertieft. In einem zweiten Teil werden die Themen „Lernen mit KI“ und „Lernen über KI“ und ethische und gesellschaftliche Fragestellungen diskutiert. In der ersten und in der letzten Sitzung des Seminars wurde jeweils die Prä- bzw. Post-Erhebung mit den oben erwähnten Fragebögen durchgeführt (Hornberger et al., 2023; Mesenhöller & Böhme, 2024).

Ergebnisse und ihre Bedeutung

Da sich das Ausgangsniveau der AI Literacy zwischen den beiden Kohorten stark unterscheidet ($M = 12.7, SD = 3.4$ vs. $M = 17.5, SD = 4.9$, maximale Punktzahl: 30), wurden die Daten getrennt analysiert. Im Wintersemester zeigt sich eine signifikante Zunahme der AI Literacy vom Prätest ($M = 12.7, SD = 3.4$) zum Posttest ($M = 15.9, SD = 3.6, t(11) = 4.2, p = .002$). Im Sommersemester gibt es eine kleinere, nicht statistisch signifikante Zunahme der AI Literacy ($M_{\text{prä}} = 17.5, SD_{\text{prä}} = 4.9, M_{\text{post}} = 18.1, SD_{\text{post}} = 5.2, t(11) = 2.0, p = .067$). Aus Platzgründen werden die weiteren Ergebnisse nur für die Kohorte des Wintersemesters berichtet, sie unterscheiden sich aber nur unwesentlich. Im Verlauf der Lehrveranstaltung zeigt sich eine signifikante Steigerung der Selbstwirksamkeit bzgl. KI ($M_{\text{prä}} = 2.6, SD_{\text{prä}} = 0.7, M_{\text{post}} = 4.0, SD_{\text{post}} = 0.5, t(11) = 4.5, p < .001$) und eine signifikante Verringerung der negativen Einstellung gegenüber KI ($M_{\text{prä}} = 2.9, SD_{\text{prä}} = 0.8, M_{\text{post}} = 2.2, SD_{\text{post}} = 0.5, t(11) = -3.0, p = .01$). Die positive Einstellung gegenüber KI und das Interesse an KI erhöhen sich leicht, aber nicht statistisch signifikant. Die Ergebnisse zeigen, dass sich im Verlauf der Lehrveranstaltung die AI Literacy erhöht und affektive Variablen verbessern. Die Ergebnisse deuten darauf hin, dass sich der AI Literacy Test für den Einsatz in Lehrvaluationen eignet und eine objektive Messung der Veränderung von AI Literacy ermöglicht.

Empowerment von angehenden Lehrpersonen durch Bildungsdatenkompetenz

Clara Schumacher¹, Dirk Ifenthaler²

¹Universität Potsdam, ²Universität Mannheim & Curtin University

Theoretischer Hintergrund und Fragestellung

Bildungsdaten wie Noten, Vergleichsarbeiten oder Befragungen werden heutzutage erweitert durch Daten aus digital-gestützten Lernumgebungen, welche mit künstlicher Intelligenz wie Learning Analytics erfasst, analysiert und visualisiert werden können. Die dafür nötige Bildungsdatenkompetenz beinhaltet Daten zu sammeln, zu interpretieren und unter Berücksichtigung theoretischer und ethischer Annahmen für pädagogische Entscheidungen nutzbar zu machen (Mandinach et al., 2015). Bildungsdatenkompetenz ist noch wenig erforscht und in der Ausbildung angehender Lehrpersonen nicht ausreichend gefördert (Beck et al., 2020). Daher soll zunächst erfasst werden, wie angehende Lehrpersonen ihre Bildungsdatenkompetenz einschätzen (Forschungsfrage 1) und mit welchen Daten sie in ihrer vorbereitenden Schulpraxis konfrontiert werden (Forschungsfrage 2). Learning Analytics verwenden häufig Daten, die leicht zu erfassen sind (Jivet et al., 2017), daher sollen (angehende) Lehrpersonen aktiv eingebunden werden, bei der Frage nach nützlichen Bildungsdaten (Forschungsfrage 3).

Methode

Dreißig Lehramts-Masterstudierende mit $M=6.58 (SD=2.17)$ Monaten Schulpraxis haben teilgenommen. Die Studie umfasst ein qualitatives leitfadengestütztes Video-Interview, eine Einführung zu Daten im schulischen Kontext, Prototyping eines

Lehrendendashboards und einen quantitativen Online-Fragebogen. Die Bildungsdatenkompetenz wurde mittels 17 Items (Cronbach's α = .725; Likertskala, 1=trifft nicht zu bis 5=trifft voll und ganz zu) sowie durch Selbsteinschätzung („Meine Bildungsdatenkompetenz ist...“ von 1=sehr gering bis 5=sehr hoch) erfasst.

Die Video-Interviews wurden transkribiert und mittels strukturierender Inhaltsanalyse anhand eines deduktiv und induktiv erweiterten Kategoriensystems ausgewertet.

Ergebnisse und ihre Bedeutung

Zusammenfassend wird die Bildungsdatenkompetenz mit $M=3.65$ ($SD=.40$) angegeben. Sechzig Prozent der Befragten schätzen ihre Bildungskompetenz mittel (3) ein, 20% hoch (4), 14% gering (2), und jeweils 3% sehr gering (1) oder sehr hoch (5) ein ($M=3.07$; $SD=.785$).

Die eigene Datennutzung in der Schulpraxis beinhaltet insbesondere Leistungsdaten (35 Nennungen) wie Noten, Vergleichsarbeiten oder Beteiligung an der Leistungsbewertung sowie eigene Beobachtungen der Befragten. Zugängliche Hintergrundinformationen (17 Nennungen) sind insbesondere Namen, Alter sowie Förderbedarfe. Verwendete Daten zur Organisation (17 Nennungen) sind Stundenpläne, Sitzpläne bzw. Namenslisten und Abwesenheiten. Genutzte Daten zur Unterrichtsdokumentation (8) waren Informationen über behandelte Themen, Schwierigkeiten und entsprechende Anpassungen. Unterrichtsmaterial (11) umfasste Daten für die Unterrichtsvorbereitung, wie z.B. den Rahmenlehrplan, Sammeln sowie Erstellen von Unterrichts- und Arbeitsmaterialien. Die Befragten haben Feedbackdaten zur eigenen Lehre (6) von Lehrpersonen oder Schüler:innen erhalten oder eingeholt.

Mit Blick auf die einzelnen Befragten zeigte sich, dass 17 Zugriff auf Noten hatten, 15 nannten Namen, neun Stundenpläne, acht Unterrichtsmaterial, sieben haben sich eigene Notizen zu Leistung und Verhalten der Schüler:innen erstellt. Sieben Befragte hatten Einblicke in weitere Informationen über die Schüler:innen wie Förderbedarfe und sechs hatten Zugriff auf Abwesenheiten. Eine:r hatte Einblicke in Fortschrittsdaten aus einem Lernmanagementsystem. Fünf hatten nur wenige und weitere fünf keine Daten zur Verfügung.

Die Befragten möchten zukünftig Daten haben, die die Bewertung oder Einschätzung der Schüler:innen erleichtert, kontextübergreifend ermöglicht oder verzerrungsfreier macht (18 Nennungen) sowie Daten für die Unterrichtsgestaltung (16) z.B. in Form von automatischen Unterrichtsentwürfen entsprechend der Fähigkeiten und Interessen der Schüler:innen oder Datenbanken über bewertete Bildungstechnologien. Einblicke in den Lernfortschritt (11) aber auch Learning-Analytics-Ansätze (15) wurden genannt. Gewünschte Daten zu den Schüler:innen beinhalten Hintergrundinformationen (11, Familie, Probleme, Migration, etc.), Interessen (9), Leistungsstand (9) oder die Einschätzung anderer Lehrpersonen (5). Auch Videodaten aus Klassenraumaktivitäten und KI-gestützte Auswertungen (7) wurden genannt.

Die Bildungsdatenkompetenz der Befragten bewegt sich im Mittelfeld, durch den hohen Anteil derer, die in ihrer Schulpraxis geringen Zugriff auf Daten haben, wird der Bildungsdatenkompetenzaufbau beeinträchtigt (Schildkamp et al., 2014). Hierfür könnten in der Lehrerbildung Trainingsdaten und Simulationen eingesetzt und auf Lehrendendashboards visualisiert werden. Das Bewusstsein über Bildungsdaten und ihre Relevanz könnte durch einen expliziten Fokus in den schulpraktischen Studien gesteigert werden (Kippers et al., 2018). Die Diskrepanz der verfügbaren zu gewünschten Daten erfordert die Etablierung entsprechender Technologien sowie Datenkultur an Schulen. Zu hinterfragen ist, inwieweit die Befragten die verwendeten Daten erinnern oder explizieren konnten bzw. als selbstverständlich erachteten.

Has the AI revolution reached students yet? Nutzungsverhalten von KI-Tools durch Schülerinnen und Schüler am Ende der Sekundarstufe I

Nicolas Hübner¹, Lisa Bäumle¹, Rene Kizilcec², Jennifer Meyer³, Ronny Scherer⁴

¹Universität Tübingen, ²Cornell University, ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel, ⁴University of Oslo

Theoretischer Hintergrund

Aktuelle Studien dokumentieren eindrucksvoll die zunehmende Auseinandersetzung unterschiedlicher Bildungsakteure mit neuen Möglichkeiten und Herausforderungen von KI-Systemen (Fütterer et al., 2023; Ständige Wissenschaftliche Kommission, 2024). Die Anzahl der Studien, die sich näher mit der Nutzung von KI-Systemen durch Schüler:innen auseinandergesetzt haben, ist derzeit jedoch noch sehr überschaubar. In einer Studie der Vodafone Stiftung (Vodafone Stiftung, 2024) wurde eine repräsentative Stichprobe von rund 1.600 Schüler:innen im Alter von 14 bis 20 Jahren befragt. Von den Befragten gaben rund 45% an, dass die Nutzung von KI-Systemen an ihrer Schule entweder noch kein Thema (38%) bzw. generell verboten sei (7%). Darüber hinaus zeigte sich, dass Schüler:innen KI-Systeme insb. zur Erklärung von Begriffen (50%) und Themen (45%) nutzten und am wenigsten für Feedback zu selbst erarbeiteten Texten (13%). In einer weiteren Studie beschäftigte sich Senkbeil (2017) allgemeiner mit computerbezogenen Motivations- und Nutzungsmustern von Schüler:innen und konnte dabei fünf verschiedene Nutzungstypen identifizieren (z.B. Instrumentalisten und sozial-interaktive Hedonisten). In dem in der Studie verwendeten Faktorenmodell wurden hierzu zunächst drei übergeordnete Faktoren unterschieden: Ein instrumenteller Faktor (z.B. Informationssuche und lernbezogene Verwendung), ein hedonistischer Faktor (z.B. Eskapismus) und ein sozial-interaktiver Faktor (z.B. sozialer Austausch). Trotz dieser wichtigen Erkenntnisse zur schülerseitigen Nutzung digitaler Tools ist über Profile der KI-Tool-Nutzung in der Schule bisher noch wenig bekannt. Der vorliegende Beitrag beschäftigt sich daher tiefergehend mit dem Nutzungsverhalten von KI-Tools durch Schüler:innen am Ende der Sekundarstufe I und fokussiert hierzu insb. auf unterschiedliche Bereiche der schulbezogenen instrumentellen Nutzung (vgl. Senkbeil, 2017).

Fragestellung

Es werden drei Fragestellungen bearbeitet:

- 1) Wie häufig nutzen Schüler:innen und Schüler KI-Tools in der Schule?
- 2) Wozu nutzen Schüler:innen KI-Tools und welche Nutzungsprofile lassen sich identifizieren?
- 3) Welche Schüler:innen-Merkmale stehen im Zusammenhang mit der Nutzung von KI-Tools in der Schule für bestimmte Zwecke?

Methode

Zur Untersuchung der Fragestellungen wurden repräsentative Daten von rund 4.000 Schüler:innen (von 90 Beruflichen Gymnasien, 30 Allgemeinbildenden Gymnasien und 12 Gemeinschaftsschulen) ausgewertet, die im Zuge der TOSCAneo Studie im Schuljahr 2023/2024 von der IEA erhoben wurden (<https://www.iea.nl/de/node/4327>). Zur Beantwortung der vorliegenden Fragestellungen wurden Rückmeldungen der Schüler:innen zur Nutzungshäufigkeit 1 (nie) bis 5 (jeden Schultag) und zu

insgesamt sieben unterschiedlichen schulbezogenen Nutzungsbereichen (z.B. Lösen von Aufgaben, Bearbeitung von Hausaufgaben, Weiterlernen, Übersetzen von Texten) näher ausgewertet. Zur Beantwortung der zweiten Forschungsfrage nach den schulbezogenen Nutzungsbereichen und -profilen wurden exploratorische sowie konfirmatorische Faktorenanalysen und personenzentrierte Ansätze (latente Klassenanalysen) durchgeführt und die Profile in Bezug auf verschiedene SchülerInnen-Merkmale (z.B. dem sozioökonomischen Hintergrund und den kognitiven Fähigkeiten) verglichen. Fehlende Werte wurden mit der Full Information Maximum Likelihood (FIML)-Methode behandelt. Zur Berücksichtigung der Nestung der Daten (Schüler:innen in Schulen) wurden Cluster-robuste Standardfehler berechnet.

Ergebnisse und ihre Bedeutung

Es zeigte sich, dass rund 40% der Schüler:innen angaben, dass sie KI-Tools für schulbezogene Zwecke nutzen. Insgesamt 60% gaben an, dass Sie KI-Tools nie nutzen. Bei einer näheren Auswertung der Nutzungsbereiche zeigte sich, dass KI-Tools am häufigsten zum Übersetzen von Texten verwendet wurden und am wenigsten zur Lösung von Aufgaben im Unterricht. Im Zuge der latenten Klassenanalysen erwies sich eine Lösung mit vier Nutzungsprofilen als passend: Durchgehend niedrige Nutzungswahrscheinlichkeiten in allen Bereichen (P1; 57%), durchgehend hohe Nutzungswahrscheinlichkeiten in allen Bereichen (P2; 13%), eine hohe Wahrscheinlichkeit eines Einsatzes für ein besseres Verständnis von Unterrichtsinhalten und das Weiterlernen (P3; 19%) sowie hohe Nutzungswahrscheinlichkeiten für Hausaufgaben und Referate sowie Aufsätze (P4; 11%). Weitere Analysen legen substantielle Unterschiede in Bezug auf die Komposition dieser Klassen nahe, beispielsweise Notenunterschiede zwischen P3 und P4 zugunsten von P3 und Testleistungsunterschiede zwischen P2 und P3, ebenfalls zugunsten von P3. Die Bedeutung der Befunde und weitere Forschungsperspektiven werden im Vortrag diskutiert.

S-32: Wandel von Schul- und Unterrichtskulturen über die letzten 50 Jahre – interdisziplinäre Einblicke in das Projekt „Schule im Wandel“

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 102

Chair der Sitzung: Tobias Feldhoff

Chair der Sitzung: Sabine Reh

Chair der Sitzung: Eckhard Klieme

Chair der Sitzung: Katharina Maag Merki

Wandel von Schul- und Unterrichtskulturen über die letzten 50 Jahre – interdisziplinäre Einblicke in das Projekt „Schule im Wandel“

Chair(s): **Tobias Feldhoff** (Johannes Gutenberg Universität Mainz, Deutschland), **Sabine Reh** (BBF | Bibliothek für Bildungsgeschichtliche Forschung des DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation und Humboldt-Universität, Berlin), **Eckhard Klieme** (DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Diskutant*in(nen): **Katharina Maag Merki** (Universität Zürich)

In der Öffentlichkeit wird häufig davon ausgegangen, dass sich Schul- und Unterrichtskulturen in der Bundesrepublik seit Ende der 1960er Jahre verändert haben. Heute sei die Schule weniger hierarchisch, die Zusammenarbeit aller Akteure stärker partizipativ gestaltet als früher und damit demokratischer. Von den einen wird dies als „Kuschelpädagogik“ beklagt, von anderen wird die Veränderung der Autoritätsverhältnisse als Gewinn an Humanität in Schule gesehen. Solche Positionen spiegeln nicht nur einen Teil der öffentlichen Meinung wider; auch empirische Studien (Fend & Berger, 2016) geben Hinweise darauf, dass sich Schüler:innen heutzutage an Schulen wohler fühlen und ein besseres Verhältnis zu ihren Lehrpersonen haben als vor vier Jahrzehnten. Allerdings ist unser Wissen über Veränderungen, Stabilität und Wandel von Schule, ihrer Kultur und des Unterrichtshandelns in den letzten 50 Jahren sehr begrenzt, sowohl in der Bildungshistoriographie wie auch in der empirischen Bildungsforschung.

Im Symposium möchten wir aufzeigen, wie im DFG-Projekt „Schule im Wandel“ ein möglicher längerfristiger Wandel von Schul- und Unterrichtskulturen interdisziplinär (Bildungsgeschichte, empirische Bildungsforschung) mit Hilfe verschiedener Fragestellungen und methodischer Ansätze untersucht wird. Ausgangspunkt ist die Drei-Länder-Studie (DLS) von Fend (1982), die 1978 in Hessen und Nordrhein-Westfalen sowie 1979 in Niedersachsen an 61 Schulen durchgeführt wurde. 2023 wurden 24 der 61 Schulen in einer Follow-Up-Studie (FUS) mit nur leicht angepassten Originalinstrumenten (Fragebögen der Lehrpersonen sowie Fragebögen und Fachleistungstest in Mathematik und Englisch der Schüler:innen der 9. Jahrgangsstufe) in Hinblick auf Merkmale der Schul- und Unterrichtskultur sowie Lernvoraussetzungen und -ergebnisse der Schüler:innen erneut untersucht (Schmid et al. 2019). So liegen Daten aus 24 Schulen zu zwei Messzeitpunkten 1978/79 sowie 2023 vor; darunter befinden sich fünf Gymnasien und 11 Gesamtschulen mit unveränderter Schulform sowie je vier Haupt- und Realschulen, bei denen es teilweise in den letzten Jahrzehnten zu strukturellen Veränderungen bzw. zu einem Wechsel der Schulform gekommen ist. Ergänzend werden an vier der 24 Schulen historiographische Fallstudien durchgeführt.

Der erste Beitrag beleuchtet den schulübergreifenden Wandel von Schul- und Unterrichtskulturen sowie von fachlichen und überfachlichen Leistungen in den letzten 45 Jahren auf der Basis des Schulkulturkonzepts von Helsper (2000) anhand von DLS/FUS-Daten. Auf der Basis approximativer Messinvarianz werden Mittelwertsunterschiede in Skalen zur Schulkultur (z.B. „Lehrer:innenkooperation“, „Reformbereitschaft“, „Schüler:innenzugewandtheit“), zur Unterrichtskultur (z.B. Strukturiertheit, Unterstützung, Adaptivität) sowie zum Wohlbefinden als überfachliches Merkmal und zur Fachleistung in Englisch und Mathematik analysiert. Die Ergebnisse zeigen über alle Schulen hinweg sowohl Veränderungen (z.B. weniger Kooperation, mehr Reformbereitschaft der Lehrkräfte) als auch Stabilität (z.B. Strukturiertheit des Unterrichts, Schüler:innenzugewandtheit). Wichtige Aspekte der Unterrichtskultur und Wirkungen haben sich in den beiden Fächern gegenläufig verändert: Mehr Unterstützung, weniger Zeitverschwendung, höhere fachliche Leistungen und mehr Wohlbefinden im Englischunterricht; keine oder gegenläufige Veränderungen in Mathematik.

Der zweite Beitrag analysiert die Unterschiede zwischen den Schulen, um zu prüfen, ob sich die Bedeutung der Einzelschule als „pädagogischer Handlungseinheit“, die im Anschluss an die DLS vielfach beschworen wurde, gewandelt hat. Zunächst wird festgestellt, dass die Schulkulturen sich aus Sicht der Lehrkräfte sogar angenähert haben. Allerdings finden sich 2023 andere Zusammenhänge zwischen Indikatoren der Schulkultur einerseits, kognitiven und affektiven Outcomes andererseits als 1978/79. Der Beitrag diskutiert diese Veränderungen im Kontext des Wandels der Schulformen und der kognitiven Selektivität.

Der dritte Beitrag untersucht aus einer bildungshistorischen Perspektive die Veränderungsprozesse einzelner Schulen auf der Grundlage der historiographischen Fallstudien im Projekt „Schule im Wandel“. Die Rekonstruktion der Schulgeschichte wird kontextuiert durch eine Analyse der historischen Entwicklung erziehungswissenschaftlicher Thematisierungen von Schule in der Gesellschaft. Beobachtbare Wandlungsprozesse können so auf verschiedenen Ebenen des pädagogischen Systems im Zeitraum zwischen Fends „Drei-Länder-Studie“ in der Mitte der 1970er Jahre und der TIMS-Studie in den 1990er Jahren zueinander in Beziehung gesetzt werden. Neben schulischen Quellen und Zeitzeugeninterviews dienen die Daten der quantitativen Erhebung als historische Quellen und Grundlage für die Analyse schulkulturellen Wandels.

Beiträge des Symposiums

Alles bleibt anders? Schulübergreifende Veränderungen von Schul- und Unterrichtskulturen im Projekt „Schule im Wandel“

Tobias Feldhoff¹, Sabine Reh², Eckhard Klieme³, Sebastian Wurster¹, Brigitte Steinert³

¹Johannes Gutenberg Universität Mainz, ²BBF | Bibliothek für Bildungsgeschichtliche Forschung des DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation und Humboldt-Universität, Berlin, ³DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation

Wie im Abstract des Symposiums skizziert, gibt es Hinweise darauf, dass sich in den letzten 50 Jahren nicht nur die deutsche Gesellschaft (Doering-Manteuffel & Raphael, 2016), sondern auch die Schul- und Unterrichtskulturen verändert haben. Empirisch und historiographisch abgesicherte Erkenntnisse hierzu liegen jedoch bislang nicht vor.

Fragestellung:

In unserem Beitrag gehen wir der Frage nach, inwiefern sich schulübergreifende Veränderungen von Schul- und Unterrichtskulturen sowie von fachlichen und überfachlichen Leistungen der Schüler:innen über die letzten 50 Jahre anhand der Daten der DLS/FUS identifizieren lassen.

Theoretischer Hintergrund:

Schulkulturen werden in Anlehnung an Helsper (2000) als „symbolische, sinnstrukturierte Ordnungen der Einzelschule“ verstanden. Die Kultur der Einzelschule wird als spezifische Ausgestaltung pädagogischer Prozesse, sozialer Beziehungen, geltender Normen und beobachtbarer Praktiken verstanden, die in Rezeption und Auseinandersetzung mit gesellschaftlichen Strukturvorgaben durch schulische Akteure entsteht. Helsper unterscheidet dabei vier Dimensionen: Partizipation, Leistung, pädagogische Orientierungen sowie Inhalte schulischen Unterrichts und Umgang mit Wissensbeständen (Helsper et al., 2001). Diese Dimensionen werden im Projekt als Heuristik verwendet.

Der Schulkulturansatz von Helsper expliziert keinen systematischen Unterrichtsbezug. Im Projekt werden Unterrichtskulturen in Anlehnung an Helspers Schulkulturansatz als Formen der Gestaltung von Prozessen und sozialen Beziehungen im Unterricht verstanden. Über die an einer Schule geltenden Normen und Werte, über die erwünschte Gestaltung von Lernprozessen und die Gestaltung von Unterricht und pädagogischen Beziehungen beeinflusst eine Schulkultur die Unterrichtskulturen in Fächern und Klassen (vgl. Nasir, Scott, Trujillo & Hernández, 2016, S. 358ff.).

Stichprobe:

Als Stichprobe werden Daten der Befragung der Schüler:innen der 9. Jahrgangsstufe in Mathematik und Englisch (DLS24 N=1135; FUS24 N=1105) und der Befragung der Lehrpersonen (DLS24 N=434; FUS24 N=842) genutzt.

Instrumente:

Zum Einsatz kommen leicht adaptierte Originalinstrumente der DLS (Schmid et al. 2019).

Drei der vier Dimensionen der Schulkultur nach Helsper et al. (2001) werden durch fünf Skalen der Lehrpersonenbefragung abgebildet, die sich in Modellen der Schulqualitätsforschung wiederfinden. Die Skalen „Lehrer:innenkooperation“ und „Schüler:innenzugewandtheit“ sind Indikatoren für die Helsper-Dimension „Partizipationsverhältnisse“, die „Reformbereitschaft“ ist ein Indikator für die Dimension „Pädagogische Werte und Gestaltung der pädagogischen Beziehungen“ und die Skalen „Leistungsdruck/Erwartungen an die Schüler:innen“ und „Förderung“ sind Indikatoren für die Dimension „Leistungsorientierung und Selektionspraktiken“.

Die Unterrichtskultur für die Fächer Mathematik und Englisch wird mit vier Skalen aus der Schüler:innenbefragung erfasst, die sich an den Basisdimensionen guten Unterrichts orientieren (vgl. Klieme & Rakoczy, 2008). Sie bilden Teilaspekte der Klassenführung (Strukturiertheit und Zeitverschwendung), des lernunterstützenden Klimas sowie der Adaptivität des Unterrichts ab. Die überfachlichen Ergebnisse wurden mit der Skala „Wohlfühlen im Unterricht“ erfasst. Die fachlichen Ergebnisse wurden mit Leistungstests in Mathematik und Englisch aus dem DLS erfasst, die in der FUS mit Hilfe des IQBs um Items aus dem Bildungstrend erweitert wurden.

Methode:

Zur Beantwortung der Forschungsfrage werden die Mittelwerte der Skalen aus DLS24 und FUS24 miteinander verglichen. Die Signifikanz der Mittelwertsunterschiede wird mit T-Tests überprüft und ihre praktische Bedeutung mit Cohens d quantifiziert. Die Ergebnisse basieren auf Messinvarianzanalysen über die Zeit. Aufgrund teilweise fehlender skalarer Messinvarianz wurde mit Hilfe der Alignmentmethode die approximative Messinvarianz bestimmt (Asparouhov & Muthén, 2014), mit deren Hilfe die Mittelwerte von DLS und FUS vergleichbar sind.

Ergebnisse:

Über alle Schulen hinweg sehen wir mehr Veränderung in der Schulkultur (z.B. weniger Kooperation und Leistungsdruck, mehr Reformbereitschaft der Lehrkräfte) als Stabilität (Schüler:innenzugewandtheit). Wichtige Aspekte der Unterrichtskultur und Wirkungen haben sich gegenläufig verändert: Mehr Unterstützung, weniger Zeitverschwendung, höhere fachliche Leistungen und mehr Wohlbefinden im Englischunterricht; keine oder gegenläufige Veränderungen in Mathematik (z.B. mehr Zeitverschwendung, geringere fachliche Leistungen).

Zur Erklärung der Befunde sind weitere Analysen notwendig: u.a. Varianzverteilungen auf den Ebenen, Korrelationen und Mehrebenenanalysen sowie eine Kombination mit den Ergebnissen der Fallstudien und eine Historisierung der Befunde. Die Ergebnisse können wichtige Erkenntnisse für unser heutiges Verständnis von Schule und Unterricht liefern.

Langfristiger Wandel von Einzelschulen: Zum Zusammenhang von Selektivität, pädagogischer Kultur und Ergebnisqualität im Kontext neu justierter Schulformen

Eckhard Klieme¹, Lisa Schäfer², Brigitte Steinert¹, Tobias Feldhoff²

¹DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Johannes Gutenberg Universität Mainz

Fragestellung und Forschungshintergrund:

Fends (1982) Vergleich zwischen Gesamtschulen und gegliedertem Schulsystem (Drei-Länder-Studie, DLS) führte zur Fokussierung der Schulforschung und -praxis auf die Einzelschule als „pädagogische Handlungseinheit“. Re-Analyse der DLS-Daten und Follow-Up-Untersuchung (FUS) erlauben es nun, (a) Unterschiede zwischen Schulen und Schulformen hinsichtlich Eingangsvoraussetzungen (Selektivität der Schülerschaft), Schulkultur und Outcomes zu beziffern und Zusammenhänge zwischen diesen Schulmerkmalen zu analysieren, (b) zu prüfen, ob Ausmaß der Unterschiede und Muster der Zusammenhänge einem historischen Wandel unterliegen, und (c) zu explorieren, inwiefern die Veränderung der Schulen über einen langen Zeitraum hinweg als konsistenter einzelschulischer Wandel verstanden und erklärt werden kann.

Stark vereinfacht war in den 1980er und 1990er Jahren die einzelschulische Profilierung zentrales Leitmotiv im System, danach die Sicherung des Leistungsniveaus (Terhart 2023). Daher lassen sich für 2023 stärkere Unterschiede zwischen Schulen und eine größere Bedeutung der Leistungsdimension erwarten. Beim einzelschulischen Wandel ist mit Interdependenzen von Schulform, Komposition der Schülerschaft und Leistungsniveau zu rechnen (Baumert et al. 2006, Klieme & Steinert 2008).

Stichprobe:

n=24 Schulen, die an DLS und FUS teilgenommen haben.

Instrumente:

Lehrkräfteskalen zur Schulkultur sowie Schüler:innentests für Mathematik und Englisch werden wie im Beitrag 1 genutzt. Die pädagogischen Werte des Kollegiums werden zusätzlich mit den Skalen „Kontrollorientierung“ und „Förderorientierung“ erfasst. Um Lernprozesse auch aus Schülersicht zu erfassen, verwenden wir die Skala „Ablenkbarkeit“ (Beispiel: „Im Unterricht sind meine Gedanken oft woanders.“). Als affektiver Outcome wird „Schulfreude“ gemessen. Die soziale Selektivität wird über

schulische und berufliche Abschlüssen der Eltern, die kognitive Selektivität über Subskalen des Culture-Fair-Tests (Cattell & Weiß, 1972) erfasst.

Methode:

Für die Befragungsskalen wurden mittels approximativer Messinvarianz korrigierte Skalen-Mittelwerte gebildet. Die kognitiven Tests (Mathematik, Englisch, CFT) wurden in Zusammenarbeit mit dem IQB skaliert. Alle in diesem Abstract dargestellten Analysen wurden auf Schulebene gerechnet: Vergleich der Varianzen in DLS und FUS nach Levene; Korrelationen zwischen Indikatoren der kognitiven und sozialen Selektivität, Prozessmerkmalen und Outcomes; Regressionsmodelle zur Erklärung der Veränderung einzelschulischer Outcomes. Im Vortrag werden auch Befunde aus Mehrebenenmodellen vorgestellt.

Ergebnisse und Diskussion:

Entgegen unserer Erwartung hat die Varianz auf Schulebene in fast allen (Lehrer-)Befragungsskalen zur Schulkultur signifikant abgenommen. Dies ist zum Teil auf eine Angleichung der Schulformen zurückzuführen: An den Schulen des gegliederten Systems ist die Reformbereitschaft gestiegen; gleichzeitig nahm das Ausmaß der Kooperation zwischen den Lehrkräften an Gesamtschulen ab. Unverändert ist die Zwischen-Schul-Varianz beim „Leistungsdruck“. Hier lagen in den 1970er Jahren die Schulformen gleich auf. Dieses Niveau wurde nur in den Gymnasien 2023 erneut berichtet, während die Lehrkräfte anderer Schulformen ein niedrigeres Anspruchsniveau wahrnehmen.

Die Bedeutung schulkultureller Merkmale hat sich insofern geändert, als ihre Zusammenhänge mit Outcomes anders aussehen. So gingen 1978/79 höhere Schüler:innenzugewandtheit, Reform- und Förderorientierung auf Schulebene jeweils mit einer stärkeren Schulfreude der Schüler:innen einher; dies war 2023 nicht der Fall. Die Korrelation der Kontrollorientierung mit der Schulfreude wandelte sich sogar von signifikant negativ zu signifikant positiv. Hoher Leistungsdruck war, wie erwartet, erst 2023 mit besseren Fachleistung verknüpft.

Untersuchungen der Veränderung der Mathematik- und Englischleistungen sowie der Schulfreude zeigen, dass diese Outcomes sich umso positiver veränderten, (i) je stärker die kognitive Selektivität der Schülerschaft und (ii) der Leistungsdruck stieg, und (iii) je positiver sich die selbstberichtete Ablenkbarkeit entwickelte. Bei multivariater Betrachtung war jeweils nur einer dieser Bedingungsfaktoren signifikant. Bemerkenswert ist jedoch, dass es gelingt, über einen sehr langen Zeitraum hinweg den Wandel auf Ebene der Einzelschulen zu charakterisieren. Thesenartig gebündelt: Der Wandel von Einzelschulen ereignet sich in den Dimensionen der kognitiven Selektion, des kognitiven Engagements der Schüler:innen sowie des Anspruchsniveaus der Lehrkräfte, und diese Faktoren spielen eine Rolle bei Veränderungen der Effektivität der Schulen.

Um die Auswirkungen eines Schulformwechsels zu beurteilen, sollen im Vortrag jene drei Schulen, die einen solchen Wechsel vollzogen haben, porträtiert werden.

Schulkulturen zwischen Bildungsreform-Ära und neuem Qualitätsparadigma.

Sabine Reh, Julia Kurig

BBF | Bibliothek für Bildungsgeschichtliche Forschung des DIPF | Leibniz-Institut für Bildungsforschung und
Bildungsinformation und Humboldt-Universität, Berlin

In diesem bildungshistorisch ausgerichteten Beitrag fragen wir danach, ob und wie sich schulische Wandlungsprozesse mit dem Wandel erziehungswissenschaftlicher Thematisierungen von Schule verschränken. Wir betrachten dabei den Zeitraum zwischen der Mitte der 1970er und der 1990er Jahre. In dieser Zeit führte Helmut Fend seine Gesamtschulstudien, vor allem die Drei-Länder-Studie (DLS) durch und die Ergebnisse von TIMSS, der ersten großen, international vergleichenden Schulleistungsstudie, an der Deutschland sich beteiligte, wurden veröffentlicht. Reflektierte und beförderte die DLS den Verlust des naiven Glaubens, mit Hilfe empirischer Forschung bildungspolitische Entscheidungen über strukturelle Transformationen des Bildungswesens rationalisieren und legitimieren zu können, so gehörte die TIMSS-Studie bereits in den Kontext des Aufstiegs und der Durchsetzung eines neuen Blickes auf die Binnenstruktur von Schulen und vor allem unterrichtliches Geschehen, damit eines neuen Paradigmas, das der „Qualität“.

Der untersuchte Zeitraum gehört in eine Bildungsgeschichte „nach dem Boom“, wie man diese bislang immer noch wenig erforschte Phase nach dem Ende der bundesrepublikanischen Bildungsreform-Ära in Anlehnung an die zeithistorischen Arbeiten von Anselm Doering-Manteuffel und Lutz Raphael nennen kann. Auf der Basis historischer Fallstudien und empirischer Daten zur Entwicklung von vier Schulen in Nordrhein-Westfalen und Niedersachsen (zwei Gymnasien, zwei Gesamtschule (alle vier aus dem Sample der 1978/79 und 2023 beforschten Schulen) wollen wir erste Erkenntnisse zu Wandlungsprozessen von Schulkulturen in den Kontext zeithistorischer sowie wissenschaftsgeschichtlicher Prozesse einordnen. Sowohl die schulischen Quellen – z.B. Konferenzprotokolle, Schülerzeitungen, Schulberichte, lokale Presseberichte u.a. – wie auch die quantitativen Daten der einzelnen Schulen weisen auf eine seit den 1980er Jahren beobachtbare Tendenz hin, die verschiedenen Akteure – auf vielfältige Weisen – stärker in schulische Entscheidungsstrukturen und Prozesse einzubeziehen. Dabei entstanden neue Formen der individuellen Responsibilisierung, die nachhaltigen Einfluss auf das Geschehen an den jeweiligen Schulen nahmen. So können wir zeigen, dass einer sich verschiebenden wissenschaftlichen Thematisierung der Schule – von der Analyse gesellschaftlicher Strukturen und Funktionen hin zur Beobachtung und Analyse systeminterner und individueller Prozesse im Bildungswesen, mit welcher der Focus stärker auf die Einzelschule und das lernende Individuum gerichtet wurde – möglicherweise Änderungen der Schulkulturen entsprachen.

S-33: Selbstregulation in digitalen Lernumgebungen – Operationalisierung und Validierung verschiedener Facetten

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 151

Chair der Sitzung: Lea Marie Nobbe

Chair der Sitzung: Leonard Tetzlaff

Chair der Sitzung: Laura Dörrenbächer-Ulrich

Selbstregulation in digitalen Lernumgebungen – Operationalisierung und Validierung verschiedener Facetten

Chair(s): **Lea Nobbe** (DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland), **Leonard Tetzlaff** (DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland)

Diskutant*in(nen): **Laura Dörrenbächer-Ulrich** (Universität des Saarlandes)

Selbstreguliertes Lernen (SRL) gilt als eine Kernkompetenz des 21. Jahrhunderts und ermöglicht es Lernenden, mit sozialen, technologischen und ökonomischen Veränderungen Schritt zu halten (OECD, 2019). Zyklische Modelle SRLs (z.B. Zimmermann, 2002) beschreiben den Lernprozess dabei in Phasen. Selbstreguliert Lernende planen ihr Lernen, überwachen sich beim Lernen und passen ihre Strategienutzung bei Bedarf an. Nach dem Lernen beurteilen sie, ob und wieso ihr Lernen erfolgreich war.

Um Lernende in ihrem SRL zu unterstützen, ist es maßgeblich, verschiedene Aspekte von SRL zu erfassen und zu verstehen. Oft werden diese durch Selbstberichte erhoben (Alhazbi et al., 2023). Deren Güte ist davon abhängig, wie fähig und gewillt Lernende sind, akkurate Informationen zu liefern. Die Entwicklung neuer Wege, SRL-Aspekte zu erfassen, ist daher weiterhin von großer Bedeutung. Um die Validität der Interpretation verschiedener Operationalisierungen zu ergründen, müssen sowohl die Beziehungen unterschiedlicher SRL-Aspekte untereinander als auch mit Lernverhalten und -ergebnissen betrachtet werden.

Dieses Symposium stellt empirische Forschungsarbeiten dar, die verschiedene Aspekte von SRL messen und die Interpretation dieser Messungen durch ihre Beziehungen zu weiteren SRL-Maßen und Lernergebnissen validieren. Dabei werden Lernende unterschiedlichen Alters betrachtet, die in digitalen Lern- und Experimentalumgebungen Aufgaben bearbeiten.

Beitrag 1 befasst sich mit Operationalisierungen von verteiltem Lernen in Medizinstudierenden und Schüler*innen der Sekundarstufe. Dabei wird das Lernverhalten objektiv über Logdaten erfasst. Die Arbeit untersucht, welche Operationalisierungen interpretierbare und prädiktiv valide Ergebnisse für die Stichproben und Materialien bringen. Zusätzlich wird das Lernverhalten mit Selbstberichten zu verteiltem Lernen in Bezug gesetzt.

Beitrag 2 testet ein auf der Situated Expectancy Value Theory (Eccles & Wigfield, 2020) sowie SRL-Modellen (Schmitz & Wiese, 2006; Zimmerman, 2002) basierendes Drei-Phasen-Modell in Schüler*innen der Sekundarstufe. Das Modell beschreibt dabei die Zusammenhänge zwischen der Motivation vor dem Lernen, dem objektiv erfassten Lernverhalten und der Wahrnehmung des Lernens und kombiniert dafür Experience Sampling und Logdaten.

In Beitrag 3 stehen metakognitive Fähigkeiten und ihre Entwicklung in Kindern im Fokus. Anhand des Phänomens der Wasserverdrängung wird untersucht, wie 5- bis 9-jährige Kinder durch Konflikte zwischen ihrem eigenen Konzept und neuen Belegen lernen. Die metakognitive Fähigkeit der Konflikterkennung wird mithilfe von Eyetracking und Reaktionszeitmessungen quantifiziert und für Kinder verschiedenen Alters verglichen.

Beitrag 4 untersucht Metakognitive Fähigkeiten von Neuntklässler*innen im Rahmen der innovativen PISA-Domäne „Lernen in der digitalen Welt“, welche das Ziel verfolgt, SRL-Aspekte in einem Large-Scale Assessment zu erfassen (OECD, 2023). Verschiedene Operationalisierungen für metakognitive Genauigkeit werden dabei miteinander in Bezug auf ihre psychometrischen Eigenschaften sowie ihre Korrespondenz zu Selbstberichten, Lernverhalten und Lernerfolg verglichen.

Die abschließende Diskussion integriert die Ergebnisse.

Beiträge des Symposiums

Operationalisierungen verteilten Lernens und ihre Validität in Studierenden und Schüler*innen

Lea Nobbe¹, Jasmin Breitwieser¹, Garvin Brod²

¹DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; Goethe-Universität Frankfurt am Main

Selbstreguliert Lernen beinhaltet das Planen und Einsetzen effektiver Lernstrategien. Eine solche Strategie ist verteiltes Lernen, d.h. das Verteilen der Lernzeit auf mehrere kurze Einheiten im Gegensatz zum Lernen in nur einer oder wenigen langen Einheiten (Richter et al., 2022). Die positiven Effekte verteilten Lernens, gerade für langfristigen Lernerfolg, sind dabei vielfach belegt (z.B. Carpenter et al., 2017; Dunlosky et al., 2013).

Untersuchungen zu den Effekten verteilten Lernens und seiner Zusammenhänge mit Charakteristiken von Lernenden nutzen dabei vermehrt Experimente (z.B. Cepeda et al., 2006; Moulton et al., 2006) oder Selbstberichte (Alhazbi et al., 2023). Dabei wird innerhalb von Experimenten nicht das natürliche Lernverhalten beobachtet. Selbstberichte wiederum sind nur dann akkurat, wenn Lernende auch willens und in der Lage sind, ihr Lernverhalten akkurat zu berichten. Während in den mehr oder weniger verteilten Lerneinheiten in Experimenten oft dasselbe Material immer wieder wiederholt wird, was zu ähnlich langen Lerneinheiten führt, wird in Fragebögen häufig abgefragt, an wie vielen Tagen oder Wochen gelernt wurde, wobei die Länge der Lerneinheiten vernachlässigt wird (z.B. Rodriguez et al., 2021). Die Operationalisierungen von verteiltem Lernen unterscheiden sich dementsprechend systematisch je nach Form der Erfassung.

Mit der zunehmenden Verbreitung von Lernplattformen und -apps nimmt auch die Verfügbarkeit von Logdaten Lernender zu. Diese repräsentieren, oft hochauflösend, objektiv gemessenes Lernverhalten. Durch die hohe Auflösung ermöglichen sie es, verschiedene Operationalisierungen in Bezug auf die Interpretierbarkeit ihrer Ergebnisse und prädiktive Validität für Lernergebnisse zu vergleichen.

Um einen solchen Vergleich anzustellen, nutzen wir die Daten von 339 Medizinstudierenden, die sich in wochenlanger Vorbereitung auf das zweite Staatsexamen befanden. Wir nutzten Operationalisierungen, die sowohl ähnlich lange, sehr regelmäßige Einheiten (durch Entropiewerte), als auch Lernverläufe (durch time-series clustering und growth curve models) sowie das Verteilen von Lernen auf mehr Tage reflektierten. Die Lernaktivität, die dabei betrachtet wurde, war das Beantworten von Übungsfragen. Die erreichte Prozentzahl im Staatsexamen wurde als Lernergebnis genutzt. Wir fanden, dass alle vier Operationalisierungen interpretierbare Ergebnisse lieferten. Im Anschluss wurden die Ergebnisse der Operationalisierungen (also

Werte für Entropie, Steigungsparameter des growth curve models und Anzahl der Lerntage bzw. Clusterzugehörigkeit) als Prädiktoren für die erreichte Prozentzahl im Staatsexamen genutzt. Dabei wurde für die Anzahl der insgesamt beantworteten Übungsfragen kontrolliert. Nur das Verteilen der Lernaktivität auf mehr Tage hatte einen signifikant positiven Effekt, wenn für die Anzahl der insgesamt beantworteten Fragen kontrolliert wurde ($\beta_{\text{Fragen}} = 0.07$, $p = .289$; $\beta_{\text{Tage}} = 0.13$, $p = .040$).

Aufbauend auf diesen Ergebnissen untersuchen wir nun eine weitere Stichprobe: 350 Schüler*innen der Sekundarstufe, die mit einer App Englischvokabeln lernen. Im Unterschied zur vorherigen Studie können hier die Zusammenhänge auf der Ebene einzelner Vokabeln betrachtet werden. Wir untersuchen, ob sich auch hier ein Verteilen der Lernaktivität über mehr Tage unter Kontrolle der Anzahl der Wiederholungen positiv auf die Lernergebnisse auswirkt. Außerdem testen wir, ob sich eine größere Varianz in den von der App gebotenen Lernaktivitäten (multiple choice, Gesten, Eintippen) oder das häufigere Wiederholen einer mit dem Testmodus konsistenten Lernaktivität (also das Eintippen von Vokabeln) positiv auf die Wahrscheinlichkeit, die Vokabel korrekt zu erinnern, auswirken. Außerdem berücksichtigen wir durch Fragebögen erhobene Selbstberichte der Kinder zu ihrer Strategienutzung. Die Ergebnisse werden zur Konferenz vorliegen.

Wir werden in einer Diskussion die Chancen aber auch Herausforderungen in der Arbeit mit Logdaten eruieren. Außerdem werden wir die Ergebnisse für die beiden Stichproben verglichen und einordnen. Basierend auf dieser Einordnung werden wir theoretische und praktische Implikationen ableiten.

Test eines Drei-Phasen-Modells der Beziehung zwischen motivationalen Einschätzungen, beobachtbarem Engagement und subjektivem Engagement beim Lernen

Jasmin Breitwieser¹, Garvin Brod², Lars-Erik Malmberg³

¹DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; IDEa-Center for Research on Individual Development and Adaptive Education of Children at Risk, Frankfurt am Main, ²DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; Goethe-Universität Frankfurt am Main, ³University of Oxford, Vereinigtes Königreich

Nach der Situated Expectancy-Value Theory (SEVT) werden akademische Entscheidungen und Verhaltensweisen in erster Linie dadurch bestimmt, welchen Wert Lernende der Aufgabe beimessen, welche Kosten sie mit der Aufgabebearbeitung assoziieren und welche Erwartung sie hinsichtlich ihres möglichen Erfolgs bei der Aufgabe haben (Eccles & Wigfield, 2020). In den letzten Jahren wurde der Mikroebene momentaner Veränderungen dieser motivationalen Einschätzungen mehr Aufmerksamkeit geschenkt. Studien, in denen Methoden der intensiven Längsschnittdatenerfassung eingesetzt wurden, haben gezeigt, dass sich die Motivation und das anschließende Engagement von Lernenden über kurze Zeiträume hinweg verändern (z.B. (Dietrich et al., 2017; Lo, 2024)). Dies hat zu einer stärkeren Betonung der situationsspezifischen Natur von Motivationsprozessen beim Lernen geführt.

Bislang stützt sich diese Forschung überwiegend auf die selbstberichtete Wahrnehmung akademischer Aufgaben durch Lernende. Studien, die einen Experience-Sampling-Ansatz verfolgen, fragen Lernende beispielsweise nach dem subjektiven Wert, den sie dem Lernen beimessen, und zum gleichen oder einem späteren Zeitpunkt, wie sehr sie sich beim Lernen angestrengt haben (z.B. Dietrich et al., 2017). Durch die Erfassung dieser Daten über mehrere Situationen hinweg können intra-individuelle Veränderungen des subjektiven Wertes mit intra-individuellen Veränderungen der wahrgenommenen Anstrengung in Beziehung gesetzt werden. Somit ermöglicht dieser Ansatz eine Untersuchung der von der SEVT postulierten situationsspezifischen Motivationsprozesse.

Durch die ausschließliche Fokussierung auf Selbstberichte bleibt jedoch eine wichtige Frage unbeantwortet: (Wie) hängen diese subjektiven Erfahrungen und Wahrnehmungen mit objektiven Veränderungen im Lernverhalten zusammen? In oben genanntem Beispiel: Hat ein höherer subjektiver Aufgabenwert zu messbaren Veränderungen im tatsächlichen Lernverhalten geführt? Wenn ja, wie spiegeln sich diese Veränderungen im Lernverhalten in der Art und Weise wider, wie sich die Studierenden später an diese Lernepisode erinnern, sie interpretieren und darüber berichten? Ein Verständnis dieser Zusammenhänge ist von entscheidender Bedeutung, um ein umfassenderes Bild davon zu erhalten, wie Motivationsprozesse das tatsächliche Lern-Engagement beeinflussen und wie dies mit den subjektiven Erfahrungen und Reflexionen der Lernenden über ihr Engagement zusammenhängt.

Basierend auf den Annahmen der SEVT sowie Prozessmodellen des selbstregulierten Lernens (Schmitz & Wiese, 2006; Zimmerman, 2002) postulieren wir ein Drei-Phasen-Modell, das beschreibt, wie die motivationalen Einschätzungen von Lernenden vor dem Lernen ihr beobachtbares Engagement während des Lernens und die anschließende subjektive Wahrnehmung dieses Engagements nach dem Lernen beeinflussen. Wir testeten dieses Modell in einer ersten Studie mit 189 Fünftklässlern, die über fünf Wochen hinweg selbstreguliert eine mobile Lern-App zum Vokabellernen nutzten sowie tägliche Fragebögen beantworteten (ambulatorisches Assessment). Die Logdaten gaben Aufschluss darüber, mit wie vielen Vokabeln sich die Lernenden jeden Tag in der App beschäftigten (beobachtbares Engagement). Die täglichen Fragebögen erfassten Motivationseinschätzungen vor dem Lernen (Erwartung, Wert, Kosten) sowie subjektive Berichte über das Engagement nach dem Lernen (Zufriedenheit, Anstrengung, Aufschiebeverhalten).

Die Ergebnisse zeigten, dass motivationale Einschätzungen vor dem Lernen die subjektive Wahrnehmung des Engagements nach dem Lernen vorhersagten (within-person Korrelationen zwischen .04 und .17), aber nur Erwartung und Wert waren auch mit dem beobachtbaren Engagement assoziiert (jeweils $r = .05$). Das beobachtbare Engagement medierte die Effekte von Erwartung und Wert auf die selbstberichtete Zufriedenheit und das Aufschiebeverhalten. In einem kombinierten Modell mit allen motivationalen Einschätzungen erklärte nur der subjektive Wert einzigartige Varianz im beobachtbaren Engagement (indirekter Effekt auf Zufriedenheit: $b = 0.02$; indirekter Effekt auf Aufschiebeverhalten: $b = -0.02$). Dies unterstreicht die Bedeutung dieser motivationalen Komponente für das Lern-Engagement und liefert teilweise Evidenz für das postulierte theoretische Modell. Insgesamt waren die Zusammenhänge zwischen den Selbstberichten und der objektiven Messung des Engagements auf within-person Ebene jedoch eher schwach, was darauf hindeutet, dass die subjektiven Berichte zwar wertvolle Einblicke in die Wahrnehmungen der Studierenden liefern, aber die Komplexität ihres tatsächlichen Verhaltens nicht vollständig erfassen. Das Verständnis des Zusammenspiels von beobachtbarem Lern-Engagement und dessen subjektiver Wahrnehmung sollte ein Hauptziel der Motivationsforschung sein.

Wie entwickeln sich metakognitive Fähigkeiten im Grundschulalter und welche Rolle spielen sie für das Lernen aus Konflikten?

Maria Theobald¹, Elfriede Holstein², Leonie Weindorf², Garvin Brod³

¹Universität Trier; DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; IDEa-Center for Individual Development and Adaptive Education of Children at Risk, ³DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; Goethe-Universität Frankfurt am Main

Lernende haben oft falsche Vorstellungen von naturwissenschaftlichen Phänomenen. Nach der „theory theory“ der kognitiven Entwicklung lernen Kinder, indem sie intuitive Theorien aufstellen, darauf basierende Hypothesen testen und die daraus resultierenden Ergebnisse nutzen, um ihre Theorien zu revidieren (Gopnik & Wellman, 2012). Das Lernen aus Hypothesentests und damit die Überwindung von Fehlvorstellungen setzt jedoch die Fähigkeit voraus, Konflikte zwischen der eigenen Vorhersage und dem tatsächlichen Ergebnis zu erkennen und aufzulösen. Das Erkennen von Konflikten erfordert daher metakognitive Fähigkeiten, wie das Überwachen und Reflektieren des eigenen Lernprozesses (Flavell, 1979; Nelson & Narens, 1990). Obwohl sich metakognitive Fähigkeiten im Laufe der Kindheit und Jugend verbessern (Roehrs, 2017), haben insbesondere jüngere Grundschulkinder häufig noch Schwierigkeiten, Fehler oder Konflikte zu erkennen (Finn & Metcalfe, 2014; Lipko et al., 2009). Basierend auf der Annahme, dass sich die Konflikterkennung mit zunehmendem Alter verbessert und zentral für die Überwindung von Fehlvorstellungen beim naturwissenschaftlichen Lernen ist, werden folgende Hypothesen getestet: (H1) Kinder zeigen mit zunehmendem Alter eine bessere Konflikterkennung. Zur Erfassung der Konflikterkennung werden zum einen die Reaktionszeiten (Verhaltensmaß) und zum anderen die Pupillenerweiterung (physiologisches Maß) bei falschen Vorhersagen herangezogen. (H2) Eine bessere Konflikterkennung sagt eine schnellere Überwindung von Fehlvorstellungen zum Konzept der Wasserverdrängung vorher.

Fünf- bis neunjährige Kinder ($N = 177$, $M_{\text{Alter}} = 7.02$) bearbeiteten eine Reihe von computerbasierten Lernaufgaben zum Konzept der Wasserverdrängung, zu dem viele Kinder Fehlkonzeppte haben. In jeder Aufgabe wurden den Kindern zwei Objekte präsentiert und sie sollten vorhersagen, welches der beiden Objekte mehr Wasser verdrängt, wenn es in Wasser getaucht wird. Anschließend erhielten die Kinder Feedback über das richtige Ergebnis. Eine bessere Konflikterkennung wurde durch eine stärkere Pupillenerweiterung und längere Antwortzeiten bei falschen im Vergleich zu richtigen Vorhersagen operationalisiert.

Die Ergebnisse zeigten, dass das Alter der Kinder die Konfliktüberwachung moderierte: Mit zunehmendem Alter zeigten die Kinder (1) längere Antwortzeiten bei falschen im Vergleich zu korrekten Vorhersagen ($\beta = -.04$, $p = .006$) sowie (2) eine stärkere Pupillenerweiterung als Reaktion auf unerwartete Ergebnisse ($\beta = -.04$, $p = .030$). Eine stärkere Pupillenerweiterung nach unerwarteten Ergebnissen sagte auch eher einen Wechsel zum korrekten Konzept im nachfolgenden Trial vorher ($\beta = -.13$, $p = .047$). Darüber hinaus zeigte sich mit zunehmendem Alter insgesamt ein stärkerer Lernzuwachs über die Lernphase ($\beta = .07$, $p = .008$).

Zusammenfassend deuten diese Ergebnisse darauf hin, dass sich die Fähigkeit, Konflikte zu erkennen, im Laufe der Kindheit verbessert und das Lernen aus falschen Vorhersagen erleichtert. Mit zunehmendem Alter revidierten Kinder ihre Fehlkonzeppte zur Wasserverdrängung schneller. Diese Ergebnisse unterstreichen daher die Rolle metakognitiver Fähigkeiten, wie der Konfliktüberwachung, bei der Entwicklung der Fähigkeit, aus falschen Vorhersagen zu lernen.

Metakognitive Genauigkeit im PISA-Kontext: Operationalisierungen und Zusammenhang mit Lernen

Leonard Tetzlaff¹, Lothar Persic-Beck¹, Ulf Kröhne², Carolin Hahnel³, Frank Goldhammer⁴

¹DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; Zentrum für Internationale Vergleichsstudien (ZIB), ²DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation, ³DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; Ruhr Universität Bochum, ⁴DIPF|Leibniz-Institut für Bildungsforschung und Bildungsinformation; Zentrum für Internationale Vergleichsstudien (ZIB); Goethe Universität Frankfurt am Main

Hintergrund:

Die Überwachung und Bewertung des eigenen Fortschritts und der eigenen Leistung sind wesentliche Bestandteile von selbstgesteuerten Lernprozessen (Veenman et al., 2006).

Als solche sind sie auch im PISA-Rahmenkonzeption für die innovativen Domäne "Lernen in der digitalen Welt" vertreten, mit welcher unter anderem das Ziel verfolgt wird, die Fähigkeiten zum selbstgesteuerten Lernen in einem Large-Scale Assessment zu messen (OECD, 2023).

Das gebräuchlichste Fähigkeitsmaß für die Überwachung und Bewertung von eigenen Lernprozessen ist die metakognitive Genauigkeit. Diese beschreibt den Grad, zu welchem die Lernenden die Korrektheit ihrer eigenen Lösung(en) für eine Aufgabe (oder eine Reihe von Aufgaben) richtig einschätzen können.

Es gibt verschiedene Möglichkeiten, diese metakognitive Genauigkeit zu konzeptualisieren und zu messen (Pieschl, 2009; Schraw, 2009), aber momentan noch relativ wenig Befunde über Zusammenhänge und Unterschiede der verschiedenen Maße.

Fragestellung:

Wir untersuchen Maße der absoluten und relativen Genauigkeit im Kontext der innovativen PISA-Domäne "Lernen in der digitalen Welt" und setzen sie in Beziehung zu Vorwissen, Lernergebnissen und Selbstberichten.

Methode:

Wir nutzen Daten von 737 Jugendlichen aus der deutschen Begleitstudie zu PISA 2022, die Aufgaben aus der innovativen Domäne pilotiert.

Als individuelles Maß der absoluten Genauigkeit wurde die Korrespondenz zwischen der Korrektheit der abgegebenen Antworten und Konfidenzratings zu diesen Antworten auf Aufgabenebene ermittelt und dann über Aufgaben hinweg mit einem 1PL IRT-Modell aggregiert.

Als individuelles Maß der relativen Genauigkeit wurde für jede Person die Korrelation über Aufgaben zwischen relativer Korrektheit der Antworten und relativer Konfidenz zu diesen Antworten ermittelt.

Als Maß für Domänenspezifisches Vorwissen, wurde ein in die Lernumgebung integrierter Prätest, welcher aus Aufgaben mit geschlossenem Antwortformat bestand, verwendet.

Als Selbstbericht-Maße nutzten wir die LIST Kurzska zur Nutzung metakognitiver Strategien (Klingsieck, 2019).

Ergebnisse:

Die Skala der absoluten Genauigkeit erreichte mit einer EAP Reliabilität von .65 eine akzeptable Messgenauigkeit und mit Item infit Statistiken von 0.84 bis 1.17 eine gute Modellpassung.

Die durchschnittliche relative Genauigkeit lag bei $r = .49$, was den Ergebnissen früherer Studien entspricht (e.g. Maki 1998).

Maße der absoluten und relativen Genauigkeit waren mäßig positiv korreliert ($r = .41$, $p < .001$).

Maße der absoluten Genauigkeit waren signifikant positiv mit domänenspezifischem Vorwissen assoziiert ($r = .32$ $p < .001$) und über das Vorwissen hinweg prädiktiv für Lernerfolg ($\beta = .07$ $p = .01$). Die Maße der relativen Genauigkeit zeigten keinen Zusammenhang mit dem Vorwissen oder dem Lernerfolg.

Weder absolute noch relative Genauigkeit zeigten einen Zusammenhang mit den Selbstberichtsmaßen zu metakognitiver Strategienutzung.

Die Ergebnisse wurden durch die Wahl von Gamma- oder Pearson-Korrelationsindizes (Annahme eines monotonen vs. linearen Zusammenhangs) für die relative Genauigkeit nicht beeinflusst.

Diskussion:

Die dargestellten Ergebnisse reihen sich in vorherige Arbeiten ein, welche verschiedene Maße für metakognitive Genauigkeit in Laborstudien verglichen haben:

So fanden z.B. Griffin et al. (2009) einen Zusammenhang zwischen domänenspezifischem Vorwissen und absoluter, nicht aber relativer Genauigkeit. Maki et al. (2005) berichten ebenfalls signifikante Zusammenhänge von absoluter, nicht aber relativer Genauigkeit mit aufgabenrelevanten Fähigkeiten und Aufgabenschwierigkeit.

Low-stakes Large-Scale Assessments bringen weitere potentielle Hürden mit sich - wie zum Beispiel die Prävalenz von schnellem Rateverhalten - welche die Validität der Interpretation verschiedener Maße in unterschiedlichem Ausmaß gefährden könnten.

Des Weiteren diskutieren wir unterschiedliche Interpretationen der verschiedenen Operationalisierungen, mögliche Gründe für die mangelnde Korrespondenz mit Selbstberichtsdaten, sowie Implikationen für zukünftige Forschung.

S-34: Auslandsaufenthalte während der Schulzeit – Zugang, Entwicklungsverläufe und deren Einflussfaktoren

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: SN 169

Chair der Sitzung: Matthias Kreuch

Chair der Sitzung: Annabell Daniel

Chair der Sitzung: Marita Jacob

Auslandsaufenthalte während der Schulzeit – Zugang, Entwicklungsverläufe und deren Einflussfaktoren

Chair(s): **Matthias Kreuch** (Ludwig-Maximilians-Universität München), **Annabell Daniel** (Ludwig-Maximilians-Universität München)

Diskutant*in(nen): **Marita Jacob** (Universität zu Köln)

In Deutschland bestehen für Schüler:innen verschiedene Möglichkeiten, einen Teil der Schulzeit im Ausland zu verbringen. Neben Sprachreisen und Workcamps, die vorrangig in den Ferien absolviert werden, sind es vor allem die Angebote eines Individualaustausches über einen kürzeren (bis zu einem Monat) oder längeren Zeitraum und eines Gruppenaustausches (zwischen zwei und vier Wochen), die von Schüler:innen in Deutschland genutzt werden. Oftmals sind mit der Teilnahme an einem Auslandsaufenthalt große Hoffnungen in die (Persönlichkeits-)Bildung der Teilnehmenden verbunden. Zumindest für längerfristige Auslandsaufenthalte lassen sich Effekte auf die Fremdsprachenkompetenz (Hübner et al., 2021) und das Selbstwertgefühl (Huttemann, 2014) sowie auf Merkmale interkultureller Einstellungen wie Offenheit und Verträglichkeit (Greischel et al., 2016) nachweisen. Der Zugang zu einem schulischen Auslandsaufenthalt ist jedoch sozial selektiv (Gerhards & Hans, 2013). Teilnehmende eines (längerfristigen) Auslandsaufenthalts während der Schulzeit kommen überwiegend aus einkommensstarken und bildungsnahen Familien. Zudem ist die Englischnote ein starker Prädiktor für die Teilnahme (Hübner et al., 2021). Diese Ergebnisse spiegeln vor allem eine reproduktionstheoretische Perspektive auf das Forschungsfeld wider. Die psychologische Perspektive konzentriert sich eher auf die individuellen Entwicklungsverläufe und die Effekte von schulischen Auslandsaufenthalten für den Kompetenzerwerb. Dabei hängt die Wirksamkeit des Aufenthalts von unterschiedlichen Faktoren ab. Neben strukturellen Faktoren wie der Dauer des Aufenthaltes (Wolff & Borzikowsky, 2018) scheint die Qualität, insbesondere beim Fremdsprachenerwerb, von zentraler Bedeutung zu sein (Kinging, 2011). Auf der Seite der Teilnehmenden nehmen individuelle Anpassungsprozesse eine wichtige Rolle für die Wirksamkeit von Auslandsaufenthalten ein (Serrano-Sánchez et al., 2021a).

Das Ziel des Symposiums ist es, die Chancen und Grenzen von Auslandsaufenthalten während der Schulzeit systematisch aufzuzeigen. Dazu werden der Zugang zu schulischen Auslandsaufenthalten, die individuellen Entwicklungsverläufe und die Bedeutung individueller, sozialer und kultureller Einflussfaktoren für die Effektivität von Auslandsaufenthalten aus interdisziplinärer Perspektive diskutiert. Die einzelnen Beiträge zeichnen sich dadurch aus, dass sie unterschiedliche Datengrundlagen heranziehen und mithilfe von Längsschnittdaten Veränderungen in den genannten Bereichen über die Zeit abbilden.

Der erste Beitrag untersucht aus einer soziologischen Perspektive soziale und regionale Ungleichheiten beim Zugang zu einem schulischen Auslandsaufenthalt im Kontext veränderter gesellschaftlicher Rahmenbedingungen. Der Vergleich verschiedener Kohorten des Sozioökonomischen Panels (SOEP) legt eine über die Zeit hinweg hohe Bedeutsamkeit der familiären Kapitalausstattung nahe.

Aus einer erziehungswissenschaftlichen Perspektive analysiert der zweite Beitrag über fünf Messzeitpunkte hinweg die Dynamik individueller Anpassungsprozesse zu Beginn eines längerfristigen Auslandsaufenthaltes. Während sich im Ausgangsniveau keine sozialen Ungleichheiten beobachten lassen, verweisen erste Befunde der Wachstumskurvenmodelle auf signifikante Herkunftsunterschiede in der Entwicklung des individuellen Akkulturationsverhaltens zugunsten von Schüler:innen mit einem geringeren sozioökonomischen Status.

Im dritten Beitrag wird mithilfe eines Experience Sampling Ansatzes die Entwicklung interkultureller Kompetenzen während eines schulischen Auslandsaufenthaltes untersucht. Die Ergebnisse der Latent Change Score-Modelle zeigen, dass bereits Auslandsaufenthalte von kurzer Dauer interkulturelle Kompetenzen, die wiederum im Zusammenhang mit der Häufigkeit und Qualität sozialer Kontakte und dem emotionalen Befinden stehen, fördern können.

Der vierte Beitrag widmet sich aus einer psychologischen Perspektive intraindividuellen und interindividuellen Zusammenhängen zwischen dem Schulumagement im Ausland und dem Akkulturationsverhalten jugendlicher Austauschschüler:innen. Die Random-Intercept Cross-Lagged Panel Analysen ergaben, dass das Schulumagement durch individuelle und kulturelle Merkmale beeinflusst wird und im Zusammenhang steht mit einer stärkeren Orientierung des Verhaltens am Gastland.

Beiträge des Symposiums

ENTFÄLLT Auslandsaufenthalte im Jugendalter – erleichterter Zugang für alle oder weiterbestehende soziale Ungleichheit?

Silke Hans

Georg-August-Universität Göttingen

Schulische Auslandsaufenthalte erfreuen sich großer Beliebtheit unter Jugendlichen; sie bieten zudem eine ausgezeichnete Möglichkeit, das zunehmend wichtige transnationale Humankapital zu erwerben - personengebundene Fähigkeiten und Kompetenzen, die es ermöglichen, jenseits des eigenen Nationalstaates bzw. der eigenen Herkunftsgruppe zu agieren. Hierzu gehören beispielsweise Mehrsprachigkeit und die Kenntnis anderer Kulturen und Regelsysteme. In zwei umfangreichen DFG geförderten Projekten haben wir untersucht, inwieweit transnationales Humankapital zu einer zunehmend

nachgefragten Ressource auf dem Arbeitsmarkt geworden ist, welche Effekte schulische Auslandsaufenthalte und das dabei erworbene Kapital auf den weiteren Bildungsweg und Berufseinstieg haben und insbesondere, welche Rolle soziale Ungleichheiten für den Zugang zu schulischen Auslandsaufenthalten spielen (vgl. Gerhards et al., 2017).

Unter Bezugnahme auf bildungssoziologische Erklärungsansätze sowie theoretische Überlegungen Pierre Bourdieus haben wir ein Erklärungsmodell für den Zugang zu Schüleraustauschprogrammen entwickelt und dieses mit einem Mixed-Methods-Ansatz empirisch getestet und weiterentwickelt. Als Datengrundlage dienten uns dabei u.a. standardisierte Umfragedaten aus dem Sozio-ökonomischen Panel (SOEP) sowie qualitative Interviews mit Eltern und mit Anbietern von Schüleraustauschprogrammen. Hierbei hat sich gezeigt, dass der Zugang zu schulischen Auslandsaufenthalten von unterschiedlichen Faktoren beeinflusst wird:

der Ausstattung der Jugendlichen und ihrer Eltern mit ökonomischem, sozialem und kulturellem Kapital, den Investitionen der Eltern in die Bildung ihrer Kinder sowie deren eigenes schulisches Engagement, der regionalen Gelegenheitsstruktur und schließlich familiären Konflikten, die Kinder dazu bewegen können, Distanz zu ihrem Elternhaus zu suchen (Gerhards & Hans, 2013; Carlson et al., 2017).

Anknüpfend an diese Erkenntnisse analysiert der geplante Symposiumsbeitrag, ob und wie sich soziale Ungleichheiten im Zugang zu schulischen Auslandsaufenthalten in den letzten Jahren verändert haben. Seit dem Zeitraum, der im Fokus der ursprünglichen Analysen stand – die 1990er bis frühen 2010er Jahre – haben umfangreiche Veränderungsprozesse stattgefunden, die auch für die soziale Strukturierung des Erwerbs von transnationalem Kapital durch Auslandsaufenthalte relevant sind. Einige dieser Prozesse betreffen die Gesellschaft insgesamt, andere speziell den Bildungsbereich bzw. den Zugang zu Schüleraustauschprogrammen. So sind durch die voranschreitende Globalisierung Fremdsprachenkenntnisse noch wichtiger geworden; diese sind aber auch ohne längere Auslandsaufenthalte leichter zu erwerben. Neueste Daten aus dem Eurobarometer zeigen, dass sich die (selbstberichtete) Fremdsprachenkompetenz in der Bevölkerung im letzten Jahrzehnt deutlich verbessert haben, gerade auch bei jüngeren Menschen. So geben mittlerweile 65 Prozent der Deutschen an, sich auf Englisch unterhalten zu können – neun Prozent mehr als noch im Jahr 2012 (Europäische Kommission, 2024).

Gleichzeitig könnten ökonomische Entwicklungen (Finanzkrise, Inflation) es gerade Familien aus der Mittelklasse erschwert haben, ihren Kindern einen Auslandsaufenthalt zu ermöglichen. Hingegen haben andere Entwicklungen – Digitalisierung, soziale Netzwerke etc. – zu einer Absenkung emotionaler Hürden für die Jugendlichen und ihre Eltern beigetragen, da es mittlerweile leichter ist, auch über große Entfernungen hinweg miteinander in Kontakt zu bleiben. Auch die Suche nach Informationen über Anbieter, Zielregionen etc. ist sicherlich leichter als vor 20 Jahren. Speziell im Bildungsbereich ist relevant, dass mittlerweile mehr Jugendliche ein Gymnasium oder eine andere Schulform, die das Abitur ermöglicht, besuchen – schließlich lässt sich ein längerer Auslandsaufenthalt leichter in eine 12- oder 13-jährige Schullaufbahn integrieren als in eine kürzere. Gleichzeitig bestehen weiterhin große soziale Ungleichheiten im Zugang zum Gymnasium, die sich somit wiederum auf die Teilnahme an Austauschprogrammen auswirken könnten.

Die ambivalenten Auswirkungen dieser sozialen und ökonomischen Veränderungen auf soziale Ungleichheiten im Zugang zu bildungsbezogenen Auslandsaufenthalten stehen im Zentrum des Symposiumsbeitrags. Hierzu werden standardisierte Umfragedaten (u.a. aus dem SOEP) zu deutschen Schüler:innen aus dem letzten Jahrzehnt mit den genannten Befunden aus dem älteren Zeitraum verglichen. Die Befunde weisen auf eine weiterhin hohe Relevanz der Kapitalausstattung der Familien für die Zugangsmöglichkeiten zu schulischen Auslandsaufenthalten hin. Ebenso spielen regionale Ungleichheiten weiter eine Rolle.

Der Einfluss der sozialen Herkunft auf die Entwicklung des Akkulturationsverhaltens während eines Schüleraustausches

Matthias Kreuch¹, Annabell Daniel¹, Kai Maaz²

¹Ludwig-Maximilians-Universität München, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Theoretischer Hintergrund:

Ein Auslandsaufenthalt während der Schulzeit wirkt sich positiv auf die persönliche Entwicklung und den Bildungserfolg von Schüler:innen aus (Hübner et al., 2021). Auslandsaufenthalte bergen daher möglicherweise Potential zur Verringerung von sozialer Bildungsungleichheit. Bisher nutzen vor allem Kinder aus sozial privilegierten Familien die Möglichkeit während der Schulzeit ins Ausland zu gehen, weshalb es an Untersuchungen zu möglichen kompensatorischen Effekten von Auslandsaufenthalten mangelt. Als wichtige Voraussetzung, um von einem Auslandsaufenthalt zu profitieren, wird eine erfolgreiche Akkulturation, also das Einleben in eine neue, womöglich fremde Kultur gesehen (Dentakos et al., 2017; Yakobov et al., 2019). Fraglich ist, ob sich Jugendliche hinsichtlich ihrer sozialen Herkunft im Akkulturationsverhalten unterscheiden.

Während Austauschschüler:innen aus sozial benachteiligten Familien häufig eine Diskrepanz zwischen dem niedrigeren sozioökonomischen Status der Herkunftsfamilie und dem höheren sozioökonomischen Status der Gastfamilie erleben (Lohmann, 2008), treffen Austauschschüler:innen aus privilegierten Familien zumeist auf Gastfamilien mit einem im Vergleich zur Herkunftsfamilie ähnlichen sozioökonomischen Status. Sie müssen daher eine geringere Anpassungsleistung erbringen als Austauschschüler:innen aus sozial benachteiligten Familien, die sich mit einem höheren Anpassungsdruck konfrontiert sehen (Koo et al., 2021). Folglich ist zu erwarten, dass Austauschschüler:innen aus sozial privilegierten Familien zumindest zu Beginn des Auslandsaufenthaltes ein höheres Akkulturationsverhalten zeigen als Austauschschüler:innen aus sozial benachteiligten Familien.

Weitgehend unbeantwortet ist die Frage, wie sich das Akkulturationsverhalten der Austauschschüler:innen über die Zeit entwickelt. Erste längsschnittliche Untersuchungen zeigen stabile Entwicklungsverläufe des Akkulturationsverhaltens (Serrano-Sánchez et al., 2021b), wobei aufgrund des sozial selektiven Zugangs zu Auslandsaufenthalten davon ausgegangen werden muss, dass es sich hier um eine Stichprobe von Austauschschüler:innen aus sozial privilegierten Familien handelt. Daher mangelt es an Befunden zur Entwicklung des Akkulturationsverhaltens von Austauschschüler:innen aus sozial benachteiligten Familien. Für den Verlauf der Entwicklung ist offen, inwiefern eine Diskrepanz durch Anpassungsanstrengungen überwunden wird oder ob sich eine mögliche soziale Benachteiligung über die Zeit stabilisiert.

Fragestellungen:

1. Gibt es Unterschiede im Akkulturationsverhalten der Austauschschüler:innen nach sozialer Herkunft?
2. Wie entwickelt sich das Akkulturationsverhalten der Austauschschüler:innen abhängig von der sozialen Herkunft im Laufe des Auslandsaufenthaltes?

Methode:

Die Datengrundlage bildet ein laufendes Projekt, das über drei Kohorten Austauschschüler:innen mit und ohne Auslandsstipendium begleitet. Die vorläufige Stichprobe umfasst N = 67 Schüler:innen der ersten Kohorte, die für mindestens drei Monate im Ausland zur Schule gehen. Die aktuell laufenden Befragungen der zweiten Kohorte (N = 81) werden im Herbst 2024 abgeschlossen sein. Die Angaben zur sozialen Herkunft, gemessen über den höchsten beruflichen Status der Eltern (HISEI; M = 70.75, SD = 15.87), stammen aus der Eingangsbefragung vor der Ausreise. Die Angaben zum Akkulturationsverhalten stammen aus den ersten fünf Wellen der längsschnittlichen Befragung während des Auslandsaufenthaltes. Die Austauschschüler:innen wurden in den ersten beiden Monaten des Auslandsaufenthaltes in einem Abstand von jeweils zwei Wochen befragt. Das Akkulturationsverhalten wurde mit der host-cultural behavioral engagement-Skala ($\alpha = .63 - .77$) erhoben (z.B. „In den letzten zwei Wochen habe ich meine Freizeit mit Freunden aus meinem Gastland verbracht“) (Serrano-Sánchez et al., 2021b). Zur Beantwortung der beiden Forschungsfragen wurde ein latentes Wachstumskurvenmodell gerechnet.

Ergebnisse und ihre Bedeutung:

Insgesamt berichten die Austauschschüler:innen ein im Mittel hohes Akkulturationsverhalten. Hierbei zeigen sich keine Unterschiede nach sozialer Herkunft (HISEI). Im Gegensatz zur bisherigen Forschung zeichnen sich signifikant positive Entwicklungsverläufe des Akkulturationsverhalten ab, die im Zusammenhang mit der sozialen Herkunft stehen. Die interindividuellen Unterschiede in der Varianz der Entwicklungsverläufe lassen sich zum Teil durch die soziale Herkunft erklären. Während das Akkulturationsverhalten der Austauschschüler:innen mit hohem HISEI in den ersten beiden Monaten des Aufenthaltes abnimmt, steigt jenes der Austauschschüler:innen mit niedrigem HISEI an. Diese erfolgreichen Anpassungsprozesse stellen eine zentrale Voraussetzung dar, damit Auslandsaufenthalte tatsächlich kompensatorische Effekte nach sich ziehen können.

Growing Abroad: Die Entwicklung interkultureller Kompetenzen von Schüler:innen während Auslandsaufenthalten und der Zusammenhang mit ihren täglichen sozialen und emotionalen Erfahrungen

Carina Schreiber¹, Michael Becker¹, Sabine Hornberg², Nadine Sonnenburg², Marion Peitz²

¹Technische Universität Dortmund - IFS, ²Technische Universität Dortmund - IADS

Globalisierung, Migration und Digitalisierung lassen unsere Welt näher zusammenrücken. Der Kontakt zu Menschen aus anderen Ländern und Kulturen ist nicht nur immer einfacher, sondern wird auch immer notwendiger und ist letztendlich unabdingbar. Gleichzeitig häufen sich weltweit kulturell bedingte Konflikte und immer mehr Länder erleben einen Rechtsruck. Sich in dieser kulturell diversen und international eng vernetzten Gesellschaft zurechtzufinden und sich erfolgreich in ihr bewegen zu können, fordert nicht nur Fremdsprachkompetenzen, sondern auch Fähigkeiten im persönlichen Umgang mit Menschen aus anderen Ländern und Kulturen sowie eine gewisse Offenheit ihnen gegenüber. Wie können Schüler:innen auf diese internationale Welt vorbereitet werden? Einen Weg diese Fähigkeiten und Einstellungen zu fördern, sehen Entscheidungstragende und vorherige Forschung in Auslandsaufenthalten (jenseits von Urlauben). So zeigten Studien mit Studierenden einen Zusammenhang zwischen der Teilnahme an Auslandsaufenthalten und einem Zuwachs in unterschiedlichen interkulturellen Kompetenzen (Genkova & Kruse, 2021; Jacobone & Moro, 2015; Roy et al., 2019; Terzuolo, 2018; Wolff & Borzikowsky, 2018; Zimmermann et al., 2020). Ob sich eine ähnliche Entwicklung bei Schüler:innen finden lässt, die sich nicht zuletzt durch ihr jüngeres Alter und einen unterschiedlich organisierten Ablauf der Austausche von Studierenden unterscheiden, ist bisher jedoch noch wenig erforscht. Außerdem ist bisher unklar, welche Prozesse und Mechanismen hinter der Entwicklung der Kompetenzen durch Auslandsaufenthalte stehen. Vorherige Forschung mit Studierenden deutet auf die Relevanz der Quantität sowie Qualität der internationalen sozialen Kontakte während Auslandsaufenthalten hin (Genkova & Kruse, 2021; Petersdotter et al., 2017). Ein weiterer entscheidender Faktor könnte außerdem das emotionale Erleben der Schüler:innen im Ausland sein. Nach der Broaden-and-Build Theory (Fredrickson, 2001) führt das Erleben von positiven Emotionen zu einer Erweiterung des Bewusstseins und Handlungsrepertoires und fördert somit die Entwicklung persönlicher sozialer, intellektueller und psychologischer Ressourcen. Negative Emotionen hingegen beschränken das Denken und behindern diese Entwicklung.

Die vorliegende Studie untersucht daher die Effekte von Erasmus+ Auslandsaufenthalten auf interkulturelle Kompetenzen von Schüler:innen sowie die Rolle, die ihre täglichen internationalen sozialen Kontakte und positiven und negativen Emotionen in dieser Entwicklung spielen. Es wird erwartet, dass die Schüler:innen nach ihrem Auslandsaufenthalt einen Zuwachs in ihren interkulturellen Kompetenzen aufweisen und dieser Zuwachs durch ihre Erfahrungen während des Auslandsaufenthalts erklärt werden kann. So wird angenommen, dass der Zuwachs durch häufigeren und qualitativ positiveren wahrgenommenen Kontakt zu ihren internationalen Peers sowie mehr positive und weniger negative Emotionen während des Auslandsaufenthalts erklärt werden kann.

In der vorliegenden Studie werden Schüler:innen mithilfe der Experience Sampling Methode (Csikszentmihalyi & Larson, 1987) täglich während ihrer ca. ein- bis zweiwöchigen Auslandsaufenthalte zu ihren sozialen Kontakten und Emotionen befragt. Im Rahmen des Prä- und Posttests sowie des sechs Wochen Follow-up Fragebogens werden zusätzlich kurz vor und nach den Auslandsaufenthalten die interkulturellen Kompetenzen (Kognitive Flexibilität, Fähigkeit zur Perspektivübernahme, Interesse an anderen Ländern und Selbstkonzept Englisch) der Schülerinnen und Schüler erfasst. Die Daten liegen somit in einer Multilevel Struktur vor, in der die täglich wiederholten Messungen (Level 1) in den Schüler:innen (Level 2) genestet sind. Latent change score-Modelle wurden herangezogen, um die Effekte der Auslandsaufenthalte auf die interkulturellen Kompetenzen der Schüler:innen zu messen und ihre täglichen sozialen Kontakte und Emotionen als Prädiktoren dieser Effekte zu prüfen.

Obgleich die Datenerhebung erst im Sommer 2024 final abgeschlossen sein wird, deuten vorläufige Analysen bereits darauf hin, dass selbst kurze Auslandsaufenthalte die Entwicklung interkultureller Kompetenzen bei Schüler:innen fördern. Die vorliegende Studie ermöglicht erstmals Einblicke in die täglichen intraindividuellen Erfahrungen von Schüler:innen während Auslandsaufenthalten und bietet Erklärungen für die damit einhergehende Entwicklung ihrer interkulturellen Kompetenzen.

Inter- und intraindividuelle Zusammenhänge zwischen Schulengagement und Akkulturationsverhalten von Austauschschüler:innen unter Berücksichtigung individueller und kontextueller Faktoren

Juan Serrano-Sánchez, Julia Zimmermann, Kathrin Jonkmann

FernUniversität in Hagen

Theoretischer Hintergrund:

In einer globalisierten Welt spielt die internationale Bildungsmobilität von Lernenden eine Schlüsselrolle für die Förderung ihrer Fähigkeit zur Partizipation in modernen globalen Gesellschaften sowie zur Überwindung kultureller Barrieren und zur Bekämpfung von Diskriminierung (Knight, 2003). Die hier betrachtete Form internationaler Bildungsmobilität betrifft adolescente Sojourners, d.h. Schüler:innen im Alter von 14 bis 18 Jahren, die ihr Heimatland für ein Schuljahr verlassen, um im Ausland die Schule zu besuchen. Um die kulturelle Transition zu bewältigen, nutzen die Schüler:innen individuelle Akkulturationsstrategien, indem sie ihr Verhalten in unterschiedlichem Maße am Gastland und am Herkunftsland (host- bzw. home-cultural behavioral engagement) orientieren. Der Erfolg dieser Transition lässt sich an verschiedenen Indikatoren der Adaptation messen. Besonders die Schule ist ein zentraler Bereich der Auslandserfahrung, da die Schüler:innen dort eine substanzielle Zeit verbringen und mit Mikrosystemen (z.B. Kommiliton:innen, Lehrkräften) und Makrosystemen (z.B. dem Schulsystem) der Gastkultur in Kontakt kommen (Hein et al., 2017; Motti-Stefanidi & Master, 2013). Schulengagement gilt daher als entscheidender Indikator für eine erfolgreiche Anpassung (Berry et al., 2006; Rienties et al., 2012).

Vor diesem Hintergrund untersuchte die vorliegende Studie die längsschnittlichen transaktionalen Zusammenhänge zwischen Schulengagement (affektiv, verhaltensbezogen und kognitiv) und Akkulturationsverhalten (host- und home-cultural behavioral engagement) von Schüler:innen auf inter- und intraindividuelle Ebene, wobei potenzielle Einflüsse individueller Merkmale

(Geschlecht, Alter, Sprachkenntnisse, Persönlichkeitsfaktoren) und kontextueller Bedingungen (geplante Aufenthaltsdauer im Ausland, kulturelle Distanz) berücksichtigt wurden. Hierbei wurde ein Random-Intercept Cross-Lagged Panel Model verwendet (RI-CLPM, Hamaker et al., 2015).

Fragestellung:

Wie gestaltet sich die Beziehung zwischen Schulengagement und Akkulturationsverhalten adoleszenter Sojourners im Verlauf der Zeit auf intra- und interindividueller Ebene? Welche individuellen und kontextuellen Merkmale spielen dabei eine Rolle?

Methode:

Die Daten stammen aus den ersten vier Wellen des längsschnittlichen Forschungsprojekts [Projektname], das deutsche Schüler:innen umfasst, die im Schuljahr 2016/17 einen Auslandsaufenthalt absolvierten. In die Analyse wurden alle Teilnehmenden einbezogen, die mindestens einen der Fragebögen zwischen t0 und t3 ausfüllten (N = 1,299). Soziodemographische und psychologische Merkmale wurden mittels etablierter Messinstrumente erhoben. Schulengagement wurde mit der Student Engagement in School Scale erfasst (Lam et al., 2014). Host- und home-cultural behavioral engagement wurden mit Bezug zu verschiedenen relevanten Lebensbereichen der adoleszenten Sojourners erfasst (Autoren, 2021). Persönlichkeit wurde vor der Abreise mit dem 21-Item Big Five Inventory gemessen (Rammstedt & John, 2005). Kulturelle Distanz wurde anhand der Landeswerte für Individualismus gemäß Hofstede et al. (2010) bestimmt.

Ergebnisse:

Das RI-CLPM zeigte eine sehr gute Anpassung an die Daten ($\chi^2(20) = 30,415$, $p = .063$, CFI = .998, RMSEA = .019, SRMR = .011).

Auf intraindividueller Ebene zeigten sich positive Zusammenhänge zwischen temporären Anstiegen des affektiven und verhaltensbezogenen Schulengagements und dem späteren host-cultural behavioral engagement, während des fortgeschrittenen Auslandsaufenthalts (t2–t3). Der temporäre Anstieg affektiven Schulengagements war hingegen über beide Intervalle (t1-t2, t2-t3) negativ mit dem späteren home-cultural behavioral engagement verbunden. Ein temporärer Anstieg im home-cultural behavioral engagement war negativ mit nachfolgendem affektiven und verhaltensbezogenen Schulengagement assoziiert (t2–t3).

Auf interindividueller Ebene bestand eine signifikante positive Beziehung zwischen host-cultural behavioral engagement und verhaltensbezogenem sowie kognitivem Schulengagement. Hingegen wurden keine signifikanten Zusammenhänge zwischen home-cultural behavioral engagement und den drei Facetten des Schulengagements identifiziert. Sprachkenntnisse des Gastlandes sagten host-cultural behavioral engagement positiv vorher. Die Persönlichkeitsmerkmale zeigten differenzierte Assoziationen mit allen Studienvariablen. Die geplante Dauer des Aufenthalts war negativ mit dem verhaltensbezogenen Schulengagement assoziiert, während der Aufenthalt in einem individualistischeren Land als Deutschland affektives und verhaltensbezogenes Schulengagement sowie home-cultural behavioral engagement positiv vorhersagte.

Diskussion:

Unsere Ergebnisse erweitern frühere Erkenntnisse, indem sie eine differenziertere Perspektive auf die längsschnittlichen Zusammenhänge zwischen Schulengagement und Akkulturationsverhalten unter Berücksichtigung von individuellen und kontextuellen Bedingungen bieten, und tragen zur Entwicklung gezielter Interventionen und Unterstützungsmechanismen bei, die die Anpassung und das Wachstum von Jugendlichen im Rahmen ihrer Auslandsaufenthalte fördern.

S-35: Unterrichtsstörungen aus Sicht von Lehrpersonen und Schüler*innen

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: EO 165

Chair der Sitzung: Sabine Glock

Chair der Sitzung: Claudia Schuchart

Unterrichtsstörungen aus Sicht von Lehrpersonen und Schüler*innen

Chair(s): **Sabine Glock** (Bergische Universität Wuppertal, Deutschland)

Diskutant*in(nen): **Claudia Schuchart** (Bergische Universität Wuppertal, Deutschland)

Unterrichtsstörungen stören mehr als nur den Unterricht und gehen nicht nur zu Lasten des Wohlbefindens von Lehrkräften, sondern auch von dem der Schüler*innen. In diesem Symposium sollen Auswirkungen von Unterrichtsstörungen sowohl auf Lehrkräfte als auch auf Grundschüler*innen und Jugendlichen untersucht werden. Der erste Beitrag in dem Symposium zeigt das Erleben von Grundschüler*innen auf. Die Autor*innen berichten von einer Studie, in der die Grundschüler*innen nach ihrer schulischen Zufriedenheit in Zusammenhang mit der Lehrer-Schüler-Beziehung und ihrer Wahrnehmung von Unterrichtsstörungen befragt wurden. Die Autor*innen zeigen eine höhere schulische Zufriedenheit von Mädchen im Vergleich zu Jungen. Während eine positive Lehrer-Schüler-Beziehung positiv mit der schulischen Zufriedenheit korreliert, ist die Wahrnehmung von Unterrichtsstörungen negativ damit verbunden. Damit leidet die schulische Zufriedenheit unter Unterrichtsstörungen. Das Geschlecht spielt nicht nur bei der schulischen Zufriedenheit eine Rolle, sondern Jungen stören generell auch häufiger den Unterricht als Mädchen (Demanet & Van Houtte, 2012). Hier wird der Druck zu geschlechterkonformen Verhalten und der Wunsch, sich als besonders maskulin zu zeigen, als Ursache diskutiert (Kessels & Heyder, 2020). Die Autor*inne des zweiten Beitrags des Symposiums erfassten über drei unterschiedliche Messzeitpunkte zu Beginn der 7. Klasse, zu Ende der 7. Klasse und der 8. Klasse das selbsteingeschätzte Fehlverhalten sowie der wahrgenommene Druck zu geschlechterkonformen Verhalten. Jungen berichteten einen Anstieg von schulischem Fehlverhalten über die drei Messzeitpunkte, welcher über den Druck, sich geschlechterkonform zu verhalten, vorhergesagt wurde. Bei Mädchen konnten diese Effekte nicht gefunden werden. Unterrichtsstörungen sind mehr als nur Störungen des Unterrichts, sie dienen auch der sozialen Identität der Schüler*innen. Solch herausfordernde Störungen und Disziplinprobleme im Unterricht tragen zum Belastungserleben von Lehrkräften bei (Skaalvik & Skaalvik, 2017). Der letzte Beitrag betrachtet nun nicht die Art oder Häufigkeit der Störung als ausschlaggebend für das Belastungserleben von Lehrkräften, sondern geht von der Annahme aus, dass die Persönlichkeit der Lehrpersonen und deren Bewältigungsmechanismen das Belastungserleben bestimmen. Die Autor*innen zeigen in einer Beobachtungsstudie, dass die Lehrkräfte, die über höheren Neurotizismus, mehr Angst vor negativer Bewertung und höherer sozialer Überlastung berichten, auch eine höhere Anzahl an Unterrichtsstörungen in ihrem Unterricht wahrnehmen. Damit können fehlende Bewältigungsmechanismen dazu führen, dass das Belastungserleben lediglich über die eigene, vielleicht verzerrte Wahrnehmung des Unterrichts erhöht wird. Insgesamt betrachtet, ist es an der Lehrkräfteausbildung angehenden Lehrpersonen bereits im Studium effektive Classroom Management Strategien an die Hand zu geben, damit Unterrichtsstörungen im Klassenzimmer minimiert werden können. Davon profitieren alle am Unterricht beteiligten Personen!

Beiträge des Symposiums

Die Wahrnehmung von Unterrichtsstörungen als Prädiktoren schulischer Zufriedenheit bei Grundschulkindern

Claudia Kastens, Martin van Wickeren

Bergische Universität Wuppertal, Deutschland

Das Erleben schulischer Zufriedenheit als affektives Merkmal ist neben leistungsbezogenen und motivationalen Merkmalen ein wichtiger Indikator für schulischen Erfolg (Klieme, 2016). Zu den Prädiktoren schulischen Wohlbefindens und Zufriedenheit bei Grundschulkindern zählen u.a. Merkmale des Lern- und Arbeitsverhaltens sowie der sozialen Integration, wie die Beziehung zu anderen Schüler*innen und Lehrkräften (Nicklaussen, 2012). Für Unterrichtsstörungen gilt, dass sie ein Prädiktor für das Belastungserleben von Lehrkräften sind (Kastens & vanWickeren, 2023; Mijakoski et al., 2022). Dieser Beitrag untersucht, ob die schulische Zufriedenheit von Grundschulkindern nicht nur von Faktoren des Lern- und Arbeitsverhaltens (Selbstregulation, Selbstwirksamkeitserwartung, Impulsivität) und der sozialen Integration (Beziehung zwischen Schülerinnen sowie zwischen Schülerinnen und Lehrperson) abhängt, sondern auch von deren Wahrnehmung von Unterrichtsstörungen. Wir erwarten, dass Selbstwirksamkeitserwartung, Selbstregulationskompetenz und geringe Impulsivität mit höherer schulischer Zufriedenheit korrelieren. Eine positive soziale Integration, besonders die Beziehung zu Lehrkräften, sollte ebenfalls mit höherer Zufriedenheit bei Grundschulkindern einhergehen. Darüber hinaus prüfen wir, ob die Wahrnehmung von Unterrichtsstörungen auch für Schüler*innen negative Zusammenhänge mit der schulischen Zufriedenheit aufweist und ob diese die zuvor beleuchteten Effekte moderiert.

Es wurden $N = 432$ nordrhein-westfälische Grundschüler*innen der vierten Jahrgangsstufe (55% Mädchen) zur schulbezogenen Zufriedenheit („In unserer Schule fühle ich mich wohl“, 4 Items, $\alpha = .87$, $M = 1.70$, $SD = 0.81$) und Wahrnehmung von Unterrichtsstörungen („Die Kinder in meiner Klasse können sich gut an die Klassenregeln halten.“, 3 Items, $\alpha = .79$, $M = 2.81$, $SD = 0.70$) befragt. Die Indikatoren des Arbeits- und Lernverhaltens umfassten die Fähigkeit zur Selbstregulation („Ich kann mich lange Zeit auf eine Sache konzentrieren, wenn es nötig ist.“, 6 Items, $\alpha = .75$, $M = 1.88$, $SD = 0.56$), Selbstwirksamkeitserwartung („Auch wenn ich im Unterricht eine schwierige Aufgabe lösen soll, glaube ich, dass ich das schaffen werde.“, 3 Items, $\alpha = .68$, $M = 1.67$, $SD = 0.57$) und Impulsivität („Ich gerate oft in Schwierigkeiten, weil ich etwas tue, ohne zu überlegen.“, 4 Items, $\alpha = .73$, $M = 2.85$, $SD = 0.76$). Die Indikatoren der sozialen Integration umfassten die Beziehung zwischen Schüler*innen („Die meisten Kinder in der Klasse sind freundlich zu mir.“, 5 Items, $\alpha = .70$, $M = 1.65$, $SD = 0.51$) und die Beziehungen zu Lehrpersonen („Mit meiner Klassenlehrerin / meinem Klassenlehrer komme ich gut aus.“, 5 Items, $\alpha = .74$, $M = 1.39$, $SD = 0.47$). Alle Items wurden auf einer vierstufigen Skala bewertet (1 = stimmt genau, 4 = stimmt nicht). Für die Analyse wurden multiple Regressionsmodelle berechnet: M1: Indikatoren des Arbeits- und Lernverhaltens, M2: Indikatoren des Arbeits- und Lernverhalten + Wahrnehmung von Störungen, M3: Merkmale der sozialen Integration, M4: Merkmale der sozialen Integration + Wahrnehmung von Störungen, M5: Alle Prädiktoren zusammen). In allen Modellen wurde um Geschlechtereffekte kontrolliert.

Die untersuchten Faktoren korrelieren erwartungskonform mit der erlebten schulischen Zufriedenheit. Die Wahrnehmung der Beziehung zwischen Schüler*innen und Lehrperson zeigt die stärkste Korrelation $r = .56$. Mädchen berichten signifikant höhere Zufriedenheit als Jungen. Die Ergebnisse der Regressionsanalysen zeigen, dass die Wahrnehmung von Unterrichtsstörungen in drei Modellen (M2, M4 und M5) einen additiven signifikant positiven Effekt auf die schulbezogene Zufriedenheit hat ($\beta = .19$, $R^2 = .26$; $\beta = .17$, $R^2 = .41$; $\beta = .14$, $R^2 = .40$), wobei in M5 der Effekt der Wahrnehmung der Beziehung zu Lehrpersonen am größten ist ($\beta = .49$). Auch Geschlechterunterschiede bleiben bestehen.

Die Ergebnisse zeigen, dass die Wahrnehmung einer positiven Beziehung zwischen Schüler*innen und Lehrperson für Grundschulkinder am stärksten mit der schulbezogenen Zufriedenheit zusammenhängt. Darüber hinaus ist erweist sich die Wahrnehmung von Unterrichtsstörungen für Grundschüler*innen als negativer Prädiktor schulbezogener Zufriedenheit.

Geschlechtsunterschiede im schulischem Fehlverhalten in der Adoleszenz: Zur Rolle des Gender Conformity Pressure

Anke Heyder¹, Margriet van Hek², Mieke Van Houtte³

¹Ruhr-Universität Bochum, ²Radboud University, ³Ghent University, Belgien,

Viele Studien haben bereits aufgezeigt, dass männliche Jugendliche mehr Störverhalten in der Schule zeigen als weibliche (z.B. Demanet & Van Houtte, 2012). Dieser Geschlechtsunterschied wächst über die Adoleszenz hinweg teilweise noch (Yunger et al., 2004) und trägt zu Geschlechtsunterschieden im Bildungserfolg zuungunsten von männlichen Jugendlichen bei (z.B. Downey & Vogt Yuan, 2005). Der vorliegende Beitrag adressiert aus psychologisch-soziologischer Perspektive die Frage, warum in der Adoleszenz männliche Jugendliche verstärkt schulisches Fehlverhalten zeigen. Im Zentrum steht dabei das Konstrukt des „Gender Conformity Pressure“, welches als Teil der Geschlechtsidentität gilt und den wahrgenommenen Druck, sich geschlechtskonform zu verhalten, beschreibt (Egan & Perry, 2001).

In der Adoleszenz stehen Jugendliche vor der Anforderung eine geschlechtliche Identität zu entwickeln (Shell Jugendstudie 2015) und erleben gleichzeitig einen hohen Druck durch Peers (Steinberg & Monohan, 2007). Frühere Studien mit kleineren Stichproben oder theoretische Arbeiten haben schulisches Störverhalten und Regelverstöße bereits als besonders maskulin - und damit die männliche Geschlechterrolle als wenig kompatibel mit der Schülerrolle als die weibliche Geschlechterrolle - beschrieben (z.B. Jackson, 2002; Kessels & Heyder, 2020; Mickelson, 1989). Längsschnittliche Untersuchungen mit großen Stichproben lagen bis jetzt jedoch kaum vor. Vor diesem Hintergrund untersucht die vorliegende Studie, ob der wahrgenommene Druck sich geschlechterstereotypenkonform verhalten zu müssen den Anstieg im Störverhalten bei männlichen Jugendlichen in der Adoleszenz vorhersagen kann.

Um unsere Annahme zu prüfen, analysierten wir Daten von 4.209 flämischen Jugendlichen (Alter T1 M = 12.12; 49.5% weiblich), die im Rahmen der Procrustes-Längsschnittstudie erhoben worden waren. Die Jugendlichen berichteten ihr Geschlecht (männlich, weiblich), schulisches Fehlverhalten (Demanet & Van Houtte, 2012; Wie häufig zeigst du folgendes Verhalten? 17 Items, z.B. Zu spät kommen, Schule schwänzen, Cronbachs Alpha $\geq .80$) und Gender Conformity Pressure (Egan & Perry, 2001; acht geschlechtsspezifische Items, Bsp. „Ich werde wütend, wenn jemand sagt ich verhalte mich wie ein Mädchen“, „Es ist wichtig für mich wie die anderen Jungen zu sein“. Cronbachs Alpha $\geq .81$) zum Anfang von Klasse 7 (T1), zum Ende von Klasse 7 (T2) und zum Ende von Klasse 8 (T3). Als Kovariate wurden zusätzlich das Alter, der sozioökonomische Status, traditionelle Geschlechterrolleneinstellungen, Wohlbefinden, Mathematikleistung, Schulform und Migrationsstatus erfasst.

Die männlichen Jugendlichen berichteten zu allen drei Messzeitpunkten einen signifikant höheren Druck sich geschlechtskonform zu verhalten sowie signifikant mehr schulisches Fehlverhalten. Zur Prüfung unserer Hypothesen zum Zusammenhang zwischen beiden Konstrukten rechneten wir 3-Ebenen-Wachstumskurvenmodellen, welche die genestete Datenstruktur (Beobachtungen genestet in Personen, genestet in Schulen) sowie die genannten Kovariaten statistisch berücksichtigten. Diese zeigten, dass männliche Jugendliche signifikant mehr eigenes Fehlverhalten in der Schule berichteten als weibliche Jugendliche und dass dieser Geschlechtsunterschied im untersuchten Zeitraum signifikant wuchs. In Übereinstimmung mit unseren Erwartungen sagte ein höherer Gender Conformity Pressure bei männlichen Jugendlichen außerdem eine Zunahme schulischen Fehlverhaltens vorher. Bei weiblichen Jugendlichen zeigte sich kein signifikanter Zusammenhang zwischen Gender Conformity Pressure und schulischen Fehlverhalten.

Die Ergebnisse deuten darauf hin, dass schulisches Fehlverhalten von männlichen Jugendlichen zu einem Teil als Ausdruck einer maskulinen Geschlechtsidentität interpretiert werden kann. Sie stützen damit frühere Befunde, dass bestimmte Männlichkeitsvorstellungen sich nachteilig auf den Bildungserfolg männlicher Jugendlicher auswirken können. Die Ergebnisse werden hinsichtlich ihrer Limitationen und möglicher Implikationen für Forschung und Praxis diskutiert.

Die Wahrnehmung von Unterrichtsstörungen vor dem Hintergrund von Neurotizismus, Angst vor negativer Bewertung und soziale Überlastung und longitudinale Folgen für die beruflichen Beschwerden der Lehrperson

Gabriel Jenni¹, Alexander Wettstein¹, Ida Schneider¹, Fabienne Kühne¹, Martin grosse Holtforth², Roberto La Marca³

¹Pädagogische Hochschule Bern, Schweiz, ²Universität Bern, Schweiz, ³Clinica Holistica Engiadina, Susch, Schweiz

Störendes Schülerverhalten gilt als Risikofaktor für ein ungünstiges Belastungserleben von Lehrpersonen (Dicke et al., 2014; Evers et al., 2004; McCormick & Barnett, 2011). Aber auch Lehrpersonen tragen ungewollt zu Störungen bei und stören den eigenen Unterricht rund zweimal pro Lektion (Wettstein & Scherzinger, 2022). Weiter können ungünstige Lehrerreaktionen das Problemverhalten weiter verstärken. Eine rein schülerzentrierte Perspektive greift deshalb zu kurz. Der interaktionale Begriff Unterrichtsstörungen geht über einzelne Kinder und Lehrpersonen hinaus und untersucht, inwieweit der Unterricht als Ganzes gestört ist (Wettstein & Scherzinger, 2022). Unterrichtsstörungen werden dabei als Resultat misslingender pädagogischer Interaktionen verstanden.

Unterrichtsstörungen sind positiv mit biologischem Stress der Lehrperson assoziiert (La Marca et al., 2023) und wirken sich ungünstig auf die Gesundheit der Lehrpersonen aus (Ingersoll, 2001). Dadurch können die Qualität des Unterrichts sowie die Motivation und Leistung der Schülerinnen und Schüler beeinträchtigt werden (Klusmann et al., 2016). Eine effektive Klassenführung und eine unterstützende Lehrer-Schüler-Beziehung sind hingegen schützende Faktoren (Wettstein & Scherzinger, 2022).

Nun zeigen neuste Studien: Es sind weniger die tatsächlich auftretenden Unterrichtsstörungen, welche Lehrpersonen psychisch belasten. Entscheidend ist vielmehr, wie Lehrpersonen diese vor dem Hintergrund ihrer Persönlichkeit und ihrer Bewältigungsstrategien wahrnehmen. Lehrpersonen nehmen Unterrichtsstörungen nicht objektiv (Scherzinger & Wettstein, 2019), sondern vermittelt durch die Brille ihrer eigenen Persönlichkeit und Bewältigungsstrategien wahr (Kokkinos et al., 2005).

Diese Beobachtungsstudie mit 42 Lehrpersonen (28 weiblich, Alter = 39.66, SD = 11.99) verglich objektiv beobachteten Unterricht, insbesondere Unterrichtsstörungen, mit dem von der Lehrperson wahrgenommenen Unterricht. Es wurde untersucht, inwieweit die Wahrnehmung von Unterrichtsstörungen durch Neurotizismus, Angst vor negativer Bewertung und soziale Überlastung beeinflusst wird. Darüber hinaus wurde über einen Zeitraum von zwei Jahren der Zusammenhang zwischen den wahrgenommenen Unterrichtsstörungen und der Zunahme beruflicher Beschwerden betrachtet.

Die Ergebnisse zeigten, dass Neurotizismus, Angst vor negativer Bewertung und soziale Überlastung positiv mit mehr wahrgenommenen Unterrichtsstörungen assoziiert waren. Neurotizismus führte aber nicht direkt zu mehr wahrgenommenen Unterrichtsstörungen, sondern diese Beziehung wurde durch die Angst vor negativer Bewertung und sozialer Überlastung mediiert. Zudem verzerrten Neurotizismus, Angst vor negativer Bewertung und soziale Überlastung die Wahrnehmung der Lehrpersonen in Form von Über- und Unterschätzung von Unterrichtsstörungen. Ausserdem waren wahrgenommene Unterrichtsstörungen mit einer Akkumulation von beruflichen Beschwerden über zwei Jahre assoziiert.

Störendes Schülerverhalten gilt als eine zentrale Ursache für ein ungünstiges Belastungserleben von Lehrpersonen. Daher wird in der Aus- und Weiterbildung von Lehrpersonen traditionell der Schwerpunkt daraufgelegt, den Umgang der Lehrperson mit herausforderndem Verhalten der Schülerinnen und Schüler zu verbessern. Dieser Ansatz greift jedoch möglicherweise zu kurz. Denn Störungen gehen nicht nur von Schülerinnen und Schülern, sondern auch von Lehrpersonen aus und im interaktionalen Verlauf entstehen Unterrichtsstörungen. Zudem nehmen Lehrpersonen Unterrichtsstörungen nicht objektiv, sondern vermittelt durch ihre Persönlichkeit und Bewältigungsstile wahr. Dies kann dazu führen, dass Lehrpersonen das Ausmaß an Unterrichtsstörungen über- oder unterschätzen. Die Untersuchung von Selbstberichten in Kombination mit Verhaltensbeobachtung ist entscheidend, um die Wahrnehmung des Unterrichts durch Lehrpersonen besser zu verstehen. Ebenso für die Ermittlung von Persönlichkeitsmerkmalen und sozialen Stressoren, die zu einer verzerrten Wahrnehmung beitragen, um gesundheitliche Beschwerden der Lehrpersonen vorzubeugen. Die Lehrerausbildung muss Lehrpersonen dabei unterstützen, mit der Angst vor negativer Bewertung und sozialer Überlastung umzugehen, und gute Lehrer-Schüler-Beziehungen sowie ein effektives Klassenmanagement aufzubauen. Es ist wichtig, die kumulativen Langzeiteffekte einer verzerrten Wahrnehmung zu erkennen, um langfristigen Stress bei Lehrpersonen zu vermeiden.

S-36: Ursachen von digitalen Ungleichheiten und Konsequenzen für den Bildungserfolg

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 148

Chair der Sitzung: Markus Lörz

Ursachen von digitalen Ungleichheiten und Konsequenzen für den Bildungserfolg

Chair(s): **Markus Lörz** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland)

Soziale Ungleichheiten in der Bildungsbeteiligung und dem Kompetenzerwerb bestehen in verschiedenen Ländern Europas bereits seit Jahrzehnten (OECD, 2023; Schütte et al., 2024). In den vergangenen Jahren schafft die Digitalisierung im Bildungsbereich zudem Dynamiken, die von den Schüler*innen zunehmend digitale Kompetenzen einfordern, deren Auswirkungen auf soziale Ungleichheiten allerdings noch nicht ausreichend erforscht sind.

Ein zentrales Ergebnis der bisherigen Forschung ist, dass die digitalen Möglichkeiten und digitalen Kompetenzen der Schüler*innen nach sozialer Herkunft variieren (Senkbeil et al., 2019; Drossel et al., 2020; Passaretta & Gil-Hernández, 2023). Die Forschung verwendet hierbei die Begriffe des first level (Zugang), second level (Nutzungs) und third level (Konsequenzen) digital divide (Van Dijk, 2000; Van Deursen & Helsper, 2015). Das vorgeschlagene Symposium befasst sich ausgehend von diesen Überlegungen mit den Fragen, wie digitale Unterschiede entstehen (Ursachen), wie sich diese im Zeitverlauf verändern (Entwicklung) und welche Auswirkungen diese auf den Bildungserfolg der Schüler*innen haben (Konsequenzen). Die Session fokussiert auf die Gruppe der Schüler*innen im Primar- und Sekundarbereich und diskutiert aus psychologischer, erziehungswissenschaftlicher und soziologischer Perspektive vier verschiedene Forschungsbeiträge. Die Forschungsbeiträge verwenden hierbei jeweils unterschiedliche Datenzugänge und Analysemethoden, setzen sich aber gleichermaßen empirisch und theoretisch mit Fragen der sozialen Ungleichheit im Zuge der Digitalisierung auseinander.

Der *erste Beitrag* beschreibt aus soziologischer Perspektive wie sich in den vergangenen zwanzig Jahren (PISA 2000-2022) der Zugang und die Nutzung digitaler Medien verändert haben und in welchen Aspekten Unterschiede nach sozialer Herkunft bestehen. Hierbei wird aus Perspektive der kulturellen Reproduktionstheorie (Bourdieu & Passeron, 1971; DiMaggio, 2004) und aus Statuserhaltungsperspektive (Boudon, 1974) untersucht, ob mit der zunehmenden Bedeutung digitaler Medien, die digitalen Ungleichheiten nach sozialer Herkunft im Zeitverlauf zu- oder abnehmen.

Im *zweiten Beitrag* wird aus erziehungswissenschaftlicher Perspektive anhand aktueller ICILS 2023-Daten (Eickelmann et al., in Druck) der Status quo schulischer Merkmale des Erwerbs digitaler Kompetenzen bestimmt, wobei erfolgreiche Schulen mit einem hohen Anteil sozioökonomisch benachteiligter Schüler*innen – sog. resiliente Schulen (Drossel et al. 2020) – mit nicht-resilienten Schulen kontrastiert werden. Hierbei steht die Frage im Vordergrund, inwiefern sich verschiedene schulische Voraussetzungs- und Prozessmerkmale (Rožman et al., 2023) des Erwerbs digitaler Kompetenzen an resilienten und nicht resilienten Schulen unterscheiden.

Im *dritten Beitrag* wird der Fokus auf die zugrunde liegenden Mechanismen gerichtet und aus entwicklungspsychologischer Perspektive der Frage nachgegangen, in welchen Phasen des individuellen Lebensverlaufs motivationsbezogene Unterschiede in Bezug auf die Nutzung digitaler Medien entstehen und inwieweit herkunftsspezifische Unterschiede in der ICT Literacy darüber vermittelt werden. Ausgehend von der Scholary Culture Theorie (Sikora et al., 2019) werden die verschiedenen Aspekte der sozialen Herkunft als potentielle Lernanregung verstanden und anhand der NEPS-Daten (SC 4) analysiert, welche Merkmale der sozialen Herkunft besonders einflussreich sind.

Im *vierten Beitrag* wird schließlich aus sozialwissenschaftlicher Perspektive der Fokus auf die Nutzung digitaler Leseapps in der Schule und in der Freizeit gelegt. Zentral ist hierbei die Frage, inwieweit die Schule zur Kompensation oder zur Verstärkung sozialer Ungleichheiten beiträgt. Hierzu wird auf Basis einer dänischen Fallstudie und anhand einer Lese-App analysiert, inwieweit Schüler*innen eine solche Lese-App Zuhause oder in der Freizeit nutzen und inwieweit herkunftsspezifische Unterschiede bestehen.

Die vier Beiträge greifen verschiedene Fragen von der Entwicklung digitaler Ungleichheiten im Kohortenvergleich (Beitrag 1), über die aktuelle schulische Situation (Beitrag 2) bis hin zu zugrundeliegenden Mechanismen digitaler Ungleichheiten im individuellen Lebensverlauf (Beitrag 3) sowie potentielle Interventionsmöglichkeiten (Beitrag 4) auf. Organisatorisch ist das Symposium so angelegt, dass zu jedem Beitrag die Möglichkeit besteht Rückfragen zu stellen. Es wird aber auch ein Zeitfenster eingeräumt, um die Forschungsergebnisse aus soziologischer, erziehungswissenschaftlicher und psychologischer Perspektive einzuordnen und hinsichtlich der Entstehung digitaler Ungleichheiten, deren Konsequenzen und potentielle Interventionsmöglichkeiten zu diskutieren.

Beiträge des Symposiums

Entwicklung sozialer Ungleichheiten im Zeitverlauf im Zugang und der Nutzung digitaler Medien

Birgit Becker¹, Markus Lörz²

¹Goethe-Universität Frankfurt, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Theoretischer Hintergrund:

In Deutschland sind soziale Ungleichheiten in der Bildungsbeteiligung und im Kompetenzerwerb stark ausgeprägt (Autor*innengruppe Bildungsberichterstattung, 2022; OECD, 2023). Wenngleich die Ungleichheiten am Übergang ins Gymnasium über die Zeit tendenziell abnehmen (Breen et al., 2010; Pollak & Müller, 2020), finden sich Hinweise, dass sich die Ungleichheitsmechanismen zunehmend auch auf andere bildungsrelevante Aspekte verlagern (Lörz, 2017). Als Erklärung für veränderte Ungleichheitsmechanismen werden in der Ungleichheitsforschung oftmals die Überlegungen der Statuserhalttheorie (Boudon, 1974) und der kulturellen Reproduktionstheorie (Bourdieu & Passeron, 1971) herangezogen. Im Rahmen der Effectively-Maintained-Inequality-These (Lucas, 2001) wird argumentiert, dass sobald Bildung als exklusives Gut für die intergenerationale Statusreproduktion an Bedeutung verliert, investieren Schüler*innen aus privilegierten Familien zunehmend in andere (bildungs-)relevante Aspekte. Auf diese Weise versuchen sie sich einen Wettbewerbsvorteil zu verschaffen und die Chancen auf die Reproduktion des Bildungs- und Berufsstatus ihrer Eltern zu erhöhen. Mit der zunehmenden Bedeutung digitaler Medien und digitaler Kompetenzen (Hargittai & Hsieh, 2013; Blossfeld et al., 2018) wäre dementsprechend zu erwarten, dass diese Aspekte über die Zeit zunehmend von den Schüler*innen aus privilegierten Familien genutzt und entsprechende Kompetenzen eingeübt werden (DiMaggio et al., 2004).

Fragestellung:

Der vorliegende Beitrag greift diese veränderten Bedingungen auf und setzt sich aus soziologischer Ungleichheitsperspektive mit der Frage auseinander, ob im Zugang und der Nutzung digitaler Medien herkunftsspezifische Unterschiede bestehen und wie sich diese in den vergangenen 20 Jahren verändert haben. Ausgehend von der Unterscheidung zwischen first level (Zugang zu digitalen Medien) und second level (Nutzung digitaler Medien) digital divide (Van Dijk, 2000), wird argumentiert, dass sowohl im Zugang als auch in der Nutzung digitaler Medien herkunftsspezifische Unterschiede bestehen. Mit Blick auf die zeitliche Entwicklung wird davon ausgegangen, dass mit der zunehmenden Digitalisierung die Unterschiede im Zugang zu digitalen Medien zwar über die Zeit abnehmen (first level), aber in Anlehnung an die Überlegungen der Effectively-Maintained-Inequality-These sich die sozialen Ungleichheiten über die Zeit zunehmend auf die Art der Nutzung digitaler Medien verlagern (second level).

Daten und Methoden

Zur Überprüfung der theoretischen Überlegungen ziehen wir die PISA-Daten der OECD (2023) von 2000 bis 2022 heran. Bei PISA handelt es sich zwar um eine international vergleichende Schulleistungsstudie, aber im Rahmen dieser Befragungen werden auch Informationen über das Medienverhalten von 15-jährigen Schüler*innen erhoben, sowie deren sozialen und finanziellen Bedingungen im Elternhaus. Der Fokus liegt in diesem Beitrag auf Deutschland und den Veränderungen im Zeitverlauf. Die PISA-Daten der Jahre 2000-2022 werden für die Analyse synchronisiert und in einen gemeinsamen Analysedatensatz integriert. In der Analyse wird in zwei Schritten vorgegangen: Zunächst wird im ersten Analyseschritt deskriptiv betrachtet, wie sich die herkunftsspezifischen Unterschiede im Zugang und der Nutzung digitaler Medien zwischen 2000 und 2022 verändert haben. Im zweiten Schritt wird anhand logistischer Regressionsanalysen untersucht, inwieweit die beobachtbaren Veränderungen über die Zeit statistisch signifikant ausfallen. Für diese Zwecke werden die Regressionskoeffizienten der verschiedenen Modelle miteinander verglichen (Chi2-Test) und eine Aussage getroffen, ob auf Ebene des Zugangs (first level) oder auf Ebene der Nutzung (second level) digitaler Medien eine Zu- oder Abnahme sozialer Ungleichheiten zu beobachten ist.

Ergebnisse und Bedeutung:

Erste Ergebnisse zeigen, dass die sozialen Ungleichheiten beim Zugang zu Internet und Hardware bereits in den 2000er Jahren massiv abgenommen haben und inzwischen bei 15-Jährigen ein nahezu universaler Zugang zu ICT besteht. Allerdings finden sich selbst auf der Dimension „Zugang“ (first level) noch (geringe) Ungleichheiten, wie etwa bei der Anzahl an digitalen Geräten zu Hause. Beim Nutzungsverhalten (second level) differieren die Ergebnisse, je nachdem, welche Art der Nutzung betrachtet wird. Bei der Nutzung von ICT zur Kommunikation gab es Anfang der 2000er Jahre massive soziale Unterschiede, die sich jedoch sehr schnell abgebaut haben. Bei der Nutzung von ICTs zu informations- und bildungsbezogenen Zwecken ist das Ergebnis hingegen weniger eindeutig.

Digitale Ungleichheiten überwinden – Voraussetzungs- und Prozessmerkmale des Erwerbs digitaler Kompetenzen an resilienten und nicht resilienten Schulen in Deutschland

Jan Niemann, Kerstin Drossel, Birgit Eickelmann
Universität Paderborn

Theoretischer Hintergrund:

Mit der zunehmenden gesellschaftlichen digitalen Transformation gewinnt der Erwerb von digitalen Kompetenzen an Bedeutung. Der daran anknüpfende schulische Bildungsauftrag fokussiert auf die gezielte Förderung dieser Schlüsselkompetenz für alle Schüler*innen (Scheiter, 2021). Trotz dieses Auftrags zeigt sich, dass sozioökonomisch benachteiligte Schüler*innen geringere digitale Kompetenzen aufweisen (Eickelmann et al., 2019). Allerdings finden sich auch Schulen mit herausfordernder Schüler*innenschaft, deren Schüler*innen überdurchschnittliche digitale Kompetenzen aufweisen (sog. resiliente Schulen; Drossel et al., 2020).

Der Erwerb digitaler Kompetenzen ist entsprechend theoretischer Annahmen und empirischer Befunde u.a. abhängig von schulischen Merkmalen auf der Voraussetzungs- und Prozessebene. Bedeutsame Merkmale, die sich in theoretischer Hinsicht auf der Ebene der Voraussetzungen verorten lassen (Eickelmann et al., in Druck; Rožman et al., 2023) und die entsprechend empirischer Studien für digitale Kompetenzen von Relevanz sind, sind beispielsweise die Ausstattung der Schulen mit digitalen Medien, Einstellungen schulischer Akteur*innen gegenüber digitalen Medien und der pädagogische Support der Lehrkräfte für die Nutzung digitaler Medien im Unterricht (Gerick, Eickelmann & Bos, 2017). Als relevante Merkmale auf der Prozessebene gelten insbesondere die digitalisierungsbezogene Kooperation der Lehrkräfte, die Nutzung digitaler Medien im Unterricht sowie die digitalisierungsbezogene Professionalisierung schulischer Akteur*innen (Gerick, Eickelmann & Bos, 2017).

Unklar ist jedoch, inwiefern sich aktuell Unterschiede dieser Merkmale – insbesondere vor dem Hintergrund technologischer und gesellschaftlicher Entwicklungen – zwischen resilienten und nicht resilienten Schulen in Deutschland zeigen.

Fragestellung:

Ausgehend vom theoretischen Hintergrund und der empirischen Befundlage wird mit diesem Beitrag folgende Forschungsfrage bearbeitet:

1. Inwiefern unterscheiden sich aktuell schulische Voraussetzungs- und Prozessmerkmale des Erwerbs digitaler Kompetenzen an resilienten und nicht-resilienten Schulen in Deutschland?

Daten und Methoden:

Die Forschungsfrage wird auf Basis der Daten der International Computer and Information Literacy Study (ICILS 2023), die aktuelle Ergebnisse für Deutschland auf Basis repräsentativer computerbasierter Kompetenztests (N = 5065; Eickelmann et al., in Druck) und Befragungen der Lehrkräfte (N=2302) bereitstellt, beantwortet. Schulen werden dabei – angelehnt an Drossel et al. (2020) – dann als resilient bezeichnet, sofern sich der mittlere auf Schulebene aggregierte HISEI (Highest International Socio-Economic Index of Occupational Status) im unteren Drittel und die mittleren digitalen Kompetenzen der Schüler*innen im oberen Drittel der Stichprobe befinden. Um die Populationsschätzwerte der digitalen Kompetenzen zu ermitteln, wurde der Plausible-Values-Ansatz gewählt, um aus den individuellen Testergebnissen der Schüler*innen möglichst genaue Rückschlüsse auf die Merkmale der gesamten Population zu ziehen (Vennemann et al., 2021).

Mithilfe von deskriptiven Analysen und t-Tests werden Unterschiede von schulischen Merkmalen auf der Voraussetzungs- und Prozessebene zwischen resilienten und nicht-resilienten Schulen untersucht. Dabei werden u.a. einschlägige international im Kontext der ICILS-Studien validierte Skalen (Fraillon et al., 2020) zur Operationalisierung der unterschiedlichen Merkmale genutzt und die Daten zum Ausgleich von Stichprobenverzerrungen gewichtet.

Ergebnisse:

Hinweis für die Reviewer*innen: Zum Zeitpunkt der Deadline zur Einreichung von Abstracts für die Tagung der GEBF 2025 dürfen die Ergebnisse der Studie ICILS 2023, die im Rahmen des eingereichten ICILS-2023-Beitrags in diesem Symposium präsentiert werden sollen, noch nicht in den Abstracts ausgeführt werden, da sie einer internationalen Sperrfrist (12.11.2024 – 10 Uhr) unterliegen. Die Ergebnisse liegen den Autor*innen des Beitrags jedoch vor.

Die Ergebnisse werden vor dem Hintergrund des Forschungsstandes sowie hinsichtlich ihrer Relevanz für die empirische Bildungsforschung diskutiert.

Soziale Disparitäten in der ICT Literacy: Die Rolle verschiedener Indikatoren des familiären Hintergrunds und ICT-bezogener Nutzungsmotive. Ergebnisse aus einer Längsschnittstudie des NEPS

Martin Senkbeil

IPN - Leibniz Institute for Science and Mathematics Education

Theoretischer Hintergrund:

Kinder nutzen digitalen Technologien bereits in sehr jungem Alter und weitaus häufiger im häuslichen Umfeld als in der Schule (Godaert et al., 2022). Faktoren wie der sozioökonomische Status (SES) haben daher einen größeren Einfluss auf die ICT Literacy als schulische Faktoren (Aesaert et al., 2015). Entsprechend erweist sich der SES in der International Computer and Information Literacy Study (ICILS) als der stärkste Prädiktor für die ICT Literacy ($r = .31$ in Deutschland; Hatlevik et al., 2018). Neuere Studien weisen jedoch ebenso auf zahlreiche offene Fragestellungen hin. So zeigt eine Meta-Analyse (Scherer & Siddiq, 2019), dass die Stärke des Zusammenhangs zwischen SES und ICT Literacy durch die Art der verwendeten SES-Indikatoren moderiert wird (mit kulturellem Kapital als stärkerem Prädiktor im Vergleich zum elterlichen Bildungsabschluss und beruflichen Status) und wenig darüber bekannt ist, welche Merkmale diesen Zusammenhang erklären bzw. medieren. Zudem wurden bisher nahezu ausschließlich Querschnittsstudien durchgeführt, die aber ungeeignet sind, um Prädiktoren und Mediatoren für die Entwicklung der ICT Literacy zu identifizieren. Die Kenntnis über solche Prädiktoren und Mediatoren ist aber Voraussetzung für die Entwicklung von Interventionen zur Förderung der ICT Literacy (Hübner et al., 2023).

Fragestellung und Methode:

Zur Beantwortung dieser offenen Fragen werden längsschnittliche Daten der Startkohorte 4 des Nationalen Bildungspanels (NEPS) in den Klassen 9 (T1: $N = 14436$ Jugendliche) und 12 (T2: $N = 5545$ Jugendlichen) analysiert. Teilnahmeausfälle werden statistisch modelliert.

Das erste Ziel war, die Vorhersagekraft verschiedener SES-Indikatoren (kulturelles Kapital, höchster elterlicher Bildungsabschluss und beruflicher Status) auf die ICT Literacy zu bestimmen und sie im Hinblick auf vermutete Unterschiede miteinander zu vergleichen (Scherer & Siddiq, 2019). Die häusliche Anzahl der Bücher wurde als Indikator für das kulturelle Kapital verwendet. In Anlehnung an die Scholarly culture theory (SCT) ist Anzahl der Bücher als potenzielle Lernanregung und als Quelle kognitiver Kompetenzen zu verstehen (Sikora et al., 2019) und dementsprechend der höchste Zusammenhang zu erwarten.

Das zweite Ziel war festzustellen, inwieweit instrumentelle (auf informations- und bildungsbezogene Aktivitäten ausgerichtet) und sozial interaktive ICT-Nutzungsmotive (auf sozialen Austausch und Selbstdarstellung ausgerichtet) den Zusammenhang zwischen den SES-Indikatoren und der ICT Literacy vermitteln (Scheerder et al., 2020; Senkbeil, 2018). Als Kontrollvariablen wurden Intelligenz sowie andere etablierte Mediatoren des familiären Hintergrunds (Leseinteresse, familiäre Kommunikation; Kraaykamp & Notten, 2016; McElvany et al., 2009) berücksichtigt, um zu prüfen, ob die untersuchten Konstrukte (SES-Indikatoren, ICT-Nutzungsmotive) einen unabhängigen Beitrag zur Vorhersage von ICT Literacy leisten.

Ergebnisse:

Erwartungsgemäß erwies sich die Anzahl der Bücher längsschnittlich unter Einbezug der Kontrollvariablen als stärkerer Prädiktor für die ICT Literacy ($\beta = .09$; $p < .001$) als die anderen SES-Indikatoren (Bildungsabschluss: $\beta = .05$; $p < .001$; Beruflicher Status: $\beta = .03$; $p < .001$).

Darüber hinaus mediierten die ICT-bezogenen Nutzungsmotive (wiederum unter Einbezug der Kontrollvariablen) partiell den Zusammenhang zwischen SES und ICT Literacy, wenn die Anzahl der Bücher als SES-Indikator verwendet wird. Korrespondierend mit früheren Befunden (Senkbeil, 2018) lässt sich für die instrumentelle ICT-Motivation ein positiver indirekter Effekt ($\beta = .021$; $p < .001$) und für die sozial-interaktive Motivation ($\beta = .017$; $p < .001$) ein negativer indirekter Effekt ermitteln (d.h., instrumentelle Motive sind positiv mit Anzahl der Bücher und ICT Literacy korreliert, für sozial-interaktive Motive gilt das Gegenteil).

Die Ergebnisse stützen demnach zentrale Annahmen der SCT, wobei weiterführend vermutet wird, dass die Größe der häuslichen Bibliothek eng mit unterstützenden elterlichen Aktivitäten und Motivationen (z.B. eigene instrumentelle ICT-Nutzungsmotive, Unterstützung beim Kompetenzerwerb) verbunden ist, die im Laufe der Sozialisation habituell von den Kindern übernommen werden (Loh et al., 2023). Limitierend ist anzumerken, dass die Mediationseffekte für die ICT-Nutzungsmotive relativ gering ausfallen und der angedeutete Wirkmechanismus der SCT noch der empirischen Prüfung bedarf.

Wenn die Schule aus ist, lesen Jungen nicht. Geschlechtsspezifische und sozioökonomische Schichtung beim Lesen von Kindern – eine Analyse von digitalen Lesedaten.

David Reimer

University of Iceland, Island

Theoretischer Hintergrund:

In den meisten europäischen Ländern erzielen Mädchen bessere schulischer Leistungen als Jungen. Dies gilt insbesondere im Bereich der Lesekompetenzen: Der Vorsprung der Mädchen im Bereich Lesen ist deutlich größer als der Vorteil der Jungen im Bereich Mathematik (Logan & Johnston, 2010; Marks, 2008; Mullis et al., 2017; OECD, 2019). Die unterschiedliche Lesefähigkeit ist ein zentraler Faktor für die Unterrepräsentation der Jungen im Studium (Stoet & Geary, 2020). Eine mögliche Erklärung für diesen Befund ist die Feminisierung der Schulen. Schulen sind diesen Überlegungen zufolge stärker auf die Eigenschaften von Mädchen zugeschnitten (Neugebauer et al. 2011). Die meisten Studien deuten jedoch darauf hin, dass die besseren Leseleistungen der Mädchen darauf zurückzuführen sind, dass sie mehr Zeit mit Lesen verbringen und damit eine für den Schulerfolg zentrale Fähigkeit trainieren (Smith & Reimer 2023). Angesichts der vorliegenden Forschung ist es aber schwierig das Leseverhalten über Selbstberichte zu messen, da diese oftmals nicht reliabel und u.U. ein „reporting bias“ vorliegen kann.

Fragestellung:

Diese Studie ergänzt die Forschung, indem sie Geschlechter- und soziale Herkunftsunterschiede beim Lesen von Kindern untersucht und die Unterschiede danach aufschlüsselt, ob das Leseverhalten von den Kindern selbst oder von der Schule oder von den Eltern initiiert wird. Wir kombinieren zwei digitale Datenquellen aus dem Grundschulbereich in Dänemark, die es ermöglichen das Leseverhalten von Kindern wesentlich genauer zu verfolgen, als es bisher möglich war: Daten aus einer Lese-App („BookBites“), die in vielen dänischen Grundschulen genutzt wird und vollständige Daten über die Ausleihe von Büchern aus öffentlichen Bibliotheken durch Kinder und Eltern. Diese Daten ermöglichen es, sozioökonomische Faktoren, die das Leseverhalten beeinflussen von geschlechtsspezifischen Faktoren zu trennen und gleichzeitig in ihrer intersektionalen Wirkung zu untersuchen.

Daten und Methoden:

Wir analysieren Bibliotheksdaten über das Ausleihverhalten von Erwachsenen und Kindern. Wir unterscheiden zwischen Ausleihen von Kinderbüchern, die über das Konto der Kinder erfolgen, und Ausleihen, die über das Konto der Eltern erfolgen (Blaabeck 2022). Diese Daten werden mit Verhaltensdaten aus der Lesapp Bookbites kombiniert, die an vielen Grundschulen zur Verfügung gestellt wird. Die Schüler*innen können diese App auch zum Lesen in ihrer Freizeit verwenden. Sobald die Schüler*innen eine Seite lesen, generiert die Lese-App einen Dateneintrag. Wie bei digital trace Daten üblich, erfordert die Verwendung dieser Daten intensives Datenmanagement – führt aber zu validen Ergebnissen (Sortkær et al. 2022). Darüber hinaus führen wir diese Datensätze mit administrativen Registerdaten zusammen, um den Bildungsstand und das Einkommen der Eltern zu ermitteln. Wir messen das von der Schule initiierte Lesen mit einer in die Schule integrierten Lese-App. Das elterngesteuerte Lesen messen wir über die Ausleihe von Kinderbüchern durch die Eltern und das kindgesteuerte Lesen als die eigene Ausleihe von Kinderbüchern durch die Kinder. Da die Bookbites-App nicht in allen Schuljahrgängen genutzt wird, fokussieren wir unsere Analysen auf Schüler*innen in der 6. Klasse und deren Leseverhalten im Zeitverlauf.

Ergebnisse:

Unsere Analysen zeigen: (1) Im Allgemeinen sind Unterschiede im Ausleihverhalten und Leseverhalten nach Bildungsniveau der Eltern stärker ausgeprägt als die Geschlechterunterschiede. Während bildungsspezifische Gradienten die geschlechtsspezifischen Unterschiede bei der Ausleihe von Kinderbüchern durch die Eltern dominieren, sind die geschlechtsspezifischen und Bildungsherkunftsunterschiede bei der Ausleihe von Bibliotheksbüchern durch die Kinder von ähnlicher Größenordnung. (2) Die Unterschiede nach Geschlecht und Bildungsherkunft sind am größten, wenn die Schule geschlossen ist (entweder ferienbedingt oder aufgrund von Lockdowns). Dies deutet darauf hin, dass sowohl geschlechtsspezifische als auch sozioökonomische Lücken beim Lesen (und bei Lesefähigkeiten) davon abhängen, in welchem Ausmaß Eltern und Kinder in ihrer Freizeit lesen. (3) Schulen scheinen die geschlechtsspezifischen Unterschiede beim Lesen zu verringern. In der Schulzeit sind die geschlechtsspezifischen Unterschiede beim Ausleihen von E-Books und der mit Lesen verbrachten Zeit durchweg geringer, und oft schneiden Jungen sogar besser ab als Mädchen.

S-37: Modeling Learners Beyond the Mean: Erhebung und Statistische Modellierung intraindividuellder Dynamiken im Hochschulkontext

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 145

Chair der Sitzung: Peter Adriaan Edelsbrunner

Chair der Sitzung: Julia Dietrich

Modeling Learners Beyond the Mean: Erhebung und Statistische Modellierung intraindividuellder Dynamiken im Hochschulkontext

Chair(s): **Peter Adriaan Edelsbrunner** (ETH Zürich, Switzerland)

Diskutant*in(nen): **Friederike Bume** (DIPF Frankfurt)

Studien mit Methodologien der intensiven Datenerhebung wie Experience Sampling werden im Hochschulkontext vermehrt eingesetzt, um intraindividuelle Zusammenhänge und Dynamiken von Konstrukten wie Affekt, Emotion und Leistungsmotivation, deren zeitliche Entwicklung sowie Zusammenhänge mit externen Kriteriumsvariablen zu untersuchen (Ketonen et al., 2018; Moeller et al., 2022). Damit bieten intraindividuelle Modellierungen solcher Daten die wichtige Möglichkeit, Prozesse auf Individualebene zu untersuchen, auf die anhand interindividueller Information nicht empirisch geschlussfolgert werden kann (Molenaar, 2004). Es bestehen jedoch offene Fragen in Bezug auf die Erhebung und statistische Modellierung solcher intraindividuellen Fragestellungen, bei denen neben Mittelwerten (Location) auch Varianzen (Scale) und weitere Verteilungsparameter (Shape) stärker im Fokus sind als bei traditionellen Methoden (Umlauf et al., 2018). Das vorliegende Symposium beschäftigt sich mit vier zentralen, jedoch teils ungeklärten Fragen zur optimalen Erhebung und Modellierung intraindividuellder Zusammenhänge und Dynamiken bei Hochschulstudierenden. Durch die Verbindung von Fragen zu Erhebung und Modellierung im Kontext inhaltlicher Fragestellungen bietet das Symposium die Möglichkeit, ein umfassenderes Bild über typische Problematiken und Lösungsansätze in der Erforschung intraindividuellder Fragestellungen bei Studierenden zu erhalten.

Der erste Beitrag beschäftigt sich sowohl mit der Erhebung als auch der Modellierung. Er untersucht experimentell den Einfluss der Antwortskala (7-stufig Likert oder visuelle Analogskala) auf Verteilungseigenschaften und Zusammenhänge bei der Erhebung selbsteingeschätzten Affektes bei Studierenden. Die Verwendung Bayesianischer generalisierter Mehrebenen-Verteilungsmodelle erlaubt es, Einflüsse der Antwortskala nicht nur auf Mittelwerte, sondern auch auf Varianzen sowie Extremwertantworten zu modellieren. Der Beitrag zeigt, dass sich die zweitrangig erscheinende Entscheidung der Antwortskala auf intraindividuelle Verteilungseigenschaften und besonders auf geschätzte Zusammenhänge mit externen Kriteriumsvariablen (allgemeiner Psychopathologie) auswirken kann.

Der zweite Beitrag schließt mit einer location-scale Modellierung an, in der Mittelwert und Varianz des akademischen Selbstkonzepts im Zusammenhang mit dem Selbstmitgefühl und den Big Five untersucht werden. Die simultane location- und scale-Modellierung erlaubt, interindividuelle Unterschiede im Mittelwert und der intraindividuellen Variabilität gemeinsam durch Mittelwerte anderer Konstrukte zu erklären. Damit schlägt diese Studie die Brücke zum Fokus der dritten Studie und der Fragestellung «wie gehen intraindividuelle Dynamiken mit dem Ausmass interindividueller Trait-artiger Unterschiede einher?»

Der dritte Beitrag betrachtet aus Sicht der Theorie emergenter Prozesse die Modellierung von Übergängen von State- in Trait-Dynamiken. Inhaltlich verfolgt er die Fragen, ob und in welchem Ausmaß sich emotionale Prozesse bei Studierenden im Verlauf des zweiten Studiensemesters stabilisieren und welche statistischen Parameter dafür geeignet sein könnten. Die Modellierung von Autokorrelations- und Varianzausprägungen über die Zeit und rollierender Parameter weisen darauf hin, dass sich Veränderungen in Autokorrelationen und Varianzen zwischen Personen unterscheiden. Die Studie hebt damit hervor, dass theoretische Modelle und methodische Ansätze Vorhersagen über interindividuelle Variation in der Entwicklung von States und Traits miteinbeziehen sollten.

Der vierte Beitrag untersucht, inwieweit die nomothetischen, stichprobenzusammenfassenden intraindividuellen Ergebnisse aus Mehrebenenmodellen die intraindividuellen Muster in realen Individuen abbilden. Dazu werden idiografische Netzwerke intraindividuellder Interkorrelationsmuster zwischen Lernemotionen und Lernmotivationsvariablen zwischen Studierenden verglichen. Die Studie zeigt, dass sich Studierende hinsichtlich ihrer intraindividuellen Motivations- und Emotions-Korrelationsmuster stark unterscheiden und die Verteilungen dieser idiografischen Muster heterogen und in mehrerlei Hinsicht multimodal sind, sodass ein einziges Ergebnis aus einer Mehrebenenanalyse diese Heterogenität nur unzureichend beschreibt. Damit ergänzt diese Studie die univariaten Verteilungseigenschaften, die im Fokus der anderen Studien stehen, um wichtige Erkenntnisse der Modellierung intraindividuellder multivariater Zusammenhänge.

Insgesamt tragen die Studien dazu bei, dass wichtige, jedoch häufig unzureichend begründete oder unbeachtete Problematiken (Haslbeck et al., 2023) wie die Auswahl der Antwortskala und die Modellierung unterschiedlicher Verteilungseigenschaften 1) Einfluss auf Schlussfolgerungen über intraindividuelle Prozesse und Dynamiken haben, sowie 2) Quellen für Antworten auf Fragestellungen bieten, die in üblichen («location-»/Mittelwerts-) Modellen nicht untersucht werden können. Auf dieser Grundlage bietet das Symposium Diskussionsraum darüber, wie spezifische Eigenschaften von Erhebungs- und Modellierungsmethoden im Kontext von Intensivdatenerhebungen mit der Untersuchung spezifischer Fragen im Bildungskontext verbunden sind.

Beiträge des Symposiums

Experimenteller Vergleich von Likert Skala und visueller Analogskala im Ecological Momentary Assessment von Affekt bei Studierenden

Peter Adriaan Edelsbrunner¹, Jonas Haslbeck², Alberto Jover-Martínez², Eiko Fried³, Lotte Lemmens², Esmee Groot², Anne Roefs²

¹ETH Zürich, ²Maastricht University, ³Leiden University

Zwei häufig eingesetzte Antwortskalen bei der Erhebung von Affekt-Selbsteinschätzungen (z.B. «glücklich», «traurig») sind die siebenstufige Likertskala und die visuelle Analogskala (z.B. 0–100), wobei die Auswahl zwischen diesen Skalen üblicherweise unbegründet bleibt (Haslbeck et al., 2023). Aus der experimentellen Forschung in querschnittlichen Studien ist bekannt, dass Antwortskalen mit mehr als fünf Ausprägungen Reliabilität erhöhen können, jedoch sind weitere Befunde wie Auswirkungen auf Mittelwerte und weitere Verteilungseigenschaften uneindeutig (Simms et al., 2019). Zudem ist unklar, in welchem Ausmass sich Ergebnisse aus querschnittlichen Befragungen auf intraindividuelle Intensivdesigns generalisieren lassen (Molenaar, 2004).

Forschungsfrage

In der vorliegenden Studie wird untersucht, wie sich die Antwortskala auf uni- und multivariate Verteilungseigenschaften von Items innerhalb von Einzelpersonen sowie deren Dynamiken und Zusammenhänge mit externen Kriteriumsvariablen auswirkt.

Methode

Im Rahmen eines between-person experimentellen Designs füllten Studierende (Alter=21.73, SD=3.94, 100 weiblich) über Apps drei Mal pro Tag für zwei Wochen getriggert Selbstberichte zum erlebten momentanen Affekt auf 14 Items mit positiver (z.B. glücklich, Vorfreude, zufrieden) oder negativer (z.B. traurig, besorgt) Valenz entweder mittels siebenstufiger Likertskala (Likert-Bedingung; n = 73) oder per visueller Analogskala von 0-100 (VAS-Bedingung; n = 65) aus. Als externe Kriteriumsvariablen wurden der BSI (generelle Psychopathologie) sowie die DASS (depression anxiety stress scales) zu Beginn der Studie ausgefüllt.

Uni- und bivariate Verteilungseigenschaften wurden pro Person deskriptiv berechnet und dann mittels Bayesianischer Mehrebenenmodellierung mit fixierten Effekten für Bedingung und Itemvalenz, deren zufällige Effekte über Items bzw. Personen, deren Interaktion und zufälligen Intercepts für Items und Personen geschätzt. Fixierte Effekte wurden ausserdem auf sämtliche Verteilungsparameter (Varianzen, Null-Inflation) spezifiziert, wodurch die Modelle als location-scale-shape Regressionsmodelle implementiert waren. Konvergenz wurde anhand $R_{hat} < 1.02$ und Posterior-Traceplots sichergestellt, effektive Posterior-Ziehungen $> 1,000$ erreicht und Modellfit wurde anhand von posterior predictive checks geprüft. Für die Modellierungen und Interpretationen wurden beide Skalen auf 0-1 normalisiert. Im Folgenden wird Evidenz für das Vorliegen von Effekten der Antwortskala berichtet, wenn 90% Kreditintervalle für Unterschiede zwischen den beiden Bedingungen 0 ausschlossen. Für Zusammenhänge mit den externen Kriteriumsvariablen wurden mittels multivariaten Bayesianischen Modellen über Items gemittelte absolute Korrelationen von interindividuellen Unterschieden in den Mittelwerten der zweiwöchigen Selbstberichte mit den BSI und DASS-Skalen geschätzt.

Ergebnisse und Implikationen

In der VAS Bedingung zeigten sich höhere within-person Mittelwerte (VAS: 0.26, Likert: 0.20), Zwischenitemkorrelationen (VAS: 0.26, Likert: 0.23), lag-1 Autokorrelationen (VAS: 0.26, Likert: 0.20), niedrigere Schiefen (VAS: 0.87, Likert: 1.37) und weniger häufige Auswahl der niedrigsten Antwortkategorie. Der grösste Unterschied zeigte sich in den Zusammenhängen mit den externen Kriteriumsvariablen: Generelle Psychopathologie (BSI): VAS: $r=.54$, Likert: $r=.30$; Depression: VAS: $r=.52$; Likert: $r=.22$; Anxiety: VAS: $r=.34$, Likert: $r=.24$; Stress: VAS: $r=.48$; Likert: $r=.19$.

Diese Ergebnisse deuten darauf hin, dass die visuelle Analogskala geeigneter für Intensivdatenerhebungen von selbstberichtetem Affekt sein könnte, wenn Boden- oder Deckeneffekte erwartet werden oder wenn Information im Fokus ist, die mit Psychopathologie in Verbindung steht. Wir stellen verschiedene Interpretationsmöglichkeiten der Ergebnisse vor und wie dieses in zukünftiger diagnostischer Forschung untersucht sowie die vorliegenden Befunde repliziert und erweitert werden können.

Eine Experience-Sampling Studie zum akademischen Selbstkonzept von Studierenden: Der Einfluss von Selbstmitgefühl und den Big Five Persönlichkeitsfaktoren

Lena Kegel, Lucas Wloch, Carola Grunschel
Universität Münster

Theoretischer Hintergrund und Fragestellung

Das akademische Selbstkonzept (ASK) von Studierenden, definiert als die subjektive Einschätzung der eigenen akademischen Fähigkeiten, ist ein zentraler Prädiktor für ihren Studienerfolg (z.B. Baulke et al., 2021; Wach et al., 2016). Seit einigen Jahren wird der traditionelle Blick auf das ASK als Trait-Variable erweitert, indem das ASK auch bezüglich situativer Schwankungen als State-Variable untersucht wird. Experience-Sampling-Studien zeigten bereits, dass das ASK von Schüler*innen von Moment zu Moment variierte (Niepel et al., 2021; Tsai et al., 2008) und dass sich Lernende in dieser intraindividuellen Variabilität ihres momentanen ASK unterschieden (Hausen et al., 2022). Im Hinblick auf eine mögliche Variabilität des ASK von Studierenden stehen Ergebnisse noch aus.

Inter- und intraindividuelle Unterschiede im ASK könnten auf das Selbstmitgefühl von Individuen zurückzuführen sein. Personen mit hohem Selbstmitgefühl bewerten ihre Probleme und Fehler akkurat, aber begegnen sich dabei freundlich statt mit Härte (Neff, 2003). Das Selbstmitgefühl könnte positive und stabile selbstbezogene Bewertungen bedingen (Leary et al., 2009). Da die Abgrenzung zwischen dem Selbstmitgefühl und Neurotizismus intensiv diskutiert wird, wird empfohlen die prädiktive Rolle beider Konstrukte zu untersuchen (Pfattheicher et al., 2017). Im Schulkontext zeigte Neurotizismus einen negativen Zusammenhang mit der mittleren Ausprägung und der Stabilität des momentanen ASK (Hausen et al., 2022).

Mit diesem Beitrag sollte die Variabilität des ASK auf inter- und intraindividuelle Ebene untersucht und der Einfluss des Selbstmitgefühls sowie der Big Five Persönlichkeitseigenschaften geklärt werden. Das Selbstmitgefühl sollte positiv mit der mittleren Ausprägung des ASK (ASK-Level) und der Stabilität des ASK über die Zeit zusammenhängen (H1), während Neurotizismus negativ mit dem ASK-Level und der Stabilität des ASK assoziiert sein sollte (H2).

Methode

Die Datenerhebung fand mit Studierenden verschiedener deutscher Universitäten, Studiengänge (z.B. Psychologie, Jura, Lehramt) und Semester ($M(SD) = 3.10 (2.38)$) statt. In einer Online-Umfrage gaben sie Auskunft zu soziodemografischen Merkmalen, ihrem Selbstmitgefühl (Hupfeld & Ruffieux, 2011) und den Big Five Persönlichkeitsfaktoren (Rammstedt & John, 2005). In den vierzehn Tagen vor ihrer individuell ersten Prüfung des Wintersemesters 2024/2025 erfolgte drei Mal täglich eine ESM-Erhebung mittels der Software formr (Arslan et al., 2020), in der Teilnehmende mittels fünf Items ihr momentanes ASK einschätzten (Schöne et al., 2002).

Insgesamt wurden Angaben von $N = 262$ Studierenden (84% weiblich; $M(SD) = 25.29 (8.31)$ Jahre) mit $N = 7,422$ ESM-Teilnahmen analysiert, die an der Online-Umfrage und mindestens drei ESM-Messzeitpunkten teilnahmen. Als Analysen wurden Mixed-Effects Location-Scale Modelle mit z-standardisierten Prädiktoren durchgeführt (Dzibur et al., 2020; Hedeker et al., 2008).

Ergebnisse und Diskussion

Für das ASK lag 16% der Gesamtvarianz (inklusive Messfehler) innerhalb der Studierenden (Intraklassenkorrelation .84) und fiel damit geringer aus als in früherer (Schul-)Forschung (25%; Hausen et al., 2022; 34%; Witte et al., 2024). Zusätzlich unterschieden sich die Teilnehmenden signifikant in ihrer intraindividuellen Variabilität des ASK ($SD = 1.78, p < .001$).

Teilnehmende mit hohem Selbstmitgefühl berichteten ein höheres ASK-Level ($\beta = 0.37, p < .001$). Zudem berichteten Teilnehmende mit um eine Standardabweichung höherem Selbstmitgefühl eine um 32% geringere intraindividuelle Variabilität des ASK (H1). Das Selbstmitgefühl zeigte sich somit als Schutzfaktor für ein positives und stabiles ASK.

Studierende mit höherem Neurotizismus berichteten ein signifikant geringeres ASK-Level ($\beta = -0.16$, $p < .001$) und eine um 38% höhere intraindividuelle Variabilität (H2). Studierende mit höherer Gewissenhaftigkeit ($\beta = 0.17$, $p < .001$), höherer Extraversion ($\beta = 0.35$, $p < .001$), und höherer Offenheit für Erfahrungen ($\beta = 0.17$, $p < .001$) berichteten dagegen ein signifikant höheres ASK-Level. Zudem berichteten Studierende mit höherer Extraversion eine um 8% höhere und mit höherer Gewissenhaftigkeit eine um 30% geringere intraindividuelle Variabilität des ASK. Diese Ergebnisse unterstreichen die Bedeutung der Big Five für das ASK, deren Zusammenhänge mit der Variabilität des State-ASK weiter untersucht werden sollte.

Die intraindividuelle Dynamik von Emotionen: Untersuchung entwicklungsbedingter Veränderungen der emotionalen Variabilität

Anna-Maria Mayer¹, Miriam Francesca Jähne², Jessica Baars³, Julia Moeller³, Julia Dietrich¹

¹Katholische Universität Eichstätt-Ingolstadt, ²Friedrich-Schiller-Universität Jena, ³Universität Leipzig

Theoretischer Hintergrund und Fragestellung

Die Veränderung von Emotionen im Hochschulstudium ist zentral für den Lernerfolg von Hochschulstudierenden. Obwohl Emotionen grundlegend situativ erlebt werden, untersuchen wenige Studien die Situationsebene (Dietrich et al., 2022). Um diese Forschungslücke zu adressieren, fokussiert das auf Dynamischen Systemtheorien basierende DYNAMICS-Modell (Moeller et al., 2022) die kurzfristige Entwicklung von einer Lernsituation zur nächsten auf Ebene emotionaler State-Prozesse und deren Zusammenhänge mit langfristig stabileren emotionalen Dispositionen.

State-Prozesse beschreiben einerseits relativ stabile Muster wiederkehrender Erlebenszustände (z.B. häufige Langeweile in der Vorlesung) oder Kontingenzen (z.B. häufige Langeweile infolge geringen Werterlebens). State-Prozesse können sich andererseits hinsichtlich ihrer Häufigkeit oder rekursiven Zusammenhänge verändern. Beispielsweise kann die Freude in einer Lernsituation zu stärkerer Anstrengung mit einer intensiveren Beschäftigung mit dem Lernstoff einhergehen (Pekrun & Linnenbrink-Garcia, 2014), dadurch einen Lernerfolg auslösen, woraufhin verstärkte Freude in der nächsten Lernveranstaltung folgt. Die stabiler werdende Freude am Lernen wäre Ausdruck eines emergenten Prozesses (Moeller et al., 2022).

Operational werden emergente Prozesse als solche definiert, deren Parameter sich über die Zeit verändern (sog. rollierende Parameter). Die vorliegende Studie untersucht, inwieweit sich emotionale States im Laufe eines Semesters zunehmend stabilisieren. Sie trägt damit zur Identifikation emotionaler Verläufe und Mikroprozesse, die sich zu emotionalen Dispositionen verfestigen, bei. Zudem diskutieren wir zwei statistische Stabilitätsparameter und welche Rückschlüsse diese auf die angenommene Stabilitätszunahme zulassen.

Methode

Mittels Experience-Sampling-Methode wurden Daten von N=7 Studierenden im 2. Fachsemester ihres Psychologiestudiums (n=6 Bachelor, n=1 Master, 100% weiblich, MAlter=20) erhoben. Die Studierenden wurden über den Verlauf eines Semesters hinweg (~11 Wochen, 46-285 MZP/Person) via Smartphone-App in ihren wöchentlich stattfindenden Lehrveranstaltungen zu ihren emotionalen States befragt (vgl. Moeller et al., 2018; Pekrun, 2006). Als positiv aktivierende Emotion wurde Begeisterung, als negativ aktivierende Emotion Frustration und als negativ deaktivierende Emotion Langeweile gemessen. Die Instruktion, die den Items vorangestellt wurde, lautete: "Ich fühle mich ...". Die Antwortskala für alle Items war: 0 = überhaupt nicht, 1 = kaum, 2 = etwas, 3 = ziemlich, 4 = sehr stark.

Auf Basis von Zeitreihenanalysen über Einzelpersonen und zwei Stabilitätsmaßen (Dejonckheere et al., 2019) überprüfen wir folgende Annahmen: 1) Zunehmende Stabilität drückt sich darin aus, dass die Variabilität emotionaler States kleiner wird. Die Varianz in den Zeitreihen sollte sich damit über die Zeit verringern. 2) Zunehmende Stabilität drückt sich darin aus, dass Vorhersagestärke (Autokorrelation) von einem Messzeitpunkt zum nächsten über die Zeit größer wird. Die sehr intensive Erhebung erlaubt eine detaillierte Modellierung emotionaler States und deren Stabilisierung.

Ergebnisse und Diskussion

Erste Ergebnisse auf Basis der Standardabweichung zeigen eine Stabilisierung von Begeisterung, Frustration und Langeweile bei jeweils zwei Personen. Bei jeweils ein bis zwei Personen zeigt sich eine Zunahme an Variabilität der betreffenden Emotion über die Zeit. Auf Basis der Autoregression zeigt sich eine Stabilisierung aller drei Emotionsfacetten bei jeweils einer Person. Eine Zunahme an Variabilität zeigt sich bei jeweils ein bis zwei Personen. Weitere Ergebnisse werden es aufgrund der Länge der erfassten Zeitreihen erlauben, die Dynamiken und Entwicklungsprozesse von alltäglichen emotionalen Erlebnissen stärker in die Tiefe zu verstehen, als es in bisherigen Feldstudien möglich war.

In unserem Vortrag werden wir diskutieren, welche Rückschlüsse die Modellierung emotionaler Verläufe auf die Identifikation von Stabilisierungsprozessen emotionaler Dispositionen ermöglicht. Zudem stellen wir zwei mögliche Parameter zur Erfassung intraindividuelle Entwicklung von emergenten Prozessen vor.

Within-Person Associations Among Academic Emotions and Situated Expectancy-Value Facets: Deriving Generalizable Insights from Idiographic Networks

Jessica Baars¹, Michael Grimmer¹, Thu Thuy Nguyen¹, Miriam Francesca Jähne², Anna-Maria Mayer³, Julia Dietrich³, Jacquelynne S. Eccles⁴, Alexander Kreiß¹, Reinhard Pekrun⁵, ManyMoments Consortium¹, Julia Moeller¹

¹Universität Leipzig, ²Friedrich-Schiller-Universität Jena, ³Katholische Universität Eichstätt-Ingolstadt, ⁴University of California, Irvine, ⁵University of Essex

Research on learning-relevant motivation and emotions increasingly uses situated measures, requiring intensive longitudinal methods and methods distinguishing within-person and between-person variance. The DYNAMICS framework proposed by Moeller et al. (2022) calls for an integration of measures capturing facets of the situated expectancy-value Theory (SEVT; Eccles & Wigfield, 2020) with measures of academic emotions (see e.g., Pekrun, 2006). The DYNAMICS framework hypothesizes that students may differ in their within-person covariance patterns, and that sample-level summary statistics, such as nomothetic summaries of within-person inter-correlation matrices, should be derived in a bottom-up way by systematically comparing these within-person correlations across individuals and by systematically examining the between-person distributions of within-person correlations. Most previous studies rely on the untested assumption that within-person patterns (e.g., within-person correlations) be homogeneous across individuals, typically describing a sample of individuals with one overall within-person effect. However, Haslbeck et al. (2023) found that many emotions measures with intensive longitudinal methods are bi- or multimodally distributed and it is unclear whether and how often bi- or multimodal between-person distributions can be expected for within-person correlation effects. If there were bi- or multimodal between-person distributions of within-person correlations, then the sample-level average or sample-level maximum likelihood estimation of that within-person effect may be biased and a misleading estimate of the sample's measure of the central tendency (Derrible & Ahmad, 2015; Wirtz & Nachtigall, 1998).

Research Questions

We examine between-person heterogeneity in within-person associations and propose a bottom-up approach to generalizability, by testing how often heterogeneous (e.g., bi- or multimodal) distributions were observed across different items, individuals, and levels of analysis.

Research question (RQ) 1a examines the nomothetic, sample-level estimate of the within-person inter-correlation patterns (Level 1 estimates) that results from the frequently used multilevel models with repeated measures (Level 1) nested in individuals (Level 2). RQ1b asks how the idiographic within-person inter-correlation patterns look for each student. RQ1c asks whether it is plausible to assume that all idiographic within-person correlation patterns originate from a homogeneous population.

RQ2 addresses the homogeneity assumptions underlying the frequently used summary statistics (such as the multilevel models mentioned above), asking whether and how often each item is distributed unimodally within individuals (RQ2a) and between individuals (RQ2b). RQ2c asks whether and how often the within-person correlations were unimodally distributed between individuals.

Methods

1,028 experience sampling method surveys from 37 students were collected over one semester in one university lecture. Twenty-one items measured motivational components specified in the SEVT and academic emotions.

Multilevel analyses estimated an overall within-person and an overall between-person network of zero-order correlations among motivation and emotion facets (RQ1a). Idiographic networks depicted the individual within-person correlations among the measures for each of the 37 students in the sample (RQ1b). Multi-sample test for high-dimensional covariance matrices (Zhang et al., 2018) were used to determine whether it was plausible that the 37 idiographic covariance matrices originated from the same population (RQ1c).

Unimodality tests and distribution plots were examined for the within-person distribution of each item (RQ2a), the between-person distribution of each item (RQ2b) and the between-person distribution of each within-person correlation (RQ2c).

Results and their implications

Idiographic within-person networks of the 37 participants revealed substantial heterogeneity in their within-person inter-correlation patterns. Heterogeneity was further indicated by one third of the within-person correlations being bi- or multimodally distributed across individuals. We conclude that heterogeneity in within-person patterns must be examined, otherwise, methods relying on homogeneity assumptions, including multilevel analyses, may fail to describe many if not all individuals' idiographic correlation patterns. This is crucial for the trustworthiness of measurement and structure models, the generalizability of ESM research, and any research aiming for personalized diagnostics and interventions.

S-38: From Data to Performance: Understanding and Enhancing Student Learning Behaviors

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 142

Chair der Sitzung: Jakob Schwerter

Chair der Sitzung: Carolin Baumann

Chair der Sitzung: Markus Dresel

English Track

From Data to Performance: Understanding and Enhancing Student Learning Behaviors

Chair(s): **Jakob Schwerter** (TU Dortmund University, Germany), **Carolin Baumann** (Karlsruhe University of Education, Germany)

Diskutant*in(nen): **Markus Dresel** (University of Augsburg)

Student performance is critical to academic success at various stages of education, yet many students struggle to perform well. Learning analytics, which use data on learning behaviors, can provide critical insight into underperformance, highlighting misapplied strategies and indicating a lack of engagement or motivation. Research shows that differences in learning behaviors, as measured in learning management systems, correlate with academic performance (Bernacki et al., 2020; Deininger et al., 2024; Schwerter et al., 2022). However, the analysis of such data is often system-specific, and its generalizability across contexts is not well understood (Nkomo et al., 2021). Therefore, more evidence is needed to know which features of digital learning environments optimally support student learning. Finally, a better understanding of how to interpret the measures generated is important for working with learning behavior data (Du et al., 2023; von Keyserlingk et al., 2023). Therefore, this symposium will shed light on these issues by presenting cutting-edge research on learning behavior in different educational contexts.

Studies 1 and 3 explore predictive analytics in different contexts, while **Study 2** uses an intervention and **Study 4** extends motivational theory based on ecological momentary assessments. More specifically, **Study 1** (three computer science courses at two universities) focusses on the generalizability of learning behavior data for predicting exam performance under different incentive structures, the minimum data required for a robust estimation, and how these structures influence the importance of variables in the predicting performance. **Study 2** (second semester teacher students from a pedagogical university) examines how instructional design (video formats) influences learning behavior and performance. **Study 3** (seventh-grade English tutoring systems) uses performance prediction results to cluster the learning behavior data and explores which student characteristics (e.g., prior knowledge, motivation, personality) explain these clusters. **Study 4** (applied statistics course for psychology university students), complements the other studies by emphasizing the role of situated motivation on students' learning behaviors.

Together, these studies provide a comprehensive view of how different aspects of learning behaviors can be understood and used to predict and enhance student performance across various educational settings. The symposium features findings for the design of learning environments and incentives, and strategies to help students learn more effectively across different educational contexts. The contributions will be discussed by an expert in the field of self-regulated learning.

Beiträge des Symposiums

Generalizability of Digital Learning Behavior Predictions Across University Courses: An Empirical Study

Jakob Schwerter¹, Judith Bose¹, Marko Schmellenkamp², Thomas Zeume², Philipp Doebler¹

¹TU Dortmund University, Germany, ²RUB Bochum, Germany

Theoretical background:

STEM dropout rates at German universities are alarmingly high, with computer science experiencing a 44% dropout rate (Heublein et al., 2022). Many students drop out due to performance difficulties, especially in theory-heavy courses like formal logic (Soll et al., 2023). Digital learning environments have emerged as a tool for supporting students' learning processes (Geck et al., 2019), and digital learning behavior predicts student performance (Deininger et al., 2023; K. Schneider et al., 2019). In addition to survey data, data on digital learning behavior is an important additional data source and crucial for understanding students' learning processes. However, its generalizability beyond the specific course and university context has not yet been sufficiently researched. It also needs to be clarified whether data that is only collected at the beginning of the semester is sufficient for predictions in order to be able to intervene in time (cf. Coglianò et al., 2022).

We examine whether data from a different university course "in the wild" can be used for out-of-sample prediction to verify generalizability of prediction models beyond the original context. Additionally, we examine which learning behavior variables are key in the prediction. Lastly, we progressively omit data from the end of the semester to the beginning of the semester to determine how early in the semester we can predict student success.

Method:

We use data from three computer science courses (N1 = 137, N2 = 104, N3 = 148). Two courses (theoretical computer sciences and logic) were taught at university A, and the third course (logic) at university B. Both universities are of the same metropolitan region. All courses used the same digital learning environment from which we collected the data to predict students' exam performance. Within the courses, students could participate in incentivized and non-incentivized digital practice and incentivized paper-based practice. The main difference between the two universities is that at university A students obtained extra points for the exam, while at university B students need to obtain enough points to register for the exam. From the digital learning environment data, we generated various variables that capture student participation (Förster et al., 2018), performance (Schwerter et al., 2022) and engagement (Ifenthaler et al., 2022; M. Schneider & Preckel, 2017), which can be related to self-regulated learning. To predict student performance, we defined students' exam success if students achieved a grade of 3.7 or better. Student survey responses on motivation and learning strategies are not yet included in the analysis.

Results:

Students complete almost all incentivized assignments across all courses, but the number of attempts decreases overtime. Both incentivized assignments (digital and paper-based) have similar completion rates across courses. At university A, over 70% of the students submitted at least one unincentivized assignment per week, compared to only 20% at university B.

Though Random Forest is best at predicting exam success, Elastic Net performs better than Random Forest and XGBoost in the out-of-sample predictions (XGBoost: -0.40 accuracy; Random Forest: -0.30 accuracy, Elastic Net: -0.07 accuracy). In some cases, the out-of-sample prediction shows improved prediction metrics over a prediction model trained on the original sample.

Typically, 4 to 5 weeks of data are sufficient to achieve 95% of accuracy. Key predictors include variables related to incentivized digital and paper-based practice that reflect students' average engagement or performance during the semester.

The study shows that student success can be predicted across courses and universities, suggesting that interventions do not necessarily require baseline measurements per course if the course design is similar. Differences between universities are greater than between courses in the same discipline. The accuracy of early semester predictions allows for timely, data-driven interventions.

Enhancing Learning Through Segmented Instructional Videos and Active Retrieval: Examining Indirect Effects on Navigation and Task Performance

Carolyn Baumann, Samuel Merk

Karlsruhe University of Education, Germany

Theoretical background:

Active retrieval has been established as an effective learning method through numerous meta-analyses (Adesope et al., 2017; Rowland, 2014; Yang et al., 2021). While its direct benefits, such as improved long-term retention and deeper cognitive processing, are well-documented, the indirect effects, like reducing overestimation by making learners aware of their knowledge gaps, are harder to observe and measure (Endres & Renkl, 2022).

Indirect effects refer to the motivational and metacognitive processes of learning that can be elicited by active retrieval and are particularly important for guiding subsequent learning efforts. The process of retrieval can e.g. adjust learners' judgments of their own learning, making them more accurate in predicting their future performance and recognizing areas needing further study (Roediger III et al., 2011).

For such additional learning efforts to be initiated, it is essential that self-regulation is functioning effectively and sufficient motivation is present (Schunk & Zimmerman, 2011). In the present study, we tested whether reducing navigational efforts within the learning materials helps foster consecutive learning after active retrieval attempts. We did this by comparing segmented instructional videos, interspersed with active retrieval tasks, to a single instructional video (all segments combined) followed by a large task block. The following two research questions were addressed:

(1) Do segmented videos lead to more indirect effects, specifically backsteps from active retrieval tasks to the instructional video, compared to a combined (long) instructional video?

(2) Is the initial success rate for active retrieval tasks higher when they are offered between shorter video segments rather than at the end of a longer video?

Methods:

In a cross-classified experimental study, 300 university students were assigned to receive either segmented or complete instructional videos across five topics in quantitative research methods. Specifically, the format of the videos (segmented or complete) varied between topics/weeks and hence within students. Log data from the learning management system were analyzed to examine usage patterns, navigation behaviors, and results from the associated active retrieval tasks. Backsteps in navigation were defined as clicks on previous videos segments/the complete video after working on corresponding active retrieval tasks. The analysis was conducted using Bayesian mixed models with the probabilistic programming language Stan, via the R package "brms" (Bürkner, 2018).

Results:

A Bayesian negative-binomial mixed model, was fitted to analyze the effects of the treatment video condition (segmented vs. complete video) on the number of backsteps, incorporating random intercepts for both topics and participants. The model demonstrated substantial explanatory power ($R^2 = .71$, 95% CI [.64, .76]), with a moderate marginal R^2 of .15 (95% CI [.09, .22]) for the fixed effects only. The conditional effects analysis showed a large effect: participants in the segmented video condition made, on average, 5.58 more backsteps per topic, with a 95% credibility interval of [4.28, 7.13], compared to the complete video condition.

Our second research question was addressed using a zero-one inflated beta regression model, with the probability of correctly solving the active retrieval task on the first attempt as dependent variable, and the independent variable being the treatment video condition (segmented vs. complete). Random effects for topics and participants were also included. The explanatory power of the model was ($R^2 = .19$, 95% CI [.16, .22]). A closer look at the conditional effects reveals (still) inconclusive results.

The findings indicate that segmented videos promote greater navigation and may enhance engagement with the material, which indicates evidence for the assumption of indirect effect of active retrieval. However, the impact on the probability of achieving initial success in active retrieval tasks is complex and warrants further examination until the GEBF conference in January.

Who Did What to Succeed? Individual Differences in Which Learning Behaviors are Relevant for Achievement

Hannah Deininger¹, Cora Parrisius², Rosa Lavelle-Hill³, Detmar Meurers⁴, Gjergji Kasneci⁵, Ulrich Trautwein¹, Benjamin Nagengast¹

¹University of Tübingen, Germany, ²Karlsruhe University of Education, Germany, ³University of Copenhagen, Denmark, ⁴Leibniz Institute für Wissensmedien (IWM), Germany, ⁵Technical University of Munich, Munich, Germany

Theoretical Background:

Previous research has linked learning behaviors to achievement (Authors, 2023; Bernacki et al., 2020), but inconsistencies remain whether specific learning behaviors, such as spending more time on learning activities, are positively or negatively related to achievement (Hendriks et al., 2014; Hui et al., 2023). While the amount of practice is not that important for learners with high prior knowledge, it may be crucial if prior knowledge is low. Similarly, both achievement motivation and personality traits are known to moderate the link between learning behaviors and achievement (Matcha et al., 2020; Richardson et al., 2012; Schneider & Preckel, 2017). What specific learning behaviors (e.g., time spent on learning activities or the number of tasks submitted) are relevant for a student's achievement might, therefore, differ depending on their motivation, personality, and prior knowledge levels.

In a two-step procedure, we thus aimed to (1) cluster students based on the relevance of their learning behaviors within an intelligent tutoring system (ITS) for English language learning for their later English performance and (2) investigate if students in these clusters differ regarding their motivation, personality, and prior knowledge.

Methods:

We analyzed data from a cluster-randomized controlled field trial investigating language learning through individualized practicing in an ITS for English as a foreign language. A total of 507 seventh-grade students (55.8% female, Mage = 12.50 years, SDage = 0.41) used the ITS over three weeks in the school year 2021/2022. Students completed a prior knowledge test and self-reports before and a post-test after this phase. During their engagement with the ITS, behavioral trace data were collected. Several machine learning models (i.e., linear regression, elastic net regression, support vector regression, decision tree, random forest, and extreme gradient boosting (XGBoost)) were trained to predict performance based on learning behavior indicators. For the best-performing model, SHAP values (Lundberg & Lee, 2017) were generated to quantify the relevance of different variables for the prediction in the overall sample (i.e., the impact of students' learning behaviors on performance). Based on these SHAP values, we clustered students using k-means (RQ1). Silhouette score analysis determined the optimal number of clusters. We employed a MANOVA to examine differences in personality traits (here: conscientiousness and openness), motivation (here: self-concept, intrinsic value, attainment value, utility value, and cost), and prior knowledge among clusters (RQ2).

Results:

Among all trained models, XGBoost showed superior performance by explaining 41% of the variance in the post-test score. K-means clustering of the generated SHAP values for this model revealed three student clusters (NC1 = 38, NC2 = 48, NC3 = 16). Figure 1 presents the top ten relevant variables for students in each cluster in predicting their achievement. For students in Cluster 1, it was most important whether or not they performed corrections after feedback and the amount of time spent on feedback; for students in Cluster 2, the ratio of directly correct attempts and time on task was decisive for their achievement; and for students in Cluster 3, the most critical learning behaviors for positive learning outcomes were time spent on feedback and time on task.

In Clusters 1 and 2, students differed only regarding their prior knowledge level (see Figure 2), with high prior knowledge in Cluster 1 and low prior knowledge in Cluster 2. Students in both clusters showed average motivation and personality levels. However, students in Cluster 3 had less prior knowledge and reported lower self-concept and utility value.

In summary, we identified clusters of students who differed in the learning behaviors relevant to their later performance. By exploring the interplay between individual differences, learning behavior, and achievement, we gained valuable insights into individual learning processes.

Situated Achievement Motivation and E-Learning Behavior: Theory-Integration with Ecological Momentary Assessments and Objective Logfile Data

Marc Philipp Janson¹, Theresa Wenker¹, Lisa Bäumke²

¹University of Mannheim, Germany, ²University of Tübingen, Germany

Theoretical background:

Recent research on achievement motivation highlights the importance of situational factors affecting learners' motivation and behavior (Moeller et al., 2022; Murayama et al., 2017). For example, the situated expectancy-value theory (SEVT; Eccles & Wigfield, 2020, 2024) proposes that learners' beliefs about their ability to succeed (expectancy) and the value of the task (value) influence their motivation, but those beliefs can change from moment-to-moment.

One important situational factor is time and for learners especially time until the exam (Capelle et al., 2022). The temporal motivation theory (TMT, Steel, 2007) provides a formalized approach to predict achievement motivation based on SEVT as well as the time till deadline (delay) and individuals' sensitivity to delay (procrastination tendency):

$$\text{Achievement Motivation}_t = (\text{Expectancy} * \text{Value}) / (\text{Sensitivity to delay} * \text{Delay}_t) \text{ (F1)}$$

However, this traditional TMT does not fully incorporate within-person variability (Dishop, 2020). We propose an extension, the Situated Temporal Motivation Theory (STMT), which accounts for moment-to-moment changes in motivation influenced by variant motivation beliefs as well as proximity of deadlines:

$$\text{Achievement Motivation}_t = (\text{Expectancy}_t * \text{Value}_t) / (\text{Sensitivity-to-delay}_t * \text{Delay}_t) \text{ (F2)}$$

In our preregistered study (https://aspredicted.org/R26_WMG) we investigated whether the STMT predicts achievement motivation and behavior. Higher STMT compound scores should reflect higher planned learning time (H1a) and higher actual learning time (H1b). Furthermore, to explore the factors driving the STMT we predicted that the single components would also predict planned (H2a) and actual learning time (H2b). On an exploratory note, we wanted to investigate whether the STMT outperforms the traditional TMT approach (F2 vs. F1).

Methods:

We conducted a field study involving 1005 observed learning days of N = 98 first-semester psychology students preparing for a statistics exam using an intelligent tutoring system. Data were collected through experience sampling and logfiles. Students' expectancy, value beliefs, sensitivity to delay, and planned learning time were assessed daily at first login. Actual learning time was tracked by the tutoring system. We used multilevel regression analyses (L1: day, L2: person) to predict daily (planned) learning time.

Results:

The STMT showed significant predictive power for both planned (H1a) and actual learning time (H1b). The model explained a substantial amount of variance in actual learning time (marginal $R^2 = .44$) but less so for planned learning time (marginal $R^2 = .03$). Out of the components, only delay (negatively) and value beliefs predicted both planned (H2a) and actual learning time (H2b) significantly.

Compared to the explained variance for planned learning time (marginal $R^2 = .08$) and actual learning time (marginal $R^2 = .12$) using the traditional TMT approach, we observed a slightly smaller amount of explained variance for planned learning time (marginal $R^2 = -.04$) and large amount of additionally explained variance (marginal $R^2 = .31$) for actual learning time with our STMT approach.

The STMT provides a nuanced understanding of achievement motivation, highlighting the importance of temporal factors (Capelle et al., 2022) and within-person variability (Eccles & Wigfield, 2020; Murayama et al., 2017). It significantly improves the prediction of actual learning behaviors under high-stakes conditions, demonstrating the dynamic nature of achievement motivation.

However, the unexpectedly lower predictive power for planned learning time suggests that further research is needed to explore the factors influencing students' intentions.

This study underlines the tremendous opportunities for theory testing with advanced methods from learning analytics and experience sampling methods. The proposed theory was able to predict achievement-motivation-related learning behavior with a parsimonious theory-based formula.

Promising further avenues of research are the incorporation of costs as negative value components (Flake et al., 2015), which were not considered yet. Moreover, the integration of achievement-related emotions (Pekrun, 2006) might also leverage our understanding of achievement motivation in high-stakes situations.

S-39: Forschungssynthesen und Digitalisierung: Empirische Befunde und methodische Ansätze

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 133

Chair der Sitzung: Lisa Ziernwald

Chair der Sitzung: Marcel Capparozza

Chair der Sitzung: Frank Fischer

Forschungssynthesen und Digitalisierung: Empirische Befunde und methodische Ansätze

Chair(s): **Lisa Ziernwald** (Technische Universität München, Deutschland), **Marcel Capparozza** (Eberhard Karls Universität Tübingen)

Diskutant*in(nen): **Frank Fischer** (Ludwig-Maximilians-Universität München)

Die Digitalisierung führt in der Bildungspraxis und der Bildungsforschung zu Veränderungen und neuen Entwicklungen. So haben digitale Medien das Potenzial, das Lernen im Unterricht positiv zu beeinflussen, für mehr Bildungsgerechtigkeit zu sorgen und die Art und Weise, wie Fortbildungen für Lehrkräfte gestaltet werden, zu verändern (Timotheu et al., 2023). Auch für die Bildungsforschung ergeben sich durch die Digitalisierung, zum Beispiel durch die Nutzung von künstlicher Intelligenz, neue Möglichkeiten (Campos et al., 2024). Dabei nehmen moderne Large Language Models durch ihre Fähigkeiten des Sprachverstehens und der Sprachgenerierung (Naveed, 2023) eine besondere Rolle ein. Im vorliegenden interdisziplinären Symposium werden beide Perspektiven aufgegriffen.

Um evidenzbasierte Praktiken (zum Beispiel in der Professionalisierung von Lehrkräften; Sims et al., 2023) zu unterstützen, bietet insbesondere die Methode der Forschungssynthese Vorteile gegenüber Primärstudien. So gibt diese einen umfassenden Überblick über ein Forschungsfeld, bestehend aus zahlreichen unterschiedlichen Einzelstudien (Woerfel et al., 2021). Des Weiteren können durch den Zugang über Forschungssynthesen zumindest teilweise Einschränkungen von Primärstudien (z. B. mangelnde Generalisierbarkeit) überwunden, heterogene Studienergebnisse verglichen und somit ein Beitrag zur Theoriebildung geleistet werden (Beelmann, 2014).

Gleichzeitig ist eine häufige Schwierigkeit bei der Erstellung von Forschungssynthesen der langwierige Erstellungsprozess (Woerfel et al., 2021), welcher insbesondere durch das Screening von meist mehreren Hundert bis Tausend Treffern aus Literaturdatenbanken verursacht wird. Um diesem Problem zu entgegnen, spielt die Nutzung neuer Möglichkeiten der Digitalisierung insbesondere im Feld der Forschungssynthesen eine zentrale Rolle. So wurden in den letzten Jahren Algorithmen entwickelt, welche bei der Beurteilung der Literatur unterstützen und damit die Effizienz und Genauigkeit von Forschungssynthesen verbessern sollen (Chernikova, 2024).

Die vier Beiträge in diesem Symposium zeigen verschiedene interdisziplinäre Perspektiven auf das Zusammenspiel von Digitalisierung und Forschungssynthesen sowohl aus einer inhaltlichen als auch einer methodischen Perspektive auf. Dabei liegt ein Fokus auf der Beschreibung der Veränderung der Fortbildungskultur von Lehrkräften durch die digitale Transformation in der Bildungspraxis. Der erste Beitrag des Symposiums befasst sich mit den potenziell unterschiedlichen Auswirkungen von Online- und Face-to-Face-Kooperationsaktivitäten in der Lehrkräftefortbildung. Mit einem Systematic Review wurden die Operationalisierung und theoretischen Grundlagen von Kollaboration von Lehrkräften in Fortbildungen sowie die Auswirkung auf Lehrkräfte thematisiert. Auch im zweiten Beitrag wird mit einem Systematic Review an die Veränderung der Lern- und Arbeitskultur von Lehrkräften angeknüpft, indem der Frage nachgegangen wird, welche Faktoren die Nutzung von offenen Bildungsmaterialien (Open Educational Resources; OER) durch Lehrkräfte positiv bzw. negativ beeinflussen. Die Nutzung digitaler Medien im Schulkontext und die dafür benötigten Kompetenzen der Lehrkräfte werden im dritten Beitrag im Rahmen eines Rapid Reviews aufgegriffen. Es wird die Frage beantwortet, mit welchen Testinstrumenten die Erfassung von digitalen Kompetenzen von Lehramtsstudierenden und Lehrkräften gelingen kann, ohne auf Selbstberichtsdaten zurückzugreifen. Dazu wird unter Einbezug relevanter Stakeholder aus Bildungspraxis und -forschung eine narrative Synthese verfügbarerer Erhebungsinstrumente erstellt, um künftig Forschende und Fortbildende bei der Auswahl von Instrumenten beispielsweise zur Evaluation des Fortbildungserfolgs zu unterstützen. Der vierte Beitrag beschäftigt sich abrundend mit der Digitalisierung aus Perspektive der Bildungsforschung, indem untersucht wird, inwieweit Large Language Models (LLMs) die Erstellung von Forschungssynthesen und hier insbesondere den Screening-Schritt unterstützen können. Durch diese digitale Unterstützung soll die Effizienz und Genauigkeit der Durchführung von Forschungssynthesen verbessert werden, bei gleichzeitigem Erhalt von Nachvollziehbarkeit und Transparenz des Prozesses. Dies ist insbesondere im Kontext der schnelllebigen Veränderungen der Forschung im Kontext der Digitalisierung notwendig.

Zusammengefasst verfolgt das Symposium das Ziel, anhand der gemeinsamen Methode der Forschungssynthese verschiedene evidenzbasierte Perspektiven auf Lehrkräfte in der digitalen Transformation aufzuzeigen und dabei durch eine interdisziplinäre Herangehensweise sowohl die Bildungspraxis als auch die Bildungsforschung im Kontext der Digitalisierung voranzubringen. Im Symposium wird das Potential von Forschungssynthesen für Bildungspraxis und -forschung im Kontext der Digitalisierung deutlich gemacht und abschließend über alle vier Beiträge hinweg diskutiert.

Beiträge des Symposiums

Teacher Collaboration in Professional Development Interventions – A Systematic Review

Georgios Mitsostergios, Lisa Ziernwald, Doris Holzberger

Technische Universität München

Based on the latest TALIS survey in 2018 (OECD, 2019), many challenges, with the increased digitalization and use of ICT tools for classwork included, require teachers to constantly update their skills and practices with Continuous Professional Development (CPD). The same survey confirms that teachers who participate in CPD display high levels of self-efficacy and job satisfaction, with teachers rating collaboration with peers as a CPD characteristic that positively impacts their instructional practices (OECD, 2019). The effectiveness of teacher collaboration in CPD, despite its straightforward appeal and the fact that relevant research consistently frames it as an effective CPD element (see Darling-Hammond et al., 2017; Desimone, 2009 for more), still lacks evidential warrant (Sims & Fletcher-Wood, 2021). Many CPD programs extensively use collaborative activities for teacher learning; however, the question arises as to whether the invaluable resources that countries allocate to CPD are well spent with respect to its effectiveness (Sims & Fletcher-Wood, 2021). Further, collaboration in CPD can take several forms, from face-to-face activities to social media discussions and interactive webinars, which may differ in their effectiveness (Damkuvieni et al., 2023; Quinn et al., 2019). A recent systematic review of online CPD programs concluded that it is still an open question how online CPD formats foster collaboration (Capparozza et al., 2023) thus, not only do we not know whether collaboration indeed

renders CPD effective for teachers, but we also don't know whether the impact of collaboration can be influenced from the format of the CPD. To answer these questions, we conduct a systematic review of teacher collaboration in CPD effectiveness studies with three primary research questions:

1. How is teacher collaboration operationalized in CPD effectiveness studies, and what collaborative activities are identified?
2. Are collaborative activities in CPD effectiveness studies grounded in theoretical frameworks that explain how these activities can affect teacher outcomes?
3. Which teacher outcomes are measured in the identified CPD effectiveness studies?

Depending on the number of identified studies and the availability of quantitative data (effect sizes), the effectiveness of CPD collaborative activities will be examined meta-analytically by asking the following question:

4. Do CPD collaborative activities affect the identified teacher outcomes, and do web-based or face-to-face CPD formats differentially influence their effectiveness?

Our systematic review follows the PRISMA guidelines (Page et al., 2021) and includes the typical stages of a systematic review: formulating research questions using the PICO framework, identifying relevant studies through predefined criteria and online databases, deduplicating and screening studies with at least two independent reviewers, extracting key information, appraising the quality of the evidence, and synthesizing the results narratively and statistically if effect sizes are available (Cooper, 2017; Grant & Booth, 2009). We conducted a systematic literature search in ERIC, PsycINFO, and Web of Science in May 2024. We identified 1,126 abstracts (after deduplication) that are currently screened in Rayyan software (Ouzzani et al., 2016) by two independent reviewers. We aim to finish screening and start with information extraction by October 2024. We will narratively synthesize all information with tabular accompaniment. The included studies and the adequacy of the quantitative data will determine the eventual meta-analysis of the CPD collaboration effects that we will perform in R-Studio (Posit team, 2024) with the packages metafor (Viechtbauer, 2010) and clubSandwich (Pustejovsky, 2024). The study is pre-registered for transparency at the Open Science Framework (OSF).

At the time of submission, we have screened 26% (n = 300) of our abstracts and included n = 60 potential studies for information extraction. We will be ready to present the first results of our systematic review at GEBF in January 2025.

Barrieren und Unterstützungsmöglichkeiten für Lehrkräfte bei der Nutzung offener Bildungsmaterialien (OER): Ein Systematic Review

Marcel Capparoza¹, Salome Wagner², Andreas Lachner¹

¹Eberhard Karls Universität Tübingen, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Offene Lizenzen (Bissell, 2009) von offenen Bildungsmaterialien (Open Educational Resources, OER) ermöglichen es, dass Lehrkräfte Unterrichtsmaterialien an die Bedürfnisse ihrer Klassen anpassen können und fördern die Zusammenarbeit zwischen Lehrkräften (Butcher & Wilson-Strydom, 2008; Jensen & Kimmons, 2022). Experimentelle Studien zeigen darüber hinaus, dass qualitativ hochwertige OER traditionellem Material ebenbürtig sein können (z. B. Hilton et al., 2019; Robinson et al., 2014). Die offenen Lizenzen von OER erlauben Lehrkräften fünf Aktivitäten, die auch als 5R-Aktivitäten bezeichnet werden (Wiley, o. D.; Wiley & Hilton, 2018): (1) Erstellung von Kopien sowie das Recht, OER zu besitzen und zu kontrollieren (retain), (2) unterrichtliche Nutzung von OER (reuse), (3) Anpassung der OER (revise), (4) Kombination von OER mit anderen Materialien (remix) und (5) Teilen von OER mit anderen Lehrkräften (redistribute).

Nicht OER per se, sondern die Nutzung von OER durch Lehrkräfte hat das Potenzial, die Bildungslandschaft zu verändern und für qualitativ hochwertige Bildung zu sorgen (Mishra, 2017). Deshalb ist das Ziel dieses Beitrags, durch ein Systematic Review in der Forschung beschriebene unterstützende und hemmende Faktoren für die fünf verschiedenen 5R-Aktivitäten zu identifizieren.

Das präregistrierte Systematic Review orientiert sich in der Durchführung an den PRISMA-Empfehlungen (Page et al., 2021). Das Ziel der Literaturrecherche und des anschließenden Auswahlprozesses ist die Identifikation von peer-reviewten empirischen Studien in englischer Sprache, die passend den Fragestellungen waren. Zunächst wurde in den Datenbanken ERIC, Google Scholar, PsycInfo, Scopus, ScienceDirect und Web of Science nach relevanten Studien im Zeitraum 2002 bis 2023 gesucht. Als Suchterm diente dabei eine Kombination aus Suchbegriffen zu Lehrkräften (z. B. teacher) und zu OER (z. B. „open educational resources“). Vorwärts- und Rückwärtssuchen dienten zur Sicherstellung einer breiten Abdeckung des Forschungsfelds. In den Suchergebnissen wurden anschließend im webbasierten Tool Rayyan (Ouzzani et al., 2016) Dubletten entfernt; im verbleibenden Korpus von 2910 Suchergebnissen wurden im ersten Screening die Titel und Abstracts der Studien (20 % der deduplizierten Suchkorpus wurden von zwei unabhängigen Ratern bewertet: $\kappa = .852$) und im zweiten Screening die Volltexte hinsichtlich thematischer Passung und methodischer Qualität überprüft. Die aufgenommenen Studien werden mit Blick auf die mit OER verknüpften 5R-Aktivitäten ausgewertet.

Bei der derzeit noch nicht abgeschlossenen Auswertung fällt eine große Vielfalt an unterschiedlichen Bildungssystemen, involvierten Materialien und Studiendesigns in den aufgenommenen Studien auf. Ein häufig auftretendes Problem, das in vielen Studien für die Wiederverwendung im Unterricht genannt wird, ist die niedrige Qualität der OER (z. B. de los Arcos et al., 2016). Auch zeitliche Schwierigkeiten bei der Überarbeitung und Nutzung von OER (z. B. Clements & Pawlowski, 2012), die schwere Auffindbarkeit von geeigneten OER (z. B. Tang & Bao, 2021) sowie eine negative Arbeitskultur (z. B. Bradley & Vigmo, 2016) sind häufig genannte Barrieren für Lehrkräfte. Als Anreize für Lehrkräfte erweisen sich vor allem Fortbildungsaktivitäten, zusätzlich zur hohen Qualität der OER (z. B. Kimmons, 2016) und der Unterstützung bei der Auffindbarkeit (z. B. Tang et al., 2020). Studien zu Fortbildungsaktivitäten zeigen, dass diese Lehrkräfte dabei unterstützen, OER sowohl pädagogisch sinnvoll zu nutzen als auch selbst zu bearbeiten (z. B. Maina et al., 2020; Tang et al., 2020).

Die aus dem Systematic Review gewonnenen Erkenntnisse zeigen Perspektiven für weitere Forschung über OER auf, z. B. die verstärkte Nutzung von Logdaten von OER-Repositorien und ein größerer Fokus der Forschung auf die Wirksamkeit von Fortbildungen für Lehrkräfte bzw. deren Unterricht. Auf der anderen Seite lassen sich Maßnahmen sowohl für Personen aus der Bildungsadministration als auch für Fortbildende und Lehrkräfte ableiten, durch die OER stärker in der schulischen Bildung verankert werden kann. Beispiele hierfür stellen die Erhöhung der Aufmerksamkeit auf das Thema OER in der Aus- und Fortbildung von Lehrkräften sowie die Bereitstellung und nutzerfreundliche Gestaltung von OER-Repositorien dar.

Digitalisierungsbezogene Kompetenzen von (angehenden) Lehrkräften unabhängig von ihrer Selbsteinschätzung erheben – Ein Rapid Review zur Übersicht verfügbarer Testinstrumente

Lisa Ziernwald¹, Frauke Voitle², Meg Farrell¹, Jakob Thomas Fischer³, Ben Heinitz⁴, Linda von Sobbe⁵, Moritz Krell², Doris Holzberger¹

¹Technische Universität München, ²IPN –Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik,

³Universität Potsdam, ⁴Universität Münster, ⁵Deutsches Institut für Erwachsenenbildung

In den letzten Jahren hat die Förderung digitalisierungsbezogener Kompetenzen in der Lehrkräfteaus- und -fortbildung zunehmend an Bedeutung gewonnen (Diepolder et al., 2021; van Ackeren et al., 2019). Bislang werden zur Erfassung von Kompetenzen im Rahmen der Evaluation von Fortbildungen häufig Selbsteinschätzungsinstrumente eingesetzt (Lachner et al., 2021). Es wird jedoch kritisiert, dass diese von der Fähigkeit der Lehrkräfte abhängen, ihr Wissen richtig einzuschätzen, und daher zu Verzerrungen neigen (Claro et al., 2024; Kopcha & Sullivan, 2007). Entsprechend zeigt sich über Fachgebiete hinweg nur eine mäßige Korrelation zwischen selbstberichteten Fähigkeiten und objektiven Leistungsmaßen (Zell & Krizan, 2014). Daher ist es wichtig, auch Testinstrumente, welche nicht auf Selbsteinschätzung beruhen, zur Erfassung digitalisierungsbezogener Kompetenzen von Lehrkräften zu nutzen.

Es wird ein Rapid Review vorgestellt, das die Erstellung einer narrativen Synthese zum Ziel hat, die sowohl Forschende als auch Lehrkräftebildende bei der Auswahl geeigneter Testinstrumente zur Erfassung digitalisierungsbezogener Kompetenzen von Lehrkräften unterstützt. Es werden verfügbare, nicht auf Selbsteinschätzungen beruhende Testinstrumente dargestellt. Zusätzlich werden ihre theoretischen Grundlagen, strukturellen und psychometrischen Eigenschaften zusammengefasst, sodass die in der Forschung entwickelten Testinstrumente inklusive deren Eigenschaften für die digitale Transformation in der (Fortbildungs-)Praxis leichter nutzbar gemacht werden.

Die narrative Synthese beantwortet die folgenden drei Forschungsfragen:

Welche Instrumente zur Erhebung der digitalisierungsbezogenen Kompetenzen von (angehenden) Lehrkräften, die nicht auf Selbsteinschätzungen beruhen, sind derzeit verfügbar?

Welche theoretischen Hintergründe, Operationalisierungen und Kontextualisierungen der digitalisierungsbezogenen Kompetenzen von (angehenden) Lehrkräften werden in diesen Instrumenten verwendet?

Welche strukturellen (d. h. Instrumententyp, Antwortformat, Umfang) und psychometrischen Eigenschaften besitzen diese Instrumente?

Zur Beantwortung der Forschungsfragen und mit Blick auf die aktuellen Bedarfe der Bildungsforschung und -praxis wird ein Rapid Review als ressourcenschonende, strikte und transparente Form der Forschungssynthese durchgeführt (Hamel et al., 2020). Ein Rapid Review erlaubt einerseits die gewünschte schnelle Bereitstellung einer Instrumentenübersicht und andererseits für eine zielgerichtete Ausrichtung und Durchführung des Rapid Reviews zudem die Einbindung von Stakeholdern in den Prozess. Da bei Rapid Reviews jedoch auch ein erhöhtes Risiko für Verzerrungen besteht (Garrity et al., 2024), ist die Transparenz getroffener Entscheidungen besonders wichtig, weswegen das vorgestellte Rapid Review im Open Science Framework (OSF) präregistriert wurde.

Für die Literatursuche wurden die Datenbanken ERIC und PsycInfo mit Hilfe von EBSCOhost nach Artikeln aus Fachzeitschriften mit peer-review Verfahren durchsucht. Die genutzte englischsprachige Syntax kombinierte dabei (synonyme) Begriffe zu den vier Bereichen Zielgruppe (Lehrkräfte), Erhebungsinstrumente, Digitalisierungskontext und Zielkonstrukt (Kompetenz).

Nach Ausschluss von Duplikaten wurden 2.874 Artikel in das Screening von Titeln und Abstracts aufgenommen. 20 % der Titel und Abstracts wurden doppelt gescreent, wobei Interrater-Übereinstimmungen von Cohen's Kappa = 0.64 und Gwet's AC1 = 0.82 erreicht wurden, welche als gut bis sehr gut einzustufen sind (Cicchetti, 1994). Für das Screening mit Volltext, welches zum Zeitpunkt der Beitragseinreichung gestartet ist, wurden 528 Artikel eingeschlossen. Die Feinkodierung der relevanten Treffer, welche parallel zum Volltext-Screening erfolgen wird, steht unmittelbar bevor.

Die Befragung von N = 28 Stakeholdern über einen Online-Fragebogen, bei der unter anderem Journals für die Handsuche vorgeschlagen wurden, wurde bereits erfolgreich abgeschlossen. Das Screening der darauf basierenden Handsuchetreffer wird demnächst abgeschlossen sein.

Auf der Tagung können sowohl die (vorläufigen) Ergebnisse der Feinkodierung sowie der Stakeholder-Befragung präsentiert werden. Vor dem finalen Abschluss des Rapid Reviews soll die Tagung genutzt werden, um die Instrumenten-Übersicht vorzustellen und zu diskutieren, um so die Perspektive der Tagungsteilnehmenden als weitere relevante Stakeholder einzuholen. Im Sinne eines Rapid Reviews sollen die Rückmeldungen für eine möglichst adressatengerechte und zielführende Aufbereitung der Ergebnisse genutzt werden.

Durch die angestrebte Übersicht über verfügbare Instrumente zur Erfassung digitalisierungsbezogener Kompetenzen von Lehrkräften, die nicht auf Selbsteinschätzungen beruhen, schafft das vorgestellte Rapid Review eine wichtige Voraussetzung für die wissenschaftlich fundierte und evidenzbasierte digitale Transformation von Schule und Lehrkräftebildung.

Einsatz von Large Language Models für die Literaturrecherche in der Bildungsforschung: Eine Evaluation mit GPT-4

Dennis Kasch, Tamara Heck

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Forschungssynthesen (FS) sind ein wichtiges Werkzeug für die Erschließung wissenschaftlicher Erkenntnisse. Sie ermöglichen es, einen umfassenden Überblick über den aktuellen Forschungsstand bezüglich einer wissenschaftlichen Fragestellung zu gewinnen, Muster und Widersprüche zu identifizieren, evidenzbasierte Schlussfolgerungen zu ziehen und Empfehlungen zu formulieren (Gurevitch, 2018). Der mit der Erstellung von FS verbundene Aufwand ist jedoch beträchtlich, da eine gründliche Literaturrecherche, eine sorgfältige Bewertung der Studien und eine systematische Analyse der Ergebnisse erforderlich sind. In der Literaturrecherche für FS müssen oft große Mengen von Studien bewertet werden, um relevante Literatur zu identifizieren (Woerfel, 2021).

Es existieren bereits verschiedene, auf Künstlicher Intelligenz (KI) und Machine Learning (ML) basierende Ansätze zur Unterstützung und Automatisierung des Screenings in FS (Chernikova, 2024; Kebede et al., 2023; Zhang et al., 2022). Durch den Aufstieg von Large Language Models (LLMs) entstehen jedoch neue Möglichkeiten für den Einsatz von KI im Screeningprozess, besonders hinsichtlich natürlichsprachlicher Interaktion mit einem auf LLMs basierendem System. LLMs zeigen unter anderem beachtliche Fähigkeiten in den Bereichen Sprachverständnis, Übersetzung, Textgenerierung und Textklassifikation. Sie besitzen durch den großen Umfang der zugeführten Trainingsdaten zudem eine große Wissensbasis und sind in der Lage natürlichsprachliche Begründung für getroffene Entscheidungen zu erstellen (Naveed, 2023). Diese Fähigkeit bildet ein Alleinstellungsmerkmal, da klassische ML-Verfahren meist auf ein reines Ranking von relevanter Literatur abzielen

(Marshall, 2019). Es ermöglicht Wissenschaftler*innen durch die nachträgliche Überprüfbarkeit von getroffenen Entscheidungen zusätzliche Transparenz.

Zusammen bilden diese Eigenschaften eine vielversprechende Grundlage für die Analyse und Auswahl relevanter Literatur zu wissenschaftlichen Fragestellungen in der Bildungsforschung.

Die deutlich geringere Bearbeitungszeit durch LLMs, ermöglicht es in kurzer Zeit große Mengen von potenziell relevanten Arbeiten zu überprüfen und damit die Durchführung einer Literaturrecherche für FS signifikant zu beschleunigen. Inzwischen besitzt GPT-4 ein Kontext-Fenster von bis zu 128.000 Tokens (OpenAI, 2024), was sogar die Analyse von Volltexten ermöglicht. Dadurch besteht das Potenzial für LLM-basierte Verfahren, zukünftig eine wesentliche Rolle bei der wissenschaftlichen Literaturrecherche für FS in der Bildungsforschung zu spielen.

In dieser Arbeit wird exemplarisch anhand von GPT-4 untersucht, wie gut LLMs in der Lage sind, die Relevanz von wissenschaftlichen Arbeiten für eine Forschungsfrage einzuschätzen. Dafür wird ein Verfahren entwickelt, bei dem anhand einer natürlichsprachlichen Beschreibung der Forschungsfrage(n), sowie Ein- und Ausschlusskriterien für das Screening in einer FS, ein Prompt für das Modell erstellt wird, welches die Identifikation von relevanter Literatur ermöglicht. Das LLM-basierte Verfahren greift dabei primär auf Titel, Abstract und Schlagworte der Literatur zurück, was der gängigen Praxis im manuellen Prozess folgt (Gupta, 2018). Für das Verfahren wird untersucht, welchen Einfluss die verschiedenen Merkmale und ihre Kombinationen auf die Genauigkeit des Ergebnisses haben. Erste Pilotierungen zeigen dabei, dass bereits ein ausschließlich auf Titeln von Arbeiten basierende Analysen gute Ergebnisse erzielen können. Zusätzlich werden Stichproben von Volltextanalysen untersucht, um auch diesen Aspekt zu untersuchen und zu vergleichen.

Das Verfahren wird anhand etablierter Leistungsmetriken (Erickson, 2021) und auf Daten aus bereits durchgeführten FS in der Bildungsforschung (siehe Datensätze von Campos, 2024) evaluiert, um die quantitative Leistungsfähigkeit des Verfahrens zu bewerten. Zusätzlich werden qualitative Untersuchungen durchgeführt, um die vom System gelieferten Begründungen für die getroffenen Entscheidungen zu überprüfen. Dafür werden die Begründungen des LLM-Systems mit denen von Forschenden aus FS der Bildungsforschung verglichen.

Die Ergebnisse zeigen zum einen auf, wie gut aktuelle LLMs, selbst ohne spezifische Anpassung (Zero-Shot) an den Kontext der Bildungsforschung, in der Lage sind FS zu unterstützen und Prozesse so effektiv zu gestalten, dass sie es ermöglichen, kontinuierlich aktuelle Erkenntnisse zu relevanten Themen in der Bildungsforschung zur Verfügung zu stellen. Zum anderen geben sie Aufschlüsse darüber, wie Transparenz und Nachvollziehbarkeit bei entsprechenden Verfahren gewährleistet werden können.

S-40: Application of Generative Large Language Models in Teacher-Centered Educational Research

Zeit: Dienstag, 28.01.2025: 9:00 - 10:45 · Ort: O 135

Chair der Sitzung: Yizhen Huang

Chair der Sitzung: Michael Sailer

English Track

Application of Generative Large Language Models in Teacher-Centered Educational Research

Chair(s): **Yizhen Huang** (University of Potsdam, Germany)

Diskutant*in(nen): **Michael Sailer** (University of Augsburg)

Large language models (LLM) is the artificial intelligence systems that process and generate human language (Naveed et al., 2024). LLM excels in natural language tasks, in many cases to human-level. For instance, natural language understanding and generation, text classification, sentiment analysis, and speech recognition.

The rise of consumer-friendly LLM services like OpenAI's GPT and open-source LLM platforms such as Meta's Llama has revolutionized the accessibility of AI technology. While the field of instructional technology has embraced LLM in areas such as intelligent learning agents Field (Dai et al., 2024) and adaptive feedback Field (Sailer et al., 2022), there remains a noticeable gap in the integration of LLM into teachers' professional learning.

How to effectively make use of LLM for training and studying teachers' professional learning is an untapped question that the present symposium hopes to shed light on. The symposium will present three diverse research projects with the aim to examine LLM's potential both as research and teacher training tool.

Contribution 1 explored preservice teachers' talk moves during a classroom debate within a fully immersive virtual reality classroom, where the teacher and the LLM-powered virtual students had authentic classroom dialogs. Contribution 2 also focused on teacher talk, but LLM was not used as part of simulation technology, but an intelligent coder more efficient than human coders, to classify the types of motivational messages in preservice teachers' real-life classroom teaching transcripts. Contribution 3 compared and prompt-engineered three market-leading LLM (Claude3, ChatGPT-4 and Gemini Advanced) to be used for high-quality feedback for teachers.

This symposium will follow the standard 105-minute format: the chair's introductory remarks will take 5 minutes, followed by 20 minutes for each contribution (60 in total), 20 minutes of discussant critique, and 15 minutes of discussions that involve all contributors and audiences.

Beiträge des Symposiums

Explore Teacher Talk Moves in a Generative AI-Powered Virtual Reality Classroom Dialog Simulation

Yizhen Huang¹, Nico Klausner-Thimm¹, Eric Richter¹, Katharina Scheiter¹, Thilo Kleickmann², Dirk Richter¹

¹University of Potsdam, ²Kiel University

Introduction

Teaching simulation as an approximation for professional practice (Duchatelet et al., 2022), provides a contextually rich environment for practicing dialogic pedagogy. Large language models (LLM)—the artificial intelligence systems that effectively process and generate human language (Naveed et al., 2024) shed new light to classroom dialog simulations. The present study seeks to explore preservice teachers' talk moves during a classroom debate within a fully immersive VR classroom, where the user and the LLM agents each embodied as the teacher and students in one authentic classroom scene. In particular, we explored the following research questions:

RQ1: How did preservice teachers perceive the simulation environment and task (leading discussion) in terms of authenticity, situational interest, and self-efficacy in classroom discussion?

RQ2: What was the quantity and length of teacher and student utterances?

RQ3: What were the types of talk moves preservice teachers used in the simulation?

Methods

Twenty-three preservice teachers were recruited from the initial teacher education program of a public German university (Mage = 24.2 years; 53.8% cisgender female; 57.7% enrolled in the Bachelor program, Msemester = 6.1).

We developed a simulation system called VRAISE (Virtual Reality and AI-powered Simulation for Educators)—a completely locally-run conversation simulation system that adjoin a VR classroom environment with a chain of generative AI software. The VR classroom was designed to resemble classrooms in German secondary public schools and was deemed realistic by participants with similar demographics in other research utilizing similar immersive VR classroom systems (e.g., Authors et al., 2022).

During interactions between the teacher and virtual students in the classroom, the speech of the teacher was transcribed automatically by Whisper. This transcription served as the user prompt and was sent as input to a Llama 3 instance operating on the same local machine. The LLM instance was instructed to speak in the voice of the student the user spoke to, with varying levels of argumentation strategies. The generated text response was converted into students' speech by voice models trained on real-human voice, known as voice clones, using a local neural text-to-speech engine (Rhasspy, 2024).

Immediately after the simulation, we measured participants' authenticity (Codreanu et al., 2020), task situational interest (Knogler et al., 2015), and self-efficacy in classroom discussion (Waag et al., 2021).

Results and Discussion

Participants perceived the simulation to be authentic ($M = 4.22$). They also reported high situational interest ($M = 3.53$) and self-efficacy ($M = 4.19$) in leading classroom discussion after the simulation. For RQ2, a paired samples t-test revealed a significantly higher number of teacher utterances ($t(22) = 2.12$, $p = .046$; $M = 23.65$) than the student utterances ($M = 21.39$) with a medium effect size ($d = 0.44$). Yet the total number of words spoken by teachers were significantly fewer ($t(22) = -4.19$, $p < .001$; $M = 435.70$) than students ($M = 654.83$) with a large effect size ($d = -0.87$). This result partially align with findings from past classroom

dialogue research that teachers without dialogic pedagogy training may talk more frequently than students (e.g., Wang et al., 2013). RQ3 is yet to be examined as the coding is still undergoing.

With these findings, we were able to show that preservice teachers' talk moves in a generative AI-powered VR simulation and actual classrooms were comparable in its manifestation. The potential of the VRAISE system would directly address the lack of training opportunities for preservice teachers' dialogical expertise in promoting students' argumentation skills.

Assessing Pre-service Teachers' Motivational Messages with a Large Language Model

Olivia Metzner¹, Yindong Wang², Wendy Symes¹, Yizhen Huang¹, Lena Keller¹, Gerard de Melo², Rebecca Lazarides¹

¹University of Potsdam, ²Hasso-Plattner-Institute

Theoretical background

Recent studies have examined the relation between teacher motivational messages, teacher motivation, and student learning (Putwain & von der Embse, 2018; Symes & Putwain, 2016). However, a limitation of these studies is that motivational messages are predominantly evaluated using students' self-reports (Putwain & Symes, 2011; Santana-Monagas et al., 2022). This approach is problematic due to the susceptibility to biases in research findings (Howard, 1980; Rosenman et al., 2011). Falcon et al. (2023) utilized teacher classroom transcripts to manually classify motivational messages, but this method is resource-intensive, requiring human rater training and manual data annotation (Macanovic & Przepiorka, 2024). To address these limitations, we propose developing an AI-based method to objectively assess teacher motivational messages. In this study, we aim to fine-tune and evaluate a Large Language Model (LLM), which is an artificial intelligence system trained on extensive corpora of parameters to understand, generate, and process human language (Myers et al., 2024) and can be adapted to specific tasks through fine-tuning (Demszky et al., 2023).

Research question

In the current study we examine the following research question: Can pre-service teachers' motivational messages be reliably assessed via AI-based methods, namely LLMs?

Method

We used video data from 122 pre-service teachers who participated in a 15-week practice-oriented course during their bachelor studies. Pre-service teachers were on average in their fifth semester ($SD = 1.72$, $Min = 2$, $Max = 15$) and the three most studied subjects were German (11.97 %), English (13.25 %) and Biology (9.83 %). First, we transcribed the video recordings using AI-based software, DaVinci Resolve Studio and Panopto (Blackmagic Design, 2023; Panopto, 2024). Second, we developed a coding system for motivational messages based on self-determination theory (SDT) (Ahmadi et al, 2023; Deci & Ryan, 2002) and annotated the transcripts with two human raters ($\kappa = .72$). Third, we used these annotations to fine-tune the LLM Gemma-7B (Banks & Warkentin, 2024). We split the annotated examples into a training and evaluation data set. Next, a prompt-based fine-tuning approach was applied to adapt the LLM to assess teachers' message behavior (Schick & Schütze, 2020). Finally, we evaluated the model's predictions by comparing them to human rater annotations.

Results and their significance

In the evaluation phase, our models achieved an average accuracy of 72.27% across all categories. The message categories of autonomy support, competence support, relatedness support, and relatedness thwart resulted in improved F1-scores over 70% compared to the LLM without fine-tuning. However, the categories of autonomy thwart, competence thwart, and out-of-domain data did not reach acceptable F1-values.

The LLM demonstrated high adaptability and effectiveness in handling multi-class classification tasks for supportive and relatedness thwarting messages. These findings suggest that the fine-tuned LLM can reliably classify teacher motivational messages. However, pre-service teachers in our sample used thwarting messages five times less frequently than supportive messages, resulting in the LLM being trained with an unequal amount of data for all categories. We recommend further fine-tuning the LLM on a more balanced and larger data set to improve accuracy for thwarting messages and out-of-domain data.

ChatGPT & Co.: Feedback as a Use Case for LLM's in Higher Education

Lucas Jasper Jacobsen¹, Kira Elena Weber²

¹Leuphana Universität Lüneburg, ²Universität Hamburg

Theoretical Background

The integration of large language models (LLMs) into educational contexts is increasing, yet few empirical studies evaluate these systems' quality and effectiveness in higher education (HE) (Crompton & Burke, 2023). Previous research has shown that AI-generated formative feedback can significantly increase student engagement and satisfaction (Demszky et al., 2023). High-quality feedback is characterized by criteria such as concreteness, activation and empathy (Prilop et al., 2019), but the provision of such feedback is often constrained by limited resources (Demszky et al., 2023). Preliminary results suggest that AI feedback can outperform that of novices and even experts in certain scenarios (Jacobsen & Weber, 2023). Building on these findings, this research examines the criteria for optimal prompt engineering in different LLMs in two studies and evaluates their effectiveness in providing feedback for different student learning goals.

Research Questions

The research questions of the first study are: Do different LLMs require different criteria for high-quality prompts? Which criteria have the greatest influence on the generation of high-quality feedback with LLMs? In study two, the guiding research question was: What are the differences between Claude3, ChatGPT-4 and Gemini Advanced in terms of feedback quality?

Method

For the first study, we selected three LLMs based on criteria of accessibility, ease of use and anticipated output quality. Specifically, these were ChatGPT-4, Claude 3, and Gemini Advanced. An empirically validated feedback prompt (Jacobsen & Weber, 2023) was revised based on the categories examples, conciseness, and domain specificity, resulting in three new prompts. All four prompts were used in the three LLMs and yielded feedback on 20 learning goals from 20 pre-service teachers, totaling 240 pieces of feedback. Three trained coders coded the quality of the feedback. For the second study, the learning goals from 153 pre-service teachers were collected during their initial teaching practicum. Using the optimal prompts identified in study one, feedback was generated for these learning goals using the LLMs Claude3, ChatGPT-4, and Gemini Advanced. A total of 459 pieces of feedback were then coded for quality by the same three trained coders from study number one. In both studies feedback quality was compared using ANOVA.

Results & Significance

The results of the first study indicate that high-quality prompts require consistent criteria, which were validated with a prompting manual from a previous study. Domain specificity, i.e. the correct use of technical terms, proved to be the most important factor for generating expert-level feedback. In addition, the results of the second study show that ChatGPT-4 performed best in six categories of feedback quality, with specific, concrete, and activating responses. Claude 3 had the fewest errors and performed well, while Gemini provided balanced feedback in terms of valence but was inferior compared to ChatGPT-4 and Claude 3.

This study shows that LLMs are effective feedback tools when combined with expertly crafted prompts and integrated 'human-in-the-loop' supervision, underscoring the necessity of expert knowledge for producing high-quality output. Furthermore, the results have shown that there are objective quality criteria for prompt engineering, making prompting a learnable skill that should be taught to both teachers and students.

IS-04: Invited Symposium: Programme for the International Assessment of Adult Competencies (PIAAC 2023): Aktuelle Ergebnisse in Deutschland und im internationalen Vergleich

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 102

Chair der Sitzung: Debora Maehler

Chair der Sitzung: Beatrice Rammstedt

Programme for the International Assessment of Adult Competencies (PIAAC 2023): Aktuelle Ergebnisse in Deutschland und im internationalen Vergleich

Chair(s): **Debora Maehler** (GESIS - Leibniz-Institut für Sozialwissenschaften, Deutschland), **Beatrice Rammstedt** (GESIS - Leibniz-Institut für Sozialwissenschaften, Deutschland)

Das Symposium präsentiert und diskutiert erste Ergebnisse aus dem zweiten Zyklus von PIAAC (2022/ 2023) – Programme for the International Assessment of Adult Competencies. PIAAC, auch bekannt als Survey of Adult Skills, wurde von der OECD initiiert und vergleicht international die kognitiven Kompetenzen der erwachsenen Bevölkerung in 10-Jahres Abständen. Die erfassten grundlegenden Kompetenzen werden als entscheidend für eine erfolgreiche Informationsverarbeitung im Alltag und für ein erfolgreiches Berufsleben betrachtet (Rammstedt et al., 2013). Diese Kompetenzen sind erforderlich, um die eigenen Ziele zu erreichen, das eigene Wissen und Potenzial zu entwickeln und an der Gesellschaft teilzunehmen; dazu zählen die Lesekompetenz, die alltagsmathematische Kompetenz und das adaptive Problemlösen. Unter Lesekompetenz wird zum Beispiel die Fähigkeit erfasst, auf geschriebene Texte zuzugreifen, sie zu verstehen, zu bewerten und zu reflektieren (Rammstedt et al., im Druck).

Die Erhebung der Hintergrundinformationen in PIAAC Zyklus 2 erfolgte 2022/2023 als persönliche Befragung durch Interviewer; während die Kompetenzaufgaben selbstständig von den befragten Personen auf einem Tablet bearbeitet wurden. Die Large-Scale-Untersuchung erlaubt Schätzungen auf Bevölkerungsebene, jedoch keine Aussagen über interindividuelle Unterschiede. Die Schätzung der Kompetenzen erfolgt auf der Grundlage von Modellen der Item-Response-Theorie (IRT). Bei der Nutzung von IRT-Modellen können Kompetenzwerte geschätzt werden, auch wenn die Befragungspersonen jeweils unterschiedliche Aufgaben bearbeitet haben. Um den daraus resultierenden Schätzfehler zu berücksichtigen, wird aus der geschätzten Verteilung der Kompetenzwerte jedoch nicht nur ein einzelner Wert je Kompetenzbereich verwendet, sondern es werden pro Person mehrere sogenannte Plausible Values (PV) gezogen. Im Falle von PIAAC wurden 10 PVs verwendet (Maehler & Rammstedt, 2020; Martin et al., 2013; OECD, 2013; 2021).

Ziel des Symposiums ist es, einen Überblick über die Ergebnisse von PIAAC 2022/2023 zu geben, indem die kognitiven Kompetenzen (1) im Zusammenhang mit verschiedenen Faktoren auf nationaler Ebene betrachtet werden, (2) mit Ergebnissen aus der letzten Dekade verglichen werden (PIAAC 2011/2012) sowie (3) im internationalen Kontext vergleichend untersucht werden. Das Symposium beginnt mit einer Einführung in die Struktur sowie die Kompetenzmessung bei PIAAC und gibt einen Überblick über grundlegende Kompetenzen im Ländervergleich. Die aktuellen Ergebnisse werden außerdem vergleichend zur ersten repräsentativen Kompetenzmessung der Bevölkerung vor zehn Jahren dargestellt (Martin et al.). Der zweite Vortrag beschäftigt sich mit den Kompetenzen von spezifischen Bevölkerungsgruppen (Gauly et al.). Hierbei werden die Kompetenzen getrennt nach Geschlechtern, nach Geburtskohorten sowie nach der Teilhabe am Arbeitsmarkt untersucht. Die Ergebnisse werden ebenso im Hinblick auf Veränderungen in der letzten Dekade (PIAAC Zyklus I) betrachtet und in die Literatur eingebettet. Der dritte Beitrag (Massing & Schneider) fokussiert auf die grundlegenden Kompetenzen von Personen mit verschiedenen Bildungsabschlüssen in Deutschland und im internationalen Vergleich. Kompetenzen werden außerdem innerhalb von Bildungsgruppen über die Zeit verglichen und die Bildungsinflation diskutiert. Das Symposium endet mit einem Beitrag zum Zusammenhang von sozio-emotionalen Kompetenzen mit in PIAAC erfassten Lese- und/oder mathematischen Kompetenzen (Rammstedt et al.). Die Zusammenhangsmuster werden im internationalen Vergleich erörtert.

Das Symposium gibt einen spannenden Einblick in die derzeitigen kognitiven Kompetenzen der Bevölkerung (Humankapital) und damit einhergehenden Faktoren, sowie Veränderungen über die Zeit. Die Präsentationen liefern somit wertvolle Erkenntnisse für Politik, Bildung und Forschung, und sollen weiterführende wissenschaftliche Arbeiten und Diskussionen anregen.

Beiträge des Symposiums

Grundlegende Kompetenzen in Deutschland im internationalen Vergleich

Silke Martin, Anouk Zabal, Britta Gauly, Sanja Kapidzic, Natascha Massing

GESIS - Leibniz-Institut für Sozialwissenschaften, Deutschland

Das Programme for the International Assessment of Adult Competencies (PIAAC) ist eine multi-zyklisch angelegte Studie der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD), die die grundlegenden Kompetenzen von Erwachsenen zwischen 16 und 65 Jahren im internationalen Vergleich misst. Am ersten Zyklus von PIAAC nahmen zwischen 2011 und 2018 insgesamt 39 Länder – über drei Erhebungsrunden verteilt – teil (OECD, 2019). Die Datenerhebung des zweiten Zyklus mit 31 Ländern fand in 2022/2023 statt. Für Zyklus 2 werden erste Ergebnisse Ende 2024 vorliegen.

Zu den in PIAAC gemessenen grundlegenden Kompetenzen, die als Grundvoraussetzung angesehen werden, um in Beruf und Alltag teilhaben zu können, zählen in Zyklus 2 die Lesekompetenz, die alltagsmathematische Kompetenz sowie das adaptive Problemlösen. Die theoretischen Rahmenkonzeptionen der Domänen wurden von internationalen Expertengruppen entwickelt und lassen sich wie folgt zusammenfassen (vgl. Greiff et al., 2017; OECD, 2021, S. 42, 93, 159):

- **Lesekompetenz:** Fähigkeit, sich Zugang zu schriftlichen Texten zu verschaffen, sie zu verstehen, zu bewerten und über sie zu reflektieren, um seine Ziele zu erreichen, sein Wissen und Potential zu entwickeln und an der Gesellschaft teilzuhaben.
- **Alltagsmathematische Kompetenz:** Fähigkeit, sich mathematische Inhalte, Informationen und Ideen, die auf vielfältige Weise dargestellt werden, zugänglich zu machen, diese anzuwenden und kritisch damit umzugehen, um die mathematischen Anforderungen in einer Reihe von Situationen im Erwachsenenleben zu bewältigen.
- **Adaptives Problemlösen:** Fähigkeit, in einer dynamischen Situation, in der ein Lösungsweg nicht sofort verfügbar ist, die eigenen Ziele zu erreichen. Es erfordert kognitive und metakognitive Prozesse, um das Problem zu definieren, nach Informationen zu suchen und eine Lösung in einer Vielzahl von Informationsumgebungen und Kontexten anzuwenden.

Deutschland hat sowohl im ersten als auch im zweiten Zyklus von PIAAC teilgenommen. In Zyklus 1 erzielten Erwachsene in Bezug auf die Lesekompetenz im Mittel 270 Punkte und in Bezug auf die alltagsmathematische Kompetenz 272 Punkte (Rammstedt, 2013). Mit diesen Ergebnissen lag Deutschland im Vergleich zu den an Runde 1 teilnehmenden Ländern im Mittelfeld, nahe dem OECD-Durchschnitt.

Im Rahmen des vorliegenden Beitrags werden zur Einführung in PIAAC zunächst zentrale Merkmale sowie notwendige Rahmeninformationen zur besseren Einordnung und Verständlichkeit der Ergebnisse vermittelt. Wir werden dann darstellen, wie sich die an PIAAC Zyklus 2 teilnehmenden Länder inkl. Deutschland im Hinblick auf die untersuchten Kompetenzen ihrer erwachsenen Bevölkerung unterscheiden. Dazu werden wir auf der höchsten Aggregationsebene – der Gesamtbevölkerung der 16- bis 65-Jährigen – erste zentrale Ergebnisse aus Zyklus 2 (zum Beispiel Verteilungen nach Mittelwerten, Kompetenzstufen) vorstellen und beschreiben, inwieweit Veränderungen der Ergebnisse zu Zyklus 1 zu verzeichnen sind. Die PIAAC Daten dienen Politik, Öffentlichkeit und Wissenschaft als reichhaltige Informationsquelle und die wissenschaftlich fundierten Ergebnisse können politische Entscheidungsträgerinnen und -träger darin unterstützen, Verbesserungsbedarfe im Bildungs- und Weiterbildungsbereich sowie am Arbeitsmarkt zu identifizieren.

Grundlegende Kompetenzen in verschiedenen Bevölkerungsgruppen

Britta Gauly, Sanja Kapidzic, Silke Martin, Natascha Massing, Anouk Zabal

GESIS - Leibniz-Institut für Sozialwissenschaften, Deutschland

In Zeiten von komplexer werdenden Anforderungen zur gesellschaftlichen Teilhabe und am Arbeitsmarkt, spielen grundlegende Kompetenzen – wie Lese- oder alltagsmathematische Kompetenzen – eine immer wichtigere Rolle.

Im Alltag müssen sich Personen mit Texten aus den unterschiedlichsten Quellen auseinandersetzen: Dabei werden Informationen, z.B. zu politischen oder gesundheitlichen Themen, nicht mehr nur aus traditionellen Nachrichtenmedien, sondern auch aus anderen Quellen wie etwa sozialen Medien oder Diskussionsforen bezogen (Greenberg & Feinberg, 2019). Um den Wahrheitsgehalt dieser Informationen bewerten zu können und sich kritisch damit auseinander zu setzen, sind ausreichende Lesekompetenzen erforderlich (OECD, 2019). Zudem ist ein grundlegendes alltagsmathematisches Verständnis notwendig, um die in diesen Informationen präsentierte Zahlen und Statistiken richtig einordnen zu können (Geiger et al., 2015).

Frühere Studien haben gezeigt, dass sich die grundlegenden Kompetenzen zwischen verschiedenen Bevölkerungsgruppen teilweise signifikant unterscheiden (Rammstedt et al., 2013; OECD, 2019). So haben jüngere Geburtskohorten in Deutschland höhere Lese- und alltagsmathematischen Kompetenzen als ältere Geburtskohorten. Ebenso weisen Männer höhere Kompetenzen auf als Frauen. Auch die Arbeitsmarktteilhabe hängt mit dem individuellen Kompetenzniveau zusammen: erwerbstätige Personen zeigen höhere Kompetenzen als diejenigen, die nicht am Arbeitsmarkt aktiv sind.

Aufgrund der Bedeutung grundlegender Kompetenzen für fast alle Lebensbereiche und unter sich ständig verändernden gesellschaftlichen Anforderungen, ist es notwendig Erkenntnisse zum Stand der Kompetenzen in der Bevölkerung periodisch neu zu untersuchen. PIAAC 2022/2023 liefert ein aktuelles Bild davon, in welchem Ausmaß Kompetenzen in Lesen und Alltagsmathematik in verschiedenen Bevölkerungsgruppen vorliegen.

Der vorliegende Beitrag stellt erste deskriptive Ergebnisse aus dem aktuellen PIAAC-Zyklus vor und untersucht potenzielle Unterschiede nach Geschlecht, Geburtskohorten und Arbeitsmarktteilhabe für die in Deutschland lebenden Bevölkerung. Der Fokus liegt auf den Mittelwerten und der Streuung der Kompetenzen in den verschiedenen Gruppen und auf statistisch signifikanten Unterschieden. Zudem werden die aktuellen Daten, mit denen des ersten PIAAC-Zyklus aus 2012 verglichen und eingeordnet. Die Erkenntnisse können zum Verständnis beitragen, ob sich die Kompetenzen in der Bevölkerung in der letzten Dekade verändert haben und ob Defizite in einzelnen Bevölkerungsgruppen vorliegen.

Bildungserfolg und grundlegende Kompetenzen

Natascha Massing, Silke Schneider

GESIS - Leibniz-Institut für Sozialwissenschaften, Deutschland

Grundlegende Kompetenzen, wie Lesekompetenzen, mathematische Kompetenzen sowie Problemlösekompetenzen stellen zentrale Kompetenzen dar, auf welchen weitere Lernprozesse aufbauen. Ein zentrales Ziel des Bildungssystems ist es, Personen mit Kompetenzen auszustatten, die für die Gesellschaft und das weitere Leben relevant sind. Damit ist zu erwarten, dass das erreichte Bildungsniveau einer Person eng mit ihren Grundkompetenzen zusammenhängt (z.B. Rammstedt, 2013). In der Tat ist formale Bildung der bedeutsamste Einflussfaktor bei der Erklärung von Unterschieden in Grundkompetenzen (Maehler et al., 2013).

Bildung ist somit ein zentrales Element von Humankapital. Bildungsabschlüsse werden häufig sogar mit Kompetenzen gleichgesetzt. Der Zusammenhang zwischen Bildungsabschlüssen und Kompetenzen ist jedoch komplexer (Massing & Schneider, 2017). Bildungssysteme sind nicht vollständig hierarchisch, sondern weisen auch eine horizontale Differenzierung auf, wie etwa die Unterscheidung zwischen allgemeiner und beruflicher Bildung. Auch kann der gleiche Abschluss auf unterschiedlichen Bildungswegen erreicht werden. Es ist davon auszugehen, dass sowohl die Selektion von Individuen in verschiedene Bildungsgänge sowie der Kompetenzerwerb je nach Bildungsweg unterschiedlich ist. Darüber hinaus erfordert der Erwerb von Bildungsabschlüssen nicht nur Wissen und das Erlernen von Kompetenzen, sondern auch Motivation und verschiedene Arten von Ressourcen. Darüber hinaus endet der Kompetenzerwerb nicht mit dem Abschluss der formalen Bildung, sondern setzt sich im Laufe des Lebens fort, insbesondere durch Arbeitsmarkterfahrungen und Möglichkeiten des lebenslangen Lernens.

Unser Ziel ist es zu beschreiben, wie verschiedene Bildungsabschlüsse mit grundlegenden Kompetenzen in der Erwachsenenbevölkerung zusammenhängen, und ob sich dies in den vergangenen zehn Jahren geändert hat. Eine Fragestellung dabei ist auch, ob sich die Bildungsexpansion in steigenden Grundkompetenzen niederschlägt, oder ob eher eine Bildungsinflation zu beobachten ist (z.B. Meierkord et al., 2015). Wir gehen insbesondere auf den Zusammenhang zwischen Grundkompetenzen und Bildungsabschlüssen in Deutschland ein, ziehen aber auch den internationalen Vergleich heran, indem wir äquivalente Abschlüsse in verschiedenen Ländern vergleichen.

Für unsere Analysen verwenden wir Daten des Programme for the International Assessment of Adult Competencies (PIAAC). Bei den Analysen differenzieren wir zwischen Bildungsabschlüssen, die beruflich oder akademisch orientiert sind. Darüber hinaus adjustieren wir für Merkmale, wie Geschlecht, Alter, Zuwanderung und Erwerbsstatus, um den Zusammenhang zwischen Bildungsabschlüssen und Kompetenzen besser zu verstehen. In den internationalen Analysen vergleichen wir Abschlüsse, die anhand der International Standard Classification of Education (ISCED) 2011 vergleichbar gemacht wurden. Unsere

Untersuchung fokussiert auf die Daten des zweiten Zyklus von PIAAC (2023), wir vergleichen sie jedoch auch mit den Daten aus dem ersten Zyklus von PIAAC (2012).

Bei der Interpretation der Ergebnisse gehen wir auf die Unterschiede zwischen den verschiedenen Arten von Bildungsabschlüssen ein und identifizieren Gruppen, die besonders unter geringen Grundkompetenzen „leiden“. Darüber hinaus vergleichen wir Grundkompetenzen innerhalb von Bildungsgruppen über die Zeit, wodurch mögliche problematische Entwicklungen im Sinne einer Bildungsinflation sichtbar gemacht werden können. Außerdem beschreiben wir Kompetenzunterschiede zwischen Ländern innerhalb vergleichbarer Bildungsgruppen, wodurch mögliche Perspektiven zur Förderung von Grundkompetenzen im deutschen Bildungssystem angedeutet werden können.

Der Zusammenhang zwischen kognitiven Kompetenzen und den Big Five Persönlichkeitseigenschaften im internationalen Vergleich

Beatrice Rammstedt, Matthias Roth, Lena Roemer, Clemens Lechner

GESIS - Leibniz-Institut für Sozialwissenschaften, Deutschland

Kognitive Fähigkeiten und Persönlichkeitseigenschaften gehören zu den grundlegendsten Bereichen individueller Unterschiede (z.B. Anglim, 2022; Deary, 2012). In der Regel weisen sie nur geringe Zusammenhänge auf, mit einer geteilten Varianz von maximal 10% (z.B. Furnham et al., 2007). In Bezug auf die Big Five sind diese Zusammenhänge im Erwachsenenalter typischerweise am deutlichsten für Offenheit gefolgt von Emotionaler Stabilität (Anglim et al., 2022).

Auch wenn die Zusammenhänge zwischen kognitiven Fähigkeiten und Persönlichkeitseigenschaften seit bereits mehr als einem Jahrhundert beforscht werden, ist bisher nicht hinreichend geprüft, inwiefern diese Zusammenhänge durch nicht-repräsentative Stichproben - wie in psychologischen Studien üblich - geprägt sind. Unklar ist jedoch insbesondere ob diese Zusammenhänge defacto universell sind oder ob sie sich systematisch zwischen Kulturen unterscheiden.

Im zweiten Zyklus des Programme for the International Assessment of Adult Competencies (PIAAC) wurden erstmals neben kognitiven Fähigkeiten auch die Big Five Persönlichkeitsmerkmale als sozio-emotionale Fähigkeiten erhoben. Dies ermöglicht die oben beschriebenen Zusammenhänge basierend auf populationsrepräsentativen Daten zu untersuchen. Erstmals können durch den Vergleich der rund 30 teilnehmenden Länder auch kulturelle Unterschiede in den Zusammenhängen beleuchtet werden.

In dem Vortrag werden Zusammenhänge zwischen den kognitiven Kompetenzen Literacy und Numeracy und den Big Five sowohl aggregiert über alle teilnehmenden Länder wie auch vergleichend zwischen den Ländern dargestellt. Unter Berücksichtigung der Besonderheiten des PIAAC Datensatzes (z.B., Messung von Literacy und Numeracy als plausible values; Erzielen bevölkerungsrepräsentativer Ergebnisse durch die Verwendung von replicate weights) präsentieren wir bivariate Korrelationen zwischen kognitiven Kompetenzen mit den Big Five Persönlichkeitsmerkmalen.

Die Ergebnisse erlauben breit verallgemeinerbare Schlussfolgerungen dazu, wie kognitive Fähigkeiten und Persönlichkeitseigenschaften gemeinsam zentrale Aspekte menschlicher Individualität ausmachen.

OF-4: Special session: Care-Arbeit in der Wissenschaft: Herausforderungen und Perspektiven auf allen Karrierestufen

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 145

Chair der Sitzung: Meike Bonefeld

Chair der Sitzung: Merle Thielmann

In der Wissenschaft trifft intellektuelle Neugier auf die Realität der Vereinbarkeit von Care-Arbeit und Karriere. In dieser Podiumsdiskussion möchten wir Menschen, die auf unterschiedlichen Karrierestufen in der Wissenschaft tätig sind und Care-Verantwortung tragen, die Möglichkeit geben, offen über ihre Erfahrungen, Herausforderungen und Bewältigungsstrategien zu sprechen. Von Studierenden bis hin zu Professor*innen – das Podium vereint eine diverse Gruppe, die über verschiedene Lebensphasen hinweg einen Einblick in die Frage gibt, wie Care-Arbeit die wissenschaftliche Karriere beeinflusst und welche strukturellen Unterstützungen notwendig sind, um bessere Vereinbarkeit zu schaffen. Als interaktive Veranstaltung sind die Teilnehmenden herzlich eingeladen, bereits im Vorfeld Fragen an die Podiumsgäste einzureichen. Diese werden von der Moderation gesammelt und im Verlauf der Diskussion gezielt eingebracht, um die verschiedenen Perspektiven und Bedürfnisse der Zuhörer*innen aufzugreifen.

Die Podiumsdiskussion soll das Bewusstsein für die Bedeutung von Care-Tätigkeiten in der Wissenschaft schärfen, Perspektiven für eine bessere Vereinbarkeit aufzeigen und den Austausch zwischen den unterschiedlichen Karrierestufen fördern.

Teilnehmende:

- Kira Weber
- Hanna Beißert
- Benedict Fehring
- Nora Ries
- Carlotta Dilcher
- Merle Thielmann
- Meike Bonefeld

Care-Arbeit in der Wissenschaft: Herausforderungen und Perspektiven auf allen Karrierestufen

Meike Bonefeld

Universität Freiburg, Institut für Erziehungswissenschaft, Deutschland

In der Wissenschaft trifft intellektuelle Neugier auf die Realität der Vereinbarkeit von Care-Arbeit und Karriere. In dieser Podiumsdiskussion möchten wir Menschen, die auf unterschiedlichen Karrierestufen in der Wissenschaft tätig sind und Care-Verantwortung tragen, die Möglichkeit geben, offen über ihre Erfahrungen, Herausforderungen und Bewältigungsstrategien zu sprechen. Von Studierenden bis hin zu Professor*innen – das Podium vereint eine diverse Gruppe, die über verschiedene Lebensphasen hinweg einen Einblick in die Frage gibt, wie Care-Arbeit die wissenschaftliche Karriere beeinflusst und welche strukturellen Unterstützungen notwendig sind, um bessere Vereinbarkeit zu schaffen. Als interaktive Veranstaltung sind die Teilnehmenden herzlich eingeladen, bereits im Vorfeld Fragen an die Podiumsgäste einzureichen. Diese werden von der Moderation gesammelt und im Verlauf der Diskussion gezielt eingebracht, um die verschiedenen Perspektiven und Bedürfnisse der Zuhörer*innen aufzugreifen. Die Podiumsdiskussion soll das Bewusstsein für die Bedeutung von Care-Tätigkeiten in der Wissenschaft schärfen, Perspektiven für eine bessere Vereinbarkeit aufzeigen und den Austausch zwischen den unterschiedlichen Karrierestufen fördern.

OF-5: World Café: GEBF-AG zur sinkenden Teilnahmebereitschaft an Erhebungen im Schulkontext – Gemeinsame Diskussion der Leitlinien des künftigen GEBF-Positionspapiers

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · *Ort:* EO 159

Chair der Sitzung: Jutta von Maurice

Chair der Sitzung: Jasmin Decristan

Chair der Sitzung: Roman Auriga

Chair der Sitzung: Bayer Sonja

Chair der Sitzung: Karina Karst

Die Bereitschaft von Schulleitungen, Lehrkräften, Schülerinnen und Schülern sowie deren Eltern zur Teilnahme an wissenschaftlichen Studien sinkt zunehmend. Wenngleich die Studiendesigns der Mitglieder unserer Fachgesellschaft sehr unterschiedlich sind, bilden die im Schulkontext erhobenen Daten eine wesentliche Grundlage für ihre Forschung und die Genese empirischer Evidenz. Vor diesem Hintergrund hat der GEBF-Vorstand im Frühjahr 2023 die Arbeitsgruppe „Schulteilnahme“ eingerichtet, die es sich zur Aufgabe gemacht hat, die gegenwärtige Situation und möglichen Ursachen für eine sinkende Teilnahmebereitschaft schulischer Akteurinnen und Akteure eingehender zu beleuchten und hieran anknüpfend herauszuarbeiten, welche Handlungsoptionen für die Empirische Bildungsforschung bestehen und auf welchen Prinzipien das eigene Handeln beruhen sollte. Diese Positionierung dient nicht nur als interne Leitlinie der Fachgesellschaft, sondern auch als Grundlage für den Austausch mit anderen Fachgesellschaften, wissenschaftsnahen Gremien und politisch-administrativen Stakeholdern. In die Erstellung des Positionspapiers und die Formulierung der Leitlinien sind die Ergebnisse von zwei von der Arbeitsgruppe „Schulteilnahme“ im Frühjahr 2024 durchgeführten GEBF-Mitgliederbefragungen zu den Erhebungs- und Forschungspraktiken im schulischen Kontext eingeflossen. Ebenso berücksichtigt wurden die Ergebnisse einer Podiumsdiskussion im Rahmen der GEBF-Jahrestagung 2024.

Die Veranstaltung auf der GEBF-Jahrestagung 2025 gibt den Teilnehmenden die Möglichkeit, detaillierte Informationen zu den gegenwärtig sich in Bearbeitung befindlichen Leitlinien und deren Hintergründen zu erhalten und die hierin enthaltenen Ideen für Empfehlungen gemeinsam zu diskutieren. Einleitend werden dazu die Kernaussagen der von der Arbeitsgruppe vorgeschlagenen Leitlinien vorgestellt. Im Format eines World Cafés haben die Teilnehmenden anschließend die Möglichkeit, die Inhalte zu diskutieren, weiter auszudifferenzieren sowie fehlende Aspekte und Fokussierungen für die Finalisierung der Leitlinien einzubringen.

Die gegenwärtige Entwurfsfassung des Positionspapiers umfasst drei Leitlinien:

1. Schulische Akteurinnen und Akteure als Partnerinnen und Partner im Forschungsprozess verstehen Die erste Leitlinie umfasst vier Aspekte, die sich auf (1) das Rollenverständnis der Forschenden im Rahmen von Erhebungen im Schulkontext, (2) den wechselseitigen Mehrwert und die Nützlichkeit von Datenerhebungen und datenbasierten Befunden für die Schulen, (3) die Erfordernis einer transparenten Kommunikation über die Studie und die damit verbundenen Datenerhebungen sowie (4) die Notwendigkeit ökonomischer und mit dem Schulalltag abgestimmter Datenerhebungen beziehen.
2. Zugang zu und Nutzung von vorhandenen Forschungsdaten stärken Die zweite Leitlinie bezieht sich auf eine stärkere Nutzung von Sekundärdaten und die Selbstverpflichtung zur Bereitstellung eigener Daten. Damit ist auch die Honorierung der Erhebung und Bereitstellung von Primärdaten als wissenschaftliche Leistungen bspw. über die Zitation von Sekundärdaten oder bei der Bewertung von Forschungsleistungen verknüpft. Insbesondere wird empfohlen, die Notwendigkeit von Datenerhebungen im Rahmen studentischer Qualifikationsarbeiten kritisch zu hinterfragen und diese auf ein Minimum zu reduzieren, um die Erhebungslast im Schulkontext zu reduzieren.
3. Genehmigungsverfahren professionell handhaben Die dritte Leitlinie setzt den Fokus auf die administrativen Abläufe vor der Studiendurchführung. So sollten Forschende im Rahmen der Anmeldung von Studien bei den Kultusministerien die unterschiedlichen Bedarfe, Logiken und Verantwortlichkeiten auch der prüfenden Stellen kennen und Unterlagen so einreichen, dass die inhaltliche und die datenschutzrechtliche Prüfung für alle Beteiligten möglichst effizient erfolgen können.

Die Ergebnisse des World Cafés fließen in die Überarbeitung der Leitlinien ein. Abschließend wird auch das weitere (fach-)politische Vorgehen diskutiert.

GEBF-AG zur sinkenden Teilnahmebereitschaft an Erhebungen im Schulkontext – Gemeinsame Diskussion der Leitlinien des künftigen GEBF-Positionspapiers

Jutta von Maurice¹, Jasmin Decristan², Roman Auriga¹, Sonja Bayer³, Karina Karst⁴

¹Leibniz-Institut für Bildungsverläufe, Bamberg; ²Bergische Universität Wuppertal; ³DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; ⁴Universität Mannheim

Die Bereitschaft von Schulleitungen, Lehrkräften, Schülerinnen und Schülern sowie deren Eltern zur Teilnahme an wissenschaftlichen Studien sinkt zunehmend. Wenngleich die Studiendesigns der Mitglieder unserer Fachgesellschaft sehr unterschiedlich sind, bilden die im Schulkontext erhobenen Daten eine wesentliche Grundlage für ihre Forschung und die Genese empirischer Evidenz. Vor diesem Hintergrund hat der GEBF-Vorstand im Frühjahr 2023 die Arbeitsgruppe „Schulteilnahme“ eingerichtet, die es sich zur Aufgabe gemacht hat, die gegenwärtige Situation und möglichen Ursachen für eine sinkende Teilnahmebereitschaft schulischer Akteurinnen und Akteure eingehender zu beleuchten und hieran anknüpfend herauszuarbeiten, welche Handlungsoptionen für die Empirische Bildungsforschung bestehen und auf welchen Prinzipien das eigene Handeln beruhen sollte. Diese Positionierung dient nicht nur als interne Leitlinie der Fachgesellschaft, sondern auch als Grundlage für den Austausch mit anderen Fachgesellschaften, wissenschaftsnahen Gremien und politisch-administrativen Stakeholdern. In die Erstellung des Positionspapiers und die Formulierung der Leitlinien sind die Ergebnisse von zwei von der Arbeitsgruppe „Schulteilnahme“ im Frühjahr 2024 durchgeführten GEBF-Mitgliederbefragungen zu den Erhebungs- und Forschungspraktiken im schulischen Kontext eingeflossen. Ebenso berücksichtigt wurden die Ergebnisse einer Podiumsdiskussion im Rahmen der GEBF-Jahrestagung 2024.

Die Veranstaltung auf der GEBF-Jahrestagung 2025 gibt den Teilnehmenden die Möglichkeit, detaillierte Informationen zu den gegenwärtig sich in Bearbeitung befindlichen Leitlinien und deren Hintergründen zu erhalten und die hierin enthaltenen Ideen für Empfehlungen gemeinsam zu diskutieren. Einleitend werden dazu die Kernaussagen der von der Arbeitsgruppe vorgeschlagenen Leitlinien vorgestellt. Im Format eines World Cafés haben die Teilnehmenden anschließend die Möglichkeit, die Inhalte zu diskutieren, weiter auszudifferenzieren sowie fehlende Aspekte und Fokussierungen für die Finalisierung der Leitlinien einzubringen.

Die gegenwärtige Entwurfsfassung des Positionspapiers umfasst drei Leitlinien:

1. Schulische Akteurinnen und Akteure als Partnerinnen und Partner im Forschungsprozess verstehen Die erste Leitlinie umfasst vier Aspekte, die sich auf (1) das Rollenverständnis der Forschenden im Rahmen von Erhebungen im Schulkontext, (2) den wechselseitigen Mehrwert und die Nützlichkeit von Datenerhebungen und datenbasierten Befunden für die Schulen, (3) die Erfordernis einer transparenten Kommunikation über die Studie und die damit verbundenen Datenerhebungen sowie (4) die Notwendigkeit ökonomischer und mit dem Schulalltag abgestimmter Datenerhebungen beziehen.

2. Zugang zu und Nutzung von vorhandenen Forschungsdaten stärken Die zweite Leitlinie bezieht sich auf eine stärkere Nutzung von Sekundärdaten und die Selbstverpflichtung zur Bereitstellung eigener Daten. Damit ist auch die Honorierung der Erhebung und Bereitstellung von Primärdaten als wissenschaftliche Leistungen bspw. über die Zitation von Sekundärdaten oder bei der Bewertung von Forschungsleistungen verknüpft. Insbesondere wird empfohlen, die Notwendigkeit von Datenerhebungen im Rahmen studentischer Qualifikationsarbeiten kritisch zu hinterfragen und diese auf ein Minimum zu reduzieren, um die Erhebungslast im Schulkontext zu reduzieren.

3. Genehmigungsverfahren professionell handhaben Die dritte Leitlinie setzt den Fokus auf die administrativen Abläufe vor der Studiendurchführung. So sollten Forschende im Rahmen der Anmeldung von Studien bei den Kultusministerien die unterschiedlichen Bedarfe, Logiken und Verantwortlichkeiten auch der prüfenden Stellen kennen und Unterlagen so einreichen, dass die inhaltliche und die datenschutzrechtliche Prüfung für alle Beteiligten möglichst effizient erfolgen können.

Die Ergebnisse des World Cafés fließen in die Überarbeitung der Leitlinien ein. Abschließend wird auch das weitere (fach-)politische Vorgehen diskutiert.

PS-32: Bildungsdisparitäten

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 382

Chair der Sitzung: Katharina Fink

Qualitätsaspekte multiprofessioneller Kooperation zugunsten der Förderung sozial bildungsbenachteiligter Schüler:innen mit Migrationshintergrund: Konkretisierung (inter-)nationaler wissenschaftlicher Impulse für die Ganztagsgrundschulpraxis in Deutschland

Haigün Ning¹, Jule Schmidt¹, Nanine Lilla¹, Jan Willem Nieuwenboom², Marianne Schüpbach¹

¹Freie Universität Berlin, Deutschland; ²Fachhochschule Nordwestschweiz, Schweiz

Theoretischer Hintergrund

Mit dem Ausbau von Ganztagsgrundschulen in Deutschland sind Erwartungen verbunden, insbesondere hinsichtlich der Förderung von Schüler:innen mit diversen sozialen und kulturellen Hintergründen, die häufig von Bildungsbenachteiligung betroffen sind (Kielblock, 2021). Diese Erwartungen wurden bisher jedoch nicht ausreichend erfüllt. Multiprofessionelle Kooperation, das meint die Zusammenarbeit von Angehörigen von zwei oder mehreren verschiedenen Professionsgruppen, z.B. die Zusammenarbeit von Lehrkräften, Erzieher:innen, Sozialarbeiter:innen, und Psycholog:innen, wird als wichtiges Qualitätsmerkmal inklusiver Förderung von allen Schüler:innen im Ganztag gesehen (z.B. Schüpbach, Lilla & Groh, 2020). Bisherige Forschung im deutschen Kontext beschäftigt sich vor allem mit Gelingensbedingungen multiprofessioneller Kooperation (Schüpbach et al., 2024) sowie mit verschiedenen Aspekten der aktuellen Kooperationspraxis im Allgemeinen (Hochfeld & Rothland, 2022). Wie die multiprofessionelle Kooperation spezifisch zugunsten der Förderung von bildungsbenachteiligten Schüler:innen gestaltet werden sollte, ist noch unzureichend erforscht.

Im internationalen Kontext finden sich vergleichbare erweiterte Bildungsangebote (engl.: extended education), die sich als wirksam für die Förderung benachteiligter Schüler:innen erweisen (Durlak et al., 2010; Bae et al., 2010). Auch zur multiprofessionellen Kooperation in der erweiterten Bildung liegen internationale empirische Forschungsergebnisse vor, insbesondere zur Frage, wie multiprofessionelle Kooperation zugunsten der Förderung von „Risikokindern“ gestaltet werden kann (z.B. Olsson et al., 2023). Dieser Beitrag zielt deshalb darauf ab, empirisches Wissen aus internationalen Kontexten für deutsche Ganztagsgrundschulen nutzbar zu machen, damit das Potenzial multiprofessioneller Kooperation zur Förderung sozial bildungsbenachteiligter Schüler:innen mit Migrationshintergrund besser ausgeschöpft werden kann.

Fragestellung

Inwiefern lassen sich die von (inter)nationalen Expert:innen genannten Qualitätsaspekte multiprofessioneller Kooperation nach Einschätzung von Praktiker:innen an deutschen Ganztagsgrundschulen umsetzen?

Methode

Zur Erfassung der Einschätzung von (inter)nationalen Bildungsforscher:innen zur qualitativ hochwertigen Gestaltung erweiterter Bildung zugunsten der Förderung bildungsbenachteiligter Schüler:innen wurden problemzentrierte Experteninterviews (Witzel, 2000) durchgeführt (n=30). Aussagen zu Qualitätsaspekten multiprofessioneller Kooperation wurden daraufhin in Fokusgruppen (Zwick & Schröter, 2012) mit 14 offenen Ganztagsgrundschulen aus vier deutschen Bundesländern diskutiert mit dem Ziel, einen Konsens zur möglichen Umsetzung dieser Aspekte in der Praxis zu finden. In den Fokusgruppen waren verschiedene Rollen vertreten, darunter Schulleitungen, Leitungen der außerunterrichtlichen Angebote, Lehrkräfte, Pädagog:innen und Eltern.

Die anonymisierten Transkripte der Experteninterviews und Fokusgruppen wurden mithilfe der strukturierenden Inhaltsanalyse (Kuckartz, 2018), unterstützt durch MAXQDA, ausgewertet. Zuerst wurden die Daten deduktiven Kategorien zugeordnet, die auf Basis des aktuellen Forschungsstandes entwickelt wurden. Anschließend wurden induktiv Subkategorien gebildet, die konkrete Ideen zur möglichen Umsetzung umfassen.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse zeigen, dass die von den (inter)nationalen Bildungsforscher:innen genannten Qualitätsaspekte multiprofessioneller Kooperation zugunsten der Förderung sozial bildungsbenachteiligter Schüler:innen mit Migrationshintergrund an den befragten deutschen Ganztagsgrundschulen unterschiedlich vorhanden sind. Zur Konkretisierung dieser Aspekte in der Praxis erfolgt derzeit eine schulspezifische Auswertung der Aussagen aus den Fokusgruppen. Zunächst kann folgender Konsens festgehalten werden: **1) Zusammensetzung und Zuständigkeit:** Erzieher:innen mit vielfältigen sprachlichen und kulturellen Hintergründen können als Bindeglied zwischen Unterricht und außerunterrichtlichen Angeboten agieren. Eine Steuergruppe aus verschiedenen Professionen sollte bereits in der Planungsphase neuer Bildungsangebote einbezogen werden. **2) Gelegenheit und Inhalt:** Gemeinsame Konzeption und Durchführung außerunterrichtlicher Angebote, in denen die ganzheitliche Entwicklung der Schüler:innen mit diversen Hintergründen im Fokus steht. Insbesondere die Kooperation mit Eltern aus sozial benachteiligten Hintergründen wird als ein Bereich multiprofessioneller Kooperation betont. **3) Format und Tool:** Eine Mischung von einfachen und komplexen Formaten der Kooperation wird bevorzugt, wobei der Bedarf des Einsatzes digitaler und niedrigschwelliger Tools zur effektiven Kommunikation und Übergabe hervorgehoben wird. **4) Zwischenmenschliche Ebene:** Die Anerkennung und Wertschätzung verschiedener fachlicher Perspektiven sowie sprachlicher und kultureller Hintergründe sollte in einem diversitätssensiblen Kooperationssetting gestärkt werden, z.B. durch vermehrtes Teambuilding, das zum besseren persönlichen Kennenlernen des Personals beiträgt.

Die Ergebnisse verdeutlichen die Komplexität eines Forschungsansatzes, der versucht, (inter)nationale wissenschaftliche Erkenntnisse in der deutschen Praxis umzusetzen. Dennoch lassen sich wichtige Implikationen für die Förderung von Chancengerechtigkeit durch diversitätssensible multiprofessionelle Kooperation im Ganztag ableiten.

Auswirkungen von (geschlechtsstereotypen) Lehrkraftüberzeugungen in den MINT-Fächern auf die Schüler*innen Überzeugungen

Katharina Fink^{1,3}, Luke McGuire⁴, Kelly Lynn Mulvey⁵, Hanna Beißert^{1,2,3}

¹DIPF| Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ²Goethe Universität Frankfurt am Main;

³DeA - Center for Research on Individual Development and Adaptive Education of Children at Risk; ⁴Department of Psychology, University of Exeter, Exeter, United Kingdom; ⁵Department of Psychology, North Carolina State University, Raleigh, NC, United States

Im MINT-Bereich lässt sich eine erhebliche Differenz der Geschlechterverteilung finden (Brotman & Moore, 2008; Makarova et al., 2019). In vergangenen Studien konnte als ein möglicher Grund dafür vorhandene Geschlechterstereotype identifiziert werden

(Wang & Degol, 2017). Verschiedene Studien haben gezeigt, dass Lehrkräfte in den MINT-Fächern immer noch geschlechtsspezifische Vorurteile und Stereotypen haben (Beißert, McGuire & Mulvey, 2022) und sich diese auf die Lehrkraft-Schüler*innen-Interaktionen auswirken (Sarouphin & Chartouny, 2016). Schüler*innen nehmen solche Unterschiede in Einstellungen und Verhalten der Lehrkräfte wahr (Gunderson et al., 2012) und neigen dazu, diese Einstellungen vor allem von gleichgeschlechtlichen Lehrkräften zu übernehmen (Bussey et al., 1984). Dabei ist weiterhin davon auszugehen, dass Lehrkräfte durch ihr Verhalten in der Schule und ihren Umgang mit den Schülerinnen und Schülern ihre eigenen (möglicherweise stereotypen) Überzeugungen auf die Kinder übertragen und somit zur Entstehung und Aufrechterhaltung von stereotypen Sichtweisen bei den Kindern beitragen können. Dies konnte bereits im Zusammenhang von Geschlechtsstereotypen in Bezug auf Mathematikfähigkeit bei Mädchen gezeigt werden (Beilock et al., 2010).

Die nachfolgende Studie knüpft hier an und untersucht für vier MINT-Domänen (Mathematik, Informatik, Naturwissenschaft, Technik), inwiefern mögliche stereotype Überzeugungen von Lehrkräften mit denen ihrer Schüler*innen in Verbindung stehen.

Die Analysen wurden mit einer vorläufigen Stichprobe mit insgesamt 390 Schüler*innen der 4. Klasse (203 weiblich, 187 männlich, $M_{Alter} = 9.42$, $SD = .83$) aus Baden-Württemberg und Hessen und deren Klassenlehrkräfte ($N = 24$, 18 weiblich, 6 männlich, $M_{Alter} = 48.28$, $SD = 11.07$) durchgeführt. Anhand eines Fragebogens wurden die erwarteten Fähigkeiten der beiden Geschlechter in den verschiedenen MINT-Domänen von Lehrkräften und Schüler*innen durch eine selbstgenerierte Skala erfasst.

Die vorläufige Mehrebenenanalyse zeigt, dass die Fähigkeitserwartungen der Schüler*innen in Mathematik weder bei Jungen noch bei Mädchen durch die der Lehrkraft vorhergesagt werden. In Informatik konnte eine signifikante Interaktion zwischen Fähigkeitserwartungen von Lehrkraft und Schülern ($t(141.15) = 2.48$, $p = 0.014$) sowie Schülerinnen ($t(0.017) = 2.63$, $p = 0.009$) gefunden werden. Ähnliche Ergebnisse lassen sich in der Technik-Domäne finden. Auch dort fand sich eine signifikante Interaktion der Fähigkeitserwartungen von Lehrkräften und der Erwartung von Schülern ($t(156) = -4.99$, $p > 0.001$) bzw. Schülerinnen ($t(167.44) = -6.63$, $p > 0.001$). Dabei ist das Geschlecht der Lehrkraft ($t(156) = -2.53$, $p = 0.01$) und die Fähigkeitserwartungen der Lehrkraft ($t(156) = -4.97$, $p > 0.001$) ein signifikanter Prädiktor für die Erwartungen der Jungen. Auch in den Naturwissenschaften zeigt sich eine signifikante Interaktion zwischen Lehrkrafte Erwartungen und Erwartungen der Mädchen ($t(168.27) = -7.65$, $p > 0.001$) und den Erwartungen der Jungen ($t(150.24) = -6.23$, $p > 0.001$). Die Lehrkrafte Erwartungen sind außerdem ein signifikanter Prädiktor der Erwartungen der Mädchen ($t(10.88) = -3.60$, $p = 0.004$) und der Jungen ($t(66.25) = -5.18$, $p > 0.001$) in dieser Domäne.

Insgesamt weisen die Ergebnisse darauf hin, dass die Überzeugungen der Lehrkraft sich nicht in Mathematik auf die der Schüler*innen auswirken, jedoch eine stereotype Lehrkrafte Erwartung der Fähigkeiten der Jungen in den Naturwissenschaften und Technik, die Fähigkeitseinschätzungen der Schüler beeinflusst. Sind die Fähigkeitserwartungen für Mädchen in Naturwissenschaften und Technik von den Lehrkräften in einer Klasse höher, werden diese zu einem weniger starken Prädiktor als in Klassen, wo Lehrkräfte, die Fähigkeiten der Jungen höher einschätzen. Werden die Fähigkeiten der Mädchen von den Lehrkräften in Informatik in einer Klasse schlechter eingeschätzt, so sind die Auswirkungen der Lehrkrafteinschätzung auf die der Kinder jedoch stärker. Die Ergebnisse zeigen ebenfalls, dass Jungs die Fähigkeiten von Mädchen besser einschätzen, wenn sie eine weibliche Lehrkraft haben. Die Ergebnisse machen deutlich, dass die Überzeugungen von Lehrkräften, die Überzeugungen der Schüler*innen in einigen MINT-Fächern beeinflussen und daher im Unterricht besonders darauf geachtet werden sollte, was man den Schüler*innen verbal und nonverbal über die Fähigkeiten bestimmter Gruppen vermittelt.

Noten-Bias im Klassenzimmer: Wie Geschlecht, Sprachhintergrund und Klassenleistung mit der Benotung der Schülerleistung durch Lehrkräfte zusammenhängen

Chantal Oggenfuss¹, Stefan C. Wolter^{1,2}

¹Schweizerische Koordinationsstelle für Bildungsforschung, Schweiz; ²Forschungsstelle für Bildungsökonomie, Universität Bern

Die Bewertung der schulischen Leistungen durch Lehrkräfte ist entscheidend für die Bildungsverläufe und beeinflusst den Zugang zu weiterführenden Schulen, Universitäten und dem Arbeitsmarkt. Frühere Forschungen haben gezeigt, dass die Erwartungen von Lehrkräften mit sozialen Stereotypen zusammenhängen, was die Benotung der Schülerinnen und Schüler nach Geschlecht und sprachlichem Hintergrund beeinflussen kann (z. B. Rosenthal & Jacobsen 1968, Rubie-Davies, 2007). Auch aktuelle Studien weisen darauf hin, dass Stereotype die Erwartungen und Bewertungen von Lehrkräften beeinflussen (vgl. Geven et al., 2021; Muntoni & Retelsdorf, 2018; Gentrup & Rjosk, 2018; Kleen & Glock, 2018). Die Mehrheit der Studien hat den Zusammenhang zwischen dem Geschlecht und der Benotung der Schülerinnen und Schüler durch die Lehrkräfte untersucht. Dabei zeigt sich die Tendenz, dass Mädchen in Sprachfächern (für einen Überblick siehe Protivínský & Münich, 2018; neuere Erkenntnisse: Lavy & Sand, 2018; Contreras, 2023; Olczyk et al., 2023) und Jungen in mathematischen Fächern (Hofer, 2015; Lavy & Sand, 2018; Olczyk et al., 2023) tendenziell überbewertet werden. Während sich die bestehende Literatur zur Bewertung der Schülerleistungen durch Lehrkräfte meist auf einzelne Fächer oder isolierte individuelle Merkmale konzentriert, haben nur wenige Studien mehrere Faktoren gleichzeitig untersucht (bspw. Lavy & Sand, 2018). Wir analysieren, wie Geschlecht, sprachlicher Hintergrund und das Leistungsniveau der Klasse (siehe dazu Sahlström & Silliman, 2024) mit den Bewertungen der Lehrkräfte in Mathematik und Sprache zusammenhängen, und tragen zur bestehenden Forschung bei, indem wir diese individuellen Faktoren gleichzeitig berücksichtigen und in zwei Fächern untersuchen. Dadurch bietet unsere Forschung ein differenzierteres Verständnis der komplexen Effekte, die zur Über- oder Unterschätzung der Schülerleistungen führen können. Für unsere Studie nutzen wir Daten aus einer groß angelegten Erhebung von Schülern der Sekundarstufe I. Diese standardisierten Leistungstests werden bereits seit zehn Jahren durchgeführt und wurden entwickelt, um die im Lehrplan definierten Ziele zu bewerten. Wir kombinieren diesen Datensatz mit den Zeugnisnoten.

Wir schätzen Regressionsmodelle, um zu analysieren, wie Geschlecht, sprachlicher Hintergrund und das Leistungsniveau der Klasse mit der Notengebung zusammenhängen. Dabei verwenden wir verschiedene Modelle, die nacheinander Testergebnisse, Geschlecht, sprachlichen Hintergrund und Klassenleistung als erklärende Variablen einbeziehen, um den umfassenden Zusammenhang zu verstehen. Wir testen auch die Interaktion zwischen Schülermerkmalen und Testergebnissen, um festzustellen, ob bestimmte Untergruppen proportional besser benotet werden. In allen Modellen berücksichtigen wir Kohorten- und Schuleffekte, um unbeobachtete Einflüsse zu kontrollieren. Durch die Verwendung klassenbasierter Cluster-robuster Standardfehler können wir Korrelationen innerhalb der Klassen berücksichtigen und die Robustheit der Ergebnisse gegenüber Datenheterogenität und Modellfehlspezifikationen erhöhen. Zur Überprüfung der Robustheit der Ergebnisse schätzen wir ein zusätzliches Modell basierend auf den standardisierten Residuen.

Insgesamt zeigen unsere Ergebnisse signifikante Verzerrungen in den Bewertungen der Lehrkräfte sowohl in Deutsch als auch in Mathematik. Im Einklang mit früheren Forschungen finden wir in den Sprachfächern Biases zugunsten von Mädchen und nicht-fremdsprachigen Schülerinnen und Schülern. Der Effekt des Geschlechts entspricht etwa 38 % des Effekts einer

Standardabweichung bei den Testergebnissen. Der Effekt, nicht-fremdsprachig zu sein, ist ebenfalls positiv und signifikant und entspricht etwa 22 % des Effekts einer Standardabweichung bei den Testergebnissen. Beispielsweise bedeutet dies, dass ein nicht-fremdsprachiger Schüler mit gleichen Testergebnissen wie ein fremdsprachiger Schüler eine um 0,104 Notenpunkte bessere Bewertung erhält. Gleichzeitig zeigt sich eine Unterschätzung der Schülerinnen und Schüler, wenn die durchschnittliche Klassenleistung steigt – sowohl in Mathematik als auch in Sprache. In Mathematik gibt es keine signifikante Verzerrung in Bezug auf das Geschlecht oder den sprachlichen Hintergrund, mit der Ausnahme der Überbewertung von Mädchen in leistungsstärkeren Klassen. Es gibt auch keinen signifikanten Bias gegen nicht-fremdsprachige Schülerinnen und Schüler, was plausibel ist, wenn das Stereotyp der Sprachschwierigkeiten bei fremdsprachigen Schülerinnen und Schülern in Mathematik schwächer ausgeprägt ist. Unsere Robustheitsprüfungen bestätigen diese Muster, mit Ausnahme der Überschätzung von Mädchen in leistungsstärkeren Klassen, die nicht konsistent beobachtet wurde. Dies entspricht den gemischten Ergebnissen früherer Forschungen zum Geschlechterbias in Mathematik.

Soziale Disparitäten in der Entwicklung digitaler Kompetenzen während der Sekundarstufe I

Yannick Stelter

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Hintergrund: Digitale Kompetenzen gelten als eine der Schlüsselkompetenzen für das 21. Jahrhundert (OECD 2015), u.a. relevant für Bildungserfolg (Robinson et al. 2018; Pagani et al. 2016) und Arbeitsmarkterfolg (Falck et al. 2021). Vor dem Hintergrund erheblicher herkunftsspezifischer Unterschiede, zugunsten sozial privilegierter Kinder (Eickelmann et al. 2019), tragen diese Disparitäten potentiell zu einer Verschärfung bestehender Bildungsungleichheiten bei. Dabei ist wenig bekannt über die Ursachen sozialer Ungleichheit beim Erwerb digitaler Kompetenzen, deren Entwicklung über den Bildungsverlauf und die Rolle der Schule für diesen Prozess.

Theorie: Nach der aktuellen Forschung handelt es sich beim Erwerb digitaler Kompetenzen um einen sozialen Lernprozess, der überwiegend in der Familie stattfindet (Selwyn 2005; Senkbeil 2023). Unterschiede in den digitalen Kompetenzen sind dabei das Resultat unterschiedlicher Lerngelegenheiten und Förderung innerhalb der Familie (Fraillon et al. 2020) – welche mit der Bedeutung digitaler Kompetenzen für den Statuserhalt zusammenhängen können (Witte & Mannon 2010).

Die Bedeutung der Schule für den Erwerb digitaler Kompetenzen und den darin bestehenden sozialen Disparitäten, stand bislang selten im Fokus der Forschung. Dabei gilt insbesondere in Deutschland die institutionelle Ausgestaltung des Bildungssystems als wichtig für die Entwicklung sozialer Ungleichheit im Kompetenzerwerb (Maaz et al. 2008).

Mit Blick auf die Frage, wie sich soziale Ungleichheiten in der Entwicklung digitaler Kompetenzen verändern sind konzeptionell drei Verlaufstypen denkbar (Dumont & Ready 2020): Persistenz, Zunahme oder Abnahme sozialer Ungleichheit. Aus Perspektive der Statuserhalttheorie (Boudon, 1974) wäre aufgrund unterschiedlicher familiärer Lerngelegenheiten über den gesamten Bildungsverlauf eine Vergrößerung sozialer Unterschiede zu erwarten (DiPrete & Eirich 2006). Im Gegensatz dazu, wäre mit Blick auf die schulischen Lerngelegenheiten (Hallinan 1994) zu erwarten, dass sich die Entwicklungsverläufe innerhalb der Bildungsgänge über die Zeit zunehmend angleichen und soziale Unterschiede verringern (Schneider & Linberg 2021).

Wenngleich innerhalb der Bildungsgänge unterschiedliche Entwicklungen stattfinden können, wäre als Folge der sozial selektiven Aufteilung der Schüler*innen nach der vierten Klasse in verschiedene Bildungsgänge davon auszugehen, dass aufgrund bildungsgangspezifische Entwicklungsverläufe die sozialen Unterschiede in den digitalen Kompetenzen insgesamt zunehmen (Neumann et al. 2014; Baumert et al. 2012).

Fragestellung: Vor dem Hintergrund des Spannungsfeldes familiärer und schulischer Einflüsse auf die Entwicklung sozialer Disparitäten in digitalen Kompetenzen während der Sekundarstufe I, analysiert der vorliegende Beitrag die Entwicklung sozialer Ungleichheiten in digitalen Kompetenzen in Abhängigkeit des Bildungsganges. Damit gehen wir auch der Frage nach, inwieweit sich Hinweise auf mögliche ungleichheitsverstärkende oder ungleichheitsreduzierende Wirkung der Schule im Kontext digitaler Kompetenzen finden lassen. Der Beitrag gliedert sich dabei in zwei Teilfragen:

RQ 1: Wie entwickeln sich die digitalen Kompetenzen von Schüler*innen verschiedener sozialer Herkunft insgesamt?

RQ 2: Wie entwickeln sich die digitalen Kompetenzen von Schüler*innen verschiedener sozialer Herkunft zwischen und innerhalb der Bildungsgänge?

Daten & Methode: Datenbasis bilden N=1846 Schüler*innen der Startkohorte 3 des Nationalen Bildungspanels (NEPS-SC3; Blossfeld & Roßbach 2019). Zunächst analysieren wir deskriptiv die Entwicklung der Mittelwerte von Schüler*innen unterschiedlicher sozialer Herkunft in Hinblick auf die oben beschriebenen Prozesse. Im Anschluss analysieren wir - getrennt für gymnasiale vs. nicht-gymnasiale Schulformen - mittels latenten Change-score Modellen (Köhler et al. 2021) die individuelle Entwicklung digitaler Kompetenzen von Klasse 6 zu 9 in Abhängigkeit des SES und Bildungsgangs.

Vorläufige Ergebnisse: Die ersten Ergebnisse deuten darauf hin, dass sich insgesamt die sozialen Disparitäten in digitalen Kompetenzen über den Verlauf der Sekundarstufe I leicht zu verstärken scheinen. Dabei zeigen sich vor allem Entwicklungsunterschiede zwischen den Bildungsgängen. Innerhalb der Bildungsgänge deuten sich an nicht gymnasialen Bildungsgängen ungleichheitsvergrößernde Prozess an, während sich am Gymnasium tendenziell kompensatorische Entwicklungen zeigen. Damit weisen die Ergebnisse auf die Bedeutung soziale Ungleichheiten kontextspezifisch zu analysieren. Die Entwicklung sozialer Ungleichheit bei digitalen Kompetenzen scheint dabei das Resultat des Zusammenspiels individueller als auch von Faktoren zu sein, die eng mit der Aufteilung der Schüler*innen zusammenhängen

PS-33: Digitale Medien: Herausforderungen für die Entwicklung von Jugendlichen im digitalen Zeitalter

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SO 422

Chair der Sitzung: Mareike Thumel

„Sense of Agency“ und Soziale Medien. Wie nehmen Jugendliche eigene Agency wahr, wenn sie Soziale Medien nutzen?

Jane Müller, Mareike Thumel, Julia Eichelmann

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Der Begriff der digitalen Souveränität wird in den letzten Jahren immer häufiger verwendet. Souveränität zielt in diesem Zusammenhang entweder auf Staaten(-verbünde), Organisationen oder Personen. Die Verwobenheit der Ebenen und die Beteiligung diverser Akteur*innen werden immer wieder betont (Gräf et al., 2018; Pohle und Thiel, 2020). Verschiedene systematische Literaturreviews zeigen eine Vielfalt unterschiedlicher Begriffsverwendungen (Couture und Toupin, 2019; Lambach und Oppermann, 2023; Müller et al., 2022a; auch Hummel et al., 2021). Im Zusammenhang mit Personen erfordert die Verwendung digitaler Souveränität eine begriffliche Schärfung. In der pädagogischen Auseinandersetzung taucht Souveränität nur selten auf, knüpft aber an klassische Bildungsziele von Selbstbestimmung, Autonomie oder Freiheit an (Klafki, 2007; ausführlicher hierzu Müller und Kammerl, 2022). Daran angelehnt wird auch der Begriff der digitalen Souveränität häufig als „neues“ Bildungsziel angenommen, dabei jedoch von etablierten und systematisch ausgearbeiteten Konzepten wie Medienkompetenz (Hugger, 2022) oder Medienbildung (Bettinger und Jörissen, 2022) nicht ausreichend abgegrenzt. Im Forschungsprojekt Digitale Souveränität Jugendlicher wird der Begriff demgegenüber als Analysezugang genutzt, um die Beziehungen zwischen Akteur*innen in einem tiefgreifend mediatisierten Alltag (Hepp 2018) in den Blick zu nehmen. Besonders aussichtsreich für Fragen digitaler Souveränität kann dabei der Blick auf Fragen von Agency sein: Wer oder was kann in einer gegebenen Situation eigene Interessen wirksam verfolgen? Der englischsprachige Begriff der *Agency* verweist auf eine umfassende Diskussion rund um einen vielschichtigen Phänomenbereich (Emirbayer und Mische, 1998). Dennoch greift keine der gebräuchlichen Übersetzungen ins Deutsche das Konzept in angemessener Breite auf (vgl. auch Hepp et al., 2022). Agency soll hier angelehnt an Emirbayer und Mische (1998) verstanden werden als die situativ und relational immer wieder neu ausgehandelte Beteiligung in unterschiedlichen temporär strukturierten Kontexten, um den historisch erwachsenden Herausforderungen des Zusammenlebens zu begegnen (Emirbayer und Mische, 1998, S. 970; vgl. auch Raitelhuber, 2012). In einer umfangreichen explorativen Untersuchung mit 28 Jugendlichen im Alter von 14 bis 19 Jahren (Müller, 2024) konnte die Konstruktion von Agency als relevante Bezugsgröße in der Alltagsgestaltung in einer tiefgreifend mediatisierten Gesellschaft herausgearbeitet werden (Müller, in Druck). Im Anschluss daran, wurden in einer repräsentativen Befragung deutschsprachiger Jugendlicher zwischen 14 und 19 Jahren (n=616; Repräsentativ in Bezug auf Alter, Geschlecht, besuchte Schulform/Schulabschluss, Bundesland) die Zusammenhänge zwischen verschiedenen Differenzvariablen (Alter, Geschlecht, Bildungshintergrund, Herkunftsland, Muttersprache) und der wahrgenommenen Agency bei der Benutzung Sozialer Medien in den Blick genommen. Neben soziodemographischen Merkmalen wurde zudem der Zusammenhang zur genutzten Anwendung (Snapchat, Instagram, YouTube, TikTok), der Art der Benutzung, der Sorge um die eigene Privatsphäre und der Datenkompetenz der Jugendlichen geprüft. Die Erhebung der Daten ist seit ca. einer Woche abgeschlossen und die Analyse erfolgt in den kommenden zwei Monaten, so dass zur Tagung die Ergebnisse vollständig vorliegen werden.

Angsterleben, Einstellungen gegenüber und Erfahrungen mit Cybermobbing bei Primarschüler:innen – Ergebnisse eines Mediationsmodells

Arvid Nagel¹, Felix Kruse^{1,2}

¹PH NMS Bern, Schweiz; ²Pädagogische Hochschule St.Gallen

Interaktive Medien nehmen innerhalb der Lebenswirklichkeiten von Kindern, neben der Familie, Schule und Peers, eine zentrale Sozialisationsinstanz ein (Davis, 2013; Falikman, 2021). Digitale soziale Netzwerke, Smartphones und Computerspiele zählen mittlerweile zu den beliebtesten Freizeitaktivitäten von Kindern und Jugendlichen (Suter et al., 2023). Die JIM-Studie (für Deutschland; Feierabend et al., 2023) und die JAMES-Studie (für die Schweiz; Külling et al., 2022) bestätigen einen stetigen Anstieg der Ausstattungsquoten mit Mediengeräten im Kindes und Jugendalter. Entsprechend zeigt sich basierend auf dieser Gegebenheit auch ein verändertes Mediennutzungsverhalten bei jungen Menschen. In der Schweiz verbringen Kinder und Jugendliche im Alter von 11 bis 15 Jahren durchschnittlich 4,5 Stunden an Schultagen sowie etwa 8 Stunden am Wochenendtagen in ihrer Freizeit vor Bildschirmen (Delgrande Jordan & Masseroni, 2020) – diese Werte überschreiten die empfohlenen Richtlinien für freizeitliche Bildschirmzeit von maximal 2 Stunden pro Tag deutlich (Marciano et al., 2021). Insbesondere die sozialen Medien prägen den Alltag vieler junger Menschen derart, dass der Tag mit einem Blick auf die sozialen Netzwerke beginnt, der Tagesverlauf durch die mit neuen Medien verbundene Unmittelbarkeit (mit-)geleitet wird und der Tagesabschluss durch einen letzten Blick oder Eintrag in den Netzwerken Ritualisierung findet (Frison & Eggermont, 2015). Mit der Mike-Studie (Suter et al., 2023) wurde weiterhin erstmals eine repräsentative Stichprobe von Kindern unter 12 Jahren in der Schweiz zu ihrer Medienausstattung und ihrem Mediennutzungsverhalten realisiert. Aufgrund der Möglichkeiten dieser technologischen Revolution und der ständigen Mediatisierung des Alltags erfolgen soziale Interaktionen heute meist über Chatportale, die jedoch auch negative Auswirkungen wie Ausgrenzung, Gewalt und Mobbing mit sich bringen können (Berne et al., 2013; Katzer, 2014; Sticca et al., 2013). Neben Aussagen zur Medienausstattung (z.B. Besitz eines Smartphones = 65%; Anteil im Alter von 12-13 Jahren = 93%), hält die Studie weiterhin einige problematische Aspekte der Mediennutzung fest. So gaben etwa 26% der Kinder an, dass sie Cybermobbing mitbekommen hätten sowie 15% selbst Ziel von Cybermobbing geworden sind. Das Thema Cybermobbing gewinnt sowohl national als auch international stark an öffentlichem und wissenschaftlichem Interesse (u.a. Fenaughty & Harré, 2013; Festl, 2015; Pieschl & Porsch, 2014), wobei die empirischen Belege bisher eher spärlich sind (auch in der Schweiz und generell im deutschsprachigen Raum).

Diese Problematik aufnehmend interessiert in diesem Beitrag (1) wie verbreitet Erfahrungen mit Cybermobbing bei Primarschüler:innen der Klassenstufen 5 und 6 in der Schweiz sind, (2) welche Einstellungen diese gegenüber Cybermobbing haben und (3) in welchem Zusammenhang Angsterleben sowie Einstellungen gegenüber Cybermobbing die aktive Cybermobbingverfälschung auf Seiten der Täter:innen erklären.

Die Stichprobe entstammt einer im Frühjahr 2024 durchgeführten Fragebogenstudie an 14 Schulen im Kanton Bern und umfasst 906 Primarschüler:innen (M=11.3 Jahre, SD=.82; 446 Mädchen, 460 Jungen). Die analytischen Betrachtungen differenzieren hierbei zwischen (a) Beobachtenden (sog. Bystander), (b) Opfern, (c) Tätern und (d) Kindern die sowohl Opfer als auch

Täter:innen von Cybermobbing waren. Die Einstellung gegenüber Cybermobbing wurde in Anlehnung an Biedermann et al. (2018) mittels fünf Items erhoben (g-Faktormodell; Modellfit: $\chi^2=7.45$; $df=5$; $p=.00$; CFI=.99; RMSEA=.02; SRMR=.02).

Bezogen auf die Gesamtstichprobe berichten 2-20 Prozent der Primarschüler:innen von Erfahrungen mit Cybermobbing, sei es als Beobachtende oder als direkt Beteiligte (Beobachtende [Bystander]: 20%, Täter:innen: 4%, Opfer: 11.7%, sowohl Opfer als auch Täter:innen: 2.1%). Im Rahmen eines SEM wurde ein Mediationsmodell Angst (UV) – Einstellung gegenüber Cybermobbing (IV) – aktive Teilnahme Cybermobbing (AV) spezifiziert. Im Ergebnis zeigt sich ein Effekt von Angst auf ECM ($\beta=.23$; $p<.001$). Eine befürwortende Einstellung gegenüber Cybermobbing prädiziert die aktive Teilnahme am Cybermobbing ($\beta=.33$; $p<.001$). Es gibt einen direkten Effekt vom Angsterleben auf die aktive Teilnahme an Cybermobbing ($\beta=.09$; $p<.05$) sowie einen indirekten Effekt von Angst über ECM ($\beta=.08$; $p<.01$). Der totale Effekt von Angst auf Teilnahme an Cybermobbing liegt bei .16 ($p<.001$).

Der Zusammenhang zwischen digitalen und sozialen Kompetenzen von Fünftklässler*innen an Gesamtschulen

Wilhelmine Berger, Sonja Nonte, Tobias Koch, Christian Reintjes, Ferdinand Stebner

Universität Osnabrück, Deutschland

Die fortschreitende Digitalisierung betrifft sämtliche Lebensbereiche unserer Gesellschaft. Studien zum Medienalltag von Kindern und Jugendlichen (JIM 2023 und KIM 2021) zeigen, dass digitale Medien zum einen omnipräsent sind und die Freizeitgestaltung Heranwachsender massiv prägen.

Der Begriffswechsel von Medienbildung und Medienkompetenz (z.B. Hugger, 2021) zu digitalen Kompetenzen in den 1990er Jahren brachte eine Vielzahl an teilweise synonymen Begrifflichkeiten mit sich: z.B. media competence/media literacy, digital literacy, ICT- und Computerskills (Kerres 2023). In aktuellen Diskursen wird zudem die Frage aufgeworfen, in welchem Verhältnis Medienhandeln und soziales Handeln bzw. Medienkompetenz und soziale Kompetenz stehen (Cousseran & Brüggem, 2022).

In den Bundesländern finden sich eine Vielzahl an Richtlinien, Rahmenmodellen aber auch konkrete Kompetenzmodelle, die unter Bezugnahme auf den Europäischen Referenzrahmen DigCom 2.2 (Vuorikari et al., 2022) entwickelt wurden oder sich darauf beziehen. Hier wird digitale Kompetenz definiert als „confident, critical and responsible use of, and engagement with, digital technologies for learning, at work, and for participation in society“ (S. 3). Digitale Kompetenz wird als eine Kombination aus Wissen, Fähigkeiten und Einstellungen verstanden, welche eine Schlüsselkompetenz des lebenslangen Lernens ist und sich stetig verändert und entwickelt (Vuorikari et al., 2022). Sie wird bisher vorwiegend bei älteren Schüler*innen (ICILS, Eickelmann et al., 2019; NEPS, Senkbeil et al., 2013) oder anhand von Selbsteinschätzungen (PISA, Mang et al., 2023) erfasst.

Studien zeigen positive Zusammenhänge zwischen Lese- und Mathematikkompetenzen, dem Bildungsniveau (Schulform) und der Grundintelligenz: je höher das Bildungsniveau und die Intelligenz desto schneller und flexibler können Fähigkeiten im Umgang mit alten und neuen Medien angeeignet werden (Senkbeil et al., 2013; Senkbeil, 2022; Treumann et al., 2007). Für die Entwicklung von Medienkompetenz ist zudem die soziale Umwelt besonders relevant, ein hoher Bildungshintergrund der Eltern sowie ein hohes familiäres kulturelles Kapital (Anzahl der Bücher) haben einen positiven Einfluss auf die Kompetenzen (Senkbeil et al., 2013; Siddiq et al., 2017). Digitale Kompetenzen können zudem nach Tso et al. (2022) ein besseres psychisches Wohlbefinden bedingen, aber als Mediator zwischen Mediennutzungszeit und Spielsucht interpretiert werden. Keine eindeutigen Befunde, beziehungsweise nur sehr geringe Effekte, zeigten sich hingegen in Bezug auf Geschlechtsunterschiede (vgl. Gnams, 2021; Treumann et al., 2007; Siddiq & Scherer, 2019). Die Mediennutzungszeit und Verhaltensauffälligkeiten von Kindern und Jugendlichen stehen in einem signifikanten negativen Zusammenhang, insbesondere wenn die tägliche Medienzeit über 180 Minuten beträgt (Cartanya-Hueso et al. 2021).

Die starken Ausdifferenzierungen technischer Endgeräte und Dienste (Hugger, 2021) führen dazu, dass die Validität von Aufgabeninhalten bei digitalen Kompetenztests aufgrund mangelnder Aktualität und Relevanz beeinträchtigt wird. Zudem erfassen bisherige Studien vor allem digitale Kompetenzen von älteren Schüler*innen. Im Kontext der hier vorgestellten Studie wurde daher erstmalig ein digitaler Kompetenztest eingesetzt, der an Lazonder et al. (2020) angelehnt ist und aus dem Niederländischen übersetzt, anschließend adaptiert und in Bezug zum Niedersächsischen Medienkompetenzrahmen (MK 2020) weiterentwickelt wurde. Im Rahmen eines multimethodischen Kohorten-Längsschnittdesigns wurde dieser zum ersten Messzeitpunkt an 16 Gesamtschulen in Niedersachsen und Nordrhein-Westfalen in 45 Klassen bei insgesamt $n=967$ Schüler*innen eingesetzt. Zudem bearbeiteten die Schüler*innen einen standardisierten Fragebogen mit Angaben u.a. zum soziodemographischen Hintergrund und zum Mediennutzungsverhalten in der Schule und in der Freizeit. Zusammenhänge zwischen dem Mediennutzungsverhalten (u.a. Zeit, die mit digitalen Medien verbracht wird; Eigenentwicklung), den digitalen Kompetenzen (s.o., WLE Reliabilität: .61) sowie den selbsteingeschätzten sozialen Kompetenzen (SDQ; Goodman et al., 1998) deuten auf signifikante geringe Effekte im Kontext von multiplen linearen Regressionsanalysen hin. In einem nächsten Schritt werden die zentralen Konstrukte latent modelliert und im Rahmen von Mehrgruppenanalysen (Asparouhov & Muthén, 2012) unter Kontrolle der Klassenzugehörigkeit und weiterer Hintergrundmerkmale sowie unter Berücksichtigung der Medienzeit als Gruppierungsmerkmal ausgewertet. Die Ergebnisse werden vorgestellt und hinsichtlich ihrer praktischen und wissenschaftlichen Relevanz eingeordnet und diskutiert.

PS-34: Inklusion und sozial-emotionale Entwicklung in der Grundschule

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SO 322

Chair der Sitzung: Stefan Kulakow

Zum längsschnittlichen Zusammenhang von internalisierenden und externalisierenden Verhaltensproblemen und schulischer Leistung in der Grundschule

Stefan Kulakow¹, Kathrin Mahlau², Aleksander Kocaj³

¹Institut für Psychologie, Universität Greifswald; ²Institut für Erziehungswissenschaft, Universität Greifswald; ³Institut für Qualitätsentwicklung im Bildungswesen, Humboldt-Universität zu Berlin

Theoretischer

Das Verständnis der Beziehung zwischen internalisierenden und externalisierenden Verhaltensproblemen und schulischen Leistungen in der Grundschule ist für die Entwicklung von Schüler*innen von entscheidender Bedeutung, um adäquate und frühzeitige Förderansätze festzulegen. Kaskadenmodelle (Masten et al., 2005) gehen davon aus, dass sich Verhaltensprobleme und schulische Leistung gegenseitig wechselseitig beeinflussen und dadurch Spiralwirkungen und gegenseitige Verstärkungen bedingen. Beispielsweise kann die Leistung bei störendem Verhalten während des Unterrichts leiden und das Absinken der Leistung dann wiederum zu einer Verstärkung von Verhaltensproblemen beitragen. Allerdings ist die empirische Ergebnislage zu diesen Beziehungen uneindeutig und hängen von den konkreten Verhaltensauffälligkeiten (z. B. internalisierend vs. externalisierend) ab. So zeigen einige Studien reziproke Effekte in Einklang mit dem Kaskadenmodell auf, während andere Studien nur unidirektionale Effekte nahelegen und wiederum andere gar keine Effekte identifizieren (Chen et al., 2022; Metsäpelto et al., 2015; Zimmermann et al., 2013).

Forschungsziele

Diese Studie zielt darauf ab, das dynamische Zusammenspiel zwischen Verhaltensproblemen (internalisierend und externalisierend) und schulischen Leistungen zu untersuchen. Basierend auf dem Modell der reziproken Effekte werden dabei folgende Hypothesen aufgestellt: H1) Frühe Verhaltensprobleme wirken sich negativ auf spätere akademische Leistungen aus. H2) Schulische Leistungen wirken sich negativ auf Verhaltensprobleme aus.

Stichprobe

Die Stichprobe der Studie umfasste N = 461 Grundschüler*innen, welche zu fünf Messzeitpunkten untersucht wurden (Beginn 1. Klasse, dann jeweils Ende der 2., 3., 4. Klassenstufe) und bei denen sowohl die schulischen Leistungen (Lesen, Schreiben, Mathematik) als auch die Verhaltensprobleme (internalisierende und externalisierende) erfasst wurden. Dabei wurde sowohl auf standardisierte Leistungstests (Lesefähigkeit mittels Würzburger Leise Leseprobe-Revision [Schneider et al., 2011], Rechtschreibfähigkeit mittels der Reihe Deutscher Rechtschreibtests [u.a. Stock & Schneider, 2008], Reihe Deutscher Mathematiktests [u.a. Roick et al., 2004]) als auch etablierte Lehrkräfteeinschätzungen zum Sozialverhalten (Strengths and Difficulties Questionnaire [Goodman, 2005]) zurückgegriffen. Darüber hinaus wurden kognitive Fähigkeiten, Aufmerksamkeit und Geschlecht als potentielle Kontrollvariablen in die Analysen einbezogen. Die Forschungsfragen und Hypothesen wurden mithilfe einer Reihe von Cross-Lagged Panel Modelle untersucht, wobei die schulische Entwicklung der Schüler*innen in der gesamten Grundschulzeit betrachtet wurde.

Ergebnisse

Entgegen unserer Hypothesen deuten die Ergebnisse unserer Analysen nur auf sporadische reziproke Effekte hin. Dabei ist anzumerken, dass sich die Effekte sowohl hinsichtlich der Leistungsdomäne (Lesen, Schreiben, Mathematik), hinsichtlich der betrachteten Facetten internalisierender (Probleme mit Gleichaltrigen, emotionale Probleme) und externalisierender Probleme (Verhaltensprobleme, Hyperaktivität) als auch hinsichtlich der Klassenstufe der Schüler*innen teils erheblich unterscheiden. So finden sich reziproke Effekte ausschließlich in der mathematischen Domäne, während die sprachlichen Domänen (Rechtschreibung, Lesen) eher von unidirektionalen Effekten geprägt sind. Unabhängig von der Signifikanz dieser Pfade ist allerdings festzuhalten, dass entsprechende Effektschätzungen klein ausfallen. Dagegen finden sich moderate bis hohe Rangreihenstabilitäten der untersuchten Konstrukte über alle vier Grundschuljahre. Diese Ergebnisse deuten darauf hin, dass sich Verhaltensprobleme und geringe Schulleistungen bereits vor dem Schuleintritt manifestieren und in der Grundschulzeit eine hohe Stabilität aufweisen. Die Ergebnisse werden vor dem Hintergrund von pädagogisch-psychologischer Diagnostik und der Früherkennung maladaptiver Entwicklungen sowie der unterrichtlichen (Differenzierungsmaßnahmen und Unterrichtsführung) und schulsystemischen Interventionsmöglichkeiten (u.a. Response-to-Intervention Ansätze) diskutiert.

Determinanten für prosoziales Verhalten und Einstellungen von Grundschulkindern gegenüber Peers mit emotional-sozialem Förderbedarf

Marwin Felix Löper, Frank Hellmich

Universität Paderborn, Deutschland

Seit der Ratifizierung der UN-BRK (2006) sehen sich Grundschullehrkräfte mit der Herausforderung konfrontiert, inklusiven Unterricht zu gestalten, an dem alle Kinder sozial teilhaben können. Soziale Teilhabe im inklusiven Unterricht kann dann als gelungen bezeichnet werden, wenn gegenseitige Freundschaften und positive soziale Interaktionen zwischen den Kindern bestehen, sich alle Kinder in ihrer Klassengemeinschaft eingebunden und sich von ihren Mitschüler*innen akzeptiert fühlen (Koster et al., 2009). Aktuelle Forschungsbefunde zeigen jedoch, dass Kinder mit einem sonderpädagogischen Förderbedarf (SPF) in der emotional-sozialen Entwicklung im inklusiven Unterricht häufig sozial ausgegrenzt werden (Krull et al., 2018; Tometten et al., 2021), was bei anhaltender Dauer zu schwerwiegenden Problemen wie beispielsweise Einsamkeit, einem geringen Selbstwertgefühl, Depressionen und schlechten Schulleistungen führen kann (Juvonen et al., 2019). Als zentrale Ansatzpunkte, um der sozialen Ausgrenzung von Kindern mit sozial-emotionalem SPF begegnen und ihnen eine gelungene Teilhabe im inklusiven Unterricht ermöglichen zu können, konnten die Einstellungen und das prosoziale Verhalten von Grundschulkindern identifiziert werden (Freer, 2020). Einstellungen sind dabei als psychologische Tendenzen zu verstehen, die Aufschluss darüber geben, ob ein Grundschulkind einem Peer mit sozial-emotionalem SPF positiv oder negativ begegnet (Eagly & Chaiken, 1993). Prosoziales Verhalten meint die Hilfe oder Unterstützung eines Peers mit SPF, ohne dass dafür eine unmittelbare Gegenleistung erwartet wird (Eisenberg & Spinrad, 2014). Während bereits bekannt ist, dass Grundschulkindern gegenüber Peers mit sozial-emotionalem SPF negativer eingestellt sind als gegenüber Peers ohne oder mit anderem SPF (Freer, 2021), liegen aktuell keine Befunde zu den Ausprägungen von prosozialem Verhalten von Grundschulkindern gegenüber Peers mit emotional-sozialem SPF vor. Auch ist gegenwärtig weitgehend unerforscht, wie sich die Einstellungen von Grundschulkindern gegenüber Peers mit emotional-sozialem SPF und ihr prosoziales Verhalten erklären lassen. Bereits vorliegende Forschungsergebnisse deuten allerdings darauf hin, dass die eigenen Kontakterfahrungen von Grundschulkindern mit Peers mit SPF (Allport, 1954; Armstrong et al., 2017) sowie das von ihnen wahrgenommene Lehrkraftverhalten (Feinman, 1992; Hendrickx

et al., 2017) wichtige Determinanten bei der Entwicklung positiver Einstellungen und prosozialen Verhaltens sind. Darüber hinaus gibt es Hinweise darauf, dass die empathischen Fähigkeiten von Grundschulkindern, beispielsweise ihre Fähigkeit, die Perspektiven anderer Kinder zu übernehmen, eine wichtige Rolle bei der Entwicklung positiver Einstellungen und bei der Genese ihres prosozialen Verhaltens spielen (Armstrong et al., 2016; Davis, 2018). Vor diesem Hintergrund nehmen wir an, dass sich die Einstellungen von Grundschulkindern gegenüber Peers mit emotional-sozialem SPF und ihr prosoziales Verhalten durch ihre bisherigen Kontakterfahrungen, das von ihnen wahrgenommene Lehrkraftverhalten im Umgang mit Kindern mit SPF und ihre Fähigkeit zur Perspektivübernahme erklären lassen.

Zur Überprüfung dieser Forschungshypothese haben wir insgesamt $N=577$ Grundschul Kinder der dritten und vierten Jahrgangsstufe ($M=8,66$ Jahre) darum gebeten, Fragebogens Aussagen zu ihren Einstellungen gegenüber Peers mit emotional-sozialem SPF, ihrem prosozialem Verhalten, ihrer Perspektivübernahme, ihren bisherigen Kontakterfahrungen und dem von ihnen wahrgenommenen Lehrkraftverhalten im Umgang mit Kindern mit SPF zu beantworten. Die Ergebnisse aus einem in Mplus berechneten Strukturgleichungsmodell ($\chi^2=1260,27$; $df=685$; $\chi^2/df=1,84$; $p\leq .001$; CFI=.931; TLI=.925; RMSEA=.038 [.035–.042]) zeigen, dass die Einstellungen der beteiligten Kinder gegenüber Peers mit emotional-sozialem SPF ($R^2=.56$; $p\leq .001$) zwar durch ihre bisherigen Kontakterfahrungen ($\beta=.16$; $p\leq .01$) und das von ihnen wahrgenommene Lehrkraftverhalten ($\beta=.43$; $p\leq .001$) erklärt werden können, nicht aber durch ihre Perspektivübernahme ($\beta=-.01$; $p=.84$). Für das prosoziale Verhalten ($R^2=.54$; $p\leq .001$) sind sowohl die Perspektivübernahme ($\beta=.62$; $p\leq .001$) als auch das wahrgenommene Lehrkraftverhalten prädiktiv ($\beta=.45$; $p\leq .001$). Zwischen dem prosozialem Verhalten und den Kontakterfahrungen besteht hingegen kein signifikanter Zusammenhang ($\beta=.08$; $p=.07$). Mit dem Blick auf die inklusive Schulpraxis verdeutlichen unsere Ergebnisse, dass Lehrkräfte vermutlich eine zentrale Vorbildfunktion in Bezug auf die Einstellungen und das prosoziale Verhalten ihrer Schüler*innen gegenüber Peers mit SPF haben und ihr Verhalten im Unterricht nachhaltige Auswirkungen auf die soziale Teilhabe von Kindern mit SPF im inklusiven Unterricht haben kann.

Der Effekt unterschwelliger Classroom Managements auf die Feedbackwahrnehmung und soziale Akzeptanz von Grundschulkindern

Raphael Plutz, Markus Spilles, Philipp Nicolay, Christian Huber

Bergische Universität Wuppertal, Deutschland

Es wird angenommen, dass Kinder mit Verhaltensproblemen bis zu dreimal häufiger von sozialer Ausgrenzung in inklusiven Bildungskontexten betroffen sind als ihre Mitschüler*innen (Bless, 2007; Huber et al., 2022). Damit ist eine Vielzahl schulbezogener Risikofaktoren verbunden, darunter Aggression, Angststörungen, depressive Symptome, schlechte schulische Leistungen und Schulabsentismus (Flook et al., 2005; Newcomb et al., 1993; Wentzel et al., 2021). Die *social referencing theory* nach Feinman (1992) legt nahe, dass Kinder das Verhalten ihrer Bezugspersonen in unsicheren oder neuen Situationen beobachten, um Hinweise darauf zu erhalten, wie sie reagieren sollen. In einem schulischen Kontext könnte dies bedeuten, dass Kinder, die häufig beobachten, wie andere negatives Feedback erhalten, eine negative Einstellung diesen gegenüber entwickeln (Garrote et al., 2020; Huber, 2019; Huber et al., 2022; Nicolay & Huber, 2023; Spilles et al., 2023).

Vor allem in Klassen, in denen Lehrkräfte ein reaktives Classroom Management zeigen, wird dementsprechend ein Zusammenhang zwischen sozialer Akzeptanz und Verhaltensproblemen vermutet (Garrote et al., 2020). Ein Ansatz, der diesem entgegenwirken könnte, ist sog. "Low Profile Classroom Management" (LPCM) (Borich, 2016; Rinne, 1984). Das Prinzip des LPCM besteht darin, Fehlverhalten ruhig und deeskalierend zu korrigieren. Den theoretischen Annahmen entsprechend kann LPCM dazu führen, dass Mitschüler*innen weniger negatives Lehrer*innen-Feedback gegenüber Schüler*innen mit Verhaltensproblemen wahrnehmen, was deren soziale Akzeptanz erhöhen könnte. Darüber hinaus könnte LPCM einen abschwächenden Effekt auf die Beziehung zwischen Verhaltensproblemen und sozialer Akzeptanz haben. Der vorliegende Beitrag behandelt die folgenden Forschungsfragen:

1. Wird der positive Zusammenhang zwischen Verhaltensproblemen und von Mitschüler*innen wahrgenommenem negativem Lehrer*innen-Feedback durch die Häufigkeit des Einsatzes von LPCM-Techniken moderiert?
2. Wird der negative Zusammenhang zwischen Verhaltensproblemen und sozialer Akzeptanz durch den Einsatz von LPCM-Techniken mediert?

Die Studie umfasste eine Befragung von 690 Schulkindern (54 % weiblich) aus 35 dritten und vierten Klassen ($M_{\text{Alter}} = 9.19$; $SD_{\text{Alter}} = 0.88$) in Nordrhein-Westfalen. Zusätzlich wurden die Klassenlehrkräfte (91 % weiblich; $M_{\text{Alter}} = 42.12$; $SD_{\text{Alter}} = 9.70$) während ihres Unterrichts beobachtet und in ihrem Classroom Management von zuvor geschulten Rater*innen bewertet. Für jedes Kind wurde das negative Lehrkraft-Feedback von allen Mitschüler*innen anhand einer Likert-Skala bewertet. Alle Peer-Bewertungen in einer Klasse wurden zu einem individuellen Mittelwert für die einzelnen Schüler*in aggregiert. Die soziale Akzeptanz wurde mittels eines soziometrischen Fragebogens (Moreno, 1953) erfasst, bei dem die Schüler*innen angaben, wie gerne sie neben ihren Mitschüler*innen sitzen möchten. Die soziale Akzeptanz wurde anschließend standardisiert, um Unterschiede in der Anzahl der Schüler*innen pro Klasse zu berücksichtigen. Um störendes Verhalten zu messen, wurde die deutsche Kurzversion der Skala *Oppositional and Disruptive Behavior* der Integrated Teacher Report Form (ITRF; Casale et al., 2019) verwendet. Die Häufigkeit von 11 LPCM-Techniken wurde von jeweils zwei Rater*innen während drei Unterrichtsstunden eingeschätzt. Die Items wurden für diese Studie eigens entwickelt und basieren auf theoretischen Überlegungen von Borich (2016), Eichhorn (2015), Seethaler et al. (2021), Syring (2016) und Piwowar (2013a, 2013b).

Die Datenanalyse wurde anhand von Mehrebenenregressionsmodellen mit den R-Paketen *lme4* (Bates et al., 2015) beziehungsweise *brms* (Bürkner, 2017) durchgeführt. Die individuellen Schüler*innenmerkmale befinden sich auf Ebene 1, während die Nutzung von LPCM auf Ebene 2 angesiedelt ist. Aufgrund der geringen Stichprobengröße auf Ebene 2 wurde zur Beantwortung der zweiten Forschungsfrage ein bayesianisches Modell verwendet. Ziel ist es, zu untersuchen, inwieweit der moderierende Effekt des LPCM auf den Zusammenhang zwischen Verhaltensproblemen und negativem Lehrerfeedback einen vermittelnden Effekt auf die soziale Akzeptanz der feedbackempfangenden Kinder hat.

Die Analysen zeigen, dass der positive Zusammenhang zwischen störendem Verhalten und negativem Feedback in Klassen mit häufigem Einsatz von LPCM-Techniken hypothesenkonform geringer ausgeprägt ist ($\beta = -0.01$; $p = .034$). Hinsichtlich der sozialen Akzeptanz konnte jedoch kein Mediationseffekt festgestellt werden.

Mögliche inhaltliche und methodische Limitationen des Forschungsvorhabens sollen während des Beitrages diskutiert werden.

Entwicklung inklusionsbezogener Kompetenzen von angehenden Grundschullehrkräften: eine Längsschnittstudie zu Studienbeginn

Theoretisch-empirische Einbettung

Der konstruktive Umgang mit Vielfalt und Inklusion führt mittlerweile seit mehr als 15 Jahren zu einer immer noch andauernden Transformation des deutschen Schulsystems. Dabei war die Grundschule als erste verpflichtende Bildungseinrichtung für alle Kinder ideengeschichtlich bereits als inklusives Setting im weiteren Sinne konzipiert (Götz et al. 2019). Aus diesem Gründungsanspruch und der Transformation zu einem auch im engeren Sinne inklusiven Setting ergeben sich erweiterte Anforderungen an die professionelle Kompetenz von Grundschullehrkräften: Um konstruktiv mit den besonderen, nun auch sonderpädagogischen Bedarfen aller Schüler:innen umzugehen, benötigen Grundschullehrkräfte inklusionsbezogene professionelle Kompetenzen (Wissen, Überzeugungen, Orientierungen, Selbstregulationsfähigkeiten; König et al. 2019). Für deren Entwicklung sind im Studium inklusionsbezogene Lehrveranstaltungen nachweislich entscheidend (Elting et al. erscheint 2024; Menge et al. 2021). Auch konnten Zusammenhänge zwischen verschiedenen inklusionsbezogenen Kompetenzfacetten, z.B. inklusionsbezogenes professionelles Wissen, Einstellungen oder Selbstwirksamkeitserwartungen, nachgewiesen werden (Bosse et al. 2017; Hellmich et al. 2016; Liebner et al. erscheint 2024). Insbesondere das inklusionsbezogene professionelle Wissen angehender Grundschullehrkräfte ist dabei jedoch weitgehend unerforscht. Es fehlen v.a. längsschnittliche Studien, die Entwicklungen des inklusionsbezogenen pädagogischen Wissens in Abhängigkeit vom Vorwissen zu Studienbeginn, von wahrgenommenen Lerngelegenheiten im Studium sowie im Zusammenspiel mit weiteren inklusionsbezogenen Kompetenzfacetten nachzuzeichnen vermögen.

Forschungsfragen

Diesen Desiderata widmet sich das Projekt InkuWi (Inklusionsbezogene professionelle Kompetenzen angehender Grundschullehrkräfte) mit folgenden ausgewählten Forschungsfragen:

1. Wie entwickeln sich die inklusionsbezogenen Kompetenzen (pädagogisches Wissen, Überzeugungen, Orientierungen, Selbstregulationsfähigkeiten) von Grundschullehramtsstudierenden im ersten Fachsemester?
 1. Wodurch kann das inklusionsbezogene pädagogische Wissen am Semesterende erklärt werden?

Methoden

Nach querschnittlichen Vorarbeiten (Elting et al. erscheint 2024) wurden im Wintersemester 2023/24 in grundschulpädagogischen und -didaktischen Einführungsveranstaltungen an vier bayerischen Universitätsstandorten inklusionsbezogene Kompetenzen von Grundschullehramtsstudierenden ($N=144$; weiblich: 88.2%; Alter: $M=19.95$, $SD=3.15$) zu Beginn (t_1) und Ende (t_2) der Vorlesungszeit des ersten Fachsemesters per Online-Fragebogen erhoben (ausführlich Liebner et al. erscheint 2024). Das inklusionsbezogene pädagogische Wissen wurde anhand eines adaptierten Multiple-Choice-Tests (30 Items zu je 0-2 Punkten; Zumbos $\alpha=.60-.72$) erfasst. Die inklusionsbezogenen Einstellungen, Selbstwirksamkeitserwartungen, motivationalen Orientierungen und Selbstregulationsfähigkeiten (beruflicher Ehrgeiz, Distanzierungsfähigkeit, offensive Problembewältigung und Berufszufriedenheit) schätzten die Studierenden anhand mehrerer adaptierter Subskalen (je 4-8 Items; vierstufige Antwortskala; Cronbachs $\alpha=.60-.94$) ein. Daneben wurden diverse Hintergrundmerkmale der Person und zum Studium erhoben.

Ergebnisse

T-Tests für abhängige Daten zeigten, dass die Studierenden – gemessen am theoretischen Mittel des Wissenstests von 30 Punkten – mit moderatem inklusionsbezogenem Wissen in das Grundschullehramtsstudium gestartet waren, welches sich zum Ende des ersten Semesters auf ein hohes Niveau steigerte ($M_{t1}=33.00$, $SD_{t1}=5.45$; $M_{t2}=52.98$, $SD_{t2}=7.26$; $T=35.35$, $p<.001$, $d=2.95$). Inklusionsbezogene Einstellungen ($M_{t1}=2.96$, $SD_{t1}=0.52$; $M_{t2}=3.10$, $SD_{t2}=0.52$; $T=2.60$, $p=.005$, $d=0.22$), beruflicher Ehrgeiz ($M_{t1}=2.59$, $SD_{t1}=0.52$; $M_{t2}=2.72$, $SD_{t2}=0.53$; $T=3.15$, $p<.001$, $d=0.26$) sowie die Berufszufriedenheit ($M_{t1}=3.03$, $SD_{t1}=0.46$; $M_{t2}=2.93$, $SD_{t2}=0.54$; $T=-2.29$, $p=.012$, $d=0.19$) veränderten sich bei einem moderaten bis eher hohen Ausgangsniveau nur geringfügig. In den weiteren Kompetenzfacetten waren keine signifikanten Veränderungen zu verzeichnen.

In multiplen linearen Regressionsanalysen konnte das inklusionsbezogene pädagogische Wissen am Ende des Semesters unter Kontrolle der Ausgangswerte ($\beta=.35$, $p<.001$) durch die Note des höchsten Bildungsabschlusses der Studierenden ($\beta=-.31$, $p<.001$) erklärt werden. Daneben erwiesen sich ein vorheriges Studium ($\beta=.20$, $p<.001$) und inklusionsbezogene Einstellungen ($\beta=.27$, $p<.001$) als relevante Prädiktoren (korrigiertes $R^2=.41$). Diverse Merkmale der Person (z.B. Geschlecht, Familiensprache), Lernerfahrungen in- und außerhalb des Studiums (z.B. Kontakt zu Menschen mit Behinderung/Förderbedarf) wie auch die weiteren inklusionsbezogenen Kompetenzfacetten leisteten hingegen keinen signifikanten Erklärungsbeitrag.

Diskussion

Die Befunde zeigen, dass grundschulpädagogische und -didaktische Einführungsveranstaltungen des ersten Fachsemesters wesentliche Lernorte für den Erwerb inklusionsbezogenen pädagogischen Wissens darstellen, jedoch nur bedingt zur Entwicklung weiterer inklusionsbezogener Kompetenzfacetten beizutragen scheinen. Mit Blick auf inklusionspädagogische Bezüge im Grundschullehramtsstudium wird das teils erwartungswidrige Ausbleiben von Effekten weiterer inner- und außeruniversitärer Lerngelegenheiten diskutiert. Weiterhin werden angesichts des Einflusses der Note des höchsten Schulabschlusses Fragen der Chancengerechtigkeit im Grundschullehramtsstudium aus hochschuldidaktischer Perspektive aufgegriffen.

PS-35: Motivation und Emotionen im Studium

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 169

Chair der Sitzung: Robert Grassinger

Emotionen bei der Studienwahl

Robin Göller¹, Larissa Altenburger², Dominik Schlüter², Michael Besser²

¹Universität Klagenfurt, Österreich; ²Leuphana Universität Lüneburg

Theoretischer Hintergrund

Studienwahlmotive und Emotionen haben eine hohe Bedeutung für Erfolg und Zufriedenheit im Studium und späteren Beruf (Biermann et al., 2019; Camacho-Morles et al., 2021; Hanfstingl & Mayr, 2007; Wach et al., 2016). Jedoch sind Emotionen, die mit der Studienwahl einhergehen, bisher kaum untersucht. In gängigen Erhebungsinstrumenten sind Studienwahlmotive oft auf Basis der Expectancy-Value-Cost-Theorie (Eccles & Wigfield, 2020) konzipiert, indem beispielsweise Expectancy als Fachliche Fähigkeitsüberzeugung, Intrinsic Value als Fachliches Interesse, Attainment Value als Gesellschaftlicher Wert des Studiums, Utility Value als Finanzielle Nützlichkeit oder die Vereinbarkeit von Familie und Beruf sowie (eine Vermeidung von) Kosten als geringe Schwierigkeit des Studiums operationalisiert sind (Göller & Besser, 2021; Pohlmann & Möller, 2010; Watt & Richardson, 2007). Aus Perspektive der Control-Value-Theorie (Pekrun, 2006) ist zu erwarten, dass solche Kontrollüberzeugungen und Werte mit Emotionen bzgl. des Studiums und der Studieninhalte verbunden sind. Dies soll im vorliegenden Beitrag anhand folgender Fragestellung näher untersucht werden:

Fragestellung

Welche Bedeutung haben verschiedene Studienwahlmotive für die mit den Studieninhalten verbundenen Emotionen?

Methode

Im Sommer 2022 haben im Rahmen des Auswahlverfahrens der Leuphana Universität Lüneburg insgesamt 544 (399 weiblich, 137 männlich, 5 divers, 3 ohne Angabe) Bewerber:innen auf verschiedene Studiengänge freiwillig an einer papierbasierten Befragung teilgenommen, in der Studienwahlmotive (20 Items; Fachliche Fähigkeitsüberzeugung, Fachliches Interesse, Gesellschaftlicher Wert des Studiums, Finanzielle Nützlichkeit, Vereinbarkeit von Familie und Beruf, geringe Schwierigkeit des Studiums; Göller & Besser, 2021) sowie Emotionen (12 Items; Neugier, Freude, Angst, Langweile; Vogl et al., 2018) mithilfe standardisierter und validierter Instrumente über fünfstufige Likert-Skalen erhoben wurden. Zur Beantwortung der Forschungsfrage wurde ein Strukturgleichungsmodell mit den vier Emotionen als endogenen Variablen und den sechs Studienwahlmotiven als exogenen Variablen in Jamovi (auf Basis des R-Packages „lavaan“) berechnet.

Ergebnisse und ihre Bedeutung

Das Strukturgleichungsmodell zeigt insgesamt einen guten Modell-Fit ($\chi^2=707$, $df=414$, $p<.001$, CFI=.951, TLI=.941, RMSEA=.036, SRMR=.040). Signifikante Effekte der Studienwahlmotive auf die Neugierde finden sich für die Fachliche Fähigkeitsüberzeugung (standardisiertes $\beta=-0.24$, $p=.003$) sowie das Fachliche Interesse ($\beta=0.75$, $p<.001$), auf die Freude für das Fachliche Interesse ($\beta=0.36$, $p<.011$), den Gesellschaftlichen Wert des Studiums ($\beta=0.16$, $p=.014$) sowie die Vereinbarkeit von Familie und Beruf ($\beta=0.37$, $p=.036$). Angst wird lediglich durch die fachliche Fähigkeitsüberzeugung prädiziert ($\beta=-0.26$, $p<.001$). Signifikante Effekte der Studienwahlmotive auf die Langweile zeigen sich für das Fachliche Interesse ($\beta=-0.24$, $p=.011$), den Gesellschaftlichen Wert des Studiums ($\beta=-0.20$, $p=.004$) sowie die geringe Schwierigkeit des Studiums ($\beta=0.27$, $p<.001$).

Insgesamt zeigt sich (und dies ist auch von theoretischem Interesse) die große Bedeutung der in den Studienwahlmotiven operationalisierten verschiedenen Werte, auf deren Grundlage Studienwahlentscheidungen getroffen werden, für die untersuchten Emotionen. Dabei erweist sich vor allem eine Studienwahl, die auf Basis von fachlichem Interesse (Intrinsic Value) und der ihr zugeschriebenen gesellschaftlichen Bedeutung (Attainment Value) getroffen wurde, als positiv für die mit den Studieninhalten verbundenen Emotionen (höhere Neugier und Freude, geringere Langweile), wie auch den weiteren Studienverlauf (Biermann et al., 2019; Hanfstingl & Mayr, 2007; Wach et al., 2016). Bemerkenswert ist hier zudem die Bedeutung der Vereinbarkeit von Familie und Beruf (Utility Value) für die mit den Studieninhalten verbundene Freude, wobei die Finanzielle Nützlichkeit des Studiums (Utility Value) kaum Bedeutung für die hier untersuchten Emotionen hat. Bemerkenswert ist auch, dass die Fachliche Fähigkeitsüberzeugung (Expectancy/Control), die in der Regel mit einem positiven Studienverlauf assoziiert ist (Theurer et al., 2021), hier negativ mit den aktivierenden Emotionen (Neugier, Angst) zusammenhängt. Eine Interpretation hiervon könnte sein, dass Neugierde an gewissen Studieninhalten teilweise auch mit einer (persönlich) herausfordernderen Studienwahl einhergehen kann, wohingegen eine Vermeidung solcher Herausforderungen (geringere Kosten, geringe Schwierigkeit des Studiums) eher mit Langweile in Bezug auf die Studieninhalte verbunden ist.

Auch aus emotionaler Perspektive ist also eine Studienwahl auf Grundlage persönlicher Interessen sowie gesellschaftlicher Bedeutung und weniger auf Basis einer Vermeidung von Herausforderungen empfehlenswert.

Exploring Pathways to Thriving: Latent Profile Analysis (LPA) of College Students' Control Strategy Use and Associations with Success and Wellbeing

Fanyi Zeng, Luise von Keyserlingk

University of Tübingen, Germany

Theoretical Background

In college, many students encounter academic challenges, such as failing exams or missing self-set performance goals. The Motivational Theory of Life-Span Development (MTD; Heckhausen et al., 2010) proposes various control strategies (CS) that help individuals self-regulate behaviors and goals after experiencing challenges or setbacks. Using this framework, our study looked at whether and how U.S. college students reported using control strategies after midterms to improve their final exam performance. These strategies could be differentiated between *goal-engaging strategies*, e.g., working harder, avoiding distractions, or seeking help, and *goal-adjustment or self-protection strategies*, e.g., lowering grade aspirations or using self-serving attributions. *Goal-engaging strategies* have been linked with higher academic performance, while *goal-adjustment* and *self-protection strategies* have been associated with better mental health outcomes (Daniels et al., 2014; Hall et al., 2006). Prior findings suggest that different combinations of CS lead to optimal outcomes for different individuals and circumstances (e.g., initially unsuccessful students benefitted from both *goal-engaging* and *adjustment* CS; Hall, Perry, et al., 2006). Hence, an adaptive combination of CS based on individual experiences and dispositions should lead to optimal academic and psychological outcomes.

Research Questions

We used a person-centered approach to examine which combinations of CS college students reported, and how these combinations of CS were associated with course experiences, course-specific motivational beliefs, course performance, and well-being.

Method

We used data from an ongoing longitudinal multi-cohort study at a public university in California. The sample consisted of 545 students who completed surveys across the academic year 2022/23 (fall, winter, and spring quarters). In each quarter, after the midterm exam of their most challenging course, students reported midterm grades, motivations, and intended control strategy use for their final course exams. At the end of each quarter, students reported their overall positive and negative affect throughout the quarter. Demographics and final grades were obtained from the university's administrative data. We used latent profile analysis in Mplus and ANOVAs to address our research questions.

Results

The LPA consistently yielded four profiles across fall, winter, and spring quarters, which were labelled as: (1) *High CS* ($n = 87-157$), (2) *Medium CS* ($n = 94-135$), (3) *Low CS* ($n = 29-66$), and (4) *Goal-Focused* ($n = 47-51$). Students in the *High CS* profile reported high intended *goal-engaging strategies* (e.g., working harder, avoiding distractions, expecting success) and *goal adjustment strategies* (e.g., lowering aspirations) and medium to high intended *self-protection strategies*. The *Medium CS* and *Low CS* profiles were characterized by similar patterns on lower levels of intended strategy use. The *Goal-Focused* profile was characterized by high intended *goal engagement strategies*, low intended *self-protection strategies*, and, unlike others, low intended *goal adjustment strategies*. One-way ANOVAs and post-hoc tests indicated that profiles differed in course experiences, motivations, and outcomes. Compared to others, *High CS* profile students consistently underperformed in midterms but held higher course values than most groups ($p < .05$). Hence, after struggling in midterms, they intended to use highly *goal engagement* and *adjustment strategies*. They received lower final course grades than others in fall but consistently experienced more positive affect than some groups ($p < .05$). *Goal-Focused* profile students consistently outperformed most groups in midterms, held higher ability beliefs and course values than others in winter, obtained better final course grades than others in two quarters, and experienced more positive affect than others in two quarters ($p < .05$). *Low CS* profile students consistently reported lower course values than others ($p < .05$). These results help us understand adaptive patterns and consequences of control strategy use in response to difficult midterm exams among college students.

Veränderung des Erlebens motivationaler Erwartungsdiskrepanz zu Studienbeginn

Robert Grassinger¹, Denise Reisch¹, Olga Kunina-Habenicht², Steffen Wild²

¹Pädagogische Hochschule Weingarten, Deutschland; ²Technische Universität Dortmund

Theoretischer Hintergrund

Die Entscheidung für ein Studium ist beeinflusst von Erfolgserwartungen und antizipierten Studienwerten (Eccles & Wigfield, 2020; Janke et al., 2021). Diese motivationalen Erwartungen erfüllen sich im Studium nicht umfassend, was mit einem Erleben sowohl motivationaler diskrepanter Erfolgserwartung als auch diskrepanter Studienwerte einhergeht (Grassinger & Schnebel, 2024). Dieses Erleben basiert auf einem retrospektiven Vergleich des aktuellen motivationalen Erlebens mit vormaligen Erwartungen hieran. Gedanken wie „die Studienanforderungen sind schwerer als erwartet“ oder „die Studieninhalte sind nützlicher als gedacht“ sind Indikatoren hierfür.

Bisherige empirische Befunde verweisen darauf, dass das Erleben motivationaler Erwartungsdiskrepanz sich faktoriell von der aktuellen Studienmotivation abgrenzen lässt und mit der Veränderung der Studienmotivation korrespondiert (Grassinger, 2018). Ein Erleben negativer diskrepanter Erfolgserwartung ist assoziiert mit einer Verunsicherung ob der passenden Studienwahl (Grassinger & Schnebel, 2024). Ein Erleben negativer diskrepanter Studienwerte ist assoziiert mit weniger Freude im Studium sowie einer Intention zum Studienabbruch oder Studiengangwechsel (Grassinger, 2018).

Wenig ist darüber bekannt, (1) wie intraindividuell stabil sich ein Erleben motivationaler Erwartungsdiskrepanz erweist sowie (2) inwiefern ein Erleben diskrepanter Erfolgserwartung und ein Erleben diskrepanter Studienwerte sich gegenseitig beeinflussen. (3) Zudem geben die bisherigen Arbeiten keine Hinweise darauf, inwiefern Hochschullehrende oder weitere soziale Ressourcen ein Erleben motivationaler Erwartungsdiskrepanz reduzieren können.

Hypothesen

Per Definition basiert ein Erleben motivationaler Erwartungsdiskrepanz auf einem Abgleich der aktuellen motivationalen Situation mit vormaligen motivationalen Erwartungen an diese Situation. Da sich die Anforderungen und die Lehrinhalte im Studium verändern, ändert sich auch die aktuelle motivationale Situation und in Folge das Erleben motivationaler Erwartungsdiskrepanz. Entsprechend wird angenommen, dass sich intraindividuell das Erleben motivationaler Erwartungsdiskrepanz ändert (H1).

Die Wahrnehmung, dass etwas schwerer als erwartet ist, begünstigt – so unsere Argumentation – dass Studierende erleben, dass die Studiensituation ihnen weniger Freude als gedacht bereitet. Entsprechend wird erwartet, dass ein Erleben diskrepanter Erfolgserwartung positiv mit einem Erleben diskrepanter Studienwerte assoziiert ist (H2)

Eccles und Wigfield (2020) postulieren die Relevanz von Unterstützung für motivationale Haltungen. Empirische Befunde bestärken und spezifizieren diese Annahmen. So zeigte sich, dass Unterstützung durch Lehrende Einfluss auf das motivationale Erleben von Studierenden hat (Lazarides et al., 2019; Wild et al., 2024). Als Folge wird angenommen, dass eine ausgeprägtere Unterstützung durch Hochschullehrende und vermehrte soziale Ressourcen ein Erleben einer negativer motivationaler Erwartungsdiskrepanz reduzieren (H3).

Methode

$N = 203$ Studienanfänger:innen im Lehramt wurden in deren ersten und deren zweiten Semester befragt ($M_{it} = 21.23$, $SD_{it} = 4.22$). Das Erleben motivationaler Erwartungsdiskrepanz (5 Items; $\alpha = .80 - .82$) und das Erleben diskrepanter Studienwerte (6 Items, $\alpha = .88 - .92$) wurden nach Grassinger (2018) gemessen. Die Operationalisierung der Unterstützung durch die Hochschullehrkraft (2 Items; $\alpha = .80$) basierte auf Heublein et al. (2010). Das Vorhandensein sozialer Ressourcen wurde nach Isphording und Wozny (2018) gemessen (2 Items, $\alpha = .73$). Die erste Hypothese wurde mit einer Varianzanalyse mit Messwiederholung geprüft. Die zweite und dritte Hypothese mit einem latenten Change-Modell.

Ergebnisse und Diskussion

Das Erleben negativer diskrepanter Erfolgserwartung stieg an ($F(1,202) = 4.997, p = .026$; Cohens' $f = .024$), während das Erleben diskrepanter Studienwerte im Mittel intraindividuell stabil blieb ($F(1,202) = 3.594, p = .059$). Die Veränderung des Erlebens diskrepanter Erfolgserwartung war nicht assoziiert mit der Unterstützung durch die Hochschullehrkraft und sozialer Ressourcen (Modellgüte: $\chi^2 = 102.326; df = 92; \chi^2 / df = 1.112; p = .217; CFI = .993; TLI = .991; RMSEA = .024; SRMR = .047$). Veränderungen im Erleben diskrepanter Studienwerte hingegen waren assoziiert mit dem Erleben diskrepanter Erfolgserwartung ($\beta = .86, p < .001$) und der Unterstützung durch die Hochschullehrkraft ($\beta = .25, p < .001$). Die Befunde bekräftigen die Relevanz motivationaler Erwartungen an das Studium und die Bedeutung einer motivationalen Unterstützung durch Hochschullehrende.

PS-36: Soziale Ungleichheit und Lesekompetenz – Herausforderungen, Interventionen

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 256

Chair der Sitzung: Josefine Lühse

Eine Betrachtung der Gruppe leseschwacher Schüler:innen: Zusammensetzung und Risikofaktoren

Josefine Lühse¹, Sebastian Weirich²

¹DIPF, Deutschland; ²IQB, Deutschland

Ein zunehmender Anteil der Kinder und Jugendlichen erreicht im Lesen nicht den Mindeststandard – in der 9. Jahrgangsstufe verfehlte 2022 jede:r Dritte den Mindeststandard für den Mittleren Schulabschluss. Die in Large-Scale-Assessments erzielten Testergebnissen stehen jedoch in einem Zusammenhang mit dem künftigen Bildungserfolg und den Arbeitsmarktchancen (Jindra et al. 2024, OECD 2018). Bildung beeinflusst darüber hinaus nichtmonetäre Bereiche wie politische Partizipation und Gesundheit, weshalb der steigende Anteil kompetenzarmer Schüler:innen von besonderer Relevanz ist. Vor diesem Hintergrund geht der Beitrag folgenden Fragen nach: Wer sind diese Schüler:innen, auf deren Förderung und Begleitung ein besonderes Augenmerk gelegt werden müsste und was sind Risikofaktoren für schwache Leseleistungen?

Die Entstehung von Unterschieden in der Leseleistung wird über das Zusammenspiel zwischen familiären Struktur- und Prozessmerkmalen erklärt. Dabei stellen der sozioökonomische Status, das Bildungsniveau sowie die Migrationsgeschichte einer Familie die bedeutsamsten Strukturmerkmale für die Leistungsentwicklung dar (Baumert et al. 2003). Die jeweilige Ausprägung der Strukturmerkmale spiegelt sich in einer spezifischen Ausstattung der Familien mit ökonomischem, kulturellem und sozialem Kapital (Bourdieu 1983) wider. Diese korrespondiert dann auf Prozessebene mit unterschiedlichen familialen Bildungspraxen (Lareau 2003), welche wiederum die Kompetenzentwicklung beeinflussen.

Die Untersuchung basiert auf Daten des IQB-Bildungstrends für die 9. Jahrgangsstufe (2015, 2022). Die abhängige Variable stellt die Zugehörigkeit zur Gruppe der leseschwachen Jugendlichen dar. Die Lesekompetenz dieser Gruppe überschreitet nicht die Kompetenzstufen Ia und Ib und verfehlt den Mindeststandard für den Mittleren Schulabschluss. Die unabhängigen Variablen sind das Geschlecht, die soziale Herkunft, abgebildet mittels des International Socio-Economic Index (ISEI), der Bildungsabschluss der Eltern (ISCED), die Zahl der Bücher im Haushalt, der Migrationshintergrund (kein Elternteil im Ausland geboren, ein Elternteil im Ausland geboren, 2. Generation, 1. Generation) und die Zuhause gesprochene Sprache. Die Fragestellungen werden mittels logistischer Regressionsanalysen beantwortet.

Die deskriptiven Befunde zeigen für den Vergleich der Zusammensetzung der Risikogruppe mit der Nicht-Risikogruppe nach den verschiedenen (Herkunfts-)Merkmalen folgendes Bild: In der Risikogruppe finden sich mehr Jungen als in der Nicht-Risikogruppe, der Anteil Jugendlicher aus Familien mit niedrigem SES ist rund doppelt so hoch ist wie in der Nicht-Risikogruppe. Ein vergleichbares Bild lässt sich bei Betrachtung des ISCED der Eltern erkennen. Zudem haben in der Risikogruppe mehr Kinder einen Migrationshintergrund und sprechen seltener Zuhause Deutsch als Kinder in der Nicht-Risikogruppe.

Die Ergebnisse der logistischen Regressionsanalyse zeigen, dass sämtliche betrachtete Faktoren in erwarteter Richtung in einem statistisch bedeutsamen Zusammenhang mit der Wahrscheinlichkeit stehen, zur Risikogruppe zu gehören. Unter Kontrolle aller im Modell enthaltenen Variablen weisen Kinder aus Familien mit Migrationshintergrund eine höhere Wahrscheinlichkeit auf, zur Risikogruppe zu gehören – wobei hier der Effekt für Zugewanderte der ersten Generation am stärksten ist. Für Mädchen und Kinder aus Familien mit höherem SES bzw. ISCED besteht hingegen eine geringere Wahrscheinlichkeit, der Risikogruppe anzugehören. Die Befunde stehen in Einklang mit älteren Forschungsbefunden zur „Risikogruppe“, die auf Grundlage von PISA-Daten gewonnen wurden (Stanat & Schneider 2004).

Erste Ergebnisse für die Trendbetrachtung zeigen, dass die Wahrscheinlichkeit, zur Risikogruppe zu gehören, unter Kontrolle der übrigen Variablen 2022 nicht größer ist als 2015. Die relative Bedeutung der Risikofaktoren hat sich zwischen 2015 und 2022 nur vereinzelt geändert: So bedeutet die Zugehörigkeit zur Zuwanderungsgruppe der ersten Generation 2022 ein höheres Risiko, die Mindeststandards zu verfehlen, als 2015. Vermutlich liegt das jedoch daran, dass der Anteil Geflüchteter innerhalb dieser Gruppe 2022 größer ist als 2015. Zuhause manchmal oder nie Deutsch zu sprechen, ist 2022 dagegen mit einem geringeren Risiko assoziiert, die Mindeststandards zu verfehlen, als 2015. Dennoch ist auch 2022 das Risiko für diese Personen größer als für Schüler:innen aus deutschsprachigen Elternhäusern.

Kompensation und Akkumulation von sozialer Benachteiligung beim Erwerb von Lesekompetenz im Europäischen Vergleich

Thomas Brüggemann¹, Ulrich Ludwig¹, Jakob Schwerter^{1,2}, Philipp Doeblér², Nele McElvany¹

¹TU Dortmund, Institut für Schulentwicklungsforschung, Deutschland; ²TU Dortmund, Statistische Methoden in der Sozialwissenschaften

Die Lesekompetenz ist ein zentraler Prädiktor für den weiteren Lern- und Lebensverlauf (Dale, 2023). Allerdings unterscheiden sich Lesefähigkeiten von Kindern zum Schuleintritt erheblich (Kieffer 2012). Diese Unterschiede können sich langfristig auf die Lesekompetenz im weiteren Schulverlauf auswirken (Burchinal et al., 2020; Claessens et al., 2009). Eine wichtige Rolle bei der Entwicklung der Lesekompetenz von Kindern spielen Einflussfaktoren des *home literacy environments* (HLE). Das HLE beschreibt Überzeugungen und Ressourcen des Elternhauses, die Kinder bei der Entwicklung ihrer Lesefähigkeiten unterstützen können (Dong et al., 2020). Diese Ressourcen und Überzeugungen sind zeitlich weitgehend stabil (Sieben & Lechner, 2019; Yeoung et al., 2000). Effekte des HLE sind sowohl auf die Lesefähigkeiten von Kindern zum Schuleintritt, als auch in der vierten Klasse bekannt (Cameron et al., 2023; Claessens et al., 2009; Dong et al., 2020). Allerdings ist noch nicht hinreichend geklärt, zu welchem Zeitpunkt das HLE sich am stärksten auf die Entwicklung der Lesekompetenz von Kindern auswirken (Kieffer et al., 2012; Liu & Chung, 2022). Ein Ansatz ist, dass Kinder vor der Einschulung im Durchschnitt mehr Zeit im HLE verbringen, als als Grundschüler*innen, wodurch sich Einflüsse des HLE stärker entfalten können (Kieffer, 2012). Dadurch lässt sich annehmen, dass die Schule soziale Ungleichheiten im HLE **kompensieren** kann (z.B. Peng et al., 2019; Quinn et al., 2016). Auf der anderen Seite können sich Unterschiede im HLE sowohl während der Vorschulzeit, als auch weiterhin während der Grundschulzeit auf die Entwicklung der Lesefähigkeiten auswirken, wodurch am Ende der Grundschule ein **kumulativer Vor- bzw. Nachteil** entsteht (DiPrete & Eirich, 2006).

Vor dem Hintergrund von verschiedenen frühkindlichen Bildungssystemen in Europa (z.B. Eurydice, 2014; 2019) wird in diesem Beitrag untersucht, wie sich Aspekte des HLE auf die Lesefähigkeiten von Kindern zum Schuleintritt und in der vierten Klasse auswirken. Weiterhin werden Effekte betrachtet, die durch die COVID-19 Pandemie und die damit verbundenen

Schulschließungen entstanden sind (z.B. Freundl et al., 2021; Ludewig et al., 2022). Um herauszufinden, ob während der Grundschulzeit ein Akkumulations- oder Kompensationseffekt des HLEs vorliegt, wird untersucht inwiefern sich der indirekte Effekt der HLE über die Lesefähigkeiten zum Schuleintritt auf die Lesekompetenz in der vierten Klasse von dem direkten Effekt der HLE auf die Lesekompetenz in der vierten Klasse unterscheidet. Zusätzlich wird angenommen, dass sich die Effekte des HLE auf die Lesefähigkeiten zum Schuleintritt und in der vierten Klasse zwischen verschiedenen Europäischen Ländern aufgrund von unterschiedlichen frühkindlichen Bildungssystemen unterscheiden. Ebenfalls werden Unterschiede zwischen 2016 und 2021 aufgrund der pandemiebedingten Schulschließungen erwartet.

Die Forschungsfragen wurden mithilfe von Daten aus der IGLU-Studie bearbeitet. Die Stichprobe umfasst insgesamt 177,386 Schüler*innen aus IGLU 2016 ($N = 91,632$) und IGLU 2021 ($N = 95,754$) aus 17 Europäischen Ländern. Fehlende Werte wurden mit missranger (Meyer, 2019) imputiert. Es wurden die Konstrukte Heimatsprache, Lesefähigkeit zum Schuleintritt, Bücher im Haushalt und Wertschätzung des Lesens der Eltern untersucht und mit 2-PL Modellen skaliert. Mit Strukturgleichungsmodellen wurde der Einfluss der Heimatsprache, der Anzahl der Bücher und der Elterlichen Wertschätzung des Lesens auf die Lesefähigkeit zum Schuleintritt und die Lesekompetenz in der vierten Klasse geschätzt.

Die vorläufigen Ergebnisse legen nahe, dass soziale Ungleichheiten auch während der Grundschulzeit in den untersuchten Europäischen Ländern zunehmen. Die Befunde unterscheiden sich nicht zwischen 2016 und 2021. Die Ergebnisse stützen die Ansicht, dass sich Vor- und Nachteile, die durch soziale Ungleichheiten entstehen, auch während der Grundschulzeit weiter akkumulieren. Unterschiede im Ausmaß der Akkumulation zwischen den Ländern werden diskutiert.

Vergleichendes Rating mit Benchmarktexten: Lassen sich damit Präsentationseffekte verhindern?

Afra Sturm, Nadja Lindauer, Markus Neuenschwander

Pädagogische Hochschule FHNW, Schweiz

Dass bestimmte Aspekte eines Textes die Rater*innen beim Beurteilen des Textes beeinflussen, ist empirisch gut belegt: Graham et al. (2011) zeigen in Bezug auf die Leserlichkeit der Handschrift, in Bezug auf Rechtschreib- und Grammatikfehler und den Produktionsmodus (handschriftlich vs. mit einem Textverarbeitungsprogramm verfasst) einen Präsentationseffekt von $d = -.38$ bis -1.03 . Solche Effekte sind für analytische und holistische Ratings nachgewiesen (z.B. Canz et al., 2020). Des Weiteren kann beobachtet werden, dass aktivierte Stereotypen dazu führen, dass die Beurteilung individueller Leistungen so verändert wird, dass sie eher dem Stereotyp entspricht: Im Bereich Schreiben gilt dies etwa bezogen auf Gender (Peterson, Childs & Kennedy, 2004), den sprachlichen Hintergrund / Nationalität (Rubin & Williams-James, 1997), den sozio-ökonomischen Status (Verheyden et al., 2012) oder Leistungsgruppen (Kuiken & Veder, 2014). Allerdings sind solche Effekte bisher noch nicht im Zusammenhang mit Präsentationseffekten untersucht worden.

Uns interessiert, ob Präsentations- und Stereotypeneffekte auch dann zu beobachten sind, wenn Benchmark-Texte den Ratingprozess anleiten (Feenstra, 2014): Wie für holistische Verfahren typisch, bilden Benchmark-Ratings ein globales Urteil über die Qualität des ganzen Textes. Die Besonderheit von Benchmark-Ratings besteht darin, dass dieses Gesamturteil vergleichend gefällt wird, d.h. jeder zu bewertende Text wird mit Benchmark-Texten verglichen, die unterschiedliche Qualitätsstufen repräsentieren. Diese Vergleichstexte werden anhand von Kriterien ausgewählt und mit Deskriptoren versehen, die auf diesen Kriterien basieren. Schließlich zeichnen sich Benchmark-Ratings dadurch aus, dass alle Texte von mehreren Rater*innen beurteilt werden und dann der Mittelwert der verschiedenen Beurteilungen als Maß für die Schreibleistung berechnet wird.

Da die Vergleichstexte mit einer Bewertungsskala verknüpft sind und somit spezifische Schreibleistungen darstellen, könnte man annehmen, dass der Präsentationseffekt keinen Einfluss hat. Mit Blick auf bisherige Studien zu analytischen und holistischen Beurteilungsverfahren gehen wir allerdings von folgenden Hypothesen aus:

H1: Jungen erhalten höhere Punktzahlen als Mädchen und im Vergleich zu den Originaltexten.

H2: Schüler*innen mit niedrigem Sozialstatus erzielen im Vergleich zu Schüler*innenn mit hohem Sozialstatus und im Vergleich zu den Originaltexten höhere Punktzahlen.

H3: Schüler*innen, die die Schulsprache nicht sprechen, erhalten höhere Punktzahlen im Vergleich zu Schüler*innen, die die Schulsprache sprechen, und im Vergleich zu den Originaltexten.

H4: Schüler*innen mit geringen Schreibleistungen erzielen höhere Werte als Schüler*innen mit hohen Schreibleistungen und im Vergleich zu den Originaltexten.

Die in unserer Studie untersuchten Texte umfassen narrative und instruktive Texte ($237 < N < 283$), die von Schüler*innen der Primarstufe (4. Klasse) in der deutschsprachigen Schweiz geschrieben wurden. Um mögliche Präsentationseffekte zu ermitteln, wurde jeder der handgeschriebenen Texte neu abgetippt, auf Rechtschreibung, Zeichensetzung und Grammatik korrigiert, anonymisiert und zufällig geordnet. Die Qualität sowohl der handschriftlichen Originaltexte als auch der abgetippten und sprachformal bereinigten Texte wurde mit fünf Vergleichstexten holistisch beurteilt ($ICC(C,3) = .94$ bis $.99$).

Die Hypothesen wurden mittels ANOVA mit Messwiederholungen unter Kontrolle der Klassenzugehörigkeit getestet. Die Hypothesen 1–3 konnten nur für die narrativen Texte, Hypothese 4 für narrative wie auch instruktive Texte bestätigt werden: Bei Jungen war der Präsentationseffekt stärker als bei Mädchen (narrativ: $F[1, 252] = 10.7$, $p < .001$, instruktiv: ns). Bei monolingualen Schüler*innen war der Präsentationseffekt stärker als bei multilingualen Schüler*innen (narrativ: $F[1, 231] = 4.6$, $p < .05$, instruktiv: ns). Bei Schüler*innen mit niedrigem SES war der Präsentationseffekt stärker als bei Schüler*innen mit hohem SES (narrativ: $F[1, 200] = 13.2$, $p < .001$, instruktiv: ns). Bei Schüler*innen mit geringer Schreibleistung war der Präsentationseffekt stärker als bei Schüler*innen mit hoher Schreibleistung (narrativ: $F[1, 252] = 107.4$, $p < .001$, instruktiv: $F[1, 231] = 56.8$, $p < .001$).

Das Abtippen und sprachformale Bereinigen von Texten sind also nicht nur für eine valide Bewertung wichtig, sondern auch für eine gleichberechtigte Bewertung aller Schüler*innen.

A Library in the Palm of your Hand? A Reading Intervention with Low-Income Children

Silke Anger^{1,2,6}, Bernhard Christoph¹, Galkiewicz Agata^{1,3}, Margaryan Shushanik³, Peter Frauke⁴, Sandner Malte^{1,5,6}, Siedler Thomas^{3,6}

¹Institut für Arbeitsmarkt- und Berufsforschung; ²Universität Bamberg; ³Universität Potsdam; ⁴DZHW; ⁵TH Nürnberg; ⁶IZA

Multiple studies document early emergence and persistence of educational achievement gaps by socioeconomic background. A key area in which students from lower-income backgrounds perform worse than students

from higher-income backgrounds is reading. Reading skills and text comprehension are crucial, as they predict overall academic achievement and further educational attainment. Moreover, they are predictive of long-term labor market success and wages.

This paper evaluates the effects of a randomized reading intervention among disadvantaged children across Germany on reading behavior, academic achievement, socioemotional well-being and non-cognitive skills. We provided e-book readers that include a one-year activated subscription to more than 1,000 age-appropriate books to children aged 11-12. Our study has several attractive features. First of all, it is highly scalable due to the straightforward delivery of e-readers with an activated library subscription. Second, it does not rely on parents or teachers. Third, we have a geographically representative sample of disadvantaged households, which mutes worries about spillover effects.

Our analysis one year after the start of the treatment shows that the children who received the e-book readers are in fact more likely to read. The treated children are 17 percentage points more likely to have read an e-book in the last month, compared to those in the comparison group. This corresponds to an increase of nearly 140 percent, evaluated at the mean of 12 percent of students in the control group who report having read an ebook in the last four weeks. We further find that this is not simply because children substitute printed books with e-books, as we find that they increase their overall reading. Our findings also show that the treatment group engages in reading more frequently, with children in this group being more likely to read on more than one day a week compared to the control children.

The intervention positively affects children's academic achievement and socio-emotional well-being. To measure academic achievement, we construct an index incorporating reading comprehension, German and math grades, and academic aspirations. The estimated intention-to-treat effect suggests a significant increase of 11 percent of a standard deviation. A closer examination of the individual components of this index reveals that the improvements are driven primarily by improvements in reading comprehension and math grades. Specifically, we find an 11 percent of a standard deviation increase in reading comprehension and a two percent increase in the probability of receiving a good math grade.

Additionally, we assess the impact of the intervention on children's socio-emotional wellbeing using the Strengths and Difficulties Questionnaire (SDQ). Socio-emotional skills are closely linked to student performance on standardized tests and are strong predictors of future educational attainment and labor market outcomes. Our analysis indicates that children who received the intervention scored ten percent of a standard deviation lower on the SDQ, indicating an improvement in their socio-emotional well-being.

Our findings on reading comprehension are on par with the findings in the meta study of summer reading interventions in Kim and Quinn (2013), who find an effect of 11–22 percent of a standard deviation and in Abeberese et al. (2014), who document an immediate improvement of 13 percent of a standard deviation in reading comprehension through a reading program. Compared to the literature on class size, our effects appear to be large. Most papers find that a reduction in class size by one person increases academic achievement between 1–3 percent of a standard deviation (Leuven and Oosterbeek 2018). This implies that the class size would have to be reduced by about four to 11 students to reach the same effect.

PS-37: Motivation in Hochschule und Freizeit

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 242

Chair der Sitzung: Martin Gartmeier

Kognitive Mechanismen der Lernvorbereitung: Offenes Problemlösen versus Selbsterklärung von Lösungsbeispielen

Katharina Ockl, Inga Glogger-Frey

Universität Erfurt, Deutschland

Die Bearbeitung einer Aufgabe zur Lernvorbereitung zu Beginn einer Lernsequenz kann Lernende auf anschließende Instruktionen vorbereiten und ihren Lernerfolg verbessern. Gerade in Bezug auf konzeptuelles Wissen und Transfer kann eine Kombination von offenem Problemlösen und anschließender Instruktion lernförderlich wirken (Sinha & Kapur, 2021). Die Wirkung der Lernvorbereitung kann durch drei kognitive Mechanismen (Vorwissensaktivierung, Wahrnehmung von Wissenslücken und Erkennen der Tiefenstruktur des Lerninhalts) erklärt werden. (Loibl et al., 2017). Der Vergleich verschiedener Lernvorbereitungsaktivitäten zeigt, dass offenes Problemlösen insbesondere die Wahrnehmung von Wissenslücken fördert (Glogger-Frey et al., 2015; Newman & DeCaro, 2019), welche bei der anschließenden Instruktion gezielt adressiert und geschlossen werden können. Andererseits kann das Erkennen der Tiefenstruktur durch das Nachvollziehen und Selbsterklären der ausgearbeiteten Version des Problems (ähnlich einem Lösungsbeispiel) besser gefördert werden als wenn Lösungen selbst generiert werden müssen, was ebenfalls Lernprozesse während der anschließenden Instruktion vertiefen kann (Glogger-Frey et al., 2022; Roelle & Berthold, 2015). Zusammengefasst können sowohl offenes Problemlösen als auch ausgearbeitete Lösungsbeispiele lernrelevante Prozesse unterstützen, wobei sie zum Teil durch verschiedene Mechanismen wirken.

Die Ziele dieser Studie sind, die während der Lernvorbereitung stattfindenden kognitiven Prozesse möglichst nah am Verhalten zu erfassen, Unterschiede zwischen Lernvorbereitungsaktivitäten zu betrachten (offenes Problemlösen vs. Selbsterklärung der Lösung), sowie Lernergebnisunterschiede durch verschiedene Lernvorbereitungsprozesse zu erklären. Wir erwarten mehr wahrgenommene Wissenslücken (H1a) und weniger erkannte Tiefenstruktur (H1b) beim offenen Problemlösen im Vergleich zum Selbsterklären der Lösung. Außerdem erwarten wir, dass die wahrgenommenen Wissenslücken und die erkannte Tiefenstruktur den Effekt der Lernvorbereitung auf die Lernergebnisse medieren (H2).

Methode. An der Studie nahmen 65 Erstsemesterstudierende (91% weiblich; Alter: $M = 20.1$, $SD = 2.8$) teil und bearbeiteten entweder eine offene Problemlöseaufgabe („Inventing“, $n = 30$) oder die ausgearbeitete Lösung („Lösungsbeispiel“, $n = 35$). Das Lernmaterial bestand aus kontrastierenden Beispielen (contrasting cases) von Verteilungen, die sich systematisch in ihrer Varianz unterschieden (basierend auf Wiedmann et al., 2012). Beim Inventing sollte zunächst intuitiv die Varianz eingeschätzt und anschließend eine Formel zur Berechnung entwickelt werden, während beim Lösungsbeispiel ein vorgegebener Lösungsversuch nachvollzogen werden sollte. Zur Erfassung der Prozessvariablen sprachen die Versuchspersonen während der Bearbeitung alle Gedanken laut aus (think-aloud nach Ericsson & Simon, 1980). Anschließend wurde ein Instruktionsvideo gezeigt und mithilfe eines Posttests (basierend auf Items von Kapur, 2012; Trninic et al., 2022) naher und weiter Transfer erfasst.

Ergebnisse. Wie erwartet nahm die Inventinggruppe signifikant mehr Wissenslücken wahr, $t(63) = 4.86$, $p < .001$, $d = 1.21$ und verarbeitete die Tiefenstruktur weniger tief, $t(63) = -8.14$, $p < .001$, $d = -2.02$. Beim weiten Transfer war die Inventinggruppe signifikant besser, $t(63) = 3.20$, $p = .002$ (zweiseitig), $d = .80$, nicht aber beim nahen Transfer, $t(63) = .69$, $p = .494$ (zweiseitig), $d = .17$. Bei den erwarteten Mediationen zeigte sich ein indirekter Effekt der Wissenslücken, $ab = .54$, $SE = .23$, 95% CI [0.14, 1.04], nicht aber der Tiefenstruktur, $ab = -.22$, $SE = .43$, 95% CI [-1.09, 0.62] auf den weiten Transfer. Für den nahen Transfer zeigten sich keine indirekten Effekte, Wissenslücken: $ab = -.42$, $SE = .17$, 95% CI [-0.90, 0.06]; Tiefenstruktur: $ab = .49$, $SE = .24$, 95% CI [-0.15, 1.22].

Zusammengefasst zeigten sich die erwarteten Unterschiede in den Prozessvariablen. Verschiedene Lernvorbereitungsaktivitäten regen folglich unterschiedliche kognitive Prozesse besonders an, was bei der Auswahl und Gestaltung von Aufgaben berücksichtigt werden sollte, um gezielt für den Lerninhalt besonders relevante Prozesse zu fördern. Obwohl Tiefenstrukturaspekte durch das Lösungsbeispiel tiefer verarbeitet wurden, war dies kein Mediator für spätere Lernergebnisse. Eine genauere Betrachtung der Think-Aloud-Daten zeigte, dass Tiefenstrukturaspekte durch das Lösungsbeispiel zwar besser wahrgenommen, aber nicht elaboriert wurden. Möglicherweise wurde deshalb keine Auswirkung auf die Lernergebnisse gefunden. Dies legt eine sorgfältige Gestaltung der Lösungsbeispiele nahe, um Elaborationsprozesse stärker anzuregen.

Ohne Fleiß kein Preis: Eine empirische Studie zur Komplexität wünschenswerter Erschwernisse

Ines Zeitlhofer, Joerg Zumbach

Paris Lodron Universität Salzburg, Österreich

Theoretischer Hintergrund und Fragestellung

Das Hauptaugenmerk der Cognitive Load Theory (CLT; Sweller, 2020, 2023) liegt auf dem Instruktionsdesign und der dadurch entstehenden kognitiven Belastung. Die Gestaltungsrichtlinien der CLT basieren auf der Annahme, dass die Kapazität des Arbeitsgedächtnisses für die Verarbeitung neuer Informationen begrenzt ist. Wird dieses Kapazitätslimit überschritten, kann sich das negativ auf die Lernergebnisse auswirken (van Gog et al., 2011). Um erfolgreiches Lernen optimal zu fördern, geht die CLT davon aus, dass die Komplexität von Aufgaben sukzessive gesteigert werden sollte (Kalyuga, 2014). Dennoch scheint eine gewisse mentale Anstrengung notwendig zu sein, um Informationen langfristig zu speichern (Seufert, 2018). Bjork und Bjork (2020) postulieren im Rahmen des Desirable Difficulty Frameworks, dass Aufgaben, die einen aktiven und kognitiv anspruchsvollen Abruf von Informationen aus dem Arbeitsgedächtnis fördern, den Einsatz metakognitiver Strategien erhöhen und in Folge zu besseren Lernergebnissen führen. Eine Möglichkeit, dies zu erreichen, ist der Einsatz wünschenswerter Erschwernisse, wie das verschachtelte Lernen. Das verschachtelte Lernen (ABCABCABC) wird meist dem geblockten Lernen (AAABBBCCC) gegenübergestellt. Da wünschenswerte Erschwernisse jedoch die kognitive Belastung erhöhen können, nimmt die Forschung an, dass deren Anwendung bei Aufgaben mit konstant hoher Komplexität zur kognitiven Überlastung führt (Chen et al., 2018).

Die vorliegende Studie testete diese Annahme und untersuchte, wie die Komplexität von Aufgaben mit wünschenswerten Erschwernissen gestaltet werden soll. Genauer wurde analysiert, wie sich die Interaktion von Komplexität (konstant hoch vs. sukzessive ansteigend) und wünschenswerter Erschwernis (verschachtelt vs. geblockt) auf die Lernergebnisse und die Selbstwirksamkeitserwartung auswirkt. Zudem wurde untersucht, ob der Einsatz metakognitiver Strategien mit der kognitiven Belastung im Zusammenhang steht.

Methode

An der Studie nahmen 132 Studierende teil. Die Untersuchung umfasste eine Hauptsitzung und einen Nachtest, der zwei Tage später durchgeführt wurde. Die Teilnehmenden wurden per Zufallsprinzip in vier Gruppen eingeteilt: verschachtelt-hoch, verschachtelt-ansteigend, geblockt-hoch und geblockt-ansteigend. Als Intervention diente eine Konjugationsaufgabe zum italienischen Verb „parlare“ (Deutsch: sprechen). In der Hauptsitzung lernten die Teilnehmenden, das italienische Verb in sechs Temporal- und Modalformen zu konjugieren. Alle Gruppen übten die Konjugation in sechs Runden. Die Komplexität wurde durch die Anzahl der zu ergänzenden Endungen manipuliert. Die Manipulation der wünschenswerten Erschwernis erfolgte durch die Reihenfolge, in der die Temporal- und Modalformen geübt wurden. Im Nachtest ergänzten alle Teilnehmenden die Endungen aller Temporal- und Modalformen einmal. Die Lernergebnisse wurden anhand der im Nachtest benötigten Zeit und der Anzahl der Versuche gemessen. Die kognitive Belastung, der Einsatz metakognitiver Strategien und die Selbstwirksamkeitserwartung wurden nach jeder Runde der Haupttestung sowie nach dem Nachtest durch Selbstevaluierungsfragebögen erhoben.

Ergebnisse und ihre Bedeutung

Eine multivariate Varianzanalyse ergab einen signifikanten Effekt von Komplexität (Zeit: $\eta_p^2 = .09$; Versuche: $\eta_p^2 = .27$) und wünschenswerter Erschwernis (Zeit: $\eta_p^2 = .09$; Versuche: $\eta_p^2 = .54$) auf die Lernergebnisse, nicht jedoch deren Interaktion. Auch der Einfluss von Komplexität, wünschenswerter Erschwernis sowie deren Interaktion auf die Selbstwirksamkeitserwartung war nicht signifikant. Eine lineare Regressionsanalyse ergab, einen signifikanten Zusammenhang zwischen der kognitiven Belastung und dem Einsatz metakognitiver Strategien ($p = 0,001$). Die deskriptiven Daten zeigten, dass die Kombination von konstant hoher Komplexität und wünschenswerter Erschwernis zu den besten Lernergebnissen führt, ohne sich negativ auf die Selbstwirksamkeitserwartung auszuwirken. Eine weitere deskriptive Analyse ergab, dass die kognitive Belastung und der Einsatz metakognitiver Strategien bei sukzessiv ansteigender Komplexität insgesamt höher war als bei konstant hoher Komplexität.

Zusammengefasst deuten die Ergebnisse der Studie darauf hin, dass sich wünschenswerte Erschwernisse und eine konstant hohe Komplexität positiv auf die Lernergebnisse auswirken können, ohne die Selbstwirksamkeitserwartung zu beeinflussen. Auch die Interaktion von konstant hoher Komplexität und wünschenswerter Erschwernis kann die Lernergebnisse verbessern, wenn auch nicht signifikant. Des Weiteren wurde veranschaulicht, dass eine konstant hohe Komplexität insgesamt zu einer geringeren kognitiven Belastung und damit auch zu einem niedrigeren Einsatz metakognitiver Strategien führt.

Förderung des Engagements von Making Aktivitäten durch Autonomie und Nutzen: Eine experimentelle Video-Vignetten-Studie

Sophie Uhing, Stephan Kröner

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Während jüngere Kinder noch relativ leicht für Naturwissenschaften und Technik zu begeistern sind, nimmt dieses Interesse ab einem Alter von 12 Jahren kontinuierlich ab, insbesondere bei Mädchen (Sadler et al., 2012; Trotman, 2017). Die Förderung von Freizeitaktivitäten im MINT-Bereich (Mathematik, Informatik, Naturwissenschaften und Technik) könnte dem entgegenwirken (Posner & Vandell, 1999). Dazu gehört auch das sogenannte Making, d.h. das Nutzen von 3D-Druckern, Lasercuttern oder Mikrocontrollern zur Herstellung eigener individueller Objekte. Entsprechende Aktivitäten werden in Makerspaces bzw. Fab Labs ermöglicht. Sie werden vielfach als besonders geeignet dafür betrachtet, um Jugendlichen MINT-Bildungserfahrungen zu ermöglichen (vgl. die Übersichtsarbeit von Smolarczyk & Kröner, 2023). Insofern stellt sich die Frage, wie Kinder und Jugendliche Making-Aktivitäten (nicht) spannend finden und solchen nachgehen möchten, wurde basierend auf der Theorie des geplanten Verhaltens eine experimentelle Videovignettenstudie durchgeführt (TPB; Ajzen, 1991). In dem 2 x 2 faktoriellen Between-Subjects-Design wurden Videos über Making-Angebote präsentiert, welche unterschiedliche Ausprägungen auf den Faktoren Nutzen (niedrig vs. hoch) und Autonomie (niedrig vs. hoch) widerspiegeln. In der präregistrierten Studie wurde erwartet, dass eine hohe Ausprägung der Faktoren Nutzen und Autonomie zu höheren Ausprägungen in den Variablen intrinsischer Wert, persönliche Wichtigkeit und Teilnahmeabsicht führen würde.

Methode

An der Studie nahmen im Juli 2024 $N = 223$ Schülerinnen und Schüler aus drei bayrischen Mittelschulen der 5. bis 8. Klassen teil (Alter $M = 12.33$, $SD = 1.30$). Von den Schülern gaben je $n = 96$ an, weiblich bzw. männlich zu sein, $n = 31$ machten keine Angabe zum Geschlecht.

Ergebnisse

Ein Manipulationscheck zeigte auf, dass die Manipulation des Faktors Nutzen für verschiedene Facetten der Variablen funktionierte (Nutzen für die Zukunft $F(2, 217) = 4.36$, $p = .006$; Nutzen für den Alltag $F(2, 217) = 4.37$, $p = .003$; Nutzen in Bezug auf die Kreation von Gegenständen $F(2, 217) = 4.41$, $p = .009$), für die Manipulation des Faktors Autonomie konnte dies nicht belegt werden ($F(2, 217) = 2.34$, $p = .28$).

Weitere vorläufige Analysen zeigten, dass weder die Kovariablen Geschlecht noch Alter einen signifikanten Erklärungswert für die abhängigen Variablen intrinsischer Wert ($F(4, 175) = 1.15$, $p_{\text{Geschlecht}} = .38$, $p_{\text{Alter}} = .59$), persönliche Wichtigkeit ($F(4, 175) = 2.43$, $p_{\text{Geschlecht}} = .61$, $p_{\text{Alter}} = .09$) und Teilnahmeabsicht ($F(4, 175) = 1.33$, $p_{\text{Geschlecht}} = .56$, $p_{\text{Alter}} = .46$) aufwiesen. In einer weiteren Regressionsanalyse, in der intrinsischer Wert, persönliche Wichtigkeit und Teilnahmeabsicht zu t_1 berücksichtigt wurden, zeigte sich, dass der Faktor Nutzen signifikant zur Erklärung der Varianz in der persönlichen Wichtigkeit zu t_2 beiträgt ($F(11, 208) = 39.47$, $p = .04$, $b = 0.59$ $SD = 0.28$, $R^2 = .68$).

Diskussion

Obwohl Autonomie eine wichtige Stellschraube zur Erhöhung der Motivation und des Engagements von Jugendlichen darstellt (Reeve & Jang, 2006), scheint dies bei dem lediglich kurzen Anschauen eines Videos zu einem Making-Workshop nicht ausschlaggebend für die Erhöhung der abhängigen Variablen persönlichen Wichtigkeit, des intrinsischen Wertes oder der Teilnahmeabsicht zu sein. In Anbetracht der Tatsache, dass ein Gender Gap in den MINT-Fächern besteht, ist es ferner erstaunlich, dass Geschlecht keinen Erklärungswert für die abhängigen Variablen aufwies. Dahingegen konnte gezeigt werden, dass das explizite Aufzeigen des Nutzens von Making-Aktivitäten zusätzlich Varianz erklären konnte in der Variablen persönliche Wichtigkeit. Dies deckt sich mit Ergebnissen aus Studien, welche die Bedeutung des Nutzens für die Motivation in naturwissenschaftlichen Fächern hervorheben (Asher et al., 2023). Inwiefern eine einmalige Intervention in Form eines Videos tatsächlich zu höherem Engagement führt, müsste in folgenden Studien evaluiert werden. Limitationen in Bezug auf die Videos sowie Implikationen für die Werbung von Maker-Aktivitäten werden diskutiert.

Wartezeit nach Dozierendenfragen im Kontext fallbasierter medizinischer Seminare: Neue Perspektiven durch Kaplan-Meier Analysen

Martin Gartmeier, Janina Häusler, Marc Grünewald, Alexander Hapfelmeier, Theresa Pfurtscheller, Tina Seidel, Pascal Berberat

TU München, Deutschland

Theoretischer Hintergrund

Interaktive Seminare, zumal in Kleingruppen, sind ein gängiges Format in der medizinischen Hochschullehre, das durch curriculare Reformen eher noch an Bedeutung gewinnt. Daher ist es für die medizindidaktische Bildungsforschung eine wichtige Zielsetzung besser zu verstehen, wie interaktive medizinische Lehre effektiv gestaltet werden kann. Unser Fokus liegt hier auf Frage-Antwort Interaktionen zwischen Dozierenden und Studierenden. Kognitiv anspruchsvolle, an das Kompetenzniveau von Lernenden angepasste Fragen sind ein einfaches und wichtiges Mittel zur kognitiven Aktivierung der Lernenden als zentrale Qualitätsdimension von Unterricht.

Ein anerkanntes Qualitätsmerkmal „guter“ Fragen in interaktivem Unterricht sind Wartezeiten, die Lehrende gewähren nachdem sie Lernenden eine Frage gestellt haben (Teacher Wait Time, TWT). Durch TWT bekommen Studierende Zeit um Fragen kognitiv zu verarbeiten und – gerade bei komplexen Fragen – Ideen für mögl. Antworten zu entwickeln. Zudem entsteht durch TWT ein gewisser sozialer Druck, auch Antworten zu teilen bzgl. derer man sich nicht 100% sicher ist. Studien aus dem Hochschulkontext zeigen, dass Lehrende häufig nur ca. zwei Sekunden TWT gewähren, bevor sie eine Frage z.B. umformulieren oder selbst beantworten; entsprechende Empfehlungen gehen eher in die Richtung von 5 Sekunden und mehr. Für die Medizindidaktik liegt bislang quasi keine Evidenz für gängige Lehrformate vor. Daher berichten wir eine Videostudie im Bereich Fallbasierter Lehre im klinischen Teil des Medizinstudiums, in den Fachrichtungen Chirurgie und Innere Medizin. Wir fokussieren folgende Fragestellung:

RQ 1: Auf wie viele Lehrerfragen folgte eine TWT und wie lange dauerte die TWT im Median?

RQ 2: Wie viel TWT sollten klinische Lehrkräfte nach initialen Fragen im Vergleich zu wiederholten/Folgefragen bereitstellen, um hohe Antwortraten zu erreichen?

Methode

Im Rahmen der Studie wurden 32 fallbasierte klinische Seminare aus den Bereichen Innere Medizin und Chirurgie videographiert. Die Seminare wurden von 21 verschiedenen Lehrenden durchgeführt, insgesamt nahmen 230 Studierende des siebten / achten Semesters daran teil. Die Videoaufnahmen wurden mithilfe von Videoanalyse-Software von vier Personen auf Basis eines theoriebasierten Kategoriensystems mit zufriedenstellender Reliabilität codiert (Cohen's kappa = 0.65.). Wartezeiten wurden erst ab einer Dauer von einer Sekunde nach einer Lehrendenfrage als solche codiert. Anders als bestehende Studien zu TWT nutzten wir Kaplan-Meier-Analysen um die Frage die hypothetischen Wahrscheinlichkeiten zu beschreiben, dass die TWT von einem Studierenden beendet wird.

Ergebnisse

RQ 1: Insgesamt beobachteten wir $n = 719$ Lehrerfragen, gefolgt von einer Wartezeit. Die mittlere Wartezeit lag bei $Mdn = 3,40$ s ($M = 4,50$ s; $IQR = 2,22\text{--}5,68$ s; $SD = 3,49$ s). Wir fokussieren die Abfolgen von (initialen) Dozierendenfragen – Wartezeit – Schülerantworten, dieses Muster beobachteten wir in 328 Fällen und die mittlere Wartezeit betrug hier $Mdn = 3,18$ s ($M = 4,41$ s, $IQR = 2,32\text{--}6,26$ s, $SD = 3,74$ s).

RQ 2: Die Wahrscheinlichkeit, dass eine initiale Dozierendenfrage nach einer TWT von fünf Sekunden beantwortet wurde lag bei 36,4 % (vgl. Abb. 1). Nach einer TWT von zehn Sekunden stieg dieser Wert weiter auf 65,8 %. Zudem stieg die anfängliche Antwortrate auf Lehrerfragen in kontinuierlich bis etwa zur 12-Sekunden-Schwelle an (wenn etwa 75,4 % der Fragen beantwortet wurden), danach beginnt die Kurve abzufallen. Nach einer TWT von ca. 12 Sekunden erhöht sich die Wahrscheinlichkeit, dass ein Studierender antwortet immer noch – aber in viel geringerem Maße als zuvor.

Diskussion

Die Ergebnisse geben für den Kontext der Medizindidaktik TWT einen erstmaligen Einblick in die Rolle und den Umfang von TWT, hier im Kontext fallbasierter Seminare. Der methodische Zugang der Studie über Kaplan-Meier-Analysen ist innovativ im Vergleich zu bestehenden Studien, da er es erlaubt Zusammenhänge zwischen verschiedenen langen Wartezeiten und entsprechenden Antwortwahrscheinlichkeiten im Sinne von Wahrscheinlichkeiten zu quantifizieren.

PS-38: Motivation und Selbstwirksamkeit im Mathematiklernen: Interventionen und ihre Wirkung

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 157

Chair der Sitzung: Janina Krawitz

Steigert das Entwickeln eigener Mathematikaufgaben die Motivation? Effekte einer Unterrichtsintervention auf Autonomieerleben, Kompetenzerleben, soziale Eingebundenheit und Selbstwirksamkeit

Janina Krawitz¹, Lars Meyer-Jenßen², Katharina Wiehe³, Katrin Rakoczy⁴

¹Universität Köln, Deutschland; ²Humboldt-Universität zu Berlin; ³Universität Münster; ⁴Universität Gießen

Laut der Selbstbestimmungstheorie ist die Erfüllung der psychologischen Grundbedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit entscheidend für selbstbestimmte Motivation (Ryan & Deci, 2020) und motivationale Faktoren wie die Selbstwirksamkeit (Bandura, 1997). Unterricht sollte daher die Erfüllung von Grundbedürfnissen fördern. Ein Interventionsansatz im Fach Mathematik besteht darin, Lernende eigene Aufgaben entwickeln zu lassen. Erste Ergebnisse weisen auf eine motivationssteigernde Wirkung dieses Ansatzes hin (Cai & Leikin, 2020). Unklar ist jedoch, wie sich das Entwickeln eigener Aufgaben auf psychologische Grundbedürfnisse und Selbstwirksamkeit auswirkt und ob die Grundbedürfnisse den Effekt auf die Selbstwirksamkeit vermitteln.

Hypothesen

1. Das Entwickeln eigener Aufgaben wirkt sich positiv auf das Erleben von Autonomie, Kompetenz, sozialer Eingebundenheit und Selbstwirksamkeit aus.
2. Das Erleben von Autonomie, Kompetenz und sozialer Eingebundenheit vermittelt den positiven Effekt des Entwickelns eigener Aufgaben auf die Selbstwirksamkeit.

Positive Effekte sind zu erwarten, da das Entwickeln eigener Aufgaben die Möglichkeit bietet, eigene Lernprozesse selbst zu gestalten (Silver, 1994), was das Autonomieerleben steigern könnte. Zudem kann das Entwickeln eigener Aufgaben auf verschiedenen Kompetenzniveaus erfolgen (Walkington et al., 2022), was sich positiv auf das Kompetenzerleben auswirken könnte. Das Erleben sozialer Eingebundenheit könnte profitieren, wenn Lernende Aufgaben für Mitschüler:innen stellen oder deren Aufgaben bearbeiten. Ein positiver Effekt auf die Selbstwirksamkeit wird durch die Möglichkeit angenommen, sich beim Entwickeln eigener Aufgaben eigene Ziele zu setzen und persönliche Erfolgserfahrungen zu haben (Bonotto & Santo, 2015). Nach der Selbstbestimmungstheorie (Ryan & Deci, 2020) und der Selbstwirksamkeitstheorie (Bandura, 1997) ist zu erwarten, dass die Erfüllung psychologischer Grundbedürfnisse den positiven Effekt des Entwickelns eigener Aufgaben auf die Selbstwirksamkeit vermittelt.

Methode

Insgesamt wurden $N=194$ Zehntklässler:innen (89 weiblich, 103 männlich, 2 divers, $M=15.47$ Jahre) randomisiert zwei Bedingungen zugewiesen. Die Experimentalgruppe (EG) erhielt eine vierstündige Unterrichtsintervention zum eigenständigen Entwickeln und Lösen realitätsbezogener Mathematikaufgaben, während die Kontrollgruppe (KG) nur das Lösen realitätsbezogener Aufgaben behandelte. Aufgabenkontexte, Interventionsdauer und Unterrichtsmethoden waren identisch. Der Unterricht wurde von geschulten Masterstudierenden nach standardisierten Instruktionsmanualen durchgeführt. Jede:r Masterstudierende unterrichtete beide Bedingungen gleich häufig. Die adäquate Durchführung der Intervention wurde durch Beobachtungsbögen überprüft. Das Erleben von Autonomie, Kompetenz und sozialer Eingebundenheit wurde mittels erprobter Skalen während der Unterrichtsintervention mit je 3 Items zu vier Messzeitpunkten erfasst und danach über die Messzeitpunkte hinweg aggregiert. Die Selbstwirksamkeit wurde vor und nach der Unterrichtsintervention gemessen (4 Items). Die internen Konsistenzen der eingesetzten Skalen waren zufriedenstellend (Cronbachs $\alpha > .72$). Die Ergebnisse wurden mit Pfadmodellen (MPlus8.10) berechnet (MLR-Schätzer, CFI > 0.957 , SRMR < 0.029). Fehlende Werte wurden mittels FIML geschätzt.

Ergebnisse und Diskussion

Entgegen den Erwartungen hatte das Entwickeln eigener Aufgaben keinen positiven Effekt auf die Selbstwirksamkeit im Nachtest unter Vortestkontrolle (totaler Effekt, $\beta = -0.04$, $p = .37$). Allerdings wirkte sich das Entwickeln eigener Aufgaben positiv auf das Erleben von Autonomie ($\beta = 0.34$, $p = .01$) und sozialer Eingebundenheit ($\beta = 0.28$, $p = .03$), jedoch nicht auf das Kompetenzerleben ($\beta = 0.05$, $p = .36$) aus. Das Erleben von Autonomie ($\beta = 0.14$, $p < .05$), Kompetenz ($\beta = 0.23$, $p < .001$) und sozialer Eingebundenheit ($\beta = 0.14$, $p < .001$) beeinflusste wiederum die Selbstwirksamkeit positiv. Das Autonomieerleben, nicht jedoch Kompetenzerleben oder soziale Eingebundenheit, erwies sich als Mediator, indem es den positiven Effekt des Entwickelns eigener Aufgaben auf die Selbstwirksamkeit vermittelte (indirekter Effekt, $\beta = 0.05$, $p = .04$).

Die Ergebnisse zeigen, dass das Entwickeln eigener Aufgaben Autonomieerleben und soziale Eingebundenheit steigern kann. Der fehlende Effekt auf das Kompetenzerleben weist auf die Herausforderungen dieser Tätigkeit hin. Im Einklang mit der Selbstbestimmungstheorie erwies sich Autonomieerleben als wichtige vermittelnde Variable. Eine praktische Implikation der Studie ist, dass durch das Entwickeln eigener Aufgaben im Mathematikunterricht die Autonomie der Lernenden gefördert wird, was die Selbstwirksamkeit positiv beeinflussen kann. Diese Erkenntnisse tragen zur Forschung über Motivation und motivationsfördernde Interventionen bei, indem Beziehungen zwischen dem Entwickeln eigener Aufgaben, der Erfüllung psychologischer Grundbedürfnisse und Selbstwirksamkeit aufgezeigt werden.

Mathe-Lust statt Mathe-Frust? Überprüfung der Wirksamkeit einer Intervention zur Motivationsförderung im Mathematikunterricht

Gabriele Brasser, Christina Drüke-Noe, Sonja Bieg

Pädagogische Hochschule Weingarten, Deutschland

Motivationale Defizite bei Lernenden gelten als besonders problematisch, um die kontinuierlichen Herausforderungen an Hochschulen und Schulen zu bewältigen (Zepeda et al., 2020). In jüngerer Zeit mehrte sich die Anzahl an Studien, welche diesbezüglich die motivationale Kompetenz von Lernenden thematisieren (Satka & Garneva, 2024; Schaller & Spinath, 2017). Die aktuelle Studie greift dieses Thema auf und konzipiert ein Training für die Schule, das verschiedene motivationale Ansätze kombiniert, um vielfältige Formen der adaptiven Motivation zu unterstützen. Der Fokus liegt auf der Integration von drei sich ergänzenden Motivationstheorien: Dem Konzept der motivationalen Kompetenz (Spinath et al., 2018), das motivationsbezogene Kompetenzen (motivationale, volitionale, bewertende) bei Lernenden thematisiert, der Selbstbestimmungstheorie (Ryan & Deci,

2000) sowie dem Erwartung x Wert-Modell (Eccles et al., 1983). Auch wenn motivationale Fördermaßnahmen zu einzelnen Theorien Erfolge verbuchen (z.B. Hulleman et al., 2010; Mittag et al., 2009) bleibt bislang ungeklärt, ob eine Intervention, die zentrale Komponenten verschiedener Ansätze integriert sowie die motivationalen Kompetenzen von Lernenden stärken möchte, erfolgreich Motivation und Leistung von Lernenden fördern kann.

Die Studie verfolgt das Ziel über eine Intervention im regulären Matheunterricht, die motivationalen Kompetenzen der Lernenden zu fördern (Hypothese 1). Zudem sollen die Grundbedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit unterstützt werden (Hypothese 2). Durch die Intervention sollen die Lernenden eine positivere Einstellung zur Schule entwickeln (Hypothese 3) und eine bessere Leistung in Mathematik zeigen als Lernende in der Kontrollgruppe (Hypothese 4).

Methode

Im Rahmen einer quasi-experimentellen Studie in Mathematik mit insgesamt 17 achten Klassen der Realschule wurde über einen Zeitraum von 10 Wochen eine Intervention mit zwei Unterrichtseinheiten (Prozentrechnen und Prismen mit je 20 Mathestunden) durchgeführt. Die Evaluation erfolgte in einem Treatment-Kontrollgruppen-Design (sieben Kontrollklassen). Die Stichprobe setzte sich aus insgesamt 235 Schüler*innen (43% weiblich) an Realschulen in Baden-Württemberg im Alter von $M=13.56$, ($SD=0.61$, Min 13, Max 16 Jahren) zusammen. Anhand von drei Messzeitpunkten (Prätest, Zwischenevaluation, Posttest) sowie einer Follow-up Erhebung sechs Wochen nach der zweiten Unterrichtseinheit zur Leistungsüberprüfung, wurden die Daten mit einer zweifaktoriellen Varianzanalyse mit Messwiederholung ausgewertet. Um motivationale Kompetenzen zu erfassen, wurde das Messinstrument MOBEKO (Schaller & Spinath, 2017, 59 Items, $\alpha = .62 - .91$) in einer adaptierten Form eingesetzt. Für Autonomieunterstützung (Black & Deci, 2000; 6 Items, $\alpha = .79$), soziale Eingebundenheit (Kunter et al., 2003; 6 Items, $\alpha = .82$), Selbstwirksamkeit (Jerusalem & Satow, 1999; 4 Items, $\alpha = .69 - .72$) und positive Einstellung zur Schule (Hascher & Hagenauer, 2018; 16 Items, $\alpha = .72 - .83$).

Ergebnisse

Es zeigten sich positive Effekte zugunsten der Treatmentklassen bezüglich der motivationalen Kompetenz zur *Herstellung von Passung* $F(2, 466) = 5.478$, $p = .005$ und der volitionalen Kompetenzen *Atmung und Konzentration* $F(2, 464) = 3.180$, $p = .042$ sowie *Zukunftsgedanken* $F(2, 464) = 4.448$, $p = .014$. Mit Bezug auf die Grundbedürfnisse der Selbstbestimmungstheorie zeigte sich lediglich ein positiver Effekt der Kompetenz (Selbstwirksamkeit) zugunsten der Mädchen der Treatmentgruppe. Die Schüler*innen der Treatmentgruppe berichteten über eine positivere Einstellung zur Schule $F(1, 233) = 8.428$, $p = .04$ im Vergleich zur Kontrollgruppe. Die Überprüfung der mathematischen Leistungen im Follow-up Test verweist auf bessere Leistung in den Treatmentklassen als in den Kontrollklassen $t(370.01) = -10.13$, $p < .001$, $d = 1.01$.

Diskussion

Das Ziel der Studie, die Motivation der Schüler*innen im Mathematikunterricht der achten Klasse zu fördern, konnte nur bedingt erreicht werden. Obwohl sich einzelne Aspekte der motivationalen Kompetenz bei Schüler*innen der Treatmentgruppe positiv entwickelten, blieb ein allgemeiner Effekt der Motivation aus. Dennoch trug die Intervention dazu bei, dass die Schüler*innen Schule positiver wahrnehmen und bessere Leistungen zeigten als die Kontrollgruppe. Diskutiert wird, ob die vielfältigen Maßnahmen im Unterricht zu einer Überforderung der Schüler*innen beitragen und inwiefern die eingesetzten Instrumente, die Motivation der Lernenden valide erfassen.

Autonomieerleben, Kompetenzerleben und ihre Wechselwirkungen: Die Bedeutung von Unterricht mit offenen Aufgaben

Jascha Quarder¹, Stanislaw Schukajlow¹, Janina Krawitz², Katharina Wiehe¹, Katrin Rakoczy³

¹Universität Münster, Deutschland; ²Universität zu Köln, Deutschland; ³Universität Gießen, Deutschland

Theorie

Die Selbstbestimmungstheorie betont die Bedeutung von Autonomie und Kompetenz als zentrale psychologische Grundbedürfnisse des Menschen (Ryan & Deci, 2020). Für den Mathematikunterricht wurde festgestellt, dass Autonomieerleben (AE) und Kompetenzerleben (KE) während des Lernens Effekte der Instruktion auf die Motivation der Schüler:innen vermitteln (Schukajlow & Krug, 2014). Es ist jedoch wenig darüber bekannt, wie sich AE und KE während des Lernprozesses gegenseitig beeinflussen und welche spezifischen Faktoren diese vorhersagen. Solch ein wichtiger Faktor kann der Unterricht mit offenen Aufgaben sein. Offene Aufgaben kommen häufig in der Realität vor und zeichnen sich durch das Fehlen lösungsrelevanter Informationen aus. Sie bieten die Möglichkeit, individuelle Annahmen zu treffen (Yeo, 2017), so dass das Bearbeiten von offenen Aufgaben AE und KE positiv beeinflussen kann. Es bedarf jedoch weiterer Forschung, um die Wechselwirkung von AE und KE sowie den Einfluss des Unterrichts mit offenen Aufgaben auf das Erleben der Grundbedürfnisse besser zu verstehen.

Forschungsfragen

FF1: Welche Wechselwirkungen bestehen zwischen AE und KE während des Lernens?

FF2: Inwieweit wird das AE und KE durch den Unterricht mit offenen Aufgaben beeinflusst?

Methode

Zur Beantwortung der Forschungsfragen wurde eine experimentelle Interventionsstudie mit 259 Neuntklässler:innen (103 weiblich, $M=14.5$ Jahre) aus Realschulen und einem Gymnasium durchgeführt. Innerhalb jeder Klasse wurden die Schüler:innen randomisiert einer Experimentalgruppe (EG) und einer Kontrollgruppe (KG) zugeordnet. In beiden Gruppen wurde über vier Mathematikstunden eine Einheit zum Satz des Pythagoras unterrichtet, wobei in der EG offene Aufgaben und in der KG geschlossene Aufgaben eingesetzt wurden (bei geschlossenen Aufgaben sind alle lösungsrelevanten Informationen vorhanden). Während des Unterrichts wurden das AE und das KE der Lernenden in beiden Gruppen zu fünf Messzeitpunkten mit jeweils einer Skala (AE: drei Items; KE: zwei Items) erhoben (Interne Konsistenz für AE: .770 bis .918 und für KE: .690 bis .845). Die erhobenen Daten wurden anschließend mittels eines Cross-Lagged-Panel-Modells ausgewertet. Dabei wurden AE und KE zu den fünf Messzeitpunkten als latente Variablen modelliert und neben Synchronkorrelation, Autoregression und Cross-Lagged-Effekten auch verzögerte Cross-Lagged-Effekte (Lüdtke & Robitzsch, 2021) sowie Regressionseffekte der Kovariate *Gruppe* berücksichtigt. Die Analysen wurden mit dem Package *lavaan* in *RStudio* durchgeführt, wobei fehlende Werte mit FIML geschätzt wurden. Die Güte des spezifizierten Modells war ausreichend hoch ($CFI = .964$; $SRMR = .043$).

Ergebnisse und Diskussion

Wie erwartet, lagen zu den Messzeitpunkten T1, T3, T4 und T5 signifikante und positive Synchronkorrelationen zwischen AE und KE vor (.247 - .499). Darüber hinaus zeigten sich für beide Konstrukte zu allen Messzeitpunkten signifikante und positive

Autoregressionen, d.h. AE bzw. KE zu einem bestimmten Zeitpunkt ist starker Prädiktor für AE bzw. KE zu einem späteren Zeitpunkt (AE: .345 - .416; KE: .415 - .566). Zudem gab es für beide Konstrukte signifikante (verzögerte) Cross-Lagged-Effekte, d.h. KE zu einem Zeitpunkt beeinflusst AE zu einem späteren Zeitpunkt und umgekehrt (AE -> KE: .106 - .200; KE -> AE: .155 - .254). Es zeigte sich, dass Unterricht mit offenen Aufgaben (EG) ein signifikanter und positiver Prädiktor für AE4 (.130) und KE1 (.206) ist. Während für andere Messzeitpunkte keine Effekte der Intervention auf AE und KE festgestellt werden konnte.

Im Einklang mit der Selbstbestimmungstheorie deuten die Ergebnisse auf eine wechselseitige Beziehung zwischen AE und KE während des Lernens hin. Dies unterstreicht die Bedeutung eines Unterrichts, der sowohl AE als auch KE fördert, sodass das Erleben der Grundbedürfnisse gestärkt wird. Die hohe Stabilität über die Zeit macht deutlich, dass die Stärkung des Erlebens der Grundbedürfnisse potenziell zu nachhaltigen Effekten führen kann. Darüber hinaus konnte gezeigt werden, dass Unterricht mit offenen Aufgaben (EG) einen positiven Einfluss auf AE und KE haben kann. Dieser Einfluss zeigt sich jedoch nur zu einzelnen Messzeitpunkten und scheint somit auch von der konkreten Unterrichtsgestaltung abhängig zu sein.

Strategiebezogene Motivation und Textverstehen: Mediert die Qualität der selbstgenerierten Fragen die Effekte der Selbstwirksamkeit, Wertüberzeugungen und Kosten auf Textverstehen in Mathematik?

Stanislaw Schukajlow, Pia Köhler, Alexander Westhölter

Universität Münster, Deutschland

Theorie

In Erwartungs-Wert Theorien wird angenommen, dass Selbstwirksamkeitserwartungen, Wertüberzeugungen und Kosten eine wichtige Rolle für Lernen, Leistung und Textverstehen spielen (Eccles & Wigfield, 2020; Wigfield et al., 2016). Eine bedeutsame Leistung in Mathematik ist das Verstehen der Aufgabestellungen, welches das Bearbeiten von mathematischen Modellierungsaufgaben (anspruchsvolle realitätsbezogene Aufgaben) beeinflusst (Leiss et al., 2019). Strategiebezogene Motivation (z.B. Selbstwirksamkeitserwartungen bzgl. einer Strategie) kann die Qualität der Strategienutzung und darüber das Textverstehen erhöhen. Positive Effekte der strategiebezogenen Motivation auf Strategiequalität und Leistungen wurden für das Zeichnen einer Skizze festgestellt (Schukajlow et al., 2021). In dieser Untersuchung werden diese Effekte für die Strategie „Fragengenerieren“ untersucht, deren positive Wirkungen auf das Textverstehen nachgewiesen wurden (Song, 2016).

Hypothesen

H1: Selbstwirksamkeitserwartungen, Wertüberzeugungen und Kosten sagen die Qualität der selbstgenerierten Fragen und das Textverstehen beim Bearbeiten von Modellierungsaufgaben vorher.

H2: Die Qualität der selbstgenerierten Fragen mediert die Effekte von Selbstwirksamkeitserwartungen, Wertüberzeugungen und Kosten auf das Textverstehen.

Methode

121 Neuntklässler:innen (81 männlich, $M=14.3$ Jahre) aus zwei Gesamtschulen haben Fragen zu Modellierungsaufgaben generiert, ihr Verstehen der Aufgabenstellungen eingeschätzt und ihre Motivation zur Strategie Fragengenerieren angegeben.

Für die Erfassung der strategiebezogenen Motivation beim Fragengenerieren wurden 5-stufige Skalen zur Motivation bei selbsterstellten Visualisierungen (Schukajlow et al., 2021) adaptiert (1=stimmt gar nicht; 5=stimmt genau). Selbstwirksamkeitserwartungen, Wertüberzeugungen und Kosten wurden jeweils mit 3 Items (z.B. „Ich bin zuversichtlich, dass es mir immer gelingt, zu schwierigen Textaufgaben hilfreiche Fragen zu stellen.“), 7 Items (2 attainment, 2 utility und 3 intrinsic Wertüberzeugungen-Items; z.B. „Das Stellen von Fragen finde ich wichtig, weil es mir hilft, schwierige Textaufgaben zu lösen.“) und 3 Items (z.B. „Es kostet mich viel Mühe, zu einer schwierigen Textaufgabe hilfreiche Fragen zu stellen.“) erfasst. Die Qualität der Strategienutzung wurde mit Hilfe eines 6-stufigen Rankings der selbstgenerierten Fragen zu den 7 Modellierungsaufgaben erhoben. Das Generieren von Fragen wurde höher gewertet, wenn diese auch für die Bearbeitung der Aufgabe wichtig waren (7 Items; 0=nicht bearbeitet/nicht nachvollziehbar; 5=die gestellte Frage ist eine Verständnisfrage bezogen auf Informationen, die für das Bearbeiten der Aufgabe relevant sind). Das Textverständnis wurde mithilfe der Selbstauskünfte zu den 7 Modellierungsaufgaben/Situationen erfasst (5-stufige Skala; 14 Items; z.B. „Ich habe die beschriebene Situation „Feuerwehr“ gut verstanden.“). Interne Konsistenzen für Selbstwirksamkeitserwartungen und Kosten waren niedrig: jeweils .49 und .53 und für anderen Skalen befriedigend (.71) bis sehr gut (.90). Die Daten wurden mittels eines saturierten Pfadmodells in MPlus 8.9 ausgewertet. Fehlende Werte wurden mit dem FIML-Algorithmus geschätzt.

Ergebnisse und Diskussion

Wie erwartet, hat Selbstwirksamkeitserwartungen die Qualität der selbstgenerierten Fragen und das Textverstehen vorhergesagt ($\beta=.247$ und $.252$; $p<.05$). Wertüberzeugungen und Kosten haben die Qualität der Strategie und Textverstehen nicht vorhergesagt ($p>.10$). Wie angenommen, hat die Qualität der selbstgenerierten Fragen den Effekt der Selbstwirksamkeitserwartungen auf das Textverstehen vermittelt ($\beta=.059$; $p<.05$). Es wurden keine indirekten Effekte von Wertüberzeugungen und Kosten auf Textverstehen über die Qualität der Strategienutzung festgestellt.

Positive Effekte von Selbstwirksamkeitserwartungen auf die Strategienutzung und das Textverstehen bestätigen die Annahmen der Erwartungs-Wert-Theorie (Eccles & Wigfield, 2020) insbesondere für das Textverstehen (Wigfield et al., 2016). Die Neuigkeit besteht darin, dass die Bedeutung der Erwartungen auch für den Bereich der Strategienutzung bestätigt wurde, welcher bisher selten untersucht wurde (Schukajlow et al., 2021). Die Effekte der Wertüberzeugungen und Kosten auf die Qualität der Strategienutzung und Textverstehen konnten nicht bestätigt werden. Eine Erklärung ist, dass Lernende in dieser Untersuchung aufgefordert wurden, die Strategie anzuwenden. Da Wertüberzeugungen und Kosten für eigene Entscheidungen prädiktiv sind (Eccles & Wigfield, 2020), haben sie in diesem Setting womöglich eine untergeordnete Rolle gespielt. Eine theoretische Implikation dieser Studie ist die Notwendigkeit die prädiktive Kraft der motivationalen Konstrukte im Hinblick auf die Outcomes (Leistungen vs. Entscheidungen) zu differenzieren.

PS-39: Digitale Transformation und Lehrerbildung

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SO 418

Chair der Sitzung: Christian Seyferth-Zapf

TPACK in Aktion: Kontextabhängige Auswirkungen auf die Wissensintegration von Lehrkräften und Lehramtsstudierenden

Franziska Tschönhens¹, Iris Backfisch¹, Tim Fütterer², Andreas Lachner¹

¹Tübingen Center for Digital Education, Deutschland; ²Hector-Institut für Empirische Bildungsforschung und Pädagogische Psychologie, Universität Tübingen

Theoretischer Hintergrund. Professionelles Wissen von Lehrkräften ist für die effektive Integration von Technologien in den Unterricht unerlässlich. Um dieses Wissen zu beschreiben, kann das *Technological-Pedagogical-Content-Knowledge* (TPACK)-Modell von Mishra und Koehler (2006) herangezogen werden. Das TPACK-Modell beschreibt drei miteinander verflochtene Wissenskomponenten, die für das Unterrichten mit Technologien wesentlich sind: technologisches (TK), pädagogisches (PK) und inhaltliches (CK) Wissen und wurde später um den Kontext erweitert, um den situativen Charakter des professionellen Wissens darzustellen (Mishra, 2019). Es wird angenommen, dass kontextuelle Faktoren, wie verschiedene Unterrichtskontexte (Porras-Hernández & Salinas-Amescua, 2013), die Anwendung und Integration des technologisch-pädagogischen Inhaltswissens (TPCK) beeinflussen (see Gollwitzer, 2012). Dabei scheint die Lehrerfahrung eine weitere wichtige Rolle zu spielen, da erst durch die tatsächliche Unterrichtspraxis und absichtliches Üben anspruchsvolle integrierte Wissensstrukturen entwickelt werden (Ifenthaler & Seel, 2011; Berliner, 2004). Wie Kontextvariablen und Lehrerfahrung die Wissensintegration beeinflussen, bleibt jedoch eine offene Frage.

Fragestellungen. In der vorliegenden Studie wird untersucht, wie Lehrkräfte die TPACK-Komponenten integrieren und wie die Wissensintegration von verschiedenen Unterrichtskontexten abhängt. Zudem wird der Einfluss der Lehrerfahrung mittels eines relativen Expertise-Ansatzes untersucht. Die Fragestellungen sind:

(F1) Welche Wissenskomponenten (TK, PK, CK) wenden Lehrkräfte bei der Reflexion über Unterrichtsstrategien an? Gibt es Unterschiede zwischen (a) verschiedenen Unterrichtskontexten und (b) der Lehrerfahrung (Lehrkräfte im Dienst versus Lehramtsstudierende)?

(F2) Wie treten die angewandten Wissenskomponenten gemeinsam auf? Gibt es Unterschiede (a) je nach Unterrichtskontext und (b) je nach Lehrerfahrungserfahrung (Lehrkräfte im Dienst versus Lehramtsstudierende)?

Methode. Die Studie verfolgt einen quasi-experimentellen-Ansatz mit Unterrichtskontext (Innersubjektfaktor: drei Unterrichtskontexte) und Lehrerfahrung (Zwischensubjektfaktor: Lehrkräfte im Dienst oder Lehramtsstudierende) als unabhängige Variablen. Als abhängige Variable wurden die Anwendung der Wissenskomponenten (TK, PK, CK) erfasst. Die vorliegenden sekundär analysierten Daten umfassten $N = 441$ Lehrkräfte ($n = 243$ Lehrkräfte im Dienst, $n = 198$ Lehramtsstudierende). Alle Lehrkräfte wurden gebeten, über die mögliche Integration von Technologien in drei Unterrichtskontexten (individualisiertes Lernen, kollaboratives Lernen, und Übungsaufgaben) zu reflektieren. Die Antworten wurden in Ideeneinheiten aufgeteilt und in Anwendung von TK, PK und CK kodiert. Zunächst wurde die Häufigkeit der angewandten Wissenskomponente analysiert. Logistische Regressionen wurden durchgeführt, um die Auswirkungen der Unterrichtskontexte und der Lehrerfahrung auf die Anwendung der Wissenskomponenten zu ermitteln. Paarweise Vergleiche dienten zur Erklärung der Haupteffekte (F1). Zur Beantwortung von F2 wurde eine Epistemische Netzwerkanalyse (ENA; Shaffer, 2017) durchgeführt, um Kookkurrenzen der Wissenskomponenten zu analysieren und zu visualisieren.

Ergebnisse und ihre Bedeutung. Insgesamt gab es signifikante Unterschiede bei der Anwendung der Wissenskomponenten zwischen den Unterrichtskontexten, $\chi^2(2) = 1392.190$, $p < .001$ (F1a). Logistische Regressionsmodelle zeigten keine Auswirkungen des Unterrichtskontexts oder der Lehrerfahrung auf die Anwendung von TK, jedoch signifikante Haupteffekte von Unterrichtskontext auf die Anwendung von PK ($B = -0.82$, $SE = 0.25$, $p < .001$). Für CK waren beide Haupteffekte (Unterrichtskontext, $B = 1.02$, $SE = 0.40$, $p = .011$ und Lehrerfahrung, $B = 1.63$, $SE = 0.78$, $p = .037$) signifikant (F1b). Die ENA zeigte unterschiedliche Muster von Ko-Vorkommen unter den TPACK-Komponenten. Der Unterrichtskontext Übungsaufgaben zeigte signifikante Unterschiede zum Unterrichtskontext individualisiertes Lernen, $t(866.97) = -3.06$, $p < .001$ und kollaboratives Lernen, $t(635.76) = -8.10$, $p < .001$. TK und PK traten häufiger gemeinsam im Unterrichtskontext individualisiertes Lernen und kooperatives Lernen auf, während CK häufiger bei Übungsaufgaben angewandt wurde (F2a). Bei Lehrkräften im Dienst war CK in Verbindung mit TK und PK häufiger, während TK und PK bei Lehramtsstudierenden öfter gemeinsam vorkamen. Statistisch signifikante Unterschiede wurden bei individualisierten Lernen, $t(437.22) = 2.80$, $p = .01$ und kollaborativen Lernen, $t(376.12) = 2.24$, $p = .03$ gefunden (F2b). Die Ergebnisse deuten darauf hin, dass TPACK dynamisch und kontextabhängig ist und nicht als statische Integration seiner Komponenten verstanden werden sollte. Weitere Studien, beispielsweise durch den Think-Aloud-Ansatz, sind notwendig, um tiefere Einblicke in die Wissensintegration von Lehrkräften zu erhalten.

Medienpädagogische Qualifizierung von Lehramtsstudierenden durch Selbsteinschätzung ihrer digitalen Kompetenzen? Überlegungen zur Weiterentwicklung des SELFIEforTEACHERS

Cindy Bärnreuther¹, Melanie Stephan¹, Christian Seyferth-Zapf², Matthias Ehmann²

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland; ²Universität Bayreuth, Deutschland

Die Bedeutung digitaler Medien für Schülerinnen und Schüler sowie Lehrkräfte ist in den letzten Jahren stetig gewachsen. Im internationalen Vergleich rangieren deutsche Lehrkräfte unterdurchschnittlich, sowohl in der Nutzung digitaler Medien für Lehren und Lernen als auch in der Förderung von ICT-Kompetenzen (Drossel et al., 2019). Die Herausforderungen der COVID-19-Pandemie haben zu einem verstärkten Einsatz digitaler Medien im Schulkontext geführt. Doch allein die Nutzung digitaler Medien garantiert noch keine erfolgreiche Lehre, da dies stark von der digitalen Kompetenz der Lehrpersonen abhängt (Endberg & Lorenz, 2022; Lorenz et al., 2022).

Theoretischer Hintergrund

Auf theoretischer Ebene gibt es seit langem einen etablierten nationalen und internationalen Diskurs über medienbezogene digitale Kompetenzen von Lehrkräften, der zu zahlreichen Modellen und Konzeptualisierungen geführt hat. Zu den prominentesten Modellen gehören TPACK (Mishra & Koehler, 2006), DigCompEdu (Redecker, 2017), das Modell der Medienpädagogischen Kompetenz (Herzig & Martin, 2018) und das K19-Framework (DCB, 2017). Diese Modelle bieten unterschiedliche Perspektiven auf die Anforderungen und Fähigkeiten, die Lehrkräfte im digitalen Zeitalter benötigen.

Einige dieser Modelle wurden in den letzten Jahren operationalisiert. Diese Instrumente helfen dabei, digitale Kompetenzen von Lehrkräften zu bewerten und zu fördern. Jedoch fehlt es vielen dieser Instrumente an einer expliziten Ausrichtung auf angehende Lehrkräfte, was eine Lücke in der Lehrkräftebildung darstellt.

Der DigCompEdu der Europäischen Kommission (Redecker, 2017) ist eines der am häufigsten rezipierten internationalen Modelle in Bezug auf die Erfassung von digitalen Kompetenzen von Lehrkräften. Dieses Modell besteht aus 22 Kompetenzen, die in sechs Bereiche gegliedert sind und berufliche sowie pädagogische und didaktische Kompetenzen der Lehrenden sowie zu fördernde Kompetenzen auf Seiten der Lernenden berücksichtigen. Das darauf aufbauende SelfieforTEACHERS (Economou, 2023) gibt personalisiertes Feedback über das individuell erreichte Kompetenzniveau (A1 bis C2).

Fragestellung

Das SELFIEforTEACHERS-Selbsteinschätzungstool zielt vor allem auf etablierte Lehrkräfte in der Praxis ab. Allerdings beschreibt Economou (2023) darüber hinaus die Verwendung des Tools im Kontext der ersten Phase der Lehrkräftebildung. Diese Studie untersucht die Anwendbarkeit und Validität des SELFIEforTEACHERS-Instruments für Lehramtsstudierende.

Methode

Die Studie wurde mit Studierenden der Universität Bayreuth und der Friedrich-Alexander-Universität Erlangen-Nürnberg durchgeführt. An der Erhebung durch Onlinefragebogen nahmen 127 Studierende teil. Zusätzlich wurden 9 Studierende mittels Methode des Lauten Denkens befragt, um tiefere Einblicke in ihre Einschätzungen und die Verständlichkeit der Items zu erhalten. Der Fragebogen umfasste die verschiedenen Kompetenzbereiche des DigCompEdu-Modells, und die Teilnehmenden wurden gebeten, ihre eigenen Fähigkeiten und Erfahrungen in diesen Bereichen zu bewerten.

Ergebnisse und ihre Bedeutung

Die Forschungsergebnisse verdeutlichten zentrale Herausforderungen. Insbesondere erweisen sich die Items für die Lehramtsstudierenden als unverständlich, zu lang und gegenüber ihrem Erfahrungshorizont unangemessen. Dies beeinträchtigt die praktische Anwendbarkeit des Instruments. Darüber hinaus mangelt es an theoretischer und statistischer Validität, was die Zuverlässigkeit und Aussagekraft der Ergebnisse in Frage stellt.

Die Untersuchung zeigt, dass das SELFIEforTEACHERS-Instrument in seiner aktuellen Form nur begrenzt für die Verwendung mit Lehramtsstudierenden geeignet ist. Es bestehen jedoch verschiedene Adaptionmöglichkeiten, um Anwendbarkeit und Validität des Instruments zu verbessern. Dazu gehören Anpassungen der Itemformulierungen an die Bedürfnisse und den Erfahrungshorizont von Lehramtsstudierenden, Enthierarchisierung durch die Verwendung einer Likert-Skala und Diskussion über das zugrunde liegende Messmodell.

Der vorgesehene Tagungsbeitrag lädt dazu ein, Überarbeitungsansätze aus einer praxisbezogenen, theoretischen und empirischen Perspektive zu diskutieren. Präsentation und Diskussion der Ergebnisse sowie der vorgeschlagenen Überarbeitungen auf der Tagung bieten Gelegenheit zur Weiterentwicklung des SELFIEforTEACHERS-Instruments. Langfristig könnte dies zur Unterstützung der medienpädagogischen Qualifizierung pädagogischer Fachkräfte beitragen und somit die digitale Kompetenz von Lehrkräften und die Qualität des digitalen Unterrichts verbessern.

Durch Diskussion und Weiterentwicklung solcher Instrumente kann die Lehrkräftebildung verbessert werden, was letztlich zur Förderung der digitalen Kompetenzen von Schülerinnen und Schülern beiträgt. Dies ist ein entscheidender Schritt, um die Potenziale der digitalen Bildung voll auszuschöpfen und gesellschaftliche Herausforderungen durch Bildung und Bildungsforschung anzugehen.

Facetten professioneller Lehrkompetenz für Bildung in einer digitalen Welt: Erkenntnisse einer Mixed-Methods-Studie zum Erklärvideo-Einsatz im Naturwissenschaftsunterricht

Michelle Hermann^{1,2}, Markus Wilhelm¹, Markus Rehm², Dorothee Brovelli¹

¹Pädagogische Hochschule Luzern, Schweiz; ²Pädagogische Hochschule Heidelberg, Deutschland

In einer zunehmend digitalisierten Welt gewinnt die Beurteilung und Nutzung digitaler Medien aus verschiedensten Quellen in vielen gesellschaftlichen Bereichen laufend an Bedeutung (z.B. van Ackeren et al. 2019). Dies zeigt sich unter anderem beim Einsatz digitaler Bildungsmedien, die Lehrpersonen vor eine doppelte Herausforderung stellen: Es wird einerseits erwartet, dass sie unter Einbezug digitaler Medien Unterricht gestalten, der bei Schülerinnen und Schülern zu einem Zuwachs an fachbezogenen Kompetenzen beiträgt und darüber hinaus den Erwerb überfachlicher und medienbezogener Kompetenzen begünstigt. Andererseits sind die Lehrpersonen in der Unterrichtsplanung selbst in diesen Kompetenzbereichen gefordert, wenn sie – gestützt auf ihr Professionswissen und ihre Erfahrung – aus der Fülle digitaler Lernangebote jene Angebote auswählen sollen, die den angestrebten fachlichen und überfachlichen Kompetenzzuwachs potenziell am besten unterstützen (Scheiter 2021). Exemplarisch zeigt sich letzteres im Zusammenhang mit Erklärvideos von partizipativen Plattformen wie YouTube. Diese unterliegen – anders als etablierte Bildungsmedien – keinen Qualitätskontrollen und sind daher insgesamt von heterogener Qualität (Kulgemeyer 2018). Entsprechend sind Lehrpersonen in der Unterrichtsvorbereitung als Gatekeeper mit der Qualitätsbeurteilung von Erklärvideos bei der Auswahl und der adäquaten unterrichtlichen Einbettung gefordert.

Der Beitrag präsentiert vor diesem Hintergrund eine vom Schweizerischen Nationalfonds unterstützte Mixed-Methods-Studie, die situationsspezifische fachdidaktische Facetten professioneller Lehrkompetenz für den lernwirksamen Einsatz von Erklärvideos im integrierten Fach Natur und Technik bei (angehenden) Lehrpersonen der Sekundarstufe I untersucht. Erhoben wurde die professionelle Wahrnehmung (van Es und Sherin 2008) von Lernunterstützungsmerkmalen in Erklärvideos zu naturwissenschaftlichen Fachkonzepten, wobei die Merkmale der Lernunterstützung als Maß für die Qualität von Erklärvideos gesehen werden. Zur Operationalisierung der Qualität der als Stimuli eingesetzten Erklärvideos wurde aus der Verknüpfung von etablierten Qualitätsmerkmalen von naturwissenschaftlichem Unterricht (Heinitz und Nehring 2020) mit Qualitätsmerkmalen für Erklärvideos (u.a. Kulgemeyer 2018) ein fachdidaktisch geprägtes Kriterienraster ausgearbeitet, das auf Erklärvideos aus den Fächern Chemie, Physik und Biologie anwendbar ist (Hermann et al. eingereicht). Im Rahmen eines Expertenratings charakterisierten 17 Fachpersonen die als Stimuli eingesetzten Erklärvideos entlang des Kriterienrasters. Die dabei generierte Expertennorm bildete die Basis für die inhaltsanalytische Auswertung der von den Studienteilnehmenden (N = 330) vorgenommenen Einschätzungen der eingesetzten Erklärvideos. Ergänzend wurden für die Fähigkeit zur professionellen Wahrnehmung potenziell erklärende Faktoren erhoben (unter anderem Facetten naturwissenschaftsdidaktischen Wissens, der Ausbildungsstand, Praxiserfahrungen und Einstellungen gegenüber digitalen Medien).

Die Ergebnisse der Studie zeigen, dass die Teilnehmenden die Qualität der Erklärvideos teilweise überschätzten, was gut zu Resultaten aus situierter Erhebung der professionellen Wahrnehmung mittels Videovignetten passt, wonach angehende

Lehrpersonen im Gegensatz zu erfahrenen Lehrpersonen lernunterstützende Maßnahmen der kognitiven Aktivierung oder inhaltlichen Strukturierung in präsentierten Unterrichtsmitschnitten häufig übersehen (Meschede et al. 2017). In vielen Fällen verzichteten die Teilnehmenden auf eine von fachdidaktischen Kriterien geleitete Einordnung ihrer Wahrnehmung von Merkmalen der Lernunterstützung und schlugen konsequenterweise mehrheitlich unterrichtliche Einbettungen vor, die den im Rahmen des Expertenratings identifizierten Defiziten in den Erklärvideos kaum Rechnung trugen. Diese Ergebnisse legen nahe, dass (angehende) Lehrpersonen das in Erklärvideos als digitale Medien vorhandene lernunterstützende Potenzial (Hillmayr et al. 2020; Fey 2021) häufig nicht ausschöpfen, da sie ihr Professionswissen nicht für eine situativ gelungene Einbettung von Erklärvideos – beispielsweise in Form einer Kombination der Erklärvideos mit passgenauen Aufgabenstellungen und einordnenden Klassengesprächen – nutzbar machen können. Vor diesem Hintergrund beleuchtet der Beitrag die Beziehungen zwischen den erhobenen Kompetenzfacetten und weiteren erklärenden Faktoren. Ausgehend davon werden Ansätze für die Aus- und Weiterbildung von Naturwissenschaftslehrpersonen zur Sensibilisierung für Qualitätsaspekte beim Einsatz von Erklärvideos abgeleitet. Ziel ist dabei eine vollständigere Ausschöpfung des Potenzials des Mediums durch kriteriengeleitete Beurteilung und der daraus abgeleiteten adäquaten unterrichtlichen Einbettung durch die Lehrpersonen.

Lerngelegenheiten und Entwicklung technologiebezogenen fachdidaktischen Wissens (TPACK) von Lehramtsanwärter:innen – eine empirische Studie auf Basis von Selbsteinschätzungen

Fabian Grünig¹, Tobias Dörfler², Markus Vogel²

¹Pädagogische Hochschule St.Gallen, Schweiz; ²Pädagogische Hochschule Heidelberg, Deutschland

Die Studie untersuchte den selbsteingeschätzten TPACK-Wissenszuwachs von 91 Lehramtsanwärter:innen während des Vorbereitungsdiensts 2017-2018 in Baden-Württemberg und identifizierte förderliche Einflussfaktoren. Es zeigte sich ein kleiner bis mittlerer Effekt für den Wissenszuwachs, wobei die Anzahl der Seminarveranstaltungen und die eigene Unterrichtserfahrung als signifikante Einflussfaktoren identifiziert wurden.

Theoretischer Hintergrund

Die Integration von digitalen Medien in den Fachunterricht stellt eine zentrale Herausforderung für angehende Lehrpersonen dar, wobei technologiebezogenes fachdidaktisches Wissen (TPACK; Mishra & Koehler, 2006) als relevanter Gelingensfaktor gilt. Der Vorbereitungsdienst soll auf solche Herausforderungen der Unterrichtspraxis vorbereiten (Eikermann et al., 2016) und wird an Ausbildungsschulen durch Mentor:innen sowie an Seminaren begleitet (KM Ba-Wü, 2014).

Als relevante Entwicklungsfaktoren für TPACK identifizieren Zhang und Tang (2021) die kollaborative Zusammenarbeit bei der Entwicklung und Reflexion von technologieunterstütztem Unterricht sowie das technologiebezogene Grundlagenwissen (TK). Lerngelegenheiten im Vorbereitungsdienst umfassen die Thematisierung von fachspezifischen Medien im Seminar, die kollaborative Reflexion von Unterricht, die eigene Unterrichtserfahrung und den Austausch mit Mentor:innen (Lohse-Bossenz & Vogel, 2021).

Für die vorliegende Studie wurden die folgenden Forschungsfragen formuliert: (1) Führt der Vorbereitungsdienst zu einem TPACK-Wissenszuwachs? (2) Welche Einflussfaktoren erklären interindividuelle Unterschiede beim Wissenserwerb?

Methoden

Für die längsschnittliche Analyse konnten Daten von $n=91$ Mathematiklehrpersonen von zwei Messzeitpunkten ausgewertet werden. Sie stammen aus dem Vorbereitungsdienst 2017–2018 für das WHRS-Lehramt im Land Baden-Württemberg.

Zur Erhebung von TPACK wurden sieben übersetzte Items aus den Selbsteinschätzungsskalen von Handal et al. (2012) verwendet; TK wurde am ersten Messzeitpunkt mit sieben übersetzten Items aus Schmidt et al. (2009) erhoben.

Mit Blick auf die Einflussfaktoren wurde zu fachbezogenen Medien jeweils erhoben, (1) in wie vielen Seminarsitzungen diese behandelt wurden, (2) wie viele Minuten diese in einer typischen Unterrichtseinheit eingesetzt wurden (Ostermann et al., 2021) und (3) wie häufig intensiver Austausch mit dem:der Mentor:in diesbezüglich stattgefunden hat. Über die Technologien hinweg wurde je Einflussfaktor ein Score aggregiert. Zudem wurde dichotom erhoben, ob am Seminar in kollaborativer Fallarbeit Unterricht der Anwärter:innen reflektiert wurde.

Es wurden lineare Regressionsanalysen mit dem TPACK-Score zum zweiten Messzeitpunkt als abhängiger Variable durchgeführt. Zur Modellierung der Veränderung wurde der TPACK-Score zum ersten Messzeitpunkt als Prädiktor modelliert. Ausgehend von diesem Basismodell wurden die Einflussvariablen als unabhängige Variablen in das Modell aufgenommen und eine schrittweise Rückwärtsselektion durchgeführt.

Ergebnisse

Vom ersten ($M = 3.70$, $SD = 0.94$, $\alpha = 0.89$) zum zweiten Messzeitpunkt ($M = 4.13$, $SD = 0.93$, $\alpha = 0.88$) ergab sich ein kleiner bis mittlerer Wissenszuwachs, $t(80) = -4.06$, $p < 0.001$, Cohens $d = 0.45$.

Digitale Medien wurden im Median zwei Mal in den Seminarsitzungen behandelt ($MD = 2.00$, $M = 2.49$, $SD = 2.86$), zwei Mal mit den Mentor:innen diskutiert ($MD = 2.00$, $M = 3.27$, $SD = 4.13$) und in einer typischen Unterrichtseinheit von 180 Minuten im Mittel $M = 37.45$ Minuten ($SD = 39.24$ Minuten) eingesetzt. Es gaben 78.78 % der Anwärter:innen an, dass ihr Unterricht in den Seminaren für Fallarbeit aufgegriffen wurde.

Das Basismodell der TPACK-Veränderung klärte $R^2 = 24.93$ %, alle Einflussvariablen zusammen weitere $\Delta R^2 = 18.80$ % der Varianz auf. Nach Rückwärtsselektion verringerte sich die Varianzaufklärung um $\Delta R^2 = 1.80$ % und es blieben die Seminarsitzungen und der Einsatz digitaler Medien als signifikante Einflussfaktoren übrig, $R^2 = 41.88$ %.

Diskussion

Die identifizierten Einflussfaktoren sind anschlussfähig an etablierte Annahmen aus TPACK-Entwicklungsmodellen und Modellen der professionellen Handlungskompetenz. Entgegen der Erwartung aus bisherigen empirischen Befunden (Zhang und Tang, 2021) zeigten sich das technologiebezogene Grundlagenwissen, der Austausch mit Mentor:innen und die Fallarbeit nicht prädiktiv. Die Fallarbeit wurde ohne Technologiebezug erhoben, was den ausbleibenden Effekt erklären könnte. Ein Effekt durch die Mentor:innen könnte ausbleiben, da Gespräche hier situativ stattfinden und sich so nicht auf die übergreifende Wissensselbstzuschreibung auswirken. Zentrale Limitationen der Studie werden im Vortrag diskutiert.

PS-40: Gender and Equality in the Classroom

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 154

Chair der Sitzung: Till Stefes

English Track

Fehlerklima und Geschlecht als Einflussfaktoren auf Fehlerreaktionen bei Grundschulkindern

Carolin Burmeister, Kim Beck, Robert Grassinger

Pädagogische Hochschule Weingarten, Deutschland

Theoretischer Hintergrund

Fehler sind ein integraler Bestandteil von Lernprozessen in der Schule und können wertvolle Einblicke in das aktuelle Verständnis von Lehrplaninhalten liefern (Hascher & Hagenauer, 2010). Nach dem von Tulis et al. (2016) vorgeschlagenen Prozessmodell eines individuellen Lernens aus Fehlern gelingt ein Lernen aus Fehlern, wenn Schüler*innen affektiv-motivational adaptiv und handlungsadaptiv darauf reagieren (Dresel et al., 2013; Steuer et al., 2013). Empirische Befunde aus Stichproben mit Sekundarstufenschüler*innen deuten darauf hin, dass diese adaptiven Fehlerreaktionen u.a. durch das Fehlerklima im Klassenzimmer – konzeptualisiert als generalisierter Faktor acht verschiedener Dimensionen – beeinflusst werden (Grassinger et al., 2018; Soncini et al., 2022; Steuer, 2014; Steuer et al., 2013). Darüber hinaus sprechen erste empirische Befunde von Steuer et al. (2021) dafür, dass die verschiedenen Fehlerklimadimensionen ('Analyse von Fehlern', 'Fehlerrisiko eingehen', 'Fehlertoleranz der Lehrkraft', 'Lernfunktionalität von Fehlern', 'Bewertungsirrelevanz von Fehlern', 'Abwesenheit negativer Mitschüler*innenreaktionen', 'Abwesenheit negativer Lehrkraftreaktionen', 'Lehrkraftunterstützung nach Fehlern') differentiell mit affektiv-motivational adaptiven oder handlungsadaptiven Fehlerreaktionen assoziiert sind. Des Weiteren gibt es erste Hinweise auf geschlechtsspezifische Unterschiede in den Fehlerreaktionen der Schüler*innen (Dresel et al., 2013; Grassinger et al., 2015) und dass diese unterschiedlich mit den Fehlerklimadimensionen assoziiert sind (Steuer, 2014).

Weniger untersucht ist, (1) inwiefern bisherige Befunde aus Sekundarstufenchüchproben zur Relevanz des Fehlerklimas und des Geschlechts für ein Lernen aus Fehlern auf Grundschüler*innen übertragen werden können. Aufgrund entwicklungsbedingter Veränderungen von beispielsweise selbstregulativen Kompetenzen (Fox & Riconscente, 2008; Holodynski et al., 2013) sowie motivationalen Bedingungen (Filipp, 2006; Finsterwald et al., 2009; Schwinger & Wild, 2006) ist dies nicht per se gegeben. (2) Zudem wurden die differentiellen Zusammenhangsmuster des Fehlerklimas mit affektiv-motivational adaptiven und handlungsadaptiven Fehlerreaktionen bis dato nicht repliziert. (3) Schließlich liegen keine Kenntnisse darüber vor, inwiefern Geschlechtsunterschiede in der subjektiven Wahrnehmung des Fehlerklimas bestehen und diese Geschlechtsunterschiede in den adaptiven Fehlerreaktionen (anteilig) medieren.

Hypothesen

Ziel der Studie ist es, die genannten Forschungsdesiderate zu adressieren:

H1a) Schüler*innen der Klassenstufen 3 und 4 reagieren adaptiver auf Fehler, wenn sie ein positiveres Fehlerklima wahrnehmen.

H1b) Die Fehlerklimadimensionen 'Fehlerrisiko eingehen', 'Fehlertoleranz der Lehrkraft', 'Bewertungsirrelevanz von Fehlern', 'Abwesenheit negativer Mitschüler*innenreaktionen' und 'Abwesenheit negativer Lehrkraftreaktionen' sind stärker mit affektiv-motivational adaptiven Fehlerreaktionen assoziiert und die Fehlerklimadimensionen 'Analyse von Fehlern', 'Lernfunktionalität von Fehlern' und 'Lehrkraftunterstützung nach Fehlern' sind stärker mit handlungsadaptiven Fehlerreaktionen assoziiert.

H2a) Es gibt Geschlechterunterschiede in den adaptiven Fehlerreaktionen bei Grundschulkindern.

H2b) Es gibt bei Grundschulkindern Geschlechtsunterschiede in der Wahrnehmung des Fehlerklimas im Klassenzimmer.

H2c) Die Wahrnehmung des Fehlerklimas durch die Grundschüler*innen vermittelt Geschlechtsunterschiede in den adaptiven Fehlerreaktionen der Lernenden.

Methode

Es handelt sich um eine querschnittliche Fragebogenerhebung mit $N=675$ Grundschulkindern aus dritten und vierten Klassen in den Schulfächern Deutsch und Mathematik. Das Durchschnittsalter der Gesamtstichprobe betrug $M=9.15$ Jahre ($SD=.85$). Erfasst wurden das subjektiv wahrgenommene Fehlerklima im Klassenzimmer anhand von 31 Items ($\alpha=.89$; Steuer, 2014) sowie affektiv-motivational adaptive und handlungsadaptive Fehlerreaktionen anhand von insgesamt 13 Items ($\alpha=.61-.75$; Dresel et al., 2013).

Ergebnisse und Diskussion

Die Analysen auf Individual- und Klassenebene replizierten Ergebnisse zur Relevanz des Fehlerklimas für ein Lernen aus Fehlern auch für Primarstufenschüler*innen (H1a). Hierbei replizierte sich der Befund, dass einzelne Fehlerklimadimensionen stärker mit affektiv-motivational adaptiven Fehlerreaktionen und andere Fehlerklimadimensionen stärker mit handlungsadaptiven Fehlerreaktionen assoziiert sind (H1b). Zudem zeigten sich geschlechtsspezifische Unterschiede im Lernen aus Fehlern (H2a), die mit geschlechtsspezifischen Unterschieden in der Wahrnehmung des Fehlerklimas (H2b) erklärt werden konnten (H2c). Limitierend ist zu beachten, dass die Studie als Querschnittsstudie durchgeführt wurde und folglich keine Kausalität in den Beziehungen interpretiert werden kann. Nichtsdestotrotz liefern die Ergebnisse wertvolle Erkenntnisse darüber, wie das Lernen aus Fehlern von Dritt- und Viertklässler*innen verstanden werden kann und unterstützen generell die Idee, dass frühere Erkenntnisse mit älteren Schüler*innen auf die in dieser Studie untersuchte Altersgruppe übertragen werden können (Soncini et al., 2022).

Exploring the Impact of Teacher-Student Similarity on Academic Outcomes: Evidence from Germany

Richard Nennstiel

Universität Bern, Schweiz

This paper addresses the important but understudied issue of teacher-student demographic congruence and its impact on educational outcomes in the German education system. The following mechanisms have been discussed in the literature to explain possible effects of teacher-student similarity on academic achievement. First, teachers act as role models, so that students need teachers of the same sex or ethnic minority to imitate and orient themselves to their behavior. Second, stereotype threat may be activated by teachers of the opposite sex or ethnic minority, which could hinder academic achievement. Third, teachers may make gender- or ethnic-minority-specific evaluations of students according to their own ascriptive characteristics and thus influence students' self-concept, for example, through grading - net of academic ability. Using data from several cohorts

(SC2 [primary school students], SC3 [lower secondary school students], and SC4 [upper secondary school students]) the National Educational Panel Study (NEPS), we examine whether primary, lower secondary, and upper secondary students who are taught by teachers of the same gender or migration background show differences in academic achievement and grading fairness. The central aim of the paper is to extend existing research, which is often cross-sectional, with a longitudinal perspective to assess the influence of teacher-student similarity on changes in educational outcomes. Our analysis uses two primary indicators: (1) achievement growth in mathematics and reading, and (2) grades in mathematics and German, adjusted for test score differences. This study covers three different cohorts of students at different stages of education in Germany, providing a comprehensive view of the educational landscape. Preliminary results suggest a correlation between demographic similarity between teachers and students and better educational outcomes. Specifically, we observe evidence that having a teacher of the same gender or immigrant background as the student may (slightly) improve equity in grading and contribute to higher learning gains. These findings could have profound implications for educational policy and practice, highlighting the potential benefits of fostering teacher-student demographic congruence as a means of promoting equity and inclusivity in educational settings. The study's findings contribute to the broader discourse on social justice in education, suggesting that teacher-student similarity might be an influential factor in enhancing academic success and ensuring fair treatment across diverse student populations.

Echoes of Aggression. The gendered Impact of School Bullying on the Well-Being of Non-Targeted Students

Till Stefes

Ruhr Universität Bochum, Deutschland

Peer relationships are central in adolescent life, profoundly influencing social integration, development, and subjective well-being (SWB). Engaging with peers is vital for identity formation (Eder 1985; Larson & Richards 1991; Swanson et al. 1998; Ragelienė 2016) and overall well-being (Brown & Larson 2009; Guhn et al. 2013; Cuadros & Berger 2016; Appau et al. 2019). Since much adolescent social life occurs in school, its climate significantly impacts SWB (Katsantonis et al. 2024; Stefes 2024).

Bullying, a detrimental aspect of peer relations, can severely undermine SWB. Unlike typical peer conflicts, bullying is persistent and involves power imbalance (Espelage et al. 2013), leading to long-term physical and psychological issues (Vanderbilt & Augustyn 2010; Moore et al. 2017; Bradshaw et al. 2017), such as poor social adjustment, psychosomatic symptoms (Rigby 2003; Wolke & Lereya 2015), and increased anxiety, depression (Romano et al. 2020), and suicidal ideation (Kelly et al. 2015; Galindo-Domínguez & Losado Iglesias 2023). Observing bullying often leads to internalizing symptoms like anxiety, depression, and suicidal ideation (Rivers et al. 2009; Rivers & Noret 2010; Rivers & Noret 2013; Werth et al. 2015; Wu et al. 2016; Lambe et al. 2017). While the direct impact of bullying on victims' well-being is well-documented (Cook et al. 2010; Moore et al. 2017; Katsantonis et al. 2024), the significant effects on bystanders are recognized in psychology (Midgett & Dumas 2019) but lack sociological analysis.

Adolescents' SWB varies by gender, with girls often reporting lower levels than boys (González-Carrasco et al. 2017; Knüttel et al. 2021; Stefes 2024b; Katsantonis et al. 2024). Bullying experiences differ due to gender roles (Hellström & Beckmann 2020): boys report less victimization (Rivers & Smith 1994; Cowie 2000) and face more physical violence (Scheithauer et al. 2006; Iossi Silva et al. 2013; Cosma et al. 2022), while girls are more often targeted indirectly (Rivers & Smith 1994) or online (Smith et al. 2019), and more affected by sexual cyberbullying (Divers & Duncan 2013; Copp et al. 2021). Thus, research on adolescent bullying must consider gender differences.

This article fills a sociological research gap by investigating the question, "Does witnessing bullying in the classroom affect adolescents' subjective well-being?"

Multilevel regression analyses are employed using a classroom-based, standardized survey conducted in 2021 among students aged 12-16 (Stefes 2023; Stefes et al. 2023). This study addresses limitations of prior research, which often used small samples, by analysing data from 121 classrooms with over 1600 students. Classrooms with at least one bullying report identify students as victims or witnesses without explicit questioning about being a witness, analysing four types: physical, verbal, social, and cyberbullying. Unlike many bullying-studies that use single-item analyses susceptible to random errors (Sandvik et al. 2009; Katsantonis et al. 2024), this study employs a multi-dimensional index to measure SWB across six dimensions: self-esteem, optimism, absence of sadness, absence of worries, body image, and life satisfaction. Each dimension is assessed through multiple items, following established practices (Pollock et al. 2018; Moore 2020).

Significant differences in SWB are found among targeted students, non-targeted witnesses, and non-targeted/non-witnessing students. Students in classrooms with at least one bullying report show significant SWB gaps compared to others, affecting all types of bullying. Girls are more strongly affected, experiencing significant SWB loss regardless of bullying type, whereas boys show differences only with physical violence or cyberbullying reports. Classrooms, rather than individuals, offer potential for effective interventions. Identifying affected classrooms is easier than individual victims, and classroom-level interventions can aid both targets and witnesses. These insights can help schools and education professionals intervene early, potentially breaking the cycle of violence (Walters 2021; Falla et al. 2022; Espino et al. 2023).

Gendergerecht oder Gendergaga? Eine Metaanalyse über den Einfluss von geschlechtergerechter Sprache im Vergleich zum generischen Maskulinum auf die subjektive Wahrnehmung von Texten, das objektive Textverständnis und die Lesedauer

Katharina M. Bach¹, Veronika Wachholz¹, Amadeus J. Pickal², Constanze Richters³, Julia Murböck¹, Laura Brandl¹, Matthias Stadler³, Michael Sailer²

¹Department Psychologie, Ludwig-Maximilians-Universität München; ²Learning Analytics and Educational Data Mining, Universität Augsburg; ³Institut für Didaktik und Ausbildungsforschung in der Medizin (DAM), LMU Klinikum, LMU München

Geschlechtergerechte Sprache wird medial, politisch und gesellschaftlich kontrovers diskutiert (Diewald, 2018). Befürworter*innen möchten alle Geschlechter sprachlich repräsentieren und dem "Male Bias" durch das generische Maskulinum entgegenwirken (Scezesny et al., 2016). Kritiker*innen befürchten negative Auswirkungen auf das Lesen (Braun et al., 2007; Payr, 2021) und einige Medien sogar einen Sprachverfall (Eisenberg, 2021). Bisher fehlen eindeutige Forschungsergebnisse zum Einfluss von geschlechtergerechter Sprache auf Textverständnis, Lesedauer und die subjektive Wahrnehmung von Texten (Friedrich et al., 2021). Pöschko und Prieler (2018) fanden in ihrer Studie mit Schüler*innen keinen signifikanten Unterschied zwischen geschlechtergerechten Formulierungen und dem generischen Maskulinum bezüglich der Erinnerungsleistung,

Wahrnehmung von Verständlichkeit und Güte der Formulierungen. Die Lesbarkeit wurde jedoch bei der geschlechtergerechten Textvariante schlechter bewertet.

Die vorliegenden Metaanalyse aggregierte Daten von 24 bisher unpublizierten Replikationsstudien basierend auf Pöschko und Prieler (2018), welche den Einfluss von verschiedenen geschlechtergerechten Formulierungen im Vergleich zum generischen Maskulinum auf die subjektive Textwahrnehmung (Güte der Formulierungen, Lesbarkeit, Verständlichkeit, kognitive Belastung) sowie die objektiven Kriterien Textverständnis und Lesedauer erfassten. Zusätzlich untersuchte sie, ob die Art der geschlechtergerechten Formulierung die Effekte moderiert. Im Gegensatz zur Originalstudie war die verwendete Stichprobe größer und altersheterogener.

Methode

Die 922 Teilnehmer*innen (nweiblich= 590, ndivers = 14; Alter: 18-70, M = 28,7, SD = 5.66) der k = 24 Replikationsstudien wurden randomisiert auf die Kontroll- (generisches Maskulinum) oder Experimentalgruppe (Gendersternchen: k = 13, Doppelpunkt: k = 4, Binnenl: k = 5 und Beidnennung: k = 2) aufgeteilt. Die Teilnehmer*innen lasen einen Text, der jeweils 20 geschlechtergerechte Formulierungen oder Formulierungen im generischen Maskulinum enthielt. Subjektive Wahrnehmung wurde mittels Fragebögen (Friedrich, 2017), Erinnerungsleistung mit einem Wissenstest und die Lesedauer computergesteuert erfasst. Für alle Replikationsstudien wurden Mittelwert und Standardabweichung der abhängigen Variablen gebildet. Der Effekt wurde mittels eines Random Effects Models mit der Art der geschlechtergerechten Formulierung als Moderator berechnet.

Ergebnisse

Für Lesedauer, objektives Textverständnis und alle Variablen der subjektiven Wahrnehmung, mit Ausnahme der wahrgenommenen Verständlichkeit, zeigten sich keine signifikanten Unterschiede zwischen den geschlechtergerechten Formulierungen und denen im generischen Maskulinum. Bei der wahrgenommene Verständlichkeit zeigte sich ein kleiner negativer Effekt der geschlechtergerechten Varianten von $g = -0.281$ (SE = 0.080, $p < .001$, 95% CI [- 0.439, - 0.124]). Die Gruppe mit den geschlechtergerechten Formulierungen bewertete die Verständlichkeit entsprechend vergleichsweise schlechter als die Gruppe mit dem generischen Maskulinum. Der Moderator hatte für keine der abhängigen Variablen einen signifikanten Einfluss.

Diskussion

Die vorliegenden Ergebnisse sind vergleichbar mit denen der Originalstudie und anderer Studien, die kaum messbaren Beeinträchtigungen durch geschlechtergerechte Sprache zeigen konnten (Blake & Klimmt, 2010, Friedrich & Heise, 2019). Vor allem die objektiv gemessenen Variablen unterschieden sich sowohl in der Originalstudie als auch in der vorliegenden Metaanalyse nicht signifikant. Auch die subjektive Wahrnehmung blieb durch die geschlechtergerechten Formulierungen weitestgehend unbeeinträchtigt. Die gesamtgesellschaftliche Skepsis gegenüber geschlechtergerechter Formulierungen und Angst vor möglichen Beeinträchtigungen des Textverständnisses (Infratest dimap, 2022; Weidner et al., 2021), zeigte sich möglicherweise in der schlechteren subjektiven Bewertung der Verständlichkeit. Diese kann jedoch auch an den ungewohnten und teilweise komplizierten Formulierungen liegen (Diewald & Nübling, 2022). Menschen gewöhnen sich allerdings schnell an sprachliche Änderungen; eine regelmäßige Konfrontation kann den Gewöhnungsprozess positiv beeinflussen (Gustafsson Sendén et al., 2015).

Während es in unserer Studie keine Unterschiede für die verschiedenen Arten der geschlechtergerechten Formulierungen gab, wurde bei Pöschko und Prieler (2018) die Schrägstrich-Schreibweise schlechter bewertet als die Neutralisierung.

Da geschlechtergerechte Sprache zu einer höheren Repräsentation von Frauen und nicht-binären Menschen führt (Flaherty, 2001; Formanowicz et al., 2012) ohne das objektive Textverständnis und die subjektive Wahrnehmung des Textes (weitestgehend) zu beeinträchtigen, gibt es aus Forschungsperspektive derzeit kaum Gründe gegen deren Nutzung. Die Nutzung könnte zudem im Sinne des Effekts der Darbietungshäufigkeit die Bewertung der Verständlichkeit positiv beeinflussen (Zajonc, 1968).

PS-41: Familie, Bildung und Lebensverläufe

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SO 318

Chair der Sitzung: Maximilian Seitz

ENTFÄLLT The untold story of internal migration in Germany: Life-cycle patterns, developments, and the role of education

Sebastian Vogler¹, Kamila Cygan-Rehm², Guido Heineck³, Anton Barabasch⁴

¹Leibniz-Institut für Bildungsverläufe (LIfBi), Deutschland; ²Dresden University of Technology (TUD); ³Otto-Friedrich-University Bamberg; ⁴Friedrich-Alexander University Erlangen-Nuremberg (FAU)

Regional mobility is an essential driver of economic growth and technological development, as both depend on the ability and willingness of workers to relocate to more innovative and productive sectors and labor markets (e.g., Blanchard & Katz 1992, Caselli & Coleman 2001, Amior & Manning 2018). Thus, an important consequence of geographic mobility is a more efficient match between workers and firms (Dauth et al. 2022), but it can also drive spatial income inequalities (e.g., Gaubert et al. 2021). At the individual level, regional mobility is often seen as a means to improve the economic situation and well-being of households and individuals (e.g., Briggs & Kuhn 2008, Deryugina et al. 2018, Groen et al. 2020). On the other hand, it can also lead to unintended consequences arising, for example, from the disruptive nature of relocation (Nassal & Paul 2022) or residential segregation (Derenoncourt 2022).

This paper presents a comprehensive and detailed analysis of regional mobility patterns in Germany from a lifetime perspective. Apart from the well-documented East-West and gender gaps, our focus is on the role of education, as it has been long recognized as the key factor for understanding why some individuals move across regions and others do not (known as the “positive skill selection”). For this purpose, we use data from the National Education Panel Study: Starting Cohort Adults (NEPS-SC6). The unique feature of these data is the availability of both detailed biographical information on residential moves and educational paths over the life cycle. Specifically, for a representative sample of nearly 13,000 individuals born in Germany between 1944 and 1986, we construct a balanced panel that tracks their geographic mobility at monthly intervals starting from birth until 2020 (i.e., until the age of 34 to 76, depending on the birth cohort). We then turn to the question of whether there is a causal component in the significant educational mobility gradient. We do so by exploiting two arguably exogenous sources of variation, each inducing a shift at a different margin of the educational distribution. First, we exploit a post-World War II compulsory schooling reform that aimed to increase the duration of schooling for students at the bottom of the ability distribution (e.g., Pischke & von Wachter 2008). Second, for the same school cohorts, we study the mobility responses to statutory cutoff rules for school enrollment, which have been shown to increase the probability of attending the highest ability track in secondary school (e.g., Dustmann et al. 2017).

Using a difference-in-differences and regression discontinuity design, we find no effect of these policies on internal mobility. Reassuringly, both panels yield very similar results. In particular, all point estimates are relatively small in magnitude and statistically insignificant. Thus, despite the substantial effect on schooling duration, the reform did not significantly affect individuals' mobility behavior. This holds for both cross-state and cross-county mobility. We conclude that none of the two policies had a meaningful impact on important channels through which education could impact regional mobility such as cognitive skills and academic credentials.

Interdisziplinärer Blick auf den Zusammenhang elterlicher und kindlicher Aspirationen – spielen innerfamiliäre Beziehungen eine Rolle?

Kerstin Schörner, Monja Schmitt, Felix Bittmann

Leibniz-Institute for Educational Trajectories

Theoretischer Hintergrund Bildungsaspirationen gelten als entscheidender Faktor für individuellen Bildungserfolg und die Entstehung von Bildungsungleichheiten (Bittmann, 2022; Ditton et al., 2005). Bisherige Forschung weiß über die bedeutende Rolle von elterlichen Aspirationen für die Ausbildung kindlicher Aspirationen (Schoon et al., 2007; Sewell & Shah, 1968; Zimmermann, 2018). Insbesondere im Rahmen der soziologisch dominierten Forschung um das Wisconsin Social Psychological Model of Status Attainment (WIM) (Sewell et al., 1969, 1970) wurde dies häufig aufgezeigt (Hauser, 1972, S. 172; Sewell & Shah, 1968). Fragestellung Wir stellen die bislang wenig untersuchte Frage, ob der Einfluss der Eltern und ihrer Erwartungen auf ihre Kinder so universell ist, wie ursprünglich vom WIM angenommen, oder ob es bestimmte Bedingungen gibt, unter denen die elterlichen Aspirationen einen besonderen Einfluss auf die Aspirationen ihrer Kinder haben. Aufbauend auf psychologisch dominierter Literatur, die sich mit der intergenerationalen Weitergabe von Werten befasst, nehmen wir an, dass innerfamiliäres Sozialkapital hier eine Rolle spielt - als Kanal für die Weitergabe von Aspirationen. Wir übertragen dabei Colemans Idee der Interaktion von Humankapital und Sozialkapital (Coleman, 1990) auf eine notwendige Interaktion von elterlichen Aspirationen und innerfamiliärem Sozialkapital, um einen Einfluss elterlicher Aspirationen auf die Aspirationen ihrer Kinder zu verzeichnen. Folglich gehen wir davon aus, dass nur das zeitgleiche Vorhandensein hoher elterlicher Aspirationen und guter innerfamiliärer Beziehungen zu hohen kindlichen Aspirationen führt. Aufbauend auf Arbeiten von Coleman unterscheiden wir verschiedene Aspekte von innerfamiliärem Sozialkapital, namentlich die allgemeine elterliche Unterstützung, das Familienklima, allgemeine schulische Unterstützung und soziale Unterstützung in der Schule. Methode Wir nutzen Daten des Nationalen Bildungspanels (NEPS Startkohorte 3), um den angenommenen Interaktionseffekt von elterlichen Aspirationen und innerfamiliärem Sozialkapital auf kindliche Aspirationen zu untersuchen. Nach multipler Imputation (25 Datensätze, unter Nutzung weiterer Prädiktoren) verbleiben 4210 Fälle im Analysesample. Mit Logit-Modellen erklären wir kindliche Schulabschlusssaspirationen in Klasse 8 (Abitur vs. anderer Abschluss) durch zunächst elterliche Aspirationen in Klasse 7 und bedeutsame Kontrollvariablen, die den sozialen Hintergrund und Charakteristika des Kindes aufgreifen. In einem weiteren Schritt nehmen wir die vier verschiedenen oben angesprochenen Aspekte innerfamiliärem Sozialkapitals in das Modell mit auf und formulieren den Interaktionsterm aus elterlichen Aspirationen und jeweiliger Facette innerfamiliärem Sozialkapital. Ergebnisse Unsere Analysen zeigen, dass hohe elterlichen Aspirationen nur dann mit hohen kindlichen Aspirationen in Zusammenhang stehen, wenn Eltern ihre Kinder in schulischen Belangen unterstützen. Wir diskutieren, warum andere Facetten einer guten Eltern-Kind-Beziehung für die Weitergabe von hohen Aspirationen scheinbar nicht entscheidend sind. Die Arbeit zeigt, wie wissenschaftliche Erkenntnisse und Theorien disziplinübergreifend genutzt werden können, und trägt dabei in zweierlei Hinsicht zu bisheriger Literatur bei. Der erste Beitrag ist theoretischer Natur, da die im Wisconsin Model angelegte Universalität des elterlichen Einflusses erstmalig hinterfragt wird und dadurch neue konzeptionell wichtige Fragen aufgeworfen und interdisziplinär erörtert werden. Der zweite Beitrag ist ein empirischer. Anhand eines qualitativ hochwertigen deutschen Datensatzes untersuchen wir die von Coleman angelegten verschiedenen Facetten des innerfamiliären Sozialkapitals und ihre Auswirkungen auf die Beziehung zwischen elterlichen und

kindlichen Aspirationen. Damit leisten wir einen empirischen Beitrag zur Klärung der Frage, unter welchen Bedingungen Eltern ihre eigenen Aspirationen auf ihre Kinder übertragen.

Unter Druck: Wie beeinflusst mütterlicher Stress frühe kognitive Fähigkeiten von Kindern?

Maximilian Seitz

Leibniz-Institut für Bildungsverläufe

Die direkte Umwelt hat einen bedeutsamen Einfluss darauf, wie Kinder sich in den ersten Lebensjahren entwickeln (proximale Faktoren; Bronfenbrenner & Morris, 2006), was wiederum eine wichtige Grundlage für den späteren Schulerfolg ist (z.B. Hofer & Clouston, 2014). Aus theoretischer Sicht kann argumentiert werden, dass elterlicher Stress aufgrund der dadurch beeinträchtigten Qualität und Quantität elterlicher Investitionen sich negativ auf verschiedene Domänen der Kindesentwicklung auswirkt (für eine Übersicht, vgl. Masarik & Conger, 2017). So wird z.B. im Bereich der kognitiven Fähigkeiten davon ausgegangen, dass Eltern durch erhöhte Stresslevel weniger emotionale Unterstützung und sprachlich-kognitive Anregung im Umgang mit ihren Kindern zeigen (Nievar et al., 2014). Gerade diese Dimensionen elterlichen Interaktionsverhaltens erweisen sich dabei aber als prädiktiv für spätere kognitive Fähigkeiten der Kinder (z.B. Farah et al., 2008). In den ersten Lebensjahren gibt es in diesem Bereich bisher nur wenig Forschung, die auf die direkte Erhebung von kognitiven Fähigkeiten durch Beobachtungsverfahren setzt. Daher untersucht der Vortrag die Zusammenhänge von mütterlichem Stress und frühen kognitiven Fähigkeiten in den ersten beiden Lebensjahren. Zudem soll explorativ die Rolle von mütterlicher emotionaler Unterstützung und sprachlich-kognitiver Anregung als mögliche Mediatoren untersucht werden. Es werden Daten der Neugeborenenkohorte des Nationalen Bildungspanels genutzt (NEPS SC1; Blossfeld & Roßbach, 2019). Konkret werden Erhebungswelle 1 (n = 2552; Alter = 6.96 Monate, SD = 0.72, Min = 5.77, Max = 11.03; 48.28% weiblich) und Erhebungswelle 2 (n = 1218; Alter = 17.05 Monate, SD = 0.61, Min = 15.77, Max = 20.36; 49.96% weiblich) untersucht. Frühe kognitive Fähigkeiten wurden mit einer Habitationsaufgabe in beiden Erhebungswellen erhoben, die basale visuelle Informationsverarbeitung erfasst (Colombo et al., 2020). Mütterlicher Stress war ein Selbstbericht zu Erhebungswelle 1 mit drei Items zur Depressivität und Belastung in den ersten Lebensmonaten des Kindes. Mithilfe von Strukturgleichungsmodellen wurden die Korrelationen zwischen frühen kognitiven Fähigkeiten und mütterlichem Stress berechnet (jeweils latente Faktoren). Die explorativen Analysen zum Interaktionsverhalten basieren auf Einzelindikatoren einer halbstandardisierten Spielsituation (vgl. Freund, 2019) in Erhebungswelle 1. Die Modelle zeigen, dass im Alter von 7 Monaten Säuglinge von Müttern mit einem hohen Stresslevel eine niedrigere Aufmerksamkeit während der Kognitionsaufgabe haben (kleine Effektstärke). Dieser Befund bleibt bestehen, wenn für den sozioökonomischen Hintergrund des Kindes kontrolliert wird. Explorative Berechnungen legen nahe, dass Aspekte der elterlichen Interaktion (d.h. emotionale Unterstützung und sprachlich-kognitive Anregung) diesen Effekt zumindest partiell vermitteln. Im Alter von 17 Monaten zeigen sich dagegen keine Effekte des mütterlichen Stresslevels – auch unter Kontrolle der kognitiven Fähigkeiten mit 7 Monaten. Dies weist darauf hin, dass es möglicherweise keine langfristigen bzw. kumulativen Effekte mütterlichen Stresses in den ersten Lebensmonaten des Kindes auf basale kognitive Fähigkeiten gibt. Die Ergebnisse werden vor dem Hintergrund sozioökonomischer Ungleichheiten und den Sample-Eigenschaften von NEPS SC1 diskutiert.

Zur Interdependenz von idealistischen und realistischen Berufsaspirationen in der Sekundarstufe

Victoria Zeddies¹, Annabell Daniel^{2,1}

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; ²Ludwig-Maximilians-Universität München

Berufsaspirationen gelten interdisziplinär als zentraler Prädiktor für den späteren Bildungs- und Berufserfolg (z.B. Ashby & Schoon, 2010; Rojewski, 2005). Theoretisch lässt sich zwischen idealistischen und realistischen Aspirationen unterscheiden (Haller, 1968; Stocké, 2012a,b). Während idealistische Aspirationen eher die beruflichen Wünsche und Interessen von Schüler:innen beschreiben – unabhängig bestehender Möglichkeiten und Grenzen – bilden die realistischen Aspirationen erwartbare bzw. erreichbare berufliche Ziele ab. Die beiden Konstrukte werden daher vielfach nicht nur theoretisch, sondern auch empirisch unabhängig voneinander betrachtet (Wicht & Ludwig-Mayerhofer, 2014).

Aus einer entwicklungstheoretischen Perspektive deutet sich jedoch speziell im Verlauf der Adoleszenz eine Abhängigkeit beider Konstrukte an: Folgt man Gottfredsons (1981) Berufswahltheorie, entwickeln sich Aspirationen von Berufswünschen hin zu realistischeren beruflichen Optionen. Gottfredson beschreibt dabei die Berufswahl als einen Kompromissbildungsprozess basierend auf hinzukommenden, altersabhängigen Erfahrungen und Erkenntnissen, wobei Wünsche zugunsten erreichbarer und realistischerer Optionen aufgegeben werden. Es stellt sich demnach die Frage, inwieweit es sich in der Entwicklung tatsächlich um getrennte Konstrukte handelt oder ob sie nicht vielmehr als gemeinsame Dimensionen der Berufswahl zu interpretieren sind (Wicht & Ludwig-Mayerhofer, 2014). Entsprechend Gottfredsons Ansatz müsste es zu einer Angleichung beider Konstrukte im Verlauf der Adoleszenz kommen. Eine übergreifende Perspektive blieb bisher jedoch weitgehend unberücksichtigt. So fehlt es bislang an Erkenntnissen zur längsschnittlichen Beziehung und Interdependenz beider Konstrukte. Es ist zudem anzunehmen, dass aufgrund der unterschiedlichen theoretisch intendierten Ausrichtung differenzielle Entwicklungsmechanismen und Wirkungsweisen greifen, die insbesondere mit Blick auf die Bedeutung für den späteren Bildungs- und Berufserfolg weiterer Untersuchung bedürfen.

Übereinstimmend mit Gottfredsons Berufswahltheorie verweisen empirische Befunde auf altersabhängige Veränderungen und individuelle Unterschiede in der Entwicklung idealistischer Berufsaspirationen im Jugendalter, welche u.a. auf individuelle Merkmale wie das Geschlecht oder die soziale Herkunft zurückgeführt werden (z.B. Gao & Eccles, 2020; Miyamoto & Wicht, 2020). Es bleibt jedoch unklar, inwieweit sich theoretisch erwartbare Zusammenhänge für die Entwicklung beider Aspirationen zeigen.

Ziel der vorliegenden Studie ist es daher, zum einen (nach bisherigem Kenntnisstand) erstmalig die längsschnittliche Interdependenz idealistischer und realistischer Aspirationen zu untersuchen, und zum anderen potenziell differenzielle Wirkzusammenhänge mit theorielevanten Entwicklungsmechanismen zu prüfen. Dabei wird angenommen, dass sich theoriekonform kurz vor dem Übergang von der Schule in den Beruf idealistische und realistische Aspirationen angleichen und sich insbesondere die realistischen Aspirationen zum Ende der Schulzeit verstärkt änderungssensitiv zeigen (Gottfredson, 1981; Heckhausen & Tomasik, 2002).

Als Datengrundlage dient die Startkohorte 4 des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019). Die Längsschnittstichprobe umfasst N=2.865 Sekundarschüler:innen von Klassenstufe 9 bis 12 und verfolgt so die Aspirationen über vier Messzeitpunkte bis kurz vor dem Abschluss am Gymnasium, mit welchem alle beruflichen Optionen angestrebt werden können. Idealistische und realistische Aspirationen werden jeweils über die kontinuierliche ISEI-Skala operationalisiert (Ganzeboom et al., 1992). Neben der Analyse von Mittelwertdifferenzen über die Zeit zwischen den idealistischen und

realistischen Aspirationen sollen darüber hinaus Cross-Lagged-Panel-Modelle spezifiziert werden. Diese ermöglichen, die Kompromissbildung in der Aspirationsentwicklung nachzuzeichnen, sowie die Interdependenz idealistischer und realistischer Aspirationen zu untersuchen.

Erste Ergebnisse zeigen einen reziproken, positiven Zusammenhang zwischen idealistischen und realistischen Aspirationen über die Zeit. Es wird deutlich, dass Schüler:innen zu allen Messzeitpunkten höhere idealistische als realistische Aspirationen aufweisen, sich jedoch beide Aspirationen mit Herannahen des Abschlusses signifikant negativ entwickeln. Entgegen der Erwartung zeichnet sich keine Annäherung der Aspirationen ab. Vielmehr deuten die Befunde auf eine parallele Entwicklung idealistischer und realistischer Aspirationen hin, wobei die Diskrepanz zum Ende der Oberstufe tendenziell größer zu werden scheint. Mit Blick auf die relevanten Prädiktoren der Entwicklung beruflicher Aspirationen deuten sich erste differenzielle Zusammenhangsmuster zwischen den Aspirationen an.

Es zeigen sich demnach Hinweise darauf, dass die gemeinsame Analyse idealistischer und realistischer Aspirationen wichtige Erkenntnisse zu zugrundeliegenden Entwicklungsmechanismen für den Übergang von der Schule in den Beruf liefern kann. Vertiefende Befunde und Implikationen werden im Vortrag diskutiert.

S-41: Geschlechtstypische Berufswahlprozesse: neue Befunde zu Mechanismen und Interventionsmöglichkeiten

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SO 108

Chair der Sitzung: Brigitte Schels

Chair der Sitzung: Corinna Kleinert

Chair der Sitzung: Kathrin Leuze

Geschlechtstypische Berufswahlprozesse: neue Befunde zu Mechanismen und Interventionsmöglichkeiten

Chair(s): **Brigitte Schels** (Paris-Lodron-Universität Salzburg), **Kleinert Corinna** (Leibniz-Institut für Bildungsverläufe)

Diskutant*in(nen): **Kathrin Leuze** (Friedrich-Schiller-Universität Jena)

Die geschlechtstypische Berufswahl junger Männer und Frauen ist eine der zentralen Dimensionen geschlechtsspezifischer Bildungsungleichheiten in unserer Gesellschaft (Solga/Becker 2012) und ein Schlüsselfaktor, der zur Persistenz des geschlechtersegregierten Ausbildungssystems und Arbeitsmarktes beiträgt (Bächmann et al. 2024; Buchmann/Kriesi 2012; Reisel et al. 2015). Daran hat sich auch angesichts des Wandels am Arbeitsmarkt, der zunehmenden Höherqualifizierung insbesondere junger Frauen und des Ausbaus von Berufsorientierungsprogrammen wenig verändert. So bleibt die Frage bedeutsam, unter welchen Voraussetzungen sich eine stereotype Berufswahl von jungen Männern und Frauen beobachten lässt und an welchen Stellschrauben eine geschlechtersensible berufliche Orientierung ansetzen kann (Makarova et al. 2017).

Während es viele Befunde zu den Determinanten geschlechts(un)typischer Aspirationen und Berufswahlentscheidungen gibt, sind die Mechanismen der Umsetzung von Aspirationen in Entscheidungen und damit einhergehenden Potentiale für Interventionen weiterhin unterbelichtet (Volodina & Nagy 2017). Vor diesem Hintergrund blicken wir in diesem Symposium auf junge Menschen in Deutschland und der Schweiz am Ende der Schulzeit. Beide Länder zeichnen sich durch ein stark beruflich strukturiertes Ausbildungssystem aus, in dem die frühe Berufswahl am Übergang in die Sekundarstufe II mit dazu beiträgt, dass Geschlechtergrenzen relativ starr sind (Buchmann/Kriesi 2012). In dieser kritischen Phase im Berufswahlprozess, die mit einer ersten Berufswahlentscheidung (zumindest vorläufig) endet, konkretisieren junge Menschen ihre Aspirationen (Kleinert/Schels 2020). Daher konzentrieren sich auch viele Orientierungs- und Beratungsprogramme auf diese Phase.

In Konzepten des Berufswahlprozesses wird angenommen, dass sich Jugendliche in dieser Phase aktiv mit ihren Interessen, Fähigkeiten, den sozialen Erwartungen in ihrem Umfeld sowie den wahrgenommenen Möglichkeitsräumen auseinandersetzen (Gottfredson 2002, Heckhausen 2000). Welche Berufe sie für sich als geeignet erachten, ist dabei zum einen durch kulturell verankerte Geschlechterstereotype bestimmt, die bereits ab dem Kindesalter wirksam sind und über die Schulzeit hinweg die Entwicklung von Interessen, Neigungen und Leistungen beeinflussen (Correll 2004; Hannover & Wolter 2017). Zum anderen interagieren Jugendliche im Berufswahlprozess mit Eltern, Lehrkräften, Berufsberater:innen und Ausbildungsentscheider:innen (Buchmann/Kriesi 2011). Ausgehend davon kann erwartet werden, dass in der Auseinandersetzung mit den Möglichkeitsräumen Rollen(vor)bilder aus der Alltagswelt der Jugendlichen und Berufsorientierungsprogramme wirksam sind und das Potential haben, Jugendlichen auch geschlechtsuntypische Berufsalternativen aufzuzeigen.

Das Symposium bringt vier empirische Beiträge zusammen, die sich aus einer soziologischen, psychologischen oder pädagogischen Perspektive mit Mechanismen vergeschlechtlichter Berufswahlentscheidungen sowie mit möglichen Stellschrauben für Veränderungen beschäftigen. Die ersten beiden Beiträge des Symposiums nehmen eine längsschnittliche Perspektive ein und gehen der Frage nach, wie Berufsaspirationen und berufliche Interessen geschlechts(un)typische Ausbildungsentscheidungen im Übergang von der Schule in Berufsausbildung strukturieren. Der erste Beitrag betrachtet Bildungspfade von Jugendlichen in männlich-dominierten MINT-Fächern und Berufen. Anhand dieses Beispiels wird die Bedeutung der Formation von Berufsaspirationen illustriert: Insbesondere unklare Aspirationen in der Schulzeit machen eine geschlechtstypische Berufswahl junger Männer wahrscheinlicher. Der zweite Beitrag verweist auf die Bedeutung beruflicher Interessensprofile, die sich in der Geschlechtstypik von idealistischen Berufswünschen und realistischen Erwartungen ebenso niederschlagen wie in den gewählten Ausbildungsberufen. Bei Mädchen und Jungen zeigen sich allerdings unterschiedliche Zusammenhänge und Dynamiken im Berufswahlprozess.

Die weiteren beiden Beiträge lenken das Augenmerk auf Ansatzpunkte einer geschlechtersensiblen Berufsorientierung. Der dritte Beitrag blickt auf Potentiale von Game-Based Learning für eine Auseinandersetzung mit Geschlechternormen. Die Studie zeigt, dass der Einsatz von Videospielen in der Berufsorientierung die Wahrscheinlichkeit fördert, dass junge Menschen Geschlechtergrenzen im Berufswahlprozess überschreiten. Der vierte Beitrag evaluiert in einer Interventionsstudie, ob und wie Ausbildungsbotschafter:innen als Rollenvorbilder wirksam sein können und die beruflichen Vorstellungen von Teilnehmer:innen an einem Workshopprogramm beeinflussen. Die Studie weist nach, dass Rollenvorbilder aus der Peergruppe im Berufsorientierungsprozess einen Beitrag leisten können, dass geschlechtsuntypische Berufe für junge Menschen an Attraktivität gewinnen.

Zum Schluss werden alle vier Beiträge zusammenfassend diskutiert, insbesondere mit Blick auf interdisziplinäre Anschlüsse, künftige Forschungsbedarfe und die zielgerichtete Implementierung von Interventionen.

Beiträge des Symposiums

Von Aspirationen zur Realisierung – Geschlechtsspezifische Entscheidungsprozesse bei MINT-Bildungskarrieren

Ralf Minor

Friedrich-Schiller-Universität Jena

Forschungsfrage

Die horizontale Segregation ist eine der Hauptursachen für berufliche Ungleichheit (Mann & DiPrete, 2013). Diese ungleiche Verteilung von Männern und Frauen auf die Berufe zeigt sich bereits im Jugendalter in Form von segregierten Berufswünschen. Obwohl solche Präferenzen während der Adoleszenz recht volatil sind, ist die Geschlechterverteilung bei MINT- und Nicht-MINT-Aspirationen bis zum Eintritt in die Hochschule ähnlich. Ein wesentlicher Einfluss späterer segregierter Bildungswege besteht darin, dass männliche Studierende ihre Präferenzen beim Studieneintritt überproportional in Richtung technologischer Studienfächer anpassen, während weibliche Studierende dazu neigen, Nicht-STEM-Wege zu verfolgen (Stefani et al., 2021). Um die Ursachen für diese geschlechtsspezifischen Verläufe zu entschlüsseln, frage ich:

Wie können geschlechtsspezifische Fachwechsel, Persistenz und späte Einstiege in MINT-Fächer beim Übergang deutscher Jugendlicher von der Schule in die Hochschule erklärt werden?

Theoretischer Hintergrund

Um der Unbeständigkeit von Lebensverläufen Rechnung zu tragen und methodische und kontextuelle Lücken zu schließen liefert meine Studie drei Beiträge zum Forschungsstand: Neben der Untersuchung am Beispiel von Deutschland und dem Einbezug der empirisch relevanten Gruppe von Schüler:innen mit unsicheren Aspirationen wird hier erstmals die Volatilität von Lebensverläufen während des Übergangs von der Schule zur Hochschule konzeptualisiert, indem Heckhausens (2000) „life-span theory of control“ angewendet wird. Damit können Persistenz, Veränderungen in den Aspirationen und späte Entscheidungen theoretisch durch soziodemographische, motivationale und umweltbedingte Mechanismen erklärt werden.

Methode

Um herauszufinden, welche Faktoren diese Werdegänge beeinflussen, verwende ich ein balanciertes Panel der Startkohorte Neuntklässler:innen (SC4) des NEPS. Meine Analysestichprobe enthält Informationen von 2.553 Schüler:innen zu ihren realistischen Berufswünschen in der zehnten Klasse und zum ersten Studienfach. Die Ergebnisse basieren auf der Passung zwischen diesen beiden Variablen, die mit multinomialen logistischen Regressionen geschätzt wurde. Neben den Hauptmodellen wurden Robustheitstests mit unterschiedlichen Messzeitpunkten, Aggregationsniveaus und Operationalisierungen durchgeführt.

Ergebnisse

Deskriptive Ergebnisse zeigen insbesondere einen geschlechtersegegrierten Zustrom junger Männer ohne klare vorherige Berufsaspirationen. Da dieses Muster in der Forschung bisher kaum berücksichtigt wurde, wurden im weiteren Verlauf der Studie die Mechanismen dieser geschlechtsspezifischen Pfade analysiert. Auf der Grundlage der theoretischen Überlegungen wurden dazu drei relevante Pfade für MINT-Karrieren unterschieden: „Persistenz“ (Übereinstimmung zwischen MINT-Aspirationen und Studienfachwahl), „Zieländerungen“ (Nicht-Übereinstimmung) und „späte Entscheidungen“ (MINT-Wahl ohne klare vorherige Aspiration). Der stärkste geschlechtsspezifische Effekt fand sich in der Gruppe der späten Entscheidungen. Frauen schlugen mit einer um 14,2 Prozentpunkte geringeren Wahrscheinlichkeit diesen Weg ein als Männer. Partiiell konnte dieser Geschlechterunterschied mit dem vorteilhaften allgemeinen und mathematikspezifischen relativen Leistungs- und Selbstkonzept von Männern erklärt werden. Die theoretisch plausible Erklärung, dass ein ambitioniertes Umfeld insbesondere späte Entscheidungen von Frauen für ein MINT-Studium negativ beeinträchtigt, ließ sich hingegen nicht bestätigen. Der Pfad Persistenz war für Frauen 4,8 Prozentpunkte weniger wahrscheinlich als für Männer. Auch hier fanden sich partielle Mediationseffekte des domänenspezifischen Selbstkonzepts, während sich die erwarteten Geschlechterunterschiede bei motivationalen Faktoren und sozialem Umfeld nicht bestätigten. Schließlich konnte der Geschlechterunterschied bei den Zieländerungen mit Ausnahme der Drop-Out-Kategorie teilweise durch niedrigere Werte im mathematischen Selbstkonzept der Frauen erklärt werden. Das Fehlen eines Geschlechtereffekts auf den Drop-Out-Pfad in MINT unterstreicht die geringe Relevanz dieses Pfads im deutschen Kontext bei der Studienfachwahl und legt nahe, dass das Modell der „leaky pipeline“ (Berryman, 1983) hier keine adäquate Beschreibung liefert. Vielmehr wird die Geschlechterungleichheit in MINT-Studienfächern wesentlich durch den überproportionalen Zustrom junger Männer verursacht. Der starke Einfluss domänenspezifischer Selbstkonzepte zeigt die Bedeutung subjektiv wahrgenommener Stärken für den beruflichen Werdegang und die an dieser Stelle relativ geringe Relevanz motivationaler Faktoren. Folglich sind theoretische Ansätze, die davon ausgehen, dass junge Menschen ihnen leicht zugängliche Wege wählen, um ihre Aspirationen zu verwirklichen (vgl. Gottfredson 1981, S. 549) angemessener. Da in Deutschland nach wie vor geschlechtsspezifische Unterschiede in Selbstkonzepten von Schüler:innen bestehen (Lewalter et al., 2023), erscheinen Maßnahmen, die an diesen Unterschieden in der Sekundarstufe ansetzen, von großem Nutzen.

Die Rolle von Interessenprofilen für geschlechtlich segregierte Berufswahlentscheidungen

Corinna Kleinert¹, Brigitte Schels²

¹Leibniz-Institut für Bildungsverläufe, ²Paris Lodron Universität Salzburg, Institut für Arbeitsmarkt- und Berufsforschung (IAB)

Fragestellung

Berufswahlprozesse sind bis heute durch eine ausgeprägte Geschlechtersegregation gekennzeichnet (Achatz, 2018). Studien zeigen, dass die beruflichen Interessen ein bedeutsamer Prädiktor der Berufswahl sind (Päßler & Hell, 2012; Volodina & Nagy, 2016) und Geschlechterunterschiede in den beruflichen Interessen zur Erklärung der Geschlechtersegregation auf dem Arbeitsmarkt beitragen (Kuhn & Wolter, 2022). Gleichzeitig sind die Interessen Jugendlicher vielschichtig, teilweise unidifferenziert und formieren sich zu Interessenprofilen, die in unterschiedlichem Ausmaß geschlechtstypisch geprägt sind (Slot et al., 2021). Vor diesem Hintergrund stellen sich die Fragen, inwieweit Interessenprofile die Geschlechtstypik der Berufswahl beeinflussen und ob es dabei systematische Unterschiede zwischen Mädchen und Jungen gibt. Uns interessiert insbesondere, wie Interessenprofile im Berufswahlprozess im Übergang Schule – Berufsausbildung wirken, wenn junge Menschen zunächst idealistische und realistische Aspirationen formieren, bevor sie in einen Ausbildungsberuf eintreten.

Theoretischer Hintergrund

Berufswahltheorien betonen, dass der Berufswahlprozess eng mit beruflichen Interessen zusammenhängt. So nimmt Gottfredson (2002) an, dass Jugendliche ihre Wunschberufe sukzessive entlang ihres Geschlechts, ihrer sozialen Herkunft sowie ihrer Interessen eingrenzen. Lent et al. (1994, 2000) postulieren, dass Jugendliche sich in Umgebungen selektieren, in denen sie ihre beruflichen Interessen verwirklichen können. Nach Holland (1997) lassen sich unterschiedliche Interessenorientierungen unterscheiden, die hexagonal entlang ihrer Ähnlichkeit zueinander angeordnet sind. Neuere Studien betonen dagegen, dass Jugendliche oftmals multiple Orientierungen gleichzeitig verfolgen, die sich in unterschiedlichen Konfigurationen manifestieren (McLarnon et al., 2015; Slot et al., 2021; Tay et al., 2011). Zudem finden sich starke Geschlechterdifferenzen in den beruflichen Interessen (Su et al., 2009) und entsprechend eine unterschiedliche Prävalenz fokussierter beruflicher Interessenprofile bei Mädchen und Jungen (Ehrmann et al., 2019; Slot et al., 2021). Ausgehend davon nehmen wir an, dass geschlechtstypische, fokussierte Interessenprofile einen größeren Einfluss auf die Geschlechtstypik der Berufswahl ausüben als unfokussierte breite. Der Einfluss der Interessenprofile auf die Geschlechtstypik der Berufswahl dürfte bei Mädchen und Jungen ähnlich stark ausgeprägt sein, die Profile selbst jedoch sollten unterschiedliche Zusammenhänge aufweisen. Für die Umsetzung von Interessen in einen Ausbildungsberuf spielen ferner Selektionsprozesse am Ausbildungsmarkt eine Rolle, da Arbeitgeber das Geschlecht der Bewerbenden als Passungsmerkmal heranziehen und geschlechtstypische Stellenbesetzungen bevorzugen (Imdorf 2012). So nehmen wir weiter an, dass junge Menschen mit geschlechtsuntypischen oder unfokussierten Interessensprofilen eine Anpassungsdynamik hin zu geschlechtstypischen Ausbildungsberufen aufweisen.

Methode

Unsere Studie nutzt Paneldaten der Neuntklässlerkohorte des NEPS (SC4). Sie enthalten Informationen zu beruflichen Interessen (IILS-I, 18-Item-Kurzform der RIASEC-Skala), idealistischen und realistischen Wunschberufen in der 9. Klasse sowie den später realisierten Ausbildungsberufen. Diese Berufangaben wurden mit Mikrozensus-Daten zu ihrer Geschlechterzusammensetzung angereichert. Unser Analysesample umfasst 9.905 Jugendliche, die frühestens nach der 10. Klasse die Schule verlassen haben (52.0% Mädchen). Interessenprofile wurden mittels latenter Profilanalysen generiert. Anschließend wurden deren Einflüsse auf die Geschlechtertypik der Berufswünsche und Ausbildungsberufe unter Kontrolle konfundierender Variablen mittels linearer Regressionen getrennt für Mädchen und Jungen geschätzt.

Ergebnisse

Die Analysen bestätigen für Jungen und Mädchen eine Konfiguration unterschiedlicher geschlechts(un)typisch fokussierter und unfokussierter breiter Interessensprofile. Demgegenüber sind die Berufswünsche und Ausbildungsberufe der Jugendlichen stärker geschlechtlich geprägt. Die idealistischen Wunschberufe von Mädchen sind insgesamt weniger geschlechtstypisch als die von Jungen. Allerdings steigt bei ihnen die Geschlechtstypik des Berufs in den realistischen Aspirationen bis zum Ausbildungsberuf immer weiter an, und zwar gleichermaßen über alle Interessenprofile hinweg. Bei Jungen finden sich dagegen differenzielle Veränderungen in der Geschlechtstypik des Berufs in unterschiedliche Richtungen je nach Interessenprofil. Vorläufige multivariate Ergebnisse zeigen, dass Interessenprofile nach Geschlecht tatsächlich unterschiedliche Einflüsse auf die Geschlechtstypik der Berufswahl ausüben. Dabei beeinflussen sie idealistische Berufswünsche am wenigsten, realistische Berufswünsche etwas stärker und die realisierten Ausbildungsberufe am deutlichsten. Diese Effekte sind bei Jungen ausgeprägter als bei Mädchen. Insgesamt legen diese Ergebnisse nahe, dass sich Mädchen im Laufe des Berufswahlprozesses stärker an vorhandene Berufsstrukturen und deren Segregationsmuster anpassen, während Jungen ihre beruflichen Interessen besser umsetzen können.

Spielend anders: Das Potenzial von Videospielen zur Förderung gendersensibler Berufsorientierung

Christine Hoffelner¹, Elena Makarova²

¹PH FHNW Murtens, ²Universität Basel

Theoretischer Hintergrund

Das Ende der obligatorischen Schulzeit markiert einen entscheidenden Moment in der Bildungsbiografie von Jugendlichen, da sie an diesem Punkt eine wichtige Weichenstellung vornehmen und sich für eine Ausbildungsoption entscheiden müssen. In der Schweiz ist der Unterricht der beruflichen Orientierung daher institutionell in der Sekundarstufe I verankert. Als wichtige Zielsetzung wird u.a. definiert, dass der Unterricht bei den Jugendlichen eine Auseinandersetzung mit sich und geschlechtsspezifischen Normen fördern soll (D-EDK, 2016). Trotz unterschiedlicher Angebote, die eine solche Reflexion anregen können, entscheiden sich die jungen Menschen weiterhin stark für gender-typische Berufs- und Ausbildungswege (Hupka-Brunner, 2022). Bei der Berufs- und Studienwahl besteht somit eine horizontale Geschlechtersegregation, in der die berufliche Geschlechtstypik die Wahl der Berufsausbildung als auch die spätere Berufslaufbahn erheblich beeinflusst (Teuscher et al., 2019). Eine gendersensible berufliche Orientierung ist demnach notwendig und fordert eine kritische Reflexion pädagogischen Handelns hinsichtlich der Frage, ob Lerngelegenheiten bestehende Geschlechterverhältnisse manifestieren oder Veränderungen zur Förderung von Geschlechtergerechtigkeit bewirken können (Makarova et al., 2017).

Generell eignet sich Game-Based Learning (Gee, 2003), um eine Erhöhung der intrinsischen Motivation und des Engagements in herausfordernden Lern- und Reflexionsprozessen, wie es die Auseinandersetzung zu Gendernormen in der Berufswahl sein kann, zu erzielen. Der aktuelle Forschungsstand liefert einige Hinweise zur Wirksamkeit von Serious Games in der beruflichen Orientierung (Keller et al., 2023; Spangenberg et al., 2022). Dagegen fehlt die empirische Evidenz zum Einsatz von kommerziellen Videospielen in der beruflichen Orientierung. Die vorliegende Studie hat zum Ziel das Forschungsdesiderat zu adressieren, indem Spielverhalten und Erfahrungen Jugendlicher in kommerziellen Videospielen erhoben und analysiert werden. Die Studie orientiert sich am Life Design Ansatz (Savickas, 2012), wobei mittels Rückbezüge zur Gamingwelt Reflexionsprozesse von Jugendlichen gefördert werden sollen (Rochat & Armengol, 2020; Rochat & Borgen, 2023).

Fragestellung

Welche Erkenntnisse lassen sich aus den Spielpräferenzen, -interessen und -strategien Jugendlicher für eine gendersensible berufliche Orientierung auf der Sekundarstufe I ableiten?

Methode

Die Studie wird im Rahmen des Forschungsprojekts "digibe" – Digitale Begleitung im Berufswahlprozess (Nägele & Hoffelner, 2023) durchgeführt. In einem eigens entwickelten Online-Tool werden dazu Jugendliche der Sekundarstufe I in der Deutschschweiz durch digital-stattfindende Reflexion im Berufswahlprozess unterstützt (n=275). Das Modul «Terra Ludus» umfasst dabei Aufgaben, welches ein Reflektieren zu Videospielen und Berufs- und Ausbildungswahl ermöglicht. Mittels eines Tests zum «challenge mindset» (McGonigal, 2016) wird mit 20 Items abgefragt, wie spielerisch Jugendliche ihren Berufswahlprozess bereits bewältigen, z.B. «Ich habe den Mut, mich allen Herausforderungen im Berufswahlprozess zu stellen.» In einer weiteren Aufgabe ordnen Jugendliche ihre Lieblingsspiele einer RIASEC-Kategorie (Holland, 1959) zu, die sich auf realistische, investigative, künstlerische, soziale, unternehmerische und konventionelle Präferenzen beziehen. Weiters geben sie an, welche der vier Spieltypen, Killer, Socializer, Achiever und Explorer (Bartle, 2003), ihnen entspricht.

Die Datenerhebung erfolgte durch standardisierte Fragebögen und Auswertung der Reflexionsaufgaben des Online-Tools. Die Daten wurden mittels t-Test und ANOVA-Analysen mit dem Programm JASP (2023) analysiert.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse zeigen, dass männliche und non-binäre Jugendliche Videospiele bevorzugen und Hindernisse im Berufswahlprozess spielerischer und selbstwirksamer überwinden. Mit zunehmendem Alter bewältigen Jugendliche den Berufswahlprozess weniger spielerisch. Jugendliche mit einer Vorliebe für kompetitive Spieltypen («Killer») vertrauen weniger auf ihre Fähigkeiten im Berufswahlprozess und erhalten weniger soziale Unterstützung. Weibliche Jugendliche spiegeln ihre Spielpräferenzen nach RIASEC im Berufswahlprozess wider. Jugendliche mit sozialen und künstlerischen Spielpräferenzen sehen diese Interessen weniger im Zusammenhang mit ihrer Berufswahl.

Zusammengefasst kann eine Auseinandersetzung mit dem eigenen Spielverhalten neue Erkenntnisse für die Berufs- und Ausbildungswahl liefern und wichtige Rückschlüsse für den Berufswahlprozess per se gezogen werden. Ein spielerischer Ansatz im Berufswahlprozess bedeutet auch, festgefahrene Bahnen wie Gendernormen leichter verlassen zu können. Videospiele bieten demnach ein hohes Potenzial, um als Türöffner für wichtige Reflexionsprozesse zu fungieren und einen gendersensiblen Unterricht der beruflichen Orientierung zu ermöglichen.

Berufliche Rollenmodelle und die Veränderung Beruflicher Aspirationen – Die Bedeutsamkeit von Geschlecht und sozialem Status

Janina Beckmann¹, Mona Granato²

¹Bundesinstitut für Berufsbildung (BIBB), ²Bundesinstitut für Berufsbildung

Theoretischer Hintergrund

Die Konkretisierung beruflicher Aspirationen ist eine zentrale Entwicklungsaufgabe im Jugendalter. Ein Großteil junger Menschen kann zwar Berufswünsche benennen, diese konzentrieren sich jedoch häufig auf wenige, aus ihrer Sicht zumeist statusreiche Berufe und Berufe, die dem eigenen Geschlecht entsprechen (Gottfredson 2002; Mann et al. 2020). Folgen eines eingeschränkten Aspirationsfeld sind Friktionen am Übergang Schule-Beruf, wenn z.B. der Übergang in berufliche Ausbildung nicht gelingt oder angesichts eingeschränkter Aspirationen berufliche Kompromisse eingegangen werden (müssen), die letztlich zu negativen Konsequenzen im Ausbildungsverlauf führen können (Holtmann und Solga 2023; Kleinert & Schels, 2020).

Neben soziostrukturellen Merkmalen wie Familie und Schule, die die Entwicklung von Berufsaspirationen langfristig mitstrukturieren, betonen neuere theoretische Ansätze darüberhinausgehend die Bedeutung kurzfristiger Zufallsereignisse für Berufswahlentscheidungen (z.B. Krumboltz, 2009). Ein Beispiel solcher kurzfristigen Begegnungen sind berufliche Rollenmodelle, die in ihrer informierenden und motivationalen Funktion zur Veränderung beruflicher Aspirationen beitragen können (Morgenroth 2009). Während die Wirksamkeit beruflicher Rollenmodelle in (vorrangig akademischen) MINT-Berufen recht gut belegt ist, fehlen (vergleichende) Ergebnisse zu Ausbildungsberufen und inwieweit sich Rollenmodelleffekte verschiedener Ausbildungsberufe unterscheiden und dadurch soziale Ungleichheiten - wie z.B. Geschlechtersegregation - abgeschwächt oder möglicherweise verstärkt werden könnten. Dieser Beitrag fragt daher, inwiefern sich Jugendliche in ihrer Berufswahl an beruflichen Rollenmodellen orientieren und welche Rolle dabei Geschlechtstypik und sozialer Status der vorgestellten Ausbildungsberufe spielen.

Fragestellung

Dieser Beitrag untersucht, inwieweit Schülerinnen und Schüler Ausbildungsberufe, die durch berufliche Rollenmodelle vorgestellt werden, in ihr Aspirationsfeld mit aufnehmen. Der Fokus liegt (1) auf der Wirksamkeit von Rollenmodellen unterschiedlicher Berufe und (2) wie Geschlecht und Prestige des Berufs diesen Zusammenhang moderieren. In Anlehnung an die Berufswahltheorie nach Gottfredson (2002) wird angenommen, dass für junge Menschen insbesondere Geschlechtstypik und berufliches Prestige für die Wahl eines Berufs bedeutsam sind, sodass Rollenmodelleffekte für geschlechtstypische und statusreiche Berufe besonders ausgeprägt sein sollten.

Methodik

Empirische Basis der Analysen ist eine regionale Studie in NRW, in welcher in einem quasi-experimentellen Forschungsdesign ca. 2.000 Schüler*innen der Jahrgangsstufen 9-11 an allgemeinbildenden Schulen befragt wurden. Eine Treatmentgruppe, die an Vorträgen von beruflichen Rollenmodellen im Klassenraum teilgenommen hat, und eine Kontrollgruppe wurden im Abstand von ca. 4-8 Wochen zu ihren beruflichen Aspirationen befragt.

Nach der Gruppierung der Ausbildungsberufe nach Geschlecht und Prestige werden anhand von Prä-Post-Vergleichen Unterschiede in der Entwicklung beruflicher Aspirationen in Abhängigkeit des Treatmentstatus untersucht. Dabei werden lineare Regressionsmodelle geschätzt, um (1) den Haupteffekt der Teilnahme an Vorträgen beruflicher Rollenmodelle auf berufliche Aspirationen zu analysieren und (2) Effektheterogenitäten nach den berufsstrukturellen Merkmalen des vorgestellten Ausbildungsberufs danach analysiert, wie Geschlechtstypik und sozialer Status den Treatmenteffekt moderieren. Getrennte Modelle für Jungen und Mädchen tragen dazu bei, Geschlechterunterschiede in diesem Zusammenhang herauszuarbeiten.

Ergebnisse und ihre Bedeutung

Berufliche Rollenmodelle können ersten Ergebnissen zu Folge berufliche Aspirationen unabhängig vom vorgestellten Ausbildungsberuf im Mittel leicht erhöhen. Bei der Analyse berufsstruktureller Merkmale zeigen sich bedeutsame Unterschiede nach der Geschlechtstypik des Ausbildungsberufs. Berufe mit einem niedrigeren Anteil des eigenen Geschlechts führen zu größeren Veränderungen der Berufsaspirationen im Vergleich zu geschlechtstypischen Berufen. Diese Befunde sind für Mädchen noch ausgeprägter als für Jungen. Auch der soziale Status eines Ausbildungsberufs beeinflusst den Treatmenteffekt dergestalt, dass ein geringerer sozialer Status nach der Intervention zu höheren Berufsaspirationen für den vorgestellten Beruf führt. Dieser Zusammenhang trifft auf beide Geschlechter gleichermaßen zu.

Die Ergebnisse verdeutlichen, dass auch kurzfristige und ungeplante Ereignisse, wie Begegnungen mit Rollenmodellen, im Berufswahlprozess bedeutsam werden. Gleichzeitig widersprechen sie den formulierten theoretischen Erwartungen, dass insbesondere Rollenmodelle in geschlechtstypischen und statusreichen Berufen eine Veränderung von Berufsaspirationen bewirken. Im Gegenteil können berufliche Rollenmodelle durchaus einen Ansatz darstellen, um geschlechtsuntypische Berufe mit einem geringeren sozialen Status in das Interesse von Jugendlichen zu rücken, und dadurch ggf. falsche Vorstellungen oder Stereotype zur Geschlechtersegregation verändern.

S-42: Gesundheitsfördernde Führung in der Schule – Drei Perspektiven

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 165

Chair der Sitzung: Anita Sandmeier

Gesundheitsfördernde Führung in der Schule – Drei Perspektiven

Chair(s): Anita Sandmeier (Pädagogische Hochschule Schwyz, Schweiz)

Die Gesundheit und das Wohlbefinden von Lehrpersonen hängt in vielfacher Weise mit Indikatoren der Schul- und Unterrichtsqualität (Hascher und Waber 2021; Klusmann et al. 2016; Madigan und Kim 2021) zusammen, weshalb sich die Frage stellt, wie eine ganzheitliche Gesundheitsförderung am Arbeitsplatz Schule aussehen kann und welche Rolle Schulleitungen darin spielen. Diese werden als wesentlich für die Qualität von Bildung und Schule sowie für die Initiierung und nachhaltige Umsetzung von Innovationen in Schulen gesehen (Huber und Schneider 2022; Leithwood et al. 2020). Aktuelle Überblicksartikel (Grissom et al. 2021; Liebowitz und Porter 2019) zeigen klare Evidenz für den Zusammenhang zwischen der Schulleitung und der Gesundheit der Lehrpersonen und der Organisationalen Gesundheit (u.a. Schulklima, Zufriedenheit der Schüler und der Lehrpersonen). Schulleitungen beeinflussen die Rahmenbedingungen ihrer Schulen, indem sie motivationale, organisatorische und inhaltliche Aspekte gestalten (Dadaczynski et al. 2020; Harazd et al. 2009; Samdal und Rowling 2011). Bislang ist die Erforschung der Frage, welche Einstellungen, Kompetenzen und Verhaltensweisen der Schulleitungen die Gesundheit beeinflussen, noch relativ unsystematisch (Leksy et al. 2023; Liebowitz und Porter 2019).

Dadaczynski, Carlsson und Gu (2022) schlagen zur Systematisierung der Forschung drei Perspektiven auf die gesundheitsfördernde Führung vor: a. Gesundheitliche Belastungen und Beanspruchungen von Schulleitungen angesichts hoher beruflicher Anforderungen und die Relevanz selbstbezogener gesundheitsfördernden Führung, b. Schulleitungshandeln und Gesundheit von Schulpersonal und c. Schulleitungen als Gatekeeper von gesundheitsförderlichen Veränderungsprozessen.

Das Symposium hat das Ziel, die Rolle von Schulleitungen in schulischer Gesundheitsförderung aus diesen drei Perspektiven zu betrachten.

Der erste Beitrag fasst basierend auf einer narrativen Systematisierung der internationalen Forschung die Bedeutung von Schulleitungen im Kontext des Themas Gesundheit in den drei Perspektiven zusammen.

Der zweite Beitrag nimmt die selbstbezogene gesundheitsfördernde Führung in den Blick und berichtet Ergebnisse einer Mixed-Methods-Studie an österreichischen Schulleitungen, in der die Dynamik von Anforderungen und Ressourcen im Schulleitungshandeln und die damit verbundenen gesundheitsgefährdenden bzw. -förderlichen Prozesse beschrieben werden.

Im dritten Beitrag wird das gesundheitsfördernde Schulleitungshandeln in den Blick genommen. Anhand einer quantitativen Längsschnittstudie wird der Frage nachgegangen, durch welche vermittelnden Mechanismen transformationales Führungsverhalten das Wohlbefinden der Lehrpersonen stärken kann.

Der vierte Beitrag fokussiert die Interventionsbezogene gesundheitsfördernde Führung anhand einer längsschnittlichen mixed-methods Fallstudie einer organisationalen Gesundheitsintervention mit der Frage, welche Faktoren auf der Ebene der Organisation und der Schulleitung die Ergebnisse der Intervention beeinflussen.

Die Diskussion der Beiträge erfolgt gelenkt von Leitfragen mit dem Publikum.

Beiträge des Symposiums

Schulleitungen im Kontext von Gesundheit. Überblick über Themen und Perspektiven einer vernachlässigten Zielgruppe der schulischen Gesundheitsförderung

Kevin Dadaczynski

Fachbereich Gesundheitswissenschaften, Hochschule Fulda und Zentrum für Angewandte Gesundheitswissenschaften, Leuphana Universität Lüneburg

Hintergrund: Seit Verabschiedung der Ottawa Charta im Jahr 1986 sind international wie auch national zahlreiche Bemühungen der Entwicklung und Verankerung von schulischer Gesundheitsförderung unternommen wurden. Diese reichen von einfachen, zeitlich begrenzten Maßnahmen mit verhaltenspräventivem Fokus bis hin zur Etablierung ganzheitlicher Strategien der gesundheitsförderlichen Schulentwicklung auf Ebene verschiedener Bundesländer (Paulus & Dadaczynski, 2020). Trotz unterschiedlicher konzeptioneller Ausrichtung und Tiefe ist diesen Ansätzen gemein, dass sich diese mehrheitlich an Schülerinnen und Schüler und/oder Lehrkräften richten, während Schulleitungen bislang kaum adressiert werden. Studienbefunde weisen jedoch darauf hin, dass Leitungen nicht nur bei der Umsetzung gesundheitsbezogener Veränderungsprozesse eine zentrale Gatekeeper-Funktion einnehmen, sondern selbst auch von verschiedenen Belastungen und Beanspruchungen betroffen sind (Dadaczynski, in Druck).

Fragestellung: Vor dem Hintergrund des beschriebenen Desiderats in Forschung und Praxis geht der Beitrag der Frage nach, welche Bedeutung Schulleitungen im Kontext des Themas Gesundheit haben und welche Themen und Perspektiven sich hierbei als relevant erweisen.

Methoden: Zur narrativen Systematisierung der internationalen Forschung wird auf einen Rahmen von Dadaczynski, Carlsson und Gu (2022) zurückgegriffen, der drei grundlegende Themenperspektiven unterscheidet: (1) Gesundheitliche Belastungen und Beanspruchungen von Schulleitungen, (2) Schulleitungshandeln und Gesundheit von Schulpersonal und (3) Schulleitungen als Gatekeeper von gesundheitsförderlichen Veränderungsprozessen. Entgegen einer isolierten Betrachtung wird davon ausgegangen, dass sich diese Perspektiven wechselseitig beeinflussen (z. B. schulleitungsbezogene Belastungen und Beanspruchungen behindern die Umsetzung eines gesundheitsförderlichen Führungsstils) und nicht unabhängig von den strukturellen und gesetzlichen Rahmenbedingungen auf Schulebene sowie kommunaler und Landesebene zu sehen sind.

Ergebnisse und ihre Bedeutung:

Im Hinblick auf die gesundheitliche Situation weist die Studienlage auf ein hohes und zum Teil zunehmendes Belastungspotential und auf kurz- und mittelfristige gesundheitliche (vor allem psychische) Beanspruchungen von Schulleitungen, die das Niveau anderer Berufsgruppen übersteigt. Diese scheinen tendenziell für weibliche und Grundschulleitungen höher ausgeprägt zu sein. Für das Führungshandeln erweist sich ein unterstützender und fürsorgender Führungsstil als gesundheitsförderlich, während es hinsichtlich der gesundheitlichen Bedeutung des transaktionalen Führungsstils bislang noch Unklarheiten gibt. Schließlich konnte herausgestellt werden, dass die Umsetzung und nachhaltige Verankerung von Maßnahmen der schulischen Gesundheitsförderung u. a. durch Einstellungen, Kompetenzen und managementbezogene Fähigkeiten der Schulleitung

gefördert wird. Die vorgestellten Befunde legen einen eindeutigen Handlungsbedarf nahe, Schulleitungen systematische als eigenständige Zielgruppe in der Gesundheitsförderung und Prävention zu adressieren. Dies umfasst auch explizit die Förderung der Gesundheit von Schulleitungen, dies nicht zuletzt auch, um angesichts des bestehenden Schulleitungsmangels positive Sekundäreffekte auf die Berufsattraktivität zu erzeugen. Da viele Studien außerhalb Deutschlands durchgeführt wurden, ist eine stärkere Belebung nationaler Forschungsaktivitäten angezeigt (u. a. durch Einbezug des Themas in bereits bestehende und kontinuierlich umgesetzte Studien).

Selbstbezogene gesundheitsfördernde Führung: Herausforderungen und Strategien österreichischer Schulleitungen

Jonas Pfurtscheller, Jana Groß Ophoff, Sophia Richter
Pädagogische Hochschule Vorarlberg

Theoretischer Hintergrund: Die selbstbezogene gesundheitsfördernde Führung bezieht sich auf die Maßnahmen und Strategien, die Schulleitungen anwenden, um ihre eigene Gesundheit zu bewahren und zu fördern (Dadaczynski et al., 2022), insbesondere angesichts der hohen beruflichen Anforderungen und Belastungen, die ihre Rolle mit sich bringt (Cramer et al., 2021). Bakker (2011) unterscheidet in dem sogenannten Job-Demands-Resources-Modell (JDR) zwischen Arbeitsanforderungen und Arbeitsressourcen, die über sowohl motivationale als auch gesundheitsgefährdende Prozesse auf das Arbeitsengagement wirken und darüber vermittelt die Gesundheit von Schulleitungen beeinflussen. Zu den Anforderungen zählen physische und psychische Kosten verursachende Merkmale der Berufstätigkeit wie hohe Arbeitsbelastung, Zeitdruck oder auch eingeschränkte Entscheidungsbefugnisse. Im Mittelpunkt des Vortrags stehen österreichische Schulleiterinnen und Schulleiter, die im internationalen Vergleich (ähnlich wie deutsche Schulleitungen) vergleichsweise wenig Autonomie bei der Einstellung, Ernennung, Vergütung und Beförderung von Lehrkräften aufweisen, was sich u.a. darin ausdrückt, dass sie weniger als 30 Prozent der schulischen Ressourcen kontrollieren (OECD-Durchschnitt: 39%, vgl. OECD, 2016). Außerdem erhalten österreichische Schulleitungen nur begrenzt Unterstützung durch administratives Personal und sind zum Unterrichten verpflichtet (Kemethofer, 2022; Tulowitzki et al., 2019). Um den mit hoher Belastung einhergehenden gesundheitsgefährdenden Prozess "mildern" zu können, sind Bakker und Demerouti (2017) zufolge berufliche (z.B. Unterstützung durch Kolleg:innen) wie auch persönliche Ressourcen im beruflichen Umfeld (z.B. Selbstwirksamkeitsüberzeugungen) von Bedeutung.

Fragestellungen

1. Inwiefern stehen motivationale und gesundheitsgefährdende Prozesse in Zusammenhang mit den für (österreichische) Schulleitungen typischen Arbeitsanforderungen und -ressourcen und können dadurch berufliche, insbesondere gesundheitliche Auswirkungen vorhergesagt werden?
2. Was wird über die als eingeschränkt erlebte Autonomie hinaus als Belastung thematisiert und von welchen Formen des Umgangs damit berichten österreichische Schulleitungen?

Methoden: Grundlage für die vorgestellten Ergebnisse ist der sog. Schulleitungsmonitor, der als Surveystudie in regelmäßigen Abständen im deutschsprachigen Raum durchgeführt wird, u.a. auch in Österreich im Jahr 2022/23 (N = 2621 Teilnehmer:innen). Im Anschluss an die eigentliche Fragebogenerhebung wurden problemzentrierte Interviews (N = 24; sog. Vertiefungsdesign, Hagenauer et al., 2023) durchgeführt, um die Dynamik von Anforderungen und Ressourcen im Schulleitungshandeln und damit verbundenen gesundheitsgefährdenden bzw. -förderlichen Prozessen besser zu verstehen. Die quantitativ-methodischen Daten wurden mittels latenter Strukturgleichungsmodelle (R-Package lavaan, Rosseel et al., 2023), die Interview-Transkripte inhaltsanalytisch ausgewertet und die Ergebnisse anschließend trianguliert.

Ergebnisse: Das entsprechend der JDR-Theorie modellierte latente Pfadmodell weist einen akzeptablen bis guten Modell-Fit auf ($\chi^2/df = 4.00$; CFI = .95; TLI = .94; RMSEA = .04; SRMR = .04). Die auf Innovieren bezogene Selbstwirksamkeitserwartungen (persönliche Ressource) haben einen signifikant positiven Effekt ($\beta = .44$), während sich eine als gering erlebte Entscheidungsautonomie als negativer Prädiktor des Arbeitsengagements bewährt ($\beta = -.32$). Arbeitsbezogene Anspannung erweist sich außerdem modellkonform als Mediator für den arbeitsbezogenen Stress ($\beta = .51$). Erste Analysen der Interviews deuten darauf hin, dass Mangel auf verschiedenen Ebenen und (angesichts ständiger Neuerungen) nicht greifende Routinen als belastend thematisiert werden. Hierbei scheint ein Spannungsverhältnis zwischen dem Bedürfnis nach Autonomie einerseits und nach Klarheit der Vorgaben andererseits zu bestehen, was sowohl ein Ohnmachtsgefühl hervorrufen, aber auch als Herausforderung angenommen werden kann.

Diskussion: In dem Vortrag werden die Befunde vorgestellt, trianguliert und daraus Schlussfolgerungen abgeleitet. Ein besonderes Augenmerk wird auf den Reformbemühungen im österreichischen Bildungswesen der letzten Jahre liegen, wie z.B. der Bildungsreform 2017, die österreichischen Schulen erweiterte Gestaltungsmöglichkeiten, aber auch Verantwortung in Schul- und Unterrichtsentwicklungsprozessen gegeben hat. Die damit verbundene Neuordnung der Akteurskonstellation in einer als fragmentiert, widersprüchlich und komplex beschriebenen Governance-Struktur hat vermutlich dazu geführt, dass den inhaltlich differenzierten wie pointierten professionellen Belangen von z.B. Schulleitungen derzeit schwer Gehör zu verschaffen ist (Groß Ophoff et al., in Druck). Zugleich besteht in Österreich ein anhaltender Mangel an Schulleitungen, weshalb in Qualifizierungs- und Fortbildungsangeboten ein besonderes Augenmerk auf der Resilienzförderung liegen sollte, damit Schulleitungen möglichst lange und gesund ihr Amt ausüben können.

Transformationale Führung und Wohlbefinden der Lehrpersonen

Simone Schoch, Roger Keller
Pädagogische Hochschule Zürich

Hintergrund: Die wachsenden Anforderungen an Lehrpersonen haben in den letzten Jahren zu einem Anstieg von Stress und emotionaler Erschöpfung in diesem Berufsfeld geführt (Iriarte Redín & Erro-Garcés, 2020; Sandmeier et al., 2017). Daher gewann die Gesundheitsförderung, verbunden mit der Frage, was Lehrpersonen gesund erhält, zunehmend an Bedeutung (Naidoo & Wills, 2019). Basierend auf dem Job Demands-Resources Modell (JD-R Modell) (Bakker et al., 2023) wird im Beitrag der Zusammenhang zwischen dem Führungsverhalten der Schulleitung und dem Wohlbefinden von Lehrpersonen in den Blick genommen. Bisherige Forschung hat gezeigt, dass insbesondere die transformationale Führung mit positiven Ergebnissen verbunden ist (Berkovich, 2018). In den letzten Jahren hat sich zudem ein Forschungsstrang entwickelt, der zwischen der individuellen und der kollektiven Dimension von transformationaler Führung unterscheidet (Schoch et al., 2021; Wang & Howell, 2010). Unklar ist allerdings, was die vermittelnden Mechanismen zwischen der individuellen und kollektiven transformationalen Führung und dem Wohlbefinden der Lehrpersonen sind.

Fragestellung: Im Vortrag werden die Ergebnisse zu zwei zentralen Fragestellungen präsentiert. Fragestellung 1 fokussiert den Motivationsprozess des JD-R Modells. Es wird untersucht, inwiefern die Zusammenarbeit im Team sowie die Befriedigung der Bedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit gemäss Self-determination Theory (SDT) (Deci et al., 2017) den Zusammenhang zwischen der transformationalen Führung und dem Erleben positiver Emotionen von Lehrpersonen vermitteln. Fragestellung 2 beleuchtet die Mechanismen des Gesundheitsgefährdungsprozesses des JD-R Modells und es wird das Zusammenwirken von Anforderungen (unklare organisatorische Rahmenbedingungen) und Ressourcen (individuelle transformationale Führung und Bedürfnisbefriedigung) in Bezug auf die emotionale Erschöpfung der Lehrpersonen analysiert.

Methode: Zur Beantwortung der Fragestellungen wurde eine quantitative Längsschnittstudie mit N = 1271 Lehrpersonen in der Deutschschweiz zu drei Zeitpunkten durchgeführt. Die angenommenen Wirkzusammenhänge wurden mittels Strukturgleichungsmodell sowie serieller multipler Mediationsanalyse überprüft.

Ergebnisse und ihre Bedeutung: Die Ergebnisse zu Fragestellung 1 zeigten, dass die individuelle transformationale Führung die Bedürfnisse nach Autonomie und sozialer Eingebundenheit befriedigen kann, was wiederum zu einer Erhöhung von positiven Emotionen bei Lehrpersonen führt. Bezüglich kollektiver transformationaler Führung waren die Ergebnisse weniger eindeutig, deuten jedoch darauf hin, dass sie zu mehr sozialer Unterstützung innerhalb des Teams führt. Dadurch werden die Bedürfnisse nach sozialer Eingebundenheit und Autonomie stärker befriedigt, was in der Folge zur Steigerung positiver Emotionen bei Lehrpersonen beiträgt. Die Ergebnisse zu Fragestellung 2 zeigten, dass Schulleitungen durch transformationale Führung zur Reduktion der emotionalen Erschöpfung der Lehrpersonen beitragen können. Der Zusammenhang wird grösstenteils durch die Wahrnehmung der organisatorischen Rahmenbedingungen und die Befriedigung des Autonomiebedürfnisses der Lehrpersonen erklärt.

Die Resultate tragen zu einem besseren Verständnis des Motivations- und Gesundheitsprozesses im JD-R Modell im Schulkontext bei. Es kann davon ausgegangen werden, dass keine direkten Zusammenhänge zwischen transformationaler Führung und positiven Emotionen beziehungsweise emotionaler Erschöpfung von Lehrpersonen bestehen. Schulleitungen können das Wohlbefinden der Lehrpersonen fördern, indem sie durch transformationale Führung (1) die psychologischen Grundbedürfnisse der Lehrpersonen befriedigen, (2) günstige Voraussetzungen für ein unterstützendes Teamklima schaffen und (3) ein unterstützendes Arbeitsumfeld gestalten, das durch klare Rahmenbedingungen und Verantwortungsbereiche gekennzeichnet und auf die individuellen Ressourcen und Fähigkeiten der Mitarbeitenden ausgerichtet ist. Auf diese Weise bieten sie den Lehrpersonen Sicherheit und den notwendigen Handlungsspielraum für eigenverantwortliches Arbeiten.

Führung einer organisationalen Gesundheitsintervention – eine längsschnittliche mixed methods Fallstudie

Anita Sandmeier, Laura Koch
Pädagogische Hochschule Schwyz

Hintergrund: Der Bedarf an effektiven Interventionen zur Förderung der Gesundheit des Schulpersonals hat angesichts der Studienlage zur Gesundheit von Lehrkräften zugenommen. Dabei wurde vor allem in Trainings und Weiterbildungen für Lehrpersonen investiert. Maßnahmen welche die organisationalen Verhältnisse adressieren, sind bislang selten, da sie komplex sind in der Umsetzung (McIntyre et al. 2017; Bauer und Jenny 2014). Im Fokus dieses Beitrags steht eine Intervention, in welcher ausgehend von einer Mitarbeitendenbefragung partizipativ mit dem ganzen Kollegium passgenaue Maßnahmen abgeleitet werden, um arbeitsbezogene Belastungen und Ressourcen auf der Schulebene zu stärken.

Die Evaluation von organisationalen Interventionen zur Verbesserung der Arbeitsbedingungen ist ein spezifischer Typus der Schulentwicklungsforschung (Maag Merki und Werner 2013). Klieme (2016) schlägt vor, verstärkt Ansätze der Innovations- und Implementationsforschung zu nutzen, um Schulentwicklung wissenschaftlich zu evaluieren. Anknüpfend an diesen Vorschlag wird im Beitrag ein Forschungsdesign aus der Organisationalen Gesundheitsinterventionsforschung genutzt.

Fragestellungen: 1. Wie wird die Intervention geplant und organisiert? 2. Welche intendierten und nicht intendierten Entwicklungen hat die Intervention ausgelöst? 3. Welche Führungsaspekte sind für die Implementation und den Entwicklungsprozess relevant?

Methode: Die Studie basiert auf einem mixed-methods Design mit multiplen Messzeitpunkten und Datenquellen: Ergebnisse der Mitarbeiterbefragung mit dem standardisierten Befragungsinstrument „Friendly Work Space Job-Stress-Analyse“ (Elfering et al., 2020) zu zwei Messzeitpunkten (Januar 2023 und 2024), relevante Dokumente (z.B. Schulprogramm, Schulstrategie, Philosophie), halbstrukturierte Interviews mit der Schulleitung vor Beginn der Intervention, Beobachtungen von Teamentwicklungstagen, Feldnotizen aus weiteren Beobachtungen ausgewählter Besprechungen, Telefonaten mit der Schulleitung, Gruppeninterviews mit Mitarbeitern (im Verlauf der Intervention) und halbstrukturierte Interviews mit der Schulleitung (ein Jahr nach Beginn der Intervention).

Die für den Beitrag analysierten Fälle umfassen ein Gymnasium (N= 124) mit sieben Teams und eine Förderschule (N= 123) mit vier Teams.

Das analytische Framework der Studie basiert auf dem Job Demands-Resources Model (Bakker und Demerouti 2023), das den grundlegenden Wirkungszusammenhang der Intervention konzeptualisiert und analytisch fassbar macht und dem Framework for Evaluating Organizational-level Interventions (Nielsen und Randall 2013), das hilft die Wirkprozesse komplexer Interventionen strukturiert zu analysieren. Die Auswertung erfolgt durch eine inhaltlich strukturierende qualitative Inhaltsanalyse gemäß Kuckartz & Rädiker (2022). Die Integration der verschiedenen Datenquellen in die Analyse folgt dabei sowohl dem Prinzip der Konvergenz als auch dem Prinzip der Erweiterung (Yin 2018, S. 129). Die Analysen werden dabei zuerst für beide Fälle durchgeführt und anschließend in einer cross-case Analyse über die Fälle hinweg kontrastiert (Yin 2018, S. 196).

Ergebnisse und ihre Bedeutung: Die Analyse des ersten Falls zeigt die hohe Relevanz der Schulleitung und deren Kapazität und Bereitschaft, genügend Ressourcen und Zeit zur Verfügung zu stellen, um die identifizierten Belastungen nachhaltig zu bearbeiten.

Die größten Belastungen der Schule liegen im sozialen Bereich des Teams. So zeigt die Personalbefragung tiefe Werte in gemeinsamen pädagogischen Vorstellungen und in partizipativer Sicherheit im Team bei gleichzeitig sozialen Belastungen durch Kolleg:innen. Die größten Ressourcen des Personals liegen in der hohen Arbeitszufriedenheit, im positiven Erleben der Arbeit, in der emotionalen Bindung zur Schule und in der positiven Einschätzung der Schulleitung bezüglich Unterstützung und Wertschätzung.

Die Schulleitung beschließt einfach umzusetzende, punktuelle Maßnahmen im Bereich Infrastruktur und Arbeitsorganisation. Belastungen im Bereich der Schulkultur (Negativität im Lehrkräftezimmer, gemeinsame pädagogische Vorstellungen bezüglich

Beurteilung), die vom Kollegium als wichtig eingeschätzt wurde, werden nicht mit weiterverfolgt. Dies führte zu unerwünschten Veränderungsprozessen wie Frustration bei einzelnen Gruppen und unbefriedigenden Ergebnisse der Intervention.

Diese Ergebnisse werden mit dem zweiten Fall kontrastierend ergänzt mit dem Ziel hinderliche und förderliche Faktoren für effektive organisationalen (Gesundheits-)Interventionen zu identifizieren und damit einen Beitrag zum Verständnis von komplexen Entwicklungsprozessen zu leisten.

Die Studie leistet mit dem längsschnittlichen mixed methods Ansatz praktische Erkenntnisse für die Führung von Schulentwicklungsprojekten und prüft die Anwendbarkeit von Forschungsdesigns aus verwandten Disziplinen für die Schulentwicklungsforschung.

S-43: Open Science in der (quantitativen) Bildungsforschung: Herausforderungen, Werkzeuge und Perspektiven

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SN 163

Chair der Sitzung: Jürgen Schneider

Open Science in der (quantitativen) Bildungsforschung: Herausforderungen, Werkzeuge und Perspektiven

Chair(s): **Jürgen Schneider** (DIPF, Deutschland)

Mit der Open Science Bewegung sind hohe Erwartungen verknüpft: Durch sie soll Transparenz erhöht, Forschungsprozesse optimiert und Bildungsgerechtigkeit gefördert werden (DFG, 2022). In vielen Bereichen des Wissenschaftssystems wird deshalb ein Kulturwandel gefordert (UNESCO 2022). Forschungsförderer wie die DFG (2015, 2019), das BMBF und das ERC (2022, 2023) sprechen sich im Zuge der Vergabe von Mitteln für die Anwendung offener Praktiken aus. Infolgedessen ist in den letzten Jahren eine bedeutende Zunahme der Umsetzung offener Forschungspraktiken zu verzeichnen (Cao et al., 2023; UNESCO, 2023). Darüber hinaus wenden sich die Einstellungen und Überzeugungen der wissenschaftlich Forschenden über die Jahre hinweg immer mehr Open Science zu (Borycz et al., 2023).

Auch in der deutschsprachigen Bildungsforschung ist Open Science seit mehreren Jahren ein Teil des wissenschaftlichen Diskurses (DGfE, 2018, 2020; Krammer & Svecnik, 2020). Um die erfolgreiche Umsetzung von Open Science Praktiken zu unterstützen, konnten bereits fachspezifische Infrastrukturen für Forschungsdaten (Verbund Forschungsdaten Bildung) und Open Access Publikationen (Repositorium peDOCS) etabliert werden. Darüber hinaus begannen wissenschaftliche Fachzeitschriften Beitragsformate anzubieten, die spezifisch auf die Umsetzung von Open Science Praktiken ausgelegt sind (s. Registered Reports bei z.B. Learning and Instruction oder der Zeitschrift für Bildungsforschung). Trotz dieser Fortschritte bestehen weiterhin Herausforderungen und Barrieren, die eine breite Implementierung offener Praktiken erschweren (Huff & Bongartz, 2023).

Das Symposium zielt darauf ab, aktuelle Themen und Ansätze zur Förderung der Offenheit und Reproduzierbarkeit in der quantitativ-empirischen Bildungsforschung zu diskutieren. Im Fokus stehen praxisnahe Beiträge, die sich mit den Herausforderungen und Möglichkeiten des Teilens von Forschungsdaten, der Steigerung der Reproduzierbarkeit von Analysen und der Implementierung von Präregistrierungen und Registered Reports in der Qualifikationsphase beschäftigen. Durch den Austausch möchten wir uns über praxistaugliche und akzeptanzfähige Umsetzungen von Open Science Praktiken verständigen.

1. Beitrag: Geringe Verfügbarkeit von Forschungsdaten in pädagogisch-psychologischen Journals: Keine Hinweise auf wirksame Forschungsdaten-Policies

Der erste Beitrag widmet sich der Verfügbarkeit von Forschungsdaten in sechs pädagogisch-psychologischen Zeitschriften und vergleicht sie mit der Zeitschrift Cognition. Auf Basis einer Inhaltsanalyse von N = 1.242 Publikationen aus sechs pädagogisch-psychologischen Zeitschriften wird der Zusammenhang von Zeitschriftenrichtlinien und institutionellen Vorgaben auf die Datenverfügbarkeit untersucht. Trotz erhöhter Sensibilisierung zeigt sich die Verfügbarkeit allerdings weiterhin als gering ausgeprägt.

2. Beitrag: Open Data in Large-Scale Assessments: Chancen und Herausforderungen am Beispiel von TALIS

Der zweite Beitrag fokussiert auf die bisherigen Praktiken im Bereich Open Data am Beispiel des Internationalen Large-Scale-Assessments TALIS der OECD. Dieser bietet zugängliche Daten von Lehrpersonen in über 50 Ländern. Der Beitrag beleuchtet die Herausforderungen bei der Nutzung von TALIS-Daten im Kontext von Open Science Praktiken, wie Datenschutz und komplexe Auswertungsanforderungen. Er diskutiert, wie Open Science-Praktiken in Sekundäranalysen gefördert werden können, um transparente und praxisorientierte Nutzung zu gewährleisten.

3. Beitrag: Einfache Möglichkeiten die Reproduzierbarkeit der eigenen Analysen zu steigern

Der dritte Beitrag präsentiert einfache und effektive Werkzeuge zur Steigerung der Reproduzierbarkeit von Datenanalysen. Im Mittelpunkt stehen verschiedene R Pakete, die eine einfache Verwaltung von Abhängigkeiten und die Erstellung reproduzierbarer Umgebungen ermöglichen. Ergänzend werden die Open-Source-Statistiksoftware (z.B. JASP) als benutzerfreundliche Alternativen zu proprietären Lösungen präsentiert. Durch Demonstrationen und Anwendungsbeispiele sollen die Teilnehmenden Anregungen für die eigene Umsetzung erhalten.

4. Beitrag: Peer-Review-Prozesse anders gedacht: Registered Reports als Formate in der Qualifikationsphase und darüber hinaus

Der vierte Beitrag beleuchtet Herausforderungen des klassischen Peer-Review-Prozesses und stellt vor diesem Hintergrund Registered Reports vor. Der Beitrag geht dabei generell auf wissenschaftliche Qualitätssicherung ein und fokussiert das Potential von Registered Reports und der damit verbundenen Adaption des Peer-Review-Prozesses. Dabei werden sowohl die Perspektiven von Editor:innen wissenschaftlicher Journals als auch die Rahmenbedingungen von Forschenden in Qualifizierungsphasen thematisiert. Gemeinsam mit den Anwesenden wird diskutiert, welche Herausforderungen und Vorteile mit Registered Reports im Publikationsalltag und speziell in Qualifikationsphasen verbunden sind.

Beiträge des Symposiums

Geringe Verfügbarkeit von Forschungsdaten in pädagogisch-psychologischen Journals: Keine Hinweise auf wirksame Forschungsdaten-Policies

Markus Huff¹, Elke C. Bongartz²

¹Leibniz-Institut für Wissensmedien, Tübingen; Eberhard Karls Universität Tübingen, ²Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen

Theoretischer Hintergrund: Mittlerweile ist Open Science als Basis guter wissenschaftlicher Praxis breit akzeptiert. Die einzelnen Dimensionen von Open Science erfordern jedoch individuelle Anpassungen an die jeweilige Fachkultur. Der Beitrag konzentriert sich auf die Open Science-Dimension Forschungsdaten. In den letzten Jahren gab es zahlreiche Initiativen, die darauf abzielen, die Verfügbarkeit und Qualität von Forschungsdaten zu erhöhen (z.B. FAIR data; Wilkinson et al., 2016). Diese Initiativen haben das Bewusstsein für die Bedeutung von Datenverfügbarkeit gesteigert und dazu geführt, dass viele Forschungseinrichtungen entsprechende Richtlinien zum Umgang mit Forschungsdaten eingeführt haben. Der theoretische Hintergrund dieser Studie basiert auf der Annahme, dass eine erhöhte Sensibilisierung und institutionelle Richtlinien zur Forschungsdaten zu einer höheren Verfügbarkeit von Forschungsdaten führen könnten. Vor diesem Hintergrund ist es wichtig, den aktuellen Stand der

Datenverfügbarkeit in der Pädagogischen Psychologie zu bewerten und die Auswirkungen institutioneller Richtlinien und journalbezogener Transparenzanforderungen (i.e. data transparency level; Center for Open Science, 2020) zu untersuchen.

Fragestellung: Die zentrale Fragestellung dieser Studie lautet: Hat die erhöhte Sensibilisierung für Forschungsdatenverfügbarkeit und die Einführung institutioneller Richtlinien tatsächlich zu einer höheren Verfügbarkeit von Forschungsdaten in empirischen Artikeln geführt? Um diese Frage zu beantworten, wurden zwei spezifische Forschungsfragen formuliert: 1. Wie hoch ist die Verfügbarkeit von Forschungsdaten in empirischen Artikeln in pädagogisch-psychologischen Zeitschriften im Vergleich zu der Referenzzeitschrift *Cognition* aus der kognitiven Psychologie (Hardwicke et al., 2018)? 2. Inwiefern beeinflussen das data transparency level der Zeitschriften und die Existenz offizieller Forschungsdatenrichtlinien auf der Ebene der Institution des korrespondierenden Autors die Verfügbarkeit von Forschungsdaten?

Methode: Zur Beantwortung der Forschungsfragen wurde eine Inhaltsanalyse von N = 1.242 Publikationen aus sechs pädagogisch-psychologischen Zeitschriften sowie der Zeitschrift *Cognition* durchgeführt. Die Auswahl der Artikel erfolgte für die Jahre 2018 und 2020, um mögliche zeitliche Veränderungen zu erfassen. Die Verfügbarkeit von Forschungsdaten wurde durch eine systematische Kodierung der Artikel ermittelt. Dabei wurden zwei Hauptmerkmale untersucht: Erstens, ob die Artikel Forschungsdaten bereitstellen, und zweitens, ob eine offizielle Richtlinie zur Forschungsdatenverfügbarkeit an der Institution des korrespondierenden Autors existiert. Zusätzlich wurde das data transparency level der jeweiligen Zeitschriften berücksichtigt.

Ergebnisse und ihre Bedeutung: Die Analyse ergab, dass die Verfügbarkeit von Forschungsdaten in den pädagogisch-psychologischen Zeitschriften insgesamt niedrig war (Huff & Bongartz, 2023). Im Jahr 2018 lag die Datenverfügbarkeit bei lediglich 0,32%, stieg jedoch bis 2020 auf 7,16%. Im Vergleich dazu zeigte die Referenzzeitschrift *Cognition* eine deutlich höhere Datenverfügbarkeit von 62,74%. Trotz des Anstiegs in den pädagogisch-psychologischen Zeitschriften bleibt die Datenverfügbarkeit im Vergleich zu *Cognition* sehr gering. Ein weiteres zentrales Ergebnis war, dass weder das data transparency level der Zeitschriften noch die Existenz offizieller Forschungsdatenrichtlinien auf der Ebene der Institution des korrespondierenden Autors einen signifikanten Einfluss auf die Verfügbarkeit von Forschungsdaten hatten. Diese Ergebnisse legen nahe, dass institutionelle Richtlinien und journalbezogene Transparenzanforderungen allein nicht ausreichen, um eine signifikante Erhöhung der Datenverfügbarkeit zu bewirken.

Erstens weisen die Ergebnisse darauf hin, dass weitere Maßnahmen erforderlich sind, um die Verfügbarkeit von Forschungsdaten in der Bildungspsychologie zu erhöhen. Dies könnte beispielsweise durch die Einführung strengerer Richtlinien und Anreize auf der Ebene der Zeitschriften und Institutionen geschehen. Zweitens unterstreichen die Ergebnisse die Notwendigkeit einer verstärkten Sensibilisierung und Schulung von Forschenden hinsichtlich der Bedeutung und Vorteile der Datenfreigabe (z.B. Nachnutzung). Vor dem Hintergrund dieser Ergebnisse diskutieren wir konkrete Empfehlungen für notwendige Weiterentwicklungen im Forschungsdatenmanagement und betonen die Rolle von Forschungseinrichtungen und wissenschaftlichen Zeitschriften bei der Förderung von Transparenz und Reproduzierbarkeit in der Wissenschaft.

Insgesamt zeigt die Studie, dass trotz eines Anstiegs der Datenverfügbarkeit in den letzten Jahren noch erheblicher Verbesserungsbedarf besteht. Die Ergebnisse liefern wichtige Hinweise für zukünftige Maßnahmen zur Förderung der Datenverfügbarkeit und zur Verbesserung der Glaubwürdigkeit und Transparenz in der Bildungspsychologie und der Wissenschaft im Allgemeinen.

Open Data in Large-Scale Assessments: Chancen und Herausforderungen am Beispiel von TALIS

Corinna Koschmieder¹, Marie Bilalovic², Manfred Herzog¹, Martin Auferbauer³, Georg Krammer⁴

¹Zentrum für Forschungssupport, Pädagogische Hochschule Steiermark, Österreich, ²Institut für Mehrsprachigkeit und Transkulturelle Bildung, Pädagogische Hochschule Kärnten, Österreich, ³Institut für Bildungswissenschaften, Pädagogische Hochschule Steiermark, Österreich, ⁴Institut für Wirtschafts- und Berufspädagogik, Johannes Kepler Universität Linz, Österreich

Open Data, ist ein wachsender Trend in der wissenschaftlichen Praxis, der darauf abzielt, Forschungsergebnisse und -daten frei zugänglich und wiederverwendbar zu machen. Dies fördert Transparenz, Reproduzierbarkeit und eine breitere Nutzung wissenschaftlicher Erkenntnisse. In groß angelegten internationalen Studien ist Open Data besonders wertvoll, da es Forschenden ermöglicht, auf repräsentative Datensätze vieler verschiedener Länder zuzugreifen und diese für sekundäre Analysen zu nutzen.

Unter diesen ist der Teaching and Learning International Survey (TALIS) der OECD ist eine zentrale Ressource für Bildungsforschende weltweit. Er bietet umfassende Einblicke in die Bedingungen des Lehrberufs, die Lernumgebungen an Schulen und aktuelle Themen wie multiprofessionelle Zusammenarbeit, Nachhaltigkeit und Arbeitszufriedenheit von Lehrkräften. TALIS basiert auf einem stratifizierten Sample mit Daten aus drei Erhebungszyklen und über 50 teilnehmenden Ländern, die offen zugänglich sind.

Ein systematischer Review von Veletic et al. (2024) hebt das große Forschungspotential von TALIS-Daten hervor, insbesondere in den Bereichen Schulleitungsforschung, Autonomie, Zusammenarbeit, Feedback-Management sowie Aus- und Weiterbildung von Lehrkräften. Trotz der Erfüllung vieler Prinzipien und Praktiken im Kontext von Open Science fehlt bisher eine direkte wissenschaftliche Diskussion, die Open Science mit internationalen Large-Scale Assessments (ILSA) verbindet.

Dieser Beitrag zielt darauf ab, die Lücke in der wissenschaftlichen Diskussion zu schließen und beleuchtet die praktischen und methodischen Herausforderungen bei der Nutzung von TALIS-Daten unter den Aspekten der Open Science Prinzipien. Dabei werden aktuelle Praktiken in der Forschungspraxis vorgestellt und bestehende Spannungsfelder diskutiert.

So ist es nicht ausreichend die Daten einfach zu veröffentlichen. Für die Auswertung von Daten der TALIS Studie gelten komplexe Auswertungsvoraussetzungen. Um die öffentliche Nutzung der Daten zu ermöglichen braucht es Strategien, die anspruchsvollen Anforderungen der ILSA-Datenverarbeitung Forscher*innen zugänglich machen (Rutkowski 2010).

Des Weiteren beinhalten die Datensätze von TALIS viele Informationen über die teilnehmenden Lehrpersonen. Vor der Veröffentlichung der Daten müssen geeignete Maßnahmen zur Wahrung der Datenschutzbestimmungen ergriffen werden.

Auch bei der Veröffentlichung steht der Anspruch einer möglichst transparenten und umfassenden Dissemination im Konflikt mit praxisorientierter Berichterstattung für eine überwiegend nicht-wissenschaftliche Zielgruppe.

Durch diesen Beitrag sollen Forschende und Praktiker*innen für die Potenziale der TALIS-Daten sensibilisiert und zur aktiven Nutzung sowie Weiterentwicklung von Open Science-Forschungspraktiken in Sekundäranalysen ermutigt werden. Zudem wird eine Möglichkeit aufgezeigt, wie aus den bisherigen Praktiken Internationaler Large-Scale Assessments gelernt werden kann.

Einfache Möglichkeiten die Reproduzierbarkeit der eigenen Analysen zu steigern

Jürgen Schneider

DIPF

Analysen zu reproduzieren oder zu erweitern gehört zum Arbeitsalltag von Bildungsforschenden. Sei es, dass im Review eines Manuskripts um weiterführende explorative Analysen angeregt werden, oder die Reviewenden zur Qualitätsprüfung um die Bereitstellung der Analysen bitten. Ebenso können andere Forschende im Rahmen einer Meta-Analyse die Analysen erneut ausführen wollen, um die für sie notwendigen Kennwerte zu extrahieren.

Der vorliegende Beitrag betrachtet die Reproduzierbarkeit von Datenanalysen im Kontext der quantitativen Bildungsforschung unter dem Gesichtspunkt der Open-Science-Praktik „open code“. Befunde sind reproduzierbar, wenn unabhängige Forschende, die dieselben Daten und Analyseverfahren verwenden, zu identischen Resultaten gelangen (dieselben Daten, dieselben Analysen; Kitzes et al., 2018). Dieser Aspekt wird auch als computational reproducibility bezeichnet. Mit denselben Daten und Analyseverfahren zu identischen Ergebnissen zu gelangen, mag auf den ersten Blick wie eine selbstverständliche Banalität erscheinen. In der Praxis stellt sich jedoch oft heraus, dass Ergebnisse, wenn überhaupt, nur mit erheblichem Aufwand reproduziert werden können. Artner et al. (2021) konnten auch mit „tedious and time-consuming work“ (S. 12) gerade einmal 51 % von 232 untersuchten statistischen Befunden aus Artikeln der Psychologie genau reproduzieren, wenn sie das im Manuskript beschriebene Vorgehen anwendeten. Ihr Fazit fiel entsprechend deutlich aus: „Vagueness makes assessing reproducibility a nightmare.“ (Artner et al., 2021, S. 12)

Neben vermeidbaren Fehlern, wie dem Umbenennen von referenzierten Dateien, hardcoding von Dateipfaden, Kopierfehlern von Kennwerten und falschen Rundungen, stellen Updates von Paketversionen und Unterschiede in den computativen Umgebungen (Systemvoraussetzungen) zentrale Herausforderungen dar. Forschende, die Datenauswertungen reproduzieren, verwenden möglicherweise andere Betriebssysteme, Softwareversionen und Hardware als bei der Erstellung des Analyseskripts durch die Autor*innen. Diese Unterschiede in den Systemanforderungen können dazu führen, dass die fehlerfreie Ausführbarkeit eines Analyseskripts nicht immer gewährleistet ist. Auch die Autor*innen selbst könnten wenige Zeit nach der Erstellung eines Analyseskripts aufgrund von Softwareupdates auf Fehlermeldungen bei dessen Ausführung stoßen. Passende Ansätze müssen also zwei grundsätzliche Probleme lösen: (1) Unterschiede in den Systemvoraussetzungen und (2) Unterschiede in den darin genutzten Paketversionen (Versionsmanagement).

Der Beitrag stellt konkrete Werkzeuge und Strategien vor, die eine Steigerung der Reproduzierbarkeit bezüglich dieser Herausforderungen ermöglicht. Dabei liegt der Fokus auf der Nutzung verschiedener R-Pakete, die speziell für ein erfolgreiches Package Management und die Herstellung reproduzierbarer computativer Umgebungen entwickelt wurden. Es wird aufgezeigt, wie die Nutzung dieser Pakete sich ohne weiteres in den eigenen Workflow integrieren lässt. Als Lösungsansätze wird (1) die Nutzung der R-Pakete {renv} gemeinsam mit {holepunch}, (2) die Nutzung von Nix anhand der Unterstützung des R-Pakets {rix} sowie (3) die Nutzung von Input-Output-Dokumenten (wie Quarto Markdown oder RMarkdown) als niederschwellige Alternative demonstriert.

Ergänzend zu den R-Paketen werden die Open-Source-Statistiksoftware JASP als benutzerfreundliche Alternativen zu proprietären Lösungen wie SPSS vorgestellt. JASP zeichnet sich durch ihre intuitive Benutzeroberfläche und die Möglichkeit aus, Analysen ohne Programmierkenntnisse durchzuführen. Diese Softwarelösung liefert automatisch umfangreiche Dokumentationen mit, die eine Nachvollziehbarkeit der durchgeführten Analysen erhöhen.

Peer-Review-Prozesse anders gedacht: Registered Reports als Formate in der Qualifikationsphase und darüber hinaus

Georg Krammer

Institut für Wirtschafts- und Berufspädagogik, Johannes Kepler Universität Linz, Österreich

Damit Bildung ihr Potential als Schlüssel für gesellschaftliche Herausforderungen entfalten kann ist es unter anderem notwendig, dass Erkenntnisse (quantitativer) Bildungsforschung auf robuster, belastbarer und reproduzierbarer Forschung beruhen. Um dies in der publizierten Fachliteratur zu gewährleisten, verlassen sich Produzent:innen und Rezipient:innen von Forschung auf den Peer-Review-Prozess. Klassisch sieht der Peer-Review-Prozess vor, dass eine abgeschlossene Studie in ihrer final verschriftlichten Form diesen Schritt der Qualitätssicherung durchläuft.

Doch was bedeutet diese Art der Qualitätssicherung für das Gewinnen wissenschaftlicher Erkenntnisse, wenn untersuchte Effekte nicht die magische Grenze von $p < .05$ bewältigen? Man bedenke, dass Widerstände gegen das Publizieren von Null-Effekten nichts Neues sind (Sterling, 1959; Sterling et al., 1995) und dass ohne publizierte Null-Effekte ein verzerrtes Abbild der Wirklichkeit entsteht (Greenwald, 1975).

Und was bedeutet es wenn junge Forscher:innen in prekären Anstellungsverhältnissen und zeitlich begrenzten Forschungsprojekten Publikationen generieren sollen oder gar ihre kumulative Dissertationen abschließen wollen? Man versetze sich dabei in folgende Situation: nach akribischer Planung, aufwendiger Datenerhebung und statistischen Auswertungen geprägt von „researcher degrees of freedom“ (Simmons et al., 2011) stellt sich heraus, dass Daten nicht die erhofften signifikanten Effekte beinhalten. Wenn das ein Versagen der Studie war, kann die Expertise von Editor:innen und Gutachter:innen im klassischen Peer-Review-Prozess nicht zur Qualitätssteigerung genutzt werden, sondern wird zur Ablehnung führen. Wenn es sich jedoch um einen Stichprobenfehler handelt oder tatsächlich ein Null-Effekt vorliegt, womöglich sogar im Widerspruch zur publizierten Literatur (s. bspw. Open Science Collaboration, 2015), wird es dadurch nicht wahrscheinlicher, dass die Studie veröffentlicht wird.

Vor diesem Hintergrund widmet sich dieser Beitrag der Frage, wie Peer-Review-Prozesse anders gestaltet werden können und geht dafür auf Registered Reports ein. Registered Reports (Nosek & Lakens, 2014) sind ein Beitragsformat wissenschaftlicher Fachzeitschriften mit einem Peer-Review-Prozess in zwei Phasen. Es soll dabei der Wert eines Forschungsvorhabens von dessen Ergebnis entkoppelt werden. Um dieses Ziel zu erreichen, wird in der ersten Phase ein Manuskript, bestehend aus Einleitung, Methode einschließlich vorgeschlagener Analysen und (optional) Pilotdaten oder „Mock Data“, begutachtet. Hochwertige Manuskripte werden nach dieser ersten Phase – noch vor der Datenerhebung! – vorläufig zur Veröffentlichung akzeptiert („in-principle acceptance“). Nachdem die Daten anschließend erhoben und die durchgeführte Studie verschriftlicht wurde, erfolgt die zweite Phase des Peer-Review-Prozesses. Dabei werden Ergebnisdarstellung einschließlich Interpretation und Diskussion begutachtet und es wird sichergestellt, dass das Forschungsvorhaben wie zuvor beschrieben umgesetzt wurde. Registered Reports betonen damit die Bedeutung der Forschungsfrage, die Qualität der Methodik und den angestrebten Erkenntnisgewinn statt Stärke und Richtung der Ergebnisse. Dieses Format ist so konzipiert, dass Publikationsbias und andere Verzerrungen der Forschung minimiert werden.

Laut dem Center for Open Science (<http://cos.io/rr/> s. Reiter „Participating Journals“) gibt es derzeit (Stand 19.7.2024) 280 Journals, welche Registered Reports als reguläres Beitragsformat anbieten. Darunter fallen u.a.: British Journal of Educational Psychology; Journal of Educational Psychology; Learning and Instruction; Mind, Brain and Education; Frontiers in Education;

Zeitschrift für Bildungsforschung. Erste Studien zur Qualität von Studien, welche als Registered Reports publiziert wurden, zeigen eine vielversprechende Qualitätssteigerung (Soderberg et al., 2021).

Über dieses Potential für eine robustere, belastbarere und reproduzierbarere Forschung hinaus, bergen Registered Reports auch das Potential wissenschaftliche Integrität in Qualifikationsphase zu steigern. Einerseits kann hier der Peer-Review-Prozess konstruktiv genutzt werden, da das Fachwissen von Editor:innen und Gutachter:innen die Qualität von Forschungsvorhaben selbst im Fall einer Ablehnung steigern kann. Andererseits können es Registered Reports ermöglichen bereits früher in Projekten Manuskripte einzureichen und vorläufig zur Veröffentlichung akzeptiert zu bekommen, wodurch eine Basis für kumulative Dissertationen geschaffen wird, die auf der Qualität des Forschungsvorhabens und nicht auf einer krankhaften Vorliebe für signifikante Effekte (s. Gullo & O’Gorman, 2012) fußt.

S-44: Augmented Reality beim Lernen durch Experimentieren

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 151

Chair der Sitzung: Sarah Malone

Chair der Sitzung: Jochen Kuhn

Chair der Sitzung: Andreas Lachner

Augmented Reality beim Lernen durch Experimentieren

Chair(s): **Sarah Malone** (Universität des Saarlandes, Deutschland), **Jochen Kuhn** (Ludwig-Maximilians-Universität München)

Diskutant*in(nen): **Andreas Lachner** (Universität Tübingen)

Dieses Symposium stellt drei empirische Studien aus einem laufenden, interdisziplinären Verbundprojekt vor, das Bildungswissenschaften und Physikdidaktiken aus Deutschland und der Schweiz gemeinsam bearbeiten. Das Projekt befasst sich mit dem Einsatz von Augmented Reality (AR) zum Einblenden essenzieller Informationen in Echtzeit, während Schülerinnen und Schüler Physikexperimente zum Elektromagnetismus durchführen.

Die Forschung im Bereich von AR in Lernkontexten nimmt stetig zu und zeigt vielversprechende Ergebnisse. Systematische Reviews und Metaanalysen attestieren AR ein hohes Potenzial, Lernprozesse und das Lernerlebnis insgesamt zu verbessern (Chang et al., 2022; Na & Yun, 2024). Insbesondere in den Naturwissenschaften wurde AR erfolgreich beim Lernen durch Experimentieren verwendet, bei dem der Einsatz von AR positive Auswirkungen auf die kognitive Belastung (z.B. Thees et al., 2020) und das konzeptuelle Lernen (z.B. Altmeyer et al., 2020) hatte.

Häufig werden Studien berichtet, die AR-gestütztes Lernen zum Beispiel mit traditionellen Unterrichtsmethoden vergleichen (*Media-Comparison-Studien*; Buchner & Kerres, 2022). Dagegen zielen wir mit unserem Projekt darauf ab, Lernprozesse innerhalb von AR-Umgebungen genau zu beobachten, zu analysieren und gezielt durch eine effektive Gestaltung dieser Umgebungen zu fördern.

Zu diesem Zweck haben wir eine interaktive Lernumgebung zur Lorentzkraft entwickelt, bei der verschiedene Darstellungen der entstehenden elektromagnetischen Felder und wirkenden Kräfte in Echtzeit eingeblendet werden, während Schülerinnen und Schülern damit experimentieren und reale, physikalische Phänomene beobachten.

Im Rahmen des Projekts haben wir zunächst in einer Reihe von Studien Instrumente zur Messung des relevanten Konzeptwissens und der Repräsentationskompetenz entwickelt, auf deren Erwerb sich unsere Lerneinheiten auswirken sollten. Anschließend wurden die Lerneinheiten in Pilotstudien hinsichtlich ihrer generellen Nutzbarkeit und Lerneffektivität überprüft.

Die drei Beiträge dieses Symposiums stellen weiterführende Studien vor, die systematisch verschiedene Gestaltungsvarianten der AR-gestützten Lernumgebung, das Zusammenspiel des Erwerbs von Konzeptwissen und Repräsentationskompetenz sowie die ablaufenden kognitiven Prozesse näher untersuchen. Das übergeordnete Ziel dieses Symposiums ist es, die Ergebnisse der einzelnen Studien zu integrieren, um zur Theoriebildung über die Interaktionen zwischen Lernendeneigenschaften (individuelle Unterschiede) und Details der Lernumgebung (z.B. Prompts) sowie der Nutzung augmentierter Repräsentationen (Prozessdaten wie z.B. Eye-Tracking) beizutragen.

Beitrag 1: Der erste Beitrag untersucht den Nutzen der Verwendung verschiedener Arten externaler Repräsentationen, die Schüler*innen nacheinander bei Experimenten zur Lorentzkraft mithilfe von AR einblenden können. In einer laufenden Studie (N=67) werden verschiedene virtuelle Darstellungen von Magnetfeldern in reale Experimente integriert, um das Konzeptwissen und die Repräsentationskompetenz zu fördern. Erste Ergebnisse zur Wirksamkeit stehen derzeit noch aus, werden aber auf der Konferenz präsentiert. Diese Studie zielt darauf ab, zu identifizieren, welche und wie viele externe Repräsentationen sich in AR-gestützten Experimenten zur Lorentzkraft als lernförderlich erweisen.

Beitrag 2: Der zweite Beitrag widmet sich der Untersuchung des Zusammenspiels von Konzeptwissen und Repräsentationskompetenz beim AR-gestützten Experimentieren. Die Studie mit 126 Sekundarschüler*innen zeigt, dass Repräsentationskompetenz zu Beginn des Lernens die Entwicklung von Konzeptwissen beim AR-gestützten Experimentieren signifikant vorhersagt, selbst wenn individuelle Unterschiede in kognitiven Fähigkeiten und Modellverständnis kontrolliert werden.

Beitrag 3: In dieser Studie (N=78) wurden erstmals sämtliche digitale Darstellungsformen (Feldlinien, Vektorpfeile, Vektordreibein) beim Experimentieren gleichzeitig eingeblendet. Um das Lernen mit diesen multiplen externalen Repräsentationen zu unterstützen, wurden der Hälfte der Lernenden auditive Selbsterklärungsprompts eingespielt, um mentale Integrationsprozesse anzuregen. Erste Ergebnisse zeigen Verbesserungen in Repräsentationskompetenz und Konzeptwissen nach der Lernphase in beiden Gruppen – jedoch ohne signifikante Vorteile für die Gruppe mit Selbsterklärungsprompts. Aufgezeichnete Augenbewegungen sollen Hinweise darauf geben, inwieweit die Prompts tatsächlich zum gewünschten Integrationsverhalten beigetragen haben.

Der projektexterne Diskutant, wird abschließend die Ergebnisse der drei Beiträge kritisch reflektieren und Implikationen für die Theorie, Praxis und weitere Forschung innerhalb des Projekts und darüber hinaus aufzeigen.

Beiträge des Symposiums

Augmented Reality und Elektromagnetismus: Welche virtuellen Darstellungen unterstützen das Lernen der Lorentzkraft am besten?

Bermann Steinmacher, Barbara Gränz, Andreas Lichtenberger, Andreas Vaterlaus
ETH Zürich

Theoretischer Hintergrund

Das Verständnis abstrakter wissenschaftlicher Konzepte ist ein wichtiger Bestandteil der physikalischen Ausbildung. Ein gegebenes Konzept kann auf verschiedenen Wegen zugänglich gemacht und betrachtet werden. Beispielsweise kann es in Form von Text beschrieben oder durch eine Grafik illustriert werden. Auch die reale Beobachtung eines Phänomens durch ein Experiment sowie die Verwendung mathematischer Gleichungen stellen Arten der Repräsentation dar. Werden verschiedene Repräsentationen verwendet, spricht man von multiplen externen Repräsentationen (MER). Die Nutzung von MER hat sich für den Erwerb von Konzeptwissen als hilfreich erwiesen, da unterschiedliche Repräsentationen verschiedene Informationen übermitteln können (Treagust & Fischer, 2017). Dabei ist es entscheidend, dass die verschiedenen Repräsentationen sinnvoll miteinander verknüpft werden, um ein tieferes Verständnis zu ermöglichen (Ainsworth, 2006). Mitunter stellt die Verwendung von MER jedoch eine kognitive Mehrbelastung für Lernende dar (beispielsweise Split-Attention-Effekt; Chandler & Sweller, 1992). Augmented Reality (AR) bietet den Vorteil, die verwendeten Repräsentationen unmittelbar in die Lernumgebung zu integrieren,

wodurch die von der Kognitiven Theorie des multimedialen Lernens (Mayer, 2009) vorgeschlagene räumliche und zeitliche Kontiguität von digitalen und realen Inhalten erreicht wird, die zu einer Reduktion der kognitiven Belastung führen kann (Thees et al., 2020).

Fragestellung

Die Verwendung von grafischen Repräsentationen in traditionellen multimedialen Lernumgebungen wurde bereits ausgiebig erforscht (Rau et al., 2015). Hingegen bleibt die sinnvolle Auswahl AR-basierter Repräsentationen weiterhin Gegenstand der Forschung, den wir mit unserer Studie adressieren. Der für diese Studie gewählte Anwendungsbereich ist der Elektromagnetismus mit Fokus auf der Bestimmung der Richtung der Lorentzkraft und der Überlagerung von Magnetfeldern. Unsere Forschungsfragen sind:

1. Welche virtuellen Repräsentationen eignen sich am besten für den Erwerb von Konzeptwissen bezüglich der Festlegung der Richtung der Lorentzkraft und der Feldüberlagerung.
2. Beeinflusst die Anzahl der dargestellten virtuellen Repräsentationen den Lernerfolg?

Methode

In der noch laufenden experimentellen Unterrichtsstudie mit einem Prä-Post-Test-Design nahmen bisher N=67 Schüler*innen (66% weiblich, Alter: M=17.14, SD=.82) der Jahrgangsstufe 11 an verschiedenen Schweizer Gymnasien teil. Im ersten Teil der Studie erhalten die Schüler*innen durch zwei Lernvideos das für die Studie nötige Vorwissen (Magnetfelder, Feldüberlagerung, Feldvektor- und Feldliniendarstellung, Lorentzkraft) und bearbeiten Fragebögen zur Repräsentationskompetenz von Vektorfeldern und zum räumlichen Vorstellungsvermögen. Im zweiten Teil wird zu Beginn ein Fragenbogen zum Konzeptverständnis (Lorentzkraft, Feldüberlagerung, Magnetismus) bearbeitet. Es folgt eine 60-minütige Experimentierphase, in der die Schüler*innen in Einzelarbeit fünf Experimente zur Lorentzkraft mit AR-Unterstützung durchführen. Die AR-Smartglasses können dabei virtuelle Darstellungen des Magnetfeldes (entweder Feldvektoren oder Feldlinien) sowie das Vektordreibein anzeigen. Die hieraus folgenden sieben kombinatorisch möglichen Anzeigemöglichkeiten stellen die unabhängigen Versuchsgruppen der Studie dar, denen die Schüler*innen zufällig zugeteilt werden. Abschließend bearbeiten die Schüler*innen Fragebögen zur Repräsentationskompetenz von Vektorfeldern, zum Konzeptverständnis sowie zur Systemnutzung und kognitiven Belastung.

Ergebnisse und ihre Bedeutung

Als Maß für das erlangte Konzeptwissen bezüglich der Festlegung der Richtung der Lorentzkraft und der Feldüberlagerung dienen zwei Subscores aus dem Posttest zum Konzeptverständnis. Um für diese beiden Subscores Gruppenunterschiede zu messen, wird eine ANCOVA durchgeführt, welche den Einfluss des Vorwissens (Repräsentationskompetenz und Konzeptwissen) eliminiert. Momentan werden hierzu noch keine Ergebnisse berichtet, da der bisher erreichte Stichprobenumfang noch nicht ausreichend ist. Der angestrebte Stichprobenumfang der Studie liegt bei N=300 und wird planmäßig bis Mitte 2025 erreicht. Dieser Wert orientiert sich an einer Stichprobenumfangsplanung mittels der Software G*Power für das Signifikanzniveau $\alpha=0.05$, die Teststärke $1-\beta=0.9$ und eine Effektstärke $f=0.25$.

Mit den Ergebnissen unserer Studie leisten wir einen wichtigen Beitrag für die physikalische Unterrichtsforschung. Die Verwendung moderner Technologien wie AR ermöglicht neue Zugänge zu physikalischen Themen, die im traditionellen Unterricht so nicht möglich sind. Sinnvolle virtuelle Darstellungen, die lernwirksam sind und kognitiv nicht überlasten, bilden die Grundlage für einen nachhaltigen Lernerfolg. Unsere Ergebnisse werden dazu beitragen, Kriterien für eine optimierte digitale Lernumgebung zu identifizieren.

Developing Conceptual Knowledge and Representational Competence in Learning with Augmented Reality: Disentangling the Roles of Prior Knowledge and Cognitive Abilities

Zoya Kozlova¹, Peter Edelsbrunner², Sarah Hofer¹, Ralph Schumacher², Elsbeth Stern²

¹Ludwig-Maximilians-Universität München, ²ETH Zürich

Theoretical Background

Previous research has highlighted the potential of augmented reality (AR) for the development of conceptual knowledge in various STEM fields (e.g., Akçayır et al., 2016; Tarng et al., 2022; Chao et al., 2016). Due to its affordances, such as visualizing invisible phenomena, AR can enable learners to develop a deeper understanding of abstract concepts by providing external visual representations (Wu et al., 2013).

In STEM topics, conceptual knowledge is closely linked to representational competence (e.g., understanding the principles of the Lorentz force with analyzing field lines or field vectors) and it is commonly assumed that conceptual knowledge and representational knowledge mutually benefit each other's development (Edelsbrunner et al., 2023; Rau, 2017). Yet, evidence on the mutual developmental effects between the two constructs is limited (Edelsbrunner et al., 2023), and current research indicates that relational reasoning (Alexander, 2017), working memory capacity (Berkowitz et al., 2022), and understanding of models (Gobert et al., 2011) are important confounding variables that affect knowledge acquisition in STEM.

In our study we examined the following research questions:

RQ1: To what extent do representational competence and conceptual knowledge at T1 (pre-learning) predict each other at T2 (post-learning) when controlling for prior knowledge in the respective variables in a cross-lagged design?

RQ2: Do any mutual developmental relations that we find hold when controlling for relational reasoning (analogy and anomaly facets), understanding of models, and working memory capacity?

Methods

N=126 German secondary school students (based on power analysis) were recruited to participate in the study. Students (Age: 15.67, SD=.75; 66 males, 53 females, 4 diverse, 5 unknown) individually performed experiments on electromagnetism in an AR environment for 60 minutes, where the real experimental setup was digitally superimposed with representations of vector fields, field lines, and the vector tripod.

To assess representational competence, we used the inventory for representational competence of fields (RCFI; Kuechemann et al., 2021; internal consistency: $\omega=0.83$), and adapted the instrument from Edelsbrunner et al. (2023) for conceptual knowledge ($\omega=0.82$). For the control variables, we utilized the analogy and anomaly scales from the Test of Relational Reasoning by Alexander et al. (2016), for understanding of models a questionnaire adapted from Treagust et al. (2002) and a Binding Task for working memory capacity (Wilhelm et al., 2013).

To examine the research questions, we employed a cross-lagged path model to analyze the predictive relationships without (RQ1) and with (RQ2) covariates.

Results

RQ1: The model indicated evidence for the cross-lagged path from representational competence at T1 predicting conceptual knowledge at T2 ($\beta = 0.28$, $p < .001$), and a positive yet non-significant cross-lagged path from conceptual knowledge at T1 to representational competence at T2 ($\beta = 0.08$, $p = .327$).

RQ2: When including the four covariates in the model, the predictive strength in the cross-lag of representational competence at T1 on conceptual knowledge at T2 diminished, yet only moderately ($\beta = 0.21$, $p = .008$).

Discussion

In accordance with a commonly stated theoretical assumption that has however seldom been tested empirically, our findings indicate predictive strength of individual differences in representational competence for gains in conceptual knowledge in learning a STEM topic with AR. Furthermore, this predictive strength mostly held up when controlling for covariates covering theoretically important individual differences in general cognitive abilities and understanding of models. There were no visible effects of prior conceptual knowledge on gains in representational competence, yet the positive estimate indicates need for replication. We conclude that representational competence might indeed affect conceptual knowledge acquisition in STEM learning with AR. We will discuss how these results inform STEM learning environments employing AR and whether these findings might hold in non-AR settings.

Selbsterklärungs-prompts zur Unterstützung von Schüler*innenexperimenten in Augmented Reality

Kristin Altmeyer¹, Max Warkentin², Christoph Hoyer², Stefan Küchemann², Jochen Kuhn², Roland Brünken¹, Sarah Malone¹

¹Universität des Saarlandes, ²Ludwig-Maximilians-Universität München

Theoretischer Hintergrund

Visuell-grafische Repräsentationen wissenschaftlicher Modelle können Lernende bei Schüler*innenexperimenten unterstützen, indem sie unsichtbare oder abstrakte Grundlagen von beobachtbaren physikalischen Phänomenen veranschaulichen. Werden Repräsentationen verschiedener Modelle kombiniert, entstehen Multiple Externale (Modell-)Repräsentationen (MER), die sich sinnvoll ergänzen und das konzeptuelle Verständnis vertiefen können (Opfermann et al., 2017). Damit MER den erwarteten Nutzen erzielen, müssen Lernende einzelne Repräsentationen mental miteinander verknüpfen und abgleichen (Berthold & Renkl, 2009). Diese Fähigkeit ist ein wesentlicher Aspekt von Repräsentationskompetenz (Kozma et al., 2000) und stellt für viele Lernende eine Herausforderung dar (Ainsworth et al., 2002). Im Sinne der Cognitive Load Theory (CLT; Sweller et al., 1998) kann Augmented Reality (AR) die mentalen Abgleichprozesse während des Experimentierens unterstützen, indem multiple virtuelle Modellrepräsentationen zeitlich und räumlich in die reale Experimentierumgebung integriert und somit Suchprozesse erleichtert werden. Dennoch zeigen erste Untersuchungen, dass die komplexen virtuell-real hybriden MER in AR-unterstützten Experimentierumgebungen hohe Anforderungen an Schüler*innen stellen und nicht automatisch zu besseren Lernergebnissen führen (Altmeyer et al., 2024).

Prompts, die Lernende dazu auffordern, Repräsentationen zu vergleichen und Selbsterklärungen zu deren konzeptioneller Verknüpfung abzugeben, haben sich als effektive instruktionale Unterstützung erwiesen, um Lernprozesse zu verbessern (Berthold et al., 2009). Die aktive Verarbeitung von multiplen Modellrepräsentationen fördert die Repräsentationskompetenz (Stieff et al., 2016) und unterstützt der CLT zufolge die Schemakonstruktion. Dies sollte sich in einer erhöhten lernförderlichen kognitiven Belastung (Germane Cognitive Load, GCL) widerspiegeln und damit auch den Erwerb von Konzeptwissen mit MER steigern (Rau et al., 2015).

Fragestellung

Ziel dieser Studie ist die Untersuchung der Wirkung von Integrationsprompts in einer AR-unterstützten Experimentierumgebung. Die aktivere Verarbeitung von MER durch Integrationsprompts sollte in mehr Blickwechseln zwischen den Repräsentationen und erhöhter Fixationsdauer für die einzelnen Repräsentationen resultieren. Weiterhin sollten sich Integrationsprompts positiv auf GCL, den Repräsentationskompetenzerwerb und Konzeptwissen auswirken.

Methode

N=78 Schüler*innen (79% weiblich, Alter: M=15.03, SD=.62) führten drei Experimente zum curricular-validen Thema Lorentzkraft im Physikunterricht der Sekundarstufe 1 durch. Mithilfe von AR-Smartglasses wurden ihnen virtuelle, dynamische Modellrepräsentationen von magnetischen Feldlinien, Feldvektoren und eines Vektordreibehns in die reale Experimentierumgebung eingeblendet. Alle drei simultan präsentierten Modellrepräsentationen passten sich in Echtzeit an, während die Schüler*innen die Ausrichtung eines Magneten und die Stromrichtung beim Experimentieren variierten.

In einem Between-Subjects-Design wurden die Schüler*innen zufällig zwei Bedingungen zugeteilt. Während der Interventionsgruppe auditive Selbsterklärungs-prompts zur Förderung mentaler integrativer Prozesse zwischen Repräsentationen präsentiert wurden, erhielten die Probanden der Kontrollgruppe Audioprompts, die sie lediglich aufforderten, sich das Experiment erneut anzuschauen. In der Experimentierphase wurden die Blickbewegungen aller Schüler*innen erfasst. Beide Gruppen bearbeiteten einen Cognitive Load Fragebogen nach dem Experimentieren sowie Prä- und Posttests zu Repräsentationskompetenz und Konzeptwissen.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse weisen darauf hin, dass die Schüler*innen sich unabhängig der Gruppe in Repräsentationskompetenz ($F(1,76)=5.62$, $p=.019$, $\eta^2=.008$) und Konzeptwissen ($F(1,76)=11.08$, $p=.001$, $\eta^2=.04$) verbessert haben. Allerdings unterstützen die Daten die Hypothesen nicht und zeigen keinen signifikanten Effekt der Bedingung auf GCL ($t(75)=-1.08$, $p=.142$, $d=.246$), den Zuwachs an Repräsentationskompetenz ($F(1,76)=2.85$, $p=.095$, $\eta^2=.004$) und Konzeptwissen ($F(1,76)=1.15$, $p=.287$, $\eta^2=.005$). Deskriptiv zeigte die Interventionsgruppe höhere GCL-Werte und Zugewinne in Repräsentationskompetenz. Die Auswertung der Blickbewegungen zur Analyse der Lernprozesse während des Experimentierens steht noch aus.

Die vorläufigen Ergebnisse legen die Vermutung nahe, dass die Integrationsprompts die Lernergebnisse nicht signifikant verbesserten. In der vorliegenden Studie deutet die Auswertung der Prätests auf sehr geringes Vorwissen der Proband*innen hin. Vorausgehende Forschung nimmt an, dass Lernende mit mittlerem Vorwissensniveau am besten von instruktionaler Integrationshilfe profitieren (Seufert & Brünken, 2004), während sie Novizen womöglich überfordert. Weitere Analysen mit größerer Stichprobe sollen den Einfluss des Vorwissens näher beleuchten. Außerdem werden Analysen der Blickdaten tiefere

Einblicke in den Einfluss von Integrationsprompts auf die Verarbeitung von MER in AR-unterstützten Experimentierumgebungen geben.

S-45: Reflexion im Fokus: Einblicke in den Reflexionsprozess von der Initiierung bis zur Förderung

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 148

Chair der Sitzung: Katrin Arendt

Chair der Sitzung: Anja Friedrich

Chair der Sitzung: Frank Fischer

Reflexion im Fokus: Einblicke in den Reflexionsprozess von der Initiierung bis zur Förderung

Chair(s): **Katrin Arendt** (Universität des Saarlandes), **Anja Friedrich** (Universität des Saarlandes)

Diskutant*in(nen): **Frank Fischer** (Ludwig-Maximilians-Universität München)

Im Rahmen der Professionalisierung von Lehrkräften hat die Reflexion von Unterricht in den vergangenen Jahren stark an Beachtung gewonnen (z.B. von Aufschnaiter et al., 2019). Die Verankerung von Reflexion in der Lehramtsausbildung wird verstärkt gefordert und auch verschiedene empirische Untersuchungen belegen die Annahme, dass Reflexion positive Auswirkungen auf die professionelle Kompetenz von Lehrkräften hat (z.B. Roters, 2019). Dennoch existieren noch wenig Befunde zum Reflexionsprozess und seiner Struktur (z.B. Aeppli & Lötscher, 2016; Göbel et al., 2022; Kager et al., 2022) sowie zu Möglichkeiten der systematischen Initiierung und Förderung von Reflexionsprozessen (Göbel et al., 2022; Kager et al., 2022).

In diesem Symposium werden vier Beiträge vorgestellt, die beginnend bei der Initiierung des Reflexionsprozesses, Bestandteile von Reflexionsprozessen genauer betrachten und Möglichkeiten zur Unterstützung dieser in den Blick nehmen. Die Fragestellung, die in diesem Symposium adressiert wird, lautet: Wie können reflexionsbezogene Prozesse konzeptualisiert, initiiert und gefördert werden?

Im ersten Beitrag wird die Initiierung von Reflexionsprozessen fokussiert. Dabei wird der Zusammenhang der wahrgenommenen Klassenführung im eigenen Unterricht mit der anlassbezogenen Motivationsstärke untersucht. Die Ergebnisse zeigen, dass die Motivation von Lehrkräften zur Reflexion der Unterrichtsstunde insbesondere nach Unterrichtsstunden erhöht war, die nicht optimal verliefen. Der zweite Beitrag setzt nach der Initiierung des Reflexionsprozesses an und versucht mithilfe eines skripttheoretischen Ansatzes den Reflexionsprozess zu konzeptualisieren. Hier werden in einem induktiven qualitativen Vorgehen mündliche Reflexionen hinsichtlich der Bestandteile eines möglichen Reflexionsskripts ausgewertet. Es zeigt sich, dass die Beschreibung und die Bewertung einer Unterrichtssituation sowie das Nennen von Handlungsalternativen als Skriptbestandteile (Scriptlets) identifiziert werden können. Auch im dritten Beitrag wird der Reflexionsprozess qualitativ untersucht: Hier werden epistemische Prozesse beim Reflektieren in den Blick genommen. Das Ziel dieses Beitrags besteht darin, diese Prozesse sowie die zur Reflexion herangezogenen Wissensarten zu identifizieren und zu untersuchen, welche Unterschiede in der Anwendung der epistemischen Prozesse vorliegen, wenn Fehlkonzepte Lehrender erkannt werden. Die Ergebnisse zeigen, dass Studierende verschiedene epistemische Prozesse nutzen, wie beispielsweise die Identifikation eigener Wissenslücken oder von Leerstellen in Unterrichtsplanungen. Der letzte Beitrag legt einen Fokus auf die Förderung von Reflexion durch computergestütztes elaboriertes Feedback. Es werden Effekte von computergestütztem Feedback auf die Bewertung des Feedbacks sowie hinsichtlich Qualitätsunterschieden schriftlicher Reflexionen und Veränderungen während einer Praxisphase betrachtet. Hier zeigt sich, dass individuelles, computergestütztes Feedback mit strukturierteren Reflexionen einhergeht, jedoch nicht zu einer signifikanten Erhöhung der Reflexionskompetenz beiträgt.

Nimmt man die Erkenntnisse der vorliegenden Beiträge zusammen, wird deutlich, dass zukünftige Forschung die spezifischen Bedingungen und Mechanismen, die Lehrkräfte zur Reflexion motivieren und die Analyse von Skripten für den Reflexionsprozess sowie epistemischer Prozesse bei der Reflexion weiter untersuchen sollte. Ein weiteres Forschungsdesiderat ist die Optimierung von computergestütztem Feedback, um nicht nur die Struktur, sondern auch die Qualität und Tiefe von Reflexion zu verbessern. Die Beiträge des Symposiums liefern hierfür erste wichtige Erkenntnisse und somit eine zentrale Grundlage; auch um die Ergebnisse in die Praxis der Lehrer*innenaus- und -weiterbildung zu transferieren.

Beiträge des Symposiums

„Darüber muss ich mal nachdenken.“ – Anlassbezogene Motivationsstärke als Teil des Reflexionsprozesses

Hendrik Lohse-Bossenz¹, Andrea Westphal¹, Gerlinde Lense²

¹Universität Greifswald, ²RPTU Kaiserslautern-Landau

Theoretischer Hintergrund

Verschiedene empirische Studien belegen die theoretische Annahme, dass die Reflexion pädagogischen Handelns zur professionellen Entwicklung von (angehenden) Lehrpersonen beiträgt (Aufschnaiter et al., 2019; Lense & Lohse-Bossenz, 2023; Weber et al., 2020). Während zu mentalen Prozessen beim Reflektieren bereits empirische Befunde vorliegen (z.B. Kleinknecht & Gröschner, 2016), ist bislang unklar, in welchen Situationen (angehende) Lehrkräfte Reflexionsprozesse initiieren. Ausgehend von Überlegungen von Leonhard und Abels (2017) definieren Lense und Lohse-Bossenz (2023) Reflexion als einen anlassbezogenen mentalen Prozess und konzeptualisieren damit ein situatives Merkmal des Reflexionsprozesses, das wir als anlassbezogene Motivationsstärke bezeichnen. Ziel des Beitrags ist es, empirisch zu untersuchen, wie die anlassbezogene Motivationsstärke mit der wahrgenommenen Klassenführung im eigenen Unterricht zusammenhängt. Es wird angenommen, dass negativ erlebte Unterrichtsereignisse die Motivationsstärke zum Reflektieren erhöhen.

Methode

Zur Beantwortung der Fragestellung schätzten 81 Grundschullehrkräfte (Alter: M = 41.63, SD = 10.70; Berufserfahrung: M = 14.31, SD = 10.39) an 10 aufeinander folgenden Schultagen klassenführungsrelevante Aspekte (Umgang mit Unterrichtsstörungen, Monitoring und effektive Zeitzuwendung; Thiel et al., 2013) im Anschluss an eine Unterrichtsstunde ein. Die anlassbezogene Motivationsstärke wurde mit einem Item erfasst („Ich möchte über diese Stunde vertieft nachdenken, um daraus etwas für mein eigenes Handeln als Lehrkraft zu lernen.“; Antwortskala von 1 = trifft überhaupt nicht zu bis 7 = trifft völlig zu). Bei der Analyse der Daten wurde die Mehrebenenstruktur der Daten mit Angaben auf Tages- (Level 1) und Lehrkräftebene (Level 2) berücksichtigt.

Ergebnisse

Die Intraklassenkorrelation der Analysevariablen lag zwischen 0.22 und 0.27 für klassenführungsrelevante Aspekte und 0.46 für die anlassbezogene Motivationsstärke für das Reflektieren. Mehrebenenanalysen bestätigten, dass Lehrkräfte nach

Unterrichtsstunden, in denen sie Unterrichtsstörungen nicht angemessen adressieren konnten, eine höhere anlassbezogene Motivationsstärke für das Reflektieren berichteten. Lehrpersonen verspürten also ein verstärktes Bedürfnis, über die jeweilige Unterrichtsstunde vertieft nachzudenken.

Diskussion

Zusammenfassend lässt sich festhalten, dass die anlassbezogene Motivationsstärke zum Reflektieren nach negativen Ereignissen im eigenen Unterricht besonders hoch ausgeprägt ist. Mit Blick auf den Einsatz von Unterrichtsvideos als Reflexionsanlässe stellt sich die Frage, ob Videos mit nicht gelungenen Unterrichtssituationen Studierende stärker zum Reflektieren motivieren als Videos von gelungenem Unterricht.

Reflexion aufschlüsseln: Empirische Prüfung eines skripttheoretischen Ansatzes zur Konzeptualisierung des Reflexionsprozesses

Katrin Arendt, Lisa Stark, Anja Friedrich, Roland Brünken, Robin Stark

Universität des Saarlandes

Die Relevanz von Reflexion für die Professionalisierung von Lehrkräften ist in den letzten Jahren immer mehr in den Fokus der Forschung gerückt (z.B. Weber et al., 2023). Während einige aktuelle Arbeiten (z.B. Arendt et al., 2023; Lenske & Lohse-Bossenz, 2023) dem mangelnden gemeinsamen Begriffsverständnis von Reflexion im pädagogischen Kontext entgegengewirkt haben, fehlt es weiterhin an einer empirisch geprüften Konzeptualisierung des Reflexionsprozesses. Reflexion wird definiert als ein kognitiver Prozess der strukturierten und (selbst-)kritischen Auseinandersetzung mit (eigenen und fremden) pädagogischen Erfahrungen unter Bezugnahme auf das eigene pädagogische, fachdidaktische und fachliche Wissen sowie eigene Fähigkeiten und Überzeugungen, um die professionelle Praxis und sich selbst weiterzuentwickeln und zu verbessern (Arendt et al., 2023). Zu einzelnen Teilprozessen des Reflektierens pädagogischer Erfahrungen liegen zwar bereits Ergebnisse vor (z.B. Kleinknecht & Gröschner, 2016), die genaue Struktur des gesamten Reflexionsprozesses muss jedoch noch weiter aufgeklärt werden (Göbel & Neuber, 2022). Komplexe kognitive Prozesse können in Form von Handlungsschemata oder Skripts konzeptualisiert werden; einzelne Teilprozesse der Skripts können als Scriptlets bezeichnet werden (Fischer et al., 2013). Da auch Reflexion als komplexer kognitiver Prozess mit einzelnen Teilprozessen verstanden werden kann, wurde in der vorliegenden Arbeit der skripttheoretische Ansatz herangezogen, um den Reflexionsprozess empirisch aufzuschlüsseln und davon ausgehend theoretisch weiter zu fundieren.

In dieser Studie wurde untersucht, inwiefern sich bei der Reflexion pädagogischer Erfahrungen der Reflexionsprozess als Skript mit einzelnen Scriptlets konzeptualisieren lässt. Weiter wird betrachtet, inwiefern sich in den einzelnen Reflexionsscriptlets ein Selbstbezug identifizieren lässt und inwieweit verschiedene Begründungen herangezogen werden.

In einem explorativen, qualitativen Untersuchungsdesign wurden N = 29 Lehrkräfte (Alter: M = 45.60, SD = 12.82; Geschlecht: 43,33 % weiblich) im Rahmen online durchgeführter Interviews befragt. Dabei wurden zwei authentische Unterrichtsvideos in ausbalancierter Reihenfolge gezeigt, verbunden mit der Aufforderung, diese im Anschluss mündlich zu reflektieren. Die Reflexionen der Lehrkräfte wurden aufgezeichnet und zur Datenauswertung transkribiert. Die Auswertung erfolgte inhaltsanalytisch durch Kategorienbildung (Mayring, 2022) mit MAXQDA 2022. Das Kodierschema wurde zunächst anhand einer ersten Auswahl von Transkripten induktiv aus dem Material entwickelt. Für den finalen Kodierdurchgang wurde bei der Benennung der Kategorien auch auf Literatur zurückgegriffen.

Erste Ergebnisse zeigen, dass als gemeinsame Scriptlets des Reflexionsskripts das Beschreiben („Hier kamen die Instruktionen. Dann konnten die SuS die Arbeitsaufträge anschauen.“) und das Bewerten der betrachteten Unterrichtssituation („Also ich fand gut, dass sie wie im ersten Video auch die Arbeitsaufträge visualisiert hat.“) sowie das Nennen von Handlungsalternativen („Also ich wäre kleinschrittiger vorgegangen.“) identifiziert werden konnten. Die Scriptlets können dabei sowohl begründet als auch unbegründet vorliegen. Als Grundlage der angeführten Begründungen in den Scriptlets konnten eigene pädagogische Erfahrungen der Versuchspersonen ausgemacht werden („Einfach, weil ich weiß, dass nicht jeder zuhört, wenn die Seitenzahl fällt.“). Besonders zeigte sich dies in Verbindung mit dem Scriptlet der Bewertung der betrachteten Unterrichtssituation. Das Herstellen eines Selbstbezugs zeigte sich vor allem im Scriptlet zum Nennen von Handlungsalternativen („Hier hätte ich doch genauere Angaben dazu gemacht [...]“).

Die Ergebnisse legen nahe, dass die Struktur des Reflexionsprozesses mithilfe des skripttheoretischen Ansatzes weiter in seine Bestandteile (Scriptlets) aufgeschlüsselt werden kann. Reflexionsskripts bieten somit eine Grundlage, um Reflexionen im Kontext der Lehrer*innenaus- und -weiterbildung zu analysieren. Es bedarf jedoch noch weiterer Konkretisierung beispielsweise mit Blick auf Variablen wie die Berufserfahrung oder reflexionsbezogene Dispositionen. Für künftige Forschung stellt sich die Frage, inwiefern der entwickelte Ansatz auch zur Unterstützung von Reflexionsprozessen und zur Bewertung der Qualität von Reflexionen genutzt werden kann.

Epistemische Prozesse bei der wissensbasierten Reflexion von Lehramtsstudierenden

Anna-Lena Molitor, Judith Schellenbach-Zell, Michael Rochnia

Bergische Universität Wuppertal

Reflexion stellt eine wesentliche Komponente in der Expertiseentwicklung von angehenden Lehrkräften dar (Gruber, 2021). Vor dem Anspruch einer evidenzorientierten Praxis (Stark, 2017) kann Reflexion angehende Lehrkräfte unterstützen, wissenschaftliches Wissen mit konkreten praktischen Situationen – wie der Unterrichtsplanung – zu verbinden (Cramer et al., 2019). Vor diesem Hintergrund ist auch die Identifikation von Fehlkonzepthen relevant, da diese im Sinne anekdotischer Evidenz Einzug in die pädagogische Praxis erhalten (Hendriks et al. 2021). Falls angehende Lehrkräfte diese Fehlkonzepthe nicht identifizieren, besteht die Gefahr, dass nicht veridikales Wissen in der Unterrichtsplanung genutzt wird. Reflexion kommt daher neben der Verbindung von Theorie und Praxis die Funktion zu, fragwürdige Wissensbestände zu erkennen und kritisch mit ihnen umzugehen.

Wir verstehen Reflexion prozesshaft und gegliedert in verschiedene Phasen (EDAMA, Aeppli & Lötscher, 2016). Angelehnt an Csanadi und Kollegen (2021) betrachten wir innerhalb der Reflexion eine Inhalts- und eine Prozessebene. Auf Inhaltsebene ist bislang wenig untersucht, auf welche Wissensbestände Lehramtsstudierende zurückgreifen und ob Fehlkonzepthe erkannt werden (Patry, 2014). Die Prozessebene kann angelehnt an Forschung zu epistemischer Kognition mithilfe der Betrachtung epistemischer Prozesse (AIR-Modell, Chinn et al., 2014) operationalisiert werden. Diese informationsverarbeitenden Prozesse werden genutzt, um ein bestimmtes epistemisches Ziel wie z.B. das Verstehen einer Situation zu erreichen. Welche dieser Prozesse von angehenden Lehrkräften genutzt werden, wurde bereits untersucht (Schellenbach-Zell et al., 2023). Im Fokus standen dabei Reflexionsprodukte (Molitor et al., 2022; Schellenbach-Zell et al. 2023) oder Think-Aloud-Daten zu Reflexionen

ohne Literaturbezug (Csanadi et al., 2021; Barnes et al., 2020). Jedoch wurden bislang keine Prozesse untersucht, die von der Erstellung eines Reflexionsproduktes bis hin zur Literatursuche reichen. Offen bleibt auch, ob bei der Reflexion Fehlkonzepte identifiziert werden.

Vor diesem Hintergrund gehen wir den folgenden Fragestellungen nach:

1. Welche epistemischen Prozesse nutzen Lehramtsstudierende bei der Reflexion einer Unterrichtsplanung und auf welches Wissen greifen sie zurück?
2. Gibt es Variationen in der Anwendung dieser Prozesse abhängig davon, ob Fehlkonzepte erkannt wurden oder nicht?

In einem qualitativen Zugriff werden Masterstudierende des Grundschullehramts mit dem Fach Sachunterricht untersucht. Zur Erfassung von Prozessen bei der Auseinandersetzung mit pädagogischen Situationen erscheint das Verfahren des Think-Alouds (Ericsson, 2006) zielführend. Entsprechend fertigten acht Studierende Think-Aloud-Protokolle bei der Reflexion einer praktischen Problemlage an: Sie gaben einer fiktiven Lehrkraft im Vorbereitungsdienst Feedback zu ihrer Unterrichtsplanung und nutzten dabei auch – wie bereits im zugehörigen Seminar eingeübt – wissenschaftliche Quellen. Die Unterrichtsplanung enthielt verschiedene Fehlkonzepte aus den drei Professionswissensbereichen (Shulman, 1986). Die Think-Aloud-Protokolle der Studierenden wurden zum gesamten Reflexionsprozess (vom Lesen der Situation, über die Quellensuche, bis hin zum Verschriftlichen der Reflexion) angefertigt. Diese wurden transkribiert und im Anschluss nach Sinneinheiten segmentiert (O'Connor & Joffe, 2020). Das Kodierschema basiert auf den Vorarbeiten von Schellenbach-Zell und Kolleg*innen (2023) und wurde auf der Grundlage des Kompetenzmodells von Blömeke und Shavelson (2015) sowie am Material induktiv erweitert. Die Kodierung erfolgte zunächst konsensuell durch die drei Autor*innen an einem Teil des Materials. Bei ausreichender Interrater-Reliabilität nach eigenständiger Kodierung wurde der Rest des Materials getrennt kodiert, uneindeutige Fälle wurden besprochen und konsensuell gelöst (Hopf & Schmidt, 1993).

Erste Erkenntnisse zeigen, dass Studierende eine Vielzahl an epistemischen Prozessen nutzen, die sich den von Blömeke und Kollegen (2015) formulierten Skills Perception, Interpretation und Decision Making zuordnen lassen. Ein wesentlicher Prozess im Bereich Perception scheint beispielsweise die ‚Identifikation von Leerstellen‘ zu sein, bei dem Studierende wahrnehmen, dass die gegebene Planung nicht vollständig durchdacht ist oder Fehlkonzepte vorliegen. Im Bereich Interpretation zeigen sich metakognitive Prozesse, wie die ‚Identifikation eigener Wissenslücken‘.

Ziel ist es, alle Reflexionen vollständig im Hinblick auf die Prozesse zu kodieren und diese in Bezug zu den erkannten Fehlkonzepten zu setzen. Die Ergebnisse werden bezüglich ihrer Relevanz zur Erfassung und Förderung von Reflexionsfähigkeiten in der Lehrkräftebildung diskutiert.

KI-gestützte Lehrkräftebildung in den Naturwissenschaften? – Strukturiertes und inhaltliches Feedback zu schriftlichen Reflexionen im Praxissemester Physik

Lukas Mientus¹, Anna Nowak¹, Peter Wulff², Andreas Borowski¹

¹Universität Potsdam, ²Pädagogische Hochschule Heidelberg

Theoretischer Hintergrund

Insbesondere in der ersten Phase der Lehrkräftebildung bieten Praxisphasen angehenden Lehrkräften Lerngelegenheiten zur Entwicklung professioneller Handlungskompetenz, unter anderem durch Möglichkeiten der Reflexion des eigenen Handelns (Brouwer & Korthagen, 2005; Gröschner et al., 2013; Schubarth et al., 2009). Systematisches Feedback kann für die Entwicklung von Reflexion eine entscheidende Rolle spielen (Hattie & Timperley, 2007). Feedback zu konkreten inhaltlichen Aspekten einer Reflexion wird hierbei als besonders wertvoll wahrgenommen (Moreno, 2004; Chase & Houmanfar, 2009). Strukturelles Feedback kann helfen, die Aufgabenanforderungen besser zu verstehen (Kost, 2019) und als Gerüst insbesondere Noviz:innen in ihrer Reflexionskompetenzentwicklung unterstützen (Lai & Calandra, 2010). Da sowohl strukturelles als auch inhaltliches Feedback ressourcenintensiv ist, erfolgt oft nur holistisch Feedback (Poldner et al., 2014). Feedback sollte individualisiert und zeitnah nach der Handlung erfolgen (Russell & Korthagen, 2013). Um dies zu erreichen können computergestützte Tools und KI-Methoden Dozierende in der Begleitung von Studierenden unterstützen. Mithilfe von KI können beispielsweise schriftliche Reflexionen analysiert werden und sofortiges Feedback gegeben werden. Inwieweit diese Mensch-Maschine-Kollaboration nicht nur zeitnahes Feedback skalierbar ermöglicht, sondern auch wirksam ist, wird in der vorliegenden Studie untersucht.

Fragestellung

Diese Interventionsstudie untersucht die Effekte der Mensch-Maschine-Kollaboration im Kontrollgruppendesign im Praxissemester Physik und verfolgt die folgenden Forschungsfragen:

1. Wie bewerten angehende Physiklehrkräfte das strukturelle und das inhaltliche Feedback?
2. Welche Unterschiede gibt es in den schriftlichen Reflexionen zwischen den Gruppen?
3. Welche Veränderungen sind bei den Reflexionen während der Praxisphase feststellbar?

Methode

Das Feedbackkonzept ReFeed (Mientus et al., 2021), bei welchem Studierende im Praxissemester dreimalig jeweils ein strukturelles und ein inhaltliches Feedback zu einer Reflexion erhalten, stellt mit N = 34 Teilnehmenden mit insgesamt 102 schriftlichen Selbstreflexionen die Interventionsgruppe dar. Die Datengrundlage der Kontrollgruppe sind schriftliche Reflexionen von N = 22 angehenden Physiklehrkräften mit insgesamt 132 Reflexionstexten.

Als Kontrolluntersuchung wurde zunächst ein halbstrukturiertes Interview mit dem Dozierenden der Kontrollgruppe durchgeführt. Das Interviewtranskript wurde qualitativ analysiert und von zwei unabhängigen Ratern kodiert, um die Seminarstrukturen der Kontroll- und Interventionsgruppe zu vergleichen. Zur Beantwortung von FF 1 wurde in der Interventionsgruppe nach jedem der beiden Feedbacks eine Akzeptanzbefragung mit Likert-Skala zur Selbsteinschätzung in den Dimensionen Wirksamkeit, Nützlichkeit, Individualität, subjektive Richtigkeit und Wertigkeit durchgeführt (vgl. Wulff et al. 2021 und Mientus et al., 2023a). Unter Verwendung der von Wulff et al. (2020) und Mientus et al. (2023b) entwickelten computerbasierten Algorithmen wurden alle schriftlichen Reflexionen beider Kohorten hinsichtlich Textlänge, deskriptiver und diskursiver Elemente sowie deren Positionen im Text analysiert, um FF 2 und 3 zu untersuchen. Diese Algorithmen klassifizierten jede Satzzuordnung nach dem Reflexionsmodell von Nowak et al. (2019) und standardisierten die Verteilung der Diskurselemente. Für jede Reflexion wurden Textlänge, absolute und relative Anteile der Diskurselemente sowie der Grad der Strukturierung analysiert, wobei t-Tests und ANOVA zur Vergleichsanalyse beider Kohorten verwendet wurden.

Ergebnisse und ihre Bedeutung

Es wurden keine relevanten Unterschiede in den Rahmenbedingungen des Praxissemesters oder den zugehörigen Seminaren festgestellt. Die Kontrollgruppe schrieb mehr Texte, was mehr Übungsmöglichkeiten bot, aber auch herausfordernder war. Die Interventionsgruppe erhielt explizites, individuelles Feedback und nutzte das Rahmenmodell von Nowak et al. (2019) zur Strukturierung. In der Akzeptanzbefragung wurde strukturelles Feedback als weniger effektiv, nützlich und persönlich bewertet. Signifikante Unterschiede zeigten, dass die Interventionsgruppe weniger deskriptive und mehr diskursive Texte sowie eine bessere Strukturierung aufwies ($dif = 5\%$, $W \leq 4675$, $p < .001$), was gegenüber der Kontrollgruppe auf eine höhere Qualität der Reflexion hindeutet. Die Mensch-Maschine-Kollaboration führte jedoch nicht zu einer signifikant anderen Entwicklung der Reflexionskompetenz ($F = 2.41$, $p > .440$). Automatisiertes Feedback kann zwar zeitnah erfolgen, bietet zum aktuellen Zeitpunkt jedoch nur strukturelles Feedback. Der Vortrag präsentiert die Studie und zeigt Perspektiven für eine bedarfsgerechte Unterstützung individueller Reflexionsprozesse im Praxissemester Physik.

S-46: KI zur Förderung und Analyse von Lehr- und Lernprozessen

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: M 003

Chair der Sitzung: Ulrike Cress

Chair der Sitzung: Marcus Specht

KI zur Förderung und Analyse von Lehr- und Lernprozessen

Chair(s): **Ulrike Cress** (Leibniz-Institut für Wissensmedien, Deutschland)

Diskutant*in(nen): **Marcus Specht** (TU Delft)

Kein gesellschaftlicher Bereich wurde durch das Erscheinen von ChatGPT so stark aufgerüttelt wie das Bildungssystem (Fütterer et al., 2023). Damit KI die Bildung jedoch tatsächlich verbessert, muss sie spezifische Lehr- und Lernprozesse unterstützen, die den Wissens- und Kompetenzerwerb fördern und auf die jeweiligen Domänen angepasst sind. Ein effektiver Einsatz von KI im Bildungsbereich, sei sie generativ oder allgemein, muss daher auf intendierte Lernprozesse ausgerichtet sein und diese durch entsprechende Lehrprozesse unterstützen (Cress & Kimmeler, 2023; Kasneci, 2023; SWK, 2024).

Das Symposium „KI zur Förderung und Analyse von Lehr- und Lernprozessen“ stellt diese Prozessebene und deren empirische Operationalisierung in den Mittelpunkt. In drei Beiträgen aus dem Leibniz-Forschungsverbund LERN werden interdisziplinäre Projekte vorgestellt, die die Funktion von KI in Lern- und Wissensprozessen auf verschiedenen Granularitätsebenen beleuchten.

Der erste Beitrag beschäftigt sich mit Lernverläufen (learning progressions) in MINT, wo durch Learning Analytics individuelle Lernverläufe modelliert und auf ihre Produktivität hin untersucht werden. Der zweite Beitrag stellt die Frage, wie Lehrprozesse variabel gestaltet werden können, um individuelle Lernprozesse adaptiv zu ermöglichen – und weist dies in einer randomisierten Feldstudie in einem sozialwissenschaftlichen Schulfach nach. Der dritte Beitrag schließlich untersucht konstruktive Prozesse, insbesondere das argumentative Schreiben und die Analyse von Revisionsprozessen nach automatisiertem Feedback.

Alle drei Beiträge folgen einer einheitlichen Struktur: Sie (1) beschreiben die relevanten Prozesse, die es zu unterstützen gilt, (2) reflektieren, welche Daten zur Prozessbeschreibung generiert werden und wie diese in die KI einfließen, (3) zeigen, welche KI-Werkzeuge entwickelt wurden, um die im Fokus stehenden Prozesse zu unterstützen, und (4) analysieren die im Lernprozess resultierenden und ihn beschreibenden Daten hinsichtlich der Effektivität der KI-unterstützenden Werkzeuge.

Diese Projekte und Beiträge stammen aus interdisziplinären Projekten des Leibniz-Forschungsnetzwerks LERN, wo Fachdidaktik, Instruktionspsychologie und Informatik zusammenarbeiten. Auf Grundlage von Befunden aus der Lernforschung werden KI tools entwickelt, in formellen Bildungsprozessen eingesetzt und ihre Wirkung analysiert. Die Beiträge decken verschiedene Domänen wie Spracherwerb, Naturwissenschaften, Programmieren und Wirtschaftswissenschaften ab.

Nach der Einleitung sind für die Präsentationen der drei Beiträge jeweils 30 Minuten Vortrag mit direkt anschließendem Zeitslot für Nachfragen und Diskussion vorgesehen. Ein Diskutant leitet die Schlussdiskussion mit dem Publikum durch eine kommentierende Betrachtung der Beiträge und den Bezug zur eigenen Forschung ein.

Beiträge des Symposiums

Ursachen unproduktiver Lernverläufe im naturwissenschaftlichen Unterricht identifizieren und adressieren

Marcus Kubsch¹, Knut Neumann², Ulrike Cress³, Hendrik Drachler⁴

¹Freie Universität Berlin, ²IPN, ³IWM, ⁴DIPF

Naturwissenschaftliche Bildung ist zentral für das Verständnis und die Adressierung von Herausforderungen wie der Energiewende. Entsprechende Bildungsziele werden jedoch häufig nicht erreicht (Lewalter et al., 2023). KI-basierten, adaptiven Unterstützungsmaßnahmen wird großes Potenzial zugesprochen, diesem Defizit entgegenzuwirken (Department of Education, 2023). Damit adaptive Unterstützungsmaßnahmen wirken können, sind Erkenntnisse über die Determinanten von (un)produktiven Lernverläufen nötig. Die Kompetenzentwicklungsforschung (z.B. Neumann et al., 2013) hat begonnen, domänenspezifische Modelle fachlicher Lernverläufe zu entwickeln. Methodische Herausforderungen – vor allem in Bezug auf eher niedrige temporale Auflösung und hauptsächlich querschnittliche Daten – limitieren dabei die Möglichkeiten, belastbar konkrete Interventionen abzuleiten, um unproduktive in produktive Lernverläufe zu überführen. Dies ist jedoch eine notwendige Voraussetzung, um das Potenzial adaptiver Unterstützungsmaßnahmen auszuschöpfen.

Das zunehmende Aufkommen von digitalen Technologien im naturwissenschaftlichen Unterricht (z.B. digitale Messwerterfassung) erlaubt im Zusammenspiel mit Methoden der Learning Analytics, die genannten methodischen Herausforderungen zu adressieren (Kubsch et al., 2022): Wenn SchülerInnen mit digitalen Technologien interagieren und Lernprodukte erzeugen, fallen Daten über die Lernprodukte sowie Lernprozesse (z.B. Schreibprozess beim naturwissenschaftlichen Bewerten) an. Diese lassen sich auswerten, um zeitlich hoch aufgelöste Lernverläufe zu modellieren und Rückschlüsse über die Determinanten von deren Produktivität zu erlangen.

Fragestellung

Im Beitrag wird am Beispiel einer 10-wöchigen Unterrichtseinheit der Frage nachgegangen, wie Lernverläufe auf Basis der Interaktion von SchülerInnen mit digitalen Technologien modelliert und Determinanten für deren Produktivität identifiziert werden können.

Methode

Die Unterrichtseinheit, auf die im Beitrag fokussiert wird, adressiert das Themenfeld Energie anhand der Leitfrage „Was ist die beste Energieversorgung für unsere Schule“ entlang von vier Sinnabschnitten in der Klassenstufe 10. Die Unterrichtseinheit wurde in einem auf evidence-centered Design (Pellegrino et al., 2015) beruhenden Prozess (Kubsch et al., 2022) entwickelt, welche von vornherein die Nutzung von digitalen Technologien zur Modellierung von Lernverläufen in den Blick nimmt. Die Unterrichtseinheit wurde mit insgesamt ca. 500 SchülerInnen an Sekundarschulen durchgeführt. Neben den während der Interaktion mit den digitalen Technologien anfallenden Daten wurden auch Daten über den Lernerfolg mit Hilfe eines Energiekompetenztest in einem Prä-Post-Design erfasst.

Zur Modellierung der Lernverläufe wurden anhand des Lernerfolgs Gruppen gebildet, wobei ein Lernzuwachs als unproduktiv gewertet wurde, wenn er kleiner als der mittlere Lernzuwachs der Gesamtstichprobe war. Mit Hilfe von Random-Forest-Modellen und unter Variation der genutzten Daten (Nutzung von zunehmend mehr Daten entlang der Sinnabschnitte der Unterrichtseinheit) wurde die Gruppenzugehörigkeit vorhergesagt. Mit Hilfe von Methoden der explainable AI wurde untersucht, welche der zur

Vorhersage der Produktivität der Lerntrajektorien genutzten Daten besonders prädiktiv sind. In diesen so identifizierten Daten wurde dann mit Hilfe von Clustering-Verfahren nach Mustern gesucht, um besser zu verstehen, warum diese Daten hochprädiktiv sind und um Determinanten für die Produktivität von Lernverläufen zu identifizieren.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass bereits mit den Daten des ersten Sinnabschnitts die Produktivität der Lernverläufe mit hoher Präzision ($F1 \sim 0,85$) vorhergesagt werden kann. Wenn zunehmend Daten von weiteren Sinnabschnitten hinzugenommen werden, steigert sich die Präzision kontinuierlich weiter bis zu $F1 \sim 0,94$. Als hochprädiktiv erwiesen sich dabei überwiegend Daten aus Aufgaben, welche die Verknüpfung zentraler Wissens Elemente erfordern. Dies stützt Befunde, dass die Ausbildung eines Wissens organisiert um zentrale Konzepte der Domäne, entscheidend für das Lernen über Physik ist (Bransford et al., 2000; Kubsch, 2024). Die weiterführende Analyse der Antworten auf entsprechende Aufgaben stützt und ergänzt diese Ergebnisse: Neben der Verknüpfung zentraler Wissens Elemente unterscheiden sich Antworten von SchülerInnen mit produktiven Lernverläufen gegenüber Antworten von SchülerInnen mit unproduktiven Lernverläufen durch wohlstrukturierte Begründungen.

Diese Ergebnisse demonstrieren, dass eine frühzeitige Vorhersage der Produktivität von Lernverläufen mit Learning Analytics Verfahren möglich ist und liefern konkrete Hinweise, wie adaptive Unterstützungsmaßnahmen aussehen können. Die Wirksamkeit solcher adaptiver Unterstützungsmaßnahmen hinsichtlich der Überführung von unproduktiven in produktive Lernverläufe wird in zukünftigen Forschungsarbeiten untersucht werden.

Adaptive Lernverläufe benötigen differenzierte Lehrprozesse

Detmar Meurers¹, Kordula De Kuthy¹, Kai Neubauer², Ulf Brefeld³, Felix Dietrich⁴, Michael Koch⁴, Dirk Loerwald⁴

¹IWM, ²Leuphana Universität Lüneburg, ³Universität Lüneburg, ⁴Universität Oldenburg

Theoretischer Hintergrund

Lernende unterscheiden sich substanziell, unter anderem in ihrem fachlichen Wissen und Kompetenzen, in ihren bildungssprachlichen Fähigkeiten und kognitiven Eigenschaften (vgl. u.a. Höhmann 2009; Trautmann & Wischer 2011). Diese individuellen Unterschiede sind von großer Relevanz für das Bildungssystem, wie Baumert schon 2002 hervorhob: "In der Verbesserung des Umgangs mit Differenz liegt vermutlich die eigentliche Herausforderung der Modernisierung des Systems." Für die Schulpraxis erwächst daraus die zentrale Frage, wie angesichts der vielschichtigen Heterogenität individuelle Lernprozesse adaptiv erfolgreich gestaltet werden können. So wird zum Beispiel in einer Studie an 18 Sekundarschulen (Kunze & Solzbacher 2008) solche individuelle Förderung als wichtigstes anzustrebendes Ziel von praktisch allen Lehrern genannt, aber 90% halten dies in der Praxis für unmöglich umsetzbar und fühlen sich dem nicht gewachsen. Die Entwicklung intelligenter Tutorsysteme verspricht hier einerseits eine individuelle Lernunterstützung durch makroadaptive Aufgabensequenzierung und mikroadaptive Unterstützung der Lernenden in ihrer entwicklungsproximalen Zone (Vygotsky 1986), allerdings wird die Heterogenität der Lernenden dabei praktisch durchgängig reduziert auf Unterschiede im Fachwissen, obwohl aktuelle Taxonomien des Adaptiven Lernens wie Plass & Pawar (2020) prinzipiell eine mehrdimensionale Adaptivität vorsehen..

Fragestellung

Da ein reduziertes Verständnis von Adaptivität auf die eindimensionale Passung von Fachkompetenz der Lernenden zu Aufgabenschwierigkeit der mehrdimensionalen Heterogenität in der Bildungsrealität nicht gerecht wird (Park et al, 2019), wird in dem hier vorgestellten Ansatz die Komplexität der Lernaufgaben für ein sozialwissenschaftliches Schulfach mehrdimensional parametrisiert und in einer randomisierten Feldstudie im authentischen Schulkontext nachgewiesen, dass die variablen Lehrprozesse vielfältige adaptive Lernpfade ermöglichen, die sich positiv auf Lernfortschritt, Motivation und Selbstwirksamkeit auswirken. Die hier vorgestellte Forschung untersucht konkret, wie sich mehrdimensional parametrisierte Lehrmaterialien und -aufgaben erstellen und KI-basiert adaptiv zuweisen lassen, die sowohl fachliche, als auch sprachliche und kognitive Dimensionen berücksichtigen. Auf Basis der Meta-Informationen zu den Parametern der Aufgaben und der Lernprozessdaten aus dem von uns entwickelten KI-basierten Lernsystem für die adaptive Zuweisung untersuchen und vergleichen wir in einer randomisierten Feldstudie die individuell adaptiv zugewiesenen Lernpfade mit den synchronen Lernschritten einer Kontrollgruppe mit festgelegter Lernsequenz.

Methode

In einer interdisziplinären Kooperation mit Partnern aus der Fachdidaktik Ökonomische Bildung, dem Maschinellen Lernen und der Schnittstelle von KI und Bildungsforschung haben wir ausgehend von einer Modellierung des in den Lehrplänen der Länder abgedeckten Inhalte der Themenbereiche Preisbildung, Marktformen und Wettbewerb systematisch Aufgaben entwickelt und hinsichtlich fachlicher, sprachlicher und kognitiver Faktoren in ihrer Komplexität parametrisiert. Nach einer Pilotierungsphase im realen Schulkontext wurde der Ansatz in drei Bundesländern in einer randomisierten Feldstudie in 5 Wochen regulärem Unterricht im Fach Wirtschaft mit Sek.I-Lernenden ($n = 150$) in einer randomisierten Feldstudie untersucht. Innerhalb der Klassen erfolgte eine Randomisierung hinsichtlich der Aufgabenzuweisung, die entweder adaptiv nach dem individuellen Lernverlauf oder entsprechend einer traditionellen statischen Aufgabensequenz von der Lernplattform realisiert wurde.

Ergebnisse und ihre Bedeutung

Die Daten aus der Feldstudie zeigen auf der Grundlagenseite, dass die mehrdimensional parametrisierten Aufgaben empirisch wie intendiert auch zu unterschiedlichen Lösungswahrscheinlichkeiten führen und der entwickelte KI-Ansatz nach wenigen Interaktionen in der Lage ist, auf dieser Basis Aufgaben adaptiv zuzuweisen. Anhand der detailliert aufgezeichneten Lernprozessdaten sehen wir, dass die Lernenden in der adaptiven Gruppe sehr unterschiedlich lange Lernpfade vom System zugewiesen bekommen und im Verhältnis zu den Lernenden mit einer traditionellen Standardsequenz weiter im Lernstoff kommen. Auf Basis der erhobenen Fragebogendaten sind sie dabei motivierter, fühlen sie sich dabei signifikant weniger unter Druck und sind mit der eigenen Leistung zufriedener.

Während die meisten Adaptivitäts-Ansätze für den MINT-Kontext ausgearbeitet wurden und sich auf eine eindimensionale Kompetenz-Adaptivität konzentrieren, zeigt die hier vorgestellte Studie, dass ein mehrdimensional parametrisiertes intelligentes Tutorsystem in einem gesellschaftswissenschaftlichen Fach wie der ökonomischen Bildung im realen Schulkontext realisiert und evaluiert werden kann und sich die durch KI ermöglichte Funktionalität positiv auf Motivation, Selbstwirksamkeit und Lernfortschritt auswirkt.

Fairness von KI-gestützter Analyse und leistungsadaptive Förderung beim Schreiben argumentativer Texte

Theoretischer Hintergrund

Die Fähigkeit, sich eine fundierte Meinung zu gesellschaftlichen Themen wie dem Klimawandel zu bilden und diese schriftlich zu kommunizieren ist eine Schlüsselkompetenz für gesellschaftliche und berufliche Teilhabe in globalisierten Informations- und Netzwerkgesellschaften. Beim Schreiben argumentativer Texte kann eine formative Beurteilung mit leistungsadaptivem Feedback eine wirksame Lernintervention darstellen (Graham et al., 2024). Allerdings ist die Erstellung dieser Lerngelegenheit komplex. Bei der Analyse der Leistung müssen verschiedene inhaltliche und strukturelle Merkmale der Argumentation berücksichtigt werden (Ferretti & Graham, 2019). Für die Wirksamkeit der Förderung, z.B. durch Feedback, spielen sowohl Eigenschaften des Feedbacks selbst (Jansen et al., 2024), insbesondere der Inhalt und die Passung zum Text, als auch Eigenschaften der Schüler:innen, wie der allgemeine Umgang mit Feedback und die wahrgenommene Nützlichkeit des spezifischen Feedbacks, eine wichtige Rolle (Lipnevich & Smith, 2022). Zur Unterstützung von Lehrkräften in der Erstellung von Feedback wurden Algorithmen der künstlichen Intelligenz untersucht (Stahl et al., 2024) und die Wirksamkeit von automatisiertem Feedback gezeigt (z.B. Fleckenstein et al., 2023). Jedoch zeigten die Studien auch große Heterogenität und ließen die für die Bildungsgerechtigkeit zentrale Frage unbeantwortet, ob die gefundene Heterogenität systematisch Gruppen von Schüler:innen bevorteilt oder sie benachteiligt (Lee et al. 2024).

Fragestellung

Der Beitrag adressiert die übergreifende Frage: Wie fair funktioniert die automatisierte Beurteilung und leistungsadaptive Förderung von schriftlichen Argumentationskompetenzen? Dazu wird die Genauigkeit der Beurteilung und die Effekte von Feedback in einer diversen Gruppe von Schüler:innen in Abhängigkeit zu ihren individuellen Eigenschaften untersucht.

Methode

In einer experimentellen Studie mit Schülerinnen und Schülern der Klassenstufen 9 bis 13 (N = 1836; 42,7 % weiblich, Durchschnittsalter 16,36 Jahre [SD = 1,47]; HISECD: M = 5.13 [SD = 1.25]) im Between-Subject-Design schrieben die Teilnehmenden innerhalb von 60 min einen argumentativen Text über den Klimawandel und überarbeiteten ihre Argumentation mithilfe von automatisiertem, computergestütztem Feedback. Die randomisierte Variation zwischen den Teilnehmenden waren die Inhalte des Feedbacks, welche entweder strukturelle Eigenschaften der Argumentation, die Qualität der Argumente oder die Inhalte fördern sollten. Auf der Ebene der Teilnehmenden wurden Hintergrundvariablen (Alter, Geschlecht, kognitive Fähigkeiten, Schulform), die Rezeptivität der Schüler:innen für Feedback (Bahr et al., 2024) und die wahrgenommene Nützlichkeit des spezifischen Feedbacks (Srijbos et al., 2021) gemessen.

Nach der Studie beurteilten zwei geschulte Personen die Texte und Überarbeitungen der Schüler:innen anhand von inhaltlichen und strukturellen Argumentationseigenschaften (siehe Schaller et al., 2024 für das Beurteilungsraster). Aufbauend auf den von Menschen vergebenen Urteilen wurden KI-Modelle genutzt (shallow learning, deep learning, large language model; vgl. Lohmann et al., 2024), um die menschlichen Urteile vorherzusagen. Analysiert wurde, ob die Urteile von verschiedenen KI-Modellen fair für alle inkludierten Gruppen von Schülerinnen und Schülern waren und welche Teilnehmenden besonders von dem vergebenen Feedback profitierten.

Ergebnisse und ihre Bedeutung

Die Urteile der trainierten Personen konnten von den untersuchten KI-Modellen, einschließlich ChatGPT, mit einer ähnlichen Übereinstimmung zwischen menschlichen und KI-generierten Urteilen vorhergesagt werden wie die Übereinstimmung zwischen den beiden Personen. Es konnte keine Evidenz für Urteilsbias der KI-Modelle in Abhängigkeit vom Hintergrund der Schüler:innen gefunden werden, was für eine faire Beurteilung spricht. Von dem automatisierten Feedback profitierten besonders Schüler:innen mit einer hohen Rezeptivität für Feedback und jene, die das Feedback als besonders nützlich empfanden, da sie dadurch motivierter waren, die Aufgabe zu bearbeiten. Zusammenfassend zeigen die Ergebnisse, dass KI-Algorithmen genau genug sind, um Urteile auf der Ebene pädagogischer Ideen in Feedback umzusetzen, und legen nahe, dass die Wirksamkeit von Feedback eher vom beabsichtigten Inhalt des Feedbacks und der Wahrnehmung der Empfänger abhängt als von der Wahl des KI-Algorithmus. Der Beitrag unterstreicht damit, dass zur Weiterentwicklung von KI basierten Fördermöglichkeiten ein Zusammenspiel der Forschung in der Fachdidaktik, Instruktionspsychologie und Informatik notwendig ist.

S-47: Potenziale besser erkennen und fördern: Neue Ansätze in der Begabungsforschung

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 145

Chair der Sitzung: Anne-Kathrin Stiller

Chair der Sitzung: Tanja Gabriele Baudson

Potenziale besser erkennen und fördern: Neue Ansätze in der Begabungsforschung

Chair(s): **Anne-Kathrin Stiller** (Karg-Stiftung, Deutschland)

Diskutant*in(nen): **Tanja Gabriele Baudson** (Vincenz Palotti University)

Unterschiede in der intellektuellen Begabung stellen eine Dimension von Schüler:innenheterogenität dar, auf die Lehrkräfte eingehen müssen. Auch hochbegabte Schüler:innen, die etwa 2 % aller Schüler:innen ausmachen, haben ein Recht auf eine ihren Begabungen entsprechende Förderung. Bleibt diese aus, können sie ihre Potenziale weniger entfalten, Leistungen ggf. nicht zeigen (Underachievement) und in Persönlichkeitsentwicklung und Wohlbefinden beeinträchtigt werden (z.B. Preckel & Vock, 2020). Eine begabungsgerechte Förderung bildet damit eine Facette von Bildungs- und Chancengerechtigkeit. Aber auch innerhalb des Kontextes der Begabtenförderung sind bestimmte Gruppen von Schüler:innen traditionell unterrepräsentiert und sind bei der Erkennung und Förderung von Begabungen systematisch benachteiligt. Dazu zählen unter anderem Mädchen, Schüler:innen aus Einwandererfamilien sowie Schüler:innen, die mit herausfordernden sozioökonomischen Bedingungen konfrontiert sind (z. B. Worrell & Dixon, 2021). Erkenntnisse darüber, wie Lehrkräfte Begabungen erkennen und Schüler:innen differenzierter fördern und in ihrer Potenzialentfaltung unterstützen können – idealerweise unabhängig von Geschlecht und Herkunft –, leisten einen Beitrag zu mehr Bildungsgerechtigkeit.

Beitrag 1 adressiert Sprachlernbegabungen in der fremdsprachlichen Bildung und richtet den Blick damit auf eine domänenspezifische Begabung. Die Autor:innen stellen die Entwicklung und Validierung eines neuen digitalen Diagnoseinstruments zur Erfassung von Sprachlernbegabungen (6. Klasse) von Schüler:innen mit Deutsch als Erstsprache oder als einer unter mehreren Muttersprachen (mehrsprachige Schüler:innen) vor. Das mehrdimensionale Instrument ist für den Einsatz durch Lehrkräfte geeignet und kann zur Unterstützung potenzialorientierter Selektions- und Förderentscheidungen genutzt werden, indem bislang unentdeckte Sprachlernpotenziale sichtbar gemacht werden.

In Beitrag 2 wird ein Lerntagebuch zur Förderung des selbstregulierten Lernens bei hoch- und normalbegabten Schüler:innen in der 6. und 8. Klasse eingesetzt. Die Daten aus dem längsschnittlichen Design geben Aufschluss über Unterschiede in Aspekten des selbstregulierten Lernens im Vergleich zu nicht hochbegabten Schüler:innen sowie über die Lernprozesse und -erfahrungen der Schüler:innen. Durch die digitale Form, sowie den geringen zeitlichen Aufwand soll ein möglichst niedrigschwelliges Angebot geschaffen werden, das es Lehrkräften erlaubt, hoch- und normalbegabte Schüler:innen in ihren Lernprozessen individuell zu fördern.

Beitrag 3 beleuchtet die Bedeutung von kognitiven und metakognitiven Fähigkeiten sowie von schulischen Rahmenbedingungen für das Lernen mit Texten bei hochbegabten Schüler:innen der 6.-8. Klasse in Modell- und Regelklassen. Die Ergebnisse weisen darauf hin, dass hochbegabte Schüler:innen auch bei guten Lesefähigkeiten ihren Verstehensprozess beim Lernen aus Texten nicht immer ausreichend überwachen. Durch eine Förderung des Einsatzes metakognitiver Strategien und eine verstärkte Ermöglichung von selbstreguliertem Lernen auch in Regelklassen könnten Lehrkräfte das Lernen mit Texten unterstützen.

Alle vorgestellten Studien nehmen Diagnose- bzw. Förderansätze in den Blick, die niedrigschwellig eingesetzt werden können, um alle Schüler:innen unabhängig von Geschlecht und sozioökonomischer oder kultureller Herkunft in ihrer Potenzialentfaltung zu unterstützen. Die aus der Psychologie, der Erziehungswissenschaft sowie der Fachdidaktik stammenden Beiträge werden abschließend gemeinsam diskutiert.

Beiträge des Symposiums

Sprachlernbegabung in der fremdsprachlichen Bildung – Entwicklung und Validierung eines Diagnoseinstruments

Alicia Battenfeld¹, Till Utesch², Ulrike Gut¹, Frauke Matz¹

¹Universität Münster, ²Universität Lübeck

Sprachlernbegabung ist als Bereich der Potentialentwicklung von Schüler*innen im deutschsprachigen Raum bisher kaum erforscht (Gut & Matz, 2022), obwohl er eine wichtige Variable darstellt, die die Unterschiedlichkeit von Lernenden in Bezug auf Lernerfolg/-prozesse im Fremdsprachenunterricht zuverlässig erklären kann (z.B. Ehrmann & Oxford, 1995; Sparks et al., 2009). Die bisher bekannteste und viel diskutierte Konzeptualisierung von Sprachlernbegabung von John B. Carroll unterscheidet zwischen phonetic coding ability, grammatical sensitivity, inductive language learning ability und rote learning ability (Carroll, 1981). Auf dieser Konzeptualisierung von Sprachlernbegabung basiert auch das derzeit immer noch international als Standardtest zur Sprachlernbegabung genutzte Diagnostikinstrument MLAT (Modern Language Aptitude Test) von Carroll und Sapon (1959). Weitere bekannte Tests sind z.B. LLAMA (Meara, 2005), der weitgehend an den MLAT angelehnt ist, sowie der PLAB (Pimsleur, 1966), CANAL-FT (Grigorenko et al., 2000) und der Hi-LAB (Linck et al. 2013), die sich vom MLAT und der damit einhergehenden Konzeptualisierung von Sprachlernbegabung unterscheiden. Einzig der LLAMA, dessen Validierung noch aussteht, kann allerdings in kurzer Zeit durchgeführt werden und ist sprachunabhängig. Alle anderen Tests setzen Englisch als Testsprache voraus, sind komplex in der Durchführung und stehen nicht frei zur Verfügung (Pan & Marsden, 2024).

Während Tests zur Sprachlernbegabung international eine lange Tradition haben (z.B. Berthele & Udry, 2021; Carroll & Sapon, 1959; Kiss, 2009; Linck et al., 2013), gibt es bisher kein zugängliches Diagnoseinstrument für die Erfassung von Sprachlernbegabung von Personen mit Deutsch als (einer ihrer) Muttersprache(n) (Gut & Matz, 2022). Durch diese Forschungslücke ist die empirische Diagnostik von Sprachlernbegabung mithilfe eines validierten Tests im deutschsprachigen Raum derzeit nicht möglich (Schlak, 2008).

Im Rahmen unseres interdisziplinären Forschungsprojekts widmen wir uns diesem Desiderat durch die Entwicklung eines entsprechenden Tests. Dieser soll zentrale Dimensionen von Sprachlernbegabung messen, die Testgütekriterien erfüllen und durch Methoden der probabilistischen Testtheorie validiert werden. Aktuell umfasst das sich in der Entwicklung und Validierung befindende Diagnoseinstrument sechs Komponenten bestehend aus insgesamt 95 Items: zwei Subtests zu sprachanalytischen Fähigkeiten (5 Items und 14 Items), je ein Subtest zum lexikalischen (15 Items) und zum phonetischen Gedächtnis (16 Items) sowie zwei Subtests zu implizitem Sprachlernen (30 Items und 15 Items). Das sich in der Entwicklung befindende Diagnoseinstrument, das digital und frei verfügbar sein wird, ist zunächst auf Lernende der 6. Klasse ausgerichtet und so

konzipiert, dass es gesamte Klassen ohne vorheriges Screening innerhalb einer Schulstunde von circa 45 Minuten durchlaufen können.

Die systematische Identifikation einer Sprachbegabung durch ein standardisiertes Diagnoseinstrument ist besonders für die gezielte Förderung bzw. Forderung von einzelnen Schüler*innen wünschenswert (Prat et al., 2020). So soll der Test beispielsweise Stärken und Schwächen von Schüler*innen in den getesteten Dimensionen aufzeigen können. Darüber hinaus

soll das Diagnoseinstrument Lehrkräfte bei fundierten und potenzialorientierten Selektionsentscheidungen unterstützen, sodass kognitive Sprachbegabung neben weiteren relevanten Variablen (wie z.B. Motivation, Interesse) etwa bei der Entscheidung für bilinguale Beschulung oder der Wahl einer dritten Fremdsprache berücksichtigt werden kann. Eine standardisierte Potenzialdiagnostik stellt im Kontext von Bildungsgerechtigkeit einen wichtigen Ansatzpunkt dar, indem Begabungen von Schüler*innen sichtbar werden, die derzeit unentdeckt bleiben. Dies betrifft insbesondere systematisch benachteiligte mehrsprachige Schüler*innen, die jedoch über eine ausgeprägte Begabung im Sprachenlernen verfügen können (Ma, Yao & Zhang, 2018). Durch eine testbasierte Diagnostik sollen solche Potenziale aufgedeckt werden, was zu einer verstärkten Anerkennung individueller Mehrsprachigkeit als Ressource für den Fremdsprachenunterricht führen kann und darin zur Chancengleichheit beiträgt.

Unser Beitrag ordnet das in der Entwicklung befindliche Diagnoseinstrument zunächst in die verschiedenen Konzeptualisierungen von Sprachlernbegabung ein und präsentiert die bisherigen Entwicklungsschritte des Tests. Er stellt erste Ergebnisse aus der bisherigen Validierungsstudie mit einer Stichprobe von N = 600 Schüler*innen der Jahrgangsstufe 6 aus Gymnasien und Gesamtschulen in NRW vor und schließt mit einem Ausblick auf den weiteren Projektverlauf und stellt offene Fragen zur Diskussion.

Projekt DigiLIFE: Förderung des selbstregulierten Lernens von Hochbegabten durch digitale Lerntagebücher

Franziska Zmatlik, Henrik Bellhäuser, Maximilian Resch, Marlene Heilmann
Johannes Gutenberg-Universität Mainz

Lerntagebücher sind eine Form der Intervention zur Förderung des selbstregulierten Lernens (SRL). Indem sie zu täglichen Reflexionen von Lernverhaltensaspekten (z.B. Zielsetzung, Planung, Motivation, Prokrastination oder Stress) anregen, werden reflektive Prozesse verfestigt sowie Lerngewohnheiten verbessert (Dignath et al., 2023). Dieser positive Effekt kann gesteigert werden, indem automatisch generiertes, individuelles Feedback in das digitale Lerntagebuch eingebaut wird (Bellhäuser et al., 2023; Theobald & Bellhäuser, 2022).

In dem Projekt „DigiLIFE“, gefördert durch die Karg-Stiftung, wird diese Form der Intervention erstmals spezifisch für hochbegabte Kinder und Jugendliche eingesetzt. Ziel ist es besser, effektiver, zeitsparender und zufriedener zu lernen.

SRL bei hochbegabten und hochleistenden Schüler*innen (Stöger & Zeidner, 2019) wird aufgrund der allgemeinen Meinung, dass diese Gruppe naturgemäß ohne spezifische Lernstrategien herausragende Leistungen erzielt, vernachlässigt (Stöger, Balestrini & Ziegler, 2021). Die Expertiseforschung betont jedoch insbesondere die Bedeutung von intensivem Üben (deliberate practice) von etwa 10.000 Stunden, um Exzellenzstatus in einer spezifischen Domäne zu erreichen (Preckel & Vock 2020). Interventionsstudien zeigen, dass besseres SRL sich sowohl positiv auf kognitive Prozesse als auch auf das Wohlbefinden hochbegabter Schüler*innen auswirkt (Efklides, 2019).

In der vorliegenden Studie wird eine erste Version des Lerntagebuchs in einem längsschnittlichen Design pilotiert, ohne schon den Interventionseffekt mit einer Kontrollgruppe zu untersuchen. Stattdessen besteht die Zielsetzung darin, ein solides Messinstrument zu entwickeln, das mit altersgemäßen Formulierungen die für das Projekt relevanten Konstrukte reliabel und valide erhebt und dabei in zumutbar kurzer Zeit auszufüllen ist. Zudem werden intraindividuelle Zusammenhänge zwischen unterschiedlichen SRL-Komponenten untersucht, die Veränderungen des Lernverhaltens von einem auf den nächsten Lerntag beschreiben (Bellhäuser et al., 2021).

Der erste Teilerhebung unserer Untersuchung fand an einem Gymnasium mit Begabtenklassen und regulären Klassen zu Beginn des Schuljahres statt. Hierfür wurde in einer vierwöchigen Erhebungsphase jeden Tag ein kurzes Lerntagebuch, sowie zu Beginn und Ende ein ausführlicher Fragebogen von den Schüler*innen ausgefüllt. Es nahmen Schüler*innen der 6. (ngifted=22; nnormal=116) und 8. Klasse (ngifted= 20; nnormal= 113) teil.

Erste Analysen zeigen signifikante Unterschiede im Bereich der Motivation, Prokrastination und des Stresserleben zwischen den Begabten und Regelklassen. In weiteren Analysen untersuchen wir mit Hilfe von Multilevel-Regressionen, inwiefern erfolgreiches Lernen an Tag x (z.B. Konzentration, Vermeidung von Prokrastination, Zielerreichung, Zufriedenheit) zu verbesserten Lernstrategien am Tag x+1 führt (z.B. Zielsetzung, Planung, Motivation, Selbstwirksamkeit). Qualitative Auswertungen der offenen Antworten im Tagebuch werden zudem Aufschluss geben über die unterschiedlichen Lernerfahrungen von hochbegabten und normalbegabten Schüler*innen im schulischen Alltag.

Im zweiten Teil der Untersuchung sollen diese Befunde auch im Verlauf des Schuljahres untersucht werden. Auf Grundlage von Feedback aus der ersten Erhebung, wurden Veränderungen im Design des Tagebuches vorgenommen. So erproben wir erstmals eine Variante des Tagebuchs, bei der jeden Tag unterschiedliche Items präsentiert werden, um Langeweile seitens der Schüler*innen zu vermeiden und den Interventionscharakter des Tagebuches zu verstärken.

Lerntagebücher können durch ihren niedrigschwelligen und zeitsparenden Charakter ressourcensparend für verschiedene Gruppen von Schüler*innen eingesetzt werden. Die Möglichkeiten, das eigene Lernen und Lernmuster durch Reflektion zu verstehen, trägt zur Autonomie, Potenzialentfaltung und Chancengleichheit der Kinder und Jugendlichen bei. Dies steht im Einklang mit den Erkenntnissen von Tibken et al. (2022), die die Bedeutung metakognitiver Kompetenzen für die schulische Entwicklung und die Vermeidung von Underachievement bei Hochbegabten hervorheben.

Lernen mit Texten im Schulalltag: Effekte der Beschulungsform bei hochbegabten Jugendlichen

Catharina Tibken, Tobias Richter
Julius-Maximilians-Universität Würzburg

Das Lernen aus Texten ist eine anspruchsvolle Tätigkeit, bei der Leser(innen) eine kohärente mentale Repräsentation des im Text beschriebenen Lerninhalts konstruieren müssen (Zwaan & Radvansky, 1998). Automatisierte Prozesse können ins Bewusstsein der Leserinnen und Leser gelangen, wenn wie in Lernsituationen der Text komplex und die Ansprüche an das eigene Textverständnis hoch sind (van den Broek & Helder, 2017). Schüler(innen) müssen ihren Verstehensprozess dann

metakognitiv planen, überwachen und regulieren. Studien zeigen, dass eine gute metakognitive Überwachung bei hochbegabten Schüler(inne)n der Entstehung und Verfestigung von Underachievement vorbeugen kann (Tibken et al., 2022).

Unklar ist bisher, von welchen individuellen Fähigkeiten und schulischen Rahmenbedingungen der Einsatz metakognitiver Kompetenzen im Schulalltag abhängt und wie sich diese von Voraussetzungen für allgemeine Lesefähigkeiten unterscheiden. Daher haben wir den Effekt von metakognitiven Kompetenzen beim Textverstehen, Intelligenz und Beschulungsform auf die allgemeine Lesefähigkeit und die Anwendung metakognitiver Strategien beim Lernen aus Texten im Schulalltag untersucht.

Die Untersuchung ist Teil eines größeren Längsschnittprojekts, in dessen Rahmen 86 (hoch)begabte Gymnasiast(inn)en der Jahrgangsstufen 6 bis 8 ($IQ \geq 120$) untersucht wurden. Von diesen Schüler(inne)n ($MIQ = 129,05$, $SD = 6,62$) besuchten 69 eine reguläre Gymnasialklasse und 17 eine Modellklasse für Hochbegabte. Die besuchte Klassenstufe und die Intelligenz unterschieden sich nicht signifikant zwischen den beiden Beschulungsformen. Zu Beginn der Studie bearbeiteten die Jugendlichen einen Intelligenztest (PSB-R 6-13; Horn et al., 2004), eine Inkonsistenzaufgabe zur Erfassung der metakognitiven Überwachung beim Lernen aus Texten (Tibken et al., 2022), einen Test zur Erfassung der allgemeinen Lesefähigkeiten (LGVT 5-12+; Schneider et al., 2017) und einen Fragebogen zur gewohnheitsmäßigen Anwendung metakognitiver Strategien beim Lernen aus Texten im Schulalltag. In diesem Fragebogen gaben die Schüler(innen) beispielsweise an, wie häufig sie beim Lernen für die Schule konkrete Strategien (z.B. Erstellen von Mindmaps zu Texten) nutzten, ihren Lese- und Lernprozess im Vorfeld planen, sich bei Verständnisschwierigkeiten Unterstützung suchten oder versuchten, vermeintliche Widersprüche zwischen Texten oder mit dem eigenen Vorwissen aktiv aufzulösen.

Die Ergebnisse der Analysen zeigen, dass die allgemeinen Lesefähigkeiten und die Anwendung metakognitiver Strategien beim Lernen aus Texten im Schulalltag nicht signifikant korrelierten ($r = -.115$, $p = .349$). Die Leistungen in den beiden Bereichen ließen sich zudem in Teilen durch unterschiedliche zugrundeliegende Fähigkeiten erklären. Prozedurale metakognitive Kompetenzen hatten lediglich einen Effekt auf die Anwendung metakognitiver Strategien ($\beta = .441$, $p < .001$), nicht auf die allgemeinen Lesefähigkeiten ($\beta = .038$, $p = .380$). Die Intelligenz hatte einen Effekt auf die Lesefähigkeiten ($\beta = .203$, $p = .045$), nicht aber auf die Anwendung metakognitiver Strategien ($\beta = -.023$, $p = .416$). Die Beschulungsform (0 = gymnasiale Regelklasse, 1 = Modellklasse für Hochbegabte) hatte sowohl einen positiven Effekt auf die Lesefähigkeiten ($\beta = .217$, $p = .034$) als auch auf die Anwendung metakognitiver Strategien beim Lernen aus Texten im Schulalltag ($\beta = .265$, $p = .006$).

Die Ergebnisse deuten darauf hin, dass allgemeine Lesefähigkeiten und der Umgang mit Lerntexten unterschiedliche Kompetenzbereiche darstellen: Auch hochbegabte Schüler(innen), die über gute Lesefähigkeiten verfügen, scheinen nicht immer aus eigenem Antrieb metakognitive Strategien in Lernsituationen anzuwenden. Eine Förderung metakognitiver Kompetenzen könnte daher die Potenzialentfaltung begabter Schüler(innen) unterstützen. Auf den Umgang mit Texten im Alltag hat neben metakognitiven Kompetenzen auch die Unterrichtung in Modellklassen für Hochbegabte einen positiven Einfluss. Teil des Konzepts von Modellklassen ist es, den Schüler(inne)n im Vergleich zu Regelklassen mehr Freiraum für projektbasiertes und selbstreguliertes Lernen zu gewähren (Stumpf & Schneider, 2008). Dies könnte das Einüben metakognitiver Strategien unterstützen. Im Sinne der Chancengerechtigkeit wäre es wünschenswert, leistungsstarken Schüler(inne)n auch in Regelklassen mehr selbstreguliertes Lernen zu ermöglichen.

S-48: Ein guter Start ins Studium? Eingangsvoraussetzungen und Erwartungen der Studierenden als Gelingensbedingungen eines Hochschulstudiums

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 142

Chair der Sitzung: Svenja Kaiser

Chair der Sitzung: Katharina Vogel

Chair der Sitzung: Leif Döring

Chair der Sitzung: Stefan Münzer

Chair der Sitzung: Stefan Janke

Ein guter Start ins Studium? Eingangsvoraussetzungen und Erwartungen der Studierenden als Gelingensbedingungen eines Hochschulstudiums

Chair(s): **Svenja Kaiser** (Pädagogische Hochschule Heidelberg, Deutschland), **Markus Vogel** (Pädagogische Hochschule Heidelberg, Deutschland), **Leif Döring** (Universität Mannheim, Deutschland), **Stefan Münzer** (Universität Mannheim, Deutschland)

Diskutant*in(nen): **Stefan Janke** (Universität Mannheim, Deutschland)

Der Fachkräftemangel in Deutschland im akademischen Bereich sorgt für ein wachsendes gesellschaftliches Interesse daran, dass Hochschulen ihre Studierenden nicht nur hochwertig qualifizieren, sondern dass auch möglichst viele Studierende ihr Hochschulstudium erfolgreich abschließen (Wissenschaftsrat, 2015). Zuletzt lag die Studienabbruchsquote im Bachelor an Universitäten bei 35%, in den Geistes- und Naturwissenschaften sogar bei knapp 50% (Heublein 2022).

Für den Studienerfolg sind nicht nur Studienbedingungen, Studier- und Lernverhalten sowie Kontextbedingungen relevant, sondern auch die Eingangsvoraussetzungen der Studierenden (Blüthmann et al., 2008). Der Einfluss verschiedener Eingangsvoraussetzungen auf Studienabbruchstentionen und Studienabbrüche konnte ebenfalls nachgewiesen werden (Pelz et al., 2020).

Bei der Studienwahl erhöhen extrinsische Studienwahlmotive die Wahrscheinlichkeit für Studienabbrüche (Heublein, 2010). Für den Studienerfolg ist stattdessen eine gute Passung von persönlichen Merkmalen mit dem gewählten Studienfach entscheidend (Le et al., 2014). Auch eine schlechte Passung der Erwartungen von Studierenden kann im Studium zu höheren Studienabbruchstentionen führen (Eccles, 1993).

Diese Gelingensbedingungen für eine erfolgreiche Studieneingangsphase stehen im Mittelpunkt des Symposiums. Die Beiträge stellen aktuelle Studien aus den Bereichen Psychologie und Mathematik sowie fächerübergreifend in den Lehramtsstudiengängen von Grundschule und Gymnasium vor.

Der erste Beitrag setzt bereits vor dem Studium an und untersucht für den Studiengang Psychologie sowohl die subjektive Erfolgs- und Passungseinschätzung von Studieninteressierten als auch deren objektive Passung in verschiedenen Faktoren. Es zeigten sich Auswirkungen der Passung auf Studienerfolg und Studienabbruchstentionen, wobei die objektive Studienfach-Passung eine präzisere Vorhersage des Studienerfolgs ermöglichte als die subjektive Einschätzung der Studieninteressierten.

Der zweite Beitrag identifiziert mithilfe von Längsschnittdaten mögliche Prädiktoren von Studienabbrüchen in der Mathematik. Dazu wurde zunächst untersucht, mit welchen verschiedenen Eingangsvoraussetzungen Mathematikstudierende ihr Studium beginnen. Neben der Abiturnote erwiesen sich auch die Arbeitsvermeidungsmotivation und das Beweisverständnis als signifikante Prädiktoren für einen Studienabbruch im ersten Studienjahr.

Im dritten Beitrag wird die Statistikangst als eine der Eingangsvoraussetzungen von Lehramtsstudierenden näher untersucht. Es konnten Gruppenunterschiede zwischen Studierenden des gymnasialen Lehramts mit verschiedenen Fächerkombinationen identifiziert werden, die tatsächlichen statistischen Leistungen dieser Fächergruppen unterschieden sich jedoch nicht signifikant. Abschließend werden daher Überlegungen zum förderlichen Umgang mit Statistikangst in der Studieneingangsphase angestellt.

Der letzte Beitrag thematisiert Studienabbruchstentionen von Grundschullehramtsstudierenden im Kontext schulischer Praxisphasen. Es wurden der Einfluss (un-)passender Erwartungen und unterrichtsbezogener Selbstwirksamkeitsüberzeugungen untersucht. Eine geringe Passung der Erwartungen führte zu einer höheren Studienabbruchstention, während die Selbstwirksamkeitsüberzeugungen sich auf die Passung der Erwartungen und somit indirekt auf die Studienabbruchstentionen auswirkten.

Im Symposium werden diese Befunde interdisziplinär diskutiert und mögliche Implikationen für die Praxis abgeleitet, um somit zu einer erfolgreichen Hochschulbildung beizutragen.

Beiträge des Symposiums

Werde ich in diesem Studienfach erfolgreich sein? Objektive fachspezifische Passung hinsichtlich Interessen, Fähigkeiten und Erwartungen sagt Studienerfolg über subjektive Einschätzungen der Studieninteressierten vorher

Belinda Merkle, Oliver Dickhäuser
Universität Mannheim, Deutschland

Theoretischer Hintergrund und Fragestellung

Die Wahl eines passenden Studienfachs ist eine zentrale Bedingung für Studienerfolg und stellt gleichzeitig für viele Studieninteressierte eine große Herausforderung dar. Bislang gibt es jedoch wenig Forschung zu Faktoren, die vor Studienbeginn gemessen wurden, um späteren Studienerfolg vorherzusagen.

Auf der Argumentation von Author et al. (2024) aufbauend, welche auf der Affective Forecasting Literatur beruht (Wilson & Gilbert, 2003), könnten subjektive Einschätzungen von Studieninteressierten zu ihrer Passung und ihrem zukünftigen Erfolg in einem spezifischen Studienfach ein solcher Faktor sein. Allerdings können solche subjektiven Einschätzungen durch inakurate Theorien darüber, welche Faktoren zu Erfolg führen und durch Fehlvorstellungen von der zukünftigen Situation (bspw. Studienfach) verzerrt sein. Objektive studienfachspezifische Passung hingegen beschreibt beispielsweise das Interesse von Studieninteressierten an validen Inhalten des jeweiligen Studienfachs beziehungsweise deren Fähigkeiten hinsichtlich valider Anforderungen (z.B. auf Basis von Experteneinschätzungen). Diese objektiven Passungsindikatoren beruhen somit auf einem

empirisch nachgewiesenen Faktor für späteren Studienerfolg, nämlich der Passung (Person-Umwelt-Passungs-Theorie, e.g., Le et al., 2014) und messen diesen Faktor objektiv auf der Basis von validen Inhalten und Anforderungen des jeweiligen Studienfachs. Dadurch sollten objektive Faktoren beispielsweise inakkurate Theorien und Fehlvorstellungen reduzieren und somit Verzerrungen in subjektiven Einschätzungen potenziell reduzieren.

Eine aktuelle Studie zeigte, dass eine höhere objektive Interessen-Studienfach-Passung Motivation und Wohlbefinden im ersten Semester über die subjektive Einschätzung von Studieninteressierten hinaus vorhersagte (Author et al., 2024). Während die prädiktive Validität von studienfachspezifischen Zulassungstests (objektive Fähigkeits-Studienfach-Passung) für späteren Studienerfolg von vielen Studien gezeigt wird (z. B. Julian, 2005; Niessen et al., 2016), gibt es keine Belege für die Robustheit von objektiver Fähigkeits-Studienfach-Passung über subjektive Einschätzungen der Studieninteressierten hinaus. Weiterhin wurde objektive Erwartungs-Studienfach-Passung (Diskrepanz zwischen Studieninteressierten-Erwartungen hinsichtlich der Studienfachinhalte und tatsächlichen Studienfach-Inhalten basierend auf Expertenschätzungen) mit Studienzufriedenheit in Verbindung gebracht (Hasenberg & Schmidt-Atzert, 2013). Allerdings wurde die Passung nicht vor Studienbeginn gemessen und die Valenz (z. B. ob Studieninteressierte an einem Inhalt interessiert sind oder nicht) der Passung vernachlässigt. Diese Valenz der objektiven Erwartungs-Studienfach-Passung bestimmt, ob Erwartungen übertroffen werden (positiv; z.B. mehr als erwartet von einem Inhalt, der interessant ist) oder enttäuscht werden (negativ; z.B. mehr als erwartet von einem Inhalt, der nicht interessiert; Author et al., 2021). Die objektive Valenz der Erwartungs-Studienfach-Passung sollte auch zur Vorhersage des Studienerfolgs beitragen, da eine größere negative Valenz der Erwartungs-Studienfach-Passung zu einer höheren Wahrscheinlichkeit von verzerrten Entscheidungen und mehr Enttäuschungen im Studium führen sollte.

Zusammengefasst postulieren wir, dass objektive Studienfach-Passung hinsichtlich Interessen (H1), Erwartungen (H2) und Fähigkeiten (H3) über die subjektive Einschätzung der Studieninteressierten hinaus Studienerfolg vorhersagt, da diese Faktoren auf einer fundierten Theorie beruhen (Person-Umwelt-Passungs-Theorie) und auf einer validen Konstruktion der zukünftigen Situation (Experteneinschätzungen eines Studienfachs).

Methode

Wir testeten unsere Hypothesen mittels multipler Regressionsanalysen in einer Längsschnitt-Feldstudie. Über 4 Jahre wurden mehr als 4000 Studieninteressierte nach ihrer subjektiven Erfolgs- und Passungseinschätzung hinsichtlich des Psychologiestudiums befragt, bevor mittels eines Online-Self Assessments ihre objektive Passung (Interessen-, Fähigkeits-, Erwartungspassung) ermittelt wurde. Die Analysen basieren auf den Studieninteressierten, welche später das Psychologiestudium ergriffen und an einer Umfrage zu ihrem Studienerfolg (Motivation, Zufriedenheit, Abbruchintention, Leistung) teilnahmen (nErstsemester=396, nDrittsemester=136).

Ergebnisse

Über die subjektiven Einschätzungen der Studieninteressierten hinaus sagte höhere objektive Interessen-Studienfach-Passung vor Studienbeginn höhere Motivation ($\beta = .27^{***}$) und Zufriedenheit ($\beta = .19^{***}$) im ersten Semester vorher. Die Fähigkeits-Studienfach-Passung hingegen sagte bessere Leistung ($\beta = .43^{***}$) und die Erwartungs-Studienfach-Passung neben besserer Leistung ($\beta = .21^{**}$) zusätzlich höhere Zufriedenheit ($\beta = .17^*$) im dritten Semester vorher.

Die Ergebnisse liefern erste Hinweise darauf, dass objektive Studienfach-Passung potenziell Verzerrungen in den subjektiven Einschätzungen von Studieninteressierten reduzieren kann. Objektive Indikatoren können somit potenziell bei der Wahl eines passenden Studienfachs unterstützen und somit zu Studienerfolg beitragen. Weitere theoretische Implikationen für Forecasting-Theorien und die Person-Umwelt-Passungs-Theorie sowie praktische Implikationen für die Studienorientierung werden diskutiert.

Identifikation möglicher Prädiktoren von Studienabbrüchen in der Mathematik –Zwischenergebnisse einer Längsschnittstudie in der Studieneingangsphase

Svenja Kaiser¹, Markus Vogel¹, Leif Döring², Stefan Münzer²

¹Pädagogische Hochschule Heidelberg, Deutschland, ²Universität Mannheim, Deutschland

Die Studienabbruchquote in der Mathematik und den Naturwissenschaften lag 2020 bei 50% (Heublein et al., 2022). Einige Ursachen sind bereits identifiziert worden: Inhaltliche Überforderung in der Studieneingangsphase, Probleme mit der mathematischen Fachsprache, große Heterogenität bei den schulischen Vorkenntnissen sowie Probleme bei der Studienfinanzierung (Heublein et al., 2010; Blömeke, 2016; Hefendehl-Hebeker, 2016; Frischmeier et al., 2016). Die Studierenden unterscheiden sich in ihren Eingangsvoraussetzungen, wie dem schulischen Vorwissen, ihrem akademischen Selbstkonzept oder dem gewählten Studiengang (Fischer & Biehler, 2011).

Insbesondere die Beweise bereiten den Studienanfänger*innen in der Mathematik Schwierigkeiten, sowohl das eigene Beweisen als auch das Verstehen gegebener Beweise (Neuhaus-Eckhardt, 2022). Der Einfluss des (fehlenden) Beweisverständnisses auf Studienabbrüche in der Mathematik wurde bisher nicht untersucht.

Die vorliegende Längsschnittstudie adressiert die Frage, mit welchen Vorkenntnissen hinsichtlich mathematischen Vorwissens und des Beweisverständnisses, mit welcher Lern- und Leistungsmotivation und mit welchem akademischem Selbstkonzept Mathematikstudierende in die Studieneingangsphase starten und wie sich diese Eingangsvoraussetzungen zum einen innerhalb eines Semesters und zum anderen innerhalb der ersten drei Semester entwickeln. Es soll außerdem der Einfluss dieser Eingangsvoraussetzungen auf die Variable Studienabbruch untersucht werden.

Jeweils im Herbst-/Wintersemester 2022 und 2023 wurden sowohl Erst- als auch Drittsemesterstudierende der Universität Mannheim zu Beginn und zum Ende des Semesters mittels Fragebogen zu Lern- und Leistungsmotivation (Lernziele (LZ), Annäherungsleistungsziele (AL), Vermeidungsleistungsziele (VL) und Arbeitsvermeidungsziele (AV)), akademischem Selbstkonzept und demographischen Daten befragt. Zusätzlich wurden jeweils ein Leistungstest zum Beweisverständnis sowie zu Beginn des Semesters ein Test des mathematischen Vorwissens durchgeführt. Bereits 2021 wurden Drittsemesterstudierende zu Beginn und zum Ende des Semesters zu Lern- und Leistungsmotivation und akademischem Selbstkonzept befragt.

Auswertungen innerhalb eines Semesters beziehen sich auf die durch Pseudonym zugeordneten Fälle, bei denen sowohl Daten vom Beginn als auch vom Ende des Semesters vorlagen. Beispielsweise nahmen 2022 zu Beginn des Semesters 98 Erstsemesterstudierende an der Befragung teil, am Ende des Semesters 44. Daraus konnten 37 vollständige Datensätze zugeordnet werden.

Für die Längsschnittbetrachtung über die ersten drei Semester konnten die Datensätze der Erstsemesterstudierenden aus 2022 und der Drittsemesterstudierenden aus 2023 in 50 Fällen durch das Pseudonym verknüpft werden. Die Pseudonyme der anderen Erstsemesterstudierenden von 2022 tauchten in den Befragungen von 2023 nicht wieder auf, diese sind also Kandidaten für einen möglichen Studienabbruch.

Von den insgesamt 491 Studierenden der verschiedenen Jahrgänge waren 36% weiblich und 56% männlich (übrige: keine Angabe), 73% studierten Wirtschaftsmathematik, 14% Lehramt, 5% Wirtschaftspädagogik und 4% VWL.

Beim Vergleich von Semesterbeginn und Semesterende 2021 durch einen t-Test für abhängige Stichproben konnten bei AL und VL signifikante Abnahmen ($p=0.00023$ und $p=0.0047$) und bei der AV eine Zunahme ($p=0.0010$) nachgewiesen werden. Beim Selbstkonzept bezüglich der Kriterialnorm war eine Abnahme bei den Erstsemesterstudierenden 2022 ($p=0.0014$) zu beobachten, bei den Erstsemesterstudierenden 2023 jedoch nicht. Bei den Drittsemesterstudierenden gab es keine signifikante Entwicklung des Selbstkonzepts.

Beim Beweisverständnis wurde durch einen t-Test für abhängige Stichproben die Leistung der Erstsemesterstudierenden 2022 mit ihrer Leistung ein Jahr später (2023) verglichen und es zeigte sich keine signifikante Verbesserung.

Um die binäre abhängige Variable des Verbleibs im Studiengang vorherzusagen, wurde eine binär logistische Regression durchgeführt. Alle gemessenen Eingangsvoraussetzungen wurden als unabhängige Variablen aufgenommen. Dabei erwiesen sich Abiturnote ($p=0.04$), Arbeitsvermeidungsmotivation ($p=0.009$) und die Punktzahl im Beweisverständnistest ($p=0.009$) als signifikante Prädiktoren, das Regressionsmodell war signifikant mit $p<0.001$ und Nagelkerke- $R^2=0.453$. Das Beweisverständnis als mathematikspezifischer Prädiktor hat somit über die Abiturnote und das mathematische Vorwissen hinaus eine eigene Aussagekraft.

In der Studie wurden potenzielle Studienabbrüche durch das Nicht-Wiederauftauchen der Pseudonyme identifiziert. Dabei könnten Studierende, die im dritten Semester beide Befragungen verpassten, fälschlicherweise als Studienabbrecher*innen eingeordnet worden sein. Im kommenden Herbst-/Wintersemester 2024 werden daher zusätzlich wöchentlich Daten zu Studienabbruchintentionen per Online-Umfrage erhoben und potenzielle Studienabbrüche systematisch nachverfolgt. Diese Daten werden voraussichtlich bis zur Tagung vorliegen.

Fachbezogene Unterschiede der Statistikangst von Lehramtsstudierenden zu Studienbeginn

Andreas Just, Linda Schirle, Christine Sälzer
Universität Stuttgart, Deutschland

Wer ein Studium beginnt, startet mit verschiedenen Eingangsvoraussetzungen, sowohl in Bezug auf Vorwissen als auch auf Einstellungen, Motivation und Überzeugungen (Klusmann, 2009). Dies zeigt sich besonders in anspruchsvollen Pflichtveranstaltungen der ersten Semester. Für Lehramtsstudierende sind dies in den Bildungswissenschaften häufig Kurse mit einem hohen Anteil an Statistik, die als Hürde erlebt werden. Dabei starten Studierende teils bereits mit einer Angst vor statistischen Inhalten, was sich gleichzeitig negativ auf den Lernerfolg und den Zugang zu den Themen auswirken kann (Bandalos, Finney & Geske, 2003).

Gerade im gymnasialen Lehramtsstudium sind zudem Gruppenunterschiede zwischen verschiedenen Fächerkombinationen erkennbar (z.B. Roloff Henoch et al., 2015; Glutsch, 2019). Bezüglich der zwei studierten Fächer gibt es in Baden-Württemberg keine curricularen Vorgaben des Landes in Bezug auf deren Kombination. Die Fächer sind jedoch unterschiedlich nah an statistischen Inhalten. Wir differenzieren naturwissenschaftlich-technische (NW), geistes- und sozialwissenschaftliche (GS) sowie sprachliche (SP) Fächer, die auch in Kombination gewählt werden können (NW+GS, NW+SP, GS+SP).

Wir untersuchen zwei Forschungsfragen:

Inwieweit hängt Statistikangst zu Studienbeginn mit den gemessenen statistischen Leistungen zusammen?

Wie unterscheiden sich die Fächergruppen bezüglich ihrer Statistikangst?

Dazu werden die Daten von 207 Studierenden des Lehramts Gymnasium zu Beginn ihres ersten Semesters an der Universität Stuttgart aus den Jahren 2020-2022 analysiert. Die Studierenden waren im Schnitt 20 Jahre alt, 73 % waren weiblich. Je nach Fachwahl wurden die Studierenden in die oben genannten Fächergruppen ($N(NW) = 43$; $N(GS) = 11$; $N(SP) = 28$; $N(NW+GS) = 32$; $N(NW+SP) = 16$; $N(GS+SP) = 69$) eingeordnet. Die Daten, die im Rahmen des Projektes EA-MINT BW erhoben wurden, beleuchten neben mathematischen Fähigkeiten die Eingangsvoraussetzungen von Lehramtsstudierenden. Dies geschieht über Skalen zur Statistikangst nach Faber et al. (2018), ergänzt um weitere Eingangsvoraussetzungen, z.B. die Leistung bezüglich der Mittelstufenmathematik (Behrendt et al., 2023). Die statistische Leistung wird über die Note der Veranstaltung „Erziehungswissenschaftliches Arbeiten“ am Ende des ersten Semesters erfasst.

Die Skalen zur Statistikangst mit den Dimensionen Besorgtheit, Aufgeregtheit und Vermeidungsvorstellungen, die von Faber et al. (2018) im ersten Mastersemester erfasst wurden, weisen eine ausreichende Güte für den Einsatz bei Studierenden zu Studienbeginn auf ($.89 < \alpha$), korrelieren jedoch hoch miteinander ($r = .86$). Bezüglich der Prädiktion der statistischen Leistung am Ende des ersten Semesters zeigt sich, dass lediglich Aufgeregtheit signifikant wird ($r = -.40$). Die Skalen der Statistikangst korrelieren signifikant hoch mit mathematischen Basisfähigkeiten der Mittelstufenmathematik im Studieneingangstest ($-.34 < r < -.55$), welche auch mit den statistischen Leistungen zusammenhängen ($r = -.24$).

Die Fächergruppen unterscheiden sich in Bezug auf ihre Statistikangst, nicht jedoch hinsichtlich ihrer Leistungen in Statistik. Für die Subdimensionen der Statistikangst, Besorgtheit (BES: $F(4,31) = 4.864$, $p < 0,01$), Aufgeregtheit (AUF: $F(4,31) = 4.244$, $p < 0,01$) und Vermeidungsvorstellungen (VER: $F(4,31) = 5.315$, $p < 0,01$), resultieren signifikante Mittelwertsunterschiede zwischen einzelnen Fächergruppen. Der Tukey-Test zeigt für alle drei Subdimensionen der Statistikangst, dass sich Studierende mit rein naturwissenschaftlicher Fächerkombination (NW) signifikant von Studierenden rein sprachlicher Fächer (SP) unterscheiden. Dies gilt auch für Studierende mit rein sprachlichen Fächern (SP) im Vergleich zu der Kombination aus einem naturwissenschaftlichen und einem geisteswissenschaftlichen Fach (NW+GS). Die Studierenden der übrigen Fächerkombinationen (z.B. NW mit GS+SP), unterscheiden sich nicht signifikant in ihrer Statistikangst. Es zeigt sich also nicht nur eine pauschal niedrigere Statistikangst bei Lehramtsstudierenden mit MINT-Fächern. Stattdessen variiert die Statistikangst bei Lehramtsstudierenden je nach Kombination ihrer Studienfächer.

Die Ergebnisse sprechen für eine genauere und fachgruppenbezogene Betrachtung der Statistikangst. Dies ist auch deshalb bedeutsam, weil sich die Mittelwerte der statistischen Leistung zwischen den Fächergruppen nicht signifikant voneinander unterscheiden. Der Beitrag schließt mit Überlegungen zum Umgang mit Statistikangst in Ergänzung zu Förderangeboten bezüglich der Statistikleistung. Dadurch soll die Studieneingangsphase im Lehramtsstudium für möglichst viele Studierende erfolgreich gestaltet werden.

Unpassende Erwartungen an schulische Praxisphasen begünstigen Studienabbruchintention von Lehramtsstudierenden

Theoretischer Hintergrund und Fragestellung

Die Stage-Environment-Fit-Theorie (Eccles, 1993) hebt hervor, wie bedeutsam die Passung zwischen individuellen Erwartungen und Umgebungsbedingungen für den Studienfolg ist. Eine unzureichende Passung kann zu einer geringeren Lernmotivation beitragen (Eccles, 1993) und einen Studienabbruch erklären. Da Praxisphasen im Lehramtsstudium zur Überprüfung der Berufswahl zentral sind (Porsch, 2019; Ulrich et al., 2020), könnte eine fehlende Passung zwischen Erwartungen der Studierenden und dem Praktikum in besonderem Maße zu Studienabbruchintentionen beitragen. Empirische Belege dafür fehlen weitgehend. Hinweise aus einer Interviewstudie (Meens & Bakx, 2019) deuten jedoch darauf hin, dass unerfüllte Erwartungen an frühe Praxiserfahrungen zu Studienabbrüchen bzw. Studienwechseln führen können.

Die sozialkognitive Karrieretheorie der Performanz und Persistenz (Lent et al., 1994; Lent & Brown, 2019) unterstreicht zudem die Bedeutung von Selbstwirksamkeitsüberzeugungen für Studienabbruchintentionen: Selbstwirksamkeitsüberzeugungen beeinflussen die Ergebniserwartungen der Studierenden, ihre Zielsetzungen sowie ihre Entscheidungen, das Studium fortzusetzen oder abzubrechen. Die vorliegende Studie kombiniert die beiden theoretischen Ansätze und untersucht, inwiefern Studienabbruchintention durch eine geringere Passung der Erwartungen ans Praktikum sowie durch unterrichtsbezogene Selbstwirksamkeitsüberzeugungen erklärt werden können.

Methode

Befragt wurden 180 Grundschullehramtsstudierende der Universität Greifswald am Ende ihres ersten, dritten bzw. fünften Fachsemesters im Wintersemester 2022/23 und 2023/24. Die Studierenden absolvieren bereits ab dem ersten Fachsemester semesterbegleitend über die ersten vier Studienjahre hinweg einen Praxistag pro Woche an einer Grundschule. Unterrichtsbezogene Selbstwirksamkeitsüberzeugungen wurden mit Subskalen (jeweils 4 Items) zu Instruktionsstrategien ($\alpha = .66$), Klassenmanagement ($\alpha = .79$) und Schüler*innen-Engagement ($\alpha = .76$) erfasst (Pfitzner-Eden et al., 2014). Die Passung der Erwartungen an den Praxistag ($\alpha = .72$) wurde mit einer eigens entwickelten Skala erhoben. Die Skala umfasste fünf Items, die erfragten, inwiefern die Erwartungen zur Einfindung in die Lehrkraftrolle, Berufseignung, Unterrichtsgestaltung, Umgang mit Heterogenität und Aufbau eines Vertrauensverhältnisses zu Schüler*innen im Praktikum erfüllt wurden. Die Abbruchintention ($\alpha = .84$) wurde mit fünf Items erhoben (Dresel & Grassinger, 2013). Zur Berechnung der Ergebnisse wurde eine Pfadanalyse durchgeführt, welche den Zusammenhang der Passung der Erwartungen an den Praxistag sowie der drei Subdimensionen der unterrichtsbezogenen Selbstwirksamkeitsüberzeugungen mit der Studienabbruchintention untersucht. Auf Grundlage des theoretischen Hintergrunds wurden zudem Pfade der unterrichtsbezogenen Selbstwirksamkeitsüberzeugungen auf die Passung der Erwartungen geschätzt.

Ergebnisse und Diskussion

Die Ergebnisse zeigten, dass eine geringere Passung der Erwartungen ans Praktikum mit einer höheren Studienabbruchintention einherging ($\beta = -0.26$; $p < .001$). Unterrichtsbezogene Selbstwirksamkeitsüberzeugungen waren nicht bedeutsam mit der Studienabbruchintention assoziiert. Es zeigte sich aber ein Zusammenhang der Selbstwirksamkeitsüberzeugungen fürs Klassenmanagement mit der Passung der Erwartungen ans Praktikum ($\beta = 0.26$; $p = .001$). Studierende mit geringeren Selbstwirksamkeitsüberzeugungen fürs Klassenmanagement berichteten demnach eine geringere Passung der Erwartungen, die letztendlich mit einer höheren Studienabbruchintention einherging. Dieser Zusammenhang lag auch für Selbstwirksamkeitsüberzeugungen fürs Schüler*innen-Engagement und der Passung der Erwartungen vor ($\beta = 0.22$; $p = .010$). Die Ergebnisse unterstreichen die Relevanz, in praktikumsvorbereitenden Veranstaltungen, eine Erwartungsklärung mit den Studierenden vorzunehmen.

S-49: Meta-Meta: Die Rolle von systematischen Übersichten zu Meta-Analysen in der aktuellen Bildungsforschung

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: SN 169

Chair der Sitzung: Maximilian Knogler

Chair der Sitzung: Valentin Emslander

Chair der Sitzung: Doris Holzberger

Meta-Meta: Die Rolle von systematischen Übersichten zu Meta-Analysen in der aktuellen Bildungsforschung

Chair(s): **Maximilian Knogler** (Technische Universität München, Deutschland), **Valentin Emslander** (Justus-Liebig-Universität Gießen)

Diskutant*in(nen): **Doris Holzberger** (Technische Universität München)

Die empirische Bildungsforschung hat in den letzten Jahren und Jahrzehnten enorme Steigerungen an Output in Form wissenschaftlicher Publikationen produziert. Entsprechend werden systematische Überblicksarbeiten, die zahlreiche Einzelbefunde integrieren, immer bedeutender und zahlreicher. Die zunehmende Verfügbarkeit von Meta-Analysen von Primärforschung erlaubt nun auch die Erstellung von systematischen Übersichten über Meta-Analysen in der Bildungsforschung. Derartige Übersichten auf einer Meta-Meta Ebene basieren auf einer breiten Datenbasis und ermöglichen Analysen und Erkenntnisse, die auf Ebene von Primärstudien oder Meta-Analysen nicht möglich sind (Polanin, Maynard, & Dell, 2017). Dieses Symposium widmet sich dem besonderen Mehrwert dieses sich derzeit etablierenden Formats und zeigt anhand von vier Beispielen, wie sich dieser Mehrwert jeweils realisieren lässt.

Ein Schwerpunkt des Symposiums liegt auf der detaillierten Darstellung der methodischen Vorgehensweisen. Darunter fällt unter anderem die Auswahl von Studien, die angewandten Analyseverfahren, sowie die Methoden zur Bewertung der Qualität und Heterogenität der eingeschlossenen Meta-Analysen. Neben einer kurzen Darstellung der jeweiligen Ergebnisse der Überblicksarbeit, zeigen die Beiträge welche Implikationen sich auf Basis der Befunde für die Weiterentwicklung des Forschungsfeldes und Umsetzungen in praktischen Zusammenhängen ableiten lassen.

Der erste Beitrag gibt einen Überblick über die Korrelate von Selbstkonzepten im K12-Kontext. Die Befunde deuten auf die Bedeutung persönlicher Merkmale für das Selbstkonzept hin. Besondere Relevanz kommt Leistungsindikatoren, aber auch sozio-emotionalen Aspekten im Zusammenhang mit positiven Selbstkonzepten zu. Die Literaturübersicht zeigt, dass (meta-analytischer) Forschungsbedarf insbesondere in Bezug auf die Rolle situativer Aspekte für die Selbstkonzepte von Kindern und Jugendlichen besteht.

Der zweite Beitrag zeigt die hohe Relevanz positiver Lehrkraft-Schüler:innen-Beziehungen für Leistung, Wohlbefinden, Motivation, Emotion, Kognition und Verhalten von Schüler:innen. Die Überblicksarbeit betont die Wichtigkeit von Open-Science Praktiken in Meta-Analysen, um klare Implikationen für die Schulpraxis ableiten zu können.

Der dritte Beitrag beschreibt die Selektion und Analyse spezifischer aggregierter Effektstärkeninformation zu effektiven instruktionalen Interventionen für das Praxisfeld des mathematisch-naturwissenschaftlichen Unterrichts. Die Ergebnisse zeigen ein umfangreiches Angebot von belastbaren und relevanten Forschungsbefunden für die evidenzorientierte Gestaltung des Unterrichts. Die Befunde unterstreichen die Bedeutung kontextspezifischer Effektstärkeninformation in der Beurteilung der Effektivität von Interventionen.

Der vierte Beitrag untersucht die Rolle von Lernaktivitäten beim Lernen mit digitaler Technologie an der Hochschule. Das systematische Review und die Second-Order Meta-Analyse zeigen auf, dass es darauf ankommt, wie digitale Technologien an der Hochschule eingesetzt werden. Die Orchestrierung von eher qualitativen (systematisches Review) und quantitativen Methoden (Second-Order Meta-Analyse) bei Überblicksarbeiten hat sich hierbei als vielversprechend erwiesen.

Insgesamt liefern die vorgestellten Überblicksarbeiten wichtige Erkenntnisse zum Status Quo und für die Weiterentwicklung unterschiedlicher Forschungsfelder, konkreter Modelle und evidenzorientierter Praxis in der Bildungsforschung. Sie zeigen, dass fundierte Entscheidungen in der Bildungspolitik und -praxis durch systematische Integration von meta-analytischer Forschung gestützt werden können. Darüber hinaus wird die Bedeutung methodischer Weiterentwicklungen betont und fundierte Vorschläge präsentiert wie Aussagekraft und Anwendbarkeit von Forschungsergebnissen erhöht werden kann. Die Diskussionen und Erkenntnisse tragen dazu bei, das Verständnis für effektive pädagogische Ansätze zu vertiefen und die Grundlage für evidenzorientierte Entscheidungen in der Bildung zu stärken.

Beiträge des Symposiums

Who am I? Eine Meta-Analyse zweiter Ordnung zu den Korrelaten des Selbst in Kindheit und Jugend

Jennifer Meyer¹, Thorben Jansen¹, John Hattie², Jens Möller¹

¹Leibniz Institute for Science and Mathematics Education, ²University of Melbourne

Theoretischer Hintergrund

Positive Selbstüberzeugungen bzw. Selbstkonzepte beeinflussen Gefühle, Gedanken und Verhalten (Oyserman, 2001) und sind somit wichtig für viele Lebensbereiche, beispielsweise für schulischen Erfolg (Möller et al., 2020) und Wohlbefinden (Extremera & Rey, 2018). Insbesondere in der Kindheit und Jugend wird die Grundlage für positive Selbstkonzepte gelegt, die durch eine Vielzahl von Faktoren beeinflusst werden. Relevante Einflussfaktoren für die Selbstkonzeptentwicklung werden in unterschiedlichen theoretischen Modellen beschrieben, beispielsweise sozial-kognitive Theorien (z.B. Bandura & Walters, 1977; Bandura, 2001) Erwartungs-Wert-Modelle (z.B. Atkinson, 1957; Wigfield & Eccles, 2002; Eccles & Wigfield, 2020) und Soziometertheorie (Leary & Baumeister, 2000). In diesen theoretischen Modellen werden unterschiedliche Schwerpunkte gelegt, grundsätzlich werden jedoch Aspekte auf Seiten des Individuums, beispielweise Leistungsaspekte (Möller et al. 2020), kognitive und nicht-kognitive Faktoren, sowie Hintergrundvariablen (z.B. Gender) und situative Aspekte (z.B. kulturelles Milieu, Annahmen von Bezugspersonen) als Faktoren beschrieben, die bei der Entwicklung von Selbstkonzepten eine Rolle spielen. Auf Basis dieser theoretischen Grundlagen haben viele empirische Studien und darauf aufbauende Meta-Analysen die Relevanz unterschiedlicher Korrelate für die Ausprägung von Selbstüberzeugungen untersucht.

Ziele/Fragstellungen

Die vorliegende Zusammenfassung meta-analytischer Befunde integriert den umfangreichen Forschungsstand und verfolgt zwei Forschungsziele: 1) Integration und Vergleich der meta-analytischen Befunde zur Frage, welche Variablen am stärksten mit Selbstüberzeugungen im K12 Kontext zusammenhängen 2) Kodierung methodischer Charakteristika der inkludierten Meta-Analysen, insbesondere in Bezug auf Transparenz und Reproduzierbarkeit und so eine Einschätzung über das Risiko von Verzerrungen bei der Interpretation der Ergebnisse in diesem Forschungsbereich.

Methoden

Diese Meta-Analyse zweiter Ordnung fasst 105 Meta-Analysen zusammen, die vor Juli 2023 veröffentlicht wurden und Korrelate von Selbstkonzepten in K12 berichten. Diese umfassen 493 Effektgrößen erster Ordnung, die auf mehr als 8.500 Primärstudien und über 16 Millionen Kindern und Jugendlichen basieren. Wir berechneten standardisierte mittlere Unterschiede (SMD) mittels robuster Varianzschätzung, um die Schachtelung von Effektstärken innerhalb von Meta-Analysen zu berücksichtigen. Effektstärken wurden auf Basis theoretischen Annahmen und der gefunden Effektstärken in einem iterativen Verfahren gruppiert.

Ergebnisse/Diskussion

Individuelle Merkmale der Schüler:innen (SMD = 0.50) zeigten stärkere Zusammenhänge mit Selbstüberzeugungen als Interventionen (SMD = 0.27). Leistung (SMD = 0.66) und nicht-kognitive Variablen (SMD = 0.67) waren die individuellen Merkmale, die am stärksten mit den Selbstüberzeugungen in Verbindung standen, verglichen mit kognitiven Fähigkeiten (SMD = 0.30) und Hintergrundvariablen (SMD = 0.21). Interventionen, die auf individuelle Eigenschaften (SMD = 0.35) und insbesondere auf die Verbesserung von Selbstüberzeugungen direkt (SMD = 0.52) abzielten, zeigten größere Effektgrößen als Interventionen, die sich auf die Verbesserung des Unterrichts und der Klassenstruktur konzentrierten (SMD = 0.20). Nur wenige Meta-Analysen untersuchten situative Aspekte im Zusammenhang mit Selbstüberzeugungen, wie beispielsweise den geografischen Ursprung der Stichprobe. Diese Meta-Analyse zweiter Ordnung integriert meta-analytische Forschung über unterschiedliche theoretische Ansätze und terminologische Unterschiede und bietet somit einen grundlegenden Überblick und Systematisierung des meta-analytischen Forschungsstandes zum Thema Selbstkonzept. Forschungsdesiderata in der meta-analytischen Literatur sowie Potentiale und Grenzen des Ansatzes werden diskutiert.

Lehrkraft-Schüler:innen-Beziehungen: Neun Meta-Analysen zweiter Ordnung zu Zusammenhängen mit Schüler:innen-Variablen

Valentin Emslander¹, Doris Holzberger², Sverre Berg Ofstad³, Antoine Fischbach⁴, Ronny Scherer³

¹Justus-Liebig-Universität Gießen, ²Technische Universität München, ³University of Oslo, ⁴University of Luxembourg

Theoretischer Hintergrund

Lehrkraft-Schüler:innen-Beziehungen (LSB) sind ein wichtiger Bestandteil des Lernens und hängen mit akademischem Erfolg und Wohlbefinden zusammen. LSB werden als Interaktionen definiert, die durch emotionale Wärme (positive LSB), Konflikte und Abhängigkeit (negative LSB; Hamre & Pianta, 2001; Pianta, 1999) gekennzeichnet sind. Sowohl positive als auch negative LSB sind gut erforscht, und Metaanalysen haben Erkenntnisse zu ihren Zusammenhängen mit Schüler:innen-Variablen zusammengefasst. Diese metaanalytische Forschung hat aufgezeigt, wie positive LSB beispielsweise mit Schulengagement (Roorda et al., 2017), guten Peer-Beziehungen (Endedijk et al., 2022), kognitive Funktionen (Vandenbroucke et al., 2018) und allgemeinem Wohlbefinden (Chu et al., 2010) korreliert sind. Metaanalysen zu negativen LSB haben Korrelationen mit negativen schulbezogenen Emotionen (Nurmi, 2012) und Mobbing (Krause & Smith, 2022) berichtet.

Diese Metaanalysen unterscheiden sich jedoch erheblich in ihren Gesamtergebnissen. Die Ergebnisse sind auch uneindeutig darüber, welche Moderatoren die Beziehung zwischen LSB und Schüler:innen-Variablen beeinflussen und in welche Richtung. Obwohl Geschlecht und Alter am häufigsten untersucht wurden, sind die metaanalytischen Ergebnisse inkonsistent. Meta-Meta- und Moderatoranalysen können hier zu einem nuancierten Verständnis beitragen, wann und für wen die LSB-Variablen-Zusammenhänge am stärksten ausgeprägt sind. Zudem schwankt die methodische Qualität stark zwischen den bisherigen Metaanalysen, was die Interpretation der Ergebnisse erschwert.

Fragestellung

Mit Hilfe eines systematischen Reviews von Metaanalysen und mehreren Meta-Metaanalysen untersuchen wir

- (a) die meta-analytische Evidenzen zum Zusammenhang zwischen LSB und Schüler:innen-Variablen,
- (b) den Einfluss von Moderatoren auf diese Zusammenhänge und
- (c) die methodische Qualität der Metaanalysen.

Methode

Nach der Präregistrierung erfolgte eine systematische Literatursuche. Mit Hilfe unserer Ein- und Ausschlusskriterien identifizierten wir 26 passende Metaanalysen, aus denen wir 119 meta-analytische Effektstärken auf der Grundlage von über 2,5 Millionen (Vor-)Schüler:innen extrahierten. Für Forschungsfrage 1 verwendeten wir Mehrebenen-Metaanalysen zweiter Ordnung (Cooper & Koenka, 2012). Forschungsfrage 2 zu den möglichen Moderatorvariablen untersuchten wir durch Meta-Regressionen. Für Forschungsfrage 3 bewerteten wir die Qualität der Metaanalysen mit Hilfe der AMSTAR-Skala (angepasst an korrelative Studien in der Psychologie und Bildungsforschung; Shea et al., 2007) und verglichen diese samt der statistischen Power.

Ergebnisse und ihre Bedeutung

Wir fanden ähnliche große signifikante Beziehungen zwischen LSB und acht Clustern von Schüler:innen-Variablen: akademische Leistungen, schulbezogene Emotionen, angemessenes Verhalten, Verhaltensprobleme, exekutive Funktionen und Selbstkontrolle, Motivation, Gefühl der Zugehörigkeit zur Schule und Engagement sowie Wohlbefinden. Der Zusammenhang mit Mobbing war nur marginal signifikant. Diese Ergebnisse weisen auf eine breite Generalisierbarkeit positiver Zusammenhänge zwischen LSB und Schüler:innen-Variablen hin.

Unsere Moderatoranalysen deuteten auf eine größere Effektstärken bei Schüler:innen in der Mittel- und Oberstufe als bei Vor- und Grundschüler:innen hin. Dies ist überraschend, da es entgegen Bowlby's (1982) Bindungstheorie eine größere Wichtigkeit von LSB bei älteren Schüler:innen anzeigt. Da nur ein kleiner Teil der getesteten Variablen einen moderierenden Einfluss aufzeigte, scheint der Zusammenhang zwischen LSB und Schüler:innen-Variablen robust über viele Situationen und Merkmale hinweg zu sein. Weitere mögliche Moderatoren werden diskutiert.

Außerdem stellten wir große Unterschiede in der methodischen Qualität der Metaanalysen fest, die aber keinen moderierenden Effekt aufwiesen. Vor allem bei der Anwendung von Open-Science-Praktiken fanden wir großen Nachholbedarf. Ebenso wie die generelle Qualität der Meta-Analysen, moderierten die Unterschiede in statistischer Power die Ergebnisse nicht.

Diese Befunde geben einen breiten Überblick über das Feld der LSB-Forschung; präsentieren ihre Beziehungen, Moderatoren und methodologische Qualität in Meta-Analysen; und zeigen, wie LSB für eine breite Palette von Schüler:innen-Variablen und -stichproben gleichermaßen wichtig sind. So sollten Lehrkräfte für die Wichtigkeit von LSB und deren Zusammenhänge sensibilisiert werden. Es bedarf neuer theoretischer Modelle, welche die Wichtigkeit von LSB auch bei älteren Schüler:innen abbilden. Zudem sollte zukünftige metaanalytische Forschung nach besserer Qualität und Reproduzierbarkeit der Ergebnisse streben, z.B. durch die Anwendung von Open-Science-Praktiken.

Ein systematisches Review zu Metaanalysen im mathematisch-naturwissenschaftlichen Unterricht

Maximilian Knogler, Andreas Hetmanek, Tina Seidel
Technische Universität München

Theoretischer Hintergrund

In vielen Ländern besteht ein zunehmender Konsens über die Ziele naturwissenschaftlich-mathematischen Unterrichts. Diese Ziele sind u.a. in Bildungsstandards festgehalten. Diese Standards definieren jedoch keine spezifischen Lehr-Lern-Strategien, wie Lehrkräfte diese Ziele effektiv im Unterricht ansteuern können. Die Identifikation wirksamer Lehrstrategien gehört zu den zentralen Aufgaben der empirischen Bildungsforschung (Slavin, 2020; Shavelson & Towne, 2004). Der enorm steigende Forschungs-Output relevanter Forschungszweige lässt ein umfangreiches Angebot an relevanten Befunden vermuten, das allerdings nicht in selektierter Form vorliegt (Hedges, 2018). Ziel dieses systematischen Reviews ist es, die beste verfügbare Evidenz für effektive Lehrstrategien für den naturwissenschaftlich-mathematischen Unterricht zu selektieren und zu analysieren (Ahn, Ames, & Myers, 2012). Dazu entwickelten wir eine spezifische Auswahlheuristik, die zwei wesentliche Kriterien berücksichtigt: A.) möglichst hohe Validität der Wirksamkeitsbehauptungen durch aggregierte experimentelle Befunde B.) möglichst gute Übertragbarkeit der Ergebnisse für den spezifischen Kontext des mathematisch-naturwissenschaftlichen Unterrichts der Sekundarstufe.

Forschungsfragen

Drei Forschungsfragen leiten diese Untersuchung:

1. Wie viele kontextspezifische aggregierte Effektstärken ergibt eine systematische Suche? Wie lassen sich diese Befunde klassifizieren?
2. Inwiefern unterscheiden sich diese kontextspezifischen Effektstärken von den Gesamteffektstärken, die in den Metaanalysen berichtet werden?
3. Welche Qualität haben die selektierten Metaanalysen?

Methode

Die systematische Suche fokussierte gängige wissenschaftliche Datenbanken und Fachzeitschriften im Zeitraum von ca. 20 Jahren (2000-2019). Es wurden nur Studien ausgewählt, die eine aggregierte Effektgröße auf der Basis experimenteller für Schüler:innenoutcomes im Bereich des mathematisch-naturwissenschaftlichen Unterrichts der Sekundarstufe bereitstellten. Auswahl und Kodierung erfolgte unabhängig durch zwei Autoren. Nicht-übereinstimmungen wurden durch Diskussion geklärt. Die Studie wurde präregistriert und das Kodierschema ist auf OSF einsehbar (https://osf.io/9n99n/?view_only=bb30c83e9bf34d73a79138ddcf91da5c). Aus selektierten Metanalysen wurden relevante Daten extrahiert und analysiert (siehe Polanin, Maynard & Dell, 2017).

Ergebnisse und ihre Bedeutung

Forschungsfrage 1: Es wurden 41 Meta-Analysen mit 78 kontextspezifischen Effektstärken identifiziert, die unsere Einschlusskriterien erfüllten. Die Mehrheit (N=32) fokussierte sich auf die Effektivität spezifischer Lehrstrategien wie Inquiry-Based Learning oder Flipped Classroom. Die verbleibenden Analysen (N=9) verglichen verschiedene Strategien hinsichtlich spezifischer Outcomes, wie z.B. kritisches Denken.

Forschungsfrage 2: In der Analyse der Unterschiede zwischen allgemeinen und kontextspezifischen Effektgrößen stellten wir fest, dass in 58% der Fälle keine oder nur geringe Unterschiede bestanden. In 29% der Fälle wurden moderate Unterschiede festgestellt, während 13% der Vergleiche signifikante Unterschiede aufwiesen.

Forschungsfrage 3: Die Qualität der Meta-Analysen wurde anhand von 37 Kriterien bewertet. Im Durchschnitt erfüllten die Studien 68% der Qualitätskriterien. Die Qualität variierte jedoch deutlich zwischen den Publikationen und einige wichtige Kriterien, wie die genaue Angabe von Effektgrößenformeln, wurden häufig nicht erfüllt.

Insgesamt zeigt dieses systematische Review, dass eine substantielle Anzahl an spezifischen aggregierten Effektgrößen vorliegt, die verschiedene Arten von Interventionen und Outcomes im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe betreffen. Häufige Unterschiede zwischen kontextspezifischen Effektstärken und Gesamteffekten in den Metaanalysen unterstreichen die Bedeutung des Kontexts für die Wirksamkeit. Zukünftige Forschung sollte verstärkt kontextspezifische Wirksamkeitsinformationen und die wissenschaftliche Qualität der Studien berücksichtigen, um eine fundierte Grundlage für die evidenzorientierte Gestaltung von Unterricht zu schaffen.

Digitales Lehren und Lernen an der Hochschule – ein systematisches Review und Second-Order Meta-Analyse

Michael Sailer¹, Rebecca Maier², Sonja Berger², Tamara Kastorff³, Karsten Stegmann²

¹Universität Augsburg, ²Ludwig-Maximilians-Universität München, ³Technische Universität München

Theoretischer Hintergrund

Technologieunterstütztes Lernen (Technology-Enhanced Learning, TEL) spielt eine zunehmend wichtige Rolle in der Hochschulbildung. Beim konkreten Einsatz von Technologie in der Hochschullehre, stellt sich jedoch für viele Dozierende häufig die Frage nach dem „Wie“. Obwohl es eine Vielzahl an Primärstudien und Meta-Analysen gibt (Tamim et al., 2011), die eine spezifische Antwort für den Einsatz bestimmter Technologien unter bestimmten Bedingungen bereithalten (z.B. Bernard, 2004), sind Untersuchungen der zentralen Mechanismen von TEL in der Forschung bisher ausgeblieben. Aktuelle Befunde deuten darauf hin, dass es auf die Art und Weise des digitalen Technologieeinsatzes ankommt und weniger auf die Häufigkeit (Fütterer et al., 2022). Dazu passend postulieren Modelle wie ICAP (Interactive, Constructive, Active, Passive), dass die Effekte von TEL durch die Art der Lernaktivität, die durch die Technologie ermöglicht wird, vermittelt werden (Chi & Wiley, 2014). Darüber hinaus legen Forschungsarbeiten nahe, dass auch die Unterstützung spezifischer Lernaktivitäten, wie die Sequenzierung von

Lernmaterial bei passiven Lernaktivitäten (Sweller, 2024) oder Kollaborationsskripte bei interaktiven Lernaktivitäten (Fischer et al., 2013) den Lernerfolg steigern können.

Fragestellung

In dieser Forschungsarbeit sollen die zentralen Mechanismen von TEL untersucht werden. Auf der Grundlage theoretischer Überlegungen und empirischer Befunde, wird vermutet, dass TEL einen positiven Effekt auf die Lernergebnisse von Studierenden hat, insofern die kognitive Auseinandersetzung der Studierenden mit dem Lerninhalt erhöht wird. Diese positive Auswirkung kann einerseits (a) durch die Bereitstellung kognitiver Unterstützung für die Ausführung kognitiver Prozesse erreicht werden oder andererseits (b) durch die Initiierung eines höheren Lernaktivitätsniveaus, das mit einer erhöhten Wahrscheinlichkeit für eine Vielzahl von kognitiven Prozessen (z.B. aktivieren, verknüpfen, schlussfolgern oder kombinieren; Chi & Wiley, 2014) einhergeht.

Methode

Um dies zu untersuchen, wurde ein systematisches Review von Meta-Analysen und eine Second-Order Meta-Analyse durchgeführt. Spezifisch wurde nach Effektstärken aus Meta-Analysen gesucht, welche sich auf eine statistische Synthese von Effekten von Lernen mit und ohne Technologien in (quasi-)experimentellen Designs in der Hochschulbildung beziehen. In beiden Bedingungen der meta-analytischen Effektstärken (Lernen mit und ohne Technologie) wurden die Lernaktivitäten der Studierenden in den Technologie- und Nicht-Technologie-Bedingungen berücksichtigt und kodiert. In der Studie wurden lediglich Effektstärken aufgenommen, die sich auf objektive Lernergebnisse beziehen. Die finale Stichprobe bestand aus 28 Meta-Analysen mit 45 Effektstärken aus Meta-Analysen, die insgesamt 1286 Effektstärken aus Primärstudien umfassen. Zum einen werden die Befunde in Anlehnung an Stegmann (2020) zunächst qualitativ berichtet und im zweiten Schritt quantitativ in einer Second-Order Meta-Analyse zusammengefasst.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass der Einsatz digitaler Technologie als Substitution für nicht-technologiebasierten Unterricht keine wesentliche Veränderung der kognitiven Lernergebnisse der Studierenden zur Folge hatte. Allerdings verbesserten sich die kognitiven Lernergebnisse signifikant, wenn die Technologie spezifische Unterstützung für Lernaktivitäten bot (siehe Mechanismus b). Digitale Technologien, die fortgeschrittenere Lernaktivitäten ermöglichten, führten zu höheren kognitiven Lernergebnissen der Studierenden (siehe Mechanismus a).

Diese Ergebnisse deuten darauf hin, dass die Effekte von TEL maßgeblich durch die durch Technologie geförderten Lernaktivitäten vermittelt werden. Für die Praxis bedeutet dies, dass die gezielte Förderung aktiver, konstruktiver und interaktiver Lernaktivitäten durch digitale Technologien das Lernen in der Hochschulbildung erheblich verbessern kann. Weiterhin können digitale Technologien, die kognitive Unterstützung (z.B. Scaffolding, Sequenzierung) bieten, die Lernergebnisse positiv beeinflussen.

Methodisch zeigte sich die Kombination aus qualitativen und quantitativen Darstellungen der Befunde und deren Interpretation vor dem Hintergrund dieser Analyse als sehr hilfreich und passend. Außerdem zeigt die Studie den Mehrwert von Second-Order Meta-Analysen auf, insofern sie auf einer fundierten theoretischen Basis beruhen und somit die Interpretation vor dem Hintergrund der entsprechenden Theorien ermöglichen.

S-50: Wie Lehrkräfte Familien sehen: Erkenntnisse aus Vignetten-Studien über Stereotype von Lehrkräften auf familiäre sozioökonomische Herkunft

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 135

Chair der Sitzung: Janina Dickert

Chair der Sitzung: Oscar Yendell

Chair der Sitzung: Melanie Olczyk

Wie Lehrkräfte Familien sehen: Erkenntnisse aus Vignetten-Studien über Stereotype von Lehrkräften auf familiäre sozioökonomische Herkunft

Chair(s): **Janina Dickert** (Bergische Universität Wuppertal, Deutschland), **Oscar Yendell** (Universität Mannheim)

Diskutant*in(nen): **Melanie Olczyk** (MLU Halle-Wittenberg)

Bildungsverläufe hängen in Deutschland in einem hohen Maße mit dem sozioökonomischen Status (SES) von Schüler*innen zusammen, indem Schüler*innen mit niedrigem SES selbst bei gleicher Leistung wie Schüler*innen mit höherem SES seltener gymnasiale Schulformen besuchen, das Abitur absolvieren oder ein Studium aufnehmen (Autor:innengruppe Bildungsberichterstattung, 2024; Dräger, 2022). Diese Ungleichheit steht dem Motto der diesjährigen GEBF „Bildung als Schlüssel für gesellschaftliche Herausforderungen“ entgegen, da soziale Ungleichheiten hierdurch nicht verringert, sondern fortgeschrieben werden. Ein Ansatz, diese - selbst bei gleicher Leistung existierenden – ungleichen Bildungsverläufe zu erklären, sind Stereotype von Lehrkräften, verstanden als generalisierte Überzeugungen über Angehörige sozialer Gruppen (Tobisch & Dresel, 2020). Diese können zu verzerrten Leistungserwartungen oder Bewertungen bei Lehrkräften führen, aber auch Attributionen für gute bzw. schlechte Schulleistungen beeinflussen (Gentrup, 2018; Wang & Hall, 2018; Reyna, 2008). Bisherige quantitative (sozialpsychologische) Studien fokussieren sich überwiegend auf SES-bezogene Stereotypisierungen gegenüber den Schüler*innen selbst (bspw. Glock & Kleen, 2020; Tobisch & Dresel, 2020). Dabei zeigen qualitativ und ethnographisch orientierte Studien, dass Lehrkräfte in Bezug auf schulische Leistung oftmals auf die Familien entsprechender Schüler*innen rekurrieren und das Verhalten sowie die Leistung von Schüler*innen in Bezug zum familiären SES-Hintergrund setzen (Budde et al., 2023; Weitkämper, 2022). Im Rahmen dieses Symposiums werden daher sozialpsychologische Studien präsentiert, die Stereotype von Lehrkräften auf Familien in Abhängigkeit unterschiedlicher familiärer SES-Herkünfte untersuchen. Die verschiedenen Beiträge zeigen, in welchem Ausmaß und über welche Mechanismen Informationen über das Elternhaus die Wahrnehmung von Lehrkräften beeinflussen können. Beitragsübergreifend werden hierfür Vignetten verwendet, die sich hinsichtlich Erhebungsarten (offene und geschlossene Angaben) sowie Auswertungsformen (quantitative Inhaltsanalyse und Varianzanalysen) unterscheiden, um einen Überblick über die Vielfalt von Vignettenforschung zu geben.

Im ersten Beitrag wurde untersucht, ob sich Lehrkraft-Attributionen für schlechte Schulleistungen je nachdem unterscheiden, ob die Eltern eines Schülers der Mittelschicht angehören, Einkommensarm sind oder Sozialhilfe beziehen. Es wird gezeigt, dass vielfältige inhaltliche Kategorien (11) für die Attribution schlechter Schulleistung existieren. Darüber hinaus wird gezeigt, dass bei Schüler*innen in Einkommensarmut und Transferleistungsbezug häufiger die Familie als Grund für schlechte Schulleistungen angegeben wird als bei Eltern, die der Mittelschicht angehören.

Der zweite Beitrag geht der Frage nach, inwiefern Informationen über den Bildungshintergrund der Eltern die Leistungsbeurteilung und die Attributionen der Leistung durch angehende Lehrkräfte beeinflussen. Dabei zeigte sich, dass angehende Lehrkräfte die Leistung sowie die Wahrscheinlichkeit, dass ein Schüler das Abitur erreichen wird, höher einschätzen, wenn die Eltern des Schülers über einen Masterabschluss verfügen, als wenn sie über einen Hauptschulabschluss verfügen. Zudem wird die Leistung des Schülers von den Teilnehmenden weniger auf elternbezogene Ursachen zurückgeführt, wenn die Eltern einen Masterabschluss haben, als wenn sie einen Hauptschulabschluss haben.

Der dritte Beitrag befasst sich mit dem Einfluss von Informationen über den familiären SES (Beruf des Vaters) auf die Einschätzung von akademischen Leistungen, Engagement, Beliebtheit und Störungstendenz eines fiktiven Schülers durch angehende Lehrkräfte. Dabei wurde zwischen hohem SES, niedrigem SES und Sozialhilfebezug variiert. Es konnte festgestellt werden, dass Schüler mit hohem und niedrigem SES im Vergleich zu Schülern, deren Eltern Sozialleistungen beziehen, besser in der akademischen Leistung und Beliebtheit bewertet wurden, wobei es keine Unterschiede zwischen hohem und niedrigem SES gab.

Alle Beiträge zeigen den Bedarf auf, in zukünftiger sozialpsychologischer Stereotyp-Forschung nicht mehr lediglich auf Lehrkraft-Stereotype gegenüber Schüler*inneneigenschaften direkt, sondern auch verstärkt auf Stereotype gegenüber den Familien der Schüler*innen zu fokussieren. Neben einer Diskussion der methodischen Umsetzungen, wird die Diskussion die Ergebnisse zudem aus soziologischer Perspektive in einen breiteren gesellschaftlichen Kontext setzen. Dies ermöglicht eine interdisziplinäre Auseinandersetzung mit den Ergebnissen der vorrangig sozialpsychologischen Studien.

Beiträge des Symposiums

Stereotype Attributionen von Lehrkräften auf schlechte Schulleistung von Schüler*innen unterschiedlicher sozioökonomischer Herkunft – Die Wahrnehmung familiärer Unterstützung

Oscar Yendell¹, Hannah Kleen², Meike Bonefeld³, Karina Karst¹

¹Universität Mannheim, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ³Albert-Ludwigs-Universität Freiburg

Stereotype können interpersonale Attributionen – verstanden als Ursachenzuschreibung für Leistung – verzerren (Reyna, 2008). (Angehende) Lehrkräfte stereotypisieren Schüler*innen eines niedrigen sozioökonomischen Status (SES) bezüglich Leistungsbereitschaft negativer als Schüler*innen höherer SES (Shevchuk & Glock, 2022). Außerdem werden verschiedene niedrige SES unterschiedlich stereotypisiert, indem Lehramtsstudierende Bürgergeld-Empfangende hinsichtlich Leistungsbereitschaft negativer stereotypisieren als Einkommensarme (Yendell et al., 2023). Entsprechende Stereotype können dazu führen, dass Schüler*innen die Kontrolle für schlechte Schulleistung zugeschrieben wird (Glock & Kleen, 2020). Kontrolle ist eine von drei Attributionsdimensionen und bei schlechter Schulleistung hängt ein kontrollierbarer Attributionsstil von Lehrkräften u.a. mit lehrkraftseitiger Wut zusammen (Wang & Hall, 2018; Weiner, 1985). Lehrkräfte attribuieren schlechte Schulleistungen oftmals auf für Schüler*innen kontrollierbare Gründe, nennen aber auch (unkontrollierbare) familiäre Einflüsse (Wang & Hall, 2018). Eine Studie zeigt, dass Lehramtsstudierende schlechte Schulleistung von Schüler*innen niedriger SES auf unkontrollierbare Ursachen attribuieren, eine Attribution auf Familien wurde jedoch nicht untersucht (Glock & Kleen, 2020). Weitere Studien zeigen defizitäre Lehrkraft-Perspektiven auf Familien niedriger SES, hierbei fehlt jedoch eine Perspektive auf

Attributionen (Budde, 2023). Zudem bleibt unklar, ob schlechte Schulleistung bei Schüler*innen unterschiedlicher niedriger SES unterschiedlich attribuiert wird (Yendell et al., 2023).

Daher wird erstens explorativ untersucht, welche Gründe Lehrkräfte für schlechte Schulleistungen in Abhängigkeit des Schüler-SES nennen und zweitens, ob sich der Attributionsstil auf der Dimension Kontrolle unterscheidet. Hierfür werden Unterschiedshypothesen formuliert:

- Der Attributionsstil auf der Dimension Kontrolle unterscheidet sich je nach SES des Schülers (niedriger vs. mittlerer SES).
- Der Attributionsstil auf der Dimension Kontrolle unterscheidet sich je nach niedrigem SES des Schülers (Bürgergeldbezug vs. Einkommensarmut).

Drei Vignetten beschrieben identisch die schlechte Schulleistung eines Schülers unterschiedlicher SES: Eltern als Bürgergeldempfänger*innen (niedriger SES), Reinigungskräfte (niedriger SES) oder Verwaltungsmitarbeitende (mittlerer SES). An der Online-Studie (Between-Subjects-Design) nahmen 149 Lehrkräfte (17 Sekundarschulen) des Projekts „Schule macht stark“ teil (Maaz & Marx, 2024). Sie notierten fünf Gründe, wählten den wichtigsten Grund und füllten dazu Skalen zur Kontrolle des Schülers ($\Omega = .7$) und anderer Personen ($\Omega = .8$) aus (McAuley et al., 1992). Alle Gründe wurden durch zwei Rater*innen in elf induktive Kategorien (u.a. „Förderbedarfe“, „Schulbezogene Kompetenzen“, „innerfamiliäre Probleme“) kodiert ($\kappa = .91$) und mittels MANOVA analysiert (Coe & Scacco, 2017). Zur Analyse des wichtigsten Grundes wurden vier Oberkategorien (Familienbezogene, Schüler*innenbezogene, Schulbezogene, Systembezogene Attribution) gebildet, die mittels Chi-Quadrat-Test und Odds Ratio analysiert wurden. Die Angaben auf den Items zur Kontrolle wurden mittels MANOVA analysiert.

In der Angabe der Gründe sowie auf der Dimension der Kontrolle wurden zwischen den niedrigen SES keine Unterschiede gezeigt. Die MANOVA bezüglich der Gründe zeigte, dass Lehrkräfte je nach Schüler-SES bestimmte Gründe unterschiedlich häufig nannten, $F(22, 272) = 3.04$, $p < .001$, partielles $\eta^2 = .2$, Wilk's $\Lambda = .64$. Post-Hoc-ANOVAs und Games-Howell-Tests zeigten, dass Lehrkräfte bei den Experimentalbedingungen mit niedrigem SES signifikant mehr Gründe in den Kategorien „Familiales schulbezogenes Verhalten“ und „Externe Rahmenbedingungen“ nannten, während sie beim Schüler mit mittlerem SES signifikant mehr Gründe in den Kategorien „Unpassender Unterricht“ und „Pubertätsbedingte Einflüsse“ angaben. Bei der Angabe zum wichtigsten Grund wurde lediglich der Chi-Quadrat-Test zur „Familienbezogenen Attribution“ signifikant, $\chi^2(2) = 7.02$, $p = .03$, $\phi = .3$. Odds Ratio zeigte, dass die Chance einer familienbezogenen Attribution gegenüber dem transferleistungsbeziehenden Schüler 3.38-mal (Einkommensarm: 3.01) größer war als beim Schüler des mittleren SES. Die MANOVA zu den Kontrolle-Skalen zeigte einen signifikanten Unterschied zwischen den Experimentalbedingungen, $F(4, 290) = 4.43$, $p = .002$, partielles $\eta^2 = .06$, Wilk's $\Lambda = .89$. Nachfolgende ANOVAs und Games-Howell-Tests zeigten, dass Lehrkräfte bei den niedrigen SES-Bedingungen eine signifikant höhere Kontrolle durch andere Personen angaben, als beim mittleren Schüler-SES. Insgesamt wurde schlechte Schulleistung beim Schüler der niedrigen SES-Bedingungen stärker auf familiäre Gründe attribuiert als beim Schüler des mittleren SES. Entsprechend wird diskutiert, wie dies die Zusammenarbeit mit Familien beeinflussen kann.

Die Rolle des Bildungshintergrunds der Eltern für die Leistungsbeurteilung und Ursachenzuschreibung von Schüler*innen durch angehende Lehrkräfte

Janina Dickert, Sabine Glock
Bergische Universität Wuppertal

Der akademische Erfolg von Schüler*innen steht in einem engen Zusammenhang mit ihrem sozioökonomischen Status (SES). Dies gilt auch nach Kontrolle der tatsächlichen akademischen Leistung der Schüler*innen (Timmermans et al., 2015). Als Ursachen dafür werden Stereotype diskutiert, die die Beurteilung von Lehrkräften so beeinflussen können, dass diese nicht nur auf leistungsbezogenen Variablen beruht. Da Lehrkräfte für die regelmäßige Bewertung ihrer Schüler*innen verantwortlich sind, stellen sie nicht nur Vermutungen über die akademischen Leistungen der Schüler*innen an, sondern auch über die Ursachenzuschreibung der Leistungen, sodass auch diese durch Stereotype verzerrt werden können.

In einigen Studien konnte gezeigt werden, dass Informationen über den SES der Schüler*innen zur Aktivierung von Stereotypen führen können, was wiederum die Beurteilung der Lehrkräfte beeinflusst (Glock & Kleen, 2023). Der SES eines/einer Schüler*in kann anhand verschiedener Faktoren gemessen werden. Bislang wurde zur Operationalisierung des SES aber nur in wenigen Studien der Bildungshintergrund der Eltern herangezogen, welcher allerdings durch den Zusammenhang zur schulischen Leistung der Kinder in verschiedenen Phasen des Bildungsweges einen bedeutsamen Einflussfaktor darstellt (Rodríguez-Hernández et al., 2020). Vor diesem Hintergrund wird im vorliegenden Beitrag untersucht, inwiefern Informationen über den Bildungshintergrund der Eltern die Leistungsbeurteilung und die Ursachenzuschreibung von Lehramtsstudierenden beeinflussen können.

Dazu wurden 120 Lehramtsstudierende (58,33 % weiblich) befragt. Als Basis für die Beurteilung diente eine Vignette, in der ein Schüler beschrieben wurde, der geringe Motivation und Konzentration sowie durchschnittliche bis niedrige schulische Leistungen zeigt. Der Bildungshintergrund der Eltern wurde dabei variiert, sodass die Eltern in der einen Variante beide über einen Masterabschluss, in der anderen Variante beide über einen Hauptschulabschluss verfügen. Die Proband*innen schätzten die schulische Leistung des Schülers in Deutsch, Englisch und Mathematik auf einer 7-Punkt Likert-Skala von 1 (sehr niedrig) bis 7 (sehr hoch), sowie die Wahrscheinlichkeit des Erreichens des Abiturs auf einer 4-Punkt Likert-Skala von 1 (sehr unwahrscheinlich) bis 4 (sehr wahrscheinlich) ein. Darüber hinaus wurden die Ursachenzuschreibungen bezüglich mathematischer Fähigkeiten und Deutschkenntnisse als fähigkeitsbezogen, anstrengungsbezogen und elternbezogen auf einer 5-Punkt Likert-Skala von 1 (stimme überhaupt nicht zu) bis 5 (stimme vollkommen zu) erhoben.

Die Ergebnisse der t-Tests zeigten, dass die angehenden Lehrkräfte die schulischen Leistungen des Schülers, dessen Eltern einen Master-Abschluss haben, signifikant höher bewerteten ($M = 3.99$, $SD = 0.78$) als die des Schülers, dessen Eltern einen Hauptschulabschluss haben ($M = 3.39$, $SD = 0.72$). Gleiches zeigte sich für die Einschätzung der Wahrscheinlichkeit, dass der Schüler das Abitur erreichen wird. Außerdem ergab sich eine signifikante Interaktion zwischen Ursachenzuschreibung der Leistung und Bildungshintergrund der Eltern, $F(2, 228) = 12.20$, $p < .001$, $\eta^2 = 0.10$, wobei die Teilnehmenden den Eltern weniger Ursachen zuschrieben, wenn diese über einen Master-Abschluss verfügten ($M = 2.09$, $SD = 0.76$) im Vergleich zu den Eltern mit Hauptschulabschluss ($M = 2.72$, $SD = 0.81$), $t(116) = 4.32$, $p < .001$, $d = 0.80$.

Insgesamt zeigte sich, dass der Bildungshintergrund der Eltern als ein Faktor des SES zu Verzerrungen in den Urteilen von angehenden Lehrkräften führen kann. Die Ergebnisse stehen somit einerseits im Einklang mit Studien zur Urteilsverzerrung von Lehrkräften gegenüber Schüler*innen mit niedrigem und hohem SES (Brandmiller et al., 2020). Andererseits schließen die Ergebnisse auch an Studien zur Relevanz des elterlichen Bildungshintergrundes in der schulischen Interaktion an: So fanden beispielsweise Pit-ten Cate und Glock (2018) heraus, dass Lehrkräfte negativere implizite Einstellungen gegenüber Schüler*innen haben, deren Eltern einen niedrigen Bildungshintergrund aufweisen, als gegenüber Schüler*innen, deren Eltern einen hohen Bildungshintergrund aufweisen. Die Ergebnisse zur Interaktion zwischen der Ursachenzuschreibung und dem

Bildungshintergrund der Eltern haben unsere Hypothesen nur teilweise bestätigt. Dennoch legt die Studie damit nahe, dass Informationen über den Bildungshintergrund der Eltern als Einflussfaktor im Klassenzimmer und als Quelle sozialer Ungleichheiten in Bildungseinrichtungen künftig genauer untersucht werden sollten.

Urteile von angehenden Lehrkräften für Schüler mit schlechtem Arbeitsverhalten in Abhängigkeit des familiären sozioökonomischen Status

Anna Shevchuk¹, Ineke Pit-ten Cate², Janina Dickert¹, Sabine Glock¹

¹Bergische Universität Wuppertal, ²Universität Luxemburg

Theoretischer Hintergrund und Fragestellung

Der sozioökonomische Status (SES) der Familie kann Erwartungen und Urteile von Lehrkräften hinsichtlich schulrelevanter Fähigkeiten beeinflussen, häufig zugunsten von Schülern mit hohem SES (z.B. Glock & Kleen, 2020). Solche Erwartungen und Urteile können auf Stereotypen basieren (Jussim et al., 1996), d.h. auf verallgemeinerten Überzeugungen über Mitglieder sozialer Gruppen unabhängig von individuellen Unterschieden, und die Urteile von (angehenden) Lehrkräften beeinflussen. Das kann dazu führen, dass Schüler*innen mit hohem SES kompetenter wahrgenommen werden verglichen mit Schülern*innen mit niedrigem SES. Bei Betrachtung eines niedrigen SES zeigt sich überdies, dass Empfänger*innen von Sozialleistungen besonders negativ stereotypisiert werden, auch von angehenden Lehrkräften (Yendell et al., 2023). Der Einfluss von niedrigem SES im Vergleich zum Sozialhilfebezug auf die Urteile von Lehrkräften gegenüber Schüler*innen ist bisher kaum erforscht. In einer experimentellen Vignettenstudie untersuchten wir, wie angehende Lehrkräfte einen Schüler, dessen Eltern einen hohen vs. niedrigen SES vs. Sozialhilfebezug aufweisen, beurteilen. Wir erwarteten, dass die Urteile für den Schüler mit hohem SES positiver ausfallen würden als für den Schüler mit niedrigem SES und mit Sozialhilfebezug.

Methode

159 angehende Lehrer (142 weiblich) nahmen an dieser Studie teil. Ein fiktiver Schüler, der häufig seine Hausaufgaben nicht erledigte, wurde beschrieben. Der Beruf der Väter variierte zwischen hohem SES (Schulleiter), niedrigem SES (Taxifahrer) und Sozialhilfebezug. Die angehenden Lehrkräfte beurteilten den Schüler bezüglich akademischer Leistung, Engagement, Beliebtheit und Störungstendenz.

Ergebnisse und ihre Bedeutung

Eine MANOVA mit den vier Urteilsdimensionen als abhängige Variablen und dem elterlichen Status der Schüler (hoher SES vs. niedriger SES vs. Sozialleistungsbezug) als Zwischensubjektfaktor ergab einen signifikanten Haupteffekt. Separate ANOVAs für jede Dimension ergaben signifikante Ergebnisse für die akademische Leistung, $F(2, 192) = 14.10$, $p < .001$, $\eta^2 = 0.13$, und die Beliebtheit, $F(2, 192) = 15.25$, $p < .001$,

$\eta^2 = 0.26$. Die angehenden Lehrkräfte beurteilten die schulischen Leistungen des Schülers mit hohem SES ($M = 4.12$, $SD = 0.72$) und die des Schülers mit niedrigem SES ($M = 3.98$, $SD = 0.61$) höher verglichen mit dem Schüler mit Sozialhilfebezug ($M = 3.55$, $SD = 0.59$), ohne Unterschied zwischen hohem und niedrigem SES. Bezüglich der Beliebtheit wurde der Schüler mit hohem SES ($M = 4.55$, $SD = 0.89$) positiver beurteilt als der Schüler mit niedrigem SES ($M = 3.96$, $SD = 0.87$) und mit Sozialhilfebezug ($M = 3.68$, $SD = 1.02$), ohne Unterschied zwischen letzteren beiden. Keine signifikanten Unterschiede wurden beim Engagement oder der Störungstendenz festgestellt.

Die Ergebnisse liefern mögliche Hinweise darauf, dass der Sozialhilfebezug stärker mit negativeren Stereotypen behaftet ist als ein niedriger SES (Yendell et al., 2023). Insbesondere ein Sozialhilfebezug der Eltern kann bei Schüler*innen mit ungünstigem Arbeitsverhalten durch Aktivierung negativer Stereotype zu deutlich negativ verzerrten Beurteilungen schulischer Leistungen führen. Die Ergebnisse unterstreichen die Notwendigkeit einer differenzierten Betrachtung des elterlichen SES im schulischen Kontext und weiterer Forschung zu möglichen Urteilsverzerrungen bei Lehrkräften im Hinblick auf Familien mit Sozialhilfebezug.

S-51: Interdisziplinäre Transferstrategien im Kompetenzverbund lernen:digital

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 133

Chair der Sitzung: Luisa Scherzinger

Chair der Sitzung: Annika Goeze

Interdisziplinäre Transferstrategien im Kompetenzverbund lernen:digital

Chair(s): **Luisa Scherzinger** (Eberhard Karls Universität Tübingen, Deutschland)

Diskutant*in(nen): **Annika Goeze** (Eberhard Karls Universität Tübingen, Deutschland)

Im Kompetenzverbund lernen:digital entstehen in über 200 Forschungs- und Entwicklungsprojekten evidenzbasierte Fort- und Weiterbildungen, Materialien sowie Konzepte für die Schul- und Unterrichtsentwicklung in einer Kultur der Digitalität. In dieser sind kulturelle und soziale Aktivitäten vermehrt durch komplexe Technologien geprägt (Stalder, 2016). Der Verbund gestaltet aktiv den Dialog zwischen Wissenschaft und Praxis und begegnet der Digitalität als gesellschaftliche Herausforderung mit vielseitigen Professionalisierungsangeboten für Lehrkräfte, Schulleitungen und das pädagogische Personal in ganz Deutschland. Der Verbund wird durch die Europäische Union – NextGenerationEU finanziert und durch das Bundesministerium für Bildung und Forschung gefördert.

Damit die Angebote nachhaltig zur Verfügung gestellt und auch die entstehenden Strukturen und Netzwerke langfristig erhalten werden können, bedarf es ausgereifter Transferstrategien – sowohl für den Kompetenzverbund selbst als auch für die verschiedenen Projektverbünde. Dabei müssen die heterogenen Perspektiven der Transferbeteiligten (Kerres et al., 2022) und die häufig fehlende Infrastruktur für die Erkenntnisimplementation in der Praxis (OECD, 2022) berücksichtigt werden. In diesem Symposium werden ineinandergreifende materiale, personale und systemische Transferansätze von der Transferstelle als Schnittstelle zwischen Forschung, Lehrkräftebildung und Schule sowie von zwei Projektverbünden unterschiedlicher Fachrichtungen aufgezeigt. Die folgenden Fragen werden dabei sowohl datenbasiert wie auch konzeptionell beantwortet:

- Welche (nachhaltige) Transferstrategie verfolgen wir?
- Wie kann eine geeignete Transferstrategie zu einem gelungenen Wissenschafts-Praxis-Transfer beitragen?
- Wie können verschiedene Akteur:innen des Bildungssystems gewinnbringend adressiert werden?
- Welche Erkenntnisse und Implikationen im Hinblick auf Transfer lassen sich zum aktuellen Zeitpunkt aus lernen:digital ableiten?

Trotz der unterschiedlichen Ausgestaltung der Transferprozesse eint die vier Beiträge ein gemeinsames Grundverständnis von Transfer als ko-konstruktiver, kontext- und bedarfsbezogener sowie evidenzbasierter Prozess der (gemeinsamen) Wissensgenerierung von Bildungsforschung, Bildungspolitik, Bildungsadministration und Bildungspraxis, in dem die jeweiligen Handlungslogiken, Perspektiven und Expertisen transparent und reflexiv aufeinander bezogen werden (Bieber et al., 2018; Gräsel, 2010; Kerres et al., 2022; Prediger, 2022; Schmiedl, 2022). Transfer bietet daher einen Raum für zielgerichtete und konkrete Zusammenarbeit. Im Detail umfasst das Symposium folgende Transferstrategien und deren Diskussion.

Die **Transferstrategie der Vernetzungs- und Transferstelle** umfasst aufbereitete Forschungssynthesen, Praxishandreichungen, Multiplikator:innenfortbildungen und ko-konstruktive Kooperationen mit Landesinstituten. Die **Broker:innen** der Transferstelle haben durch ihre Vernetzungs- und Vermittlungsangebote sowohl eine Schlüsselrolle beim Wissenstransfer innerhalb der Wissenschaft als auch zwischen Bildungsforschung und Schulpraxis. Damit wirken sie der unzureichenden Verknüpfung zwischen den Domänen Wissenschaft und Praxis (Farley-Ripple et al., 2018; Munk et al., 2023) aktiv entgegen.

Durch die Erforschung beteiligter Stakeholder sowie förderlicher und hinderlicher Bedingungen für Transfer wird im Arbeitsbereich **Verständnis von Transfer der Transferstelle** ein besseres Verständnis von Transferprozessen in der digitalisierungsbezogenen Lehrkräftebildung und Schulentwicklung anvisiert und an der Entwicklung einer systemischen Transferstrategie mitgewirkt.

Der **Projektverbund WÖRLD** fokussiert auf die Entwicklungs- und Implementationsprozesse digitalisierungsbezogener Fortbildungsangebote und Unterrichtsmaterialien für Lehrkräfte im allgemeinbildenden und beruflichen Schulwesen im Bereich Wirtschaft. Durch die evaluative Begleitung der Transferstrategien und -formate der Teilprojekte, werden Merkmale und Bedingungen für die Konstituierung von Transfer-House-Strukturen entwickelt und diskutiert.

Landesinstitute und Qualitätseinrichtungen an der Schnittstelle zwischen Wissenschaft und Praxis sind ein wichtiger Baustein im Transferprozess. Der **Projektverbund DigiProMIN** entwickelt eine mehrdimensionalen Transfer- und Implementierungsstrategie zur systematischen Einbindung der Landesinstitute und Qualitätseinrichtungen bei der Erstellung und Durchführung von Lehrkräftefortbildungen.

Aus den Einzelbeiträgen lassen sich übergreifende Maßnahmen für nachhaltige Transferprozesse ableiten. Diese werden im Rahmen des Symposiums unter besonderer Berücksichtigung der genannten Fragestellungen ausführlich dargestellt und diskutiert.

Beiträge des Symposiums

Transferstrategie der Vernetzungs- und Transferstelle sowie der Broker:innen

Luisa Scherzinger¹, Ayşegül Serpemen², Till Schmäing³

¹Eberhard Karls Universität Tübingen, Deutschland, ²Universität Duisburg-Essen, Deutschland, ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Deutschland

Damit die im Kompetenzverbund entstehenden Fortbildungen und Produkte der Bildungspraxis nachhaltig und breitflächig zur Verfügung stehen, wurde als übergeordnete Struktur eine Vernetzungs- und Transferstelle mitgefördert (Lazarides et al., 2024). Diese umfasst die vier Handlungsfelder Vernetzung, Forschung, Transfer sowie Wissenschaftskommunikation. Folgend wird zunächst die übergreifende multidimensionale Transferstrategie vorgestellt, bevor im Detail auf das innovative Handlungsfeld Vernetzung eingegangen wird.

Neben den digitalisierungsbezogenen Materialien und Konzepten aus den Projektverbünden bereitet die Transferstelle im Rahmen der materialen Strategie relevante Forschungssynthesen und Handreichungen zur digital gestützten Schul- und Unterrichtsentwicklung adressatengerecht für die Praxis auf. Zudem werden Qualifizierungskonzepte für Multiplikator:innen und Leitlinien für digital gestützte Lehrkräftebildung erarbeitet. Zusätzlich zu den Lehrkräftefortbildungen der Projektverbünde umfasst

die personale Strategie die Weiterbildung von Multiplikator:innen. Zuletzt werden über die systemische Strategie die Landesinstitute, Qualitätseinrichtungen, Schulverwaltungen und die Bildungspolitik zum Dialog und Austausch über eine nachhaltige und konstruktive Kultur der Zusammenarbeit angeregt. Beispielsweise haben alle Bundesländer einer Nutzung der gemeinsamen Plattform ComPLeTT (Kiesler & Schiffner, 2022) als fortwährendem Kanal für die Verbreitung der Fort- und Weiterbildungen zugestimmt.

Bei dieser übergreifenden Transferstrategie nehmen die Broker:innen des Handlungsfelds Vernetzung eine Schlüsselrolle ein. Grundsätzlich sind Wissensbroker:innen intermediäre Akteure, die gezielt Wissen und Informationen sammeln, für unterschiedliche Personen- und Interessengruppen aufbereiten und zur Verfügung stellen (Meyer, 2010). In anderen Ländern wie Großbritannien, Canada oder den Niederlanden wird das Brokering im Bildungsbereich bereits genutzt, um wissenschaftliche Evidenz für die Politik und die Praxis sichtbar und nutzbar zu machen (Wollscheid et al., 2019). Für die deutsche Bildungslandschaft stellt der Einsatz der lernen:digital Broker:innen ein Novum dar.

Abgesehen von dem stetigen Kontakt zu den Projektverbünden, identifizierten die lernen:digital Broker:innen sowohl Alleinstellungsmerkmale als auch Schnittmengen im Hinblick auf Themenschwerpunkte und Transfermaßnahmen der involvierten Forschungs- und Entwicklungsvorhaben. Aufbauend auf diesen Synergien übernehmen die Broker:innen vielfältige Vernetzungs- und Vermittlungsaufgaben, indem sie Expertise bündeln, themenübergreifende Publikationen und Handreichungen anregen und mitgestalten. Zum Beispiel wird verbundübergreifend an Leitlinien für digital gestützte inklusive Materialien und Fortbildungen gearbeitet. Weiterhin organisieren die Broker:innen für die Projektverbünde Impulse, Workshops, Austauschformate und Peer-Feedbacks zu den Produkten. Dadurch sollen sowohl vorhandene Konzepte und Expertise zwischen unterschiedlichen Akteursgruppen gestreut als auch langfristige wissenschaftliche Partnerschaften aufgebaut und vertieft werden. Darüber hinaus ermöglichen die Broker:innen durch das Angebot von wissenschaftlichen Tagungen, Landesinstitutsbesuchen oder Regionalmessen auch den aktiven Dialog und die ko-konstruktive Zusammenarbeit mit den Landesinstituten. Schließlich sind die Einbeziehung der vielfältigen Perspektiven (Kerres et al., 2022) und der Aufbau interpersonaler Beziehungen essenziell für den Erfolg der Wissensvermittlung (Rycroft-Smith, 2022) und damit auch für die Transferstrategie als Ganzes.

Vonseiten der Broker:innen sind die Auswertungsergebnisse der angebotenen Vernetzungsformate, die Fortsetzung der Kommunikations- und Kooperationsnetzwerke, die deutschlandweite und nachhaltige Implementation und Nutzung der entstehenden Produkte sowie eine Anerkennung und Wertschätzung der Broker:innenrolle für den Bildungsbereich maßgebliche Kriterien für gelungenen Transfer. Durch eine Endevaluation des Gesamtprojekts werden auch Aussagen dazu getroffen werden können, ob die Förderung von Kollaborationen und die Lockerung von institutionellen Grenzen durch die Transferstelle erfolgreich war. Dies wäre ein höchstrelevantes Ergebnis, denn bisher gibt es wenig Beispiele für den Erfolg des Brokerings in diesen Bereichen (Chew et al., 2022).

Im Beitrag wird die Transferstrategie mit allen Angeboten der Transferstelle und der Broker:innen, deren konzeptioneller Hintergrund sowie erste deskriptive Auswertungen des Veranstaltungsangebots und dessen Reichweite in der Bildungspraxis vorgestellt. Außerdem werden vorläufige Erkenntnisse der Transferstelle aufgezeigt, um Folgeprojekten Anhaltspunkte für einen optimierten Organisationsablauf zu geben. Exemplarisch hat sich die zeitlich vor- und nachgelagerte Implementierung der Transferstelle und der Broker:innen als hilfreich für die Projektverbünde erwiesen. Allerdings gibt es durch die sowieso schon ausgelastete Zielgruppe der Lehrkräfte sowie teilweise unterschiedliche Haltungen der Projektverbünde und Landesinstitute auch Herausforderungen zu meistern. Die personelle und finanzielle Einbindung der Landesinstitute ist daher eine unverzichtbare Voraussetzung für kommende Projekte.

Ko-konstruktive Transferprozesse gestalten. Erste Erkenntnisse aus der Evaluation eines neuen Veranstaltungsformats der lernen:digital Transferstelle

Linda von Sobbe¹, Claire Förster², Vanessa Pieper³, Torben Bennink⁴, Therese Gesswein⁴, Hannes Schröter¹

¹Deutsches Institut für Erwachsenenbildung (DIE), Deutschland, ²Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Deutschland, ³Bergische Universität Wuppertal, Deutschland, ⁴Universität Potsdam, Deutschland

Die im Kompetenzverbund lernen:digital entstehenden Produkte zielen darauf ab, die Lehrkräftebildung in Bezug auf das Thema Digitalisierung weiterzuentwickeln. Um den Transfer der im Kompetenzverbund entstehenden Produkte und Wissen zu gewährleisten, ist ein Austausch zwischen Bildungspraxis, -administration und dem Kompetenzverbund fest im Vorhaben und insbesondere in den Aufgabenbereichen der lernen:digital Transferstelle verankert. Zu diesem Zweck werden vom Forum Bildung Digitalisierung, dem zivilgesellschaftlichen Projektpartner der lernen:digital Transferstelle, neue Veranstaltungsformate entwickelt, die einen kontinuierlichen Austausch anregen sollen.

Eines dieser neu entwickelten Veranstaltungsformate ist die „Roadshow“. Diese zielt darauf ab, ko-konstruktive Aushandlungsprozesse zwischen den Stakeholdern zu erzeugen (Bieber et al., 2018). Es handelt sich dabei um ein 3,5 bis 6-stündiges Workshopformat, das einmal in jedem der 16 Bundesländer mit jeweils ca. 15 Teilnehmenden stattfinden soll. Bei den Roadshows reisen Vertreter:innen des Kompetenzverbund lernen:digital in die Landesinstitute bzw. Qualitätsagenturen der einzelnen Bundesländer, um die Forschungsvorhaben von lernen:digital vorzustellen und mit den Vertreter:innen der Landesinstitute bzw. Qualitätsagenturen in Austausch zu kommen, einen Dialog auf Augenhöhe zu führen und somit die Basis für eine erfolgreiche und nachhaltige Kooperation zu schaffen. Die mit diesem ko-konstruktiven Aushandlungsprozess verbundene Nutzung der Wissensbestände aus allen beteiligten Bezugssystemen soll zum Transfererfolg beitragen (Bremm et al., 2018; Schuster et al., 2024).

Im vorliegenden Beitrag soll anhand der Evaluationsdaten der Roadshows diskutiert werden, inwieweit das Erzeugen ko-konstruktiver Prozesse im Rahmen des Veranstaltungsformats gelingt. Ko-Konstruktion wird hierbei als Prozess der Wissensgenerierung verstanden, für den die Schaffung sozialer Gelingensbedingungen eine zentrale Voraussetzung ist und dessen Ergebnis zum Beispiel neue Erkenntnisse oder Veränderungen von Handlungen, Haltungen oder Einstellungen sein können (Autor:innengruppe Metavorhaben Digitalisierung im Bildungsbereich, in Vorb.). Die Evaluation im Mixed-Method-Design beinhaltet eine Online-Befragung der Veranstaltungsteilnehmenden und eine teilnehmende Beobachtungen der Veranstaltungen. Mithilfe des Online-Fragebogens werden soziale Gelingensbedingungen wie z. B. Kommunikation auf Augenhöhe oder das Veranstaltungsklima in den Roadshows aus Teilnehmendensicht erhoben. Auch mögliche Outcomes eines ko-konstruktiven Prozesses, wie der wahrgenommene Mehrwert der Veranstaltung für die Teilnehmenden oder die Zufriedenheit mit den Diskussionsergebnissen, werden im Fragebogen erfragt. Um das Verständnis ko-konstruktiver Prozesse zu vertiefen, wird der Zusammenhang zwischen selbsteingeschätzten sozialen Gelingensbedingungen und wahrgenommenen Outcomes der Roadshows untersucht.

Erkenntnisse aus der Fragebogenauswertung werden durch qualitative Daten einer teilnehmenden Beobachtung komplementiert, die an ausgewählten Roadshow-Terminen zusätzlich zur Fragebogenerhebung erfolgt. Dabei liegt der Fokus auf den während der Veranstaltung stattfindenden Interaktionsprozessen zwischen den Teilnehmenden, sowie der Rolle des Forum Bildung Digitalisierung als zivilgesellschaftlichem Mediator zwischen Bildungspraxis bzw. -administration und Bildungsforschung. Es wird empirisch hinterfragt, wie Gestaltungsaspekte der Roadshows die Kommunikation zwischen Teilnehmenden als soziale Gelingensbedingung für Ko-Konstruktion beeinflussen können. Hierbei wird zudem das Wechselspiel zwischen Gestaltungsmacht der Eingeladenen und der Organisierenden beleuchtet.

Da die Roadshows ein iteratives Veranstaltungsformat darstellen, das von Juni 2024 bis Mitte 2025 durch das Forum Bildung Digitalisierung durchgeführt wird, werden zum Tagungszeitpunkt voraussichtlich Daten aus der Hälfte der Roadshows vorliegen. Die auf deren Auswertung beruhenden Untersuchungsergebnisse sollen dabei helfen, zukünftige ko-konstruktive Transferprozesse zwischen Bildungspraxis, -administration und Bildungsforschung effektiver zu gestalten. Weiterhin können sie einen Beitrag zur Operationalisierung der Beforschung von Transferprozessen leisten.

Konstituierung von Transfer House Strukturen durch (Programm-)Evaluation multizentrischer, multiperspektivischer Entwicklungs- und Implementationsarbeit im Verbundprojekt WÖRLD

Sina Schadow, Marian Thiel de Gafenco, Jens Klusmeyer

Universität Kassel, Deutschland

Zur Überwindung der „Lücke“ zwischen Bildungsforschung und Bildungspraxis (Vanderlinde & van Braak, 2010) werden Konzepte zur Intensivierung und nachhaltigen Verzahnung beider Bereiche durch wirksame Strukturen des Wissenschafts-Praxis-Transfers diskutiert.

Zu diesen gehören unter anderem die bereits existierenden Clearing und Leading Houses, aber auch die bisher nur skizzenhaft vorliegenden Entwürfe zur Etablierung sogenannter Transfer Houses (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz, 2022). Ausgehend von einem weitem Transferverständnis, welches sowohl die Übertragung als auch die Verbreitung evidenzbasierter (Bildungs-)Innovationen in unterschiedliche Praxisfelder einschließt (Bosse et al., 2019; Gräsel, 2010; Prenzel, 2010), finden Forschungs- und Entwicklungsaktivitäten zunehmend in einem gleichberechtigten, teils ko-kreativen Zusammenspiel verschiedener Akteure und Akteurskonstellationen der Lehrer:innenbildung statt. Dies spiegelt sich unter anderem im Diskurs zu Transfermodellen (Becheikh et al., 2010; OECD, 2022; Schmiedl, 2022) als Ausgangspunkt der Planung, Durchführung und Evaluation strategisch agierender Akteursnetzwerke der Lehrer:innenbildung wider, in dem ein Trend zu Transformations- und Relationsmodellen zu erkennen ist, die sich besonders für diese Entwicklungsaktivitäten zu eignen scheinen.

Im Verbundprojekt WÖRLD (Wirtschaftspädagogik und Ökonomische Bildung: Lehrkräftebildung und Unterricht digital) werden digitalisierungsbezogene Fortbildungsangebote und Unterrichtsmaterialien für Lehrpersonen des allgemeinbildenden, als auch des beruflichen Schulwesens entwickelt. Die 14 Teilprojekte aus sieben Bundesländern, die gemeinsam mit der Verbundleitung an der Konstituierung eines WÖRLD-House betitelten Transfer House Konzeptes mitwirken, betreiben auf Basis eigener Vorarbeiten die multiperspektivische Entwicklungs- und Implementationsarbeit in ihren spezifischen Kontexten. Sie nutzen hierfür den methodischen Ausgestaltungsspielraum des Design-Based Research-Ansatzes (Euler, 2014; Feulner, 2020; The Design-Based Research Collective, 2003) zur Erbringung der Innovationen und Produkte.

In diesem Beitrag werden die Ergebnisse der ersten, internen Evaluation des Verbundes vorgestellt und vor diesem Hintergrund die theoretischen Annahmen zur Konstituierung einer Transfer House Struktur in WÖRLD reflektiert. Hierbei wird mit Fokus auf die Prozessebene der Akteurskonstellationen und Aktivitäten der Teilprojekte der Frage nachgegangen, inwiefern die gewählten Transferstrategien und -formate der Konsortialpartner dazu beitragen, einen Innovations- und Produkttransfer für und mit den Zielgruppen der Entwicklungsforschung zu erreichen.

Zur Beantwortung dieser Frage werden ausgehend vom Ansatz der Cluster Evaluation (Haubrich, 2009; Koch, 2011; Worthen & Schmitz, 1997) die iterativen Entwicklungsprozesse durch problemzentrierte Gruppeninterviews (Witzel & Reiter, 2012) der Teilprojektleitungen (N=16) und -beschäftigten (N=20) in drei Erhebungswellen innerhalb der Projektlaufzeit (06/2023 – 12/2025) begleitet, die qualitativ mittels eines inhaltsanalytischen Vorgehens in Anlehnung an Kuckartz (2012) ausgewertet werden.

Im ersten Interview (1. Quartal 2024) wurden zunächst Ausgangssituation, Einordnung der Teilprojekte in die Verbundstruktur sowie Vernetzungs- und Kooperationsaktivitäten (1), Zielgruppen, Innovationsziele und Projektoutcomes (2) sowie Transferstrategie und Herausforderungen (3) als teilprojektspezifische Hauptkategorien angelegt.

(1) Als herausfordernd für die Forschungs- und Entwicklungsarbeit der Teilprojekte stellte sich die komplexe Struktur und die Verantwortlichkeiten in lernen:digital dar, die von Problemen bei der Besetzung offener Stellen aufgrund des kurzfristigen Projektstarts begleitet wurden. Heterogene „Standorttraditionen“ und daraus erwachsenen Kooperationsmöglichkeiten mit der zweiten und dritten Phase treffen auf ohnehin im Verbund als schwierig bezeichnete, föderale Struktur im Bildungssystem, die die Koordination im Transfer House erschwert.

(2) Als Innovationsziele setzen sich die Teilprojekte die Einbindung/Verstetigung der Projekthinhalte in den jeweiligen Studienstandorten, die (Weiter)Entwicklung fachdidaktischer Konzepte und die Professionalisierung von (angehenden) Lehrpersonen im Bereich der Digitalisierung. Dies ergibt sich aus dem iterativen Vorgehen der Entwicklungsarbeit, in der die Konsortialpartner ihre Pilotierungen häufig auf ihre hochschulischen Lehrveranstaltungen begrenzen.

(3) Induktiv wurden Subkategorien für die Beschreibung der Transferstrategien (stufenweises, phasenübergreifendes Vorgehen; paralleles, phasenverbindendes Vorgehen; direktes, phasenspezifisches Vorgehen) entwickelt. Die Teilprojekte von WÖRLD gehen in der Planung ihrer Entwicklungs- und Implementationsprozesse primär stufenweise und phasenübergreifend vor, welches durch sukzessives Vorarbeiten zeitaufwändiger als alternative Vorgehen sein kann, jedoch die Berücksichtigung phasenspezifischer Bedürfnisse und Präferenzen stärker beleuchtet.

Diese Ergebnisse fließen in die Konkretisierung der Transfer House Strukturen ein, die im Beitrag weiter diskutiert werden.

Brücke zwischen Wissenschaft und Praxis: Die Einbindung der Landesinstitute und Qualitätseinrichtungen in den Transferprozess am Beispiel von DigiProMIN

Simon Munk¹, Tanja Mutschler², Andreas Borowski², Doris Holzberger¹

¹Technische Universität München, Deutschland, ²Universität Potsdam, Deutschland

Die Digitalisierung des Bildungssystems und die Professionalisierung der Lehrkräfte im Umgang mit digitalen Medien gelten als zentrale Herausforderungen für Schule, Lehrkräfte und Unterricht (Blossfeld et al., 2018; Kultusministerkonferenz, 2021). Bei der

Bewältigung dieser Aufgabe kann die Bildungsforschung Evidenz für den erfolgreichen Einsatz digitaler Medien liefern (z. B. Hillmayr et al., 2017; Hillmayr et al., 2020). Trotz des Anspruchs der Bildungsforschung, Implikationen für die Praxis zu liefern (Gräsel, 2010), finden wissenschaftliche Erkenntnisse nicht immer den Weg ins Klassenzimmer (Broekkamp & van Hout-Wolters, 2007; Joyce & Cartwright, 2020). Die Gründe für diese Lücke zwischen Wissenschaft und Praxis (researcher-practitioner-gap; Farley-Ripple et al., 2018; Munk et al., 2023; Vanderlinde & van Braak, 2010) sind vielfältig: Zeitmangel der Akteure aus Praxis und Wissenschaft (Ratcliffe et al., 2005; Thomm et al., 2021), Sprachunterschiede (Vanderlinde & van Braak, 2010), mangelndes gegenseitiges Verständnis und daraus resultierende unterschiedliche Einstellungen gegenüber der jeweils anderen Seite (Gräsel, 2010; Thomm et al., 2021).

Beim Abbau dieser Hürden können Landesinstitute und Qualitätseinrichtungen eine zentrale Rolle spielen. Sie verstehen sich als „ebenso praxisnaher wie wissenschaftsnaher Mittler zwischen den an Transferprozessen Beteiligten“ (Bieber et al., 2018, S. 5). Der Projektverbund DigiProMIN in lernen:digital entwickelt daher eine Strategie zur Einbindung der Landesinstitute und Qualitätseinrichtungen in die Projektarbeit.

DigiProMIN erstellt Fortbildungen zur digitalisierungsbezogenen Professionalisierung von MIN-Lehrkräften. Neben der Entwicklung der Fortbildungen steht die Entwicklung einer mehrdimensionalen Transfer- und Implementierungsstrategie im Vordergrund. Ziel dieses Beitrags ist es, diese Transferstrategie des Projektverbundes DigiProMIN im Hinblick auf die Einbindung der Landesinstitute und Qualitätseinrichtungen vorzustellen. Dabei stehen folgende Fragen im Vordergrund:

1. Welche Strategie verfolgt DigiProMIN, um Landesinstitute und Qualitätseinrichtungen in den Transferprozess einzubinden?
2. Welche Erfolgsfaktoren für zukünftige Projekte zeigen sich bei der Umsetzung dieser Strategie?

Hinsichtlich der Strategie im Umgang mit den Landesinstituten und Qualitätseinrichtungen verfolgt DigiProMIN unterschiedliche Ansätze. Die Kommunikation zwischen dem Projektverbund und den Landesinstituten und Qualitätseinrichtungen findet vorrangig über sogenannte Regionalkoordinator:innen statt (ein:e Projektmitarbeiter:in in jedem teilnehmenden Bundesland). Damit haben die Landesinstitute konkrete Ansprechpartner:innen, die mit den föderalen Besonderheiten (z. B. Plattformen oder Rahmenbedingungen für Fortbildungen wie Dauer oder Ort) vertraut sind. Auf Seiten der Landesinstitute findet die Kooperation gemäß der hierarchischen Struktur der Landesinstitute auf unterschiedlichen Ebenen statt (Direktor:innen, Abteilungsleiter:innen, Referent:innen). Durch eine Vielzahl von Veranstaltungen (z. B. Interviews mit den Landesinstituten, fachliche Abstimmungsgespräche, Abstimmung der Landesinstitutsleitungen mit der Leitung von lernen:digital, Vor-Ort-Präsentationen der Fortbildungen) können die Bedarfe und Wünsche der Landesinstitute ermittelt werden.

Die Erfahrungen bei der Umsetzung dieser Strategie zeigen insbesondere zwei Erfolgsfaktoren auf: Erstens ist die Einbindung der Landesinstitute während der gesamten Projektlaufzeit (und nicht nur am Ende) unabdingbar, wenn die Fortbildungen den Bedürfnissen der Praxis entsprechen sollen (Gräsel & Parchmann, 2004). Nur durch die Einbindung der Perspektive der Landesinstitute in den Erstellungsprozess kann sichergestellt werden, dass sie mit den vor Ort üblichen Rahmenbedingungen und Plattformen kompatibel sind. Zweitens erweist sich ein regelmäßiger Kontakt über feste Ansprechpersonen als gewinnbringend. Etablierte Gesprächskanäle sichern hier eine vertrauensvolle Zusammenarbeit auf Augenhöhe.

Insgesamt erweisen sich die Landesinstitute und Qualitätseinrichtungen und die institutsübergreifende Zusammenarbeit der Akteur:innen der Lehrerbildung als wichtige Bausteine für einen erfolgreichen Transferprozess (van Ackeren et al., 2019). Durch die Stärkung der Erfolgsfaktoren stellen sie eine wichtige Brücke zwischen Wissenschaft und Praxis im Projektprozess dar.

S-52: Stimmgesundheit bei Lehrkräften: Innovative Ansätze zur Messung und Prävention

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: O 129

Chair der Sitzung: Christian Gegner

Chair der Sitzung: Uta Klusmann

Stimmgesundheit bei Lehrkräften: Innovative Ansätze zur Messung und Prävention

Chair(s): **Christian Gegner** (Universität Regensburg)

Diskutant*in(nen): **Uta Klusmann** (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik)

Die Gesundheit von Lehrkräften ist von herausragender Bedeutung. Nur gesunde und motivierte Lehrkräfte sind imstande, kontinuierlich guten Unterricht zu gewährleisten. Hierfür ist im Besonderen eine gesunde und tragfähige Stimme für Lehrkräfte als „professional voice users“ essentiell. Im Rahmen dieses Symposium werden (innovative) Ansätze zur Messung von Stimmgesundheit bei Lehrkräften vorgestellt und Ergebnisse diskutiert, wie diese gefördert werden kann. Die vier Vorträge basieren auf unterschiedlichen Forschungsdesigns und Datenmodellen.

Im ersten Vortrag werden die gesundheitsbezogenen Belastungen von Lehrkräften sowie deren Konsequenzen für die Arbeitsfähigkeit beleuchtet. Zu diesem Zweck wurde seitens des Arbeitsmedizinischen Instituts für Schulen in Bayern (AMIS-Bayern) ein Befragungsinstrument entwickelt, welches an 27 Schulen mit insgesamt 851 Lehrkräften pilotiert wurde. Die Ergebnisse ordnen auch die von Lehrkräften empfundene Bedeutung der Stimmgesundheit im Vergleich zu anderen Belastungsfaktoren ein.

Der zweite Vortrag präsentiert die Ergebnisse einer erfolgreichen Methode zur Prävention von Stimmstörungen im Lehrberuf an der Universität Regensburg. Im Rahmen des Regensburger Stimmtrainings wurde ein App-basiertes Stimmtraining entwickelt und mit 404 Lehrkräften sowohl durch Selbsteinschätzungen wie auch objektive Messmethoden evaluiert. Die Ergebnisse zeigen, dass eine Verbesserung der Stimmgesundheit möglich ist, werfen jedoch auch die Frage nach einer geeigneten Operationalisierung von Stimmgesundheit auf.

Im dritten Vortrag werden die Ergebnisse des ambulanten Assessments im Kontext des Regensburger Stimmtrainings präsentiert. Das Verfahren basiert auf einer wiederholten Echtzeiterfassung der stimmlichen Belastung von Lehrkräften zu drei Messzeitpunkten mit jeweils mehreren Abfragen pro Messzeitpunkt. Dadurch lassen sich präzisere Daten zur Wirkung der Intervention gewinnen, was einen innovativen Ansatz für die Evaluierung von Interventionen darstellt und die vorherigen Ergebnisse differenzierter beleuchtet.

Der letzte Vortrag thematisiert die Entwicklung eines Langzeitmonitorings der Stimme am Fraunhofer-Institut für Digitale Medientechnologie (IDMT) mit fortlaufender Messung der Stimmgesundheit im Unterricht. Das komplexe Verfahren ermöglicht damit die Dokumentation stimmlicher Veränderungen innerhalb eines Tages.

Beiträge des Symposiums

Belastungen am Arbeitsort Schule beurteilen – eine Pilotstudie zur Gefährdungsbeurteilung psychischer Belastung an bayerischen Schulen

Sarah Stöhr, Victoria Heumann, Meike Sons, Daniel Ossenschmidt

Arbeitsmedizinisches Institut für Schulen in Bayern

Theoretischer Hintergrund

Aktuelle Studien zeigen, dass Lehrkräfte mit unterschiedlichen Belastungen am Arbeitsort konfrontiert sind (Wischnitzki et al., 2020; Robert Bosch Stiftung 2024). Das spiegelt sich auch in aktuellen Statistiken zum vorzeitigen Ausscheiden aus dem Berufsleben wider. Psychische Erkrankungen sind z. B. laut Auswertung der amtsärztlichen Untersuchungen des Gesundheitsamtes Bremen eine der Hauptursachen für Dienstunfähigkeit bei Lehrkräften (Spieß, 2023).

Im Rahmen des § 5 des Arbeitsschutzgesetzes wird die Berücksichtigung psychischer Belastungen bei der Gefährdungsbeurteilung (GBU) explizit gefordert. Gemäß der internationalen Norm DIN EN ISO 10075-1 handelt es sich bei psychischen Belastungen um alle erfassbaren Einflüsse am Arbeitsplatz, die von außen auf die Beschäftigten einwirken, also Reaktionen im Denken, Fühlen, Wahrnehmen, Erinnern usw. hervorrufen. Ziel der GBU psychischer Belastung ist es, Arbeitsbedingungen zu verbessern, um die Gesundheit und das Wohlbefinden zu erhalten und zu fördern. Vor diesem Hintergrund sieht die GBU psychischer Belastung mehrere Prozessschritte vor, die von der Ermittlung von Belastungen über die Ableitung und Umsetzung von Maßnahmen bis zur Wirksamkeitsevaluation reichen. Trotz der gesetzlichen Verpflichtung fand die GBU psychischer Belastung an bayerischen Schulen bisher kaum Beachtung, da adäquate Prozesse und Instrumente fehlten.

Fragestellung und Methode

Das Arbeitsmedizinische Institut für Schulen in Bayern (AMIS-Bayern), beauftragt von der bayerischen Staatsregierung, hat einen Prozess zur Erfassung, Beurteilung und Veränderung von Arbeitsbedingungen entwickelt und pilotiert. Dieser Prozess beinhaltet unter anderem ein Befragungsinstrument zur Erhebung von Belastungen in fünf Merkmalsbereichen, die in den Leitlinien der Gemeinsamen Deutschen Arbeitsschutzstrategie (GDA, Arbeitsprogramm Psyche, 2022) festgelegt sind. Dabei handelt es sich um (1) die Arbeitsorganisation, (2) den Arbeitsinhalt, (3) die sozialen Beziehungen bei der Arbeit, (4) die Arbeitsmittel und Arbeitsumgebung sowie (5) neue Arbeitsformen (z. B. Digitalisierung und Inklusion). Das Instrument umfasst 127 Items, wurde als digitale Umfrage konzipiert und in einer Pilotierungsphase an 27 bayerischen Schulen verschiedener Schularten eingesetzt. Daraus ergab sich eine Stichprobe von N=851 Lehrkräften. Aufgrund der hohen datenschutzrechtlichen Anforderungen an Gesundheitsdaten wurden keine personenbezogenen Daten erhoben. Die Daten wurden inferenzstatistisch analysiert.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen zunächst wenige schulartspezifische Unterschiede hinsichtlich der Belastungsfaktoren in den fünf Merkmalsbereichen. Im Bereich „Arbeitsorganisation“ zeigen sich hohe Belastungen in den Subkategorien Arbeitszeit und Arbeitsabläufe. Im Bereich „Arbeitsinhalt“ deuten die Ergebnisse auf einen hohen Belastungswert in der emotionalen Inanspruchnahme hin. Dieser Wert ist insbesondere bei Grund- und Förderschullehrkräften ausgeprägt. Dieses Ergebnis spiegelt sich auch im Bereich „Soziale Beziehungen“ wider. Dort kann gezeigt werden, dass Verhaltensauffälligkeiten von Schülerinnen und Schülern mit hoher Belastung einhergeht. Im Bereich „Arbeitsmittel und Arbeitsumgebung“ werden Belastungen vor allem hinsichtlich mangelnder Rückzugs- und Arbeitsmöglichkeiten und ungünstiger akustischer Bedingungen genannt. In der

Subkategorie der körperlichen Belastungen sind die stimmlichen Herausforderungen besonders ausgeprägt. Neben generellen Belastungen können anhand der Ergebnisse auch Ressourcen identifiziert werden. Ein gutes Schulklima, Handlungsspielraum, Abwechslungsreichtum und gute soziale Beziehungen innerhalb des Kollegiums sowie zur Schulleitung werden als positiv wahrgenommen. Die Ergebnisse der Pilotierungsstudie decken sich mit dem Ergebnis, dass insbesondere psychische Arbeitsbelastungen zu einer möglichen Dienstunfähigkeit führen können.

Da die Ergebnisse hohe Belastungen im Bereich Stimmgesundheit zeigen, lässt sich der Bedarf für ein stimmliches Präventionsprogramm ableiten. Als ein Resultat wurde das Projekt „Regensburger Stimmtraining“ vom AMIS-Bayern beauftragt, dessen Ergebnisse in zwei weiteren Vorträgen dieses Symposiums vorgestellt werden.

Fit für den Klassenraum – das App-basierte Präventionsprogramm „Regensburger Stimmtraining“

Jonas Hauck, Christian Gegner, Marina Giglberger, Sven Hilbert
Universität Regensburg

Theoretischer Hintergrund

Wie in der GBU psychischer Belastung durch AMIS-Bayern festgestellt, sind Lehrkräfte bei ihrer Berufsausübung ständig hohen Anforderungen an ihren Stimm- und Sprechapparat ausgesetzt (Schneider-Stickler & Bigenzahn, 2013). Die Folge ist, dass die Prävalenz von Stimmstörungen in der Berufsgruppe der Lehrkräfte im Vergleich zu Nicht-Lehrkräften deutlich erhöht ist (Van Houtte et al., 2011). Diese führen bei Lehrkräften zu erheblichen beruflichen Einschränkungen (Behlau et al., 2012), hohem Leidensdruck (Sapir et al., 1993) und können sogar den Wunsch nach einer beruflichen Neuorientierung auslösen (Behlau et al., 2012). Zusätzlich beeinträchtigen Stimmstörungen die zwischenmenschliche Kommunikation und können die Motivation und Leistung von Schülerinnen und Schülern negativ beeinflussen (Voigt-Zimmermann, 2017).

Lehrkräfte werden in den seltensten Fällen ausreichend auf ihren stimm- und sprechintensiven Beruf vorbereitet. Der Bedarf an Stimmpräventionsprogrammen ist daher immens, wie die Nachfrage an Stimmworkshops z. B. des AMIS-Bayern bestätigen. Um möglichst alle Lehrkräfte zu erreichen, ist ein niederschwelliges und flächendeckendes Angebot vonnöten. Im Forschungsprojekt „Regensburger Stimmtraining (ReSt)“ wurde ein App-basiertes Präventionsprogramm entwickelt, welches Selbstreflexions- und Trainingsmöglichkeiten zu den Bereichen Körper und Haltung, Atmung, Phonation, Artikulation und Stimmhygiene beinhaltet.

Fragestellung

Im Rahmen der App-Evaluation stellen sich folgende Forschungsfragen:

- (1) Lässt sich mit dem App-basierten Stimmtraining die Stimmgesundheit von Lehrkräften verbessern?
- (2) Lassen sich Unterschiede hinsichtlich des stimmlichen Selbstkonzepts, der selbstberichteten und der objektiv gemessenen Stimmgesundheit feststellen?
- (3) Existieren Nutzungsmuster, die eine größere Verbesserung nach sich ziehen als andere?

Methode

Zur Evaluation wurde ein längsschnittliches Kontrollgruppen-Design mit Eingangs- und Ausgangsuntersuchung bei bayerischen Lehrkräften (N=404) aller Schularten umgesetzt. Zu beiden Messzeitpunkten wurde Stimmgesundheit sowohl objektiv im Rahmen computergestützter Analysen der Stimme (Dysphonia Severity Index, Acoustic Voice Quality Index, Stimmfeldgröße) als auch durch die Fragebögen Voice Handicap Index (VHI-12), Stimmprofil für Berufssprecher (SPBS) und Fragebogen zur Erfassung des Stimmlichen Selbstkonzepts (FESS) erhoben.

Zur Beurteilung der sieben-wöchigen Intervention wurde ein Radom-Intercept-Mixed-Model (Hilbert et al., 2017) geschätzt. Die Identifikation von Nutzungsprofilen wird mit Hilfe eines Clusterings von Markov-Ketten-Modellen (Stragier et al., 2019) erfolgen. Die Auswirkungen der verschiedenen Profile werden wiederum mittels Mixed-Models geschätzt.

Ergebnisse und ihre Bedeutung

Die bisherigen Befunde bezüglich der z-standardisierten Interaktion zwischen Gruppenzugehörigkeit und Messzeitpunkt weisen auf eine positive Wirkung des app-basierten Stimmtrainings auf das stimmliche Selbstkonzept hin ($\beta=0.11$, $SE=0.05$, $p=.020$). Des Weiteren konnte ein positiver Einfluss auf die objektiv gemessene Stimmgesundheit festgestellt werden (DSI: $\beta=0.63$, $SE=0.09$, $p<.001$; Stimmfeldgröße: $\beta=0.47$, $SE=0.08$, $p<.001$). Im Gegensatz dazu konnte bei der selbstberichteten Stimmgesundheit (VHI, SPBS) kein Einfluss der Intervention festgestellt werden. Hinsichtlich der Einflussfaktoren zeigen sich ebenfalls deutliche Unterschiede zwischen objektiven und subjektiven Maßen. Während der Messzeitpunkt und die App-Nutzung einen signifikanten Einfluss auf die beiden objektiven Maße der gemessenen Stimmgesundheit aufweisen, lassen sich diese Effekte weder für den VHI noch für den SPBS nachweisen. Die Auswertung der Nutzungsanalysen ist zum gegenwärtigen Zeitpunkt noch nicht abgeschlossen, die Ergebnisse werden jedoch bis zum Beginn der Tagung vorliegen.

In der Gesamtschau lässt sich festhalten, dass das entwickelte app-basierte Präventionsprogramm aufgrund seiner zeit- und ortsunabhängigen Einsatzmöglichkeiten einen wertvollen Beitrag zur Stärkung der Stimmgesundheit bei Lehrkräften leisten kann. Insofern ist es für den Einsatz in der Lehrkräfteaus- und -weiterbildung geeignet. Überdies liefert das Forschungsprojekt detaillierte Erkenntnisse über die stimmlichen Herausforderungen bei Lehrkräften und leistet zudem einen Beitrag zur Validierung stimmdiagnostischer Messinstrumente.

Von Belastung zu Balance – Untersuchung der Effekte des Regensburger Stimmtrainings auf die wahrgenommene Stimmbelastung im Alltag

Marina Giglberger¹, Jonas Hauck¹, Christian Gegner¹, Brigitte M. Kudielka¹, Stefan Wüst¹, Richard Göllner²

¹Universität Regensburg, ²Universität Tübingen

Theoretischer Hintergrund

Im Hinblick auf die Genese von Stimmbeeinträchtigungen ist von einem Wechselspiel zwischen Arbeitsbelastungen (z.B. lautstarke Klasse) und Stimmreaktionen (z.B. Erhöhung der Sprechlautstärke) im Sinne einer reziproken Beziehung auszugehen (Hammer, 2023). Die Stimme ist im Tagesverlauf somit unterschiedlichen und sich ändernden Belastungen ausgesetzt. Gerade bei Lehrkräften scheint die Stimmanstrengung im Verlauf der Unterrichtszeit zuzunehmen (Kristiansen et al., 2014; Rantala et al., 2002). Entsprechend sollten Unterstützungsmaßnahmen Strategien bereitstellen, durch die in einer konkreten Situation Stimmbelastungen verringert sowie Stimmstörungen vermieden werden können. Im Rahmen des „Regensburger Stimmtrainings (ReSt)“ werden App-basiert solche Strategien vermittelt.

Vor diesem Hintergrund untersucht die vorliegende Studie mittels des Ambulanten Assessments (AA), inwiefern Lehrkräfte von der Nutzung der App im Verlauf ihres Schulalltags profitieren. Die Vorteile eines AA sind eine erhöhte ökologische Validität durch Messungen einer konkreten Situation im Alltag, eine höhere Reliabilität aufgrund der wiederholten Messungen und einer Reduzierung des retrospektiven Bias durch die Erfassung in der Situation (Conner & Mehl, 2015; Trull & Ebner-Priemer, 2020). Im Rahmen des ReSt-Projekts wurde neben klassischen Testverfahren (siehe vorheriger Vortrag) das AA benutzt, um die subjektive stimmliche Belastung und Störung der Stimme der Lehrkräfte im Alltag zu messen.

Fragestellung

Im Rahmen der App-Evaluation stellen sich folgende Forschungsfragen:

- (1) Lässt sich mit dem App-basierten Stimmtraining das wahrgenommene, momentane stimmliche Befinden von Lehrkräften verbessern?
- (2) Hat die App-basierte Intervention einen Einfluss auf die zeitliche Stabilität des wahrgenommenen, momentanen stimmlichen Befindens von Lehrkräften?

Methode

In einer Teilstichprobe des ReSt-Projekts ($n=79$) wurde ein längsschnittliches Interventions-Kontrollgruppen-Design mit drei Messzeitpunkten umgesetzt. Zu jedem Messzeitpunkt erhielten die Teilnehmenden 16 Abfragen zu ihrem momentanen, stimmlichen Befinden gemessen über die Variablen "Stimmstörung" und "Stimmbelastung". Die erste Messung erfolgte als Eingangsuntersuchung vor der App-Nutzung, die zweite fand in der Mitte des sieben-wöchigen Interventionszeitraums und der dritte als Ausgangsuntersuchung im Anschluss an die App-Nutzung statt. Die Kontrollgruppe durchlief dasselbe Studiendesign wie die Interventionsgruppe, allerdings wurde dieser Gruppe die App erst nach Abschluss der Studie zur Verfügung gestellt.

Zur Auswertung der Daten wurde ein 3-Ebenenregressionsmodell (Erhebungssituation genestet innerhalb der Messzeitpunkten und Messzeitpunkte innerhalb von Personen). Das Modell wurde als Längsschnittmodell mit autoregressiven Effekten spezifiziert, indem zur Vorhersage jeder Erhebungssituation jeweils die vorausgegangene Erhebungssituation als unabhängige Variable diente. Weiterhin wurden die Gruppe (d.h. Interventions- vs. Kontrollgruppe) und der Messzeitpunkt als unabhängige Variablen im Modell genutzt. Für die beiden Zielkriterien Stimmbelastung und Stimmstörung wurden getrennte Modelle gerechnet.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigten keinen Gruppenunterschied im zeitlichen Verlauf der Zielvariablen über die drei Messzeitpunkte hinweg. Lediglich zum dritten Messzeitpunkt ließ sich ein Trend hinsichtlich einer geringeren Ausprägung der Stimmbelastung in der Interventionsgruppe erkennen ($\beta=-0.15$, $SE=0.10$, $p=.139$).

Bezüglich der zeitlichen Stabilität der wahrgenommenen stimmlichen Belastung zeigte sich insgesamt, dass Angaben zur Stimmbelastung und Stimmstörung in einer Situation durch die jeweils vorausgegangene Situation bestimmt werden (Stimmbelastung: $\beta=0.32$, $SE=0.05$, $p<.001$; Stimmstörung: $\beta=0.35$, $SE=0.05$, $p<.001$). Zum dritten Messzeitpunkte konnte in der Kontrollgruppe eine Zunahme dieser Stabilität gefunden werden (Stimmbelastung: $\beta=0.22$, $SE=0.07$, $p=.001$; Stimmstörung: $\beta=0.26$, $SE=0.06$, $p<.001$). In der Interventionsgruppe hingegen war eine solche Zunahme des Zusammenhangs nicht zu erkennen (Stimmbelastung: $\beta=-0.22$, $SE=0.10$, $p=.029$; Stimmstörung: $\beta=-0.24$, $SE=0.10$, $p=.020$).

Die Ergebnisse unserer Studie legen einen präventiven Wert des Stimmtrainings nahe. Während die Kontrollgruppe im Verlauf der Untersuchung eine Zunahme der Stabilität der wahrgenommenen Stimmbelastung im Alltag aufweist, können Lehrkräfte der Interventionsgruppe diese möglicherweise mittels erlernter Strategien durchbrechen. Der fehlende Gruppenunterschied im generellen Verlauf des stimmlichen Befindens könnte durch die eingeschränkte Nutzungszeit von sieben Wochen bedingt sein. Da es sich um die subjektive Wahrnehmung des stimmlichen Befindens handelt, könnten längere Zeiträume für die Manifestation einer positiven Veränderung notwendig sein. Weiterführende Analysen, welche eine detaillierte Untersuchung der wahrgenommenen Stressoren berücksichtigen, können einen möglichen Beitrag zur Erklärung der Wirkmechanismen der Intervention leisten.

Der Weg zum ökologisch validen Langzeitmonitoring der Stimmbelastung im Alltag

Alina Ernst, Christian Rollwage

Fraunhofer-Institut für Digitale Medientechnologie IDMT

Theoretischer Hintergrund

Wie bereits im vorherigen Vortrag dargelegt, sind Stimmbelastungen im Allgemeinen mit alltäglichen Verhaltensweisen assoziiert und können daher in ihrem zeitlichen Verlauf erheblich variieren. Die Messung der stimmlichen Einschränkung einer Person zu einem einzelnen Messzeitpunkt weist somit keine hinlängliche ökologische Validität auf. Dies gilt sowohl für selbstberichtete Maße der Stimmgesundheit wie sie beispielsweise in Fragebögen erfasst werden, als auch für objektive Messverfahren, die im Rahmen von Stimmscreenings zum Einsatz kommen.

Daher wurden bereits seit den 1970er Jahren mehrere Versuche unternommen, Geräte für ein Langzeitmonitoring der Stimmbelastung zu entwickeln (vgl. Ohlsson et al., 1989). Diese Ansätze konnten den wissenschaftlichen Nutzen eines solchen Monitorings belegen, jedoch lieferten sie aufgrund der technischen Umsetzung lediglich unzureichende Daten für eine ganzheitliche Erfassung der Stimmgesundheit. Zu nennen ist in dem Zuge die Anfälligkeit der akustischen Stimmanalyse vor dem Hintergrund von hohem Umgebungslärm und minderer raumakustischer Qualität innerhalb der Stimm aufnehmen (Marsano-Cornejo et al., 2023; Bottalico et al., 2020). Bis heute existieren keine validen und datenschutzkonformen Instrumente für ein solches Langzeitmonitoring, obwohl der wissenschaftliche Nutzen für Interventionsprogramme als immens zu erachten ist. Dies belegen auch die Ergebnisse des Ambulanten Assessments des vorherigen Vortrags. Aus diesem Grund wird am Fraunhofer-Institut für Digitale Medientechnologie (IDMT) in der Arbeitsgruppe „Assistive Sprech- und Sprachanalyse“ ein Langzeitmonitoring für die Stimmgesundheit von Lehrkräften entwickelt. Die potenziellen Anwendungsfelder eines solchen Systems sind vielfältig. Diesbezüglich ist zunächst die Steigerung der Messgenauigkeit zur Erfassung der Stimmgesundheit im Rahmen wissenschaftlicher Studien zu nennen. Darüber hinaus ist auch eine eigenständige Nutzung des Systems durch die Lehrkräfte denkbar, wobei diese Rückmeldungen zum Einsatz der Stimme erhalten. Schließlich ist auch eine Anwendung im therapeutischen Bereich zur Verbesserung von Diagnose und Therapie denkbar.

Fragestellung

Im Rahmen der Entwicklung stellt sich die Frage, welche Indikatoren es für eine valide und datenschutzkonforme Messung der Stimmgesundheit im laufenden Unterricht gibt und wie sich daraus ableitend ein entsprechendes System hardware- und softwareseitig umsetzen lässt.

Methode

Bestehende Daten der im Rahmen des Regensburger Stimmtrainings aufgenommenen Stimmaufnahmen dienen der Entwicklung eines gegen Umweltfaktoren (Umgebungs­lärm, Raumakustik) robusten Stimmbewertungsalgorithmus auf Basis von Deep Neural Networks (DNNs). Die Aufnahmen und ihre Stimmqualitätsreferenzwerte werden zum Training eines DNN genutzt. Die Robustheit des DNN wird anschließend durch akustische Degradierung der Stimmaufnahmen durch Störeinflüsse evaluiert und optimiert.

Für das Proof-of-concept des Langzeitmonitorings müssen sowohl die Umsetzung des Datenschutzes als auch die Robustheit der Stimmanalyse im Feld evaluiert werden. Für die Evaluation und Erstellung eines Hardware- und Softwaredesigns ist eine Datenerhebung im Anwendungsfeld des Schulalltags geplant. Unterrichtseinheiten werden dabei von einer Aufnahmeeinheit bestehend aus Aufnahmegerät und Mikrofon aufgenommen. Die Sprachdaten werden hinsichtlich des aktiven, alleinigen Sprechanteils der Zielperson annotiert sowie hinsichtlich der auditiv-perzeptiven Stimmqualitätsbeurteilung der Stimme der Zielperson anhand einer logopädischen Fachkraft. Zusätzlich werden raumakustische Parameter extrahiert.

Anhand dieser Annotationen können die bestehenden und entwickelten Algorithmen zur Einhaltung des Datenschutzes und der robusten Stimmanalyse anwendungsfeldbasierter Aufnahmen als erstes Benchmark evaluiert werden. Die Hardwareanforderungen können angelehnt an die weitere notwendige Performanzsteigerung der Software im Zuge der Benchmarkergebnisse konzipiert werden.

Ergebnisse und ihre Bedeutung

Zum gegenwärtigen Zeitpunkt befindet sich die Entwicklung noch in einem laufenden Prozess, sodass kontinuierlich neue Ergebnisse generiert werden. Im Rahmen der GEBF sollen die zu diesem Zeitpunkt aktuellen Ergebnisse vorgestellt und diskutiert werden.

S-53: Motivationale Wirkung von Feedback: Welche Rolle spielen Heterogenitätsmerkmale der Schüler*innen?

Zeit: Dienstag, 28.01.2025: 13:15 - 15:00 · Ort: EO 162

Chair der Sitzung: Stefanie van Ophuysen

Chair der Sitzung: Gruehn Sabine

Chair der Sitzung: Lintorf Katrin

Chair der Sitzung: Susanne Narciss

Motivationale Wirkung von Feedback: Welche Rolle spielen Heterogenitätsmerkmale der Schüler*innen?

Chair(s): **Stefanie van Ophuysen** (Universität Münster, Deutschland), **Sabine Gruehn** (Universität Münster), **Katrin Lintorf** (Universität zu Köln)

Diskutant*in(nen): **Susanne Narciss** (TU Dresden)

Als Feedback gelten Informationen, die Lernende dabei unterstützen, die Diskrepanz zwischen Lernstand und Lernziel zu reduzieren (Hattie & Timperley, 2007; Narciss, 2008). Es kann gemäß des Feedback-Modells von Narciss (2020) auf kognitive, metakognitive und motivationale Wirkungen abzielen. Wie wirksam das Feedback hinsichtlich dieser Dimensionen ist, hängt einerseits von Feedbackmerkmalen (Funktion, Form und Inhalt) und andererseits von situativen und individuellen Bedingungen ab. Das geplante Symposium rückt die motivationale Wirkung in Abhängigkeit von schulisch relevanten Heterogenitätsdimensionen als Merkmale der Lernenden in den Mittelpunkt.

Zahlreiche Studien belegen förderliche Effekte von Feedback(merkmalen) auf diverse motivationale Konstrukte (z. B. Attributionen: Knickenberg, 2018; Selbstwirksamkeit: Hoya, 2019; intrinsische Motivation: Wollenschläger, Möller & Harms, 2011; time-on-task: Narciss, 2004), wenngleich je nach Umsetzung des Feedbacks auch negative Effekte auftreten können (Wisniewski, Zierer & Hattie, 2019). Die Wirksamkeit von Feedback wird typischerweise in Abhängigkeit verschiedener Feedbackmerkmale seltener jedoch in Abhängigkeit von *Lernendenmerkmalen* untersucht. In einem zunehmend inklusiven Schulsystem wird jedoch die Frage nach differenziellen Feedbackeffekten in Abhängigkeit von Heterogenitätsdimensionen salient. Diesbezügliche Befunde sind selten und beziehen sich zumeist auf ergebnisbezogenes, evaluatives Feedback. So belegen Koenka et al. (2021) einen negativen Effekt von evaluativem Feedback insbesondere für leistungsschwache Schüler*innen. Schwab et al. (2022) weisen für Jungen (vs. Mädchen) und für Kinder mit nicht-deutscher (vs. deutscher) Familiensprache höhere Korrelationen zwischen leistungsbezogenem Feedback und motivational relevanten Leistungsemotionen nach.

Die Beiträge des Symposiums überprüfen in unterschiedlichen Studiendesigns moderierende Effekte verschiedener Heterogenitätsdimensionen auf den Zusammenhang von Feedbackgestaltung und motivationalen Variablen im Unterricht in sprachlichen Fächern in verschiedenen Schulstufen.

Beitrag 1 untersucht mit Daten einer Fragebogenstudie mit Grundschüler*innen (N = 257) den Zusammenhang von dialogischem Feedback und Student Engagement. Dabei werden die Qualität der Schüler-Lehrkraft-Beziehung sowie die wahrgenommene Nützlichkeit des Feedbacks als Mediatoren und der Migrationshintergrund als Moderator modelliert.

Beitrag 2 stellt Befunde einer Interventionsstudie zum argumentativen Schreiben mit Schüler*innen an Gesamtschulen der Jahrgangsstufe 9 (N = 202) dar. Die Wirkung unterschiedlicher schriftlicher Feedbackvarianten auf Selbstwirksamkeit sowie Schreibangst wird analysiert, wobei die Leistungsstärke vor der Intervention als Moderator dient.

Beitrag 3 betrachtet in einer Längsschnittstudie mit Oberstufenschüler*innen (N = 696) den Zusammenhang verschiedener Feedbackmerkmale (Ergebnis-, Prozess-, Selbstregulations- und Dialogorientierung) mit nachfolgenden Erfolgsattributionen. Mittels Strukturgleichungsmodellen wird die Mediation über das Autonomie- und Kompetenzerleben überprüft, wobei ein Multigruppenvergleich auf geschlechtsspezifische Musterunterschiede testet.

Durch die Kombination von Beiträgen aus erziehungswissenschaftlicher, fachdidaktischer und psychologischer Perspektiven und die Diskussion durch eine ausgewiesene Expertin im Themenfeld Feedback trägt das Symposium zu einem genaueren Verständnis der Feedbackwirksamkeit und der Bedeutsamkeit einer adaptiven Feedbackpraxis bei.

Beiträge des Symposiums

Understanding associations between dialogic feedback and schoolwork engagement – the role of teacher-student relationship, feedback usefulness and students' migration background

McCullagh Laura¹, Gruehn Sabine¹, Lintorf Katrin², van Ophuysen Stefanie¹

¹Universität Münster, ²Universität zu Köln

Dialogic Feedback has been defined as "interactive exchanges [related to the quality of student work] in which interpretations are shared, meanings negotiated and expectations clarified" (Carless et al., 2011). Closely related to the concept of scaffolding, dialogic feedback is argued to be particularly effective in promoting learning because it requires the student to actively use and make sense of feedback, instead of merely "receiving" it (van der Kleij et al., 2019). Furthermore, dialogic feedback practices, when applied appropriately, are predictive of a trusting student-teacher relationship (Carless, 2012), which may have direct consequences for student achievement, engagement and motivation (Roorda et al., 2011).

Research explicitly examining dialogic feedback tends to focus on secondary to higher educational settings (Adie et al., 2018; van der Kleij, 2023; Winstone & Carless, 2019). However, even young children have a need for feedback to be negotiable and adaptive to their interests, rather than merely prescriptive (Hargreaves, 2011). Initial research concerning student participation in class dialogue indeed indicates some positive effects on learning, particularly when the conversation is scaffolded by the teacher (Bürgermeister et al., 2019). However, results are mixed, with Herrmann et al. (2021) recently finding such benefits only for classrooms with a high ratio of students from a migration background. Thus, the function and benefit of dialogic practices in the early years of schooling are far from settled, and effects may differ due to individual characteristics.

Focusing on text production in 3rd and 4th grade German classes, the present study examines the relationship between dialogic feedback practices and schoolwork engagement. Engagement is conceptualised as the multidimensional (emotional, cognitive, behavioural; Fredricks et al., 2004), "observable manifestation of motivation" (Skinner et al., 2009). Based on prior literature, it is hypothesized that dialogic feedback predicts the student-reported (a) utility of feedback and (b) quality of teacher-student relationships. These are postulated as mediating pathways in the association between dialogic feedback and engagement dimensions. Students' migration background is explored as a potential moderator of feedback effects.

Questionnaires were administered to 3rd and 4th grade students (N=257) in May and June of 2024. Instructions and items were read aloud during sessions. Children self-reported their cognitive-behavioural (16 items, e.g. "Beim Lernen im Deutschunterricht passe ich gut auf"; $\alpha = .91$) and emotional engagement (8 items, e.g. "Im Deutschunterricht arbeite ich voller Energie"; $\alpha = .83$) on a 5-point Likert Scale (fast nie - fast immer). Teacher-student relationship (7 items, e.g. "Meine Deutschlehrkraft... interessiert sich dafür, wie es mir geht"; $\alpha = .89$) and feedback variables were rated on a 4-point Likert scale (stimmt gar nicht – stimmt genau). With respect to feedback, students reported whether teachers sought their input during feedback conversations (7 items, e.g. "wir besprechen... wie ich meinen Text selbst finde"; $\alpha = .83$) and whether they find the feedback useful (5 items, e.g. "das Feedback... hilft mir, mich zu verbessern"; $\alpha = .75$). Additionally, participants indicated their most recent grades, gender, age and language spoken at home.

Preliminary analyses show that the effect of dialogic feedback on emotional engagement with schoolwork is mediated through increased quality of teacher-student relationship ($b = 0.115$, Bootstrapped 95% CI [.051, .188]) and feedback utility ($b = 0.169$, Bootstrapped 95% CI [0.083, 0.273]). Effects on cognitive-behavioural engagement are mediated only by feedback utility ($b = 0.15$, Bootstrapped 95% CI [0.074, 0.246]). Neither the indirect nor the direct pathways significantly differ based on students' migration background. Results are discussed in the context of previous findings.

Differential effects of formative feedback on argumentative writing and affective-motivational variables among less-proficient writers

Katrin Peltzer¹, Steve Graham², Vera Busse¹

¹Universität Münster, ²Arizona State University

Effective argumentation is crucial for societal participation, making the teaching of argumentative writing a key part of the English curriculum in Germany (Peltzer et al., 2022). However, many foreign language learners struggle to establish structure and coherence in their essays (Siekmann et al., 2022). A process-oriented approach with revision cycles (Hyland, 2010; Graham & Sandmel, 2011) and feedback (Scherer et al., 2024) helps support students' argumentative writing. Feedback based on rubrics and exemplars is effective for learning (Lipnevich et al., 2014), far more time-efficient than in-text comments, and is also suitable for promoting deep-level features of texts, such as structure and coherence (Peltzer et al., 2024). However, little is known whether such feedback also caters to the needs of at-risk writers and how it affects affective-motivational factors such as writing self-efficacy and writing enjoyment. Yet, the latter may influence learner engagement and perseverance in the challenging writing process (Graham, 2018). Our paper responds to calls for more research on feedback to deep-level features of writing among language learners (Scherer et al., 2024), takes feedback perceptions into account (Berry et al., accepted for publication) and addresses the scarcity of research investigating the motivational impact of different feedback practices (Camacho et al., 2021).

We conducted a randomized controlled intervention study to examine differential effects of formative teacher feedback (EG1: rubric + exemplar; EG2: in-text comments; EG3: rubric, exemplar, in-text comments; CG1: learning unit without additional feedback) on argumentative writing in English, writing self-efficacy, and writing enjoyment. Feedback was implemented within a learning unit on argumentative writing (see also Authors, 2024). The sample consisted of 202 secondary school students at comprehensive schools (Year 9) who generally struggle with writing. In our analyses, we divided the sample into at-risk writers (below median) and developing writers (median and up), based on pre-test holistic writing scores assessed via comparative judgment using the software Comprovid. We measured writing self-efficacy and writing enjoyment at pre- and posttest via questionnaires and assessed deep-level features (structure and coherence) via an analytic rating of 808 essays that students wrote during the intervention.

Our findings indicate that rubric + exemplar feedback not only improves holistic writing quality (Authors, 2024) but also deep-level features and is beneficial for at-risk writers. Results show differential effects of feedback for at-risk or developing writers with post-hoc analyses revealing that EG1 (rubric + exemplar) was particularly helpful for at-risk writers. Importantly, both at-risk and developing writers perceived feedback during the intervention as positive with no significant differences between groups. However, the overall quality of the feedback was rated higher than its motivational impact (at-risk writers: $t(66) = 10.790$, $p < .001$, $d = 1.318$, 95%-CI [.987, 1.644]; developing writers: $t(67) = 9.423$, $p < .001$, $d = 1.143$, 95%-CI [.834, 1.446]). In general, at-risk writers felt less self-efficacious and showed less writing enjoyment than their peers (self-efficacy: $t(194) = -4.345$, $p < .001$, $d = -.621$, 95%-CI [-.907, .333]; writing enjoyment: $t(197) = -5.266$, $p < .001$, $d = -.747$, 95%-CI [-1.033, -.458]), although both groups showed overall medium to high scores not only in enjoyment but also in self-efficacy, with the latter indicating potentially uncalibrated perceptions (see also Busse et al., 2023). There were no significant changes over time in writing self-efficacy and writing enjoyment among at-risk students. Our results suggest that time-efficient feedback methods such as rubrics and exemplars are also suitable for at-risk writers and can enhance writing achievement. Rubric + exemplar feedback is also perceived positively by students but did not change writing self-efficacy or writing enjoyment. However, calibration between self-efficacy and writing achievement was better at T2 in EG1. Implications for teaching are discussed.

Moderiert das Geschlecht den Einfluss des wahrgenommenen Feedbacks auf die Erfolgsattribution von Schülerinnen und Schülern in der gymnasialen Oberstufe?

Ariane S. Willems, Katharina Dreiling

Georg-August-Universität Göttingen

Lern- und leistungsbezogenes Feedback, das Schülerinnen und Schüler (SuS) von Lehrkräften erhalten, beeinflusst sowohl ihre Kompetenz- als auch die Motivationsentwicklung (Hattie & Clarke, 2018; Hattie & Timperley, 2007; Rakoczy et al., 2008; Vollmeyer & Rheinberg, 2005). Allerdings belegen neuere Studien differenzielle Effekte für die Feedbackwahrnehmung und -wirkung, etwa je nach Feedbackdimension oder in Abhängigkeit individueller Merkmale der Lernenden, wie ihrem Leistungsniveau oder dem Geschlecht (Dreiling, 2023; Willems & Dreiling, 2022; Wisniewski et al., 2020).

Basierend auf der Selbstbestimmungstheorie (Deci & Ryan, 2000) wird dabei häufig der Zusammenhang zwischen Feedback und intrinsischer Motivation sowie der medierende Effekt des motivationsunterstützenden Erlebens von Kompetenz und Autonomie analysiert. Deutlich weniger Befunde liegen zu der Frage vor, wie sich Lehrkraftfeedback auf die Leistungsattribution von SuS – insbesondere vermittelt über das motivationsfördernde Erleben – auswirkt (Graham, 1991; Zhou & Urhahne, 2012). Diese Frage erscheint besonders relevant vor dem Hintergrund vergleichsweise stabiler Befunde, die zeigen, dass Jungen (zumindest in mathematisch-naturwissenschaftlichen Domänen) selbstwertdienlicher attribuieren als Mädchen (Benölken, 2014; Dickhäuser & Meyer, 2006).

Für den Deutschunterricht (der Grundschule) verweist Knickenberg (2018) auf positive Zusammenhänge zwischen einem positiven Lehrkraftfeedback und der Attribution von Erfolgen auf Fähigkeiten sowie Anstrengungen der Lernenden. Aus differenzieller Perspektive wird deutlich, dass die Zusammenhänge für Jungen stärker sind als für Mädchen. Offen bleibt

allerdings, über welche motivationsunterstützenden Mechanismen die Feedbackwahrnehmung auf die Erfolgsattribution wirkt und inwieweit die berichteten Befundmuster auch bei SuS der Sekundarstufe auftreten.

Unser Beitrag geht dementsprechend folgenden Fragen nach: (F1) Wie beeinflusst das wahrgenommene Lehrkräftefeedback die Erfolgsattribution von SuS im Deutschunterricht der gymnasialen Oberstufe? (F2) Inwieweit mediert das motivationsfördernde Kompetenz- und Autonomieerleben den Zusammenhang von Feedback und Erfolgsattribution? (F3) Lassen sich geschlechtsspezifische Zusammenhangsmuster nachweisen?

Die Analysen basieren auf Daten der Messwiederholungsstudie FeeHe – Feedback im Kontext von Heterogenität, in der differenzielle Wahrnehmungen und Wirkungen von Feedback im Deutschunterricht der gymnasialen Oberstufe untersucht werden. Zu Schuljahresbeginn (t1) nahmen N=807 SuS an der Studie teil (53% weiblich). Die Folgerhebung (t2) wurde am Ende des Schulhalbjahres mit N=696 Teilnehmenden durchgeführt. In den Analysen werden die Feedbackwahrnehmung sowie das Kompetenz- und Autonomieerleben zu t1 genutzt, um die Erfolgsattribution zu t2 vorherzusagen (Für ausführliche Informationen zu allen eingesetzten Instrumenten: Willems et al., 2020).

Zur Erfassung der Feedbackwahrnehmung wurden vier Subskalen entwickelt, die mit je vier vierstufig skalierten Items unterschiedliche Feedbackdimensionen abbilden (ergebnisorientiert, prozessorientiert, selbstregulationsorientiert, dialogorientiert). Konfirmatorische Faktoranalysen bestätigen die vierfaktorielle Struktur der Daten (latente Interkorrelationen: $.62^{**} \leq r_{t1} \leq .84^{**}$). Die Reliabilitäten der Subskalen sind zufriedenstellend ($.72 \leq \alpha_{t1} \leq .73$). Das Autonomieerleben wurde über eine Skala zur erlebten kognitiven Selbstständigkeit im Deutschunterricht erfasst (Stefanou et al., 2004). Die Skala umfasst vier vierstufig skalierte Items und weist eine einfaktorielle Struktur auf ($\alpha_{t1} = .73$). Das Kompetenzerleben wurde mit einer sechs Items (vierstufig) umfassenden Skala erfasst (Williams & Deci, 1995), die eine einfaktorielle Struktur aufweist ($\alpha_{t1} = .77$). Hinsichtlich die Erfolgsattribution werden die Dimensionen Anstrengung und Begabung fokussiert, die jeweils mit drei vierstufig skalierten Items erfasst wurden (Rakoczy et al., 2005). Die Reliabilitäten der Skalen sind zufriedenstellend ($\alpha_{t2} = .83$ bzw. $\alpha_{t2} = .82$), wobei konfirmatorische Faktorenanalysen die zweifaktorielle Struktur bestätigen (latente Interkorrelation: $r_{t2} = .20^{**}$).

Zur Beantwortung der Fragestellungen wird für alle Instrumente die geschlechtsspezifische Messinvarianz geprüft. Anschließend werden Strukturgleichungsmodelle unter Berücksichtigung der Mehrebenenstruktur der Daten (F1/F2) sowie Multigruppenvergleiche (F3) zur Analyse eines potenziellen Moderatoreffektes des Geschlechts berechnet. Die Attributionsdimensionen werden in einem gemeinsamen Modell als (korrelierte) abhängige Variablen modelliert. Dabei wird zunächst ein Modell ohne Mediationseffekte und anschließend ein Modell mit den angenommenen Mediationseffekten berechnet.

Die Befunde zeigen, dass alle eingesetzten Instrumente über die nötige geschlechtsspezifische Messinvarianz verfügen. Zusammenhangsanalysen zeigen zunächst, dass alle Feedbackdimensionen positiv mit beiden Dimensionen der Erfolgsattribution der SuS korrelieren ($.14^{**} \leq r \leq .20^{**}$), wobei die Zusammenhänge jeweils vollständig über das Kompetenz- bzw. Autonomieerleben vermittelt werden. Die Multigruppenanalysen weisen dabei nicht auf signifikante geschlechtsspezifische Befundmuster hin.

IS-05: Invited Symposium: Üben im Unterricht: Theoretische und empirische Überlegungen zu einer zentralen Dimension von Unterrichtsqualität

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: SO 108

Chair der Sitzung: Stefan Keller

Chair der Sitzung: Erin Gerlach

Üben im Unterricht: Theoretische und empirische Überlegungen zu einer zentralen Dimension von Unterrichtsqualität

Chair(s): **Stefan Keller** (PH Zürich, Schweiz)

Diskutant*in(nen): **Erin Gerlach** (Uni Hamburg)

<p>Ein zentrales Anliegen der Schule besteht darin, den nachhaltigen Erwerb von Wissen und Kompetenzen bei den Schülerinnen und Schülern zu förüdern. Um dieses Ziel zu erreichen, ist es entscheidend, dass das Erlernte fest in ihrem Gedüchtnis verankert wird und bei Bedarf rasch und sicher abgerufen werden kann. Dieser Verankerungsprozess wird vorrangig durch regelmüssiges Üben realisiert, sodass das Üben eine universelle Kernpraxis schulischen Arbeitens darstellt (Brinkmann, 2012; Stern et al., 2017). Nachhaltige Lernprozesse gehen über die reine Aufnahme von Informationen hinaus und erfordern wiederholte Bewusstmachung, Verknüpfung mit Vorwissen sowie die praktische Anwendung des erworbenen Wissens. Dabei ist es zentral, das erworbene Verstündnis durch aktive Wiederholung und praktische Anwendung zu festigen, um sicherzustellen, dass es nachhaltig verfügbar ist (Helmke, 2022).</p>

<p>Wührend also die Bedeutung des Übens für das (fachliche) Lernen theoretisch gut verankert ist, gibt es noch wenig empirische Forschung dazu in konkreten Fachkontexten. Die empirische Erforschung von Faktoren der Unterrichtsqualitüt hat sich bisher stark auf die «drei Basisdimensionen» von Unterricht fokussiert (Klieme, 2018), wührend das Üben als zentrale Determinante des Lernerfolgs ein blinder Fleck auf der Forschungslandkarte darstellt (Wemmer Rogh et al., 2024). Dabei geht es einerseits um die Frage, in welchen fachlichen Kontexten des Lehrens und Lernens welche Art von Üben Sinn macht, und welche empirisch gesicherten Wirkungen (im Sinne einer Probabilitüt des Lernerfolgs) davon zu erwarten sind. Diese Frage kann nicht unabhüngig von bestimmten Füchern und Lerngegenstünden beantwortet werden und hat deshalb in den unterschiedlichen Fachdidaktiken besondere Bedeutung (Keller, Steffensky, Winkler et al., im Druck). Andererseits kann man aus interdisziplinürer Perspektive fragen, ob es bestimmte generische (d.h. fücherübergreifend vergleichbare) Formen des intelligenten Übens gibt, und wo die Grenze zu fachspezifischen Formaten des Übens zu ziehen ist.</p>

<p>Diesen Fragen wird in diesem Symposium aus drei unterschiedlichen Fachperspektiven nachgegangen. Im ersten Beitrag behandeln Friederike Korneck, Julia Arnold, Livia Murer, Charlotte Schneider & Markus Wilhelm das Üben im Kontext der Naturwissenschaften. Sie stellen Projekte zur Aufgabenqualitüt, zum intelligenten Üben sowie zur Lernunterstützung beim Erwerb und Festigen von Methodenwissen beim Experimentieren vor. Im zweiten Beitrag betrachten Timo Leuders & Katharina Loibl das Üben im Kontext des Fachs Mathematik. Sie zeigen auf, in welcher Weise empirische Befunde der Lehr-Lernforschung und begründete Praxisvorschlüge im Mathematikunterricht aufeinander bezogen werden. Dabei identifizieren sie Gemeinsamkeiten und Forschungslücken. Im dritten Beitrag betrachten im dritten Beitrag betrachten Jochen Heins & Dorothee Wieser das Üben im Kontext des Fachs Deutsch. Sie weisen darauf hin, dass für das Üben im Kontext des sprachlichen und literarischen Lernens das (insbesondere zeitliche) Verhütnis von Wissen und Können zentral ist, welches in Abhüngigkeit von den Lernzielen und Lerngegenstünden variiert. Diskutiert werden zudem empirische Befunde zum Üben, die allerdings vorrangig aus Interventionsstudien zur Einführung von Strategietrainings resultieren und insofern hinsichtlich ihrer Übertragbarkeit zu reflektieren sind.</p>

<p>Im vierten Slot des Symposiums ist eine vertiefte, interdisziplinür angelegte Diskussion der Thematik auf der Basis der drei Fachbeitrüge vorgesehen. Unter der Moderation des Chairs und des Diskutanten sollen gemeinsam mit den anwesenden Kolleginnen und Kollegen Forschungs- und Entwicklungsstrategien diskutiert werden, welche dazu beitragen können, intelligente und lernwirksame Übungsformen zu etablieren und damit zu einer Verbesserung der Lernprozesse in unterschiedlichen Fachkontexten beizutragen.</p>

<p>Dieses Symposium wird gemeinsam von der GEBF und der GFD im Rahmen der gegenseitig an den jeweiligen Tagungen angebotenen der „Invited Symposia“ organisiert.</p>

Beitrüge des Symposiums

Üben im Kontext der Naturwissenschaften

Friederike Korneck¹, Julia Arnold², Charlotte Schneider², Murer Livia³, Markus Wilhelm⁴

¹Uni Frankfurt, ²PH FHNW, ³PH Zürich, ⁴PH Luzern

Ausgehend von den Ergebnissen der Third International Mathematics and Science Study (TIMSS, Baumert et al., 1999) stand die Aufgabenkultur im Zentrum der Weiterentwicklung des mathematisch-naturwissenschaftlichen Unterrichts im Rahmen verschiedener fachdidaktischer Projekte. Beispielsweise förderte die Bund-Lünder-Kommission für Bildungsplanung und Forschungsförderung den Modellversuch SINUS (Steigerung der Effizienz des Unterrichts der MINT-Füchern). Die dem Projekt zugrundeliegende Expertise beschreibt Aufgaben als „wohldefinierte Probleme, die (mindestens) eine Lösung haben und deren Bearbeitung in relativ kurzer Zeit möglich ist“ (Hüußler & Lind, 1998, S. 3). Sie ging davon aus, dass kontextorientierte, variantenreiche Anwendungsaufgaben „dem Üben Reiz und Bedeutung geben und zur Konsolidierung des Wissens beitragen“ können (ebd.). Im Rahmen des Nachfolgeprojekts SINUS-Transfer entwickeln und erproben Netzwerke aus Schulen, Hochschulen und Landesinstitutionen - inzwischen dezentral in der Verantwortung der Lünder (z.B. SINUS-NRW oder SINUS-SH) - unter anderem Aufgaben für intelligentes, kompetenzorientiertes Üben (SINUS-NRW, 2015, Kapitel 5).

Das Luzerner Modell zur Entwicklung kompetenzfördernder Aufgabensets (LUKAS, Luthiger et al. 2018, S. 42 ff) unterscheidet Konfrontations-, Erarbeitungs-, Vertiefungs- und Übungsaufgaben, die verschiedene Funktionen im kompetenzorientierten Unterricht einnehmen. Dabei fokussieren Übungsaufgaben auf das Automatisieren, das Konsolidieren sowie auf das Bilden von Routinen mit dem Ziel, deklaratives Wissen und Fertigkeiten zukünftig rasch anwenden zu können. Vertiefungsaufgaben hingegen verknüpfen durch durcharbeitendes, elaboriertes Üben die zuvor erarbeiteten naturwissenschaftlichen Denk-, Arbeits- und Handlungsweisen zum Lerngegenstand variantenreicher. Zur Charakterisierung lernförderlicher Übungsaufgaben dienen folgende vier Merkmalsbereiche (in Klammern die jeweiligen Ausprügungen für Übungs- und Vertiefungsaufgaben): Authentizitüt

(Abbildung einer Kompetenz, geringe Lebensnähe), Komplexität (vorstrukturiert, wenige Repräsentationsformen) Kognition (implizite Arbeit mit Präkonzepten, Reproduktion und naher Transfer) und Differenzierung (Lernunterstützung).

Im Rahmen eines Konzepts zur Choreografie des Unterrichts schlugen Oser und Baeriswyl (2001) Basismodelle zur Strukturierung von Lehr-Lern-Prozessen vor, die durch Lernzieltypen bestimmt wurden. Für den naturwissenschaftlichen Unterricht sind die Basismodelle „Konzeptbildung“, „Lernen durch Eigenerfahrung“ und „Problemlösen“ zentral (Krabbe et al., 2015). Die Theorie sequenziert den Unterricht durch festgelegte Handlungskettenschritte, die in ihrer Reihenfolge eingehalten werden sollen. Insbesondere Unterricht mit dem Ziel der Einführung, Festigung und Konsolidierung von neuen Konzepten beinhaltet Übungsaufgaben des nahen Transfers. Die Unterrichtsplanung zur Konzeptbildung orientiert sich an den fünf Handlungskettenschritten „Bewusstmachen des Vorwissens“, „Durcharbeiten eines Prototyps“, „Beschreiben der wichtigsten Merkmale des neuen Konzepts“, „Aktiver Umgang mit dem neuen Konzept“ sowie „Anwendung des neuen Konzepts in anderen Kontexten“ (Krabbe et al., S. 49 ff). Damit ähneln diese Handlungskettenschritte den Phasen des PADUA-Modells von Aebli (1983). Im Zentrum des Basismodells steht somit ein sorgfältig gewählter Prototyp, anhand dessen die Elemente des neuen Konzepts strukturiert und unter Berücksichtigung des Vorwissens der Lernenden eingeführt werden. Anhand von Beispielaufgaben, die sich zunächst stark an dem Muster des Prototyps orientieren und dann zunehmend komplexer werden, wird das Konzept gesichert, vernetzt und dessen Anwendung flexibilisiert. Für den Physikunterricht dienten die Basismodelle als Grundlage für eine Lehrkräftefortbildung. Es zeigten sich Effekte einer besseren Zielorientierung und Strukturierung der Unterrichtsplanung und -durchführung sowie positive Zusammenhänge mit dem Fachwissen insbesondere der schwächeren Schüler:innen (Krabbe et al, S. 101 ff).

Eine Besonderheit der naturwissenschaftlichen Fächer ist der Fokus auf Kompetenzen des Erkenntnisgewinns und somit auf das Üben und Entwickeln wissenschaftlicher Denk- und Arbeitsweisen, meist im Rahmen des Forschenden Lernens. Mit dem Ziel der Festigung und Unterstützung des Erwerbs von Methodenwissen beim Experimentieren wird in diesem Beitrag exemplarisch die Lernunterstützung durch Concept Cartoons als diskursiv-reflexive Szenarien vorgestellt. Diese wurden im Fach Biologie eingesetzt, um die Schüler:innen in Gruppen zur Diskussion über den Zweck beispielsweise von Hypothesenbildung, der Planung von Experimenten, der Variablenkontrolle, usw. und zur Begründung dieser Schritte anzuregen. Im Zentrum standen die Gütekriterien wissenschaftlichen Arbeitens, die von der durch Concept Cartoons unterstützten Kontrollgruppe, konkreter zur Begründung der einzelnen Experimentierschritte genutzt wurden. Zudem führte die Lernunterstützung zu erhöhten Lernleistungen im Methoden- und Fachwissen (Arnold, 2015; Arnold et al., 2016).

Üben im Kontext des Fachs Mathematik

Timo Leuders, Katharina Loibl

PH Freiburg

Seit vielen Jahrzehnten findet man im deutschsprachigen Raum eine Vielfalt an Konzepten für das Üben im Mathematikunterricht, die man als begründete Praxisvorschläge bezeichnen kann. Unter Bezeichnungen wie „produktives Üben“ oder „intelligentes Üben“ findet man fachlich konkretisierte Ansätze vor allem der Gestaltung von Übungsaufgaben (z. B. Winter, 1984, Leuders, 2014). Daneben gibt es konkrete Übeformate mit speziellem Fokus, wie z.B. regelmäßige Kopfrechenübungen oder Vorstellungsbildungen. Solche Konzepte werden typischerweise von Personen an Hochschulen mit wissenschaftlichem Hintergrund und engem Praxisbezug entwickelt und halten immer mehr Einzug in die breite Praxis, z. B. über Fortbildungen oder Schulbücher. Ein Vorzug solcher Konzepte ist, dass sie sich in konkrete Heuristiken für die Gestaltung von Übeformaten oder Typen von Übungsaufgaben übersetzen lassen. Ihre Begründung finden solche Übekonzepte in lerntheoretisch durchaus plausiblen Annahmen. Allerdings fußen diese Begründungen kaum explizit auf allgemeinen Lerntheorien und sind nur punktuell durch systematische empirische Forschung untermauert.

Auf der anderen Seite wird in der empirischen Lehr-Lernforschung die Frage des Übens seit vielen Jahrzehnten untersucht (Carpenter et al. 2022). Das geschieht oft aus einem übergreifenden, d.h. vom Fach oder Lerngegenstand unabhängigen Erkenntnisinteresse. Sehr oft bezieht sich diese Forschung, auch wenn sie sich nicht als fachdidaktische Forschung versteht, aber dennoch auf ganz konkrete Gegenstände des Mathematikunterrichts (Booth et al, 2017). Viele Lerntheorien und Überprinzipien der empirischen Lehr-Lernforschung beanspruchen einen gewissen Grad an Allgemeinheit, sind oft aber nicht sehr explizit darin, wie breit ihr Geltungsbereich angenommen wird, d.h. unter welchen Rahmenbedingungen und in welchen Umsetzungsvarianten sie herangezogen werden können.

Dieser Beitrag diskutiert in dieser Situation die folgenden Fragen: Inwiefern können empirische Forschung und Praxis im Mathematikunterricht aufeinander bezogen werden? Welche Konsequenzen kann die mathematikdidaktische Praxis aus den Erkenntnissen der empirischen Forschung ziehen? Welche weitere Forschung zum Üben wäre hierfür wünschenswert?

1. Beispielsweise zeigt die empirische Forschung, dass verteiltes Üben („spaced practice“) einem geballten Üben überlegen ist, fokussiert dabei aber vor allem auf das Behalten von deklarativem Faktenwissen oder Fertigkeiten. Verteiltes Üben wird in diesem Sinne auch im Mathematikunterricht in Form von regelmäßigen Wiederholungen empfohlen und umgesetzt, und zwar bei Fakten und Fertigkeiten, die Grundlage für das Weiterlernen sind, wie beim Kopfrechnen oder dem kleinen Einmaleins. Ebenfalls findet man immer wieder Kritik am geballten Üben, das sich in der Praxis insbesondere vor Leistungsüberprüfungen einstellt.

2. Verschachteltes Üben („interleaved practice“) von verschiedenen Lernzielen ist im Gegensatz zu geblocktem Üben in empirischen Studien oftmals überlegen. Der zugrundeliegende Mechanismus des Vergleichens wird dabei oft als implizit ablaufend angenommen. In den Übeformaten, die für das Fach Mathematik vorgeschlagen (und seltener empirisch untersucht) werden, steht eher das explizite Vergleichen (z.B. von Lösungsstrategien) im Vordergrund. Ein zentrales Ziel des verschachtelten Übens ist das Vermeiden von Verwechslungen, in Mathematik sind das oft Oberflächenstrategien bei der Auswahl und Anwendung einer Prozedur. Diese „Kraft des Vergleichens“ ist in der Praxis des Übens (z.B. in Lehrwerken) noch wenig ausgeprägt.

3. Die Lernwirksamkeit von Abrufübungen („retrieval practice“) zeigt sich zumindest bei einfachen Lerninhalten als förderlich, bei komplexeren Lernzielen sind die Befunde eher uneinheitlich. In der Praxis des Mathematikunterrichts findet man Umsetzung von Abrufübungen z.B. als „Übekarteien“ und übergreifenden „Wissenschecks“, bei denen die Grundidee der günstigen Erschwernis durch die Anforderung des Gedächtnisabrufes und die der wechselnden Abrufkontexte Berücksichtigung finden.

4. In der mathematikdidaktischen Praxis findet sich zudem verknüpfendes Üben, bei dem mehrere Lernziele synergetisch in geeigneten Aufgabenformaten verbunden werden („produktives Üben“). In der Praxis wird dieses Übeformat auch damit begründet, verschiedenen Lernziele (z.B. konzeptuelles Wissen und Automatisierung) in der verfügbaren Zeit zu erreichen. In der Forschung ist die Annahme des synergetischen Übeformats jedoch nicht systematisch untersucht. Da sich die Prinzipien des produktiven Übens jedoch lerntheoretisch begründen lassen, kann sich Forschung in diesem Bereich lohnen.

Üben im Kontext des Fachs Deutsch

Dorothee Wieser¹, Jochen Heins²

¹TU Dresden, ²Uni Hildesheim

Es besteht innerhalb der Deutschdidaktik kein Zweifel, dass die im Deutschunterricht zu erwerbenden Kompetenzen geübt werden müssen, um sie sicher und flexibel realisieren zu können. Dass das Üben auch hier eine zentrale Qualitätsdimension ist, dürfte darum unstrittig sein. Darüber hinaus aber ist mehr als offen denn als geklärt zu markieren. Der Beitrag möchte deshalb theoretisch-konzeptionellen Fragen nachgehen und die wenigen vorliegenden empirischen Befunde zum Üben im Deutschunterricht systematisieren und reflektieren.

Eine mit Blick auf das Üben bedeutsame Besonderheit des sprachlichen und literarischen Lernens im Deutschunterricht ist, dass es an den außerschulischen Sprach- und Literaturerwerb anschließt bzw. darauf basiert, der nicht beim expliziten Erwerb von Wissen ansetzt, um dieses anschließend zu konsolidieren. Außerschulischer Sprach- und Literaturerwerb erfolgt über Erwerbsprinzipien, die einer mimetischen Angleichung an Muster folgen und als Einüben von Können zu bezeichnet sind.

In dem Beitrag wird darum an die Unterscheidung zwischen einem Üben als primäre oder sekundäre Lernform (vgl. Brinkmann, 2021, S. 29) angeschlossen, die sich insbesondere auf das (auch zeitliche) Verhältnis von Wissen und Können im Sprach- und Literaturerwerb bezieht. Das Üben als primäre Lernform zielt auf den Erwerb von Können bzw. auf die Verbesserung eines Könnens. In dieser Weise eingeübt werden sowohl komplexe Fähigkeiten und Fertigkeiten, aber auch Einstellungen und Haltungen (Brinkmann, 2011, S. 140). Ein Üben als sekundäre Lernform hingegen schließt an den Erwerb von Wissen oder Fertigkeiten an und zielt auf eine Prozeduralisierung, Automatisierung, Festigung und Vertiefung von erworbenem Wissen oder erworbenen Fähigkeiten. Dieses Verständnis von Üben ist in der pädagogischen Psychologie sowie in der aktuellen Lehr-Lern-Forschung prominent (Lipowsky, 2020, S. 90; siehe auch Praetorius et al., 2020; Wemmer-Rogh et al., 2023, S. 47) und wird auch in der Deutschdidaktik mit dem Ziel einer fehlerfreien, schnellen, unbewussten und unangestrengten Ausführung von Handlungen verbunden (vgl. Müller, 2021, S. 5; Brand, 2010, S. 119). Für den nicht-institutionalisierten Sprach- und Literaturerwerb ist das konsolidierende Üben jedoch ein eher untypisches Erwerbsprinzip.

Daraus resultieren für die Deutschdidaktik u. a. Fragen danach, (1) wann es in Abhängigkeit vom Lerngegenstand zunächst ein Wissen braucht, d. h. wann Üben als primäre Lernform nicht geeignet erscheint, (2) wann ein dem Können inhärentes Wissen expliziert werden muss, um einen Lernfortschritt zu erzielen und (3) welche Lernendenfaktoren dabei zu berücksichtigen sind (z. B. Spracherfahrungen, Fähigkeit zur Abstraktion). Vermutlich sind die Antworten abhängig vom Lerngegenstand und Lernziel zu finden und häufig dürfte gerade die Verbindung zum Ziel führen.

Wie viel allerdings im Deutschunterricht geübt wird und welche Formen des Übens in der unterrichtlichen Praxis vorkommen, ist eine Frage, zu der man kaum empirisch fundierte Aussagen findet. In den wenigen vorliegenden Studien, welche alltäglichen Deutschunterricht untersuchen, wurde dieser Frage bisher nicht dezidiert nachgegangen. Und auch empirische Studien zur Wirksamkeit von Übungsformaten in verschiedenen Lernbereichen sind nur insofern auszumachen, dass aus Interventionsstudien, insbesondere im Kontext der Einführung von Strategietrainings oder Förderkonzepten, Rückschlüsse auf bestimmte Aspekte des konsolidierenden Übens gezogen werden können.

Für den Beitrag wurden darum Studien zur Lese- und Schreibförderung gesichtet und die zentralen Befunde entlang der gängigen Kategorien der Übe-Forschung (u. a. Passung der Übeformate, Dosierung des Übens sowie Begleitung durch die Lehrenden) gegliedert. Zu diskutieren ist allerdings, inwiefern diese Befunde auf andere Lernbereiche und -gegenstände des Faches übertragen werden können

Der Beitrag schließt darum mit Vorschlägen zur künftigen Erforschung des Übens, die zum einen das grundsätzliche Desiderat einer dezidierten Übe-Forschung betreffen und zum anderen die Frage nach der Fachspezifik des Übens mit Blick auf das Verhältnis von Wissen und Können sowie die Bedeutung diagnostischer Kompetenzen der Lehrpersonen für das Gelingen Übeprozesse berühren.

Diskussion im Plenum: Forschungs- und Implementationsstrategien zum Üben im Unterricht

Stefan Keller¹, Erin Gerlach²

¹PH Zürich, ²Uni Hamburg

Anstelle eines vierten Fachbeitrags sollen in diesem Teil zentrale Fragen rund um das Thema Üben mit den anwesenden Kolleginnen und Kollegen diskutiert werden. Dieses Format eines freien, erweiterten Austauschs hatte sich an der GEBF Tagung 2024 in Potsdam sehr bewährt und soll deshalb hier wieder aufgenommen werden.

Die Diskussion wird vom Chair des Symposiums sowie vom Diskutanten moderiert und sowohl Austausch in Gruppen wie auch im Plenum enthalten. Die Rolle des Diskutanten dabei ist es, zentrale Fäden der Diskussion zu bündeln und auf die einleitenden Beiträge zu beziehen. Dieser Prozess wird über ein Padlet festgehalten und den Teilnehmenden anschließend zur Verfügung gestellt.

Zentrale Fragen werden sein:

- o Welche Relevanz hat das Thema des Übens für die anwesenden Kolleginnen und Kollegen in ihren Communities?
- o Wo liegt die Schnittstelle zwischen allgemeinen (generischen) und fachspezifischen Aspekten des guten Übens
- o Welche Fragestellungen und methodischen Zugänge sind sinnvoll, um in der empirischen Erforschung des Übens im Unterricht weiter zu kommen?
- o Wie können Wissenschaft und Schulpraxis bei dieser Thematik aufeinander bezogen werden?
- o Wie können Studierende lernen, sinnvolle Übungsformen für die ihre Lernziele zu planen und umzusetzen?

Ziel dieser Diskussion ist es, konkrete Projektideen und Kooperationen anzuregen, welche über die Tagung hinaus weiterbestehen und umgesetzt werden.

IS-06: Invited Symposium: Wissen oder Annahmen? Beiträge zur Validierung von Interpretationen aus Logdaten

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: SN 163

Chair der Sitzung: Carolin Hahnel

Wissen oder Annahmen? Beiträge zur Validierung von Interpretationen aus Logdaten

Chair(s): **Carolin Hahnel** (Ruhr-Universität Bochum, Deutschland)

Diskutant*in(nen): **Ulf Kröhne** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Seit Computer zur Messung von Kompetenzen und anderen bildungsbezogenen Merkmalen eingesetzt werden, stehen Logdaten, die u.a. durch Mausklicks und Tastenanschläge erzeugt werden, in dem Ruf, Kompetenzmessung als leicht zugängliche Datenquelle prozessbezogen informieren und verbessern zu können. Obwohl inzwischen zahlreiche Studien hierzu vielversprechende Erkenntnisse generiert haben, scheint eine praktische Anwendung im Bildungskontext dennoch eher die Ausnahme zu sein (Stadler et al., 2023). Ein Grund für diese Diskrepanz könnte in der Interpretierbarkeit von Logdatenindikatoren liegen, die oft nicht hinreichend definiert und/oder gesichert zu sein scheint (vgl. Goldhammer et al., 2021; Hahnel et al., 2023). Erschwert wird dieser Umstand beispielsweise dann, wenn bildungsbezogene Konstrukte wenig konkretisieren, welche Handlungen für oder gegen eine bestimmte Kompetenzausprägung oder Strategieanwendung sprechen, oder wenn sich Designs und Darstellungsformate von Testaufgaben derart verändern, dass die Übertragung von abgeleiteten Logdatenindikatoren für diese veränderten Formate problematisch erscheint.

Die vier Beiträge des Symposiums verdeutlichen diese Herausforderungen bei der Verwendung von Logdaten und vollführen in ihren jeweiligen Forschungskontexten eine anspruchsvolle Pionierarbeit. Sie präsentieren verschiedene Aspekte der Kompetenzmodellierung, Diagnostik und Psychometrie, an denen verschiedene Disziplinen zusammenkommen müssen, um valide und belastbare Schlüsse über die Interpretation von Logdatenindikatoren und ihre Verortung in Kompetenzmodellen zu ziehen.

Der erste Beitrag von Illmann et al. fokussiert die Frage, wie Logdaten dazu beitragen können, die informationsbezogenen Kompetenzen von Studierenden abzubilden, die im Internet Informationen zu verschiedenen wirtschaftswissenschaftlichen oder medizinischen sowie domänenunspezifischen Problemen recherchieren. Besondere Herausforderungen, denen sich der Beitrag widmet, liegen in der Verwendung des realen Internets als Informationssuchraum und der konzeptuellen Frage, welche Verhaltensweisen für ein kompetentes Suchen und Bewerten von Informationen sprechen.

Hartmann et al. konzentrieren sich im zweiten Beitrag ebenfalls darauf, logdatenbasierte Indikatoren zu identifizieren, die für kompetentes Handeln sprechen. Die Autoren untersuchen allerdings Kompetenzen im Bereich der Kfz-Störungsdiagnose und vergleichen die aus den Logdaten abgeleiteten Informationen mit den Beobachtungsergebnissen von geschultem Prüfungspersonal, um die Verwendung eines computerbasierten Tests für die Prüfungspraxis von Kfz-Mechatroniker:innen zu evaluieren.

Der dritte Beitrag von Deribo et al. untersucht die Frage, ob Logdaten Aussagen über Time-on-Task erlauben, wenn Testaufgaben so präsentiert werden, dass mehrere Items zeitgleich auf einer Seite dargeboten werden. Im Beitrag werden dazu logdatenbasierte Time-on-Task-Maße mit Blickbewegungsdaten aus der Bearbeitung eines fluiden Intelligenztests miteinander verglichen und Abweichungen zwischen den untersuchten Maßen durch verschiedene Personen- oder Itemeigenschaften erklärt.

Im vierten Beitrag befassen sich Mertens & Lindner ebenfalls mit der Analyse von logdatenbasierten Zeitmaßen und untersuchen, wie sich diese im Zusammenhang mit Fehlerfeedback beim Bearbeiten naturwissenschaftlicher Multiple-Choice-Aufgaben verändern. Die Autorinnen zeigen anhand eines within-subjects Designs unter anderem, dass sich die Zeit bis zur Antwortabgabe signifikant verkürzt, wenn Lernende ein Feedback erwarten. Die Ergebnisse implizieren somit, dass Lernende zwischen den verschiedenen Bedingungen andere Antwortstrategien anwenden, die zu Veränderungen im logdatenbasierten Time-on-Task-Indikator führen.

Beiträge des Symposiums

Die Definition und Validierung von Logdaten-Indikatoren für die Messung von Critical Online Reasoning

Jannick Illmann¹, Daniel Baumartz², Philine Drake¹, Carolin Hahnel³, Johannes Hartig¹, Carmen Köhler¹, Marcus Schrickel¹, Alisa Stroganova⁴, Frank Goldhammer⁵

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Goethe-Universität Frankfurt a.M., ³Ruhr-Universität Bochum, ⁴Ludwig-Maximilians-Universität München, ⁵DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; Zentrum für internationale Vergleichsstudien (ZIB)

Theoretischer Hintergrund. Das Internet zählt zu den wichtigsten Informationsquellen von Studierenden, um für ihr Studium zu lernen (Maurer et al., 2020). Durch die sich ständig erweiternde Informationsfülle sind sie dabei auf ihre Fähigkeiten angewiesen, Informationen zu filtern, sie kritisch zu hinterfragen und kohärente Argumente aus ihnen zu bilden. Critical Online Reasoning (COR; Molerov et al., 2020) ist ein Konstrukt, das diese Fähigkeiten anhand von drei Facetten beschreibt: Online Information Acquisition (OIA), Critical Information Evaluation (CIE) und Reasoning with Evidence, Argumentation and Synthesis (REAS). Trotz solcher Fortschritte in der theoretischen Beschreibung stellt das Assessment der Fähigkeiten von Studierenden, das Internet zur Lösung von Informationsproblemen zu nutzen, nach wie vor eine Herausforderung für Wissenschaft und Praxis dar. Viele Untersuchungen, die sich mit vergleichbaren Fähigkeiten und Konstrukten beschäftigen, bilden COR-nahe Fähigkeiten ab, indem sie Informationen nur in einem sehr eingeschränkten Suchraum präsentieren (z.B. Breakstone et al., 2021). Demgegenüber stehen offene Suchräume wie das Internet, in dem Logdaten einen detaillierten Einblick in den Such- und Arbeitsprozess von Studierenden bieten und dadurch zu einer valideren Messung von COR-Fähigkeiten beitragen könnten.

Fragestellung. Der vorgestellte Beitrag setzt an dieser Stelle an und beleuchtet, wie verschiedene Logdatenindikatoren, die aus der Bearbeitung mehrerer Informationsprobleme (Brand-Gruwel et al., 2009) mithilfe des freien Internets gewonnen werden, zur Messung der COR-Fähigkeiten von Studierenden beitragen können.

Methode. Wir präsentierten N = 423 Studierenden aus vier Fachrichtungen (Medizin, Soziologie, Wirtschaftswissenschaften und Physik) verschiedene Informationsprobleme, die entweder domänenspezifische Szenarien aus der Medizin oder den Wirtschaftswissenschaften darstellten oder die ohne ein spezifisches Vorwissen bearbeitet werden konnten. Beispielsweise forderte eines der generischen Informationsprobleme die Studierenden zu einer kurzen Recherche und einer begründeten Stellungnahme zu der Frage auf, inwiefern Mobilfunkmasten, die in der Nähe von Wohngebieten aufgestellt sind, eine

Gesundheitsgefährdung darstellen können. Im Rahmen der Aufgabe wurden die Studierenden zusätzlich gebeten, zwei URLs zu benennen, die für sie wichtige Informationsquellen darstellten, und zu begründen, warum sie diese Informationsquellen herangezogen haben. Neben den Lösungen der Studierenden der jeweiligen Informationsprobleme, die über offene Textantworten erfasst wurden, sammelten wir hochaufgelöste Logdaten (z.B. geöffnete Webseiten, Tastenanschläge, Mauspositionen), um ihren Weg durch den Informationsraum exakt nachzuverfolgen. Insbesondere für die COR-Facetten OIA und CIE adaptierten wir Indikatoren vergleichbarer Kompetenzkonstrukte und entwarfen neue Indikatoren mittels eines Bottom-Up-Ansatzes basierend auf den beobachteten Daten und der Konstruktdefinition. Die resultierenden Indikatoren bewerten wir hinsichtlich ihrer Varianz, ihres Beitrags zur Bildung von Skalen für OIA und CIE und ihrer validen Interpretierbarkeit im Sinne des COR-Kompetenzkonstrukts.

Vorläufige Ergebnisse und Diskussion. Wir konzentrieren uns im aktuellen Beitrag auf Indikatoren, die vollautomatisiert berechnet werden können, sodass sie leicht auf andere Untersuchungen und praktische Anwendungen übertragbar sind, sofern die benötigten Daten vorliegen. Beispielsweise bezieht sich der Indikator „Click Restraint“ darauf, ob Personen die Ergebnisse ihrer Suche betrachteten und eine bewusste Entscheidung für eine oder mehrere der angebotenen Webseiten trafen oder ob sie „blind“ auf das erstbeste Suchergebnis klickten. Zurzeit laufen die Analysen der Indikatoren, deren Ergebnisse auf der GEBF-Tagung vorgestellt werden. Dazu stellen wir die Definitionen und die intendierten Interpretationen der Indikatoren vor, berichten über empirische Ergebnisse hinsichtlich individueller Unterschiede zwischen Studierenden verschiedener Fachrichtungen und darüber, inwiefern die Indikatoren die beabsichtigten Konstrukte valide erfassen. Die vorgestellte Untersuchung erweitert das Verständnis von COR und von Fähigkeiten zur Nutzung des Internets für die Informationssuche im Allgemeinen, indem sie verallgemeinerbare Indikatoren definiert und ein psychometrisches Modell für COR aufstellt, mit dessen Hilfe diese wichtigen Fähigkeiten gemessen, verglichen und trainiert werden können.

Digitalisierung praktischer Prüfungen im Handwerk: Chancen und Herausforderungen am Beispiel einer Kfz-Computersimulation

Peter Hesse, Stefan Hartmann, Stephan Abele
Technische Universität Dresden

In der beruflichen Bildung birgt der Einsatz von Computersimulationen als innovatives Verfahren in fachpraktischen Abschlussprüfungen große Potentiale in den Bereichen Objektivität, Testfairness, Reliabilität und Testökonomie (Kirschner, Park, Malone, & Jarodzka, 2017). Computersimulationen stellen daher einen vielversprechenden Ansatz zur nachhaltigen Lösung zunehmender technischer und personeller Herausforderungen im Prüfungswesen dar (Deutscher & Winther, 2019).

In einem interdisziplinären Forschungs- und Transferprojekt entwickeln Akteur:innen aus Fachdidaktik, pädagogischer Psychologie und beruflicher Bildung gemeinsam digitale Prüfungsstationen zur Erfassung von Kompetenzen im Bereich der Kfz-Störungsdiagnose. Kernstück ist eine Kfz-Computersimulation, mit der Defekte in elektronischen Kraftfahrzeugsystemen simuliert werden (Hesse, Kaseler, Meier, Müller, & Abele, 2022). Ziel ist ein Einsatz der Simulation in der praktischen Gesellenprüfung für Kfz-Mechatroniker:innen. Um aussagekräftige und belastbare Ergebnisse zu liefern sowie die notwendige Akzeptanz bei den Anwender:innen zu erhalten, muss eine gute Validität der Testwertinterpretationen sichergestellt werden: Inwieweit führt eine Auswertung von Klicks am Computer zu vergleichbaren Interpretationen wie eine strukturierte Beobachtung des Verhaltens Auszubildender bei der Störungsdiagnose an realen Fahrzeugen durch geschultes Prüfungspersonal?

Um diese Frage zu beantworten, ist ein mehrschrittiges Vorgehen erforderlich, bei dem Validitätsevidenzen aus unterschiedlichen Quellen (siehe unten) systematisch gesammelt und ausgewertet werden (AERA, APA, & NCME, 2014). In einem ersten Schritt werden Aufgabenstellungen aus bestehenden Prüfungen am realen Fahrzeug möglichst präzise in der Kfz-Computersimulation nachgestellt („Evidence based on test content“). Im zweiten Schritt wird überprüft, wie die in den zugehörigen Prüfungsprotokollen dargelegten Bewertungskriterien und ihre Gewichtung auf Logdaten aus der Kfz-Simulation angewendet werden können („Evidence based on internal structure“). In einem dritten Schritt werden Auswertungsalgorithmen entwickelt, die entsprechende Informationen aus den Logdaten extrahieren und in Punktzahlen überführen („Evidence based on response processes“). In einem vierten Schritt wird zu überprüfen sein, inwieweit die auf diese Weise generierten Punktzahlen mit den Ergebnissen an klassischen Prüfungsstationen korrespondieren, in denen die Auszubildenden identische Störungsszenarien an realen Fahrzeugen zu bearbeiten haben („Evidence based on relations to other variables“).

Unstrukturierte Interviews mit Prüfungsverantwortlichen und die Analyse bestehender Prüfungsprotokolle ergaben, dass sich die Aufgabenstellungen und Inhalte bestehender Prüfungsstationen zur Kfz-Störungsdiagnose sehr gut in der Kfz-Computersimulation abbilden lassen. Ein Abgleich der in den Prüfungsprotokollen beschriebenen Kriterien und ihrer Gewichtung mit den Logdaten der Computersimulation zeigte, dass eine Bewertung fachlich relevanter Aktionen anhand der Logdaten möglich ist – beispielsweise das Nachvollziehen von Funktionsstörungen am Kfz, das Nutzen digitaler Diagnosesoftware und das Planen, Durchführen und Interpretieren von Messungen mit Messmitteln wie Multimeter und Oszilloskop. Hierzu mussten die ereignisbasierten Log-Einträge in zustandsbasierte Variablen überführt werden. Einschränkungen ergaben sich hingegen bei nicht anhand einzelner Aktionen beobachtbarer Meta-Kompetenzen (z.B. „strategisches Vorgehen“) sowie bei nicht-fachlichen Kriterien (z.B. „Ordnung und Sauberkeit am Arbeitsplatz“).

Time-on-Task aus Log- und Blickbewegungsdaten: Ein Vergleich

Tobias Deribo¹, Ulf Kröhne¹, Carolin Hahnel², Frank Goldhammer³

¹DIPF – Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Ruhr-Universität Bochum, ³DIPF – Leibniz-Institut für Bildungsforschung und Bildungsinformation; Zentrum für internationale Vergleichsstudien (ZIB)

Zeitinformationen aus computerbasierten Assessments können hilfreich sein, um z.B. kognitive Strategien zu untersuchen (Scherer et al., 2015) oder die Messeffizienz von Leistungstests zu steigern (De Boeck & Jeon, 2019). Wenn jedoch mehrere Items zeitgleich auf einer Seite präsentiert werden, erscheint eine einfache und direkte Zuordnung der Zeitinformationen zu einem Item erschwert. Hierbei kann es für die Testteilnehmenden zu Verzerrungen kommen, wenn diese z.B. verschiedene Aufgaben explorieren oder ein Assessment nicht linear bearbeiten.

Um diese Problemstellung zu untersuchen wurden in diesem Beitrag Time-on-Task Maße (z.B., Goldhammer et al., 2014) aus Blickbewegungs- und Logdaten verglichen und deren Gemeinsamkeiten und Unterschiede beleuchtet. Darüber hinaus wurde untersucht, ob bestimmte Personen- oder Itemeigenschaften die Unterschiede zwischen den gewonnenen Maßen erklären können.

Dafür verwendet der Beitrag vorverarbeitete Blickbewegungs- und Logdaten von N = 172 Studierenden aus dem TüEyeQ-Datensatz (Kasneji et al., 2021). Die Stichprobe nahm dabei an einem Assessment der fluiden Intelligenz teil (CFT 20-R; Weiß, 2006). Alle Aufgaben wurden in einem Design mit mehreren Items pro Seite präsentiert. Da sowohl Blickbewegungs- als auch

Logdaten Quellen von Prozessdaten darstellen, wurde für beide ein Finite-State-Machine Ansatz (Kroehne & Goldhammer, 2018) verwendet, um den Testablauf auf Grundlage der jeweiligen Datenquelle zu rekonstruieren. Für das blickbewegungs-basierte Time-on-Task Maß wurde die aufsummierte Fixationszeit auf die Stimuli und Antwortoptionen einer spezifischen Aufgabe (vor Abgabe der finalen Antwort) verwendet. Dieser Ansatz nimmt an, dass Testteilnehmende ihren Blick auf den Aufgabenteil fixieren den sie aktuell bearbeiten (z.B., Just & Carpenter, 1980). Für das logdatenbasierte Time-on-Task Maß wurden die Zeitabstände zwischen Antwortänderungsereignissen verwendet. Dieser Ansatz geht davon aus, dass die Zeit vor einer finale Antwort mit der Lösung der spezifischen Aufgabe verbracht wurde (z.B., Deribo et al., 2021).

Die Zeitmaße wurden auf der Ebene der rohen Time-on-Task Maße sowie nach der Modellierung von latenter Geschwindigkeit (Van Der Linden, 2006) verglichen. Dabei wurden Rohdifferenzwerte, die Mittlere Absolute Differenz (MAD) und Cohens d (Cohen, 1988) zwischen den beiden Zeitmaße und den latenten Geschwindigkeitsparametern betrachtet. Weiterhin wurden bayesianische Regressionsmodelle verwendet, um zu untersuchen, ob Item- oder Stichprobenmerkmale (Geschlecht, Durchschnittsnote und Bildungshintergrund) die Unterschiede bedingen.

Die vermutete, nicht-lineare Aufgabenbearbeitung betreffend fixierten die Testteilnehmenden für 4.48s (SD = 7.96; Min = 0.004; Max = 50.63) eine andere Aufgabe als die, auf der sie danach eine Antwort abgegeben haben. Dabei fixierten n = 138 Personen der Stichprobe (80.25%) mindestens einmal für 5 Sekunden eine andere Aufgabe vor Antwortabgabe. Dieses Verhalten trat für das Subsample im Mittel 2.06-mal (SD = 1.31, Min = 1, Max = 8) auf. Die Unterschiede in den rohen Time-on-Task Werten variierten dabei zwischen -38.98s und 105.75s pro Antwort zwischen beiden Maßen, mit einem mittleren Bias von 1.9s (d.h. das blickbewegungs-basierte Zeitmaß war im Mittel kürzer) und einer MAD von 2.46s. Weiterhin zeigten sich Unterschiede in der latenten Geschwindigkeit von -0.259 bis 0.288 (MAD = 0.077).

Nach Cohen (1988) scheinen die gefundenen Unterschiede zwischen den beiden Time-on-Task Maßen ($d = 0.166$) und den daraus gewonnen latenten Geschwindigkeitsschätzern ($d = 0.026$) mit $d < 0.2$ vernachlässigbar. Auch die untersuchten Personenmerkmale zeigten keinen nachweisbaren Einfluss auf die Unterschiede zwischen den Maßen. Jedoch zeigte sich, dass ein Anstieg der Zeitintensität eines Items zu einem höheren Time-on-Task-MAD führte. Zusammenfassend deuten die Ergebnisse der Studie darauf hin, dass, trotz der Komplexität, die durch ein Design mit mehreren Items pro Seite entsteht, Unterschiede zwischen den untersuchten Maßen vernachlässigbar scheinen. Das weniger aufwändige, logdatenbasierte Time-on-Task-Maß kann potenziell zu einer gleichwertigen Messung der Time-on-Task verwendet werden. Die gefundenen Unterschiede zwischen den Maßen scheinen nicht durch das nicht-lineare Verhalten zu selten aufzutreten und/oder durch die Aggregation der Maße (z.B. über die latente Modellierung) überdeckt zu werden.

Effekte von direktem und multiple-try Feedback auf die Bearbeitungszeiten von naturwissenschaftlichen Testaufgaben

Ute Mertens¹, Marlit Annalena Lindner²

¹IPN – Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik, ²IPN – Leibniz Institut für die Pädagogik der Naturwissenschaften und Mathematik; EUF – Europa Universität Flensburg

Digitale Testumgebungen ermöglichen die Erfassung des Antwortprozesses von Teilnehmenden und liefern Einblicke in ihr Wissen sowie in ihre Interaktion mit den Testaufgaben (z.B. durch Bearbeitungszeiten; Anghel et al., 2024). Die Bearbeitungszeit kann als kognitiver Indikator für die Geschwindigkeit der Aufgabenverarbeitung betrachtet werden und sowohl auf (weniger) sorgfältige Bearbeitung oder das Fähigkeitsniveau hinweisen (Goldhammer et al., 2014; Lindner et al., 2017). Studien zu den Auswirkungen von Feedback auf Bearbeitungszeiten zeigen uneinheitliche Ergebnisse: Einerseits antworten Teilnehmende schneller bei unmittelbarem Feedback (Candel et al., 2020; Ling et al., 2017); andererseits kann Feedback die Bearbeitungszeit bei leistungsschwachen Teilnehmenden verlängern (Ling et al., 2017). Hinsichtlich der Zeit auf dem Feedback scheinen die Antwortrichtigkeit sowie die Feedbackvariante entscheidende Faktoren für die Interaktion mit der Feedbacknachricht zu sein. Frühere Studien zeigen, dass Teilnehmende nach einem Misserfolg dem Feedback mehr Aufmerksamkeit widmen als nach einem Erfolg (Heckler & Mikula, 2016; Máñez et al., 2019) sowie umfangreichere oder unerwartete Informationen häufig mehr Zeit erfordern (Gal & Hershkovitz, 2019; Kuklick et al., 2023). Die vorliegende Studie fokussiert auf zwei typische computerbasierte Feedbackvarianten, die inhaltlich gleiche Informationen bieten, sich jedoch in ihrer Präsentation unterscheiden: einmaliges direktes Feedback und wiederholtes multiple-try Feedback (Mertens et al., 2022; Shute, 2008).

Diese Studie untersucht, ob und welchen Einfluss die Präsentation dieser Feedbackvarianten im direkten Vergleich zu keinem Feedback auf die Bearbeitungszeit der Aufgaben (Time on Task; TOT) hat und inwiefern sich die Verarbeitungszeiten der Feedbacknachrichten (Time on Feedback; TOF) in Abhängigkeit der Antwortrichtigkeit unterscheiden.

Methode. Die Analysen basieren auf 335 Studierenden (65% weiblich, MAlter = 23.04, SDAlter = 4.30), die 36 schwierige naturwissenschaftliche Multiple-Choice-Aufgaben (Lösungserfolg: 13%–62%) bearbeiteten. Die Testaufgaben wurden personenweise vollständig randomisiert und in einem within-subject Design in drei Blöcken ohne Feedback, mit direktem Feedback und mit multiple-try Feedback präsentiert (je 12 Aufgaben je Bedingung).

Die Bearbeitungszeiten wurden anhand des Klickverhaltens der Teilnehmenden erfasst, ohne die Bearbeitung der Testaufgaben zu stören (Anghel et al., 2024). TOT dient als Indikator für die Gesamtbearbeitungszeit, die vom Beginn der Aufgabe bis zur Abgabe einer Antwort gemessen wird. In der multiple-try-Feedback-Bedingung wird die TOT als die Zeit vom Beginn der Aufgabe bis zur Abgabe der ersten Antwort definiert. Zudem wurde die mit dem Feedback verbrachte Zeit (TOF) erfasst, die vom Beginn der ersten Feedbacknachricht bis zum Schließen der letzten Feedbacknachricht gemessen wird. Die Daten wurden für die Analyse log-transformiert.

Ergebnisse & Diskussion. Die Studierenden zeigten signifikant kürzere TOT bei Aufgaben, die mit Feedback präsentiert wurden als bei Aufgaben ohne Feedback (direktes Feedback: $b = -0.04$, $p < .001$, $\eta^2_{\text{part}} = 0.004$; multiple-try Feedback: $b = -0.07$, $p < .001$, $\eta^2_{\text{part}} = 0.01$). Möglicherweise verleitete die Erwartung der Rückmeldung die Studierenden zur Anwendung von schnellen Antwortstrategien (z.B. Heuristiken, Intuitionen; Gigerenzer & Gaissmaier, 2011), um in dieser low-stakes Testsituation schneller an die Feedbackinformationen zu gelangen. Die Erkenntnisse stützen somit bisherige Befunde zu verkürzter TOT durch Feedback (Ling et al., 2017).

Hinsichtlich der TOF zeigte sich, dass die Studierenden mehr Zeit auf dem multiple-try Feedback als auf dem direkten Feedback ($b = 0.33$, $p < .001$, $\eta^2_{\text{part}} = 0.18$) verbrachten; insbesondere nach falschen Antworten (direkt vs. multiple-try: $ES = -0.52$, $p < .001$, $d = -0.92$). Auch unter Berücksichtigung der Anzahl der Lösungsversuche blieb dieser Zusammenhang bestehen ($ES = -0.07$, $p < .001$, $d = -0.12$). Diese Befunde deuten darauf, dass die zeitliche Auseinandersetzung mit schwierigen Testaufgaben zu naturwissenschaftlichem Faktenwissen durch das Feedback zwar reduziert wurde, nach falschen Antworten aber dennoch eine Auseinandersetzung mit dem Feedback und der Aufgabe stattzufinden scheint. Die Ergebnisse werden mit Blick auf Befunde

einer verbesserten Testleistung im Rahmen eines Recall-Posttests in der Studie ($p_s < .001$, $\eta^2_{\text{part}} \leq 0.05$) auf der Tagung weitergehend diskutiert.

OF-6: Diskussion: Was bringen die Öffnungen bei der Differenzierung des Bildungssystems?

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 102

Chair der Sitzung: Steffen Schindler

Chair der Sitzung: Hans-Peter Blossfeld

Schulsysteme unterscheiden sich, in der Regel im Sekundarbereich, danach, inwieweit und zu welchem Zeitpunkt Schülerinnen und Schüler gemeinsam beschult oder auf unterschiedliche Schulformen oder Lernkontexte aufgeteilt werden. Über die Vor- und Nachteile dieser Varianten von Bildungssystemen entzünden sich seit Jahrzehnten regelmäßig wissenschaftliche und bildungspolitische Debatten. Diese drehen sich um Fragen der Effizienz des Schulsystems und der sozialen Bildungsungleichheit und Bildungsgerechtigkeit. Auch gegenwärtig wird über diese Fragen wieder sowohl in wissenschaftlichen als auch medialen Auseinandersetzungen teils heftig debattiert. Die Befunde aus den internationalen Vergleichsstudien unterstützen eher die Position, wonach es in früh differenzierenden Bildungssystemen im Vergleich zu integrativen Systemen keine Verbesserungen des Leistungsniveaus gebe, wohl aber verstärkte Einflüsse der sozialen Herkunft (Dräger et al. 2023, Hanushek und Wößmann 2006, Skopek et al. 2019, Strello et al. 2021, Terrin und Triventi 2022, van de Werfhorst und Mijs 2010). An den betreffenden Analysen hat es jedoch immer schon auch Kritik gegeben (Brunello und Checchi 2007, Dronkers et al. 2011, Duflo et al. 2010, Korthals und Dronkers 2016, Waldinger 2007). Insbesondere wurde moniert, dass für die Identifikation der Effekte wichtige Einflüsse nicht erfasst seien (Dronkers et al. 2011, Esser 2021, Marks 2014, Marks 2020) und es bei den differenzierenden Systemen selbst wieder Unterschiede in den Regelungen der Implementation geben kann (Bol et al. 2014, Checchi and Flabbi 2007, Hallinan 1994, Sörensen 1970), etwa regionaler Art, die in den internationalen Vergleichen aber meist nicht berücksichtigt werden.

Vor diesem Hintergrund will sich das Symposium gezielt mit aktuelleren Entwicklungen im deutschen Schulsystem auseinandersetzen. Die Sekundarschulsysteme der deutschen Bundesländer sind nach wie vor gegliedert, und der Übergang aus der Primarstufe erfolgt im internationalen Vergleich eher früh. Einige Bundesländer sind dabei in den Leistungen und der sozialen Durchlässigkeit deutlich besser als andere und kommen international z.T. bis an die Spitzengruppe heran. Zur weiteren Verbesserung der Situation hat es nun seit einiger Zeit eine Reihe von Reformen mit einem allgemeinen Trend zur Öffnung und Entdifferenzierung gegeben. Dazu zählen der Wechsel von der Mehr- zur Zweigleedrigkeit, die zunehmende Etablierung integrativer Schulen unterhalb des Gymnasiums sowie Lockerungen bei der Verbindlichkeit der Schulempfehlungen am Ende der Primarstufe und bei den Anforderungen für Leistungen und Noten.

Damit stellt sich die Frage, wie diese Entwicklungen mit Blick auf die Effizienz des Lernens und die sozialen Ungleichheiten zu bewerten sind. Bisher hat sich hierzu kein eindeutiger Erkenntnisstand herausgebildet. Einige Analysen zeigen, dass bestimmte Formen der stärkeren Integration, wie die Zusammenlegung der unteren Bildungswege oder die zeitliche Verschiebung des Übergangs auf die weiterführenden Schulen, eher günstige Ergebnisse mit sich bringen – und das gerade auch für Kinder in den schwierigen Lagen und in den unteren Bildungswegen (Matthewes 2021, Meghir und Palme 2005, Piopiunik 2014, Wößmann 2010). Es gibt aber auch Hinweise, dass Lockerungen des Zugangs und der schulischen Organisation zu geringeren Leistungen führen, auch schon in der Grundschule, Kinder in den unteren Leistungsbereichen eher benachteiligen und sich die sozialen Ungleichheiten nicht verringern (Bach und Fischer 2020, Cord und Giuliano 2016, Fend 2017, Grewenig 2021, Horn 2013).

Um die bisherigen empirischen Erkenntnisse nochmals gemeinsam zu strukturieren sowie aus der bisherigen Debatte um das integrierte oder differenzierte Lernen informierte Erwartungen zu den Auswirkungen der Öffnungs- und Entdifferenzierungsprozesse abzuleiten, soll das Symposium gezielt verschiedene Perspektiven aus unterschiedlichen Disziplinen der empirischen Bildungsforschung zusammenbringen. Dabei soll es ausdrücklich *nicht* zu einer weiteren Gegenüberstellung der Positionen kommen, sondern zu einem offenen Austausch über die Frage, was aus der Sicht der unterschiedlichen, mit den Fragen befassten Disziplinen, wie Erziehungswissenschaft, Pädagogische Psychologie, Bildungsökonomie und Bildungssoziologie zu den Entwicklungen der zunehmenden bzw. weiter verzögerten Öffnung der Bildungswege zu sagen ist, gerade auch mit einem speziellen Blick auf die deutschen Bundesländer. Dazu soll es möglichst prägnante und kurze Stellungnahmen geben, die dann ausführlicher im Plenum und auch untereinander diskutiert werden können. Diese Stellungnahmen sollen einen bestimmten Aspekt der Einflüsse des differenzierten oder integrierten Lernens herausstellen und im Kontext der aktuellen Öffnungsprozesse bewerten, am besten gestützt auf möglichst belastbare (neue) Forschungsbefunde. In der Diskussion soll dann selbstkritisch die jeweils eigene Position reflektiert und in Bezug zu den Arbeiten der anderen gesetzt werden. Aus der Veranstaltung werden Impulse erwartet, welche dazu beitragen sollen, die aktuelle Debatte stärker zu ordnen und voranzubringen.

Das Symposium beginnt mit einer kurzen Einleitung von Hans-Peter Blossfeld, der in die aktuelle Debatte um die Differenzierung vs. Integration in Schulsystemen einführt. Daran schließen sich vier kurze Impulsvorträge von Kai Maaz, Sönke Matthewes, Sarah Lenz und Hartmut Esser an, bevor eine moderierte Diskussion eröffnet wird. An der Diskussion nehmen als weitere Discussants Hanna Dumont, Jan Skopek und Georg Lorenz teil. Auch das Publikum ist eingeladen, sich an der Diskussion zu beteiligen.

Ablaufplan

Moderation: Steffen Schindler

Einführung (10 Min): Hans-Peter Blossfeld

Impulsvorträge (jeweils 10 Min)

- Kai Maaz

- Sönke Matthewes

- Sarah Lenz

- Hartmut Esser

Diskussion (50 Min)

- Hans-Peter Blossfeld

- Hanna Dumont

- Hartmut Esser

- Sarah Lenz

- Georg Lorenz

- Kai Maaz
- Sönke Mathewes
- Jan Skopek

Was bringen die Öffnungen bei der Differenzierung des Bildungssystems?

Steffen Schindler¹, Hans-Peter Blossfeld¹, Hannah Dumont², Hartmut Esser³, Sarah Lenz⁸, Georg Lorenz⁴, Kai Maaz⁵, Sönke Mathewes⁶, Jan Skopek⁷

¹Universität Bamberg, Deutschland; ²Universität Potsdam, Deutschland; ³Universität Mannheim, Deutschland; ⁴Universität Leipzig, Deutschland; ⁵DIPF Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ⁶Utrecht University, Netherlands; ⁷Trinity College Dublin, Ireland; ⁸andere

Schulsysteme unterscheiden sich, in der Regel im Sekundarbereich, danach, inwieweit und zu welchem Zeitpunkt Schülerinnen und Schüler gemeinsam beschult oder auf unterschiedliche Schulformen oder Lernkontexte aufgeteilt werden. Über die Vor- und Nachteile dieser Varianten von Bildungssystemen entzündeten sich seit Jahrzehnten regelmäßig wissenschaftliche und bildungspolitische Debatten. Diese drehen sich um Fragen der Effizienz des Schulsystems und der sozialen Bildungsungleichheit und Bildungsgerechtigkeit. Auch gegenwärtig wird über diese Fragen wieder sowohl in wissenschaftlichen als auch medialen Auseinandersetzungen teils heftig debattiert. Die Befunde aus den internationalen Vergleichsstudien unterstützen eher die Position, wonach es in früh differenzierenden Bildungssystemen im Vergleich zu integrativen Systemen keine Verbesserungen des Leistungsniveaus gebe, wohl aber verstärkte Einflüsse der sozialen Herkunft (Dräger et al. 2023, Hanushek und Wößmann 2006, Skopek et al. 2019, Strello et al. 2021, Terrin und Triventi 2022, van de Werfhorst und Mijs 2010). An den betreffenden Analysen hat es jedoch immer schon auch Kritik gegeben (Brunello und Checchi 2007, Dronkers et al. 2011, Duflo et al. 2010, Korthals und Dronkers 2016, Waldinger 2007). Insbesondere wurde moniert, dass für die Identifikation der Effekte wichtige Einflüsse nicht erfasst seien (Dronkers et al. 2011, Esser 2021, Marks 2014, Marks 2020) und es bei den differenzierenden Systemen selbst wieder Unterschiede in den Regelungen der Implementation geben kann (Bol et al. 2014, Checchi and Flabbi 2007, Hallinan 1994, Sörensen 1970), etwa regionaler Art, die in den internationalen Vergleichen aber meist nicht berücksichtigt werden.

Vor diesem Hintergrund will sich das Symposium gezielt mit aktuelleren Entwicklungen im deutschen Schulsystem auseinandersetzen. Die Sekundarschulsysteme der deutschen Bundesländer sind nach wie vor gegliedert, und der Übergang aus der Primarstufe erfolgt im internationalen Vergleich eher früh. Einige Bundesländer sind dabei in den Leistungen und der sozialen Durchlässigkeit deutlich besser als andere und kommen international z.T. bis an die Spitzengruppe heran. Zur weiteren Verbesserung der Situation hat es nun seit einiger Zeit eine Reihe von Reformen mit einem allgemeinen Trend zur Öffnung und Entdifferenzierung gegeben. Dazu zählen der Wechsel von der Mehr- zur Zweigliedrigkeit, die zunehmende Etablierung integrativer Schulen unterhalb des Gymnasiums sowie Lockerungen bei der Verbindlichkeit der Schulempfehlungen am Ende der Primarstufe und bei den Anforderungen für Leistungen und Noten.

Damit stellt sich die Frage, wie diese Entwicklungen mit Blick auf die Effizienz des Lernens und die sozialen Ungleichheiten zu bewerten sind. Bisher hat sich hierzu kein eindeutiger Erkenntnisstand herausgebildet. Einige Analysen zeigen, dass bestimmte Formen der stärkeren Integration, wie die Zusammenlegung der unteren Bildungswege oder die zeitliche Verschiebung des Übergangs auf die weiterführenden Schulen, eher günstige Ergebnisse mit sich bringen – und das gerade auch für Kinder in den schwierigen Lagen und in den unteren Bildungswegen (Mathewes 2021, Meghir und Palme 2005, Piopiunik 2014, Wößmann 2010). Es gibt aber auch Hinweise, dass Lockerungen des Zugangs und der schulischen Organisation zu geringeren Leistungen führen, auch schon in der Grundschule, Kinder in den unteren Leistungsbereichen eher benachteiligen und sich die sozialen Ungleichheiten nicht verringern (Bach und Fischer 2020, Cord und Giuliano 2016, Fend 2017, Grewenig 2021, Horn 2013).

Um die bisherigen empirischen Erkenntnisse nochmals gemeinsam zu strukturieren sowie aus der bisherigen Debatte um das integrierte oder differenzierte Lernen informierte Erwartungen zu den Auswirkungen der Öffnungs- und Entdifferenzierungsprozesse abzuleiten, soll das Symposium gezielt verschiedene Perspektiven aus unterschiedlichen Disziplinen der empirischen Bildungsforschung zusammenbringen. Dabei soll es ausdrücklich *nicht* zu einer weiteren Gegenüberstellung der Positionen kommen, sondern zu einem offenen Austausch über die Frage, was aus der der Sicht der unterschiedlichen, mit den Fragen befassten Disziplinen, wie Erziehungswissenschaft, Pädagogische Psychologie, Bildungsökonomie und Bildungssoziologie zu den Entwicklungen der zunehmenden bzw. weiter verzögerten Öffnung der Bildungswege zu sagen ist, gerade auch mit einem speziellen Blick auf die deutschen Bundesländer. Dazu soll es möglichst prägnante und kurze Stellungnahmen geben, die dann ausführlicher im Plenum und auch untereinander diskutiert werden können. Diese Stellungnahmen sollen einen bestimmten Aspekt der Einflüsse des differenzierten oder integrierten Lernens herausstellen und im Kontext der aktuellen Öffnungsprozesse bewerten, am besten gestützt auf möglichst belastbare (neue) Forschungsbefunde. In der Diskussion soll dann selbstkritisch die jeweils eigene Position reflektiert und in Bezug zu den Arbeiten der anderen gesetzt werden. Aus der Veranstaltung werden Impulse erwartet, welche dazu beitragen sollen, die aktuelle Debatte stärker zu ordnen und voranzubringen.

Das Symposium beginnt mit einer kurzen Einleitung von Hans-Peter Blossfeld, der in die aktuelle Debatte um die Differenzierung vs. Integration in Schulsystemen einführt. Daran schließen sich vier kurze Impulsvorträge von Kai Maaz, Sönke Mathewes, Sarah Lenz und Hartmut Esser an, bevor eine moderierte Diskussion eröffnet wird. An der Diskussion nehmen als weitere Discussants Hanna Dumont, Jan Skopek und Georg Lorenz teil. Auch das Publikum ist eingeladen, sich an der Diskussion zu beteiligen.

Ablaufplan

Moderation: Steffen Schindler

Einführung (10 Min): Hans-Peter Blossfeld

Impulsvorträge (jeweils 10 Min)

- Kai Maaz
- Sönke Mathewes
- Sarah Lenz
- Hartmut Esser

Diskussion (50 Min)

- Hans-Peter Blossfeld

- Hannah Dumont
- Hartmut Esser
- Sarah Lenz
- Georg Lorenz
- Kai Maaz
- Sönke Matthewes
- Jan Skopek

PS-42: Sprachentwicklung und -Förderung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 154

Chair der Sitzung: Steffi Sachse

Effekte durchgängiger, institutionell koordinierter Sprachförderung am Übergang Kita-Grundschule

Teresa Patzek¹, Holger Hopp², Dieter Thoma³

¹MAZEM gGmbH, Deutschland; ²TU Braunschweig, Deutschland; ³Universität Mannheim, Deutschland

Sprache ist die Schlüsselkompetenz zum Zugang und zur Teilhabe an Bildung. Insbesondere vor dem Hintergrund migrationsbedingter Mehrsprachigkeit werden seit Jahren Sprachförderprogrammen in allen Bildungsetappen eingesetzt, um sprachliche Entwicklungsdefizite gerade zu Anfang von Bildungsverläufen zu kompensieren (als Überblick: Deutscher Bildungsserver, 2024). Die Effekte von (vor-)schulischer Sprachförderung sind jedoch nur unzureichend belegt, und bestehende Befunde zeigen ein uneinheitliches Bild mit zum Teil nur geringen Erfolgsquoten von sprachlichen Unterstützungsangeboten. Ein wesentliches Merkmal, das geringen Effekten zu Grunde liegen kann, ist die Gebundenheit der meisten bisherigen Förderprogramme an einzelne Bildungsetappen, die eine inhaltliche wie personelle Kontinuität von sprachlicher Förderung verhindern.

In diesem Vortrag berichten wir erste Ergebnisse eines breiten Modellprojekts zur kontinuierlichen Sprachförderung in KiTa-Schulverbünden, um Potenziale einer institutionenübergreifenden Förderung zu messen. Das Projekt verfolgt insgesamt das Ziel einer individuellen, subjektzentrierten Sprachförderbegleitung, um durch den Aufbau adäquater Alltagssprachlicher und dann bildungssprachlicher Kompetenzen die sozial-kommunikativen Fertigkeiten sowie die Selbstwirksamkeit der Kinder in Schule zu stärken. Von 2020-2024 wurden an 6 KiTa-Schulverbünden in Baden-Württemberg 140 Schülerinnen und Schüler mit hohem sprachlichen Förderbedarf im Deutschen additiv von Beginn des letzten KiTa-Jahres bis Ende der 2. Grundschulklasse von den gleichen Sprachlernbegleitern im Umfang von 2-3 Stunden pro Woche unterstützt. Die Unterstützung erfolgte durch qualifizierte Sprachlernbegleitende, die den KiTa-Alltag und die Schulstunden flankierten, um den jeweiligen Bezugskindern eine individualisierte Sprachförderung zukommen zu lassen. Zudem fand ein kontinuierlicher Austausch zwischen Sprachlernbegleitenden und KiTa- bzw. Lehrkräften statt.

In einer quantitativen Wirksamkeitsstudie wurde die Entwicklung von sprachlichen Kompetenzen der Kinder in Bezug zu den Testnormwerten und in Bezug zu Kindern mit diskontinuierlicher Förderung gemessen. Sprachliche und (vor-)schulische Kompetenzen aller teilnehmenden Kinder wurden zu Beginn der Förderung am Start des letzten KiTa-Jahres und zu Ende jedes KiTa bzw. Schuljahres mittels standardisierter Tests, die Altersnormwerte bieten, erhoben. Neben sprachlichen und mathematischen Kompetenzen wurden auch Daten zu metasprachlicher Bewusstheit, Entwicklung des Sozialverhaltens und Lernkompetenzen gesammelt.

Die ersten Ergebnisse zeigen statistisch signifikante Normwertannäherungen in den Bereichen rezeptiver und expressiver Wortschatz (WWT; Glück, 2011) sowie Grammatik (LiSe-DaZ; Schulz & Tracy, 2011) sowohl über den Verlauf der ersten beiden Förderjahre am Übergang von KiTa und Grundschule als auch bis zum Ende der zweiten Klasse. Kompetenzvergleiche mit Schülerinnen und Schülern in Vergleichseinrichtungen, in denen die Sprachförderung diskontinuierlich pro Institution verläuft, belegen, dass kontinuierlich und individuell von KiTa zu Grundschule geförderte Kinder höhere sprachliche Kompetenzzuwächse haben: Institutionenübergreifend geförderte Kinder weisen bereits zu Ende der 1. Grundschulklasse vergleichbare Kompetenzen in der Wortschatzentwicklung auf wie Kinder zu Ende der 2. Grundschulklasse, deren Sprachförderung erst in der Grundschule begann bzw. wieder aufgenommen wurde.

Ein einjähriger Vorsprung in der sprachlichen Entwicklung zeigt deutliche Mehrwerte institutionenübergreifender, kontinuierlicher Förderung und belegt die Notwendigkeit zu trägerübergreifender bildungspolitischer Zusammenarbeit, um Schlüsselqualifikationen zur Bildungsteilhabe frühzeitig und nachhaltig zu fördern. Wir diskutieren die Ergebnisse vor dem Hintergrund geplanter Initiativen und Programme zur institutionenübergreifenden (Sprach-)Förderung am Übergang KiTa-Grundschule und skizzieren bildungspolitische Herausforderungen und Umsetzungsmöglichkeiten im deutschen Bildungssystem

ENTFÄLLT Doppelte Benachteiligung: Erfahren Kinder mit sprachlichen Defiziten eine geringere pädagogische Interaktionsqualität in Kindertagesstätten?

Elisabeth Resa, Nicole Voss, Katharina Kluczniok

Freie Universität Berlin, Deutschland

Ergebnisse empirischer Studien der letzten 20 Jahre zeigen wiederholt Zusammenhänge zwischen der Migrationsgeschichte von Schüler*innen in der Grundschule und deren Bildungserfolg auf (McElvany et al., 2023). Neben dem sozioökonomischen Status (Wendt et al., 2020) leistet die Familiensprache der Kinder einen wichtigen Beitrag zur Erklärung dieser Befunde (Haag et al., 2016), denn die Kompetenzen in der Instruktionssprache sind für den Bildungserfolg von mehrsprachig aufwachsenden Kindern mit Migrationsgeschichte in Deutschland prägend (Edele et al., 2020).

Es existieren Hinweise, dass eine Bildungsbenachteiligung bereits im vorschulischen Alter einsetzt (Weinert et al., 2010). Der bioökologischen Theorie von Bronfenbrenner & Morris (1998) folgend ist die frühkindliche Entwicklung durch das Zusammenspiel der Anregungen und Interaktionen zwischen dem Kind und seinen Lernumgebungen geprägt. Neben der Familie stellt die Kita die zentrale Lernumgebung der meisten Kinder im vorschulischen Alter dar. Die dort ablaufenden Prozesse werden im in der frühkindlichen Bildungsforschung gängigen Struktur-Prozess-Modell pädagogischer Qualität als Teil der Prozessqualität verstanden (Kluczniok, 2018). Zahlreiche Studien deuten auf die hohe Bedeutung von Anregungsprozessen in der Kita und deren Qualität für die Entwicklung und den Bildungserfolg der Kinder hin (Ulferts & Anders, 2016; von Suchodoletz et al., 2023). Insbesondere Kinder mit Migrationsgeschichte profitieren in ihrer sprachlichen Entwicklung von einer hohen Interaktionsqualität in den Einrichtungen (Anders, 2013; Beckh et al., 2014). Allerdings liegen auch Befunde vor, wonach ein hoher Anteil von Kindern mit Migrationsgeschichte in Kitagruppen mit niedrigerer Interaktionsqualität in diesen Gruppen assoziiert ist (Anders et al., 2021; Karoly & Gonzalez, 2011; Kuger & Kluczniok, 2008; Schober et al., 2017). Daraus ergibt sich die Frage, ob Kinder mit geringeren Sprachkenntnissen in der Bildungssprache in unterschiedlichem Ausmaß vom Kita-Angebot profitieren. Hierzu liegen bislang wenig Befunde vor, die in der Kita auf Ebene einzelner Kinder zwischen deren Sprachkenntnissen differenzieren und diese mit der erfahrenen pädagogischen Interaktionsqualität in Beziehung setzen. Die Befunde von Kluczniok und Schmidt (2020) deuten darauf hin, dass die kindliche Interaktionsqualität mit Bezug auf Peers bei Kindern mit nicht-deutscher Familiensprache niedriger ausfällt als bei Kindern mit deutscher Familiensprache. Allerdings wird hier nicht zwischen den verschiedenen sprachlichen Kompetenzen im Deutschen unterschieden.

Mit der vorliegenden Studie soll deshalb untersucht werden, ob sich die von Kindern erfahrene pädagogische Interaktionsqualität in Kitas in Abhängigkeit von ihren sprachlichen Kompetenzen in der/den Bildungssprache/n unterscheidet. Es wird vermutet, dass Kinder mit geringen oder nicht vorhandenen Kenntnissen der in der Kita verwendeten Sprachen eine vergleichsweise geringere pädagogische Interaktionsqualität erfahren und insofern doppelt benachteiligt sind.

Die Datenbasis bildet das Forschungsprojekt „Qualität in bilingualen Kindertagesstätten“. In dem Projekt wird die pädagogische Interaktionsqualität in den Bereichen Sprache und sozial-emotionale Unterstützung in Kitas zweier bilingual (Deutsch-Englisch bzw. Deutsch-Spanisch) arbeitender Träger in Berlin erfasst und mit den über eine Elternbefragung im Reportverfahren erhobenen sprachlichen und sozial-emotionalen Kompetenzen der Kinder in Verbindung gebracht. Die Stichprobe umfasst N = 200 Kinder aus N = 10 Kitas.

Für die Beantwortung der Forschungsfrage eignen sich die Daten aufgrund zweierlei Besonderheiten sehr gut. Zum einen wird die pädagogische Interaktionsqualität in den Kitas nicht wie sonst in der Forschung zur pädagogischen Qualität von Kitas (und Schulen/Unterricht) üblich auf Ebene der gesamten Kita-Gruppe (Klasse) erhoben, sondern für jedes Kind individuell (sog. Zielkindebene). Dadurch lassen sich die sprachlichen Kompetenzen jedes Kindes mit der von diesem Kind individuell erfahrenen Interaktionsqualität in Verbindung setzen. Darüber hinaus weist die Stichprobe wegen des bilingualen Konzeptes der Träger und der internationalen Klientel eine große Varianz in Bezug auf die sprachlichen Fähigkeiten der Kinder in den Kitas gesprochenen Sprachen auf.

Die Beantwortung der Forschungsfrage soll anhand von varianzanalytischen Verfahren erfolgen.

Die Befunde werden vor dem Hintergrund von (doppelter) Bildungsbenachteiligung von Kindern diskutiert. Es werden praktische Implikationen für den Umgang mit Kindern mit geringeren Kenntnissen der in der Kita erfahrenen Bildungssprache(n) abgeleitet.

ENTFÄLLT: Sprachliche und emotionale Kompetenzen bei Kindern mit Sprachentwicklungsstörungen

Christina Kauschke

Philipps-Universität Marburg

Wechselwirkungen zwischen Sprache und Emotion kommen besonders dann zum Tragen, wenn die Fähigkeiten in einem dieser Entwicklungsbereiche nicht altersentsprechend ausgebildet sind. Zum einen können sich Störungen des emotionalen Erlebens im Sprachverhalten niederschlagen, zum anderen können sich sprachlich-kommunikative Einschränkungen negativ auf die sozio-emotionalen Kompetenzen auswirken (Kauschke 2019, 2023). Letzteres ist bei Sprachentwicklungsstörungen (SES) der Fall, die häufig von sozio-emotionalen Auffälligkeiten begleitet werden (Lautenschläger et al. 2020). Ob dies als reine Folgeproblematik der Sprachstörung zu werten ist oder ob die Verarbeitung emotionaler Information bereits auf einer basalen Ebene auffällig ist, ist Gegenstand von Diskussionen. In meinem Beitrag wird ein Forschungsüberblick über sozio-emotionale Probleme, Emotionsverarbeitung und sprachlichen Ausdruck von Emotionen bei Kindern mit SES gegeben. Anhand einer Studie (Bahn et al. 2021) wird aufgezeigt, dass Grundschul Kinder mit SES nicht nur bei der Verarbeitung sprachlich vermittelter emotionaler Information, sondern auch bei nonverbalen Aufgaben zur Emotionsverarbeitung schlechter abschneiden als gleichaltrige sprachunauffällige Kinder. Die Ergebnisse sprechen dafür, dass sich nuancierte Emotionskonzepte in Wechselwirkung mit Sprache herausbilden (Lindquist et al. 2015) und somit bei SES anfällig sind. Darüber hinaus können sich Kinder mit SES aufgrund ihrer sprachlichen Restriktionen weniger differenziert über innere Zustände austauschen. Ob die verbal geäußerte emotionale Granularität durch eine spezifische Anreicherung des Sprachangebots mit Emotionswörtern gesteigert werden kann, wurde in einer Pilotstudie mit 30 Grundschulkindern untersucht, deren Methodik und erste Ergebnisse abschließend präsentiert werden.

Einschätzung von Sprachfähigkeiten ein- und mehrsprachiger Kinder - Welche Normwerte sind geeignet?

Steffi Sachse¹, Christina Kauschke², Julia Siegmüller³, Anne Tenhagen⁴, Tobias Dörfler¹

¹Pädagogische Hochschule Heidelberg, Institut für Psychologie; ²Philipps-Universität Marburg, Institut für Germanistische Sprachwissenschaft; ³Rostock; ⁴Inthera – Interdisziplinäre Lehrpraxis GmbH

Hintergrund

In Deutschland wächst mittlerweile ein substanzieller Anteil an Kindern mehrsprachig auf (Autor:innengruppe Bildungsberichterstattung, 2024). Eine Herausforderung stellt dabei die Diagnostik des sprachlichen Entwicklungsstandes dar, um zwischen Kindern mit Sprachförderbedarf und Kindern mit Sprachentwicklungsstörungen zu differenzieren und geeignete Maßnahmen zu implementieren. In diesem Rahmen stellt sich die Frage, welche Normen angewendet werden sollen, um den Sprachentwicklungsstand ein- und mehrsprachiger Kinder angemessen zu beurteilen (Ehl & Grosche 2021). Vorhandene Sprachentwicklungstests gehen bei der Normierung sehr unterschiedlich mit mehrsprachigen Kindern um: oft wird ein unterschiedlich großer Anteil von Kindern mit Migrationshintergrund pauschal in die Normierungstichproben einbezogen („gemischt-sprachige Normen“), wobei meist unklar ist, in welchem Erwerbskontext diese Kinder aufwachsen und welche Folgen dies nach sich zieht. In anderen Tests basieren die Normwerte nur auf den Daten einsprachig aufwachsender oder ausschließlich auf den Daten mehrsprachiger Kinder. Ob normierte Testverfahren mehrsprachigen Kindern gerecht werden und wie mehrsprachigerwerbssensible Normwerte gestaltet und interpretiert werden sollten, ist Gegenstand kontroverser Diskussionen.

Fragestellung

Es wird untersucht, welche Auswirkungen einsprachige vs. gemischt-sprachige Normwerte auf die diagnostischen Aussagen hinsichtlich des sprachlichen Entwicklungsstandes haben.

Methodik

Im Rahmen der Normierungsstudie der „Patholinguistischen Diagnostik für Sprachentwicklungsstörungen“ (PDSS; Kauschke et al., 2023) wurden 697 einsprachig sowie 241 mehrsprachig aufwachsende Kinder im Alter zwischen 2;6 und 6;11 einbezogen und mit den einzelnen Subtests der PDSS getestet. Auf dieser Datengrundlage wurden zweimal kontinuierliche Normwerte (Lenhard et al., 2021) mit dem R-Package cNORM berechnet: Normen für die Gesamtgruppe der Kinder (gemischt-sprachig) und Normen für Kinder mit der Muttersprache Deutsch (einsprachig). Um die Auswirkungen beider Varianten (gemischt-sprachig vs. einsprachig) zu vergleichen, wurden die Leistungen einsprachiger Kinder mit beiden Normierungsarten bestimmt.

Ergebnisse und ihre Bedeutung

Es fanden sich keine systematischen Gruppenunterschiede zwischen eher simultan und eher sukzessiv mehrsprachig aufwachsenden Kindern über die Altersgruppen und Subtests hinweg. Im Vergleich gemischt-sprachiger vs. einsprachiger

Normen ergab sich, dass die T-Werte für einsprachige Kinder bei Anwendung der gemischt-sprachigen Normwerte im unteren Leistungsbereich deutlich absanken. In einzelnen Subtest betrug die Differenz mehr als eine Standardabweichung. Einsprachige Kinder mit schwachen Sprachleistungen werden anhand einer gemischten Norm somit teilweise nicht mehr als sprachauffällig erkannt.

Die Ergebnisse zeigen, dass die pauschale Einbeziehung mehrsprachiger Kinder in die Normierungsstichprobe eines Sprachtests gravierende Auswirkungen auf die Diagnostik im Einzelfall hat. Für die Identifikation einsprachiger sprachentwicklungsgestörter Kinder sind einsprachige Normen notwendig. „Normen“ in eigentlichen Sinn erscheinen vor dem Hintergrund der Heterogenität mehrsprachigen Aufwachsens eher weniger geeignet. Für die Beurteilung mehrsprachiger Kinder mit der PDSS wurden deshalb „Orientierungswerte“ angegeben. Diese können in den mehrschrittigen gesamt-diagnostischen Prozess bei mehrsprachigen Kindern integriert werden.

Im Vortrag werden die Verzerrungen auf die einzelnen Subtests bei der Anwendung mehrsprachiger Normen diskutiert.

PS-43: Herausforderungen und Ressourcen im Lehrberuf

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: SO 318

Chair der Sitzung: Carmela Aprea

Herausforderungen von Lehrkräften an beruflichen Schulen: Ausprägungen sowie Auswirkungen auf Burnoutgefährdung, Wohlbefinden und Berufszufriedenheit

Nina Marlene Sarochan, Carmela Aprea

Universität Mannheim, Deutschland

Lehrkräfte sind täglich mit zahlreichen Herausforderungen konfrontiert. Die Bewältigung dieser Herausforderungen führt häufig zu gesundheitlichen und psychischen Problemen, da Lehrkräfte hohe Burnout-Raten, ein erhöhtes Stressniveau und gesundheitliche Beeinträchtigungen erleben (Aprea & Lohner, 2019). Während dies für Lehrkräfte an allgemeinbildenden Schulen berichtet wurde, wurde das Arbeitserleben von Lehrkräften im beruflichen Bereich in der Forschung bisher wenig berücksichtigt. Um das Wohlbefinden, die Berufszufriedenheit und das Burnout-Risiko dieser Gruppe von Lehrkräften besser zu verstehen, verwenden wir das „Job Demands-Resources“-Modell (Schaufeli & Taris, 2014). Während das ursprüngliche JD-R-Modell nur berufliche Ressourcen umfasst, werden in der in diesem Beitrag vorgestellten Studie auch persönliche Ressourcen einbezogen. Die Studie adressiert die folgenden Forschungsfragen:

- (1) Welche Herausforderungen nehmen Lehrkräfte an beruflichen Schulen als besonders ausgeprägt in ihrem Arbeitsalltag wahr?
- (2) Welche Auswirkungen haben diese Herausforderungen auf das Wohlbefinden, die Berufszufriedenheit und das Burnout-Risiko der Lehrkräfte?
- (3) Medieren persönliche Ressourcen und Unterstützung aus dem privaten Umfeld sowie berufliche Ressourcen wie die Unterstützung durch Schulleitung und Kollegen die Beziehung zwischen beruflichen Anforderungen und Erlebensmaßen?

Zur Beantwortung dieser Fragen stützen wir uns auf Daten aus einer Online-Befragung, die von Februar bis November 2022 durchgeführt wurde. Die Stichprobe ist repräsentativ für Lehrkräfte an beruflichen Schulen in Baden-Württemberg und umfasst 1.470 Lehrpersonen an beruflichen Schulen (56,9% weiblich, $M_{Alter} = 44,89$, 60,3% Vollzeitlehrkräfte).

Zur Erfassung der Herausforderungen im Arbeitsalltag wurde eine Skala mit 24 Items zu verschiedenen Arten von Herausforderungen, unterteilt in sieben Kategorien, eingesetzt. Als persönliche Ressourcen wurden die Stresserholungsfähigkeit (BRS, Chmitorz et al., 2018), und die Unterstützung durch das private Umfeld betrachtet (Ainsworth & Oldfield, 2019; 5-Punkte-Likert-Skala von „trifft überhaupt nicht zu“ bis „trifft komplett zu“). Die Unterstützung durch Kollegen und die Schulleitung (Ainsworth & Oldfield, 2019) sind die beiden berücksichtigten beruflichen Ressourcen und werden auf einer 5-Punkte-Likert-Skala von „trifft überhaupt nicht zu“ bis „trifft komplett zu“. Als Erlebensmaße wurden das Wohlbefinden (WHO-5, WHO, 1998), die Berufszufriedenheit (gemessen mit einem einzelnen Item; 5-Punkte-Likert-Skala von „sehr unzufrieden“ bis „sehr zufrieden“) und das Burnout-Risiko (BAT, Schaufeli et al., 2020) berücksichtigt.

Die Herausforderungen im Arbeitsalltag wurden zunächst deskriptiv analysiert. Anschließend wurden die Auswirkungen dieser Herausforderungen auf das Wohlbefinden, die Berufszufriedenheit und das Burnout-Risiko mithilfe linearer Regressionen ermittelt. Schließlich wurde mit Hilfe von Mediationsanalysen und Bootstrapping geprüft, welche Rolle die persönlichen und beruflichen Ressourcen in der Beziehung zwischen den beruflichen Herausforderungen und den Erlebensmaßen spielen.

Die Analysen zeigen, dass die am stärksten wahrgenommene Herausforderung für Lehrkräfte die Heterogenität der Leistungsvoraussetzungen der Lernenden ist ($M=4,45$, $SD=.85$). Auf der Ebene der Herausforderungskategorien weist die Kategorie „Aufgabenumfang und -komplexität“ ($M=3,96$, $SD=0,73$) den höchsten Mittelwert auf. Die Regressionsanalysen zeigen für alle Herausforderungen signifikante Effekte auf alle Erlebensmaße. Die Effektstärken sind jedoch meist klein bis mittel ($R^2=.004-.260$). Die Wahrnehmung von Herausforderungen ist negativ mit dem Wohlbefinden ($\beta=-.35$, $p<.001$) und der Berufszufriedenheit ($\beta=-.41$, $p<.001$) von Lehrkräften korreliert und steht in einem positiven Zusammenhang mit dem Burnout-Risiko ($\beta=.51$, $p<.001$). Die Mediationsanalysen ergeben mehrere signifikante Mediationseffekte. Für das Wohlbefinden sind alle vier hypothetisierten Mediatoren signifikant, wobei die Unterstützung durch die Schulleitung die Beziehung vollständig mediert ($\beta_{indir}=-0,28$, $p<.001$, $[-0,430;-0,126]$). Die Analysen zur Berufszufriedenheit zeigen ähnliche Ergebnisse. Für Burnout erweisen sich nur persönliche Ressourcen als signifikante Mediatoren. Die Mediationsanalysen zeigen die Bedeutung der Stresserholungsfähigkeit und Unterstützung durch das private Umfeld. Darüber hinaus spielt die Unterstützung durch die Schulleitung eine wichtige Rolle.

Obwohl diese Studie Einschränkungen aufweist, wie z. B. das Querschnittsdesign, tragen ihre Ergebnisse dazu bei, die Lücke im Verständnis der Arbeitsrealität von Lehrkräften an beruflichen Schulen zu schließen, indem sie relevante Herausforderungen und deren Auswirkungen auf das Wohlbefinden der Lehrkräfte, die Berufszufriedenheit und das Burnout-Risiko aufzeigt. Außerdem lassen sich daraus Schlussfolgerungen für die Ausbildung sowie die Arbeitsgestaltung und -organisation von Berufsschullehrkräften ziehen.

Lehrkräftewohlbefinden – ein Conceptual Systematic Review (2020-2023)

Laura Kurrle, Julia Warwas

Universität Hohenheim, Deutschland

Dass Lehrkräftewohlbefinden relevant für das gesamte Schulwesen ist, legen Forschungsergebnisse nahe, die dieses Konzept mit der Effizienz der Lehre, der Bleibeabsicht der Lehrkräfte sowie dem Lernerfolg und dem Wohlbefinden der Schüler:innen in Zusammenhang bringen (Beames et al., 2023; Duckworth et al., 2009; Hascher & Waber, 2021, S. 1; McCallum et al., 2017; Sisask et al., 2014; Sutton & Wheatley, 2003; Watt & Richardson, 2013). Diese Relevanzzuschreibung spiegelt sich in einer seit zwei Jahrzehnten steigenden Publikationszahl wider, ging aber nicht mit wachsender konzeptioneller Klarheit einher. Anstelle einer einheitlichen Definition des Lehrkräftewohlbefindens besteht lediglich Konsens in dessen Multidimensionalität, wobei auch hier vielfältige Annahmen bezüglich seiner konstitutiven Merkmale und Beziehungen zu anderen Konzepten bestehen (Dreer, 2023; Hascher & Waber, 2021; McCallum et al., 2017). In Studien kommt die theoretische Fundierung der jeweils betrachteten Subdimensionen von Lehrkräftewohlbefinden häufig zu kurz, während die auf Basis eines spezifischen dimensionalen Schwerpunktes ermittelten Ergebnisse oft als gültig für das Gesamtkonzept unterstellt werden (Hascher & Waber, 2021). Die Vielfalt der theoretischen (und operationalen) Zugänge führt aber de facto zu einer mangelnden Vergleichbarkeit der Ergebnisse (Cumming, 2017; Hascher & Waber, 2021).

Trotz bestehender, durchaus kritischer, Forschungssichtungen zum Lehrkräftewohlbefinden (Dreer, 2023; McCallum et al., 2017), steht eine systematische Übersichtsarbeit aus, die sich tiefgehend mit theoriegeleiteten Begrifflichkeiten und Konnotationen

sowie Dimensionsstrukturen und Determinanten von Lehrkräftewohlbefinden beschäftigt. Diese Lücke möchte das vorliegende Forschungsvorhaben mit der Fragestellung schließen: *Welche konzeptuellen Modelle von Lehrkräftewohlbefinden liegen veröffentlichten Studien zugrunde?*

Zur Beantwortung wird ein *Conceptual Systematic Review* nach Schreiber und Cramer (2022) durchgeführt. Hierfür wurde zuerst ein *Coding Framework* deduktiv auf Basis des OECD-Rahmens von Viac und Fraser (2020) entwickelt, welcher zwischen *Arbeitsbedingungen* von Lehrkräften, *Komponenten* ihres Wohlbefindens sowie *Ergebnissen* hohen bzw. mangelnden Wohlbefindens differenziert. Im Zuge induktiver Erweiterungen wurden u.a. der Bereich der Arbeitsbedingungen auf *Rahmenbedingungen* ausgeweitet und in die Dimensionen der *Qualität der Arbeitsumgebung*, des *Bildungssystems* und der *sozio-kulturellen Bedingungen* untergliedert. Die internalen Komponenten umfassen nun die Dimensionen: *selbstbezogene Bewertung*, *kontextbezogene Bewertung*, *beziehungsbezogene Bewertung*, *emotionale Befindlichkeit* und *psychosomatische Symptome*. Unter Ergebnisse lassen sich wiederum *inwendige Ergebnisse* sowie *äußere Ergebnisse* subsumieren.

Die Literaturrecherche über die Datenbanken EBSCOhost und peDOCS ergab auf Basis des Suchstrings (*Lehrkräfte OR Lehrpersonen OR Lehr**) AND (*Wohlbefinden OR Lehrkräftewohlbefinden*) und dem englischen Äquivalent (*teacher OR educator*) AND (*wellbeing OR well-being*) 1716 Treffer für die Jahre 2020 bis 2023. Der Zeitraum orientiert sich am Veröffentlichungsdatum des OECD-Rahmens. Eingeschlossen wurden auch Dissertationen und Artikel ohne peer-review-Verfahren, da primär theoriegeleitete Konzepte und Modellannahmen zum Lehrkräftewohlbefinden interessieren und insbes. in thematisch fokussierten Dissertationen im Regelfall ausführlich dargelegt werden. Durch Ausschluss aller Werke, die nicht deutsch- oder englischsprachig verfasst sind, das Konstrukt nicht explizit beschreiben oder nicht im schulischen Primar- und Sekundarbereich lokalisiert sind, werden 187 Arbeiten näher betrachtet.

Im Zuge der laufenden Kodierarbeiten deuten erste Erkenntnisse auf eine Dominanz des *Job-Demands-Resources Modells* (JD-R), des *PERMA-Modells* sowie der definitorisch etablierten Konzepte *subjektives Wohlbefinden* (SWB) und *psychologisches Wohlbefinden* (PWB) als theoretische Grundlagen der berücksichtigten Studien hin. Dementsprechend bewegt sich ein Großteil der Forschung im Bereich der (annahmegemäß determinierenden) Rahmenbedingungen (JD-R) sowie in den Bewertungsdimensionen innerhalb des Bereichs der internalen Komponenten (PERMA, SWB und PWB). Dahingegen scheint es noch wenig Studien zu den psychosomatischen Symptomen des Lehrkräftewohlbefindens zu geben.

Mithilfe des *Conceptual Systematic Reviews* wird letztlich eine Übersicht über dominierende und randständige Verwendungen des vielschichtigen Konzepts des Lehrkräftewohlbefindens entstehen, klassifiziert in analytische Bereiche, zentrale Dimensionen und hierarchisch ausdifferenzierte Subdimensionen. Dies erleichtert nicht nur die Bewertung der Vergleichbarkeit empirischer Befunde (bezogen auf deren konzeptionelle und operationale Basis). Vielmehr ermöglicht es eine erhöhte konzeptionelle Klarheit den Forschenden, Forschungslücken und interessante Spezialbereiche für Meta-Analysen zu identifizieren sowie bei der Planung eigener Studien einen klaren und konsistenten konzeptionellen Zuschnitt zu verfolgen.

Selbstbewusst ins Risiko: Kurvilineare Effekte der Selbstwirksamkeit für das Beanspruchungserleben und die Unterrichtsqualität von Lehrkräften im Vorbereitungsdienst

Hannah Kirschning, Friederike Zimmermann

Christian-Albrechts-Universität zu Kiel, Deutschland

Theoretischer Hintergrund

Kompetente Lehrkräfte sind essentiell für guten Unterricht (Hattie, 2009). Eine relevante motivationale Komponente der professionellen Kompetenz ist die Selbstwirksamkeit von Lehrkräften, die die Überzeugung umfasst, auch unter herausfordernden Bedingungen erforderliche Handlungen der zukünftigen Tätigkeit erfolgreich ausführen zu können (Bandura, 1997; Schwarzer & Hallum, 2008). Dabei müssen Wirkmechanismen der Selbstwirksamkeitserwartungen vertieft untersucht werden: Denn einerseits zeigen Studien linear positive Assoziationen zwischen Selbstwirksamkeit und einem ausgewogenen Beanspruchungserleben der Lehrkräfte sowie längerfristiger Aufrechterhaltung qualitativen Unterrichts (Aloe et al., 2014; Klassen & Tze, 2014; Zee & Koomen, 2016). Andererseits weisen Studien darauf hin, dass Selbstwirksamkeitserwartungen psychologische Stressreaktionen folgen können sowie leistungsmindernde Effekte möglich sind, die beispielsweise durch reduzierte Anstrengungsbereitschaft erklärbar sind (Schönfeld et al., 2017; Sitzmann & Yeo, 2013; Vancouver et al., 2014). Diese uneinheitlichen Befunde werfen die Frage auf, ob positive Effekte auf Beanspruchungserleben und Unterrichtsqualität gemindert sind, wenn Lehrkräfte über ein zu hohes Maß an Selbstwirksamkeit verfügen und somit ihre Fähigkeiten, zukünftigen Herausforderungen zu begegnen, überschätzen. Die vorliegende Studie untersucht mögliche kurvilineare Effekte der Selbstwirksamkeitserwartungen von Lehrkräften am Ausbildungsende für deren Beanspruchungserleben und Unterrichtsqualität am Berufseinstieg.

Fragestellung

Inwiefern zeigen sich kurvilineare Effekte der Selbstwirksamkeitserwartungen von Lehrkräften am Ausbildungsende auf deren Beanspruchungserleben (Work Engagement, Burnout) und Unterrichtsqualität (kognitive Aktivierung, Klassenführung, konstruktive Unterstützung) zu Beginn ihrer beruflichen Laufbahn?

Methode

Datengrundlage sind Erhebungen des Panels zum Lehramtsstudium (PaLea). Die bundesweite Stichprobe umfasst $N = 55$ Lehrkräfte (80% weiblich) und ihre $N = 2453$ Schüler:innen (66% weiblich; Klassenstufe: $M = 8.57$; $SD = 1.98$; Range = 5-12). Die längsschnittliche Panel-Befragung der (angehenden) Lehrkräfte von Studienbeginn bis zur Berufstätigkeit wurde auf Personen mit erfolgter Teilnahme an der angeschlossenen Schülerbefragung (mit ihren Klassen) in ihrer Berufstätigkeit begrenzt. Pro Lehrkraft waren dies durchschnittlich 45 Schüler:innen, verteilt auf ein bis drei Klassen der Klassenstufen 5-12; insgesamt 155 Klassen aus 55 weiterführenden Schulen; davon 65% Gymnasien. Zwischen beiden Messzeitpunkten (Ende Vorbereitungsdienst; erstes Berufsjahr) lagen durchschnittlich 20 Monate.

Lehrkräfte schätzen die Lehrkraft-Selbstwirksamkeit am Ausbildungsende ein (Schwarzer & Jerusalem, 1999).

Basisdimensionen der Unterrichtsqualität (kognitive Aktivierung, Klassenführung, konstruktive Unterstützung) bewerteten die Schüler:innen (Roloff et al., 2016).

Für berufliches Beanspruchungserleben der tätigen Lehrkräfte wurden Burnout und Work Engagement im ersten Berufsjahr erfasst (Maslach & Jackson, 1996; Schaufeli & Bakker, 2003).

Ergebnisse und ihre Bedeutung

Latente (Mehrebenen-) Regressionsmodelle kontrolliert für Kovariaten (Geschlecht, Alter) zeigten wesentlich, dass die Beziehung von Selbstwirksamkeit und Work Engagement, Burnout sowie Klassenführung kurvilinear sind. Selbstwirksamkeit sagte positiv Work Engagement und Klassenführung und negativ Burnout vorher; $\beta = [.27 \text{ bis } .45]$, $p < .01$. Der positive Zusammenhang zwischen Selbstwirksamkeit und Work Engagement sowie Klassenführung nimmt ab, wenn die quadratischen Koeffizienten berücksichtigt werden; wohingegen der negative Zusammenhang zwischen Selbstwirksamkeit und Burnout zunimmt; $B = [.0.99 \text{ bis } .1.44]$, $p < .05$. Die D-Test-Statistik zeigte, dass das kurvilineare Model von Selbstwirksamkeit signifikant mehr Varianz von Work Engagement sowie Burnout aufklärte; so nicht bei Klassenführung.

Die kurvilineare Beziehung von Selbstwirksamkeit mit Work Engagement, Burnout und Klassenführung zeigt, dass ein zu hohes Ausmaß an Selbstwirksamkeitserwartungen negative Folgen haben kann. Es ist zu vermuten, dass Lehrkräfte bis zu einem Schwellenwert von einer realistischen Selbstwirksamkeitseinschätzung profitieren, jedoch bei zu hoher Ausprägung die förderlichen Effekte auf das Beanspruchungserleben und die Klassenführung abnehmen. Dies beeinflusst die wahrgenommene Unterrichtsqualität der Schüler:innen nicht. Diese Arbeit leistet einen Beitrag zur Forschungsfrage, ob eine Überschätzung der Fähigkeiten frisch ausgebildeter Lehrkräfte für deren Unterrichtsqualität und Beanspruchungserleben am Berufseinstieg eine Rolle spielt. Die Ergebnisse stellen die Linearitätsannahme in Frage und tragen dazu bei, sich widersprüchlichen Ergebnissen früherer Untersuchungen anzunähern. Weitere Analysen zur Feststellung erhöhter Selbstwirksamkeitsmaße im Ausbildungsverlauf der Lehrkräfte sind ausstehend. Diese Studie gibt einen Anstoß, berufliche Anforderungen und Selbstwahrnehmung der persönlichen Fähigkeiten in Lehrveranstaltungen und im Vorbereitungsdienst der Lehrkräftebildung zu reflektieren.

Förderung der Entwicklung von Handlungsstrategien bei herausforderndem Schüler*innenverhalten

Sarah Burkhardt, Anja Böhnke, Felicitas Thiel

Freie Universität Berlin, Deutschland

Theoretischer Hintergrund

Der Umgang mit Unterrichtsstörungen ist elementarer Bestandteil eines guten Klassenmanagements von Lehrkräften (Ophardt & Thiel, 2013; Thiel, 2016). Dazu gehört der Umgang mit individuellem herausforderndem Schüler*innenverhalten, was sich in verschiedenen Verhaltensweisen zeigt, z.B. in Form einer „Abwertung der Lehrkraft“ oder der „Vermeidung von Herausforderungen“ (Thiel, 2016). Den Verhaltensweisen können unterschiedliche Ursachen und Motive zugrunde liegen (Thiel, 2016). Die Verweigerung einer Aufgabebearbeitung kann z.B. als Folge der Vermeidung eines antizipierten Misserfolgs auftreten (Thiel, 2016). Die Funktion des Verhaltens, also der Nutzen, den das Verhalten für Schüler*innen hat, bedingt, wie mit Unterrichtsstörungen umgegangen werden kann (Gresham et al., 2001). Die funktionale Verhaltensanalyse (FVA) nach Kanfer und Saslow (1969) kann (angehenden) Lehrkräften helfen, herausforderndes Verhalten hinsichtlich der bedingenden und aufrechterhaltenden Faktoren (z.B. situative Auslöser) sowie dessen Funktion zu beurteilen (Borgmeier et al., 2015; Brodersen et al., 2022). Die FVA ermöglicht außerdem die Entwicklung funktionsbasierter Handlungsstrategien, die direkt an den störungskritischen Faktoren ansetzen (Brodersen et al., 2022; Gresham et al., 2001). Die Anpassung einer Aufgabenstellung verändert z.B. den situativen Auslöser eines Verhaltens von Schüler*innen, die eine herausfordernde Aufgabe aufgrund eines befürchteten Misserfolgs verweigern. Ein Klassenverweis würde hingegen das dysfunktionale Verhalten aufrechterhalten, da die Schüler*innen die Aufgabe dadurch weiter vermeiden können (Borgmeier et al., 2015). Obwohl es eine Vielfalt an Handlungsstrategien bei herausforderndem Schüler*innenverhalten gibt (Landrum & Kauffman, 2006), fällt es insbesondere Noviz*innen schwer, adäquate Strategien auf Basis der identifizierten Funktion des Verhaltens einzusetzen (Freeman et al., 2014; Gallagher, 2009). Für angehende Lehrkräfte fehlen bisher niederschwellige Angebote zur Förderung der Kompetenzen, herausforderndes Schüler*innenverhalten hinsichtlich der störungskritischen Faktoren zu beurteilen und funktionsbasierte Handlungsstrategien zu entwickeln (McCahill et al., 2014). Dieser Beitrag fokussiert die Frage, inwiefern diese Kompetenzen von Lehramtsstudierenden durch den Einsatz der FVA im Rahmen eines digitalen Selbstlernmoduls gefördert werden können.

Methode

317 Bachelor-Lehramtsstudierende im zweiten Semester bearbeiteten eine Online-Selbstlernumgebung, in der Inhalte zur FVA sowie zur Entwicklung von funktionsbasierten Handlungsstrategien vermittelt und eingeübt wurden. Zur Erfassung der Kompetenzen wurden fünf *staged* Unterrichtsvideofälle in randomisierter Reihenfolge bearbeitet, in denen individuelle herausfordernde Schüler*innenverhaltensweisen (z.B. „Vermeidung von Herausforderungen“) dargestellt waren (Kilbury et al., 2023). Nach jedem Videofall beschrieben die Studierenden in einem offenen Antwortformat die wahrgenommenen störungskritischen Faktoren des gezeigten Verhaltens („Auslöser“, „Reaktion“, „Konsequenzen“, „Funktion“) sowie die von Ihnen am ehesten gewählte Handlungsstrategie in der gezeigten Unterrichtssituation. Das Scoring der Antworten erfolgte mittels eines Kodiermanuals durch geschulte Rater:innen (Cohen's Kappa $\kappa = .79$, $p < .011$). Das Kodiermanual wurde von Expert*innen validiert und enthielt Indikatoren für relevante störungskritische Faktoren (z.B. „Finn stellt Sinnhaftigkeit der Aufgabe in Frage“) sowie für funktionsbasierte Handlungsstrategien (z.B. „Adaption der Aufgabenstellung“) in den gezeigten Unterrichtsvideofällen. Nach einer allgemeinen theoretischen Einführung zum Klassenmanagement und zu Unterrichtsstörungen für alle Teilnehmenden erfolgte die Erhebung in einem randomisiert kontrollierten Design, sodass die Interventionsgruppe zunächst die Online-Selbstlernumgebung zur FVA durchlief und anschließend die Videofälle bearbeitete und die Kontrollgruppe zuerst die Videofälle bearbeitete und anschließend die Online-Selbstlernumgebung zur FVA durchlief.

Ergebnisse

Für vier von fünf herausfordernden Schüler*innenverhaltensweisen liegen bereits Ergebnisse hinsichtlich der identifizierten störungskritischen Faktoren vor. Für die Verhaltensweisen „Abwertung der Lehrkraft“ (EG: $M=3.73$; $SD=2.01$; KG: $M=2.53$; $SD=1.67$), „Suche nach Anerkennung von Peers“ (EG: $M=3.61$; $SD=1.78$; KG: $M=2.37$; $SD=1.97$), „Vermeidung von Herausforderungen“ (EG: $M=4.16$; $SD=1.71$; KG: $M=2.23$; $SD=1.64$) und „Reaktantes Verhalten“ (EG: $M=3.49$; $SD=2.03$; KG: $M=2.22$; $SD=1.62$) fand sich ein Haupteffekt der Gruppen ($F(1, 288)=74.71$, $p < .001$; partielles $\eta^2 = .21$) sowie ein Interaktionseffekt der Gruppen mit den herausfordernden Verhaltensweisen ($F(3, 864)=4.92$, $p = .002$, partielles $\eta^2 = .02$). Aktuell werden die Ergebnisse der entwickelten funktionsbasierten Handlungsstrategien ausgewertet. Die vorläufigen Ergebnisse zeigen, dass bereits angehende Lehrkräfte durch den Einsatz der FVA herausforderndes Schüler*innenverhalten besser reflektieren können. Alle Ergebnisse werden in Hinblick auf die Kompetenzförderung zur Beurteilung störungskritischer Faktoren sowie zur Entwicklung funktionsbasierter Handlungsstrategien bei herausforderndem Schüler*innenverhalten diskutiert.

PS-44: Wahrnehmung und Wirkung von digitalem Feedback im Lernprozess

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: SO 418

Chair der Sitzung: Marlit Annalena Lindner

Visuelle Verarbeitungsprozesse von elaborierten Feedbackbotschaften nach richtigen und falschen Antworten: Eine Blickbewegungsstudie

Thérèse Eder¹, Livia Kuklick², Marlit Annalena Lindner^{1,2,3}

¹Leibniz-Institut für Wissensmedien, Deutschland; ²Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel; ³Europa-Universität Flensburg

Hintergrund und Hypothesen

Computerbasiertes Feedback ist ein zentraler Faktor für adaptives Lehren und Lernen. In digitalen Kontexten hat sich elaboriertes Aufgabenfeedback (EF) als besonders lernförderlich erwiesen (Mertens et al., 2022; Van der Kleij et al., 2015), da EF neben der korrekten Antwort weitere nützliche Informationen enthält, wie beispielsweise einen Lösungsweg oder bildliches Material. Jedoch lassen sich nicht immer positive Effekte von EF auf das Lernen nachweisen (Kuklick et al., 2023), was daran liegen könnte, dass Lernende die bereitgestellten Informationen nicht immer hinreichend verarbeiten.

Die kognitiven Verarbeitungsprozesse von Feedback wurden bisher kaum detailliert untersucht. Cutumisu et al. (2019) zeigten via Eyetracking im Kontext eines Lernspiels mit Feedback, dass sich Aufmerksamkeitsprozesse bei der Feedbackverarbeitung unterschieden, je nachdem, ob Lernende eine Aufgabe richtig oder falsch bearbeiteten. Zentrale Informationen des Feedbacks wurden nach falschen Antworten länger (Fixationen) und häufiger (Revisits) betrachtet. In unserer Studie knüpfen wir daran an und untersuchen im Kontext von Geometrieaufgaben wie sich das Blickverhalten von Testbearbeitenden bei der Verarbeitung von elaborierten Feedbackbotschaften in Abhängigkeit von der individuellen Antwortrichtigkeit verändert und inwiefern dies das Lernen aus Fehlern vorhersagen kann.

Wir erwarten unter anderem (präregistriert), dass die Versuchspersonen den bereitgestellten Lösungsweg insgesamt länger (H1) anschauen, wenn sie die Aufgabe falsch anstatt richtig beantworteten. Ferner gehen wir davon aus, dass die Versuchspersonen nach falschen Antworten häufiger zum Lösungsweg zurückblicken (Revisits), als nach richtigen Antworten (H2). Wir erwarten zudem, dass die Fixationsdauer des Lösungswegs und die Betrachtung der Visualisierung (in Form des repräsentationalen Bildes) in der Feedbackbotschaft nach falschen Antworten positiv mit der Fehlerkorrektur im Recall Post-Test (H3a) und der Leistung im Transfertest (H3b) zusammenhängt.

Studiensetting und Methode

Wir untersuchten 84 Studierende im Labor, die einen computerbasierten Test mit 12 Geometrieaufgaben und automatisiertem Feedback bearbeiteten. Unmittelbar nach jeder Aufgabe wurde eine elaborierte Feedbacknachricht mit mehreren Elementen (z.B. geometrische Visualisierung, Rückmeldungssatz, Lösungsweg) angezeigt. Das Feedback war entweder positiv (richtige Antwort) oder negativ (falsche Antwort). Die Variation der Antwortrichtigkeit im Treatment-Test erlaubte eine differenzierte Analyse der Verarbeitung des positiven und negativen Feedbacks im Sinne eines Ex-Post Facto quasi-experimentellen Within-Subject-Designs. Nach dem Treatment-Test mit Feedback wurde ein Post-Test mit 12 identischen Recall-Aufgaben und 12 Transfer-Aufgaben bearbeitet. Die Blickbewegungen der Studierenden wurden via videobasiertem Remote Eye Tracking (250Hz) aufgezeichnet.

Ergebnisse und ihre Interpretation

Zur Analyse der Daten nutzten wir (generalized) linear mixed-effects Modelle. Dabei betrachten wir das Blickverhalten auf den Feedbackbotschaften getrennt für richtige und falsche Testantworten bzw. inwiefern die Aufmerksamkeitsverteilung nach falschen Antworten mit dem Abschneiden im Recall- und Transfer-Posttest zusammenhängt.

Für das Blickverhalten zeigte sich erwartungsgemäß, dass die Studierenden Informationen zum Lösungsweg nach falschen Antworten länger betrachteten (H1: $t(681) = -14.59, p < .001$) und öfter dorthin zurück blickten (H2: $t(610) = -12.22, p < .001$) als nach richtigen Antworten. Eine falsche Beantwortung ging demnach mit einer intensiveren Verarbeitung der Information zum Lösungsweg der Geometrieaufgaben einher. Wir fanden allerdings keinen signifikanten Zusammenhang der Fixationszeit des Lösungswegs und der Visualisierung nach falschen Antworten mit der Recall-Leistung (H3a: $z = -1.59, p = .11$), jedoch einen überraschenden negativen Zusammenhang mit der Transferleistung (H3b: $z = -2.29, p = .02$). Längere Fixationszeiten auf dem Lösungsweg und der Abbildung waren mit einer schlechteren Transferleistung assoziiert. Dieser Zusammenhang ist jedoch nicht kausal mit Bezug auf die Feedbackverarbeitung zu interpretieren. Er deutet vielmehr darauf hin, dass besonders fähige Personen die Feedbackinhalte effizienter verarbeiteten, die aufgrund ihrer höheren Fähigkeit auch bessere Transferleistungen erreichten. Allgemein indizieren unsere Befunde, dass die Aufmerksamkeitsverteilung auf den Feedbackelementen von der Antwortrichtigkeit abhängt. Eine längere Betrachtung relevanter Elemente der Feedbackbotschaft führte aber nicht unmittelbar zu einer besseren Leistung im Post-Test, was neue Forschungsfragen aufwirft. Zusätzliche differenzierte Befunde und weitere Interpretationen der Daten werden auf der Konferenz vorgestellt.

Zum Wunsch nach computergestütztem Leistungsfeedback in Low-Stakes Assessments: Die Perspektive von Schülerinnen und Schülern

Marlit Annalena Lindner^{1,2}, Ute Mertens¹, Livia Kuklick¹, Christian Schöber^{3,4}, Steve Wise⁵

¹IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Kiel; ²EUF - Europa-Universität Flensburg;

³IfBQ - Institut für Bildungsmonitoring und Qualitätsentwicklung, Hamburg; ⁴IQSH - Institut für Qualitätsentwicklung an Schulen Schleswig-Holstein, Kiel; ⁵Measurement Consultant, Portland, Oregon, USA

Hintergrund und Fragestellung

Viele Studien und Metaanalysen zeigen positive kognitive Effekte von computerbasierten Leistungsrückmeldungen (Mertens et al., 2022; van der Kleij, 2015). Automatisiertes Feedback wird bisher allerdings überwiegend in formativen Tests und Tutoring-Systemen eingesetzt, um Lernprozesse zu unterstützen. Automatisiertes Leistungsfeedback könnte jedoch gerade in Large-Scale-Assessments (LSA) wertvoll sein, in denen Schülerinnen und Schüler (SuS) meist keine Rückmeldung zu ihren Testergebnissen erhalten. Die Vertraulichkeit der Testaufgaben in großen Bildungsstudien zum Leistungsmonitoring stellt ein zentrales Problem dar, das es erschwert, SuS nach dem Test aufgabenbezogenes Feedback zu geben. Computergestütztes Feedback könnte hingegen nach jeder Testaufgabe, verzögert am Ende eines Testabschnitts oder am Ende des gesamten Tests, bereitgestellt werden. Diese Form des Feedbacks wäre vor allem geeignet, um zu verhindern, dass sich SuS beispielsweise durch

Distraktoren in Multiple-Choice-Aufgaben falsche Antworten merken oder Fehlkonzepte entwickeln (Butler & Roediger, 2008), was zukünftiges Lernen behindern könnte. Feedback könnte auch das situative Interesse der SuS fördern (Finn, 2015), da die Rückmeldung verdeutlicht, dass die Antworten von Bedeutung sind und eine metakognitive Reflektion anstoßen. Dennoch kann Feedback unterschiedlich auf SuS wirken, abhängig von Vorwissen und Testleistung. Studien zeigen, dass negatives Feedback nach falschen Antworten den affektiv-motivationalen Zustand negativ beeinflussen kann, auch wenn es die spätere Leistung verbessert (z.B. Kuklick & Lindner, 2021, 2023). Dies wirft die Frage auf, inwieweit automatisiertes Feedback für SuS vorteilhaft wäre und wie Feedback bereitgestellt werden sollte, um negative affektive Reaktionen zu vermeiden. Hierfür ist es unerlässlich, die Perspektive der SuS zu verstehen und ihren wahrgenommenen Feedbackbedarf, ihre Gründe für oder gegen den Wunsch nach Feedback und ihre bevorzugte Art der Feedbackbereitstellung zu ergründen. Hierfür haben wir Daten aus Selbstberichten und Leistungsmaßen in zwei verschiedenen Low-Stakes-Assessments erhoben und untersucht, um entsprechende Einblicke in Präferenzen von SuS hinsichtlich des Erhalts von automatisiertem Feedback zu gewinnen.

Methode

Basierend auf einem spezifisch hierfür konstruierten Fragebogen gaben SuS Auskunft zu ihrer Feedbackpräferenz nach Teilnahme an einem der zwei unabhängigen computergestützten Low-Stakes-Assessments (ohne Feedback). Ziel war es, (a) das hypothetische Interesse an Leistungsfeedback, (b) die affektiv-motivationalen und kognitiven Nutzenwahrnehmungen von Feedback und (c) die bevorzugte Art des computergestützten Feedbacks (z.B. sofort, verzögert) besser zu verstehen. In Stichprobe 1 füllten $n=285$ SuS (Klasse 5–7; $M_{\text{age}}=11.21$; $SD_{\text{age}}=0.89$; 46.8% weiblich) aus verschiedenen Schulen den Fragebogen zum Feedbackwunsch aus, nachdem sie mathematische und naturwissenschaftliche Aufgaben in einer Forschungsstudie bearbeiteten. In Stichprobe 2 füllten $n=731$ zufällig ausgewählte SuS (Klasse 5; 47.3% weiblich) aus mehreren Schulen in einem deutschen Stadtstaat den Fragebogen aus, nachdem sie an einem flächendeckenden Low-Stakes-Assessment teilnahmen, das ihren Lernfortschritt in verschiedenen Bildungsbereichen erfasste. Weitere Hintergrundfaktoren (u.a. Schulform, Geschlecht) und Testergebnisse wurden erhoben.

Ausgewählte Ergebnisse und Diskussion

Insgesamt ergab sich ein homogenes Bild über beide Stichproben hinweg. Die Mehrheit der SuS ($S1=66.3\%$; $S2=74.9\%$) äußerte den Wunsch, Feedback zu erhalten und mehr über ihre Testergebnisse zu erfahren. Nur wenige SuS ($S1=8.5\%$; $S2=6.9\%$) bevorzugten kein Feedback; eine kleinere Gruppe war indifferent gegenüber dem Erhalt von Feedback ($S1=24.9\%$; $S2=18.3\%$). Insgesamt gaben die meisten SuS an, dass sie sich durch das Feedback gut fühlten und kognitiv davon profitieren würden. Erwartungsgemäß war der Wunsch nach Feedback bei leistungsstärkeren SuS an Gymnasien etwas stärker ($ps<.05$). Eine signifikante, aber geringe Korrelation zeigte in beiden Stichproben, dass SuS mit niedrigeren Leistungen im Test tendenziell weniger Interesse an dem Feedback hatten ($r_{S1}=.19$, $p=.002$ / $r_{S2}=.12$, $p=.001$). Ein Großteil der SuS bevorzugte Feedback nach jeder Aufgabe, ein weiterer Teil am Ende des Tests. Trotz gewisser Unterschiede waren die Präferenzen der SuS zugunsten von Feedback relativ konsistent über Leistungsniveaus, Geschlecht, Schultypen und Datensätze hinweg. Detaillierte Analysen zu den Schülerhintergrundfaktoren werden auf der Konferenz vorgestellt. Die Daten legen insgesamt nahe, dass Testkonstrukteure und Low-Stakes-Administratoren den Wunsch nach Leistungsrückmeldungen in computergestützten Tests ernsthaft berücksichtigen sollten.

Lernen mit Tutorials: Retrospektive Prompts fördern Lernprozesse und Lernerfolg

Simon A. Schriek, Kirsten Berthold, Markus H. Hefter

Universität Bielefeld, Deutschland

Theoretischer Hintergrund

Videotutorials beinhalten typischerweise Schritt-für-Schritt-Anweisungen und Erklärungen (van der Meij & van der Meij, 2015). Ein Vorteil von Videotutorials ist, dass Lernende eine Grundlage von deklarativem Wissen über Arbeitsschritte erhalten, bevor sie eine bestimmte Fähigkeit erwerben. Gemäß der ACT*-Theorie ist diese Grundlage von deklarativem Wissen essenziell für den Fähigkeitserwerb (Anderson, 1982). Videotutorials können jedoch insbesondere Lernende ohne Vorkenntnisse kognitiv überfordern, da vermittelte Informationen flüchtig sind (Singh et al., 2012) und die Videotutorials verführerische Details beinhalten können (Garner et al., 1989). Eine textbasierte Darstellung von Informationen hingegen ist weniger komplex und könnte der Gefahr einer kognitiven Überfrachtung entgegenwirken (Kalyuga, 2012). Vorherige Forschung zeigte jedoch keinen Unterschied im Lernerfolg beim Vergleich von video- und textbasierter Vermittlung von Lerninhalten, wenn Lernende durch Leitfragen zu einer tiefen Verarbeitung der Lerninhalte angeregt werden (Heffer et al., 2019). Leitfragen (oder Prompts) werden Lernenden typischerweise nach einem Lerninhalt, also als retrospektiver Prompt, zur Verfügung gestellt (z.B. Hefter & Berthold, 2020). Eine tiefe Verarbeitung kann auch durch eine Instruktion zur Antizipation hervorgerufen werden, als antizipatorischer Prompt (Meier et al., 2023). Vorherige Forschung deutet darauf hin, dass Lernende Vorwissen brauchen, um antizipieren zu können (Renkl, 1997; Meier et al. 2023). Vor diesem Hintergrund wollten wir analysieren, welchen Einfluss die Präsentationsform (Video vs. Text und Screenshots) und unterschiedliche Prompttypen auf Lernprozesse und Lernerfolg beim Lernen mit Tutorials nehmen.

Fragestellung

Folgende Hypothesen wurden aufgestellt:

- H1: Retrospektive Prompts führen zu ausgeprägteren Lernprozessen im Vergleich zu antizipatorischen Prompts und allgemeinen Notizprompts.
- H2: Retrospektive Prompts führen zu einem größeren Lernerfolg im Vergleich zu antizipatorischen Prompts, allgemeinen Notizprompts und keinen Prompts.
- H3: Die Ausprägung des Lernprozesses mediiert den Effekt des Prompttyps auf den Lernerfolg.

Zusätzlich wurde die folgende Forschungsfrage formuliert:

- Nimmt die Präsentationsform (Videos vs. Text und Screenshots) Einfluss auf Lernprozesse und Lernerfolg?

Methode

Unsere digitale Trainingsintervention befasste sich mit dem Thema „Arbeiten mit einer industriellen Bohrmaschine“. Wir stellten dieses Training insgesamt 206 Schüler*innen deutscher weiterführenden Schulen zur Verfügung ($M_{\text{Alter}} = 15.94$, $SD_{\text{Alter}} = 1.02$). Die Trainingsintervention fand in den Räumlichkeiten der Schulen statt und bestand aus drei Multimedia-Tutorials in denen ein Experte erklärt, wie man mit einer industriellen Bohrmaschine arbeitet. Erklärungen zur Bedienung der Bohrmaschine wurden entweder in Form von Videos oder in Form von Screenshots und Text präsentiert. Wir kombinierten die Tutorials entweder mit retrospektiven Prompts, antizipatorischen Prompts, allgemeinen Notizprompts oder gar keinen Prompts. Die Schüler*innen

wurden in diesem 2x4-Design randomisiert einer der Bedingungen zugewiesen. Die Ausprägung der Lernprozesse wurde durch die Anzahl richtig genannter Arbeitsschritte in den Antworten auf die Prompts bewertet. Der Lernerfolg wurde durch einen freien Abruf der Arbeitsschritte in einem unmittelbaren sowie verzögerten Nachtest erfasst.

Ergebnisse und ihre Bedeutung

Für H1 und H2 führten wir ANOVAS und Kontrastanalysen durch und fanden die erwarteten Effekte: retrospektive Prompts führen zu den ausgeprägtesten Lernprozessen ($d = 0.87$) und dem größten Lernerfolg im verzögerten Nachtest ($d = 0.46$). Die Mediationsanalyse für H3 zeigte, dass die Lernprozesse den Effekt des Prompttyps auf den Lernerfolg medieren. In Bezug auf unsere Forschungsfrage zeigten sich keine signifikanten Unterschiede zwischen den beiden Präsentationsformen für Lernprozesse und Lernerfolg. Die Ergebnisse zeigen, dass sich in einer Stichprobe von Lernenden ohne Vorwissen retrospektive Prompts am besten eignen, um eine tiefe Verarbeitung von Lerninhalten und einen nachhaltigen Lernerfolg hervorzurufen. Unsere Ergebnisse deuten außerdem darauf hin, dass Antizipation für Lernende ohne Vorwissen keine geeignete Instruktion ist. Zukünftige Studien sollten die Rolle von Vorwissen bei Antizipation weiter untersuchen, z.B. durch einen Vergleich von Novizen und Experten. Zusätzlich konnten wir zeigen, dass auch beim Lernen mit Tutorials die Präsentationsform eine untergeordnete Rolle spielt, solange Lerninhalte tief verarbeitet werden. Dies steht im Einklang mit früheren Forschungsarbeiten (z.B. Hefter et al., 2019).

Selbstständige Feedbackwahl in computerbasierten Lernangeboten – Welche Faktoren haben einen Einfluss auf die Feedbackwahl?

Friederike von Kienitz, Sebastian Krusekamp, Michael Besser

Leuphana Universität Lüneburg, Deutschland

Theorie

Feedback gilt als einer der wichtigsten Faktoren, die das Lernen beeinflussen (Hattie & Timperley, 2007). Insbesondere in computerbasierten Lernangeboten stellt die automatisierte Darbietung von Feedback ein Potenzial dar, Lernprozesse zu unterstützen (Mertens et al., 2022). Bezüglich der Lernwirksamkeit scheint elaboriertes Feedback, das beispielweise eine Erklärung bereitstellt (*Erklärungsfeedback*) oder Hinweise liefert, mit denen Lernende ihre Fehler selbst korrigieren können (*Hint & Try-again Feedback*), lernförderlicher zu sein als einfaches Feedback, das nur die richtige Antwort bereitstellt (*KCR-Feedback*) (Mertens et al., 2022). Die Voraussetzung für dessen Gelingen ist jedoch, dass das Feedback auch von den Lernenden verarbeitet und genutzt wird (Jonsson & Panadero, 2018; Van der Kleij & Lipnevich, 2021). Aktuelle Studien zeigen jedoch, dass dies oft nicht der Fall ist (z.B. Woods, 2022). Entscheidende Erklärungsfaktoren hierfür sind insbesondere motivationale Faktoren der Lernenden (Lui & Andrade, 2022; Fong & Schallert, 2023). Bisherige Forschung konzentriert sich insbesondere auf Feedback, das Lernenden zugewiesen wird. Gemäß der Selbstbestimmungstheorie (Ryan & Deci, 2002) könnte aber gerade das Angebot einer selbstständigen Wahl eines automatisierten Feedbacks die intrinsische Motivation und daraus resultierend die Auseinandersetzung mit dem Feedback fördern.

Empirische Evidenzen, insbesondere welches (computerbasierte) Feedback Lernende wählen und welche motivationalen Faktoren diese Wahl beeinflussen, fehlen bisher weitestgehend. Allerdings lassen sich aus der Erwartungs-Wert-Theorie (Eccles et al., 1983) der zugemessene subjektive Aufgabenwert sowie das Selbstkonzept als mögliche Einflussfaktoren auf die Feedbackwahl ableiten.

Forschungsfragen

Aus obigem Desideratum abgeleitet ergeben sich die folgenden Forschungsfragen:

- (1) Welches Feedback wählen Lernende, wenn sie die Wahl zwischen KCR-Feedback, Erklärungsfeedback und Hint & Try-again Feedback haben?
- (2) Haben der subjektive Aufgabenwert und das Selbstkonzept einen Effekt auf die Wahl des Feedbacks?

Methode

Im Oktober 2023 wurde eine computerbasierte quasi-experimentelle Studie mit 153 Studierenden (75% weiblich, $M_{\text{Alter}}=21.67$, $SD_{\text{Alter}}=3.62$) durchgeführt. Zunächst wurden das Selbstkonzept sowie der subjektive Aufgabenwert (gemessen durch die wahrgenommene Nützlichkeit, Wichtigkeit, Kosten und den intrinsischen Wert) in Bezug auf Mathematik durch etablierte Skalen (Mang et al., 2018; Steinmayr & Spinath, 2010; Trautwein et al., 2012, Cronbachs $\alpha=.74-.90$) erhoben. Anschließend bearbeiteten die Teilnehmenden Aufgaben zur Potenzrechnung (Single-Choice Format). Bei Falschantwort hatten die Teilnehmenden jeweils die Möglichkeit, eine von drei angebotenen Feedbackarten auszuwählen: (1) Der richtigen Antwort (KCR-Feedback), (2) der richtigen Antwort mit einer Erklärung (Erklärungsfeedback) oder (3) einem Hinweis, um die Aufgabe nochmal zu lösen (Hint & Try-again Feedback). Für die Datenauswertung wurden insbesondere einfache und multiple robuste Regressionen in R gerechnet.

Ergebnisse

(1) Über alle Aufgaben hinweg wählte der Großteil der Teilnehmenden (zwischen 46% und 61%) das Erklärungsfeedback aus. Etwa 50% der Teilnehmenden wechselten im Verlauf der Aufgabenbearbeitung jedoch ihre Feedbackwahl zwischen den Aufgaben.

(2) Selbstkonzept, wahrgenommene Nützlichkeit, Wichtigkeit und intrinsischer Wert haben einen signifikanten Einfluss auf die relative Häufigkeit der Wahl des Erklärungsfeedbacks. Multiple Regressionsanalysen zeigen, dass die meiste Varianz (9%) durch ein Modell aus Selbstkonzept und der wahrgenommenen Nützlichkeit aufgeklärt wird. Dies ändert sich nicht, wenn für die Anzahl der insgesamt falschen Antworten kontrolliert wird. Signifikante Effekte auf die Häufigkeit der Wahl der Feedbackarten KCR und Hint & Try-again liegen hingegen nicht vor. Korrelationsanalysen zeigen jedoch, dass eine höhere relative Häufigkeit des Hint & Try-again Feedbacks mit einem höheren Selbstkonzept (Spearman's $\rho=.18^*$) und einem höheren intrinsischem Wert (Spearman's $\rho=.17^*$) korrelieren.

Diskussion

Die Studie verdeutlicht, dass viele Lernende bei einem entsprechenden Angebot verschiedene Feedbackarten nutzen und dass das Selbstkonzept sowie Aspekte des subjektiven Aufgabenwertes einen signifikanten Einfluss auf die Wahl des Erklärungsfeedbacks haben. Demnach legen die Ergebnisse ein anderes Vorgehen beim Geben von Feedback nahe, als dies in der derzeitigen Forschungspraxis sowie im Kontext digitaler Lernangebote üblich ist. In weiterführenden Studien sollen daher

insbesondere die Auswirkungen der selbstständigen Feedbackwahl auf motivationale, kognitive und behaviorale Prozesse näher untersucht werden.

PS-45: Innovationen und Professionalisierung in der Lehrkräftebildung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 382

Chair der Sitzung: Sara Großbruchhaus

Lehrkräfte implementieren praktische DNA-Analysen im Unterricht: Gelingensfaktoren im individuellen Fokus

Sara Großbruchhaus, Patricia Schöppner, Claudia Nerdel

Technische Universität München, Deutschland

Lehrerfortbildungen (LFB) ermöglichen Lehrkräften mit stetig wachsenden Wissen und technologischen Neuerungen Schritt zu halten, indem sie bestehende Lehrpraxis mit Innovationen verknüpfen (Merchie et al. 2018). Generell werden LFB Potential zur Verbesserung der Unterrichtsqualität zugesprochen und Determinanten ihrer Wirksamkeit erforscht (Brühwiler et al. 2017). Dazu zählen u. A.: Teilnahme an der LFB mit Kolleg*innen, unterstützende Schulleitung, Passung des dargestellten Inhalts zur existierenden Praxis, Möglichkeiten Feedback zu erhalten (u.A. Gräsel und Parchmann 2004; Schrader und Hasselhorn 2020). Außerdem spielt die LFB-Dauer eine Rolle, weshalb die Wirksamkeit einmaliger Angebote angezweifelt wird. Weiterhin wird kritisiert, dass die Evaluation vieler LFB-Angebote bei der direkten Reaktion der Lehrkräfte endet, da alles darüber hinaus methodisch anspruchsvoller ist (Lipowsky und Rzejak 2020). Die vorliegende Studie nimmt sich dessen an und adressiert die Entscheidungsfindung von Lehrkräften Inhalte einer molekular-biologischen LFB praktisch in den Unterricht zu implementieren:

- Welche Argumente ziehen Lehrkräfte für oder gegen eine Umsetzung heran?
- Wie gehen Lehrkräfte bei der praktischen Implementation vor?

Die LFB fokussiert den Prozess der DNA-Analyse mittel Polymerase-Ketten-Reaktion, und setzt diese in verschiedenen curricular valide Kontexten, beispielsweise bittere Geschmackswahrnehmung (ISB 2015). Die Konzeption anhand oben genannter Gelingensfaktoren und die Evaluation unmittelbar nach Teilnahme wurden bereits beschrieben (Nerdel und Schöppner 2021; Schöppner et al. 2022). Im Anschluss an die LFB-Teilnahme können alle Lehrkräfte die benötigten Geräte und Reagenzien, wie Enzyme und Thermozykler, kostenlos ausleihen. Um einen ersten Einblick in die Forschungsfragen zu erhalten, wurde eine leitfadengestützte Interview-Studie angelegt. Zum Zeitpunkt der Studie nahmen 289 Lehrkräfte aus 98 Schulen an der LFB teil. Die Auswahl der Interviewpartner erfolgte theoriegeleitet, um eine möglichst vollständige Abdeckung bekannter Variablen und deren Kombinationen zu erreichen (Flick 2006): (1) Schulform, (2) Teilnahme (alleine, mit Kolleg*innen), (3) Ort der LFB (Schule, Universität), (4) Implementation (alleine, mit Kolleg*innen). Es wurden $N=39$ Lehrkräfte interviewt, von diesen implementierten $n=20$ und $n=19$ nicht. Die Transkripte (Dresing und Pehl 2020) wurden in fünf Schritten analysiert: (1) Fallzusammenfassungen (Kuckartz et al. 2008), (2) Typisierung der Entscheidungsfindung und Implementation (Mayring 2015), (3) (Inter)-Kodierung in MaxQDA (30% der Interviews, $\kappa=0,89$), theoriegeleitetes Kategoriensystem aus sieben Hauptkategorien: LFB, Persönlichkeitsmerkmale, Schulorganisation, Systemmerkmale, Innovation, Kooperation, Dissimulation, (5) induktive Kodierung.

Für die erste Forschungsfrage ergab sich, dass Lehrkräfte unabhängig von ihrer Implementations-Entscheidung häufig gleicher Meinung sind. So beurteilen alle die LFB als „sehr gut“, den Lehrplan als „zu voll“, damit die Zeit im Unterricht als „zu knapp“, sowie Praxiseinheiten als „wichtig“. Abweichende Meinungen beziehen sich meist auf die Umsetzung der Praxiseinheit, bei der beispielsweise die Instruktion von Schüler*innen als stressig und mit mangelnder Routine beschrieben wurde. Basierend auf den Argumenten können keine Vorhersage zur Implementation gemacht werden, da auch implementierende Lehrkräfte diese Aussagen trafen. Entsprechend müssen bei der Entscheidung individuelle Grenzwerte vorliegen, beispielsweise kann die Wichtigkeit von Praxiseinheiten den Stress bei der praktischen Umsetzung „überbieten“. Außerdem scheinen implementierende Lehrkräfte eine größere Toleranz gegenüber eigener Fehler und Übungs-Bereitschaft zu haben.

Bei der zweiten Forschungsfrage ergaben sich drei Implementationsstrategien mit steigendem Organisationsaufwand: Regelunterricht, Blockunterricht, Tagesevent. Die meisten Lehrkräfte wählten Blockunterricht oder wechselten bei Wiederholung in dieses Format, obwohl es Stundentauch oder Freistellung von Schüler*innen bedarf. Eine Schule integrierte die praktische Umsetzung von DNA-Analysen in ihr Schulcurriculum und implementiert jährlich, wobei die Strategie im Turnus verfeinert wurde (Großbruchhaus et al. 2023).

Die Ergebnisse verdeutlichen, dass diese einmalige LFB geeignet ist, Implementation zu initiieren. Obwohl aus der Literatur bekannte Gelingensfaktoren wiedergefunden wurden (u. A. Schrader und Hasselhorn 2020; Lipowsky und Rzejak 2021), scheint bei der Entscheidungsfindung vielmehr die individuelle Gewichtung bedeutsam zu sein. In der Unterrichtspraxis gilt, dass Methodenvielfalt Schüler*innen-Lernen unterstützt. Dieses Konzept könnte, wie das Angebot-Nutzen-Modell auf LFB übertragen werden. Folgestudien könnten sich entsprechend spezifisch mit der Frage auseinandersetzen, welche LFB-Angebote für welchen Lehrkräfte hilfreich sind

Validierung einer rollenspielbasierten Simulationsprüfung für angehende Englischlehrkräfte – Validitätsargumente zum Scoring und zur Testwertinterpretation

Thomas Janzen, Lea Grotegut, Philipp Wotschel, Christoph Vogelsang

PLAZ - Professional School of Education, Universität Paderborn, Deutschland

Da es für angehende (Englisch-)Lehrkräfte eine Herausforderung ist, ihr im Studium erworbenes Wissen in typischen Unterrichtssituationen einzusetzen (Zeichner, 2010), sollten handlungsnahen Fähigkeiten bereits in der ersten Phase der Lehrkräftebildung stärker einbezogen und geprüft werden (Ständige Wissenschaftliche Kommission, 2023). Eine Möglichkeit, solche Prüfungen zu gestalten, stellen rollenspielbasierte Simulationsprüfungen (RobS) dar. Diese sind im Medizinstudium seit Längerem etabliert, aber in der Lehrkräftebildung noch wenig verbreitet (Miller, 1990; Meier et al., 2023). Vor diesem Hintergrund wurde ein Prototyp für eine RobS als summatives Prüfungsformat für die Englischlehrkräftebildung konzipiert. Orientiert am Konzept der „core practices“ (Grossman, 2018) und dem Kontinuumsansatz von Blömeke et al. (2015) misst die RobS die performative Facette der Englischlehrkräftekompetenz. Den inhaltlichen Fokus bildet das mündliche Feedback zu Lernendentexten im Rahmen des Englischunterrichts der 9. Klasse.

Zur Validierung wurde der „argument-based approach to validation“ von Kane (2013) verfolgt. In diesem Beitrag werden aktuelle Ergebnisse der Validierung in Bezug auf Annahmen zum Scoring (Reliabilität des Ratings) und der Testwertinterpretation (externe Validierung) vorgestellt. Grundlage für die Untersuchungen bildeten videographierte RobS-Durchführungen sowie unterschiedliche Wissens- und Kompetenztests der Zieldomäne „Englischlehrkräftekompetenz“ von $N = 51$ Lehramtsstudierenden.

Scoring (Reliabilität des Ratings)

Die RobS-Aufnahmen wurden mit Hilfe eines literaturbasierten und durch Expert*inneninterviews ($N = 5$) validierten Rasters (27 Items) von Rater 1 bewertet (Janzen et al., 2024). Die interne Konsistenz betrug nach Ausschluss von $n = 3$ Items $\alpha = .81$, sodass 24 Items für das finale Scoringinstrument übernommen wurden.

Zur Messung der Interrater-Reliabilität wurden $n = 10$ Videos (ca. 20%) von zwei weiteren unabhängigen und zuvor geschulten Rater*innen bewertet. Nach dem ersten vollständigen Rating der Videos (Scoring 1) wurden evtl. Diskrepanzen zwischen den Rater*innen geklärt. Anschließend wurden die Videos von Rater*in 2 und 3 erneut bewertet (Scoring 2). Die Reliabilität der Summenscores wurde mittels Intraklassenkorrelation (unjust, two way, mixed) ermittelt und beträgt $ICC_{S1} = .91$ (95%-CI [.77, .98]) bzw. $ICC_{S2} = .96$ (95%-CI [.90, .99]). Normalverteilung (Shapiro-Wilk-Test, $p > .05$) und Varianzhomogenität (Levene-Test, $p > .05$) wurden für beide Scorings aller Rater*innen bestätigt. Die Summenscores der Rater*innen korrelieren für beide Scorings hoch (Cicchetti, 1994). Dies ist für den Einsatz als summatives Prüfungsverfahren von hoher Relevanz, da auf Basis des Testscores Entscheidungen über z.B. Noten oder Credit Points getroffen werden. Die Reliabilitätswerte für die Gesamtheit der Itemscores ergaben Fleiss' $\kappa_{S1} = .60^{**}$ bzw. $\kappa_{S2} = .82^{**}$ und sind damit für Scoring 1 moderat und für Scoring 2 nahezu perfekt (Landis & Koch, 1977). Dies deutet auf eine akzeptable Reliabilität der einzelnen Items des Scoringinstruments hin.

Testwertinterpretation (externe Validierung)

Für die externe Validierung im Sinne der Generalisierung bearbeiteten die Teilnehmenden Tests zum fachdidaktischen Wissen im Fach Englisch (Kirchhoff, 2017), Feedbackwissen (Schütze et al., 2017) und der Sprachkompetenz (Lemhöfer & Broersma, 2012), um die Korrelationen zu eher kognitiv-dispositionalen Konstrukten der Zieldomäne zu messen. Es zeigt sich, dass englischdidaktisches Wissen ($p = .37^{**}$) und Feedbackwissen ($p = .42^{**}$) moderat (Cohen, 1988) mit den Leistungen in den RobS korrelieren, Sprachkompetenz ($p = .23$) allerdings nicht. Die Werte deuten darauf hin, dass die RobS ebenfalls ein Konstrukt der Zieldomäne „Engischlehrkräftekompetenz“ erfasst, wobei Korrelationen nur für didaktische bzw. pädagogische Facetten bestehen. Sprachkompetenzen allein scheinen in keinem Zusammenhang zur Performanz in den RobS zu stehen.

Es lässt sich zusammenfassen, dass Reliabilitäts- und Übereinstimmungswerte der Summenscores und Itemscores die Validität des Scorings stärken, wobei limitierend anzumerken ist, dass nur ein Teil des Gesamtsamples für das Interrating herangezogen werden konnte. Ebenso deuten die Ergebnisse der externen Validierung darauf hin, dass die Testwertinterpretation in der Zieldomäne anzusiedeln ist.

Interventionsstudie zur Professionalisierung der Assessment Literacy von Lehrpersonen: Erfassung mittels Text- und Videovignette

Selin S. Oeksuez¹, Sandra Zulliger¹, Alois Buholzer¹, Karolin Maskos², Ladina Feucht², Andreas Schulz²

¹Pädagogische Hochschule Luzern, Schweiz; ²Pädagogische Hochschule Zürich, Schweiz

Theoretischer Hintergrund und Fragestellung

In der aktuellen Diskussion über formative Assessments werden mit Peer- und Self-Assessment (PASA) zwei grundlegende Ansätze hervorgehoben, die ein erhebliches Potenzial zur Unterstützung sowohl fachlichen als auch überfachlichen Lernens bieten (Buholzer et al., 2020; Decristan et al., 2015; Hondrich et al., 2018; Panadero et al., 2017; Pinger et al., 2018). Ein entscheidender Faktor für die erfolgreiche Implementierung von formativen (Peer- und Self-) Assessments ist die Assessment Literacy (AL) von Lehrpersonen. Diese umfasst das notwendige Wissen und die Fähigkeiten, die für den gesamten formativen Beurteilungsprozess erforderlich sind – von der Erhebung von Informationen über den Lernstand der Schüler:innen bis hin zur mikro- und makroadaptiven Gestaltung des Unterrichts, die auf diesen Informationen basiert (Schildkamp et al., 2020).

Trotz der zentralen Rolle, die formative Beurteilung im schulischen Lernen spielt, zeigt sich, dass Lehrpersonen eher über ein geringes Mass an formativer AL verfügen (Atjonen et al., 2022). Insbesondere wird vermutet, dass das unzureichend entwickelte Professionswissen zu formativem Assessment, welches gemäss Xu & Brown (2016) als Grundlage von AL gilt, ein wesentlicher Grund dafür ist, dass Lehrpersonen formative Assessments oft nicht lernwirksam im Unterricht umsetzen (Buholzer et al., 2020; Oswald, 2013). Diese Erkenntnisse verdeutlichen die Notwendigkeit, die formative AL von Lehrpersonen zu fördern, um die Wirksamkeit von PASA im Unterricht zu steigern und somit das Lernen der Schüler:innen nachhaltig zu unterstützen.

Die vorliegende Studie ist Teil der vom Schweizerischen Nationalfonds finanzierten Interventionsstudie FORMA (FORMATIVES Assessment im MATHematikunterricht). Sie untersucht, inwiefern bei Lehrpersonen die AL zu formativem Peer- und Self-Assessment mit einer Weiterbildung weiterentwickelt werden kann und welche Faktoren (wie z.B. epistemologische Überzeugungen) diesen Prozess beeinflussen.

Methodische Zugänge

Die Interventionsstudie FORMA zielt darauf ab, die Wirksamkeit einer Weiterbildung zu formativem Peer- und Self-Assessment anhand eines randomisierten Prä-Post-Follow-Up-Wartekontrollgruppendesign zu erforschen. Im Zusammenhang mit einer Unterrichtseinheit zur halbschriftlichen Multiplikation und Division wurden Lehrpersonen in den beiden Interventionsgruppen (Peer-Assessment und Self-Assessment) theoretisch fundiert und praxisorientiert zur lernwirksamen Gestaltung von formativem PASA weitergebildet. Insgesamt 90 Lehrpersonen (je 30 pro Interventionsgruppe PA oder SA, Wartekontrollgruppe) haben an der Studie teilgenommen. Nebst anderen Erhebungsverfahren wurde für die Erfassung der AL vor und nach der Weiterbildung ein online Survey zu ihrem Professionswissen hinsichtlich des formativen PASA durchgeführt.

Um das situationsspezifische Professionswissen der Lehrpersonen gemäss Blömeke et al. (2015) zu ermitteln, wurden drei offene Fragen mit Video- und Textvignetten zur Umsetzung von PASA vorgelegt. Durch ihre Antworten zeigen die teilnehmenden Lehrpersonen, wie sie die dargestellte Situation deuten. Zur Analyse der Antworten wurde unter Verwendung des TALiP-Modells von Xu und Brown (2016) zur AL deduktiv Oberkategorien gebildet, welche induktiv zu Unterkategorien differenziert wurden.

Nebst der inhaltlichen Zuordnung werden die Antworten der Lehrpersonen zusätzlich geratet (Santagata & Sandholtz, 2019). Auf einer dreistufigen Ratingskala wird die Qualität des Interpretationsniveaus (von rein deskriptiv bis analytisch) der einzelnen Antworten beurteilt. Diese Herangehensweise mit inhaltlichen und Rating-Kategorien erwies sich in anderen Studien als reliabel und valide (Kersting et al., 2010).

Die Ergebnisse zeigen, inwieweit die Lehrpersonen der beiden Interventionsgruppen ihr Professionswissen zu PASA quantitativ und qualitativ stärker erweitern können als die Lehrpersonen der Wartekontrollgruppe.

Ergebnisse und ihre Bedeutung

Derzeit werden die offenen Antworten noch kodiert und geratet. Bis zur GEBF-Tagung werden die vollständigen Analysen zur Beantwortung der Forschungsfrage vorliegen. Erste Analysen, die bei der Entwicklung des Kategorien- und Ratingsystems durchgeführt wurden, deuten darauf hin, dass Lehrpersonen der beiden Interventionsgruppen (PA und SA) im Vergleich zur

Wartekontrollgruppe beim Posttest vermutlich mehr tiefenstrukturelle Elemente bei der Gestaltung von PASA ansprechen. Ein Beispiel hierfür ist das vertiefte Nachdenken und Begründen der Peer- und Selbstbeurteilung. Es darf daher angenommen werden, dass mit den Ergebnissen aufgezeigt werden kann, wie eine Weiterbildung die AL von Lehrpersonen positiv beeinflussen kann.

Wie sollte ein Core Practice-Training aufgebaut sein, das den Erwerb von Handlungskompetenz optimal fördert?

Anna Holstein¹, Hadmut Hipp², Marc Kleinknecht¹, Matthias Nückles²

¹Leuphana Universität Lüneburg, Deutschland; ²Universität Freiburg, Deutschland

Theoretischer Hintergrund und Fragestellung

Die Kluft zwischen Theorie und Praxis ist die größte Herausforderung in der Ausbildung von Lehrkräften (Voss et al., 2019). Lehramtsstudierende erwerben in der akademischen Phase viel wissenschaftliches Wissen, welches später kaum noch für ihr unterrichtliches Handeln eine Rolle spielt. Der Core-Practice-Ansatz ist eine vielversprechende Möglichkeit, um wissenschaftliches Wissen mit praktischem Handeln zu verknüpfen (Fraefel & Grossman, 2023; Nückles & Kleinknecht, 2023). Core Practices sind häufig vorkommende, forschungsbasierte Lehrtätigkeiten, die auch von Lehramtsstudierenden erlernt werden können. Beispiele für Core Practices sind z.B. das *Erklären*, das *kognitive Modellieren* sowie das *Scaffolding* von Lesestrategien, wie sie etwa in dem Lesestrategietraining Reciprocal Teaching auf wissenschaftlich fundierte Weise dargelegt werden (Palinscar & Brown, 1984).

Ein aus der Praxis der Lehrkräftebildung stammender Vermittlungsansatz für Core Practices ist der Learning Cycle von McDonald et al. (2013). Er beinhaltet Phasen wie die Modellierung einer Core Practice oder das Üben im geschützten Rahmen. Dem Learning Cycle fehlt jedoch eine kognitionspsychologische Fundierung, die wir in unserem DFG-Projekt ACTion leisten wollen. Anhand einer Gegenüberstellung des Learning Cycle und der kognitionspsychologischen Fertigkeitserwerbstheorie ACT-R (Anderson, 1982) schlussfolgern wir, dass eine prinzipienbasierte Einführung der Core Practice, eine Modellierung derselben, und ein selbständiges Ausprobieren mit Schüler*innen wichtige Phasen eines Core Practice Training sind. Wir untersuchen vor diesem Hintergrund, welche Kombination von Trainingsphasen am besten den Erwerb von Core Practices fördert.

In unserer Studie variierten wir systematisch, welche Kombinationen von Trainingsphasen ein Training umfasste. Wir erwarteten, dass die Kombination aus prinzipienbasierter Einführung, anschaulicher Modellierung und eigenständigem Ausprobieren mit Schüler*innen die Handlungskompetenz am besten fördern würde (*Performanzhypothese*). Wir vermuteten weiterhin, dass die Kombination aus prinzipienbasierter Einführung und anschaulicher Modellierung ein tiefes Verständnis der Core Practices fördert und daher eine zentrale Voraussetzung für kompetentes Handeln ist. Insofern sollte die Fähigkeit, Videosequenzen zu Core Practices zu analysieren (Professionelle Wahrnehmung), die Fähigkeit, Core Practices kompetent auszuführen medieren (*Mediationshypothese*, Blömeke et al., 2015).

Methode

$N = 128$ Lehramtsstudierende der Universitäten Freiburg und Lüneburg nahmen an der Trainingsstudie teil. Die Core Practices *Erklären*, *kognitives Modellieren* und *Scaffolding* wurden anhand von Reciprocal Teaching aufgezeigt. Gruppe 1 studierte eine prinzipienbasierte Einführung zu Reciprocal Teaching sowie eine Video-Modellierung und setzte Reciprocal Teaching mit Schüler*innen um. In den anderen Versuchsgruppen fehlte jeweils eine Trainingsphase. Um die Time-on-Task gleich zu halten, wurde die Erklär- oder Modellierungsphase zweimal durchlaufen. In der Bedingung ohne Ausprobieren wurde eine schriftliche Reflexion erstellt. Wir verwirklichten also folgende Trainingsbedingungen: E-M-A: Erklärung, Modellierung, Ausprobieren; E-E-A: Erklärung, Erklärung, Ausprobieren; M-M-A: Modellierung, Modellierung, Ausprobieren und E-M-R: Erklärung, Modellierung, Reflexion. Die Handlungskompetenz erhoben wir in einem halbstandardisierten Rollenspiel, indem wir die Teilnehmenden beim Erklären, Modellieren und Unterstützen von Schüler:innen beim Ausführen der Lesestrategien filmten und die Aufnahmen kodierten (ICCs: 0.7–0.9). Die Teilnehmenden bearbeiteten außerdem einen Professionellen-Wahrnehmungs-Test ($\kappa = .84$).

Ergebnisse und Diskussion

Entsprechend der *Performanzhypothese* erzielte die E-M-A-Gruppe die besten Ergebnisse im Rollenspieltest, $F(1, 119) = 7.53$, $p = .004$, $\eta^2 = .06$. Diese und die E-M-R-Gruppe schnitten im Professionellen-Wahrnehmungs-Test signifikant besser ab als die anderen Gruppen, $F(1, 119) = 8.28$, $p = .003$, $\eta^2 = .065$. Bestätigt wurde auch die *Mediationshypothese*, da der positive Effekt der E-M-A und E-M-R Bedingungen auf Handlungskompetenz mediert wurde durch die professionelle Wahrnehmung, indirekter Effekt $ab = .0227$, 95% CI [.0001, .0515].

Core Practice Trainings, die das Studium von prinzipienbasierter Erklärung und Modellierung mit einer praxisnahen Ausprobierphase kombinierten, förderten also die Handlungskompetenz angehender Lehrkräfte am besten. Das Studium von prinzipienbasierter Einführung und anschaulicher Modellierung förderte besonders das analytische Verständnis der Core Practices, und medierte daher den positiven Effekt des Trainings auf die Handlungskompetenz. Die spezifische Kombination aus abstrakt-wissenschaftlicher Einführung und konkret-anschaulicher Modellierung ist also entscheidend für die Gestaltung von Core Practice Trainings, die effektiv Handlungskompetenz fördern.

PS-46: Bildung und Berufseinstieg unter den Bedingungen der COVID-19-Pandemie

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: SO 322
Chair der Sitzung: Wilhelm Koschel

Graduating in times of crisis? The role of type of education for school-to-work transitions of VET graduates during Covid-19

Miriam Hänni¹, David Glauser², Irene Kriesi¹

¹Eidgenössische Hochschule für Berufsbildung, Schweiz; ²Universität Bern, Schweiz

The Covid-19-pandemic and the measures to contain it led to economic downturns in most economies. Economic recessions hit young people at the start of their careers particularly hard and make school-to-work transitions more vulnerable (Brzinsky-Fay, 2007; Cockx & Ghirelli, 2016). Young people graduating during economic downturns are more likely to become unemployed (Rothstein, 2023) or to continue in education (Hampf, Piopiunik & Wiederhold, 2020), and earn less when entering the labour market (Kahn, 2010; van den Berge, 2018).

However, there is some evidence that for labour market entrants the economic repercussions of the Covid-19-pandemic differed from those of other economic crises. The crises hit economic sectors unevenly. Although unemployment rose in all sectors, graduates from educational programmes in the service sector were affected most, due to mandatory business closures over prolonged periods of time (SECO, 2021a). In addition, the effect of the crisis on new entrants to the labour market may have been lower than in previous recessions due to the strong protective measures that were taken (Bussink, Vervliet & Weel, 2022). We are thus interested *to what extent, if at all, the Covid-19-pandemic has affected the labour market status (in employment, in training, NEET) of recent graduates and whether the economic repercussions of the pandemic differed between graduates from different educational programmes.*

To theorize school-to-work transitions the labour queue model is useful (Thurow, 1979). It assumes that employers rank job applicants in a queue and that those with the most sought-after characteristics receive job offers first. Young graduates thus face strong competition from more experienced workers when searching for a job (Salvisberg & Sacchi, 2014), particularly during economic recessions and low labour demand (Brzinsky-Fay, 2007; Cockx & Ghirelli, 2016). Difficult school-to-work transitions may place graduates in jobs that do not match their qualifications or in jobs with fewer training and promotion opportunities (Altonji, Kahn & Speer, 2016), and may cause long-lasting scarring effects on an individual's career (Cockx & Ghirelli, 2016). Such scarring effects are particularly strong among graduates from vocational education and training (VET) (Shi & Di Stasio, 2022). We thus focus on school-to-work transitions of dual VET graduates and hypothesize *that the economic downturn due to the pandemic affected apprentices' school-to-work transition by prolonging the time-span needed to find a first job and by increasing the risk of unemployment. Taking into account the uneven impact of the pandemic across the economy, we expect differences between different VET programmes.*

We use register data on education trajectories of Swiss VET graduates and compare the trajectories of full cohorts of apprentices graduating before (2016 – 2019) during (2020 – 2022) and after (2023) the pandemic. Switzerland is an ideal case to study the economic repercussions of the pandemic among VET graduates because VET is the dominant upper-secondary track in Switzerland and because the Swiss VET-system is closely linked to the economy and, therefore, vulnerable to economic fluctuations. In addition, the 250 different VET programmes allow us to study differences across occupations and occupational subsegments.

We use multivariate regression and event-history models to analyse our research question. The main dependent variables are (i) time to first employment as well as (ii) the probability of different labour market outcomes (in employment, in education, NEET).

Results indicate that, overall, it took VET graduates only slightly longer to transition into their first job and that unemployment increased only marginally among VET graduates during the pandemic. However, we observe a clear increase in the share of individuals who remained in education after graduating from a VET programme. Distinguishing between graduates from different VET programmes points to heterogeneous effects of the economic crisis.

Bildungsgerechtigkeit in Zeiten der Pandemie – Einflüsse auf die Vergabe von Abschlüssen und Abschlussnoten

Tobias Brändle, Britta Pohlmann, Markus Lücken

Institut für Bildungsmonitoring und Qualitätsentwicklung (IfBQ), Deutschland

Theoretischer Hintergrund

Die Steigerung der Bildungsgerechtigkeit war und ist zentrales Ziel bildungspolitischer Bemühungen. Jedoch ist die Kopplung zwischen Leistung und sozialer Herkunft weiterhin sehr hoch (Autor:innengruppe Bildungsberichterstattung, 2024; Lewalter et al., 2024; Stanat et al., 2023). Im Zuge der Corona-Pandemie scheinen die Bildungsungleichheiten sogar zugenommen zu haben (Schult et al., 2022; Schulze, Heuling & Helm, 2022).

Unter dem Stichwort der Bildungsgerechtigkeit (Stojanov, 2011) analysieren wir geschlechterbezogene, soziale und zugewanderungsbezogene Disparitäten. Dabei fokussieren wir nicht nur die Frage der Chancengerechtigkeit - also ob Schüler*innen unabhängig von Herkunft und Geschlecht vergleichbare Bildungserfolge erzielen -, sondern auch die weniger untersuchte Frage der Leistungsgerechtigkeit - als ob Schüler*innen mit gleichen Kompetenzen unabhängig von Herkunft und Geschlecht gleiche Noten und Abschlüsse erwerben.

Beide Gerechtigkeitsdimensionen werden mit bildungspolitischen Maßnahmen adressiert. Zur Erhöhung der Chancengerechtigkeit wurden diverse Maßnahmen etabliert, die insbesondere Schüler*innen mit ungünstigen Ausgangslagen zugutekommen sollen. Zur Steigerung der Leistungsgerechtigkeit wurden zentrale Prüfungen eingeführt sowie Bildungsstandards und Vergleichsarbeiten implementiert.

Vor diesem Hintergrund untersuchen wir, wie es um die Bildungsgerechtigkeit bestellt ist und welche Entwicklungen sich in den letzten Jahren hinsichtlich der Chancen- und Leistungsgerechtigkeit abgezeichnet haben. Dabei soll ein besonderer Blick auf die Entwicklung während der Corona-Pandemie gerichtet werden.

Fragestellungen

1) Wie hoch ist die Chancengerechtigkeit und Leistungsgerechtigkeit bei der Vergabe von Schulabschlüssen und Abschlussnoten?

2) Welche Trends lassen sich über die letzten Jahre beobachten? Hat die Corona-Pandemie zu vergrößerten Ungleichheiten geführt?

Methode

Zur Untersuchung der Fragestellungen werden 8 Schülerkohorten aus Hamburg untersucht, die in den Schuljahren 2015/16 bis 2022/23 Prüfungen zum mittleren Schulabschluss (MSA) an nicht-gymnasialen Schulformen absolviert haben. Für diesen Schüler*innen liegen amtliche Angaben zum soziodemographischen Hintergrund sowie Ergebnisse der Abschlussprüfungen in den geprüften Fächern mit den entsprechenden Unterrichtsnoten vor. Außerdem werden Kompetenzdaten aus den regelhaften Kompetenzerhebungen KERMIT (Lücken et al., 2014) genutzt.

Zur Analyse der Chancengerechtigkeit werden Mittelwerts- und Anteilsvergleiche zwischen verschiedenen Schülergruppen im Hinblick auf die Abschlüsse und Abschlussnoten berechnet und die Ergebnisse über die Jahre verglichen. Zur Analyse der Leistungsgerechtigkeit werden schrittweise Regressionsmodelle gerechnet. Dabei werden zunächst die Haupteffekte der soziodemografischen Merkmale auf die Abschlussergebnisse modelliert und anschließend überprüft, wie sich diese unter der Kontrolle der Kompetenzen verhalten. Im dritten Schritt wird die Kohortenzugehörigkeit in das Modell aufgenommen, um Veränderungen infolge der Pandemie zu untersuchen.

Ergebnisse und ihre Bedeutung

Sowohl hinsichtlich der Abschlüsse als auch hinsichtlich der Abschlussnoten sind deutliche Unterschiede in Abhängigkeit der betrachteten soziodemographischen Merkmale der Schüler*innen feststellbar. In Übereinstimmung mit den regelhaften Befunden internationaler und nationaler Schulleistungsstudien schneiden Schüler*innen mit geringerem sozioökonomischem Status und Schüler*innen mit Zuwanderungshintergrund durchweg schlechter ab. Geschlechterbezogene Disparitäten lassen sich ebenfalls beobachten und entsprechen dem bekannten Muster mit Vorteilen der Mädchen im sprachlichen Bereich und Vorteilen der Jungen im mathematischen Bereich.

Wie die Regressionsanalysen zeigen, werden die Abschlüsse und Abschlussnoten jedoch weitgehend leistungsgerecht vergeben. So lassen sich die Unterschiede in den Abschlüssen und Abschlussnoten zwischen den verschiedenen Schüler*innen-Gruppen im Wesentlichen auf die Kompetenzunterschiede zurückführen. Allerdings vergrößern sich die Effekte des Geschlechts unter Kontrolle der Kompetenzen. Mädchen erreichen im Vergleich zu Jungen somit bessere Abschlussergebnisse, als es aufgrund der Kompetenztests zu erwarten gewesen wäre. Dies kann durch bessere Prüfungsvorbereitungen und günstigere Prüfungsstrategien oder bessere Bewertungen von Mädchen bedingt sein.

Das Ausmaß der Chancen- und Leistungsgerechtigkeit scheint dabei über die letzten Jahre relativ stabil zu sein. Die Pandemie hat bei den Abschlüssen und Abschlussergebnissen nicht zu größeren Ungleichheiten geführt. Ob sich dies erst bei nachfolgenden Schülerkohorten zeigt, bleibt abzuwarten und weiter zu analysieren.

Auch wenn die Abschlussvergabe weitgehend leistungsgerecht erfolgt, darf dies nicht darüber hinwegtäuschen, dass weiterhin erhebliche Ungleichheiten bestehen und die Steigerung der Chancengerechtigkeit zentrales Ziel von Bildungspolitik bleiben (wie aktuell beim Startchancenprogramm) und Thema der Bildungsforschung sein muss.

Haben unsere Medizinabsolvent*innen ihre Resilienz und ihre Kompetenzen verloren? Einfluss der Covid-19 Pandemie auf die Resilienz und die ärztlichen Kompetenzen ärztlicher Berufsanfänger*innen in Bayern

Nicola Kolb¹, Stephanie Keil², Johanna Huber¹

¹Institut für Didaktik und Ausbildungsforschung in der Medizin, LMU Klinikum; ²Universität Regensburg

Resilienz beschreibt die eigene Widerstandsfähigkeit gegenüber psychischen Belastungen und den angemessenen Umgang mit Stresssituationen (Quintiliani et al., 2022; Southwick et al., 2014). Für Studierende der Humanmedizin ist diese Kompetenz von besonderer Bedeutung, da sie ihre Ausbildung häufig als stressreich und herausfordernd wahrnehmen und z.B. ein erhöhtes Burnout-Risiko aufweisen (Dyrbye et al., 2014). Resilienz kann daher einen Beitrag zur Gesunderhaltung von Medizinstudierenden leisten. Daneben spielt auch der Erwerb fachlicher, wissenschaftlicher und kommunikativer Kompetenzen eine wichtige Rolle auf dem Weg zu einer guten Ärztin bzw. einem guten Arzt.

Eine Herausforderung, die die Resilienz und die zu erlernenden Kompetenzen beeinflusst haben, waren die Einschränkungen im Rahmen der COVID-19 Pandemie. Im Zuge dieser Pandemie hat sich die Lehre im Medizinstudium und die Ausbildung im praktischen Jahr (PJ) stark verändert. Es erfolgte eine vollständige Umstellung bisheriger Präsenzlehrformate auf reine Online-Lehre. Auch Studierende im PJ waren mit einer stark veränderten Ausbildung konfrontiert, z.B. durch umfassenden Hygieneregeln auf Station, die starke Limitierung von PJ-begleitenden Seminaren, sowie unzureichende praktische Übungsmöglichkeiten und eine höhere Arbeitsbelastung (Chen et al., 2021). Diese umfassenden Änderungen und Herausforderungen in der theoretischen und praktischen Ausbildung von Medizinstudierenden können auf psychologische Faktoren, wie Resilienz, und auf die zu erlernenden Kompetenzen Einfluss nehmen (Quintiliani et al., 2022).

Ziel der Studie ist es deshalb zu untersuchen, ob sich Medizinabsolvent*innen, die vor der COVID-19 Pandemie ihr PJ absolvierten, von Medizinabsolvent*innen, die ihr PJ während der Pandemie absolviert haben, in der Resilienz und den ärztlichen Kompetenzen (Fach-, Kommunikations- und Wissenschaftskompetenz) unterscheiden.

Die Daten, der seit 2014 im dreijährigen Turnus durchgeführten Bayerischen Absolventenstudie Medizin (MediBAS), bilden die Grundlage zur Überprüfung der Hypothesen. In die Analysen wurden die Daten der zweiten (Messzeitpunkt 2018/19) und dritten (Messzeitpunkt 2022/23) Erhebungswelle von Befragten der Humanmedizin eingeschlossen. Die befragten Absolvent*innen der dritten Erhebungswelle haben unter COVID-19 Bedingungen ihr PJ durchlaufen.

Die subjektive Resilienz wurde mit der deutschen Version der 10 Item, 5-stufig Likert-skalierten Connor-Davidson Resilience Scale (CD-RISC) erhoben. Die Fachkompetenz wurde mit 13 5-stufig Likert-skalierten Items (1 „gar nicht“ bis 5 „in sehr hohem Maße“) des Freiburger Kompetenzfragebogen gemessen (Giesler et al., 2011). Die Kommunikationskompetenz wurde mithilfe einer neu entwickelten Skala erhoben, die 9 5-stufig Likert-skalierten Items (1 „gar nicht“ bis 5 „in sehr hohem Maße“) enthält. Die Wissenschaftskompetenz wurde mit 10 5-stufig Likert-skalierten Items (1 „gar nicht“, 5 „in sehr hohem Maße“) der Munich Research Competence Scale ermittelt.

Von Messzeitpunkt zwei zu drei vergrößerte sich die Resilienz signifikant von einem Mittelwert von 3.4 vor COVID-19 zu Mittelwert von 3.7 während COVID-19 ($U = 100579.5$, $p < .001$, $r = .154$). In Bezug auf die Fachkompetenz zeigte sich kein signifikanter

Unterschied ($t(1080) = 1.960$, $p = .050$, $d = .601$), wobei die Mittelwerte sich von Welle zwei ($M = 3.37$) auf Welle drei ($M = 3.30$) verschlechterten. Auch bei der Kommunikationskompetenz ($U = 133142$, $p = .117$, $r = .048$) und der Wissenschaftskompetenz ($U = 126248$, $p = .103$, $r = .051$) gab es keine signifikanten Unterschiede.

Die COVID-19 Pandemie hatte keinen negativen Einfluss auf die Resilienz, sondern führte zu einer Steigerung der Resilienz bei den Medizinabsolvent*innen. Demnach verbessert sich die Widerstandsfähigkeit von Medizinabsolvent*innen und deren Fähigkeit, herausfordernde Situation trotz widriger Bedingungen zu bewältigen, obwohl Studien bereits vor der Pandemie auf das Burnout-Risiko von Medizinstudierenden hinwiesen. Zudem zeigte sich, dass die Fachkompetenz sich zwischen den zwei Zeitpunkten zwar nicht statistisch unterschied, aber sich ein mittlerer Effekt zeigte. Das lässt darauf schließen, dass die Fachkompetenz durch die Pandemie und die veränderte Lernsituation zu einer schlechteren Einschätzung führte. In Bezug auf die Kommunikationskompetenz und Wissenschaftskompetenz zeigten sich keine Veränderung, sodass die Pandemie keinen Einfluss hatte und somit die Kompetenz weiterhin trotz veränderter Lernsituation gut vermittelt werden konnten.

Empirische Befunde zu pandemiebedingten Limitationen im betrieblichen Bildungskontext des Pflegeberufs

Wilhelm Koschel, Ulrike Weyland, Katharina Kettler

Universität Münster, Deutschland

Der Pflegeberuf ist bereits seit Jahren durch zahlreiche gesellschaftliche Herausforderungen gekennzeichnet. Hierzu gehören z. B. die Ökonomisierung der Pflege oder der Fachkräftemangel. Klie, Remmers & Manzeschke (2021) verweisen zugleich darauf, dass mit der Einführung des neuen Pflegeberufegesetzes die Herausforderungen für das Fachpersonal aufgrund der Sicherstellung der damit verbundenen Ansprüche an die Qualität pflegeberuflichen Handelns bei gleichzeitigem Fachkräftengpass zugenommen haben. Mit der Novellierung des Pflegeberufegesetzes stellen sich auch weitere Ansprüche an die betriebliche Bildungsarbeit, was u. a. mit einer Anhebung des Qualifikationsumfangs von 200 auf 300 Std. Fortbildung für Pflegekräfte, die als betriebliches Bildungspersonal bzw. i.e.S. als Praxisanleitende bezeichnet werden, einhergeht (Weyland, 2023).

Mit Beginn der Corona-Pandemie wurde das betriebliche Bildungspersonal in Krankenhäusern und Pflegeeinrichtungen zusätzlich herausgefordert (Kaufhold & Wesselborg, 2022; Hinterbuchner et al. 2022). Dies zeigte sich in unterschiedlicher Weise: Aufgrund der Doppelrolle, bestehend aus Pflegefachkraft und Praxisanleitende/r blieb pandemiebedingt weniger Zeit für die betriebliche Bildungsarbeit. Neben den quantitativen Limitationen mussten aber ebenso auf qualitativer Ebene Einschränkungen in Kauf genommen werden. Diese ergaben sich z. B. durch die pandemiebedingten Schulschließungen. Die Auszubildenden konnten zwar in dieser Zeit nicht mehr die Schule besuchen, mussten jedoch weiter in den Krankenhäusern und Pflegeeinrichtungen arbeiten. Die Theorie-Praxis-Vernetzung war aufgrund des begrenzten schulischen Unterrichts also nur sehr erschwert möglich (Hinterbuchner et al. 2022; Wöhlke & Ruwe 2022, S. 57–61; Dziurka et al. 2022, S. 10). Ebenso zeigten sich aufgrund zahlreicher Personalausfälle deutliche Engpässe in der Praxisanleitung und somit auch in der Qualität der praktischen Ausbildung (Kaufhold & Wesselborg, 2022).

Diese Problematik aufgreifend richtet sich das BMBF-geförderte Projekt LimCare (Laufzeit bis 08/25) auf die empirische Untersuchung von coronabedingten Limitationen in der betrieblichen Bildungsarbeit, einschließlich ihrer Auswirkungen in der post-pandemischen Phase. Folgende übergeordnete Fragestellungen werden bearbeitet: (1) Welche Auswirkungen sind pandemiebedingt in der praktischen Ausbildung von Pflegefachkräften entstanden? (2) Welche Schulungsbedarfe zur Förderung heterogenitätssensibler betrieblicher Bildungsprozesse liegen bei Praxisanleitenden in der Pflege vor? Die Studienergebnisse werden vor dem Hintergrund berufspädagogischer und pflegewissenschaftlicher Erkenntnisse sowie aus pädagogisch-psychologischer Sicht interpretiert und in diesen Feldern verortet. Das Forschungsdesign gliedert sich in ein hypothesengenerierendes und ein hypothesenprüfendes Verfahren. Im Rahmen des hypothesengenerierenden Abschnitts wurde auf der Grundlage eines PRISMA geleiteten Reviews zunächst die aktuelle Befundlage eingeschätzt. Es folgten Fokusgruppeninterviews mit u. a. Praxisanleitenden sowie eine standardisierte Erhebung in den Einrichtungen der externen Kooperationspartner. Die Hypothesengenerierung sowie die Modulation latenter Konstrukte und ihrer Zusammenhänge sind bereits abgeschlossen und wurden extern durch Wissenschaft und Praxis validiert. Es wurde deutlich, dass die Limitationen während der Pandemie sich hypothetisch in drei Bereiche zusammenfassen lassen: (1) Die fehlende Fähigkeit der Auszubildenden, Pflegesituationen zusammenfassend einschätzen und interpretieren zu können, (2) Eine Zunahme der Rollenkonflikte aufseiten der Pflegekräfte, die gleichzeitig Praxisanleiter:innen sind sowie (3) die fehlende Fähigkeit, die Auszubildenden zum eigenständigen Lernen in der Praxis anzuleiten. Diese Modulation, die auf der Grundlage des hypothesengenerierenden Abschnitts erfolgte, stellt die Basis für die derzeit laufende Hauptstudie dar. Die Hypothesenprüfung erfolgt über die Aufstellung und Überprüfung von Strukturgleichungsmodellen. Es wird in einer bundesweiten Studie überprüft, welche Limitationen in der betrieblichen Bildungsarbeit pandemiebedingt entstanden sind und wie sich diese im Zuge aktueller Ausbildungsprozesse auswirken. Auf der Grundlage der gewonnenen Studienergebnisse wird ein Schulungsangebot für das betriebliche Bildungspersonal in der Pflege konzipiert. Ziel ist es, die betriebliche Bildungsarbeit krisenfester werden zu lassen. Das Schulungsangebot wird als Massive-Open-Online-Course (MOOC) angelegt.

Im Rahmen dieses Vortrags werden nach Darlegung der Ausgangslage zunächst der Forschungsgegenstand sowie das Forschungsdesign präsentiert. Mit Bezugnahme auf den derzeitigen Projektstand werden anschließend die Ergebnisse des hypothesengenerierenden Abschnitts aus der Vorstudie sowie erste Ergebnisse aus der Hypothesenprüfung der Hauptstudie über eine Strukturgleichungsmodellierung vorgestellt und diskutiert. Der Beitrag schließt mit ersten Implikationen für die Entwicklung des MOOC und Reflexionen zu möglichen Limitationen.

PS-47: Kooperation und Problemlösen im Unterricht

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 256

Chair der Sitzung: Christina Wekerle

Decoding Dynamics: Potentiale der Advanced State Space Grids Methode zur Analyse dynamischer Interaktionen im Unterricht

Niklas Litzenberger, Sebastian Wurster, Andreas Pysik

Johannes Gutenberg-Universität, Deutschland

Theoretischer Hintergrund

Unterricht zeichnet sich durch komplexe dynamische Interaktionen aus (Maskus, 1976). Dies stellt die Unterrichtsforschung vor eine methodische Herausforderung, wenn diese Dynamik erfasst werden soll. Häufig werden bislang Methoden eingesetzt, die auf globale Einschätzungen des Unterrichts abzielen (Praetorius et al, 2018), wobei Aussagen über zeitabhängige dynamische Prozesse undokumentiert bleiben (Bell et al. 2019) oder nicht Bestandteil der Analysen sind (Praetorius et al, 2020). Aus diesem Grund haben jüngst einige Forschungsgruppen den Wunsch nach einer Erweiterung existierender Modelle und Methoden formuliert (Dietrich et al, 2022; Dirk & Nett, 2022; Pekrun & Marsh, 2022; Moeller et al., 2022).

Einen vielversprechenden Ansatz zur Analyse dieser zeitabhängigen Interaktionen im Unterricht bieten die sogenannten State Space Grids (SSGs) (Hollenstein, 2007). Dabei werden die Bewertungen zweier Indikatoren in festen Zeitabständen in einem zweidimensionalen Koordinatensystem visualisiert (Lewis et al., 1999; Hollenstein, 2007). Durch diese Darstellung können insbesondere auch nicht-lineare Zusammenhänge gefunden werden (Smit et al., 2016) und Unterricht als dynamisches System untersucht werden (Hollenstein, 2013). Mit dieser Methode konnten bereits einige Forschungsgruppen Rückschlüsse über die Gründe von Veränderungen in der Interaktion zwischen Lehrkräften und Lernenden ziehen (Scherzinger et al., 2020; Pennings & Hollenstein, 2019; Mainhard et al., 2012; Smit et al., 2016; Turner & Christens, 2020). Jedoch beschränkt sich die Analyse von Unterricht mit Hilfe von SSGs meist nur auf die (qualitative) Beschreibung der SSG Grafiken (Smit et al., 2016). Zudem gebendie von Hollenstein (2013) entwickelten Maße nur wenig zusätzliche Analysemöglichkeiten und sind schwer zu interpretieren, da es dazu keine Referenzwerte gibt.

Fragestellung

Wir bauen die Idee der SSGs in der Advanced State Space Grid (ASSG) Methode weiter aus. Unsere Fragestellung dabei ist, ob es möglich ist mit dieser Methodenerweiterung tiefergehende Analysemöglichkeiten zu Interaktionen im Unterricht zu erhalten. Dafür haben wir neue Parameter entwickelt und grafische Elemente um dynamische Prozesse „messbar“ zu machen und nicht nur visuell zu analysieren. Dazu gehören neue Zusammenhangsmaße zwischen den erfassten Indikatoren, Maße zur zeitabhängigen Interraterreliabilität, Gütetests, sowie Parameter zu Streudichten und Abweichungen. Basierend auf den neuen Parametern wurden verschiedene Kategorien von bestimmten Zusammenhängen der Indikatoren gebildet, um die inhaltliche Interpretation zu erleichtern. Zudem haben wir untersucht, ob sich für diese neuen Parameter Referenzwerte bilden lassen, um die ASSG-Grafiken zu kategorisieren.

Methode

Für diese neuen (nicht-)linearen Zusammenhangskategorien haben wir mit Hilfe von über 42 Millionen computergestützten Simulationen Referenzwerte bestimmt. Diese wurden in einem zweiten Schritt durch Bewertungen von 17 Ratern von fünf Unterrichtseinheiten validiert.

Basierend auf den Referenzwerten wurde ein Algorithmus entwickelt, der eine automatisierte Einordnung der grafischen Darstellung der ASSGs liefert und auf Grund der Referenzwerte transparent darlegt. Die Ergebnisse der Kategorisierung des Algorithmus wurde anhand von 300 Beispiel ASSGs mit insgesamt 1200 Eigenschaften mit fünf trainierten Rater verglichen und validiert.

Ergebnisse und ihre Bedeutung

Durch unsere Untersuchung können wir zeigen, dass die neuen Parameter einen tieferen Einblick in sieben verschiedene Zusammenhangskategorien zwischen Unterrichtsaspekten ermöglichen. Beispielsweise lassen sich zeitabhängige Verhaltensänderungen in der Interaktion von Lehrenden und Lernenden numerisch erfassen und (neue) Zusammenhänge zwischen verschiedenen Verhaltensaspekten finden.

Es zeigt sich zudem, dass der Auswertungsalgorithmus wesentlich schneller (1107% schneller) und zuverlässiger zu korrekten Ergebnissen kommt (96,75% korrekt) und damit tiefergehende und zuverlässigere Analysen ermöglicht als menschliche Rater (93,5% korrekt). Insbesondere kann der Algorithmus weitere dynamische Effekte im Unterricht finden, die selbst Expert:innen in der Methode nicht gefunden haben. Dadurch kann diese Form von Analyseunterstützung zeitsparend zusätzliche Einordnungsvorschläge liefern, die zudem zusätzliche tiefere Einblicke ermöglicht.

Im Vortrag wird die neue ASSG-Methode vorgestellt und mögliche Implikationen aufgezeigt. Zudem wird der computergestützte Simulationsprozess, der zu der Bestimmung von Referenzwerte geführt hat, diskutiert. Weitergehend werden die Resultate des Vergleichs zwischen menschlichen Rater und dem computergestützten Algorithmus vorgestellt und die Potentiale einer solchen automatisierten Analyseunterstützung kritisch beleuchtet.

Effekte kooperativen Lernens mit dynamischen multimedialen Lernmaterialien: eine Metaanalyse

Laura Schultze, Nicoletta Bürger, Christof Wecker

Stiftung Universität Hildesheim, Deutschland

Theoretischer Hintergrund

Während in der Multimediaforschung viel Wert auf die Gestaltung von Multimedia gelegt wurde, um sachfremde kognitive Belastung zu reduzieren und generative Verarbeitung zu erleichtern (Mayer & Fiorella, 2021), wurde nicht intensiv untersucht, wie der Einsatz von dynamischem Multimedia in verschiedene Lernarrangements in einem Klassenzimmer eingebettet werden kann. Eine naheliegende und im Schulalltag verbreitete Möglichkeit des Einsatzes dynamischer multimedialer Lernmaterialien im Unterricht besteht darin, Schülerinnen und Schüler mit diesen in kooperativen Settings lernen zu lassen. Laut der *Collaborative Cognitive Load Theory* (Janssen & Kirschner, 2020) kann Kooperation dazu beitragen, kognitive Ressourcen für die individuelle generative Informationsverarbeitung freizusetzen, indem das individuelle Arbeitsgedächtnis jeder beteiligten Person einen Teil des kollektiven Arbeitsgedächtnisses der Gruppe bildet. Eine kooperative Lernumgebung mit dynamischen multimedialen

Lernmaterialien könnte daher Lernergebnisse wie Wissens- und Kompetenzerwerb fördern. Bislang scheint allerdings keine systematische Überblicksarbeit zu den Effekten von Kooperation beim Lernen mit dynamischen multimedialen Lernmaterialien (sowie spiegelbildlich von dynamischen multimedialen Lernmaterialien beim kooperativen Lernen) auf die kognitive Belastung sowie auf Lernergebnisse vorzuliegen (Janssen und Kirschner, 2020). Der vorliegende Beitrag soll diese Lücke mit metaanalytischen Methoden schließen.

Fragestellung

Welche durchschnittlichen Effekte haben *Kooperation* beim Lernen mit dynamischen multimedialen Lernmaterialien sowie *dynamische multimediale Lernmaterialien* beim kooperativen Lernen jeweils auf die *kognitive Belastung* und auf *Lernergebnisse*?

Methodik

Identifikation der Primärstudien. Als *Einschlusskriterien* wurde festgelegt, dass (1) Primärstudien als Zeitschriftenartikel oder Dissertationen in englischer oder deutscher Sprache vorliegen müssen, (2) als unabhängige Variable Kooperation beim Lernen mit dynamischen multimedialen Lernmaterialien oder dynamische multimediale Lernmaterialien beim kooperativen Lernen sowie (3) als abhängige Variablen Lernergebnisse oder kognitive Belastung untersuchen müssen. Die *Literaturrecherche* erfolgte in den bibliografischen Datenbanken *APA PsycInfo* und *ERIC* anhand eines booleschen Suchausdrucks auf der Grundlage der Einschlusskriterien. Für die *Auswahl der Primärstudien* wurden im ersten Schritt die Abstracts und Titel bezüglich der (mutmaßlichen) Erfüllung der Einschlusskriterien sowie der Verwendung einer quantitativen Forschungsmethodik codiert (Cohen's $k = .93$). Die dadurch identifizierten 155 Publikationen wurden im zweiten Schritt anhand der Volltexte auf die gleichzeitige Rezeption des dynamischen Lernmaterials in einer kooperativen Lernumgebung, berichtete Lernergebnisse und berichtete kognitive Belastung codiert. Auf diese Weise wurden 17 Publikationen für die Aufnahme in die Metaanalyse ausgewählt.

Extraktion von Informationen aus den Primärstudien. Auf der Grundlage der Volltexte wurden die unabhängigen Variablen (dynamische multimediale Lernmaterialien vs. Kooperation) codiert und die Effektstärken ermittelt.

Statistische Analysen. Zur Berücksichtigung untereinander abhängiger Effektstärken innerhalb von Untersuchungen wurden *Drei-Ebenen-random-effects-Metaanalysen* durchgeführt (Harrer et al., 2021).

Ergebnisse

Die durchschnittlichen Effekte von *Kooperation* beim Lernen mit dynamischen multimedialen Lerneinheiten auf die *Lernergebnisse* ($g = 0.164$; $k = 31$; $p = 0.239$; $CI_{95\%} = [-0.115; 0.443]$) und die *kognitive Belastung* ($g = -0.130$; $k = 14$; $p = 0.360$; $CI_{95\%} = [-0.426; 0.166]$) sind vernachlässigbar und statistisch nicht signifikant. Dagegen ist der durchschnittliche Effekt von *dynamischen multimedialen Lernmaterialien* beim kooperativen Lernen auf die *Lernergebnisse* mittelgroß bis groß und statistisch signifikant ($g = 0.711$; $k = 28$; $p = 0.033$; $CI_{95\%} = [0.064; 1.358]$), während diese keinerlei Effekt auf die *kognitive Belastung* zu haben scheinen ($g = -0.005$; $k = 15$; $p = 0.985$; $CI_{95\%} = [-0.607; 0.597]$).

Diskussion

Der zwar negative, jedoch vernachlässigbare und statistisch nicht abgesicherte Effekt von Kooperation beim Lernen mit dynamischen multimedialen Lernmaterialien auf die kognitive Belastung liefert keine klare Evidenz für die Kernannahme der *Collaborative Cognitive Load Theory*. Entsprechend findet sich auch kein statistisch abgesicherter Effekt von Kooperation auf Lernergebnisse. Die relativ hohe Effektstärke für dynamische multimediale Lernmaterialien beim kooperativen Lernen steht im Einklang mit Befunden aus früheren Metaanalysen mit einem Fokus auf individuelle Lernsettings (Höfler, 2008). Wie – und auf welcher theoretischen Grundlage – durch eine kooperative Einbettung dynamischer multimedialer Lernmaterialien im Unterricht Lernerfolge gesteigert werden können, bleibt weiterhin ein Thema für die Forschung.

Förderung der professionellen Wahrnehmung von kooperativem Lernen im Lehramtsstudium – Wirkt Problem-Solving Prior to Instruction?

Christina Wekerle, Lara Tar, Özün Keskin, Ingo Kollar
Universität Augsburg, Deutschland

Theoretischer Hintergrund

Kooperatives Lernen erweist sich im Vergleich zum individuellen Lernen als hochwirksam mit Blick auf die Lernergebnisse von Lernenden (Kyndt et al., 2013). Allerdings hängen die Effekte kooperativen Lernens von der Fähigkeit der Lehrkraft ab, die Zusammenarbeit angemessen zu unterstützen (Kaendler et al., 2014). Entsprechend benötigen Lehrkräfte professionelle Unterrichtswahrnehmung, um kooperative Lernprozesse effektiv zu unterstützen. Professionelle Unterrichtswahrnehmung beschreibt die Fähigkeit von Lehrpersonen, relevante (kooperative) Unterrichtssituationen (z.B. aktive Teilnahme der Schüler*innen) basierend auf wissenschaftlichen Erkenntnissen wahrzunehmen und zu interpretieren (Blomberg et al., 2011). Allerdings scheint diese Fähigkeit bei Lehramtsstudierenden oftmals gering ausgeprägt zu sein (Modell, 2017). Eine instruktionale Förderung kann unterschiedliche Formen auf dem Kontinuum zwischen instruktionalen Erklärungen und Problemlöseaktivitäten annehmen. Als vielversprechend hat sich dabei im MINT-Bereich der Ansatz erwiesen, diese beiden Aktivitäten zu kombinieren. Hier zeigen Studien, dass Problemlösen gefolgt von instruktionalen Erklärungen (problem-solving prior to instruction; PS-I) im Vergleich zu instruktionalen Erklärungen gefolgt von Problemlösen (instruction prior to problem-solving; I-PS) zu einem erhöhten konzeptionellen Verständnis und Wissenstransfer führt (Sinha & Kapur, 2021). Jedoch gibt es in weniger strukturierten Bereichen wie der Lehrer*innenbildung bisher kaum und uneinheitliche Forschungsergebnisse zu den Effekten von PS-I vs. I-PS (Krause-Wichmann et al., 2023).

Deshalb wurde in der vorliegenden Interventionsstudie der Effekt von PS-I vs. I-PS auf die professionelle Wahrnehmung von kooperativem Lernen untersucht. Es wurde angenommen, dass PS-I zu einer verbesserten professionellen Wahrnehmung von kooperativem Lernen im Vergleich zu I-PS und einer Kontrollbedingung (PS-PS) führt.

Methode

$N = 151$ Lehramtsstudierende ($M_{Alter} = 21.23$, $SD = 3.89$; 79.5% weiblich) wurden zufällig einer von drei Bedingungen zugewiesen, bestehend aus einer Kontrollbedingung (PS-PS) und zwei Interventionsbedingungen (I-PS, PS-I).

Nach einem Wissenstest zu kooperativem Lernen (16 Items) wurde die Kontrollgruppe (PS-PS, $N = 52$) in einer ersten Runde gebeten, drei Videosequenzen von kooperativen Lernsituationen hinsichtlich der aktiven Beteiligung, des Respekts, der Elaboration und der gegenseitigen Bezugnahme zu bewerten, gefolgt von einer zweiten Bewertungsrunde mit den identischen drei Videos. Die Studierenden sollten dann ihre Bewertungen der ersten und zweiten Runde für jede der drei Videosequenzen

vergleichen und Gemeinsamkeiten und Unterschiede herausarbeiten. In der I-PS-Bedingung ($N = 49$) erhielten die Studierenden zunächst Expert*innenbewertungen der drei Videosequenzen und wurden dann in einer zweiten Runde gebeten, diese Videosequenzen selbst zu bewerten. In der abschließenden Phase sollten die Studierenden ihre Lösungen mit den Expert*innenlösungen vergleichen. In der PS-I-Bedingung ($N = 50$) wurden die Studierenden gebeten, die drei Videosequenzen in der ersten Runde zu bewerten und erhielten dann in der zweiten Runde die Expert*innenbewertungen, mit denen sie ebenfalls ihre eigenen Lösungen vergleichen sollten. Im Nachtest wurden alle Studierenden gebeten, zwei weitere Videos von Kooperationssituationen basierend auf vier Likert-skalierten-Items zu bewerten, die die aktive Teilnahme der Gruppenmitglieder, ihre Reaktionen auf die Beiträge anderer, ihre Elaboration von Aussagen und den Grad, in dem sie aufeinander aufbauen, adressierten (1 'nicht vorhanden' bis 4 'vorhanden'; Cronbach's $\alpha = .70$). Die durchschnittliche Übereinstimmung mit der Expert*innenbewertung bildete den professionellen Wahrnehmungsscore (gleiche Bewertung wie Expert*in = 1 Punkt, andere Bewertung als Expert*in = 0 Punkte).

Ergebnisse und Diskussion

Eine MANCOVA mit professioneller Wahrnehmung von kooperativem Lernen als abhängiger Variable (Nachtest), der experimentellen Bedingung als unabhängiger Variable und dem Vorwissen über kooperatives Lernen als Kovariate zeigte keine signifikanten Unterschiede zwischen den drei Bedingungen ($F(3,149) = 3.15, p = .27$). Die fehlenden Effekte von PS-I auf die professionelle Wahrnehmung von kooperativem Lernen von Lehramtsstudierenden im Vergleich zu den anderen Bedingungen könnten (1) auf Variationen im Design von Problemlöse- und Instruktionsaktivitäten in den verschiedenen PS-I Studien sowie (2) die höhere Komplexität des Settings im Vergleich zu anderen PS-I Studien (Buszard et al., 2017) zurückzuführen sein.

Kollaboratives Problemlösen in wissensreichen Domänen: Ein Multi-Studien-Strukturgleichungsmodell

Laura Brandl¹, Matthias Stadler², Constanze Richters², Anika Radkowsch³, Ralf Schmidmaier⁴, Martin Fischer², Frank Fischer¹

¹Department Psychologie, Ludwig-Maximilians-Universität München, Deutschland; ²Institut für Didaktik und Ausbildungsforschung in der Medizin (DAM), LMU Klinikum, Ludwig-Maximilians-Universität München, Deutschland; ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland; ⁴Medizinische Klinik und Poliklinik IV, LMU Klinikum, Ludwig-Maximilians-Universität München, Deutschland

Kollaborative Fähigkeiten sind beim kollaborativen Lernen und kollaborativen Problemlösen in der beruflichen Praxis relevant (Fiore et al., 2018). Bisherige Modelle des kollaborativen Problemlösens (z.B. OECD, 2017) berücksichtigen bisher kaum wissensreiche Probleme. Eine Ausnahme ist das CDR-Modell (Radkowsch et al., 2022), welches interdisziplinäres kollaboratives Diagnostizieren in wissensreichen Domänen wie die Medizin fokussiert. Nach diesem Modell erfordert es u.a. kollaborativer Diagnoseaktivitäten zur gemeinsamen Informationsverarbeitung um akkurate, effiziente und ausreichend begründete diagnostische Ergebnisse zu erzielen: eine klare Kommunikation über benötigte Informationen für sich (Evidenzen elizitieren) und den Kollaborationspartner:innen (Evidenzen teilen) sowie das Austauschen relevanter Verdachtsdiagnosen (Hypothesen teilen). Das Modell nimmt an, dass individuelle Voraussetzungen wie soziale Fähigkeiten, Inhaltswissen und Wissen über die Domäne des Kollaborationspartners (Kollaborationswissen) positiv mit der Qualität kollaborativer Diagnoseaktivitäten und diese wiederum positiv mit diagnostischen Ergebnissen zusammenhängen. Zudem werden indirekte Effekte von den individuellen Voraussetzungen auf das Diagnoseergebnis über die kollaborativen Diagnoseaktivitäten angenommen. Bisherige empirische Studien (z.B. Radkowsch et al., 2021; Richters et al., 2024) haben nur einzelne Aspekte des CDR-Modells in simulationsbasierten Lernumgebungen überprüft.

Das Ziel ist es, dass CDR-Modell umfassend in einer Studie mit Daten von 504 Medizinstudierenden aus drei Erhebungsmodalitäten, Labor ($n = 157$), Onlineerhebung ($n = 155$) sowie im Feld des Medizinstudiums ($n = 192$) mittels Strukturgleichungsmodellen zu überprüfen. Zunächst beantworten die Medizinstudierenden Wissenstest (Inhalts- und Kollaborationswissen) und Fragen zur Selbsteinschätzung der sozialen Fähigkeiten. Dann diagnostizieren sie einen fiktiven aber realitätsnahen Patient:innenfall in der Rolle eines/r Internist:in in Kollaboration mit einem agentenbasierten Radiologin in einer simulationsbasierten Lernumgebung. Evidenzen elizitieren, Evidenzen teilen, Hypothesen teilen werden mittels Qualitätsindikatoren basierend auf Prozessdaten erfasst. Die Akkuratheit des diagnostischen Ergebnisses wird durch den Grad der Übereinstimmung der angegebenen Diagnose mit einer Expertenlösung (Akkuratheit), die Begründung durch die Verwendung relevanter Informationen und die Effizienz durch die benötigte Zeit operationalisiert.

Für jede Erhebungsmodalität wurde ein Strukturgleichungsmodell in MPlus geschätzt, das eine gute Modellpassung zeigte, $\chi^2(75) = 68.31-93.82$, RMSEA = 0.00–0.04, CFI = 1.00. Es werden nur Pfadkoeffizienten interpretiert, die in mindestens zwei Studien bedeutsam sind ($\beta > 0,1$). Entsprechend der Vorhersagen des CDR Modells gibt es einen Effekt des Inhaltswissen auf das Elizitieren und Teilen von Evidenzen. Kollaborationswissen zeigt nur einen Effekt auf das Elizitieren von Evidenzen, während soziale Fähigkeiten nur einen Effekt auf das Teilen von Evidenzen zeigt. Evidenzen elizitieren zeigt einen Effekt auf die Akkuratheit und Effizienz; Evidenzen teilen auf die Begründung und Hypothesen teilen auf die Effizienz. Entgegen der Vorhersagen des CDR Modells finden wir keine Unterstützung für indirekte Effekte.

Die angenommenen Zusammenhänge im CDR-Modell (Radkowsch et al. 2022) lassen sich teilweise studienübergreifend anwenden. Die Ergebnisse liefern allerdings auch Anregungen für die Weiterentwicklung des CDR-Modells und bieten Erkenntnisse für das Lernen von kollaborativem Problemlösen in Simulationen. Erstens ist es wichtig, Kollaborationswissen zu haben, um gezielt Evidenzen teilen zu können: Für eine erfolgreiche Zusammenarbeit scheint entscheidend, zu antizipieren, welchen Informationsbedarf der Kollaborationspartner hat. Zweitens ist das Inhaltswissen in simulationsbasierten Lernumgebungen nicht von so großer Bedeutung, wie in anderen Diagnosekontexten (Boshuizen et al. 2020), wahrscheinlich, weil hier diagnostische Aktivitäten beliebig wiederholt werden können und inhaltliches Wissen auch über andere Quellen erschlossen werden kann. Drittens reichen Inhaltswissen und Kollaborationswissen nicht aus für erfolgreiches kollaboratives Diagnostizieren: Die Fähigkeit, sich in kollaborativen Diagnoseaktivitäten wie dem Evidenzen teilen zu engagieren, ist ein weiterer wichtiger Bestandteil. Die hohe Kohärenz der Studien, die nur in der Art der Datenerhebung variieren, spricht für die Robustheit der Ergebnisse, lässt aber offen, ob diese auch auf andere ähnlich komplexe Domänen wie z.B. die Lehrerbildung, übertragbar sind. Zusammenfassend lässt sich schlussfolgern, dass (1) für erfolgreiches kollaboratives Problemlösen in wissensreichen Domänen neben dem Inhaltswissen das Wissen über die Domäne des Kollaborationspartners und (2) die kollaborativen Diagnoseaktivitäten, eine entscheidende Rolle spielen.

PS-48: Vertrauen in Wissenschaft und die Beurteilung wissenschaftlicher Argumente

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 242
Chair der Sitzung: Nina Udvardi-Lakos

Science Communication: How Do Controversial Topics Affect The Evaluation Of Science-Based Arguments?

Nina Udvardi-Lakos¹, Abrama Ella Nikrandt², Louisa Vankann², Alexander Renkl²

¹Eberhard-Karls Universität Tübingen, Germany; ²Albert-Ludwigs-Universität Freiburg, Germany

Background

People encounter a variety of science-based arguments online, meant to influence political, economic, and personal choices and behavior. Evaluating the reliability of such arguments is often relevant for making informed choices (e.g., whether to get vaccinated). However, people may show biases when evaluating arguments (Larson et al., 2009). For example, participants seem to prefer explanation-based arguments over arguments supported with evidence (Glassner et al., 2005; Slusher & Anderson, 1996), even though evidence represents a stronger support for scientific claims (Keil, 2006, 2010). This preference could depend on pragmatic factors, such as the availability of further information (Brem & Rips, 2000). In a high information-availability context, people may prefer evidence over explanations (Brem & Rips, 2000).

Although offering high information availability, the current online information environment also contains increasing amounts of misinformation and disinformation (Barzilai & Chinn, 2020). Additionally, many discussions of controversial topics seem to become more and more one-sided (Karimi & Richter, 2021; O'Keefe, 1999; Wolfe et al., 2013), offering little reliable information from different perspectives. Particularly for controversial topics, people should be able to evaluate arguments. Thus, we will examine how the controversiality of arguments affects their evaluation. We assume that topics seen as more controversial will be perceived as more relevant, leading to a higher motivation to engage with the topic and a higher perceived prior knowledge. Following the elaboration-likelihood model, these factors should lead to a higher level of elaboration, which could support a thorough evaluation of arguments (Petty & Cacioppo, 1988). The evaluation of arguments of controversial topics could lead to a preference for evidence-based arguments, due to the high information-availability context, and a higher perceived need for further information.

Methods

We conducted two studies to test how the controversiality of arguments (high, medium or low; tested in pilot studies) affects argument evaluation. In the laboratory-based Study 1, 143 participants (mean age = 23.8, 64% female) were presented with six arguments sequentially. For each argument, participants rated the reliability of two support types for the claim, an evidence-based and an explanation-based support. We assessed motivation and mental effort invested, and perceived controversiality as manipulation check. Study 2 was conducted online study with 141 participants (mean age = 23.2, 81% female) evaluating six arguments sequentially, where the arguments contained both evidence and explanations, and differed only in their level of controversiality. We measured perceived reliability and need for further information.

Results and Discussion

ANOVAs with planned contrasts were used to analyze the data. In Study 1, arguments with high controversiality led to a stronger preference for evidence-based supports compared to arguments with medium and low controversiality. Higher controversiality was associated with higher motivation and higher invested mental effort, but these factors did not mediate the effect of controversiality on the evaluation of the support type.

In Study 2, low-controversiality arguments were perceived as less reliable than medium- and high-controversiality arguments. High-controversiality arguments led to a higher need for further information compared to medium- and low-controversiality arguments. The controversiality effects on perceived reliability and on need for information were both mediated by the motivation to engage with the topic.

Findings suggest that more controversial topics lead to a higher level of elaboration, indicated by a higher mental effort and need for further information. This elaboration could support the stronger preference for evidence over explanation observed for more controversial arguments. However, more controversial arguments were also seen as more reliable than less controversial arguments, suggesting that people still show biases in argument evaluation even with controversial topics. Further studies should investigate how people could be supported to overcome biases when evaluating science-based arguments to make informed choices.

Was führt zum Glauben an Neuromythen? Initiale Befunde einer explorativen Studie über die Prädiktoren der Zustimmung zu Neuromythen und Neurofakten

Eric Klopp, Robin Stark

Universität des Saarlandes, Deutschland

Theoretischer

Neuromythen sind Fehlkonzepte, die entstehen, wenn neurowissenschaftlich fundierte Fakten falsch verstanden, missverstanden oder falsch zitiert werden, um diese in Bildungskontexten zu nutzen[1]. Neben Neuromythen existieren auch Neurofakten, d.h. Wissen über gültige wissenschaftliche Konzepte, die in Bildungskontexten berücksichtigt werden können. Nur wenige Studien haben untersucht, welche Faktoren Neuromythen und -fakten bedingen (z.B. [2,3,4,5,6,7,8]), wobei solche Evidenz die Entwicklung von Interventionen zur Korrektur von Neuromythen unterstützen kann. Mögliche Faktoren, die aktuell diskutiert werden, sind bspw. die Nutzung von Medien (insb. Alltagsmedien oder wissenschaftliche Medien), Persönlichkeitsmerkmale wie Fixed Mindset, der Glaube an Intuition oder motivationale Variablen wie das Bedürfnis nach kognitiver Geschlossenheit, oder wissensbezogene Variablen wie epistemologische Überzeugungen. Darüber hinaus können auch das Geschlecht und das Studienfach für die Befürwortung von Neuromythen oder Neurofakten relevant sein (z.B. [3,9]). Weitere potenzielle Faktoren wurden in der aktuellen Forschung entweder nicht, z.B. Persönlichkeitsmerkmale wie Offenheit für Erfahrungen oder Spiritualität, oder nur unzureichend berücksichtigt, wie bspw. epistemologische Überzeugungen. Die vorliegende Pilotstudie zielt darauf ab, explorativ mögliche Prädiktoren für Neuromythen bzw. Neurofakten zu ermitteln, um Ansatzpunkte für die weitere Forschung bzw. Interventionen zu ermitteln.

Hintergrund

Methoden

Insgesamt nahmen 258 Studierende teil, davon 76 Psychologiestudierende, der Rest Lehramtsstudierende (MAlter=22.7, sdAlter=4.5; MSemester=4.6, sdSemester=2.8). Die Teilnehmenden beantworteten einen Fragebogen mit 25 Neuomythen und 15 Neurofakten, es wurde ein Summenscore für Neuomythen ($\lambda_6=.67$) und Neurofakten ($\lambda_6=.65$) berechnet, der die Zustimmung zu Neuomythen bzw. Neurofakten angibt. Erfasst wurde die Häufigkeit der Nutzung wissenschaftlicher Originalliteratur (z.B. Lehrbücher oder Fachzeitschriften), populärwissenschaftlicher Medien, Fernsehdokumentationen, Radiosendungen und sozialer Medien mit Bildungsbezug. Weiterhin wurden folgende Konstrukte aus den oben genannten Bereichen erfasst: Spiritualität ($\alpha=.78$), Fixed Mindset ($\alpha=.50$), Offenheit ($\alpha=.74$), intuitives Denken ($\alpha=.60$), Bedürfnis nach kognitiver Geschlossenheit ($\alpha=.69$), Evaluationsbedürfnis ($\alpha=.49$), wissenschaftliches Interesse ($\alpha=.78$), Persönliche Rechtfertigung ($\alpha=.75$), Rechtfertigung durch Autorität ($\alpha=.74$), Rechtfertigung durch mehrere Quellen ($\alpha=.64$), Rechtfertigung durch die wissenschaftliche Gemeinschaft ($\alpha=.61$) und die Sicherheit des Wissens ($\alpha=.68$).

Zur Bestimmung der Prädiktoren von Neuomythen bzw. -fakten wurde eine best-subset-Regression unter Verwendung des AIC als Auswahlkriterium durchgeführt mit anschließender Bestimmung der relativen Wichtigkeit der Prädiktoren[10,11]. Das Geschlecht der Teilnehmer sowie deren Studienfach wurden als Dummy-Variablen (Referenzkategorien: weiblich bzw. Lehramtsstudium) im Modell erfasst.

Ergebnisse

Das selektierte Regressionsmodell für Neuomythen erklärt 19.70% der Varianz, die Prädiktoren sind in absteigender Folge der relativen Wichtigkeit Studienfach ($b=-2.39$, $CI=[-2.83;-1.96]$), Geschlecht ($b=-1.25$, $CI=[-1.72;-0.77]$), Spiritualität ($b=0.08$, $CI=[0.05;0.11]$), Nutzung bildungsbezogener Social-Media-Kanäle ($b=0.38$, $CI=[0.23;0.54]$) sowie wissenschaftliches Interesse ($b=-0.07$, $CI=[-0.11;-0.03]$). Das für Neurofakten selektierte Regressionsmodell besitzt eine Varianzaufklärung von 15.58%, die Prädiktoren sind in absteigender Folge der relativen Wichtigkeit Persönliche Rechtfertigung ($b=-0.20$, $CI=[-0.26;-0.14]$), Nutzung wissenschaftlicher Originalliteratur ($b=0.40$, $CI=[0.24;0.57]$), Nutzung von Dokumentationssendungen ($b=0.26$, $CI=[0.13;0.40]$), wissenschaftliches Interesse ($b=0.07$, $CI=[0.04;0.11]$), Rechtfertigung durch Autoritäten ($b=0.10$, $CI=[0.04;0.16]$), Spiritualität ($b=0.04$, $CI=[0.02;0.07]$), Studienfach ($b=0.55$, $CI=[0.20;0.91]$), sowie Fixed Mindset ($b=-0.12$, $CI=[-0.12;-0.05]$).

Diskussion

Eine exemplarische Betrachtung der Befunde zeigt, dass für Neuomythen das Studienfach der wichtigste Prädiktor ist, vermutlich bedingt durch eine tiefere Elaboration entsprechender neurowissenschaftlicher Inhalte im Psychologiestudium. Hinsichtlich der Wichtigkeit des Geschlechts für die Zustimmung zu Neuomythen ist zu vermuten, dass damit zusammenhängende Drittvariablen von Bedeutung sind. In Bezug auf Neurofakten zeigt sich, dass eine geringere Ausprägung der epistemologischen Überzeugung Persönliche Rechtfertigung, d.h. eine geringe Überzeugung, dass Wissen aus persönlichen Meinungen besteht, dem Erwerb korrekter wissenschaftlicher Konzepte zuträglich ist, was auch die Relevanz wissensbezogener Variablen belegt. Weiterhin zeigt sich die Wichtigkeit der Nutzung wissenschaftlicher Originalliteratur hinsichtlich des Wissens über Neurofakten. Auch zeigt sich, dass das motivationale Konstrukt wissenschaftliches Interesse sowohl für Neuomythen als auch Neurofakten von Bedeutung ist. Weiterer Forschungsbedarf das Persönlichkeitsmerkmal Spiritualität, das sowohl für Neuomythen als auch Neurofakten ein positiver Prädiktor ist. Exemplarische Limitationen der Studie sind die ungleiche Verteilung des Studienfachs sowie die geringe Reliabilität einiger Skalen.

Assessing Trust in Science: Development and Validation of a Short Scale for Adolescents and Adults

Diana Steger¹, Astrid Schütz²

¹Leibniz-Institut für Bildungsverläufe, Deutschland; ²Otto-Friedrich Universität Bamberg

Theoretical Background

Trust in science is a critical factor in promoting informed decision-making and societal progress - especially in a time of rapid technological change, 'alternative facts', misinformation, and related social challenges (Nadelson et al., 2014; Schoor & Schütz, 2021). Trust in science is not only relevant to public discourse but also has major impact on people's decisions, preferences, and behaviors in education, e.g. the choice of subjects at school or university or the motivation to learn about science and invest in such learning (Gauchat, 2012). However, the construct, which includes cognitive as well as emotional aspects, is not yet well-understood. Particularly, evidence on developmental trajectories and its predictors and precursors is scarce. To this end, we aimed to develop a short scale to be used in panel studies and thus gain more insight into the respective developmental trajectories and its influencing factors.

Study 1

In study 1 we analyzed cross-sectional data from 449 adults (16 – 75 years, mean = 37.66 years, SD = 16.97; 52.33% women) with an 8-item version to assess trust in science (Schoor & Schütz, 2021), demographic data, the Big Five (BFI-2-XS; Rammstedt et al., 2020; Soto & John, 2017), a 3-item measure of conspiracy mentality (adapted from Imhoff, 2013), and a 3-item measure of law-and-order positions (Ulbrich-Herrmann, 2001). We computed all possible item-sets and compared model fit (CFI, RMSEA, SRMR), factor saturation (McDonald's ω), and factor loading patterns to derive a psychometric sound short scale of trust in science. We identified a 4-item version with good fit in a unidimensional model (CFI = .997, RMSEA = .046, SRMR = .013), excellent reliability ($\omega = .835$), and homogeneous factor loadings ($\min_{\lambda} = .730$). The scale had small positive correlations with agreeableness ($r = .19$, $p < .05$) and conscientiousness ($r = .10$, $p < .05$), and a small negative correlation with neuroticism ($r = -.13$, $p < .05$). It showed negative correlations with conspiracy mentality ($r = -.31$, $p < .05$) and law-and-order positions ($r = -.10$, $p < .05$). There was no significant correlation with age ($r = .06$, $p = .24$), nor were there systematic differences between participants from different educational backgrounds. We found small but significant gender differences, with men reporting slightly higher scores ($t = -2.790$, $df = 423$, $p < .05$).

Study 2

In study 2 we analyzed data from 298 parent-adolescent dyads (children's age ranged from 13 – 15 years, mean = 13.77 years, SD = 0.71; 47.9% female) and replicated the model fit in both adults (CFI = 1.00, RMSEA = .000, SRMR = .008, $\omega = .861$) and adolescents (CFI = 1.00, RMSEA = .000, SRMR = .003, $\omega = .886$) for our short measure. The scale was invariant across age groups, gender, and educational groups, which renders it suitable for assessing trust in science in heterogeneous samples across the lifespan. We also took the dyadic structure into account and estimated a combined model of children's and parent's trust in science (CFI = .987, RMSEA = .063, SRMR = .023), which revealed a latent correlation of .664, indicating a substantial relationship between parents' and children's trust in science.

Relevance

With the present study, we offer a reliable tool to measure trust in science in different demographic groups and show that there seems to be some generational transmission of that belief. Regarding validity there is initial evidence of expected correlations. The scale can be a valuable resource for studies that inform educators and policymakers and can be used in efforts to foster trust in science within various learning environments, including families and formal educational settings.

ENTFÄLLT Thinking fast and slow in der Beurteilung von Effektstärken - verzerrte Relevanzeinschätzung trotz akkurater Urteile

Kristina Bohrer, Florian Kühlwein, Kirstin Schmidt, Samuel Merk

Pädagogische Hochschule Karlsruhe, Deutschland

Theoretischer Hintergrund und Fragestellung

Mithilfe von Evidenz können Lehrpersonen ihre eigene Praxis professionalisieren sowie Schul- und Unterrichtsentwicklung vorantreiben (z.B. Bauer & Kollar, 2023; Bromme et al., 2014). Ein solcher gelingender Prozess evidenzinformierten Handelns ist allerdings voraussetzungsreich, wie idealtypisch von Groß Ophoff und Kolleg:innen dargestellt (2023): Nach der 1) Zielsetzung, der 2) Datenerhebung sowie dem 3) Verständnis der konsultierten wissenschaftlichen Evidenz folgt eine 4) Entscheidung, inwiefern die vorhandene Evidenz für die Praxis genutzt wird. Dabei sind die Charakteristika der Nutzenden, also der Lehrpersonen, wesentlich (vgl. Groß Ophoff et al., 2023). Nicht nur ein möglichst korrektes Verständnis statistischer Kennwerte wie bspw. Effektstärken (Kraft, 2020) ist von Bedeutung, sondern ebenso, für wie relevant die Forschungsergebnisse erachtet werden (Thomm et al., 2021): Angelehnt an die theory of planned behavior (Ajzen, 1991) können solche Relevanzeinschätzungen, vermittelt über eine Verhaltensabsicht, als Einstellungen auf das tatsächliche Verhalten einwirken. Dementsprechend kann die Wahrnehmung der Relevanz beeinflussen, inwiefern die Evidenz in der Praxis genutzt wird (vgl. Prenger & Schildkamp, 2018). Eine Möglichkeit, die Relevanzeinschätzung zu messen, bietet die Erfassung der *Willingness to Pay* (WTP; Brooks et al., 2014; Li et al., 2021) - die ökonomische Bewertung einer Intervention, eines Gegenstandes oder einer Dienstleistung.

In einer experimentellen Studie mit der Fragestellung, ob Lehrpersonen Effektstärken besser visuell oder textuell rezipieren können, wurde unter anderem die WTP als eine abhängige Variable (AV) erhoben. Dabei zeigte sich innerhalb sequentieller WTP Urteile einer Person eine hohe Intraklassenkorrelation ($ICC = 0.62$). Daher und da es sich bei der WTP um eine numerische Schätzung handelt, wird eine Verzerrung der Urteile durch sequentielle Ankereffekte (= Anpassung eines numerischen Wertes an einen bekannten Vergleichsmaßstab; vgl. Bahnik et al., 2019; Li et al., 2021; Mochon & Frederick, 2013; Tversky & Kahneman, 1974) angenommen. Folglich wird im vorliegenden Forschungsvorhaben in einer Sekundäranalyse untersucht, inwieweit Ankereffekte die durch Lehrpersonen wahrgenommenen Effektstärken (AV1) und ihre Relevanzeinschätzungen (AV2) von Evidenz im Bildungsbereich beeinflussen.

Methode

Es wurden $N = 20$ englischsprachige Lehrpersonen via Prolific in einem kontrollierten randomisierten Online-Experiment befragt: Jede:r Teilnehmende erhielt einen Forschungsbericht und anschließend fiktive Ergebnisse (textuell vs. visuell, sechs Effektstärken within-person randomisiert), sodass alle Teilnehmenden in sequentieller Folge insgesamt 12 unterschiedliche Effektstärkedarstellungen beurteilt haben. Als AV wurden u.a. die *Probability of Superiority* (PoS) als Maß für die wahrgenommene Effektstärke erfasst sowie die WTP (Relevanzeinschätzung) über ein offenes Antwortfeld abgefragt.

Um den Ankereffekt in den AV zu untersuchen, werden jeweils die Intraklassenkorrelationen für beide AV berechnet sowie mithilfe von *brms* (Bürkner, 2017) hierarchische Modelle mit random intercepts geschätzt, wobei die jeweils erste Bewertung die darauffolgenden Bewertungen prädiziert.

Ergebnisse

Während für die PoS der $ICC = 0.03$ (95% CI: [0.004, 0.153]) und das Random-Intercept-Modell weder höhere between-person Varianz noch die Abhängigkeit der darauffolgenden Bewertungen von der ersten wahrgenommenen Effektstärke induzieren ($R^2 = .05$), deuten für die Relevanzeinschätzung der $ICC = 0.62$ (95% CI: [0.308, 0.765]) sowie das Random-Intercept-Modell auf einen Ankereffekt in den Relevanzeinschätzungen hin. Die erste Relevanzeinschätzung prädiziert signifikant die darauffolgenden ($R^2 = .60$). Gleichzeitig liegen substantielle Korrelationen der within-person zentrierten WTP ($r = 0.53$) sowie der PoS ($r = 0.60$) mit der wahren Effektstärke vor. Lehrpersonen diskriminieren dementsprechend zwischen den unterschiedlichen Effektstärken und nehmen sie akkurat wahr, die Relevanzeinschätzungen sind absolut gesehen jedoch konstant sehr ähnlich pro Person.

Diskussion

Bei denselben Ergebnissen und präsentierten Effektstärken entstehen unter der Annahme eines sequentiellen Ankereffekts (Bahnik et al., 2019; Mochon & Frederick, 2013) unterschiedliche Relevanzeinschätzungen. Dadurch entstehen heterogene Bedingungen für den Nutzen von Evidenz trotz akkurater Einschätzung der Effektstärke. Die Ergebnisse können darüber hinaus mit Blick auf System 1 (verzerrte Relevanzeinschätzungen) und 2 (nicht verzerrt wahrgenommene Effektstärken; Kahneman & Frederick, 2005) diskutiert werden.

Kritischer Blick gegen Fehlinformationen – Untersuchung von Trainingselementen in Cognitive-Apprenticeship-basierten Lateral-Reading-Trainings

Marvin Fendt, Christian Scheibenzuber

Ludwig-Maximilians-Universität München, Germany

Theoretischer Hintergrund

Fehlinformationen stellen in unserer heutigen Gesellschaft professionsübergreifend ein großes Problem dar (Ecker et al., 2022). Der Glaube an Fehlinformationen kann beispielsweise in klimaschädlichem oder in Extremfällen antidemokratischem Verhalten resultieren. Es gibt eine Vielzahl psychologischer Interventionen gegen Fehlinformationen, die jedoch mitunter in der Kritik stehen, generelle Skepsis statt akkurater Glaubwürdigkeitseinschätzungen zu fördern (Modirrousta-Galian & Higham, 2023). Lateral Reading, ein pädagogischer Ansatz zur Quellenkritik, könnte diesen Einschränkungen entgegenwirken, erfordert jedoch größeren Zeit- und Ressourceneinsatz (Wineburg et al., 2022). Als komplexe Strategie scheint Lateral Reading besonders von instruktionalen Frameworks zu profitieren, beispielsweise Cognitive Apprenticeship, in dem die Technik zuerst modelliert, dann

initial mit intensiver Unterstützung geübt und anschließend zunehmend selbstständig durchgeführt wird (Collins et al., 1991; Fendt et al., 2023).

Fragestellung

Während die Wirksamkeit abgegrenzter Cognitive-Apprenticeship-Elemente bereits erforscht wurde (Matsuo & Tsukube, 2020), wurden unseres Wissens Lateral-Reading- und Cognitive-Apprenticeship-Elemente nicht im Zusammenspiel untersucht. Hierdurch könnten sich insbesondere für großflächige Studien wichtige Implikationen ergeben, welche Faktoren für den Trainingserfolg wichtig sind und gegebenenfalls zumindest substituiert erhalten bleiben sollten.

Infolgedessen untersuchen wir in unserer Studie die Wirksamkeit eines Lateral-Reading-Trainings mithilfe von Lernsequenzanalysen und Performanzmessungen. Es wurde untersucht, inwieweit der Trainingserfolg durch 1) Teilnehmenden-Feedbacks 2a) Lateral-Reading- und 2b) Cognitive-Apprenticeship-Elemente vorhergesagt werden kann.

Methode

$N=241$ Teilnehmende (61% weiblich, 39% männlich, 1% divers; $M(\text{Alter})=28.64$, $SD=13.28$) wurden zufällig entweder einer Vergleichsbedingung ($n=109$) zugewiesen in der sie selbstständig Scaffolding-Material über Lateral Reading lasen, oder einer Apprenticeship-Bedingung ($n=132$), in welcher in etwa einstündigen Lateral-Reading-Trainings das Erkennen erst modelliert, dann gecoached und schließlich selbstständig geübt wurde. Zusätzlich wurden 57 (53% weiblich, 47% männlich; $M(\text{Alter})=24.61$, $SD=9.86$) Trainingsaufnahmen sequenziell kodiert ($\kappa=.74$) auf Lateral-Reading-Bestandteile (Vorbereitung, $M=.08$, $SD=.05$; Webseitenprüfung, $M=.23$, $SD=.12$, Verfassenüberprüfung, $M=.27$, $SD=.09$, Plattformüberprüfung, $M=.33$, $SD=.13$; Fazit, $M=.09$, $SD=.06$). und Cognitive-Apprenticeship-Bestandteile (Modeling $M=.38$, $SD=.22$; Coaching/Scaffolding, $M=.54$, $SD=.23$; Articulation/Reflection, $M=.05$, $SD=.05$; Exploration, $M=.03$, $SD=.06$). Die Teilnehmenden stuften in einem Glaubwürdigkeitstest 6 Quellen (3 verlässliche, 3 unlässliche) ein; Glaubwürdigkeits-Scores wurden anhand des Unterschiedes in der mittleren Zustimmung am Pretest bzw. Posttest zu den beiden Arten von Quellen errechnet ($\Omega=.68/.86$). Als Moderatoren wurden die Teilnehmenden gefragt, ob sie das Training verständlich, interessant, nachvollziehbar, kurzweilig, hilfreich und wiederholungswürdig fanden.

Ergebnisse

Frage 1 wurde mithilfe einer Moderationsanalyse beantwortet. Die Glaubwürdigkeitseinschätzung stieg insgesamt vom Pretest ($M=5.43$, $SD=4.59$) zum Posttest ($M=7.57$, $SD=5.21$), in der Apprenticeship-Gruppe von $M=6.03$ ($SD=4.52$) auf $M=8.71$ ($SD=4.95$) und in der Vergleichsgruppe von $M=4.71$ ($SD=4.60$) auf $M=6.19$ ($SD=5.20$). Im Nullmodell sagte die Kondition Trainingserfolg vorher ($\beta=0.26$, $p=.042$, $R^2=.017$). Dieser Effekt war im Moderationsmodell ($F(13,227)=2.41$, $p=.005$, $R^2=.005$) nicht mehr signifikant, jedoch der Einfluss von Interessantheit ($\beta=1.59$, $p<.001$) und Kurzweiligkeit ($\beta=-1.00$, $p=.032$). Um Frage 2 mittels multipler Regressionsmodelle zu beantworten, wurden als Prädiktoren die Lateral-Reading-Elemente ($F(4,52)=0.70$, $p=.593$) und Cognitive-Apprenticeship-Elemente verwendet ($F(3,53)=1.79$, $p=.160$, $R^2_{adj}=.041$). Coaching/Scaffolding sagte signifikant die Trainingseffektivität vorher ($\beta=0.31$, $p=.025$).

Diskussion

Die Ergebnisse zeigen, dass Lateral Reading eine effektive Intervention ist, die von instruktionalen Frameworks wie Cognitive Apprenticeship profitiert. Der Zusammenhang zwischen Coaching/Scaffolding und Trainingserfolg deutet darauf hin, dass die intensivere Unterstützung förderlich ist. Zusätzlich erscheint das Training die Teilnehmenden stärker aktiviert zu haben, wodurch wiederum der Trainingserfolg verbessert wurde. Insbesondere motivationale Bestandteile scheinen somit wichtig zu beachten zu sein.

Zusätzlich erscheint Cognitive Apprenticeship förderlich, um motivierende lernförderliche Interventionen zu entwerfen. Hierfür erscheint die Teilnehmendenwahrnehmung eine zeit- und kostengünstige Annäherung an Erkenntnisse aufwändigerer Analysen, beispielsweise der Lernsequenzen, zu sein. Für die großflächige, kostenökonomische Implementierung solcher Trainings ergeben sich Implikationen, beispielsweise der Einsatz von KI-Tutoren (beispielhaft implementiert von Zhang & Pradeep, 2023), um weiterhin individualisiertes Coaching bereitzustellen.

Im Vortrag wird zusätzlich auf den Einfluss der Teilnehmendenmerkmale eingegangen. Zusätzlich werden die Implikationen der Ergebnisse hinsichtlich großflächiger Implementationen von Fehlinformationstrainings beispielsweise mit individualisiertem KI-Feedback diskutiert.

PS-49: KI und sprachliche Bildung: Innovationen in der Lernförderung und Leistungsbewertung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 159
Chair der Sitzung: Fabian Zehner

Innovationen beim automatischen Bewerten von Kurztextantworten in PISA: von Fuzzy Lexical Matching für das operative Kodieren bis Large Language Models

Fabian Zehner^{1,2}, Nico Andersen¹, Andrea Horbach³, Hyo Jeong Shin⁴, Emily Kerzabi⁵, Frank Goldhammer^{1,2}

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ²Zentrum für internationale Vergleichsstudien (ZIB); ³FernUniversität Hagen; ⁴Sogang University; ⁵Educational Testing Service

Bereits seit der ersten PISA-Studie werden substanzielle Anteile ihrer Kompetenztests in offenem Antwortformat administriert und dies zumeist in Form offener Textantworten (OECD, 2019b). Seit dem Vorliegen computerbasierter Assessmentdaten können resultierende offene Antworten vermehrt automatisch mithilfe natürlicher Sprachverarbeitung und Maschinenlernens ausgewertet werden, um menschliche Kodierung zu unterstützen und die Kodierkonsistenz zu verbessern.

Dieser Beitrag berichtet von zwei zusammenhängenden Studien. (1) Zunächst wird von der Entwicklung und Evaluation einer unterstützenden Methode zur automatischen Auswertung von Kurztextantworten aus einem Projekt des *PISA Research, Development and Innovation Programme* berichtet. Entwickelt wurde hierfür ein flexibles Methodenframework namens Fuzzy Lexical Matching (FLM), das in der operativen Kodierung der PISA-Erhebungen Einzug erhalten soll. (2) Zweitens berichtet der Beitrag von der anschließenden Integration dieser lexikalischen Methode in ein mehrstufiges Rückfallmodell, das zusätzlich einen modernen state-of-the-art Ansatz inkludiert, bei dem ausgehend von einem vortrainierten Large Language Model itemweise Klassifikatoren trainiert werden. Damit werden die Vorteile generisch einsetzbarer lexikalischer Techniken mit jenen moderner semantikbasierter Methoden vereint. Beide Methoden orientieren sich an den Rahmenbedingungen internationaler Large-Scale Assessments, bei denen automatische Bewertung mit einer Vielzahl von Herausforderungen konfrontiert ist, gleichzeitig aber hohe Anforderungen an Kodiergenauigkeit und Erklärbarkeit bestehen, weshalb nicht einfach auf state-of-the-art-Methoden (vgl. Gombert et al., 2023; Gong & Yao, 2019; Sung, Dhamecha & Mukhi, 2019; Whitmer et al., 2023) zurückgegriffen werden kann.

Fuzzy Lexical Matching baut auf dem seit PISA 2018 operativ implementierten *Machine-Supported Coding System* (Yamamoto, He, Shin & von Davier, 2018) auf. Dieses basiert auf dem trivialen Fakt, dass alle Antworten, die exakt identisch sind, denselben Score zugewiesen bekommen können, und ist somit für alle Testsprachen einsetzbar. Während die Anforderung exakter Gleichheit für einige Items bereits große Einsparungen bei der manuellen Kodierung erbringt, ist die sprachliche Antwortvielfalt bei vielen anderen Items hingegen zu groß. An dieser Stelle greift FLM ein, das Normalisierungstechniken der natürlichen Sprachverarbeitung einsetzt, um leicht verschiedene Antworten zu homogenisieren; so werden zum Beispiel großgeschriebene Buchstaben zu kleinen transformiert, Umlaute normalisiert sowie Affixe und Stoppwörter entfernt. Somit können Antworten wie *der Hund rannte* und *die Hündin rannten* [sic!] als [hund, rannt] zusammengruppiert werden, sofern die manuelle Kodierung (lt. Trainingsdaten) diese als homogen kodiert indiziert. Zentral für FLM ist jedoch, dass je nach Sprache und Item verschiedene Normalisierungsschritte automatisch angepasst werden können: Wird erkannt, dass ein Schritt relevante sprachliche Informationen entfernt (d.h. Kodiergenauigkeit reduziert), wird diese Normalisierung zurückgenommen oder adaptiert.

Bei der empirischen Evaluation wurden Daten der PISA-Lesetests von Erhebungen von 2018 bis 2022 verwendet. Abgedeckt wurden zwölf diverse Testsprachen (von Arabisch über Kasachisch, Finnisch, Hebräisch, Englisch zu Chinesisch), resultierend in $n = 2.820.018$ Antworten. Insgesamt konnten im Durchschnitt 36 Prozent des Kodieraufwands ohne Verlust in der Genauigkeit (-0,05%) eingespart werden.

In einem anschließenden, aktuell noch laufenden Projekt, wurde FLM in ein mehrstufiges Rückfallmodell integriert. Antworten werden dabei zunächst durch FLM als Methode mit höchster Genauigkeit verarbeitet. Falls FLM keine Bewertung vornehmen kann, da keine gleiche, bereits manuell kodierte Antwort gefunden wird, wird die Antwort an einen testsprachen- und itemspezifisch Klassifikator weitergereicht, der auf der Basis von XLM-RoBERTa (Conneau, Lample, Ranzato, Denoyer & Jégou, 2017) optimiert wurde. Ist die Klassifikationssicherheit für eine bestimmte Antwort zu niedrig, wird die Antwort letztlich menschlich kodiert.

XLM-RoBERTa ist ein vortrainiertes vielsprachiges Deep-Learning-Modell, das um einen Klassifikationskopf im neuronalen Netz für die Studie erweitert und damit auf die manuellen Kodierungen trainiert wurde. Erste Ergebnisse zeigen, dass die automatische Bewertung nur auf dieser Basis über die allermeisten Sprachen hinweg gut funktioniert (durchschnittliche Übereinstimmung Mensch-Maschine: $\kappa = .78$), wobei die Übereinstimmung für einzelne Sprachen signifikant schwächer ist (etwa Indonesisch mit $\kappa = .58$), was die Notwendigkeit des mehrstufigen Modells unterstreicht. Die Präsentation wird die finalen Evaluationsergebnisse berichten und zeigen, welche Aufwandsreduktion durch das mehrstufige Modell bei welcher resultierenden Genauigkeit erreicht wird.

Personalisierte Lesematerialien mit generativer KI in der Grundschule: Textinteresse, -schwierigkeit und Aufgabenerfolg in einer Leseförderung mit maßgeschneiderten Inhalten

Ulrich Ludewig¹, Andreas Dengel², Nele McElvany¹, Nell Royal¹, Linder Rustemeier², Sarah Sommer³, Ester Spang³, Caro Zierer²

¹Institut für Schulentwicklungsforschung, Deutschland; ²Lehrstuhl für Informatikdidaktik an der Goethe-Universität Frankfurt am Main; ³Vodafone Stiftung Deutschland

Theoretischer Hintergrund

Lesekompetenz wird als die aktive Generierung von Bedeutung aus dem Geschriebenen definiert (Mullis et al., 2019). Die aktuellen Entwicklungen im Bereich der generativen Künstlichen Intelligenz (KI) bieten große Potenziale für die Förderung der Lesekompetenz durch die Möglichkeit, personalisierte Lernmaterialien zu erstellen (Holmes, & Tuomi, 2022). Generative KI kann Texte basierend auf den persönlichen Interessen von Schüler*innen erstellen, was die Motivation und das Engagement erhöhen kann. Die Passung zwischen Inhalt und Interessen/Vorwissen ist entscheidend für den Erwerb von Lesekompetenz (Leppert & McElvany, 2021). Zusätzlich kann die Länge und Komplexität von Texten angepasst werden. So kann die Textschwierigkeit an das aktuelle Lesekompetenzniveau angepasst werden, um eine optimale Herausforderung zu bieten. Studien zeigen, dass schwierigkeitsangepasste Lesebücher die Lesekompetenz von Schüler*innen verbessern können (Seals, 2013). Texte, die mit einer Genauigkeit von 95 Prozent oder höher gelesen werden, fördern die Leseflüssigkeit signifikant, während Texte, die mit einer wesentlich niedrigeren oder höheren Genauigkeit gelesen werden, weniger positive Effekte zeigen (Allington, McCuiston &

Billen, 2015). Generative KI kann durch Personalisierung Texte und Aufgaben bereitstellen, die für eine Schülerschaft mit heterogenen Interessen und Lesekompetenzniveaus interessant sind und eine angemessene Schwierigkeit haben.

Obwohl die Personalisierung vielversprechende Potenziale bietet, zeigen Meta-Analysen zu Leseförderprogrammen für Kinder mit Schwierigkeiten, dass die entscheidenden Faktoren für den Erfolg einer Leseförderung die Dosis (d.h. die Dauer aktiver Lernzeit), explizite Inhalte (d.h. die Vermittlung von Lesestrategien) und ein systematischer Aufbau des Programms sind (Al Otaiba et al., 2022). Ein Beispiel für einen systematischen Aufbau sind ansteigende Aufgabenschwierigkeiten (Souvignier & Ehliert, 2024).

Daher wird in der vorliegenden Studie eine Webplattform erprobt, die es Schulklassen der vierten Klasse ermöglicht im Rahmen eines evidenzbasierten Leseförderprogramms unter Einbezug verschiedener expliziter Inhalte mit angemessener Dosis, personalisierte Geschichten zu generieren. Diese Geschichten können inhaltlich sowie in ihrer Länge und sprachlichen Schwierigkeit angepasst werden. Die personalisierten Texte sollen dabei als interessant und angemessen schwierig bewertet werden sowie Leseaufgaben mit einer hinreichenden Wahrscheinlichkeit richtig beantwortet werden können, ohne zu einfach zu sein.

Fragestellung

1. Textinteresse: Werden die Texte über den Verlauf des Leseförderprogramms als interessant bewertet?

Eine stetig hohe Bewertung ist wünschenswert.

2. Schwierigkeitsanpassung: Werden die Texte als angemessen schwierig bewertet?

Die Schwierigkeitsbewertung der Texte sollte im mittleren Bereich liegen und bei Anpassung der Textschwierigkeit durch Lehrkräfte entsprechend variieren.

3. Aufgabenerfolg: Erreichen die Schüler*innen angemessene Aufgabenlösungsraten?

Die Aufgaben sollten eine Lösungshäufigkeit aufweisen, die in einem hohen Bereich liegt.

4. Dosis: Gelingt es bei der Implementierung des Leseförderprogramms mit KI-generierten Geschichten, dass Schüler*innen eine ausreichende Menge an Texten lesen und bearbeiten?

Effekte einer Leseförderung sind ab einer Dauer von 20 Stunden aktiver Lesezeit mit hinreichend hoher Wahrscheinlichkeit nachweisbar (Roberts et al., 2022).

Methode

Datengrundlage ist eine Pilotierungsstudie als Vorstufe zu einem großangelegten randomisierten Kontrollgruppenversuch. Die Datenerhebung startet Ende August 2024. Zehn Lehrkräfte und etwa 200 Schüler*innen nehmen an der Pilotierung des Leseförderprogramms teil. Die Schüler*innen lesen verschiedene Texte, bewerten ihr Textinteresse und beantworten pro Text vier Fragen. Lehrkräfte werden in Gruppendiskussionen zur Nützlichkeit der generativen KI und Usability der Anwendung befragt. Quantitative Analysen untersuchen die präsentierten Fragestellungen.

Ergebnisse und ihre Bedeutung

In dem Vortrag werden die Ergebnisse der Pilotierungsstudie vorgestellt. Erste Ergebnisse zur Beteiligung, Bewertung des Textinteresses und der Textschwierigkeit sollen wertvolle Informationen zur Verbesserung KI-generierter Geschichten liefern. Die vorliegende Studie zielt zudem darauf ab, fundierte Daten zu sammeln, die zur Optimierung der Leseförderung beitragen. Durch die Untersuchung der Passungsdimensionen soll ein besseres Verständnis dafür entwickelt werden, wie personalisierte und auf die Fähigkeiten und Interessen der Schüler*innen abgestimmte Leseförderprogramme gestaltet werden können. Die Ergebnisse der Pilotierungsstudie dienen als Grundlage für die Weiterentwicklung und Implementierung großangelegter Fördermaßnahmen mit Evaluation im randomisierten Kontrollgruppendesign, die langfristig die Lesekompetenz und -motivation der Schüler*innen verbessern sollen.

How can AI improve L1, L2, and FL students' writing? Results from meta-analyses

Sina Verena Scherer¹, Steve Graham², Vera Busse¹

¹Universität Münster, Deutschland; ²Arizona State University, USA

In most educational institutions worldwide, the need to enhance writing support has grown for both first language (L1) writers and language learners due to migration movements. Most migrants must learn a new language upon school arrival, usually called a second language (L2). Additionally, globalization and work mobility demands have increased the necessity of learning foreign languages (FLs), particularly the lingua franca English. Providing timely feedback plays a vital role in improving students' writing. However, providing feedback to each student individually and frequently is hardly feasible in large classes. Algorithm-based feedback based on artificial intelligence (AI) can assist teachers in providing feedback and provide information on how to improve both surface- (e.g., grammar, spelling) and deep-level features (e.g., content, structure, and coherence) of texts. Previous meta-analytical evidence suggests medium effects of algorithm-based feedback (e.g., Fleckenstein et al., 2023; Mohsen, 2022; Zhai & Ma, 2021). However, usually, meta-analyses do not clearly distinguish between surface- and deep-level writing outcomes and between different kinds of learners, such as L1, L2, and FL learners, although they tend to have different learning experiences. L2 learners, such as first-generation migrants but also international students, can engage in more natural communication situations within the immediate environment (Ferris & Hedgcock, 2023) and are usually integrated into regular classrooms with different writing goals and feedback approaches from FL classes, which take place in students' home countries and typically emphasize language development. To clearly distinguish between different learner groups (L1, L2, and FL learners), we wanted to answer the overarching research question with our meta-analytical research to what extent algorithm-based feedback can foster L1, L2, and FL learners' writing on a surface- and deep-level. Therefore, we conducted two meta-analyses, one including algorithm-based feedback aside other feedback agents (instructors, peers, self) and one delving deeper into the effects of solely algorithm-based feedback on different learners' writing.

The search for experimental and quasi-experimental studies was conducted in four different databases and in five different languages (English, German, French, Italian, and Spanish). We followed meta-analytical guidelines and utilized the Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) (Page et al., 2021). Statistical analyses were conducted with the software Comprehensive Meta-analyses (CMA) (Borenstein et al., 2005).

In the first meta-analysis that included publications until 2022 (Authors, 2024), we calculated effect sizes for deep—and surface-level outcomes of algorithm-based feedback for different kinds of learners. However, none of the calculated effects proved significant; only the moderate effect of algorithm-based feedback on surface-level features for FL learners ($g = 0.53$) bordered on

significance. Results suggest that algorithm-based feedback may be beneficial to improve grammar and spelling in language students' writing but results need further investigation regarding deep-level features and different kinds of learners.

In order to account for the increasing number of recent publications on algorithm-based feedback, our new meta-analysis explores the impact of surface-level and deep-level AWE feedback on students' surface- and deep-level writing outcomes across 28 studies, 10 of which were published after 2022. We additionally examine how language proficiency, the duration of the intervention, the educational level, and the writing genre influence surface- and deep-level outcomes. Moreover, we distinguish between outcome measures that look at revision and those that focus on transfer to a new task.

Analyses of the second meta-analysis are still ongoing. First results suggest that in the majority of studies, AWE tools such as Pigai, Grammarly, ChatGPT, and PEG Write are used providing both feedback on surface- and deep-level features. The majority of new studies was conducted in a foreign language university setting. Further results are discussed against the backdrop of fostering literacy with digital technologies.

PS-50: Psychosoziale Herausforderungen und schulisches Wohlbefinden von Jugendlichen

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 157

Chair der Sitzung: Tim Hamer

Stolz, Angst und Hoffnungslosigkeit von Schüler:innen nach dem Übergang in die Sekundarschule – Wachstumskurvenmodelle zur Prüfung von individuellen und schulischen Prädiktoren

Simon Meyer¹, Juliane Schlesier², Michaela Gläser-Zikuda¹

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland; ²Carl von Ossietzky Universität Oldenburg

Theoretischer Hintergrund und Fragestellung

Im schulischen Kontext wirken sich Stolz, Angst und Hoffnungslosigkeit häufig auf das Lernen und die Leistung von Schüler:innen aus (Pekrun, 2016). Im Kontext des Übergangs von der Grundschule in die weiterführende Schule sehen sich Schüler:innen mit einer Reihe von spezifischen Veränderungen konfrontiert, wie einem neuen sozialen Umfeld, höheren Leistungsanforderungen und homogenen Leistungsgruppen. Der schulische Übergang wird daher oft als emotional herausfordernd erlebt und z. B. als „kritisches Lebensereignis“ (Filipp, 2001) bezeichnet.

Bislang liegen allerdings nur wenige Studien vor, die Lern- und Leistungsempfindungen von Schüler:innen nach dem Übergang in die Sekundarstufe untersuchen. Insbesondere systematische Längsschnittstudien fehlen, die potenzielle Einflussfaktoren auf Entwicklungsverläufe positiver und negativer Emotionen berücksichtigen (Meyer & Schlesier, 2021; Pekrun & Stephens, 2010). Einige Studien legen nahe, dass sich das Erleben von Lern- und Leistungsempfindungen in den frühen Sekundarschuljahren eher ungünstig entwickelt. Dabei zeigt sich, dass positive Emotionen im Verlauf der Sekundarschule abnehmen, während negative Emotionen zunehmen (vgl. Hagenauer & Hascher, 2010; Meyer & Schlesier, 2021; Pekrun et al., 2007; Vierhaus, Lohaus & Wild, 2016).

Die vorliegende Längsschnittstudie hat daher zum Ziel, den Einfluss lernrelevanter individueller Aspekte (z. B. schulbezogene Selbstwirksamkeitserwartung, Motivation) sowie kontextueller Determinanten auf Klassenebene (z. B. Unterrichtsqualität) auf die Entwicklung von Stolz, Angst und Hoffnungslosigkeit der Schüler:innen nach dem Übergang in die Sekundarstufe zu analysieren. Aufgrund früherer Querschnittsstudien lässt sich annehmen, dass höhere schulbezogene Selbstwirksamkeitserwartungen sowie eine höhere Motivation das Erleben positiver Emotionen positiv beeinflussen und zu einem geringeren Niveau an negativen Emotionen führen können (z.B. Meyer & Gläser-Zikuda, 2020; Pekrun, 2017). Des Weiteren wird postuliert, dass auch Aspekte der Unterrichtsqualität (z.B. eine hohe Klarheit, viel Struktur) bei Lernenden zu einer Zunahme positiver und einer Abnahme negativer Emotionen führt (z.B. Lei et al., 2018; Meyer & Gläser-Zikuda, 2020).

Methode

Die Längsschnittstudie wurde vom Beginn der fünften Klassenstufe bis zum Beginn der sechsten Klassenstufe (drei Messzeitpunkte) durchgeführt. Die Stichprobe umfasst $N = 584$ Schüler:innen (84.9 % weiblich), die im Durchschnitt $M = 10.19$ (SD = .45) Jahre alt sind. Sie verteilen sich zu 56.2 % auf Realschulen und zu 43.8 % auf Gymnasien.

Die Längsschnittdaten wurden mittels Wachstumskurvenmodellierung mit dem Programm R (Version 4.3.2) analysiert. Der Modellaufbau erfolgte dabei schrittweise (Basismodell, Random-Intercept-Modell, Random-Intercept-Modell mit Prädiktor ‚Zeit‘, Random-Intercept-Random-Slope-Modell). Ausgehend vom Random-Intercept-Random-Slope-Modell wurden schrittweise weitere Prädiktoren (schulbezogene Selbstwirksamkeitserwartungen, Motivation, Klassenführung, Klarheit und Struktur) aufgenommen.

Ergebnisse und ihre Bedeutung

Die Modelle für die einzelnen Emotionen zeigen, dass jeweils das Random-Intercept-Random-Slope-Modell mit allen Prädiktoren im Vergleich zu den anderen Modellen den besten Modellfit aufweist (*Stolz*: AIC = 3069.94, BIC = 3129.48, L.Ratio = 43.83, df = 11, $p < 0.001$; *Angst*: AIC = 2765.02, BIC = 2824.55, L.Ratio = 105.00, df = 11, $p < 0.001$; *Hoffnungslosigkeit*: AIC = 3174.54, BIC = 3234.07, L.Ratio = 79.36, df = 11, $p < 0.001$). Die Ergebnisse der Modellvergleiche legen nahe, dass insbesondere der Rückgang schulbezogener Selbstwirksamkeitserwartungen den Rückgang der positiven Emotion *Stolz* ($\beta = .52$ (.03), $t(1012) = 18.45$, $p < .001$) und der negativen Emotionen *Angst* ($\beta = -.43$ (.03), $t(1012) = -13.33$, $p < .001$) und *Hoffnungslosigkeit* ($\beta = -.44$ (.03), $t(1012) = -15.14$, $p < .001$) erklärt.

Auf Klassenebene zeigen sich darüber hinaus signifikante Effekte der Unterrichtsqualität auf die Entwicklung der untersuchten positiven und negativen Emotionen.

Die Resultate der Studie sind mit früheren Befunden konsistent. Die Entwicklung der untersuchten Emotionen wird sowohl durch individuelle Faktoren (v.a. schulbezogene Selbstwirksamkeitserwartungen), als auch durch kontextuelle Determinanten auf Klassenebene (v.a. Klassenführung) erklärt. Daher sollten emotionsbezogene Interventionen in Schulen insbesondere nach dem Übergang von der Primar- zur Sekundarstufe implementiert werden, um eine Steigerung positiver und eine Verringerung negativer Emotionen zu fördern. Im Vortrag werden weitere Implikationen sowie Limitationen der Studie erörtert.

Klimasorgen unter Schülerinnen und Schülern – folgenswer und förderlich?

Tim Hamer¹, Rebekka Schröder¹, Lars König^{1,2}

¹Stiftung Gesundheitswissen, Berlin, Deutschland; ²Charité – Universitätsmedizin Berlin, Institut für Medizinische Soziologie und Rehabilitationswissenschaft, Berlin, Deutschland

Theoretischer Hintergrund und Fragestellung

Der globale Klimawandel stellt ein ernstzunehmendes Risiko für die psychische Gesundheit dar (Walinski et al., 2023). Neben Extremwetterereignissen und damit verbundenen Folgen kann auch die Sorge um die zukünftigen Auswirkungen des Klimawandels die psychische Gesundheit beeinträchtigen (Cianconi et al., 2020).

Klimasorgen können durch Medienberichte über die Auswirkungen des Klimawandels ausgelöst werden (Crandon et al., 2022; Ogunbode et al., 2022). Durch die teilweise drastischen Veränderungen, welche der Klimawandel mit sich bringt, scheinen Klimasorgen als emotionale Reaktion nachvollziehbar zu sein (Hickman et al., 2021). Ferner existieren Befunde, nach denen Klimasorgen wie Wut oder Trauer in Bezug auf den Klimawandel mit umweltfreundlichem Verhalten in Zusammenhang stehen (Stanley et al., 2021; Whitmarsh et al., 2022).

Wenn Sorgen um den Klimawandel zu stark werden, können sie allerdings zu Beeinträchtigungen des alltäglichen Lebens führen (Clayton & Karazsia, 2020; Hepp et al., 2023; Hickman et al., 2021). Daher scheint es ratsam, zwischen negativen Emotionen (z.B. Wut, Angst, Trauer) und psychischen Beeinträchtigungen (z.B. Energielosigkeit, Schlafprobleme, Kopf- & Bauchschmerzen) aufgrund des Klimawandels zu unterscheiden (Hepp et al., 2023).

Kinder und Jugendliche sind hierbei als Risikogruppe für psychische Beeinträchtigungen besonders zu beachten (Hickman et al., 2021). Daher wurde in der vorliegenden Studie untersucht, wie negative Emotionen und psychische Beeinträchtigungen aufgrund des Klimawandels unter Jugendlichen verteilt sind. Des Weiteren wurde die subjektiv eingeschätzte Kompetenz im Umgang mit dem Klimawandel und die Motivation zu umweltbewussten Verhalten erfasst. Bei der Auswertung wurden verschiedene demographische Faktoren (z.B. Alter, Geschlecht, Gesundheitskompetenz) berücksichtigt. Die Ergebnisse sollen der Identifikation vulnerabler Gruppen dienen und, darauf aufbauend, der Entwicklung zielgruppenspezifischer Interventionsangebote.

Methode

Es wurde eine repräsentative Schülerbefragung ($N = 1021$) unter 12- bis 17-Jährigen mittels einer Fragebogenbatterie durchgeführt. Zur Messung der Klimasorgen wurde die Climate Change Distress and Impairment Scale (Hepp et al., 2023) eingesetzt, welche die Unterscheidung zwischen negativen Emotionen und psychischen Beeinträchtigungen aufgrund des Klimawandels ermöglicht. Gesundheitskompetenz wurde durch den Einsatz der Kurzform des European Health Literacy Questionnaires (Jordan & Hoebel, 2015) erhoben, der die Einteilung der Versuchspersonen in Personen mit ausreichender, inadäquater und problematischer Gesundheitskompetenz ermöglicht. Zudem wurden sowohl demographische Daten als auch weitere Konstrukte erhoben. Zur Analyse der Daten wurden verschiedene inferenzstatistische Analysen durchgeführt.

Ergebnisse und ihre Bedeutung

Negative Emotionen waren im Vergleich zu psychischen Beeinträchtigungen signifikant stärker ausgeprägt. Hierbei zeigte sich ein Effekt für Schulformen. So berichteten Gymnasiastinnen und Gymnasiasten über stärkere negative Emotionen, jedoch über weniger psychische Beeinträchtigungen im Vergleich zu Schülerinnen und Schülern anderer Sekundarstufen. Negative Emotionen korrelierten insgesamt signifikant positiv sowohl mit der eigenen Überzeugung zu wissen, was man gegen den Klimawandel tun kann, als auch mit der Bereitschaft sich selbst für den Klimaschutz einzuschränken. Es zeigte sich zudem ein Zusammenhang zwischen Gesundheitskompetenz und klimabedingten Beeinträchtigungen. Schülerinnen und Schüler mit ausreichender Gesundheitskompetenz berichteten über weniger psychische Beeinträchtigungen aufgrund von Klimasorgen als Personen mit problematischer bzw. inadäquater Gesundheitskompetenz.

Die vorliegenden Ergebnisse zeigen, dass Klimasorgen unterschiedlich stark in der Schülerschaft ausgeprägt sind. Negative Emotionen aufgrund des Klimawandels scheinen zudem mit einer eher adaptiven Kognition im Hinblick auf den Klimawandel zusammenzuhängen. Es lassen sich zudem Zusammenhänge zu anderen Faktoren wie der individuellen Gesundheitskompetenz und weiteren demographischen Eigenschaften erkennen. Demzufolge sollten in Präventionsprogrammen zum Thema Klimawandel und Gesundheit besonders vulnerable Personengruppen berücksichtigt sowie die Stärkung der Gesundheitskompetenz als Ansatz der Prävention von psychischen Beeinträchtigungen aufgrund von Klimasorgen getestet werden.

Es gilt zu beachten, dass aufgrund des Querschnittsdesigns keine kausalen Schlüsse gezogen werden können.

Überwindung von „Jingle-Jangle Fallacies“ in der Forschung zu Selbstmitgefühl: Ein systematisches Literaturreview und Konzeptualisierungsvorschlag

Ramona Schöne-Hoffmann¹, Sarah K. Schäfer², Christina David¹, Dorota Reis³, Manuela Benick¹, Franziska Perels¹

¹Department of Educational Sciences, Saarland University, Saarbrücken, Germany; ²Department of Clinical Psychology, Psychotherapy and Diagnostics - Child and Adolescent Psychology and Psychotherapy, Technische Universität Braunschweig, Braunschweig, Germany; ³Department of Psychology, Saarland University, Saarbrücken, Germany

Misserfolge sind ein unvermeidbarer Bestandteil unseres Lebens. Gerade im Bildungsbereich sind Fehler von besonderer Bedeutung, da sie Lernenden helfen, ihr Lernverhalten zu reflektieren und vorhandene Kompetenzen zielgerichteter zu nutzen bzw. zu erweitern (Spychiger et al., 2006). Diese positiven Effekte von Fehlern können jedoch nur erzielt werden, wenn Lernende nicht durch negative Emotionen belastet oder entmutigt werden. Denn wenn das Scheitern bei Lernenden so große Unzufriedenheit auslöst, dass sie sich gezwungen sehen, sich vor den starken negativen Emotionen zu schützen, so kommt es zu defensiven Selbstreaktionen wie Vermeidungsverhalten oder Prokrastination (Zimmerman & Kitsantas, 1996). Selbstmitgefühl spielt hier eine entscheidende Rolle, indem es hilft, negative Emotionen anzuerkennen, und Maßnahmen zu ergreifen, um negative Emotionen zu überwinden und das persönliche Wohlbefinden wiederherzustellen (Dodson & Heng, 2021). Das Konstrukt „Selbstmitgefühl“ wurde erstmals im Jahr 2003 von Kristin Neff definiert. Laut Neff (2003) besteht Selbstmitgefühl aus drei Bestandteilen: (1) Selbstfreundlichkeit, einer wohlwollenden und verständnisvollen Haltung gegenüber der eigenen Person, (2) Achtsamkeit, d.h. einer nicht-wertenden, distanzierten Wahrnehmung der eigenen Gedanken und Emotionen sowie (3) einem Verständnis gemeinsamer Menschlichkeit. Damit ist das Verstehen des Erlebens negativer, belastender Emotionen als fester Bestandteil der menschlichen Existenz gemeint.

Eine hohe Ausprägung von Selbstmitgefühl hängt nicht nur mit einem gesteigerten, subjektiven Wohlbefinden zusammen (Zessin et al., 2015), sondern auch mit einer verringerten Ängstlichkeit und weniger Stress (Marsh et al., 2017). Bezogen auf den Bildungsbereich zeigen sich positive Zusammenhänge mit Selbstwirksamkeit (Liao et al., 2021) sowie einer stärkeren Lernzielorientierung anstelle einer Leistungszielorientierung (Leary et al., 2007). Selbstmitfühlende Lernende neigen weiterhin weniger zu Prokrastination (Sirois, 2014) sowie akademischen Burnout-Symptomen (Kyeong, 2013).

Seit den Anfängen der psychologischen Forschung zu Selbstmitgefühl kamen diverse weitere allgemeine Konzeptualisierungen des Konstrukts hinzu (z.B. Dodson & Heng, 2021; Gilbert et al., 2017; Strauss et al., 2016). Die verschiedenen Konzeptualisierungen unterscheiden sich hinsichtlich ihrer Struktur und ihrer Bestandteile voneinander (Strauss et al., 2016). Eine einheitliche Definition allgemeinen Selbstmitgefühls fehlt bisher (Halamová et al., 2022).

Es wurde zudem zunächst angenommen, dass es sich bei Selbstmitgefühl um eine Art von Disposition handle, die über die Zeit und verschiedene Kontexte hinweg stabil ist (Zuroff et al., 2021). Neuere Befunde zeigen jedoch, dass es Unterschiede in den selbstmitfühlenden Reaktionen von Personen gibt, je nachdem welcher Lebensbereich untersucht wird (Zuroff et al., 2021). Dies haben einige Autoren zum Anlass genommen, bereichsspezifische Konzeptualisierungen von Selbstmitgefühl für soziale Beziehungen (Rose & Kocovski, 2020), den Arbeitskontext (Jennings et al., 2022;) und die Beziehung zum eigenen Körper (Altman et al., 2017) zu postulieren.

Insgesamt kann die Vielfalt an Konzeptualisierungen dieses Konstrukts zu Fehlschlüssen (sogenannten „Jingle-Jangle Fallacies“) führen: Verschiedene Begriffe werden zur Beschreibung desselben Phänomens verwendet bzw. der gleiche Begriff benennt unterschiedliche Phänomene (Porter, 2023).

Zielsetzung dieses Beitrages ist es daher, ein systematisches Review zur Konzeptualisierung von Selbstmitgefühl vorzustellen, das einen Überblick über bestehende Konzeptualisierungen allgemeinen und bereichsspezifischen Selbstmitgefühls gibt und die Bestandteile bisheriger allgemeiner Konzeptualisierungen in eine umfassende Definition und ein kohärentes Modell integriert. Hierfür werden im Anschluss an die systematische Sichtung bisheriger Konzeptualisierungen, mithilfe des Vorgehens der qualitativen Inhaltsanalyse nach Mayring (2022) deren Gemeinsamkeiten und Unterschiede sowie ihre zentralen theoretischen Hintergründe herausgearbeitet.

Ein weiteres Ziel dieser Arbeit ist es, die integrative, allgemeine Definition zu einer bereichsspezifischen Version ‚akademischen Selbstmitgefühls‘ weiterzuentwickeln, um das Konstrukt perspektivisch im Bildungsbereich systematisch untersuchen zu können.

Die Studie wird in Form eines Registered Reports durchgeführt. Der initiale Suchprozess in insgesamt dreizehn Datenbanken ergab $k = 5700$ Treffer. Diese werden aktuell einem Screening durch zwei unabhängige Beurteilende unterzogen, welche im Anschluss daran relevante Daten aus den identifizierten Zielstudien extrahieren. Die Ergebnisse der Analysen sollen im Rahmen des Kongresses vorgestellt werden.

Esstörungen und Gesundheitskompetenz bei Jugendlichen in Deutschland

Rebekka Schröder¹, Tim Hamer¹, Lars König^{1,2}

¹Stiftung Gesundheitswissen, Deutschland; ²Charité – Universitätsmedizin Berlin, Institut für Medizinische Soziologie und Rehabilitationswissenschaft, Berlin, Deutschland

Theoretischer Hintergrund und Fragestellung

Esstörungen wie Anorexia und Bulimia nervosa sind schwerwiegende psychische Erkrankungen, die mit substanziellen gesundheitlichen Folgen für die betroffenen Individuen einhergehen und in der Regel im Jugendalter zum ersten Mal auftreten (Solmi et al., 2022; Treasure et al., 2020). Neben der hohen Krankheitslast ist vor allem die erhöhte Mortalität von Patient:innen mit Essstörungen besorgniserregend (Arcelus et al., 2011; GBD 2019 Mental Disorders Collaborators, 2022). In Ergänzung zur effektiven Behandlung bereits bestehender Essstörungen ist die Prävention neuer Essstörungsfälle ein wichtiger Ansatz, um die Krankheitslast insgesamt zu reduzieren. Allerdings haben bisher evaluierte Präventionsprogramme für Essstörungen keine eindeutigen Ergebnisse bezüglich ihrer Effizienz geliefert (Koreshe et al., 2023; Le et al., 2017). Ein bisher in Präventionsprogrammen noch nicht berücksichtigter Faktor, der mit Essstörungen zusammenhängen könnte, ist die Gesundheitskompetenz der Betroffenen (Boberová & Husárová, 2021; Campanino et al., 2023). Gesundheitskompetenz bezieht sich auf die kognitiven und motivationalen Kompetenzen, die die Motivation und die Fähigkeit eines Individuums bestimmen, sich Zugang zu Informationen zu verschaffen, diese zu verstehen und anzuwenden, um die eigene Gesundheit erhalten und zu fördern (Nutbeam, 1998). Die vorgestellte Studie geht der Frage nach, inwiefern bei Jugendlichen Zusammenhänge zwischen möglichen Essstörungen und ihrer Gesundheitskompetenz sowie weiteren soziodemographischen Faktoren bestehen. Die Ergebnisse können darüber hinaus dazu genutzt werden, Risikofaktoren für Essstörungen zu identifizieren, um Präventionsprogramme weiter zu verbessern.

Methode

Die Datenerhebung fand in einer großen repräsentativen Stichprobe ($N = 1021$) von deutschsprachigen und in Deutschland lebenden Jugendlichen im Alter von 12-17 Jahren mithilfe von online-Fragebögen und face-to-face-Befragungen statt. Mögliche Fälle von Essstörungen wurden mit dem SCOFF-Fragebogen identifiziert (Morgan et al., 1999). Wenn die Versuchspersonen zwei oder mehr der fünf Fragebogen-Items mit „ja“ beantworteten, wurden sie als Fälle mit einer möglichen Essstörung klassifiziert. Zusätzlich wurde die Gesundheitskompetenz der Versuchspersonen mit dem HLS-EU-Q16 Fragebogen erfasst (Pelikan et al., 2014; Sørensen et al., 2013) und zwei Kategorien zugeordnet (inadäquate und problematische vs. adäquate Gesundheitskompetenz). Weitere Merkmale wie Körperbild, Geschlecht, Alter, Bildungsstand und Sozialstatus wurden im Selbstbericht erhoben. Assoziationen zwischen möglichen Fällen von Essstörungen und Gesundheitskompetenz, Körperbild und den soziodemographischen Variablen wurden mithilfe von χ^2 -Unabhängigkeitstests statistisch überprüft.

Ergebnisse und ihre Bedeutung

Mögliche Fälle von Essstörungen wurden bei über 19% der untersuchten Jugendlichen identifiziert. Dabei waren mögliche Essstörungsfälle häufiger bei Mädchen als bei Jungen und nahmen mit zunehmendem Alter zu. Bildungsstand und Sozialstatus hatte keinen signifikanten Einfluss auf die Raten von möglichen Essstörungen. Die Wahrnehmung des eigenen Körpers als ein bisschen oder viel zu dick oder zu dünn ging mit höheren Raten von möglichen Essstörungen einher. Mögliche Fälle von Essstörungen traten doppelt so häufig bei Jugendlichen mit inadäquater oder problematischer Gesundheitskompetenz als bei Jugendlichen mit adäquater Gesundheitskompetenz. Die Ergebnisse legen nahe, dass zukünftige Präventionsprogramme für Essstörungen neben bereits bekannten Risikofaktoren wie Geschlecht und Alter besonders auch die Gesundheitskompetenz der Zielgruppen adressieren sollten. Erste Ergebnisse aus Evaluationsstudien zeigen, dass sich die Gesundheitskompetenz durch gezielte Interventionen fördern lässt (Berkman et al., 2011; Visscher et al., 2018). Inwieweit sich dies auch in der Reduktion von Essstörungssymptomen oder -fällen widerspiegeln kann, sollte in zukünftigen Studien untersucht werden.

PS-51: Schulleitung und digitale Transformation: Impulse aus der Bildungsforschung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 169

Chair der Sitzung: Andreas Harder

Digitalisierungsbezogener Entwicklungsstand beruflicher Schulen in der Schweiz – Bedeutung des Schulleitungshandelns und struktureller Faktoren

Andreas Harder¹, Stephan Schumann¹, Serge Imboden²

¹Universität Konstanz, Deutschland; ²HES-SO Valais-Wallis, Schweiz

Theoretischer Hintergrund

Verschiedene theoretische Ansätze verdeutlichen, dass es sich bei der digitalen Transformation von Schulen um eine Innovation im Sinne der Schulentwicklung handelt (u.a. Ilomäki & Lakkala, 2018; Eickelmann & Gerick, 2017). Petko et al. (2018) unterscheiden hierbei zwischen schulbezogenen (*school readiness*) und lehrpersonenbezogenen Merkmalen (*teacher readiness*), die für die Integration von digitalen Medien in Lehr- und Lernprozesse entscheidend sind. Unter *school readiness* fallen beispielsweise Aspekte wie die technische Infrastruktur und Ausstattung oder die digitalisierungsbezogenen Ziele der Schule. *Teacher readiness* umfasst die digitalen Kompetenzen und Einstellungen der Lehrkräfte. Wenn alle schulischen Akteur*innen zur erfolgreichen Gestaltung digitaler Schulentwicklungsprozesse beitragen sollten, wird der Schulleitung eine zentrale Rolle bei der Initiierung und Begleitung dieser Prozesse zugesprochen (u.a. Gerick et al., 2023; Tulowitzki & Gerick, 2022; Tulowitzki et al., 2021; Waffner, 2021).

In der Schweiz zeigen sich hinsichtlich des digitalisierungsbezogenen Entwicklungsstands von Schulen sprachregionale Unterschiede; Schulen in der Deutschschweiz sind digital tendenziell besser aufgestellt als Schulen in der französisch- und italienischsprachigen Schweiz (u.a. Educa, 2021; Oggenfuss & Wolter, 2021; Suter et al., 2023). Ob diese Erkenntnisse aus den allgemeinbildenden Schulen auch auf die berufliche Bildung übertragbar sind, ist jedoch unklar.

Fragestellungen

Vor diesem Hintergrund untersucht dieser Beitrag den digitalisierungsbezogenen Entwicklungsstand beruflicher Schulen in der Schweiz und analysiert, welche Bedeutung das Schulleitungshandeln sowie strukturelle Merkmale hierbei haben. Somit können zwei zentrale Fragestellungen abgeleitet werden:

- (1) Wie ist der digitalisierungsbezogene Entwicklungsstand beruflicher Schulen in der Schweiz ausgeprägt?
- (2) In welchem Zusammenhang steht dieser Entwicklungsstand mit dem Schulleitungshandeln und strukturellen Merkmalen der Schule?

Methode

Zur Beantwortung der Fragestellungen werden Daten von n = 320 Schulleitungsmitgliedern beruflicher Schulen aus der Schweiz herangezogen, die im Februar und März 2023 erhoben wurden. Der digitalisierungsbezogene Entwicklungsstand der Schulen wurde – basierend auf dem Modell von Petko et al. (2018) – mit acht Skalen erfasst. Das Schulleitungshandeln wurde anhand von drei Skalen gemessen (Diel & Steffens, 2010): *Innovative Grundhaltung* ($\alpha = .89$), *Innovative Prozesse anstoßen und gestalten* ($\alpha = .88$) sowie *Strukturen legen für Entwicklungsprozesse* ($\alpha = .87$). Die sprachregionale Zugehörigkeit sowie die Bewertung der zur Verfügung stehenden finanziellen Ressourcen wurden manifest erfasst.

Zur Untersuchung des digitalisierungsbezogenen Entwicklungsstands und möglicher Cluster wurde mit der Software Mplus 8.7 zunächst eine latente Profilanalyse (LPA) durchgeführt. Darauf aufbauend wurden mithilfe von ANOVAs und Chi-Quadrat-Tests potenzielle Zusammenhänge bzw. Unterschiede zwischen dem digitalisierungsbezogenen Entwicklungsstand, dem Schulleitungshandeln und den strukturellen Merkmalen untersucht.

Ergebnisse und ihre Bedeutung

Die Befunde zeigen, dass die befragten Schulleitungsmitglieder drei Schulprofilen zugeordnet werden können: *digital fortschrittlichen Schulen* (n = 61), *digital durchschnittlichen Schulen* (n = 228) und *Schulen mit digitalem Entwicklungspotenzial* (n = 31). Die Profile wurden sowohl anhand inhaltlicher Überlegungen sowie unter Einbezug statistischer Kennzahlen (AIC, BIC, Entropy) abgeleitet. Es zeigt sich, dass alle drei Skalen des Schulleitungshandelns in den übergeordneten Profilen positiver ausgeprägt sind als in den untergeordneten ($p < .001$; $\eta^2 < .244$). Im Hinblick auf die strukturellen Faktoren geht aus den Analysen hervor, dass zwischen den Profilen und der Sprachregion ein Zusammenhang existiert: So befinden sich *digital fortschrittliche Schulen* signifikant häufiger in der Deutschschweiz als in der französisch- und italienischsprachigen Schweiz ($p < .001$; Cramer's $V = .267$). Zudem werden die finanziellen Ressourcen in *digital fortschrittlichen Schulen* und *digital durchschnittlichen Schulen* besser bewertet als in *Schulen mit digitalem Entwicklungspotenzial* ($p < .001$; $\eta^2 = .092$).

Die Befunde deuten darauf hin, dass sowohl das Schulleitungshandeln als auch strukturelle Faktoren wie beispielsweise die finanziellen Ressourcen den digitalisierungsbezogenen Entwicklungsstand von Schulen beeinflussen. Die Ergebnisse legen zudem nahe, dass die Befundlage aus dem allgemeinbildenden Bereich auf berufsbildende Schulen übertragen werden kann. Die Bildungspolitik könnte die Erkenntnisse als Anlass nehmen, sprachregionale Förderangebote zu entwickeln und *Schulen mit digitalem Entwicklungspotenzial* gezielt zu unterstützen.

Zur Bedeutung von Absorptive Capacity für ICT-Wissensintegration in Schulen

Sandra Fischer-Schöneborn¹, Marcus Pietsch¹, Chris Brown², Burak Aydin^{1,3}, Stephen MacGregor⁴

¹Leuphana Universität Lüneburg, Deutschland; ²University of Southampton, UK; ³Ege Üniversitesi, Izmir, Turkey; ⁴University of Calgary, Canada

Theoretischer Hintergrund

Technologieintegration und -nutzung ist für die Bildung im 21. Jahrhundert von entscheidender Bedeutung. Unterricht kann durch den Einsatz von Technologien verbessert oder verändert werden (Ventista et al., 2024). Wie Autor*innen (2023, 2024) feststellten, benötigen Schulen für die Umsetzung von Innovationen externes Wissen und sektorübergreifende Zusammenarbeit, insbesondere wenn es sich um komplexe Innovationen handelt, die Wissen erfordern, das nicht in ausreichendem Maße innerhalb der Schule oder zwischen Lehrkräften generiert werden kann. Dies zeigt sich besonders bei technologischen Innovationen.

Absorptive Capacity (ACAP) wird definiert als organisatorische Routinen und Prozesse, durch die Organisationen Wissen erwerben, assimilieren, transformieren und nutzen, um dynamische Fähigkeiten zu erzeugen (Zahra & George, 2002). Zahra & George (2002) weisen dabei auf verschiedene Kontingenzfaktoren hin, die ACAP beeinflussen: *activation triggers*, *social integration mechanisms* und *regimes of appropriability*. Während ACAP als organisationales Lernen und Innovationsmotor bisher vor allem im Kontext von Managementstudien untersucht wurde (Zahra & George, 2002; Cohen & Levinthal, 1990; Sancho-Zamora et al., 2021; Lane et al., 2001), sind Studien zu Effekten in Bildungskontexten rar (z.B. Crain-Dorough & Elder, 2021).

Ziele und Methode

Ziel dieser Studie ist es, zur internationalen Debatte über ACAP als kritischer Faktor für organisationales Lernen von Schulen und für die Integration von Wissen in Schulen zu *Information and Communication Technology* (ICT) beizutragen, indem folgende Forschungsfragen (FF) beantwortet werden:

- 1) Gibt es einen positiven Zusammenhang zwischen ICT-ACAP und der Technologieintegration in Schulen?
- 2) Beeinflusst ICT-ACAP den Zusammenhang zwischen externem Wissen und Technologieintegration in Schulen?
- 3) Hängt das Ausmaß von ICT-ACAP von Kontingenzfaktoren ab?

Zur Beantwortung dieser Fragen wurden Daten aus dem Schulleitungsmonitor von 2021, durchgeführt vom forsa Institut für Sozialforschung und statistische Analysen, ausgewertet (n=411 Schulleitungen; bereinigter Datensatz n=247; Rücklaufquote 63.3%). In einem ersten Schritt wurde die Angemessenheit des Modells und die Beziehung zwischen externem Wissen, ACAP und Technologieintegration in Schulen mit Hilfe eines Strukturgleichungsmodells (SEM) in Mplus 8.6 (Muthén & Muthén, 2021) untersucht. Im Anschluss wurde die Rolle der Kontingenzfaktoren in der Beziehung zwischen externem Wissen, ACAP und Technologieintegration in Schulen unter Anwendung von Klassifikations- und Regressionsbäumen (CART) (Breiman et al., 1984) als machine learning in R analysiert.

Ergebnisse und Implikationen

Die Ergebnisse deuten darauf hin, dass sich ICT-ACAP positiv auf die Technologieintegration in Schulen auswirkt (FF1) und dass ICT-ACAP als beeinflussender Faktor zwischen externem Wissen (modelliert über die Variablen *Open Innovation Depth* und *Open Innovation Breadth*) und Technologieintegration wirkt (FF2). Ein höherer Wert für ACAP war mit einem höheren Wert für Technologieintegration verbunden ($r=0,139$, $p<.01$). Die SEM-Ergebnisse zeigten einen statistisch signifikanten Zusammenhang zwischen *Open Innovation Depth* und ICT-ACAP ($\beta=.382$, $SE=.093$, $p<.001$) sowie zwischen ICT-ACAP und Technologieintegration in Schulen ($\beta=.152$, $SE=.055$, $p<.01$), während die direkten Pfade sowohl von *Open Innovation Depth* ($\beta=.014$, $SE=.091$, $p>.05$) als auch *Open Innovation Breadth* ($\beta=-.017$, $SE=.089$, $p>.05$) auf die Technologieintegration in Schulen keine Zusammenhänge zeigten.

Die Ergebnisse zu Forschungsfrage 3 deuten darauf hin, dass die Auswirkungen von ACAP auf die Technologieintegration vom Vorhandensein und der Wirksamkeit von Mechanismen für den Wissensaustausch innerhalb der Schulen abhängen sowie von dem Ausmaß, in dem sich Schulen an der Zusammenarbeit mit Konkurrenten beteiligen. CART zeigte beispielsweise, dass wenn ICT-ACAP tendenziell hoch war ($ACAP \geq 14$) und die Heterogenität der ICT-Fähigkeiten der Lehrkräfte nicht als extreme Herausforderung angesehen wurde ($CHAL < 3,5$), die Kombination aus hohem Wissensaustausch und hoher ICT-ACAP gemeinsam zu einer höheren Technologieintegration in Schulen führte.

Die Ergebnisse der Studie deuten auf ICT-ACAP als relevantes Konzept in der empirischen Bildungsforschung hin mit Desideraten für weitere Forschung. Die Ergebnisse geben auch praktische Hinweise für die Einführung oder Verbesserung der Technologieintegration in Schulen. So kann ICT-ACAP durch Gründung oder Intensivierung interner und externer Netzwerke verbessert werden, bspw. durch Kooperationen mit anderen (öffentlichen) Organisationen.

Die Position der Schulleitung im sozialen Netzwerk und ihr Zusammenhang mit dem Konsens im Kollegium über die digitale Schulentwicklung

Amelie Sprenger, Nina Werth, Cornelia Gräsel, Kathrin Fussangel

Bergische Universität Wuppertal, Deutschland

Die Einigung aller Bundesländer auf die Umsetzung der KMK Strategie „Bildung in der digitalen Welt“ (2016) fordert offiziell eine uneingeschränkte Priorisierung des Themas *digitale Schulentwicklung*. Die Forschung zur Educational Governance hebt die Bedeutung der Schulleitung und deren Kooperationen für den Erfolg digitaler Schulentwicklung hervor (Dexter, 2018; Gerick et al., 2016). Zudem wird die Relevanz einer klaren und transparenten Vision für das Gelingen digitaler Schulentwicklung in empirischen Studien zunehmend betont (Brauckmann, 2014; Tulowski & Gerick, 2020). Diese zwei wesentlichen Kernaspekte werden in dieser Studie tiefergehend betrachtet. Anders als in vorangegangenen Studien, die bereits einen der beiden Aspekte näher untersucht haben (vgl. Waffner, 2021; Eickelmann, & Gerick, 2017), erweitern wir den aktuellen Forschungsstand durch die Verknüpfung beider Elemente.

Schulleitungen spielen eine entscheidende Rolle bei der digitalen Transformation von Schulen (DuFour & Mattos, 2013). Dabei sind ihre Kooperationen mit anderen wesentlichen Akteuren der digitalen Schulentwicklung – sogenannten Promotor*innen – ebenso bedeutsam wie die Schulleitung selbst (Wagner & Gerholz, 2022; Huber, 2013). Als Promotor*innen lassen sich Personen bezeichnen, die durch innovationsfördernde Aktivitäten den Veränderungsprozess unterstützen. Trotz der anerkannten Bedeutung von Schulleitungen und Promotor*innen in der digitalen Schulentwicklungsforschung gibt es wenige Studien, die deren Kooperationen detaillierter untersuchen. Die vorliegende Studie analysiert die Netzwerkposition der Schulleitung sowie die Häufigkeit und Muster ihrer Kooperation mit Promotor*innen. Das „Innovative Digital School Model“ von Ilomäki & Lakkala (2018) nimmt einen Zusammenhang zwischen den Kooperationen der Schulleitung in ihrem Schulnetzwerk und dem Konsens über die Vision digitaler Schulentwicklung an. Obwohl der Zusammenhang zwischen dem Konsens über die Vision und den Netzwerkaktivitäten der Schulleitung theoretisch angenommen wird, ist er bislang kaum empirisch erforscht worden. In der vorliegenden Studie soll untersucht werden, inwiefern ein solcher Zusammenhang festzustellen ist.

Die zugrundeliegende Forschungsfrage lautet: Wie beeinflusst die Position der Schulleitung unter Berücksichtigung des Faktors „Zentralität“ und ihre Kooperationsstrukturen innerhalb des schulischen Netzwerks den Konsens über die Ziele der digitalen Schulentwicklung? Es lassen sich zwei Hypothesen formulieren: (H1) Wenn Schulleitungen eine zentrale Position im schulischen Netzwerk innehaben, ist der Konsens über die Vision digitaler Schulentwicklung im Kollegium höher. (H2) Wenn Schulleitungen eng mit anderen Promotoren vernetzt sind, ist der Konsens über die Vision digitaler Schulentwicklung im Kollegium höher. Datengrundlage ist eine Stichprobe mit ca. 500 Lehrkräften von 24 Gymnasien und Gesamtschulen aus Nordrhein-Westfalen, Bayern und Hamburg. Die Daten werden mithilfe von Fragebögen und Interviews erhoben. Die Auswertung erfolgt mit Hilfe einer sozialen Netzwerkanalyse, welche das soziale Netzwerk der einzelnen Kollegien visuell abbildet. Dies

ermöglicht die Identifikation der Zentralität der Schulleitung) und eine Analyse ihrer Kooperationsstrukturen. So kann ein tieferes Verständnis der Einflussfaktoren auf die digitale Schulentwicklung gewonnen werden.

Die Ergebnisse aus ersten vorliegenden Netzwerkanalysen mit einer Teilstichprobe belegen die Hypothesen. Sie lassen darauf schließen, dass Schulleitungen, die eine zentrale Position im schulischen Netzwerk einnehmen (H1) und solche, die gut mit anderen Promotor*innen vernetzt sind (H2), einen höheren Konsens über die Vision der digitalen Schulentwicklung im Kollegium erzielen können. Dies deutet aus der Perspektive der Distributed Leadership darauf hin, dass die soziale Einbindung der Schulleitung und deren Kooperationsstrukturen entscheidende Parameter für den Erfolg der digitalen Schulentwicklung sind. Die Ergebnisse sollen Aufschluss darüber geben, wie die Zentralität und die Kooperationsstrukturen der Schulleitung den Konsens der Vision und damit den Erfolg der digitalen Schulentwicklung beeinflussen können.

Wir diskutieren dies als Implikationen für die Gestaltung digitaler Schulentwicklung, sowie Gelingensbedingungen in Form von Strategien und Handlungsansätzen, die Schulleitungen aufgreifen können, sollten sie die digitale Schulentwicklung ihrer Schule vorantreiben wollen.

Setting the Direction for Digital Transformation – Systematisches Review zu individuellen und verhaltensbezogenen Merkmalen effektiver Schulleitungen

Florian Sievert, Katharina Scheiter
Universität Potsdam, Deutschland

Theoretischer Hintergrund

Dass Schulleitungen einen bedeutenden Einfluss auf die Entwicklungen an ihrer Schule, das Schulklima, die Arbeitsbedingungen des Personals und die Qualität der Lehre haben, wurde mehrfach nachgewiesen (Fullan, 2002; Huber, 2009; Navaridas-Nalda et al., 2020; Kareem et al., 2023). Die komplexen Anforderungen der digitalen Transformation (Shyr, 2017; Håkansson et al., 2019; Tulowitzki et al., 2020; McCarthy et al., 2023), erfordern ein tieferes Verständnis über Faktoren, die in diesem Kontext zu einer erfolgreichen Leitungspraxis führen. Auch wenn das Forschungsinteresse am Schulleitungshandeln insgesamt zunimmt (Leithwood & Riehl, 2003; Fullan & Kirtman, 2019; Day et al., 2020) und bereits Zusammenhänge zwischen den Merkmalen einer Schulleitungen und deren Outcome festgestellt wurden (Tulowitzki & Gerick, 2020; Navaridas-Nalda et al., 2020; Ruloff & Petko 2021), mangelt es an empirischen Studien, die personenbezogene Merkmale systematisch erfassen (Daniels et al., 2019; Khan, 2021; Waffner, 2021; Johannmeyer & Cramer, 2023). Die Lehrkräfte- und allgemeine Führungsforschung ist hier weiter (Costa & McCrea, 1992; Baumert & Kunter, 2006; Baumert et al., 2009; Kauffeld, 2019).

Vorgestellt werden Befunde einer laufenden systematischen Literaturrecherche zu individuellen und verhaltensbezogenen Merkmalen erfolgreicher Schulleitungen. Die Synthese der Befunde wird in einem Modell zusammengefasst, das Schlüsselfaktoren professioneller und effektiver Schulführung im Kontext der digitalen Transformation darstellt.

Forschungsfrage

Welche individuellen und verhaltensbezogenen Merkmale tragen zu erfolgreichem Schulleitungshandeln bei – im Allgemeinen sowie im speziellen Kontext der digitalen Transformation?

Methode

Die Literaturrecherche erfolgt anhand der PRISMA-Richtlinien (Moher et al., 2015) und den Richtlinien nach Newman und Gough (2020). Folgende Fragen werden untersucht: (1) *What individual characteristics are associated with successful school leadership in general?* (2) *What behavioral characteristics are associated with successful school leadership in general?* (3) *How does digital transformation affect the relationship between individual and behavioral characteristics and successful school leadership?*

Wir nutzen die Datenbanken *Web of Science*, *SCOPUS*, *EBSCO*, *ERIC*, *Fachportal Pädagogik* sowie *Google Scholar*. Berücksichtigt werden empirische, englisch- oder deutschsprachige Studien von 2000 bis 2025 mit folgenden Kriterien: qualitatives, quantitatives oder mixed method Studiendesign; Längs- oder Querschnittsstudie; experimentelle, quasi-experimentelle oder Korrelationsstudie; Peer-reviewed Journals sowie Buchkapitel, Dissertationen und Konferenzberichte.

Der Fokus liegt auf Studien mit Schulleitungen von Primar- und Sekundarschulen und die Forschungsfragen beantworten.

Ergebnisse

Bisherige Befunde zeigen, dass Schulleitungen, die die Herausforderungen der digitalen Transformation erfolgreich bewältigen, selbst über Kompetenzen im Sinne einer Digital Literacy verfügen (Mohamad Said et al., 2015; Tulowitzki et al., 2021; Marmoah et al., 2023). Sie sind zudem flexibel, offen und optimistisch gegenüber neuen Entwicklungen (Jacobsen, 2005; Zadok et al., 2024) und nehmen Trends der digitalen Transformation (wie künstliche Intelligenz und Metaverse) frühzeitig wahr und ernst (Hao et al., 2021; Dagli et al., 2023). Besonders motiviert und professionell agieren Schulleitungen mit einer guten Ausbildung und umfangreichen Vorkenntnissen in diesem Kontext (Mohamad Said et al., 2015). Sie praktizieren zudem einen eher transformationalen Führungsstil (Berkovich & Hassan, 2023; Schmitz et al., 2023), setzen sich für die Professionalisierung ihrer Lehrer:innen ein und etablieren kooperative Arbeitsformate (Polizzi, 2011; Wulandari & Arifin, 2020). Insgesamt befassen sich die meisten Schulleitungsstudien noch mit Aspekten einer erfolgreichen Technologisierung von Schule, einige wenige untersuchen das Schulleitungsverhalten in sozialen Netzwerken (Bunnel, 2018; Prasajo & Yuliana, 2021), kaum untersucht werden bisher sozio-kulturelle Auswirkungen der digitalen Transformation auf das Schulleitungshandeln (Krein, 2023). Zudem mangelt es, verglichen mit Studien der allgemeinen Führungsforschung, an Untersuchungen zur Selbst-Regulation und Selbst-Entwicklung in diesem Kontext.

Schulleitungen müssen als Promotoren der digitalen Transformation umfassende Bewältigungsstrategien entwickeln. Die Ergebnisse des Reviews sollen dazu beitragen, ein tieferes Verständnis zum professionellen Leitungshandeln zu erhalten. Die identifizierten Merkmale tragen zur Weiterentwicklung digitalisierungsbezogener Trainingsprogramme bei, ermöglichen Aussagen über das individuelle Potenzial zur Bewältigung digitalisierungsbezogener Herausforderungen und unterstützen eine Schulleitung darin, die Richtung für die digitale Transformation ihrer Schule vorzugeben.

S-54: Demokratie lernen: Schulische Einflüsse auf die politische Entwicklung junger Menschen

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 162

Chair der Sitzung: Katharina Eckstein

Chair der Sitzung: Miriam Schwarzenthal

Chair der Sitzung: Peter Noack

Demokratie lernen: Schulische Einflüsse auf die politische Entwicklung junger Menschen

Chair(s): **Katharina Eckstein** (Friedrich-Schiller-Universität Jena, Deutschland), **Miriam Schwarzenthal** (Bergische Universität Wuppertal)

Diskutant*in(nen): **Peter Noack** (Friedrich-Schiller-Universität Jena)

Junge Menschen verbringen an kaum einem Ort so viel Zeit wie in der Schule, zu deren Bildungsauftrag es auch gehört, offene, reflektierte und aktive Bürger:innen zu erziehen (z.B. SWK, 2024). Schule beeinflusst Lernen und Entwicklung auf unterschiedliche Art und Weise, wobei vor allem formal-curriculare von informellen Lernerfahrungen unterschieden werden. Empirische Befunde haben wiederholt gezeigt, dass neben der curricularen Vermittlung von gesellschaftspolitischen Inhalten und demokratischen Prinzipien, vor allem strukturelle und klimatische Merkmale des Schulkontextes einen wichtigen Einflussfaktor bilden (Neundorff & Smets, 2017). Im Einklang mit Sozialisations- und sozialen Lerntheorien (z.B. Bandura, 1986), kann Schule als Mikrokosmos betrachtet werden, der das Lernen über demokratische Prozesse im Kleinen ermöglicht. So können Erfahrungen, die auf den ersten Blick nicht politisch erscheinen, wie das Verteidigen der eigenen Ansichten gegenüber der Klasse oder das Aushandeln eines Kompromisses, das Denken über und die Auseinandersetzung mit politischen/demokratischen Prozessen beeinflussen. Des Weiteren spiegelt Schule auch das breitere soziale Umfeld, in dem junge Menschen aufwachsen, wider (z.B. ihre Nachbarschaft oder ihren Stadtteil) und ermöglicht es dadurch, gesellschaftliche Prozesse im Kleinen zu erfahren. In dieser 'Miniaturgesellschaft' der Schule sammeln junge Menschen somit unter anderem Erfahrungen im Umgang mit Diversität und Offenheit oder gegebenenfalls auch Diskriminierung und Intoleranz (Eckstein et al., 2024).

Dieses Symposium widmet sich der Frage, inwiefern schulische Erfahrungen die politische Entwicklung junger Menschen beeinflussen. Ein besonderer Schwerpunkt liegt dabei auf schulischen Faktoren, die außerhalb des formalen Fachunterrichts liegen, wie schulische Mitbestimmungsmöglichkeiten (Studie 1), Sozialbeziehungen (Studie 2) und das wahrgenommene Klassenklima (Studie 3). Das Symposium betrachtet neben dem kritischen Bewusstsein junger Menschen für soziale Ungleichheiten (Studie 2 & 3), die Entwicklung demokratischer Werte (Studie 1) und politischer Selbstwirksamkeitserfahrungen (Studie 4). Dabei berücksichtigt es einen breiten Altersrange, der Schüler:innen der Grundschule (Studie 1) bis hin zur Sekundarstufe umfasst (Studie 2-4). Darüber hinaus werden vielfältige methodische Ansätze präsentiert, die auf mehrbenenanalytische Auswertungen zurückgreifen (Studie 1 & 2), Fallstudien betrachten (Studie 3) und meta-analytische Auswertungen heranziehen (Studie 4).

Die Beiträge des Symposiums werden abschließend vor dem Hintergrund der aktuellen Befundlage sowie ihrer Implikationen für Forschung und Praxis diskutiert.

Beiträge des Symposiums

Der Einfluss schulischer Partizipationsmöglichkeiten auf die (demokratischen) Werte von Grundschüler*innen

Florian Monstadt

Bergische Universität Wuppertal

Theoretischer Hintergrund

Die Stabilität von Demokratien hängt maßgeblich von der Entwicklung demokratischer Werte ihrer Bürger*innen ab (Easton, 1975). Die Kindheit hat sich dabei als entscheidende Phase in der politischen Sozialisation herausgestellt (Rippl, 2015). Auch wenn demokratische Werte und Einstellungen bereits bei Kindern beobachtet wurden, gibt es nach wie vor ein Forschungsdesiderat, insbesondere hinsichtlich ihrer Internalisierung und möglicher Einflussfaktoren (Abendschön, 2010). Als eine zentrale Methode zur Förderung von Werten haben sich Mitbestimmungsmöglichkeiten für Kinder und Jugendliche etabliert (Himmelmann, 2017). Partizipation in der Schule/Klasse hat demnach einen positiven Effekt auf bspw. die politische Selbstwirksamkeit und Kompetenz (Jungkunz, 2023). Gerade für Kinder in der Grundschule wurde dieser Zusammenhang allerdings bisher kaum untersucht.

Fragestellung

1. Inwiefern begünstigt die Mitbestimmung von Schüler*innen innerhalb der eignen Klasse und darüber hinaus die Internalisierung demokratischer Werte? Bisherige Forschung deutet auf einen positiven Effekt von Partizipation auf die Entwicklung von demokratischen Werten hin.
2. Welche Rolle spielen Grundwerte hinsichtlich dieses Zusammenhangs? Bisherige Forschung zu Grundwerten konnte starke Korrelationen zu politischen Werten und Einstellungen feststellen, die ggf. auch für diesen Zusammenhang relevant sein könnten (Schwartz et al., 2013).

Methoden

Um sich der Beantwortung dieser Fragen zu nähern, wurden im Rahmen einer standardisierten Befragung 429 Grundschüler*innen in der 3. und 4. Klasse sowohl zu ihren Grundwerten, ihren demokratischen Werten als auch ihren wahrgenommenen Partizipationsmöglichkeiten innerhalb der Klasse befragt. Die Grundwerte wurden mittels Picture-Based Value Survey for Children (PBVS-C; Döring et al., 2010) erfasst. Für die Erhebung demokratischer Werte wurde u.a. auf Messinstrumente aus der Demokratie-Leben-Lernen Studie von van Deth und Kolleg*innen (2007) zurückgegriffen. Die wahrgenommenen Partizipationsmöglichkeiten wurden mit einer Skala von Gerbeshi & Ertl (2023) erfasst, die sowohl die Ebenen Unterricht und Klasse umfasst, als auch verschiedene Formen von Mitbestimmung. Zusätzlich wurden von 20 Lehrkräften die Grundwerte, demokratischen Einstellungen und ihre Perspektiven auf die Partizipationsmöglichkeiten der Kinder erhoben. Ausgewertet wurden die gesammelten Daten mit linearen Mehrebenenmodellen, wobei die Schüler*innen die erste Ebene bilden und die Klassen inkl. der Angaben der Lehrkraft die Zweite. So kann im Detail untersucht werden, welchen Effekt wahrgenommene und tatsächliche Partizipationsmöglichkeiten haben.

Ergebnisse und Diskussion

Die Analysen der vorliegenden Daten bestätigen, dass Kinder über demokratische Wertvorstellungen verfügen und dass diese eng mit den Grundwerten zusammenhängen. Darüber hinaus können die bisher gefundenen positiven Effekte von Mitbestimmungsmöglichkeiten in der Schule bestätigt werden. Die Daten zeigen, dass die wahrgenommenen Partizipationsmöglichkeiten einen positiven und höchst signifikanten Effekt darauf haben, in welchem Ausmaß die Grundschüler*innen demokratische Wertorientierungen äußern. Dieser Effekt ist äußerst stabil auch unter Berücksichtigung weiterer Kontrollvariablen und Prädiktoren. Kinder, die ihre Partizipationsmöglichkeiten im Unterricht und darüber hinaus als hoch einschätzen, vertreten stärkere demokratische Werte als Kinder, die dies nicht tun. Dabei scheint die Wahrnehmung der Schüler*innen zu ihren Mitbestimmungsmöglichkeiten gegenüber den Angaben der Lehrkräfte die entscheidendere Rolle zu spielen. Letztere haben keine signifikanten Effekte auf die demokratischen Werte der Schüler*innen.

Die Ergebnisse der Studie legen nahe, dass die Förderung von Partizipationsmöglichkeiten im Unterricht sowie die Betonung von Grundwerten wie Universalismus und Benevolenz wesentlich zur Entwicklung und Festigung demokratischer Einstellungen bei Kindern beitragen können.

Schulische Bedingungen von Diskriminierungswahrnehmungen im internationalen Vergleich

Johanna Ziemes
Universität Duisburg-Essen

Theoretischer Hintergrund

Diskriminierungswahrnehmungen, also das Urteil, dass eine oder mehrere Gruppen in der Gesellschaft ungerecht behandelt werden, lassen sich als Demokratiedefizit (Norris, 2011) interpretieren. Um sich gegen Diskriminierung einzusetzen, muss diese zunächst wahrgenommen werden. International wird bereits untersucht, wie Schüler*innen dazu befähigt werden können soziale Ungleichheiten wahrzunehmen und sich gegen diese einzusetzen (Diemer et al., 2019; Schwarzenenthal et al., 2024). Weitere Analysen konnten zeigen, dass ein offenes Unterrichtsklima und positive schulische Sozialbeziehungen mit der Unterstützung gleicher Rechte verschiedener marginalisierter Gruppen zusammenhängen (Ziemes, 2024). Bisherige Analysen von Diskriminierungswahrnehmungen haben sich bislang auf ein Bildungssystem fokussiert und hatten damit einen deutlich begrenzten Verallgemeinerungspotential. Die vorliegende Studie soll nun die Bedeutung von Aspekten des Klassenklimas für die Wahrnehmung von Diskriminierung im internationalen Vergleich untersuchen.

Fragestellung

- (1) Inwieweit hängen Diskriminierungswahrnehmungen mit soziodemographischen Variablen zusammen? Frühe Studien legen nahe, dass Mädchen und Schüler*innen, die als fremd wahrgenommen werden sensibler gegenüber Diskriminierung sind.
- (2) Welche Aspekte der Beschulung hängen mit Ungleichheitswahrnehmungen zusammen? Ein Zusammenhang zwischen positiven Sozialbeziehungen und Toleranz konnte in der Vergangenheit nachgewiesen werden, jedoch stärken Sozialbeziehungen auch das Vertrauen in politische Institutionen (Stals & Ziemes, 2024).
- (3) Inwieweit unterscheiden sich die Effekte auf Individual- und auf Klassenebene? Teilweise entfalten Klassenklimavariablen ihre Bedeutung erst auf Klassenebene.

Methoden

Stichprobe. Genutzt wurden die Daten der International Civic and Citizenship Education Study (ICCS) 2022 (Schulz et al., 2023). An ICCS 2022 haben insgesamt 24 Bildungssysteme teilgenommen, davon zwei in Deutschland. Für diese Untersuchung werden die Bildungssysteme in Dänemark (N=4769), den Niederlanden (N=2609), Nordrhein-Westfalen (NRW, N=3269), und Schleswig-Holstein (SH, N=1488) miteinander verglichen. Diese Kombination hat den Vorteil, dass intranationale Variationen mit internationalen verglichen werden können. Schüler*innen waren zum Zeitpunkt der Befragung etwa 14 Jahre alt. Durch für Large-Scale Assessments übliche Formen der Stratifizierung, Zufallsauswahlen, sowie Gewichtungen sind die Daten geeignet um repräsentative Aussagen zu machen.

Instrumente. Schüler*innen wurden gebeten auf einer vierstufigen Skala einzuschätzen, inwieweit „Religiöse Minderheiten“, „Menschen mit Migrationsgeschichte“ und „Menschen, die nicht weiß sind (z. B. Schwarze Menschen, people of color)“ Diskriminierung erleben. Die internationale Vergleichbarkeit wurde durch eine Multigruppen-CFA bestätigt.

Als soziodemographische Variablen wurden Geschlecht (Junge vs. Mädchen), sozioökonomischer Status (Indikator auf Grundlage von Büchern im Haushalt, elterlicher Ausbildung und Berufen) und familiäre Migrationsgeschichte (Schüler*in oder beide Elternteile im Ausland geboren) eingesetzt.

Für die Messung des Klassenklimas wurden drei Skalen aus dem ICCS Datensatz eingesetzt: Offenes Unterrichtsklima für Diskussionen („Lehrkräfte ermutigen Schüler*innen, mit Menschen zu diskutieren, die andere Meinungen haben.“), Bewertung der Beziehungen der Schüler*innen (z.B. „Die meisten Schüler*innen behandeln einander mit Respekt.“) und die Bewertung der Schüler*innen mit Lehrkräften (z.B. „Die Schüler*innen kommen mit den meisten Lehrkräften gut aus.“). Auch hier gab es je vierstufige Antwortformate.

Die Analysen wurden mit Mplus 8.7 und der twolevel Option durchgeführt.

Ergebnisse und ihre Bedeutung

Ergebnisse auf Individualebene:

In den Niederlanden, NRW und SH hatten Mädchen eine höhere Diskriminierungswahrnehmung. In keinem Bildungssystem zeigte der sozioökonomische Status oder die Migrationsgeschichte einen bedeutsamen Zusammenhang mit der Diskriminierungswahrnehmung. In allen Bildungssystemen war das offene Unterrichtsklima signifikant mit einer höheren Diskriminierungswahrnehmung verbunden und positive Sozialbeziehungen der Schüler*innen untereinander mit einer niedrigeren.

Ergebnisse auf Klassenebene: 5-7% der Varianz der Diskriminierungswahrnehmung lag auf der Klassenebene. In NRW und SH fanden sich negative Zusammenhänge zwischen der Einschätzung der Sozialbeziehungen der Schüler*innen mit den Diskriminierungswahrnehmungen ($\beta_{NRW} = -.802^{**}$; $\beta_{SH} = -.542$, $p = .070$). In Dänemark fand sich überraschend ein negativer Effekt des offenen Unterrichtsklimas für Diskussionen ($\beta = -.564^{*}$).

Eine Wahrnehmungslücke (Perry et al., 2021) findet sich nur für Mädchen, nicht für Schüler*innen mit Migrationsgeschichte. Positive Sozialbeziehungen scheinen eher den Eindruck einer diskriminierungsarmen Gesellschaft zu wecken. Ein offenes

Unterrichtsklima für Diskussionen hängt nur auf Individualebene mit Diskriminierungswahrnehmungen zusammen. Wenn alle Schüler*innen sich gehört fühlen, scheint dies dem Eindruck der Diskriminierungsfreiheit beizutragen.

Erfassung der kritischen Reflexion sozialer Ungleichheiten bei Jugendlichen mit Fallstudien – Zusammenhänge mit Klassenklima und schulischem Engagement

Miriam Schwarzenthal

Bergische Universität Wuppertal

Theoretischer Hintergrund

Eine kritische Reflexion sozialer Ungleichheiten kann eine Entwicklungsressource für Jugendliche in von sozialer Ungleichheit und Diskriminierung geprägten Gesellschaften sein, und hängt mit schulischem Engagement zusammen (Heberle et al., 2020; McWhirter & McWhirter, 2016; Schwarzenthal et al., 2024). Eine Klassenklima, in dem soziale Ungleichheiten offen diskutiert werden, wurde als bedeutsamer Prädiktor identifiziert (Schwarzenthal et al., 2022; Seider & Graves, 2020).

Quantitative Skalen zur Erfassung der kritischen Reflexion sozialer Ungleichheiten bei Jugendlichen (z.B. Diemer et al., 2022) beruhen rein auf Selbsteinschätzungen und erfassen die Komplexität von Reflexionsprozessen nicht. Qualitative Methoden können komplexe Reflexionsprozesse besser erfassen, wurden aber hauptsächlich mit kleineren Fallzahlen im Rahmen von Interviews angewendet (Golden et al., 2023; Hope & Bañales, 2018). Diese Studie hat zum Ziel, ein Messinstrument zur Erfassung der kritischen Reflexion sozialer Ungleichheiten zu erstellen, welches auf Fallstudien basiert und auch in Fragebogenerhebungen eingesetzt werden kann.

Zwei Fallstudien, welche soziale Ungleichheit in der Schule darstellen, wurden basierend auf Interviews mit 16 Jugendlichen (Morales et al., 2024) erstellt und durch Expert*innen-feedbacks und kognitive Vortests überprüft. Die Fallstudien sollen nun anhand einer Fragebogenuntersuchung validiert und in Zusammenhang mit zentralen Prädiktor- und Outcomevariablen (z.B. Klassenklima, schulisches Engagement) gesetzt werden.

Fragestellung

1. Wie erklären Jugendliche soziale Ungleichheiten und wie komplex sind ihre Erklärungen?
2. Sind die neu entwickelten Fallstudien ein reliables und valides Messinstrument zur Erfassung der kritischen Reflexion sozialer Ungleichheiten bei Jugendlichen?
3. Wie hängt die über Fallstudien erfasste kritische Reflexion sozialer Ungleichheiten bei Jugendlichen mit einem Klassenklima, in dem soziale Ungleichheiten offen diskutiert werden, sowie mit schulischem Engagement zusammen?

Methode

Die Hypothesen und der Analyseplan für diese Studie wurden präregistriert (<https://osf.io/vuy2x>). Schüler*innen der 9. bis 11. Klasse (N = 234, MAlter = 15,87, SDAlter = 1,21, 46% männlich, 50% weiblich, 3% nicht-binär, 56% selbst oder min. ein Elternteil im Ausland geboren) wurden mit einem Fragebogen befragt. Dieser enthielt die beiden neu entwickelten Fallstudien gefolgt von offenen Fragen, in denen die Jugendlichen gebeten wurden, das Verhalten der Personen in den Fallstudien zu erklären, Validierungsmaße (z.B. kritische Reflexionsskala, Diemer et al., 2022), Skalen zum Klassenklima (Schachner et al., 2021) sowie zum schulischen Engagement (Skinner et al., 2009).

Die offenen Antworten der Jugendlichen auf die Fallstudien werden aktuell manuell anhand eines Kodiermanuals analysiert, z.B. in Hinblick auf individuelle vs. strukturelle Erklärungen für soziale Ungleichheiten. Zudem wurde mit einem auf machine learning basierenden Tool („AutoIC“) die integrative Komplexität der Antworten kodiert, d.h. ob die Schüler*innen mehrere Perspektiven berücksichtigten und diese zu einem kohärenten Verständnis zusammenführten (Conway et al., 2020; Suedfeld et al., 1992). Anhand von Korrelations- und Pfadanalysen werden die Reliabilität und Validität des neuen Messinstruments sowie Zusammenhänge zwischen den kodierten Antworten, Klassenklima und schulischem Engagement untersucht.

Ergebnisse und Diskussion

Schüler*innen erklärten soziale Ungleichheiten mit individuellen Faktoren (z.B. Motivation), strukturellen Faktoren (z.B. Diskriminierung), sowie elterlichem Kapital (z.B. Mangel an Geld). Die Antworten wiesen niedrige bis mittlere integrative Komplexität auf.

Erste Korrelationsanalysen zeigen positive Zusammenhänge zwischen der kodierten integrativen Komplexität der Antworten auf die beiden Fallstudien ($r = .37, p < .001$), sowie zwischen der kodierten integrativen Komplexität und der selbstberichteten kritischen Reflexion ($r = .14, p = .04$) und stützen daher die Reliabilität und Validität des neu entwickelten Instruments. Pfadanalysen in Mplus zeigten, dass integrative Komplexität weder vom Klassenklima vorhergesagt wurde ($\beta = -.03, p = .46$), noch ein signifikanter Prädiktor des schulischen Engagements war ($\beta = .08, p = .36$). Weiterführende Analysen basierend auf den finalen Ergebnissen der manuellen Kodierung stehen aus.

Die Ergebnisse der Studie legen nahe, dass Fallstudien eine aufschlussreiche Methode zur Erfassung der kritischen Reflexion sozialer Ungleichheiten bei Sekundarschüler*innen darstellen. Zudem zeigen sie Forschungsbedarf hinsichtlich der Identifikation relevanter Prädiktoren der komplexen Analyse sozialer Ungleichheiten bei Jugendlichen auf.

Eine metaanalytische Untersuchung des Zusammenhangs zwischen schulischen Erfahrungen und politischer Selbstwirksamkeit im Jugendalter

Katharina Eckstein

Friedrich-Schiller-Universität Jena

Theoretischer Hintergrund

Mit dem Ziel, aktive und verantwortungsbewusste Bürger:innen zu erziehen, sind Schulen ein einflussreicher Kontext der politischen Sozialisation. Das Erleben von politischer Selbstwirksamkeit (auch Internal Political Efficacy, politisches Selbstkonzept) bildet dabei ein wichtiges Ergebnis dieses Sozialisationsprozesses, da das Vertrauen von Personen in ihre Fähigkeit, politische Prozesse zu verstehen und zu gestalten eine wesentliche Voraussetzung für eine aktive Auseinandersetzung mit gesellschaftlichen Themen darstellt (Niemi et al., 1991).

Der Zusammenhang zwischen schulischen Erfahrungen und politischer Selbstwirksamkeit wurde empirisch mehrfach bestätigt (Allen & Bang, 2015; Jagers et al., 2017). Jedoch bietet der schulische Kontext eine Vielzahl von Erfahrungsmöglichkeiten, die von curricularen bis hin zu klimatischen Merkmalen reichen. Aufgrund der Fülle an Indikatoren ist es daher schwierig, allgemeine Schlussfolgerungen darüber zu ziehen, inwieweit allgemein "günstige" schulische Bedingungen (z.B. das allgemeine Schulklima)

oder sehr spezifische Schulbedingungen (z.B. konkrete curriculare Initiativen) mit dem Erleben politischer Selbstwirksamkeit bei Jugendlichen in Verbindung stehen.

Fragestellung

Die vorliegende Studie präsentiert Ergebnisse einer aktuellen Metaanalyse, die quantitative Befunde zum Zusammenhang zwischen Schulerfahrungen und Indikatoren politischer Entwicklung von Jugendlichen untersucht. Im Fokus der Studie stehen die folgenden Ziele:

(1) Die Testung der Signifikanz und Generalisierbarkeit schulischer Effekte auf inhaltlicher Ebene (Vergleich verschiedener schulischer Indikatoren), auf kontextueller Ebene (Vergleich von Effekten auf individueller und Klassenebene) und auf messzeitlicher Ebene (Vergleich von quer- und längsschnittlichen Mustern).

(2) Die Untersuchung, inwiefern der Zusammenhang zwischen schulischen Faktoren und politischer Selbstwirksamkeit durch sozio-strukturelle Faktoren (z.B. Alter, Geschlecht) moderiert wird.

Methode

Die Meta-Analyse betrachtet Studien zwischen 2000 und 2024. Insgesamt konnten 105 unabhängige Stichproben aus 64 Publikationen zusammengetragen werden. Auf Seiten schulischer Faktoren wurden folgende Facetten unterschieden: (1) Curriculare Aspekte (1.1 Struktur, 1.2 Inhalt, 1.3 Motivational) und (2) Non-Curriculare Aspekte (2.1 Demokratisches Schul- bzw. Klassenklima, 2.2 Schulische Partizipationsmöglichkeiten, 2.3 Beziehungsklima).

Der Datenpool umfasst insgesamt 248.560 Teilnehmer:innen aus 40 Ländern. Ein Großteil des Samples stammt aus den USA oder Europa. Die meisten Studien basieren auf querschnittlichen Daten ($k = 91$; 86.7%), nur ein Siebtel des Datenpools beruht auf längsschnittlichen Designs ($k = 14$). Mehrebenenanalytische Auswertungen werden von ca. einem Drittel der eingeschlossenen Stichproben berichtet ($k = 33$, 31.4%).

Ergebnisse und ihre Bedeutung

Die Auswertungen wurden mit R (Package Metafor; Viechtbauer, 2010) sowie SPSS (Wilson's Meta-Analysis Macros for SPSS; Wilson, 2005) durchgeführt.

Mit Blick auf die erste Forschungsfrage verwiesen die Ergebnisse auf einen kleinen, jedoch signifikanten, Effekt ($r = 0.11$, $p < 0.001$). Schulische Erfahrungen hängen entsprechend signifikant mit politischer Selbstwirksamkeit zusammen. Die mittlere Effektstärke fällt höher aus, wenn nur bivariate Zusammenhänge in die Aggregation eingingen ($r = 0.17$, $p < 0.001$). Auch für die wenigen längsschnittlichen Studien ($k = 14$) konnte eine signifikante mittlere Effektstärke identifiziert werden ($r = 0.07$, $p = 0.001$). Ein Vergleich schulischer Indikatoren zeigte, dass die mittlere Effektstärke zwischen den schulischen Facetten schwankte. So ergab sich für curriculare motivationale Faktoren beispielsweise ein besonders markanter Zusammenhang ($r = 0.28$, $p < 0.001$, $k = 14$).

Mit Blick auf die zweite Forschungsfrage verwiesen die Ergebnisse auf eine signifikante Moderation durch Alter und Diversität der Stichprobe. Schulische Faktoren zeigten einen stärkeren Zusammenhang mit politischer Selbstwirksamkeit bei jüngeren und kulturell diverseren Stichproben.

Insgesamt zeigt die Studie, dass eine metaanalytische Integration empirischer Forschungsbefunde zu einem tieferen Verständnis der Rolle von Schulen in der politischen Entwicklung von Jugendlichen beiträgt. Zudem verweist sie systematisch auf existierende Lücken in der Literatur (z.B. längsschnittliche Studien, Studien außerhalb Europas und den USA).

S-55: Vielfalt in der Hochschule: Chrononormativität an Deutschlands Hochschulen – oder doch nicht?

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 129

Chair der Sitzung: Jessica Ordemann

Chair der Sitzung: Laura Naegele

Vielfalt in der Hochschule: Chrononormativität an Deutschlands Hochschulen – oder doch nicht?

Chair(s): **Jessica Ordemann** (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), Deutschland), **Laura Naegele** (TU Dortmund/BIBB)

Der deutsche Arbeitsmarkt steht vor grundlegenden Herausforderungen durch den demografischen und technischen Wandel, einen wachsenden Fachkräftebedarf sowie Migrations- und Fluchtbewegungen. Um diesen Herausforderungen zu begegnen, wird häufig auf die Notwendigkeit einer verstärkten Bildungsbeteiligung und der Implementierung einer „Kultur des lebenslangen Lernens“ hingewiesen. Gleichzeitig ist der gesellschaftliche Blick auf Bildung stark geprägt von Chrononormativität, also der Vorstellung, zu welchem Zeitpunkt und in welchem Alter Menschen im Lebensverlauf in Bildungsmaßnahmen eingebunden sind und wann diese „Bildungsphase“ abgeschlossen ist (Freeman, 2010; Wanka, 2020). Wurde Bildung früher meist als vorgelagert zum Erwerbsleben wahrgenommen (Kohli, 1985), ist sie inzwischen als ein lebenslanger, dynamischer Prozess zu verstehen, in dem Menschen (im Idealfall) mit vielfältigen Lebensläufen, aus diversen Lebenskontexten und zu unterschiedlichen Zeitpunkten im Lebensverlauf partizipieren.

Hochschulische Bildung nimmt im Kontext des lebenslangen Lernens häufig eine besondere Bedeutung ein, denn sie vermittelt nicht nur innovatives Wissen sondern auch Problemlösekompetenzen, die die gesellschaftliche Entwicklung nachhaltig prägen können. Partizipation an hochschulischer Bildung ist zudem oftmals mit dem Ziel verbunden, vergangene – sozial selektive – Bildungsentscheidungen zu revidieren und so die eigene Arbeitsmarktposition nachhaltig zu verbessern. Bedingt durch immer diversere Lebensverläufe hat die Vielfalt in der Studierendenschaft in Deutschland in den vergangenen Jahren deutlich zugenommen (Kerst et al., 2023). Studierende mit einer größeren Altersspanne und unterschiedlichen Vorerfahrungen, z.B. mit Pflegeverantwortung (Wazinski et al., 2022), mit Migrations- und Fluchtgeschichte (Grüttner 2022), aus der LGBTQI*-Community (O'Neill et al., 2022) oder beruflich Vorqualifizierte (Tieben, 2020), prägen inzwischen das Bild der Studierendenschaft an den Hochschulen. Ihnen ist gemeinsam, dass sie sich entgegen dem Mandat der Chrononormativität zu einem späteren Zeitpunkt in ihrem Leben für hochschulische Bildung entscheiden und dies häufig aus Lebenssituationen tun, die oftmals als wenig „typisch“ für Studierende angesehen werden.

Die Vielfalt an der Hochschule durch die Aufweichung von chrononormativen Lebensverläufen steht daher im Mittelpunkt des Symposiums. Das Symposium beginnt mit einem Beitrag zur biografischen Vielfalt an Deutschlands Hochschulen mit dem Schwerpunkt auf beruflich vorqualifizierte Studierende und thematisiert dann im folgenden Beitrag die Gruppe der häufig „unsichtbaren Studierenden“ mit Pflegeverantwortung, d.h. diejenigen Studierenden die neben dem Studium eine*n Angehörige*n pflegen. Der dritte Beitrag geht der Frage nach, ob die Aufweichung der Vorstellungen von Chrononormativität mit Diskriminierungserfahrungen im Studium einhergeht. Werden die Leistungen älterer Studierender beispielsweise anders bewertet als die von Jüngeren? Der letzte Beitrag thematisiert die Gruppe der älteren Promovierenden (40+ Jahre) und deren diverse Motivlagen zur Aufnahme eines Studiums bzw. einer anschließenden Promotion. Abschließend sollen aktuelle Entwicklungen der zunehmenden Heterogenität in der Studierendenschaft im Kontext diskutiert und der Frage nachgegangen werden und was diese für künftige (Hochschul-)Forschung bedeutet.

Beiträge des Symposiums

Biographische Diversität an der Hochschule

Nicole Tieben

Leibniz University Hannover

Hochschulzugänge haben sich in den letzten Jahrzehnten deutlich verändert: Zugangswege zur Studienberechtigung und Übergänge in die Hochschule werden komplexer und diverser. So erreicht beispielsweise ein wachsender Anteil an Studierenden die Hochschule nicht direkt über die allgemeinbildende gymnasiale Oberstufe, sondern über alternative schulische und berufsbildende Umwege, wie z.B. berufliche Schulen, Ausbildungen und Erwerbstätigkeit.

Die Vielfalt der Studierendenpopulationen und die Anerkennung früher erworbener Kenntnisse gewinnen an Bedeutung für Zulassungsverfahren und die Qualitätssicherung im Hochschulbereich. Ebenso sind Fragen der Studienverläufe und Erfolgsquoten nicht-traditioneller Studierender seit langem ein Thema in der Hochschulforschung verschiedener Disziplinen. Der Vortrag gibt eine Einführung in Konzepte, Begriffe und Maße biografischer Diversität und verknüpft diese mit der Entwicklung spezifischer institutioneller Bedingungen in verschiedenen Hochschulsystemen. Neben einem allgemeineren Überblick über den nationalen und internationalen hochschulpolitischen Hintergrund wird die Situation in Deutschland diskutiert und die verschiedenen „nichtlinearen“ Wege in die Hochschulbildung mit theoretischen und empirischen Fragen zu individuellen Lebensverläufen und Studienbedingungen verknüpft.

Mit Daten des Nationalen Bildungspanels (NEPS) werden die Wege in die Hochschule nachgezeichnet. Die Analysen identifizieren fünf Gruppen, deren Zugangswege zur Hochschule anhand des Typs der Hochschulzugangsberechtigung und Umwege durch die Sekundarstufe und das berufliche Bildungssystem unterscheiden. Der Beitrag zeigt die sozio-demographische und sozio-ökonomische Zusammensetzung der Gruppen sowie Unterschiede in den Studienbedingungen, akademischen Selbstkonzepten, Kompetenzen und Erfolgserwartungen.

Pflegende Studierende: eine wenig erforschte Gruppe

Moritz Hess¹, Anna Wanka²

¹Hochschule Niederrhein (HSNR), ²Goethe-Universität Frankfurt am Main

Die Zahl der Studierenden des Archetyps des jungen, männlichen Vollzeit-Studenten ohne Betreuungspflichten ist rückläufig, und die Zahl der Studierenden, die vor der Herausforderung stehen, Studium und Pflege zu vereinbaren, nimmt zu (Moreau & Robertson, 2019). Vor dem Hintergrund des demographischen Wandels kann erwartet werden, dass die Gruppe der Studierende, die informelle Pflege für erwachsene Familienmitglieder und Freunde übernehmen – pflegende Studierende – stark wachsen wird (Wazinski et al., 2022).

Bisherige Studien zu pflegenden Studierenden sind dadurch limitiert, dass sie sich auf nur eine Universität oder sogar nur ein Studiengang beschränken. Das vorliegende Papier schließt diese Lücke und beantwortet drei Fragen:

- Wie hoch ist der Anteil der Pflegenden unter den Studierenden an deutschen Hochschulen?
- Welche soziodemografischen Merkmale korrelieren mit der Wahrscheinlichkeit, ein pflegender Studierender zu sein?
- Welche Konsequenzen hat die Vereinbarkeit von Pflege und Studium?

Zur Beantwortung unserer Forschungsfragen haben wir Daten aus der deutschen Studierendenbefragung „Die Studierendenbefragung in Deutschland“ analysiert. Die Daten der Umfrage wurden 2021 unter 107.488 Studierenden erhoben. Unter Verwendung von Regressionsmethoden fanden wir heraus, dass ältere Studierende und solche, deren Eltern keine höhere Bildung haben, eine signifikant höhere Wahrscheinlichkeit haben, pflegende Studierende zu sein. Darüber hinaus zeigen die Ergebnisse, dass pflegende Studierende, kontrolliert für Alter, Geschlecht und Bildungshintergrund der Eltern, signifikant mehr Stress empfinden, eine geringere Lebenszufriedenheit berichten und weniger Vertrauen in ihren Studienerfolg haben als Studierende ohne Betreuungspflichten.

Da die Daten im Querschnittsdaten erhoben wurden, sollte man vorsichtig sein, kausale Schlussfolgerungen zu ziehen. Daher sollte zukünftige Forschung auf Längsschnittsdaten zurückgreifen, um Lebenslauf-Effekte von Studium und Pflege zu untersuchen: Zum Beispiel ob junge pflegende Erwachsene seltener ein Studium anfangen oder abbrechen, ob negative Auswirkungen auf ihre Gesundheit langfristig bestehen bleiben und wie sich ihre Arbeitsmarktteilnahme entwickelt.

Praktische Implikationen schließen die Sensibilisierung für pflegenden Studierenden und deren Herausforderungen bei Lehrenden und Verwaltungspersonal an den Universitäten sowie bei politischen Entscheidungsträgern und weiteren Stakeholdern. Darüber hinaus sollten Maßnahmen zur Unterstützung dieser Gruppe im Rahmen von ‚Caring Universities‘ in einem partizipativen Verfahren auf nationaler Ebene und an den Hochschulen selbst entwickelt werden. Dies ist von hoher gesellschaftlicher Relevanz, da man angesichts der demografischen Alterung und der wachsenden Zahl älterer Menschen, die Pflege benötigen, davon ausgehen kann, dass die Zahl der pflegenden Studierenden in vielen Ländern in Zukunft zunehmen wird.

Ageism in German higher education

Frederike Gerdes¹, Laura Naegele², Moritz Hess³, Jessica Ordemann¹

¹Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), ²TU Dortmund/BIBB, ³Hochschule Niederrhein (HSNR)

Age and ageism in German higher education – meaning discrimination of students based on their (older) chronological age – is a seldom discussed topic in literature. This is primarily due to the understanding inherent in the German education system that education prepares mainly young people for the labor market. When looking at participation rates this seems to hold true as the average age in 2022/23 of all freshmen in higher education was 18.9 years, that of all students 23.3 years and that of graduates 26.6 years (Destatis 2023). Nevertheless, 8.6 per cent of all freshmen begin their studies aged 25 and older, which is app. four years after the average student in Germany, showing that student bodies become increasingly age-divers.

This share of students aged 25 and older is unsurprising as students come to higher education from different stages in life (Kerst, Kroher, and Steinkühler 2023), e.g. entering higher education with vocational and occupational qualifications (Ordemann 2019, 2023; Wolter and Kerst 2022) or migration or refugee backgrounds (Mentges 2019; Struchholz 2021). Despite each social group contributing its unique set of knowledge and competencies to higher education, older students may encounter ageist perceptions and potential discrimination based on their age. Ageist stereotypes on older students can include negative stereotypes (e.g. being less innovative) and positive ones (e.g. being more socially competent) (Naegele et al. 2018; Marques et al. 2020). These stereotypes are often based on chrononormative expectations, which assume a "right age" for certain life phases or transitions (Wanka 2020), such as taking up formal higher education, thus overlooking people's individuality and increasingly destandardized life courses. If acted upon, these stereotypes may lead to discrimination and unequal treatment of older students, ultimately affecting their studies and overall student experience.

However, to our very best knowledge no previous work has explored in which areas older students do report ageism in the context of higher education. Thus we ask: What types of discrimination do older students experience?

We investigate the question of perceived discrimination by older students in higher education using representative data from the German Student Survey 2021 (Kroher et al. 2023). Our sample consists of 3.229 freshmen and 5.652 students who study in the fourth to sixth semester. For our freshmen sample we build an age indicator for those who are younger than 25 years of age (0) and those who are 25 and older (1) – thus being beyond the average age of freshman students in Germany. For the second sample, our age cut is 30 years and older. The first sample consists of 755 older students and the second of 488 older than average students. We adjust our analysis for known dimensions of discrimination such as gender and migration background. Due to potential cultural age differences in each subject, we control our analysis for fields of study. We will equalize the mean and variance of all adjustment variables in the logistic analysis using a weight created by entropy balancing (Hainmueller 2012).

Preliminary results show that freshmen aged 25 and older experience more often that they were not credited the ability to achieve, that they were excluded or marginalized and in general had more discriminatory experiences than their younger counterparts. Once they progress in their studies, these discriminatory effects vanish and they report they less often than their younger counterparts experience discrimination. We interpret this as an adaptation process not only to the study environment but also to being an older student in a group of younger ones (Old fish younger pond-effect).

Am I too old? Motives and perceived ageism in the "race for tenure"

Selina Staniczek¹, Laura Naegele², Jessica Ordemann³

¹Bundesinstitut für Berufsbildung (BIBB), ²TU Dortmund/BIBB, ³Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW)

A scholar's age is a seldom researched topic. However, besides one's chronological age, every Ph.D. graduate also has an academic age; that is the time that has passed since the attainment of the Ph.D. A scholar's academic age is not mandatorily linked to chronological age, making it possible that two Ph.D. graduates have the same academic age, but were born years apart. When looking at chronological age as a source of unequal treatment in labour markets ('ageism') (Butler, 1969), a rich body of research can be found (Naegele et al., 2018, 2020). Its effect on the individual are well documented, especially with regard to motivation (Rothermund et al., 2021) and career development (Swift, 2017). Ageism can reduce job satisfaction and undermine people's motivation to depart from or stay in career tracks, which are perceived to be reserved for a specific age group. The latter

relates to chrononormative expectations of what career step is appropriate at what age and which competencies are assigned to a specific age group. Paradoxically, although a higher age is generally associated with a higher level of competency, scholars who finish their Ph.D.s in their 40s are often viewed as too old to start a research career afterwards (Ordemann & Naegele, 2024). As discrimination based on age can comprise all kinds of practices and institutional structures, some more obvious than others, being exposed to ageist stereotyping could affect one's motivation to stay in academia, ultimately hampering career advancement. Against this background, the following research questions arise: What are the motives to start and stay in a scientific career? How was this affected by social origin?

To further investigate these age group specific differences and since the DZWH Ph.D. Panel does not enquire further about experience made with ageist stereotyping and unequal treatment based on age, we qualitatively explore the role of chronological age in the "race for tenure". The study, which we will present, focusses on scholars' motives to pursue, and stay in a scientific career. The study utilising sequential explanatory mixed method design (quant + qual). The objective of the quantitative study is to investigate the motive to pursue a Ph.D. and to remain in science afterward for Ph.D. graduates aged 40 years and older compared to those who are younger. Based on data from the initial cohort of doctoral candidates and the initial cohort of Ph.D.s of the National Academics Panel Study (nacaps) we first estimate multinomial logistic regression on a battery of motives why a person chooses to attain a Ph.D. followed by logistic panel regressions on how the motive to remain in science changes over time. Our analyses are adjusted by gender, migration background, Ph.D. grade, and academic productivity factors. Preliminary findings show a) small differences in the extrinsic motivation of both groups and that older b) are slightly more intrinsically motivated and c) that Ph.D. graduates have a stronger motive to remain in science. The subsequent qualitative study applies focus group interviews, with scholars who received their Ph.D. after forty years of age and now work in academia.

S-56: Individuelle Unterschiede im selbstregulierten Lernen

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 133

Chair der Sitzung: Charlotte Dignath

Chair der Sitzung: Olena Kryshko

Chair der Sitzung: Franziska Perels

Individuelle Unterschiede im selbstregulierten Lernen

Chair(s): **Charlotte Dignath** (Goethe Universität Frankfurt), **Olena Kryshko** (TU Dortmund), **Franziska Perels** (Universität des Saarlandes)

Diskutant*in(nen): **Franziska Perels** (Universität des Saarlandes)

Die Beiträge des Symposiums "Individuelle Unterschiede im selbstregulierten Lernen" bieten einen umfassenden Einblick in die vielfältigen Facetten der Selbstregulation bei Schüler*innen sowie Studierenden, insbesondere unter Berücksichtigung individueller Unterschiede und der Auswirkungen externer Einflüsse wie der Covid-19-Pandemie.

Der erste Beitrag beschäftigt sich mit der metakognitiven Überwachung in der Klausurvorbereitungsphase bei Lehramtsstudierenden. Im Rahmen einer Cross-Lagged-Panel-Studie mit 273 Lehramtsstudierenden und fünf Messzeitpunkten wird mithilfe eines Online-Tagebuchs das Zusammenspiel von Strategien und Urteilen untersucht, und wie Lernende ihre metakognitiven Fähigkeiten entwickeln und anwenden, um effektiv zu lernen bzw. ihre Leistungen zu optimieren. Die Ergebnisse tragen zu einem besseren Verständnis der komplexen Beziehung zwischen verschiedenen Komponenten der metakognitiven Überwachung während einer intensiven Lernphase bei.

Der zweite Beitrag beleuchtet die motivationale Selbstregulation im Studium vor und während der pandemiebedingten Campusschließung. Bei 922 Erstsemesterstudierenden werden interindividuelle Unterschiede hinsichtlich diverser Personenmerkmale und ab dem vierten Semester intraindividuelle Entwicklungsverläufe in Motivationsregulation über sieben Messzeitpunkte hinweg betrachtet. Es wird aufgezeigt, wie die Nutzung von Motivationsregulationsstrategien sowie motivationsregulatorische Selbstwirksamkeitsüberzeugung sich im Laufe des Studiums bzw. unter den Pandemiebedingungen verändern und welche sozio-demographischen, kognitiven, motivationalen oder persönlichkeitsbezogenen Faktoren dabei eine prädiktive Rolle spielen.

Der dritte Beitrag vergleicht das selbstregulierte Lernen im pandemiebedingten Distanzunterricht von 413 Schüler*innen der Grundschule, Gesamtschule und des Gymnasiums. Mit Hilfe einer inhaltlich-strukturierenden qualitativen Inhaltsanalyse werden sowohl die Herausforderungen als auch die Chancen untersucht, die sich aus der Situation des Distanzunterrichts ergeben. Dabei zeigten sich insbesondere Schwierigkeiten beim Einsatz von Selbstregulationsstrategien, während mögliche positive Aspekte eines gesteigerten Autonomieerlebens und einer individuellen Lernzeitgestaltung nicht gefunden werden konnten.

Im vierten Beitrag wird die individuelle Entwicklung von Strategienutzung und metakognitivem Strategiewissen bei 596 Schüler*innen der Sekundarstufe über zwei Messzeitpunkte hinweg erfasst. Es wird analysiert, welche Faktoren der Schüler*innen (z. B. Mindset und Motivation), sowie die Merkmale der familiären (z. B. elterliche Unterstützung) und schulischen Umgebung (z. B. Rückmeldung zum Lernen) relevant sind, um zu verstehen, wie Schüler*innen ihr Lernverhalten über die Zeit hinweg anpassen und verbessern. Die Ergebnisse tragen dazu bei, die individuellen Unterschiede in der Entwicklung und Anwendung von Lernstrategien besser zu erfassen.

Zusammen bieten die Beiträge dieses Symposiums wertvolle Einblicke in die Mechanismen und Faktoren, die das selbstregulierte Lernen beeinflussen. Sie zeigen auf, wie unterschiedliche Bedingungen und individuelle Unterschiede die Entwicklung und Anwendung von Selbstregulationsstrategien prägen. Das Symposium trägt dazu bei, ein tieferes Verständnis für die Komplexität des selbstregulierten Lernens zu entwickeln und liefert wichtige Impulse für die Förderung individueller Lernprozesse in verschiedenen Bildungskontexten. Dabei werden unterschiedliche Lerngruppen (Schüler*innen verschiedener Schulformen; Studierende) in verschiedenen Lernkontexten (pandemiebedingter Fernunterricht; regulärer Unterricht; Studium) herangezogen und multi-methodale und komplexe Designs (Cross-Lagged-Panel; Längsschnitt; Qualitative Inhaltsanalyse) verwendet, um diese Fragestellungen angemessen zu adressieren. Im abschließenden Diskussionsbeitrag werden die vier inhaltlichen Beiträge des Symposiums kritisch gewürdigt und ihre Befunde sowie Implikationen zusammenfassend eingeordnet.

Beiträge des Symposiums

Wie wird in der Klausurvorbereitungsphase metakognitiv überwacht? Mikrogenetische Entwicklung und Zusammenspiel von Strategien und Urteilen

Marion Händel¹, Ulrike Nett², Donna Bryce², Markus Dresel²

¹Hochschule Ansbach, ²Universität Augsburg

Theoretischer Hintergrund

Metakognitive Überwachung ist eine zentrale Komponente des selbstregulierten Lernens (z. B. Winne & Hadwin, 2008). Während des Lernens ist es wichtig, das eigene Lernverhalten und die Lernergebnisse zu überwachen, um den Lernprozess bei Bedarf regulierend zu steuern. Die Forschung zur metakognitiven Überwachung konzentrierte sich bislang häufig entweder auf den Einsatz von Überwachungsstrategien oder auf metakognitive Urteile und deren Genauigkeit. Beide Komponenten und interindividuelle Unterschiede dabei werden jeweils als bedeutsam betrachtet für die Qualität von Lernprozessen und den Lernerfolg (z. B. Panadero, 2017). Allerdings waren diese beiden Forschungsperspektiven bislang wenig miteinander verbunden und es ist unklar, wie Überwachungsstrategien und metakognitive Urteile interagieren und wie sie sich gemeinsam entwickeln (vgl. Dörrenbächer-Ulrich & Perels, 2023; Golke et al., 2022; Händel & Dresel, 2022; Händel et al., 2023).

Theoretisch können mehrere Objekte und Phasen der Überwachung unterschieden werden. Einerseits ist der offensichtlichste Gegenstand von Überwachung der Lernprozess selbst – daher wurde bislang die Anwendung von Überwachungsstrategien hauptsächlich in der Lernphase analysiert. Andererseits beziehen sich metakognitive Urteile oft auf die Ergebnisse des Lernens – in der bisherigen Forschung wurden sie meist im Hinblick auf die Testergebnisse explizit erfragt und analysiert (z. B. die Korrektheit der verschiedenen Aufgaben in einem Test). Unklar ist jedoch, ob und in welchem Ausmaß Lernende auch während der Lernphase spontane Urteile bilden. Zudem waren bisherige Studien zu den Zusammenhängen zwischen Überwachungsstrategien und metakognitiven Urteilen überwiegend querschnittlich angelegt und geben daher keine Auskunft darüber, wie sich verschiedene Komponenten der metakognitiven Überwachung im Lernprozess entwickeln und gegenseitig beeinflussen.

Fragestellung

Daher werden im Folgenden drei Fragestellungen untersucht:

Q1: Wie hängen spontane metakognitive Urteile und Urteilsgenauigkeit zusammen?

Q2: Wie verändern sich metakognitive Überwachungsstrategien und metakognitive Urteile während einer intensiven Lernphase?

Q3: Wie beeinflussen sich metakognitive Überwachungsstrategien und metakognitive Urteile (spontane Urteile und Urteilsgenauigkeit) im Laufe der Lernphase gegenseitig?

Methode

Um zu untersuchen, wie sich Überwachungsstrategien und metakognitive Urteile mikrogenetisch entwickeln und sich gegenseitig beeinflussen, wurde eine längsschnittliche, Cross-Lagged-Panel-Studie realisiert (vgl. Siegler & Crowley, 1991). Die Online-Tagebuchstudie mit $N = 273$ Lehramtsstudierenden (Fachsemester $M = 3.2$, $SD = 1.5$; 74 % weiblich, 25 % männlich, 1 % divers) fand eine Woche vor einer abschließenden Modulklausur statt und umfasste fünf aufeinanderfolgende Lern- und Prüfungstage mit jeweils einem Wissenstest (eingeführt als „Probewissenstest“) und einem Begleitfragebogen. Die Qualität eingesetzter Überwachungsstrategien sowie die Bildung spontaner metakognitiver Urteile wurden täglich in Bezug auf die Lernsituation (d. h. tägliche Vorbereitung auf den Wissenstests) erfragt, während explizite metakognitive Urteile jeweils zu den Items im Wissenstests erhoben wurden. Als Maße der Urteilsgenauigkeit wurden Sensitivität und Spezifität analysiert. Insgesamt vier Random-Intercept-Cross-Lagged-Panel-Modelle (RI-CLPM; Hamaker et al., 2015) wurden berechnet, wobei jeweils einerseits die Überwachungsstrategien und andererseits die Urteile und die Urteilsgenauigkeit in die Modelle eingingen.

Ergebnisse und ihre Bedeutung

Die Korrelationen zwischen spontanen Urteilen und der Urteilsgenauigkeit (Sensitivität bzw. Spezifität) waren entweder nicht signifikant oder von geringer Stärke (Q1), was darauf hindeutet, dass nur geringe Zusammenhänge zwischen impliziten und expliziten Urteilen bestehen.

Zur Untersuchung von Q2 und Q3 zeigen RI-CLPM einen guten Modellfit für alle untersuchten Modelle ($.940 \leq CFI \leq .966$). Während die autoregressiven Pfade der Überwachungsstrategien über die fünf Erhebungstage hinweg signifikant und mäßig bis hoch ausfielen, zeigten sich keine signifikanten autoregressiven Pfade für spontane Urteile sowie Sensitivität und Spezifität der explizit erfragten Urteile. Bei der Betrachtung der Cross-Lagged-Pfade zeigt sich, dass die Anzahl der Spontanurteile zu einer höheren Qualität der Überwachungsstrategien führt, nicht aber umgekehrt.

Die Ergebnisse tragen dazu bei, die Beziehung zwischen verschiedenen Komponenten der metakognitiven Überwachung während des Lernprozesses zu verstehen.

Motivationale Selbstregulation im Studium vor und während der pandemiebedingten Campusschließung: Interindividuelle Unterschiede und intraindividuelle Entwicklungsverläufe

Olena Kryshko¹, Michael Becker¹, Jens Fleischer², Detlev Leutner³

¹TU Dortmund, ²Ruhr-Universität Bochum, ³Universität Duisburg-Essen

Erfolgreiches Studieren erfordert ein hohes Maß an Selbstregulation beim Lernen (Dresel et al., 2015). Neben der eigenen kognitiven Verarbeitung soll dabei u.a. auch die eigene Motivation überwacht und bei Bedarf gesteuert werden (Wolters, 2003). Studien zufolge sind Motivationsprobleme bzw. maladaptive Motivationsveränderungen im Studium weitverbreitet und können zu unerwünschten akademischen Konsequenzen führen (Benden & Lauermann, 2022; Heublein, 2014; Kosovich et al., 2017; Robinson et al., 2019; Schnettler et al., 2020). Durch effektive motivationale Selbstregulation kann es Studierenden allerdings gelingen, ihr lernbezogenes Motivationsniveau aufrechtzuerhalten bzw. zu erhöhen (Daumiller & Dresel, 2019; Engelschalk et al., 2015, 2016; Leutner et al., 2001).

Laut aktuellen Rahmenmodellen besteht der komplexe Prozess motivationaler Selbstregulation aus zwei metamotivationalen Subprozessen, Selbst-Monitoring sowie strategische Kontrolle, die sich wechselseitig bedingen und deren Abläufe jeweils von unterschiedlichen Formen metamotivationalen Wissens bzw. metamotivationaler Überzeugungen abhängen (Miele et al., 2024; Miele & Scholer, 2018). Während bereits viele Befunde zur Relevanz unterschiedlicher Aspekte motivationaler Selbstregulation für den akademischen Erfolg im Studium vorliegen (z. B. Bülke et al., 2018; Grunschel et al., 2016; Kryshko et al., 2020; Trautner & Schwinger, 2020), wissen wir bisher nur wenig darüber, 1) wie sich diese im Verlauf des Studiums oder unter besonderen Lernbedingungen entwickeln sowie 2) ob und in welchem Ausmaß verschiedene soziodemografische, kognitive, motivationale oder persönlichkeitsbezogene Faktoren die jeweiligen Entwicklungsverläufe vorhersagen. Außerdem ist Längsschnittforschung zur Motivationsregulation generell sehr selten.

Angesichts dieser Forschungslücken, besteht das Ziel der vorliegenden Studie darin, potenzielle interindividuelle Unterschiede (d.h. zwischen Personen) in motivationaler Selbstregulation und ihren intraindividuellen Entwicklungsverläufen (d.h. innerhalb von Personen über die Zeit) bei Studierenden zu untersuchen. Dabei sollen folgende Forschungsfragen adressiert werden: Wie entwickelt sich die Motivationsregulation im Laufe der Zeit? Sagen die interindividuellen Unterschiede hinsichtlich einer Reihe von Personenmerkmalen (Geschlecht, Bildungsherkunft, Abiturdurchschnittsnote, kognitive Grundfähigkeiten, motivationale Zielorientierungen und Big Five Persönlichkeitsfaktoren) das Anfangsniveau sowie die intraindividuellen Entwicklungsverläufe von Motivationsregulation vorher? Zur Beantwortung dieser Forschungsfragen werden Längsschnittdaten eines großen, abgeschlossenen Forschungsprojekts herangezogen.

Die für diese Studie relevanten Datenerhebungen umfassen acht Messzeitpunkte. Zu Beginn ihres Studiums gaben $N = 922$ Erstsemesterstudierende (49.6% weiblich; 65.6% Studierende erster Generation) unterschiedlicher Fächer an zwei großen Universitäten im Rahmen einer Paper-Pencil-Befragung Auskunft zu den oben genannten Personenmerkmalen ($t_1 = WS 2016/2017$). Zur Erfassung pädagogisch-psychologischer Konstrukte wurden dabei (Sub)Skalen etablierter Messinstrumente eingesetzt (kognitive Grundfähigkeiten, KFT 4-12+R, Heller & Perleth, 2000; motivationale Zielorientierungen, SELLMO, Spinath et al. 2002; Persönlichkeitsfaktoren, NEO-FFI, Borkenau & Ostendorf, 1993; NEO-PI-R, Ostendorf & Angleitner, 2004). Ab dem vierten bis zum zehnten Studiensemester ($t_2 = SoSe 2018$ – $t_8 = SoSe 2021$) wurden die Studierenden jeweils mithilfe einer Online-Datenerhebung in der vorlesungsfreien Zeit zu ihrer Motivationsregulation beim Lernen anhand von zwei Skalen befragt (Nutzung von Motivationsregulationsstrategien, Schwinger et al., 2007; motivationsregulatorische Selbstwirksamkeitsüberzeugung, Kryshko et al., 2023). Die Stichprobe umfasste $N = 445$ Studierende zu t_2 , $N = 455$ zu t_3 , $N = 414$ zu t_4 , $N = 386$ zu t_5 , $N = 338$ zu t_6 , $N = 368$ zu t_7 und $N = 360$ zu t_8 . Dabei fanden die drei letzten Messzeitpunkte während der pandemiebedingten Campusschließungen statt und bieten somit die Möglichkeit, Motivationsregulationsprozesse in einem speziellen akademischen Lernkontext zu untersuchen, der Studierende sowohl motivational als auch selbstregulatorisch vor großen Herausforderungen stellte (Biwer et al., 2021; Daniels et al., 2021; Janke et al., 2022; Naujoks et al., 2021).

Die Datenauswertungen, bestehend aus deskriptiven und korrelativen Analysen, Überprüfung der Messinvarianz der beiden Motivationsregulationsskalen über die Zeit, Analyse fehlender Werte sowie Modellierung von (konditionalen) Wachstumskurven, dauern derzeit noch an und die Ergebnisse werden bis zur Tagung vorliegen. Insgesamt werden Befunde dieser Längsschnittstudie nicht nur unser theoretisches Verständnis von Motivationsregulationsprozessen erweitern, sondern können auch dazu beitragen, zielgerichtete Interventionen für Studierende zu entwickeln, die häufig mit Motivationsproblemen zu kämpfen haben.

Selbstreguliertes Lernen im Distanzunterricht: Schwierigkeiten und Chancen aus Perspektive der Schüler:innen

Tobias Koch, Christian Reintjes, Sonja Nonte, Ferdinand Stebner
Universität Osnabrück

Distanzunterricht während der Corona-Pandemie stellte alle Akteur:innen des Bildungssystems vor große Herausforderungen. Wichtige Strukturen und Gewohnheiten brachen weg (Reintjes, 2022), die für das schulische Lernen wichtig sind. Lehrkräfte konnten ihre Schüler:innen nicht in gewohnter Weise unterstützen, wodurch die Kompetenz des selbstregulierten Lernens noch wichtiger wurde. Selbstreguliertes Lernen ist nicht gleichbedeutend mit offenem Lernen, aber es ist eine Voraussetzung dafür, dass offenes Lernen gelingt. Das Lernen während der Schulschließungen war zumeist offeneres Lernen (Wößmann et al., 2021). Es ist anzunehmen, dass Lernende im Distanzunterricht vor allem dann gut zurechtkamen, wenn sie in der Lage waren, selbstreguliert zu lernen. Studien zeigen, dass selbstreguliertes Lernen selten gefördert wird (Dignath & Büttner, 2018). Deutsche Schüler:innen besitzen nicht unbedingt das notwendige prozedurale und konditionale Strategiewissen, das sie befähigt, zu wissen, wie, warum und wann sie im Sinne des selbstregulierten Lernens entsprechende Lernstrategien einsetzen. Darüber hinaus fehlt ihnen Übung, was zu zusätzlicher kognitiver, lernhinderlicher Belastung führt (Wirth et al., 2020).

Es ist anzunehmen, dass deutsche Schüler:innen vorwiegend Probleme hatten, im Distanzunterricht zurechtkommen und erfolgreich zu lernen. Und weil die Offenheit des Lernens auch Vorteile mit sich bringt, insbesondere im Bereich der Selbstbestimmung (Ryan & Deci, 2017), sollen auch positive Aspekte während des Distanzunterrichts betrachtet werden. Es stellt sich die Frage, welche Schwierigkeiten und Chancen Schüler:innen während des Distanzunterrichts wahrgenommen haben und wie diese mit selbstreguliertem Lernen zusammenhängen. Zudem soll untersucht werden, ob es Unterschiede zwischen Schulformen, Geschlecht und Alter gibt.

In der vorliegenden Studie werden die offenen Antworten von $N = 413$ Schüler:innen (Alter $M = 13.27$; $SD = 3.02$; $MIN = 6$; $MAX = 21$; 55.4% weiblich) auf die Frage nach individuellen Schwierigkeiten und Chancen während der Schulschließungen ausgewertet. Die Antworten werden mithilfe einer inhaltlich-strukturierenden qualitativen Inhaltsanalyse analysiert, was eine inhaltliche Strukturierung der Daten ermöglicht (Kuckartz & Rädiker, 2022). Zum Vergleich der Schulformen, Geschlecht und Alter werden uni- und multivariate Varianzanalysen sowie Korrelationsanalysen durchgeführt.

Die Ergebnisse zeigen, dass viele der genannten Schwierigkeiten und Chancen mit den Strategien des selbstregulierten Lernens in Verbindung stehen. Aussagen zu kognitiven, metakognitiven und ressourcenbezogenen Strategien werden codiert. In den quantitativen Analysen zeigen sich signifikante Unterschiede zwischen den Schulformen, $F(2,284) = 4.04$, $p < .05$, $\eta^2 = .03$. Gesamtschüler:innen haben allgemein signifikant mehr Schwierigkeiten als Gymnasiast:innen ($p < .05$, $\eta^2 = .03$), während Gymnasiast:innen ($M = 0.26$) im Vergleich zu Grundschüler:innen ($M = 0.08$) signifikant mehr Schwierigkeiten in der Kategorie der internen ressourcenbasierten Lernstrategien melden ($p < .05$, $\eta^2 = .05$). Die Nennung der Schwierigkeiten unterscheidet sich nicht hinsichtlich des Geschlechts ($p > .05$). Das Alter korreliert nur mit Schwierigkeiten, die den internen ressourcenbasierten Lernstrategien ($r = .22$; $p < .001$) zugeordnet werden konnten. Je älter die Schüler:innen sind, desto mehr Schwierigkeiten werden in diesem Bereich genannt. Im Bereich der Chancen lassen sich vergleichsweise weniger Aussagen mit den Strategien des selbstregulierten Lernens in Verbindung bringen.

Fazit: Viele Schüler:innen erlebten während der Schulschließungen Schwierigkeiten, die mit den Strategien des selbstregulierten Lernens zusammenhängen. Dies untermauert womöglich die Notwendigkeit eines Paradigmenwechsels hin zur Etablierung einer Schulkultur des selbstregulierten Lernens, um Schüler:innen dauerhaft die Kompetenzen des selbstregulierten Lernens zu vermitteln und sie für zukünftige Lernherausforderungen zu wappnen.

Individuelle Entwicklung von Strategienutzung und metakognitivem Strategiewissen – welche Faktoren sind relevant?

Sabrina N. Brunner, Yves Karlen
Universität Zürich

Selbstreguliertes Lernen (SRL) ist bedeutsam für erfolgreiches Lernen in der Schule (Richardson et al., 2012) und darüber hinaus (Taranto & Buchanan, 2020). SRL beinhaltet kognitive, metakognitive, ressourcen-orientierte und motivationale Komponenten, die miteinander interagieren und für den gesamten Lernprozess relevant sind (Pintrich, 2000). Selbstregulierte Lernende sind in der Lage durch die Wahl von geeigneten Strategien ihren Lernprozess erfolgreich zu planen, zu regulieren und zu reflektieren (Ertmer & Newby, 1996). Dies erfordert metakognitives Strategiewissen (metacognitive strategy knowledge; MSK) über die Umsetzung und den gezielten Nutzen von SRL-Strategien (Pressley et al., 1987). Im Jugendalter zeigt sich eine Abnahme in der SRL-Strategienutzung (van der Veen & Peetsma, 2009) während MSK stabil bleibt (Karlen et al., 2014). Dabei gibt es interindividuelle Unterschiede in den Entwicklungsverläufen, abhängig von Charakteristika der Lernenden (Jeong & Feldon, 2023; van der Veen & Peetsma, 2009) sowie von Faktoren der familiären (Karlen et al., 2014; Xu et al., 2010) und schulischen Umgebung (Schuitema et al., 2016). Eine besondere Rolle kommt dabei den Lehrpersonen zu, ihren Schüler:innen SRL näherzubringen und MSK zu vermitteln (Kramarski, 2018). Allerdings haben Lehrpersonen diesbezüglich Mühe bei der Diagnose und Förderung (Dignath & Veenman, 2021; Karlen et al., 2023). Daher haben wir für die Sekundarschule ein digitales SRL-Tool entwickelt. In diesem Tool können Schüler:innen für verschiedene SRL-Komponenten MSK-Tests ausfüllen sowie ihre Strategienutzung anhand eines Fragebogens beurteilen. Die Schüler:innen erhalten danach eine automatisierte Auswertung sowie adaptive Lerntipps zur Verbesserung ihres SRLs. Die Lehrpersonen haben Zugriff auf die Resultate ihrer Schüler:innen und erhalten zusätzliches Fördermaterial. Das Ziel der aktuellen Studie war es, den Effekt des digitalen Tools auf die individuelle Entwicklung von MSK und Strategienutzung zu überprüfen. Zudem wurden Charakteristika der Schüler:innen, sowie ihrer familiären und schulischen Umgebung als mögliche Einflussfaktoren untersucht. Berücksichtigt wurde eine Längsschnittstichprobe von $N = 596$ Schüler:innen mit einem mittleren Alter von $M = 13.93$ ($SD = 0.91$), die an zwei Messzeitpunkten im Abstand von etwa drei Monaten für fünf SRL-Komponenten (kognitive Strategien, metakognitive Strategien, Zeitplanung, Management der Lernumgebung, Motivationsregulation) MSK-Tests ausgefüllt und ihre Strategienutzung beurteilt haben. Die Mehrheit der Schüler:innen ($n = 502$) hat hierfür das neu entwickelte Tool benutzt und die damit verbundenen

Auswertungen und Lerntipps erhalten. Bei einer Subgruppe der Schüler:innen ($n = 360$) erhielten die Lehrpersonen zudem eine kurze Weiterbildung zum Thema SRL. Zusätzlich gab es eine Kontrollgruppe von $n = 94$ Schüler:innen, welche die MSK-Tests und die selbstbeurteilte Strategienutzung in einem Online-Fragebogen ohne Rückmeldung ausgefüllt hat. Somit hat die Studie ein längsschnittliches Design mit drei Gruppen. Schüler:innen aller Gruppen haben zusätzlich einen Fragebogen ausgefüllt, in dem Charakteristika der Schüler:innen (z. B. Mindset-Variablen, Motivation zur Verhaltensänderung) sowie der familiären (sozioökonomischer Status, elterliche Unterstützung) und schulischen Umgebung (z. B. Rückmeldung zum Lernen) erhoben wurden. Um den Effekt der Gruppe zu testen, wurden ANOVAs berechnet mit der Differenz zwischen den Messzeitpunkten ($t_2 - t_1$) als abhängige Variable. Mittels Multilevelmodellen wurden die Effekte der verschiedenen Variablen auf die individuelle Entwicklung von t_1 zu t_2 angesehen, wobei für die Klassenzugehörigkeit kontrolliert wurde. Hierfür wurde der t_2 -Messwert unter Kontrolle für den t_1 -Messwert durch die verschiedenen Variablen vorhergesagt. Erste Ergebnisse zeigen, dass die Nutzung des digitalen Tools mehrheitlich keine Effekte auf die individuelle Entwicklung der Schüler:innen hatte. Schüler:innen-Charakteristika hatten insbesondere einen Effekt auf die Entwicklung der Strategienutzung, während die wahrgenommene elterliche Unterstützung die Entwicklung der Strategienutzung und des MSK für die meisten SRL-Komponenten vorhersagte. Effekte der schulischen Umgebung wurden nur für das Leistungsniveau gefunden, wobei ein höheres Leistungsniveau mit besserem Entwicklungsverlauf einherging. Die Ergebnisse weisen darauf hin, dass es nicht ausreicht, SRL in der Schule zu fördern. Auch im Jugendalter scheint die elterliche Unterstützung relevant zu sein, was die Bedeutung der Zusammenarbeit der Schule mit den Eltern hervorhebt.

S-57: Forschungsdaten und Open Science – Praktische Hinweise zum Erheben, Teilen und Nachnutzen von Forschungsdaten und Angebote des Verbund Forschungsdaten Bildung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 135

Chair der Sitzung: Aleksander Kocaj

Chair der Sitzung: Bayer Sonja

Forschungsdaten und Open Science – Praktische Hinweise zum Erheben, Teilen und Nachnutzen von Forschungsdaten und Angebote des Verbund Forschungsdaten Bildung

Chair(s): **Aleksander Kocaj** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB)), **Sonja Bayer** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation)

In der empirischen Bildungsforschung zeigt sich eine veränderte Forschungskultur, die zunehmend Open Science Prinzipien und die Bereitstellung von Daten für Re- und Sekundäranalysen betont. Um diese Potenziale auszuschöpfen, müssen Forschungsdaten nach ethisch-rechtlichen Standards erhoben sowie auffindbar und nachvollziehbar dokumentiert sein (Wilkinson et al., 2016). Hierbei unterstützen Forschungsdateninfrastrukturen mit passenden Angeboten zur Archivierung und Suche von Forschungsdaten. Für die Bildungsforschung vereint der Verbund Forschungsdaten Bildung (VerbundFDB) seit 2013 nationale Einrichtungen, um entsprechende Angebote und Beratung für das Forschungsdatenmanagement und das Teilen von Forschungsdaten bereitzustellen. Der VerbundFDB unterstützt Forscher*innen, das Potenzial der generierten Forschungsdaten umfassend zu nutzen und eine Kultur des Data Sharing zu etablieren, wobei die Arbeitsprozesse durch Open Science Prinzipien im gesamten Data Life Cycle geleitet werden.

Das Symposium illustriert die Angebote des VerbundFDB für Forscher*innen entlang des Data Life Cycle und bezieht ausgewählte Open Science Praktiken ein. Der Fokus liegt auf praktischen Hinweisen sowie aktuellen Herausforderungen im Hinblick auf die ethisch-rechtlichen Anforderungen bei der Datenerhebung, des Forschungsdatenmanagements in Einklang mit Förderrichtlinien, dem Datenteilen und Datenfinden sowie der Durchführung von Sekundäranalysen. Das Symposium gliedert sich in vier Beiträge:

Der erste Beitrag behandelt rechtliche und ethische Aspekte bei der Erhebung, Speicherung, Weitergabe und Nachnutzung von Forschungsdaten. Dabei sind insbesondere datenschutzrechtliche und urheberrechtliche Themen in der Bildungsforschung im gesamten Datenlebenszyklus relevant. Besondere Bedeutung kommen der informierten Einwilligung sowie den Vorgaben genehmigender Gremien zu, die Rahmenbedingungen für den Umgang mit den Daten festlegen.

Der zweite Beitrag behandelt die Datenmanagementplanung in Förderanträgen. Der Fokus liegt darauf, Forscher*innen praktische Hinweise zu geben, welche Anforderungen und Möglichkeiten Forschungsförderer hinsichtlich der Datenaufbereitung und Datendokumentation bereithalten und warum das Forschungsdatenmanagement schon bei der Projektbeantragung bedacht werden sollte. Dabei wird u. a. der standardisierte Datenmanagementplan für die Bildungsforschung (Stamp; Künstler-Sment et al., 2024) vorgestellt.

Der dritte Beitrag gibt eine Übersicht zur Publikation von Forschungsdaten und stellt Infrastrukturen (z. B. Repositorien, Forschungsdatenzentren, Konsortien) vor, über die Forschungsdaten archiviert und geteilt werden können. Zudem gibt er Entscheidungshilfen zur Auswahl eines geeigneten Archivierungsortes und Hinweise zum Finden von Forschungsdaten für Sekundäranalysen.

Der vierte Beitrag fokussiert die Nachnutzung von Forschungsdaten und beschreibt, wie Open Science Praktiken bei der Durchführung von Sekundäranalysen umgesetzt werden können. Dabei wird auf die Präregistrierung von Sekundäranalysen, Möglichkeiten der Transparenz und Nachvollziehbarkeit bei der Datenaufbereitung, Datenauswertung und Datensatzzitation sowie die Robustheit von Forschungsergebnissen eingegangen.

Abschließend werden alle Beiträge in den Forschungskreislauf eingebettet und durch Leitfragen sowie Fragen aus dem Plenum diskutiert.

Beiträge des Symposiums

Forschungsdaten erheben – Ethische und rechtliche Voraussetzungen zur (Nach-)Nutzung von Forschungsdaten

Reiner Mauer¹, Sonja Bayer², Karoline Harzenetter¹

¹GESIS – Leibniz-Institut für Sozialwissenschaften, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Ethisches Handeln ist ein essenzieller Bestandteil guter wissenschaftlicher Praxis und betrifft den gesamten Lebenszyklus von Forschungsdaten – von der Erhebung, über die Analyse bis hin zur Archivierung und Nachnutzung der Daten (Deutsche Forschungsgemeinschaft, 2022). Den rechtlichen Rahmen für die Erhebung von Forschungsdaten bestimmen maßgeblich das Datenschutzrecht und das Urheberrecht. Wissenschaftliche Untersuchungen an Schulen bedürfen zudem in der Regel einer Genehmigung (Meyermann & Porzelt, 2019). Weitere Auflagen, die beispielsweise von prüfenden Gremien oder Förderer gestellt werden, können den Umgang mit Forschungsdaten zusätzlich regeln.

Vor dem Hintergrund einer offenen Forschung und Datensparsamkeit verlangen Drittmittelgeber, wie die Europäische Kommission im Rahmen von Horizon Europe (European Commission, 2021) oder das BMBF im Bereich der empirischen Bildungsforschung (BMBF, 2022), dass Forschungsdaten aus geförderten Projekten zugänglich gemacht werden. Dies steht jedoch im Konflikt mit dem Recht von Studienteilnehmer*innen, grundsätzlich selbst über die Preisgabe und Verwendung der eigenen personenbezogenen Daten zu bestimmen oder der Pflicht, geistiges Eigentum zu schützen.

Strukturiertes FDM erleichtert den ethisch und rechtlich sicheren Umgang mit Forschungsdaten und gewährleistet die Einhaltung von Standards guter wissenschaftlicher Praxis sowie möglichen Auflagen durch Genehmigungsbehörden, Gremien und Förderer. Zu den wesentlichen Aspekten des FDM zählen der Umgang mit personenbezogenen Daten, von der informierten Einwilligung (Verbund Forschungsdaten Bildung, 2018) und forschungsethischen Prüfung (RatSWD, 2017) zur Erhebung personenbezogener Daten über Maßnahmen zur Pseudonymisierung und Anonymisierung bis hin zur geschützten Speicherung personenbezogener Informationen (RatSWD, 2020).

Darüber hinaus müssen die Nutzungsrechte der verwendeten Forschungsmaterialien geklärt sein. Dies betrifft sowohl urheberrechtlich geschützte Materialien anderer, die im Forschungsprojekt genutzt werden, als auch die Lizenzierung der im

Forschungsprojekt generierten Daten. Eigenentwickeltes Material kann lizenziert im Sinne einer offenen Forschungspraxis für eine Weiterverwendung bereitgestellt werden (bspw. Creative Commons Lizenzen).

Der Vortrag adressiert den ethisch-rechtlichen Umgang mit Forschungsdaten von der Studienplanung bis zur Aufbewahrung und zeigt Möglichkeiten auf, wie Daten für die wissenschaftliche Weiternutzung bereitgestellt werden können. Dabei werden auch Möglichkeiten und Grenzen aufgezeigt, Daten mit Personenbezug für wissenschaftliche Zwecke zu teilen. Die Zuhörer*innen werden in die Lage versetzt, (1) ein grundlegendes Verständnis für das Forschungsdatenmanagement zu entwickeln, (2) zentrale ethisch-rechtliche Aspekte adäquat zu adressieren und (3) Forschungsdaten rechtskonform zu verarbeiten und für die Nachnutzung vorzubereiten.

Datenmanagement für Förderanträge planen

Salome Wagner, Sonja Bayer

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Der adäquate Umgang mit Forschungsdaten bekommt immer mehr Bedeutung und wird von Forschungsförderern explizit gefordert. Dies stellt Wissenschaftler*innen vor neue Herausforderungen bei der Planung und Umsetzung ihrer Forschungsprojekte. Sowohl Fachgesellschaften, Universitäten und Forschungseinrichtungen als auch Forschungsförderer stellen Anforderungen, Richtlinien und Empfehlungen für den adäquaten Umgang mit Forschungsdaten bereit (z. B. Allianz der deutschen Wissenschaftsorganisationen, 2010; DGfE, GEBF & GFD, 2020). Neben der Sensibilisierung Forschender für einen ethisch-rechtlich fundierten Umgang mit Forschungsdaten sollen diese Regelungen auch eine Auseinandersetzung mit Themen wie Datensicherung, Datensicherheit, Datenteilen, langfristiger Datenarchivierung, Transparenz und Nachvollziehbarkeit schon zu Beginn eines Forschungsprojektes initiieren. Forschungsförderer setzen zudem Vorgaben bei der Projektbeantragung, die Nutzung bereits vorhandener Daten zu prüfen, die Notwendigkeit eigener Datenerhebungen zu begründen und ein strukturiertes Forschungsdatenmanagement im Projekt zu verankern (siehe z. B. BMBF, 2022; DFG, 2022). Teilweise ist es verpflichtend, erhobene Daten und entwickeltes Forschungsmaterial für die wissenschaftliche Weiternutzung durch Dritte verfügbar zu machen, soweit dies rechtlich möglich ist. Damit wird Forschungsdatenmanagement auch Teil der Berichtspflicht (BMBF, 2022). Um den Prozess des Forschungsdatenmanagements in Projekten sowie den zielorientierten und effizienten Einsatz der notwendigen Tätigkeiten über den gesamten Lebenszyklus der Daten zu unterstützen, verlangen Forschungsförderer zunehmend Angaben von den Antragstellenden zum geplanten Forschungsdatenmanagement. Die konkreten Auflagen zum Datenmanagement und zur Datenmanagementplanung variieren jedoch je nach Förderer bzw. Institution hinsichtlich des Umfangs und des Inhalts. Eine Beschreibung des Umgangs mit den zu erhebenden Forschungsdaten und die Auseinandersetzung mit deren Archivierungsmöglichkeiten wird aber übereinstimmend als bedeutsamer Teil des Förderantrags empirischer Forschung erachtet.

In diesem Beitrag werden die Planung, Umsetzung und Dokumentation des Forschungsdatenmanagement als zentraler Bestandteil eines Forschungsprojekts adressiert (BMBF, 2021). Dabei wird ein besonderer Fokus auf die verpflichtenden Anforderungen bezüglich des Forschungsdatenmanagements bei der Beantragung von Fördermitteln gelegt. Andererseits werden die Besonderheiten und Herausforderungen des Datenmanagements für Verbundprojekte oder Projekte mit Partner*innen diskutiert. Der Beitrag geht auch auf Möglichkeiten ein, zusätzliche Mittel für die Aufbereitung und Bereitstellung von Daten für die Nachnutzung sowie die Publikation von Ergebnissen im Open Access zu beantragen. Ziel dieses Beitrags ist es, Bildungsforschenden praktische Hinweise für die Planung und Umsetzung des Forschungsdatenmanagements an die Hand zu geben und das Unterstützungsangebot des Verbund Forschungsdaten Bildung zu Datenmanagementplänen vorzustellen. Dabei werden auch die Checklisten, Planungshilfen und Fallbeispiele des Standardisierten Datenmanagementplans Stamp (Künstler-Sment et al., 2024) vorgestellt, welche ein strukturiertes Forschungsdatenmanagement unterstützen (DDP Bildung & VerbundFDB, 2024). Der Beitrag bietet Wissenschaftler*innen der Bildungsforschung einerseits einen Überblick über die Richtlinien verschiedener Forschungsförderer und andererseits Unterstützung, wie sie diese Richtlinien erfüllen und adäquates Forschungsdatenmanagement bei der Erstellung und Umsetzung von Förderanträgen realisieren können.

Forschungsdaten teilen und finden: Eine Übersicht über Repositorien und Forschungsdatenzentren

Thomas Lösch¹, Aleksander Kocaj²

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Im Zuge der Open Science Bewegung sind in der heutigen Wissenschaftslandschaft zahlreiche Infrastrukturen zur Archivierung und Bereitstellung von Forschungsdaten entstanden. Diese Vielfalt bietet Vorteile, kann aber auch dazu führen, dass Forschende den Überblick verlieren und Schwierigkeiten haben, die passende Plattform für ihre Bedürfnisse zu finden und auszuwählen. Der Beitrag adressiert diese Herausforderungen und bietet eine Übersicht über Repositorien (Stall et al., 2020) und Forschungsdatenzentren (Buck et al., 2022) im Bereich der empirischen Bildungsforschung. Diese Infrastrukturen werden aus zwei Perspektiven betrachtet. Einerseits wird die Frage verfolgt, welche Repositorien geeignet sind, um eigene Daten zu teilen. Andererseits wird betrachtet, wie passende Daten für die eigene Forschung gefunden werden können. Für beide Fragen stellt der Verbund Forschungsdaten Bildung (VerbundFDB) ein zentrales Unterstützungsangebot dar.

Um eigene Daten zu teilen, können Forschende zwischen generischen Repositorien wie dem Open Science Framework (OSF) und fachspezifischen Forschungsdatenzentren (FDZen) wie dem FDZ Bildung oder dem FDZ des Instituts zur Qualitätsentwicklung im Bildungswesen (FDZ am IQB) wählen (Logan et al., 2021). In diesem Beitrag vergleichen wir diese Angebote anhand von Kriterien wie den FAIR-Prinzipien (findable, accessible, interoperable, reusable; Wilkinson et al., 2016), Möglichkeiten für Zugriffsbeschränkungen und der langfristigen Nutzbarkeit von Daten. Ausgewählte Forschungsdatenzentren werden vorgestellt, gefolgt von konkreten Empfehlungen für verschiedene Anwendungsszenarien. Zusätzlich werden praktische Hinweise zur verteilten Archivierung oder zur Dauer der Datenaufbereitung in FDZen gegeben.

Um Daten zur Sekundärnutzung zu finden, gibt es zahlreiche Angebote, die oft nicht vollständig genutzt werden (Jansen et al., 2022). Eine Passung zwischen Daten und Forschungsfragen ist entscheidend für deren Nutzung in sekundäranalytisch ausgerichteten Forschungsprojekten (Houben & Eckert, 2022). Der VerbundFDB bietet eine Datensuche, in der Datensätze mehrerer FDZen wie beispielsweise den FDZen von IQB, DIPF, GESIS, DZHW oder ZPID in Bezug auf Studieninformationen sowie Variablen und Konstrukten durchsucht werden können. Wir stellen diese Suchfunktion vor und geben Hinweise zu ihrer zielführenden Nutzung (Gregory et al., 2020). Zusätzlich weisen wir auf übergreifende Angebote zur Suche geeigneter Forschungsdaten für Sekundäranalysen wie KonsortSWD, QualidataNet, CESSDA und re3data hin. Darüber hinaus gehen wir auf praktisch-organisatorische Fragen der Datennutzung ein. Dazu zählen insbesondere Zugangswege zu Daten, die vom freien Zugang über Nutzungsanträge bis zu Remote Access reichen.

Insgesamt bietet unser Vortrag praxisnahe Hinweise zur Nutzung von Repositorien und Forschungsdatenzentren. Der VerbundFDB kann mit seinen Infrastrukturen und Beratungsangeboten eine zentrale und maßgeschneiderte Unterstützung bieten.

Forschungsdaten nachnutzen: Aktuelle Perspektiven und Entwicklungen bei der Durchführung von Sekundäranalysen in Einklang mit Open Science-Praktiken

Aleksander Kocaj, Annegret Rucker, Christin Rüdiger, Benjamin Becker, Lisa Pegelow
Institut zur Qualitätsentwicklung im Bildungswesen (IQB)

Die Bereitstellung gut dokumentierter Forschungsdaten ist ein zentraler Baustein der Open Science-Bewegung und eine Voraussetzung für die Reproduzierbarkeit und Replizierbarkeit von Forschungsergebnissen. In den letzten 15 Jahren hat sich auch in der empirischen Bildungsforschung ein Kulturwandel hin zu stärkerem Datenteilen vollzogen, was neue Möglichkeiten für Sekundäranalysen geschaffen hat (Hopfenbeck et al., 2018; Lenkeit et al., 2015). Sekundäranalysen zeichnen sich durch die Nutzung bereits vorhandener Datensätze zur Beantwortung neuer Forschungsfragen durch Forscher*innen, die in der Regel nicht an der Datenerhebung beteiligt waren, aus (Jansen et al., 2023). Voraussetzung für Sekundäranalysen ist, dass Datensätze gut dokumentiert und leicht auffindbar sind (Wilkinson et al., 2016). Dabei gewinnt auch die Reproduzierbarkeit und Replizierbarkeit von Forschungsergebnissen auf Basis von Sekundäranalysen zunehmend an Bedeutung. Dieser Beitrag zeigt, wie Open Science-Praktiken bei Sekundäranalysen umgesetzt werden können. Zunächst wird erläutert, wie sich Sekundäranalysen von Primäranalysen unterscheiden, welche praktischen Schritte erforderlich sind und welche Herausforderungen sich im Hinblick auf Open Science-Praktiken ergeben. Die Schritte umfassen die Datensatzauswahl, Datenbeantragung, Zugangswege, Datenaufbereitung, Datenauswertung, Einhaltung datenschutzrechtlicher Bestimmungen, Zitationsrichtlinien und die Dokumentation des Forschungsprozesses. Diese werden anhand von Beispielen aus dem Forschungsdatenzentrum am Institut zur Qualitätsentwicklung im Bildungswesen (FDZ am IQB) illustriert. Dabei wird auf spezifische Kernbereiche von Open Science fokussiert. Diese umfassen erstens die Präregistrierung von Sekundäranalysen, wobei hilfreiche Templates (Mertens & Kryptos, 2019; Van den Akker et al., 2021) vorgestellt werden. Zweitens wird auf Möglichkeiten der Transparenz und Nachvollziehbarkeit bei der Datenaufbereitung, Datenauswertung und Datensatzzitation eingegangen. Abschließend wird die Bedeutung von Robustheitschecks bei der Durchführung von Sekundäranalysen diskutiert, insbesondere, wenn sie auf Basis großer Schulleistungsstudien durchgeführt werden. Aufgrund der großen Stichprobe und Variablenanzahl bestehen für Sekundärforscher*innen viele Freiheitsgrade hinsichtlich der Datenaufbereitung und -auswertung (Weston et al., 2019). Wir stellen zwei Möglichkeiten vor, systematisch die Robustheit von Forschungsergebnissen über verschiedene plausible Analysestrategien und geeignete Datensätze zu untersuchen: Multiversumsanalysen (Steege et al., 2016) und integrative Datenanalysen (Curran & Hussong, 2009). Diese Strategien werden konzeptuell erläutert und mit empirischen Beispielen aus großen Schulleistungsstudien illustriert. Wir empfehlen, bei Sekundäranalysen die Reproduzierbarkeit zu berücksichtigen und deren Potenziale zur Untersuchung der Robustheit von Forschungsergebnissen zu nutzen.

S-58: Mehrdimensionale Kompetenzorientierung im mathematisch-naturwissenschaftlichen Unterricht – Ergebnisse aus PISA-Ceco 2022

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: SN 169

Chair der Sitzung: Doris Lewalter

Chair der Sitzung: Anja Schiepe-Tiska

Chair der Sitzung: Samuel Greiff

Mehrdimensionale Kompetenzorientierung im mathematisch-naturwissenschaftlichen Unterricht – Ergebnisse aus PISA-Ceco 2022

Chair(s): **Doris Lewalter** (Zentrum für internationale Vergleichsstudien (ZIB), Technische Universität München), **Anja Schiepe-Tiska** (Universität Koblenz)

Diskutant*in(nen): **Samuel Greiff** (Goethe-Universität Frankfurt)

Schulische Bildung soll neben der Vermittlung von (Fach-)Wissen und kognitiven Fertigkeiten auch das Entwickeln und Herausbilden positiver schul- und fachbezogener Einstellungen, Werthaltungen und motivationaler Orientierungen verfolgen (Schiepe-Tiska et al., 2016a; OECD, 2018). Das Verfolgen mehrdimensionaler Bildungsziele ist wichtig, damit Schüler*innen Vorstellungen von der eigenen Person entwickeln, sich beruflich orientieren und Beziehungen zu anderen und der Welt im Allgemeinen klären können (Aktionsrat Bildung, 2015).

Eine der zentralen Lernumgebungen zur Förderung des Erreichens mehrdimensionaler Bildungsziele ist der Unterricht. Die Umstellung auf kompetenzorientierte Lehrpläne in Deutschland ermöglichte eine klarere Fokussierung mehrdimensionaler Bildungsziele (KMK, 2010), da die primär fachlich orientierten Bildungsstandards um die Perspektive motivational-emotionaler Lernergebnisse erweitert werden können (KMK, 2022; KMK, 2024a, b, c; Schiepe-Tiska et al., 2021).

Die Ergebnisse der aktuellen PISA-Studie 2022 zeigen jedoch, dass die mathematischen und naturwissenschaftlichen Kompetenzen der 15-jährigen Schüler*innen in Deutschland im Vergleich zu vorherigen PISA-Erhebungsrunden signifikant gesunken sind (Diedrich et al., 2023b; Kastorff et al., 2023). Außerdem nahmen sowohl die Freude und das Interesse am Mathematikunterricht als auch die instrumentelle Motivation ab (Diedrich et al., 2023a). Somit zeigt sich in Deutschland ein rückläufiger Trend in der Entwicklung mehrdimensionaler mathematischer und naturwissenschaftlicher Kompetenzen.

Mit der an PISA 2022 angeknüpften Begleitstudie PISA-Ceco (Classroom experience, characteristics and outcome: multidimensional educational goals and the views of students and teachers) wird ein wichtiger Beitrag geleistet, um genauere Einblicke in das Zusammenspiel von Kompetenzen und Einstellungen der Lehrkräfte zu mehrdimensionalen Bildungszielen und dem Erreichen dieser mehrdimensionalen Bildungsziele durch die Lernenden zu geben.

Im Rahmen des Symposiums wird der gemeinsamen Forschungsfrage nachgegangen, welchen Beitrag eine an mehrdimensionalen Kompetenzen orientierte Unterrichtsgestaltung leisten kann, um das Erreichen kognitiver und motivationaler Bildungsziele in der Mathematik und den Naturwissenschaften in Deutschland zu erklären.

Zur Beantwortung der Forschungsfrage greift der erste Beitrag die Einstellungen von Mathematiklehrkräften in Bezug auf die wahrgenommene Nützlichkeit der Kompetenzorientierung auf und untersucht den Zusammenhang mit der Wahrnehmung eines kognitiv aktivierenden Unterrichts und der Lernmotivation der Schüler*innen.

Der zweite Beitrag nimmt das fachdidaktische Wissen als Aspekt professioneller Kompetenz von Mathematik- und Naturwissenschaftslehrkräften als Prädiktor für Unterrichtsqualität und Lernerfolg von Schülerinnen und Schülern in den Blick und stellt anschließend die Entwicklung und Validierung eines Tests zur Erfassung der mathematischen und naturwissenschaftsdidaktischen Kompetenz von Lehrkräften vor.

Ein abschließender Beitrag untersucht den Zusammenhang zwischen fachbezogener und motivationaler Prüfungsaufgabengestaltung und der naturwissenschaftlichen Kompetenz der Lernenden sowie deren Freude und Interesse und der instrumentellen Motivation im Naturwissenschaftsunterricht.

Beiträge des Symposiums

Kompetenzorientierung von Lehrkräften, kognitive Aktivierung und Motivation von Schüler*innen in Mathematik

Mirjam Weis¹, Anja Schiepe-Tiska², Pia Todtenhöfer¹, Frank Reinhold³, Doris Lewalter¹

¹Zentrum für internationale Vergleichsstudien (ZIB), Technische Universität München, ²Universität Koblenz,

³Pädagogische Hochschule Freiburg

Theoretischer Hintergrund

Obwohl die kognitive Aktivierung als eine Basisdimension von qualitativem Unterricht gilt (Lanski et al., 2017; Klieme et al., 2009; Praetorius et al., 2018), wurden bisher selten verschiedene Aspekte der kognitiven Aktivierung im Zusammenhang mit Unterrichtsergebnissen betrachtet. Bisherige Studien zeigten, dass kognitive Aktivierung im Mathematikunterricht einen positiven Einfluss auf die Motivation der Schüler*innen haben kann (Lazarides & Buchholz, 2019; Schiepe-Tiska et al., 2016b) und motivationale Merkmale eher mit der individuellen Sichtweise des Unterrichts zusammenhängen (Kunter, Baumert & Köller, 2007). In der vorliegenden Studie untersuchten wir zwei Aspekte der kognitiven Aktivierung im Mathematikunterricht (Ermutigung zum mathematischen Denken; mathematisches Argumentieren) im Zusammenhang mit der Motivation der Schüler*innen auf Individual- und Klassenebene. Kompetenzorientierter Unterricht bietet den Lehrkräften möglicherweise mehr Freiraum für die Gestaltung des Unterrichts und die Fokussierung auf mehrere Lernziele, beispielsweise auch auf die Motivation (Schiepe-Tiska et al., 2021). Obwohl kompetenzorientierter Unterricht als eine Dimension „guten“ Unterrichts beschrieben wird (Helmke, 2015), gibt es bisher wenige Studien zu Zusammenhängen zwischen kompetenzorientiertem Unterricht und Unterrichtsergebnissen (Schiepe-Tiska et al., 2021). Die vorliegende Studie setzt hier an und untersucht Zusammenhänge zwischen der Kompetenzorientierung von Lehrkräften, der kognitiven Aktivierung und der Motivation von Schüler*innen.

Fragestellungen

- Gibt es positive Zusammenhänge zwischen der kognitiven Aktivierung (Ermutigung zum mathematischen Denken; mathematisches Argumentieren) im Mathematikunterricht und der Motivation von Schüler*innen in Mathematik (Freude und Interesse; instrumentelle Motivation) auf Individualebene und auf Klassenebene?

- Hängt die wahrgenommene Nützlichkeit der Kompetenzorientierung von Mathematiklehrkräften positiv mit der kognitiven Aktivierung im Mathematikunterricht und mit der Motivation der Schüler*innen in Mathematik zusammen?
- Gibt es indirekte Zusammenhänge zwischen der Kompetenzorientierung von Mathematiklehrkräften und der Motivation der Schüler*innen in Mathematik vermittelt über die kognitive Aktivierung?

Methode

Die Stichprobe der vorliegenden Studie umfasst 1234 Neuntklässler*innen und deren Mathematiklehrkräfte in Deutschland, die sowohl an der PISA-Studie 2022 als auch an der nationalen Ergänzungsstudie PISA-Ceco teilgenommen haben. Die wahrgenommene Nützlichkeit der Kompetenzorientierung von Mathematiklehrkräften wurde im PISA-Ceco Onlinefragebogen erfasst. Zur Erhebung der kognitiven Aktivierung gaben die Jugendlichen im PISA-Fragebogen an, wie häufig ihre Mathematiklehrkraft sie zum mathematischen Denken ermutigte beziehungsweise zum mathematischen Argumentieren aufbezog. Zudem beantworteten die Jugendlichen im PISA-Fragebogen Fragen zu Freude und Interesse an Mathematik sowie zur mathematikbezogenen instrumentellen Motivation. Wir analysierten die Daten unter Berücksichtigung der Klasse als Clustervariable sowie mit Geschlecht und Schulart als Kontrollvariablen mittels Mehrebenen-Regressionsanalysen und Strukturgleichungsmodellen in Mplus.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigten positive Zusammenhänge zwischen beiden Aspekten der kognitiven Aktivierung und der Motivation der Schüler*innen auf der Individualebene aber nicht auf der Klassenebene, wenn die Aspekte der kognitiven Aktivierung separat einbezogen wurden. Wenn beide Aspekte der kognitiven Aktivierung einbezogen wurden, zeigten sich lediglich Zusammenhänge zwischen der Ermutigung zum mathematischen Denken mit Freude und Interesse ($\beta = .15$, $SE = .04$, $p < .01$) sowie mit instrumenteller Motivation ($\beta = .14$, $SE = .04$, $p < .01$). Demnach scheint die Ermutigung zum mathematischen Denken eine größere Rolle für die Motivation zu spielen als das mathematische Argumentieren. Außerdem zeigten sich positive Zusammenhänge zwischen der Kompetenzorientierung von Mathematiklehrkräften und der kognitiven Aktivierung im Mathematikunterricht (Ermutigung zum mathematischen Denken: $\beta = .39$, $SE = .11$, $p < .01$; mathematisches Argumentieren: $\beta = .24$, $SE = .11$, $p < .05$). Es zeigten sich zwar keine direkten Zusammenhänge zwischen der Kompetenzorientierung von Mathematiklehrkräften und der Motivation der Schüler*innen aber indirekte Zusammenhänge, vermittelt über die Ermutigung zum mathematischen Denken. Die vorliegende Studie liefert Hinweise darauf, dass es wichtig ist verschiedene Aspekte der kognitiven Aktivierung differenziert zu betrachten. Darüber hinaus bestätigen die Ergebnisse, dass motivationale Aspekte stärker mit der individuellen Sichtweise auf den Unterricht zusammenhängen. Die indirekten Zusammenhänge machen die Relevanz der Kompetenzorientierung von Mathematiklehrkräften für die Motivation von Schüler*innen deutlich.

Erhebung des fachdidaktischen Wissens von Lehrkräften in den Fächern Mathematik, Biologie, Chemie und Physik im Rahmen von PISA-Ceco 2022

Oliver Tepner¹, Stefanie Hohenleitner¹, Stefan Krauss¹, Nathalie Stegmüller¹, Christina Ehrs¹, Arne Dittmer¹, Victoria Telser¹, Benjamin Münch¹, Stefan Sorge², Knut Neumann²

¹Universität Regensburg, ²IPN Kiel

Theoretischer Hintergrund

Fachdidaktisches Wissen (FDW) ist ein wesentlicher Teil professioneller Handlungskompetenz von Lehrkräften und stellt einen starken Prädiktor für Unterrichtsqualität und Lernerfolg von Schülerinnen und Schülern dar (Anderson & Taner, 2023; Baumert et al., 2010). Im Refined Consensus Model (RCM) of pedagogical content knowledge (PCK) von Carlson et al. (2019) steht die Anwendung des FDW (enacted PCK) während der Planung, der Durchführung und des Reflektierens von Unterricht im Fokus und wird als professionelle Handlungskompetenz sichtbar. Basis ist das individuelle FDW von Lehrpersonen (personal PCK), welches sich aus dem Disziplin-eigenen verfügbaren theoretischen FDW (collective PCK) und aus eigenen Erfahrungen speist. Im Modell werden weitere Aspekte wie Domänenspezifität, pädagogisches Wissen und Fachwissen sowie Kenntnisse des Curriculums berücksichtigt (Carlson et al., 2019).

Nach Baumert et al. (2001) kann ein Mangel an theoretischen Kenntnissen zur Einschränkung der Handlungsfähigkeit im Unterricht führen. Basierend auf den Arbeiten der Projekte FALKO (Krauss et al., 2017), ProWiN-Chemie (Tepner & Dollny, 2014) sowie KiL/KeiLa (Sorge et al., 2019) wurde eine gemeinsame und an COACTIV angelehnte Konzeptionalisierung des FDW entwickelt, um entsprechende Items neu zu konstruieren oder aus den o.a. Projekten zu adaptieren. Um in den PISA-Ceco-Tests die unterrichtliche Handlungsfähigkeit abzubilden, wurden offene Itemformate implementiert.

Fragestellungen

Um im Rahmen weiterer Auswertungen der PISA-Ceco-Daten Zusammenhänge zwischen Schüler*innen-, Unterrichts- und Lehrkräftevariablen messen zu können und Informationen über die Struktur des FDW zu erhalten, wird zunächst folgenden Fragestellungen nachgegangen:

1. Lässt sich das FDW in einem offenen Antwortformat objektiv und reliabel erfassen?
2. Finden sich fachspezifische Unterschiede in den Korrelationen einzelner Facetten des FDW?
3. Gibt es in Bezug auf das FDW von Mathematiklehrkräften Unterschiede zwischen COACTIV 2004 und Ceco 2022?

Methode

In den Fächern Biologie, Chemie und Physik wurden jeweils neun Items entwickelt, von denen nach der Pilotierung jeweils sechs zum Einsatz in PISA-Ceco ausgewählt wurden. Die Kodierschemata für die 18 naturwissenschaftlichen FDW-Items wurden in einem iterativen Prozess konsensual entwickelt, während in der Mathematik der Kodierleitfaden für sechs ausgewählte FDW-Items aus COACTIV weitgehend übernommen wurde.

In allen vier beteiligten Fächern wird bei PISA-Ceco das FDW jeweils durch drei Facetten repräsentiert: „Wissen über kognitive Schülerfehler und -schwierigkeiten“ („Skog“), „Wissen über Erklären und Repräsentieren von Inhalten“ („E&R“) sowie das „Wissen über das Potential von Aufgaben und der Gestaltung von Erkenntnisgewinnungsprozessen“ („Pot“). Jede Facette wurde mit zwei Items operationalisiert, sodass der FDW-Test für jedes Fach insgesamt sechs Items umfasst. Alle Items wurden in einem offenem Kurzantwortformat online administriert.

Insgesamt liegen von N=319 Lehrkräften auswertbare Daten zu einem der vier FDW-Tests vor (davon gültige in Mathematik: n=129, Biologie: n=54, Chemie: n=49 und Physik: n=44 Lehrkräfte).

Ergebnisse und ihre Bedeutung

Für die FDW-Tests ergaben sich durch Cronbachs Alpha indizierte Reliabilitäten von .65 (Mathematik), .64 (Biologie), .71 (Chemie) und .62 (Physik). Die über Cohens-Kappa bestimmte Objektivität (Interraterreliabilität bei Doppelkodierungen) der FDW-Items ergab über alle Fächer für die vergebenen Codes Werte zwischen .66 und .89. Die Rho-Werte für die Kompetenz-Scores liegen zwischen .73 und .99.

Während im Fach Mathematik FDW-Schulformunterschiede zwischen Lehrkräften an Gymnasien und Nicht-Gymnasien nachgewiesen werden können ($T=4,38$; $p<.001$; $d=0,79$), ist dies weder einzeln noch zusammengefasst in den Fächern Biologie, Chemie und Physik der Fall. Möglicherweise hat eine fachwissenschaftliche vertiefte Gymnasial-Ausbildung im Fach Mathematik einen stärkeren Einfluss auf das FDW als in den Naturwissenschaften (Kirschner et al., 2017).

Die Facetten Skog, E&R sowie Pot korrelieren mit r zwischen .25 und .44 schwach bis mittelstark miteinander: Mathematik: .26-.46, Biologie: .36-.44, Chemie: .37-.50, Physik: .19-.40.

Im Fach Mathematik liegen die Mittelwerte des FDW auf vergleichbarem Niveau zu 2004 (COACTIV).

Auf der GEBF sollen die Entwicklung der Test-Items und des gemeinsamen Kodierhandbuchs sowie weitergehende Regressionsanalysen unter Einbezug der Schüler*innen- und Unterrichtsebene vorgestellt werden.

Mehrdimensionale Bildungsziele im Naturwissenschaftsunterricht: Einfluss der Aufgabengestaltung auf kognitive und motivationale Lernergebnisse

Pia Todtenhöfer¹, Anja Schiepe-Tiska², Mirjam Weis¹, Doris Lewalter¹

¹Zentrum für internationale Vergleichsstudien (ZIB), Technische Universität München, ²Universität Koblenz

Theoretischer Hintergrund

Im Sinne mehrdimensionaler Bildungsziele sollte kompetenzorientierter Unterricht nicht nur die Förderung grundlegender Kompetenzen der Schülerinnen und Schüler in den Fokus nehmen, sondern auch sozio-emotionale Aspekte wie die Motivation der Schüler*innen abdecken (Schiepe-Tiska et al., 2016a).

Im Jahr 2004 wurden in Deutschland nationale Bildungsstandards entwickelt, um Kompetenzen zu definieren, die Schüler*innen am Ende einer bestimmten Bildungsphase erworben haben sollten (KMK, 2004a, b, c). Die Bildungsstandards können zusätzlich um die Perspektive motivationaler Lernergebnisse erweitert werden (Schiepe-Tiska et al., 2021). Ziel eines kompetenzorientierten naturwissenschaftlichen Unterrichts sollte folglich sein, sowohl die Kompetenzbereiche der Bildungsstandards (Fachwissen, Erkenntnisgewinnung, Kommunikation, Bewertung) als auch motivationale Aspekte (dargestellt durch innere Differenzierung, Lebensweltbezug, soziale Eingebundenheit, Kompetenzunterstützung, Autonomieunterstützung) in der Unterrichtsgestaltung abzudecken.

Eine Analyse der Prüfungsaufgaben ermöglicht einen konkreten Einblick in das Unterrichtsangebot, da sie den Unterrichtsstoff eines bestimmten Unterrichtszeitraums repräsentativ abbilden sollen (Lederer, 2008) und besonderes Potenzial haben, das Erreichen mehrdimensionaler Bildungsziele (hier abgedeckt durch die Naturwissenschaftsleistung sowie Freude und Interesse und instrumentelle Motivation in den Naturwissenschaften) zu fördern.

Insbesondere im MINT-Unterricht haben (Prüfungs-)Aufgaben das Potenzial, das Erreichen mehrdimensionaler Bildungsziele zu fördern, da sie zentral für den Lernprozess sind (Knoll, 2003).

Fragestellungen

1. Inwieweit adressieren Prüfungsaufgaben im Naturwissenschaftsunterricht in Deutschland Kompetenzbereiche der Bildungsstandards und beinhalten motivationales Potenzial?
2. Wie hängt das Potenzial der Prüfungsaufgaben mit der Naturwissenschaftskompetenz, Freude und Interesse sowie der instrumentellen Motivation in den Naturwissenschaften zusammen?

Methode

Die Daten basieren auf N = 693 Prüfungsanalyseeinheiten von 29 Lehrkräften, deren Klassen im Rahmen von PISA 2022 erfasst worden sind. Für die Analyse der Prüfungsaufgaben wurden im Rahmen von PISA-Ceco eingesetzte Klassenarbeiten aus dem Schuljahr 2021/2022 eingesammelt. Die Aufgaben wurden unter anderem in Bezug auf die Dimensionen „Orientierung an den Kompetenzbereichen der Bildungsstandards“ sowie in Bezug auf das motivationale Potenzial analysiert (Heinle et al., 2022).

Alle Kategorien wurden anhand dichotomer Antwortkategorien doppelt kodiert. Bei abweichenden Kodierungen einigten sich die Rater*innen auf einen Code. Diese Prüfungsaufgaben wurden mit dem klassenbasierten Datensatz der Schüler*innen aus PISA 2022 (NGymnasium = 381 aus 23 Klassen und Nicht gymnasiale Schularten = 199 aus 13 Klassen) verbunden, um Zusammenhänge zwischen mehrdimensionalen Bildungszielen (Naturwissenschaftskompetenz, Freude und Interesse sowie instrumentelle Motivation in den Naturwissenschaften) und Unterrichtsmerkmalen (Fachwissen, Erkenntnisgewinnung, Kommunikation, Bewertung sowie das motivationale Potenzial) zu untersuchen. Grundlage für die Orientierung der Prüfungsaufgaben an den Bildungsstandards waren die zum Zeitpunkt der Erhebung gültigen Bildungsstandards aus dem Jahr 2004. Aufgrund der genesteten Datenstruktur wurden für jedes Bildungsziel Mehrebenenanalysen in MPlus für beide Schularten (Gymnasium = Gym, Nicht gymnasiale Schularten = Nongym) gleichzeitig durchgeführt.

Ergebnisse und ihre Bedeutung

Bezogen auf die Orientierung der Prüfungsaufgaben an den Bildungsstandards zeigt sich, dass die Kategorie Fachkenntnisse (Gym 99.54 %, Nongym 98.84%) in nahezu allen Prüfungsaufgaben vorkommt. Erkenntnisgewinnung (Gym 67.97%, Nongym 49.80 %) ist insgesamt in etwas mehr als der Hälfte der Prüfungsaufgaben Bestandteil, während das Kommunizieren in etwas weniger als der Hälfte (Gym 40.55%, Nongym 48.26%) vorkommt. Den Kompetenzbereich Bewertung konnten wir nur selten finden (Gym 5.30%, Nongym 11.97%). Dennoch erfüllen die Prüfungsaufgaben aller Lehrkräfte Teilkompetenzen der Bildungsstandards. In Bezug auf das motivationale Potenzial ist deutlich zu erkennen, dass nur der Lebensweltbezug (Gym 36.87 %, Nongym 27.79%) ein nennenswerter Bestandteil der analysierten Prüfungsaufgaben ist.

Die Ergebnisse der Mehrebenenmodelle zeigen, dass die Schulart einen Moderatoreffekt hat und die Schularten getrennt betrachtet werden müssen. Insgesamt konnten nur an nicht gymnasialen Schularten Zusammenhänge zwischen Aufgabenmerkmalen und mehrdimensionalen Bildungszielen gefunden werden. Hier zeigt sich, dass das Aufgabenmerkmal „Kommunikation“ mit der Naturwissenschaftskompetenz positiv zusammenhängt ($\beta = 69.03$, $SE = 28.64$, $p < .05$), bei dem Aufgabenmerkmal „Bewertung“ zeigt sich ein negativer Zusammenhang ($\beta = -58.54$, $SE = 22.28$, $p < .01$). Es wurden keine Zusammenhänge zwischen Aufgabenmerkmalen und motivationalen Aspekten der Schüler*innen in den Naturwissenschaften gefunden.

S-59: „Ein soziales Merkmal kommt selten allein“: Intersektionalitätseffekte in Lehrkrafturteilen und dem Erleben von Schüler*innen

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 142

Chair der Sitzung: Ineke M. Pit-ten Cate

Chair der Sitzung: Sabine Glock

Chair der Sitzung: Katja Scharenberg

„Ein soziales Merkmal kommt selten allein“: Intersektionalitätseffekte in Lehrkrafturteilen und dem Erleben von Schüler*innen

Chair(s): Ineke Pit-ten Cate (LUCET Universität Luxemburg), Sabine Glock (Bergische Universität Wuppertal, Deutschland)

Diskutant*in(nen): Katja Scharenberg (Ludwig-Maximilians-Universität München)

Forschung zur Bildungsbenachteiligung von Schüler*innen mit bestimmten sozialen Hintergrundmerkmalen wie der Migrationshintergrund und der sozio-ökonomische Status der Familien fokussiert sich zumeist auf das eine interessierende, dominante Merkmal. Allerdings haben Schüler*innen noch mehr Merkmale, die in Kombination mit dem dominanten sozialen Merkmal interagieren können. Diese Intersektionalitätseffekte sollen in diesem Symposium näher untersucht werden. Im ersten Beitrag des Symposiums wird die Kombination zwischen sozio-ökonomischem Status und Geschlecht von Schüler*innen in der Leistungsbeurteilung von angehenden Lehrkräften untersucht. Die Autorinnen berichten von Salienzeffekten, die sich im schlechteren Leistungsurteil für Schüler*innen mit niedrigem sozialem Status verglichen mit hohem sozialem Status zeigen. Auch werden Jungen besser in Mathematik bewertet als Mädchen. Mädchen erhalten aber kein besseres Urteil in Deutsch. Der Vorteil der Jungen in Mathematik wird konsistent in der Literatur berichtet (Holder & Kessels, 2017), allerdings konnte der immer wieder für Mädchen aufgeführte Vorteil in Deutsch (Siems-Muntoni & Retelsdorf, 2023) nicht gefunden werden. Auch im zweiten Beitrag spielt das Geschlecht der Schüler*innen eine Rolle, wird hier aber mit einem türkischen Migrationshintergrund kombiniert. Ebenso berichten die Autorinnen dieses Beitrags ausschließlich von Salienzeffekten. Mädchen werden schlechter als Jungen in Mathematik bewertet und Schüler*innen mit Migrationshintergrund werden schlechter in Deutsch bewertet als diejenigen ohne Migrationshintergrund. Die Schlechterstellung in Deutsch von Schüler*innen mit Migrationshintergrund wird immer wieder repliziert (Kleen & Glock, 2018). Auch der zweite Beitrag konnte keinen Vorteil von Mädchen in Deutsch finden. Das Geschlecht und der Migrationshintergrund wird ebenfalls im dritten Beitrag untersucht, hier aber im Falle der Schullaufbahnpfehlungen. Die Autor*innen berichten, dass, auch nach Kontrolle der Schulleistungen, Schülerinnen eher auf das höchste Schulniveau empfohlen werden und die Jungs eher auf den niedrigsten Schulzweig verwiesen werden. Mädchen mit Migrationshintergrund werden weniger auf den niedrigen Schulzweig empfohlen als Jungen mit Migrationshintergrund. Insgesamt gibt es keinen Unterschied für Jungen und Mädchen ohne Migrationshintergrund. In dieser Studie werden tatsächlich Intersektionalitätseffekte und damit keine Salienzeffekte gefunden. Im vierten Beitrag des Symposiums wird die Kombination von Geschlecht, sozio-ökonomischer Status und Migrationshintergrund betrachtet, aber aus der Schülerperspektive. Auch hier werden Intersektionalitätseffekte gefunden. Jungen aus Familien mit hohem sozio-ökonomischem Status weisen ein hohes Fähigkeitsselbstkonzept auf, Mädchen mit Migrationshintergrund und niedrigem sozialen Status zeigen die größte Leistungsangst. Die Beiträge des Symposiums zeigen, dass Intersektionalitätseffekte unbedingt weiter erforscht werden sollten, damit doppelte Benachteiligungen von Schüler*innen und die daraus folgenden Disparitäten aufgedeckt werden können.

Beiträge des Symposiums

Die Rolle von Geschlecht und sozio-ökonomischen Status von Schüler*innen für die Beurteilung der Leistung in Mathe und Deutsch durch Lehramtsstudierende

Sabine Glock¹, Hannah Kleen²

¹Bergische Universität Wuppertal, Deutschland, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main

Jungen sind gut Mathe und Mädchen sind gut in Deutsch. Schüler*innen aus sozial schwachen Familien sind schlechter in der Schule als Schüler*innen aus Familien mit hohem sozio-ökonomischen Status. Diese stereotypen Aussagen sind sowohl durch Large-Scale Studien wie PISA (Lewalter et al., 2023) und IGLU (Hußmann et al., 2017) empirisch gut gesichert und auch Studien, die die Beurteilung von (angehenden) Lehrpersonen untersuchen, zeigen dies (Auwarter & Aruguete, 2008; Bonefeld et al., 2021). Wenn es um die Beurteilung durch Lehrkräfte geht, spielen Stereotype von Lehrkräften eine Rolle (Glock & Kleen, 2020b). Stereotype zu Geschlecht und zum sozio-ökonomischen Hintergrund sind in Lehrkrafturteilen mehrfach nachgewiesen worden, auch in Kombination, das heißt, dass Intersektionalitätseffekte untersucht wurden (Auwarter & Aruguete, 2008). Jedoch bleibt bisher offen, wie diese beiden Schüler*innenmerkmale mit den Hauptfächern Mathematik und Deutsch im Lehrkrafturteil interagieren. Die Intersektionalität berücksichtigt (McCall, 2005), dass sozio-ökonomische Hintergrundmerkmale nicht in Isolation, sondern ihre Interdependenz betrachtet wird. Unterschiedliche Hintergrundmerkmale können in ihrer Kombination multiple Effekte ausüben, beispielsweise könnten sich Benachteiligungen kumulieren (Cole, 2009). Eine Kumulation könnte je nach Fach unterschiedlich ausfallen. Bisherige Forschung zu Stereotypen zeigt, dass Mädchen mit schlechter Mathematikleistung, aber guter Leseleistung assoziiert werden (Muntoni et al., 2019; Muntoni & Retelsdorf, 2018). Schüler*innen mit hohem sozio-ökonomischen Status werden besser bewertet als Schüler*innen mit niedrigem sozio-ökonomischen Status, auch wenn die Leistung konstant gehalten wird (Glock & Kleen, 2020a). Entsprechend der Intersektionalität in Interaktion mit den Fächern könnte daraus resultieren, dass beispielsweise Mädchen mit niedrigem sozio-ökonomischen Status in Mathe doppelt benachteiligt nicht. Entsprechend gehen wir der Frage nach, inwiefern das Geschlecht, der sozio-ökonomische Status und die Fächer in den Urteilen von Lehrkräften interagieren.

Dieser Fragestellung wurde in einer experimentellen Vignettenstudie nachgegangen. Insgesamt haben 208 Lehramtsstudierende (151 weiblich, 55 männlich, 2 ohne Angabe; Mage = 26.31, SDage = 4.12) aller Schulformen teilgenommen. Die Proband*innen lasen entweder eine Vignette einer Schülerin oder eines Schülers, deren Vater entweder am Fließband oder als Rechtsanwalt arbeitete. Die Beurteilung der Leistung in Mathematik und in Deutsch wurde auf einer 7-Punkt Likertskala mit 1 (gering ausgeprägt) bis 7 (hoch ausgeprägt) erfasst. Eine 2 (Geschlecht: männlich vs. weiblich) x 2 (sozio-ökonomischer Status: hoch vs. niedrig) x 2 (Fach: Mathematik vs. Deutsch) mixed ANOVA mit Messwiederholung auf dem letzten Faktor ergab eine Zweifachinteraktion zwischen Geschlecht und Fach, $F(1,196) = 4.74$, $\eta^2_p = 0.02$, $p < .05$. Jungen wurden in Mathematik ($M = 3.79$; $SD = 0.94$) besser eingeschätzt als in Deutsch ($M = 3.48$, $SD = 1.05$) und besser als Mädchen in Mathematik ($M = 3.44$, $SD = 0.88$). Für Mädchen konnte hier kein Unterschied zu der Beurteilung in Deutsch ($M = 3.47$, $SD = 1.03$) gefunden werden. Auch die Zweifachinteraktion zwischen sozio-ökonomischen Status und Fach wurde signifikant, $F(1,196) = 4.03$, $\eta^2_p = 0.02$, p

< .05. Während bei Schüler*innen mit hohem sozio-ökonomischen Status kein Unterschied zwischen der Beurteilung in Deutsch ($M = 3.70$, $SD = 1.18$) und Mathematik ($M = 3.72$, $SD = 0.94$) gefunden wurde, wurden die Schüler*innen mit niedrigem sozio-ökonomischen Status schlechter in Deutsch ($M = 3.26$, $SD = 0.83$) als in Mathematik ($M = 3.62$, $SD = 0.93$) beurteilt und damit auch schlechter als die Schüler*innen mit hohem sozio-ökonomischen Status.

Die Ergebnisse zeigen, dass keine Intersektionalitätseffekte zwischen Geschlecht und sozio-ökonomischem Status gefunden wurden. Vielmehr scheint es sich um Salienzeffekte zu handeln. In Mathematik scheint das Geschlecht eine Rolle zu spielen, während in Deutsch der sozio-ökonomische Status stärker ins Gewicht fällt. Dies könnte dazu führen, dass Mädchen in Mathematik benachteiligt sind, weil sie Mädchen sind, und in Deutsch, weil sie aus einem niedrigen sozio-ökonomischen Umfeld stammen. Die Studie liefert erste Hinweise auf diese möglichen Effekte, die zukünftig stärker beachtet werden sollten.

Das Zusammenspiel von Schüler*innenmerkmalen: Migrationshintergrund und Geschlecht

Meike Bonefeld¹, Hannah Kleen², Sabine Glock³

¹Albert-Ludwigs-Universität Freiburg, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt am Main, ³Bergische Universität Wuppertal

Stereotype beziehen sich auf unterschiedlichen Schüler*innenmerkmalen, wie zum Beispiel Geschlecht oder Migrationshintergrund. Diese können im Beurteilungsprozess eine Rolle spielen und zu differenziellen Urteilen führen (Fiske & Neuberg, 1990). Diese differenziellen Urteile konnten bereits für mehrere Schüler*innenmerkmale gezeigt werden, wobei diese jeweils isoliert betrachtet wurden (Holder & Kessels, 2017; Parks & Kennedy, 2007).

Schüler*innen weisen allerdings simultan diverse Merkmale auf, die Lehrkräfte in ihren Urteilen beeinflussen können. So sind Schüler*innen männlich oder weiblich und können zusätzlich einen Migrationshintergrund aufweisen. Bei der Betrachtung von Urteilen über Schüler*innen mit Migrationshintergrund wurde bisher allerdings das Zusammenspiel verschiedener Merkmale weitgehend außer Acht gelassen. Beispielsweise wurden geschlechtsspezifische Unterschiede nicht systematisch im Zusammenhang mit der Benachteiligung von Schüler*innen mit Migrationshintergrund untersucht. So zeigen Studien zur Beurteilung von Schülerinnen (Kleen & Glock, 2018) und Schülern (Glock, 2016) mit Migrationshintergrund eine Benachteiligung im Urteil über die Sprachkompetenz, betrachten jedoch nicht das Geschlecht. Gerade wenn es um geschlechterstereotype Beurteilungsdimensionen, wie Sprache oder Mathematik geht, kann das Geschlecht eine Rolle spielen. So zeigte sich, dass Mädchen in Mathematik bei gleicher Leistung schlechter bewertet werden als Jungen (Holder & Kessels, 2017; Steffens, Jelenec, & Noack, 2010). Auch Schüler*innen mit Migrationshintergrund werden im Vergleich zu Schüler*innen ohne Migrationshintergrund schlechter in Mathematik bewertet (Bonefeld, Dickhäuser, Janke, Praetorius & Dresel, 2017). Bei Schüler*innen mit Migrationshintergrund kann das Geschlecht eine zusätzliche Rolle spielen, so dass es sein kann, dass beispielsweise Schüler mit Migrationshintergrund in ihren sprachlichen Kompetenzen doppelt—aufgrund des Migrationshintergrunds und zusätzlich des Geschlechts—benachteiligt sind.

Ziel dieser Studie war es deshalb, das Zusammenspiel der Merkmale Migrationshintergrund und Geschlecht der Schüler*innen im Hinblick auf die Beurteilung von Lehrkräften in verschiedenen Beurteilungsdimensionen (sprachliche vs. mathematische Kompetenzen) zu untersuchen.

An dieser Studie haben insgesamt 118 Lehrkräfte (93 weiblich) teilgenommen. Es wurden Schüler*innenvignetten genutzt, die einen Schüler oder eine Schülerin beschrieben, die unterdurchschnittliche Schulleistungen und schulisches Engagement zeigten. Dabei wurde sowohl die Ethnie als auch das Geschlecht variiert. Die Lehrkräfte schätzten die sprachlichen und die mathematischen Fähigkeiten ein. Die Studie hatte somit ein 2 (Schüler*innengeschlecht: weiblich vs. männlich) x 2 (Migrationshintergrund der Schüler*innen) x 2 (Beurteilungsdimension: Sprache vs. Mathe) mixed Design mit Messwiederholung auf dem letzten Faktor. Die Ergebnisse zeigen einen signifikanten Haupteffekt des Migrationshintergrunds, so dass Schüler*innen mit Migrationshintergrund schlechter beurteilt werden als Schüler*innen ohne Migrationshintergrund. Darüber hinaus ist der Haupteffekt der Beurteilungsdimension signifikant. Lehrkräfte schätzten die sprachlichen Fähigkeiten aller Schüler*innen als niedriger ein als ihre mathematischen Fähigkeiten. Ebenso gab es einen Interaktionseffekt zwischen Schüler*innengeschlecht und Beurteilungsdimension. Schülerinnen werden bezüglich ihrer mathematischen Kompetenzen niedriger eingeschätzt als Schüler, was sich bezüglich der sprachlichen Kompetenzen nicht zeigte. Gleichzeitig werden Schüler in ihren mathematischen Kompetenzen höher eingeschätzt als in ihren sprachlichen, was sich für Schülerinnen nicht zeigte. Auch zwischen Migrationshintergrund der Schüler*innen und Beurteilungsdimension gab es eine signifikante Interaktion. Schüler*innen mit Migrationshintergrund wurden bezüglich der sprachlichen Fähigkeiten als weniger kompetent eingeschätzt als Schüler*innen ohne Migrationshintergrund, bei den mathematischen Fähigkeiten zeigte sich kein Unterschied.

Die Ergebnisse bestätigen die bekannten Forschungsbefunde, dass Schüler*innen mit Migrationshintergrund schlechter bewertet werden als Schüler*innen ohne Migrationshintergrund und dass es bezüglich des Geschlechts zu differenziellen Urteilen in der Sprache und der Mathematik kommt. Schüler*innen mit Migrationshintergrund sind somit einerseits benachteiligt, weil sie der Gruppe der Migrant*innen angehören und andererseits, genau wie deutsche Schüler*innen, weil sie männlich oder weiblich sind. Auch wenn aufgrund der nicht signifikanten dreifachen Interaktion zwischen Geschlecht, Migrationshintergrund und Beurteilungsdimension nicht von einer additiven Benachteiligung von Schüler*innen mit Migrationshintergrund ausgegangen werden kann, so besitzen sie doch ein Merkmal mehr, welches zu verzerrten Urteilen von Lehrkräften führen kann. Zukünftige Forschung sollte sich daher mit der Reduktion von stereotypen Urteilen beschäftigen.

Schulübergang in Luxemburg: Intersektionalität von Geschlecht und Migrationshintergrund

Ineke Pit-ten Cate¹, Martha Ottenbacher¹, Aigul Alieva², Taylor Kroezen², Andreas Hadjar³, Antoine Fischbach¹

¹LUCET Universität Luxemburg, ²Luxembourg Institute of Socio-Economic Research, ³Universität Freiburg, Schweiz

Soziodemografische Merkmale von Schüler*innen stehen mit der Bevorzugung und Benachteiligung bestimmter Gruppen von Schüler*innen in Verbindung. Bisher haben sich die meisten Forschungsarbeiten in gegliederten Bildungssystemen—wie in Luxemburg—auf einzelne Merkmale der Schüler*innen, vor allem auf das Geschlecht und den Migrationshintergrund, konzentriert. Allerdings ist wenig über die Auswirkungen der Überschneidung verschiedener Merkmale bekannt. Theorien der Intersektionalität gehen davon aus, dass jede Person mehrere soziale Kategorien repräsentiert und dass die Kombination dieser Kategorien zu unterschiedlichen Erfahrungen und Ergebnissen für den Einzelnen führen kann (Silverstein, 2006), was auch durch empirische Studien bestätigt wurde (Ghavami & Peplau, 2012). Die aktuelle Studie erweitert die bestehende Literatur, indem sie individuelle Merkmale sowie deren Interaktion bei der Vorhersage des Schulübergangs von der Grundschule in die Sekundarstufe berücksichtigt.

In der vorliegenden Studie wurde der Einfluss verschiedener Schüler*innenmerkmale auf die Schullaufbahn untersucht. In Luxemburg spiegelt die Schullaufbahn fast zu 100% die Schullaufbahnpflicht wider, da diese für Eltern bindend sind. Wir

interessierten uns insbesondere für die prädiktiven unabhängigen und interaktiven Zusammenhänge zwischen Schüler*innenmerkmalen in Klasse 3 und dem Schulzweig in Klasse 9.

Die Daten wurden im Rahmen des luxemburgischen Schulmonitoringprogramms „Épreuves Standardisées“ (ÉpStan; Fischbach et al., 2014) erhoben. Die Kohorte umfasst alle Schüler*innen, die im November 2013 in der 3. Klasse des öffentlichen luxemburgischen Bildungssystems eingeschrieben waren, kombiniert mit Daten derselben Schüler*innen in der 9. Klasse im November 2017-2019 für Schüler*innen, die einen regulären Bildungsweg verfolgt haben. Diese werden ergänzt durch Daten von November 2020 und 2021 für Schüler*innen, die ein- oder zweimal Klassen wiederholt haben (N=3600). In der vorliegenden Studie wurden soziodemografische Merkmale wie Geschlecht, sozioökonomischer Status, zu Hause gesprochene Sprache und Migrationshintergrund sowie Leistungsdaten in Mathematik und Deutsch in der dritten Klasse zur Vorhersage des Schulzweigs in der neunten Klasse verwendet.

Um den Einfluss von Schülermerkmalen auf den Schulzweig zu untersuchen, haben wir eine mehrstufige multinomiale Regressionsanalyse mit random intercepts durchgeführt. Die Ergebnisse dieser Analysen zeigen, dass die Einstufung in die Schulzweige durch individuelle Merkmale beeinflusst wird. Der Schulzweig in der 9. Klasse wird durch das Geschlecht, den SES, den Migrationshintergrund und die bisherigen Leistungen der Schüler*innen beeinflusst. Jungen haben eine deutlich geringere Wahrscheinlichkeit, in den höchsten und eine höhere Wahrscheinlichkeit, in den niedrigsten Schulzweig orientiert zu werden, als Mädchen. Schüler*innen mit Migrationshintergrund oder aus Familien mit niedrigem Sozialstatus haben eine deutlich geringere Chance, dem höchsten Bildungszweig zugewiesen zu werden. Im Gegensatz dazu scheint die zu Hause gesprochene Sprache keinen Einfluss auf den Schulübergang zu haben.

Die Ergebnisse der intersektionalen Analyse zeigen, dass die meisten Interaktionen auf individueller Ebene nicht signifikant sind. Nur die Interaktion zwischen Geschlecht und Migrationshintergrund ist signifikant, aber die individuellen Effekte verlieren in diesem Modell ihre Bedeutung. Diese Ergebnisse spiegeln wider, dass die Haupteffekte für Geschlecht und Migration durch ihre Interaktion vermittelt werden und dass—im Vergleich zu den anderen Untergruppen—männliche und weibliche Schüler ohne Migrationshintergrund eine ähnliche Wahrscheinlichkeit haben, dem niedrigsten Schulzweig zugewiesen zu werden, während sowohl Mädchen der ersten als auch der zweiten Generation eine geringere Wahrscheinlichkeit haben, dem niedrigsten Schulzweig zugewiesen zu werden als alle anderen Untergruppen.

Im Einklang mit früheren Untersuchungen (Encinas-Martín und Cherian, 2023; Klapproth et al., 2013) sind die soziodemografischen Merkmale der Schüler*innen signifikante Prädiktoren des Schulübergangs. Das Vorhandensein signifikanter Effekte einzelner Hintergrundprädiktoren bei gleichzeitiger Kontrolle der Schülerleistungen deutet darauf hin, dass Schulübergangsentscheidungen potenziell zuungunsten benachteiligter Schüler*innen, nämlich der Schüler*innen mit Migrationshintergrund, männliche Schüler und Schüler*innen mit niedrigem SES verzerrt sein könnten (siehe Baumert et al., 2006; Timmermans, de Boer, et al., 2018). Die Ergebnisse der intersektionalen Analysen ergaben kein klares Interaktionsmuster. Für die meisten Schüler*innenmerkmale gibt es unabhängige Haupteffekte auf den Schulübergang, mit Ausnahme der Interaktion zwischen Geschlecht und Migrationshintergrund.

Unterscheiden sich Motivation und emotionales Erleben von Berufsschülerinnen und -schülern in Abhängigkeit ihrer Herkunft und ihres Geschlechts?

Anita Tobisch, Markus Dresel
Universität Augsburg

Die Erwartungs-Wert-Theorie der Leistungsmotivation (Wigfield & Eccles, 2000) und die Kontroll-Wert-Theorie der Leistungsemotionen (Pekrun, 2000) postulieren, dass Motivation und Emotionen entscheidende Faktoren für Lern- und Leistungsverhalten von Schülerinnen und Schülern sind, was bereits vielfach empirisch gestützt werden konnte. Diese Theorien legen weiterhin nahe, dass Sozialisationseinflüsse (z.B. der sozio-kulturelle Kontext) einen bedeutsamen Einfluss auf die Ausprägungen der Motivation und Emotionen von Schülerinnen und Schülern haben. Es ist daher davon auszugehen, dass sich Schülerinnen und Schüler hinsichtlich bestimmter Merkmale – wie ihrem Migrationshintergrund, ihrer Familiensprache, ihrem Geschlecht oder ihrer sozialer Herkunft – in ihrer Motivation und ihrem emotionalem Erleben unterscheiden können. Dies könnte auch eine mögliche Erklärung für die bestehenden Disparitäten im Bildungs- und Berufserfolg nach Herkunft und Geschlecht der Schülerinnen und Schüler sein. Erste empirische Studien aus dem Bereich von Grundschule und Sekundarstufe I konnten bereits zeigen, dass Schülerinnen und Schüler aus niedrigeren sozialen Schichten und insbesondere Schülerinnen und Schüler mit Migrationshintergrund ungünstigeres emotionales Erleben (z.B. Leistungsangst), aber dennoch eine hohe Motivation (z.B. Lernzielorientierung) aufweisen (z.B. Brandenberger, Hagenauer, Hascher, 2017; Tobisch, Kopp, Martschinke, Kröner & Dresel, 2016). Offen bleibt bislang allerdings, ob dies auch für spätere Bildungsphasen gilt und ob sich aus Perspektive des Intersektionalitätsansatzes bestimmte Herkunft-Geschlecht-Kombinationen identifizieren lassen, die besonders von negativem emotionalen Erleben oder ungünstiger Motivation betroffen sind. Diesen Fragen wird mit den Daten einer Studie mit N = 1382 Berufsschülerinnen und -schülern nachgegangen. Untersucht wurde dabei der Zusammenhang zwischen soziostrukturellen Hintergrundmerkmalen (Geschlecht, sozialer Status der Eltern, Migrationshintergrund, Familiensprache) sowie Motivation (subjektiver Wert, Fähigkeitsselbstkonzept, Anstrengungsbereitschaft, Zielorientierungen) und Emotionen (Lernfreude, Langeweile, Hilflosigkeitserleben, Angst, Wohlbefinden) von Berufsschülerinnen und -schülern. Die Ergebnisse deuten insgesamt auf positive Ausprägungen von Motivation und Emotion in der Schülerschaft hin. Gleichzeitig zeigten sich bedeutsame Unterschiede zwischen Berufsschülerinnen und -schüler unterschiedlicher Herkunft und unterschiedlichen Geschlechts. So wiesen Schülerinnen und Schüler mit Migrationshintergrund etwa ein durchschnittlich geringeres Fähigkeitsselbstkonzept auf. Schülerinnen und Schüler mit niedrigerem sozialem Status berichteten z.B. höhere Angst und stärkeres Hilflosigkeitserleben. Weibliche Schülerinnen gaben zudem geringeres Wohlbefinden an. Mit Blick auf die Verschränkung der unterschiedlichen Heterogenitätsdimensionen zeigten sich ebenfalls bedeutsame Effekte. So ergab sich u.a. eine Dreifach-Interaktion von Geschlecht, sozialem Status und Migrationshintergrund für das emotionale Erleben (Hilflosigkeitserleben und Wohlbefinden) sowie motivationale Konstrukte (intrinsischer Wert und Fähigkeitsselbstkonzept). Demnach weisen insbesondere männliche Schüler aus Familien mit hohem sozialen Status und ohne Migrationshintergrund ein besonders positives Fähigkeitsselbstkonzept auf. Weibliche Schülerinnen mit Migrationshintergrund aus Familien mit niedrigem sozialen Status gaben hingegen die höchste Leistungsangst an. Die Befunde deuten zum einen auf die Bedeutung von Herkunftsmerkmalen für Motivation und emotionales Erleben hin. Zum anderen zeigen sie, dass es eine bedeutsame Verschränkung der Herkunft und des Geschlechts von Berufsschülerinnen und -schülern in Bezug auf ihre Motivation und ihr emotionales Erleben gibt.

S-60: Veränderungen beruflicher Orientierungen im Kontext struktureller Herausforderungen - Die Rolle sozialer, regionaler und fluchtspezifischer Faktoren

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: M 003

Chair der Sitzung: Alexandra Wicht

Chair der Sitzung: Oliver Winkler

Chair der Sitzung: Helen Hickmann

Veränderungen beruflicher Orientierungen im Kontext struktureller Herausforderungen - Die Rolle sozialer, regionaler und fluchtspezifischer Faktoren

Chair(s): **Alexandra Wicht** (Universität Siegen, Deutschland), **Oliver Winkler** (Martin-Luther-Universität Halle-Wittenberg, Institut für Soziologie)

Diskutant*in(nen): **Helen Hickmann** (Bundesinstitut für Berufsbildung, Bonn)

Die berufliche Orientierung von Jugendlichen, d.h. ihre beruflichen Aspirationen und Entscheidungen, sind ausschlaggebend für einen erfolgreichen Ausbildungsübergang und somit ein zentraler Schlüssel zur Integration Jugendlicher in den Arbeitsmarkt und für ihren späteren Lebensverlauf (Schoon, 2021). Vielen Jugendlichen gelingt es jedoch nicht, ihre beruflichen Aspirationen umzusetzen (Nießen et al., 2022). Auch auf dem Ausbildungs- und Arbeitsmarkt zeigen sich zunehmende Passungsprobleme und ein akuter Fachkräftemangel (Oeynhausen et al., 2023). Um Zugänge zur Ausbildung zu verbessern, ist es entscheidend, den beruflichen Orientierungsprozess von Jugendlichen und die Bedeutung struktureller Kontexteffekte genauer zu betrachten.

Aus einer interdisziplinären Perspektive fokussiert das Symposium daher verschiedene Phasen der beruflichen Orientierung und geht der Frage nach, welche Rolle soziale, regionale, und fluchtspezifische Faktoren für die Formation, Anpassung und Umsetzung von beruflichen Aspirationen spielen.

Nach der Circumscription & Compromise Theorie (Gottfredson, 1981) stellt die berufliche Orientierung einen Prozess dar, über den Jugendliche ihre beruflichen Aspirationen inkrementell eingrenzen und anpassen. Hierbei spielen verschiedene Faktoren wie signifikante Andere, familiäre Ressource und Opportunitäten der räumlichen (Lern-)Umwelten, in die junge Menschen eingebunden sind, eine entscheidende Rolle für ihre beruflichen Aspirationen (Bronfenbrenner, 1979; Sewell et al., 1969). Insbesondere bei der Formation beruflicher Aspirationen greifen Jugendliche auf soziale Rollenmodelle zurück und grenzen ihre Aspirationen hinsichtlich der Passung ihres eigenen Selbstbildes und des wahrgenommenen Prestige eines Berufs sowie den Erwartungen des sozialen Umfelds ein. Für die Umsetzung beruflicher Aspirationen sind zudem die tatsächlichen Gelegenheitsstrukturen entscheidend, also welche Möglichkeiten und Hindernisse für die Jugendlichen bestehen. Hierbei gehen Jugendliche oftmals weitere Kompromisse ein und müssen ihre Aspirationen an die gegebenen Gelegenheitsstrukturen anpassen oder Strategien finden um ihre Möglichkeiten zu erweitern.

Die ersten beiden Symposiumsbeiträge nehmen die Formation und Veränderung beruflicher Aspirationen von Jugendlichen in den Fokus. Der erste Beitrag betrachtet anhand einer quasi-experimentellen Interventionsstudie, inwieweit langfristige und kurzfristige soziale Einflüsse die Aspirationen von Schüler:innen für einen Ausbildungsberuf erhöhen können und untersucht Wechselwirkungen zwischen kurzfristigen Interventionsmaßnahmen und dem langfristigen familiären und schulischen Kontext.

Der zweite Beitrag fokussiert die berufliche Orientierung junger Geflüchteter ohne Berufsabschluss nach Ankunft in Deutschland. Anhand qualitativer leitfadengestützter Interviews werden die Bedeutung struktureller Herausforderungen für die Anpassung beruflicher Aspirationen untersucht sowie differenzierte Handlungsstrategien in Abhängigkeit fluchtspezifischer Einflussfaktoren aufgedeckt.

Im dritten Beitrag werden Veränderungen der Mobilitätsbereitschaft als mögliche Anpassungsstrategie von unversorgten Jugendlichen bei der Ausbildungsplatzsuche betrachtet. Basierend auf repräsentativen Längsschnittdaten zeigt sich eine Erhöhung der Mobilitätsbereitschaft über die Zeit, sowie die Bedeutung individueller beruflicher Aspirationen und regionaler Gelegenheitsstrukturen.

Beiträge des Symposiums

Berufliche Rollenmodelle und ihr Einfluss auf die Bildungsaspirationen von Jugendlichen – welche Rolle spielen soziale Kontexte in Schule und Familie?

Eva Böhle, Janina Beckmann, Mona Granato

Bundesinstitut für Berufsbildung, Bonn

Theoretischer Hintergrund

Die Entwicklung von Bildungs- und Berufsaspirationen ist für Jugendliche ein wichtiger Schritt im Übergang Schule-Beruf (Ziegler, 2023). Zum Ende der Schulzeit haben Jugendliche die Möglichkeit, zwischen verschiedenen Bildungs- und Berufswegen zu entscheiden, um einen berufsqualifizierenden Abschluss zu erreichen. Dabei zeigt sich ein anhaltend geringes Interesse junger Menschen an Berufen des dualen Ausbildungssystems, welches sich unter anderem in Passungsproblemen auf dem Ausbildungsmarkt widerspiegelt (Oeynhausen et al., 2023). Gleichzeitig ist festzustellen, dass viele junge Menschen ihren ersten gewählten Berufs- und Bildungsweg nachträglich korrigieren. So brechen ca. 25% aller Studierenden das Studium vorzeitig ab und ein Großteil mündet anschließend in eine berufliche Ausbildung ein (Heublein et al., 2018). Angesichts des Mangels an Nachwuchskräften auf Fachkräfteebene (Bundesagentur für Arbeit, 2024), stellt sich zusätzlich die Frage, inwieweit geeignete Interventionen Aspirationen für eine berufliche Ausbildung erhöhen könnten. Ein möglicher Ansatz, der in den Vordergrund des bildungspolitischen Interesses gerückt ist, stellen berufliche Rollenmodelle dar.

Die Entwicklung von Bildungs- und Berufsaspirationen ist in verschiedene soziale Kontexte eingebettet, die nicht nur einen institutionellen Rahmen bieten, sondern auch eine Sozialisationsfunktion aufweisen, indem soziale Erwartungen an den eigenen Berufsweg durch signifikante Andere in Schule und Familie formuliert werden oder diese als Vorbildfunktion für die eigenen Berufsorientierungen dienen (Buchmann & Dalton, 2002; Stocké, 2010). Ein zentraler Bestandteil sozialer Erwartungen und Normen stellt dabei das wahrgenommene Prestige beruflicher Optionen dar (Gottfredson, 1981). Es stellt sich daher die Frage, inwieweit gegenüber diesen langfristig wirkenden sozialen Kontexten in Schule und Familie, eine kurzfristige Intervention mit Rollenmodellen die beruflichen Bildungsaspirationen von Jugendlichen verändern kann. Eine theoretische Grundlage dafür bietet die Happenstance Learning Theory (Krumboltz, 2009; Mitchell et al., 1999), die Einflüsse durch zufällige und ungeplante Ereignisse auf die Berufswahl betrachtet.

Fragestellung

Interventionen zur Veränderung von Bildungsaspirationen zielen in der Beruflichen Orientierung häufig darauf, Bildungsungleichheiten am Übergang in die Hochschule abzubauen (Herbaut & Geven, 2020). Wenngleich die Erhöhung von Aspirationen für eine Berufsausbildung gleichermaßen als bildungspolitisches Ziel formuliert wird (vgl. Bundesministerium für Arbeit und Soziales, 2022), liegen bislang kaum Befunde über die Wirksamkeit berufsorientierender Angebote mit diesem Ziel vor. Die vorliegende Interventionsstudie setzt mit Hilfe einer Intervention, die zum Ziel hat, das Interesse an beruflicher Ausbildung durch den Einsatz beruflicher Rollenmodelle an Schulen zu erhöhen, an dieser Forschungslücke an. Dabei stellt sich insbesondere die Frage, inwieweit dieses Ziel auch für Jugendliche erreicht werden kann, die in ihrer sozialen Umgebung eine geringe soziale Unterstützung für eine berufliche Ausbildung erfahren, beispielsweise Jugendliche aus akademischen Elternhäusern. Im Fokus der Studie stehen die Fragen:

1. Inwieweit können Rollenmodelle im Schulkontext dazu beitragen, die Aspirationen von Schüler:innen für einen Ausbildungsberuf zu erhöhen?
2. Inwieweit bestehen Wechselwirkungen zwischen der Intervention und bestehenden sozialen Kontexten, die den Interventionseffekt verstärken oder abschwächen könnten?

Methode

Empirische Basis der Analysen ist eine quasi-experimentelle Interventionsstudie mit ca. 2.000 Schüler:innen an allgemeinbildenden Schulen in Nordrhein-Westfalen. Anhand von hierarchischen Regressionsmodellen wird zunächst der kausale Zusammenhang zwischen dem Besuch der beruflichen Rollenmodelle und der Aspiration für den vorgestellten Ausbildungsberuf 4-8 Wochen nach der Intervention untersucht. Im nächsten Schritt wird analysiert, inwieweit die Kontextfaktoren Familie und Schule die Wirkung der Intervention moderieren. Dabei stehen sowohl deskriptive Normen (z.B. der Berufsabschluss der Eltern und Berufsaspirationen der Peers) als auch präskriptive Normen (z.B. die sozialen Erwartungen von Eltern und Peers) im Vordergrund.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse zeigen, dass berufliche Rollenmodelle die Berufsaspirationen für die vorgestellten Ausbildungsberufe im Mittel erhöhen. Weiterführende Analysen werden zeigen, wie sich dieser Effekt für Jugendliche in unterschiedlichen familiären und schulischen Kontexten unterscheidet. Berufliche Rollenmodelle könnten somit ein wirksames Angebot sein, die Berufsaspirationen für bestimmte Ausbildungsberufe zu erhöhen. Die Effektstärke wird abschließend diskutiert.

„Und dann habe ich gedacht, dass ich einen zweiten Weg finden muss“: Berufliche Orientierungen junger Geflüchteter nach Ankunft in Deutschland

Hannah Glinka

Martin-Luther-Universität Halle-Wittenberg, Institut für Soziologie

Theoretischer Hintergrund

Gottfredson (1981) geht von einem Selbstkonzept (self-concept) aus, welches die Wahrnehmung von Akteuren über sich selbst beschreibt. Das Selbstkonzept wird von verschiedenen Elementen wie Geschlecht, sozialem Status sowie Interessen und Fähigkeiten, in Abhängigkeit der kognitiven Entwicklung und damit verknüpft auch dem Alters beeinflusst. Berufe müssen mit Selbstkonzepten übereinstimmen, um in Betracht gezogen zu werden und weiter mit strukturellen Faktoren des Kontexts abgeglichen werden. Dadurch wird ersichtlich, ob berufliche Ziele unter Berücksichtigung der sozialen und ökonomischen Umwelt umsetzbar sind. Berufliche Orientierung (BO) kann somit als Anpassungsprozess an den kulturellen, sozialen und ökonomischen Kontext verstanden werden, was Flexibilität und somit die Reflektion der beruflichen Identität erfordert (Savickas, 2013). Ferner können gesellschaftliche Agenten BO beeinflussen (z.B. Lehrkräfte, Eltern) und so als Gatekeeper fungieren oder Möglichkeiten eröffnen (Daheim, 1967). BO ist somit eine Integrationsleistung und ein Kompromiss von subjektiver Lebenswelt, die eingebettet in sozialer Umwelt konstruiert wird, und kontextspezifischer, objektiver Arbeitswelt (vgl. Held et al., 2015). Ein Kontextwechsel wie im Falle von Geflüchteten bedeutet, dass BO und deren Realisierungsmöglichkeiten unter neuen Voraussetzungen der Umwelt evaluiert werden und ggf. unter Wahrung des Selbstkonzepts Kompromisse eingegangen werden müssen, wodurch sich die BO entwickeln kann.

BO, die an die strukturellen Anforderungen des Kontexts angepasst ist, ist der Schlüssel zu einer langfristigen und nachhaltigen Arbeitsmarktintegration (vgl. Frintrup & Spengler, 2018). Für Deutschland ist die Relevanz der BO junger Geflüchteter durch die junge Altersstruktur der Schutzsuchenden (Statistisches Bundesamt, 2024) zusätzlich erhöht, da in dieser Lebensphase der Grundstein für die weitere berufliche Entwicklung gelegt wird (vgl. Gottfredson, 1981).

Fragestellung

Bisher ist unbekannt, wie die BO junger Geflüchteter nach Ankunft in Deutschland verläuft, auf welche Handlungsprobleme sie stoßen und wie sie damit umgehen. Die Forschungsfrage lautet daher: Wie entwickeln sich berufliche Orientierungen junger Geflüchteter nach Ankunft in Deutschland?

Methode

In qualitativen, leitfadengestützten Interviews wurden junge Geflüchtete zwischen 18 und 30 Jahren in Deutschland befragt, die noch keinen Berufsabschluss haben. Die Analyse der Interviews erfolgt in Anlehnung an biografische Fallrekonstruktionen (Rosenthal, 1995), da die Entwicklung der BO als Prozess zu verstehen ist, der in biografische Ereignisse eingebettet ist (vgl. Rosenthal, 2015).

Ergebnisse

Für Fall Carlos zeigt sich, dass die Umsetzung von im Herkunftskontext ausgebildeten Berufszielen im Aufnahmekontext herausfordernd ist. Dies ist insbesondere auf die Entwertung von kulturellen Ressourcen wie Sprache und Studienkompetenzen zurückzuführen, die beim Kontextwechsel eintritt. Die strukturellen Anforderungen des Aufnahmekontexts erfordern eine Anpassung der ursprünglichen Berufsziele, orientiert an jenem Kontext, um erfolgreich umsetzbar zu sein. Fall Ilyas bildet seine Berufsziele in Deutschland aus. Die Entwicklungsaufgabe der BO ist jedoch vor dem Hintergrund fluchtbedingter Herausforderungen (z.B. Deutscherwerb, Schulunterbrechungen, Umzüge und damit einhergehende schulische Schwierigkeiten) erschwert, da diese die BO überlagern können. Die Einschätzung signifikanter Anderer bzgl. Bildungs- und Berufszielen wird aufgrund dieser komplexen Ausgangssituationen besonders wichtig im BO-Prozess.

In beiden Fällen treten ähnliche Handlungsprobleme auf, die die BO beeinflussen, z.B. Ausbildung bzw. Umsetzung von Berufszielen, eine subjektiv eingeschränkte Agency und Altersnormen. Auffällig ist jedoch, dass sich die Strategien, um den

Handlungsproblemen zu begegnen, unterscheiden und mögliche Erklärungen dafür im unterschiedlichen Einreisearchter der Geflüchteten zu finden sind. Ilyas scheint in der Entwicklung von Strategien von der Vertrautheit mit dem Bildungssystem des Aufnahmekontexts zu profitieren und richtet seine Orientierungen stärker am Aufnahmekontext aus, während Carlos stärker auf Erfahrungen im Herkunftskontext zurückgreift.

Aus den Ergebnissen lassen sich Rückschlüsse auf den Unterstützungsbedarf in der BO von jungen Geflüchteten ziehen. Eine Differenzierung nach (Einreise-)Alter der Geflüchteten ist zu betonen, da damit unterschiedliche Herausforderungen einhergehen. Fluchtspezifische Herausforderungen müssen bei der Beratung berücksichtigt werden ebenso wie bereits vorhandene Kompetenzen aus dem Herkunftskontext und wie diese unter neuen kontextuellen Bedingungen eingebracht werden können, insbesondere vor dem Hintergrund konstanter Berufsziele.

Die Rolle beruflicher Aspirationen und regionaler Gelegenheitsstrukturen für Anpassungen der Mobilitätsbereitschaft von unversorgten Ausbildungsplatzsuchenden in Deutschland

Linda Hoffmann¹, Alexandra Wicht²

¹Bundesinstitut für Berufsbildung, Bonn, ²Universität Siegen, Deutschland

Theoretischer Hintergrund

Vielen Jugendlichen gelingt kein unmittelbarer Übergang in die Berufsausbildung, was folgenreich für ihren weiteren Lebensverlauf sein kann (Brinsky-Fay, 2022). Auf individueller Ebene stellen berufliche Aspirationen einen wichtigen Gelingensfaktor für erfolgreiche Übergänge dar (Nießen et al., 2022). Zudem hängt die Chance von Jugendlichen ihre beruflichen Aspirationen zu realisieren von regionalen Gelegenheitsstrukturen ab. So ist der Ausbildungsmarkt in Deutschland geprägt von starken regionalen Disparitäten, die zusätzliche Hindernisse für Jugendliche aus ungünstigen Regionen darstellen (Wicht & Nonnenmacher, 2017).

In Forschung und Praxis wird räumliche Mobilität als Lösung angesehen um regionale Disparitäten auszugleichen und erfolgreiche Ausbildungsübergänge zu erleichtern (Schmidt, 2023). Empirisch ist belegt, dass ein beachtlicher Anteil Jugendlicher für die Ausbildung mobil wird, insbesondere wenn sie höhere Statusaspirationen und längere Suchdauern haben, sowie aus ungünstigen Regionen kommen (Hoffmann & Wicht, 2023). Wie sich die Mobilitätsbereitschaft über die Zeit der Ausbildungsplatzsuche entwickelt und welche Rolle hierbei individuelle Aspirationen und regionale Gelegenheitsstrukturen spielen ist jedoch noch unklar.

Nach berufspsychologischen Ansätzen wie der Circumscription and Compromise Theory (Gottfredson, 1981) gleichen Jugendliche während der Ausbildungsplatzsuche ihre beruflichen Aspirationen mit den tatsächlichen Opportunitäten ab. Die Agency-Structure Debatte (Schoon & Heckhausen, 2019) betont zudem das Zusammenspiel individueller und struktureller Faktoren für Anpassungsprozesse Jugendlicher um ihre Ziele zu erreichen. Studien zeigen, dass ausbildungsplatzsuchende Jugendliche ihre Suchstrategien anpassen und Kompromisse eingehen (Beckmann et al., 2023; Scheels & Abraham, 2023). Eine Anpassungsstrategie die bislang wenig untersucht wurde, ist die Ausweitung der Mobilitätsbereitschaft um die Chancen auf eine passende Ausbildungsstelle zu erhöhen.

Fragestellung

Basierend auf den theoretischen Überlegungen gehen wir zwei Forschungsfragen nach: (1) Verändern unversorgte Jugendliche während der Ausbildungsplatzsuche ihre Mobilitätsbereitschaft? (2) Welche Rolle spielen hierbei individuelle Aspirationen und regionale Gelegenheitsstrukturen?

Methode

Wir verwenden repräsentative Längsschnittdaten des Nationalen Bildungspanels (NEPS, Startkohorte 4) (Blossfeld & Roßbach, 2019), die jährliche Informationen über das tatsächliche Bewerberverhalten (sozio-ökonomischer Status und Segment des aspirierten Berufs, Entfernung zur weit entferntesten beworbenen Stelle) unversorgter Jugendlicher enthalten. Zur Erfassung regionaler Gelegenheitsstrukturen werden administrative Regionaldaten angespielt, die die Verfügbarkeit des aspirierten Wunschberufs in der Heimatregion abbilden (BA, 2022). Es werden Mehrebenen-Wachstumskurvenmodelle berechnet um Veränderung der Mobilitätsbereitschaft über die Zeit zu untersuchen. Um Wechselwirkungen der individuellen Statusaspirationen und regionalen Gelegenheitsstrukturen zu analysieren werden entsprechende Interaktionsmodelle geschätzt.

Ergebnisse

Unsere Ergebnisse zeigen eine signifikante Erhöhung der Mobilitätsbereitschaft unversorgter Jugendlicher während der Ausbildungsplatzsuche. Ebenso zeigt sich die Relevanz regionaler Gelegenheitsstrukturen. Ausbildungsplatzsuchende aus ungünstigeren Regionen mit einer geringen Verfügbarkeit des aspirierten Berufs weisen eine konstant höhere Mobilitätsbereitschaft auf als Jugendliche aus günstigeren Regionen. Bezüglich der Rolle individueller Statusaspirationen zeigen Jugendliche mit höheren Statusaspirationen eine signifikant höhere Mobilitätsbereitschaft zu Beginn der Suche. Über die Zeit hinweg erhöhen jedoch Jugendliche mit geringeren Statusaspirationen ihre Mobilitätsbereitschaft deutlich.

Subgruppenanalysen unterstreichen zudem das Zusammenspiel individueller Statusaspirationen und regionaler Gelegenheitsstrukturen. Während Jugendliche mit hohen Statusaspirationen, unabhängig von Suchdauer und regionaler Gelegenheitsstrukturen, eine konstant hohe Mobilitätsbereitschaft aufweisen, zeigen sich insbesondere für Jugendliche mit niedrigen Statusaspirationen differenzierte Anpassungsprozesse. In günstigen Regionen mit einer hohen Verfügbarkeit des Wunschberufs weisen Jugendliche mit niedrigen Statusaspirationen eine deutlich geringere Mobilitätsbereitschaft zu Beginn auf, und erhöhen diese erst über die Zeit. In ungünstigen Regionen mit einer geringen Verfügbarkeit des Wunschberufs hingegen sind Jugendliche mit geringen Statusaspirationen schon von Beginn an deutlich mobilitätsbereiter.

Die Ergebnisse deuten darauf hin, dass unversorgte Jugendliche ihre Mobilitätsbereitschaft während der Ausbildungsplatzsuche ausweiten um die Chancen auf einen erfolgreichen Übergang zu erhöhen. Die detaillierte Analyse des Zusammenspiels individueller Aspirationen und regionaler Gelegenheitsstrukturen offenbart differenzierte Anpassungsprozesse, insbesondere bei Jugendlichen mit niedrigen Statusaspirationen, wenn sie keine Ausbildung vor Ort oder über längere Zeit finden. Dies impliziert relevante Hinweise für die Praxis um gezielte Maßnahmen zur Förderung räumlicher Mobilität zu implementieren um erfolgreiche Übergänge zu erleichtern sowie soziale und regionale Disparitäten abzubauen.

S-61: Educational Chatbots: Ein interdisziplinärer Blick auf aktuelle Forschungsfragen zu hybriden und generativen Chatbots in Unterricht und Lehrkräftebildung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 145

Chair der Sitzung: Steven Beyer

Chair der Sitzung: Maria Klar

Chair der Sitzung: Sebastian Becker-Genschow

Educational Chatbots: Ein interdisziplinärer Blick auf aktuelle Forschungsfragen zu hybriden und generativen Chatbots in Unterricht und Lehrkräftebildung

Chair(s): **Steven Beyer** (Universität Duisburg-Essen, Lehrstuhl für Mediendidaktik und Wissensmanagement), **Maria Klar** (Universität Duisburg-Essen, Lehrstuhl für Mediendidaktik und Wissensmanagement)

Diskutant*in(nen): **Sebastian Becker-Genschow** (Universität zu Köln, Digitale Bildung mit Schwerpunkt Künstliche Intelligenz)

Dieses Symposium widmet sich den Fragestellungen, wie Designmerkmale von Educational Chatbots in Lehr-Lern-Kontexten wirksam werden können und wie sie in Wechselbeziehung zu Inhalten und Akteuren stehen. Die Beiträge reflektieren dabei die Annahme, dass spezifische Lehr-Lern-Situationen unterschiedliche didaktische Chatbot-Konfigurationen verlangen.

Jüngste Entwicklungen um generative KI-Chatbots sorgten für eine substantielle Weiterentwicklung (z.B. Neumann et al., 2024) und erneuern die Diskussion um die Potenziale von Educational Chatbots, wie Personalisierung (z.B. Wollny et al., 2021), dialogischem Lernen (Huang et al., 2022), ständige Verfügbarkeit (Okonkwo & Ade-Ibijola, 2021), erhöhte Motivation (Zhang et al., 2023) sowie leichte Nutzung aufgrund der Vertrautheit mit Messenger-Apps (Hobert & Meyer von Wolff, 2019). Auf der Unterrichtsebene konnten Effekte auf den Lernerfolg empirisch belegt werden (Wu & Yu, 2024), hinsichtlich der Lehrkräftebildung bestehen jedoch noch deutliche Evidenzdefizite (Chiu et al., 2023; Lee & Yeo, 2022).

Chatbots lassen sich anhand verschiedener Merkmale kategorisieren, wobei eine Unterscheidung nach der Inputverarbeitung und Antwortgenerierung eine hinreichende Differenzierung bietet (Adamopoulou & Moussiades, 2020; Smutny & Schreiberova, 2020). Es gibt regelbasierte, machine-learning-basierte, generative und hybride Chatbots. Regelbasierte Chatbots geben zu vordefiniertem Input einen vordefinierten Output aus und sind daher erklärbar, aber unflexibel. Machine-learning-basierte Chatbots rufen basierend auf erkannter Absicht mögliche Antwortkandidaten ab, können Aufgaben wie Absichtserkennung sowie Stimmungsanalyse durchführen und lernen aus Konversationen. Generative Chatbots erzeugen durch große Sprachmodelle hoch individualisierte Antworten und imitieren menschliche kognitive Funktionen. Hybride Chatbots kombinieren regelbasierte mit machine-learning-Ansätzen, um kontrollierbare und fähige Chatbots zu entwickeln (Adamopoulou & Moussiades, 2020; Smutny & Schreiberova, 2020; Wu & Yu, 2024). Diese architektonischen Merkmale werden durch didaktische Gestaltungsfragen ergänzt, wie z.B. formale bzw. non-formale/informelle Lernsettings, Inhalte/Wissensbasis und Sozialformen (Hobert & Meyer von Wolff, 2019).

Übergreifend bestehen zahlreiche Desiderate zu Gestaltung, Affordanzen, Nutzung und Wirkung von Educational Chatbots. So liegen die Hauptanwendungsfelder meist bei hoch strukturierten Domänen wie dem Sprachenlernen, Programmieren und Mathematik (z.B. Wollny et al., 2021), andere Domänen sind noch unterrepräsentiert. Auf der Lernendenseite besteht die Frage nach den nötigen Kompetenzen im Umgang mit Chatbots (z.B. Wu & Yu, 2024). Darüber hinaus besteht ein allgemeiner Bedarf an theoriebasierter, iterativer und lernendenzentrierter Entwicklung von Chatbots, die insbesondere Zielgruppen jenseits der Schul-/Hochschulbildung adressiert (Khosrawi-Rad et al., 2022). Diese allgemeinen Fragen werden durch spezifische Fragen ergänzt, die im Zuge von generativen KI-Chatbots hinzugekommen sind, wie die Frage nach den genuin neuen Möglichkeiten (Neumann et al., 2024).

Um tieferes Wissen über die Potenziale und Grenzen von Chatbots zu erhalten, braucht es daher Forschung, die die Bedingungen für den effektiven Einsatz sowie deren lernförderliche Gestaltung in den Blick nimmt - also das "Wann" und das "Wie" statt nur das "Ob" untersucht (Buchner & Kerres 2023; Scheiter, 2021). Dieser Annahme folgend, möchte dieses Symposium nach einem kurzen einleitenden Referat durch den Diskutant:in in drei Beiträgen diese Fragen diskutieren. Der erste Beitrag untersucht in einer experimentellen Vergleichsstudie, inwiefern Lernende (n = 214) im Physikunterricht durch KI-generierte Erklärungen gefördert werden können, um nachhaltig mathematische Kompetenzen in Lerngelegenheiten der Physik anzuwenden. Im zweiten Beitrag steht die mediendidaktische Frage im Vordergrund, wie Lernende dabei unterstützt werden können, generative KI-Chatbots effektiv zur Informationsbeschaffung zu nutzen. Eine experimentelle Studie (n = 103) beleuchtet die Wahrnehmung und das Interaktionsverhalten der Lernenden sowie die Effekte von Adaptionsunterstützung. Der dritte Beitrag hebt die Betrachtungen zu Educational Chatbots auf die Lehrkräfteebene (n = 174, für alle Teilstudien) und fragt nach dem Potential von hybriden Chatbots als Boundary Objects fachdidaktische Transferaktivitäten am Beispiel planungsbezogener bzw. schulpraktischer Aktivitäten zu unterstützen.

Das Symposium schließt mit einer Diskussion, die Gemeinsamkeiten, Innovationsgehalt und Spezifika der Beiträge thematisiert. Darüber hinaus sollen die Bedarfe und Schwerpunkte zukünftiger Forschung zu Educational Chatbots in den Teilgebieten der Bildungsforschung diskutiert werden, um den Blick auf die Breite der Thematik zu weiten.

Beiträge des Symposiums

Individuelle Unterstützung im Mathematik- und Physikunterricht durch KI-Chatbots

Julia Lademann, Prof. Dr. Sebastian Becker-Genschow

Universität zu Köln, Digitale Bildung mit Schwerpunkt Künstliche Intelligenz

Theoretischer Hintergrund

Verständnisprobleme von Schülerinnen und Schülern im Physikunterricht sind nicht selten auf unzureichende mathematische Kompetenzen zurückzuführen (Christensen & Thompson, 2012; Leinhardt et al., 1990; Pollock et al., 2007). Oft basieren sie aber auch auf der Schwierigkeit, die im Rahmen des Mathematikunterrichts erworbenen Fähigkeiten auf den physikalischen Kontext zu übertragen (Becker et al., 2023; Redish & Kuo, 2015). Die Ursachen für diese Schwierigkeiten sind mannigfaltig (Uhlen, 2016). Generative KI kann an dieser Stelle mit der individuellen Förderung durch Chatbots eine mögliche Lösung darstellen. KI-Chatbots können den Lernenden personalisiertes Feedback geben und auf individuelle Lernschwierigkeiten und Fragen eingehen (Kasneci et al., 2023; Adiguzel et al., 2023; Neumann et al., 2024). Des Weiteren lassen sich Chatbots gezielt konfigurieren, so

dass beim Erlernen physikalischer Inhalte Hilfestellungen und Feedback mit Bezug zu dem zugehörigen mathematischen Kontext bereitgestellt werden.

Fragestellung

In einer ersten Studie wurde untersucht, wie sich das Lernen mit von KI-Chatbots generierten Erklärungen auf das Lernerleben sowie die Lernleistung in einem mathematischen und physikalischen Lernkontext auswirkt, insbesondere im Hinblick auf die notwendige Transferleistung auf den physikalischen Kontext.

Methode

Zu diesem Zweck wurde ein auf ChatGPT basierender KI-Chatbot entwickelt, der die Schülerinnen und Schüler sowohl beim Erlernen mathematischer Inhalte als auch bei der Anwendung von Mathematik im Physikunterricht gezielt und individuell unterstützen soll. Der Chatbot verwendet eine altersgerechte Sprache und ist mit fachspezifischem Wissen ausgestattet.

In der randomisiert kontrollierten Studie mit $N = 214$ Schülerinnen und Schülern der sechsten Klasse an weiterführenden Schulen in Deutschland wurde untersucht, wie sich KI-generierte Erklärungen im Vergleich zu Material aus einschlägigen Schulbüchern auf das Lernen auswirken. Innerhalb der Schulklassen wurden die Kinder zufällig zwei Gruppen zugeteilt. Beide Gruppen erhielten Lernmaterialien zu dem für sie neuen mathematischen Thema „Proportionale Zuordnungen“, das einen Anwendungsbezug in der Physik hat. Die Kontrollgruppe arbeitete ausschließlich mit herkömmlichem Schulbuchmaterial. Die Experimentalgruppe erhielt Schulbuchmaterial sowie eine darauf bezogene Erklärung des Themas, die mittels eines vorher definierten Prompts von dem oben genannten KI-Chatbot erstellt wurde. Beide Gruppen arbeiteten ausschließlich mit Material in Papierform, um Verzerrungen durch den Einsatz digitaler Geräte zu vermeiden. Außerdem hatten beide Gruppen gleich viel Zeit, um sich mit den Materialien zu beschäftigen und wurden vorher nicht über die Art des vorliegenden Lernmaterials informiert.

Nach dem Lernen mit den Materialien wurden verschiedene Konstrukte anhand einer vierstufigen Likert-Skala erhoben. Erfasst wurden die wahrgenommenen Emotionen und das situationale Interesse der Teilnehmenden. Darüber hinaus wurde die intrinsische und extrinsische kognitive Belastung beim Lernen mit den Materialien sowie die Selbstwirksamkeitserwartung im Hinblick auf das Lösen einer themenbezogenen Aufgabe erhoben. Um die Lernleistung und die Fähigkeit zur Transferleistung mathematischer Inhalte in den physikalischen Kontext zu vergleichen, bearbeiteten die Kinder abschließend einen Leistungstest.

Ergebnisse und ihre Bedeutung

Erste Auswertungen zeigen signifikante Unterschiede zwischen Experimental- und Kontrollgruppe hinsichtlich des Erlebens positiver Emotionen sowie des situationalen Interesses – zugunsten des KI-gestützten Lernens. Auch die Selbstwirksamkeitserwartung in Bezug auf das Lösen einer themenbezogenen Aufgabe ist in der Experimentalgruppe signifikant höher, während die kognitive Belastung signifikant niedriger ist. Diese Ergebnisse deuten darauf hin, dass das Lernen mit KI-Chatbots einen positiven Effekt auf das Lernerleben von Schülerinnen und Schülern haben kann, und eröffnen den Raum für weiterführende Forschung.

Generative KI-Chatbots beim selbstregulierten Lernen: Wahrnehmung, Interaktionsstile und Adaptionunterstützung

Maria Klar

Universität Duisburg-Essen, Lehrstuhl für Mediendidaktik und Wissensmanagement

Theoretischer Hintergrund

Schüler:innen nutzen generative KI-Chatbots (genKI-Chatbots) vor allem zur Informationsbeschaffung und sie wünschen sich mehr Anleitung, wie man sie zum Lernen nutzen kann (Vodafone Stiftung, 2024). Dabei erscheinen genKI-Chatbots zunächst intuitiv in der Bedienung, allerdings bieten sie diverse implizite Affordanzen, also Nutzungsmöglichkeiten, die nicht eindeutig als solche entwickelt wurden (Antonenko et al., 2017). Lernende sind nicht immer in der Lage ihre Lernprozesse effektiv zu regulieren (z.B. Peverly et al., 2003), beispielsweise auch im Umgang mit Suchmaschinen (Rieh et al., 2016). Insofern ist unklar, inwiefern Schüler:innen über die nötigen Strategien verfügen, um effektiv mit genKI-Chatbots recherchieren zu können. Es besteht die Möglichkeit, dass Lernende Chatbots als „einfaches Medium“ wahrnehmen (Salomon, 1984) und dabei lernförderliche Affordanzen übersehen. Ebenso ist möglich, dass das Chatbot-Interface als unstrukturiertes Lernmedium zu wenig Steuerung bietet und somit die kognitive Last erhöht (Seufert, 2018). Das Ergänzen von Interaktionshinweisen könnte beiden Herausforderungen entgegenwirken, insbesondere wenn die Interaktionshinweise mit Instruktion kombiniert werden (z.B. Moser et al., 2017).

Fragestellungen

1. Wie nehmen Schüler:innen genKI-Chatbots wahr?
2. Wie interagieren sie mit einem genKI-Chatbot bei der Recherche?
3. Welchen Effekt haben Instruktion und unterstützende Features auf Adaptionprompts, extrinsische kognitive Last und Lernzuwachs?

Methode

An dieser Studie im experimentellen Design haben $n = 106$ Schüler:innen der 9./10. Klasse teilgenommen, die jeweils einer von zwei Bedingungen zugeordnet wurden: In beiden Bedingungen war die Aufgabe, 20 Minuten mit Hilfe eines auf ChatGPT3.5 basierenden Chatbots zum Thema „kognitive Verzerrungen“ (also systematische Verzerrungen im Wahrnehmen und Denken) zu recherchieren. In einem Prä-Fragebogen wurden Vorwissen, Interesse, bisherige Chatbot-Nutzung und eine Selbsteinschätzung zum selbstregulierten Lernen (LASSI-HS, Subskala „Informationen verarbeiten“ (Eldredge, 1990; Fehrer, 2018)) erfragt. Der Post-Fragebogen erhob Wissen, kognitive Last (Klepsch et al., 2017), Interesse und Zufriedenheit mit dem Chatbot. Alle Items waren geschlossen mit Ausnahme der Wissenstests. Während in der Kontrollbedingung ein Chatbot ohne didaktische Features verwendet wurde, erhielt die Experimentalbedingung Unterstützungsfeatures: Vorschläge für weitere Prompts (User Eingaben) sowie die Möglichkeit, Antwortlänge und Sprachniveau des Chatbots zu variieren. Darüber hinaus erhielt die Experimentalbedingung eine 2-minütige Instruktion zu adaptierenden Prompts wie „Veranschauliche dies“ oder „Erstelle eine Zusammenfassung“.

Ergebnisse und ihre Bedeutung

(1) Die Lernenden nehmen den Chatbot positiv und leicht bedienbar wahr ohne signifikante Unterschiede zwischen den Gruppen. Mit oder ohne Unterstützung gaben die Schüler:innen gleichermaßen an, dass sie immer wussten, was sie als Nächstes fragen sollten und dass die Antworten des Chatbots ihre Informationsbedürfnisse erfüllten. (Deskriptive Auswertung)

(2) Lernende der Gruppe mit Unterstützung nutzten vielfältigere adaptierende Prompts, stellten mehr Fragen und Folgefragen und beteiligten sich weniger an themenfremden Interaktionen als die Kontrollgruppe. Insgesamt zeigt sich jedoch auch, dass Lernende in beiden Gruppen eher selten ein Gespräch mit dem Chatbot eingingen. Die Lernenden stellten überwiegend mehrere Fragen hintereinander und deckten somit zentrale Informationen zum Thema ab. (Induktive Codierung der Chatverläufe)

(3) Die Unterstützung führte zu einer signifikant höheren Anzahl von adaptierenden Prompts, während die extrinsische Belastung und der Wissenszuwachs sich nicht signifikant unterschieden. (einfaktorielle MANOVA & post-hoc ANOVAS)

Die Teilnehmenden nahmen den Chatbot als „einfaches Medium“ wahr, nutzten aber dessen Nutzungsmöglichkeiten nicht aus: In der Kontrollgruppe nutzte ein Drittel der Teilnehmenden keinerlei adaptierende Prompts und es wurden wenige Nachfragen gestellt. Unterstützungsfeatures und eine kurze Instruktion konnten dem entgegenwirken.

Studie 2

Aufbauend auf diesen Ergebnissen soll Studie 2 u.a. untersuchen, ob durch den Chatbot gestellte Fragen zu höherer mentaler Anstrengung führen. Die Erhebung ist für Herbst 2024 geplant, sodass voraussichtlich erste Ergebnisse präsentiert werden können. Etwa 100 Schüler:innen sollen wiederum an einer von zwei Bedingungen teilnehmen: Beide Bedingungen recherchieren mit einem KI-Chatbot, basierend auf ChatGPT4o, zum Thema „Verschwörungstheorien“; die Experimentalbedingung erhält Verständnis- und Reflexionsfragen. Die Fragebögen entsprechen jenen in Studie 1 mit der Ergänzung von Items zur mentalen Anstrengung (Klepsch & Seufert, 2021).

Educational Chatbots als Boundary Objects: Gegenstandsbezogene Transferunterstützung aus fachdidaktischer Perspektive

Steven Beyer

Universität Duisburg-Essen, Lehrstuhl für Mediendidaktik und Wissensmanagement

Theoretischer Hintergrund

Substantielle Lernumgebungen

Substantielle Lernumgebungen verbinden empirische Forschung und Entwicklungsaktivitäten (Wittmann, 1998). Kern dieser Lernumgebungen sind mathematisch gehaltvolle Aufgaben, die durch ihren konstruktivistischen Charakter u.a. mehr Verstehensorientierung fördern sollen (Krauthausen, 2018). Auch in der Lehrkräftebildung sind Lernumgebungen wertvoll, da sie vielfältiges Wissen für den Mathematikunterricht effizient organisieren (Wittmann, 2021). Dies kann in Lerngelegenheiten erkundet und in der Schulpraxis genutzt werden, um zeitgemäßen Mathematikunterricht zu gestalten.

Jedoch ist der Transfer von Forschung in Praxis als auch der Rückfluss der Evaluationsergebnisse aus der schulpraktischen Erprobung, ein Unterfangen mit zahlreichen Hürden und Bedarfen, z.B. auf Seiten der (angehenden) Lehrkräfte (z.B. Lipowsky, 2014).

Educational Chatbots als Boundary Objects

Zur erfolgreichen Forschungs-Praxis-Zusammenarbeit ist es notwendig institutionelle/disziplinäre Grenzen zu überschreiten (Boundary Crossing). In diesem Zusammenhang können Boundary Objects, im Sinne von Produkten des Knowledge Brokering, als Unterstützungsangebote dienen (Schellenbach-Zell & Hartmann, 2022; Star & Griesemer, 1989).

Im Kontext der Transferunterstützung von (angehenden) Lehrkräften wird der förderliche Einsatz von digitalen Medien, z.B. bei der Aufgabenadaption, wiederholt postuliert. Jedoch besteht ein Evidenzdefizit, zum einen hinsichtlich förderlicher Gestaltungsmerkmale/Formate von Unterstützungsangeboten und zum anderen bzgl. deren Wirkung (Gorard et al., 2020; McKenney & Visscher, 2019).

In diesem Kontext sind ebenso neuere Entwicklungen, wie der Einsatz von Educational Chatbots zu betrachten. Auch hier werden für (angehende) Lehrkräfte zahlreiche Potentiale beschrieben, z.B. zur strukturierten Begleitung im Coaching sowie von professionellen Lerngemeinschaften (Jordan, 2023). Auf der Unterrichtsebene konnte umfangreich der Nutzen des Einsatzes von Educational Chatbots für z.B. lernbezogene Outcomes aufgezeigt werden (ausführlich Meta-Analyse von Wu & Yu, 2024). Jedoch bestehen Desiderate z.B. im Bereich Lehrkräfteprofessionalisierung und konkreter Anwendungsfälle (Chiu et al., 2023).

Fragestellung

Im Rahmen eines Design Research-Prozesses (z.B. Plomp, 2013) wurde folgende übergreifende Forschungsfrage bearbeitet: Wie kann ein Chatbot als Boundary Object konzipiert und evaluiert werden, der das Verständnis sowie die Praktiken von (angehenden) Mathematiklehrkräften in Bezug auf Transferaktivitäten zu digital unterstützten Lernumgebungen unterstützt?

Im Rahmen der Untersuchungszyklen wurden dazu iterativ Forschungsfragen bearbeitet, die im Folgenden zusammen mit den Erhebungs- und Auswertungsmethoden umrissen werden.

Untersuchungszyklen & Methoden

- Analyse des Kontextes: Erfassung der Praktiken sowie Orientierungen bei einer schulpraktischen Transferaufgabe bei n = 19 Mathematiklehrkräften (Qualitative Inhaltsanalyse von Portfolios).

- Evaluation des Prototypendesigns I: Erfassung chatbotbezogener Beliefs und Nutzungsintention sowie Einflussvariablen hinsichtlich der Chatbotnutzung bei Transferaktivitäten bei n = 11 Mathematiklehrkräften/-fortbildenden (Qualitative Inhaltsanalyse von Fokussierten Interviews).

- Evaluation des Prototypendesigns II: Intraindividuelle/interindividuelle Veränderung der chatbotbezogenen Beliefs (t-Test/Pfadanalyse) sowie gegenseitige Effekte von Chatbotnutzung und Beliefs bei n = 92 Lehramtsstudierenden (Pfadanalyse).

- Beginn summative Evaluation: Intraindividuelle/interindividuelle Veränderung der planungsbezogenen Selbstwirksamkeit (pSWE) bei Chatbot-Nutzung (ANOVA mit Messwiederholung, Mixed ANOVA/Pfadanalyse) sowie Effekte von Chatbot-Nutzung, Kontrolle, Wert und Freude auf pSWE während sowie nach Transferaktivitäten bei n = 83 Lehramtsstudierenden (Pfadanalyse).

Ergebnisse und ihre Bedeutung

Die Design Research-Projektergebnisse können als Machbarkeitsstudie (proof of concept; Bakker, 2018) betrachtet werden. Ausgangspunkt waren Lernumgebungen als praxisorientierte Form der Wissensaufbereitung, die bereits als Boundary Object gesehen werden können, jedoch Transferaktivitäten nicht ausreichend unterstützten. Dieses Projekt hat gezeigt, dass ihre Aufbereitung in einem Chatbot mit Transferunterstützungsmerkmalen zu einem Boundary Object führt, das als interaktives Werkzeug flexibel von (angehenden) Lehrkräften genutzt werden kann, um ihre Transferaktivitäten zu unterstützen. Die Untersuchungszyklen zeigen, dass es planungsbezogene/schulpraktische Herausforderungen zielgerichtet adressiert

(Nützlichkeit) und meist mit geringem Aufwand nutzbar ist (Bedienfreundlichkeit). Die begonnene summative Evaluation lieferte erste Ergebnisse: Die Chatbotnutzung zeigte keinen signifikanten Effekt auf Veränderung der pSWE, Praxisphase sowie Freude positive Effekte und Alter negativen Effekt. Mögliche Gründe sind die Komplexität der Chatbot-Effekte und unberücksichtigte Konstrukte. Daher sollten Fragen nach Effekten mit und von Medien (Salomon & Perkins, 2005) in Folgestudien weiterverfolgt werden. Dazu zählen Fragen nach der Veränderung fachdidaktischer Qualität chatbotgestützter Unterrichtsplanungen (Effekte mit Medien) sowie langfristige Veränderungen der Merkmale professioneller Kompetenz (Effekte von Medien).

S-62: Chancen und Risiken KI-gestützter Literaturüberblicke und Metaanalysen

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 145

Chair der Sitzung: Tim Fütterer

Chair der Sitzung: Olga Chernikova

Chair der Sitzung: Jürgen Schneider

Chancen und Risiken KI-gestützter Literaturüberblicke und Metaanalysen

Chair(s): **Tim Fütterer** (Universität Tübingen), **Olga Chernikova** (Ludwig-Maximilians-Universität München)

Diskutant*in(nen): **Jürgen Schneider** (Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Forschungssynthesen wie systematische Literaturüberblicke und Meta-Analysen sind methodische Ansätze zur Synthese von Forschungsergebnissen aus Primärstudien (Campos et al., 2021), die für eine evidenzbasierte Weiterentwicklung von Praxis, Forschung und Politik von zentraler Bedeutung sind (z.B. Schneider & Preckel, 2017; van de Schoot et al., 2021; van Huizen & Plantenga, 2018). Forschungssynthesen stellen Forschende jedoch vor Herausforderungen (Card, 2015). Zum Beispiel nehmen Screening und Informationsextraktion viele Ressourcen (z.B. Arbeitszeit) in Anspruch (Borah et al., 2017). Zudem können unerfahrene Reviewer einzuschließende Studien übersehen, wodurch Metaanalysenergebnisse beeinträchtigt werden können (bis zu 13%; siehe Waffenschmidt et al., 2019).

Tools, die auf Ansätzen künstlicher Intelligenz (KI) basieren (z.B. Ansätze maschinellen Lernens zur automatisierten Auswertung von Texten; kurz KI-Tools), haben das Potenzial, Forschungssynthesen zu beschleunigen und zuverlässiger werden zu lassen. Algorithmen zeigten sich bereits zuverlässiger als Menschen (Loh, 2018) und bei Technologien wie der generative KI ist das Potenzial zu erkennen, zeitnah ebenfalls zuverlässige Leistungen (bspw. bei Informationsextraktion) zu erzielen. Das Ziel ist dabei nicht, dass KI-Tools menschliche Entscheidungen im Forschungsprozess ersetzen, sondern in einem human-in-the-loop Prozess als Unterstützung genutzt werden (Van de Schoot & De Bruin, 2020). KI-Tools werden zunehmend auch im Bildungsbereich genutzt (z.B. Fütterer et al., 2023) aber es gibt viele offene Fragen:

1. Welche Tools gibt es, die Qualitätsstandards erfüllen und im Bildungsbereich genutzt werden können?
2. Wie performen KI-gestützte Tools im Vergleich zu herkömmlichen Screening-Methoden?
3. Wie können KI-Tools praktisch im Forschungsprozess eingebunden werden?

Das Symposium umfasst drei Beiträge und eine rahmende Diskussion. Im ersten Beitrag wird ein Überblick über systematisch ausgewählte KI-Tools gegeben, die zum einen Qualitätskriterien (z.B. Datenschutz) erfüllen und zugleich in der Praxis von Forscher:innen im Bildungsbereich ohne Informatikkenntnisse genutzt werden können. Es wird auf Stärken, Schwächen, Potenziale und Grenzen der Tools eingegangen. Im zweiten Beitrag evaluieren die Autor:innen die Frage, wie KI-Tools für das automatisierte Screenen von Titeln und Abstracts (z.B. ASReview) zuverlässig insbesondere in der pädagogisch-psychologischen Forschung eingesetzt werden können. Abbruchregeln wie die Knie-Methode werden anhand von Sensitivität, Spezifität und Screening-Kosten basierend auf 21 Datensätzen evaluiert und Einblicke in deren Performanz in Abhängigkeit von der Prävalenz relevanter Studien und der Wahl der Lernalgorithmen gegeben. Im dritten Beitrag wird an einem konkreten Beispiel der Einsatz eines KI-Tools im Forschungsprozess anhand einer aktuell durchgeführten Metaanalyse zu den Auswirkungen simulationsbasierten Lernens aufgezeigt und dabei praktische Herausforderungen (z.B. Prompts-Design, Interrater-Übereinstimmung) diskutiert.

Um die Anwendungen von KI-Tools für den Forschungsprozess kritisch zu reflektieren und Stärken, Schwächen, Chancen und Risiken für eine Diskussion im Plenum aufzubereiten, wird das Symposium von einem ausgewiesenen Experten für *gute wissenschaftliche Praxis* und *open science* diskutiert.

Beiträge des Symposiums

Wie kann man künstliche Intelligenz für systematische Literaturübersichten und Meta-Analysen nutzen? Ein systematischer Überblick über Tools und ein Leitfaden für Forscher:innen

Tim Fütterer¹, Diego G. Campos², Gfrörer Thomas¹, Lavelle-Hill Rosa³, Murayama Kou¹, Scherer Ronny²

¹Universität Tübingen, ²Universität Oslo, ³Universität Kopenhagen

Theoretischer Hintergrund. Forschungssynthesen wie systematische Literaturüberblicke oder Metaanalysen sind von hoher Bedeutung für den Fortschritt der Wissenschaft und wichtige Instrumente für die Förderung evidenzbasierter Entscheidungen. Beispielsweise hat die COVID-19 Pandemie gezeigt, dass informierte Entscheidungen nahezu täglich basierend auf aktuell verfügbarer Evidenz geprüft werden mussten. Schätzungen zeigen jedoch, dass hochqualifizierte Forscher:innen bis zu 72 (Smith et al., 2011) Wochen benötigen, um eine Metaanalyse zu veröffentlichen. Vor dem Hintergrund der steigenden Zahl an jährlich veröffentlichten Artikeln (Bormann & Mutz, 2015; Larsen & von Ins, 2010) kann die Verarbeitung jüngster Erkenntnisse kaum so effizient erfolgen, dass sie bei Veröffentlichung nicht wieder überholungsbedürftig sind (Marshall & Wallace, 2019).

Vor diesem Hintergrund können KI-Tools für Forschungssynthesen mehrere Vorteile bieten. Zum Beispiel können einzelne Prozesse (z.B. das Screening) beschleunigt werden. Zudem bieten KI-Tools das Potenzial, die Reproduzierbarkeit von Forschungsarbeiten zu fördern und menschliche kognitive Verzerrungen oder mentaler Ermüdung zu reduzieren. In den vergangenen Jahren wurden zahlreiche KI-Tools für die Erstellung von Forschungssynthesen entwickelt, jedoch nicht alle sind evaluiert und mit guter wissenschaftlicher Praxis vereinbar (Stichwort: Blackbox). Zudem erfordern viele Tools Programmiererfahrungen, was angewandte Forscherinnen und Forscher vor einer Nutzung abschreckend kann. Zudem ist anzunehmen, dass viele Forscherinnen und Forscher einen begrenzten Überblick über KI-Tools haben, oder wenn sie sie kennen, diese oft nicht mit ihnen vertraut sind oder für zu schwierig in der Anwendung halten (Chai et al., 2021).

Forschungsfragen. Forschungssynthesen zielen darauf ab, empirische Belege mit systematischen, replizierbaren, transparenten, zuverlässigen Methoden zusammenzutragen (Cooper et al., 2019; Moreau & Gamble, 2020) und bestehen meist aus sechs Schritten von der Definition der Forschungsfragen bis zur Kommunikation der Ergebnisse (Carrasco-Labra et al., 2021). In diesem Beitrag stellen wir verfügbare KI-Tools für unterschiedliche Phasen einer Forschungssynthese, die ohne Informatikvorkenntnisse genutzt werden können, vor. Zudem wollen wir Forscherinnen und Forscher im Rahmen einer Handreichung über evaluierte Stärken, Schwächen, Potenziale und Risiken – z.B. bezüglich Replizierbarkeit, Transparenz und Zuverlässigkeit – informieren. Wir adressieren drei Fragen:

- (1) Welche KI-Tools stehen für die einzelnen Schritte einer Forschungssynthese zur Verfügung?
- (2) Was ist mit KI-Tools bei der Durchführung einer Forschungssynthese (nicht) möglich?

(3) Wie können Forscherinnen und Forscher diese KI-Tools nutzen, um effizienter zu arbeiten und gleichzeitig wissenschaftliche Standards einzuhalten?

Insgesamt zielt unser Beitrag nicht darauf ab, einen vollständigen Überblick über alle KI-Tools zu geben, sondern diese Tools in der Community bekannt zu machen, damit sie in der Forschungspraxis eingesetzt und evaluiert werden können.

Methode. Um die Forschungsfragen zu beantworten, haben wir 254 KI-Tools aus vier Übersichtslisten (u.a. Bolanos et al., 2024; Cierco Jimenez et al., 2022; van de Schoot et al., 2021) zusammengetragen, die potenziell für Forschungssynthesen im Bildungsbereich verwendet werden können. Anschließend haben wir die Tools anhand von fünf Kriterien systematisch gefiltert. Die Tools mussten tatsächlich KI nutzen, hinreichend transparent hinsichtlich verwendeter Algorithmen sein, Datensicherheit gewährleisten, kostenlos verfügbar sein und im Bildungsbereich nutzbar sein. Acht Tools für unterschiedliche Phasen einer Forschungssynthese (z.B. Screening, Extraktion von Informationen) erfüllten alle Kriterien. Alle acht Tools wurden jeweils doppelt unter Verwendung eines Kodiermanuals bezüglich der Funktionalität der Tools, Ergebnissen von Evaluationen zur Performanz, und der Nutzerfreundlichkeit evaluiert.

Ergebnisse. Wir werden einen Überblick zu den acht KI-Tools geben und auf Stärken, Schwächen, Potenzialen und Risiken bei der Verwendung dieser Tools eingehen. Die Evaluation der Tools ergab, dass sich die Tools darin unterscheiden, wie sehr Nutzer in der Lage sind, die Tools für eigene Anwendungen zu optimieren, wie sehr sie geeignet sind, einen reproduzierbaren Workflow zu etablieren oder wie hoch das Risiko von Verzerrung von Ergebnissen ist (z.B. aufgrund von Fehlfunktionen oder aufgrund von „Blackbox“-Anteilen bei genutzten Algorithmen). Wir werden zudem kurz aufzeigen, wie man die Tools im Sinne einer guten wissenschaftlichen Praxis für den eigenen Forschungsprozess nutzen kann.

Evaluation der Leistung von Abbruchregeln im KI-gestützten Screening für die psychologische Meta-Analyseforschung

Lars König¹, Steffen Zitzmann², Tim Fütterer³, Diego G. Campos⁴, Ronny Scherer⁴, Martin Hecht¹

¹Helmut-Schmidt-Universität, Hamburg, ²Medical School Hamburg, ³Universität Tübingen, ⁴Universität Oslo

Theoretischer Hintergrund. Die stetig wachsende Anzahl an wissenschaftlichen Publikationen führt zu einer zunehmend komplexen Literaturlandschaft. Die Evidenzsynthese durch systematische Reviews und Meta-Analysen erfordert daher einen höheren Zeit- und Arbeitsaufwand (Elliott et al., 2017). Insbesondere das Screening von Abstracts macht dabei einen wesentlichen Teil der Arbeitsbelastung aus (Marshall & Wallace, 2019; Pham et al., 2018). Viele (KI-)Programme wurden daher entwickelt, um Forschende bei der Durchführung von Forschungssynthesen und den verschiedenen damit verbundenen Aufgaben zu unterstützen (Zhang & Neitzel, 2023). Dies gelingt im Wesentlichen durch die Sortierung der ungelesenen Abstracts nach der Wahrscheinlichkeit ihrer Relevanz, die durch spezielle Lernalgorithmen bestimmt wird (Burgard & Bittermann, 2023). Auf diese Weise sollten alle relevanten Artikel identifiziert werden können, bevor alle Abstracts gescreent wurden. Tatsächlich zeigten Forschungsergebnisse, dass im Durchschnitt nur etwa 50% der Abstracts gescreent werden müssen, um 95% der relevanten Artikel zu identifizieren. Jedoch wurden auch erhebliche Schwankungen in der Effizienz dieser Tools entdeckt, die sowohl auf Unterschiede in den maschinellen Lernalgorithmen als auch auf die zu screenende Literatur zurückzuführen seien (Burgard & Bittermann, 2023). Die Nutzer dieser Tools stehen daher vor der Herausforderung, zu entscheiden, wann das Screening beendet werden kann. Verschiedene Stoppregeln wurden entwickelt, um den Nutzern bei dieser Entscheidung zu helfen. Beispielsweise kann die Anzahl relevanter Studien basierend auf einer Stichprobe geschätzt und das Screening gestoppt werden, sobald die entsprechende Anzahl gefunden wurde (Prävalenzschätzung; Callaghan & Müller-Hansen, 2020; van Haastrecht et al., 2021). Bei anderen Stoppregeln kann das Screening beendet werden, wenn mehrere aufeinanderfolgende Abstracts als irrelevant eingestuft wurden (Datenbasierte Heuristik; Ros et al., 2017) oder ein bestimmter Punkt in der Lernkurve der Algorithmen überschritten wurde (Knee-Methode; Cormack & Grossman, 2016).

Forschungsfragen. Aufgrund fehlender Evaluationsstudien der Stoppregeln, kann derzeit keine robuste Schätzung für ihre Performanz und somit des Anteils nicht identifizierter relevanter Studien abgeleitet werden. Auch der Einfluss von Faktoren wie der Prävalenz relevanter Studien und der gewählten Lernalgorithmen bleibt ungewiss. Zudem können die bisherigen Forschungsergebnisse nicht auf psychologische oder pädagogische Literatur generalisiert werden. Aus diesem Grund befasste sich die vorliegende Arbeit mit den folgenden Forschungsfragen.

1. Wie performen drei Stoppregeln (Prävalenzschätzung, Knee-Methode und datenbasierte Heuristik) in verschiedenen psychologischen Literatursammlungen?
2. Inwiefern beeinflussen das Prävalenzverhältnis, der Lernalgorithmus und die zugrundeliegende Literatur die Effektivität der Stoppregeln?

Methode. Wir evaluierten die Effektivität der drei Stoppregeln (Knee-Methode, datengesteuerte Heuristik und Prävalenzschätzung) hinsichtlich ihrer Sensitivität, Spezifität und der Screening-Kosten (Anteil der gescreenten Abstracts an allen Abstracts). Zu diesem Zweck sammelten wir Abstracts, die ursprünglich für Meta-Analysen aus fünf verschiedenen psychologischen Forschungsbereichen gescreent wurden. Zusätzlich manipulierten wir die Prävalenz relevanter Artikel in jeder Abstract-Sammlung, sodass ein Verhältnis von relevanten zu irrelevanten Studien von 0,5%, 1%, 5% und 10% entstand. Auf Basis dieser manipulierten Abstract-Sammlungen simulierten wir das KI-gestützte Screening und die Anwendung der Stoppregeln unter Verwendung verschiedener Lernalgorithmen. Dabei verwendeten wir das KI-gestützte Screening-Tool ASReview (van de Schoot et al., 2021).

Ergebnisse. Unsere Studie zeigte deutliche Unterschiede in der Leistung der drei Stoppregeln, sowohl in Bezug auf die gemessenen Leistungskriterien als auch hinsichtlich ihrer Variabilität bei unterschiedlichen Prävalenzraten relevanter Studien. Obwohl der gewählte Lernalgorithmus generell einen geringeren Einfluss hatte, zeigte sich, dass spezifische Kombinationen aus Stoppregeln und Lernalgorithmen bei bestimmten Prävalenzraten von relevanten Abstracts besonders effektiv waren. Insgesamt erwies sich die Knee-Methode als besonders geeignet für Prävalenzverhältnisse von 0,5% und 1%, während die datengesteuerte Heuristik bei Prävalenzverhältnissen von 5% und 10% eine überlegene Leistung zeigte. Zwar resultierte die Schätzmethode häufig in der höchsten Sensitivität, sie birgt jedoch auch das Risiko des Oversamplings, was dazu führen kann, dass der Screening-Prozess nicht gestoppt wird. Basierend auf diesen Ergebnissen haben wir praktische Empfehlungen für die Anwendung von KI-gestützten Screening-Tools formuliert. Außerdem diskutieren wir die Implikationen unserer Befunde.

Einsatz von Large Language Models (LLMs) zur Unterstützung der systematischen Literaturrecherche und Metaanalyse in Bildungskontext

Olga Chernikova¹, Matthias Stadler²

¹Ludwig-Maximilians-Universität München, ²Klinikum der Ludwig-Maximilians-Universität München

Theoretischer Hintergrund. Forschungssynthesen, tragen zu einer erhöhten Generalisierbarkeit der Ergebnisse bei (z. B. Xiong et al., 2018) und sind ein wichtiger Bestandteil der Bildungsforschung. Die exponentielle Zunahme der Anzahl an Publikationen von Primärstudien in einigen Bereichen der Bildungsforschung (z.B., simulationsbasiertem Lernen) hat jedoch in den letzten Jahren zu einer Reihe von Komplikationen bei der systematischen und vollständigen Erfassung des Forschungsstands geführt (z. B. Ware & Mabe, 2015). Zu diesen Komplikationen gehören mögliche Aufmerksamkeitsfehler, systematischer Bias und erhöhte Kosten, die mit der manuellen Auswertung und Kodierung von Primärstudien verbunden sind. Erste Tools zur Unterstützung der Auswertung, die mit maschinellem Lernen (ML), generativer künstlicher Intelligenz (GenAI) und anderer Algorithmen zur Verarbeitung natürlicher Sprache arbeiten, wurden bereits vorgestellt (Ruiz, 2022; van de Schoot et al., 2021). Es mangelt aber noch an Informationen über die Nützlichkeit dieser Tools. Darüber hinaus entwickeln sich die verwendeten Algorithmen sehr dynamisch weiter und ihre Leistung kann je nach Bereich und Thema unterschiedlich sein (z. B. Burgard & Bittermann, 2023; Chai et al., 2021; van de Schoot et al., 2021). Der Einsatz maschineller Lernalgorithmen zur Unterstützung der Klassifizierung während des Abstract Screenings im Bildungsbereich kann hilfreich sein (z.B., Chernikova et al., 2024). ML erfordert jedoch gute Programmierkenntnisse sowie erheblichen Zeit- und Datenaufwand für das Training. Updates von ML-Bibliotheken (z. B. in R oder Python) können zudem die Reproduzierbarkeit beeinflussen. Der Einsatz von Großen, vortrainierten Sprachmodellen (Large Language Models; LLMs), wie GPT-4.0, ist auch Forscher*innen mit geringeren Programmierkenntnissen zugänglich. Dennoch ist es wichtig, den Einsatz von ML und KI-Tools im Forschungsprozess kritisch zu reflektieren, die Rolle und Verantwortung des Menschen in diesem Prozess zu definieren, um Risiken zu minimieren und gute wissenschaftliche Praxis zu fördern.

Forschungsfragen. In diesem Beitrag wird Forschungsfrage 3 des Symposiums thematisiert.

Methode. In dieser empirischen Studie haben wir die Nützlichkeit eines LLMs (GPT-4.0) untersucht, die Eignung von Abstracts für systematische Literaturübersichten und Metaanalysen zu beurteilen und verglichen die Leistung des LLM mit trainierte menschliche Kodierer. Um dieses Forschungsziel zu erfüllen, verwendeten wir Daten aus einer laufenden Metaanalyse zu den Auswirkungen simulationsbasierten Lernens in der Hochschulbildung genutzt. Als erster Schritt haben wir einen Trainingsdatensatz (basierend auf manuell kodierten Daten aus einer früheren Metaanalyse) genutzt, um unsere Prompts zu optimieren. Anschließend wurde eine neue Literaturrecherche durchgeführt, die Studien zwischen Dezember 2020 und April 2024 umfasste. Mit den Schlüsselwörtern *simulat**, *competenc**, *skill** und *higher education* erhielten wir 10.000 Treffer aus den Datenbanken PsycINFO, PsycARTICLES, ERIC und MEDLINE. Aus diesen 10.000 Treffern wurden zufällig 1.000 Abstracts ausgewählt, um die Leistung von KI und menschlichen Kodierern zu vergleichen.

Die Abstracts wurden mehreren Kodierungsrunden mit GPT-4.0 unterzogen und zusätzlich von zwei menschlichen Kodierern mit unterschiedlichem Erfahrungsniveau kodiert. Mithilfe einer binären Kodierungsaufgabe und der gewichteten Cohen-Kappa-Statistik wurden die Zuverlässigkeit und Konsistenz der Kodierung bewertet.

Ergebnisse und Diskussion. Die Ergebnisse zeigen einen hohen Grad an Übereinstimmung zwischen den LLM- und menschlichen Kodierern, was auf die Wirksamkeit von LLMs bei der Kodierung von Studienabstrakten hinweist. Unsere Forschung zeigt, dass durch die Einbindung von LLMs nicht nur die Effizienz gesteigert, sondern auch die Qualität der Kodierung verbessert werden kann, indem systematische Fehler und Biases minimiert werden. Dennoch bleibt die kritische Überprüfung und regelmäßige Aktualisierung der eingesetzten Algorithmen essenziell, um die langfristige Validität und Zuverlässigkeit der Forschungsergebnisse zu gewährleisten. Die nicht-deterministische Natur von LLMs unterstreicht jedoch die Notwendigkeit menschlicher Beteiligung bei der Interpretation von Antworten und der Lösung von Meinungsverschiedenheiten. Wir schlagen daher einen iterativen Ansatz vor, der die Expertise menschlicher Programmierer mit der Effizienz von LLMs für Anwendungen im größeren Maßstab kombiniert. Diese Studie bestätigt das vielversprechende Potenzial von LLMs bei der Unterstützung systematischer Literaturrecherchen und bietet einen vielversprechenden Weg, der einer sorgfältigen Erforschung bedarf.

S-63: Inter- und intraindividuelle Variation im Zusammenspiel von selbstreguliertem Lernen, Metakognition und exekutiven Funktionen: Erkenntnisse aus verschiedenen Altersgruppen (DFG Netzwerk SeReNe)

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 148

Chair der Sitzung: Laura Dörrenbächer-Ulrich

Chair der Sitzung: Maria Theobald

Inter- und intraindividuelle Variation im Zusammenspiel von selbstreguliertem Lernen, Metakognition und exekutiven Funktionen: Erkenntnisse aus verschiedenen Altersgruppen (DFG Netzwerk SeReNe)

Chair(s): **Laura Dörrenbächer-Ulrich** (Universität des Saarlandes)

Diskutant*in(nen): **Maria Theobald** (Universität Trier)

Selbstreguliertes Lernen (SRL) und exekutive Funktionen (EF) beschreiben zielgerichtetes Verhalten und hängen positiv mit Leistung zusammen (Cortés Pascual et al., 2019; Dent & Koenka, 2016). Während SRL kognitive, metakognitive und motivationale Prozesse umfasst und sich auf die Lernleistung bezieht (Perels et al., 2020), umfassen EF grundlegende kognitive Prozesse zur Verhaltenssteuerung (Arbeitsgedächtnis, Inhibition, Shifting; Miyake et al., 2000). Erste Studien sprechen für einen positiven Zusammenhang zwischen SRL und EF (Davis et al., 2021) und legen nahe, dass Metakognition ein zentrales Bindeglied zwischen beiden Konstrukten darstellt (Dörrenbächer-Ulrich et al., 2023). Metakognition umfasst Fähigkeiten zur Einschätzung der eigenen Lernleistung (Monitoring) und zur Anpassung des Lernprozesses (Regulation) (Nelson & Narens, 1990), die prädiktiv für den Lernerfolg sind (Ohtani & Hisaka, 2018). Neben interindividuellen Unterschieden, die z.B. durch das Arbeitsgedächtnis erklärt werden können (Bryce et al., 2015), scheint auch intraindividuelle Variation bei metakognitiven Fähigkeiten vorzuliegen (van Loon et al., 2022). Das vorliegende Symposium umfasst vier Beiträge, die den Zusammenhang von SRL bzw. Metakognition und (Komponenten von) EF fokussieren und dabei verschiedene Altersgruppen vergleichen, weitere relevante kognitive Fähigkeiten wie Intelligenz miteinbeziehen, die intraindividuelle Heterogenität der Metakognition sowie den Zusammenhang der Konstrukte mit Lernerfolg analysieren.

Der erste Beitrag untersucht im Rahmen einer Mediationsanalyse die Rolle der Metakognition für den Zusammenhang zwischen EF und SRL bei Vor- und Grundschul*innen. SRL wurde mit einem Strategiewissenstest, EF mit je zwei Aufgaben pro Dimension und die Metakognition mit metakognitiven Judgments im Rahmen einer Lernaufgabe erhoben. Vorläufige manifeste Analysen konnten keinen Hinweis auf eine medierende Rolle der Metakognition in diesem Zusammenhang finden. Bis zur Konferenz liegen Ergebnisse einer Strukturgleichungsmodellierung vor.

Unter Verwendung von Strukturgleichungsmodellen untersucht der zweite Beitrag, inwiefern sich Intelligenz und EF als Prädiktoren von SRL im Vor- und Grundschulalter zeigen. Zur Messung von SRL wurde ein Strategiewissenstest eingesetzt, EF wurde über je zwei Aufgaben pro Dimension erfasst und Intelligenz über einen Test zum logischen Schlussfolgern erfasst. Während Intelligenz und EF hoch korrelieren und EF ein signifikanter Prädiktor für SRL sind, hat Intelligenz keinen prädiktiven Wert für SRL.

Der dritte Beitrag analysiert die Rolle intraindividuelle Unterschiede in der metakognitiven Regulationsfähigkeit für das Lernen von Sekundarschüler*innen. Die Genauigkeit des Monitorings, die effektive Selbstregulation sowie die kognitive Belastung wurden im Rahmen einer Lernaufgabe zu drei Sitzungen erfasst. Eine mehrbenenanalytische Messwiederholungsmodellierung soll die Zusammenhänge zwischen der Variabilität metakognitiver Fähigkeiten und der Aufgabenleistung auf intra- und interindividueller Ebene aufdecken und dabei auch die Rolle der kognitiven Belastung für die effektive Selbstregulation untersuchen.

Auch der vierte Beitrag untersucht den Zusammenhang von Metakognition und Lernerfolg, bezieht aber das Arbeitsgedächtnis als EF-Komponente mit ein. Im Rahmen einer digitalen Lernaufgabe wurden die metakognitiven Fähigkeiten (Genauigkeit des Monitorings, strategische Regulation) von Siebtklässler*innen erhoben. Multilevel-Regressionsanalysen werden genutzt, um die Lernleistung durch metakognitive Fähigkeiten sowie Arbeitsgedächtnis vorherzusagen. Erste Ergebnisse zeigen positive Korrelationen zwischen der strategischen Regulation und der Arbeitsgedächtnisleistung.

Beiträge des Symposiums

Zusammenhang zwischen exekutiven Funktionen und selbstreguliertem Lernen im Vor- und Grundschulalter – Die Rolle der Metakognition

Christina David¹, Lena Grüneisen¹, Sophia Grobe², Laura Dörrenbächer-Ulrich¹, Julia Karbach², Franziska Perels¹

¹Universität des Saarlandes, ²RPTU Kaiserslautern

Sowohl exekutive Funktionen (EF) als auch selbstreguliertes Lernen (SRL) spielen eine zentrale Rolle für akademische Leistung (Cortés Pascual et al., 2019; Dent & Koenka, 2016). Während EF als bereichsunspezifisches Konstrukt zielgerichtetes Verhalten im Allgemeinen unterstützt (Karbach & Unger, 2014), begünstigt SRL zielgerichtetes Denken, Handeln und Fühlen im Lernkontext (Zimmerman & Schunk, 2011). EF umfassen übergeordnete kognitive Prozesse, die unter anderem zur Hemmung irrelevanter Impulse (Inhibition), Aufmerksamkeitswechsel bei Anforderungsänderung (Shifting) und zur Aufrechterhaltung von Informationen im Arbeitsgedächtnis (Updating/Working memory) notwendig sind (Karbach & Unger, 2014; Miyake et al., 2000). Sie stehen so in engem Zusammenhang zu den kognitiven, metakognitiven und motivationalen Strategien des selbstregulierten Lernens (Effeney et al., 2013) und unterstützen ihre Ausführung (Rutherford et al., 2018). Entsprechend fanden Davis et al. (2021) Hinweise, dass EF SRL voraussagt. Die prozedurale Metakognition scheint in diesem Zusammenhang ein zentrales Bindeglied zu sein (Dörrenbächer-Ulrich et al., 2023; Effeney et al., 2013), da sie zum einen eine zentrale Komponente von SRL ist (Boekaerts, 1999), zum anderen aber auch in engem Zusammenhang zu EF steht (Bryce et al., 2023; Roebbers, 2017). Zu der prozeduralen Metakognition zählen sowohl metakognitive Überwachungs- als auch Kontrollprozesse während des Lernens (Roebbers, 2017). Erste Befunde an Studierendenstichproben deuten auf eine medierende Rolle der Metakognition im Zusammenhang von EF und SRL hin (Aghdar et al., 2020; Follmer & Sperling, 2016). Für das Vor- und Grundschulalter liegen bisher jedoch keine entsprechenden Befunde vor.

Die Zeit des Vor- und Grundschulalters ist eine bei der Entwicklung von EF, SRL und Metakognition zentrale Phase. Bereits im Vorschulalter sind grundlegende Fähigkeiten exekutiver Funktionen zu finden, welche sich im Vor- und Grundschulalter weiterentwickeln (Chevalier & Clark, 2018; Dörrenbächer-Ulrich et al., 2023). Eine ähnliche Entwicklung ist auch bei SRL zu beobachten, während im Vorschulalter bereits grundlegende (Vorläufer-)Fähigkeiten zu finden sind (z.B. Bronson, 2000;

Grüneisen et al., 2023; Jeong & Fry, 2020; Neitzel & Connor, 2018), trägt z.B. die Erfahrung in Lernsituationen zur Entwicklung von SRL im Grundschulalter bei (Dörrenbächer-Ulrich et al., 2023, Schneider, 2014). Metakognitive Fähigkeiten liegen zwar im Vorschulalter bereits rudimentär vor, entwickeln sich jedoch ebenfalls vor allem nach dem Schuleintritt weiter (Roebbers, 2017). Inwiefern diese Entwicklungen die Rolle der Metakognition im Zusammenhang von EF und SRL im Vor- und Grundschulalter beeinflusst ist jedoch noch unklar.

Das Ziel der Studie ist es daher, die Rolle der Metakognition im Vor- und Grundschulalter bei der Betrachtung des Zusammenhangs von EF und SRL genauer zu untersuchen. Aufgrund der theoretischen Überlegung, dass Metakognition ein zentrales Bindeglied zwischen EF und SRL darstellt und der empirischen Befundlage in älteren Stichproben (Aghdar et al., 2020, Follmer & Sperling, 2016), wird eine medierende Rolle sowohl im Vorschulalter als auch im Grundschulalter angenommen. Ob es Unterschiede in diesem Zusammenhang zwischen den beiden Alterskohorten gibt soll explorativ untersucht werden.

EF wurde multimethodal mit je zwei Aufgaben pro Dimension erhoben (z.B. backward corsi-like Aufgabe (Updating), Go/no-go Aufgabe (Inhibition) und modifizierter Wisconsin Card Sorting Test (Shifting); Cianchetti et al., 2007; Johann et al., 2022).

Zur Erfassung von SRL wurde eine an das Alter adaptierte Version eines Strategiewissenstests für Studierende genutzt (SKT-SRL; Dörrenbächer-Ulrich et al., 2024). Prozedurale Metakognition wurde mittels judgment-of-learning-Wertungen und allocation-of-study-time-Wertungen bei einer Paar-Assoziationslernaufgabe erfasst (z.B. Nelson et al., 1994). Vorläufige Analysen auf manifester Ebene an den Kindern, die alle Aufgaben vollständig bearbeitet haben und nicht aufgrund der Ausschlusskriterien ausgeschlossen wurden, konnten sowohl im Vorschulalter ($n = 92$, $M_{\text{Alter}} = 5.77$ Jahre; $SD_{\text{Alter}} = 0.42$) als auch im Grundschulalter ($n = 101$, $M_{\text{Alter}} = 9.88$ Jahre; $SD_{\text{Alter}} = 0.45$) keine Mediation durch die prozedurale Metakognition aufdecken. Finale latente Analysen mittels Strukturgleichungsmodellen an $N = 311$ Vorschul- und $N = 293$ Grundschulkindern sollen vorgestellt und diskutiert werden.

Fluide Intelligenz und Exekutive Funktionen (EF) als Prädiktoren Selbstregulierten Lernens (SRL) im Vor- und Grundschulalter

Sophia Grobe¹, Tanja Könen¹, Christina David², Lena Grüneisen², Laura Dörrenbächer-Ulrich², Franziska Perels², Julia Karbach¹

¹RPTU Kaiserslautern, ²Universität des Saarlandes

Selbstreguliertes Lernen (SRL) umfasst Prozesse, mit denen Lernende ihre Kognitionen, Affekte und Verhaltensweisen beobachten und steuern, um ein Lernziel zu erreichen (Zimmerman und Schunk 2011). SRL integriert kognitive, metakognitive und motivationale Prozesse (vgl. Perels et al. 2020). Die Entwicklung dieser Fähigkeiten beginnt im Vorschulalter (z.B. metakognitives Monitoring; Bryce und Whitebread 2012) und wird im Grundschulalter und darüber hinaus zunehmend differenzierter (vgl. Vandevelde et al. 2015). Die empirische Evidenz zum Zusammenhang von Intelligenz und SRL ist heterogen: Während überwiegend positive Zusammenhänge zwischen Intelligenz und metakognitivem Wissen berichtet werden (Alexander et al. 1995), sind Befunde zum Zusammenhang von Intelligenz und dem Einsatz von SRL-Strategien nicht eindeutig (Stoeger und Sontag). Exekutive Funktionen als mentale Prozesse zur Steuerung und Kontrolle von Gedanken und Handlungen zur Erreichung von Zielen (Friedman und Miyake 2017) sind eng verwandt mit Intelligenz (z.B. Brydges et al. 2012) und spielen möglicherweise eine wichtige Rolle für die Entwicklung und den Einsatz von SRL. Einzelne Studien berichteten positive Zusammenhänge zwischen EF und SRL im Kindesalter (Davis et al. 2021; Rutherford et al. 2018). Die vorliegende Studie greift bisherige empirische Befunde auf und führt sie weiter, indem Intelligenz und EF als Prädiktoren von SRL in zwei Altersgruppen, nämlich bei Vor- und Grundschulkindern, untersucht werden.

Die Stichprobe umfasste $n = 262$ Vorschulkinder ($M_{\text{Alter}} = 5.8$ Jahre, $SD_{\text{Alter}} = 0.41$, $\text{Range}_{\text{Alter}} = 5.0 - 6.7$ Jahre, 54.0% weiblich) und $n = 224$ Grundschulkindern ($M_{\text{Alter}} = 9.9$ Jahre, $SD_{\text{Alter}} = 0.45$, $\text{Range}_{\text{Alter}} = 8.8 - 12.1$ Jahre, 49.5% weiblich). SRL wurde mittels eines Strategiewissenstests erfasst (SRL-SKT; adaptiert nach Dörrenbächer-Ulrich et al. 2024). Fluide Intelligenz wurde mit einem Test zum logischen Schlussfolgern erhoben (Kaufman und Kaufman 2015). Die Messinstrumente für EF umfassten je zwei Aufgaben für das Arbeitsgedächtnis (z.B. Visuell-räumliche Spannaufgabe; Kubota et al. 2020), Inhibition (z.B. Go/No-go-Aufgabe; Johann et al. 2022) und kognitive Flexibilität (z.B. Flexible Item Selection Task; Kuhn et al. 2017).

Die Untersuchung der Zusammenhänge zwischen den Konstrukten erfolgte mittels Strukturgleichungsmodellierung mit den latenten Prädiktoren fluide Intelligenz und EF und dem latenten Kriterium SRL. Bei den Vorschulkindern zeigte das Modell einen sehr guten Fit ($\chi^2(df) = 99.76$ (85), $p = .13$, CFI: .98, RMSEA: .03 (90% CI: .01 -.04), SRMR: .06), ebenso bei den Grundschulkindern ($\chi^2(df) = 93.7$ (85), $p = .24$, CFI: .98, RMSEA: .02 (90% CI: .01 -.04), SRMR: .06). Die latente Korrelation zwischen EF und fluider Intelligenz war hoch (Vorschule: $r = .55$, $p < .001$; Grundschule: $r = .70$, $p < .001$). EF war ein signifikanter Prädiktor für SRL (Vorschule: $\beta_{\text{EF}} = .43$, $p < .05$; Grundschule: $\beta_{\text{EF}} = .53$, $p < .05$), fluide Intelligenz hingegen nicht (Vorschule: $\beta_{\text{INT}} = -.12$, $p = .33$; Grundschule: $\beta_{\text{INT}} = -.24$, $p = .22$). Die Varianzaufklärung in SRL betrug in der Substichprobe der Vorschulkinder $R^2 = 14\%$ und in der Substichprobe der Grundschulkindern $R^2 = 16\%$. Es handelt sich hierbei um vorläufige Ergebnisse.

Kindern mit effizienteren EF gelang es besser, in einem Lernszenario-basierten Strategiewissenstest nützliche von weniger nützlichen SRL-Strategien zu unterscheiden. Der Befund ist im Einklang mit früheren Befunden im Vor- und Grundschulalter (Rutherford et al. 2018; Davis et al. 2021). Fluide Intelligenz leistete keinen signifikanten Beitrag zur Vorhersage SRLs in beiden Altersgruppen (Dresel und Haugwitz 2005; Sontag et al. 2012). Eine Erklärung hierfür ist, dass lediglich fluide Intelligenz als Maß für die allgemeine kognitive Fähigkeit herangezogen wurde. Angesichts von Studien, die Sprachfähigkeit als Vorläuferfähigkeit von SRL betrachten (Jacob et al. 2019), ist verbale Intelligenz eine Fähigkeit, die potenziell in engerem Zusammenhang zu SRL steht, und sollte in künftigen Studien berücksichtigt werden. Überdies ist eine multimethodale Erfassung SRLs anzustreben (vgl. Perels et al. 2020), um die Rolle kognitiver Fähigkeiten für die tatsächliche Nutzung von SRL aufzudecken.

Heterogenität in Metakognitiven Regulationsfähigkeiten und ihr Einfluss auf das Lernen in der Sekundarschule

Katharina Zimmermann, Eileen Rueegg, Mariette van Loon
Universität Zürich

Metakognition beschreibt die Fähigkeit, den eigenen Lernfortschritt einzuschätzen (Überwachungskomponente) und darauf basierend das eigene Lernen effektiv anzupassen (Selbstregulationskomponente) (Nelson, 1990). Metakognition spielt eine entscheidende Rolle beim Lernen und beim Schulerfolg (Dent & Koenka, 2016). Ohtani & Hisaska's (2018) Metaanalyse konnte darüber hinaus zeigen, dass Metakognition ein noch stärkerer Prädiktor von akademischem Erfolg ist als der Intelligenzquotient. Diese Erkenntnisse verdeutlichen damit nicht nur die Bedeutung von metakognitiven Fähigkeiten im Schul- und Lernkontext sondern auch ihre gesellschaftliche Relevanz zur Förderung von Chancengleichheit.

Während die metakognitive Überwachung bis zur Erreichung des Jugendalters relativ genau ist, legen Studien nahe, dass Jugendliche dennoch Schwierigkeiten haben, diese genauen Überwachungseinschätzungen in effektive Selbstregulation umzuwandeln, wodurch sich grosse Unterschiede in Selbstregulationsfähigkeiten zwischen Jugendlichen zeigen (van Loon & Roebbers, 2017), was darüber hinaus auch durch latente Profilanalysen verdeutlicht werden konnte (Robey, 2020). Ansätze wie die Agendabasierte Regulation (ABR) (Ariel et al., 2009) und das Effort Monitoring Regulation (EMR) Framework (de Bruin et al., 2020) versuchen zu verstehen, welche Strategien zur Selbstregulation verwendet werden, wie die Variabilität dieser zu erklären ist und inwiefern beispielsweise kognitive Belastung oder motivationale Faktoren die Selbstregulation beeinflussen.

Variabilität in metakognitiven Fähigkeiten zeigt sich jedoch nicht nur auf interpersoneller Ebene sondern bildet sich auch innerhalb von Personen ab. So konnten neuere Untersuchungen mit Längsschnittdaten zeigen, dass Metakognition nur moderat stabil über die Zeit hinweg ist und zwischen Messzeitpunkten schwankt (van Loon et al., 2022). Darüber hinaus zeigen Studien, die Selbstregulation via Selbstbeurteilung erfasst haben, dass starke tägliche Schwankungen in der Selbstregulation mit niedrigerem akademischen Erfolg assoziiert sind und die Untersuchung von Schwankungen in Lernfähigkeiten auf intraindividuelle Ebene ebenfalls erforderlich ist (Blume et al., 2022; Malmberg et al., 2016).

Diese Studie ist die erste, die Variabilität in metakognitiven Fähigkeiten während der direkten Bearbeitung einer Aufgabe untersucht. Dazu haben wir eine Studie mit mehreren Messzeitpunkten durchgeführt, bei der 124 Sekundarschüler (7. Schulstufe, $MA_{\text{Alter}}=13.27$, $SD_{\text{Alter}}=0.83$) aus öffentlichen Schulen der deutschsprachigen Schweiz rekrutiert wurden. Die Schüler*innen haben dabei an drei Messzeitpunkten im Abstand von drei Wochen Kanji (japanische Schriftzeichen) gelernt. Jeder Messzeitpunkt beinhaltete eine Lernphase, eine Testphase mit Überwachungseinschätzungen (für jede Antwort wurde eingeschätzt, wie sicher sich die Proband*innen sind, dass ihre Antwort korrekt ist) und einer optionalen Wiederholungsphase, bei der bereits gelernte Kanji ausgewählt werden konnten, die erneut gelernt werden möchten. Nachdem zehn Kanji gelernt wurden, konnte jedes Mal entschieden werden, ob die Kanji-Aufgabe beendet werden soll oder weitergelernt werden möchte. Dies diente zur Messung von Aufgabenpersistenz. Am Ende jedes Messzeitpunktes erfolgte die Messung der mit der Kanji-Aufgabe verbundenen kognitiven Belastung. Dabei sollten die Proband*innen die von ihnen wahrgenommene Anstrengung und Schwierigkeit der Kanji-Aufgabe sowie ihre Motivation, die Aufgabe zu bearbeiten, einschätzen.

In der Datenanalyse werden die Überwachungsgenauigkeit sowie effektive Selbstregulation (Einteilung der Lernzeit, Auswahl für Wiederholungslernen, Aufgabenpersistenz) und die Kanji-Aufgabenleistung in einer Mehrebenenmodellierung für die Analyse von Messwiederholungsdaten unter Berücksichtigung von zufälligen Item- und Messzeitpunkteffekten untersucht. Darüber hinaus wird die genestete Struktur der Daten mittels hierarchischer Mehrebenenanalyse kontrolliert. Dabei werden Zusammenhänge zwischen der Variabilität in metakognitiven Fähigkeiten und der Aufgabenleistung auf intra- und interindividuellen Ebene über drei Messzeitpunkte hinweg untersucht. Zudem soll genauer beleuchtet werden, ob ein hohes Mass an wahrgenommener Belastung und Schwierigkeit der Aufgabe mit einer weniger effizienten Selbstregulation einhergeht. Die Datenerhebung ist bereits abgeschlossen und die Daten werden derzeit aufbereitet (z. B. Kodierung der Testantworten). Auf der Konferenz werden wir die Resultate von metakognitiver Überwachungsgenauigkeit und effizienter Selbstregulation sowie der Einfluss ihrer Variabilität auf die Aufgabenleistung und das Lernen berichten. Intraindividuelle Schwankungen sowie interindividuelle Unterschiede in metakognitiven Fähigkeiten besser zu verstehen, kann dabei helfen, spezifischen Bedarf zur Förderung metakognitiver Fähigkeiten einzelner Schüler*innen zu erkennen und Strategien zu entwickeln, die diese beim Lernen bestmöglich unterstützen sollen.

Der Zusammenhang zwischen Metakognition, Arbeitsgedächtnis und Lernerfolg

Eileen Rueegg, Katharina Zimmermann, Mariette van Loon
Universität Zürich

Modelle der Metakognition gehen davon aus, dass die Einschätzung des eigenen Lernens, das metakognitive Monitoring, die strategische Regulation des Lernens beeinflusst (Nelson & Narens, 1990). Beispielsweise ist das erneute Lernen von unzureichend gelerntem Lernmaterial nur möglich, wenn man einschätzen kann, welches Lernmaterial man noch nicht beherrscht. Diese auf Monitoring basierenden strategischen Entscheidungen, Lernmaterial erneut zu lernen, wirken sich positiv auf die Lernleistung aus (van Loon & Oeri, 2023). Diese metakognitiven Fähigkeiten sind prädiktiv für den Lernerfolg, sowohl auf der Aufgabenebene als auch auf der Ebene der akademischen Leistung (Ohtani & Hisasaka, 2018; van Loon & Oeri, 2023). Metakognitive Fähigkeiten gewinnen insbesondere im Jugendalter an Bedeutung, da mit dem Übergang in die Sekundarstufe erhöhte Anforderungen an das selbstständige Lernen gestellt werden.

Die bisherige Forschung hat gezeigt, dass es große individuelle Unterschiede in den metakognitiven Fähigkeiten gibt. Unterschiede in kognitiven Fähigkeiten wie dem Arbeitsgedächtnis könnten diese individuellen Unterschiede erklären. Das Arbeitsgedächtnis wurde in mehreren Studien mit Metakognition in Verbindung gebracht (z.B. Bryce et al, 2015, Coutinho et al., 2015). Es wird angenommen, dass insbesondere die strategische Regulation, die eine exekutive Komponente beinhaltet, mit dem Arbeitsgedächtnis zusammenhängt (Kálin & Roebbers, 2022). Der vermutete Zusammenhang wurde bisher jedoch überwiegend bei jüngeren Kindern untersucht, mit uneinheitlichen Resultaten.

Da neben der Metakognition auch das Arbeitsgedächtnis einen Einfluss auf die Lernleistung hat (Swanson & Alloway, 2012), untersucht diese Studie den gemeinsamen sowie den getrennten Einfluss dieser Fähigkeiten auf die Lernleistung von Jugendlichen, um individuelle Unterschiede im Lernen besser zu verstehen.

Die folgenden Hypothesen werden untersucht: 1) Die Auswahl erneut zu lernender Items basiert auf Monitoring Judgments. 2) Die Arbeitsgedächtnisleistung erklärt Varianz in der strategischen Auswahl erneut zu lernender Items. 3) Die Genauigkeit des Monitorings, die strategische Auswahl erneut zu lernender Items und die Arbeitsgedächtnisleistung erklären Varianz in der Lernleistung.

N = 85 SchülerInnen der 7. Klasse aus sieben verschiedenen Klassen in fünf Sekundarschulen haben an der Untersuchung teilgenommen. 43.5% der Teilnehmenden sind weiblich, das Durchschnittsalter beträgt 13 Jahre. Die digitale Lernaufgabe bestand aus dem Lernen von schwierigen Konzepten und ihren Definitionen. Hierbei konnten die Teilnehmenden in einer Lernphase von vier Konzepten ihre Zeit frei einteilen und mussten in einer Testphase zusätzlich zu ihren Testantworten beurteilen (Monitoring Judgments), wie sicher sie sich waren, dass ihre Antwort richtig war. Anschließend konnten sie auswählen, welche Konzepte sie noch einmal lernen wollten, um ihre Testantwort zu verbessern. Die Genauigkeit des Monitorings (Übereinstimmung der Monitoring Judgments mit der Lernleistung auf Itemebene), sowie die individuellen, strategischen Entscheidungen, Items erneut zu lernen (Auswahl von Items mit niedrigen Monitoring Judgments) werden berechnet. Ein Operation Span Task (OST) sowie ein Reading Span Task (RST) wurden bearbeitet, um das Arbeitsgedächtnis der SchülerInnen zu messen.

Die Hypothesen werden mit Korrelationen sowie Multi Level-Regressionsanalysen untersucht. Dabei werden die Itemschwierigkeit sowie die Klassen- und Schulzugehörigkeit berücksichtigt. Vorläufige Resultate zeigen eine signifikante

Korrelation zwischen der strategischen Auswahl der erneut zu lernenden Konzepte und der Arbeitsgedächtnisleistung (OST: $r = .35$, $p < .01$; RST: $r = .31$, $p < .05$). Die Datenanalyse wird aktuell durchgeführt. Die vollständigen Ergebnisse werden auf der Konferenz vorgestellt.

Die vollständigen Ergebnisse werden Aufschluss darüber geben, wie Metakognition und Arbeitsgedächtnisleistung zusammen die Lernleistung beeinflussen. Dieses Wissen trägt zu einem umfassenderen Verständnis der Faktoren bei, die selbständiges Lernen beeinflussen können, um SchülerInnen in ihrem Lernprozess bestmöglich zu unterstützen.

S-64: Understanding Student Motivation, Emotion, and Engagement in Higher Education: Perspectives from Longitudinal and Context-specific Studies

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 165

Chair der Sitzung: Luise von Keyserlingk

Chair der Sitzung: Jakob Schwerter

Chair der Sitzung: Stefan Janke

English Track

Understanding Student Motivation, Emotion, and Engagement in Higher Education: Perspectives from Longitudinal and Context-specific Studies

Chair(s): **Luise von Keyserlingk** (Universität Tübingen, Deutschland), **Jakob Schwerter** (TU Dortmund)

Diskutant*in(nen): **Stefan Janke** (Universität Mannheim)

College is an important qualification period for individuals' educational trajectories and for the education of new generations of a society. Motivation is central to students' learning, performance, and educational choices (Wigfield & Eccles, 2020) and maintaining high levels of motivation throughout college is of great importance for students and higher education institutions. However, motivation can be subject of change over short and long time periods, and it can vary across learning situations, such as weekly sessions or exercises (Benden & Lauermann, 2022; Gaspard et al., 2017). The symposium brings together four empirical studies that examine changes and variation in students' domain-specific motivation based on the Situated-Expectancy-Value-Theory (Eccles & Wigfield, 2020) in different college settings. The studies provide different perspectives on the role of motivation for students' college success and learning processes, focusing on different educational college contexts, different timeframes, and interrelations with other central research fields, such as self-regulated learning and achievement emotions:

Two papers focus on German students in specific majors (study 2: Applied Mathematics; study 3: Educational Psychology), one on German STEM students (study 1) and one on U.S. students across majors (study 4). All contributions use longitudinal data to investigate associations of motivation and academic college outcomes. The studies examine course-specific motivation across different time points of lectures (all studies), and situation-specific motivation across exercises (study 1, 2) to advance our understanding of variation and changes in motivation across time and learning situations.

All four papers examine associations between motivation and course grades for college success. In addition, paper 1-3 examine the role of student motivation for learning processes, the engagement with course materials and the use of self-regulated learning strategies within lectures and courses. Paper 4 focuses on changes in motivation across multiple courses and investigates the role of motivational change for broader college outcomes, such as dropout intentions and college major change. Because motivation is an important, but not the only driver of learning behavior and success, the contributions also examine the interrelation of motivation with other central psychological constructs related to learning and success, such as self-regulation (study 3), engagement (study 2,3), and achievement emotions (study 2).

The symposium brings together studies that advance our understanding of variability and change of student motivation in college contexts, and how these are related to students' learning process, performance, and retention. The contributions will be discussed by an expert in the field of higher education research.

Beiträge des Symposiums

How Relevant is Math for My Field of Study? Context-Dependent Differences in Students' Motivation and Engagement With Math Problems Across STEM Fields

Daria Benden, Fani Lauermann

Universität Bonn

Theoretical Background and Research Questions

Math-related achievement difficulties significantly contribute to high student dropout rates from science, technology, engineering, and mathematics (STEM) fields (Heublein et al., 2017; Müller et al., 2018). Many postsecondary institutions and instructors offer targeted interventions to improve students' math skills at the beginning of their studies, such as diagnostic math self-tests and quizzes across the semester (Kinnear et al., 2022). However, the utility of these self-tests depends on the quality of students' test motivation and engagement, which can vary widely. Evidence suggests that students who would benefit the most from self-testing—those with low math motivation and achievement—are least likely to engage with it (Schwerter et al., 2022). Furthermore, the (perceived) relevance of the tested content may vary across courses and STEM programs, potentially affecting the context-specific utility of math self-tests.

Based on Eccles' situated expectancy-value theory (Eccles & Wigfield, 2020), this study therefore examined individual and context-dependent differences in students' test motivation and test engagement, and their links to students' performance in gateway math courses.

Method

The sample of the present study included STEM students enrolled in gateway math courses for their respective study program ($N = 2,165$; $k = 20$ math courses). The students completed a comprehensive statewide test of math prerequisites for STEM fields during class time. Students' motivational beliefs (i.e., ability self-concept, intrinsic value) and perceived relevance of the tested content were assessed four times during the test and averaged for the main analyses. Students' test performance and test engagement were operationalized as the number of correctly solved math problems and the number of math problems students attempted to solve, respectively.

Multilevel modeling was used to account for the nested structure of the data with students (Level 1) nested in math courses (Level 2). Students' gender (33% female), advanced math course participation (52%), and final math grade in high school ($M = 10.4$ points, 15-point-scale) were included as control variables in all models.

Results

Multilevel analyses revealed that 12%-28% of the variance in students' test-specific motivational beliefs, test engagement, and test performance were accounted for by differences between math courses/study programs, suggesting that students' test-specific motivational beliefs, test engagement/performance, and the perceived relevance of math for students' study programs depended on the type of math course and study program.

In addition, multilevel path analyses indicated that, controlling for differences in students' prior math achievement and test performance, students' perceived relevance of the tested content significantly predicted their ability self-concept and intrinsic interest while working on the math test. Thus, students who saw greater relevance of math for their study program than others in their math course were more optimistic about their math abilities and more interested in the tested content.

In addition, after accounting for interindividual differences in students' prior math achievement, test performance, and perceived relevance of the tested content, the course-aggregated perceived relevance of math to their study program significantly predicted students' intrinsic interest in and engagement with the math test (i.e., the number of math problems students attempted to solve). Controlling for performance differences, students' individual test engagement was higher in contexts that valued math. However, students' intrinsic interest in the tested content was significantly lower in high-math-relevance contexts, possibly due to performance pressures. Finally, students' test performance and test motivation/engagement significantly predicted their end-of-term course performance.

These findings underscore the importance of considering motivational aspects for the utility of self-tests. Ensuring that the tested content is relevant to students' specific contexts and study programs and highlighting this relevance before administering self-tests is likely to enhance students' quality of test engagement.

Situational and Longitudinal Associations Between Appraisals, Anxiety, and Learning during Self-testing

Elisabeth Graf¹, Jakob Schwerter¹, Daria Benden², Fani Lauermann²

¹TU Dortmund, ²Universität Bonn

Theoretical background

Self-testing is an effective learning strategy in college, still students often fail to implement it during their studies (Dunlosky et al., 2015). Understanding motivational and emotional processes during self-testing situations can provide crucial information on how students' self-testing behavior could be enhanced. However, little evidence exists on the complex interplay between motivation (expectancy-value appraisals), emotions (anxiety), and learning (engagement and achievement) during authentic self-testing exercises. Building on situated expectancy-value theory (Eccles & Wigfield, 2020) and control-value theory (Pekrun et al., 2023), we focus on students' control appraisal (operationalized as expectancy beliefs) and task values (including their interactions), along with anxiety to investigate their associations with students' engagement and achievement in self-testing across a semester in a gateway math course.

Research question

The aims of the study are twofold: First, we combine both theoretical perspectives to a common model testing concurrent associations between control- and value appraisals and anxiety, and their associations with engagement and achievement. Specifically, we inspect the situative (i.e., intraindividual) associations, including the interaction between control- and value appraisal proposed by both theories. Additionally, we aimed to investigate possible feedback loops, namely whether engagement and achievement predict motivation and anxiety in the upcoming week. Second, we were interested in the interindividual associations, specifically whether students' trait and aggregated state motivation, anxiety, engagement, and achievement predict students' final course performance.

Method

We use data from students of an introductory applied math course at a German university (N = 883, 51% female). At the beginning and end of the semester, students reported their social demographics, trait motivation, and anxiety. During the semester, the course offered weekly online self-testing exercises. At the end of each self-testing exercise, situative motivational beliefs (expectancy of success, intrinsic and utility value, cost) and anxiety focusing on the self-testing situation were reported. Students' achievement on each self-testing exercise was recorded. In addition, different indicators of engagement were recorded, including the percentage of processed tasks, time on task, time to deadline, and engagement with additional study material (e.g., worksheets, instructional videos, lectures), and combined to one latent engagement variable. Performance on the final course exam is used as a measure for overall course performance. To test our hypotheses, we used multilevel and dynamic structural equation modeling in Mplus.

Results

Across the semester, we found decreasing motivation, engagement, and achievement, while anxiety and cost increased. Utility value and control were positively, but intrinsic value negatively associated with engagement at the intraindividual level. Further, intrinsic-, utility value and control were negatively, and cost positively related to anxiety. The effects held when including autoregressive effects, and revealed positive feedback loops from engagement to intrinsic value, utility value, and control. The intraindividual models on achievement revealed positive associations between intrinsic-, utility value, control, and achievement, and negative associations with anxiety. Surprisingly, we found negative feedback loops from achievement to intrinsic value and control, and a positive feedback loop on cost. Thus, while participation in self-testing tends to increase motivation in the coming week, achievement seems to dampen motivation. Analyses of the hypothesized interaction effects between control and task value were largely confirmed for achievement and anxiety.

Finally, we were interested in how trait- and state-variables affect final course performance. Though utility value (both trait and state) showed negative effects, we found positive effects of state achievement and control. In the model with engagement, aggregated state engagement and cost positively predicted students' exam performance. In a final overall model, only effects from engagement and utility value remained significant. In sum, the results stress the importance of student engagement throughout the semester, both for ongoing motivation and for final course performance.

Motivation Matters - Interaction of Motivational Trajectories and the Use of Self-Regulated Learning Activities in University Students

Zoe Sander¹, Birgit Spinath²

¹Universität Heidelberg, ²Universität Heidelberg

Theoretical background

In higher education, students are independent learners and must organize and monitor their learning behavior without external supervision (Larose et al., 2005). Therefore, university represents a learning environment in which learners are required to have particularly high self-regulatory skills (Bosch et al., 2021). Based on the model of self-regulated learning strategies by Zimmerman and Martinez-Pons (1986), we implemented several evidence-based learning activities in an introductory lecture course. These included preparing lecture slides (seeking information), class attendance, reviewing lecture notes (reviewing records), reviewing

additional literature (reviewing records), practice tests (rehearsing and memorizing / monitoring), submitting practice essays (rehearsing and memorizing / monitoring), study group / consulting peers (seeking social assistance). As theorized in the component model of self-regulated learning, the application of these self-regulatory learning strategies also depends on motivational prerequisites (Schmitz & Wiese, 2006). Previous research has investigated this relationship at the sample level and cross-sectionally (e.g., Bosch et al., 2021; Putwain et al., 2019). With regard to the heterogeneity of students' motivational processes, we expanded these findings using a person-oriented approach.

Research question

This study pursued three main goals. First, we examined whether different profiles of longitudinal trajectories of expectancies and values emerged in a sample of university students and how these profile paths are characterized (RQ 1). The second aim was to investigate how profile membership relates to evidence-based self-regulated learning activities that were implemented in the lecture course (RQ 2). Finally, we investigated whether profile membership and the use of self-regulated learning strategies were related to course dropout intention, performance and perceived learning success (RQ 3).

Method

The first sample consisted of N = 338 teacher and psychology students who were enrolled in an introductory lecture course in educational psychology. Based on the model by Zimmerman and Martinez-Pons (1986), students were offered the aforementioned self-regulated learning activities. Using the theoretical framework of Situated Expectancy-Value Theory (Eccles & Wigfield, 2020), we assessed students' expectancies and task values at 5 measurement points across one academic term. In order to record the learning strategies as timely and process-oriented as possible, we also assessed learning activity use in these 5 surveys. We applied longitudinal latent class growth analysis to differentiate characteristic patterns of motivational development and their relationship with using self-regulated learning activities. Following the current semester, further N = 237 teacher students will be included in the analyses.

Results

For the expectancy component, the two-class solution showed the best model fit: Moderate stable (76.6%) and Steady decrease (23.4%). Students in the Moderate stable group were more likely to attend class ($\chi^2(1, N = 334) = 4.45, p = .035$) and to review lecture notes ($\chi^2(1, N = 334) = 3.99, p = .046$). Regarding task values, four longitudinal profile paths were identified: Stable very high (34.2%), Stable high (43.4%), Moderate early decrease (15.3%) and Sharp early decrease (7.1%). As expected, students with more favorable value trajectories were more likely to attend class ($\chi^2(3, N = 334) = 8.55, p = .036$), review lecture notes ($\chi^2(3, N = 334) = 14.03, p = .003$) and submit practice essays ($\chi^2(3, N = 334) = 16.85, p = .001$). Associations with the aforementioned learning outcomes will be analyzed when the data of the last measurement point is available.

Is Change in Motivation in College a Normative Calibration Process or an Early Warning Sign?

Luise von Keyserlingk¹, Hye Rin Lee², Charlott Rubach³, Hanna Gaspard⁴, Gabe Orona¹, Jutta Heckhausen⁵, Jacque S. Eccles⁵

¹Universität Tübingen, ²University of Delaware, ³Universität Rostock, ⁴TU Dortmund, ⁵University of California, Irvine

Theoretical background:

Students with higher academic motivation have better performance outcomes, show higher effort investment, and are less likely to drop out of courses (Benden & Lauermann, 2022; Wigfield & Eccles, 2020). However, empirical studies in higher education have shown declines in students' expectancies and subjective task values (STV), and increases in perceived costs over time, typically in the first college semester (Benden & Lauermann, 2022; Robinson et al., 2019). Shortly after the transition to college, a change in motivation could be interpreted as a normative calibration to a new and challenging learning environment in college. However, a continued decrease in motivation could be worrisome, as it is linked to undesirable academic outcomes.

Research question:

We use data from a longitudinal study from a public university in California. Students reported their expectancy value beliefs at the beginning, middle and end of their most difficult course per quarter in three consecutive quarters in their freshman year (fall, winter, spring). We examined how students' expectancies, STV, and cost values changed across the quarters and tested whether change was larger in the first college quarter, compared to their second and third quarter as a function of a normative adjustment process (RQ1). We examined person-specific, and course-specific predictors of change (RQ2), and we examined how change in motivation was related to students' performance and retention (RQ3).

Method:

Our sample consisted of 941 freshman students. We used latent change score models in Mplus.

Results:

In the three quarters of the freshman year, students' expectancies declined in the first half of the quarter (change T1-T2 expectancies: $b = -.12$ to $-.18$; $SE = .04$ to $.05$; $p < .01$) and recovered in the second half (change T2-T3 in expectancies: $b = .23$ to $.34$; $SE = .04$ to $.06$; $p < .01$). STV declined in the first half of the quarter (change T1-T2 STV: $b = -.46$ to $-.30$; $SE = .05$; $p < .01$), and had no change or a small increase in the second half (change T2-T3 STV: $b = -.002$ to $.18$; $SE = .05$ to $.06$; $p < .01$ to $.96$). Cost values declined in the first half of the quarter (change T1-T2 cost: $b = -.56$ to $-.60$; $SE = .05$ to $.06$; $p < .01$) and did not change in the second half. The decline in expectancies and cost values did not differ between the first college quarter and subsequent quarters. The decline in STV was larger in first college quarter than the third quarter. Furthermore, T1 expectancies and STV were higher and cost values were lower in the first college quarter, than in the second and third quarter.

Students' who performed better in their midterm exams had more positive changes in expectancies and STV in the first half of the quarter. A more positive change in expectancies and STV was related to better course grades and higher intentions to stay in the major. Associations between changes in cost value and intentions to stay in major and final grades were less consistent across terms, but pointed into the expected direction. Results show a trend that increases in cost values were related to higher dropout intentions and lower grades.

Findings indicate that declines in motivation are not only a result of a normative adjustment processes shortly after the transition to college, but continue across the entire academic year. While change patterns across the quarters were highly similar in all quarters, the initial motivation at the beginning of the course was lower in students' second and third quarter at college.

S-65: Partizipation und Transfer in der kulturellen Bildung

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: O 151

Chair der Sitzung: Lisa Birnbaum

Chair der Sitzung: Katrin Rakoczy

Partizipation und Transfer in der kulturellen Bildung

Chair(s): **Lisa Birnbaum** (Friedrich-Alexander-Universität Erlangen-Nürnberg)

Diskutant*in(nen): **Katrin Rakoczy** (Justus-Liebig-Universität Gießen)

Gesellschaftliche Transformationen werden in der kulturellen Bildung aus verschiedenen Perspektiven und multimethodisch betrachtet (Gaede et al., 2023), denn kulturelle Bildung zielt darauf ab, die Teilhabe an Kultur in ihrer Breite und Vielschichtigkeit zu ermöglichen. Darüber hinaus ist auch Partizipation an der Weiterentwicklung von kulturellen Bildungsangeboten möglich, welche sich aufgrund gesellschaftlicher Veränderungen kontinuierlich wandeln (Birnbaum, Klepacki & Kröner 2023; Müller & Kammerl, 2022). Hier setzt das Symposium an und nimmt in den Blick, inwiefern Partizipation an kultureller Bildung und an der Entwicklung von kulturellen Bildungsangeboten in verschiedenen Disziplinen und Professionen bislang realisiert wird.

Kulturelle Bildungsforschung erfolgt aus verschiedenen disziplinären und methodischen Perspektiven und involviert neben Wissenschaftler:innen auch häufig Praxisakteur:innen. Partizipative Ansätze werden genutzt, um als Wissenschaft und Praxis gemeinsam im Feld der kulturellen Bildung zu agieren und zur sozialen Integration und zur Stärkung demokratischer Einstellungen beizutragen. Somit bietet es sich an, in der Entwicklung von kulturellen Bildungsangeboten aber auch in der Entwicklung von Forschungsprojekten im Bereich der kulturellen Bildung zu überlegen, inwiefern partizipativ angelegte Arbeitsphasen sowohl zwischen verschiedenen Disziplinen als auch zwischen Wissenschaft und Praxis einen Mehrwert leisten können (Birnbaum, Klepacki & Kröner 2023; Senabre Hidalgo et al. 2021).

Dieses Symposium widmet sich der Frage, wie die Partizipation verschiedener Akteur:innen zur Gestaltung kultureller Bildungsprozesse beitragen können und wie durch mehrdimensionale Zusammenarbeit der Wissenschafts-Praxis-Transfer gestärkt werden kann. Die Beiträge des Symposiums liefern Erkenntnisse zur Partizipation an kulturellen Bildungsangeboten und zur Realisierung von Transfer in der kultureller Bildung zwischen unterschiedlichen Akteur:innen:

- Im ersten Beitrag wird eine Forschungssynthese vorgestellt, in der die Forschung zu kultureller Bildung und kultureller Partizipation in ländlichen Räumen synthetisiert wird und unter die Lupe nimmt, inwiefern die Praxisperspektive in Forschungsansätzen einbezogen wird.
- Der zweite Beitrag analysiert empirische Forschungsergebnisse zur kulturellen Bildung in ländlichen Räumen. Es wird analysiert, wie eine praxis- und transferorientierte Umsetzung von Forschungsergebnissen so erfolgen kann, dass auch gesellschaftliche Machtfragen, strukturelle Defizite und demokratische Herausforderungen zusammen mit kultur- und bildungspolitischen Überlegungen reflektiert werden.
- Der dritte Beitrag gibt Einblick in aktuelle Forschungsprojekte aus der kulturellen Bildung und wie von der Forschungsseite aus Projekte geplant werden, um Praxisakteur:innen einzubeziehen bzw. Partizipation und Transfer zwischen Wissenschaft und Praxis zu ermöglichen.

Basierend auf den vorliegenden Befunden werden verschiedene Kollaborationsmöglichkeiten zwischen Forschung und Praxis aufgezeigt, die dienlich sein können, um evidenzbasierte kulturelle Bildung und nachhaltig angelegte Bildungsangebote über Disziplin- und Professionsgrenzen hinweg zu stärken. Es wird diskutiert wie gelingende, transformative Bildungsprozesse innerhalb der kulturellen Bildung und darüber hinaus gestaltet werden können.

Beiträge des Symposiums

Kulturelle Partizipation in ländlichen Räumen: Eine Forschungssynthese

Jessica Fittkau, Lisa Birnbaum, Stephan Kröner

Friedrich-Alexander-Universität Erlangen-Nürnberg

Die Verfügbarkeit von Möglichkeiten zur Auseinandersetzung mit Kunst und kulturellem Erbe kann zur Lebensqualität (Tjarve & Zemite, 2016) und zur nachhaltigen Entwicklung (Deutscher Kulturrat, 2019) beitragen. Durch die Förderung der kulturellen Bildung und Partizipation können soziale Strukturen gestärkt, ökologische Bewusstseinsbildung unterstützt und wirtschaftlich Impulse gesetzt werden (Stoltenberg, 2010). Kulturelle Partizipation umfasst dabei nicht nur die Nutzung kultureller Angebote, sondern auch deren Mitgestaltung und Organisation. Gerade in ländlichen Regionen kann das Mitgestalten an Kultur und kultureller Bildung eine wichtige Rolle spielen, um den Herausforderungen solcher Räume wie der Abwanderung junger Menschen, dem demographischen Wandel und einem begrenzten Zugang zu kulturellen Angeboten zu begegnen.

Vor diesem Hintergrund stellen sich folgende Fragen: In welcher Weise kann kulturelle Partizipation verschiedene Dimensionen der nachhaltigen Entwicklung in ländlichen Räumen fördern? Und welche Forschungserkenntnisse liegen zur Mitgestaltung an kulturellen Angeboten in ländlichen Räumen vor, insbesondere auch hinsichtlich partizipativer Forschungsmethoden?

Methode

Eine Literaturrecherche in den Datenbanken Scopus und ERIC ergab n = 1165 Treffer, von denen n = 44 die Einschlusskriterien (ländliche Räume, Kultur oder kulturelle Bildung, Partizipation) erfüllten. Die eingeschlossenen Artikel wurden kategorisiert nach (a) Form der Partizipation (z. B. rezeptiv, aktiv und ko-konstruktiv), (b) angesprochenen Nachhaltigkeitsaspekten (sozial, kulturell, ökonomisch und ökologisch), (c) abgedeckten Facetten von Künsten (bildende, darstellende, literarische Kunst und Musik) oder kulturellem Erbe (natürlich, materiell und immateriell) und (d) Forschungsdesigns (quantitativ, qualitativ und Mixed-Methods).

Ergebnisse

Die Kategorisierung zeigt, dass neben kulturellen Aspekten vor allem soziale, nachhaltige Entwicklungsaspekte durch kulturelle Partizipation in ländlichen Räumen betont werden. Soziale Entwicklung wie insbesondere die Stärkung der Gemeinschaft, Förderung nach Chancengleichheit und erhöhtes Wohlbefinden durch kulturelle Partizipation können als häufig genannte Ergebnisse identifiziert werden. Zudem decken die eingeschlossenen Artikel ein breites Spektrum kultureller Aspekte ab, welche sowohl die Partizipation an kulturellen Angeboten und Aktivitäten als auch die Mitgestaltung an solchen thematisieren. Dabei fokussieren qualitative Studien (n = 22) hauptsächlich auf die Partizipation und Mitgestaltung im Bereich der Kunst, während in quantitativen Studien (n = 7) überwiegend die aktive Partizipation im Kunstbereich im Vordergrund stand. Darüber hinaus nutzten n = 12 der Studien partizipative Forschungsmethoden und thematisierten damit den Wissenschafts-Praxis-Transfer.

Diskussion

Kulturelle Partizipation und Mitgestaltung von kulturellen Angeboten und Aktivitäten kann sowohl die kulturelle sowie die soziale nachhaltige Entwicklung von ländlichen Regionen unterstützen. Ein zentraler Aspekt in diesem Kontext ist der Einsatz von kreativ-partizipativen Forschungsmethoden, welche die Personen vor Ort in die Forschung einbeziehen. Solche Methoden könnten die Akzeptanz und Wirksamkeit kultureller Bildungsangebote in ländlichen Räumen erhöhen und so zu praxisnahen und umsetzbaren Ergebnissen führen. Zudem ist die Mitgestaltung an kulturellen Angeboten und Aktivitäten ein wichtiger Faktor für die Stabilität und Vitalität ländlicher Gemeinschaften. Weitere Forschung sollte sich daher auch langfristigen Effekte dieser Form von Partizipation (quantitativ) zu untersuchen und Gelingensbedingungen identifizieren.

Zusammenwirken von Forschung und Praxis in der kulturellen Bildung

Anna Heinemann, Jens Leber

Universität Duisburg-Essen

Der Begriff des Transfers wandelt sich vermehrt von einem linearen Verständnis zwischen Wissenschaft und Praxis hin zu einer Vorstellung, die die Komplexität von Bildungssituationen stärker berücksichtigt und in diesem Sinne auch bemüht ist, die Beziehungen zwischen Forschung und Praxis zu stärken (Malin & Brown, 2020). Denn trotz anhaltender Bestrebungen nach einer evidenzorientierten Bildungspraxis (z.B. Schmiedl, 2022; Biebert et al., 2018) ist weiterhin unklar, wie diese genau aussehen kann (OECD, 2022; Farley-Ripple et al., 2018, 237; Mohajerzad, Krzeminski & Schrader, 2022). Für ein gelingendes Zusammenkommen und -wirken wird mitunter ein gegenseitiges Verständnis für beide Communities und ihre Handlungspraktiken angestrebt (Farley-Ripple et al., 2018), um praxisrelevante Evidenz erzeugen zu können (Hartmann & Kunter, 2022, S. 2; OECD, 2022, S. 11).

Farley-Ripple et al. (2018) formulieren ein konzeptuelles Framework, welches die verschiedenen Perspektiven und Entscheidungen von Forschung (Evidenzproduktion) und Praxis (Evidenznutzung) miteinander in Beziehung setzen. Sie definieren sechs Dimensionen aus der Perspektive der Forschung, um einer evidenzbasierten Bildungspraxis entgegen zu arbeiten (evidence, search/dissemination, interpretation, participation, frequency, and stage of decision-making). Diese Dimensionen stehen jeweils analog einer gleichwertigen Dimension auf der Bildungspraxis entgegen (s.a. Hartmann & Kunter, 2022).

Das Ziel dieses Beitrags ist, das Verhältnis von Forschung und Praxis dementsprechend zu reflektieren: Inwiefern können Praktiker:innen an Forschungsprojekten teilhaben, um das Wirkungspotenzial von Forschung, sowohl im Sinne evidenzorientierter Bildungspraxis als auch einer gleichberechtigten Kollaboration im Handlungsfeld der kulturellen Bildung zu stärken? Beginnend bei der Gestaltung von Transferformaten von wissenschaftlichen Erkenntnissen in die Praxis sind über partizipative Teilprozesse in Forschungsprojekten auch ko-konstruktiv angelegte Gesamtvorhaben interessant für die Ausgestaltung von Partizipation im Wissenschaft-Praxis-Verhältnis. Der Fokus wird auf die Partizipationsdimension nach Farley-Ripple et al. (2018) gelegt und somit die Frage adressiert, wie Forschungsprojekte die aktive Einbeziehung von Praxisperspektiven ermöglichen können und welche Ziele mit den verschiedenen Partizipationsformaten verfolgt werden.

Vorgestellt werden Teilergebnisse der ersten Studie einer Interviewreihe, bei der im Sommer und Herbst 2024 die (Teil-) Projektleitungen einer Förderrichtlinie im Bereich der kulturellen Bildung befragt werden. Innerhalb der Konzeptionsphase in den Forschungsprojekten werden sie dazu interviewt, wie und warum ein aktiver Einbezug von Praxisperspektiven im Forschungsprojekt geplant wird. Geplant sind 22 Interviews mit (Teil-) Projektleitungen aus 11 Forschungsprojekten.

Die zugrundeliegende Forschungsfragen für diesen Beitrag lauten:

1. Welche partizipativ angelegten Formen der Einbeziehung von Praxisperspektiven lassen sich in den untersuchten Forschungsprojekten identifizieren?
2. Welche Gelingensbedingungen und hemmende Faktoren der Partizipation von Praxis im Forschungsprozess werden von Forschungsprojektleitenden genannt?

Zudem werden Zusammenhänge zwischen den Forschungsfragen sowie den geplanten Forschungsmethodiken in den Projekten untersucht. Ausgewertet werden die Interviews nach der qualitativen Inhaltsanalyse nach Mayring (2019). Es wird sowohl mit deduktiven als auch mit induktiv gebildeten Kategorien aus dem Datensatz gearbeitet. Deduktive Kategorien stellen hierbei z.B. die Dimension der Partizipation sowie die damit in Beziehung stehenden Dimensionen Interpretation, Häufigkeit und Dissemination (Farley-Ripple et al., 2018; Hartmann & Kunter, 2022) dar. Die Daten werden sowohl in anonymisierten Projektpfaden ausgewertet, um den Verlauf darstellen zu können (drei Erhebungszeitpunkte in drei Jahren) als auch projektübergreifend nach den Ko-Konstruktions- sowie Transferformaten und identifizierten Gelingensbedingungen sowie hemmenden Faktoren.

Bisherige Ergebnisse deuten darauf hin, dass es sich lohnen kann, Stimmen aus der Praxis bereits frühzeitig in Forschungsprojekten mit einzubinden oder auch bereits am Anfang von Forschungsprojekten darüber nachzudenken, wie das Wissen für die Praxis bereitgestellt werden soll und welche Zwecke erfüllt werden sollen, um Wissen zu identifizieren, welches als praxisrelevant eingestuft werden kann. Grundlegend sollte aber auch das Verhältnis von Wissenschaft und Praxis zueinander reflektiert werden, denn die Nähe zwischen beiden Systemen kann auch eine distanzierte Forschungsperspektive erschweren. Des Weiteren erscheint es wichtig, Übertragbarkeit und Abstraktion der Ergebnisse anzustreben, zum einen aufgrund des grundlegenden Forschungsauftrags und zum anderen um Insellösungen für einzelne Bildungspraxissituationen zu vermeiden.

Kulturelle Bildung und regionale Disparitäten: Impulse aus der Sozialraumforschung für erweiterte Perspektiven auf ländliche Räume

Nils Zimmer, Felica Selz

Universität Potsdam

Theoretischer Hintergrund

Im Bildungsbericht 2018 werden die Ausprägungen regionaler Unterschiede als zentrale Herausforderung für das deutsche Bildungssystem und damit auch für den Bereich der Kulturellen Bildung benannt (vgl. Autorengruppe Bildungsberichterstattung 2018). Dies bildete den Ausgangspunkt für die Förderrichtlinie „Kulturelle Bildung in ländlichen Räumen“ des Bundesministeriums für Bildung und Forschung (BMBF) und ihren 21 Projekten. Diese Projekte befassten sich mit unterschiedlichen Forschungs- und Praxisperspektiven mit verschiedenen Aspekten Kultureller Bildung in ländlichen Regionen in ganz Deutschland. Die Ergebnisse wurden in einem Sammelband zusammengeführt (Kolleck/Fischer 2023), der eine ganze Reihe bestehende Forschungslücken in diesem Forschungsfeld (vgl. das Literaturreview in Büdel/Kolleck 2023) füllen konnte.

Die beforschten Regionen unterscheiden sich in ihren geographischen, politischen, historischen und kulturellen Gegebenheiten. Gleichzeitig existiert eine Reihe an Gemeinsamkeiten und Überschneidungen, die charakteristisch für ländliche Gebiete sind. Wir fokussieren uns dabei auf ähnliche strukturelle Gegebenheiten und setzen sie in den Kontext aktueller Sozialraumforschung. Dabei gehen wir einerseits auf zentrale Herausforderungen ein, denen sich ländliche Räume vermehrt ausgesetzt sehen. Dazu zählen etwa Defizite im infrastrukturellen Bereich, die nicht nur die Umsetzung und Etablierung kultureller Bildungsangebote erschweren, sondern auch den Boden für eine Verfestigung autoritärer und demokratiefeindlicher Strukturen in den betreffenden Regionen ebnen. Andererseits beschäftigen wir uns mit relationalen Raumtheorien (Lefebvre 1974; Löw 2001). Der Raum existiert in dieser Perspektive nicht als feste, vorgegebene Entität. Er wird durch soziale Prozesse und Interaktionen erst produziert und fortwährend neu konstituiert. Demnach ist Raum eng mit gesellschaftlichen Machtverhältnissen verknüpft.

Für eine Praxis- und Transfer-orientierte Umsetzung der Forschungsergebnisse zu Kultureller Bildung in ländlichen Räumen müssen daher gesellschaftliche Machtfragen, strukturellen Defizite und demokratische Herausforderungen zusammen mit kultur- und bildungspolitischen Überlegungen für unterschiedliche ländliche Räume reflektiert werden.

Fragestellung

Dem Vortrag liegt folgende Leitfrage zugrunde: Welches Verständnis von ländlichen Räumen hat die Forschung der BMBF-Förderrichtlinie „Kulturelle Bildung in ländlichen Räumen“ hervorgebracht und wie lassen sich diese Erkenntnisse aus der Sozialraumforschung erweitern?

Methode

Der Beitrag basiert auf einer qualitativen Analyse der Projektergebnisse der Förderrichtlinie sowie auf einer Auseinandersetzung und Verknüpfung mit theoretischen Konzepten, die vor allem in der kritischen Stadt- und Humangeographie rezipiert werden und im deutschsprachigen Raum seit kurzem zur Formierung einer kritischen Landforschung beitragen.

Ergebnisse und Diskussion

Seit einigen Jahren ist im deutschsprachigen Raum eine Kritische Landforschung im Entstehen, die sich in der Tradition kritischer Gesellschaftstheorie wiegt und die Notwendigkeit einer Betrachtung ländlicher Räume fernab eines verkürzten Stadt-Land-Dualismus betont. Dabei hilft ein relationales Raumverständnis zu verstehen, welche Rolle (kulturelle) Bildung bei der Konstitution des Raumes spielt und inwiefern diese ihrerseits wieder Einfluss auf die Bedingungen zur Umsetzung und Etablierung von Bildungsangeboten nimmt. Allerdings hat die Forschung der vergangenen Jahre auch herausgestellt, wie fragil kulturelle Bildung in ländlichen Gebieten ist. Wenn vielfältige kulturelle Bildung in ländlichen Räumen wegbricht, entstehen Lücken und es besteht die Gefahr, dass diese mit einem engen, nationalistischen Kulturverständnis gefüllt werden. Das ist nicht zuletzt angesichts eines erstarkenden Rechtstextremismus wichtig, der sich vor allem in ländlichen Regionen zu manifestieren droht (vgl. exemplarisch Belina 2022; Simon 2020). Neben der Betrachtung spezifischer lokaler Kontexte müssen solche Entwicklungen jedoch immer auch als Ausprägungen eines gesamtgesellschaftlichen Phänomens begriffen werden (vgl. Mullis/Miggelbrink 2022).

Das Ziel unseres Beitrags ist es einerseits, empirische Forschungsergebnisse der Förderrichtlinie mit einer kritischen Perspektive auf ländliche Räume zu verknüpfen. Andererseits sollen durch die Abstraktion der Projektergebnisse neue Impulse für die weitere Forschung angestoßen werden und zu einer Verstetigung der Erkenntnisse im Sinne eines nachhaltigen Transfers beitragen.

S-66: Herausforderungen im Umgang mit Belastungen durch eine Stärkung personaler Ressourcen von Lehrkräften und Schüler:innen begegnen

Zeit: Dienstag, 28.01.2025: 15:15 - 17:00 · Ort: EO 150

Chair der Sitzung: Franziska Greiner-Döchert

Chair der Sitzung: Hendrik Lohse-Bossenz

Chair der Sitzung: Gerda Hagenauer

Herausforderungen im Umgang mit Belastungen durch eine Stärkung personaler Ressourcen von Lehrkräften und Schüler:innen begegnen

Chair(s): **Franziska Greiner-Döchert** (Universität Leipzig, Deutschland), **Hendrik Lohse-Bossenz** (Universität Greifswald)

Diskutant*in(nen): **Gerda Hagenauer** (Paris Lodron Universität Salzburg)

Bereits im Studium zeichnen sich Lehramtsstudierende durch eine hohe Prävalenz problematischer Belastungsprofile mit psychosomatischen Symptomen aus (Hohensee & Schiemann, 2022). Darüber hinaus werden seit Jahrzehnten Befunde repliziert, die bestätigen, dass Lehrkräfte zu einer stark belasteten Berufsgruppe gehören (Koestner et al., 2022; Iriarte Redín & Erro-Garcés, 2020). Doch auch Schüler:innen sind insbesondere seit der COVID19-Pandemie psychisch stark belastet und geben zunehmend psychosomatische Symptome an (Kaman et al., 2023). Auch für den Zusammenhang von Lehrkräfte- und Schüler:innenbelastung gibt es inzwischen Evidenz (Oberle & Schonert-Reichl, 2016; Bilz et al., 2022).

Ob Stressoren wie hoher Leistungsdruck jedoch zu Belastungsfolgen wie negativen Emotionen, Stress oder gar Burnout führen, hängt maßgeblich von der subjektiven Bewertung durch das Individuum ab (Kaluza, 2018). Stressoren, die potenziell zu Beanspruchungen führen, können durch Ressourcen abgepuffert werden. Neben beruflichen Ressourcen wie kollegialer Unterstützung, für deren Aktivierung i. d. R. eine Personengruppe erforderlich ist, können personale Ressourcen, zu denen Professionswissen, selbstregulative Fähigkeiten, Überzeugungen und motivationale Orientierungen gehören (u.a. Cramer et al., 2018), diesen Puffereffekt ermöglichen. Das Symposium fokussiert ebendiese personalen Ressourcen und adressiert die Frage, wie sich das Belastungserleben von (angehenden) Lehrkräften und Schüler:innen unter Berücksichtigung konkreter, personaler Ressourcen verändert und wie sich eine Stärkung von personalen Ressourcen auf mögliche Beanspruchungsfolgen auswirkt.

Beitrag 1 geht der Frage nach, was Lehrkräfte im erfolgreichen Umgang mit Belastungen unterscheidet. Dafür wurde in einer Tagebuchstudie das Zusammenwirken von Klassenführung und Wohlbefinden von berufstätigen Lehrkräften unter Berücksichtigung von Selbstwirksamkeit als potenzieller Puffer für wenig gelungene Klassenführung (Stressor) untersucht. Die Studie zeigt, dass die Selbstwirksamkeit der Lehrkräfte zwischen Klassenführung (Störungsprävention und Zeitmanagement) und Wohlbefinden moderiert: Der Zusammenhang von Klassenführung und Wohlbefinden ist dann geringer, wenn Lehrkräfte eine hohe unterrichtsbezogene Selbstwirksamkeit aufweisen.

Beitrag 2 beleuchtet mit einem Pre-Post-Kontrollgruppen-Design, wie sich eine Intervention zur Stärkung der Gesundheitskompetenzen auf das subjektive Wohlbefinden von angehenden Lehrkräften auswirkt. In der Intervention werden insbesondere Wissen und Stressregulationsstrategien als personale Ressourcen für den Umgang mit Belastungen fokussiert. Die Datenauswertung wird bis Herbst 2024 abgeschlossen sein, sodass entsprechende Ergebnisse im Symposium berichtet werden können.

Beitrag 3 thematisiert ein in Deutschland bislang wenig beforschtes Konstrukt: In zwei Validierungsstudien wurde Self-Compassion, also die Haltung der Fürsorge und Freundlichkeit gegenüber der eigenen Person, als personale Ressource von Kindern und Jugendlichen untersucht. Es zeigte sich, dass Self-Compassion signifikant positiv mit Glück ($r = .57$), Lebenszufriedenheit ($r = .56$) und Resilienz ($r = .57$) sowie signifikant negativ mit Depression ($r = -.59$), Angst ($r = -.55$) und Stress ($r = -.57$) korreliert.

Die Diskussion der drei Beiträge erfolgt aus einer prozesshaften Perspektive und wird u.a. durch die Frage geleitet, ob und wie sich die personalen Ressourcen gemeinsam betrachten und erforschen sowie in den Kontexten Schule und Hochschule stärken lassen, um negative Belastungsfolgen nachhaltig zu reduzieren.

Beiträge des Symposiums

Wohlbefinden nach Unterrichtsstörungen aufrechterhalten? Unterrichtsbezogene Selbstwirksamkeit und pädagogisches Wissen stärken

Andrea Westphal, Hendrik Lohse-Bossenz

Universität Greifswald

Das Wohlbefinden von Lehrkräften gilt als ein wesentlicher Indikator für den Erfolg im Lehrerberuf. Empirische Studien belegen, dass eine ungünstige Klassenführung, die mit einem hohen Maß an Unterrichtsstörungen einhergeht, sich negativ auf das Wohlbefinden von Lehrkräften auswirken kann (Aldrup et al., 2018; Dicke et al., 2018). Das transaktionale Stressmodell (Lazarus & Folkman, 1984) oder das Demands-Ressourcen-Modell (Demerouti et al., 2001) postulieren, dass wiederholte Stressoren, die die persönlichen Ressourcen übersteigen, Stress verursachen und das Wohlbefinden beeinträchtigen können. Unterrichtsbezogene Selbstwirksamkeit und pädagogisches Wissen können demnach als persönliche Ressourcen betrachtet werden, die den Einfluss täglicher Stressoren auf das Wohlbefinden abzumildern vermögen. Die vorliegende Tagebuchstudie untersucht das Zusammenspiel zwischen der Klassenführung und dem Wohlbefinden von Lehrkräften über zehn Unterrichtsstunden hinweg. Zudem wird geprüft, inwiefern die unterrichtsbezogene Selbstwirksamkeit und das pädagogische Wissen dazu beitragen, dass Lehrkräfte ihr Wohlbefinden auch nach Unterrichtsstunden mit wenig gelungener Klassenführung aufrechterhalten können.

Die Fragestellungen wurden basierend auf einer Stichprobe von 81 Grundschullehrkräften (Alter = 41.63, SD = 10.70; Berufserfahrung = 14.31, SD = 10.39) untersucht. Die Lehrkräfte schätzten an zehn aufeinanderfolgenden Schultagen Aspekte der Klassenführung (Prävention von Unterrichtsstörungen, Monitoring und effektive Zeitnutzung, je ein Item, 5-stufiges Antwortformat; Thiel et al., 2013) und ihr Wohlbefinden (ein Item: „Hat sich ihr Allgemeinbefinden im Vergleich vor und nach der Stunde...“, 5-stufiges Antwortformat, „sehr verschlechtert...sehr verbessert“; Eigenentwicklung) nach einer selbstgewählten Unterrichtsstunde ein. Die Intraklassenkorrelationen lagen zwischen 0.22 und 0.27. Im Vorfeld der zehntägigen Befragung wurden unterrichtsbezogene Selbstwirksamkeit (acht Items, fünfstufiges Antwortformat, Cronbachs $\alpha = .85$; Tschannen-Moran & Woolfolk Hoy, 2001) und das pädagogische Wissen der Lehrkräfte (sechs Items, fünfstufiges Antwortformat, Cronbachs $\alpha = .77$; Piwowar et al., 2013) erfasst. Um die hierarchische Datenstruktur mit Angaben auf Tages- (Level 1) und Lehrkräfebene (Level 2) zu berücksichtigen, wurden Mehrebenenanalysen berechnet.

In einem ersten Modell wurde das Wohlbefinden vorhergesagt durch Aspekte der Klassenführung (Level 1) unter Kontrolle von Berufserfahrung und Geschlecht der Lehrkräfte (Level 2). Die Analyse zeigte, dass sich effektives Zeitmanagement positiv auf das Wohlbefinden der Lehrkräfte auswirkte, während eine misslungene Störungsprävention sich negativ auf das Wohlbefinden auswirkte. Das Monitoring der Lehrkräfte zeigte darüber hinaus keinen Zusammenhang mit dem Wohlbefinden. In einem weiteren Modell wurden zusätzlich unterrichtsbezogene Selbstwirksamkeit und pädagogisches Wissen (Level 2) berücksichtigt, für die sich kein Zusammenhang mit dem Wohlbefinden zeigte. In weiteren Modellen fanden wir signifikante Interaktionseffekte zwischen unterrichtsbezogener Selbstwirksamkeit und Aspekten der Klassenführung: Bei Lehrkräften mit höherer unterrichtsbezogener Selbstwirksamkeit hing das Wohlbefinden weniger eng mit den Aspekten der Klassenführung (Störungsprävention, Zeitmanagement) zusammen. Ein signifikanter Interaktionseffekt zeigte sich auch für das pädagogische Wissen: Bei Lehrkräften mit höherem pädagogischen Wissen hing das Wohlbefinden weniger eng mit dem Zeitmanagement zusammen.

Die Ergebnisse unterstreichen, wie relevant die Förderung der unterrichtsbezogenen Selbstwirksamkeit für ein stabiles Wohlbefinden von Lehrkräften ist. Wir diskutieren Implikationen für die Lehrkräftebildung.

Förderung der mentalen Gesundheit von Lehramtsstudierenden der beruflichen Bildung

Kira Weber

Universität Hamburg

Theoretischer Hintergrund und Fragestellungen

Studien zeigen, dass Studierende zum einen mehr ungesunde Verhaltensweisen aufweisen und zum anderen ihren subjektiven Gesundheitszustand schlechter bewerten als gleichaltrige Nicht-Studierende (Grützmacher et al., 2018). In der Studie von Lutz-Kopp et al. (2019) zeigte sich zudem, dass ein Viertel der befragten Studierenden stark erschöpft sind und 58 % eine geringe Resilienz berichten. Besonders Lehramtsstudierende haben oft ein problematisches Belastungsprofil. Sie berichten häufig von Symptomen wie Konzentrationsschwierigkeiten, Rückenschmerzen, Schlafstörungen und Ängsten, die das Wohlbefinden und die Studienleistungen negativ beeinflussen können (Hohensee & Schiemann, 2022). Dies ist besonders bedenklich vor dem Hintergrund, dass die Gesundheit von Lehrkräften einen Einfluss auf die Qualität des Unterrichts hat (Wartenberg et al., 2023). Um die Gesundheit von Lehramtsstudierenden und damit langfristig auch die der Lehrkräfte zu stärken, ist es notwendig, gesundheits-theoretische Inhalte bereits in das Studium zu integrieren (Bonse-Rohmann et al., 2023). Daher wurde das Seminar „Gesund bleiben im Studium durch Resilienz und Stressbewältigung“ für Lehramtsstudierende der beruflichen Bildung angeboten und im Hinblick auf folgende Fragestellungen untersucht:

1. Wie wirkt sich die Intervention auf Gesundheitskompetenzen und das subjektive Wohlbefinden der Studierenden aus?
2. Welche Hindernisse berichten die Studierenden bei der praktischen Umsetzung der erlernten Strategien in den Alltag?
3. Was benötigen die Studierenden, um die erlernten Gesundheitskompetenzen in ihrem Alltag sowohl kurz- als auch langfristig umsetzen zu können?

Methode

Stichprobe

Bei der Stichprobe handelt es sich um 27 Lehramtsstudierende der beruflichen Bildung, die sich im zweiten oder vierten Mastersemester befinden. Das Seminar wird im Rahmen eines Wahlpflichtmoduls angeboten. 95 % der Studierenden berichteten, dass sie das Seminar aus Interesse am Thema besuchen.

Beschreibung des Seminars

Das Seminar „Gesund bleiben im Studium durch Resilienz und Stressbewältigung“ umfasst acht Sitzungen, von denen sieben in Präsenz und eine im Selbststudium stattfinden. Zwischen den Sitzungen sind die Teilnehmenden aufgefordert, Reflexionsaufgaben zu bearbeiten und sich mithilfe eines Workbooks mit den Themen zu beschäftigen. Als Grundlage für das Seminar diente zum einen eine Online-Fortbildungsreihe für Lehrkräfte (Autorin, i. Dr. a & b) und zum anderen das ReMind-Projekt (<https://www.resilienz-im-studium.de/>), welches Material und Folien zur Stressprävention und Resilienzförderung von Studierenden frei zugänglich zur Verfügung stellt. Inhalte des Seminars sind unter anderem die Entstehung von Stress, Stressregulationsstrategien, Resilienz, Selbstfürsorge, Arbeitsmanagement, Achtsamkeit und Stärkenorientierung. Im Selbststudium erstellen die Teilnehmenden ein individuelles Vision Board, welches in der letzten Sitzung in einem Gallery Walk ausgestellt wird.

Forschungsdesign und Instrumente

Die Forschungsfragen wurden mithilfe eines mixed-methods-Designs untersucht. Zur Beantwortung der ersten Forschungsfrage wurde ein Prä-Post-Kontrollgruppen Design gewählt. Folgende Instrumente wurden eingesetzt: WHO-5 Well-being Index (Krieger et al., 2014), Fragebogen zur Gesundheitskompetenz (Lennartz, 2011), Copenhagen Burnout Inventory (Kristensen et al., 2005), SEK-27 (Berkling & Znoj, 2011), Selbstregulation im Bereich der Selbstfürsorge (Li Sanchez & Schwinger, 2023). Zudem wurden leitfadengestützte Interviews mit n = 5 Studierenden aus dem Seminar geführt, um die zweite und dritte Forschungsfrage zu beantworten.

Ergebnisse

Sowohl die Ergebnisse der standardisierten Befragung als auch die der Interviews liegen erst zum Ende des Semesters vor und können daher noch nicht beschrieben werden. Im WS 24/25 wird das Seminar zudem noch einmal angeboten und es werden erneut Daten erhoben. Die Ergebnisse der Evaluation des Seminars zeigen, dass die Studierenden das Seminar als sehr hilfreich empfunden haben ($M = 4.9$; $SD = 0.2$) und durch die Lehrveranstaltung ihr Wissen vertiefen konnten ($M = 4.6$; $SD = 0.7$).

Diskussion

Die Ergebnisse der Evaluation und die informellen Rückmeldungen der Studierenden zeigen, dass das Thema als sehr wichtig erachtet wird. Viele der Studierenden hätten sich gewünscht, dass die Lehrveranstaltung bereits früher im Studium angeboten wird und dass die Teilnahme für alle Lehramtsstudierenden verpflichtend ist. Auf der Konferenz werden weitere Implikationen diskutiert.

Self-Compassion als Resilienzfaktor – Validierung der deutschsprachigen Self-Compassion Scale for Youth

Friederike Carlotta Grabowski

Europa Universität Flensburg

Self-Compassion beschreibt eine Haltung der Fürsorge und Freundlichkeit gegenüber der eigenen Person (Castello, 2020). Es bedeutet, das eigene Leiden und die eigenen Fehler bewusst wahrzunehmen, ihnen verständnisvoll zu begegnen und sie als Teil des menschlichen Lebens zu betrachten (Gruber et al., 2020; Neff, 2003a). Nach Neff (2003b) können drei Kerndimensionen von Self-Compassion mit jeweils einem Antagonisten unterschieden werden: (1) Selbstfreundlichkeit vs. Selbstverurteilung, (2) verbindende Menschlichkeit vs. Isolation und (3) Achtsamkeit vs. Überidentifikation. Im Kindes- und Jugendalter kann Self-Compassion als Resilienzfaktor gegenüber psychischen Belastungen dienen. Internationale Studien geben Hinweise auf inverse Beziehungen zwischen Self-Compassion und Angst, Depression sowie Stress (z. B. Marsh et al., 2018) sowie positive Assoziationen zum Wohlbefinden (z. B. Bluth & Blanton, 2015) in dieser Altersgruppe. Angesichts gesamtgesellschaftlicher Krisenlagen und der damit einhergehenden Zunahme psychischer Belastungen (Reiß et al., 2023) sind pädagogische Präventionskonzepte erforderlich, um die Resilienz von Kindern und Jugendlichen zu stärken. Self-Compassion ist dabei eine personale Ressource, die durch Interventionen gestärkt werden kann und mit positiven psychosozialen Entwicklungen assoziiert ist (Ferrari et al., 2019). Zur Operationalisierung von Self-Compassion liegt ein englischsprachiges Instrument für das Jugendalter vor (Self-Compassion Scale for Youth (SCS-Y); Neff et al., 2021). Dieses 17 Items umfassende Instrument bildet die drei Kerndimensionen von Self-Compassion sowie ihre jeweiligen Antagonisten ab, zudem kann ein Gesamtscore gebildet werden.

Es fehlt jedoch ein deutschsprachiges Instrument, mit dem Self-Compassion im Jugendalter erhoben werden kann. Dieses ist jedoch nötig, um Self-Compassion psychometrisch valide und reliabel zu untersuchen sowie pädagogische Interventionen mit dem Ziel der Self-Compassion-Stärkung zuverlässig bewerten und verbessern zu können. Dadurch könnte Kindern und Jugendlichen geholfen werden, adaptivere Wege zur Bewältigung psychischer Belastungen und einer Verbesserung der psychischen Gesundheit zu ermöglichen.

Vor diesem Hintergrund war das Ziel der vorzustellenden Studie, die deutschsprachige SCS-Y zu validieren. Hierzu wurden zwei Studien durchgeführt. In Studie 1 wurde die Reliabilität und Konstruktvalidität mit $N = 165$ 10-14 Jährigen ($M_{\text{Alter}} = 12.01$, $SD_{\text{Alter}} = 1.27$) überprüft. Es zeigte sich eine sehr gute Reliabilität ($\alpha = .85$), zudem konnte die angenommene sechsfaktorielle Struktur bestätigt werden, wobei sich in der CFA gute Model-Fit Werte zeigten. Studie 2 untersuchte die konvergente und divergente Validität mit einer Stichprobe von 236 Jugendlichen im Alter von 10 bis 16 Jahren ($M_{\text{Alter}} = 12.63$, $SD_{\text{Alter}} = 1.36$). Es wurden Zusammenhänge zwischen Selbstmitgefühl und den Konstrukten Glück, Lebenszufriedenheit und Resilienz sowie Depression, Angst und Stress analysiert. Die Ergebnisse zeigten eine angemessene Konstruktrepräsentativität: Selbstmitgefühl korrelierte signifikant positiv mit Glück ($r = .57$), Lebenszufriedenheit ($r = .56$) und Resilienz ($r = .57$). Darüber hinaus war Selbstmitgefühl signifikant negativ mit Depression ($r = -.59$), Angst ($r = -.55$) und Stress ($r = -.57$) assoziiert.

Die Ergebnisse zeigen, dass die deutschsprachige SCS-Y ein psychometrisch valides und reliables Instrument ist, um Self-Compassion als vielversprechenden Faktor zur Stärkung der Resilienz von Kindern und Jugendlichen zu erheben. Implikationen der Ergebnisse, insbesondere hinsichtlich des Einsatzes der Skala in der Evaluation von Self-Compassion Interventionen im Kindes- und Jugendalter, sollen diskutiert werden.

IS-07: Invited Symposium: Zentrale Ergebnisse aus TIMSS 2023 im Überblick und ihre Bedeutung für die Bildungsforschung, -politik und -praxis

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 102

Chair der Sitzung: Christin Beese

Chair der Sitzung: Luise Scholz

Chair der Sitzung: Christine Sälzer

Zentrale Ergebnisse aus TIMSS 2023 im Überblick und ihre Bedeutung für die Bildungsforschung, -politik und -praxis

Chair(s): **Christin Beese** (Universität Hamburg, Deutschland), **Luise Scholz** (Universität Hamburg, Deutschland)

Diskutant*in(nen): **Christine Sälzer** (Universität Stuttgart)

Als Teil der Gesamtstrategie der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) zum Bildungsmonitoring werden mit der Trends in International Mathematics and Science Study (TIMSS) seit 2007 die mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen aus Deutschland im internationalen Vergleich untersucht (KMK, 2015). Seit Kurzem liegen nun die Ergebnisse von TIMSS 2023 und somit des fünften Zyklus der Studie vor. Dieser zeichnet sich im Vergleich zu vorherigen Zyklen durch eine nahezu vollständig digitale Erhebung sowie eine gruppenadaptive Kompetenztestung aus (Schwippert et al., 2024). Darüber hinaus bietet sich die Möglichkeit, die Kompetenzen von Grundschüler:innen unmittelbar vor (TIMSS 2019) und nach der COVID-19-Pandemie (TIMSS 2023) unter Berücksichtigung verschiedener Hintergrundinformationen gegenüberzustellen und im Trend zu betrachten (Schwippert et al., 2020). Insgesamt wurden in TIMSS 2023 ca. 4 500 Schüler:innen in 230 Schulen getestet und befragt.

Da in TIMSS nicht nur Kompetenzen von Schüler:innen, sondern auch zentrale Merkmale ihrer Lehr- und Lernumwelten umfangreich über Hintergrundfragebögen für Schüler:innen, Erziehungsberechtigte, Lehrkräfte und Schulleitungen erhoben werden, ist es möglich ein umfassendes Bild des Lernens von Schüler:innen am Ende der Grundschulzeit zu zeichnen. In insgesamt vier Einzelbeiträgen werden die zentralen Befunde aus TIMSS 2023 vorgestellt und anschließend in Hinblick auf ihre praktische Relevanz für Bildungspolitik, Bildungsplanung und auch die Praxis diskutiert.

Nach einer kurzen Einführung zur TIMS-Studie widmet sich der erste Beitrag den mathematischen und naturwissenschaftlichen Kompetenzen im internationalen Vergleich sowie im Trend seit 2007. Anschließend werden im zweiten Beitrag ausgewählte Aspekte des Unterrichtens in Grundschulen betrachtet: das Lehren und Lernen mit digitalen Medien als Schwerpunktthema in TIMSS 2023, Merkmale der Unterrichtsqualität im Mathematik- und Sachunterricht sowie die Teilnahme der unterrichtenden Lehrkräfte an Fortbildungsveranstaltungen. Der dritte Beitrag nimmt die mathematischen und naturwissenschaftlichen Kompetenzen von Schüler:innen unter Berücksichtigung individueller Merkmale im internationalen Vergleich sowie im Trend seit 2007 in den Blick. Hierbei werden das Geschlecht, die Herkunftseffekte und der Migrationshintergrund der Schüler:innen als Faktoren zur Erklärung von mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen betrachtet. Der letzte Beitrag befasst sich mit außerschulischer Nachhilfe sowie der Nutzung von Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände. Abschließend werden die zentralen Aspekte der Einzelbeiträge in Hinblick auf ihre Bedeutung für die Bildungsforschung, -politik und -praxis durch Christine Sälzer sowie das Plenum diskutiert.

Da die Ergebnisse von TIMSS 2023 zur Einreichungsfrist noch nicht berichtet werden durften, werden in den Einzelabstracts keine Ergebnisse vorgestellt.

Beiträge des Symposiums

Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich

Aiso Heinze¹, Mirjam Steffensky², Daniel Walter³, Olaf Köller¹, Christoph Selter³, Johanna Brandt³

¹IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, ²Universität Hamburg, ³Technische Universität Dortmund

Theoretischer Hintergrund

Die Entwicklung grundlegender mathematischer und naturwissenschaftlicher Kompetenzen ist eine zentrale Aufgabe der Grundschule. Hinreichende Kompetenzen in beiden Bereichen stellen notwendige Voraussetzungen für das anschlussfähige Lernen in der Sekundarstufe I dar. Dennoch zeigt der IQB-Bildungstrend 2021 deutlich sinkende Leistungen in Deutsch und Mathematik am Ende der 4. Klasse. Die im Bildungstrend 2021 getesteten Kinder waren, ebenso wie die in TIMSS 2023 getesteten Kinder, von den Schulschließungen in der Corona-Pandemie betroffen. Zu erwarten ist demnach, dass sich auch in den mathematischen und naturwissenschaftlichen Leistungen von Viertklässler:innen in TIMSS negative Effekte der Pandemie zeigen werden und die Leistungen gegenüber TIMSS 2019 zurückgegangen sind. Dabei ist eine offene Frage, ob die Rückgänge in den Naturwissenschaften (analog zu PISA 2022) geringer ausfallen als in Mathematik.

Fragestellung

Im Zentrum dieses Beitrags steht die Frage, welche Leistungen Schüler:innen in Deutschland in Mathematik und Naturwissenschaften in TIMSS 2023 zeigen und inwieweit sich im Vergleich zur ersten Erhebungsrunde (TIMSS 2007) und der vorangegangenen Erhebungsrunde (TIMSS 2019) Veränderungen beobachten lassen. Konkret wird folgenden Fragen nachgegangen:

- Welche Leistungen in Mathematik und Naturwissenschaften zeigen Schüler:innen in Deutschland am Ende der vierten Jahrgangsstufe im internationalen Vergleich?
- Wie lassen sich die Leistungskennwerte auf die Kompetenzstufen in Deutschland im internationalen Vergleich einordnen? Wie groß sind die Gruppen der auffällig leistungsstarken und leistungsschwachen Kinder?
- Welche Ergebnisse erzielen Schüler:innen differenziert nach den mathematischen Teilgebieten Arithmetik, Geometrie/Messen und Umgang mit Daten sowie den naturwissenschaftlichen Inhaltsbereichen Biologie, Physik/Chemie und Geografie im internationalen Vergleich?
- Welche Ergebnisse erzielen Schüler:innen differenziert nach den kognitiven Anforderungsbereichen Wissen, Anwenden und Problemlösen im internationalen Vergleich?

- Über welche Einstellungen zur Mathematik bzw. zu den Naturwissenschaften und über welches mathematische bzw. naturwissenschaftliche Selbstkonzept verfügen die Schüler:innen am Ende der vierten Jahrgangsstufe?

Methode

Datengrundlage ist die für Viertklässler:innen in Deutschland repräsentative Schülerstichprobe (Durchschnittsalter 10,4 Jahre) der TIMSS-Zyklen aus den Jahren 2007 (N=5 200), 2019 (N=4 942) und 2023 (N=4 442). Die Kompetenzen der Schüler:innen wurden mit standardisierten Kompetenztests erfasst, die die drei genannten fachspezifischen Inhaltsbereiche abdecken. Der Mathematiktest im Jahr 2023 umfasst 187 Aufgaben, der Naturwissenschaftstest 176 Aufgaben. Etwa die Hälfte aller Testaufgaben sind im Multiple-Choice-Format angelegt, während die übrigen Testaufgaben eine kurze schriftliche Antwort oder die Zuordnung verschiedener Begriffe erfordern. Die Verteilung der Testaufgaben auf die verschiedenen mathematischen bzw. naturwissenschaftlichen Inhaltsgebiete und kognitiven Anforderungsbereiche ist bei den Erhebungsrunden nahezu identisch. Dadurch können Vergleiche der Testergebnisse gezogen werden. Da jedem Lernenden nur ein Teil der Aufgaben vorgelegt wurde (Multi-Matrix-Design), wurden mit dem Verfahren der Item-Response-Theory für jedes Kind fünf Plausible Values ermittelt. Diese wurden in der aktuellen Studie verwendet. Um Spezifika des Studiendesigns zu berücksichtigen, wurden alle Analysen mit dem IDB Analyzer der International Association for the Evaluation of Educational Achievement (IEA; Fishbein et al., 2021) berechnet.

Unterrichten in der Grundschule: Digitale Medien, Lehrkräftefortbildung und Unterrichtsqualität – Ergebnisse aus TIMSS 2023

Donieta Jusufi, Christin Beese, Luise Scholz
Universität Hamburg

Unterricht zählt zu den einflussreichsten Faktoren für die schulischen Leistungen von Schüler:innen (Hattie, 2009; Kunter, 2016; Seidel & Shavelson, 2007). Insbesondere die Qualität von Unterricht ist seit langem Gegenstand der nationalen und internationalen Forschung. In der letzten Zeit rückt zudem aufgrund aktueller Entwicklungen die Nutzung digitaler Medien in Schule und Unterricht zunehmend in den Fokus der Forschung (Eickelmann & Gerick, 2017; Irion & Peschel, 2016). Die Gestaltung von Unterricht in Deutschland ebenso wie die Leistungen von Schüler:innen hängen u. a. davon ab, inwiefern Lehrkräfte bereit sind, ihren Unterricht weiterzuentwickeln und sich über neue Konzepte, Befunde oder Ansätze beispielsweise im Rahmen von Fortbildungen zu informieren (Hattie, 2009; Lipowsky & Rzejak, 2023; Richter, 2023). Dazu zählen neben Fortbildungen zum Einsatz digitaler Medien auch Fortbildungsveranstaltungen zu inhaltlichen, fachdidaktischen und diagnostischen Aspekten des Unterrichtens.

Mit den Daten der Trends in International Mathematics and Science Study (TIMSS) 2023 werden diese Bedingungen des Lehren und Lernens am Ende der Grundschulzeit und ihr Zusammenhang mit den mathematischen und naturwissenschaftlichen Kompetenzen der Schüler:innen näher in den Blick genommen. Auf Basis aktueller Befunde (Eickelmann & Gerick, 2020; Lorenz et al., 2023; OECD, 2023) wird in TIMSS schwerpunktmäßig die Verfügbarkeit und Nutzung digitaler Medien sowie die Entwicklungen zwischen 2019 und 2023 als Abbild der Auswirkungen der COVID-19-Pandemie betrachtet. Dabei werden in TIMSS 2023 unter digitalen Medien sowohl digitale Endgeräte als auch Software, wie z. B. Lernprogramme, zusammengefasst. Zur Darstellung der Verfügbarkeit und Nutzung digitaler Medien in der Schule und zu Hause werden Aussagen der Schüler:innen, der Lehrkräfte und der Schulleitungen genutzt. Für ein umfassenderes Bild wird darüber hinaus das subjektiv wahrgenommene Kompetenzerleben der Schüler:innen im Umgang mit digitalen Medien erfragt.

Als etablierter Prädiktor schulischer Leistungen gilt die Unterrichtsqualität. Verschiedenste Untersuchungen belegen den Zusammenhang zwischen den drei Basisdimensionen der Unterrichtsqualität Klassenführung, kognitive Aktivierung und konstruktive Unterstützung nach Klieme et al. (2001) sowie kognitiven und motivationalen Merkmalen der Schüler:innen (Fauth et al. 2019, Schiepe-Tiska, 2016). Das international etablierteste Konstrukt Klassenführung (classroom management) wird in TIMSS 2023 auf internationaler Ebene erhoben, auf nationaler Ebene lässt sich durch entsprechende Ergänzungen in den Fragebögen das vollständige Modell der drei Basisdimensionen von Unterrichtsqualität abbilden und im Zusammenhang mit den Kompetenzen betrachten.

Zur Sicherstellung der Unterrichtsqualität und einer qualitativ hochwertigen Einbindung digitaler Medien in den Unterricht hat sich gezeigt, dass die Teilnahme an Fortbildungen einen relevanten Faktor darstellt (Herzig, 2004; Lipowsky & Rzejak, 2023; Lorenz, 2018). Daher wird in diesem Zusammenhang die Teilnahme der Lehrkräfte an fachspezifischen Fortbildungen, insbesondere auch zur Integration digitaler Medien in den Unterricht, betrachtet. Als Reaktion auf Befunde aus TIMSS 2019 (Guill & Wendt, 2020) wurden die bereits bestehenden fachspezifischen Skalen zur Fortbildungsteilnahme um die Teilaspekte Angebote und Bedarfe ergänzt, um ein umfassenderes Bild zur Lehrkräftefortbildungen zeichnen zu können.

Im Rahmen des Einzelbeitrages werden die Ergebnisse von TIMSS 2023 zu den drei inhaltlichen Bereichen digitale Medien, Unterrichtsqualität und Lehrkräftefortbildung präsentiert und diskutiert. Dazu werden sowohl internationale Vergleiche mit anderen Teilnehmerstaaten, Entwicklungen im Trend, aber auch vertiefende nationale Analysen vorgestellt.

Mathematische und naturwissenschaftliche Kompetenzen von Schüler:innen unter Berücksichtigung individueller Merkmale im Trend

Eva Grommé¹, Annika Beier², Leonhard Schulz², Sonja Nonte¹, Tobias C. Stubbe²

¹Universität Osnabrück, ²Georg-August-Universität Göttingen

Disparitäten in schulischen Leistungen aufgrund individueller und struktureller Merkmale sind ein zentrales Thema der Bildungsforschung. Zahlreiche Large-Scale-Assessments (LSAs) zeigen auf nationaler (Stanat et al., 2023) wie auch auf internationaler Ebene (OECD, 2023), dass der sozioökonomische Status und der Migrationshintergrund der Familie sowie das Geschlecht der Kinder in einem erheblichen Zusammenhang mit den Kompetenzen von Schüler:innen stehen (Keller et al., 2023). Diese Ungleichheiten wirken sich langfristig auf die Bildungs- und Berufschancen der Betroffenen aus (Costache et al., 2022). In diesem Vortrag werden das Geschlecht, die Herkunftseffekte und der Migrationshintergrund als Faktoren zur Erklärung von mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen betrachtet. Dazu werden aktuelle Ergebnisse der Trends in International Mathematics and Science Study 2023 (TIMSS 2023) vorgestellt, und es wird die Notwendigkeit komplexer methodischer Herangehensweisen zur Beantwortung relevanter Forschungsfragen diskutiert.

Bezüglich des Geschlechts sind es im Bereich der Mathematik häufiger Mädchen, die geringere Kompetenzen aufweisen (Keller et al., 2022). Befunde des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) aus dem Jahr 2021 zeigen einen zwar geringen, aber signifikanten Effekt des Geschlechts von Viertklässler:innen zugunsten der Jungen (Gentrup et al., 2022). Im Bereich der Naturwissenschaften sind die Unterschiede weniger eindeutig (Nonte et al., 2020). Als ein theoretischer

Erklärungsansatz für Geschlechterunterschiede kann beispielsweise das Erwartungs-Wert-Modell (Eccles & Wigfield, 2002) dienen, welches besagt, dass das Fähigkeitsselbstkonzept sowie die intrinsische Motivation beeinflussen, wie intensiv sich Kinder mit Mathematik oder Naturwissenschaften auseinandersetzen.

Hinsichtlich des sozialen Status von Familien ist die enge Koppelung mit dem Bildungserfolg von Kindern umfangreich dokumentiert. Für die mathematischen und naturwissenschaftlichen Kompetenzen in der Grundschule zeigt TIMSS seit 2007 für alle teilnehmenden Staaten einen zum Teil erheblichen Vorsprung der Schüler:innen aus bildungsnahen Elternhäusern (Stubbe, Krieg et al., 2020). Neben diesen primären Herkunftseffekten müssen bei der Erklärung von sozialen Disparitäten auch die sekundären Herkunftseffekte berücksichtigt werden. Nach Boudon (1974) versteht man darunter Bildungsentscheidungen, die auch unter Kontrolle der primären Herkunftseffekte, einen Zusammenhang mit dem sozialen Status von Familien aufweisen. Für TIMSS bedeutet dies, dass Kinder aus bildungsnahen Familien auch unter Kontrolle der mathematischen und naturwissenschaftlichen Kompetenzen eine höhere Chance haben, nach der Grundschule auf ein Gymnasium zu wechseln (Stubbe, Kasper et al., 2020).

Das dritte individuelle Merkmal, das in der empirischen Bildungsforschung üblicherweise berücksichtigt wird, ist der Migrationshintergrund. So zeigen Kinder, bei denen ein oder beide Elternteile nicht in Deutschland geboren wurden, beispielsweise geringere mathematische beziehungsweise naturwissenschaftliche Kompetenzen als Kinder ohne Migrationshintergrund (Wendt et al., 2020). Damit zusammenhängende Erklärungen für das schlechtere Abschneiden können sprachliche Kompetenzen oder sozioökonomische Benachteiligungen sein (Stanat et al., 2010).

Bedeutsam ist, dass diese individuellen Merkmale nicht isoliert auftreten, sondern insbesondere im Zusammenspiel den Bildungserfolg erklären können (Codioli McMaster & Cook, 2019). Eine aktuelle Sekundäranalyse von PISA-2018-Daten konnte durch eine intersektionale Betrachtung komplexe Effekte auf die Lesekompetenzen von 15-Jährigen aufzeigen. So weisen bestimmte Schülergruppen (z. B. Jungen mit Migrationshintergrund aus Familien mit niedrigem sozioökonomischem Status) besonders niedrige Lesekompetenzen auf (Keller et al., 2023). Befunde in Bezug auf mathematische Kompetenzen indizieren (Cascella, 2020), dass Disparitäten auch in anderen Kompetenzbereichen nicht isoliert betrachtet werden sollten.

In diesem Vortrag werden zunächst aktuelle Befunde aus TIMSS 2023 für die einzelnen Ungleichheitsdimensionen präsentiert und in den Trend seit 2007 eingeordnet. Im Fokus stehen dabei das Geschlecht, die primären sowie die sekundären Herkunftseffekte. Anschließend werden die Entwicklungen im Trend anhand eines linearen Modells mit Zufallseffekten in Form eines Mehrebenenmodells berichtet (McCulloch et al., 2008), wobei als erklärende Faktoren der Kompetenzen der Messzeitpunkt, die soziodemografischen Merkmale und deren Interaktion berücksichtigt werden.

Privater Nachhilfeunterricht und Förderungen zum Ausgleich pandemiebedingter Lernrückstände am Ende der Grundschulzeit

Lisa Benckwitz¹, Karin Guill¹, Christin Beese²

¹IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, ²Universität Hamburg

Theoretischer Hintergrund und Fragestellungen

Zusätzlich zum regulären Schulunterricht nehmen viele Schüler:innen in Deutschland privaten Nachhilfeunterricht in Anspruch (Hille et al., 2016). Dadurch erhalten sie regelmäßigen, auf sie abgestimmten Unterricht außerhalb der Familie und der Schule. Nachhilfeunterricht wird meist von Lehrkräften, Studierenden oder älteren Schüler:innen gegeben und verfolgt das Ziel, den schulischen Lernerfolg abzusichern (Krüger, 1977; Park et al., 2016).

Den Ergebnissen aus TIMSS 2015 und TIMSS 2019 zufolge nutzen allerdings nur wenige Schüler:innen in Deutschland am Ende der Grundschulzeit privaten Nachhilfeunterricht (Guill & Wendt, 2016, 2020). Bisherige Untersuchungen zu den Folgen der Schulschließungen im Rahmen der COVID-19-Pandemie zeigen jedoch, dass vor allem die Leistung von Grundschüler:innen negativ durch die Schulschließungen beeinflusst wurde (Hammerstein et al., 2021; Helm et al., 2021; König & Frey, 2022). Diese pandemiebedingten Lernrückstände könnten zu einer verstärkten Nutzung von privatem Nachhilfeunterricht geführt haben. Folglich wird in diesem Beitrag untersucht, ob der Anteil an Grundschüler:innen, die privaten Nachhilfeunterricht in Anspruch nehmen, nach der Pandemie angestiegen ist (Fragestellung 1). Zusätzlich werden pandemiebezogene Prädiktoren der Nutzung von Nachhilfe analysiert (z. B. die Länge der Schulschließungen), um zu prüfen, ob Hinweise auf einen Zusammenhang zwischen pandemiebezogener Schulschließung und Nutzung von privatem Nachhilfeunterricht bestehen (Fragestellung 2).

Zum Ausgleich negativer Folgen der Pandemie für Schüler:innen hat die Bundesregierung darüber hinaus z. B. das Aktionsprogramm Aufholen nach Corona für die Jahre 2021 und 2022 beschlossen (BMFSFJ, 2022). Im Rahmen des Programms wurden zwei Milliarden Euro für Projekte zur Förderung der sozial-emotionalen und kognitiven Entwicklung von Schüler:innen bereitgestellt, die gezielt auch an Bildungsübergängen eingesetzt werden sollen (SWK, 2021). Dieser Beitrag untersucht daher zusätzlich den Anteil der Schüler:innen am Ende der Grundschulzeit, die an verschiedenen schulischen und außerschulischen Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände teilgenommen haben (z. B. unterrichtsbegleitende Fördermaßnahmen; Fragestellung 3).

Des Weiteren wird von der SWK (2021) empfohlen, dass sich Förderprogramme vor allem auf besonders stark von der Pandemie betroffene Schüler:innen fokussieren sollten, z. B. auf Schüler:innen, die lediglich die unteren Kompetenzstufen erreichen oder sich in benachteiligten schulischen Lernkontexten befinden. Es gibt bisher jedoch wenige empirische Befunde, inwiefern jene Schüler:innen tatsächlich von den Fördermaßnahmen erreicht wurden, da bei der Umsetzung von Aufholen nach Corona auf eine standardisierte, bundesweite Lernstandserhebung verzichtet werden musste (SWK, 2024). Dementsprechend wird in diesem Beitrag weiterhin untersucht, inwiefern Schüler:innen mit besonders hohem Unterstützungsbedarf tatsächlich von den Aufholprogrammen erreicht wurden (Fragestellung 4).

Methode

Die vier Fragestellungen werden auf Grundlage der Daten aus TIMSS 2023 deskriptiv ausgewertet. Um Fragestellungen 1 und 3 zu beantworten, werden die Häufigkeiten der Nutzung von Nachhilfe und Aufholprogrammen angegeben. Während die Nutzung von Aufholprogrammen in TIMSS 2023 erstmals erfragt wurde, wird die Häufigkeit der Nutzung von Nachhilfe auf Grundlage der Daten aus IGLU 2006 sowie TIMSS 2011 bis TIMSS 2023 im Trend abgebildet. In Bezug auf Fragestellungen 2 und 4 werden die Häufigkeiten der Nutzung differenziert nach den jeweiligen Prädiktoren (z. B. pandemiebezogene Prädiktoren) angegeben.

PS-52: Soziale Mobilität und Bildungserfolg

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 129

Chair der Sitzung: Thomas Zimmermann

Höherqualifizierung nach einem ersten Bildungsabschluss – soziale Selektivitäten und Auswirkungen auf langfristige Einkommensunterschiede

Steffen Schindler¹, Felix Bittmann²

¹Universität Bamberg, Deutschland; ²Leibniz Institut für Bildungsverläufe

In der soziologischen Bildungsforschung existieren zahlreiche Studien, die sich mit den Einflüssen der sozialen Herkunft auf das Nachholen von Bildungsabschlüssen beschäftigen (z.B. Buchholz/Pratter 2017; Hillmert/Jacob 2010; Kurz/Böhner-Taute 2016). Dabei zeigt sich, dass unter den Personen, die einen ersten Bildungsabschluss erworben haben, die Anteile des Erwerbs weiterer Bildungsabschlüsse bei Personen privilegierter sozialer Herkunft höher sind als bei Personen benachteiligter sozialer Herkunft (Schindler 2015; Buchholz/Schier 2015). Weit weniger Forschung existiert zu der Frage, wie durch diese Prozesse spätere soziale Ungleichheiten im Arbeitsmarkt beeinflusst werden. Einige wenige Arbeiten deuten darauf hin, dass bei formal gleichen Bildungsabschlüssen Personen mit nichtgeradlinigen oder nachgeholten Abschlüssen Nachteile auf dem Arbeitsmarkt erfahren (Schuchart/Schimke 2019; Schmelzer 2012). Unklar ist jedoch, inwieweit hierbei Selektionseffekte und unbeobachtete Heterogenität der Vergleichsgruppen eine Rolle spielen. Zur Frage, in welcher Weise vor dem Arbeitsmarkteintritt nachgeholte Bildungsabschlüsse den Zusammenhang zwischen der sozialen Herkunft und Arbeitsmarkterträgen beeinflussen, liegen bisher keine Befunde vor. Hier setzt das vorliegende Papier an. Es stellt erstens die Frage, inwieweit durch sozial selektives Verhalten beim Erwerb weiterer Bildungsabschlüsse soziale Ungleichheiten im späteren Einkommen erzeugt werden. Zweitens sollen die Muster der unterschiedlichen Karrierewege identifiziert werden, die mit diesen Höherqualifizierungsprozessen einhergehen und letztlich zu den Einkommensrenditen führen.

Theoretisch stützt sich die Studie auf Argumente der Ansätze des „Compensatory Advantage“ (Bernardi 2014) und der „Effectively Maintained Inequality“ (Lucas 2001). Demnach verfügen Personen aus privilegierter sozialer Herkunft sowohl über höhere Ressourcen als auch höhere Anreize (Vermeidung sozialen Abstiegs), in Bildung zu investieren – gerade auch, wenn sie mit dem ersten Bildungsabschluss noch nicht das maximal erreichbare Bildungsziel ausgeschöpft haben. Aufbauend auf dieser Argumentation wird erwartet, dass Höherqualifizierungsprozesse im allgemeinbildenden Bereich für Personen aus privilegierter sozialer Herkunft häufiger in akademische Karrierewege münden als für Personen aus benachteiligter sozialer Herkunft (Schindler 2014), welche wiederum mit höheren Einkommenserträgen verbunden sind (vgl. Piopiunik et al. 2017). Das gleiche Muster wird für Höherqualifizierungsprozesse nach Erwerb eines ersten beruflichen Abschlusses erwartet, da sich die grundlegenden Mechanismen (sozial selektive Anreize und Ressourcen) nicht unterscheiden sollten. Zusammengefasst führen diese Argumente zu folgenden Erwartungen:

- (1) Über sozial selektives Verhalten beim Erwerb weiterer Bildungsabschlüsse werden langfristig Einkommensvorteile für Personen privilegierter sozialer Herkunft generiert.
- (2) Dies gilt sowohl für Höherqualifizierungen im allgemeinbildenden als auch im berufsbildenden Bereich.
- (3) Den größten Beitrag bei der Generierung der sozialen Einkommensunterschiede leisten Karriereverläufe, die nach der Höherqualifizierung den Erwerb eines Hochschulabschlusses beinhalten.

Die empirischen Analysen stützen sich auf die Erwachsenenkohorte (SC6) des Nationalen Bildungspanels (NEPS). Diese Daten wurden erstmals umfassend und systematisch zur Analyse der enthaltenen Informationen zu Bildungsverläufen und Einkommen bereinigt, aufbereitet und multipel imputiert.

Es werden die Geburtskohorten 1970-79 herangezogen, deren Einkommen im Alter von 40 Jahren (Konzept der „occupational maturity“) im Erhebungszeitraum beobachtet werden können. Um den Einfluss der Höherqualifizierungen auf das spätere Einkommen identifizieren zu können, kommen Matchingverfahren zum Einsatz. Dabei werden Personen unterschiedlicher sozialer Herkunft (gemessen über das EGP-Klassenkonzept) miteinander verglichen, die denselben ersten allgemeinbildenden oder berufsbildenden Bildungsabschluss aufweisen und sich auf weiteren Merkmalen ähnlich sind. In einem ersten Schritt wird dann anhand von Dekompositionsanalysen untersucht, inwieweit Unterschiede in der Höherqualifizierung zur Generierung von sozialen Unterschieden im Brutto-Stundenlohn im Alter von 40 Jahren beitragen. In einem zweiten Schritt werden die Beiträge der Höherqualifizierungen weiter danach zerlegt, mit welchen nachfolgenden Karriereverläufen (Bildungs- und Berufsverläufe) sie einhergehen.

Erste Analysen stützen die theoretischen Erwartungen. Das sozial selektive Nachholen von Bildungsabschlüssen kann auch mit den vorliegenden Daten repliziert werden. Beim Vergleich von Personen mit identischen ersten Bildungsabschlüssen führt das sozial selektive Nachholverhalten zu Vorteilen im Brutto-Stundenlohn für Personen aus Dienstklassenfamilien im Vergleich zu Personen aus Arbeiterfamilien. Diese Befunde sind vorläufig und werden bis zur Konferenz durch weitere Robustheitskontrollen überprüft. Ferner werden auch weitere Analysen zur Überprüfung der restlichen Erwartungen durchgeführt.

Exploring Scholarship Effects of the Studienstiftung on Bachelor to Master Transitions

Sebastian Lang¹, Jessica Ordemann²

¹Leibniz-Institut für Bildungsverläufe (LIfBi), Deutschland; ²Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), Deutschland

As part of the Bologna reform, a two-tier higher education system with a Bachelor's and Master's program was introduced in Germany. This institutional change created another transition within tertiary education where social inequality can arise (Lörz et al., 2016; Neugebauer et al., 2016). Here, students from less privileged family background less often remain in higher education and more often transition into the labor market. This has manifold consequences as graduates with a bachelor's degree earn less than those with a Master's degree and have lower returns in other life spheres (Neugebauer & Weiss, 2017). At the same time, Germany has a whole range of organizations for the promotion of gifted students that are politically, religiously, trade unionist, entrepreneurial oriented or politically and ideologically independent. The „Studienstiftung des deutschen Volkes“ (Studienstiftung) is the oldest and largest of these scholarship foundations and aims to promote talented young people to higher education – independently of their background. This study takes a causal perspective and investigates the influence of a scholarship of the Studienstiftung on transitions from bachelor's to master's degree programs. We put an emphasis on potential heterogeneities within the transition to a master's program and in addressing social disparities. Based on capital theory, we expect the scholarship to increase the probability of starting a master's degree program by providing scholars with different types of additional capital,

such as economic capital in the form of financial support and cultural or social capital in the form of idealistic support. Therefore, our first research question is: Can we find an effect of Studienstiftung scholarships on transitions from Bachelor's to Master's degrees? Although the promotion by the Studienstiftung follows meritocratic principles, nevertheless, there could also be unintended consequences by fostering social disparities in the transitions from a Bachelor's to a Master's degree. At the same time, Studienstiftung could probably reduce existing social disparities by providing necessary support for excellent students at this stage of their educational career. Therefore, our second and third research questions are as follows: Can heterogeneous effects be found by social groups, such as migration background, social origin, and gender? And: What is the effect on known social disparities? To answer these questions, we analyze data from the Graduate Panel 2017 of the German Centre for Higher Education Research (DZHW) and an additional sample of Studienstiftung scholars from the same graduation cohort. We first estimate the probability to transition to a Master's program using a boosting algorithm (Schonlau 2005) for a propensity score before applying propensity score matching with 3 nearest neighbors with ties to examine these questions. We have 3220 untreated and 560 treated Bachelor graduates on common support. We check for heterogeneous effects for sex, parental education and migration background. Our finding indicate no or little causal effect. However, by shedding light on these dynamics, our research aims to contribute to the discourse on educational mobility and social stratification, offering insights that can inform policy interventions and institutional practices aimed at fostering a more inclusive higher education landscape.

Berufliche Aspirationen im Jugendalter: Das Zusammenspiel von Geschlecht, SES und institutionellen Strukturen

Ricarda Ullrich^{1,2}, Victoria Zeddies², Katharina Kohl¹, Michael Becker^{3,2}

¹IPN | Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik; ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; ³Institut für Schulentwicklungsforschung (IFS), TU Dortmund

Theoretische und empirische Arbeiten unterstreichen die Bedeutung beruflicher Aspirationen für den Bildungs- und Berufserfolg, weisen jedoch auf individuelle Differenzen in den beruflichen Aspirationen über die Adoleszenz hinweg hin (Gao & Eccles, 2020; Gottfredson, 1981). Basler et al. (2021, S. 18) diskutieren, dass die Entstehung und Entwicklung beruflicher Aspirationen ein Resultat des Zusammenspiels von Geschlecht, sozioökonomischem Hintergrund (SES) und institutionellen Strukturen ist.

Geschlechterdifferenzen in Berufsaspirationen werden theoretisch auf stereotype Geschlechtsidentitäten und/oder Geschlechterrollen zurückgeführt (Gottfredson, 2005; Miyamoto & Wicht, 2020). Eine mögliche Erklärung für diese Differenzen ist die Verinnerlichung sozialer Erwartungen, die sich in der unterschiedlichen Bewertung beruflicher Optionen widerspiegelt (Dicke et al., 2019). Geschlechterrolleneinstellungen bieten Einblicke in die normativen Erwartungen der Macht- und Arbeitsverteilung zwischen Männern und Frauen. Studien zu Geschlechterungleichheiten zeigten, dass weibliche Jugendliche höhere Berufsaspirationen aufwiesen als männliche Jugendliche (vgl., Miyamoto & Wicht, 2020). Dies ist insofern überraschend, als Befunde zum beruflichen Erfolg recht übereinstimmend ein tendenziell geringeres Berufsprestige sowie Einkommen von Frauen finden (Becker et al., 2019). Studien, die sich mit den zugrundeliegenden Konstrukten und Prozessen von Geschlechterdifferenzen in Berufsaspirationen beschäftigen, gibt es bisher wenige. Hinsichtlich des SES sprechen sowohl theoretische Annahmen (Gottfredson, 1981) als auch empirische Befunde (Autor:innengruppe Bildungsberichterstattung, 2022) für einen positiven Zusammenhang zwischen SES und beruflichen Aspirationen von Schüler:innen. Neben diesen individuellen Merkmalen werden Gelegenheitsstrukturen, insbesondere in stark gegliederten Schulsystemen, als wichtige soziale Umwelten für die Entwicklung von Berufsaspirationen diskutiert (Heckhausen & Tomasik, 2002), was Befunde zur unterschiedlichen Entwicklung beruflicher Aspirationen an verschiedenen Schulformen belegen (Zeddies et al., under review).

Theoretische (Gottfredson, 2005) und erste empirische Arbeiten unterstreichen zudem die Interkorrelation dieser drei zentralen Prädiktoren: So weisen Schüler:innen mit einem privilegierten SES egalitäre Geschlechterrolleneinstellungen auf (Ullrich et al., 2022) und Schüler:innen mit einem weniger privilegierten SES am Gymnasium senken ihre Berufsaspirationen (Zeddies et al., under review). Wie diese Interkorrelationen zwischen Geschlechterrollen, SES und Schulform und ihre Bedeutung für Berufsaspirationen systematisch aussehen, ist bislang nicht untersucht.

Im vorliegenden Beitrag wird daher eine kombinierte und integrierte Perspektive eingenommen, die für ein umfassendes Verständnis von Berufsaspirationen unabdingbar ist. Dabei wird – unseres Wissens erstmalig – die Bedeutung des Zusammenspiels von Geschlechterrolleneinstellungen, SES und institutionellen Strukturen für (die Entwicklung von) Berufsaspirationen während der Adoleszenz untersucht. Der Forschungsstand wird in zweierlei Hinsicht erweitert: Zum einen werden – über einfache Geschlechterunterschiede hinaus – Geschlechterrolleneinstellungen und ihr Zusammenhang mit Berufsaspirationen untersucht. Zum anderen werden Interaktionen der zentralen Prädiktoren Geschlechterrolleneinstellungen, SES und Schulform und ihre Bedeutung für berufliche Aspirationen geprüft.

Zur Beantwortung der Fragestellung werden zwei beziehungsweise drei Wellen des BIJU-Datensatzes ($N = 10.470$) verwendet. Die Schüler:innen wurden in der 7. Klasse (1992), in der 10. Klasse (1995) sowie die Schüler:innen am Gymnasium in der 12. Klasse (1997) nach ihren beruflichen Wünschen gefragt. Die offenen Berufsangaben wurden auf der Grundlage des *International Socio-Economic Index of Occupational Status* (ISEI; Ganzeboom, 2010) kodiert. Dabei bietet der Datensatz eine besondere Perspektive, da die Entwicklung beruflicher Aspirationen kurz vor dem Abschluss (10. und 12. Klasse) unter Berücksichtigung des vorherigen Berufswunsches (7. Klasse) untersucht werden kann. Neben den Berufsaspirationen liegen Informationen zu den Einstellungen zu Geschlechterrollen zu allen drei Messzeitpunkten vor, was sowohl die Untersuchung der Entwicklung von Geschlechterrolleneinstellungen als auch die Entwicklung beruflicher Aspirationen erlaubt. Informationen zum Geschlecht, dem SES sowie der besuchten Schulform liegen ebenfalls zu mehreren Messzeitpunkten vor. Aufgrund des large-scale Designs des Datensatzes kann für weitere wichtige Prädiktoren wie Migrationshintergrund, kognitive Fähigkeiten, schulische Noten sowie schulische Interessen kontrolliert werden. Als Analysemodelle werden längsschnittliche lineare Mehrgruppenregressionsmodelle sowie Cross-Lagged-Panel Modelle gerechnet, welche ermöglichen, die wechselseitige Entwicklung von Geschlechterrolleneinstellungen und Berufsaspirationen genauer zu untersuchen. Zusätzlich werden Interaktionseffekte berechnet, um das Zusammenspiel von Geschlechterrolleneinstellungen, SES und der Schulformen zu untersuchen. Ergebnisse sowie zentrale Implikationen werden auf der Konferenz präsentiert.

Das Streben nach Aufwärtsmobilität bei Jugendlichen aus Familien mit niedrigem sozioökonomischen Status in verschiedenen Länderkontexten: Ein internationaler Vergleich mit PISA 2018

Thomas Zimmermann¹, Birgit Becker^{1,2}

¹Goethe-Universität Frankfurt; ²Forschungszentrum "Individual Development and Adaptive Education of Children at Risk" (IDEA)

Einleitung und Fragestellung

Es gibt eine reichhaltige Literatur zu beruflichen Aspirationen von Schüler:innen und deren Zusammenhang mit Bildungsentscheidungen und späterem beruflichen Status. Insbesondere soziale Unterschiede und deren Einfluss auf Berufsaspirationen wurden vielfach untersucht (vgl. Schoon & Polek, 2011). Weniger beachtet wurde jedoch, wie der Zusammenhang zwischen sozialer Herkunft und beruflichen Aspirationen von Schüler:innen nach Länderkontexten variiert, was jedoch für bildungspolitische Implikationen höchst relevant ist. Wir untersuchen daher mit einer internationalen Vergleichsstudie das berufliche Statusstreben in unterschiedlichen sozialen Gruppen mittels der folgenden zwei Fragen:

- Wie unterscheiden sich Jugendliche aus Familien mit unterschiedlichem sozioökonomischem Status (SES) in ihrem Statusstreben?
- Inwiefern werden diese Zusammenhänge durch den Länderkontext moderiert?

Theoretischer Hintergrund

Soziale Unterschiede im beruflichen Statusstreben werden von verschiedenen Theorien vorhergesagt, wie etwa dem Wisconsin-Modell oder soziologischen Rational Choice Modellen (RCT) zu Bildungsentscheidungen (Sewell et al., 1969; Breen & Goldthorpe, 1997; Erikson & Jonsson, 1996). Mit dem Statuserhaltungsmotiv betont die soziologische RCT (Breen & Goldthorpe, 1997; Esser, 1999) sogar explizit soziale Ungleichheiten im beruflichen Statusstreben, da Familien nach den theoretischen Annahmen ausschließlich motiviert sind, einen intergenerationalen Statusabstieg zu vermeiden. Einem Statusaufstieg wird hingegen keine über den Statuserhalt hinausgehende Bedeutung zugesprochen. Im Gegensatz dazu sehen andere Ansätze das Streben nach Statusverbesserung als zentralen menschlichen Antrieb (Lipset & Zetterberg, 1959), der speziell bei Schüler:innen aus Familien mit niedrigem SES zu höherem Bildungsengagement führen kann (Browman et al., 2017, 2022). Die Angemessenheit dieser konkurrierenden theoretischen Annahmen sollte nach Länderkontexten variieren. Wir untersuchen dabei die Hypothese, dass das Streben nach einem hohen Berufsstatus bei Jugendlichen aus Familien mit niedrigem SES umso wahrscheinlicher ist, je mehr Opportunitäten und je mehr Anreize sie dafür in einem Land sehen.

Daten und Methode

Wir verwenden die Daten des „Programme for International Student Assessment“ (PISA) aus dem Jahr 2018. Unser Analysesample umfasst ca. 350.000 Jugendliche aus 68 Ländern.

Wir operationalisieren das berufliche Statusstreben durch die Angabe der Jugendlichen zu ihrem angestrebten Beruf im Alter von 30 Jahren, recodet nach dem "International Socio-Economic Index of Occupational Status" (ISEI) (Ganzeboom et al., 1992).

Als Makroindikatoren auf Länderebene nutzen wir den Grad der Stratifizierung des Bildungssystems (gemessen über das Alter beim ersten Tracking und die Anzahl der Tracks) sowie das Ausmaß an Bildungsmobilität im jeweiligen Land als Indikatoren für Opportunitäten bzw. Restriktionen in Bezug auf soziale Mobilität. Als Indikatoren für Anreize verwenden wir die wirtschaftliche Situation (gemessen über das Bruttoinlandsprodukt und die Arbeitslosenrate) sowie das Ausmaß an ökonomischer Ungleichheit (gemessen über den Gini-Index), da vermutet werden kann, dass die objektiven Anreize für Statusverbesserung in ärmeren Ländern größer sind und Statusverbesserungen in Ländern mit höherer Einkommensungleichheit relevanter sind (vgl. Mijs, 2019).

Zur Auswertung verwenden wir lineare Regressionen mit Fixed Effects für die Länder und Cross-level Interaktionen (d. h. Interaktionen zwischen elterlichem SES und den jeweiligen Ländermerkmalen).

Ergebnisse

Unsere Befunde zeigen, dass Schüler:innen aus Familien mit hohem SES durchschnittlich nach statushöheren Berufen streben als ihre weniger privilegierten Peers. Dieser Befund variiert jedoch mit den hier untersuchten Ländermerkmalen: Je weniger Opportunitäten und je weniger Anreize es für Schüler:innen aus Familien mit niedrigem SES für sozialen Aufstieg gibt, desto stärker fällt der Zusammenhang zwischen SES und Statusstreben aus. Besonders ausgeprägt ist der Zusammenhang daher in Ländern mit einem stärker stratifizierten Bildungssystem, geringer Bildungsmobilität, einem höheren BIP und niedrigerer Einkommensungleichheit. Mit diesen Befunden können wir erstmals zeigen, dass Schüler:innen aus Familien mit niedrigem SES umso mehr nach sozialer Aufwärtsmobilität streben, je mehr Opportunitäten und je mehr Anreize sie dafür im jeweiligen Land finden. Während statushohe Schüler:innen über unterschiedliche Kontexte hinweg nach Statuserhalt streben, erweist sich das Statusstreben von Jugendlichen aus Familien mit niedrigem SES hingegen der Annahme der soziologischen RCT als stark kontextabhängig, was wiederum Möglichkeiten für politische Interventionen bietet.

PS-53: Diversität in der Bildung

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 133

Chair der Sitzung: Diana Raufelder

„Ich glaube, dass Vielfalt uns bereichern kann“. Eine Interviewstudie zu Perspektiven von Lehrkräften auf Vielfalt in der Schule

Cindy Ballaschk, Camilla Rjosk

Universität Potsdam, Deutschland

Abstract anlässlich der 12. GEBF-Tagung vom 27.-29.01.2025 in Mannheim

Titel: „Ich glaube, dass Vielfalt uns bereichern kann“. Eine Interviewstudie zu Perspektiven von Lehrkräften auf Vielfalt in der Schule

Autor_innen: Cindy Ballaschk, Camilla Rjosk

Institution: Universität Potsdam, Professur für Schulentwicklung

Theoretischer Hintergrund

In der Bildungsforschung wird der Umgang mit Vielfalt, insbesondere im Kontext der inklusiven Bildung und Chancengleichheit, als zentrales Thema verhandelt (Tillmann, 2022). Ein differenziertes Verständnis von Vielfalt ist essenziell für die Gestaltung inklusiver Schulen (Gloystein & Barth, 2020). Vielfalt soll dabei aus einer machtkritischen und intersektionalen Perspektive als ein Zusammenwirken unterschiedlicher Heterogenitätsdimensionen verstanden werden (Rott, 2022). Die Bildungsforschung zeigt, dass Lehrkräfte dazu tendieren, Schulklassen als homogene Lerngruppen zu konzeptualisieren und ihre vielfältige Eingebundenheit in unterschiedliche Machtverhältnisse unzureichend anerkennen (Gloystein & Barth, 2020; Budde, 2013). Daher stellt sich die Frage, wie Lehrkräfte Vielfalt in Bezug auf ihre Schüler_innen wahrnehmen, welche Heterogenitätsdimensionen mitgedacht werden und welche außen vor bleiben. Dies hilft Lehrkräfte dabei zu unterstützen, ihren Schüler_innen individuelle Lernangebote zu schaffen und Diskriminierungen abzubauen sowie eine produktive und teilhabeorientierte Atmosphäre zu fördern (Ainscow et al., 2006; Gloystein & Barth, 2020).

Fragestellung

Der Vortrag zeigt Perspektiven von Lehrkräften auf Vielfalt in der Schule. Zentral ist die Frage: Wie nehmen Lehrkräfte Vielfalt in Bezug auf ihre Schüler_innen wahr, welche Heterogenitätsdimensionen berücksichtigen sie dabei, und welche bleiben möglicherweise unbeachtet?

Methode

Für die Untersuchung wurde ein qualitatives Forschungsdesign gewählt. Es wurden sieben problemzentrierte Interviews (Witzel, 2000) mit Lehrkräften an vier verschiedenen Berliner Sekundarschulen geführt. Die Fallauswahl erfolgte in Anlehnung an ein „purposeful Sampling“ (Palinkas et al., 2015), wobei gezielt kontrastreiche Kriterien berücksichtigt wurden. Die Interviews wurden transkribiert (Dresing & Pehl, 2015) und mit der Grounded Theory Methode ausgewertet. Dazu wurde das Datenmaterial computergestützt offen, axial und selektiv codiert (Corbin & Strauss, 1990). Dies ermöglichte es, zentrale Themen und Zusammenhänge in Bezug auf Perspektiven auf Vielfalt systematisch zu identifizieren und zu analysieren.

Ergebnisse

Vielfalt wird in den Interviews in verschiedenen Dimensionen diskutiert, z.B. Geschlecht, Herkunft, Sprache und Befähigung. Zugleich wird Vielfalt teilweise auf sonderpädagogische Förderbedarfe verengt, insbesondere dann, wenn Lehrkräfte ihren pädagogischen Umgang mit der Vielfalt ihrer Schüler_innen in der Schulpraxis beschreiben. Lehrkräfte artikulieren Vielfalt in der Schule als Spannungsfeld zwischen Anerkennung und Herausforderung. Vielfalt wird sowohl als wertvoll und bereichernd angesehen, aber auch als Quelle von Konflikten und Machtstrukturen. Der Anspruch, dass jede_r wahrgenommen werden sollte, wird als hoher, jedoch oft unerfüllter Standard dargestellt, der im schulischen Alltag häufig zu Frustrationen führt. Insgesamt ist auffällig, dass zwar eine positive Haltung zu Vielfalt eingenommen wird, aber zugleich die Vielfalt von Schüler_innen aus einer defizitorientierten Perspektive verhandelt wird. Dies verdeutlicht die Notwendigkeit Lehrkräfte darin zu unterstützen, konkrete Strategien und Methoden für einen inklusiveren Umgang mit Vielfalt zu erarbeiten, um die Diskrepanz zwischen Haltung und Praxis zu überwinden.

Die Wirksamkeit von Diversitätstrainings – eine Metaanalyse

Frederik Zirn¹, Susanne Wißhak²

¹Universität Konstanz; ²Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau

Im Rahmen Lebenslangen Lernens sollen Wissen, Fähigkeiten und Werthaltungen erworben werden, die für einen toleranten und konstruktiven Umgang mit Vielfalt in einer pluralistischen Gesellschaft erforderlich sind. Dies gilt auch für Bildungsprozesse in Organisationen. Folglich setzen immer mehr Organisationen auf Diversitätstrainings, welche darauf abzielen Vorurteile und Diskriminierung zu reduzieren (Bezrukova et al., 2016). Obwohl beträchtliche Ressourcen für die Fortbildung von Mitarbeitenden in diesem Bereich aufgewendet werden, ist wenig darüber bekannt, ob die Teilnehmenden das Wissen, die Fähigkeiten und die Einstellungen, welche in Diversitätstrainings vermittelt werden, am Arbeitsplatz anwenden. Bisherige Metaanalysen konzentrierten sich auf die Frage, ob Teilnehmende einen *Lernzuwachs* nach Diversitätstrainings erzielten (Bezrukova et al., 2016; Kalinoski et al., 2013). Verschiedene Befunde deuten darauf hin, dass die erfolgreiche Aneignung von Lerninhalten nicht zwangsläufig zur späteren *Anwendung des Gelernten* im beruflichen Kontext führt (Alliger et al., 1997; Karl & Ungsrithong, 1992). Der Transfererfolg einer Trainingsmaßnahme kann gemäß dem Modell von Baldwin & Ford (1988) von unterschiedlichen Faktoren wie dem Trainingsdesign, den Merkmalen der Trainingsteilnehmenden und der Arbeitsumgebung abhängen. In unserer Metaanalyse gehen wir den Fragen nach, ob der Transfer der Trainingsinhalte von Diversitätstrainings in den Arbeitsalltag gelingt und welche Faktoren den Transfererfolg in diesem Kontext beeinflussen.

Um geeignete Studien zu identifizieren, wurde eine systematische Literaturrecherche gemäß dem PRISMA-Statement durchgeführt (Page et al., 2021). Die Einschlusskriterien für unsere Meta-Analyse wurden anhand des PICOS-Frameworks (Participants, Intervention, Control, Outcome, Study design) abgeleitet (Amir-Beghadami & Janati, 2020). Anhand einer Suche in fünf verschiedenen Datenbanken wurden 4853 potenzielle Artikel identifiziert. 12 Studien erfüllten alle Einschlusskriterien und wurden in die Metaanalyse aufgenommen. Relevante Informationen wurden von den Autor*innen zunächst unabhängig voneinander kodiert. Die Interrater-Reliabilität ergab bei allen Items eine Übereinstimmung von über 90%. Die Effektstärken aus

den jeweiligen Studien, standardisierte Mittelwertsunterschiede (Hedges g), wurden mithilfe des packages „metafor“ in R kalkuliert (Viechtbauer, 2010). Bei mehreren Effektstärken pro Studie wurde eine aggregierte Effektstärke kalkuliert, um zu gewährleisten, dass die Daten unserer Metaanalyse unabhängig voneinander sind (Borenstein et al., 2021). Die Effektstärken wurden im Anschluss anhand einer Random-Effects Metaanalyse kombiniert. Wir wandten die Knapp-Hartung-Methode an, um robuste Konfidenzintervalle und Signifikanztests für den Gesamteffekt zu erhalten (Int'Hout et al., 2014). Neben dem generellen Transfereffekt von Diversitätstrainings wurde anhand von Subgruppenanalysen der Einfluss von theoretisch hergeleiteten Moderatoren untersucht (Wißhak, 2022). Die Moderatorenanalysen basierten auf einem Mixed-Effects-Modell.

Im Rahmen der Random-Effects Metaanalyse wurde ein kleiner statistisch signifikanter Effekt von Diversitätstrainings auf das Arbeitsplatzverhalten identifiziert (Hedges $g = 0.29$, 95%-CI = 0.12-0.47). Die Effektstärken variieren stark zwischen den einzelnen Studien, weshalb unsere Ergebnisse geprägt sind durch eine hohe Heterogenität ($Q = 40.48$; $p < 0.001$; $I^2 = 70.1\%$). Moderatoranalysen geben Hinweise darauf, dass eine konkrete Planung des Transferverhaltens während des Trainings in Form von Aktionsplänen den Transfererfolg von Diversitätstrainings erhöhen kann ($p = .026$). Unsere Ergebnisse zeigten keine Anhaltspunkte dafür, dass der Transfererfolg vom verpflichtenden oder freiwilligen Charakter des Trainings sowie von der vorherigen Durchführung einer Bedarfsanalyse beeinflusst wird. Bei der Interpretation der Ergebnisse gilt es zu beachten, dass Studien, welche ausschließlich auf Selbsteinschätzungen basieren, höhere Effekte berichteten, als solche, welche objektive Bewertungsmaßstäbe wählten ($p = .031$).

Die Ergebnisse der Metaanalyse legen nahe, dass Bildungsanstrengungen in Organisationen in Form von Diversitätstrainings nicht nur zu Lernerfolgen, sondern auch zu Verhaltensänderungen führen können. Darüber hinaus liefert die Moderatoranalyse erste Ansatzpunkte für eine transferförderliche Gestaltung von Diversitätstrainings. Weitere Moderatoren sollten in Zukunft untersucht werden. Zudem wären mehr Originalstudien, die sich nicht allein auf Selbstauskünfte stützen, wünschenswert.

Motivationsdynamiken von Kontext, Selbst und schulischem Engagement unter Gymnasialschüler:innen

Olga Steinberg, Diana Raufelder
Universität Greifswald, Deutschland

Theoretischer Hintergrund

Das Self-System Model of Motivational Development (SSMMD), entwickelt von Skinner et al. (2008), konzentriert sich auf das schulische Engagement als einen zentralen Bestandteil der motivationalen Erfahrungen von Schüler:innen. Dieses Modell betont die dynamischen Prozesse, die zum Engagement im Klassenzimmer beitragen und hebt die Zusammenhänge verschiedener motivationaler Variablen im Bildungskontext hervor (Dinçer et al., 2019; Pitzer, 2000; Karasinski & Anderson, 2017). Indem Faktoren des Kontexts, des Selbst und der Aktionsebene berücksichtigt werden, bietet das SSMMD ein Rahmenwerk, um zu verstehen, wie sich aus diesem Zusammenspiel das schulische Engagement generiert (Reeve, 2012; Furrer & Skinner, 2003). In dieser Studie wurde in Anlehnung an das SSMMD die Bedeutung der Unterstützung durch Lehrkräfte, Peers und Eltern (*context*) unter Berücksichtigung des akademischen Selbstkonzepts und der Befriedigung der drei motivationalen Grundbedürfnisse (Autonomie, soziale Eingebundenheit und Kompetenz) (*self*) auf die Entwicklung des emotionalen und behavioralen schulischen Engagements von Gymnasialschüler*innen (*action*) von Beginn der 8. Klasse (T1) bis zum Ende der 9. Klasse (T2) untersucht.

Fragestellung

Es wurde der Fragestellung nachgegangen, ob – entsprechend des SSMMD – verschiedene Kontext- und Selbst-Variablen in Zusammenhang mit der Entwicklung des emotionalen und behavioralen Schulengagements von Gymnasiast:innen im Verlauf von der achten bis zum Ende der neunten Klasse fungieren. Im Detail galt es folgende Hypothesen zu überprüfen:

- H1: Aufbauend auf bisherigen Befunden wurde in dieser Studie von einer Abnahme des behavioralen und emotionalen schulischen Engagements von T1 zu T2 ausgegangen.
- H2: Sowohl Kontextvariablen (wahrgenommene Unterstützung durch die Eltern, Peers und Lehrkräfte) als auch Selbstvariablen stehen in Zusammenhang mit der Veränderungen im Engagement der Schüler:innen.
- H3: Die Selbstvariablen fungieren – entsprechend des SSMMD Modells – als Mediatoren zwischen Kontext und schulischem Engagement.

Da Geschlecht und Leistung das Zusammenspiel beeinflussen können, wurden diese als Kontrollvariablen mitberücksichtigt.

Methode

Es wurden Fragebogendaten von jugendlichen Gymnasialschüler:innen ($N = 779$; $M_{Age} = 13.09$; $SD = .50$; 57% weiblich) genutzt. Die Schüler:innen stammten aus 13 Gymnasien in Brandenburg. Sie haben Skalen zur Unterstützung von Peers, Lehrkräften, Eltern, sowie ihrem akademischen Selbstkonzept, der Befriedigung der drei Grundbedürfnisse, sowie zu ihrem schulischen Engagement ausgefüllt. Diese Skalen wurden in einem Latent Change Modell inklusive Pfad-Analyse und indirekter Effekte mittels Mplus 8.8 untersucht.

Ergebnisse und ihre Bedeutung

Die Ergebnisse bestätigen H1: das behaviorale und das emotionale schulische Engagement nehmen im Laufe von 1.5 Jahren signifikant ab. Auch H2 konnte teilweise bestätigt werden: Allerdings sagt – außer dem schulischen Selbstkonzept in Bezug auf das emotionale schulische Engagement – keine der Variablen die Veränderung des schulischen Engagements vorher. D.h., ein positives schulisches Selbstkonzept kann vor einer Abnahme des emotionalen Schulengagements schützen. Allerdings fungieren fast alle Variablen des Kontexts (außer der väterlichen schulischen Unterstützung) als Prädiktoren der Ausgangswerte des emotionalen und behavioralen Schulengagements. Von den Variablen des Selbst konnte nur das schulische Selbstkonzept als Prädiktor identifiziert werden (nicht aber die Befriedigung der drei Grundbedürfnisse). Für die Befriedigung der drei Grundbedürfnisse spielt wiederum vor allem die Wahrnehmung von Lehrkräften als positive Motivatoren eine zentrale Rolle. H3 konnte nur für zwei von 32 indirekten Pfaden bestätigt werden, was wichtige Hinweise für die Weiterentwicklung des SSMMD liefert, insofern die Beziehung zwischen Kontext, Selbst und der Aktionsebene eher direkte (und nicht indirekte) Natur zu sein scheint (keine Mediation). Lediglich für den Zusammenhang zwischen mütterlicher schulischer Unterstützung und den Ausgangswerten des schulischen Engagements konnte das allgemeine schulische Selbstkonzept als Mediator identifiziert werden. Insgesamt heben die Ergebnisse die zentrale Stellung des akademischen Selbstkonzept auf der Selbstebene, sowie die wahrgenommene motivationale Unterstützung der Mutter, Peers und Lehrkräften hervor, die folglich als vielversprechende Ansatzpunkte zur Vermeidung des Absinkens des schulischen Engagements verstanden werden können.

PS-54: Führung und Entwicklung im Schulwesen: Internationale Perspektiven und aktuelle Herausforderungen

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SO 318

Chair der Sitzung: Olivia Wohlfart

Schulleitungs-Barometer Austria 2024: Aktuelle Herausforderungen in Bildung, Schule und Schulleitung in Österreich

Jane Pruitt^{1,2}, Stephan Gerhard Huber^{1,3}, Christoph Helm³

¹Arbeitsgruppe Bildungsmanagement, Schweiz; ²Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Deutschland; ³Johannes Kepler Universität Linz, Österreich

Die Bedeutung des Schulleitungshandelns für die Qualität und Entwicklung von Schulen ist wissenschaftlich gut gestützt: Umfangreiche empirische Bemühungen der Schulwirksamkeitsforschung ergaben, dass die pädagogische Steuerung von Schule durch die Schulleitung ein zentraler Faktor für die Qualität einer Schule ist. Auch Studien zu Schulentwicklung betonen die Relevanz von Schulleitung, besonders in Hinblick auf den angestrebten kontinuierlichen Verbesserungsprozess in einer jeden Schule (vgl. u. a. Autor 2010, 2016a, 2016b, 2019a, 2019b; Coelli & Green, 2012; Goldring et al., 2009; Grissom et al., 2021; Hallinger & Kovačević, 2019; Kovačević & Hallinger, 2019; Leithwood et al., 2019; Leithwood & Louis, 2012; Rhodes & Brundrett, 2009; Robinson & Gray, 2019; Sammons et al., 1995; Teddlie & Stringfield, 1993).

Angesichts der bildungspolitischen Entwicklungen, darunter vor allem der Verlagerung von Management- und Führungsaufgaben sowie Entscheidungskompetenzen auf die Einzelschulen, sehen sich Schulleiter*innen neuen und erweiterten Anforderungen gegenüber (Autor, 2005, 2007). Auch infolge allgemeingesellschaftlicher Veränderungen sind die Anforderungen an Schulleitungen und somit auch die Komplexität der Schulleitungstätigkeit über die letzten Jahrzehnte gestiegen. Dies kann u. a. zu einem erhöhten Belastungsempfinden bei Schulleitungen führen (z. B. Borg & Riding, 1993; Boyland, 2011; Dadaczynski & Paulus, 2015, 2016; Döbrich et al., 1995; Maxwell & Riley, 2016; Ostermann & Saueressig, 1994).

Die aktuellen Bildungsreformen und derzeitigen Entwicklungen im österreichischen Schulsystem erfordern daher eine Berücksichtigung der Meinungen, Erfahrungen und Perspektiven der Praxis, insbesondere von Schulleitenden. Das Schulleitungs-Barometer Austria 2024 befragt Schulleitende in ganz Österreich aus allen Bundesländern und allen Schulformen zu aktuellen bildungspolitischen und berufspraktischen Themen von Bildung, Schule und Schulleitung. Auf der Grundlage dieser Expertise aus der täglichen Führungsarbeit sollen Problemfelder ermittelt und Anregungen und Empfehlungen für die Bildungspolitik und Verwaltung/Aufsicht sowie die Schulleitungspraxis entwickelt werden.

Die zentralen Fragestellungen des Schulleitungs-Barometers Austria 2024 lauten:

1. Wie stellt sich die aktuelle Arbeitssituation und Arbeitsbelastung von Schulleitungen in Österreich dar?
2. Wie werden aktuelle Entwicklungen und Anforderungen durch die Schulleitungen erlebt und eingeschätzt wie die aktuellen Bildungsreformen, die Änderungen im Dienstrecht und der Lehrer*innenbildung, die Ressourcen zur Unterstützung, der Personalmangel und die Gestaltung der Reifeprüfung?

Die Studie wurde mittels einer standardisierten Online-Befragung durchgeführt. Im Zentrum der Befragung standen die aktuelle Arbeitssituation und Arbeitsbelastung von Schulleitungen sowie ihre Einschätzung zu aktuellen Entwicklungen und Anforderungen. Die Stichprobe besteht aus rund 40% aller Schulleitenden in Österreich.

Es werden Regressionen, Profilanalysen und Strukturgleichungsmodelle insbesondere zur Erklärung des Belastungserlebens präsentiert. So zeigt sich insgesamt eine eher hohe Arbeitszufriedenheit der Schulleitenden. Allerdings ist die Gruppe der Schulleitungen, die ein hohes Stresserleben haben, eine hohe berufliche Belastung empfinden und eine niedrig ausgeprägte Arbeitszufriedenheit aufweisen, mit 41% ausgesprochen hoch. Eine der Ursachen für diese empfundene Belastung könnte darin liegen, dass der Aufgabenbereich von Schulleitenden nach eigenen Angaben in den letzten fünf Jahren eher größer (20%) bzw. viel größer (71%) geworden ist. Auch Unterstützungsmaßnahmen und Ressourcen vor Ort spielen eine Rolle. Als Prädiktoren der Belastung lassen sich weitere Variablen identifizieren: Konflikte mit Eltern, schlechte Planbarkeit von Aufgaben, aufwändige Archivierung und Arbeiten fürs Bundesministerium, die als Gründe für Überstunden angegeben wurden. Zentrales Ergebnis ist also, dass die administrative Belastung ein zentraler Faktor für das allgemeine Belastungsempfinden ist.

Darüber hinaus wird die Arbeitssituation von Schulleitungen in Österreich anhand ausgewählter deskriptiver Befunde u. a. zu Arbeitsbelastung und Arbeitszufriedenheit, Personal- bzw. Lehrkräftemangel, Auswirkungen der Bildungsreform, Unterstützungspersonal und weitere Ressourcen, Einführung des neuen Dienstrechts sowie der Lehrer*innenausbildung beleuchtet.

Die Ergebnisse bieten einen Ansatzpunkt, um die Perspektiven der Schulleitungen sichtbar zu machen und in die bildungspolitische Diskussion einzubringen. So werden mit Blick auf die große Verantwortung, die Schulleitungen für die Qualität und Entwicklung ihrer Schule tragen, Anregungen und Empfehlungen zur Reduktion des Belastungserlebens, z.B. Etablierung von Entlastungsstrukturen, abgeleitet und mit der Schulaufsicht, den Schulerhaltern und den Schulleitungen weiterentwickelt.

Führung im Fokus - Eine Scoping-Review zur Rolle der Schulleitung in der deutschsprachigen Schulentwicklungsforschung

Olivia Wohlfart, Katja Gerstmaier, Ingo Wagner

Karlsruher Institut für Technologie (KIT), Deutschland

Als zentrale Gatekeeper der Bildungssysteme und wesentliche Treiber von Schulentwicklungsprozessen spielen Schulleitungen eine transformative Rolle bei der (Weiter-) Entwicklung von Bildungseinrichtungen (Fullan, 2016). In der Vergangenheit wurde die Forschung im Bereich des „*educational leadership*“ vorwiegend von angloamerikanischen Studien geprägt, die wiederum hauptsächlich die Effektivität von „*school leadership*“ in Bildungseinrichtungen untersuchten (z. B. Hallinger, 2013). Aber anders als im angloamerikanischen Raum übernehmen Schulleitungen in Deutschland historisch gewachsen andersartige Aufgaben, da sie oft überlappende Rollen als Teilzeitlehr- und Führungskräfte innehaben (Klein, 2016). Darüber hinaus mangelt es den vorherrschenden normativen oder normativ-analytischen Studien im Bereich der Schulleitungsforschung in deutscher Sprache häufig solide theoretische Grundlagen (Johannmeyer & Cramer, 2023). Es ist anzunehmen, dass die im Englischsprachigen vorherrschenden managementorientierten Führungstheorien nicht direkt auf den deutschen Diskurs übertragen werden können/sollten.

Diese Forschungslücke adressieren wir mit einem systematischen Scoping-Review nach Moher et al. (2009) zur deutschsprachigen, peer-reviewed Literatur zur Rolle der Schulleitung in Schulentwicklungsprozessen. Dies umfassten die Identifizierung, Auswahl, kritische Bewertung und Datenanalyse von 14 relevanten Artikeln, darunter sieben qualitative, fünf quantitative, eine Mixed-Methods-Studie sowie ein theoretischer Beitrag. Geografisch beleuchten die Studien zehn deutsche Stichproben, eine österreichische Stichprobe, eine Schweizer Stichprobe sowie eine Studie, die deutsche und schweizerische Schulleitungen vergleicht, und eine weitere, die deutsche und US-amerikanische Schulleitungen vergleicht, was somit eine vielfältige und facettenreiche Landschaft berücksichtigt. Wir analysieren derzeit die Daten mittels thematischer Analyse nach Braun und Clarke (2020), um ein tieferes Verständnis der generierten Muster und Themen innerhalb der ausgewählten Literatur zu erhalten.

Erste Ergebnisse heben eine Verschiebung des Forschungsschwerpunkts und die Anerkennung der Implementierung von Managementwerkzeugen in einem Bereich hervor, der historisch aus einer pädagogischen Rolle herausgewachsen ist. Unterschiede zwischen Schultypen, -größen und Variationen in der Autonomie einzelner Schulen unterstreichen die Komplexität und den Umfang von Aktivitäten, für die Schulleitungen in Deutschland verantwortlich sind (z. B. Klein, 2016). Die New Public Management-Bewegung und ihre damit verbundene Orientierung an evidenzbasierter Schulentwicklung (z. B. ICILS, PISA oder standardisierte nationale Tests) haben begonnen, die Forschung zur Schulleitung im deutschsprachigen Raum zu prägen (z.B. Altrichter et al., 2012). Während die Personalentwicklung im angloamerikanischen Diskurs oft klar als Teil der Verantwortlichkeiten von Schulleitungen definiert ist, haben deutsche Schulleitungen diese Aufgabe vergleichsweise erst kürzlich als eigenständigen Bereich ihrer Aufgaben wahrgenommen (z. B. Braun et al., 2018; Voigt & Engel, 2018). Darüber hinaus haben wir festgestellt, dass die wahrgenommene Akzeptanz der Schulleitung innerhalb des Kollegiums die Rolle und Handlungsmöglichkeiten der Schulleitung in Schulentwicklungsprozessen stark beeinflusst (z. B. Klein, 2016; Voigt & Engel, 2018). Angesichts gesellschaftlicher Prozesse wie der digitalen Transformation (Tulowitzki & Gerick, 2020) entwickelt sich die Rolle der Schulleitung weiter, und unsere Ergebnisse betonen die Notwendigkeit einer Professionalisierung dieser Schlüsselposition innerhalb der Schulen. In Anlehnung an Johannmeyer und Cramer (2023) diskutieren wir Fragen zum Zeitgeist der formalen Qualifikation von Schulleitungen in verschiedenen Makrokontexten und wie dies mit Schulentwicklung und Bildungsqualität zusammenhängt. Darüber hinaus reflektieren wir kritisch die Eignung der vorwiegend managementorientierten Forschungs- und Theorieansätze aus dem angloamerikanischen Kontext für die (notwendige) Professionalisierung von Schulleitungen in Deutschland für eine erfolgreiche Schulentwicklung.

ENTFÄLLT: Repräsentative Schulleitungsbefragung zu psychosozialen Infrastrukturen in 13 Bundesländern in Deutschland

Judith Bauch, Julian Schmitz

Universität Leipzig, Deutschland

Theoretischer Hintergrund: Psychische Erkrankungen im Kindes- und Jugendalter zeigen im späteren Verlauf hohe Chronifizierungs- und Persistenzraten^{1,2,3}. Neben erheblichen individuellen sowie gesellschaftlichen Folgekosten können psychische Probleme Bildungserfolge⁴ und somit Bildungswege, Anstellung und Einkommen beeinträchtigen⁵, welche wiederum elterliche Einflussfaktoren für die psychische Gesundheit von Kindern und Jugendlichen darstellen⁶. Bildung spielt im Hinblick auf diese komplexen Wechselwirkungen eine besondere Rolle und bietet in Deutschland aufgrund der Schulpflicht sowie geeigneten psychosozialen Infrastrukturen wie Schulsozialarbeit (SSA) und Schulpsychologie (SP) die Möglichkeit, psychische Probleme vorzubeugen, frühzeitig zu erkennen und geeignete Interventionen einzuleiten. Trotz hoher und über die letzten Jahre gestiegener psychischer Belastungen bei Kindern und Jugendlichen in Deutschland^{7,8}, gibt es jedoch Hinweise auf unzureichende psychosoziale Infrastrukturen an Schulen in Deutschland^{9,10}. Umfassende wissenschaftliche Daten fehlen hierzu jedoch bislang.

Fragestellung: Ziel der Studie ist es, relevante psychosoziale Infrastrukturen an Schulen in Deutschland repräsentativ über einen Zeitraum von drei Jahren zu untersuchen. Folgende Fragestellungen werden vorgestellt:

- Wie viele Schulen verfügen über psychosoziale Infrastrukturen (Angebote der SP und SSA)? Besteht ein Zusammenhang zwischen Verfügbarkeit und Schulform?
- Wie ist die Bedarfsdeckung psychosozialer Infrastrukturen an Schulen?
- Wie entwickeln sich Bedarfe an psychosozialen Infrastrukturen über die Zeit?

Methode: Dazu wird jährlich eine zufällige Stichprobe von Schulleitungen (SL) aus 13 Bundesländern mittels einer anonymen Online-Befragung zu psychosozialen Infrastrukturen an der jeweiligen Schule befragt. Alle Items der Befragung wurden aufgrund mangelnder theoretischer Fundierung mittels intuitiver Konstruktionsstrategie entworfen. Dem zugrunde lagen eine ausführliche Recherche, halbstrukturierte Interviews sowie bilaterale Reviewprozesse mit Expert:innen und eine Pilotierungsphase. Die erste Datenerhebung erfolgte basierend auf einer zufällig gezogenen Schulstichprobe (N=8414) allgemeinbildender Schulen in öffentlicher Trägerschaft (ausgenommen Bayern, Saarland, Sachsen aufgrund von Studienablehnung) nach postalischer Studieneinladung im Zeitraum 30.04.–17.06.2024. Alle Daten wurden nach Bundesland und Schulform gewichtet (Fehlertoleranz +/- 3 Prozent). Die Datenauswertung erfolgte mittels Complete-Case Analysen.

Ergebnisse: Die finale Stichprobe enthielt Daten von 486 Schulen. Von diesen verfügten 76% über Angebote der SP. Ein signifikanter Zusammenhang zwischen Verfügbarkeit und Schulform bestand nicht, $X^2(3, N = 486) = 1.63, p = .65$. 77% der SL mit verfügbaren SP-Angeboten gaben an, dass die verfügbaren Angebote den Bedarf an der Schule nicht oder eher nicht abdecken (vs. 23% Bedarfsdeckung ja oder eher ja). 66% der SL berichteten einen Bedarfsanstieg an SP-Angeboten im Vergleich zum letzten Schuljahr (vs. 2% Bedarfsabnahme an SP-Angeboten).

Über SSA-Stellen verfügten 75% aller Schulen. Es bestand ein signifikanter Zusammenhang zwischen Verfügbarkeit von SSA-Stellen und Schulform, $X^2(3, N = 486) = 36.27, p < .001$. Am häufigsten verfügten Haupt-, Real- und Gesamtschulen über SSA-Stellen (95%), am wenigsten Gymnasien (62%). Schulen mit SSA-Stellen gaben im Schuljahr 2023/24 im Durchschnitt ein vorhandenes Vollzeitäquivalent (VZÄ/500 Schüler:innen) von 1.64 ($SD = 1.72$, $Mdn = 1.28$ [0.86, 1.98]) an, während der Bedarf im Durchschnitt auf 2.89 VZÄ/500 Schüler:innen ($SD = 2.05$, $Mdn = 2.44$ [1.65, 3.51]) geschätzt wurde. 67% der SL mit SSA-Stellen gaben an, dass verfügbare SSA-Angebote den Bedarf an der Schule nicht oder eher nicht abdecken (vs. 33% Bedarfsdeckung ja oder eher ja). 72% der SL berichteten einen Bedarfsanstieg an SSA-Angeboten im Vergleich zum letzten Schuljahr (vs. 2% Bedarfsabnahme an SSA-Angeboten).

Diskussion: Die für 13 Bundesländer repräsentativen Ergebnisse zeigen, dass bislang noch nicht alle Schulen in Deutschland über psychosoziale Infrastrukturen in den Bereichen SP und SSA verfügen, vorhandene Strukturen den Bedarf an Schulen nach Einschätzung des Großteils der Schulleitungen nicht abdecken und Bedarfe im Vergleich zum letzten Schuljahr als gestiegen eingeschätzt werden. Unserer Kenntnis nach ist diese Studie der erste Versuch, psychosoziale Infrastrukturen an Deutschland

repräsentativ abzubilden. Die Ergebnisse sind von bildungspolitischer Relevanz und können dazu dienen, Infrastrukturen an Schulen bedarfsgerecht differenziert auszubauen. Limitationen werden diskutiert.

Internationale Perspektiven auf datengestützte Schulentwicklung an Grundschulen: Befunde eines Scoping Reviews

Nicol Sperling, Jana Schlöpker, Mario Mallwitz, Sonja Nonte

Universität Osnabrück, Deutschland

Aktuelle gesellschaftliche Herausforderungen, wie der stetige Rückgang basaler Kompetenzen bei Grundschüler*innen (Stanat et al., 2023) und der anstehende Rechtsanspruch auf Ganztagsbetreuung an Grundschulen (Artikel 1, § 24a Ganztagsförderungsgesetz, 2021), üben erheblichen Druck auf das Bildungssystem, insbesondere aber auf den Primarbereich aus. Parallel dazu sehen sich schulische Akteur*innen mit tiefgreifenden Veränderungen durch die digitale Transformation konfrontiert, die weite Bereiche des schulischen Alltags, einschließlich zentraler Aspekte der Schul- und Unterrichtsentwicklung betreffen.

Im Kontext eines vom BMBF geförderten Forschungsprojektes werden Grundschulen bei der Bewältigung derartiger Herausforderungen in verschiedenen Teilprojekten durch praxisorientierte Gemeinschaften unterstützt. Das Teilprojekt „Datenbasierte Schulentwicklung“ legt im Sinne der Evidenzbasierung den Fokus auf die Förderung der systematischen Nutzung von Daten, die entweder bereits vorliegen oder gesammelt werden, um sie gezielt für Schulentwicklungszwecke nutzen zu können (Demski, 2017; Sliwka, 2024). Mit dem gesamten schulischen Personal werden bedarfsorientiert Maßnahmen und Angebote an bis zu fünf Einzelschulen entwickelt, implementiert und evaluiert. Langfristig sollen diese Maßnahmen dazu dienen, Schulpersonal zu entlasten, die Arbeit in multiprofessionellen Teams sowie insgesamt schulische Entwicklungsprozesse durch eine effizientere Datennutzung zu optimieren.

Bisher ist in Deutschland wenig über die Nutzung von Daten als Informationsquelle für Schulentwicklungsvorhaben an Grundschulen bekannt. Als empirische Grundlage für das Projektvorhaben wird daher ein Scoping Review erstellt, welches die Nutzungsintensität und die Varianz der genutzten Datenquellen auf Basis der aktuellen Forschungsliteratur mit dem Fokus auf Grundschulen abbilden soll. Eine Recherche zu bereits bestehenden Reviews in diesem Forschungsfeld verdeutlicht, dass diese in Bezug auf außereuropäische Schulsysteme und der Sekundar- und Hochschulbildung bereits vorhanden sind.

Angesichts dessen adressiert dieses Review die folgende zentrale Forschungsfrage: Welche Rolle spielen Daten für die Schulentwicklung an Grundschulen? Daraus leiten sich folgende spezifische Unterfragen ab: a. Welche Daten und Informationen werden an Grundschulen zu welchem Zweck genutzt? b. Welche Auswirkungen hat die Nutzung von Daten auf die Schulentwicklung? c. Wie beeinflussen schulstrukturelle und personelle Rahmenbedingungen die Nutzung von Daten an Grundschulen?

Zur Beantwortung dieser Fragen orientiert sich das Scoping Review an der *PRISMA Extension for Scoping Reviews* (PRISMA-ScR) Checkliste von Tricco et al. (2018). Datenbanken und Suchbegriffe wurden durch die Analyse themenähnlicher Reviews und die Berücksichtigung der Empfehlungen von Gusenbauer und Haddaway (2019) festgelegt. Insgesamt wurden acht Datenbanken (u.a. FIS Bildung, ERIC und ProQuest) unter der Berücksichtigung boolescher Operatoren durchsucht. Die ausgewählten Suchbegriffe basieren auf den Forschungsfragen sowie auf konkreten Ein- und Ausschlusskriterien. Das Review befindet sich zurzeit in der Eignungsphase. Ein von drei Reviewer*innen unabhängig voneinander durchgeführtes Screening von Titeln und Abstracts von 2.658 identifizierten Artikeln ergab eine ungefähre Anzahl von 120 Artikeln. Die Eignung dieser wird nachfolgend durch ein Volltextscreening begutachtet. Die Einschlusskriterien umfassen dabei u.a. (1) empirische Untersuchungen zu datengestützter Schulentwicklung (2) eine Fokussierung auf den Primarbereich (3) sowie (4) einen Veröffentlichungszeitraum von 2015 bis 2024. Studien werden ausgeschlossen, wenn sie den Einschlusskriterien nicht entsprechen und Reviews oder Metaanalysen darstellen. Die Daten werden anhand einer Mehrfachcodierung extrahiert und analysiert.

Die Ergebnisse dieses Reviews sollen auf der Tagung unter verschiedenen Gesichtspunkten und identifizierten Forschungslücken präsentiert und diskutiert werden. Inhaltlich soll ein umfassender Überblick über den aktuellen Forschungsstand gegeben werden, indem veranschaulicht wird, welche Daten und Datenquellen zu welchen Zwecken genutzt werden und inwiefern sich Datennutzung auf Schulentwicklung auswirkt. Mögliche Einflüsse schulstruktureller und personeller Rahmenbedingungen auf die Nutzung von Daten sollen ebenfalls dargestellt werden. Einen weiteren Fokus stellt ein Überblick zu genutzten methodischen Strategien sowie Studienarten und -designs zur Untersuchung von datengestützter Schulentwicklung dar.

Die gewonnenen Erkenntnisse sollen dazu genutzt werden, zukünftige Projektmaßnahmen zu steuern und Impulse für die erfolgreiche Implementierung datengestützter Ansätze für Schulentwicklungsprozesse an Grundschulen zu geben. Die Ergebnisse tragen somit nicht nur zur wissenschaftlichen Erkenntnis bei, sondern bieten auch praktische Implikationen für schulische Akteur*innen sowie deren Ausbildung und bildungspolitische Entscheidungsträger.

PS-55: Führung und Innovation in benachteiligten Schulen: Perspektiven für eine gerechte Bildungsentwicklung

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SO 418

Chair der Sitzung: Rukiye Ates

Bildungsgerechtigkeit, schulische Partizipation und Demokratiebildung: Mixed-Methods-Befunde einer Fallstudienanalyse im Research-Practice-Partnership-Design an einer weiterführenden Schule in herausfordernder sozialer Lage in NRW

Rukiye Ates¹, Laura Beckmann¹, Christine Becks¹, Kathrin Racherbäumer², Isabell van Ackeren-Mindl¹

¹Universität Duisburg Essen, Deutschland; ²Universität Siegen, Deutschland

Theorie

Förderung von Bildungsgerechtigkeit und demokratischer Partizipation ist ein zentrales Ziel im deutschen Bildungssystem (Bellmann, 2019; Oberle, 2022). Angesichts wachsender Ungleichheit gewinnt Demokratiebildung (DB) an Schulen an Bedeutung. Sie umfasst politisches Wissen, Toleranz, Respekt und die Befähigung zur Teilhabe (KMK, 2024; SWZ, 2024). Eine inklusive, demokratische und chancengerechte Schulentwicklung ist besonders an herausfordernden Standorten notwendig (Bremm, Racherbäumer, van Ackeren, 2017). DB befähigt benachteiligte Schüler*innen (SuS), sich aktiv in schulische und gesellschaftliche Prozesse einzubringen und sich als selbstwirksam zu begreifen (Stojanov, 2013). Da Demokratie von Handlungsdispositionen lebt, ist es wichtig zu klären, welches Wissen, welche Überzeugungen und welches Zugehörigkeitsgefühl SuS entwickeln (Abs et al., 2024). Ziel des Beitrags ist es, durch interdisziplinäre Zusammenarbeit eines multiprofessionellen Teams den Wissenstransfer zwischen Forschung und Praxis zu ermöglichen. Eine Research-Practice-Partnership (RPP) fördert die Relevanz und Anwendbarkeit der Forschungsergebnisse.

Die Forschungsfrage wird durch einen längsschnittlichen, partizipativen Ansatz beantwortet, der Schul- und Unterrichtsprozesse an einem herausfordernden Schulstandort (Sozialindex 9) in NRW beobachtet. Im Rahmen einer RPP wird kooperativ an kontextsensiblen Lösungen für Chancengleichheit gearbeitet (Farrell et al., 2021; Coburn & Penuel, 2016).

Inwiefern können qualitative Deutungen aus den Gesprächen mit Lehrkräften (LK) die Ergebnisse der quantitativen Schülerdaten zu Partizipation, offenem Klassenklima, Vertrauen und sozialen Medien sowie die schulischen sozialen Beziehungen widerspiegeln?

Methode

Ein explorativer Mixed-Methods Ansatz wurde genutzt, um Einblicke in die DB und Partizipation an einer Fallschule in NRW zu gewinnen. Die Studie basiert auf zwei Untersuchungen, die im Frühjahr 2022 und 2024 an der Fallschule (Jahrgang-8), mithilfe von RPP und einem multiprofessionellen Team, durchgeführt wurden. Die quantitative Erhebung basiert auf in ICCS 2022 (Abs et al., 2024) durchgeführten Schülerbefragungen zu Einstellungen und Schulerfahrungen. Die Nutzung der ICCS-Daten erfolgte mit Zustimmung der Schule. Die qualitative Erhebung 2024 umfasste Gruppen- und Einzelgespräche mit Lehrkräften (n=77) nach Bevan (2014). Quantitative Konstrukte wurden in die Codierung der qualitativen Analyse mit MAXQDA-2024 einbezogen. Die qualitative Datenanalyse erfolgte deduktiv auf Basis der quantitativen Ergebnisse, wobei die Daten kombiniert wurden, um die Forschungsergebnisse zu erklären.

(Vorläufige) Ergebnisse

Die Ergebnisse zur Offenheit des Unterrichtsklimas liegen in der erhobenen der Fallschule bei etwa 57,5 %; im Landesvergleich sind es ca. 52 Skalenpunkte. 41 % der SuS gaben an, dass SuS oft aktuelle politische Ereignisse zur Diskussion bringen dürfen (31 % in NRW), weitere 41 % geben an, dass dies manchmal passiert (39 % in NRW). 94 % der SuS an der Fallschule berichten, dass Lehrkräfte oft oder manchmal SuS ermutigen, ihre Meinung deutlich zu machen (73 % in NRW). 88 % geben an, dass Lehrkräfte oft oder manchmal verschiedene Sichtweisen auf ein Problem erklären (76 % in NRW). 50 % der SuS berichteten von Diskriminierung (NRW: 27 %). Das Vertrauen in soziale Medien liegt nahe dem Durchschnitt (33 % vs. 32 %), während das Vertrauen in Menschen und den Bundestag niedriger ist. 78 % vertrauen ihren Lehrkräften (NRW: 65 %) und das politische Wissen der SuS an dieser Schule ist höher im Vergleich zu ähnlichen Schulen (Sozialindex 6-9).

Fast alle Lehrkraftgespräche zeigten ein ähnliches Muster: mentale Belastung des Kollegiums, abweichendes Schülerverhalten (Gewalt und Radikalisierung), sinkende Unterrichtsqualität (Lautstärke und fehlendes Personal), große Sprach- und Lernschwierigkeiten sowie fehlende Basiskompetenzen (Lesen und Schreiben) bei den SuS, Herausforderungen in der Elternarbeit und im Schulsozialraum (Sprachbarrieren) und Interessenkonflikte mit der Schulleitung. Diese Ergebnisse stehen den quantitativen SuS-Erhebungen diametral gegenüber, was die Notwendigkeit gezielter Schulentwicklung zur Verbesserung der Bildungs- und Teilhabechancen sowie der SuS-Lehrkraft-Beziehung betont.

Diskussion

Ein integrativ-kooperativer Ansatz durch Research-Practice Partnership (RPP) wird empfohlen, um Herausforderungen wie Diskriminierungserfahrungen und Vertrauen im Schulkontext zu bewältigen. Die wahrgenommenen fehlenden Basiskompetenzen der SuS, die beeinträchtigte Unterrichtsqualität und die herausfordernde Elternarbeit bestätigen dies. Ein multiprofessionelles Team und gezielte Schulentwicklung sind daher entscheidend.

Kollektive Innovationsbereitschaft an Schulen in herausfordernden Lagen – kann Führung den Unterschied machen?

Laura Beckmann¹, E. Dominique Klein²

¹Universität Duisburg-Essen, Deutschland; ²TU Dortmund, Deutschland

Theoretischer Hintergrund

Aktuelle nationale sowie internationale Forschungsarbeiten zu den Bedingungsfaktoren erfolgreicher Schul- und Unterrichtsentwicklungsprozesse an Schulen in herausfordernden Lagen heben die Bedeutsamkeit kollektiver Handlungsfähigkeit auf einzelschulischer Ebene hervor (Klein, 2018; Klein & Bremm, 2019; Zumpe, 2024). Diese ist eng verknüpft mit dem Konzept der kollektiven Innovationsbereitschaft (Buske, 2014; Masry-Herzallah & Da'as, 2021) im Sinne einer geteilten Bereitschaft des Lehrkräftekollegiums zur Übernahme von Verantwortung für Veränderungen, zur kritischen Reflexion der eigenen Praxis und zur Veränderung dieser Praxis sowie zur Suche nach alternativen Wegen, wo dies notwendig ist. Bisherige

Studien zeigen jedoch, dass Schulen in herausfordernden Lagen oftmals durch Hilflosigkeit und Pessimismus innerhalb der Schulgemeinschaft geprägt sind, was tendenziell mit einer geringen kollektiven Innovationsbereitschaft einhergeht.

Eine solche Situation stellt eine besondere Herausforderung für das Führungshandeln von Schulleitenden dar. In ihrem Konzept der *transformationalen Führung* argumentieren Bass und Avolio (1994), dass Führungskräfte durch idealisierte Beeinflussung eine Veränderung der Einstellungen der Mitglieder ihrer Organisation erleichtern können. Führungspraktiken, die die Innovationsbereitschaft von Lehrkräften insbesondere an Schulen in herausfordernden Lagen verändern können, wurden jedoch bisher hauptsächlich in qualitativen Studien untersucht (z. B. Jacobson, 2011; Muijs et al., 2010). Bisher fehlt es damit an quantitativen Studien, die Verbindungen zwischen kollektiver Innovationsbereitschaft und transformationalem Führungshandeln in Schulkontexten mit vielfältigen Widrigkeiten untersuchen.

Weiterhin legen Forschungsarbeiten zu den Prädiktoren von individueller Innovationsbereitschaft von Lehrkräften nahe, dass der Zusammenhang zwischen kollektiver Innovationsbereitschaft und Führung durch weitere organisationale Merkmale (z. B. Entwicklung von Transparenz über spezifische Ziele, kollegiale Unterstützung und Kommunikation) vermittelt werden könnte (Thurlings et al., 2015). Inwieweit diese Ergebnisse auf kollektive Überzeugungen von Lehrkräften bezüglich Veränderungen übertragen werden können, ist jedoch unklar.

Basierend auf vorgestellten empirischen Befunden kann angenommen werden, dass das Führungshandeln an Schulen in herausfordernden Lagen eine entscheidende Rolle dabei spielt, die Vision, Struktur und Kultur der Schule zu gestalten und damit die kollektive Innovationsbereitschaft im Lehrkräftekollegium zu beeinflussen.

Fragestellung

Vor diesem Hintergrund sollen folgende Forschungsfragen beantwortet werden:

1. Wie hängen verschiedene Dimensionen transformationalen Führungshandelns mit der kollektiven Innovationsbereitschaft von Lehrkräften an Schulen in herausfordernden Lagen zusammen?
2. Inwieweit vermitteln weitere organisationale Faktoren, die mit der *Vision* (Konsens über die schulischen Ziele) und der *Kultur* (interpersonelles Vertrauen im Kollegium) einer Schule zusammenhängen, diesen Zusammenhang?

Methode

Zur Beantwortung der Forschungsfragen verwenden wir Daten aus zwei standardisierten Lehrkräftebefragungen ($N = 912$ und $N = 391$), die zu Beginn und Ende des Jahres 2020 an insgesamt über 50 weiterführenden Schulen in benachteiligten Lagen in Deutschland durchgeführt wurden. Diese Schulen sind Teil eines landesweiten Schulversuchs mit dem Ziel, modellhaft Wege zu ‚gerechten Bildungschancen‘ zu erproben. Zur Analyse der Zusammenhänge wurden Mehrebenenanalysen mit Mediationseffekten mithilfe der Software Mplus 8 (Muthén & Muthén, 2017) durchgeführt. Durch die Aufnahme von zwei Mediatorvariablen mit Bezug zur organisationalen Ausgestaltung der jeweiligen Schule wurde ermittelt, inwieweit der postulierte Zusammenhang zwischen Führungshandeln und kollektiver Innovationsbereitschaft durch a) den wahrgenommenen Konsens innerhalb des Kollegiums über schulische Ziele und b) das wahrgenommene interpersonelle Vertrauen im Kollegium erklärt werden kann.

Ergebnisse und ihre Bedeutung

Zielorientierte Führung stellte sich als der stärkste Prädiktor für die wahrgenommene kollektive Innovationsbereitschaft heraus, vermittelt über den selbstberichteten Konsens der Lehrkräfte über die schulischen Ziele. Weiterhin war Führung, die die professionelle Entwicklung der Lehrkräfte in den Fokus stellt, positiv und statistisch signifikant mit der kollektiven Innovationsbereitschaft assoziiert. Zudem stärkte Führungshandeln, das ein sicheres Lern- und Arbeitsumfeld kultiviert, das Vertrauen unter den Lehrkräften, was wiederum positiv auf die Innovationsbereitschaft wirkte. Zusammenfassend betonen unsere Ergebnisse die Bedeutung zielorientierter Führungspraktiken sowie Arbeitsstrukturen und -prozesse, insbesondere Strukturen zur Unterstützung der professionellen Weiterbildung von Lehrkräften und zur Ermöglichung eines systematischen, kollaborativen und evidenzbasierten Handelns.

Mehr Ressourcen für Schulen in deprivierter Lage?! Ausgestaltungen und Intentionen einer bedarfsorientierten Ressourcensteuerung auf kommunaler Ebene

Denise Demski, Sarah Eiden, Gabriele Bellenberg

Ruhr-Universität Bochum, Deutschland

Theoretischer Hintergrund

Herkunftsbedingte Disparitäten im Bildungssystem sind für unterschiedliche Bildungsetappen und im Zeitverlauf international und national vielfach empirisch belegt, wobei für ihre Erklärung neben primären und sekundären Herkunftseffekten (Boudon 1974) auch dem Sozialraum und sozialen Segregationstendenzen Relevanz zukommt (z.B. Baumert et al. 2006; Helbig 2010; Makles et al. 2019; Jeworutzki et al. 2017). Eine bedarfsorientierte Ressourcensteuerung, im Zuge derer Bildungseinrichtungen an herausfordernden Standorten und mit einem hohen Anteil von Kindern und Jugendlichen aus benachteiligten Familien zusätzliche Ressourcen erhalten, wird als (potenzielles) Mittel zum Abbau von Bildungsbarrieren angesehen und seit einigen Jahren auch auf kommunaler Ebene verstärkt eingesetzt. Im Mehrebenensystem kommt Kommunen dabei eine entscheidende Rolle zu, da sie sowohl über eigene Ressourcen verfügen, daneben z.T. aber auch für die konkrete Verteilung von Bundes- oder Landesmitteln verantwortlich zeichnen. Ansätzen der Educational Governance (z.B. Altrichter et al. 2007) folgend ist dabei von Rekontextualisierungen (Fend 2008) und eigenlogischen Adaptionen auszugehen und Verteilungsentscheidungen können als Ergebnis sozialer Konstruktionsprozesse (Berger & Luckmann 2004) aufgefasst werden. In der Folge ist eine beträchtliche Variation hinsichtlich Umfang und Art der zugewiesenen Ressourcen (Stellen, Finanzmittel etc.) sowie der Verteilungsgrundsätze (z.B. Zuweisung aufgrund von Sozialindex oder sonstigen Daten, Anträgen oder Expertise von Zuweisenden) zwischen Kommunen zu erwarten.

Internationale Befunde deuten darauf hin, dass der Ressourcenumfang sowie die Art und Weise, wie diese zugewiesen und genutzt werden (sollen), die Wirkungen einer bedarfsorientierten Ressourcensteuerung stark beeinflussen (können) (im Überblick Franck & Nicaise 2022). Für Deutschland liegt jedoch noch kaum Evidenz in Bezug auf Ausgestaltungen, Intentionen und Wirkungen einer bedarfsorientierten Ressourcensteuerung vor.

Fragestellungen

Vor dem Hintergrund der skizzierten Desiderata werden im Beitrag folgende Fragestellungen adressiert:

Welche zusätzlichen Ressourcen werden an Schulen in deprivierter Lage zugewiesen und welche Entscheidungskriterien werden hierfür zugrunde gelegt?

Welche Handlungslogiken der beteiligten Akteur*innen lassen sich identifizieren und inwiefern beziehen sie Fragen von Wirkungen einer Bedarfsorientierung in ihre Überlegungen ein?

Wo zeigen sich im Mehrebenensystem Rekontextualisierungen sowie vertikal und/oder horizontal Wechselwirkungen in der Ausgestaltung?

Methodisches Vorgehen

In einem laufenden BMBF-Projekt wurden auf Basis einer systematischen Analyse öffentlich zugänglicher Dokumente Bundesländer und Kommunen identifiziert, die zusätzliche Ressourcen an Schulen in deprivierter Lage zuteilen. Darauf aufbauend wurden kontrastiv (hinsichtlich Regionalität, Art der Ressourcen und Einführungszeitpunkt) vier Großstädte mit Erfahrungen in Bezug auf eine bedarfsorientierte Ressourcensteuerung ausgewählt. In jeder Kommune wurden unter Berücksichtigung verschiedener Bildungsetappen und organisationaler Zuständigkeiten (Referate bzw. Fachbereiche, z.B. Schulamt, Jugendamt) durchschnittlich acht Interviews mit erzählgenerierenden sowie leitfadengestützten Passagen mit Akteur*innen durchgeführt, die direkt oder indirekt in Entscheidungsprozesse im Zuge einer bedarfsorientierten Ressourcensteuerung involviert sind. Diese wurden transkribiert und inhaltsanalytisch (Mayring 2007) ausgewertet; zur Rekonstruktion von Handlungslogiken wurden dichte Beschreibungen darüber hinaus mit der Dokumentarischen Methode (Nohl 2017; Bohnsack 2014) ausgewertet.

Ergebnisse und Implikationen

Die Ergebnisse verweisen auf differente Intentionen, Rahmenbedingungen und Entstehungsgeschichten einer bedarfsorientierten Ressourcensteuerung in den betrachteten Fallkommunen, die in verschiedenen Ausgestaltungen evident werden. Die Intensität der Datennutzung innerhalb des eigenen Fachbereichs, aber auch in der Stadt insgesamt erscheint unterschiedlich ausgeprägt. Es lassen sich sowohl Verteilungspraktiken identifizieren, die vorwiegend datenbasiert sind, als auch solche, die stärker die Expertise der lokalen Entscheidungsträger*innen berücksichtigen. Wirkungsfragen werden vor allem in den Kommunen konkreter bearbeitet, die über langjährige Erfahrungen verfügen und eine gesamtstädtische Strategie im Hinblick auf eine bedarfsorientierte Ressourcensteuerung verfolgen.

Zwischen den Bundesländern zeigen sich Unterschiede bezüglich der Mittel, die bedarfsorientiert zugewiesen werden. Ebenso werden Variationen in der Beteiligung der Kommunen im Verteilungsprozess von Landesmitteln sichtbar, was von den Befragungspersonen in den Fallkommunen unterschiedlich stark problematisiert wird.

Kommunen stellen in Bezug auf die Förderung von Bildungschancen für benachteiligte Schüler*innen eine relevante, aber in der Forschung bislang noch wenig berücksichtigte Steuerungsebene dar, und Fragen der Wirkungen wurden bislang kaum betrachtet, sodass sich weitgehende Forschungsdesiderata zeigen.

Qualität und Wirksamkeit multipler Interventionen für Schulleitungen an Schulen in herausfordernder Lage

Stephan Gerhard Huber^{1,2}, Jane Pruitt^{1,3}, Christoph Helm², Rolf Strietholt⁴

¹Arbeitsgruppe Bildungsmanagement, Schweiz; ²Johannes Kepler Universität Linz, Österreich; ³Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN); ⁴TU Dortmund

Die zentrale Rolle der Schulleitung für die Schulqualität und -entwicklung ist durch zahlreiche Studien belegt (z. B. Grissom et al., 2021; Kovačević & Hallinger, 2019; Leithwood & Louis, 2012; Rhodes & Brundrett, 2009; Sammons et al., 1995; Teddlie & Stringfield, 1993). Über die Notwendigkeit und Bedeutung entsprechender Professionalisierungs- und Qualifizierungsangebote besteht weitgehend Einigkeit (z. B. Darling-Hammond et al., 2007, 2022; Grissom & Harrington, 2010; Gümüş et al., 2024; Young et al., 2009). In diesem Sinne ist die Führungskräfteentwicklung eng mit der Verbesserung der Schulqualität verbunden. Besonders an Schulen in herausfordernden Kontexten kommt der Schulleitung eine Schlüsselrolle zu, wenngleich die Schulleitung und ihre Schule häufig weitere externe Unterstützung benötigen (z. B. Baumert et al., 2006; Creemers et al., 2022).

Bislang gibt es jedoch wenig Evidenz zur Wirksamkeit umfassender, mehrjähriger Unterstützungsprogramme für Schulleitungen solcher Schulen in Deutschland. Diese Studie untersucht ein innovatives Programm, das multiple Interventionen (Fortbildungsprogramm, Coaching, Schulentwicklungsberatung) zur Professionalisierung und Unterstützung von Schulleitungen an Schulen in herausfordernder Lage kombiniert.

Die übergeordneten Fragestellungen lauten:

1. Wie bewerten die Teilnehmenden die Qualität und den Nutzen der verschiedenen Interventionen?
2. Welche Auswirkungen haben die Interventionen auf die Schulleitung, Schulentwicklung und Schulqualität?

In einer fünfjährigen quasi-experimentellen Längsschnittstudie mit Mixed-Methods-Design wurden 75 Schulen in herausfordernden Kontexten, die am Programm teilnahmen, und 47 in den Kontextmerkmalen analoge Vergleichsschulen untersucht. Jährliche Befragungen von Schulleitungen und pädagogischem Personal erfassten die Qualität der Interventionen sowie Indikatoren der Schulqualität (z. B. Arbeitszufriedenheit, Kooperation) zu drei Messzeitpunkten. Neben einer deskriptiven Auswertung der Qualitätseinschätzungen wird mittels autoregressiven Regressionsanalysen auf Schulebene der Einfluss der wahrgenommenen Qualität der Interventionen auf die Schulqualitätsentwicklung untersucht. Hierzu ziehen wir verschiedene Indikatoren für die Ergebnisqualität der drei Interventionen heran: Zielerreichung, bewerteter Nutzen, Motivationssteigerung, Kompetenzentwicklung (in Bezug auf persönliche und schulische Ziele), Verhaltensänderung sowie Organisationsveränderung. Aufgrund der reduzierten Stichprobengröße auf Schulebene wurden die Regressionsanalysen in separaten Modellen für jede erklärende Variable geschätzt. Dabei wurde jeweils ein Merkmal der Schulqualität zum dritten Messzeitpunkt (abhängige Variable) unter Kontrolle der gleichen Variable zum ersten Messzeitpunkt (Kontrollvariable t1) sowie ein Merkmal der Ergebnisqualität (Effektvariable) einer Intervention regressiert. Dies zeigt, ob die Ergebnisqualität mit dem Merkmal der Schulqualität nach der Programmlaufzeit zusammenhängt, unabhängig vom Ausgangswert des Schulqualitätsmerkmals vor Erhalt der Interventionen. Die Kontrollvariablen bzw. abhängigen Variablen basieren auf den Schulmittelwerten der Einschätzungen der Schulqualität durch das Personal zum ersten bzw. dritten Messzeitpunkt. Die Effektvariablen basieren auf den Mittelwerten der Ergebnisqualitäten aus den Erhebungen zu allen Messzeitpunkten. Zur Erfassung der Veränderungen in den Schulen und um eine Gegenüberstellung der Programm- und Vergleichsschulen anhand der Fragebogendaten zu ermöglichen, wird die Effektstärke Cohen's d (Cohen, 1988) berechnet.

Präsentiert werden ausgewählte Ergebnisse zur Wahrnehmung der Qualität und des Nutzens der verschiedenen Interventionen sowie deren Auswirkungen auf die Schulleitung, die Schulentwicklung und die Schulqualität.

Die Qualität und der Nutzen der Interventionen wurden von den Teilnehmenden im Programmverlauf hoch eingeschätzt. Die Regressionsanalysen zeigen, dass positiv wahrgenommene Ergebnisqualitäten der Interventionen mit Verbesserungen in zahlreichen Dimensionen der Schulqualität einhergehen. So ist beispielsweise die von den Schulleitungen wahrgenommene Kompetenzentwicklung durch das Fortbildungsprogramm mit einer von den Mitarbeitenden wahrgenommenen verbesserten Handlungskoordination der Schulleitung verbunden ($\beta = .26^{**}$). Die Analyse der Effektstärken liefert Hinweise darauf, dass sich die Projektschulen in den Schulqualitätsdimensionen über die Programmlaufzeit hinweg mehrheitlich besser entwickelten als die Vergleichsschulen.

Die Ergebnisse liefern empirische Evidenz für die Wirksamkeit des multiplen Unterstützungsansatzes zur Professionalisierung von Schulleitungen und zur Schulentwicklung an belasteten Schulen. Sie unterstreichen die Bedeutung eines kohärenten, langfristigen und auf den schulischen Kontext abgestimmten Unterstützungsansatzes. Auf Basis der Ergebnisse werden die Interventionen hinsichtlich ihrer Wirkungen diskutiert. Auf diese Weise können Bedingungen für eine erfolgreiche Implementierung identifiziert und Implikationen für die Praxis zur Unterstützung von Schulen in herausfordernder Lage abgeleitet werden.

PS-56: Professionelle Perspektiven auf Inklusion

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 169

Chair der Sitzung: Veronika Dumbacher

Der Umgang mit Schülerinnen und Schüler als ein Schlüssel inklusiver Bildung? – Ein Vergleich unterschiedlicher Professionen

Josephine Laukner

Technische Universität Braunschweig, IfE, Deutschland

Schulen haben in Deutschland einen klaren Bildungsauftrag, der in den einzelnen Schulgesetzen der jeweiligen Bundesländer verankert ist. Demnach soll Schülerinnen und Schüler erfolgreiches Lernen ermöglicht und deren Lernfreude sowie Lern- und Leistungsbereitschaft angeregt werden (vgl. z. B. Niedersächsisches Kultusministerium, 2023). Neben diesen Fähigkeiten und der Vermittlung von Wissen, sollen Schülerinnen und Schüler zudem in ihrer Persönlichkeit gefördert werden (ebd.). Der Aspekt des sozialen Lernens, im Sinne eines Demokratieverständnisses, Grundrechte, Solidarität und Toleranz ebenso wie der Umgang mit Konflikten und die Übernahme von Verantwortung ist u. a. in §1 diverser Schulgesetze verankert (z. B. Landesrecht Berlin, 2023). Kommen Schulen diesem Bildungsauftrag nach, kann Bildung als Schlüssel für gesellschaftliche Herausforderungen verstanden werden.

Die gesellschaftlichen Herausforderungen für Schulen wurden spätestens mit der Ratifizierung der UN-Behindertenrechtskonvention (2006) nochmals deutlich. Seitdem sind Lehrer*innen gefordert, alle Schülerinnen und Schüler mit ihren je unterschiedlichen Lernvoraussetzungen und -bedürfnissen sowie entsprechend vielfältigen Heterogenitätsdimensionen (Hinz 2009) in ihrer Individualität eine chancengerechte Bildung zu ermöglichen. Auf Grundlage politischer und administrativer Voraussetzungen sind es insbesondere die Lehrer*innen, die die Forderungen nach inklusiver Bildung im Schulalltag und im Unterricht umsetzen müssen. Neben fachlichen Inhalten rücken dabei insbesondere soziale Kompetenzen und sowie soziales Lernen in den Fokus. Die Haltung von Lehrer*innen und deren Einstellungen im Umgang mit Vielfalt im inklusiven Unterricht kann dabei als ein entscheidendes Kriterium beschrieben werden (vgl. Booth & Ainscow, 2019; Arndt & Werning, 2018; Greve & Hauenschild, 2017; Laukner, 2024). Diese Haltung spiegelt sich insbesondere in den pädagogischen Beziehungen zwischen Lehrer*innen und Schüler*innen wieder (vgl. Kowalski, 2022). Prengel (2022) benennt pädagogischen Beziehungen als wesentliches Merkmal inklusiver Bildung, wodurch jedem Kind das Gefühl geben werden sollte, willkommen zu sein. Die Qualität pädagogischer Beziehungen ist angesichts empirischer Ergebnisse als wesentlich und gleichermaßen herausfordern zu beschreiben (Carle, 2017). Lehrer*innen können demnach auch als Vorbilder (Mendl 2014) dienen und eine Art Modell für angemessenes Sozialverhalten, emotionales Verhalten und ebenso planvolles Handeln und Problemlösen darstellen (vgl. Edelmann & Wittmann 2019).

Im inklusiven Unterricht verkörpern unterschiedliche Akteure diese Vorbildfunktion, mindestens jedoch zwei Professionen. Die Grundschulpädagogik wird um die Expertise der Sonderpädagogik ergänzt oder erweitert. Um möglichst alle Schüler*innen entsprechend ihrer Bedürfnisse fördern und fordern zu können, bedarf es keine konkurrierende, sondern gewinnbringende Kooperation von Grundschullehrer*innen und Sonderpädagog*innen, die allenfalls von Synergien geprägt ist. Ihre Expertisen, Kompetenzen und Sichtweisen sollten sich um die der jeweils anderen Profession ergänzen. Der vorliegende Beitrag möchte genau diese Faktoren in den Blick nehmen und der Frage nachgehen, welche Unterschiede und Gemeinsamkeiten der jeweiligen Professionen es auf inklusive Bildung gibt. Welche Sichtweisen sowohl auf die Schülerinnen und Schüler als auch auf das Lernen in heterogenen Lerngruppen lassen sich identifizieren?

Dazu dienen 57 problemzentrierten Interviews mit Grund- und Förderschullehrer*innen, ausgewertet mit der Qualitativen Inhaltsanalyse (vgl. Kuckartz & Radiker, 2022), aus dem interdisziplinären Forschungsprojekt „Inklusion – Denken und Gestalten“ als empirische Grundlage für den vorliegenden Vortrag. Ein Vergleich beider Professionen, der Grundschul- und Sonderpädagogik, weist Unterschiede aber auch Gemeinsamkeiten auf. So lassen sich u. a. verschiedene Sichtweisen auf Schüler*innen skizzieren. Ebenso erscheint der Umgang mit Konflikten different. Aber auch vielfältige Schwerpunkte im Unterricht lassen Unterschiede im Umgang mit Schülerinnen und Schülern erahnen.

Wenn die Art des Lehrens und Lernen und im weitesten Sinne der Umgang mit Schüler*innen, vielleicht auch der Umgang mit Menschen innerhalb der Schule, Einfluss auf Bildung hat, kann ein entsprechend motivierender, respektvoller und wertschätzender Umgang, wie er auch in inklusiver Bildung gefordert wird, ein Schlüssel für gesellschaftliche Herausforderungen sein.

Lehrer-Schüler-Beziehungen im inklusiven Unterricht – eine Fallvergleichsstudie

Lena Ohnesorge¹, Josephine Laukner²

¹Universität Hildesheim, Deutschland; ²Technische Universität Braunschweig, Deutschland

Die gesellschaftliche Herausforderung eines anerkennenden Umgangs mit Heterogenität und der Einlösung von Chancengerechtigkeit entfaltet insbesondere in der Schule angesichts u. a. sozialer, migrations- und geschlechtsbezogener Disparitäten im deutschen Bildungssystem (vgl. Lewalter et al. 2023) eine besondere Relevanz. Insofern wird an eine inklusive Schule die Zielsetzung geknüpft, Diskriminierung und Marginalisierung jedweder Heterogenitätsdimension zu vermeiden (vgl. Hinz 2015) sowie soziale Kompetenzen und wertschätzende Einstellungen gegenüber Vielfalt aufzubauen (vgl. Greve & Hauenschild 2017; KMK 2011). Zudem stellt der Umgang mit Heterogenität selbst einen Lerngegenstand dar: Die Auseinandersetzung mit gesellschaftlich produzierter Ungleichheit als epochaltypischem Schlüsselproblem in Gegenwart und Zukunft trägt zur Allgemeinbildung schon in der Grundschule bei (vgl. Klafki 2005). Aus der Einführung der inklusiven Schule in Folge der UN-Behindertenrechtskonvention (vgl. UN 2006) resultiert ein weitreichender Veränderungsprozess auf allen Ebenen des Schulsystems. Der vorliegende Beitrag nimmt vor allem die Ebene des Unterrichts (vgl. Fend 1998) in den Blick.

Lehrer*innen kommt bei der Gestaltung eines inklusiven Unterrichts eine tragende Rolle zu (vgl. Arndt & Werning 2018; Biederbeck & Rothland 2017) – sei es in Bezug auf Unterrichtsplanung und -durchführung oder die Förderung der sozialen Eingebundenheit aller Lernenden sowie einer positiven Beziehungsgestaltung (vgl. Arndt & Werning 2018, Reich 2014). In empirischer Forschung zu schulischer Inklusion überwiegt daher die Perspektive von Lehrkräften (vgl. de Boer, Pijl & Minnaert 2010; Laukner 2024; Ruberg & Porsch 2017). Doch eine Schule, die menschenrechtliche Forderungen nach Partizipation von Kindern (vgl. UN 1989) ernst nimmt und dem Anspruch folgt, die Lernenden und ihre Bedürfnisse, Interessen und Wahrnehmungen als Ausgangspunkt zu nehmen (vgl. Kahlert 2016), muss Kinder ebenso zu Wort kommen lassen. Schließlich ist das Kind der neuen Kindheitsforschung folgend seiner Lebenswelt nicht einfach ausgeliefert, sondern aktiv an der Reproduktion sozialer Wirklichkeit – und damit auch der inklusiven Schule – beteiligt (vgl. Honig 2009; Kluge 2021). Lehrkräfte

und Lernende nehmen jedoch unterschiedliche Rollen im Unterricht ein. Bisher fehlt es an Studien, die die Perspektiven beider Gruppen als zentralen Akteur*innen inklusiven Unterrichts im Wechsel- und Zusammenspiel aufeinander beziehen.

Das interdisziplinäre Forschungsprojekt „Inklusion – Denken und Gestalten“ basierend auf einer Kombination qualitativer und quantitativer Zugänge nimmt sich diesem Forschungsdesiderat an, indem Wahrnehmungen und Deutungsmuster von Grund- und Förderschullehrer*innen sowie Schüler*innen mit und ohne sonderpädagogischem Förderbedarf derselben Schulen empirisch untersucht und verglichen werden. Dabei wird der Frage nachgegangen, wie Lehrkräfte auf der einen und Kinder auf der anderen Seite das Lernen im inklusiven Unterricht in Bezug auf inklusive Unterrichtsmerkmale, Kooperation sowie soziale Interaktionen und Beziehungen wahrnehmen und inwiefern sich trotz desselben Schul- und Klassenkontextes Unterschiede, aber auch Gemeinsamkeiten ergeben.

Eine Fallstudie mit problemzentrierten Interviews (vgl. Witzel 2000), ausgewertet mit der inhaltlich strukturierenden qualitativen Inhaltsanalyse (vgl. Kuckartz & Rädiker 2022), ermöglicht es, Erfahrungen und Wahrnehmungen der beiden Befragten Gruppen in ihrer Komplexität und Ganzheitlichkeit zu analysieren und unterrichtsbezogene Phänomene unter Berücksichtigung des realen Kontextes an einzelnen ausgewählten Fällen miteinander in Beziehung zu setzen (vgl. Yin 2018).

Die Ergebnisse lassen auf Chancen einer inklusiven Schule für Toleranz schließen, bspw. in Bezug auf die Rolle von Erfahrungen gemeinsamen Unterrichts oder die Aufmerksamkeit von Grundschulkindern für die individuellen Bedürfnisse in heterogenen Klassen. Allerdings zeigen sich gleichwohl Herausforderungen durch Stigmatisierungsrisiken, (Re-)Produktion von Differenz im Sinne des doing difference (vgl. West & Fenstermaker 1995) sowie das Spannungsfeld der (De-)Thematisierung von Differenz im Unterricht (vgl. Katzenbach 2015). Der Vergleich der unterschiedlichen Perspektiven lässt Rückschlüsse auf die Bedeutung von Einstellungen und pädagogisch-didaktischen Orientierungen der Lehrenden zu (vgl. Laukner 2024). Dabei lassen sich Bedarfe der Lehrkräftebildung für einen heterogenitätssensiblen Umgang mit Lernenden mit dem Ziel der Chancengerechtigkeit (vgl. Budde 2017) durch eine Orientierung an der Individualität der Schüler*innen (vgl. Arndt & Werning 2018) sowie anerkennenden Beziehungen (vgl. Kowalski 2022; Prengel 2019) als Merkmalen inklusiver Schulen ableiten.

Überzeugungen frühpädagogischer Fachkräfte zu Schulfähigkeit bei Kindern mit Beeinträchtigungen

Veronika Dumbacher, Daniel Then, Sanna Pohlmann-Rother

Lehrstuhl für Grundschulpädagogik und Grundschuldidaktik Universität Würzburg, Deutschland

Schulfähigkeit gilt als zentrales Konzept der Übergangsforschung (Kammermeyer, 2014). Während in der Vergangenheit ein individuumszentriertes Verständnis vorherrschte, werden gegenwärtig – nach einem ökosystemischen Verständnis – zunehmend Umweltmerkmale als Schulfähigkeitskriterien diskutiert (Dockett & Perry, 2009). Dieser veränderte Blick auf Schulfähigkeit ist für die inklusive Gestaltung des Schuleintritts von besonderer Bedeutung (Pohlmann-Rother & Then, 2023). Vor allem wenn ein enges Inklusionsverständnis zugrunde gelegt (Lindmeier & Lütje-Klose, 2015) und Kinder mit Beeinträchtigungen fokussiert werden, wird dies evident. So zeigen Kinder mit Beeinträchtigungen im Übergang erhöhten Unterstützungsbedarf (Jiang et al., 2021), der eine Anpassung der Rahmenbedingungen verlangt. Inwieweit den Bedarfen eines Kindes im Übergang begegnet wird, hängt dabei entscheidend von den Überzeugungen der beteiligten AkteurInnen zu Schulfähigkeit ab (Puccioni, 2018). Als bedeutende AkteurInnen im inklusiven Übergang (Then & Pohlmann-Rother, 2023) spielen die pädagogischen Fachkräfte eine wichtige Rolle. Forschungsarbeiten, die die Überzeugungen pädagogischer Fachkräfte zur Schulfähigkeit bei Kindern mit Beeinträchtigungen fokussieren, sind bislang dennoch selten. Die wenigen vorliegende Befunde legen nahe, dass neben sozial-emotionalen Kompetenzen (Takriti et al., 2019) auch Umweltmerkmalen (z.B. familiäre Unterstützung) Bedeutung zugeschrieben wird (Chadwick & Kemp, 2002). Aktuelle Studien, welche die Überzeugungen frühpädagogischer Fachkräfte zu Schulfähigkeit von Kindern mit Beeinträchtigungen im deutschsprachigen Raum in den Blick nehmen, stehen hingegen aus. An dieser Stelle setzt die vorliegende Studie an. Im Zentrum steht die Frage, welche Kriterien pädagogische Fachkräfte berücksichtigen, um für Kinder mit Beeinträchtigungen den Besuch einer allgemeinen Grundschule in Betracht zu ziehen, und in welche Beziehung die Fachkräfte diese Kriterien setzen.

Die Daten wurden im Rahmen qualitativer, leitfadengestützter Interviews mit n=22 frühpädagogischen Fachkräften erhoben. Die Datenbasis bilden Strukturlegebilder (Scheele & Groeben, 1988), die im Rahmen der Interviews entstanden und die Schulfähigkeitskriterien sowie die Beziehungen widerspiegeln, in die die Fachkräfte die Kriterien setzen. Die Datenauswertung erfolgte mittels Verfahren der inhaltlich-strukturierenden qualitativen Inhaltsanalyse (Kuckartz & Rädiker, 2022) durch zwei unabhängige CodiererInnen (*Cohen's Kappa*=.84). Anschließend wurde ein holistischer Vergleich der Strukturbilder angestrengt. Dabei wurden die Strukturbilder daraufhin eingeschätzt, 1) inwieweit die Fachkräfte die Kriterien in dem jeweiligen Strukturbild in eine Hierarchie (vom ‚wichtigsten‘ zum ‚unwichtigsten‘) legen und 2) inwieweit die Kriterien in den Strukturbildern nach spezifischen Gesichtspunkten (z.B. Inhaltsbereichen) angeordnet werden. Die Einschätzung der Strukturbilder wurde von zwei unabhängigen CodiererInnen durchgeführt und konsensuell validiert.

Im Ergebnis zeigt sich, dass die Fachkräfte v.a. sozial-emotionale (z.B. Sozialverhalten) und kognitive Kompetenzen (z.B. Konzentrationsfähigkeit) für den Grundschulbesuch eines Kindes mit Beeinträchtigungen als relevant ansehen. Unterschiede zwischen einzelnen Beeinträchtigungsarten (z.B. körperlich-motorischen und geistigen Beeinträchtigungen) werden dabei nur vereinzelt genannt. Darüber hinaus wird verbreitet strukturellen Aspekten der schulischen Umwelt (z.B. Klassengröße) Bedeutung zugeschrieben. Merkmale der Unterrichtsgestaltung (z.B. Individualisierung) stehen seltener im Fokus, obwohl die Unterrichtsgestaltung für gelingende schulische Inklusion zentral ist (Heimlich & Bjarsch, 2020). Der holistische Vergleich der Strukturbilder zeigt, dass nur wenige Fachkräfte die Kriterien linear vom ‚wichtigsten‘ zum ‚unwichtigsten‘ Kriterium anordnen. Der Großteil der Fachkräfte wählt stattdessen alternative Formen, z.B. thematische Cluster oder Kreisformen, um die Beziehung der Kriterien zu illustrieren. Die Befunde deuten somit auf eine komplexe Struktur der Überzeugungen zur Schulfähigkeit bei Kindern mit Beeinträchtigungen hin. Welche Implikationen sich hieraus für Bildungspraxis ergeben, wird im Vortrag diskutiert.

Überzeugungen frühpädagogischer Fachkräfte zum Konzept Beeinträchtigung

Agneta Floth, Daniel Then

Universität Würzburg, Deutschland

Die Bildung und Betreuung von Kindern mit Beeinträchtigungen in allgemeinen Bildungssettings ist ein bedeutender Teil des Alltags in Kindertageseinrichtungen (Kißgen et al., 2021). Wie frühpädagogische Fachkräfte der Inklusion dieser Kinder gegenüberstehen und Inklusionsprozesse gestalten, hängt bedeutend von ihren Überzeugungen zu Beeinträchtigung ab (Tamakloe, 2018). Dabei lassen sich theoretisch zwei Grundmodelle von Beeinträchtigung unterscheiden, denen die Überzeugungen der Fachkräfte zugeordnet werden können: Das *individuelle bzw. medizinische Modell* versteht Beeinträchtigung als personale Eigenschaft, die einem Individuum anhaftet und als biologischer Defekt charakterisiert ist (Goodley, 2017). Im

sozialen Modell wird Beeinträchtigung als Folge sozialer Stigmatisierungs- und Exklusionsprozesse beschrieben, indem Personen, die den gesellschaftlichen Normvorstellungen nicht entsprechen, von der gesellschaftlichen Teilhabe ausgeschlossen werden (Reindal, 2008; Dederich, 2009). Forschungsbefunde zeigen, dass sich die Überzeugungen von pädagogischen Fachkräften zu Beeinträchtigungen häufig am medizinischen Modell orientieren (Moriña & Carnerero, 2022). Sozial akzentuierte Verständnisweisen existieren (Sandberg et al., 2010; Thornton & Underwood, 2013), sind im Vergleich aber seltener. Es stellt sich die Frage, wie die Überzeugungen pädagogischer Fachkräfte zu Beeinträchtigungen in Kindertageseinrichtungen in Deutschland ausgeprägt sind. Dem wird in der vorliegenden Studie nachgegangen. Von Interesse sind a) die explizierten Anteile der Überzeugungen (Reusser & Pauli, 2014), welche die Fachkräfte auf unmittelbare Nachfrage verbalisieren, b) die Überzeugungsanteile, die zum Ausdruck kommen, wenn die Fachkräfte Erfahrungen mit Kindern mit Beeinträchtigungen schildern und nicht explizit nach ihrem Bild von Beeinträchtigungen gefragt werden, und c) der Grad, in dem die Aussagen der Fachkräfte diesbezüglich konsistent sind.

Die Studie basiert auf qualitativen, leitfadengestützten Interviews (Reinders, 2022) mit n=22 pädagogischen Fachkräften in Kindertageseinrichtungen. Der Fokus liegt auf Interviewfragen, in denen die Fachkräfte explizit nach ihrem Verständnis von Beeinträchtigung gefragt sowie Erfahrungen mit Kindern mit Beeinträchtigungen zu schildern gebeten wurden. Die Auswertung erfolgt deduktiv-induktiv mittels Verfahren der inhaltlich strukturierenden qualitativen Inhaltsanalyse (Kuckartz & Rädiker, 2022). Zur Sicherung der Güte der Auswertung wurden die Ergebnisse von zwei Codierer:innen konsensuell validiert.

Im Ergebnis konnten zwei Sichtweisen auf Beeinträchtigung – personenbezogen, bei der ein individuelles Verständnis dominiert; und umweltbezogen, bei der ein soziales Verständnis leitend ist – identifiziert werden. Personenbezogene Überzeugungen äußerten sich in einem *defizitorientierten* Verständnis (bei dem der Fokus auf zugeschriebenen Diagnosen oder Schwächen der Kinder liegt), einer *begriffsnegierenden* Verständnisweise (bei der der Begriff Beeinträchtigung als solcher infrage gestellt wird) und/oder einem Verständnis, bei dem – in wertender Weise – zwischen *normalen und nicht-normalen* Kindern unterschieden wird. Umweltbezogene Überzeugungen integrierten *strukturelle oder gruppenbezogene Aspekte* und spiegelten ein soziales Verständnis von Beeinträchtigung wider.

Werden die Fachkräfte explizit nach ihrem Verständnis von Beeinträchtigung gefragt, fokussieren sie die äußeren Strukturen sowie die Defizite des Kindes. Ähnlich häufig wird das Konzept ‚Beeinträchtigung‘ infragegestellt. Eine wertgeleitete Differenzierung von normalen und nicht normalen Kindern wird auch, insgesamt aber selten, vorgenommen. Werden die Fachkräfte gebeten, ihre Erfahrungen mit Kindern mit Beeinträchtigungen zu schildern und folglich nicht explizit nach Beeinträchtigungen gefragt, wird über die benannten Aspekte hinaus der Zusammenhang zwischen dem Kind und der umgebenen Gruppe berücksichtigt. Die vorliegenden Befunde zeigen somit, dass – im Einklang mit dem Forschungsstand – individuell akzentuierte Verständnisweisen unter den Fachkräften etabliert sind. Allerdings fallen diese nicht durchgehend defizitorientiert aus und werden teils auch von sozial akzentuierten Verständnisweisen begleitet.

Der aktuelle Stand der Auswertungen zeigt, dass unter den Fachkräften sowohl konsistente als auch inkonsistente Überzeugungsmuster zu finden sind. Dabei bringt der geringere Anteil der Fachkräfte auf direkte Nachfrage das gleiche (z.B. defizitorientierte) Beeinträchtigungsverständnis zum Ausdruck wie in ihren weiteren Ausführungen (*konsistentes* Muster). Die Mehrzahl der Fachkräfte zeigt dagegen *Inkonsistenzen*, z.B. indem sie explizit eine umweltbezogene Perspektive verbalisieren, in den weiteren Ausführungen aber ein defizitorientiertes Verständnis zum Ausdruck bringen. Es deutet sich an, dass in den Erfahrungsberichten der Fachkräfte ein ausdifferenzierteres Verständnis von Beeinträchtigung geäußert wird als auf direkte Nachfrage.

PS-57: Bildungsqualität und Fachkräfte: Herausforderungen und Strategien im Schul- und Betreuungswesen

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 135

Chair der Sitzung: Chantal Mose

Qualitätsfragen in der Ganztagsbetreuung 2026 – Einblicke von OGS-Beschäftigten aus NRW

Chantal Mose, Iris Nieding, Katharina Schilling

Institut Arbeit und Qualifikation (Universität Duisburg-Essen)

Der aktuelle Bildungsbericht zeigt, dass Bildungs- und Teilhabechancen in Deutschland weiterhin ungleich verteilt sind (Autor:innengruppe Bildungsberichterstattung, 2024). Damit Kinder unabhängig von ihrer sozialen Herkunft an der Gesellschaft teilhaben können, sind frühzeitige Fördermaßnahmen erforderlich. Ganztagsangebote können durch individuelle Förderung dazu beitragen, soziale Ungleichheiten abzubauen und Aufstiegschancen zu schaffen. Zudem können (Grundschul-)Kinder durch qualitativ hochwertige ganztägige Bildungs- und Betreuungsangebote in ihrer sozialen, emotionalen, körperlichen und persönlichen Entwicklung ganzheitlich gefördert werden (Nolte, 2023; Bellenberg et al., 2021; BMBF, 2023; Tillmann, 2023; BMFSFJ, 2021). Eine zeitlich umfassende Ganztagsbetreuung trägt außerdem zur besseren Vereinbarkeit von Beruf und Familie sowie zur Armutsprävention bei (BMFSFJ, 2021; Graßhoff & Sauerwein, 2021). Sie ist Voraussetzung für eine Steigerung der Erwerbstätigkeit und des Erwerbsvolumens von Eltern und sichert das Familieneinkommen (Gambaro et al., 2019).

Als Reaktion darauf wurde 2021 das Ganztagsförderungsgesetz als bildungs- und sozialpolitische Stellschraube auf den Weg gebracht (Stöbe-Blossey, 2023). Ab 2026 werden sukzessive alle Grundschulkindern einen Anspruch auf einen Ganztagsplatz haben. Aus organisationssoziologischer Perspektive sind Kommunen, Schulen und Einrichtungen dazu angehalten, die Anforderungen und Bedarfe aufzunehmen und geeignete Strukturen zu schaffen. Die Umsetzung des Rechtsanspruchs ist mit der Verankerung im SGB VIII in kommunaler Verantwortung und wird dort enorme Herausforderungen mit sich bringen. Für Kommunen mit prekären Haushaltslagen wird der Ausbau besonders mühsam: geringe Personalausstattungen, niedrige Versorgungsquoten und hohe Bedarfe prägen die Ausgangslage. Für die Erfüllung des Rechtsanspruchs ist in den meisten Bundesländern – trotz bereits gestiegenen Betreuungsquoten – ein massiver Ausbau des Ganztagsangebots und Personalbestands nötig (ebd.; Autor:innengruppe Bildungsberichterstattung, 2024). Dies wird den vorhandenen Fachkräftemangel im pädagogischen Bereich weiter verschärfen. Vielerorts werden deshalb vermehrt Quereinsteiger*innen im offenen Ganztags eingesetzt. Eine weitere Herausforderung sind die (kaum veränderbaren) knappen räumlichen Ressourcen. In den fachlichen Debatten zum Ganztagsausbau wird das Risiko des primär quantitativ verfolgten Ausbaus unter Verlust von Qualität der pädagogischen Betreuung und Förderung von Grundschulkindern diskutiert.

Um herauszufinden, wie eine pädagogisch qualitativ hochwertige Betreuung und Förderung dennoch funktionieren kann, erscheint es wichtig, die Perspektive des Personals einzubeziehen, weshalb folgende Fragestellung im Zentrum steht: **Wie blicken Beschäftigte von offenen Ganztagsschulen in Nordrhein-Westfalen dem Rechtsanspruch auf Ganztagsbetreuung von Grundschulkindern entgegen? Welche Chancen und Herausforderungen werden gesehen?** Vor diesem Hintergrund wurden am Beispiel zweier Kommunen in NRW Ausbaustrategien hinsichtlich der Potenziale und Gelingensbedingungen bei fünf Trägern von offenen Ganztagsschulen erforscht. Die Studie ist an ein Lehrforschungsprojekt im Bachelorstudiengang Politikwissenschaften angebunden; so konnten mithilfe der Studierenden 53 qualitative Interviews mit Beschäftigten in der Ganztagsförderung an 11 Grundschulen geführt werden. Im Rahmen des Einzelbeitrags werden einige zentrale Erkenntnisse auf Grundlage der Interviews präsentiert, welche anhand der qualitativen Inhaltsanalyse nach Kuckartz & Rädiker (2022) ausgewertet wurden.

Insgesamt blickt das Personal dem nahenden Rechtsanspruch 2026 mit großer Sorge entgegen. Unter den aktuell vorherrschenden Rahmenbedingungen wird eine vollständige Versorgungsquote in der Ganztagsförderung an vielen Schulen nur mit einem erheblichen Qualitätsverlust einhergehen können. Die Herausforderungen und Engpassfaktoren sind stark miteinander verwoben: Räumlichkeiten in der Ganztagschule werden als zu klein für eine höhere Betreuungsquote beschrieben, wobei gleichzeitig die Gruppengröße kritisiert wird. Fehlende Vorgaben zu Personalschlüsseln führen zu einem Ungleichgewicht zwischen Gruppengröße und Betreuungskraft. Dies führt zu weniger individueller Betreuung und Förderung der Kinder und zu einer „Abfertigung“ im zeitlichen Ablauf des Ganztagsangebots. Der Quereinstieg von Personal ohne einschlägige Qualifikation wird in der Praxis sowohl als Chance als auch Herausforderung gesehen. Fachfremdes Personal kann die Personallücke im Ganztagsangebot schließen, weist aber kein „Professionswissen“ auf – für die Bewältigung des Alltagsgeschäfts sind laut des OGS-Personals aber andere Kompetenzen entscheidender. Veränderte Teamgefüge sowie ungleiche Bezahlungen können jedoch innerhalb der Ganztagschulen zu Spannungen führen. Ohne grundlegende Veränderungen der Rahmenbedingungen fällt das Zukunftsbild im Ganztags für das interviewte Personal eher negativ aus und wird mit dem Verlust jeglicher Qualität und Förderung betitelt.

ENTFÄLLT: Potential pensionierter Lehrpersonen: Bildungspolitische Reaktionen auf die gesellschaftliche Herausforderung des akuten Lehrer*innenmangels in Österreich

Katharina Resch

Pädagogische Hochschule Oberösterreich, Österreich

Theoretischer Hintergrund

Der Lehrer*innenmangel gilt in Österreich als eine aktuelle, gesellschaftliche und bildungspolitische Herausforderung, auf die dringend Antworten gefunden werden müssen. Sollte dem Personalmangel nicht begegnet werden können, so drohen den Schulen fachliche Engpässe und den Lehrpersonen an Schulen Überstunden in einem nicht abschätzbaren Ausmaß. Für die Berufsgruppe der Lehrer*innen würde dies eine generell abnehmende Arbeitszufriedenheit bedeuten, die mit frühzeitigen Pensionierungen vor dem gesetzlichen Pensionsalter einhergehen können (Ingersoll, 2001). Internationale Studien, etwa aus Belgien, Deutschland oder Spanien, weisen nach, dass die Berufsgruppe der Lehrer*innen aufgrund ihrer belastenden Arbeitsbedingungen früher als üblich in Pension gehen und das Potential ihrer Arbeitskraft nicht vollständig genutzt werden kann (García-Hernández et al., 2022; Seibt & Kreuzfeld, 2023).

Der Lehrer*innenmangel in Österreich ist durch gesellschaftliche Herausforderungen, wie etwa den Ukrainekrieg und COVID-19, noch weiter verschärft worden. Als Reaktion darauf werden derzeit verschiedene bildungspolitische Maßnahmen gegen den akuten Lehrer*innenmangel umgesetzt. Diese sind die Verkürzung der Lehramtsausbildung, der frühzeitige Einstieg von nicht vollständig ausgebildeten Lehrpersonen in den Schuldienst, das Zulassen von Quereinsteiger*innen in den Beruf sowie die Rückholung pensionierter Lehrpersonen in den aktiven Schuldienst. All diese Maßnahmen sollen dem akuten Personalmangel

entgegenwirken, der durch die Pensionierungen der Baby Boomer Generation entsteht. Über die Maßnahme der Rückholung pensionierter Lehrpersonen in den aktiven Schuldienst ist wenig bekannt und noch weniger dokumentiert, sie ist jedoch hinsichtlich des akuten Personalmangels vielversprechend (Droogenbroeck & Spruyt, 2014). Dem widmet sich der vorliegende Beitrag.

Fragestellung

Ziel der Studie ist die Exploration der bisherigen Erfahrungen mit dem bildungspolitischen Modell der Rückholung pensionierter Lehrpersonen in den aktiven Schuldienst aus zwei Perspektiven: aus der Sicht von Expert*innen und der Sicht von pensionierten Lehrpersonen. Die Forschungsfrage lautet: Wie wird der akute Lehrer*innenmangel von Expert*innen aus der Bildungspolitik sowie pensionierten Lehrpersonen wahrgenommen und eingedämmt?

Methode

Auf Basis einer qualitativen Studie (N=14) mit Expert*inneninterviews nach Bogner et al. (2002) mit bildungspolitisch Verantwortlichen und episodischen Interviews nach Flick (1995) mit pensionierten Lehrpersonen zeigt die Studie das Potential und die Stolpersteine der Rückholung pensionierter Lehrpersonen in den aktiven Schuldienst als bildungspolitische Maßnahme zur Bekämpfung des akuten Lehrer*innenmangels auf.

Ergebnisse

Die Ergebnisse zeigen, dass die Entscheidung, aus der Pension in den aktiven Schuldienst zurückzukehren ein multi-faktorieller Prozess ist, der sich aus gesundheitlichen, arbeitsplatzbezogenen, familiären und finanziellen Aspekten zusammensetzt. Zudem zeigen die Ergebnisse, dass pensionierte Lehrpersonen mit einer hohen Identifikation mit ihrem Beruf die Fortführung beruflicher Rollen in der Pension begrüßen – allerdings nur, wenn die Rahmenbedingungen flexibel gestaltbar sind. Kritik an den von der Bildungspolitik gesetzten Rahmenbedingungen üben die Expert*innen: Es werde von pensionierten Lehrpersonen erwartet, für ein Einstiegsgehalt zu arbeiten, das ihre Vordienstzeiten nicht berücksichtigt. Dennoch werden diese in schwierigen Klassen und kritischen, gesellschaftspolitischen Situationen, wie dem Ukrainekrieg, eingesetzt, ohne deren Verantwortung und Professionalität entsprechend abzugelten. Das Modell sei vielversprechend, jedoch müsse an den Rahmenbedingungen gearbeitet werden.

Conclusio

Die Studie leitet entsprechende Empfehlungen für die Bildungspolitik für die Gestaltung weiterer, ähnlicher Maßnahmen gegen den akuten Lehrer*innenmangel ab und zeigt erstmals für Österreich das Potential der Arbeitskraft pensionierter Lehrpersonen auf.

Effekte der KMK-Sondermaßnahmen auf die Unterrichtsqualität und Leistungen von Schülerinnen und Schülern an der Grundschule. Befunde aus dem IQB-Bildungstrend

Eric Richter¹, Lars Hoffmann², Stefan Schipolowski², Sofie Henschel², Sebastian Weirich², Dirk Richter¹

¹Universität Potsdam, Deutschland; ²Institut zur Qualitätsentwicklung im Bildungswesen, Deutschland

Theoretischer Hintergrund

Das deutsche Bildungssystem steht vor großen Herausforderungen, insbesondere vor einem ausgeprägten Lehrkräftemangel. Um diesem Mangel zu begegnen, hat die Kultusministerkonferenz (KMK) verschiedene Sondermaßnahmen eingeführt. Dazu gehören alternative Wege in den Schuldienst wie der Quer- und Seiteneinstieg oder der fach- und schulfremde Einsatz von Lehrkräften (KMK, 2013). Die Auswirkungen dieser Maßnahmen auf die Unterrichtsqualität und die Schülerleistungen im Primarbereich sind bislang nur unzureichend erforscht. Erste querschnittliche Befunde für die Sekundarstufe I deuten darauf hin, dass es keine systematischen Unterschiede zwischen traditionell und nicht-traditionell ausgebildeten Lehrkräften in der Unterrichtsqualität gibt (Lucksnat et al., 2024). Ähnliche Befunde zeigen sich auch für die Kompetenzen der Schülerinnen und Schüler (SuS) in der Sekundarstufe I, wobei die zugrunde liegenden Analysemodelle nicht immer identisch sind (Hoffmann & Richter, 2016; Richter et al. 2019). Die vorliegende Studie greift diese Forschungslücke auf und untersucht die Zusammenhänge zwischen den KMK-Sondermaßnahmen und der Unterrichtsqualität sowie den Leistungen von SuS an Grundschulen.

Fragestellungen

1. Inwieweit hängt die Ausbildung von Grundschullehrkräften mit der durch die Schülerinnen und Schüler wahrgenommenen Unterrichtsqualität in den Fächern Deutsch und Mathematik zusammen?
2. Inwieweit hängt die Ausbildung der Grundschullehrkräfte mit den Kompetenzen der Schülerinnen und Schüler am Ende der 4. Klassenstufe in den Fächern Deutsch und Mathematik zusammen?

Methode

Die Untersuchung basiert auf einer Sekundärdatenanalyse der Daten des IQB-Bildungstrends 2021 für die Fächer Deutsch und Mathematik in der vierten Jahrgangsstufe.

Die Stichprobe umfasst auf Ebene der Lehrkräfte 630 (Deutsch) bzw. 563 (Mathematik) regulär ausgebildete Lehrkräfte, 186 (Deutsch) bzw. 257 (Mathematik) fachfremd unterrichtende Lehrkräfte, 79 (Deutsch) bzw. 70 (Mathematik) schulfremd unterrichtende Lehrkräfte, 21 (Deutsch) bzw. 24 (Mathematik) Quereinsteigende sowie 27 (Deutsch) bzw. 31 (Mathematik) Seiteneinsteigende. Auf der Ebene der SuS umfasst die Stichprobe 11.778 (Deutsch) bzw. 10.610 (Mathematik) SuS bei regulär ausgebildeten Lehrkräften, 3.463 (Deutsch) bzw. 4.663 (Mathematik) SuS bei fachfremd unterrichtenden Lehrkräften, 1.432 (Deutsch) bzw. 1.331 (Mathematik) SuS bei schulfremd unterrichtenden Lehrkräften, 378 (Deutsch) bzw. 442 (Mathematik) SuS bei Quereinsteigenden sowie 516 (Deutsch) bzw. 585 (Mathematik) SuS bei Seiteneinsteigenden.

Die Unterrichtsqualität wurde anhand etablierter Skalen aus Sicht der SuS bezogen auf kognitive Aktivierung, konstruktive Unterstützung und Klassenführung erfasst ($.67 \leq \alpha \leq .90$; $0.07 \leq ICC1 \leq 0.28$). Die Kompetenzen der SuS wurden mit standardisierten Kompetenztests in den Fächern Deutsch (Lesen, Zuhören, Rechtschreiben) und Mathematik (Globalskala) erfasst (Sachse et al., 2022).

Die Datenanalysen wurden mittels 1:1-Multilevel Propensity Score Matching und Mehrebenenregressionen durchgeführt, um die Effekte der verschiedenen Sondermaßnahmen (Quereinstieg, Seiteneinstieg, fachfremd, schulfremd) jeweils im Vergleich zu einer traditionellen Lehramtsausbildung auf die Unterrichtsqualität und die Kompetenzen der SuS zu testen.

Ergebnisse

Die Ergebnisse zeigen, dass es keine signifikanten Unterschiede in den untersuchten Dimensionen der Unterrichtsqualität (kognitive Aktivierung, konstruktive Unterstützung, Klassenführung) zwischen traditionell ausgebildeten Lehrkräften und Lehrkräften mit einer anderen Ausbildung (Quereinstieg, Seiteneinstieg, fachfremd oder schulfremd) gibt. Hinsichtlich der Kompetenzen der SuS zeigen sich signifikante Unterschiede mit geringer Effektstärke zwischen regulär ausgebildeten und schulartfremden Lehrkräften im Fach Deutsch in den Bereichen Rechtschreibung und Zuhören zu Ungunsten der schulartfremden Lehrkräfte. Für die anderen Kompetenzfacetten in den Fächern Deutsch und im Fach Mathematik zeigen sich hingegen keine signifikanten Unterschiede für Klassen von regulär ausgebildeten Lehrkräften und Lehrkräften, die entweder im Quereinstieg, im Seiteneinstieg, fachfremd oder schulartfremd unterrichten.

Diskussion

Die Ergebnisse liefern erste Hinweise darauf, dass die Sondermaßnahmen der KMK eine wirksame Strategie zur Bewältigung des Lehrkräftemangels in der Primarstufe darstellen können. Die Ergebnisse unterstützen die Integration alternativer Zugangswege in den Schuldienst und den flexiblen Einsatz von Lehrkräften. Gleichzeitig werfen sie die Frage nach den nachhaltigen Wirkungen der Lehrkräftebildung auf. Mögliche Ursachen für die fehlenden bzw. geringen Unterschiede in der Unterrichtsqualität und den Schülerkompetenzen könnten auch in der Messung der Konstrukte und in den teilweise geringen Fallzahlen liegen.

Binnendifferenzierendes Handeln von grundständig qualifizierten Lehrkräften und Seiteneinsteiger:innen an Förderschulen im Schwerpunkt Geistige Entwicklung (SPS GE)

Anna Seifart, Nadine Poltz, Oliver Wendt, Antje Ehlert, Katrin Böhme

Universität Potsdam, Deutschland

Gegenwärtig stellt der Lehrkräftemangel eine zentrale Herausforderung im deutschen Bildungssystem dar (Porsch & Reintjes, 2023). Um Personaldefizite auszugleichen wird bundesweit – so auch in Brandenburg – vermehrt auf Seiteneinsteiger:innen zurückgegriffen, insbesondere an Förderschulen (Statistisches Bundesamt, 2023). Bisher veröffentlichte Studien im Forschungsfeld vergleichen mehrheitlich Selbstwirksamkeitserwartungen und Unterrichtsenthusiasmus von Seiteneinsteiger:innen mit denen grundständig qualifizierter Lehrkräfte und finden meist keine signifikanten Unterschiede (Ziegler et al., 2022). Im deutschsprachigen Raum untersuchten bislang lediglich Lucksnat et al. (2024) das konkrete unterrichtliche Handeln beider Lehrkräftegruppen an Sekundarschulen anhand der Basisdimensionen guten Unterrichts. Dabei finden sich keine signifikanten Unterschiede zwischen Seiteneinsteiger:innen und grundständig qualifizierten Lehrkräften.

Angeichts der geringen empirischen Befundlage zielt der Beitrag darauf ab, das unterrichtliche Handeln von grundständig qualifizierten Lehrkräften und Seiteneinsteiger:innen im SPS GE zu vergleichen. Da binnendifferenzierende Förderung – und damit einhergehend die Differenzierung von Aufgaben und Lernzielen – ein zentrales Charakteristikum des Unterrichts im SPS GE darstellt (Stöppler & Wachsmuth, 2010) bearbeiten wir folgende Fragestellung:

F1: Unterscheiden sich Seiteneinsteiger:innen von grundständig qualifizierten Lehrkräften hinsichtlich *a) der Einsatzhäufigkeit binnendifferenzierender Aufgaben* und *b) dem Umgang mit Lernzieldifferenzierung* an Schulen im SPS GE?

Bislang wurden im sonderpädagogischen Kontext kaum Faktoren eruiert, die das Differenzierungshandeln von Lehrkräften beeinflussen. Ziel ist daher außerdem, zu untersuchen, inwiefern neben der Qualifikation andere Faktoren – hier die Berufserfahrung und Lehrkräfte-Kooperation im Rahmen verschiedener Teamteaching-Formen – auf die Dimensionen *a)* und *b)* wirken. Die zweite Forschungsfrage lautet entsprechend:

F2: Inwiefern beeinflussen die berufliche Erfahrung sowie kooperative Teamteaching-Strukturen *a) die Einsatzhäufigkeit binnendifferenzierender Aufgaben* und *b) den Umgang mit Lernzieldifferenzierung* beider Lehrkräftegruppen an Schulen im SPS GE?

Zur Beantwortung der Fragestellungen wurde auf Daten der Studie „Evaluation der Förderschulen mit dem Sonderpädagogischen Förderschwerpunkt Geistige Entwicklung im Land Brandenburg“ zurückgegriffen. Die Stichprobe des onlinebasierten Lehrkräfte-Fragebogens umfasst 179 Lehrkräfte (81% weiblich, 19% männlich) aus allen 30 öffentlichen Förderschulen mit SPS GE in Brandenburg. Hiervon gaben 51 Personen (28.5%) an, Seiteneinsteiger:in zu sein.

Zur Erfassung der *a) Häufigkeit binnendifferenzierender Aufgaben* diente eine adaptierte Version einer PISA-Skala (Frey et al., 2009) (Cronbachs $\alpha = .77$). Zur Erfassung des *b) Umgangs mit binnendifferenzierenden Lernzielen* bewerteten die Lehrkräfte die Gestaltung ihres Unterrichts hinsichtlich des Umgangs mit Lernzieldifferenzierung (Cronbachs $\alpha = .81$).

Die Ergebnisse zeigen, dass beide Lehrkräftegruppen in den Dimensionen *a)* ($M_{\text{Seiteneinsteiger:innen}} = 3.26$, $M_{\text{grundständig ausgebildete Lehrkräfte}} = 3.34$) und *b)* ($M_{\text{Seiteneinsteiger:innen}} = 3.08$, $M_{\text{grundständig ausgebildete Lehrkräfte}} = 3.14$) ähnlich binnendifferenzierend handeln. Dabei kann sowohl für *a)* ($t(176) = -1.55$, $p = 0.250$) als auch für *b)* ($t(177) = -0.634$, $p = 0.527$) kein signifikanter Unterschied zwischen den Gruppen festgestellt werden.

Zur Erfassung der Faktoren, die auf das binnendifferenzierende Handeln wirken, wurden unabhängige multivariate Varianzanalysen (MANOVA) durchgeführt. Als abhängige Variablen gingen *a)* und *b)* in die Berechnungen ein, wobei Qualifikation, Teamteaching-Formen und Berufserfahrung die unabhängigen Variablen darstellten:

Die Qualifikation hat keinen signifikanten Einfluss auf beide untersuchten Variablen. Berufserfahrung beeinflusst die Dimension *a)* nicht signifikant, jedoch die Dimension *b)* ($F(3, 170) = 3.99$, $p < .05$, $\eta_p^2 = .066$). Dabei klärt die Berufserfahrung 6.6% der Varianz auf, was einem mittleren Effekt entspricht.

Die Teamteaching-Formen zeigen sowohl in Bezug auf die Dimension *a)* ($F(1, 159) = 6.22$, $p < .05$, $\eta_p^2 = .038$) als auch auf die Dimension *b)* ($F(1, 159) = 4.02$, $p < .05$, $\eta_p^2 = .025$) einen signifikanten Effekt. Die kleinen Effektstärken weisen darauf hin, dass nur ein geringer Anteil der Gesamtvarianz aufgeklärt wird. Ferner konnten weder Interaktionseffekte zwischen Qualifikation und Berufserfahrung noch zwischen Qualifikation und verwendeter Team-Teaching-Form festgestellt werden.

Die Ergebnisse zeigen, dass Seiteneinsteiger:innen an Schulen im SPS GE ähnlich binnendifferenzierend unterrichten wie grundständig qualifizierte Lehrkräfte. Die Analysen der MANOVA verdeutlichen, dass die Qualifikation nicht als entscheidender Einflussfaktor auf das Differenzierungshandeln wirkt. Die geringe Varianzaufklärung der signifikanten Effekte der Teamteaching-Formen und der Berufserfahrung lassen vermuten, dass andere Faktoren das Differenzierungshandeln der Lehrkräfte stärker beeinflussen. Die Ergebnisse werden im Beitrag vor dem Hintergrund des bestehenden Lehrkräftemangels kritisch diskutiert.

PS-58: Herausforderung digitale Lehre: Perspektiven auf Lernen und Zusammenarbeit

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SO 422
Chair der Sitzung: Maik Beege

Eine Meta-Analytische Überprüfung des Instructor Presence Effekts und potenzieller Moderatoren

Maik Beege

Pädagogische Hochschule Freiburg, Deutschland

1 Theoretischer Hintergrund

Der *Instructor Presence Effect* besagt, dass sichtbare Instruktor:innen in Lehrvideos die soziale Präsenz sowie die Motivation und affektive Bewertungen erhöhen, jedoch keine Verbesserung der Lernergebnisse bewirken. Basierend auf der *Cognitive-Affective-Social Theory of Learning in digital Environments* (CASTLE; Schneider et al., 2021), aktivieren soziale Hinweise soziale Schemata im Langzeitgedächtnis. Folglich werden soziale Prozesse ausgelöst, die wiederum das Lernen fördern sollten. Allerdings muss gemäß der *Cognitive Theory of Multimedia Learning* (CTML; Mayer, 2014) die Instruktor:in zusätzlich zu den Lerninhalten verarbeitet werden, was Arbeitsgedächtnisressourcen erfordert. Daher können soziale Hinweise das Arbeitsgedächtnis überlasten und das Lernen hemmen. Zusammengefasst sind die theoretischen Positionen widersprüchlich, weshalb eine Meta-Analyse durchgeführt wurde, um die Auswirkungen der *Instructor Presence* auf das Lernen zu beleuchten.

2 Methode

Insgesamt erfüllten 35 Studien die folgenden Einschlusskriterien: (1) Einbeziehung *realer Menschen*; (2) Einbeziehung einer *Kontrollgruppe* ohne sichtbare Instruktor:in; (3) die Studien müssen *videobasierte Instruktionen* untersuchen; (4) mindestens muss der *sich bewegende Kopf* einer sprechenden Instruktor:in gezeigt werden; (5) die Studien müssen *Lernergebnisse* berichten; und (6) es mussten ausreichend *statistische Daten* berichtet werden.

Lernleistungsmessungen wurden in *Behaltens-* und *Transferleistungen* unterteilt. Darüber hinaus wurden subjektive Bewertungen einbezogen. Separate Analysen wurden für *soziale Präsenz*, *affektive Bewertungen*, *motivationale Bewertungen* und *extraneous cognitive load* durchgeführt. Moderatorvariablen wurden in Merkmale der Lernenden (z. B. Vorwissen), Studienmerkmale (z. B. Lernzeit) und Merkmale des Lehrers (z. B. Größe des Lehrers) unterteilt.

3 Ergebnisse

Bezüglich der Behaltensleistung betrug die gewichtete mittlere Effektstärke $g^+ = 0,22$, $p = 0,004$. Hinsichtlich der Transferleistung war die berechnete Effektstärke nicht signifikant, $g^+ = 0,06$, $p = 0,46$. In Bezug auf die soziale Präsenz betrug die gewichtete mittlere Effektstärke für die soziale Präsenz $g^+ = 0,35$, $p = 0,01$. Bezüglich der affektiven Bewertung war die berechnete signifikante Effektstärke moderat, $g^+ = 0,43$, $p < 0,001$. In Bezug auf die motivationale Bewertung war die berechnete signifikante Effektstärke moderat, $g^+ = 0,41$, $p = 0,003$. Hinsichtlich der extrinsischen kognitiven Belastung war die berechnete signifikante Effektstärke klein, $g^+ = 0,31$, $p = 0,003$. Allerdings zeigten die Q-Statistiken, dass alle Effektstärken heterogen waren.

Bezüglich der Behaltensleistung war der einzige signifikante Moderator der Lerninhalt. Ein sichtbarer Lehrer war vorteilhaft für den Unterricht in Sozialwissenschaften sowie Informatik. Ein negativer Trend wurde für Geschichtswissenschaften festgestellt.

In Bezug auf die Transferleistung war das Vorwissen ein signifikanter Moderator. Ein positiver Effekt wurde bei Lernenden ohne Vorwissen festgestellt, während ein nicht signifikanter Effekt bei Lernenden mit geringem Vorwissen gefunden wurde. Darüber hinaus erwies sich die Größe der Instruktor:in im Bild als signifikanter Moderator. Mit zunehmender Größe der Instruktor:in sanken die Transferleistungen.

4 Schlussfolgerungen Bezüglich der CASTLE (Schneider et al., 2021) fungierten sichtbare Lehrpersonen als soziale Hinweise, die den Eindruck einer sozialen Präsenz aktivieren und die motivationalen und affektiven Bewertungen weiter verbesserten. Allerdings wurde die Behaltensleistung, aber nicht die Transferleistung gefördert. Die extrinsische kognitive Belastung (ECL) wurde reduziert, wenn ein sichtbarer Lehrer eingesetzt wurde. Folglich schien eine sichtbare Instruktor:in den visuellen Kanal nicht per se zu überlasten. Lehrpersonen, die im Video wenig Platz einnehmen, führten tatsächlich zu positiven Effekten. Diese Effekte scheinen mit zunehmender Größe des Lehrers abzunehmen. Im Einklang mit der CTML (Mayer et al., 2014) beeinträchtigte eine zunehmende Größe eines potenziellen *Seductive Detail* das Lernen. Darüber hinaus war Vorwissen ein signifikanter Moderator, was impliziert, dass Lernende, die sich mit einem völlig neuen Thema befassen, von der sozialen Führung durch eine sichtbare Instruktor:in profitieren.

Die Sicht von Schüler:innen auf hybriden und digitalen Unterricht – Vergleich und Identifikation von Prädiktoren

Julian Brauchle², Laura Schmidberger¹, Valentin Unger², Fabian Grünig², Jan Hochweber²

¹Universität Stuttgart, Deutschland; ²Pädagogische Hochschule St.Gallen, Schweiz

Hintergrund:

Die Maßnahmen zur Eindämmung der Covid-19-Pandemie führten zu erheblichen Einschränkungen im Leben von Jugendlichen und stellten Schüler:innen, beispielsweise durch Schulschließungen und eingeschränkten Kontakt zu Gleichaltrigen, vor Herausforderungen.¹ Studien berichteten über einen Rückgang des Lernerfolgs², der Lebenszufriedenheit³ und einer Zunahme von Angst und Depression⁴ bei Schüler:innen während der Pandemie. Die Schüler:innen schätzten jedoch auch Aspekte wie verstärkt selbstständige Arbeit⁵ und mehr Schlaf während der Schulschließungen zu Beginn der Pandemie⁶. Die Eindämmungsmaßnahmen entwickelten sich mit dem Fortschreiten der Pandemie weiter und führten, durch individuelle Quarantänemaßnahmen, zu verschiedenen Formen des hybriden Fernunterrichts, bei dem ein Teil der Klasse vor Ort und ein anderer Teil per Fernunterricht unterrichtet wurde. Diese Quarantäne kann einen großen Einfluss auf die Psyche haben⁷ und zu Stress, Wut, Angst, Depression und Schlafstörungen führen.^{8,9} Für Schüler:innen, die unter Quarantäne standen, aber nicht ernsthaft erkrankt waren, konnte es vorteilhaft sein, weiterhin in Form eines hybriden Fernunterrichts am Unterricht zu partizipieren, um nicht zu viel Lernstoff zu versäumen sowie andererseits, um die gesellschaftliche Bildungsfunktion der Schule zu erhalten. Während sich zahlreiche Studien bereits mit den Auswirkungen des generellen Fernunterrichts zu Beginn der Pandemie auseinandersetzen, stellen die Auswirkungen der Quarantänebedingungen auf die Wahrnehmung der Schüler:innen während des hybriden Fernunterrichts weitestgehend ein Desiderat dar. Auch der differenzielle Einfluss individueller Schüler:innenmerkmale auf die Wahrnehmung der Formen des Fernunterrichts wurde bislang kaum untersucht. Dies ist aber

zentral, um Fernunterricht möglichst wenig belastend zu gestalten, sollte es erneut zu vergleichbaren Situationen kommen. Zudem können Ideen abgeleitet werden, wie hybride Settings in den regulären Unterricht implementiert werden können.

Zielsetzung:

Diese Studie untersucht kontrastierend, wie Schüler:innen die positiven und negativen Aspekte des allgemeinen sowie hybriden Fernunterrichts während der Quarantäne wahrnehmen. Außerdem wird der Zusammenhang zwischen den Persönlichkeitsmerkmalen der Schüler:innen und der Wahrscheinlichkeit, dass sie diese Aspekte nennen, untersucht.

Methode:

Die Stichprobe der Studie umfasste $N=721$ Schüler:innen (Deutschland und Schweiz). Eine quantitative Inhaltsanalyse¹⁰ kategorisierte die Erfahrungen während des allgemeinen und hybriden Fernunterrichts. Die Häufigkeiten der genannten Aspekte in den beiden Befragungsgruppen wurden verglichen und die Unterschiede mittels Chi-Quadrat-Tests auf Signifikanz geprüft.¹¹

In einem weiteren Schritt wurde der Einfluss der Persönlichkeitsmerkmale Geschlecht, Klassenstufe, Muttersprache, Schulland und selbsteingeschätzte Selbstregulierung der Schüler:innen auf die Häufigkeit der genannten positiven und negativen Aspekte untersucht, da frühere Studien darauf hinweisen, dass diese Prädiktoren die Wahrnehmung der Covid-19-Pandemie der Schüler:innen beeinflussen können.^{12,13,14}

Für die vier am häufigsten genannten Aspekte pro offener Frage wurden dichotome Kodierungen durchgeführt. Um wichtige Vorhersagevariablen zu identifizieren, wurden Random-Forest-Modelle berechnet.^{15,16} Schiefe abhängige Variablen wurden gewichtet¹⁷ und die Hyperparameter wurden für jedes berechnete Modell angepasst.¹⁸ Die Interpretation der identifizierten Prädiktoren erfolgte mittels ICE-Plots.^{19,20}

Ergebnisse:

Als häufigste positive Aspekte beider Lernformate nannten die Schüler:innen die Vorteile, zu Hause zu sein und flexibler arbeiten zu können sowie ein höherer Anteil an Freizeit, während der Mangel an sozialen Kontakten und direkter Lernunterstützung häufig als negative Aspekte genannt wurden. Die Chi-Quadrat-Tests konnten nur geringe Unterschiede in der Wahrnehmung zwischen den Formaten des allgemeinen und hybriden Fernunterrichts identifizieren. Die Random-Forest-Modelle wiesen darauf hin, dass das Schulland und die Klassenstufe (Primar- vs. Sekundarstufe) sowie das Geschlecht der Schüler:innen wichtige Prädiktoren für die Nennung der oben genannten Aspekte in der untersuchten Stichprobe waren. So gaben Schüler:innen in der Primarstufe mit höherer Wahrscheinlichkeit an, dass sie die Vorteile des Zuhause-seins genossen sowie, dass sie soziale Kontakte vermissten, während Schüler:innen aus der Schweiz mit höherer Wahrscheinlichkeit angaben, mehr Freizeit zu haben.

Schlussfolgerungen:

Die Studie bietet differenzierte Einblicke in die Auswirkungen von Covid-19-Eingrenzungsmaßnahmen auf Schüler:innen und ermöglicht es, den unterschiedlichen Bedürfnissen der Schüler:innen in zukünftigen ähnlichen Situationen besser gerecht zu werden. Die Ergebnisse umfassen auch positive Aspekte des Fernunterrichts, die zur Verbesserung des Unterrichts unter normalen Bedingungen genutzt werden können.

Technological change and social inequality in online education at home: New empirical evidence from 29 OECD countries

Rona Geffen, Birgit Becker
Goethe University, Germany

1. Introduction

Internet access across countries has created new learning opportunities for students (González-Betancora et al., 2021). However, it is still unclear how these changes in technology have affected social inequality in online education at home. We investigated two questions. Does the diffusion of the Internet compensate for or increase social inequality in online education use at home? Does this depend on the level of use (high vs. low frequency use)?

2. Theoretical framework

There are several competing theses regarding the relationship between technological advancement and social inequality in online education at home. The compensatory thesis (Erola and Kilpi-Jakonen, 2017) suggests that technological advancement and the spread of Internet use across the population reduce social inequality in online education at home because they compensate for students' limited educational resources at home. Even students who do not have books at home that can help them with their school work have access to numerous sources of information at home from the Internet. For students with more educational resources, however, technological advancement does not matter much for their online education at home, because they already have good access to numerous educational and information sources at home.

The second hypothesis is the cumulative advantages and disadvantages thesis (Erola & Kilpi-Jakonen, 2017). This hypothesis suggests that the diffusion of the Internet increases social inequality in online education at home because those with more educational resources are better equipped to make use of this new opportunity. This difference between the groups is explained by the better access of advantaged students to educational resources at home compared to less advantaged students. Better access to educational resources provides advantaged students with the skills, information and orientations that positively affect the further accumulation of cultural capital – offline as well as online.

A third hypothesis is the effectively maintained inequality thesis (Lucas, 2001), which argues that compensation and cumulative advantages and disadvantages operate in parallel, depending on the level of use of the Internet. Students with limited educational resources at home might get better access to instructional materials due to the increasing spread of the Internet. However, unlike students with more educational resources, they are not in an advantageous position to use the new technology for learning. Therefore, this theory maintains that while increasing the spread of the Internet reduces social inequality in access to online educational resources, what emerges instead is inequality in the frequency of using the Internet for educational purposes. Thus, disadvantaged students are less likely than their advantaged counterparts to use the Internet for learning on a high level.

3. Methods

Using multilevel models with cross-level interactions and country-fixed effects for a sample of students from 29 OECD countries based on harmonized PISA datasets (2009-2022), we examined the association between a country's Internet diffusion and social inequality in online education use at home. As a measure of social inequality, we compared the differences in online education use for students with more and fewer educational resources at home, because these resources are considered important cultural capital.

4. Results

Our findings show that Internet diffusion both compensates for and increases social inequality in online education use, depending on the level of use. Specifically, Internet diffusion in a country reduces inequality in the basic use of online education at home, but it increases inequality at the highest level of use. Students with more educational resources at home are better positioned than those with limited resources to use the new technological learning opportunities. Thus, they spend more time using the Internet for schoolwork as opposed to other purposes. In order to equalize educational opportunities, policymakers should provide better educational resources at home for disadvantaged students.

Digitale Kommunikation und Kooperation unter Lehrkräften – Eine vergleichende Analyse von Messengern und Videokonferenzen

Simon Biller¹, Lena Groß-Mlynek², Jasmin Bastian², Marion Händel¹

¹Hochschule Ansbach; ²Johannes Gutenberg-Universität Mainz

Theoretischer Hintergrund

Digitale Kompetenzentwicklung ist ein wichtiger Bestandteil des Bildungsauftrags in Deutschland (Kultusministerkonferenz, 2020). Während der Einsatz digitaler Medien im Schulunterricht ein breit untersuchtes Forschungsfeld darstellt (Eickelmann & Drossel, 2020; Lorenz et al., 2022) und auch Lehrkräftekooperation im Fokus der Forschung steht (Drossel et al., 2017; Fraillon et al., 2020), wird die Nutzung digitaler Medien zur Kommunikation und Kooperation unter Lehrkräften in empirischen Studien bisher weniger intensiv betrachtet. Es deutet sich jedoch an, dass das Potential digitaler Medien zur Kooperationen und Kommunikation unter Lehrkräften häufig nicht umfassend ausgeschöpft wird (vgl. Drossel & Heldt, 2022).

Technologieakzeptanzmodelle bilden einen Ansatzpunkt ausgehend von persönlichen Einstellungen und Erwartungen das mediale Nutzungsverhalten von Lehrkräften zu untersuchen. Hierbei vereint die Unified Theory of Acceptance and Use of Technology - UTAUT (Venkatesh et al., 2003, 2012) mehrere vorangehende Modelle und betrachtet den Einfluss von Selbstwirksamkeitserwartung, Leistungserwartung an die Technologie, sozialen Aspekten sowie erleichternden Bedingungen auf die Akzeptanz neuer Technologien. Neben der grundsätzlichen Technologieakzeptanz spielt in diesem Zusammenhang die Kommunikationskompetenz von Lehrkräften in Bezug auf digitale Medien ebenfalls eine Rolle (Schauffel et al., 2021).

Fragestellung

Das zentrale Forschungsinteresse dieser Studie liegt auf der schulinternen digitalen Kommunikation und Kooperation von Lehrkräften in außerunterrichtlichen Situationen. Anhand der vergleichenden Betrachtung des Nutzungsverhaltens und der Technologieakzeptanz von Lehrkräften in Bezug auf zwei konkrete Medien, Messenger und Videokonferenzen, soll der Blick auf Unterschiede zwischen synchronen und asynchronen Kommunikationsmedien gerichtet werden.

Zunächst wird die tatsächliche Nutzung von Messengern (inkl. konkreter Anwendungen) und Videokonferenzen von Lehrkräften beleuchtet:

- F1: Ist die Nutzung von Messengern und Videokonferenzen an der eigenen Schule erwünscht und für welche Kommunikationsanlässe werden die Medien jeweils genutzt?
- F2: Gibt es Nutzungsunterschiede zwischen Lehrkräften unterschiedlicher Schulformen?
- F3: Wie unterscheiden sich die Nutzungshäufigkeiten von Messengern und Videokonferenzen im Kontext schulinterner Kommunikation und Kooperation?

Weiterhin werden Einstellungen und Erwartungen von Lehrkräften in Bezug auf die Technologieakzeptanz betrachtet:

- F4: Welche Einstellungen und Erwartungen haben Lehrkräfte gegenüber der Nutzung von Messengern und Videokonferenzen?
- F5: Gibt es Unterschiede bei der Akzeptanz von Messengern und Videokonferenzen?

Methode

Im Juli 2024 wurde eine quantitative Online-Befragung an Grund- und weiterführenden Schulen im Regierungsbezirk Mittelfranken (Bayern), im Kreis Mainz-Bingen (Rheinland-Pfalz) und in Wiesbaden (Hessen) durchgeführt. Insgesamt wurde der Fragebogen von 250 Lehrkräften (72.0% weiblich, 28.0% männlich; Berufserfahrung in Jahren: $M = 18.3$; $SD = 10.3$) vollständig abgeschlossen. Neben der Erfassung demographischer Daten wurden basierend auf dem UTAUT-Modell von Venkatesh et al. (2003, 2012) Fragen zur Technologieakzeptanz von Messengern und Videokonferenzen im Rahmen schulinterner Kommunikation und Kooperation gestellt. Zudem gaben die Lehrkräfte eine Selbsteinschätzung ihrer medialen Kommunikationskompetenz ab (Schauffel et al., 2021).

Die Daten werden aktuell ausgewertet. Dazu werden neben deskriptiven Statistiken (F1, F4), Mittelwertvergleiche anhand gepaarter t -Tests (F3, F5) mit den abhängigen Variablen Messenger (AV1) und Videokonferenzen (AV2) sowie Varianzanalysen mit der unabhängigen Variablen Schulform und den abhängigen Variablen zu Nutzungshäufigkeiten und -anlässen angestellt (F2).

Ergebnisse und ihre Bedeutung

Die Studienergebnisse geben einen Einblick in das aktuelle Nutzungsverhalten von Lehrkräften und zeigen, welche Rolle die Nutzung von Messengern und Videokonferenzen im Rahmen schulinterner Kommunikation und Kooperation spielt. Vor dem Hintergrund, dass die Lebenswelt der Lehrkräfte von Kommunikationsmedien verschiedenster Art durchdrungen ist, können die Studienergebnisse zu einem erweiterten Verständnis der Nutzung digitaler Medien im beruflichen Schulalltag beitragen. Um digitale Kommunikation und Kooperation im Kollegium perspektivisch zu fördern, bildet die Erfassung des Nutzungsverhaltens sowie möglicher Hindernisse für die Akzeptanz synchroner und asynchroner Kanäle eine wichtige Grundlage. Basierend auf den Studienergebnissen zur Nutzung niedrigschwelliger Messengersysteme für asynchrone Kommunikation und Videokonferenzen als synchrones Kommunikationsmedium und eventueller Nutzungsunterschiede bezüglich Schulformen werden Implikationen und passgenaue Förderansätze zur Nutzung digitaler Kommunikations- und Kooperationsformate in der Schule abgeleitet.

PS-59: spezielle Themen der Motivation & Metakognition

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 256

Chair der Sitzung: Miriam Wünsch

Was uns die Handfläche verrät: Neue Einsichten zum Zusammenhang zwischen Psychophysiologie und subjektiv erlebten Leistungsemotionen

Miriam Wünsch, Ji-Young Min, Franziska Eckerskorn, Anne C. Frenzel

LMU München, Deutschland

Leistungsemotionen sind mehrdimensionale Konstrukte, die affektive, kognitive, physiologische, motivationale und expressiv-verhaltensbezogene Komponenten umfassen (Pekrun et al., 2023). Während in der Forschung bisher vorwiegend Selbstberichte zur Erfassung von Leistungsemotionen verwendet wurde, wird dieser in vereinzelteren neueren Studien durch psychophysiologische Maße ergänzt (z.B. Harley et al., 2019; Kiuru et al., 2022). Diese Maße gelten als geeignet, um einen objektiveren, weniger verzerrten Einblick in die emotionalen Zustände von Individuen zu gewinnen (Pekrun, 2023). Jedoch mangelt es noch an einem umfassenden Verständnis davon, wie verschiedene psychophysiologische Maße mit der subjektiven emotionalen Erfahrung in Leistungssituationen zusammenhängen. Die bisherigen Ergebnisse zu Assoziationen zwischen selbstberichteten Leistungsemotionen und physiologischen Maßen wie Hautleitfähigkeit oder Herzrate sind inkonsistent und Zusammenhänge bestenfalls von moderater Stärke (für einen Überblick siehe Horvers et al., 2021; Roos et al., 2021). Ein möglicher Grund für diese inkonsistente Befundlage liegt in der unzureichenden Unterscheidung zwischen inter- bzw. intraindividuellen Zusammenhängen: Auf der interindividuellen Ebene zeigt sich, ob Unterschiede zwischen Personen auf einer Variablen mit Unterschieden zwischen Personen auf einer anderen Variablen zusammenhängen. Die intraindividuelle Ebene hingegen bildet ab, ob Veränderungen auf einer Variable mit Veränderungen auf einer anderen Variable innerhalb derselben Person einhergehen, wie etwa zu verschiedenen Messzeitpunkten. Die von Kiuru et al. (2022) berichteten Korrelationen unterstreichen die Bedeutung dieser Unterscheidung: Nur auf der intraindividuellen, nicht aber auf der interindividuellen Ebene zeigten sich signifikante Korrelationen zwischen Herzrate und selbstberichteten Leistungsemotionen.

Mit der vorliegenden Studie setzen wir uns zum Ziel, ein tieferes Verständnis der Zusammenhänge zwischen verschiedenen psychophysiologischen und Selbstberichtsmaßen im Kontext von leistungsbezogenen Emotionen zu erlangen. Hierzu verwenden wir psychometrische Netzwerkanalyse, bei welcher auf Basis von Partialkorrelationen das Zusammenspiel verschiedener Variablen modelliert wird (Epskamp, 2020). Konkret modellieren wir die Beziehungen zwischen mehreren Maßen der Hautleitfähigkeit (Frequenz von Ausschlägen, Amplitude dieser Ausschläge, tonisches Level) und Selbstberichtsmaßen (Arousal, Valenz, Freude, Stolz, Ärger, Angst, Frustration, Langeweile) sowohl auf inter- als auch auf intraindividuelle Ebene. Hierzu bearbeiteten $N=159$ Teilnehmende ($M_{\text{Alter}}=24.06$, $SD_{\text{Alter}}=7.81$; 81% weiblich, 16% männlich, 1% divers; 3 Personen fehlende Angabe) jeweils fünf bzw. sechs Aufgabenblöcke zum numerischen Denkvermögen. Während der Bearbeitung wurde kontinuierlich die Hautleitfähigkeit der Teilnehmenden gemessen und nach jedem Aufgabenblock berichteten sie ihre aktuellen subjektiven Leistungsemotionen. Mithilfe des psychometrics package (Epskamp, 2024) wurden intraindividuelle Netzwerke durch Grafische Vektor-Autoregressionsmodelle, sowie interindividuelle Netzwerke durch Gaußische Grafische Modelle auf Basis der Mittelwerte für jede Variable (physiologische Maße mit Baseline-Korrektur) berechnet. Jedes Netzwerk beinhaltete je eine selbstberichtete Leistungsemotion, selbstberichtetes Arousal und Valenz, sowie die drei Hautleitfähigkeitsmaße.

Die nach Pruning ($\alpha=.05$) und Modell-Suche (BIC-Kriterium) resultierenden Modelle zeigen auf intraindividuelle Ebene einen signifikanten Zusammenhang zwischen selbstberichteter Valenz und der Frequenz von Ausschlägen in der Hautleitfähigkeit (Kanten $.14-.18$, $p<.001$). Auf interindividuelle Ebene hingegen gab es, trotz der Baselinekorrektur, keine Assoziationen zwischen den physiologischen und den Selbstberichtsmaßen. Das bedeutet, dass Teilnehmende mehr Ausschläge in der Hautleitfähigkeit während Phasen zeigten, die sie persönlich als angenehmer empfanden. Im Gegensatz dazu unterschieden sich Teilnehmende, die im Vergleich zur restlichen Stichprobe generell mehr Ausschläge in der Hautleitfähigkeit zeigten, nicht systematisch von anderen Teilnehmenden in ihrem subjektiven emotionalen Erleben. Zudem widerspricht der positive Zusammenhang zwischen Valenz und physiologischer Reaktion auf der intraindividuellen Ebene der Erwartung, dass negative Emotionen ebenfalls von einer erhöhten physiologischen Aktivität begleitet werden, in welchem Fall sich kein linearer Zusammenhang zwischen Valenz und physiologischer Erregung zeigen würde. Das Finden eines signifikanten Zusammenhangs könnte darauf zurückzuführen sein, dass eine weitere, nicht im Modell enthaltene Variable diesen mediiert. Denkbar ist etwa, dass positiver Valenz mit mehr Engagement und damit erhöhter physiologischer Aktivität einhergeht. Insgesamt liefern die aktuellen Ergebnisse neue Einblicke in das Zusammenspiel von psychophysiologischen Maßen und subjektiver Erfahrung von Leistungsemotionen und geben wichtige Hinweise für zukünftige Forschung und praktische Anwendungen.

Training spatial orientation and navigation with the map-based geogame "GeoGami"

Helen Müller, Angela Schwering

Institut für Geoinformatik - Universität Münster, Deutschland

Theoretical background

Spatial orientation is defined by Maier (1996) as "the ability to physically or mentally orientate oneself in space" and is crucial for navigating environments or when using maps and other geospatial information (Carbonell-Carrera et al. 2020). The ability to orientate and navigate spatially is not only essential for successfully mastering everyday tasks, but also plays an important role in education, particularly in STEM subjects as it is evidenced to be a predictor for STEM performance (cf. Carbonell-Carrera et al. 2020, Ventura et al. 2013).

Individuals demonstrate considerable variability in spatial orientation and navigation abilities, affecting their navigation strategies and performance (Hegarty et al. 2002). While the majority is able to navigate their surroundings with relative ease, challenges do arise, especially in cognitive mapping of large spaces, due to the need for a comprehensive view (Ishikawa & Newcombe, 2021). This process requires a considerable cognitive effort, which frequently results in navigational errors and inefficiencies (cf. Holmes et al. 2018). Nevertheless, research indicates that spatial training significantly enhances spatial ability (Hawes et al. 2022, National Research Council 2006, Smith et al. 2021). However, to date, there remains a lack of interventions integrated into education that aim to train these skills, despite the availability of various methods, including digital media. In this paper we propose a new intervention to train navigational map reading in order to improve spatial orientation. Given the discrepancies in individuals spatial ability and the evidence of the positive influence of spatial training, this study evaluates an intervention of two training sessions with a digital, map-based geogame in spatial ability improvement.

Method

In the first study, we compared a digital map training with a paper map training among primary school students. The second study focused on the evaluation of a teaching unit that integrated digital map training into a theoretical framework of regular geography lessons in a German secondary school. Both studies consisted of a pre- and post-test before and after the training. The tests and trainings included small tasks focusing on different subskills of spatial orientation and navigation, such as spatial configuration and path integration. The digital map-based geogame used in both studies is the GeoGami app (Fig. 1) developed by the Institute for Geoinformatics at the University of Münster (www.geogami.org). The app allows users to create map-based geogames that can be played in the real world (Fig. 2) or in a virtual reality environment on a computer or mobile device. Its purpose is to improve users' map understanding, orientation and wayfinding skills while fostering the development of spatial ability in a playful environment.

Results

In both studies, participants showed significant improvement in their post-test scores after a digital map training with GeoGami. In the first study, the experimental group even outperformed the control group that received paper map training. The calculated effect sizes in the second study for the subskills Spatial Configuration ($d = 0.98$, $p = 2,83 \cdot 10^{-6}$), Landmarks ($d = 0.88$, $p = 2,4334 \cdot 10^{-6}$) and Path Integration ($d = 0.44$, $p = 0,0077$) indicate a large and medium effect. Evaluating the second study in terms of gender differences, although both genders improved their skills, there were noticeable differences depending on the subskills, with girls improving more often overall. In summary, the studies show that training spatial skills, especially spatial orientation and navigation, with a digital map-based geogame that focuses on the practical use of these skills in the real environment is effective and could be a promising method for enhancing spatial ability in education.

Der Einfluss räumlicher Fähigkeiten auf die Effizienz und Lösungsqualität im Problemlösen von Grundschulkindern

Jonas Schäfer, Timo Reuter, Miriam Leuchter

Rheinland-Pfälzische Universität Kaiserslautern-Landau, Deutschland

Räumliche Fähigkeiten sind wesentliche kognitive Fähigkeiten, die sich im Vor- und Grundschulalter entwickeln (Carr et al., 2018). Die Fähigkeiten, sich Objektdrehungen vorzustellen (mentale Rotation), räumliche Informationen zu manipulieren (visuelles Arbeitsgedächtnis) und sich räumliche Informationen zu merken (visuelles Kurzzeitgedächtnis), sind zentrale räumliche Fähigkeiten (Gathercole et al., 2004; Shepard&Metzler, 1971). Die mentale Rotation wird als Komponente der bildlichen Vorstellungskraft betrachtet (Zacks, 2008), während das visuell-räumliche Gedächtnis das Einprägen von wahrgenommenen Informationen ermöglicht, z.B. von Positionen, Formen und Merkmalen von Objekten (Astle et al., 2009). Räumliche Fähigkeiten können nachweislich zu den mathematischen Leistungen von Kindern beitragen (Gilligan et al., 2017), während entsprechende Forschung in anderen MINT-Feldern spärlich ist.

Eine essentielle Fähigkeit in der kindlichen MINT-Entwicklung ist das Problemlösen (John et al., 2018; OECD, 2014). Systematisches Testen und Optimieren wird als effiziente Top-Down-Methode des Problemlösens angesehen (NRC, 2013), die in Kontrast zum Trial-and-Error-Vorgehen steht. Typische Untersuchungsmethoden des Problemlösens von Kindern sind Konstruktionsaufgaben, z.B. mit Zahnrädern (Reuter&Leuchter, 2022; Schäfer et al., 2023).

Es wurde festgestellt, dass räumliche Fähigkeiten mit der mentalen und physischen Manipulation von Objekten bei Kindern assoziiert sind (Töugu et al., 2017). Daher könnten räumliche Fähigkeiten von Kindern ihr Problemlösen beeinflussen, wobei dies im Hinblick auf Effizienz und Lösungsqualität nicht umfassend untersucht wurde. Außerdem gibt es nur wenige Erkenntnisse über altersbedingte Veränderungen dieser Problemlöseaspekte. Daher zielt unsere Studie darauf ab, folgende Forschungsfragen bei Sechs- bis Achtjährigen zu untersuchen:

- 1) Sind räumliche Fähigkeiten mit Effizienz und Lösungsqualität im Problemlösen assoziiert?
- 2) Beeinflusst das Alter die räumlichen Fähigkeiten, sowie die Effizienz und die Lösungsqualität im Problemlösen?

In unserer Studie bearbeiteten 476 Erst- und ZweitklässlerInnen (M_{age}=7.40 Jahre, SD_{age}=0.64 Jahre; 219 weiblich) Aufgaben zur mentalen Rotation (Picture rotation test; Quaiser-Pohl, 2003), visuell-räumlichem Arbeitsgedächtnis (Corsi blocks backward-recall task; Gathercole et al., 2004) und visuell-räumlichem Kurzzeitgedächtnis (Corsi blocks forward-recall task). Außerdem führten die TeilnehmerInnen zwei Problemlösungsaufgaben durch: Karussell (Aufgabenziel: Verbinde zwei Zahnräder auf einem Zahnradbrett miteinander, sodass sich beide in die gleiche Richtung drehen), Propeller (Aufgabenziel: Befestige Propeller auf Zahnrädern, sodass sich ein Propeller möglichst schnell und der andere möglichst langsam dreht). Die TeilnehmerInnen hatten eine ausreichende Anzahl von Zahnrädern, die sie verschieben und drehen konnten. Die Aufgabe wurde nach 3 Minuten beendet oder sobald das Kind angab, mit seiner/ihrer Lösung zufrieden zu sein. Die Effizienz im Problemlösen wurde anhand der Anzahl der durchgeführten Operationen (d.h. Verschieben und Drehen von Zahnrädern) gemessen, wobei weniger Operationen eine höhere Effizienz indizierten. Die Lösungsqualität wurde anhand der Erfüllung der Aufgabenziele der Problemlösungsaufgaben bewertet. Die Aufgaben wurden tabletbasiert mit standardisierten Anweisungen über Kopfhörer durchgeführt (Schäfer et al., 2024).

Strukturgleichungsmodelle mit Alter als manifestem Indikator, sowie räumlichen Fähigkeiten und Problemlösen als latenten Faktoren zeigten gute Modellpassungen. Regressionspfade zwischen den latenten Faktoren ergaben, dass in beiden Problemlösungsaufgaben räumliche Fähigkeiten eine signifikante Prädiktorvariable für die Lösungsqualität waren (Karussell: $\beta = .27^{**}$; Propeller: $\beta = .39^{**}$). Zudem waren räumliche Fähigkeiten teilweise positiv mit Effizienz assoziiert (Karussell: $\beta = .16^{*}$; Propeller: $\beta = .06$).

Das Alter zeigte als Prädiktorvariable einen positiven Einfluss auf die räumlichen Fähigkeiten ($\beta = .42^{**}$) und teilweise auf die Effizienz (Propeller: $\beta = .21^{**}$; Karussell: $\beta = .13$). Alter und Lösungsqualität waren nicht signifikant assoziiert (Karussell: $\beta = .01$; Propeller: $\beta = .06$).

Unsere Studie zeigt, dass das Merken von Positionen, Größen und Merkmalen von Objekten, wie Zahnrädern und Propellern, beim Lösen von MINT-Problemen von wesentlicher Bedeutung ist. Die Ergebnisse deuten zudem darauf hin, dass die bildliche Vorstellung von Objektbewegungen, z.B. Rotationsmechanismen von Zahnrädern, in kindlichen Problemlöseprozessen involviert ist. Die positiven Assoziationen von räumlichen Fähigkeiten mit der Effizienz und der Lösungsqualität im Problemlösen lassen darauf schließen, dass solide räumliche Fähigkeiten das Trial-and-Error-Vorgehen reduzieren und mehr Top-Down-Prozesse ermöglichen. Darüber hinaus zeigen unsere Ergebnisse, dass Kinder mit zunehmendem Alter effizienter beim Problemlösen werden, was in zukünftigen Bildungsstudien näher untersucht werden sollte.

Multimethodale Erfassung von selbstreguliertem Lernen in einer Grundschulstichprobe

Selbstreguliertes Lernen (SRL) wird als zyklischer Prozess gesehen und definiert als "self-generated thoughts, feelings, and actions for attaining academic goals" (Zimmerman, 1998). Die drei Komponenten Kognition, Metakognition und Motivation werden dabei als zentral für die Definition des Konstrukts angesehen (z.B. Perels et al., 2020). SRL zeigt für den gesamten Bildungsweg von der Grundschule bis hin zum Studium Zusammenhänge zu wichtigen Outcomevariablen wie akademischer Leistung (Kitsantas et al., 2008; Throndsen, 2011) oder Wohlbefinden (Fomina et al., 2020) und besitzt damit eine hohe Relevanz für Lernende. Eine möglichst reliable und valide Erfassung von SRL ist unabdingbar, um Interventionen systematisch evaluieren zu können.

Methoden zur Erfassung von SRL lassen sich in Online-/Event- oder Offline-/Aptitude-Instrumente unterteilen (Cleary & Callan, 2018; Wirth & Leutner, 2008): Während erstere sich auf eine bestimmte Aufgabe beziehen und den SRL-Prozess synchron erfassen, messen letztere SRL eher als ein stabiles Konstrukt und retrospektiv. Offline-Instrumente, insbesondere Selbstberichtsfragebogen, sind die am häufigsten genutzten Instrumente, auch wenn ihre Nutzung häufig kritisiert wird (Endedijk et al., 2016). Aufgrund ihres verallgemeinernden Ansatzes eignen sie sich nicht dafür, "tatsächliches" SRL zu erfassen und es bleibt unklar, auf welche Situationen Lernende sich beim Ausfüllen beziehen. Online-Instrumente hingegen können die konkrete SRL-Nutzung in Echtzeit während einer spezifischen Aufgabe erfassen, eignen sich aber weniger dazu, Aussagen zum allgemeinen SRL-Verhalten einer Person zu machen (Dinsmore et al., 2008).

Es gibt Bestrebungen, SRL multimethodal zu erfassen, indem entweder mehrere Instrumente einer Messkategorie oder sowohl Online- als auch Offline-Instrumenten kombiniert werden (z.B., Callan & Cleary, 2018; Dörrenbächer-Ulrich et al., 2021). Durch einen solchen multimethodalen Ansatz soll die Reliabilität der Messung gesteigert (Perry & Rahim, 2011; Veenman, 2011) und die Vor- und Nachteile der verschiedenen Instrumententypen ausgeglichen werden. Erste Studien zeigen Zusammenhänge zwischen SRL-Komponenten, die mit unterschiedlichen Instrumenten einer Messkategorie erhoben wurden (Cleary et al., 2015; Dörrenbächer-Ulrich et al., 2021), allerdings nur wenig Zusammenhänge zwischen Online- und Offline-Instrumenten (Cleary et al., 2015).

Ziel der vorliegenden Studie war es daher, SRL mithilfe verschiedener Instrumente (Fragebogen, Strategiewissenstest, Mikroanalyse) bei Grundschüler*innen zu erfassen und deren Zusammenhänge sowie die Zusammenhänge mit verschiedenen Leistungsvariablen zu untersuchen. Ein weiteres Ziel war die latente Modellierung von SRL durch die drei Instrumente und ihre jeweiligen Komponenten. Im Rahmen der Studie bearbeiteten $N = 107$ Viertklässler*innen die verschiedenen Aufgaben und Fragebögen am Computer. Bei der Mikroanalyse sollten sie einen Text lesen und lernen sowie zwischendrin verschiedene Fragen zu ihrem Lernverhalten beantworten. Die Reliabilität des Verfahrens liegt mit $\alpha = .87$ in einem hohen Bereich (Blaž, 2021). Ähnlich zufriedenstellende Reliabilitäten zeigen sich auch für den Strategiewissenstest ($\alpha = .88$) und den SRL-Fragebogen ($\alpha = .92$). Zudem wurden die Noten in Mathematik und Deutsch als Leistungsmaße erhoben.

Die Ergebnisse zeigen moderate bis hohe Zusammenhänge der verschiedenen SRL-Komponenten innerhalb eines Instrumentes. Über die Instrumente hinweg zeigten sich lediglich zwei signifikante Korrelationen. Während die Mathematiknoten schwach mit der metakognitiven Komponente der Mikroanalyse korrelierten, zeigten sich für die Deutschnoten keine Zusammenhänge. Die Performanz im studieneigenen Wissenstest hängt moderat mit Metakognition zusammen, allerdings nur gemessen mit dem Strategiewissenstest oder der Mikroanalyse. Die latente Modellierung von SRL durch die drei Instrumente und ihre Komponenten zeigt gute Modellfits, allerdings weist der Strategiewissenstest im Vergleich zu den anderen Instrumenten nur sehr geringe Ladungen auf SRL auf.

Ein möglicher Grund für die geringen Ladungen des Strategiewissenstest könnte sein, dass hier das Wissen über nützliche Strategien abgefragt wird und nicht das tatsächliche SRL-Verhalten. Die fehlenden Korrelationen mit den Leistungsvariablen lassen sich womöglich durch eine geringe Varianz der Noten in der Stichprobe erklären.

Die Ergebnisse der vorliegenden multimethodalen Studie zeigen, dass SRL-Instrumente miteinander kombiniert werden sollten, um das Konstrukt möglichst breit zu erfassen und den Stärken und Schwächen der einzelnen Instrumente Rechnung zu tragen.

PS-60: Diversität im MINT-Bereich als Herausforderung für Bildung und Chancengleichheit

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 142
Chair der Sitzung: Charlotte Popp

Just an Alibi or a Real Concern: Results of a Scoping Review on Diversity Studies in STEM

Charlotte Popp¹, Heidrun Stöger², Albert Ziegler¹

¹FAU Erlangen-Nürnberg, Deutschland; ²Universität Regensburg, Deutschland

Diversity usually refers to the variety of different characteristics in a group (Sanger & Gleason, 2020). In school education in particular, classroom diversity is predominantly defined by student characteristics, such as social background, race, gender, or aptitude (Pozas et al., 2020).

As part of the German Mesh_MINT project, this scoping review examines second-order articles on diversity in mathematics, engineering, science, and technology (STEM), focusing on gender, ethnicity, and special needs. By analyzing research through these three lenses, we take an essential step toward synthesizing knowledge about diversity in STEM education. We also gain insights into how seriously the topic of diversity is taken in STEM education research regarding the three student characteristics.

A comprehensive literature search resulted in a database of 1578 potentially relevant articles. After a close inspection, 205 articles were included in the final analysis. The results show that while gender and ethnicity were mentioned with similar frequency, special needs, in particular learning disabilities and difficulties, were particularly prominent. Mathematics was the most frequently analyzed in terms of subjects, followed by the other STEM subjects and science in general.

The results of our analyses led to an exciting, multi-faceted picture of STEM education research considering diversity. It reveals both strengths and severe weaknesses. For example, regarding outcomes, special needs studies focus almost exclusively on attainment, similar to ethnicity studies, while neglecting many potentially influential factors. Gender studies, on the other hand, include far more aspects such as beliefs, interests, and career choices. By and large, although diversity in STEM education has been increasingly addressed in recent years, much more research is still needed, and we will address these weak spots in our contribution.

Geschlecht, MINT und mündliche Beteiligung: Einblicke aus einer Feldstudie mit 6. und 8. Klassen während ihres Mathematik- und Deutschunterrichts

Lukas Mundelsee, Katharina Reschke

Universität Heidelberg, Deutschland

Zahlreiche Studien konnten den Zusammenhang zwischen einer höheren mündlichen Unterrichtsbeteiligung und dem Lernerfolg sowie der Schulleistung von Schüler:innen belegen (für einen Überblick siehe Rocca, 2010). Der Zusammenhang mit der Schulleistung ergibt sich in Deutschland schon alleine dadurch, dass die mündliche Beteiligung zwischen 30% und – je nach Bundesland und Schulfach – teilweise sogar bis zu 100% in die Gesamtnote einfließt (Krüll, 2023). Da Jungen ein tendenziell größeres Interesse für MINT-Fächer zeigen, Mädchen dagegen für Sprachenlernen (Herwartz-Emden et al., 2012; Schurtz & Artelt, 2014) und das Fachinteresse mit der mündlichen Beteiligung im jeweiligen Fach korreliert (Böheim et al., 2020; Wang & Eccles, 2013), sollten sich Jungen in Mathematik und Mädchen in Deutsch mündlich häufiger beteiligen. Damit könnte die mündliche Beteiligung eine Erklärvariable sein, die zu den oftmals gefundenen Leistungsunterschieden zwischen Jungen und Mädchen in MINT-Fächern beiträgt (z.B. Spinath et al., 2014); diese fand jedoch bislang in der Gender- und MINT-Forschung wenig Beachtung. Die vorliegende Feldstudie hatte daher das Ziel zu beleuchten, ob sich Jungen und Mädchen in ihrer mündlichen Beteiligung im Mathematik- und Deutschunterricht unterscheiden.

Darüber hinaus stellen sowohl die Peer-Beziehung als auch die Schüler-Lehrer-Beziehung vielversprechende Moderatorvariablen für diese Zusammenhänge dar. Beispielsweise konnte gezeigt werden, dass Lehrkräfte insbesondere das Arbeitsverhalten von Mädchen begünstigen (z.B. Koepke & Harkins, 2008), während die Peers einen stärkeren Effekt auf Jungen ausüben (z.B. Stürzer, 2003). Somit könnten die genannten Zusammenhänge durch die beiden Beziehungsvariablen beeinflusst werden.

Wir formulierten zusammenfassend folgende Hypothesen: (1) Mädchen und Jungen unterscheiden sich in Deutsch gegenüber Mathematik in ihrer Meldehäufigkeit, Aufrufhäufigkeit und Sprechzeit mit höheren Werten für Mädchen in Deutsch und Jungen in Mathematik. (2) Die Zusammenhänge werden bei Mädchen stärker von der Schüler-Lehrer-Beziehung und bei Jungen stärker von der Peer-Beziehung moderiert.

Wir testeten diese Hypothesen anhand einer Stichprobe von 204 Schüler:innen (54 % Mädchen; $M_{\text{Alter}} = 11.85$, $SD_{\text{Alter}} = 1.05$) in jeweils vier 6. und 8. Klassen einer Realschule. Die Studie speiste sich aus drei Datenquellen: (1) Das Schülergeschlecht, das Fachinteresse sowie die Fachnoten in Deutsch und Mathematik wurden mit Schülerfragebögen erhoben; (2) die Aufrufhäufigkeit und Sprechzeit der Schüler:innen wurden durch Beobachtungen in insgesamt 58 Unterrichtsstunden erfasst; (3) während gleichzeitig die Meldehäufigkeit der Schüler:innen mittels eines Accelerometers ähnlich einem Lineal, das jede:r Schüler:in zum Melden verwendete, ermittelt wurde.

Ergebnisse der Voranalysen bestätigen bisherige Befunde. So fanden wir, dass das Fachinteresse und die Fachnoten signifikant korrelierten; lineare gemischte Modelle, die die Schachtelung der Daten berücksichtigen, zeigten zudem, dass die Mädchen im Vergleich zu den Jungen bessere Noten und ein höheres Interesse am Deutsch- als am Matheunterricht berichteten (und umgekehrt). Um zu untersuchen, ob sich diese Unterschiede auch in den drei Variablen der mündlichen Beteiligung zeigen, wurden generalisierte lineare gemischte Modelle (GLM) gerechnet. Als Kontrollvariable wurde die Anzahl der Lehrerfragen sowie die Wartezeit nach Lehrerfragen aufgenommen. Wir fanden keine Unterschiede zwischen Mädchen und Jungen in Bezug auf die Meldehäufigkeit, Aufrufhäufigkeit oder Sprechzeit, unabhängig davon, ob das Schulfach als Moderator in das Modell mit aufgenommen wurde oder nicht. Um schließlich zu untersuchen, ob die Peer- oder Schüler-Lehrer-Beziehung Einfluss auf diese Zusammenhänge haben, rechneten wir weitere GLM, in die wir die beiden Beziehungsvariablen als zusätzliche Moderatoren aufnahmen. Die Ergebnisse stützten bisherige Befunde: Jungen mit guter Peer-Beziehung meldeten sich in unserer Stichprobe häufiger in Mathematik, während umgekehrt Mädchen sich tendenziell mehr in Deutsch meldeten und in Mathematik mehr sprachen, wenn sie von einer guten Schüler-Lehrer-Beziehung berichteten.

Die Ergebnisse deuten darauf hin, dass sich das höhere Fachinteresse der Jungen in Mathematik und der Mädchen in Deutsch nicht in deren mündlicher Beteiligung niederschlägt und somit die Geschlechtsunterschiede in den Fachnoten nicht erklären kann.

Dagegen scheinen sowohl die Peer- als auch Schüler-Lehrer-Beziehung vielversprechende Erklärvariablen zu sein, die die Effekte abschwächen bzw. verstärken.

Soziale Vererbung der MINT-Brille: Welchen Vorteil haben Kinder aus MINT-Familien?

Johanna Pfeuffer¹, Heidrun Stöger², Albert Ziegler¹

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland; ²Universität Regensburg, Deutschland

Der Beitrag beschäftigt sich mit der beruflichen Tätigkeit der Eltern und deren Bedeutung für die individuelle Identifikation und die Wahlentscheidungen ihrer Kinder im MINT-Bereich. Vor dem Hintergrund der Chancengleichheit im Bildungsbereich und der weltweit forcierten MINT-Förderung von Mädchen und Jungen, ist eine Untersuchung dieses familiären Vorteils von Belang.

Bourdieu's Habitus ist der theoretische Ausgangspunkt der Studie und wurde von Archer et al. (2012) auf Familien und den MINT-Bereich bezogen. Entsprechend dieser theoretischen Sichtweise prägt der familiäre Habitus tagtägliche Praktiken, Ressourcen, Werte und Identitätsvorstellungen von Familien (Archer et al., 2012). So ist davon auszugehen, dass sich Familien, abhängig vom MINT-Hintergrund, hinsichtlich der Geisteshaltung und der Kapitalausstattung unterscheiden (Moote et al., 2020).

Vor diesem theoretischen Hintergrund wurden in der Studie Familienkonstellationen in den Blick genommen und die MINT-Wahlintentionen von Kindern verglichen. In einem ersten Schritt wurde untersucht, ob sich die Wahlintentionen in Abhängigkeit davon, welche Elternteile im MINT-Bereich arbeiten, unterscheiden (RQ1: The more, the better?).

Angeichts der Uneinigkeit bezüglich der geschlechterspezifischen Wirkung von (MINT-)Eltern (Lee et al., 2020; Hoferichter & Raufeder, 2019; Sonnert, 2009) wurde in einem nächsten Schritt getestet, ob die Gruppenunterschiede bei der Wahlintention in Abhängigkeit des familiären MINT-Hintergrunds für Mädchen und Jungen gleichermaßen gelten. Vermutet wurde, dass ein familiärer MINT-Hintergrund für Mädchen bedeutsamer wäre als für Jungen (RQ2: Booster for Girls?).

Abschließend wurde auf Erklärungsursachen eingegangen und untersucht, zu welchem Anteil die individuelle MINT-Identifikation den Zusammenhang zwischen dem familiären MINT-Hintergrund der Eltern und den MINT-Wahlintentionen erklären kann (RQ3: Inherited Identities).

Für die Analyse wurden Befragungsdaten von Mittelstufenschüler:innen von Gymnasien, Gesamtschulen und Stadtteilschulen mit MINT-Schwerpunkt der deutschlandweiten FösaMINT-Studie ($N > 1000$) herangezogen. Zur Beantwortung der Fragestellungen wurden Varianzanalysen und Post-hoc-Tests eingesetzt sowie Mediationsmodelle geschätzt.

RQ1: Auf Basis der Datengrundlage zeigten sich signifikante Gruppenunterschiede in Abhängigkeit des familiären MINT-Hintergrunds hinsichtlich der MINT-Wahlintentionen von Schülerinnen und Schülern. Mit Ausnahme des Vergleichs von Kindern von MINT-Müttern und MINT-Vätern wurden signifikante Mittelwertsunterschiede bei den MINT-Wahlintentionen festgestellt. Die Ergebnisse zeigen an, dass Kinder, deren Eltern einen familiären MINT-Hintergrund aufweisen, häufiger zu MINT-Wahlen tendieren als Kinder, deren Eltern nicht im MINT-Bereich arbeiten.

RQ2: Bezüglich der zweiten Forschungsfrage zeigten die Analyseergebnisse an, dass kein Interaktionseffekt zwischen dem familiären MINT-Hintergrund und dem Geschlecht hinsichtlich der abhängigen Variable MINT-Wahlintentionen besteht. Auf Basis der Post-Hoc-Tests wurden für Jungen signifikante Mittelwertsunterschiede für den Gruppenvergleich von zwei MINT-Elternteilen und keinem MINT-Elternteil gefunden. Für Mädchen wurde beim Vergleich der Gruppe ohne MINT-Eltern mit den weiteren Gruppen, eine MINT-Mutter, einen MINT-Vater und zwei MINT-Elternteile zu haben, signifikante Unterschiede gefunden. Die Analyseergebnisse deuten an, dass der familiäre MINT-Hintergrund hinsichtlich der Wahlintentionen für Mädchen eher einen Unterschied als für Jungen macht.

RQ3: Die Mediationsanalyse zeigte, dass der positive Zusammenhang zwischen dem familiären MINT-Hintergrund und den MINT-Wahlintentionen teilweise durch eine stärkere individuelle Identifikation mit MINT mediert wird. Bezüglich der Wahlintentionen kann der Effekt einer MINT-Mutter oder eines MINT-Vaters im Vergleich zu Kindern, deren Eltern nicht im MINT-Bereich arbeiten, fast vollständig durch eine stärkere individuelle Identifikation mit MINT erklärt werden. Auch Kinder, deren beide Elternteile im MINT-Bereich arbeiten, weisen höhere MINT-Wahlintentionen auf. Ein großer Anteil des erklärten Effekts kann auf eine stärkere Identifikation zurückgeführt werden.

Abschließend kann auf Basis der Datengrundlage und Analyseergebnisse geschlossen werden, dass die berufliche Tätigkeit im MINT-Bereich einen Unterschied für die MINT-Wahlintentionen der Kinder macht. Die Ergebnisse sind ein Indikator für ungleiche Start- und Zugangsvoraussetzungen von Kindern, deren Eltern nicht im MINT-Bereich arbeiten. Auch wenn die Analyseergebnisse bisher nicht auf einem Längsschnitt beruhen, ist davon auszugehen, dass sich Kinder, deren Eltern nicht im MINT-Bereich arbeiten, seltener schulisch, außerschulisch, akademisch oder beruflich für MINT entscheiden. Im Hinblick auf Chancengleichheit und des Bedarfs an Fachkräften im MINT-Bereich bedarf es Maßnahmen, derartige Ungleichheiten durch identitätsstiftende Maßnahmen auszugleichen.

Genderspezifische Unterschiede im Technischen Problemlösen

Jennifer Stemann¹, Marcus Schrickel², Carolin Hahnel²

¹Pädagogische Hochschule Freiburg, Deutschland; ²Ruhr-Universität Bochum, Deutschland

Theoretischer Hintergrund

Durch Weiterentwicklungen erhalten alltagstechnische Geräte zunehmend neue Funktionen und werden komplexer (z.B. Raumthermostate; Siepmann, 2016). In der Folge kennen nur wenige Personen alle Gerätefunktionen und die Bedienung wird häufig zu einem Problem. Bedienungswissen kann prinzipiell auf verschiedene Arten erworben werden. Verbreitet sind vor allem die Nutzung schriftlicher Instruktionen (Becker-Mrotzek, 1997) und das selbstgesteuerte Explorieren eines technischen Systems (Oehme, 2014). Der selbstständige Erwerb fehlenden Bedienungswissens gelingt aber nicht allen Nutzern gleich gut. Neben dem nötigen kognitiven Aufwand erfordert es auch die Bereitschaft, sich selbstreguliert mit der technischen Problemsituation auseinanderzusetzen (Beier, 2004). Diese Bereitschaft hängt u.a. von Traits wie Technikbezogenen Selbstwirksamkeitserwartungen (SWE) ab, die das Vertrauen einer Person beschreiben, aufgrund ihrer eigenen Fähigkeiten, zukünftige Situationen im Umgang mit Technik bewältigen zu können (Autoren). Personen mit geringerer SWE (Jones et al., 2000) neigen dazu, Anweisungen Schritt für Schritt zu befolgen. Einzelne Studien zeigen, dass das Explorieren als Methode zudem häufiger von männlichen Personen eingesetzt und besser beherrscht wird (Vincent & Janneck, 2012), wohingegen Frauen eher angeben, zunächst Bedienungsanleitungen zu lesen (Beckwith et al., 2006). Frauen haben in technischen Kontexten ein konstant geringer ausgeprägtes Selbstkonzept als Männer (acatech & VDI, 2009). Die Analysen von Guedel (2014) zeigten die größten Unterschiede in SWE in den Bereichen «Technik nutzen und reparieren» ($d=1.3$) und «Technik verstehen und erklären»

($d=.80$). Bisherige Arbeiten, die den Einfluss von Einstellungen und Selbstwahrnehmungen auf das Nutzungsverhalten von Technik untersuchen, erfassen oft nur selbstberichtetes Verhalten (Karrer et al., 2009; Richter et al., 2010).

In dem vorliegenden Beitrag wird daher geprüft, ob Frauen tatsächlich häufiger zu Bedienungsanleitung greifen, ob dieses Verhalten mit ihrer SWE zusammenhängt und ob ihre Fähigkeiten im Umgang mit den Geräten ihrer berichteten SWE entspricht.

Methode

Zur Erfassung des Technischen Problemlösens (TPL) wurde ein computerbasierter Test mit 36 Items zu neun Simulationen weiterentwickelt (*Autoren*). Eine TPL-Simulation präsentierte zunächst ein technisches Gerät (z.B. ein digitales Heizungsthermostat) und erfragte das gerätespezifische Vorwissen von Probanden, die das Gerät anschließend selbstständig erkundeten und Funktionen in einem Bedienungsmanual nachschlagen konnten. Nach Abschluss dieser Wissenserwerbsphase wurden den Probanden vier Items vorgelegt, in denen sie eine bestimmte Handlung am Gerät vornehmen mussten. Anhand von Logdaten wurde anschließend bestimmt, ob, wie lange und wie umfassend Personen ein technisches Gerät in der Wissenserwerbsphase erkundet und/oder das Bedienungsmanual genutzt haben. Zur Beantwortung der Forschungsfragen wurden die Daten einer Stichprobe von $n=213$ Erwachsenen analysiert. Neben den Leistungs- und Prozessdaten aus dem TPL-Test wurden auch demografische Angaben sowie Daten zur technikbezogenen SWE erfasst.

Ergebnisse

Die Ergebnisse des t-Tests für unabhängige Stichproben zeigen erwartungskonform geschlechtsspezifische Unterschiede in der technikbezogenen SWE, $t(185)=5.25$, $p<.001$, mit einer Effektstärke $d=.73$. Ein geringer Effekt ($d=.25$) zeigt sich hinsichtlich der Unterschiede in der Präferenz des Explorierens: Männliche Probanden nutzten in der Wissenserwerbsphase eher die Möglichkeit des Explorierens als Frauen, $t(205)=1.77$, $p=.03$. Keinen Unterschied gab es dabei in der Explorationsvollständigkeit, wohl aber in der Vollständigkeit der Manualnutzung. Frauen lasen Bedienungsanleitungen tendenziell vollständiger als Männer, $t(206)=-2.40$, $p=.99$, mit $d=.34$. Die SWE hatte erwartungswidrig keinen Einfluss auf die Präferenz einer Wissenserwerbsart. Der Zusammenhang von SWE und dem Anteil des Explorierens an der Wissenserwerbsphase betrug $r=.10$, der von SWE und der Explorationsvollständigkeit $r=.16$ und der von SWE und der Manualvollständigkeit $r=-.04$. Trotz, dass die SWE einen kleinen Einfluss auf die anschließende Steuerungsleistung mit $r=.20$ hat und Frauen eben eine geringe SWE berichteten, zeigten sich nur geringe geschlechtsspezifische Unterschiede in den Steuerungsleistungen, $t(208)=1.60$, $p=.05$, mit $d=.23$. Frauen gelang es offenbar durch die Kombination von Exploration und Manualnutzung in der Wissenserwerbsphase die durch das SWE oft vorhergesagte Leistung auszugleichen. Beide Wissenserwerbsmöglichkeiten hängen positiv mit der Steuerungsleistung zusammen, die Explorationsvollständigkeit mit $r=.52$ und die Manualnutzung mit $r=.37$. Diese Annahme ist allerdings noch weiter zu prüfen.

PS-61: Eltern-LehrerInnen Kooperation & elterliches Handeln im Schulkontext

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SO 322

Chair der Sitzung: Josef Strasser

Zusammenarbeit von Erziehungsberechtigten und Lehrkräften – ein Erklärungsmodell für ein gelingendes Miteinander

Caroline Villiger, Edith Niederbacher, Sümeyra Sahbaz

Pädagogische Hochschule Bern, Schweiz

Die Familie und damit insbesondere die Erziehungsberechtigten resp. die Eltern haben einen wesentlichen Einfluss auf den Bildungserfolg von Heranwachsenden (OECD, 2023). Dieser Einfluss auf die Leistungen von Schülerinnen und Schülern wird sogar stärker eingeschätzt als derjenige der Schule, was oftmals zur Schlussfolgerung führt, dass die Elternzusammenarbeit intensiviert werden sollte, um Bildungschancen zu stärken (z.B. Sacher, 2022). Demgegenüber steht die Tatsache, dass Lehrkräfte stark belastet sind aufgrund diverser zunehmend komplexer Aufgabenstellungen, wobei nicht zuletzt auch die Elternkontakte dazugehören (Brägger et al., 2019). Davon ausgehend stellt sich die Frage, wie eine konstruktive Elternzusammenarbeit verbunden mit den erhofften positiven Effekten zustande kommt, ohne dabei eine erhöhte Belastung zu generieren. Es gibt in der Fachliteratur diverse Arbeiten, die sich mit Motiven und Hindernissen elterlicher Beteiligung in schulischen Belangen (school-based/home-based involvement) auseinandergesetzt und wichtige Erklärungsfaktoren für elterliches Verhalten geliefert haben (bspw. Eccles & Harold, 1996; Green et al., 2007; Schwanenberg, 2015; Yotodying & Wild, 2014). Diese Arbeiten bieten wertvolle Einblicke in die Frage, wie, warum und unter welchen Umständen Eltern sich beteiligen und für die schulische Entwicklung ihrer Kinder interessieren. Bisherige Arbeiten zur Zusammenarbeit von Schule und Familie im Allgemeinen haben sich darauf konzentriert, das Verhältnis der beiden Instanzen (sowie weiteren: z.B. Community) zu beschreiben (z.B. Epstein's Modell der überlappenden Sphären, Epstein et al., 2019) oder Qualitätsmerkmale der Zusammenarbeit zu definieren (z.B. Vodafone Stiftung, 2013). Letztere haben jedoch einen normativen Charakter und sind losgelöst von den agierenden Personen. Dies stellt ein Nachteil dar, wenn es darum geht gelingende Elternzusammenarbeit zu beschreiben, da Ursachen von Konflikten bei der Zusammenarbeit im Wesentlichen auf der interpersonalen Ebene verortet sind. So wird vermutet, dass Überzeugungen und Dispositionen (bspw. Rollenkonstruktion, Werte, Erwartungen) eine wichtige Rolle bei der Zusammenarbeit spielen und zu einem gelingenden Miteinander wesentlich beitragen können.

Dieser Beitrag versteht sich als Übersichtsarbeit und verfolgt das Ziel, die bisherigen Arbeiten und Modelle zur Elternzusammenarbeit systematisch aufzuarbeiten und einzuordnen. Dabei werden in der bisherigen Literatur zum Teil miteinander verstrickte Aspekte (z.B. institutionelle Strukturen und Gefässe der Zusammenarbeit, persönliche Dispositionen) entflechtet und in eine übersichtliche Struktur gebracht. Das vorgeschlagene Modell bietet eine Grundlage für empirische Studien zur Erklärung von positiv erlebter Zusammenarbeit einzelner Beteiligter. Es trägt den jeweils unterschiedlichen individuellen Merkmalen und Überzeugungen sowie Dispositionen Rechnung und berücksichtigt gleichzeitig institutionelle und gesetzliche Rahmenbedingungen der Zusammenarbeit. Das Modell dient als Grundlage für die Theoriebildung, aber auch für die Schulentwicklung (Veränderung von Praktiken der Elternzusammenarbeit) sowie für die Konzeptualisierung von Lehrveranstaltungen in der Lehrkräftebildung. Der Beitrag lädt dazu ein, den Nutzen des vorgeschlagenen Modells zu diskutieren und allfällige blinde Flecken, die das Modell nicht abzubilden vermag, zu identifizieren.

Professionelle Wahrnehmung im Kontext schulischer Elterngespräche – Erforschung expertisebedingter Wahrnehmungsunterschiede mittels Eye-Tracking

Kristina Ackel-Eisnach, Inga Wagner, Ilona Weyrauch, Josef Strasser

Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau (RPTU), Deutschland

Anders als professionelle Berater:innen in der pädagogisch-psychologischen Praxis werden Lehrkräfte hinsichtlich ihrer beratungsbezogenen Aufgaben allenfalls rudimentär ausgebildet. Dieses Manko ist nicht zuletzt auf fehlende Kenntnisse darüber zurückzuführen, worauf kompetentes Beratungshandeln im schulischen Kontext basiert (Strasser & Behr, 2022).

Es ist anzunehmen, dass die Wahrnehmung von Beratungssituationen hier eine zentrale Rolle spielt. Um ein adäquates individuelles Fallverständnis zu entwickeln, gilt es, die Situation umfassend zu beobachten und relevante Informationen angemessen zu verarbeiten. (Strasser & Gruber, 2015; Strasser et al., 2005). Es stellt sich die Frage, inwieweit schulische Beratungskompetenz auf spezifischen Mustern der Wahrnehmung gründet, oder konkreter: inwiefern und in welcher Hinsicht Beratungsexpert:innen eine Gesprächssituation anders wahrnehmen als weniger kompetente bzw. erfahrene Personen

Im Projekt „BiB – Beratung im Blick“ wird dieser Frage mit Hilfe von Eyetracking und „Retrospective thinking aloud“ nachgegangen. Neben dem Ausmaß an Beratungsexpertise werden zur Kontrolle weitere personenbezogene Variablen berücksichtigt, wie z.B. Empathie oder beratungsbezogene Selbstwirksamkeit, um Unterschiede in der Wahrnehmung von Gesprächssituationen erklären zu können.

Vor der empirischen Untersuchung dieser Fragestellung wurde der aktuelle Forschungsstand zu expertisebedingten Wahrnehmungsunterschieden von Berater:innen und Lehrkräften in Gesprächssituationen systematisch erhoben. Hierfür wurden verschiedene Datenbanken (FIS-Bildung, peDOCS, Eric, PsycINFO und Pubmed), die unterschiedliche Forschungsbereiche und Sprachräume abdecken, systematisch für den Zeitraum 2000 bis 2023 durchsucht. Die Festlegung der zweisprachigen Suchstrings erfolgte unter Anwendung des PICO-Konzeptes (Higgins et al. 2019).

Durch die Recherche konnten insgesamt 805 Beiträge in den Datenbanken ermittelt werden, die unter Anwendung spezifischer Ein- und Ausschlusskriterien (v.a. Eye-Tracking als Messmethode,) ausgeschlossen werden mussten, da keiner der Beiträge expertisebedingte Wahrnehmungsunterschiede bei Elterngesprächen durch Lehrkräfte und Berater:innen mittels Eye-Tracking-Methoden untersuchte. Im Gegensatz zu inzwischen intensiv erforschten Bereichen professionellen Lehrer:innenhandelns wie dem Unterrichten (Keskin et al., 2024), wird demnach mit dem Erforschen der professionellen Wahrnehmung in einem anderen Aufgabenbereich Neuland betreten.

In einer (noch laufenden) quasi-experimentellen Studie werden Expert:innen und Noviz:innen aus der pädagogischen/psychologischen Beratungspraxis zu ihren Wahrnehmungsprozessen von schulischen Elterngesprächen untersucht. Die Stichprobe besteht bislang aus $n = 8$ Expert:innen, 5 weibliche, 2 männliche, 1 k.A., Alter: $M_E = 32.12$ ($SD_E = 4.22$) und $n = 28$ Noviz:innen (23 weibliche, 5 männliche Lehramtsstudierende, $M_N = 22.68$ ($SD_N = 4.33$)). Die Stichprobe wird weiter ausgebaut.

Den Proband:innen wird ein authentisches videographiertes schulisches Elterngespräch gezeigt, dabei werden ihre Blickbewegungen (u.a. Fixationen, Sakkaden) mittels Eye-Tracking gemessen. Anschließend folgen Fragen zur kognitiven Belastung (Paas, 1992). Danach betrachten die Proband:innen nochmals das Video mit ihren Blickbewegungen und verbalisieren ihre Gedanken (retrospective thinking aloud). Die Personen werden mit Bild und Ton aufgenommen. Zuletzt folgt ein standardisierter Fragebogen zu soziodemographischen Merkmalen, Vorerfahrungen, Empathie (Leibetseder et al., 2001) und beraterbezogener Selbstwirksamkeit (Hertel, 2009).

Zur Auswertung wurden verschiedene Areas of Interest (zu Mimik, Gestik, salienten Merkmalen) und Phasen des Gesprächs („belebtere“ vs. ruhige Phasen) voneinander unterschieden. Aufgrund der noch kleinen Stichprobe können bislang nur vorläufige Ergebnisse auf deskriptiver Basis berichtet werden. Sie deuten auf bedeutsame expertisebedingte Unterschiede hin. Für die Expert:innen scheint v.a. die Mimik der beteiligten Personen im Mittelpunkt zu stehen (84% der Fixationszeit vs. 73%). In belebteren Situationen blicken Expert:innen häufiger auf das Gesicht der nichtsprechenden Person (58% vs. 47%), bei ruhigen Abschnitten ergibt sich kein Unterschied. Perzeptuell saliente aber irrelevante Aspekte werden von den Expert:innen so gut wie nicht beachtet (2 Details durch 1 Person) während dies bei Noviz:innen wesentlich häufiger der Fall ist.

Sollten sich die bislang identifizierten Unterschiede statistisch absichern lassen, deutet dies daraufhin, dass die umfassende Beobachtung beider Gesprächspartner:innen, insbesondere auch der Reaktionen der nichtsprechenden Person in potentiell schwierigen Gesprächsphasen, sowie das Ausblenden irrelevanter Aspekte wichtige Aspekte expertenhafter Wahrnehmung sind. Dies korrespondiert mit Ergebnissen zur professionellen Wahrnehmung von Unterricht (Wolf et al., 2016) wie auch zur Beraterexpertise (Strasser & Gruber, 2015).

[

ENTFÄLLT Die Kooperation von Lehrkräften und Eltern im Kontext unterschiedlich privilegierter Elternschaft

Sarah Eiden^{1,2}

¹Ruhr-Universität Bochum, Deutschland; ²Universität Duisburg-Essen, Deutschland

Die Kooperation von Eltern und Lehrkräften ist ein zentrales Merkmal schulischer Qualität (Muijs 2003). An Schulen mit sozial benachteiligter Schüler- und Elternschaft kommt eine Kooperation jedoch häufig nicht zustande (Gomolla, 2009; Schwaiger & Neumann, 2010). Eltern berichten, dass ihre spezifischen Bedürfnisse und Potenziale durch Schule und Lehrkräfte nicht ausreichend berücksichtigt werden. Das Selbstverständnis und die Einstellungen von Lehrkräften werden von Eltern als kooperationshinderlich wahrgenommen (Fürstenau & Hawighorst, 2008). Killus & Paseka (2016) zeigen auf, dass die wahrgenommene Adressierung von Eltern durch Lehrkräfte deren Bereitschaft und Praxis zur Kooperation beeinflusst. Während zahlreiche Studien das (defizitäre) Kooperationsverhalten von Eltern zentrieren, gibt es nur unzureichende Erkenntnisse zum Kooperationsverständnis und -handeln von Lehrkräften. Das Schulgesetz für NRW formuliert für die Kooperation von Lehrkräften und Eltern den Anspruch bei der „Verwirklichung der Bildungs- und Erziehungsziele“ (§ 2 III) „in vertrauensvoller Weise“ (§ 62 I) „partnerschaftlich zusammen“ (§ 2 III) zu wirken und eine „angemessene“ Vertretung von „Schülerinnen und Schüler[n] aus Migrantenfamilien und ihre[n] Eltern“ „in Mitwirkungsgremien“ (§62 III) zu gewährleisten (Schulgesetz NRW, Stand 15.06.2014). Trotz klarer Vorgaben zu elterlichen Partizipationsrechten und schulischen Beratungs- und Informationspflichten obliegt die konkrete Ausgestaltung der Kooperation in der Verantwortung der Schulen und Lehrkräfte. Es fehlen Erkenntnisse darüber, ob und wie Lehrkräfte diesen Gestaltungsspielraum wahrnehmen und nutzen. Je gegensätzlicher die Lebenswelt von Lehrkräften und Familien, desto erwartbarer sind Schwierigkeiten eine Partnerschaft zu entwickeln (Chrispeels, 1996). Sacher (2011) stellt zudem fest, dass Lehrkräfte die fehlende schulische Präsenz bildungsnaher Eltern weniger irritiert als die Abwesenheit von Eltern, die über geringe formale Bildungsabschlüsse verfügen (Sacher, 2011, S. 36).

Der vorliegende Beitrag präsentiert Erkenntnisse über die Kooperationshäufigkeit und -zufriedenheit von Eltern und Lehrkräften im Kontext unterschiedlich privilegierter Schüler- und Elternschaft. Zudem wird die Frage zentriert, wie Lehrkräfte die Kooperation mit Eltern gestalten und welche Rolle die sozialen Hintergründe der Schüler- und Elternschaft dabei spielen. In einem „Sequential Mixed Design“, wurden zunächst die Perspektiven von Eltern und Lehrkräften an Schulen mit unterschiedlich privilegierter Schüler- und Elternschaft kontrastiert (TP1). Die Datengrundlage bildet eine umfassende Befragung von Schulleitungen, Lehrkräften, Schüler:innen und Eltern in der Sekundarstufe I an 36 Schulen in herausfordernden Lagen, durchgeführt im Rahmen des Projektes „Potenziale entwickeln – Schulen stärken“. Es wurden deskriptive Korrelationsanalysen durchgeführt um die Perspektiven von Eltern und Lehrkräften sowie zwischen Schulen mit unterschiedlich privilegierten Elternschaften bzgl. relevanter Kooperationsaspekte zu kontrastieren. Als Maß für die soziale Herkunft der Schüler:innen (bzw. ihrer Eltern) wurde ein durchschnittlicher Sozialindex der Schulen gebildet, welcher auch die Grundlage für die Kontrastierung in TP2 bildet. Hier wurden acht Schulen für eine vertiefende Fallstudie zur Perspektive der Lehrkräfte (TP2) ausgewählt. Es liegen 32 Experteninterviews mit episodisch-narrativen Passagen zur Ausgestaltung der Kooperationsbeziehung durch Lehrkräfte vor, die mittels inhaltlich strukturierender und evaluativer Inhaltsanalysen (Kuckartz & Rädiker, 2022) ausgewertet werden. Es zeigen sich starke Zusammenhänge zwischen der Kooperationszufriedenheit von Lehrkräften und der sozialen Herkunft der Eltern. Die Kooperationserfahrungen von Lehrkräften an Schulen mit eher wenig privilegierter Elternschaft werden häufiger negativ bewertet. In vertiefenden Fallanalysen kann nachgezeichnet werden, dass Erwartungen von Lehrkräften mit universalem Gültigkeitsanspruch in die Kooperationspraxis einfließen und dort auf unterschiedlich „passende“ Elternschaften treffen. In beiden Kontrastgruppen (Schulen mit gering vs. stark privilegierter Elternschaft) werden Lehrkräfte durch Abweichungen von diesen Normalitätserwartungen herausgefordert. „Schwererreichbarkeit“ kann als ein Symptom dieser fehlenden Passung gedeutet werden. Wenn die Kooperation mit ohnehin benachteiligten Elterngruppen systematisch seltener zustande kommt, sind ungleichheitsreproduzierende Effekte zu erwarten. Das Potenzial einer konstruktiven Kooperationsbeziehung wird dann zu einem weiteren Privileg, das benachteiligten Schüler:innen verwehrt bleibt (Bourdieu, 2005; Lareau, 1989).

Veränderbarkeit von elterlichen Überzeugungen und Handlungen in der Lernbegleitung – Effekte eines Kurztrainings für Eltern

Edith Niederbacher, Caroline Villiger

PHBern, Schweiz

Die Familie hat einen wesentlichen Einfluss auf das schulische Lernen von Schüler*innen. Neben strukturellen Merkmalen (z.B. soziale Herkunft, Erstsprache) spielen elterliche Überzeugungen und Handlungen in der elterlichen Lernbegleitung eine bedeutsame Rolle (Harju-Luukkainen et al. 2020; Tazouti & Jarlégan 2019). Studien zeigen, dass beispielsweise lernförderliche Überzeugungen von Eltern (z. B. growth mindset; Dweck, 2000) positiv mit den schulischen Leistungen ihrer Kinder

zusammenhängen (Matthes & Stöger 2022). Hinsichtlich der elterlichen Handlungen in der Lernbegleitung sind die Befunde widersprüchlich – was u. a. darauf zurückzuführen ist, dass die Effekte je nach Handlung (z. B. Autonomieunterstützung, Kontrolle, etc.) positiv oder negativ ausfallen (vgl. Dumont et al. 2012; Moroni et al., 2015; Wandeler et al. 2013). Insgesamt kann festgehalten werden, dass elterliche Überzeugungen und Handlungen wichtige Prädiktoren für das schulische Lernen von Kindern darstellen.

Frühere Forschungen zeigen, dass Handlungen von Eltern im Rahmen von Interventionen verändert werden können. Bisher standen im Rahmen der Trainingsprogramme vor allem familiäre Praktiken im Fokus (z.B. gemeinsames Lesen oder Hausaufgabenunterstützung; Villiger et al., 2010). Vor dem Hintergrund, dass aber auch elterliche Überzeugungen für Schulleistungen wirksam sind (Jeynes 2005), ist es lohnenswert, diese ebenfalls in eine lernförderliche Richtung zu verändern.

In der vorliegenden Studie wurde im Rahmen eines Forschungsprojekts ein Elterntraining entwickelt und als Abendveranstaltung durchgeführt. Inhalte des Kurz-Trainings waren Ansatzpunkte zur Unterstützung des (selbstgesteuerten) Lernens zuhause mit dem Fokus auf elterliche Überzeugungen (growth mindset, positiver Umgang mit Misserfolg; Haimovitz & Dweck 2016; Matthes & Stöger 2018) und lernförderliche Handlungen im familiären Kontext (Autonomieunterstützung, Strukturgebung, geringe Kontrolle; Pino-Pasternak & Whitebread 2010; Wild & Remy 2001).

Im Rahmen des Beitrags wird untersucht, inwiefern die elterlichen Überzeugungen (growth mindset und Umgang mit Misserfolg) und Handlungen (Kontrolle, Strukturgebung und Autonomieunterstützung) durch das Elterntraining verändert wurden.

Die Daten zur Prüfung der Fragestellungen stammen aus einer Stichprobe von N=757 Schüler*innen aus 40 Schulklassen (5. und 6. Klasse, Primarstufe) und deren Eltern aus einem deutschsprachigen Kanton der Schweiz. Die Daten wurden mittels Fragebogen und standardisierten Leistungstests zu drei Messzeitpunkten innerhalb eines Schuljahres erhoben (t1: 09/2021, t2: 12/2021+01/2022, t3: 05/2022). 14 Schulklassen wurden randomisiert für das Elterntraining ausgewählt. Der Elternabend fand zwischen t1 und t2 im Oktober/November 2021 statt. Von 252 eingeladenen Erziehungsberechtigten, nahmen 157 am Elternabend teil. Die Kontrollgruppe in den Analysen besteht aus n=505 Elternteilen. Eltern, die zum Elternabend eingeladen waren, jedoch nicht daran teilnahmen (n=95), wurden von den Analysen ausgeschlossen.

Die Ergebnisse von T-Tests zeigen, dass sich die Interventions- und Kontrollgruppe (mit einer aufgrund geringer Effektstärke vernachlässigbaren Ausnahme bei der Variable Strukturgebung) in allen Variablen zum 1. Messzeitpunkt nicht unterscheiden.

Zur Beantwortung der Fragestellung wurden Regressionsanalysen in Mplus mit latent modellierten Konstrukten, unter Berücksichtigung der Mehrebenenstruktur der Daten, berechnet. Es wurden folgende Kontrollvariablen in die Modelle integriert: die Überzeugung resp. Handlung der jeweiligen abhängigen Variable zu t1, das Ausbildungsniveau, die Erstsprache sowie das Geburtsland der Eltern.

Die Ergebnisse zeigen einen signifikanten Effekt des Elterntrainings auf die Veränderung des Kontrollverhaltens von t1 zu t3. Eltern, die am Elterntraining teilnahmen, berichteten rund 7 Monate nach dem Training signifikant weniger Kontrollverhalten in der Lernbegleitung. Die übrigen untersuchten Überzeugungen und Handlungen wurden durch das Training nicht verändert. Das Ergebnis ist gleichwohl bedeutsam, zumal insbesondere Kontrolle eine sehr verbreitete elterliche Handlung in der Lernbegleitung ist, die sich nachweislich negativ auf die Leistungsentwicklung der Kinder auswirkt. Ausserdem war das Elterntraining kurz (1,5-stündiger Elternabend) und der Effekt bestand noch mehrere Monate später.

Die Studie zeigt auf, wie im Rahmen eines Elterntrainings Überzeugungen und Handlung hinsichtlich der Lernbegleitung thematisiert werden können. Trainingsoptimierungen zur Erreichung höherer Wirksamkeit werden diskutiert. Implikationen für die Praxis werden vor dem Hintergrund des Tagungsthemas diskutiert.

S-67: Diagnosegeleitete Leseförderung in der und für die Schulpraxis

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 154

Chair der Sitzung: Sarah Bez

Chair der Sitzung: Karina Karst

Chair der Sitzung: Vanessa Völlinger

Diagnosegeleitete Leseförderung in der und für die Schulpraxis

Chair(s): **Sarah Bez** (Universität Tübingen/Pädagogische Hochschule Karlsruhe), **Karina Karst** (Universität Mannheim)

Diskutant*in(nen): **Vanessa Völlinger** (Pädagogische Hochschule Freiburg)

Aktuelle Studien wie die Internationale Grundschul-Leseuntersuchung (IGLU) zeigen eine große Anzahl an Schüler*innen mit geringen Lesekompetenzen (Frey et al., 2023). Gleichzeitig sind hinreichende Lesefähigkeiten wichtig für den Erfolg des schulischen Lernens sowie langfristig für gesellschaftliche Teilhabe, weshalb eine frühzeitige, diagnosegeleitete Förderung zentral scheint, um basale Kompetenzen zu sichern (SWK, 2022).

Entsprechend wurden und werden aktuell Programme zur Lesekompetenzförderung (weiter)entwickelt und evaluiert (Blumenthal et al., 2022), die systematisch, diagnosegeleitet, digital und evidenzbasiert sind: Sie zeichnen sich dadurch aus, dass sich an eine kontinuierliche formative Diagnostik der individuellen Lernvoraussetzungen und -entwicklungen der Schüler*innen eine adaptive, auf evidenzbasierten Methoden fußende Förderung anschließt, d.h. Diagnose und Förderung eng verknüpft sind (Förster et al. 2018). Zudem bietet die Digitalität der Programme bzw. Plattformen u.a. Entlastung für Lehrkräfte.

Gleichwohl zeigt sich, dass die Umsetzung solcher systematischen, diagnosegeleiteten und evidenzbasierten Programme zur Lesekompetenzförderung im Unterrichtsalltag bislang eher selten (Jude et al., 2023; Schilcher et al., 2022) und mit Herausforderungen für Lehrkräfte verbunden ist: Hier scheinen u.a. Unterstützungssysteme relevant, die es Lehrkräften z.B. erleichtern, die Ergebnisse der Diagnostik zu interpretieren und darauf basierend adaptiv zu fördern (Souvignier & Ehler, 2024).

Hier setzen die im Rahmen dieses Symposiums vorgestellten Programme bzw. die entsprechenden Forschungsprojekte zur diagnosegeleiteten Leseförderung in der Grundschule an. Insgesamt adressiert das Symposium die Frage der Wirksamkeit verschiedener Leseförderprogramme mittels längsschnittlicher Experimental- und Kontrollgruppendesigns sowohl entsprechende Aspekte der Professionalisierung von (angehenden) Lehrkräften. Dabei wird auch untersucht, welche Faktoren die Wirksamkeit begünstigen oder einschränken, z.B. welche Faktoren die Nützlichkeits einschätzungen von Lehrkräften beeinflussen, die u.a. zentral für die gelingende Umsetzbarkeit in der Praxis gelten (Souvignier & Ehler, 2024).

Beitrag 1 berichtet die Ergebnisse eines fachintegrierten Leseförderprogramms, welches auf Lesestrategien fokussiert und von vorab geschulten Lehrkräften im unterrichtlichen Alltag zum Einsatz kommt. Gleichzeitig wird geprüft, inwiefern die Wirksamkeit dieses systematischen Trainings von lernrelevanten Merkmalen der Schüler*innen beeinflusst wird.

Beitrag 2 überprüft die Wirksamkeit eines diagnosegeleiteten Leseförderprogramms zur Leseflüssigkeit, welches durch geschulte Lehramtsstudierende im Rahmen von Service-Learning-Seminaren in Schulen durchgeführt wird.

Beitrag 3 liefert auf der Basis einer längsschnittlichen Befragung von Lehrkräften Erkenntnisse zur Nutzung digitaler Lernverlaufsdagnostik mit Fördermaterialien in der Praxis und damit wichtige Ergebnisse für den gelingenden Transfer solcher Programme in die Schulpraxis.

Das Symposium wird insbesondere mit Blick auf die schulische Praxis diskutiert. Fragen des Erkenntnistransfers aus Wissenschaft in die Praxis und vice versa stehen ebenso im Vordergrund wie die Frage der nachhaltigen Umsetzbarkeit und Adaptierbarkeit diagnosegeleiteter Leseförderprogramme an Grundschulen mit einer großen Anzahl an Schüler*innen mit Leseförderbedarf.

Beiträge des Symposiums

Lesestrategien in der Grundschule. Ergebnisse aus der longitudinalen post-Covid-Erhebung zur Fachintegrierten Leseförderung Bayern (FiLBY) in den Jahrgangsstufen 3-4

Johannes Wild¹, Anita Schilcher¹, Sven Hilbert¹, Laura Simböck¹, Elisabeth Kraus²

¹Universität Regensburg, ²LMU München

Theoretischer Hintergrund

Obgleich die deutschen Schüler:innen in der Sekundarstufe international das höchste Lesestrategiewissen besitzen (Diedrich et al., 2019), ist der Anteil schwach Lesender sowohl in der Primar- als auch in der Sekundarstufe hoch (McElvany et al., 2023; Weis et al., 2019). Strategien werden offensichtlich zwar vermittelt, sind aber vielfach nicht als Handlungswissen verfügbar. Die Integration von Lesestrategien in den Schulalltag wird von Pearson und Cervetti (2017) deshalb als Achillesferse von Lesestrategieprogrammen bezeichnet. Bei FiLBY steht die systematische Vermittlung von Lesestrategien in der dritten und vierten Jahrgangsstufe im Fokus. Das Training wurde von geschulten Lehrkräften durchgeführt, die zudem auf Basis standardisierter Lesetests regelmäßig Rückmeldung zur Leistungsentwicklung ihrer Schüler:innen erhielten. Die Durchführung der FiLBY-Studie wurde allerdings durch die Schulschließungen während der Coronapandemie unterbrochen. Dass das FiLBY-Training auch unter den Bedingungen des Distanzunterrichts wirksam sein kann, zeigen bereits Wild et al. (2022). Befunde zum regulären Schulbetrieb stehen noch aus und sollen in der vorliegenden Studie untersucht werden.

Fragestellung

Es wird untersucht, wie sich ein systematisches Leseförderprogramm, das mehrere Jahrgangsstufen in Primarschulen umfasst, auf die Lesekompetenzentwicklung von Schüler:innen auswirkt.

- Sind die FiLBY-Trainings wirksamer als eine Förderung des Leseverstehens im Rahmen des regulären Deutschunterrichts?
- Wie entwickelt sich das Leseverstehen beider Gruppen in Abhängigkeit ausgewählter Personenmerkmale (Geschlecht, Migrationshintergrund, Bildungshintergrund)?
- Welche Fortschritte erzielen Schüler:innen beider Gruppen in Abhängigkeit von ihrem Ausgangsniveau in der Leseflüssigkeit und Leseverstehen?

Methode

In den Schuljahren 2021-2023 konnte FiLBY im Präsenzunterricht durchgeführt und Daten im regulären Schulbetrieb erhoben werden. An dieser post Covid Evaluation nehmen insgesamt ca. 1 300 bayerische Schüler:innen der dritten und vierten Klasse

teil. Davon entfallen 1 047 auf die Experimentalgruppe. Die demografischen Daten wurden mittels Fragebogen erhoben, das Leseverstehen mit dem standardisierten Leseverstehenstest BYLET (Kraus et al., 2021). Die statistischen Analysen wurden mit der Statistiksoftware R (R Core Team, 2020) durchgeführt. Schüler:innen mit fehlenden Werten zu mindestens einem Messzeitpunkt wurden von den Analysen ausgeschlossen. Zur Beantwortung der vorliegenden Forschungsfragen wurden hierarchisch lineare Regressionsmodelle geschätzt.

Ergebnisse und ihre Bedeutung

Die statistischen Analysen zeigen, dass das FiLBY-Programm auch unter Feldbedingungen wirksam ist. Beide Trainings, FiLBY-3 ($b = 6.42$, $p < .001$) und FiLBY-4 ($b = 3.44$, $p < .05$), weisen signifikante Effekte zugunsten der Experimentalgruppe auf. Differentielle Analysen in der dritten Jahrgangsstufe zeigen zudem, dass die Zuwächse der Kinder in FiLBY-3 unabhängig von Migrationshintergrund, Geschlecht und Bildungsniveau vorliegen ($b = 5.15$, $p < .001$) und auch unter Kontrolle des Ausgangsniveaus in der Leseflüssigkeit und des Leseverstehens stabil sind ($b = 3.86$, $p < .01$). Das FiLBY-4-Training zeigt ähnliche Effekte, die ebenfalls unter Kontrolle von Migrationshintergrund, Geschlecht und Bildungsniveau bestehen bleiben ($b = 4.20$, $p < .05$). Eine differenzierte Analyse der Ausgangsleistung der FiLBY-Kinder in der vierten Jahrgangsstufe weist außerdem darauf hin, dass im Vergleich zu den leistungsstarken besonders die leistungsschwachen Schüler:innen im Leseverstehen profitieren ($b_{25\%} = 20.02$, $p < .001$; $b_{50\%} = 11.78$, $p < .001$; $b_{75\%} = 9.46$, $p < .001$). Dieser Befund bleibt auch bei Kontrolle des Ausgangsniveaus in der Leseflüssigkeit erhalten.

Diese Studie ist Teil eines präregistrierten Projekts: https://osf.io/qjpxw/?view_only=a7cba20de41941ba806d2f4782c695ae

Individuelle Leseförderung durch Lehramtsstudierende auf der Basis von Lernverlaufsdiagnostik

Sarah Bez¹, Mădălina Paizan², Karina Karst², Samuel Merk³

¹Universität Tübingen/Pädagogische Hochschule Karlsruhe, ²Universität Mannheim, ³Pädagogische Hochschule Karlsruhe

Theoretischer Hintergrund

Hinreichend ausgeprägte Lesekompetenzen gelten als eine zentrale Voraussetzung für Bildungserfolg und gesellschaftliche Teilhabe, wobei einer frühen inner- und ggf. auch außerunterrichtlichen Förderung der Leseflüssigkeit bei Kindern mit Schwierigkeiten im Leseerwerb eine wichtige Bedeutung zugeschrieben wird (SWK, 2022). Hierbei gelten digitale Lernverlaufsdiagnostik und darauf basierende adaptive Förderung als effizient (z.B. Förster et al., 2018): Denn Lernverlaufsdiagnostik bietet eine valide und sensitive Erfassung (von Veränderungen) der Lesefähigkeiten, was formative Rückmeldungen sowie die adaptive Förderung auf der Basis der individuellen Bezugsnorm ermöglicht und erleichtert. Dennoch scheint die Anwendung und Implementation in der schulischen Praxis voraussetzungsreich, da bspw. die adäquate Interpretation von Lernverlaufsdiagnostik nicht unbedingt vorausgesetzt werden kann (z.B. Espin et al., 2018), wissenschaftlich evaluierte digitale Plattformen für das Lesen in Deutschland zwar vorliegen, aber (noch) nicht breit in der Praxis verankert sind (Jude et al., 2023) und Lehrkräfte subjektiv die Durchführung und Implementation als zu aufwändig wahrnehmen können (Souvignier & Ehlert, 2024). Neben einer gezielten Leseförderung im Unterricht durch Lehrkräfte scheint es daher auch relevant zu sein, inwiefern eine gezielte, zusätzliche außerunterrichtliche Förderung, z.B. durch Lehramtsstudierende, wirksam sein kann.

Fragestellung

Vor diesem Hintergrund untersucht die vorliegende Studie, inwiefern eine intensive, auf Lernverlaufsdiagnostik basierende Leseförderung durchgeführt von geschulten Lehramtsstudierenden die Lesekompetenzen von Schüler:innen in der Grundschule mit Förderbedarf im Lesen beeinflusst.

Methode

An der Studie nahmen insgesamt $N = 449$ Schüler:innen aus $N = 25$ Klassen aus den Klassenstufen 2-4 teil. Die diagnosegeleitete Leseförderung erhielten $N = 101$ Schüler:innen. In der Studie wurden zunächst die Lesekompetenzen der Schüler:innen mithilfe des ELFE II standardisiert erfasst und auf Basis der Eingangsmessung (Prätest) diejenigen Schüler:innen pro Klasse mit dem höchsten Förderbedarf ausgewählt. In einem quasiexperimentellen Prä-Post-Design führten Lehramtsstudierende über ein Schuljahr hinweg regelmäßig Lernverlaufsdiagnostik und darauf basierende Leseförderung mit der Plattform Levumi (www.levumi.de) als außerunterrichtliche Förderung in Kleingruppensettings durch. Dabei lag der Schlüssel zwischen einem Studierendentandem und der Anzahl geförderter Kinder in einer Kleingruppe zwischen 2:4 bis 2:6, um eine (möglichst) individuelle Förderung zu gewährleisten. Die Studierenden wurden durch ein Hochschulseminar, konzipiert als Service Learning (Gerholz, 2020), geschult und begleitet. In diesen Service-Learning-Seminaren wurden fachlich-theoretische Inhalte aus dem Studium (Learning), in diesem Fall konzeptuelle Grundlagen zur diagnosegeleiteten Leseförderung, mit der praktischen Anwendung in gemeinnützigen Organisationen (Service), in diesem Fall die Durchführung der diagnosegeleiteten Leseförderung in Grundschulen, verknüpft.

Ergebnisse und ihre Bedeutung

Aufgrund großer Kurtosis der abhängigen Variablen wurden sowohl für die Modellierung des Gesamttestergebnisses (Leseverständnisleistung) als auch der Ergebnisse der Subtests (Lesegeschwindigkeit und Lesegenauigkeit jeweils auf Wort-, Satz- und Textebene) Bayesian Generalized Linear Regression Models (Gelman & Hill, 2007) für t-verteilte abhängige Variablen mit Fixed Effects für die Klassen spezifiziert. Unter Annahme einer Region of Practical Equivalence (ROPE) von $-0.1 \leq d \leq 0.1$ liegen ausschließlich inkonklusive Ergebnisse vor, da 95% Highest Posterior Density Intervals (HDPI) und ROPE jeweils eine Schnittmenge aufweisen (Kruschke & Liddell, 2018). Damit liegt weder Evidenz dafür vor, dass die Intervention wirksam ist, noch dafür, dass sie nicht wirksam ist.

Digitale Lernverlaufsdiagnostik in der schulischen Praxis: Erkenntnisse zur Nutzung durch Lehrkräfte im Bereich Lesen

Stefan Blank¹, Maren Specker¹, Natalie Förster², Elmar Souvignier³, Julia Blank¹, Benjamin Fauth¹

¹Institut für Bildungsanalysen Baden-Württemberg, ²Universität Wuppertal, ³Universität Münster

Theoretischer Hintergrund

Verfahren der Lernverlaufsdiagnostik (LVD) liefern Lehrkräften hilfreiche Informationen, um Schüler:innen individuell zu fördern (vgl. Blumenthal et al. 2022). Die Nutzung von LVD hat positive Effekte auf den Lernerfolg (Förster et al., 2018; Förster & Souvignier, 2014; 2015; Schmitterer et al., 2023; Souvignier & Förster, 2011). Eine zentrale Rolle kommt dabei der Lehrkraft zu, die im Sinne datengestützter Entscheidungsprozesse (data-based decision-making; Mandinach, 2012) die diagnostischen Informationen nutzen muss, damit die LVD wirksam werden kann (Hebbecke et al., 2022).

Es stellt sich die Frage, wie Lehrkräfte die Nützlichkeit und Praktikabilität der LVD einschätzen, inwiefern deren Ergebnisse zur Unterrichtsplanung und Anpassung von Förderentscheidungen genutzt werden und welche Faktoren die Einstellungen sowie die Nutzung beeinflussen. Bisher liegen hierzu wenige Erkenntnisse vor (z.B. Förster & Souvignier, 2015; Salaschek & Souvignier, 2013), es gibt aber Hinweise, dass die thematische Fortbildung von Lehrkräften vielversprechend ist (Förster & Souvignier, 2015; Hebbecker et al., 2022).

Im Zuge der Implementation der digitalen LVD quop in Baden-Württemberg werden an einer substantiellen Stichprobe von Lehrkräften zu 2 Messzeitpunkten folgende Fragestellungen untersucht:

Fragestellungen

1. Wie schätzen Lehrkräfte a.) die fachdidaktische Passung, b.) die Wirksamkeit für den Lernerfolg, sowie c.) das Kosten/Nutzen-Verhältnis der LVD quop ein?
2. Wie werden die Ergebnisse der quop-Assessments für Unterrichtsplanung und Förderentscheidungen genutzt?
3. Welche Faktoren tragen dazu bei, dass...
 - a.) die diagnostischen Informationen im Unterrichtsaltag genutzt werden?
 - b.) die LVD quop als förderlich für die Lernentwicklung wahrgenommen wird?
4. Faktor Zeit: Welche Veränderungen in der Nutzung und Einschätzung zeigen sich zwischen erster und zweiter Erhebung?

Methode

Die erste Erhebung fand im Sommer 2023 statt. Für den Kompetenzbereich Lesen nahmen 357 Lehrkräfte teil (Berufserfahrung $M = 17.1$ Jahre). Die zweite Erhebung startete im Juli 2024. Skalen zur fachdidaktischen Passung, wahrgenommenen Wirksamkeit für den Lernerfolg, Kosten-Nutzen-Verhältnis und Nutzung von quop wurden mit einem Online-Fragebogen erfasst. Als potenziell relevante Faktoren für die Nutzung und Einschätzung von quop wurden die Bewertung des Unterstützungssystems miterhoben. Alle Items wurden auf einer vierstufigen Likertskala eingeschätzt. Es wurden (teils angepasste) etablierte Instrumente aus vorliegenden Studien genutzt (Garet et al. 2001; Chafouleas et al. 2009; Hertel et al. 2014; Ditton, H. & Merz, D. 2013). Die Reliabilität der Skalen war hinreichend hoch ($\alpha > .70$).

Ergebnisse erste Erhebung

1. quop deckt die zentralen Kompetenzen für den Lese-Unterricht aus Sicht der Lehrkräfte gut ab ($M = 3.2$; $SD = 0.7$). Im Mittel stimmen sie einer positiven Lern-/Leistungsentwicklung zu ($M = 2.6$; $SD = 0.8$). Das Kosten-Nutzen-Verhältnis wird positiv eingeschätzt ($M = 3.3$; $SD = 0.7$).
2. Die Lehrkräfte rezipieren die Ergebnisse regelmäßig ($M = 3.6$; $SD = 0.7$), ziehen diese für Förderentscheidungen heran ($M = 3.1$; $SD = 0.9$) und nutzen sie für Gespräche mit Eltern ($M = 3.1$; $SD = 0.9$). Weniger häufig werden die Ergebnisse für die Anpassung des Unterrichts der gesamten Klasse ($M = 2.48$; $SD = 0.9$) sowie für die Lerngruppenbildung genutzt ($M = 2.7$; $SD = 1.1$).
3. a.) Die Regressionsanalyse zur Vorhersage der Nutzung von quop im Unterrichtsaltag zeigt das Unterstützungssystem als Prädiktor ($R^2 = 0.102$; $p < 0.001$). Die Multivariate Regressionsanalyse zeigt, neben Innovationsbereitschaft, Beteiligungsmöglichkeiten und Klassenstufe, ebenfalls das Unterstützungssystem als stärksten Prädiktor ($\beta = 0.342$; $t(105) = 3.262$; $p = 0.001$; $R^2 = 0.14$). b.) Gleiches gilt für die wahrgenommene Wirkung auf den Lernerfolg ($\beta = 0.332$; $t(91) = 2.866$; $p = 0.05$; $R^2 = 0.149$).

Die Ergebnisse korrespondieren mit bisherigen Befunden, liefern jedoch weitere Hinweise zur Nützlichkeit und Passung sowie zum Einfluss des Unterstützungssystems.

Zweite Erhebung

Die zweite Erhebung läuft noch bis Ende Juli 2024. Aus der längsschnittlichen Beobachtung erwarten wir Hinweise, ob sich die Ergebnisse stabilisieren oder andere Tendenzen zu beobachten sind.

S-68: Übergänge im Jugendalter und der (unsichere) Blick auf die Zukunft: Berufsorientierung und Zukunftsangst von Jugendlichen in Deutschland und Österreich

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 148

Chair der Sitzung: Lisa Hasenbein

Chair der Sitzung: Stefan Hofherr

Übergänge im Jugendalter und der (unsichere) Blick auf die Zukunft: Berufsorientierung und Zukunftsangst von Jugendlichen in Deutschland und Österreich

Chair(s): **Lisa Hasenbein** (Deutsches Jugendinstitut), **Stefan Hofherr** (Deutsches Jugendinstitut)

Das Erwachsenwerden ist geprägt durch eine Vielzahl von Herausforderungen, vor denen junge Menschen stehen. Der 15. Kinder- und Jugendbericht der Bundesregierung (Deutscher Bundestag, 2017) benennt drei zentrale Kernherausforderungen der Lebensphase Jugend: Qualifizierung, Verselbstständigung und Selbstpositionierung. Der Erwerb von Bildungszertifikaten und der Übergang von der Schule in Ausbildung und Studium sowie der Einstieg in den Beruf sind damit zentrale Anforderungen an junge Menschen. Darüber hinaus zeichnet sich das Jugendalter dadurch aus, zunehmend eigenverantwortlich zu handeln und eigene Wertvorstellungen sowie ein politisches Bewusstsein zu entwickeln (Berngruber & Gaupp, 2022).

Das Symposium nimmt diese gesellschaftlich relevanten Kernherausforderungen des Jugendalters aus unterschiedlichen disziplinären Perspektiven in den Blick und geht der gemeinsamen Frage nach, wie junge Menschen diese alterstypischen Anforderungen bewältigen bzw. welche individuellen und kontextuellen Faktoren zu einer mehr oder weniger erfolgreichen Bewältigung führen. Die vier Beiträge fokussieren einzelne Aspekte der gemeinsamen Fragestellung aus einer bildungs- und erziehungswissenschaftlichen sowie einer soziologischen und allgemeineren sozialwissenschaftlichen Perspektive. Rahmgebend sind dabei das Berufswahlkompetenzmodell (Driesel-Lange et al., 2020) sowie das Verständnis des Jugendalters mit seinen Übergängen als biografisch vulnerable Phase, die mit Unsicherheiten und Ängsten einhergehen kann (v. Wyl et al., 2018). Die Beiträge stützen sich in der Beantwortung der jeweiligen Forschungsfragen auf einen quantitativen empirischen Forschungsansatz mittels längsschnittlicher sowie querschnittlicher Analysen verschiedener Datensätze.

Der erste Beitrag befasst sich mit der Entwicklung der Berufswahlkompetenz von Schüler:innen der einjährigen Polytechnischen Schule, die im österreichischen Bildungssystem mit Fokus auf Berufsorientierung und Ausbildungsvorbereitung als letztes Pflichtschuljahr vor dem Übergang in eine duale Ausbildung verankert ist. Im Beitrag werden aufbauend auf einer Validierung der Kurzversion des Fragebogens Berufswahlkompetenz (Dehne et al., 2020) an einer Längsschnittstichprobe die Zusammenhänge zwischen den Berufswahlkompetenzdimensionen und der Berufswahlklarheit von 2.433 Schüler:innen an insgesamt 48 Polytechnischen Schulen untersucht und der Frage nachgegangen, inwiefern sich die Berufswahlkompetenz und Berufswahlklarheit im Verlauf des untersuchten Schuljahrs verändern.

Der zweite Beitrag verfolgt ebenfalls einen längsschnittlichen Ansatz und nimmt die Veränderungen von nachschulischen Bildungs- und Berufsplänen im Laufe der Sekundarstufe sowie am anschließenden Übergang in den Blick. Der Beitrag wertet Daten aus drei Erhebungswellen des Surveys „Aufwachsen in Deutschland: Alltagswelten“ (AID:A 2019, 2021 und 2023; Kuger et al., 2021) aus, wobei alle Panelbefragten berücksichtigt werden, die an mindestens zwei der Erhebungszeitpunkte eine allgemeinbildende Schule besucht haben. Es wird untersucht, inwiefern sich die nachschulischen Pläne ändern, welche Planänderungen besonders häufig sind und welche soziodemographischen Merkmale mit Planänderungen assoziiert sind.

Der dritte Beitrag schließt an, indem er sich mit dem Übergang von Schule in Ausbildung und Beruf befasst. Mit einem Fokus auf junge Menschen aus nicht-gymnasialen Bildungsgängen untersucht der Beitrag insbesondere die Verunsicherungen, mit denen Übergangsentscheidungen und -erfahrungen für angehende Schulabsolvent:innen verbunden sind. Der Beitrag wertet Daten von 1.548 im Rahmen der DFG-finanzierten Panelstudie „Effekte von organisierten Freizeitaktivitäten auf den Übergang in die berufliche Ausbildung“ (Hemming & Tillmann, 2022) befragten Jugendlichen aus den Erhebungsjahren 2019/2020 sowie 2021 aus. Dabei werden sowohl diffuse als auch konkrete Verunsicherungen und Ängste bezogen auf die eigene Zukunft untersucht und der Frage nach individuellen sowie kontextuellen Einflussfaktoren nachgegangen.

Der vierte Beitrag greift das Thema der Zukunftsängste junger Menschen auf und weitet den Blick über das Thema der Berufsorientierung hinaus, indem das Vertrauen junger Menschen in politische Institutionen und die allgemeine politikbezogene Zukunftsangst näher untersucht werden. Mittels der Daten des Surveys „Aufwachsen in Deutschland: Alltagswelten“ (AID:A; Kuger et al., 2021) aus dem Jahr 2023 geht der Betrag der Frage nach, inwiefern eine mehr oder weniger pessimistische Einstellung Jugendlicher und junger Erwachsener ($n = 3.459$) eine Frage des Bildungshintergrunds ist bzw. welche anderen soziodemographischen Merkmale beeinflussen, wie positiv oder negativ junge Menschen ihre politikbezogene Zukunft bewerten.

Beiträge des Symposiums

Der Kurzfragebogen Berufswahlkompetenz – Ergebnisse einer längsschnittlichen Validierungsstudie an Polytechnischen Schulen in Österreich

Silvio Kaak¹, Julia Niederfriniger², Karin Heinrichs³, Bärbel Kracke¹

¹Friedrich-Schiller-Universität Jena, ²PH Niederösterreich, ³PH Oberösterreich

Im österreichischen Bildungssystem ist die einjährige Polytechnische Schule (PTS) mit Fokus auf Berufsorientierung und Ausbildungsvorbereitung als letztes Pflichtschuljahr im Übergang in eine duale Ausbildung verankert und intendiert, Lernende in der Gestaltung ihres weiteren Bildungs- und Berufsweges individuell zu unterstützen. Zur Wirksamkeit der an den PTSn etablierten Berufsorientierungsmaßnahmen gibt es bislang nur wenige empirische Befunde (Bauer & Kainzmeyer, 2017; Telsnig, 2020). Im Beitrag wird dieses Forschungsdesiderat adressiert und die Berufswahlkompetenz (Driesel-Lange et al., 2010) der PTS Schüler:innen im Längsschnitt über ein Schuljahr (drei Messzeitpunkte) mit der Kurzversion des Fragebogens Berufswahlkompetenz (Dehne et al., 2020) untersucht. Die vorgestellte Studie verfolgt darüber hinaus das Ziel, das in Deutschland entwickelte Diagnostikinstrument an einer österreichischen Längsschnittstichprobe zu validieren.

Der Beitrag geht folgenden Fragestellungen nach:

- FF1: Sind Faktorstruktur und Reliabilitäten der Kompetenzdimensionen und -facetten des Instruments an einer Stichprobe von PTS-Schüler:innen in Österreich im Sinne einer Kreuzvalidierung replizierbar? Ist die Modellstruktur stabil über die Zeit?
- FF2: Korreliert Berufswahlkompetenz mit dem Außenkriterium der Berufswahlklarheit?

- FF3: Wie entwickelt sich die Berufswahlkompetenz der PTS-Schüler:innen im Verlauf eines Schuljahres?

Der Stand der Berufswahlklarheit und Berufswahlkompetenz wurde an einer Stichprobe von 2.433 Schüler:innen an insgesamt 48 PTSn in Kärnten, Nieder- und Oberösterreich und Salzburg untersucht. Zu drei Messzeitpunkten wurden biographische Daten, die Berufswahlklarheit (Heinrichs & Wuttke, 2013) sowie die Berufswahlkompetenz (FBK-k) (Dehne et al., 2020; Lipowski et al., 2021) der Schüler:innen erfasst. Zur Überprüfung der Konstruktvalidität wurden konfirmatorische Faktorenanalysen (CFAs) durchgeführt und die Modellstabilität im Längsschnitt überprüft. Zur Bestimmung der Kriteriumsvalidität wurden Korrelationen zwischen der Berufswahlkompetenz und der Berufswahlklarheit berechnet.

Hinsichtlich der ersten Fragestellung zeigte ein Drei-Faktormodell 2. Ordnung im Rahmen der CFAs die beste Passung zu den Daten. Diese Faktorstruktur entspricht der theoretischen Modellierung des Berufswahlkompetenzmodells (Driesel-Lange et al., 2010). Die Modellfit Indizes bestätigen eine gute Modellpassung. Darüber hinaus erweist sich die Modellstruktur über die drei Messzeitpunkte hinweg als stabil (T1: CFI = .932, TLI = .924, RMSEA = .045, SRMR = .046; T2: CFI = .927, TLI = .919, RMSEA = .048, SRMR = .050; T3: CFI = .939, TLI = .932, RMSEA = .052, SRMR = .058).

Bezogen auf die zweite Fragestellung stützten Analysen zur Kriteriumsvalidität ebenfalls die Güte des Fragebogens Berufswahlkompetenz und zeigen erwartungskonforme Zusammenhänge zwischen den Berufswahlkompetenzdimensionen und der Berufswahlklarheit. Zwischen der Berufswahlklarheit und den Berufswahlkompetenzdimensionen Wissen ($r = 0,484$; $p < 0,001$) sowie Handlung ($r = 0,302$; $p < 0,001$) bestehen mittlere positive Zusammenhänge und zwischen Berufswahlklarheit und Motivation ein kleiner positiver Zusammenhang ($r = 0,044$; $p = 0,032$).

Erwartungskonform zeigen Schüler:innen der PTS insgesamt eine überdurchschnittliche Ausprägung der Berufswahlkompetenz, jedoch mit Differenzierung nach Dimensionen und Facetten. In Bezug auf Selbst- und Konzeptwissen sowie Betroffenheit zeigen sich sehr hohe Ausprägungen. Das berufsbezogene Explorationsverhalten ist im Vergleich dazu niedriger ausgeprägt und rangiert auf durchschnittlichem Niveau. Die Ausprägung der Berufswahlkompetenz zeigt sich mit Blick auf die dritte Fragestellung über den Zeitverlauf von einem Schuljahr als relativ stabil. Signifikante Zuwächse zeigen sich vor allem in der Dimension Wissen.

Insgesamt bestätigen die Ergebnisse die Konstrukt- und Kriteriumsvalidität des Kurzfragebogens Berufswahlkompetenz. Damit sind die Voraussetzungen für eine zuverlässige und zeitökonomische Messung der Berufswahlkompetenz bei Jugendlichen unter verschiedenen schulischen Rahmenbedingungen im deutschsprachigen Raum gegeben. Erwartungsgemäß zeigen die Schüler:innen an Polytechnischen Schulen, deren Fokus auf der Vorbereitung des Übergangs in eine duale Ausbildung liegt, eine hohe Ausprägung der Berufswahlkompetenz. Zudem ist eine deutliche Steigerung des berufswahlbezogenen Wissens beobachtbar. Auch die Berufswahlklarheit steigt über die Zeit deutlich an. Das gibt Hinweise darauf, dass die Schulart der PTS ihre intendierten Ziele erreicht.

Veränderungen nachschulischer Bildungs- und Berufspläne im Laufe der Sekundarstufe

Stefan Hofherr

Deutsches Jugendinstitut

Die Planung der eigenen beruflichen Zukunft stellt eine zentrale Entwicklungsaufgabe des Jugendalters dar. Das Berufswahlkompetenzmodell (Driesel-Lange et al., 2020) betrachtet diese Planung als einen langfristigen Prozess, welcher durch vier Phasen gekennzeichnet ist: beim Einstimmen sollen Jugendliche erkennen, dass die eigene berufliche Zukunft ein wichtiges Thema ist und sie auch bereit sein müssen, sich damit auseinanderzusetzen. Beim Erkunden sollen sie einerseits ihre eigenen Interessen, Präferenzen und Talente erkunden, andererseits aber auch Informationen über verschiedene Berufe und Ausbildungswege einholen. Beim Entscheiden können Jugendliche auf Grundlage dieser Informationen verschiedene Optionen abwägen und sich für einen ersten nachschulischen Plan entscheiden. Beim Erreichen beginnen Schüler:innen und Schüler diese Pläne umzusetzen, indem sie sich beispielsweise um Praktika und Ausbildungsstellen in ihrem Wunschberuf bewerben. Der Forschungsstand zeigt, dass diese Phasen aber nicht linear ablaufen, sondern die meisten Schüler:innen im Laufe der Sekundarstufe ursprünglich gefasste nachschulische Pläne wieder verwerfen und sich umorientieren (Wood & Lauterbach, 2013).

Der Beitrag geht folgenden Fragestellungen nach:

- Verändern sich die angestrebten nachschulischen Pläne von Jugendlichen im Laufe der Sekundarstufe?
- Welche Planänderungen treten am häufigsten auf?
- Welche soziodemografischen Merkmale sind mit einer Planänderung assoziiert?

Datengrundlage ist der Survey „Aufwachsen in Deutschland: Alltagswelten“ (AID:A) des Deutschen Jugendinstituts (Kuger et al., 2021). Berücksichtigt werden Panelbefragte, welche zu mindestens zwei Zeitpunkten der Erhebungswellen 2019, 2021 und 2023 eine allgemeinbildende Schule besucht hatten. Es werden im folgenden sechs nachschulische Pläne unterschieden: 1) weiterer Schulbesuch, 2) Studienaufnahme, 3) Ausbildungsaufnahme, 4) Arbeiten/Jobben/Praktikum, 5) sonstige Pläne (z.B. Freiwilligendienst, Work & Travel) und 6) Weiß nicht. Diese Pläne werden stark durch den aktuell besuchten Bildungsgang strukturiert (Bittmann, 2021). So setzt beispielsweise eine Studienaufnahme den Erwerb der (Fach-)Hochschulreife voraus. Aus diesem Grund werden Schüler:innen in gymnasialen (Besuch eines Gymnasiums oder des gymnasialen Zweiges an Gesamtschulen) und nicht-gymnasialen Bildungsgängen (Besuch einer Hauptschule, Realschule, Schule mit mehreren Bildungsgängen oder eines nicht-gymnasialen Zweigs an Gesamtschulen) unterschieden.

Erste Ergebnisse mit Daten aus den Jahren 2019 und 2021 ($n = 844$; 48,3% weiblich, 26,8% Schüler:innen mit Migrationshintergrund; die finalen Daten aus 2023 liegen voraussichtlich bis Ende 2024 vor), zeigen, dass 57,9% der Schüler:innen ihre nachschulischen Pläne zwischen 2019 und 2021 geändert haben (in nicht-gymnasialen Bildungsgängen zu 62,8% und in gymnasialen zu 55,5%). Die häufigste Änderung in nicht-gymnasialen Bildungsgängen war der Wechsel von einer geplanten Studienaufnahme auf einen weiteren Schulbesuch ($n = 16$) und in gymnasialen Bildungsgängen von der Studienaufnahme zu sonstigen Plänen ($n = 68$). Die häufige Abnahme des Studienwunsches in gymnasialen Bildungsgängen kann als eine Verringerung der sehr hohen Bildungsaspirationen im Laufe der Sekundarstufe interpretiert werden. In nicht-gymnasialen Bildungsgängen kann diese Abnahme eher als eine Anpassung an realistischere Anschlussoptionen bezeichnet werden (Astleithner et al., 2021). Logistische Regressionen mit multipel imputierten fehlenden Werten zeigen, dass Jungen eine tendenziell höhere ($OR = 0,263$; $p = 0,077$), Schüler:innen in gymnasialen Bildungsgängen ($OR = -0,383$; $p = 0,024$) und mit höherem standardisierten sozioökonomischen Status der Eltern (HISEI; Ganzeboom & Treiman, 2010) eine (tendenziell) geringere Chance ($OR = -0,146$; $p = 0,071$) aufweisen, ihren nachschulischen Plan innerhalb der Sekundarstufe zu verändern. Die Ergebnisse legen nahe, dass auch Jugendliche im selben Alter unterschiedlich weit gefestigt sind in ihren nachschulischen Plänen. Aufgrund dieser Heterogenität wird empfohlen, den Entwicklungsfortschritt im Berufsorientierungsprozess von Jugendlichen individuell zu messen und darauf angepasste Berufsorientierungsmaßnahmen anzubieten (Ohlmann, 2020).

ENTFÄLLT: Zukunftsbezogene Verunsicherung am Übergang Schule – Beruf unter Schulabsolvent:innen nicht-gymnasialer Bildungsgänge

Frank Tillmann¹, Karen Hemming²

¹Deutsches Jugendinstitut, ²Stadt Leipzig

Am Übergang von der Schule in Ausbildung und Beruf bewegen sich Jugendliche innerhalb einer biografisch vulnerablen Phase (v. Wyl et al., 2018). Hier drohen insbesondere Schulabsolvent:innen mit niedrigeren beziehungsweise fehlenden Bildungsabschlüssen zu scheitern. Zwar scheinen sich in den letzten Jahren die Rahmenbedingungen am Ausbildungs- und Erwerbsmarkt gerade für diese jungen Menschen positiv zu verändern, dennoch weisen Befunde der aktuellen Übergangsforschung am Deutschen Jugendinstitut (DJI) auf neue Herausforderungen hinsichtlich individueller Bewältigungsaufgaben hin – etwa die Digitalisierung oder die zunehmende Unübersichtlichkeit (Tillmann & Reißig, 2019), aber auch mit Blick auf institutionelle Unterstützungsleistungen – wie z.B. Tendenzen der Standardisierung und Verantwortisierung (Enggruber et al., 2023). Wie frühere Untersuchungen verdeutlichen, sind Übergangsentscheidungen zudem maßgeblich durch ein Ausräumen von Unabhängigkeits- und Sicherheitsbedürfnissen abhängig. Dabei ist anhand von Befragungszeitreihen aus Regionalstudien zu beobachten, dass immer mehr Jugendliche eine diffuse Besorgnis darüber äußern, nicht zu wissen, was später aus ihnen werden soll. Zudem spielen trotz der verbesserten Situation auf dem Ausbildungsstellen- und Erwerbsmarkt konkrete Ängste eine stärkere Rolle, später keinen Ausbildungs- oder Arbeitsplatz zu bekommen (Hemming et al., 2018; Reißig et al., 2018 sowie Fehser et al., 2024), was zunächst paradox erscheinen mag. Gleichzeitig treten aber mit Blick auf zurückliegende und aktuelle Krisen – wie Corona, Kriege und Klimawandel – Verunsicherungen durch äußere Rahmenbedingungen auf (Vornanen et al., 2009), deren Konsequenzen für die Jugendlichen subjektiv nur schwer abzuschätzen sind (Hemming & Hofmann-Lun, 2023).

Ausgehend von diesen Vorannahmen und Beobachtungen wendet sich der Beitrag folgenden Fragestellungen zu:

- Welche Erscheinungsformen von diffuser und konkreter Verunsicherung sind bei Jugendlichen aus nicht-gymnasialen Bildungsgängen zu beobachten?
- Welche individuellen Einflussfaktoren bzw. kontextuellen Randbedingungen können in Bezug auf ein hohes Maß an Verunsicherung identifiziert werden?
- Wie können Unterstützungssysteme so verändert werden, dass junge Menschen in ihrer Suche nach Übergangsentscheidungen und -anschlüssen besser orientiert und begleitet werden?

Für den Beitrag werden Daten der DFG-finanzierten Panelstudie „Effekte von organisierten Freizeitaktivitäten auf den Übergang in die berufliche Ausbildung“ (Hemming & Tillmann, 2022) aus den Erhebungsjahren Jahr 2019/2020 sowie 2021 herangezogen. Damit steht für die querschnittlichen Analysen ein Datensatz der Befragung von angehenden Schulabsolvent:innen aus Sachsen und Sachsen-Anhalt zur Verfügung (n = 1.548; 51,5% weiblich, 15,5% mit Migrationshintergrund). Des Weiteren kann für die Auswertungen die Ausstattung des Elternhauses mit kulturellem und ökonomischem Kapital einbezogen werden – bspw. hinsichtlich der Bildungsabschlüsse, der kulturellen Alltagspraktiken sowie der finanziellen Situation.

Die Ergebnisse verdeutlichen, dass das Ausmaß an konkreter Verunsicherung durch dasjenige an diffuser übertroffen wird. Zudem äußern sich vor allem diejenigen jungen Menschen verunsichert über ihre Zukunft, die sich auch objektiv mit riskanteren Übergangsverläufen konfrontiert sehen: diejenigen mit Migrationshintergrund, junge Frauen sowie diejenigen aus materiell schlechter gestellten Familien. Hierbei handelt es sich also um Faktoren, auf die sie selbst keinerlei Einfluss haben. Auf der Grundlage der gewonnenen Ergebnisse werden Hinweise für die Fachpraxis der Unterstützungsangebote für junge Menschen am Übergang in Ausbildung und Erwerbsarbeit abgeleitet.

Vertrauen in politische Institutionen und politikbezogene Zukunftsangst im Jugendalter – eine Frage des Bildungshintergrunds?

Lisa Hasenbein, Ina Weigelt

Deutsches Jugendinstitut

Aktuelle Jugendstudien zeichnen häufig ein pessimistisches Bild der Jugend (Schnetzer et al., 2024). Zwei wichtige Kenngrößen sind dabei das Vertrauen in politische Institutionen und politikbezogene Zukunftsangst. Studien zeigen einen Zusammenhang zwischen dem Vertrauen in politische Institutionen und äußeren Bedrohungen wie kriegerischen Auseinandersetzungen oder Gesundheitskrisen (Hrbková & Kudrnáč, 2024; Erhardt et al., 2021). Es wird angenommen, dass Angst vor zukünftigen Unsicherheiten und ungünstigen Veränderungen dazu führt, dass Menschen Stabilität und Unterstützung bei politischen Institutionen suchen (Erhardt et al., 2021). Dabei sind sowohl das Vertrauen in politische Institutionen als auch Zukunftsangst von individuellen Merkmalen abhängig. Studien zu den Folgen der Corona-Pandemie etwa konnten zeigen, dass Zukunftsangst vom Geschlecht und sozioökonomischen Status beeinflusst wird (Reitegger et al., 2023; Dadaczynski et al., 2022). Vertrauen in politische Institutionen entwickelt sich durch individuelle Erfahrungen in unterschiedlichen Sozialisationskontexten, beispielsweise innerhalb der Schule oder der Familie (Mayne & Hakhverdian, 2017), wobei insbesondere auch das individuelle politische Wissen eine Rolle spielt (Schnaudt, 2020). Während die unterschiedlichen Einflussfaktoren bekannt sind, bleiben differenzierte Analysen zum Zusammenhang vom Vertrauen in politische Institutionen, politikbezogener Zukunftsangst und individuellen Merkmalen – insbesondere bezogen auf die Altersgruppe der Jugendlichen und jungen Erwachsenen und die Rolle des Bildungshintergrunds – weitestgehend aus.

Vor diesem Hintergrund adressiert der vorliegende Beitrag die folgenden Fragestellungen:

- Welche soziodemographischen Merkmale (z.B. Geschlecht, Bildungshintergrund und sozioökonomischer Status) sind mit dem Vertrauen junger Menschen in politische Institutionen und ihrer politikbezogenen Zukunftsangst assoziiert?
- Inwiefern hängt das Vertrauen junger Menschen in politische Institutionen mit ihrer politikbezogenen Zukunftsangst zusammen?

Der Beitrag nutzt Daten des Surveys „Aufwachsen in Deutschland: Alltagswelten“ (AID:A; Kuger et al., 2021) aus dem Jahr 2023, wobei insbesondere die Stichprobe der Jugendlichen und jungen Erwachsenen zwischen 12 und 25 Jahren berücksichtigt wird (n = 3.459; 50,3% weiblich, 26,3% mit Migrationshintergrund). Zu den Befragten liegen umfassende Hintergrundinformationen vor (z.B. Bildungshintergrund, sozioökonomischer Status), die im Zusammenhang mit der politikbezogenen Zukunftsangst (erfragt mit der deutschen Übersetzung der sog. Dark Future Scale; Zaleski et al., 2019; Wieczorek & Hock, 2023) sowie mit dem politischen Vertrauen in die Bundesregierung und die EU (angelehnt an etablierte Erhebungen des politischen Institutionenvertrauens; GESIS, 2024) mittels multipler linearer Regressionsanalysen analysiert werden.

Erste Ergebnisse zeigen, dass die politikbezogene Zukunftsangst unter Jugendlichen deutlich ausgeprägt ist, wobei das Geschlecht und Alter der Befragten einen signifikanten Einfluss haben: Je älter die Jugendlichen, desto größer ihre Zukunftsangst

($\beta = 0,06$; $p < 0,001$). Außerdem äußern Mädchen und junge Frauen eine größere Zukunftsangst als Jungen und junge Männer ($\beta = 0,17$; $p = 0,031$). Der Bildungshintergrund, sozioökonomische Status und die individuelle Zuwanderungsgeschichte zeigen keine statistisch signifikanten Unterschiede bezogen auf die politikbezogene Zukunftsangst der jungen Menschen, beeinflussen jedoch maßgeblich das Vertrauen in politische Institutionen: Höhere (angestrebte) Bildungsabschlüsse gehen mit einem deutlich höheren Vertrauen in politische Institutionen einher ($\beta = 0,67$; $p < 0,001$ für einen Hochschulabschluss im Vergleich zu keinem oder Hauptschulabschluss). Dagegen äußern junge Menschen, die selbst im Ausland geboren sind, und diejenigen, die in mittel bis hoch deprivierten Haushalten aufwachsen, ein signifikant geringeres Vertrauen in politische Institutionen als junge Menschen ohne Migrationshintergrund ($\beta = 0,34$; $p = 0,005$) und Deprivation im eigenen Haushalt ($\beta = 0,51$; $p < 0,001$). Es zeigt sich über alle Befragten hinweg ein leicht negativer Zusammenhang zwischen dem Vertrauen in politische Institutionen und politikbezogener Zukunftsangst ($r = -0,15$; $p < 0,001$), was auf eine wechselseitige Verfestigung negativer bzw. positiver politischer Einstellungen gegenüber politischen Institutionen und der eigenen Zukunft unter jungen Menschen hindeutet. Der finale Datensatz liegt voraussichtlich bis Ende 2024 vor, so dass vertiefende Ergebnisse präsentiert werden können.

S-69: Methodische Innovationen bei der Konstruktion von Test- und Fragebogenverfahren

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 159

Chair der Sitzung: Andreas Frey

Chair der Sitzung: Johannes Hartig

Chair der Sitzung: Joachim Ulf Kröhne

Methodische Innovationen bei der Konstruktion von Test- und Fragebogenverfahren

Chair(s): **Andreas Frey** (Goethe-Universität Frankfurt am Main, Deutschland), **Johannes Hartig** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation), **Ulf Kroehne** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Für die empirische Bildungsforschung ist es elementar, dass die untersuchten Merkmale objektiv, präzise und valide gemessen werden. So ist sicherzustellen, dass die interessierenden Merkmale nicht etwa systematisch durch andere Variablen verzerrt werden, wie beispielsweise durch unzureichende Motivation bei der Messung maximaler Leistungsfähigkeit. Auch bei der Untersuchung kausaler Effekte von Interventionen oder Unterricht ist wichtig, dass die genutzten Instrumente etwaige Änderungen in der oder den Zielvariablen überhaupt abbilden können. Die Konstruktion angemessener Test- und Fragebogeninstrumente ist keine rein methodische Fragestellung. Vielmehr muss ein klares Verständnis dazu vorherrschen oder durch empirische Forschung geschaffen werden, welche Merkmale zu messen sind und in welchen Zusammenhängen Sie mit anderen Merkmalen stehen.

Dieses Symposium umfasst vier Beiträge zu aktuellen Forschungs- und Entwicklungsarbeiten dieses Bereichs. Zunächst wird die statistisch optimale Konstruktion von Testinstrumenten betrachtet, dann wie man unmotiviertes Verhalten während der Testung erkennen und darauf reagieren kann, gefolgt von wie man änderungssensitive Items konstruieren und erkennen kann und letztlich wie die Zusammenhangsstruktur der verwandten Konstrukte kreatives Denken, mathematische Kompetenz und allgemeine kognitive Begabung ist.

Im ersten Beitrag „eatATA: Ein R-Paket zur automatisierten Testzusammenstellung“ wird eine Methode zur computerbasierten Zusammenstellung von Instrumenten unter Nutzung der statistischen linearen Optimierung vorgestellt. Es wird ein R-Paket präsentiert, welches diese statistisch anspruchsvolle Methode auch für Anwenderinnen und Anwender aus der empirischen Bildungsforschung zugänglich macht.

Der zweite Beitrag „Messung von Unterrichtseffektivitätskriterien: Konstruktion instruktionssensitiver Testaufgaben“ widmet sich der Frage, wie Items konstruiert werden können, die sensitiv für Änderungen der mathematischen Kompetenz sind. In einer empirischen Studie werden verschiedene Aufgabenmerkmale untersucht und die gewonnenen Daten mit einem erklärendem längsschnittlichem Mehrebenen-IRT-Modell modelliert.

Im dritten Beitrag „Echtzeiterkennung von unmotiviertem Antwortverhalten in Fragebögen – Kann sofortiges Feedback das zukünftige Antwortverhalten beeinflussen?“ werden aktuelle fortgeschrittene Methoden der Analyse von Log-Daten verwendet, um automatisiertes Feedback bei schneller Bearbeitung von Fragebögen bei Large-Scale Assessments zu geben. Mit einem experimentellen Design wird in einer nationalen Ergänzung zu PISA 2022 überprüft, ob dieses Feedback zu einer Verlangsamung des Fragebogenausfüllens und zu verändertem Antwortverhalten führt.

Im vierten Beitrag „Gleich ist nicht gleich gleich: Kreatives Denken, mathematische Kompetenz und allgemeine kognitive Begabung in PISA 2022“ wird ebenfalls an der Stichprobe von PISA 2022 die Threshold-Hypothese für den Zusammenhang von mathematischer Kompetenz und Kreativität untersucht. Aufgrund dieser Hypothese ist zu erwarten, dass der Zusammenhang von mathematischer Kompetenz und Kreativität ab einer gewissen Ausprägung der mathematischen Kompetenz abflacht. Die Datenanalyse erfolgt mit segmentierter Regressionsanalyse.

Nach jedem Beitrag erfolgt eine Diskussion des jeweiligen Beitrags anhand von Leitfragen.

Beiträge des Symposiums

eatATA: Ein R-Paket zur automatisierten Testzusammenstellung

Benjamin Becker¹, Dries Debeer²

¹Institut zur Qualitätsentwicklung im Bildungswesen an der Humboldt Universität zu Berlin, ²Universität Ghent, Belgien

Theoretischer Hintergrund

Eine häufige Herausforderung in der empirischen Bildungsforschung und psychologischen Forschung besteht darin, Aufgaben aus einem bestehenden Aufgabenpool zu Testformen zusammenzustellen. Dieser Vorgang wird im Englischen als Test Assembly bezeichnet (van der Linden, 2005).

Insbesondere wenn mehrere Testformen erstellt werden, die psychometrischen und inhaltlichen Anforderungen (z.B. inhaltliche Abdeckung, durchschnittliche Schwierigkeit, Nutzung verschiedener Aufgabenformate) gerecht werden müssen, kann die resultierende Testzusammenstellung äußerst komplex werden. Händische Lösungen solcher komplexen kombinatorischen Optimierungsprobleme sind ineffizient, fehleranfällig und führen häufig zu suboptimalen Lösungen. Solche Testzusammenstellungen können jedoch als gemischt-ganzzahlige lineare Optimierungsprobleme betrachtet werden, für deren automatische Lösung es eine breite Vielzahl an Softwaretools gibt. Die Nutzung solcher Optimierungstools zur automatisierten Testzusammenstellung wird im Englischen als Automated Test Assembly bezeichnet (van der Linden, 2005).

Da die verfügbaren Softwaretools jedoch allgemeine Tools sind, ist ihre Anwendung auf psychometrische Testzusammenstellungen nicht intuitiv und selten explizit dokumentiert (Diao & van der Linden, 2011). Folglich sind diese Werkzeuge möglichen Anwender*innen der empirischen Bildungsforschung und Psychologie oft unbekannt und aufgrund langer Einarbeitungszeiten praktisch nicht zugänglich.

Fragestellung

Benötigt wird ein Softwaretool, das Forscher*innen und Anwender*innen der empirischen Bildungsforschung und Psychologie den Zugang zu Optimierungsprogrammen für Anwendungen der automatisierten Testzusammenstellung ermöglicht. Ein entsprechendes Softwaretool sollte frei verfügbar und ausführlich dokumentiert sein sowie die relevante Funktionalität für gängige Anwendungsfälle (z.B. Pilotierungsstudien, Large-Scale Assessments, lineare Tests, Multi-Stage Tests) abdecken.

Methode

Die Softwareumgebung R ist sowohl in der empirischen Bildungsforschung als auch der Psychologie weit verbreitet. R-Pakete sind ein gängiger Weg, um psychometrische Neuentwicklungen Anwender*innen und Forscher*innen zur Verfügung zu stellen.

Das R-Paket eatATA zielt darauf ab, Techniken der automatisierten Testzusammenstellung einer breiten Nutzendenschaft zugänglich zu machen. eatATA stellt dabei Zugangsmöglichkeiten zu verschiedenen Tools für die gemischt-ganzzahlige lineare Optimierungsprobleme (GLPK, IpSolve, Symphony, Gurobi) dar.

Das Paket deckt gängige Anforderungen im Rahmen von Testzusammenstellung ab (kategoriale und metrische Constraints, verschiedene Objective Functions) und ist ausführlich im Rahmen eines Tutorial-Papers dokumentiert (Becker et al., 2021). Im Rahmen des Tutorial Papers werden häufige Anwendungsfälle der automatisierten Testzusammenstellung thematisiert und durchgespielt.

Ergebnisse und ihre Bedeutung

In dieser Präsentation werden zuerst grundsätzliche Herangehensweisen der automatisierten Testzusammenstellung eingeführt. Darauf aufbauend wird die allgemeine Funktionalität des eatATA-Pakets erörtert und anhand eines praktischen Anwendungsfalls der empirischen Bildungsforschung demonstriert.

Messung von Unterrichtseffektivitätskriterien: Konstruktion instruktionssensitiver Testaufgaben

Alexander Naumann¹, Marit List², Stephanie Leininger³, Stephan Schönenberger⁴, Jan Hochweber³, Johannes Hartig²

¹Institut für Schulentwicklungsforschung (IFS), TU Dortmund, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ³Pädagogische Hochschule St. Gallen (PHSG), ⁴Pädagogische Hochschule Thurgau (PHTG)

Theoretischer Hintergrund

Schulische Leistungsmessungen dienen regelmäßig als ein zentrales Kriterium für die Beurteilung der Wirkungen und der Wirksamkeit von Unterricht (z.B. Klieme, 2019; Kultusministerkonferenz, 2006). Üblicherweise werden Unterrichtsmerkmale dann als wirksam angesehen, wenn sie mit höheren Testwerten der Schülerinnen und Schüler einhergehen, und entsprechend als unwirksam, wenn nicht. Allerdings setzt eine solche Vorgehensweise voraus, dass die eingesetzten Tests und Items auch prinzipiell dazu in der Lage sind, mögliche Effekte des Unterrichts zu erfassen, also instruktionssensitiv sind (z.B. Naumann, Musow, Aichele, Hochweber & Hartig, 2019; Popham, 2007).

Wird ein Test als Kriterium verwendet, dessen Instruktionssensitivität unbekannt ist, bleibt unklar, inwiefern ein Ausbleiben erwarteter Effekte auf die Unwirksamkeit des Unterrichts oder auf die Insensitivität des Tests gegenüber Unterricht zurückzuführen ist. Daher ist Instruktionssensitivität eine wesentliche Voraussetzung für die valide Nutzung von Testwerten als Kriterium für die Wirksamkeit des Unterrichts. Doch obwohl es fundierte und kohärente Verfahren zur Messung der Instruktionssensitivität bestehender Tests und Items gibt (Naumann, Hartig, & Hochweber, 2017; Naumann, Rieser, Musow, Hochweber, & Hartig, 2019), liegt bisher nur wenig Wissen darüber vor, wie gezielt neue instruktionssensitive Aufgaben konstruiert werden können.

Fragestellung

Daher zielt die vorliegende Studie auf die Untersuchung von Aufgabenmerkmalen ab, von denen erwartet werden kann, dass sie die Instruktionssensitivität erhöhen.

Methode

Zunächst wurden 58 Mathematik-Items zu linearen Funktionen und linearen Gleichungen konstruiert, bei denen vier Merkmale systematisch variiert wurden: (1) das Antwortformat, (2) ob ein Item einen Wechsel zwischen mathematischen Darstellungen erfordert, (3) die Art des geprüften Wissens und (4) die Anzahl der angesprochenen Grundvorstellungen. Im Schuljahr 2021/2022 wurden die Items dann 663 Schülerinnen und Schüler in 32 Klassen deutscher Gymnasien (9. Klasse) und schweizer Sekundarschulen (8. Klasse) in einem Prä-Post-Design vor und nach der Vermittlung der Inhalte im Unterricht vorgelegt. Die durchschnittliche Zeitspanne zwischen den Messungen betrug drei Monate.

Um den Zusammenhang von Aufgabenmerkmalen und Instruktionssensitivität zu prüfen, modifizierten wir das von Naumann und Kollegen (2017) vorgeschlagene längsschnittliche Mehrebenen-IRT-Modell (LMLIRT) zur Messung der globalen und differentiellen Sensitivität von Items. Da das LMLIRT-Modell rein deskriptiv ist, formulieren wir das Modell im Sinne eines erklärenden IRT-Modells mit Itemprädiktoren um (Wilson & De Boeck, 2004). Das heißt, wir kombinieren das LMLIRT-Modell mit den Merkmalen des linear-logistischen Testmodells (LLTM; Fischer, 1973). Bei dieser Erweiterung wird die Item-Schwierigkeitskomponente des LMLIRT-Modells durch eine lineare Kombination von Prädiktoren auf Item-Ebene und einem Residualterm ersetzt (LLTM+e; Janssen, Schepers, & Peres, 2004). In dem sich ergebenden LMLLLTMe ist der mittlere Effekt der Item-Merkmale über Klassen hinweg relevant für die Vorhersage der globalen Sensitivität, und die vorhergesagte differentielle Sensitivität eines Items kann als die Variation der vorhergesagten Schwierigkeit des Items über die Klassen hinweg definiert werden. Die Schätzung wurde in einem Bayesianischen Framework durchgeführt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse deuten darauf hin, dass die Aufgabenmerkmale Antwortformat, Art des geprüften Wissens und die Notwendigkeit, zwischen mathematischen Darstellungen zu wechseln, zur mittleren Veränderung der klassenspezifischen Schwierigkeit eines Items zwischen den Messzeitpunkten (globale Sensitivität) beitragen. Mit Ausnahme des Antwortformats scheinen die Aufgabenmerkmale die differentielle Sensitivität zu erhöhen. Während die Aufgabenmerkmale jedoch die globale Sensitivität gut vorhersagen, bleibt ein erheblicher Anteil der differentiellen Sensitivität der Items trotz statistisch bedeutsamer Merkmaleffekte unerklärt.

Zusammenfassend liefern die Ergebnisse Hinweise darauf, welche Aufgabenmerkmale bei der gezielten Konstruktion neuer instruktionssensitiver Mathematik-Testaufgaben von Vorteil sein können. Die Möglichkeit, gezielt instruktionssensitive Items zu konstruieren, trägt dazu bei, die Instruktionssensitivität von Tests zu gewährleisten, ohne umfangreiche Pretests durchzuführen oder sich auf bestehende Instrumente zu stützen, was ökonomisch vorteilhafter ist. Darüber hinaus wird das Wissen um die Konstruktion instruktionssensibler Items zur Stärkung von Validitätsargumenten in der Unterrichtseffektivitätsforschung beitragen.

Echtzeiterkennung von unmotiviertem Antwortverhalten in Fragebögen – Kann sofortiges Feedback das zukünftige Antwortverhalten beeinflussen?

Ulf Kroehne¹, Lothar Persic-Beck¹, Leonard Tetzlaff¹, Carolin Hahnel², Frank Goldhammer¹

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Ruhr-Universität Bochum

Theoretischer Hintergrund

Unmotivierte Antworten, die z. B. als schnelles Raten (bei kognitiven Tests als "Rapid Guessing" bezeichnet, vgl. z.B. Wise & Kong, 2005) oder als schnelles Antwortverhalten (in Fragebögen, „Rapid Responding“, als Teil des Phänomens „Careless and Insufficient Effort Responding“, C/IER) durch die Verwendung von Antwortzeiten identifiziert werden, sind eine bekannte Bedrohung der Validität (u.a. Wise, 2017). Aus der Literatur ist bekannt, dass unmotiviertes Antwortverhalten vor allem in low-stakes Assessments (Wise et al., 2009), bei männlichen Testteilnehmern oder Befragten (z.B. DeMars, Bashkov, & Socha, 2013) und mit zunehmender Item-Anzahl häufiger auftritt (Lindner et al., 2019). Eine spezifische Behandlung der schnellen und ungültigen Antworten, die auf Item-Ebene auf unmotiviertes Bearbeitungsverhalten zurückgehen, oder die Berücksichtigung des Antwortzeitaufwands als Prozessindikator („Response Time Effort“, RTE) auf Personen-Ebene, können die Datenqualität und die Validität von Messungen und Schlussfolgerungen jedoch nur indirekt verbessern, wenn diese Berücksichtigung als Teil der Datenaufbereitung oder Datenanalyse erfolgt.

Fragestellung

In dem vorliegenden Beitrag wird deshalb die Fragestellung untersucht, (1) wie eine Echtzeiterkennung von unmotiviertem Antwortverhalten anhand von Log-Daten bereits während der Datenerhebung möglich ist, und (2) ob eine unmittelbare Rückmeldung (Feedback) über das beobachtete unmotivierte Antwortverhalten zukünftiges Antwortverhalten beeinflusst.

Methode

Anhand eines experimentellen Designs mit Zwischensubjekt-Variation wird Feedback beim Verlassen einer Fragebogenseite, welches auf fehlende Antworten, monotones (i.S. von „Straightlining“) oder schnelles (i.S. von „Rapid Responding“) Antworten in einem computerisierten Fragebogen hinweist (Experimentalgruppe) verglichen mit einem Feedback, welches ausschließlich auf fehlende Antworten hinweist (Kontrollgruppe). Der computerisierte Fragebogen, welcher die für die Identifikation von Bearbeitungszeiten in Fragebögen mit mehreren Items pro Seite notwendigen Log-Daten in Echtzeit auswertete, wurde in einer Zusatzstudie zu PISA 2022 administriert (N=705). Die Position von zusammenhängenden Fragebogenseiten, die jeweils eine Skala enthielten, wurde mit Hilfe eines balancierten Designs mit 18 Heften ausgeglichen. Die Erkennung von unmotiviertem schnellem Antwortverhalten basierte auf der Methode der algorithmischen Auswertung von Log-Daten mit Finite-State Machines (Kroehne & Goldhammer, 2018) mit Hilfe der durchschnittlichen Antwortzeit („Average Answering Time“), d.h. der Zeit zwischen den Antwortänderungen bei der Bearbeitung von Fragebögen mit mehreren Fragen auf einer Seite. Von diesem Zeitmaß war aus vorherigen Analysen (Kroehne et al, in Druck) bekannt, dass es bei Vorliegen von schnellem Antwortverhalten eine bimodale Verteilung aufweist. Da für die verwendeten Fragebogeninhalte keine Bearbeitungszeitdaten vorlagen, wurde für die Erstellung des Feedbacks in der Experimentalbedingung eine konservative Zeitschwelle von 1.0 Sekunden für die Erkennung von schnellem Antwortverhalten gewählt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse bestätigen zunächst die erwartete bimodale Verteilung der durchschnittlichen Antwortzeit, d.h. in der Experimentalbedingung und in der Kontrollgruppe lassen sich zwei Antwortprozesse trennen. Sowohl für männliche als auch für weibliche 15-Jährige zeigt sich ein signifikanter Effekt des Feedbacks auf die durchschnittliche Antwortzeit und die Wahrscheinlichkeit schnelles Antwortverhalten zu zeigen (standardisierte Odds Ratio von 0.867 für Jungen und 0.734 für Mädchen). Zusätzlich zu den direkten Effekten der Mikrointervention, welche sich auch dann zeigen, wenn die Position statistisch mitberücksichtigt wird, berichten wir über weitere indirekte Effekte des Feedbacks auf die Datenqualität (Reliabilität, Differential Item Functioning und latente Korrelationen), und stellen anhand von deskriptiven Ergebnissen einer eingefügten in-situ Frage an die Testteilnehmer*innen dar, wie die identifizierten Antworten verwendet werden sollen.

Während die Echtzeiterkennung von unmotiviertem Antwortverhalten in Fragebögen mit der entwickelten Methode funktioniert, sind die Effektstärken des Feedbacks insgesamt niedrig. In der abschließenden Diskussion wird deshalb die Bedeutung der Ergebnisse für zukünftige computerisierte Befragungen dargestellt und im Kontext alternativer Ansätze verortet, um mögliche Vorteile für die Datenqualität und mögliche Nachteile, wie bspw. die Anpassung des oberflächlichen Verhaltens, abzumildern. Für die praktische Anwendung wird Bezug zu Arslan et al. (2021) hergestellt, um die Zeitschwelle für künftige Anwendungen vorab auch theoriegetrieben bestimmen zu können.

Gleich ist nicht gleich gleich: Kreatives Denken, mathematische Kompetenz und allgemeine kognitive Begabung in PISA 2022

Jennifer Diedrich, Clievin Selva

Zentrum für internationale Bildungsvergleichsstudien (ZIB), Technische Universität München, Deutschland

Theoretischer Hintergrund

„die“ Intelligenz und „die“ Kreativität zeigen sowohl in Einzelstudien als auch in Metaanalysen statistisch bedeutsame Zusammenhänge: Kim (2005) berichten eine mittlere Korrelation von $r = .17$, welche aber je nach gewähltem Kriterium für der Kreativität (Shapiro, 1970) variieren kann. Für akademische Leistungen ergaben sich Zusammenhänge um $r = .36$ (Ma, 2009) bzw. $r = .22$ (Gajda, Karwowski, & Beghetto, 2017). Metaanalytische Befunde generalisieren dabei auf mehreren Ebenen: Zum einen werden Studien aus verschiedenen Kulturen gleichermaßen gemischt, zum anderen gehen (erwachsene) Personen aller Fähigkeitsniveaus ein. Gemäß der Threshold Hypothese zeigen akademische Leistungen (Hochschulleistungen im Vergleich zu Grundbildung) und allgemeine kognitive Begabung allerdings nur bis zu einem Schwellenwert bedeutsame Zusammenhänge mit Kreativitätsmaßen: in einer segmented regression fanden Jauk, Benedek, Dunst, and Neubauer (2013) einen Threshold von $IQ = 120$ für Originalität kreativer Antworten. Karwowski et al. (2016) fanden eine Effektstärke von $d = .24$ für Intelligenz als notwendige Voraussetzung für Kreativität. Vergleichbare Analysen für Grundbildung, wie sich in Schulleistungsstudien erhoben wird, sind nicht bekannt.

In PISA 2022 wurde erstmals kreatives Denken als innovative Domäne erhoben (OECD, 2023), sodass hierfür erstmals ein internationaler Vergleich möglich ist. Die Hauptdomäne in der aktuellen Erhebungsrunde bildete die mathematische Kompetenz, sodass erstmals die Überprüfung der Threshold Hypothese für Grundbildung möglich ist. Zudem wurde in Deutschland die allgemeine kognitive Begabung (Wilhelm, Schroeders, & Schipolowski, 2014) erfasst, deren Daten bis zur Konferenz vorliegen werden.

Fragestellung

PISA 2022 erlaubt somit erstmals die Überprüfung der Threshold Hypothese bei Jugendlichen verschiedener Kulturen in Bezug auf ihre mathematische Kompetenz:

1.

a. Gibt es einen Schwellenwert der mathematischen Kompetenz (im Sinne der Threshold Hypothese) für das kreative Denken?

b. Ist dieser Schwellenwert in allen PISA-Teilnehmerstaaten vergleichbar?

Zusätzlich wird für die deutsche Stichprobe eine Analyse hinsichtlich der allgemeinen kognitiven Begabung möglich sein:

2. Gibt es einen Schwellenwert der allgemeinen kognitiven Begabung für das kreative Denken auch bei Jugendlichen in Deutschland?

Methode

Zunächst sollen mittels segmented regression (Mueggo, 2008) Schwellenwerte für die mathematische Kompetenz (international) sowie die allgemeine kognitive Begabung (in Deutschland) identifiziert werden. Anschließend sollen die Schwellenwerte der Grundbildung international verglichen werden.

Ergebnisse und ihre Bedeutung

Der Beitrag wird die Schwellenwerte der mathematischen Kompetenz sowie der allgemeinen kognitiven Begabung für das kreative Denken vorstellen und vergleichen. Dabei soll auch diskutiert werden, ob oder inwiefern es sich um methodische Artefakte wie Deckeneffekte handelt. Abschließend wird, je nach identifizierten Schwellenwerten, die Operationalisierung des kreativen Denkens in PISA basierend bestehender Literatur zur Threshold Hypothese (Jauk et al., 2013) klassifiziert: z.B. handelt es sich eher um ein Originalitäts- oder Flüssigkeitsmaß.

S-70: Erfassung musikbezogener Kompetenzen

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 242

Chair der Sitzung: Stefan Münzer

Chair der Sitzung: Johannes Hasselhorn

Chair der Sitzung: Andreas C. Lehmann

Erfassung musikbezogener Kompetenzen

Chair(s): **Münzer Stefan** (Universität Mannheim), **Johannes Hasselhorn** (Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland)

Diskutant*in(nen): **Andreas C. Lehmann** (Hochschule für Musik Würzburg)

Sowohl der schulische Musikunterricht als auch außerschulische musikalische Aktivitäten wie z. B. Gesangs- oder Instrumentalunterricht haben das Ziel, musikbezogenen Kompetenzerwerb zu ermöglichen (vgl. Hasselhorn & Knigge, 2018). Für den schulischen Bereich sind nur wenige Forschungsvorhaben durchgeführt worden, die zu allgemeingültigen Kompetenzmodellen nach dem Vorbild der empirischen Bildungsforschung (z.B. Hartig & Klieme, 2006) für das Schulfach Musik geführt haben (empirisch validierte Modelle: Jordan, 2014; Hasselhorn, 2015; Wolf, 2016; Ehninger, 2022; theoretische Modelle: Hechler & Spsychiger, 2016; Hasselhorn, 2022). Für den Instrumental- und Gesangsunterricht liegen bislang keine Modellierungen vor, die diesem Ansatz entsprechen. Hier wird als Zielkonstrukt traditionell eher auf Expertise zurückgegriffen (vgl. Hasselhorn & Knigge, 2018).

Da sich aufgrund des hohen Individualisierungsgrades in Musik Bildungs- und Unterrichtsziele im schulischen und außerschulischen Musikunterricht nur schwer kompetenzorientierungskonform verallgemeinern lassen, gibt es bislang noch kein umfassendes Rahmenmodell musikbezogener Kompetenzen (vgl. Hasselhorn & Knigge, 2018). Stattdessen ist in den letzten Jahren die Identifikation und Beforschung einzelner Kompetenzbereiche wieder verstärkt in den Fokus der musikpädagogischen Forschung gerückt. Hierbei werden für das Musiklernen wichtige Fähigkeiten und Fertigkeiten herausgegriffen, die sich sowohl auf Schülerinnen und Schüler als auch auf Lehrende beziehen. Vier dieser Forschungsvorhaben werden im interdisziplinären Symposium (Musikpädagogik, Bildungspsychologie) zusammengeführt. Ein Schwerpunkt liegt dabei auf der Erfassung der jeweiligen musikalischen Kompetenzen.

Der erste Beitrag nimmt mit Qualitätsvorstellungen von Klassensingen eine Grundlage für fachspezifische diagnostische Kompetenzen von Musiklehrkräften in den Blick. Es können anhand von Fragebogendaten deutliche Unterschiede in der Wahrnehmung der Relevanz von Kriterien gefunden werden, die auf konkretes didaktisches Handeln und damit auch auf den Lernerfolg von Schülerinnen und Schülern direkten Einfluss haben werden.

Im zweiten Beitrag wird das aus der Mathematikdidaktik bekannte Konstrukt der Grundvorstellungen auf das Fach Musik zu übertragen. Anhand von Schulbuchanalysen und Interviewdaten werden verschiedene mögliche Grundvorstellungen zum Inhaltsbereich Notenwerte herausgearbeitet und hinsichtlich ihrer möglichen Folgen für ein fachlich korrektes Handeln analysiert.

Der dritte Beitrag wendet sich Kompetenzen von Lernenden im Instrumentalspiel zu. Anhand von Experimentaldaten wird untersucht, ob die Verfügbarkeit von Chunks – Erkennen von Akkorden und Umkehrungen in der musikalischen Notenschrift – in Überprozessen von Amateuren am Klavier eine Rolle spielt. Die Ergebnisse deuten darauf hin, dass implizit gelernte, typische Akkordformen das Üben von Klavierstücken erleichtert. Die Ergebnisse lassen aber bislang noch offen, wie spezifisch sich die Verfügbarkeit von Chunks auswirkt.

Im vierten Beitrag stehen methodische Fragen im Vordergrund. Beatwahrnehmung wird seit Jahrzehnten häufig als Merkmal zur Erfassung von Musikalität, einer Art fachbezogenen Grundkompetenz, eingesetzt. Die Operationalisierung von Beatwahrnehmung kann aber unterschiedlich ausfallen. Daher werden hier in einem experimentellen Ansatz drei mögliche Operationalisierungen (tappen, klatschen, Fuß tappen) miteinander und mit den Ergebnissen eines Hörtests verglichen.

Die abschließende Diskussion ordnet die Beiträge u.A. hinsichtlich ihres Innovationspotenzials ein und diskutiert auch die Frage, ob sich aus solchen Ansätzen, die eher aus der Praxis heraus einzelne Fähigkeitsbereiche fokussieren, ein verallgemeinerbares Kompetenzrahmenmodell für das Fach Musik ableiten lässt.

Beiträge des Symposiums

Zielvorstellungen von Musiklehrkräften in Bezug auf Klassensingen

Patrick Ehrich¹, Gabriele Puffer²

¹Universität Regensburg, ²Hochschule für Musik und Theater München

Singen ist eines der ältesten Unterrichtsfächer und seit den Anfängen der Zivilisation nachweisbar (Nolte & Weyer, 2011), die damit verbundenen Funktionen und Zielsetzungen waren jedoch einem steten Wandel unterworfen (Gruhn, 2003). In der wissenschaftlichen Musikpädagogik hat Singen im Musikunterricht in jüngerer Zeit wieder verstärktes Forschungsinteresse auf sich gezogen (z. B. Blanchard, 2022; Günster, 2023; Imthurn, 2023; Puffer, 2022). Gleichzeitig sind „viele didaktische Fragen“ (Lehmann-Wermser, 2017, S. 90) in Bezug auf das schulische Singen nach wie vor ungeklärt.

Der vorliegende Beitrag befasst sich mit der Frage, welche Zielvorstellungen Musiklehrkräfte mit Blick auf das Singen im Klassenverband haben und in welcher Relation diese zu Positionen der Fachliteratur und der Bildungsadministration stehen. Dies ist von Relevanz, weil professionelle Überzeugungen von Lehrkräften als handlungsleitend gelten (z. B. Kunter et al., 2011) und eine entsprechende Wirksamkeit auf Unterrichtsqualität zu vermuten ist (z. B. Helmke, 2022).

Methoden

Zunächst wurden ein repräsentatives Korpus fachdidaktischer Schriften sowie curriculare Vorgaben für die Jahrgangsstufen 5 und 6 aller Bundesländer und Schularten mit Mitteln der Dokumentenanalyse (Hoffmann, 2018; Noetzel et al., 2018) auf Zielformulierungen zum Klassensingen untersucht. Auf Grundlage der dabei gewonnenen Systematik wurde ein Fragebogen entwickelt, um zu untersuchen, welchen Zielen von Klassensingen Musiklehrkräfte welche Relevanz und Realisierbarkeit zuerkennen. An der im Herbst 2023 durchgeführten Onlinebefragung nahmen N = 222 vorrangig aus Bayern stammende Musiklehrkräfte teil, von denen 87 % über einen Abschluss im Unterrichtsfach Musik oder für ein Lehramt an Gymnasien verfügen. Das durchschnittliche Dienstalter der Teilnehmer*innen liegt bei M = 16,11 (SD = 10,67) Jahren bei einer Spannweite von 45 Jahren.

In einem zweiteiligen Verfahren wurden diese zunächst gebeten, in Freitextfeldern die drei aus Ihrer Sicht wichtigsten Ziele von Klassensingen zu nennen. Die Auswertung dieser Daten erfolgte mittels quantitativer Inhaltsanalyse (Rössler, 2017).

Anschließend gaben die Lehrkräfte Relevanzeinschätzungen zu 31 Zielformulierungen in fünf Itemgruppen ab, die aus der Fachliteratur und curricularen Vorgaben abgeleitet sind. Die Skalenreliabilität dieser Itemgruppen ist nach Field et al. (2012) als akzeptabel bis gut einzustufen ($0,63 \leq \alpha \leq 0,82$). Anschließend wurden Zielformulierungen mit hoher bzw. geringer Relevanz mittels parametrischer und non-parametrischer Mittelwertvergleiche gegen das jeweilige empirische Mittel der Itemgruppe identifiziert.

Ergebnisse

Die Auswertung der Freitextantworten ergab, dass Lehrkräfte von sich aus u.a. besonders häufig Ziele mit Bezug zu Stimm- und Singkompetenz anführen, ein Befund, der sich mit der hohen Relevanz, die diesem Bereich in den curricularen Vorgaben zukommt, deckt. Beim Blick auf die geschlossenen Items zeigen sich allerdings interessante Differenzen. So bewerteten die Befragten beispielweise Zielformulierungen als besonders gut realisierbar, die den mittelfristigen sängerischen Kompetenzzuwachs der Klasse als Kollektiv betreffen ($p < 0,001$, $d = 1,18$), mit deutlichem Abstand gefolgt von Lernerträgen einzelner Schüler*innen ($p < 0,001$, $d = 0,44$). Gleichzeitig wurden Zielformulierungen, die konkrete, auf Ebene einzelner Lerneinheiten umzusetzende Aspekte wie rhythmisch und intonatorisch korrektes Singen in den Blick nehmen, als besonders wenig relevant eingeschätzt ($p < 0,001$, $d = -0,34$ bzw. $-0,27$).

Die sich hier zeigende Diskrepanz ist deshalb von Bedeutung, weil die von den Lehrkräften als wenig relevant eingestuft konkreten und kurzfristigen Ziele auf grundlegende Kompetenzen zielen, die in Fachliteratur und Curricula gemeinhin als Voraussetzung für den Kompetenzerwerb über längere Zeit und auf höherem Niveau angesehen werden. Das wirft unter anderem die Fragen auf, welche Vorstellungen von grundlegenden musikalischen Lernprozessen die Befragten haben, wie sich diese auf ihr Handeln im Unterricht auswirken und welche Implikationen sich daraus für die Musiklehrkräftebildung ergeben.

Grundvorstellungen in Musik

Katharina Zink, Johannes Hasselhorn

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Die Entwicklung von Modellen und Testverfahren zur Erfassung musikbezogener Kompetenzen hat in den letzten Jahren im Rahmen der Kompetenzorientierung immer mehr an Bedeutung gewonnen (vgl. Hasselhorn & Knigge, 2018). Während zwar einerseits immer besser erfasst werden kann, was die Schülerinnen und Schüler in Musik (wahrscheinlich) können sollten, bleibt andererseits nach wie vor die Frage weitgehend unbeantwortet, welche Konsequenzen daraus konkret für die Schulpraxis folgen.

Um in Unterrichtskontexten die häufig auftretende Differenz zwischen aktuellem Lernstand von Schüler:innen und Lern- und Unterrichtszielen zu überwinden, nutzen Lehrkräfte das Werkzeug der Pädagogischen Diagnostik (Hesse & Latzko, 2017; in Musik z.B. Hasselhorn, Platz & Harnischmacher, 2022). Dabei ist neben der reinen Leistungsfeststellung auch die Suche nach möglichen Gründen für erwartungswidrige Leistungsstände Teil ebendieser Diagnostik.

Um zu verstehen, warum ein bestimmtes Lernziel bislang nicht erreicht werden konnte, wird beispielsweise in der Mathematikdidaktik das Konzept der Grundvorstellungen herangezogen, wodurch Fehlvorstellungen und Missverständnisse von fachlichen Inhalten getrennt identifiziert werden können. Grundvorstellungen beschreiben dabei „Beziehungen zwischen mathematischen Inhalten und dem Phänomen der individuellen Begriffsbildung“ (vom Hofe, 1995) und verknüpfen somit fachliche Inhalte und deren Anwendungszusammenhang, sodass diese nicht beziehungslos gegenüberstehen (vom Hofe, 1996).

Ausgehend von der Annahme, dass das Musizieren (auch) auf einem gemeinsamen, bzw. kompatiblen Verständnis von musikalischen Parametern wie Metrum, Dynamik oder Tonhöhen besteht, soll daher in diesem Beitrag ein Blick über reine musikdidaktische Methoden oder auf solche Methoden fokussierte Konzepte hinaus der Frage nachgegangen werden, ob sich das in der Mathematikdidaktik bewährte Konzept der Grundvorstellungen auch für den Musikunterricht gewinnbringend anwenden lässt. Dabei ist zunächst entscheidend, ob spezifische, häufig auftretende und sich (teilweise) widersprechende Vorstellungen basaler musikbezogener Inhalte identifizieren lassen, aus denen dann in weiteren Schritten didaktische Konsequenzen abgeleitet werden können.

Die Beantwortung dieser Frage wurde zunächst auf den spezifischen Bereich der Notenwerte eingegrenzt. Bereits auf neuronaler Ebene können in diesem Bereich unterschiedliche Verarbeitungsprozesse identifiziert werden (Proksch et al., 2020). Um darüber hinaus für musikbezogenes Handeln potenziell unterschiedliche und relevante Grundvorstellungen in diesem Themenfeld zu identifizieren, wurde einerseits eine Schulbuchanalyse mit in Bayern zugelassenen Schulbüchern durchgeführt, um zu erfassen, auf welche unterschiedliche Arten Notenwerte in diesen Unterrichtsmaterialien dargestellt werden. Zum anderen wurden Interviews mit Lehrkräften verschiedener Schulformen durchgeführt, in denen Strategien thematisiert wurden, wie Notenwerte Schülerinnen und Schülern beigebracht werden können. In der Synthese beider Analysen bildeten sich vier wesentliche Kategorien heraus, denen sich alle bisher gefundenen Darstellungen und Erklärweisen zuordnen ließen: Notenwerte a) als Teil eines Ganzen b) im Verhältnis zueinander c) im Verhältnis zum Metrum und d) dargestellt als Stellvertretervorstellung.

Die Ergebnisse der Analysen werden hinsichtlich ihrer Bedeutung für musikdidaktisches Handeln diskutiert. Insbesondere die Frage, welche Grundvorstellungen werden im Musikunterricht ausgebildet werden sollten, spielt dabei eine wesentliche Rolle und kann durch diesen Beitrag vorangebracht werden.

The Role of Chunks in the Practicing Process of Piano-Playing Amateurs

Sarah Meinert, Stefan Münzer

Universität Mannheim

Cognitive chunking integrates information into patterns – so-called chunks – resulting in efficient memory storage and retrieval (Gobet et al., 2001). Chunks representing musical patterns (such as chords) facilitate reading musical notation (Lörch, 2021; Maturi & Sheridan, 2020; Perra et al., 2022; Sloboda, 2004). While chunking is predominantly studied in the expertise framework (Ericsson & Charness, 1994), it might be relevant for amateurs as well. The present study examines the availability of musical chunks in amateur piano players for practicing.

Chunk availability was assessed with a chord comparison task. Participants indicated as accurately and quickly as possible whether basic chords (e.g., C-E-G) and inversions (e.g., G-C-E) were the same irrespective of the inversion. Half of the comparisons were invalid (e.g., G-B-E). Some participants needed long reaction times and did not complete the fifty trials in the given time of eight minutes. The rate correct score (RCS; Vandierendonck, 2016) was chosen as a combined measure, representing the number of correct answers per second.

Participants practiced three versions of a short chord-based piano piece on a digital piano. There was an original and a transposed version. A varied version altered tonal relations within chords, resulting in unfamiliar-sounding harmonies. Each version existed

in three difficulty levels. The versions were practiced in counterbalanced order. Participants practiced a version as long as they wanted. Subsequently, two performances were recorded. Practicing times as well as tempo and mistakes were measured.

Thirty-three amateur piano players (students at a German university) with at least four years of lesson-based piano training participated (age $M = 23.15$, $SD = 3.68$; 22 female, 10 male, 1 divers). Participants were screened for simple musical notation reading skills and assigned a difficulty level. Five participants who were unable to practice the easiest level and one who could sight-read were excluded. Participants filled out the Gold-MSI (Gold-MSI percentiles $M = 84.56$, $SD = 19.04$, Min. = 5, Max. = 96). Chunk availability and Gold-MSI were related, $r = .44$, $p = .022$.

A repeated measures ANOVA suggested differences in practicing times between versions, $F(2,52) = 15.14$, $p < .001$. Planned contrasts revealed that times were longer for the varied version ($M = 764.00$ s ± 12 min 44 sec, $SD = 374.30$ s) than for the original version ($M = 511.91$ s ± 8 min 31 sec, $SD = 305.19$ s), $F(1,26) = 26.27$, $p < .001$. No such difference was found between the transposed ($M = 454.58$ s ± 7 min 35 sec, $SD = 285.53$ s) and original version. A corresponding result was found for self-chosen performance tempo (lower tempo specifically for the varied version). These results show that the varied version was considerably more difficult to practice. Analyses on mistakes are currently ongoing.

There were correlations between chunk availability and practicing. A substantial relation was found between chunk availability and practicing time for the varied version, $r = -.50$, $p = .008$, but not for the original ($r = -.33$, ns) or transposed version ($r = -.26$, ns). Significant correlations were found between chunk availability and performance tempo for the transposed, $r = .52$, $p = .006$, and varied versions, $r = .55$, $p = .003$, but not for the original piece ($r = .34$, ns). This pattern does not suggest that chunk availability was specifically beneficial for reading familiar chords only.

The chord comparison task may diagnose musical chunk availability. Relations with Gold-MSI and practicing measures suggest its validity. However, it might not be specific for reading familiar musical chords but may indicate other aspects of skilled piano playing as well. Research could further investigate beneficial effects of chunking for practicing of amateurs.

Musikalische Beatsynchronisation - Gibt es Unterschiede zwischen verschiedenen Operationalisierungen?

Miriam Knebusch, Andre Sept, Johannes Hasselhorn

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Die Fähigkeit, den Beat eines Musikstückes wahrzunehmen und sich dazu zu synchronisieren, gehört zu den grundlegenden musikalischen Fähigkeiten (Honing, 2012), die auch für das Schulfach Musik zentral sind (vgl. Jank, 2021). Zahlreiche Studien haben in diesem Zusammenhang Einflüsse der motorischen Komponente sowohl bei der Beatwahrnehmung als auch der –produktion untersucht (Proksch et al., 2020; Ross et al., 2016; Tranchant et al., 2021). Die spezifische Bewegung beim Mitmachen eines musikalischen Beats beeinflusst dabei auch die Synchronisationsgenauigkeit, abhängig von der motorischen Vorbildung der Versuchspersonen (Nguyen et al., 2022).

In den meisten Studien wird ein einfaches Fingertappen als Bewegung eingesetzt, ohne auf mögliche motorische Unterschiede einzugehen (Dalla Bella et al., 2017; Fujii & Schlaug, 2013; Iversen & Patel, 2008). Dieses Fingertappen stellt jedoch eine künstliche, im musikalischen Alltag nur sehr selten vorkommende Bewegung dar. Neben konkreten, instrumentenspezifischen Bewegungen oder auch Tanzbewegungen spielt Beatsynchronisation im individuellen Alltag hauptsächlich in Form von Klatschen oder Tappen mit dem Fuß eine Rolle.

Mithilfe der vorliegenden Studie soll untersucht werden, inwieweit Fingertappen geeignet ist um die Synchronisationsleistung zum Beat und die dahinterstehende kognitive Fähigkeit der Beatwahrnehmung im Vergleich mit ökologisch validen Bewegungen zu erfassen.

Dafür wurden drei Methoden der motorischen Synchronisation verglichen: Fingertappen, Klatschen als auch im Alltag vorkommende Form und Fußtippen als im Musikbereich häufig eingesetzte Methode, um sich den Beat bewusst zu machen. $N = 79$ Grundschullehramtstudierende ohne Studienfach Musik (Alter $M = 23.96$, $SD = 4.58$, 86.3 % weiblich) bearbeiteten die 12 kurzen Musikausschnitte des Beat-Alignment-Tests von Iversen und Patel (2008) in diesen drei Modi. Zusätzlich sollten sie zur Erfassung der motorisch nicht beeinflussten kognitiven Hörkomponente der Beatwahrnehmung hörend beurteilen, ob ein hinzugefügtes Klicken passend zum Beat des Stückes ist.

Die Datenanalyse läuft aktuell. Geplant ist eine varianzanalytische Auswertung (mit Messwiederholung) zum Vergleich der verschiedenen Modi, sowie eine korrelative Analyse der Zusammenhänge zwischen Tapping- und Hörtest. Erste Ergebnisse des Hörtests deuten auf größere Schwierigkeiten beim Erkennen von Phasenfehlern ($M = 63.75\%$ richtig erkannt) im Vergleich zu Tempofehlern ($M = 81.25\%$ richtig erkannt) hin.

Die Ergebnisse dieser Untersuchung können zu validen Beat-Alignment-Testverfahren beitragen und können durch genauere Einblicke in die Zusammenhänge zwischen Hörwahrnehmung und motorischer Synchronisation wichtige Grundlagen für die Entwicklung von musikpädagogischen Beatwahrnehmungstrainings liefern. Im Sinne eines kompetenzorientierten Unterrichts können wichtige Erkenntnisse für differenzierte Betrachtungen hinsichtlich motorischer Einschränkungen bei rhythmischen Leistungsproben gewonnen werden.

S-71: Einzelschulen im Kontext räumlicher und sozialer Ungleichheit: Die Rolle des schulischen Angebots und der elterlichen Schulwahl

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 151

Chair der Sitzung: Carolin Zink

Chair der Sitzung: Nicky Zunker

Chair der Sitzung: Stefan Immerfall

Einzelschulen im Kontext räumlicher und sozialer Ungleichheit: Die Rolle des schulischen Angebots und der elterlichen Schulwahl

Chair(s): **Carolin Zink** (Leibniz-Institut für Bildungsverläufe (LIfBi); Friedrich-Schiller-Universität Jena, Deutschland), **Nicky Zunker** (Humboldt Universität zu Berlin; Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF))

Diskutant*in(nen): **Stefan Immerfall** (Pädagogische Hochschule Schwäbisch Gmünd)

Seit fast 25 Jahren gilt es als stabiler Befund der Bildungsforschung, dass in Deutschland die Bildungschancen stark von der sozialen Herkunft abhängen (Lämmchen et al. 2022, Mang et al. 2023). Ein breiter Forschungsstrang zur Genese herkunftsbedingter Bildungsungleichheiten im deutschen Schulsystem richtet den Fokus vor allem auf Schulformunterschiede innerhalb der mehrgliedrigen Sekundarstufe I und untersucht elterliches Wahlverhalten am Übergang von der Grund- zur weiterführenden Schule (z.B. Dumont et al., 2014). Neben der vertikalen Differenzierung durch die Schulformgliederung, wäre gleichermaßen anzunehmen, dass auch einzelschulische Spezifika für die Entstehung von Bildungsoportunitäten relevant sind. Insbesondere vor dem Hintergrund einer erhöhten Schulautonomie wird in Deutschland diese zweite – horizontale – Differenzierungslinie zwischen Einzelschulen sichtbar, die auch der elterlichen Einzelschulwahl eine größere Bedeutsamkeit verleiht. Das Symposium widmet sich diesem bisher noch wenig erforschten Gegenstand und untersucht Unterschiede auf Ebene der Einzelschule im Kontext von Bildungsungleichheiten. Dabei werden in Anschluss an Mayer und Koinzer (2019) nicht nur Eltern und deren Bildungsentscheidungen in den Mittelpunkt gestellt, sondern ebenso die Rolle von Bildungspolitik und Sozialraum berücksichtigt.

Der erste Beitrag richtet das Augenmerk auf die Schulprofilierung in der Sekundarstufe I und geht der Frage nach, inwiefern die konkrete Ausgestaltung von Einzelschulen durch sozialräumliche Spezifika der Nachbarschaft bedingt ist. Für die Untersuchung werden Schulleitungsdaten des Nationalen Bildungspanels herangezogen und durch Regionaldaten ergänzt. Während in diesem Beitrag davon ausgegangen wird, dass Schulleitungen bestimmte Anforderungen ihrer sozialen Umwelt antizipieren und in ihr schulisches Angebot übersetzen, untersucht der zweite Beitrag dann tatsächliches elterliches Wahlverhalten im Berliner Sekundarschulwesen. Im Fokus stehen hier nicht nur organisatorisch-pädagogische Merkmale, gleichermaßen wird auch die soziale Komposition der Schülerschaft als spezifisches Einzelschulmerkmal einbezogen und bei der Untersuchung ungleicher Übergänge berücksichtigt. Dabei geht es im zweiten Beitrag auch um die Frage, inwiefern beobachtete Ungleichheiten beim Übergang auf unterschiedliche Einzelschulen Resultat eines institutionellen oder eines Selbstselektionsprozesses sind. Der dritte Beitrag knüpft unmittelbar daran an, nimmt nun allerdings die Grundschulen in den Blick und beschäftigt sich auf Grundlage einer Erfurter Elternbefragung insbesondere mit Grundschulen in privater Trägerschaft. In Anlehnung an das Schulwahlmodell von Kristen (2005) wird hier die Genese von Ungleichheiten bei der Wahl einer privaten Grundschule herausgearbeitet, wobei sowohl Selektionsprozesse elterlicher Schulwahl als auch durch schulische Fremdselektion betrachtet werden.

Insgesamt leuchten die drei Beiträge den gemeinsamen Gegenstand horizontaler Differenzen im Schulwesen und deren Beitrag zur Ungleichheitsgenese sehr breit aus, indem verschiedenartige Perspektiven zusammengeführt sowie unterschiedliche Datenquellen und analytische Ansätze genutzt werden. Das Symposium trägt damit wesentlich zur Weiterentwicklung des Forschungsfeldes bei.

Beiträge des Symposiums

Schulentwicklung im Sozialraum. Ungleichheiten schulischer Infrastruktur im deutschen Sekundarschulwesen

Carolin Zink¹, Marcel Helbig²

¹Leibniz-Institut für Bildungsverläufe (LIfBi); Friedrich-Schiller-Universität Jena, ²Leibniz-Institut für Bildungsverläufe (LIfBi)

Im deutschen Bildungssystem lassen sich seit den 1990er Jahren bildungspolitische Aktivitäten beobachten, die darauf abzielen, Bildungsinstitutionen mehr Autonomie in finanzieller, personeller, organisatorischer und pädagogischer Hinsicht zuzugestehen (Altrichter et al., 2011; Altrichter & Rürup, 2010). Jene Freiheitsgrade sollen Schulen in die Lage versetzen, ihr inhaltliches Profil zu schärfen und spezifische (außer-)unterrichtliche Angebote unter Berücksichtigung ihres sozialräumlichen Umfelds zu entwickeln (Altrichter et al., 2011; Hopf, 2014; Maag Merki, 2022). Ausgehend von diesem Anspruch geht der Beitrag der Frage nach, welche schulischen Spezifika in der konkreten Ausgestaltung von Einzelschulen zusammenwirken und inwiefern sich weiterführende Schulen an Standorten mit vergleichbaren sozialräumlichen Bedingungen in ihrer inneren Verfasstheit ähneln.

Neo-institutionalistischen Ansätzen folgend (Meyer & Rowan, 1977; DiMaggio & Powell, 1983; Beckert, 2010), sollten sich Schulen in Reaktion auf Anforderungen ihrer äußeren Umwelt konstituieren, um Legitimität zu erhalten und ihren Bestand zu sichern – insbesondere im Lichte einer allgemeinen Legitimitätskrise in Folge des PISA-Schocks 2000 (Schaefers, 2002). Anforderungen, die über einen allgemeinen Bildungsauftrag hinausgehen, könnten sich dabei in besonderen lokalen Unterstützungsbedarfen (wie z.B. einem begleiteten Übergang in die berufliche Bildung) oder spezifischen elterlichen Präferenzen (z.B. Bourdieu, 1982; Grundmann et al., 2004) ausdrücken.

Unter Rückgriff auf die Reproduktionsfunktionen des Schulsystems (Fend, 2012) ist jedoch davon auszugehen, dass Schulen nicht nur der Qualifikation dienen, sondern potentiell auch als Räume sozialer Schließung und Distinktion in Erscheinung treten können (Altrichter et al., 2011; Nonte & Keßler, 2020). Entsprechend vermuten wir zwei zentrale „Anspruchsgruppen in der Umwelt“ (Walgenbach, 2014: 296) von Einzelschulen, die die Ausgestaltung von Bildungsangeboten bewusst oder unbewusst beeinflussen: einerseits die Bildungspolitik mit einem besonderen Fokus auf Qualifikation und Kompensation, andererseits hohe Statusgruppen, die die Qualifizierungsleistung von Schule stärker vor dem Hintergrund von Selektion und kultureller Reproduktion einordnen könnten.

Die Untersuchung stützt sich auf die Schulleiter*innen-Befragung zur NEPS-Startkohorte 3 (NEPS-Netzwerk 2024), die Informationen zu organisatorischen und pädagogischen Spezifika liefert. Die Daten werden um kleinräumige Microm-Regionaldaten (Schönberger & Koberg, 2016) zur Beschreibung der Schulumgebung (z.B. Zuwanderungshintergrund, Sozialstatus, Kaufkraft, Arbeitslosenquote) ergänzt. Zur Untersuchung des Zusammenspiels schulischer Merkmale finden in

einem ersten Schritt Clusteranalysen Anwendung. Multinomiale logistische Regressionsanalysen dienen schließlich der Modellierung der Auftretenswahrscheinlichkeit bestimmter Schulcluster in unterschiedlichen Sozialräumen. Dabei können Daten von n=230 Sekundarschulen herangezogen werden.

Erste Analysen zeigen vielfältige Zusammenhänge zwischen Schulspezifika und Sozialraum. Mit Blick auf die Vernetzung mit außerschulischen Akteuren wird erkennbar, dass Schulen in herausfordernden sozialen Lagen eine insgesamt höhere Netzwerkichte aufweisen, während mit steigendem Sozialstatus zunehmend weniger Kooperationen unterhalten werden. Zudem befinden sich Schulen, die mit Einrichtungen der Jugendhilfe, Betrieben und Unternehmen sowie mit Bildungsträgern (z.B. VHS) kooperieren, tendenziell eher in einem sozialräumlichen Umfeld, welches von statusniedrigeren Gruppen bzw. überdurchschnittlicher Arbeitslosigkeit geprägt ist. Im Gegensatz dazu werden keine eindeutigen Zusammenhänge zwischen einem sozial privilegierten Umfeld und schulischen Kooperationen mit kulturellen Einrichtungen sowie Hochschulen sichtbar. Hinsichtlich des Schulprofils kann zunächst festgehalten werden, dass die soziale Lage keinen Einfluss darauf zu haben scheint, ob Schulen ein Profil ausbilden oder nicht. Jedoch ist die inhaltliche Ausrichtung zum Teil deutlich mit dem Sozialraum assoziiert. Schulen mit einem alt- und neusprachlichen sowie musischen Profil sind eher in statushöheren Gebieten anzutreffen. Dies gilt mit Ausnahme des altsprachlichen Profils, dafür jedoch für MINT-Schwerpunkte, ebenso für Nachbarschaften mit einem höheren Anteil von Einwohner*innen mit Zuwanderungshintergrund. Mit zunehmender Arbeitslosigkeit werden die genannten Schulprofile hingegen deutlich unwahrscheinlicher. Werden spezifische Ansätze sowie außerunterrichtliche Angebote betrachtet, so wird deutlich, dass spezifische Pädagogiken (z.B. Montessori) und Partizipationsangebote mit einem höheren Sozialstatus in der Umgebung einhergehen. Spannend ist in diesem Kontext, dass Projektunterricht mit einer zunehmenden Arbeitslosenquote häufiger eingesetzt wird. Des Weiteren unterbreiten jene Schulen häufiger Lernangebote in den Kernfächern, engagieren sich stärker in Netzwerken zum Übergang auf den Arbeitsmarkt und halten intensiveren Kontakt zu Berufsberatungsstellen.

Stichworte: Schulprofilierung, Schulk Kooperation, Sozialraum, Sekundarstufe

Einzelschulwahl und soziale Ungleichheit. Die Rolle primärer, sekundärer und tertiärer Herkunftseffekte

Nicky Zunker¹, Marko Neumann²

¹Humboldt Universität zu Berlin; Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF), ²Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF)

Herkunftsspezifische Disparitäten beim Übergang von der Grundschule in die weiterführenden Schulen werden in der Bildungsforschung als ein Schlüsselement zur Erklärung von sozialen Ungleichheiten im weiteren Bildungserwerb verstanden. Dieser Zusammenhang ist in Deutschland für die Wahl der Schulform umfassend theoretisch aufgearbeitet und empirisch nachgewiesen worden (zusammenfassend: Dumont et al., 2014). Theoretisch wird davon ausgegangen, dass diese Ungleichheiten einerseits Resultat herkunftsspezifischer Leistungsunterschiede der Schüler*innen (primärer Herkunftseffekt) und andererseits eines herkunftsspezifisch verzerrten Schulwahlverhaltens sind (sekundärer Herkunftseffekt) (Boudon, 1974). In den letzten Jahren wurde zunehmend auch auf herkunftsspezifisch verzerrte Bewertungspraxen der Lehrkräfte (Helbig & Morar, 2018) als weiteres Erklärungsmoment verwiesen.

Wie etwa Baumert et al. (2006) zeigen, bestehen außer zwischen Schulen unterschiedlicher Schulformen, auch zwischen Schulen derselben Schulform erhebliche Unterschiede. Dabei geht es nicht nur um die Komposition der Schülerschaft. Hinzu kommen vielfach bedeutsame Unterschiede in der inhaltlichen, organisatorischen und pädagogischen Ausgestaltung der Schulen (Zunker et al., 2018). Daraus lässt sich im Weiteren ableiten, dass mit Blick auf die sich durch den Besuch einer bestimmten Schule eröffnenden Lern- und Entwicklungschancen und damit assoziierter Ungleichheiten im Bildungserwerb, entsprechend nicht nur die Wahl der Schulform, sondern darüber hinaus auch der konkreten Einzelschule von Bedeutung sein sollte. Bis auf wenige Ausnahmen (z.B. Clausen, 2006; Klinge, 2016) wurde die Wahl konkreter Einzelschulen im deutschsprachigen Raum nicht untersucht. Der internationale Forschungsdiskurs gibt jedoch Grund zur Annahme, dass auch in Hinblick auf die Wahl der Einzelschule von herkunftsspezifischen Disparitäten auszugehen ist (Harris & Larsen, 2014; Burgess et al., 2015; Davies & Aurini, 2011).

An dieses Desiderat schließt der vorliegende Beitrag an und geht der Frage nach, inwieweit sich am Übergang in die weiterführende Schule auch herkunftsspezifische Unterschiede in Bezug auf die Einzelschule nachweisen lassen. Darüber hinaus ist zu überprüfen, inwiefern etwaige Disparitäten auf primäre, sekundäre und tertiäre Herkunftseffekte zurückzuführen sind.

Als Datengrundlage wird auf eine repräsentative Stichprobe der BERLIN-Studie (Maaz et al., 2013) von n = 3489 Schüler*innen zurückgegriffen, die zum Schuljahr 2011/12 in die weiterführenden Schulen (Gymnasien und Integrierte Sekundarschulen) übergingen. Der Datensatz umfasst neben umfangreichen Daten zu leistungsbezogenen und sozialen Hintergrundmerkmalen der Schüler*innen sowie elterlichen Schulwahlmotiven auch konkrete Angaben zu präferierten und realisierten Schulen, welche mit Informationen aus dem Berliner Schulverzeichnis zusammengeführt werden konnten. Zur Beschreibung der Einzelschulen wurde in Anlehnung an Baumert et al. (2017) ein metrischer Schulindex gebildet. Der Index setzt sich sowohl aus Schülerkompositionsmerkmalen (z.B. Anteil der Schüler*innen nicht deutscher Herkunftssprache oder mit Lernmittelzuzahlungsbefreiung) als auch Angebotsmerkmalen der Schule (z.B. Vorhandensein einer gymnasialen Oberstufe oder die Anzahl angebotener Fremdsprachen) zusammen. Mit Hilfe von Pfadanalysen wird unter Kontrolle schulischer Leistungen (Noten und Testleistungen), des Geschlechts und des Zuwanderungshintergrundes zunächst der Einfluss der sozialen Herkunft auf den Schulindex der präferierten und realisierten Schulen geschätzt. Durch die Modellierung indirekter Effekte wird in einem zweiten Schritt eine Dekomposition der Herkunftseffekte in ihre primären, sekundären und tertiären Anteile ermöglicht. Die Analysen erfolgen schulförmenspezifisch und separat für zwei verschiedene Operationalisierungen des sozialen Hintergrundes (HISEI vs. höchster allgemeinbildender Schulabschluss).

Es zeigen sich insofern klare herkunftsspezifische Unterschiede im Schulwahlverhalten, dass statushöhere Familien Schulen mit höheren Indexwerten präferieren und auch realisieren. Die Disparitäten sind unter den Familien, die Integrierte Sekundarschulen (ISS) wählen, stärker ausgeprägt als unter den Familien, die sich für ein Gymnasium entscheiden, wobei zwischen diesen beiden Gruppen bereits erhebliche Unterschiede im sozialen Status bestehen. Innerhalb der Gymnasialteilstichprobe lassen sich die Disparitäten vor allem auf sekundäre und kaum auf tertiäre Effekte zurückführen, während innerhalb der ISS-Teilstichprobe herkunftsbezogene Disparitäten mehr das Resultat primärer und tertiärer Herkunftseffekte sind. Wie die Anteile primärer, sekundärer und tertiärer Herkunftseffekte im Genauen ausfallen, variiert neben der Schulform auch danach, wie die soziale Herkunft operationalisiert wird.

Stichworte: Soziale Ungleichheit, Schulwahl, Effektdekomposition

Soziale Ungleichheiten beim Privatschulzugang. Die Wahl einer privaten Grundschule aus Elternsicht

Privatschulen müssen in Deutschland Kindern aller sozialen Schichten offenstehen (Art. 7, Abs. 4 GG). Dennoch lassen sich ausgeprägte soziale Unterschiede beim Besuch privater (Grund-)Schulen beobachten (Habeck et al., 2017; Helbig et al., 2022; Görlitz et al., 2018; Lohmann et al., 2009). Ziel des vorliegenden Beitrags ist es, Erklärungen für diese ungleichen Schulbesuchsquoten zu liefern.

Theoretisch rekurriert dieser Beitrag insbesondere auf das dreistufige Modell zur Schulwahl von Kristen (2005), das nochmal auf den speziellen Fall der Genese von Ungleichheiten bei der Wahl privater Schulen adaptiert wurde. Auf der ersten Stufe wird dabei der Frage nachgegangen, ob und warum Eltern überhaupt Alternativen zur nächstgelegenen und/oder zugewiesenen Schule wahrnehmen und warum nicht. Diese Unterscheidung zwischen wählenden und nicht-wählenden Eltern ist eng mit dem sozio-ökonomischen und kulturellem Hintergrund der Familien verbunden (Kristen, 2005; Schwarz et al., 2017). Auf der zweiten Stufe wird entsprechend den Annahmen der Werterwartungstheorie (Esser, 1999) eine bestimmte Einzelschule gewählt, wobei diese Entscheidung ebenfalls in Abhängigkeit der sozialen Herkunft getroffen wird (Kristen, 2005; Henry-Huthmacher, 2008; Drope, 2017; Stirner et al., 2019; Habeck et al., 2017; Eisinger et al., 2007; Mayer, 2017; Mayer & Koinzer, 2019). Schließlich wird auf der dritten Stufe auch vonseiten der Schulen eine Passung zu den Familien hergestellt, indem diese ihrerseits Schüler*innen auswählen (Kristen, 2005; Heinrich et al., 2011; Reimann et al., 2020; Krüger, 2015). Dies geschieht in Abhängigkeit von der Verteilung der Ressourcen der Bewerber*innen oder der Präferenz spezifischer sozialer oder religiöser Gruppen. Die drei Stufen des Modells bauen dabei aufeinander auf. Um die Entstehung von Ungleichheiten insbesondere durch die Wahl einer Privatschule anhand dieses sequenziellen Modells zu analysieren, werden folgende Fragestellungen untersucht:

Stufe 1: Welche Eltern wissen von der Möglichkeit, eine Privatschule für ihr Kind wählen zu können? Lassen sich dadurch soziale Ungleichheiten bei der Schulwahl erklären?

Stufe 2: Welche Eltern entscheiden sich für eine private Grundschule? Inwieweit lassen sich über Nutzen- und Kostenabwägungen soziale Ungleichheiten an dieser Stufe erklären?

Stufe 3: Zeigen sich soziale Ungleichheiten in Bezug auf die Fremdselektionsprozesse aufseiten der Schulen?

Als Datengrundlage wird eine Elternbefragung im Rahmen der Einschulung 2020/21 in Erfurt herangezogen. Nach Ausschluss fehlender Werte konnten n = 683 Fälle analysiert werden.

Um die Fragestellungen beantworten zu können, wird auf ein schrittweise aufgebautes lineares Wahrscheinlichkeitsmodell zurückgegriffen. Als abhängige Variablen liegen a) die Bewerbung an einer Privatschule und die b) die Annahme durch die Privatschule vor. Dies ermöglicht es je nach Analyse, zwischen Eltern zu unterscheiden, die sich a) für eine Privatschule entschieden haben, ganz unabhängig davon, ob das Kind auch an der Schule angenommen wurde und b) Eltern, die eine Privatschule gewählt haben und deren Bewerbung tatsächlich auch erfolgreich war. Als unabhängige Variablen können der Bildungsabschluss der Eltern, der Migrationshintergrund, die Intensität der Schulwahl, die wahrgenommene Qualität des öffentlichen resp. privaten Schulsystems, die Schulwahlmotive und die Entfernung der Schule in das Modell aufgenommen werden. Neben den Regressionen ergänzen deskriptive Befunde die multivariaten Ergebnisse.

Gemäß dem Modell der Privatschulwahl zeigt sich bei der ersten Stufe, dass viele Eltern mit niedrigerer Bildung Privatschulen nicht als Wahlalternativen wahrnehmen. Auf der zweiten Stufe konnte festgestellt werden, dass Eltern mit niedrigerer Bildung überdurchschnittlich häufig angeben, sich private Schulen nicht leisten zu können. Es gibt darüber hinaus unterschiedliche Ertragsenerwartungen der Eltern unterschiedlicher Bildungsniveaus im Hinblick auf die Qualität privater Schulen und bei den Schulwahlmotiven. Die-se können aber die sozialen Unterschiede bei der Privatschulbewerbung auf Stufe zwei nur in geringem Maße aufklären. Auf der dritten Stufe wird deutlich, dass sich Akademi-kereltern häufiger erfolgreich bei einer privaten Schule bewerben als Nicht-Akademikereltern. Insgesamt zeigt sich, dass auf allen drei Stufen des adaptierten Modells von Kristen soziale Ungleichheiten bei der Privatschulwahl verstärkt werden.

Stichworte: Soziale Ungleichheit, Schulwahl, Privatschule, Grundschule

S-72: Inklusive Bildung: Die Einstellungen von Lehrkräften im Mittelpunkt

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 157

Chair der Sitzung: Thorsten Fahrbach

Chair der Sitzung: Carmen Zurbriggen

Inklusive Bildung: Die Einstellungen von Lehrkräften im Mittelpunkt

Chair(s): **Thorsten Fahrbach** (Institut für Bildungsanalysen Baden-Württemberg)

Diskutant*in(nen): **Carmen Zurbriggen** (Universität Freiburg (CH))

Der Umgang, die Integration und die Inklusion von Vielfalt in das Schulsystem spielen in den drei deutschsprachigen DACH-Ländern in den letzten Jahrzehnten eine zentrale Rolle in der gesellschaftlichen, politischen und wissenschaftlichen Diskussion (Lütje-Klose et al., 2017). Alle drei Länder bemühen sich zunehmend, Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Regelschulen zu unterrichten (Autor:innengruppe Bildungsberichterstattung, 2022; BMSGPK, 2022; Der Bundesrat, 2016). Gleichzeitig nimmt die Diversität der Schülerschaft auch durch Migration und Flucht zu (BMBWF, 2019; Henschel et al., 2023; SKBF, 2023). Im Sinne eines erweiterten Inklusionsbegriffs ist das Schulsystem in den DACH-Ländern daher zunehmend von einer heterogenen Schülerschaft mit unterschiedlichen Hintergründen, Voraussetzungen und Bedürfnissen geprägt (Wischer, 2021).

Neben dem politischen Willen und der gesellschaftlichen Unterstützung kommt den Lehrkräften eine zentrale Rolle bei der Umsetzung eines inklusiveren Bildungssystems zu. Die Bereitschaft von Lehrkräften, sich auf neue Erfahrungen einzulassen und die eigenen Kompetenzen zu erweitern, hängt von ihren individuellen Einstellungen ab (Ajzen, 1991, 2012; Baumert & Kunter, 2006). Daher gelten die Einstellungen von Lehrkräften zur inklusiven Bildung als entscheidende Faktoren, die es ihnen ermöglichen, sich mit inklusivem Unterricht auseinanderzusetzen und qualitativ hochwertige Inklusion in Schulen umzusetzen (Hellmich et al., 2019; Sharma & Mannan, 2015).

Die Einstellungen von Lehrkräften zur inklusiven Bildung können als ihre Überzeugungen und Gefühle in Bezug auf die Inklusion von Kindern mit unterschiedlichen pädagogischen Bedürfnissen in Regelschulklassen verstanden werden (Yada et al., 2022). Das Themenfeld der Einstellungen von Lehrkräften gilt als eines der am intensivsten erforschten Bereiche im Bereich der inklusiven Bildung (Van Miegheem et al., 2020). Aus diesem Grund soll das Symposium verschiedene Perspektiven auf das Forschungsfeld eröffnen und wir freuen uns darüber, nicht nur Forschungsberichte aus Deutschland, sondern auch aus Österreich und der Schweiz präsentieren zu dürfen.

Der erste Beitrag untersucht, wie professionelle Dispositionen von Lehrkräften die Umsetzung inklusiven Unterrichts beeinflussen. Der Beitrag identifiziert vier Lehrkräfteprofile auf der Grundlage von Einstellungen zur Inklusion, lerntheoretischen Überzeugungen und Zufriedenheit mit der Kooperation. Die Ergebnisse zeigen, dass sonderpädagogische Lehrkräfte häufiger positive Einstellungen zur Inklusion und konstruktivistische Überzeugungen haben als Regelschullehrkräfte.

Der zweite Beitrag stellt die Entwicklung einer neuen Skala zur Erfassung der Einstellungen von Lehrkräften zur Legitimität, Effektivität und Durchführbarkeit von segregierten Deutschförderklassen in Österreich vor. Erste Ergebnisse einer Online-Befragung von 474 Lehrkräften zeigen eine hohe interne Konsistenz und Validität der Skala, die eine differenzierte Erfassung der Einstellungen von Lehrkräften ermöglicht.

Der anschließende Beitrag beschäftigt sich mit längsschnittlichen Zusammenhängen zwischen inklusionsbezogener Selbstwirksamkeit, Einstellungen zur Inklusion und inklusiven Lerngelegenheiten bei (angehenden) Lehrkräften. Erste Ergebnisse zeigen, dass höhere Selbstwirksamkeitserwartungen negativ mit späteren Einstellungen zur Inklusion zusammenhängen. Inklusive Lerngelegenheiten hatten zeitverzögert positive Effekte sowohl auf die Selbstwirksamkeit als auch auf die Einstellungen zur Inklusion.

Der letzte Beitrag untersucht die Zusammenhänge zwischen Merkmalen guter inklusiver Schulen, der Arbeitszufriedenheit und der kollektiven Selbstwirksamkeit von Lehrkräften sowie der sozialen Integration von Schülern in inklusiven Sekundarschulen. Die Daten von 28 Schulen in Nordrhein-Westfalen zeigen, dass positive Schulmerkmale und die Arbeitszufriedenheit der Lehrkräfte zusammenhängen und indirekt die soziale Integration der Schüler beeinflussen.

Beiträge des Symposiums

Welche Rolle spielt die sonderpädagogische Ausbildung für die Einstellung zur Inklusion, die Überzeugung über das Lernen und die Zufriedenheit mit der interdisziplinären Kooperation? Eine Clusteranalyse

Simon Luger¹, Susanne Schnepel², Maria Wehren-Müller¹, Elisabeth Moser Opitz¹

¹Universität Zürich, ²Universität Münster

Es wird angenommen, dass berufsbezogene Dispositionen von Lehrkräften wie Einstellungen und Überzeugungen die Umsetzung inklusiven Unterrichts beeinflussen. Lehrkräfte mit einer positiven Einstellung zum inklusiven Unterricht von Schüler:innen mit besonderem Bildungsbedarf setzen inklusive Unterrichtspraktiken häufiger um (Avramidis et al., 2019) als Lehrkräfte mit einer negativen Einstellung. Lehrkräfte mit konstruktivistischen Überzeugungen gestalten ihren Unterricht kognitiv aktivierender als Lehrkräfte mit transmissiven Überzeugungen (Blömeke, 2012). Neben berufsbezogenen Dispositionen wird auch die Bedeutung der Kooperation von Lehrkräften für eine gelingende Umsetzung von inklusivem Unterricht betont (Lütje-Klose & Urban, 2014). So gibt es etwa Hinweise, dass die Zufriedenheit mit der interdisziplinären Kooperation zwischen Regellehrkräften (RLK) und sonderpädagogische Lehrkräfte (SLK) bedeutsam sein könnte (Gebhard et al., 2014).

Studien weisen darauf hin, dass sich RLK und SLK hinsichtlich der genannten Faktoren unterscheiden können. Moser et al. (2014) zeigten auf, dass ausgebildete SLK eine positivere Einstellung zur Inklusion haben als RLK. SLK scheinen jedoch weniger zufrieden zu sein mit der interdisziplinären Kooperation als RLK (Gebhard et al., 2014). Diese und andere Studien untersuchten die einzelnen Lehrkraftvariablen jedoch meist isoliert, sodass Unklarheit über die Beziehungen zwischen den Variablen besteht – oder die Untersuchungen wurden nur mit einer Lehrkräftegruppe (RLK oder SLK) durchgeführt. Es ist wichtig, entsprechende Zusammenhänge bzw. Profile zu berücksichtigen, um Faktoren zu identifizieren, welche die Umsetzung von inklusivem Unterricht beeinflussen könnten.

Fragestellungen

FS1: Welche charakteristischen Profile von Lehrkräften, die auf den Variablen Einstellung zur inklusiven Bildung, lerntheoretische Überzeugungen und Zufriedenheit mit der Kooperation basieren, können unterschieden werden?

FS2: Wie sind Lehrkräfte mit unterschiedlicher Ausbildung (Regel- vs. Sonderpädagogik) in den Profilen vertreten?

Methoden

Stichprobe: N = 173 Lehrkräfte der Klassen 1. - 4. (93,5 % weiblich), die in 80 inklusiven Klassen arbeiteten. Alle Lehrkräfte verfügten über einen Lehramtsabschluss für die Regelschule, 34,1 % (n = 59) zusätzlich über einen Abschluss in Sonderpädagogik.

Instrumente:

Einstellung: AIS, 8 Items, $\alpha = .73$ (Sharma & Jacobs, 2016); Überzeugung: 6 Items, $\alpha = .70$ (vom Forscher entwickelte Skala), Kooperationszufriedenheit: 10 Items, $\alpha = .87$ (Gebhard et al., 2014). Alle Fragebögen wurden online ausgefüllt.

Auswertungen: Zur Beantwortung der ersten Fragestellung (Bestimmung der Anzahl und der Arten von Clustern) wurden eine hierarchische Clusteranalyse (Methode von Ward) und ANOVAs durchgeführt. Jedes Cluster stellt ein Profil dar. Um die Unterschiede zwischen den Profilen (kategoriale Variablen) und der Berufsausbildung (dichotome Variable) zu untersuchen, wurden Chi-Quadrat-Tests durchgeführt.

Ergebnisse

FS1: Die Analyse teilte die Stichprobe in vier Cluster. Das erste Cluster ('Inklusion durch Konstruktion'; n = 25) zeichnet sich durch hohe Werte bei allen drei Variablen aus. Die Lehrkräfte des zweiten Clusters ('Ausgewogene Disposition'; n = 87) stehen der Inklusion vergleichsweise kritisch gegenüber, zeigen aber durchschnittliche Werte bezüglich lerntheoretischer Überzeugung und der Kooperationszufriedenheit. Das dritte Cluster ('Inklusion durch Transmission'; n = 51) zeichnet sich durch eine Tendenz zu transmissiven Überzeugungen aus. Das vierte Cluster ('Divergenzen in der Kooperation'; n = 10) verfügt über ein geringes Maß an Kooperationszufriedenheit und ist skeptisch gegenüber der Inklusion aus.

FS2: Der Chi-Quadrat-Test zeigt einen schwachen bis mittleren Effekt für die 4-Cluster-Lösung. Das bedeutet, dass RLK und SLK über die vier Cluster hinweg ungleich verteilt sind ($w = .28^{**}$). Im Cluster 'Inklusion durch Konstruktion' ist die Anzahl von SLK signifikant höher als in den anderen Clustern. Im Cluster 'Divergenzen in der Kooperation' hat nur eine der zehn Lehrkräfte eine sonderpädagogische Ausbildung.

Interpretation

Lehrkräfte, die in inklusiven Klassen arbeiten, scheinen unterschiedliche Profile in Bezug auf die untersuchten Variablen zu haben. SLK waren in der Gruppe mit hohen Werten in allen Skalen überrepräsentiert. Die Frage, ob dies ein Ergebnis der sonderpädagogischen Ausbildung ist, oder ob Lehrkräfte in diesem Profil dazu neigen, sich für eine sonderpädagogische Ausbildung zu entscheiden, kann mit den vorliegenden Daten nicht beantwortet werden.

Einstellung von Lehrkräften gegenüber segregierenden Sprachfördermaßnahmen

Sepideh Hassani
Universität Wien

Die Beherrschung der Unterrichtssprache ist ein entscheidender Faktor für den akademischen Erfolg und die zukünftigen Karrierechancen von Schüler:innen (OECD, 2020). In Österreich wird traditionell ein einsprachiger Ansatz verfolgt (siehe z.B. Erling et al., 2021) obwohl Mehrsprachigkeit in den Klassenzimmern stark ausgeprägt ist (Statistik Austria, 2023). Dadurch werden Schüler:innen mit beginnenden Deutschsprachkenntnissen benachteiligt. Weiters initiierte die österreichische Bundesregierung eine segregierte Deutschsprachförderung (BMBWF, 2019), sogenannte Deutschförderklassen (DFK), obwohl dies empirischer Evidenz widerspricht, wenn es um das effektive Erlernen der Unterrichtssprache geht (Cummins, 2015; Erling et al., 2021; Short, 2017; Thomas & Collier, 2017). Das Modell der segregierende Beschulung von neu angekommenen Schüler:innen findet sich ebenfalls in vielen anderen europäischen Ländern wieder (European Commission, 2019).

In Österreich werden DFKs parallel zum Regelunterricht durchgeführt. Schüler:innen besuchen diese 15 Stunden pro Schulwoche in der Grundschule und 20 Stunden in den weiterführenden Schulen. Während des Besuchs der DFK ist ein Aufstieg in die höhere Klasse nicht gestattet (BMBWF, 2019). In anderen Ländern sind diese Pull-out-Klassen nur für neu angekommenen Schüler:innen vorgesehen (siehe z. B. Willkommensklassen in Berlin (Deutschland), Neumann et al., 2020). In Österreich hingegen ist mehr als die Hälfte der Schüler:innen, die die DFKs besuchen, in Österreich geboren, und etwa ein Drittel von ihnen hat zuvor einen Kindergarten in Österreich besucht (Expert Council for Integration, 2022).

Die Literatur zur inklusiven Beschulung (Avramidis & Norwich, 2002) und zur Sprachförderung (z.B. Erling et al., 2023) weißt immer wieder auf die zentrale Rolle der Einstellung der Lehrkräfte für das Lernen der Schüler:innen hin. Filsecker und Abs (2021) betonen die Bedeutung einer positiven Einstellung gegenüber Schüler:innen mit Fluchterfahrung für eine erfolgreiche Inklusion ebendieser. Darüber hinaus legen Forschungsergebnisse zu den Einstellungen von Lehrkräften Vorurteile gegenüber Schüler:innen mit Migrationsbiografie nahe (z.B. Glock & Klapproth, 2017).

Bleidick (1985; siehe auch Luger, 2023) diskutierte in seiner Arbeit, dass die Einstellung zu Inklusion oder Segregation im Hinblick auf ihre Legitimität, Wirksamkeit und Durchführbarkeit untersucht werden kann. Das Konzept der Legitimität bezieht sich dabei auf ethischen und moralischen Komponenten. Die Effektivität hängt davon ab, inwieweit sie zu besseren Lern- und psychosozialen Entwicklungsergebnissen bei Schüler:innen führt. Der Aspekt der Durchführbarkeit, bezieht sich auf die praktische Umsetzung, wie beispielsweise Fragen zur Infrastruktur, zu den Kompetenzen des Personals und zur Verfügbarkeit finanzieller Mittel.

In den letzten Jahren gab es Untersuchungen, die darauf hinwiesen, dass Lehrkräfte (zumindest einige Aspekte von) DFKs als problematisch erachten (Gitschthaler et al., 2024; Spiel et al., 2022). Die meisten dieser Studien waren qualitativ, und bisher wurde in keiner Studie eine geeignete psychometrische Skala für eine systematische Bewertung der Einstellungen von Lehrkräften zu diesen Maßnahmen verwendet. Die vorliegende Präsentation befasst sich daher mit der Entwicklung einer neu entwickelten Skala, „Attitudes towards pull-out language support classes questionnaire“ (ALSC; Schwab & Hassani, 2024), zur Bewertung der Einstellungen von Lehrkräften zur Legitimität, Effektivität und Durchführbarkeit von Sprachförderklassen. In der Präsentation werden die psychometrischen Eigenschaften der deutschen Version des Fragebogens präsentiert. Die Ergebnisse stammen aus einer laufenden Studie („Eine multiperspektivische Studie zu Deutschförderklassen“). Insgesamt 474 Lehrkräfte (313 Grundschullehrkräfte und 161 Sekundarschullehrkräfte) nahmen an einer Online-Befragung teil. Die Ergebnisse der konfirmatorischen Faktorenanalyse zeigten ein bifaktorielles Modell. Neben einem starken eindimensionalen Faktor wurden drei (stark korrelierte) Subfaktoren identifiziert: Legitimität, Effektivität und Umsetzbarkeit. Sowohl die Gesamtskala als auch jede der drei Subskalen wiesen hohe interne Konsistenzen auf. Zusätzlich wurde die Messinvarianz für Grund- und Sekundarschullehrkräfte bestätigt. Die Validität der Skala wurde durch die erwarteten Korrelationen mit zusätzlichen Items bestätigt.

Wechselwirkungen zwischen inklusionsbezogenen Einstellungen, Selbstwirksamkeitserwartungen und Lerngelegenheiten: Ergebnisse von Cross-Lagged-Panel-Analysen aus dem NEPS-Lehramtsstudierendenpanel

Thorsten Fahrbach¹, Susanne Schwab², Nicolas Hübner³

¹Institut für Bildungsanalysen Baden-Württemberg, ²Universität Wien, ³Universität Tübingen

Theoretischer Hintergrund

Den Einstellungen gegenüber Inklusion und den inklusionsbezogenen Selbstwirksamkeitserwartungen von Lehrkräften werden bei der Umsetzung eines inklusiven Bildungssystems zentrale Rollen zugesprochen (Hellmich et al., 2019; Sharma & Mannan, 2015). Die Beschäftigung mit Inklusion in der Schule sowie der Kontakt mit Schülerinnen und Schülern mit Förderbedarf können sowohl die inklusionsbezogene Selbstwirksamkeit als auch die Einstellungen zur Inklusion beeinflussen (Allport, 1954; Fernandez et al., 2016). Verschiedene Studien legen nahe, dass derartige Lerngelegenheiten einerseits als Ergebnis und andererseits als Auslöser von Einstellungen zur Inklusion und inklusionsbezogener Selbstwirksamkeit verstanden werden können (Fives & Buehl, 2012; Richardson, 1996), was Fragen zur Notwendigkeit ihrer simultanen Betrachtung aufwirft.

Die Zusammenhänge zwischen Einstellungen zur Inklusion, inklusionsbezogener Selbstwirksamkeit und inklusionsbezogenen Lerngelegenheiten in der Lehrerbildung und in Schulen wurden in verschiedenen Studien untersucht (z.B. Dignath et al., 2022; Schaub & Lütolf, 2023; Werner et al., 2021). Ein Großteil dieser Studien verwendete hierzu allerdings querschnittliche Designs, die keine Informationen zur zeitlichen Ordnung der Zusammenhänge dieser Variablen liefern können. In den wenigen vorliegenden Längsschnittstudien wurden bisher lediglich die Zusammenhänge zwischen inklusionsbezogener Selbstwirksamkeit und Einstellungen zur Inklusion im Zeitverlauf untersucht (Bosse et al., 2016; Savolainen et al., 2022). Studien, die auch die längsschnittliche Zusammenhänge mit Lerngelegenheiten berücksichtigen, fehlen bislang (Menge et al., 2021).

Fragestellung

Um das dargestellte Forschungsdesiderat zu schließen, wird der Frage nachgegangen, welche längsschnittlichen Zusammenhänge sich bei einer simultanen Betrachtung der drei Konstrukte (inklusionsbezogene Selbstwirksamkeitserwartungen, Einstellungen zur Inklusion und Lerngelegenheiten zu inklusivem Unterricht) über drei Messzeitpunkte hinweg zeigen.

Methode

Die Arbeit verwendet Daten des Lehramtsstudierendenpanels (NEPS-Netzwerk, 2023; Schaeper et al., 2023), in dem (angehende) Lehrkräfte aller Schulformen seit 2011 regelmäßig zu ihren Einstellungen und Entwicklungen befragt werden. Die relevanten Variablen wurden jeweils zu drei Messzeitpunkten (2016, $n = 1862$; 2018, $n = 1360$; 2020, $n = 1383$) erfasst. Alle Variablen wurden mit einer sechsstufigen Likert-Skala erhoben. Die Einstellungen zur Inklusion und die inklusionsbezogene Selbstwirksamkeit wurden mit sechs bzw. vier Items erfasst (Bosse & Spörer, 2014) und wiesen eine gute interne Konsistenz auf ($\alpha = .85$ bzw. $\alpha = .86$). Die Inklusionsbezogene Lernerfahrungen wurden mit fünf Items einer modifizierten Skala nach Laschke & König (2014) erfasst ($\alpha = .84$).

Die längsschnittlichen Zusammenhänge der drei Konstrukte wurden mittels verschiedener Varianten von Cross-Lagged-Panel-Modellen (CLPM) geschätzt (vgl. Hübner et al., 2023). Wir verfolgen die Modellierungsstrategie, verschiedene plausible Varianten aus dem Multiversum möglicher methodologischer Optionen darzustellen, um so einen Vergleich zu ermöglichen (Steege et al., 2016).

Ergebnisse

Erste Schätzungen aus dem CLPM und dem Full-Forward CLPM zeigen, dass höhere Selbstwirksamkeitserwartungen negativ mit Einstellungen zu späteren Messzeitpunkten zusammenhängen. Umgekehrt zeigen sich jedoch keine Zusammenhänge von Einstellungen auf die späteren Selbstwirksamkeitserwartungen. In Bezug auf die Lerngelegenheiten zeigen sich positive zeitverzögerte Zusammenhänge mit der inklusionsbezogenen Selbstwirksamkeit und den Einstellungen zur Inklusion, jedoch keine statistisch signifikanten gegenläufigen Pfade.

Diskussion

Im Gegensatz zu den wenigen vorliegenden Längsschnittstudien (Savolainen et al., 2022) zeigen sich in unserer Studie negative Zusammenhänge zwischen den Selbstwirksamkeitserwartungen und späteren Einstellungen. Die umgekehrten Zusammenhänge hingegen deckten sich mit früheren Längsschnittstudien (Bosse et al., 2016; Carew et al., 2019; Savolainen et al., 2022).

Insgesamt unterstreichen die Ergebnisse die besondere Bedeutung von inklusiven Lernerfahrungen für nachfolgende inklusionsbezogene Einstellungen und Selbstwirksamkeitserwartungen von (angehenden) Lehrkräften.

S-73: Benefits of Generative Learning and Desirable Difficulties for Meaningful Learning: Combining Learning Activities and Addressing Learner Characteristics

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SN 163

Chair der Sitzung: Veit Kubik

Chair der Sitzung: Matthias Nückles

English Track

Benefits of Generative Learning and Desirable Difficulties for Meaningful Learning: Combining Learning Activities and Addressing Learner Characteristics

Chair(s): **Veit Kubik** (University of Würzburg, Germany)

Diskutant*in(nen): **Matthias Nückles** (University of Freiburg, Germany)

A key objective in educational psychology is to examine instructional and learning techniques that help learners retain and flexibly use knowledge in authentic learning situations in school or higher education. However, much of the previous research was limited to short-term studies in lab settings using simple materials, primarily focusing on memory, which leaves a gap in understanding how to create and to foster lasting knowledge in the educational “wild” (Brod, 2020; Richter et al., 2022).

In prior research, generative learning strategies and desirable difficulties have been highlighted as effective learning activities. Generative learning activities typically prompt students to self-explain, write learning journals (Nückles et al., 2020), or explain the learned contents to fictitious others (i.e., non-interactive teaching, Fiorella & Mayer, 2013), and thereby help to build up coherent mental representations of new knowledge (Roelle et al., 2023). Desirable difficulties typically make the learning experience harder, for example, by means of retrieving information or interleaving different types of subjects, with the desirable effect of consolidating mental representations and thereby enhancing long-term learning. A scarcity of research has investigated how generative learning activities and desirable difficulties can be combined and instructionally supported to enhance meaningful learning (Roelle et al., 2023; Richter et al., 2022).

Against this background, the symposium examines how generative learning (journal writing, non-interactive teaching) and desirable difficulties (retrieval practice, interleaved practice) can be effectively implemented and combined with other learning strategies. Furthermore, this symposium addresses the question of whether generative learning activities are genuinely and equally helpful for all students. In fact, it is assumed that their efficacy can vary with characteristics of the learner, such as their academic self-concepts (Jacob et al., 2022).

Contribution 1 investigates whether retrieval practice can enhance journal writing as an effective follow-up course activity within a digital lecture among teacher students; the degree of retrieval practice was manipulated by providing the lecture script (open-book) versus not (closed-book). Specifically, the authors examined whether and why the format of journal writing and revision (closed vs. open book) enhances long-term comprehension performance.

Contribution 2 examines whether the effectiveness of non-interactive teaching after immediate and delayed (meta)cognitive learning outcomes can be improved by (i) adding drawing as a visual-spatial generative learning strategy and by (ii) spacing the learning activities across the study phase. The authors target the individual and combined benefits as well as potential Learner x Activity interactions on an inquiry-based physics unit with regards to cognitive, motivational, and personality-related individual differences.

Contribution 3 tackles the research question whether the effectiveness of non-interactive teaching differs with certain motivational learner characteristics. Applying a pre-post-test experimental design, the authors manipulated students' academic self-concept (by providing fictitious feedback on their performance), and examined whether non-interactive teaching is more beneficial for students with low rather than with high self-concept regarding their immediate and delayed comprehension and metacognitive monitoring performance.

Contribution 4 investigates whether interleaved, compared to blocked, practice is effective for learning complex physics concepts. More specifically, the authors explored whether cooperative learning strategies could instructionally support 12th-grade students in actual classroom environments and thereby enhance the interleaving effect in meaningful learning.

All four contributions feature innovative experimental designs with measures of learning outcomes at longer delays, across varied student populations and subjects. Together, these studies offer insights on fostering lasting and flexible knowledge in school and university contexts and their differential effectiveness for motivational learning characteristics. An instructional design expert will discuss these contributions, providing theoretical and practical insights into combining learning activities and addressing learner diversity. These findings and implications are relevant for informing evidence-based educational practices and policies (Slavin, 2002).

Beiträge des Symposiums

Benefits of Closed- and Open-Book Format in Journal Writing: Is Retrieval Practice a Hidden Gem?

Alina Timmermann¹, Kirsten Berthold¹, Veit Kubik²

¹Bielefeld University, Germany, ²University of Würzburg, Germany

Theoretical Background:

Prompted journal writing is a generative learning activity that encourages students to organize and elaborate on learning contents, and is known to promote students' comprehension (Berthold et al., 2007; Nückles et al., 2009). Learning journals are particularly effective when students can subsequently work on them in a subsequent revision phase (Nückles et al., 2020). Previous research has typically implemented the initial journal-writing task in a closed-book format and the revision task in an open-book format (i.e., with the lecture transcript being openly available). Note that a closed-book journal writing task requires both engaging in generative learning but also in retrieval practice. Thus, it remains an open question as to whether beneficial effects of journal writing are partly due to retrieval practice, and how to most efficiently design journal writing as a follow-up learning activity.

Research Questions:

First, we examine whether and why the book format (open vs. closed) in follow-up journal-writing and -revision tasks would affect students' learning outcomes after 1 week. Theoretically, closed-book format in journal writing is more likely to additionally prompt

retrieval practice, which is assumed to consolidate learners' mental representations and thereby to enhance long-term retention (Roelle et al., 2023). However, with the closed-book format, generative-learning benefits can only be utilized to the extent that learners successfully retrieve the relevant lecture contents. In the open-book format, the whole lecture script is openly available to be reprocessed, and thereby retrieval practice is less likely to occur.

Second, we examine whether and why a hybrid-book format is more effective, compared to a same-book format. Theoretically, a hybrid format can combine the advantages of open- and closed-book formats, that is, profit from both generative learning of potentially the complete lecture materials (open-book format) and retrieval practice (closed-book format). We hypothesized that learning outcomes would be enhanced in a hybrid-book, compared to a same-book format.

Third, we will explore which sequence of open- and closed-book journal-writing tasks would learners profit most from. Working in an open-book format first would help students to engage in generative learning and to build a sufficient knowledge base before they work in a closed-book format, requiring them to retrieve information from memory. However, the reversed sequence of working in a closed-book format first would help students to identify knowledge gaps that could be targeted efficiently for remediation in the revision phase with open-book format.

Methods:

In this study (N = 312), we employed a one-factorial, between-subjects design with four experimental groups: closed-book-only, closed-book-before-open-book, open-book-before-closed-book, and open-book-only. More specifically, participants took a pretest, watched a 20-min video lecture on Piaget and his Theory of Cognitive Development, and were subsequently instructed to write a learning journal. They were randomly assigned to either an open-book format (i.e., they could assess the lecture transcript while writing the learning journal) or to a closed-book format (i.e., they had to retrieve relevant information from memory). All participants were supported with six cognitive prompts to organize and elaborate on the learning content for at least 20 minutes, and then had the opportunity to revise their learning journals for at least 20 minutes. Again, they were randomly assigned to either an open- versus closed-book format. Finally, participants received a posttest after 1 week, involving questions on retention, comprehension, and transfer.

Results and Discussion:

We are currently in the process of data coding. Based on the predictions, we will conduct confirmatory and exploratory data analysis. The present study is of both theoretical and educational relevance, as it investigates possible boundary conditions and sheds light on the mechanisms underlying journal writing, a widely used meaningful follow-up learning activity.

Students' Academic Self-Concept and Conscientiousness Moderate the Effectiveness of Distributing Non-Interactive Teaching

Heike Russ¹, Leonie Sibley¹, Salome Flegr², Jochen Kuhn², Vincent Hoogerheide³, Katharina Scheiter⁴, Andreas Lachner¹

¹University of Tübingen, Germany, ²LMU Munich, Germany, ³University of Utrecht, Netherlands, ⁴University of Potsdam, Germany

Theoretical Background:

Non-interactive teaching is regarded as a powerful generative learning activity in which students teach the previously learned contents to a fictitious peer. Non-interactive teaching may help students construct a meaningful mental representation of the learning contents, ideally contributing to (meta-)cognitive learning outcomes (Fiorella, 2023). However, its effectiveness is relatively modest and varied across studies (Kobayashi, 2024; Lachner et al., 2021), suggesting that non-interactive teaching may require additional support to fully exploit its potential and that effectiveness may depend on students' prerequisites. Moreover, the generalizability to classroom settings and the examination of potential lasting learning effects remains uncertain.

Research Questions:

Against this background, we aimed to test whether the effectiveness of non-interactive teaching can be improved by 1) adding drawing as a visual-spatial generative activity to the act of verbal teaching, and by 2) distributing the generative activity across the study phase, both regarding immediate and lasting (meta-)cognitive outcomes (after eight weeks). Moreover, we explored potential interaction effects between the learning activity (restudy, teaching-only, teaching+drawing) and the timing factor (after-study, distributed), and explored whether cognitive, motivational, and personality differences affected effectiveness of non-interactive teaching. We tested the following preregistered hypotheses regarding (meta-)cognitive learning: First, teaching activities (teaching-only, teaching+drawing) should outperform restudying (teaching-hypothesis). Second, teaching+drawing should outperform teaching-only (drawing-hypothesis). Third, distributed conditions should outperform the after-study conditions (distributing-hypothesis).

Methods:

We conducted a controlled field experiment with secondary school students (N = 317; power > .90; M = 12.37 years; 49.83% female) attending an inquiry-based physics unit on converging lenses. Following a 2 × 3 design, students engaged in a learning activity after an entire study phase (after-study) or distributed across three time points during the study phase (distributed). For the learning activity, students either taught contents to a fictitious peer (teaching-only), taught and drew (teaching+drawing), or restudied the learning contents (restudy). Time on task was kept constant, and activities were open-book. We assessed cognitive learning outcomes using a validated conceptual knowledge test ($\alpha = .72$; Wörner et al., 2022). For metacognitive learning outcomes, monitoring accuracy was calculated as difference between students' estimated and actual performance (Schraw, 2009). We tested our preregistered hypotheses with planned contrasts, controlling for prior (meta-)cognitive knowledge: (1) restudy vs. non-interactive teaching, (2) teaching-only vs. teaching+drawing, (3) after-study vs. distributed activities. In moderation analyses with the contrast coded predictors, the potential effects of prior knowledge, academic self-concept, and conscientiousness were examined.

Results:

Analyses showed no significant differences between conditions regarding students' prerequisites, i.e., prior (meta-)cognitive knowledge, academic self-concept, conscientiousness, age, gender, native language ($p \geq .359$). For immediate cognitive outcomes, contrast analyses revealed in line with our teaching-hypothesis that non-interactive teaching (teaching-only, teaching+drawing) resulted in better learning than restudying ($\beta = .11$, $p = .006$). However, contrary to our drawing- and distributing-hypothesis, adding drawing and distributing the activities did not enhance learning ($p \geq .132$). Interestingly, teaching after the study phase was more effective than distributed teaching ($\beta = -.29$, $p = .029$). For immediate metacognitive outcomes,

contrary to our teaching-hypothesis, restudying was better than teaching ($\beta = -.09$, $p = .014$). There were no significant differences regarding the drawing- and distributing-hypothesis ($p \geq .326$). Overall, no lasting learning effects were observed ($p \geq .114$). Explorative analyses indicated that higher academic self-concept and conscientiousness in physics led to greater benefits of after-study teaching ($\beta = .25$, $p = .041$; $\beta = .32$, $p = .005$).

Theoretical and Educational Relevance:

The results provide valuable directions for optimizing non-interactive teaching and highlight the importance of considering individual differences in educational practice. They also suggest further research is needed to explore additional support and strategies to enhance non-interactive teaching.

Academic Self-Concept Determines Lasting Learning of Non-Interactive Teaching

Leonie Sibley¹, Julie Thommes¹, Julian Roelle², Kou Murayama¹, Andreas Lachner¹

¹University of Tübingen, Germany, ²Ruhr University Bochum, Germany

Theoretical Background:

Non-interactive teaching is considered an effective learning strategy to promote students' (meta-)cognitive learning (Brod, 2020; Fiorella, 2023). In line with the generative learning theory, teaching helps student to select the most important information to organize the information in a coherent and logical structure, and to integrate the previously acquired contents with their prior knowledge to be able to make sense out of the learned contents (Fiorella & Mayer, 2016). However, non-interactive teaching seems to be not effective in general. Prior studies showed mixed effects regarding students' (mega-)cognitive learning and large variances among studies (see Lachner et al., 2021, for meta-analytical evidence) which may be an indicator that specific factors determine its effectiveness.

Prior studies started to investigate potential influencing factors. Jacob et al. (2022), for instance, explored the influence of students' academic self-concept. Interestingly, only students with low academic self-concept benefitted from teaching but not students with high academic self-concept. However, this finding was of explorative nature. Additionally, no potential lasting learning effects have been examined. Research is needed that systematically examines (lasting) effects of academic self-concept on learning-by-teaching.

Research Questions:

In this study, we aimed at closing this research gap by systematically manipulating students' academic self-concept to examine whether non-interactive teaching is more beneficial for students with low than with high self-concept regarding their immediate and lasting cognition (comprehension) and meta-cognition (monitoring accuracy). We conducted an online experiment to test this pre-registered hypothesis.

Methods:

We applied a 2x2 design with academic self-concept (high versus low) and learning activity (teaching versus retrieval practice as control group) as between-subjects-factor. In total, $N = 228$ participants took part in the study (recruiting via Prolific). However, after exclusion criteria (i.e., German is native language, attention checks), $N = 216$ participants remained as the final sample ($M = 30.53$ years, 50.46 % female). The participants provided consent to participate, answered questions regarding their academic self-concept, answered a prior knowledge test, and judged their current knowledge (monitoring accuracy). After the prior knowledge test, we randomly provided the participants with fictitious feedback of their performance to either increase (Figure 1, left side) or decrease their academic self-concept (Figure 1, right side). Afterwards, they again answered questions regarding their academic self-concept (manipulation check) and read a text about atomic theory. Then, they either orally taught the learned content to a fictitious peer or to retrieve it in the same amount of time. Finally, they answered the posttest and judged their current knowledge. After two weeks, they again answered the posttest.

Results:

Preliminary results indicated no differences among conditions regarding students' pre-requisites (e.g., prior knowledge, academic self-concept; $ps \geq .229$). Our manipulation of the academic self-concept was successful, as students in the high self-concept condition judged their academic self-concept after the manipulation as higher than students in the low self-concept condition ($p < .001$).

We used contrast analyses to test our pre-registered hypothesis. Regarding students' immediate comprehension, the contrast almost reached significance ($p = .069$). Interestingly, the contrast reached significance in the delayed test ($p = .039$). An additional linear regression revealed that participants with low self-concept indeed benefitted more from teaching compared to participants with high self-concept ($p = .025$). There were no effects regarding monitoring accuracy ($ps \geq .351$).

Conclusion:

The aim of the study was to systematically examine the effect of students' academic self-concept on the non-interactive teaching effect. Results revealed that teaching is particularly helpful for students with low academic self-concept do enhance their lasting learning.

Facilitating the Interleaving Effect in Physics Through Cooperation

Maria Danzglock, Martin Hänze

University of Kassel, Germany

Theoretical Background and Research Questions:

Interleaved practice—in which practice tasks within a domain are alternated among different topics—has demonstrated significant learning success (Mayfield & Chase, 2002; Rohrer & Taylor, 2007; Rohrer et al., 2014; Sana & Yan, 2022; Taylor & Rohrer, 2010). Although this method introduces greater challenges during the learning process, it leads to superior outcomes, particularly for topics that share similarities but demand different solution strategies (Carvalho & Goldstone, 2019). However, most studies on interleaved practice have primarily focused on mathematics. Therefore, it remains unclear whether its effectiveness extends to other domains with a different knowledge complexity (e.g., physics). It is uncertain if the benefits of interleaving diminish with more complex learning materials. If so, could cooperation compensate for these potential disadvantages, given that it is an instructional design element particularly beneficial for tackling complex tasks (Laughlin et al., 2003)?

Therefore, this study aims to investigate whether the interleaving effect is present in real classroom settings when practicing complex physics concepts and to examine if cooperative learning can enhance these effects. We hypothesize that learners who

engage in interleaved practice will achieve greater learning outcomes than those who practice in blocked sequences, especially when they learn cooperatively.

Methods:

A field-experimental study was conducted with 407 12th-grade students in German schools. The study focused on the movement of charged particles in electric and magnetic fields. The design included an initial 90-minute lecture and a 20-minute follow-up test eight weeks later. In the initial session, baseline knowledge was assessed via a pretest, followed by a learning phase incorporating an educational game. In the blocked condition, learners solved 36 game scenarios, with the first 18 related to the magnetic field and the last 18 to the electric field. In the interleaved condition, learners faced 36 game scenarios for both fields in an alternating manner. Subsequently, learners completed a knowledge test (Berger et al., 2019). Eight weeks later, a follow-up test and a questionnaire on self-concept and domain interests were administered.

Results:

The participants were randomly assigned into the following groups: interleaved individual ($n = 103$), interleaved cooperative ($n = 104$), blocked individual ($n = 98$), and blocked cooperative ($n = 102$). A multiple linear regression analysis was conducted to address the research questions, with learning success as the dependent variable. Independent variables included prior knowledge, practice type (interleaved or blocked), social format (cooperative or individual), and their interaction. Prior knowledge was found to have a significant positive effect on learning success ($p = .007$). The primary effects of interleaving and cooperation were not statistically significant. However, the interaction between cooperation and interleaved practice demonstrated a notable effect ($p = .025$), indicating that interleaved practice in conjunction with cooperative learning is associated with enhanced learning success. These findings were also evident in the follow-up test conducted eight weeks later (prior knowledge: $p = .021$; interleaving*cooperation: $p = .004$).

Discussion:

The findings suggest that interleaved practice, when integrated with cooperative learning, is advantageous for comprehending the movement of charged particles in electric and magnetic fields, with these positive outcomes persisting over time. This study reveals that interleaved practice is effective for teaching complex conceptual physics knowledge in real classroom environments, as long as students are involved in cooperative learning.

S-74: Historisch denken und (mit-)fühlen lernen. Arbeit mit Zeitzeug*innen ein Beitrag für die Demokratiebildung?

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 150

Chair der Sitzung: Christiane Bertram

Chair der Sitzung: Martin Nitsche

Historisch denken und (mit-)fühlen lernen. Arbeit mit Zeitzeug*innen ein Beitrag für die Demokratiebildung?

Chair(s): **Christiane Bertram** (Universität Tübingen, Deutschland)

Diskutant*in(nen): **Martin Nitsche** (Pädagogische Hochschule FHNW)

Wie kann eine interdisziplinär angelegte empirische Bildungsforschung bei dem Umgang mit beziehungsweise der Bewältigung von multiplen gesellschaftlichen Herausforderungen helfen (CfP GEBF 2025, S. 2)? Angesichts von Ressourcenknappheit und Bedrohungsszenarien fällt es den liberalen Demokratien schwer, überzeugende Lösungen zu finden, so dass extremistische und populistische Parolen und Narrativen (z.B. zur „Remigration“ oder Leugnung des Klimawandels) verfangen (Richter & Ulrich, 2024, S. 166). Kontrovers diskutierte historische Themen, die in den sozialen Medien verstärkt werden, bieten den Raum für Fake News und Filterblasen (Wansink et al., 2023). Eine Demokratie ist auf informierte, reflektierte und empathische Bürger*innen angewiesen. Die Schulen und im Besonderen auch die Fächer Politik und Geschichte sind bei der Demokratiebildung gefordert (SWK, 2024; Wilke et al. 2022), denn die jüngsten Wahlerfolge rechtsextremistischer und populistischer Parteien sind alarmierend.

Einschätzungen von Zeitzeug*innen sind kontrovers – ob es um die deutsch-deutsche Geschichte oder um den Israel-Palästina-Konflikt geht. Gerade die Multiperspektivität, die sich in den individuellen Lebensgeschichten widerspiegelt, bietet die Chance, den Lernenden einen kritischen und differenzierenden Umgang mit persönlichen Argumenten und gleichzeitig Empathie für andere Lebensgeschichten zu vermitteln (Bertram, 2017; Bartelds, 2023). Dabei ist es wichtig, den „unnatural act“ (Wineburg, 2001) des historischen Denkens (Körber, Schreiber & Schöner, 2007) oder Abwägens (van Dri & van Bortel, 2008) explizit zu machen. Denn historisches Denken basiert auf der Rekonstruktion der Vergangenheit, ist narrativ verfasst und bedeutet, ausgehend von einer historischen Frage Quellen und Darstellungen methodisch kontrolliert auszuwerten (Rüsen 1983; 2013). Ziel des Geschichtsunterrichts ist die Förderung eines reflektierten und (selbst-)reflexiven Geschichtsbewusstseins, das bedeutet u.a. die Bereitschaft, in der Geschichte Eigenes (Identität) und Fremdes (Alterität) zu erkennen und ggf. bereit für die Veränderung von Einstellungen zu sein.

In dem Symposium wollen wir drei empirische Studien vorstellen, in denen der Zusammenhang zwischen Zeitzeug*innen und historischem Denken bzw. Empathie in einem weiten Rahmen untersucht wird. In den ersten beiden Beiträgen werden Unterrichtsstudien vorgestellt, in denen multiperspektivisch mit Zeitzeug*innen gearbeitet wurde. In Beitrag 1 (Thema: Israel-Palästina-Konflikt) steht die Förderung von Empathie im Vordergrund, in dem Beitrag 2 (deutsch-deutsche Geschichte) geht es um kognitive Outcomes (historische Kompetenzen) und um den Vergleich der Wahrnehmung von Live- und Video-Zeitzeug*innen. In Beitrag 3 werden mit den lebensgeschichtlichen Erzählungen von Erwachsenen aus Ost- und Westdeutschland die Ergebnisse eines historischen Lernprozesses hinsichtlich der spezifischen Verbindung kognitiver und emotionaler Anteile untersucht. Das Symposium wird von Martin Nitsche (Pädagogische Hochschule FHNW, Schweiz), Experte für den Zusammenhang von epistemologischen Überzeugungen und historischen Kompetenzen, abschließend kommentiert.

Beiträge des Symposiums

Förderung der Empathie mit Hilfe von Zeitzeug*innen im Geschichtsunterricht. Sechs Gestaltungsprinzipien zur Vermittlung von historischem Einfühlungsvermögen durch Lehrkräfte

Hanneke Bartelds

University of Amsterdam

Theorie

Historische Empathie wird verstanden als ein Prozess der kognitiven und affektiven Auseinandersetzung mit historischen Figuren, um deren Erfahrungen, Entscheidungen oder Handlungen besser zu verstehen und zu kontextualisieren. Historische Empathie ist wichtig, um fachspezifische Ziele (z. B. ein vertieftes Verständnis vergangener Ereignisse), aber auch um staatsbürgerliche Ziele (z. B. das Einnehmen anderer Perspektiven von Menschen in der Gegenwart) zu erreichen (Bartelds et al., 2020, 2023; Barton & Levstik, 2004; Endacott & Brooks, 2013, 2018; Karn, 2023, 2024; Seixas, 2012). Insbesondere die Methode der Oral History, in der weniger nach den Ereignissen, sondern nach ihrer Bedeutung für die Menschen gefragt wird (Portelli, 2009), eignet sich für die Förderung historischer Empathie. Allerdings kann die Vermittlung historischer Empathie mit Herausforderungen durch Präsentismus, Anachronismus und Wertetransfer einher gehen (Davison, 2017; Karn, 2023; Miles & Gibson, 2022). In der vorliegenden Fallstudie wurde untersucht, wie zwei Geschichtslehrkräfte Unterrichtseinheiten mit Zeitzeug*innen über den Konflikt Israel-Palästina umsetzen, bei denen sechs Gestaltungsprinzipien für die Vermittlung von Empathie im Geschichtsunterricht berücksichtigt wurden (Bartelds, et al., 2023): (1) historischen Kontext anbieten, (2) die eigene Position wahrnehmen, (3) Multiperspektivität ermöglichen, (4) persönliche Verbindung unterstützen, (5) kognitive und emotionale Aspekte thematisieren, (6) explizite Erklärung, was historische Empathie ist (Bartelds et al., 2020).

Forschungsfragen

(1) Wie setzen die beiden Geschichtslehrkräfte die sechs Gestaltungsprinzipien für den Unterricht mit Zeitzeug*innen um, um historische Empathie zu fördern?

(2) Welche Überlegungen und Bedenken haben die Lehrkräfte dabei, und wo sehen sie Verbesserungsbedarf?

Methode

Die Teilnehmenden dieser Studie waren die beiden Geschichtslehrkräfte, die die Unterrichtseinheiten der Interventionsstudie unterrichteten, in der in einer Bedingung mit zwei Zeitzeug*innen und in der anderen mit einem Dokumentarfilm gearbeitet wurde. Die Daten wurden – dem Mixed Method Ansatz folgend – auf drei Arten erhoben: 1. Videoaufzeichnungen der Unterrichtsstunden zur Beobachtung des Lehrerverhaltens (insgesamt 20 Unterrichtsstunden); 2. schriftliche Beobachtungsprotokolle aller Unterrichtsstunden durch die Erstautorin; und 3. Interviews mit den beiden Lehrkräften unmittelbar nach Abschluss der Unterrichtseinheiten. Wir analysierten die Videoaufnahmen, Beobachtungen und Interviewtranskripte anhand der sechs Gestaltungsprinzipien als Hauptcodes. Für jedes Gestaltungsprinzip haben wir beschrieben, was die Lehrkräfte im Hinblick auf

die Prinzipien taten und wie sie damit umgingen. Bei der Analyse der Interviewdaten wurden auch die Überlegungen und Entscheidungen der Lehrkräfte in Bezug auf die sechs Gestaltungsprinzipien, ihre Bedenken und Verbesserungsmöglichkeiten ermittelt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass die Kombination aus der Arbeit mit Zeitzeug*innen (live und im Dokumentarfilm) und den sechs Gestaltungsprinzipien den Lehrkräften konkrete Instrumente an die Hand gab, die es ihnen ermöglichten, je nach ihren eigenen Vorlieben und Zielen Schwerpunkte zu setzen und Entscheidungen zu treffen. Alle sechs Gestaltungsprinzipien wurden in die Praxis umgesetzt, wobei die Gestaltungsprinzipien der Kontextualisierung, des persönlichen Bezugs und der Multiperspektivität von den Lehrkräften als besonders wichtig erachtet wurde. Hinsichtlich der eigenen Position bezogen auf den historischen Fall erachteten die Lehrkräfte die Frage nach der eigenen Neutralität als unproblematisch. Der kognitive Ansatz zeigte sich für die Lehrkräfte häufig in der Kontextualisierung, dem Bewusstsein der eigenen Position und der Anerkennung von Multiperspektivität. Der affektive Ansatz beinhaltete hingegen die Herstellung einer persönlichen Verbindung und die Erkundung gemeinsamer menschlicher Erfahrungen und Gefühle. In Bezug auf die explizite Unterweisung ließ sich feststellen, dass es den Lehrkräften schwer fiel, die affektiven Aspekte der historischen und der „staatsbürgerlichen“ Empathie explizit zu vermitteln.

Daher schlagen wir als ein siebtes Gestaltungsprinzip die Verknüpfung von historischem Einfühlungsvermögen und Empathie vor, indem wir eine Verbindung zwischen der Perspektive von Menschen in der Vergangenheit (fachspezifisches Ziel: historische Empathie) und in der Gegenwart (politisches Ziel: staatsbürgerliche Empathie) herstellen.

Schüler*innenwahrnehmung von Zeitzeug*innen der „Generation 1975“ in einer Interventionsstudie (RCT) im Geschichtsunterricht

Katharina Totter, Wolfgang Wagner, Christiane Bertram, Lisa Zachrich, Ulrich Trautwein
Universität Tübingen

Theorie

Zeitzeug*innen im Geschichtsunterricht einzubinden birgt neben (motivationalen) Chancen (Lanman, 1987; Bartelds et al., 2023) auch Risiken für das historische Lernen (Galda, 2013; Lanman, 1987; Obens & Geißler-Jagodzinski, 2010). In einer ersten randomisierten kontrollierten Feldstudie (RCT) zeigte sich, dass Schüler*innen mit Live-Berichten im Klassenzimmer zwar mehr Spaß hatten, aber weniger lernten als die, die mit Video-/Text-Zeitzeug*innen arbeiteten (Bertram, Wagner & Trautwein, 2017). Als Risikofaktoren identifizierten die Autor*innen die wahrgenommene „Aura“ der Zeitzeug*innen, die durch den eindrücklichen Bericht eines oder einer ehemaligen Oppositionellen, der oder die live im Klassenzimmer befragt werden konnte, verstärkt wurde (Bertram, 2017).

In einer folgenden RCT-Studie der Forschungsgruppe wurde in Zusammenarbeit mit dem Zentrum für Schulqualität und Lehrerbildung eine Intervention mit Live- beziehungsweise Video-Zeitzeug*innen entwickelt, die Lehrkräfte mit ihren eigenen Klassen durchführten. Die Intervention wurde in zwei zentralen Punkten verändert: Anstatt Heldengeschichten wurden Alltagsberichte zum Thema Transformationszeit nach 1989/90 gewählt. Außerdem kamen statt einer Person nun zwei Zeitzeug*innen mit jeweils unterschiedlichen Perspektiven (Ost- und West-deutschland) zu Wort.

Differenzielle Effekte zwischen den Bedingungen zeichneten sich nur bezüglich der Evaluation der Unterrichtseinheit deutlich ab. Die Lernenden in der Live-Gruppe schätzten – analog zu den Ergebnissen in der Vorgängerstudie (Bertram et al., 2017) – das motivationale Potenzial der Arbeit mit Zeitzeug*innen signifikant höher ein als die in der Videogruppe. Im Gegensatz zur Vorgängerstudie schätzte die Live-Gruppe ihren inhaltlichen und methodischen Lernfortschritt nicht höher ein als die Video-Gruppe. Bezüglich des Komposits an eingesetzten Kompetenztests zum historischen Denken zeigten sich keine Unterlegenheit der Live-Gruppe gegenüber der Video-Gruppe, bezüglich der Relevanz von Geschichte für das eigene Leben kamen die Schüler*innen der Live-Gruppe aber auch nicht zu einer günstigeren Einschätzung im Vergleich zur Video-Gruppe.

Differenzielle Effekte zwischen der Live- und der Videogruppe waren in dieser didaktisch veränderten Unterrichtseinheit, die aus der Vorbereitung, Durchführung und Auswertung der Zeitzeug*innengespräche bezogen auf die Transformationszeit bestand, kaum zu beobachten. Doch wurde die Wahrnehmung der Begegnung mit den Zeitzeug*innen durch die Schüler*innen unmittelbar nach der Arbeit mit den Video- bzw. Live-Zeitzeug*innen mittels eines psychometrisch validierten Instruments erhoben (Zachrich et al., 2020; Zachrich et al., 2023). In dem hier eingereichten Beitrag sollen die differenziellen Effekte zwischen den Bedingungen bezogen auf die unmittelbare Wahrnehmung der Begegnung untersucht werden.

Fragestellung

1. Gibt es Unterschiede zwischen der Wahrnehmung der Zeitzeug*innen aus den neuen und alten Bundesländern (Ost versus West)?
2. Unterscheidet sich die Wahrnehmung der Zeitzeug*innen aufgrund der Interventionsbedingungen (Live versus Video)?

Methode

Eine Präregistrierung zu Effekten der Intervention ist bei der Registry of Efficacy and Effectiveness Studies (REES) abrufbar (Registry ID: 14881.1v1). Nach randomisierter Zuweisung der 50 Lehrkräfte (n = 23 Live, n = 22 Video, n = 15 Kontrolle) erhielten die Lehrkräfte der Interventionsgruppen eine Fortbildung zur Unterrichtseinheit, die sie am Schuljahresende 2022 unterrichteten. Schüler*innendaten der Interventionsklassen wurden vor sowie nach der Intervention erfasst, die Wartekontrollgruppe nach Lehrplan beschult und analog zu den Interventionsgruppen getestet. Schüler*innen der Interventionsgruppen wurden zusätzlich nach der Begegnung mit den Zeitzeug*innenberichten mit einem erprobten Fragebogen (Zachrich et al., 2020; Zachrich, et al. 2023) zu ihrer kognitiven und emotionalen Involviertheit (beispielsweise, wie sehr sie der Bericht bewegt hatte oder ihrer gefühlten Erkenntnis durch den Bericht) und zu der wahrgenommenen Authentizität befragt (k = 45 Items). Die Einschätzung wurde auf einer sechsstufigen Skala von 1 (trifft überhaupt nicht zu) bis 6 (trifft voll und ganz zu) jeweils einzeln pro Zeitzeug*in aus dem Westen bzw. Osten erfasst. Die Stichprobe umfasste insgesamt 843 Schüler*innen.

Die Auswertung zu den Fragestellungen wurden noch nicht durchgeführt, da eine Präregistrierung vorab geplant ist. Insgesamt sollen die vertiefte Analysen Anhaltspunkte für die (unterschiedliche) Wirkung der Interventionen mit Live- oder Video-Zeitzeug*innen liefern, die im Symposium diskutiert werden sollen.

Erzählweisen über die deutsch-deutsche Geschichte als Ausprägungen des "Geschichtsbewusstseins"

Christiane Bertram
Universität Tübingen

Theoretischer Hintergrund

Die deutsch-deutsche Teilung und der Verlauf der Transformationszeit wird häufig für die Erklärung von Unterschieden zwischen Ost- und Westdeutschland herangezogen (z.B. Kowalczyk, 2019; Mau, 2024; Oschmann, 2023). Während im Westen die Auswirkungen der Wiedervereinigung auf den Lebensalltag minimal waren, wurde das Leben aller in Ostdeutschland auf den Kopf gestellt (Bösch, 2015; Großbölting, 2020). Im Interviewprojekt „Generation 1975“ wurden 26 Menschen, die in der DDR bzw. BRD aufgewachsen sind, zu ihrer Erinnerung an einen geteilten und wiedervereinigten Staat befragt. Um trotz der unterschiedlichen Erfahrungen ein gemeinsames Raster für den Vergleich der „Erzählweisen“ zu finden, wurde das Konzept des „Geschichtsbewusstseins“ – d.h.: der narrativ hergestellte Zusammenhang zwischen den Zeitdimensionen Vergangenheit, Gegenwart und Zukunft – herangezogen, welches sich durch „eine komplexe Interaktion und Integration verschiedener kognitiver und affektiver Hinwendungen zum geschichtlichen Gegenstand“ (Jeismann, 1988, vgl. auch Borries, 2014; Brauer & Lücke, 2021; Zachrich et al. 2020) auszeichnet. In der vorgelegten Studie werden aus dem Datenmaterial historische Erzählweisen erarbeitet, die für die Untersuchung der gesamten Stichprobe genutzt werden können.

Forschungsfragen

- (1) Welche typischen historischen Erzählweisen zeigen sich in Bezug auf die Gesamtstichprobe?
- (2) Lassen sich hinsichtlich der Verteilung der Erzählweisen Unterschiede zwischen den Interviewpartner*innen aus Ost- und Westdeutschland erkennen?

Methode

Die Studie lässt sich der Fallanalyse und Typenbildung zuordnen, bei der „Idealtypen“ gewonnen werden sollen, die als „Konstruktionen theoretischer Art“ (Kuckartz, 2006, S. 4051) die Wirklichkeit verständlicher machen und erklären (Bohnsack, Hoffmann & Nentwig-Gesemann, 2018; Haas & Scheibelhofer, 1998; Kelle & Kluge, 2010). Als Basis für die Typenbildung wird in unserer Studie ein „Merkmalsraums“ mit zwei Dimensionen konstruiert. Diese sind dichotom ausgerichtet, so dass vier Felder entstehen, die jeweils einen „Typ“ beschreiben (Kuckartz, 2006, S. 4051-4052). In einem dreischrittigen Verfahren (offene, axiale und selektive Codierung) wurden traditionelle Dokumentationen wie Forschungstagebücher und Datensitzungsprotokolle genutzt sowie auch die Software MAXQDA, in der die Verwendung von Stereotypen basierend auf dem Stereotyp Content Model von Fiske et al. (2020) – hier wird zwischen neidischen, bewundernden, verachtenden und paternalistischen Stereotypen unterschieden – codiert wurde.

Ergebnisse und ihre Bedeutung

Die Analysekategorien der kognitiven Gestaltung (von durchdacht und strukturiert bis intuitiv und spontan) und der emotionalen Nähe versus Ferne zum Erzählgegenstand ermöglichen die Differenzierung von vier Erzählweisen, die wir im Westen und Osten gefunden haben:

Vorsichtige Erzählweise: Es wird überlegt, was wie erzählt wird, zeitliche Differenzierung, bewusstes Aussparen mancher Themen

Orientierte Erzählweise: Einbezug verschiedener Perspektiven, Mitgefühl, Hintergrundwissen, zeitliche Differenzierung, Reflexion der eigenen Entwicklung

Selbstbezogene Erzählweise: Deutliche Abgrenzung (Othering): „Wir“ – „die anderen“, eher negative Erinnerungen

Berührte Erzählweise: Erinnerungen an Gemeinsamkeiten, Offenheit den „anderen“ gegenüber, eher positive Erinnerungen

Die Interviewpartner*innen aus dem Osten wurden etwas häufiger dem Typ der orientierten (Westen: 3; Osten: 4) und die aus dem Westen der selbstbezogenen Erzählweise (Westen: 3; Osten: 2) zugeordnet. Zwei aus dem Osten und vier aus dem Westen wurden der berührten Erzählweise zugeordnet, bei der vorsichtigen Erzählweise waren es jeweils drei. Bei einem Vergleich der Stereotype fiel auf, dass die Menschen aus dem Westen mit insgesamt 166 Stereotypen (produzierte und reflektierte) deutlich mehr Stereotype nannte als die Menschen aus dem Osten (115 produzierte und reflektierte Stereotype). Im Weiteren produzierten Menschen im Westen deutlich mehr Stereotype (157) als dass sie diese reflektierten (103), während die Menschen im Osten umgekehrt mehr Stereotype reflektierten (75) als produzierten (32).

Diskussion / Transfer

Abschließend sollen das methodische Vorgehen und die Aussagekraft der Studie (N = 26) diskutiert werden. Das Ergebnis, dass die Stereotype und die distanzierten Erzählweisen eher im Westen zu finden ist, überrascht nicht wirklich, sind aber höchst relevant angesichts eines gesellschaftlichen Diskurses, in dem häufig der „Osten“ als Problem und Abweichung dargestellt wird (Oschmann, 2023).

S-75: Museen als Orte informeller Bildung

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: O 131

Chair der Sitzung: Alexandra W. Busch

Chair der Sitzung: Olaf Köller

Museen als Orte informeller Bildung

Chair(s): **Alexandra W. Busch** (Leibniz-Zentrum für Archäologie, Mainz), **Olaf Köller** (IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik, Kiel)

Museen sind nicht nur Orte von Geselligkeit und Freizeitgestaltung, sondern als informelle Lernorte eine wichtige Säule des lebenslangen Lernens. Die neuen Stichworte der Museen sind Besucher*innenorientierung, Dialog mit der Gesellschaft, Partizipation und gesellschaftliche Relevanz. Museen sind dabei, eine aktivere Rolle in der Gesellschaft einzunehmen: weg vom Musentempel und dem „neutralen“ Ort, hin zu relevanten gesellschaftlichen Akteuren, die sowohl Dialog als auch aktive Beteiligung unterschiedlicher Interessensgruppen fördern und gestalten und auch Werteorientierungen geben möchten. Mit ihren Ausstellungen und Bildungsprogrammen zielen Museen, über die Vermittlung von Sachinhalten hinaus, u. a. darauf ab, Kompetenzen wie Reflexionsfähigkeit, kritisches Denken, evidenzbasierte Meinungsbildung und in ausgewählten Fällen auch die Bildung von Verhaltensintentionen zu fördern. Um diese Ziele zu erreichen, ist eine enge, kooperative Zusammenarbeit zwischen empirischer Lehr-Lernforschung und musealer Vermittlung notwendig. Durch solche Kooperationen können theoriegeleitet und forschungsbasiert neue Standards und innovative Formate für eine nachhaltige museale Vermittlung, wie z.B. maßgeschneiderte Angebote für unterschiedliche Zielgruppen, entwickelt werden. Interessante Fragen dabei sind u.a., welche Aspekte und persönliche Merkmale dazu beitragen, dass Menschen zu (regelmäßigen) Museumsbesucher*innen werden und inwieweit Führungen von den Besucher*innen als pädagogischer Vermittlungsprozess wahr- und angenommen werden. Bei der Optimierung der Angebote durch Wirksamkeitsüberprüfung der Ausstellungsformate wiederum geht es nicht nur um das Erfassen des Wissenserwerbs, sondern z.B. auch um die Einsicht in das eigene Wissen und das Messen von Verhaltensintentionen. Die Verbindung von Theorie und Praxis spielt in der Museumsforschung eine zentrale Rolle: Durch eine enge Zusammenarbeit können Wirksamkeitsstudien vor Ort durchgeführt und diese Forschungserkenntnisse vermehrt in der Bildungspraxis in den Museen umgesetzt werden. In diesem Symposium, organisiert vom Leibniz-Kompetenzzentrum Bildung im Museum, werden unterschiedliche museale Bildungsforschungsprojekte vorgestellt: Der erste Beitrag geht anhand verschiedener Museumstypen der Frage nach, welche Faktoren die Häufigkeit beeinflussen, mit denen Menschen Museen besuchen. Hier spielen insbesondere das kulturelle Kapital und der Bildungshintergrund eine Rolle, aber auch z.B. die Frage wie oft die Personen in ihrer Kindheit schon Museen besucht haben. Im zweiten Beitrag geht es um Persönlichkeit, familiäres kulturelles Kapital und Interessen im Jugendalter als Prädiktoren des musealen Besuchsverhaltens im mittleren Erwachsenenalter. Hier zeigt sich der Einfluss des familiären kulturellen Kapitals und der Persönlichkeitsvariable Offenheit auf die Besuchshäufigkeit in Museen. Der dritte Beitrag schaut im Rahmen einer Interviewstudie auf Museumsführungen für Erwachsene als Forschungsfeld und stellt die Frage: Können Führungen als pädagogische Kommunikation beschrieben werden? Der vierte Beitrag erforscht den Zusammenhang von Besucherleben, Wissenserwerb und Verhaltensintentionen in einer Ausstellung über Korallenriffe. Er zeigt, dass eine starke Verhaltensintention direkt nach einem Ausstellungsbesuch positiv mit einer kritischen Auseinandersetzung mit den Ausstellungsinhalten und dem Gefühl eigener Verantwortung zusammenhängt. Die Stärke der Verhaltensintention weist zudem einen positiven Zusammenhang mit dem selbstberichteten Ausmaß der Umsetzung der Intention fünf bis acht Wochen nach dem Ausstellungsbesuch auf. Der fünfte Beitrag beschäftigt sich mit der verbesserten Einsicht in das eigene Wissen bzw. Nicht-Wissen durch einen Ausstellungsbesuch. Hier werden metakognitive Einflussmöglichkeiten eines Ausstellungsbesuchs erfasst. Die Ergebnisse legen nahe, dass Museumsausstellungen nicht nur Wissen, sondern auch die Einsichtsfähigkeit in das eigene Wissen erhöhen kann.

Beiträge des Symposiums

Einflussfaktoren auf die Besuchshäufigkeit in Museen – eine vergleichende Studie über verschiedene Museumstypen

Lorenz Kampschulte¹, Gun-Brit Thoma²

¹Deutsches Museum, München, ²IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik, Kiel

Museen nehmen zunehmend neue Rollen ein, wobei ein Schwerpunkt auf dem Dialog mit der Gesellschaft und ihren Partizipationsmöglichkeiten liegt. Dabei sollen sie möglichst die gesamte Gesellschaft in ihrer Vielfalt ansprechen. Entsprechend relevant ist für Museen das Wissen darum, welche Menschen Museen besuchen, wie oft und warum. Theoretische Erkenntnisse und empirische Befunde weisen auf die Bedeutung sozioökonomischer Merkmale für die Nutzung kultureller Aktivitäten wie Museumsbesuche hin. Besonders das kulturelle Kapital und der Bildungshintergrund haben einen Einfluss auf die Besuchshäufigkeit von Museen (vgl. Thoma et al., 2022), wobei sich Besuchshäufigkeit je nach Museumstyp deutlich voneinander zu unterscheiden scheint. Personen, die häufiger Kunstmuseen oder/und kulturhistorische Museen besuchen, verfügen über einen höheren Bildungshintergrund (Wegner, 2016) und ein höheres kulturelles Kapital. Darüber hinaus ist bekannt, dass das kulturelle Kapital und ein hoher Bildungshintergrund mit entwicklungsförderlichen Eltern-Kind-Aktivitäten wie Museumsbesuche einhergehen (vgl. Wößmann et al., 2023; Biedinger & Klein, 2010). Insofern kulturelles Kapital intergenerational weitergegeben wird (Burkhard et al., 2024), kann davon ausgegangen werden, dass auch Personen, die heute häufig Museen besuchen, schon als Kind häufig an kulturellen Aktivitäten wie Museumsbesuchen teilgenommen haben. Daraus ergeben sich folgende Fragestellungen:

- 1.) Unterscheiden sich verschiedene Museumstypen in ihren Besuchenden im Hinblick auf die allgemeine Besuchshäufigkeit von Museen und damit auf das kulturelle Kapital und den Bildungshintergrund?
- 2.) Haben Personen, die heute häufig Museen besuchen, auch in ihrer Kindheit öfter Museen besucht? Unterscheiden sich die Museumstypen im Hinblick auf die Häufigkeit von Kindheitsbesuchen?

Datengrundlage ist die Besucher:innenstrukturanalyse 2022/2023 des Leibniz-Kompetenzzentrums Bildung im Museum, die in 22 Museen und Science Centern in Deutschland und Österreich zu jeweils drei Erhebungszeitpunkten durchgeführt wurde. Insgesamt umfasst die Stichprobe fünf Naturkundemuseen (N = 5.391), sechs Naturwissenschaftlich-technische Museen (N = 4.246), sechs Kulturhistorisch-archäologische Museen (N = 2.574), zwei Kunstmuseen (N = 1.070) und drei Science Center (N = 1.273). Der Bildungshintergrund wurde in Anlehnung an PISA als höchster (angestrebter) Schulabschluss abgefragt (Mang et al., 2021). Das kulturelle Kapital wurde anhand der vielfach validierten „Bücherfrage“ (Sieben & Lechner, 2019), die Häufigkeit der Kindheitsbesuche anhand von vier Antwortmöglichkeiten (sehr oft / hin und wieder / nie / keine Erinnerung) erfasst. Zur

Ermittlung von Unterschieden zwischen den Museumstypen bzw. den drei Besuchshäufigkeitsgruppen (habituelle / Gelegenheits- / Einmal-Besuchende) wurden deskriptive Kennwerte sowie Chi-Quadrat-Tests nach Pearson berechnet.

Die Ergebnisse zeigen, dass Kunstmuseen von besonders vielen habituellen Besuchenden aufgesucht werden (69% der Gesamtbesuchenden), gefolgt von Kulturhistorisch-archäologischen Museen (55%) und Naturkundemuseen (47%). Der Anteil Einmal-Besuchender (erster Museumsbesuch in den letzten 12 Monaten) ist in allen Museen eher niedrig, wobei Science Center mit 16% den verhältnismäßig höchsten Anteil Einmal-Besuchender haben, gefolgt von Naturwissenschaftlich-technischen Museen (16%) und Naturkundemuseen (16%). Schulbildung und kulturelles Kapital steigen erwartungsgemäß mit zunehmender Nutzung der Museen. Interessant ist, dass das Schulbildungsniveau für alle Museumstypen ähnlich ist und jeweils in etwa gleich von Einmal-Besuchenden zu habituellen Besuchenden ansteigt, das kulturelle Kapital hingegen bei den Kulturhistorisch-archäologischen und insbesondere bei den Kunstmuseen schon bei Einmal- und Gelegenheitsbesuchenden sehr stark ausgeprägt ist (Kunstmuseen: Einmal-Besuchende 68%, Gelegenheitsbesuchende 73%, habituelle Besuchende 79% mit hohem kulturellem Kapital (>100 Bücher im Haushalt)). Bezogen auf die Museumsbesuche in der Kindheit zeigt sich, dass etwa 33-38% der habituellen Besuchenden über alle Museumstypen hinweg schon „sehr oft“ als Kind Museen besucht haben. Sowohl strukturelle Faktoren als auch Kindheitsbesuche scheinen einen Einfluss auf Museumsbesuche im Erwachsenenalter zu haben. Gerade mit Blick auf die Kindheitsbesuche wird die Bedeutung einer frühen Berührung mit Museen für eine spätere Nutzung als Informations-, Lern- und Freizeitort deutlich, um die Rolle der Museen als gesellschaftliche Akteure weiter zu stärken.

Persönlichkeit, familiäres kulturelles Kapital und Interessen im Jugendalter als Prädiktoren des musealen Besuchsverhaltens im mittleren Erwachsenenalter

Alexandra Valerie Dicks, Ricarda Ullrich, Gun-Brit Thoma, Olaf Köller

IPN Leibniz-Institut für die Pädagogik der Naturwissenschaften und der Mathematik, Kiel

Museen stellen als informelle Lernorte und Kulturstätten wesentliche Eckpfeiler lebenslangen Lernens dar. Über den inkrementellen Wert der Teilhabe an Kultur und dem materiellen wie immateriellen Kulturerbe hinaus können positive Effekte von Museumsbesuchen auf das Wohlbefinden, etwa das subjektive Wohlbefinden und die Lebenszufriedenheit, erwartet werden (Cotter & Pawelski, 2022; Šveb Dragija & Jelinčić, 2022). Allerdings zeigen Befragungen in Museen, dass Besuchende mit hohen sekundären und tertiären Bildungsabschlüssen überproportional vertreten sind (für die deutsche Museumslandschaft siehe Leibniz-Besuchendenstrukturanalyse, Thoma et al., 2022). Dieser Disparität gilt es zu begegnen, als dass Museen als Kultur- und Bildungsinstitutionen die gesamte Gesellschaft ansprechen sollten. Um ihre gesellschaftliche Relevanz zu stärken und sie zu unterstützen sich für weitere Zielgruppen zu öffnen, ist es entscheidend, diejenigen Faktoren zu verstehen, die das Besucherverhalten bestimmen. Soziologische Forschungsarbeiten betonen die Bedeutung struktureller Familienmerkmale für die intergenerationelle Weitergabe, wobei ein bestimmter kultureller Habitus, teilweise als diffuser Korpus nicht-kognitiver Fähigkeiten, von Eltern auf ihre Kinder übermittelt wird (Jæger, 2022; Kröner, 2013). Dabei spielen gemeinsame kulturelle Aktivitäten innerhalb der Familie als auch die von den Eltern vorgelebten Aktivitäten eine zentrale Rolle. Allerdings bleibt der aktive und passive Erwerb von kulturellem Kapital durch Kinder und Jugendliche selbst als Akteure in der bisherigen Forschung weitgehend unberücksichtigt (Burkhard, 2024). Dem gegenüber heben psychologische Forschungsarbeiten die Bedeutung der Persönlichkeit für das museale Besucherverhalten hervor. Die drei Facetten der Persönlichkeitsdimension Offenheit für Erfahrungen des Fünf-Faktoren-Modells zeigen sich beispielsweise unterschiedlich prädiktiv für verschiedene Museumstypen (Meyer et al., 2023). Personen, die hohe Werte in dieser Persönlichkeitsdimension aufweisen, neigen dazu, häufiger Museen zu besuchen, da sie etwa neue ästhetische, immersive Erfahrungen und intellektuelle Herausforderungen suchen. Obwohl das Niveau der Offenheit ab dem Erwachsenenalter recht zeitstabil bleibt, fanden sich in einer repräsentativen niederländischen Stichprobe über den Lebensverlauf Transaktionen von Offenheit und der Häufigkeit kultureller Aktivitäten. Für die anderen Persönlichkeitsdimensionen waren die gemeinsamen Entwicklungstendenzen weniger konsistent (Schwaba et al., 2018). Bislang sind beide Forschungsstränge jedoch wenig miteinander verknüpft (Kröner, 2013). Der vorliegende Beitrag setzt hier an und greift verschiedene Einflussfaktoren aus der (Bildungs-)Biografie auf: Mittels Strukturgleichungsmodellierung (SEM) werden die Persönlichkeitsdimension Offenheit für Erfahrungen, das familiäre kulturelle Kapital, der eigene kulturelle Wert in der Jugend sowie das verbale und naturwissenschaftliche Sachinteresse in der Jugend und das jeweils entsprechende domänenspezifische Selbstkonzept der Begabung als Prädiktoren für die Häufigkeit von Museumsbesuchen im mittleren Erwachsenenalter zusammengeführt. Abhängige Variablen sind die Besuchshäufigkeit in Museen allgemein sowie in spezifischen Museumstypen. Dazu werden vier Wellen des longitudinalen Datensatzes „Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter“ (BIJU, Baumert et al., 1996) zwischen 1992 (Klasse 7) und 2018 (MAlter = 40) verwendet (Nimputiert = 8459, für Details zum Erhebungsdesign siehe Becker et al., 2019). In allen Modellen wurde Geschlecht, Kinder und Migrationshintergrund als Kontrollvariablen aufgenommen. Offenheit für Erfahrungen zeigt sich als stabiler Prädiktor für das allgemeine Besucherverhalten sowie über verschiedene Museumstypen hinweg. Insbesondere das naturwissenschaftliche Sachinteresse zu Klasse 7 klärt eigenständige Varianz in der allgemeinen Besuchshäufigkeit sowie in der Besuchshäufigkeit von Technikmuseen über 25 Jahre später auf. Diese Ergebnisse unterstützen die Bedeutung frühzeitiger Interessenförderung und Bildung im naturwissenschaftlichen Bereich. Neben Offenheit und dem familiären kulturellen Kapital scheint für das Besucherverhalten in Kunstmuseen ebenfalls der eigene kulturelle Wert im Jugendalter bedeutsam zu sein. Diese Ergebnisse weisen weiter auf die Bedeutsamkeit der Differenzierung verschiedener Museumstypen hin, als dass verschiedene Prädiktoren je nach Museumstyp bedeutsam werden. Die Limitationen etwa der altershomogenen deutschen Stichprobe aus ursprünglich vier der 16 Bundesländer sowie des selektiven Dropouts werden diskutiert.

Museumsführungen für Erwachsene als Forschungsfeld – Ein Beispiel

Inga Specht¹, Franziska Loreit²

¹Leibniz-Institut zur Analyse des Biodiversitätswandels, Museum Koenig, Bonn, ²DIE Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen, Bonn

Museen als Orte informellen Lernens nutzen eine Vielzahl unterschiedlicher Vermittlungsangebote auf. Neben der musealen Aufarbeitung von Wissensständen anhand von, in Ausstellungen integrierte mediale Vermittlungen (z.B. Texte, Medienstationen, Dioramen) (Hof, 2003), zählen auch personale Formen der Vermittlung (Nuißl, 2004) wie z.B. ausstellungsbegleitende Bildungs- und Vermittlungsangebote in Form von Führungen, Kursen (Lewalter & Noschka-Roos, 2010; Nettke, 2016) dazu. Führungen sind dabei sowohl national als auch international das meistgenutzte Format ausstellungsbegleitender, personalbetreuer Vermittlung für Erwachsene (Institut für Museumsforschung, 2018; Sachatello-Sawyer et al., 2002). Wie Führungen jedoch konzipiert werden, ob und, wenn ja, welche vermittlungsbezogenen Überlegungen dabei eine Rolle spielen, welche Aneignungsprozesse (durch welche Vermittlungsprozesse) bei den teilnehmenden Besucher*innen intendiert werden, ist bislang nicht empirisch untersucht worden. Darüber hinaus fehlt es bis heute an einer theoretisch überzeugenden Modellierung der Vermittlungs- und Aneignungsprozesse in Führungen für Erwachsene (Monk, 2013). Ein am Deutschen Institut für

Erwachsenenbildung (DIE) durchgeführtes, von der Deutschen Forschungsgemeinschaft gefördertes Forschungsprojekt (SP 1739/1-1, 2018-2021) setzt an den genannten Desideraten an. Es untersucht, inwiefern Führungen als pädagogische Kommunikation im Sinne einer aneignungsbezogenen Wissensvermittlung (Kade & Seitter, 2003, 2007a, b) interpretiert werden können. Das systemtheoretische Konzept wurde entwickelt, um pädagogische Prozesse außerhalb organisierter Lehr-Lernprozesse in den traditionellen Institutionen des Bildungssystems erfassen zu können. Dieses Konzept scheint für Museumsführungen daher insofern anschlussfähig, als bei Führungen in Museen durchaus auch Bildungsprozesse intendiert und erwartet werden (von Freymann 2004). Können Führungen folglich als pädagogische Kommunikation beschrieben werden?

Dem Projekt liegt ein methodenkombinierendes Forschungsdesign mit zwei, aufeinander aufbauenden Phasen zugrunde. Die erste Phase umfasste Expert*inneninterviews mit 40 Vermittelnden aus vier verschiedenen Museumstypen in Deutschland. Die zweite Phase umfasste strukturierte, teilnehmende Beobachtungen von Führungen bei einem selektiven Sample der Vermittelnden, Audioaufzeichnungen der Führungen und eine Prä-Post-Befragung der Vermittelnden und der teilnehmenden Besucher*innen. Die zweite Phase wurde durch den Ausbruch der Corona- Pandemie beeinträchtigt.

Aus dem Datenmaterial der ersten Phase lässt sich vorläufig ableiten, dass Wissensvermittlung nur ein mögliches, pädagogisches Ziel von Vermittelnden ist. Weitere primär nicht kognitive, aber ebenfalls pädagogische Ziele sind beispielsweise die Sensibilisierung für ein Thema, die Förderung eines kreativ-selbstständigen Zugangs zu Museumsinhalten, Ermöglichung von Meinungsbildung, Empowerment im Umgang mit Kunst und/oder die Förderung haptischer Erfahrungen. Das zugrundeliegende Führungs- und Rollenverständnis weist zudem darauf hin, dass Vermittelnde ihre Rolle genuin pädagogisch verstehen. Dies korrespondiert mit den von den Vermittelnden antizipierten Erwartungen der Besucher*innen, in denen sich Spuren einer defizitären Adressatenkonstruktion insbesondere im Hinblick auf Aussagen über Erwartungen bzgl.

Wissensvermittlung, -erweiterung und/oder -vertiefung, finden lassen.

In dem Beitrag werden der methodische Ansatz und das Herangehen an ein sehr heterogenes Forschungsfeld präsentiert und Ergebnisse der Interviewstudie diskutiert.

Zum Zusammenhang von Besuchserleben, Wissenserwerb und Verhaltensintentionen in einer Ausstellung über Korallenriffe

Martin Merkt¹, Laura Amelung¹, Eva Roßmanith², Franziska Loreit¹

¹DIE Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen, Bonn, ²Senckenberg Gesellschaft für Naturforschung, Frankfurt am Main

Hintergrund

Als Orte des informellen Lernens (Thoma et al., 2022) sind Museen neben der Vermittlung von Sachinhalten auch Orte der Sensibilisierung für Themen der Demokratie- und Umweltbildung. Vor dem Hintergrund psychologischer Modelle ist jedoch fraglich, inwiefern ein Museumsbesuch neben kurzfristigen Effekten auf Sachwissen und der Aktivierung von Normen als Voraussetzungen einer Verhaltensänderung (vgl. Bamberg et al., 2015) längerfristige Effekte auf das Verhalten der Besuchenden hat. Neben Effekten auf Verhaltensintention wird im vorliegenden Beitrag untersucht, ob die Ergänzung von Inhalten aus unterschiedlichen Perspektiven (Kunst vs. Forschung) einen Einfluss auf das Besuchserleben, den Wissenserwerb und die Stärke von Verhaltensintentionen hat.

Methode

Der Beitrag beruht auf Daten von 200 Besucher/innen (MAlter = 36.33, SDAAlter = 15.76, 109w, 84m, 5d, 2kA), die am Eingang des Senckenberg Naturkundemuseums Frankfurt a.M. angesprochen und direkt in den Dauerausstellungsbereich Korallenriff geführt wurden. Die permanenten Ausstellungsinhalte wurden in aufeinander folgenden Zeiträumen ergänzt um Klang- und Objektinstallationen (Perspektive Kunst) sowie Informationen zu aktuellen Forschungsergebnissen und Forschungsmethoden (Perspektive Forschung). Unmittelbar nach der freien Exploration beantworteten die Teilnehmenden unter Anderem jeweils drei von Packer et al. (2018) adaptierte Fragen zu Ästhetik der Inhalte (Appreciation of Beauty, Cronbach's $\alpha = .79$), gewecktem Interesse (Engagement, $\alpha = .71$), inhaltlicher Reflexion (Reflection, $\alpha = .80$), kritischer Auseinandersetzung mit der persönlichen Verantwortung (Reflective Engagement, $\alpha = .78$) sowie zu Aktivierung von Verhaltensimpulsen (Call to Action, $\alpha = .89$) auf Skalen von 1 (gering) bis 5 (hoch). Außerdem beantworteten sie 16 richtig/falsch-Wissensfragen und gaben für eine selbst formulierte Verhaltensintention an, wie sehr sie diese umsetzen möchten. Fünf bis acht Wochen nach dem Ausstellungsbesuch wurde eine Online-Nachbefragung zur erfolgten Umsetzung der Verhaltensintention durchgeführt, an der 55 Besucher/innen teilnahmen.

Ergebnisse

Es zeigten sich positive Zusammenhänge der Stärke der Verhaltensintention mit den Erlebensdimensionen Engagement, Reflection, Reflective Engagement und Call to Action, alle $r > .20$, jedoch nicht mit Appreciation of Beauty, $r = .08$. Zusätzlich wurde ein positiver Zusammenhang zwischen dem selbst berichteten Ausmaß der Umsetzung der Verhaltensintention zum Zeitpunkt der Nacherhebung und der Stärke der Verhaltensintention direkt nach dem Ausstellungsbesuch beobachtet, $r = .27$. Allerdings wurden weder Stärke noch Ausmaß der Umsetzung der Verhaltensintention durch die unterschiedlichen Ausstellungsperspektiven beeinflusst, beide $p > .097$. Dagegen gaben Besucher/innen der „Perspektive Kunst“ höhere Werte hinsichtlich der Dimensionen Reflection und Call to Action an und schnitten besser im Wissenstest ab als Besucher/innen der „Perspektive Forschung“, alle $p < .026$, wobei der Unterschied im Wissenstest hauptsächlich auf perspektivspezifische Inhalte $p < .001$, und weniger auf perspektivenübergreifende Inhalte rückführbar ist, $p = .050$. Hinsichtlich der Erlebensdimensionen Reflective Engagement, Engagement und Appreciation of Beauty unterschieden sich die Perspektiven nicht, alle $p > .064$.

Diskussion

Dieser Beitrag zeigt, dass eine starke Verhaltensintention direkt nach einem Ausstellungsbesuch positiv mit einer kritischen Auseinandersetzung mit den Ausstellungsinhalten und dem Gefühl eigener Verantwortung zusammenhängt. Zudem besteht ein positiver Zusammenhang zwischen der Stärke der Verhaltensintention und dem selbst berichteten Ausmaß ihrer Umsetzung fünf bis acht Wochen nach dem Ausstellungsbesuch. Darüber hinaus führte eine Ausstellungsergänzung aus einer künstlerischen Perspektive zu einer intensiveren Auseinandersetzung mit den Inhalten sowie der eigenen Verantwortung und zudem zu einem besseren Abschneiden in einem Wissenstest als eine Ergänzung der Ausstellung aus einer Forschungsperspektive. Möglicherweise waren Besucher/innen durch eine erhöhte inhaltliche Dichte der Ausstellungsergänzung aus der „Perspektive Forschung“ überfordert, was sich darin widerspiegelt, dass im Wissenstest nicht nur die mutmaßlich schwierigeren Inhalte aus der Forschungsperspektive schlechter erinnert wurden, sondern auch leichte Wissensseinbußen hinsichtlich der in allen Perspektiven vergleichbaren Inhalte beobachtet wurden. Diese höhere inhaltliche Dichte könnte auch eine tiefere Auseinandersetzung mit den Inhalten verhindert haben. Vor diesem Hintergrund könnten sich zukünftige Studien systematisch

mit dem Einfluss der Informationsdichte in Ausstellungen auf das Besuchererleben, den Wissenserwerb sowie Verhaltensintentionen von Besucher/innen auseinandersetzen.

Verbesserte Einsicht in eigenes Wissen und Nicht-Wissen durch Ausstellungsbesuch

Helen Fischer¹, Lorenz Kampschulte², Fabienne Will²

¹IWM Leibniz-Institut für Wissensmedien, Tübingen, ²Deutsches Museum, München

Metakognition – die Einsicht in das, was wir wirklich wissen, im Vergleich zu dem, was wir nur zu wissen glauben – ist von entscheidender Bedeutung, insbesondere im Zeitalter von Missinformation und Fake News. Unsere Studie untersuchte, ob Ausstellungsbesuche im Museum metakognitive Einsicht in unser Wissen und Nicht-Wissen erhöhen kann. Besucher*innen einer Ausstellung (N = 165) zu Planetarer Gesundheit im Deutschen Museum in München beantworteten vor und nach dem Ausstellungsbesuch Wissensfragen zu Planetarer Gesundheit und gaben jeweils an, wie subjektiv sicher sie sich waren, richtig geantwortet zu haben.

Anfänglich beantworteten die Teilnehmenden 60 % der Wissensfragen richtig, mit einer durchschnittlichen Wahrheitsunterscheidung (d') von 0,67, was auf ein gewisses Grundwissen, aber erhebliche Schwierigkeiten bei der Unterscheidung zwischen wahren und falschen Aussagen hinweist. Zudem fanden wir eine geringe metakognitive Einsicht im Sinne einer geringen Übereinstimmung zwischen subjektiver Sicherheit und tatsächlicher Richtigkeit des Wissens. Nach der Ausstellung stiegen die korrekten Antworten der Teilnehmer auf 68 %. Insbesondere verbesserte sich auch die metakognitive Einsicht deutlich. Diese Verbesserung war auf ein erhöhtes Vertrauen in die richtigen Antworten zurückzuführen, während das Vertrauen in die falschen Antworten unverändert blieb.

Unsere Ergebnisse legen nahe, dass Museumsausstellungen nicht nur Wissen, sondern auch die Einsichtsfähigkeit in das eigene Wissen erhöhen kann. Diese Studie unterstreicht das Potenzial von Museen als wirksame Orte für die Wissenschaftskommunikation, insbesondere zur Verbesserung metakognitiver Fähigkeiten. Diese Ergebnisse zeigen praktische Anwendungen zur Stärkung des öffentlichen Verständnisses für wissenschaftliche Informationen auf.

S-76: Lernen durch Feedback von Lehrpersonen, Eltern und Schüler:innen

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 162

Chair der Sitzung: Sebastian Röhl

Chair der Sitzung: Fabian Hoya

Lernen durch Feedback von Lehrpersonen, Eltern und Schüler:innen

Chair(s): **Sebastian Röhl** (Pädagogische Hochschule Freiburg, Deutschland), **Fabian Hoya** (Pädagogische Hochschule Vorarlberg)

Feedback gilt als förderlicher Faktor für das Lernen von Schüler:innen (z.B. Brooks et al., 2021; Hattie & Timperley, 2007), von Studierenden (z.B. Carless et al., 2017), aber auch für die berufliche Weiterentwicklung von Lehrpersonen (z.B. Ridge & Lavigne, 2020; Röhl, 2021). Das Geben von Feedback erfolgt jedoch nicht nur vertikal von oben nach unten, also vom Lehrenden hin zum Lernenden, sondern es sind auch andere Konstellationen möglich: beispielsweise horizontales Peer-Feedback zwischen Lernenden (Ishchenko & Verkhovtsova, 2019) oder auch von den Lernenden zur lehrenden Person mit dem Ziel der Professionalisierung und der Verbesserung der Lernsituation (Strijbos & Müller, 2014). Dieses Symposium adressiert Effekte und Wirksamkeitsbedingungen von Feedback aus verschiedenen Quellen an Schüler:innen und (angehende) Lehrpersonen. Hierfür werden Befunde aus den unterschiedlichen Forschungsbereichen der Lehr-Lernforschung, Lehrer:innenbildung sowie der berufsbegleitenden Professionalisierung und Unterrichtsentwicklung zusammengeführt.

Im Unterricht obliegt es überwiegend den Lehrpersonen, den Schüler:innen lernförderliche Rückmeldungen zu geben. Die hierfür notwendigen professionellen Fähigkeiten sollen Lehrpersonen in den qualifizierenden Ausbildungsphasen erwerben (KMK, 2004/2019). So wendet sich der erste Beitrag des Symposiums der Vermittlung von feedbackbezogenen Fähigkeiten für angehende Lehrpersonen im Studium zu. Hierzu werden verschiedene Formen der Analyse von Unterrichtsvideos verwendet und hinsichtlich ihrer Wirksamkeit auf das feedbackbezogene konzeptuelle Wissen und die professionelle Wahrnehmung untersucht.

Hieran anschließend fokussiert der zweite Symposiumsbeitrag Effekte dieses regelmäßig im schulischen Unterricht vorkommenden Feedbacks von Lehrpersonen an die Schüler:innen und analysiert die direkten sowie indirekten Zusammenhänge mit der Mathematikleistung sowie dem Selbstkonzept und der Selbstwirksamkeit.

Der dritte Beitrag adressiert, dass Lehrpersonen im Kontext von Unterricht nicht nur in der Rolle des Feedbackgebenden, sondern auch als Empfangende von Feedback fungieren können und sollen (KMK, 2004/2019). So wird mithilfe einer repräsentativen Stichprobe untersucht, wie häufig Lehrpersonen in Deutschland unterrichtsbezogenes Feedback von ihren Schüler:innen und hospitierenden Kolleg:innen einholen und wie nützlich die beiden Feedbackformen bezüglich verschiedener Lern- und Entwicklungsbereiche bezüglich des eigenen unterrichtlichen Handelns wahrgenommen werden.

Ob und inwiefern auch Rückmeldungen aus anderen Quellen, z.B. aus dem familiären Umfeld, Wirkungen auf das Lernen der Kinder haben, ist bisher nur wenig erforscht. Der vierte Beitrag greift dieses Forschungsanliegen auf und untersucht die Rolle von Eltern als Feedbackgeber:innen für die eigenen Kinder im Grundschulalter beim Lesenlernen.

Abschließend werden die Befunde aus den unterschiedlichen Forschungsbereichen in einer gemeinsamen Diskussion hinsichtlich des aktuellen Stands der Feedbackforschung zusammengeführt.

Beiträge des Symposiums

Konzeptuelles Wissen und professionelle Wahrnehmung von Lehramtsstudierenden im Bereich Feedback – Effekte der Beobachtungsmethode bei der Analyse von Unterrichtsvideos

Fabiana Karstens¹, Miriam Hess²

¹Universität zu Köln, ²Otto-Friedrich-Universität Bamberg

Die Fähigkeit, gutes Feedback zu geben, ist ein wesentlicher Teil der professionellen Kompetenz von Lehrkräften und zentral für die Leistungsentwicklung von Schüler*innen (Hattie, 2013), stellt jedoch für viele Lehrkräfte eine Herausforderung dar (Hattie & Timperley, 2007; Lotz, 2015). Daher sollte bereits im Studium nicht nur konzeptuelles Wissen über Feedback vermittelt, sondern es sollten auch handlungsnahe Kompetenzen wie die professionelle Wahrnehmung gefördert werden. Diese ermöglicht es Lehrkräften, ihre Aufmerksamkeit auf lernrelevante Aspekte des Unterrichts zu lenken und für den weiteren Verlauf des Unterrichts wissensgeleitet weiter zu verarbeiten (Seidel & Stürmer, 2014, van Es & Sherin, 2008). Zur Förderung gilt besonders die Analyse von Unterrichtsvideos als vielversprechend (Blömeke et al., 2015; Steffensky & Kleinknecht, 2016), wobei die Gestaltung der Lernumgebung und der Einsatz zielorientierter Beobachtungsaufträge entscheidend für die Lernwirksamkeit sind (Seidel et al., 2013; Schneider et al., 2022). Unklar ist bislang jedoch, welchen Einfluss die Variation der Beobachtungsmethode bei der Videoanalyse hat.

Der vorliegende Beitrag untersucht daher die Effekte unterschiedlich stark strukturierter Beobachtungsmethoden auf (1) das konzeptuelle Wissen und (2) die professionelle Wahrnehmung im Bereich Feedback. Zudem werden Effekte auf (3) die professionelle Wahrnehmung unter Berücksichtigung des konzeptuellen Vorwissens betrachtet.

Es werden Daten aus einem Projekt genutzt, das die Wirksamkeit unterschiedlicher Seminarbedingungen für den Erwerb professioneller Kompetenzen im Bereich Feedback fokussierte (Hess, 2019). Die Studie wurde in einem quasi-experimentellen Prä-Post-Kontrollgruppendesign umgesetzt, wobei vor und nach dem Seminar u.a. jeweils das konzeptuelle Wissen und die professionelle Wahrnehmung im Bereich Feedback erfasst wurden. Alle Seminare beinhalteten eine allgemeine Einführung zum Thema Feedback, unterschieden sich jedoch in den Beobachtungsmethoden zur Videoanalyse: stark strukturierte Beobachtung durch niedrig-inferente Kodierung, strukturierte Beobachtung durch hoch-inferentes Rating und offene leitfragenorientierte Beobachtung. Zum Einsatz kamen dabei Videos für die Studierenden unbekannter Lehrpersonen. Eine Kontrollgruppe verzichtete auf den Einsatz von Videoanalysen. Die Stichprobe umfasste 226 Grundschullehramtsstudierende, die zu 88% weiblich und durchschnittlich im vierten Semester waren. Unterschiede zwischen den Interventionsgruppen gab es bezüglich ihrer vorherigen Praxiserfahrung oder ihrer thematischen Vorerfahrungen keine. Das konzeptuelle Wissen im Bereich Feedback wurde mit einem Instrument mit zwei offenen Aufgaben erfasst (Hess et al., 2017). Es sollte von den Studierenden definiert werden, was unter Feedback verstanden wird und Kriterien guten Feedbacks sollten benannt werden. Beide Aufgaben wurden anhand eines Manuals von zwei geschulten Kodierer*innen zur Bestimmung der Anzahl korrekt genannter Kriterien und zur Einschätzung der Gesamtqualität der Antworten ausgewertet. Die professionelle Wahrnehmung wurde mittels eines videobasierten Instruments erfasst, bei dem das Feedbackverhalten einer Lehrperson in einem Videoausschnitt kommentiert wurde (Hess & Deistler, 2022). Die Kommentare wurden anhand eines Ratingmanuals nach fünf Qualitätskriterien jeweils

vierstufig bewertet: Qualität der Beobachtung, Qualität der Bewertung und Begründung, Qualität der Verbesserungsvorschläge, Qualität der theoretischen Bezüge und Gesamteindruck. Zur Beantwortung der Fragestellungen des Beitrags wurden univariate Varianzanalysen mit Messwiederholungen sowie ergänzend Kovarianzanalysen durchgeführt.

Die Ergebnisse zeigen, dass die Studierenden sowohl beim konzeptuellen Wissen (Gesamtqualität: $F(1,221)=689.17$, $p < .001$, $\eta^2=.76$) als auch der professionellen Wahrnehmung profitieren (Gesamteindruck: $F(1,222)=252.73$, $p < .001$, $\eta^2=.53$). Dies trifft auf alle Interventionsbedingungen und die Kontrollgruppe zu. Gruppeneffekte treten nur bei der Qualität der theoretischen Bezüge ($F(3,222)=5.33$, $p = .001$, $\eta^2=.07$) und der Gesamtqualität ($F(3,222)=4.10$, $p < .01$, $\eta^2=.05$) der professionellen Wahrnehmung auf. Post-hoc-Analysen zeigen Vorteile für die Beobachtungsmethoden, die zu einer tieferen Auseinandersetzung mit den Unterrichtsvideos durch niedrig-inferente Kodierung oder leitfragenorientierte Beobachtung anregen. Die berichteten Gruppeneffekte bleiben auch unter Kontrolle des konzeptuellen Vorwissens bestehen (Theoretische Bezüge: $F(3,221)=9.03$, $p < .001$, $\eta^2=.11$; Gesamtqualität: $F(3,221)=6.06$, $p < .001$, $\eta^2=.08$). Die Befunde unterstreichen somit die Bedeutsamkeit zielorientierter Beobachtungsaufträge und systematischer Lernbegleitung bei der Arbeit mit Unterrichtsvideos. Im Vortrag werden zudem Folgerungen für die Gestaltung videobasierter Lehrkonzepte in der Lehrkräftebildung sowie Limitationen der Studie diskutiert.

Feedback und selbstbezogene Kognitionen als Prädiktoren für die Mathematikleistungen von Grundschulkindern

Nils Finke, Margarita Knickenberg, Frank Hellmich
Universität Paderborn

Für das Mathematiklernen von Grundschulkindern stellt das Feedback von Lehrkräften eine wichtige Voraussetzung dar. Unter Feedback werden dabei Informationen verstanden, die Kinder im Zusammenhang mit ihren Lernprozessen oder nach einer erbrachten Leistung von ihren Lehrkräften erhalten (Hattie, 2009). Lehrkräfte vergeben beispielsweise Feedback, um die Anstrengungsbereitschaft der Kinder zu erhöhen oder aber um ihnen alternative Lernwege aufzuzeigen. Kinder verarbeiten das von den Lehrkräften vergebene Feedback unterschiedlich. Folgt man dem ‚Student-Feedback Interaction Model‘ von Lipnevich und Smith (2022), so ist davon auszugehen, dass das von Kindern wahrgenommene Feedback im Zusammenhang mit ihren selbstbezogenen Kognitionen (z. B. Selbstkonzept, Selbstwirksamkeit) sowie ihren schulischen Kompetenzen steht. Die Selbstwirksamkeit von Kindern kann als die subjektive Überzeugung aufgefasst werden, neue komplexe Sachverhalte bewältigen zu können, um zuvor festgelegte Ziele zu erreichen (Bandura, 1997). Demgegenüber wird das Wissen über die eigenen Stärken und Schwächen unter dem Begriff Selbstkonzept subsumiert (Marsh et al., 2012). In den Studien von Burnett (1999) und Chen et al. (2011) konnten Hinweise für enge Korrelationen zwischen dem von Kindern wahrgenommenen Feedback und ihren Selbstkonzepten gefunden werden. Auch konnten Zusammenhänge zwischen Schulleistungen und Selbstkonzepten bei Schüler*innen nachgewiesen werden (Ehm et al., 2022). Ob und inwiefern Effekte des Feedbacks auf die Schulleistungen von Kindern durch ihre selbstbezogenen Kognitionen (Selbstkonzept, Selbstwirksamkeit) mediert werden, ist bislang ungeklärt.

Vor diesem Hintergrund gehen wir auf der Grundlage des ‚Student-Feedback Interaction Models‘ von Lipnevich und Smith (2022) der Frage nach, ob sich die Mathematikleistungen von Kindern im Grundschulalter einerseits durch das von ihnen wahrgenommene Feedback, sowie andererseits durch ihre selbstbezogenen Kognitionen erklären lassen. Darüber hinaus untersuchen wir, ob und inwiefern Effekte des wahrgenommenen Feedbacks auf die Mathematikleistungen durch die mathematikbezogenen Selbstkonzepte sowie durch die Selbstwirksamkeitsüberzeugungen der Kinder mediert werden.

Im Rahmen unserer Studie wurden $N=701$ Kinder der dritten und vierten Jahrgangsstufe anhand eines Paper-Pencil-Fragebogens befragt. Die Kinder gaben hierbei Auskunft zu ihrem wahrgenommenen anstrengungsbezogenen („Du hast dir viel Mühe gegeben!“, $\square=.87$), fähigkeitsbezogenen („Für Mathe bist du sehr talentiert!“, $\square=.92$) und allgemeinem positiven Feedback („Gut gemacht!“, $\square=.85$) im Mathematikunterricht. Darüber hinaus wurden die Kinder zu ihren Selbstwirksamkeitsüberzeugungen („Auch wenn manche Matheaufgaben schwierig sind, finde ich trotzdem immer eine Lösung.“; Jerusalem & Schwarzer, 1999, $\square=.78$) sowie ihren mathematischen Selbstkonzepten („Ich lerne schnell in Mathe.“; Wendt et al., 2017, $\square=.87$) befragt. Die Mathematikleistungen der Kinder wurden jeweils anhand des Deutschen Mathematiktests (DEMAT 3+, Roick et al., 2018; DEMAT 4, Göllitz et al., 2006) erhoben.

Die Ergebnisse aus einem Strukturgleichungsmodell (Mplus, Muthén & Muthén, 1998–2023) verdeutlichen, dass die Gesamtvarianz in den Mathematikleistungen der Kinder durch das wahrgenommene allgemeine positive Feedback ($\beta=.38$), das anstrengungsbezogene Feedback ($\beta=.64$) sowie die mathematischen Selbstkonzepte ($\beta=.34$) zu 30% erklärt werden können. Die Selbstwirksamkeitsüberzeugungen der Kinder und das fähigkeitsbezogene Feedback stellen hingegen keine signifikanten Prädiktoren für ihre Mathematikleistungen dar. Die mathematischen Selbstkonzepte der Kinder ($R^2=.53,4$) werden durch das wahrgenommene allgemeine positive, anstrengungs- und fähigkeitsbezogene Feedback erklärt. Effekte des wahrgenommenen Feedbacks auf die Mathematikleistungen der Kinder werden durch die mathematischen Selbstkonzepte mediert. Hiermit liegen Hinweise dafür vor, dass Zusammenhänge zwischen Feedback und Leistungsverhalten bei Kindern im Grundschulalter entsprechend des ‚Student-Feedback Interaction Models‘ durch selbstbezogene Kognitionen vermittelt werden.

Nutzen Lehrpersonen in Deutschland Unterrichtsfeedback von Schüler:innen und Kolleg:innen? Eine repräsentative Studie zur Nutzungshäufigkeit und eingeschätzten Nützlichkeit

Sebastian Röhl¹, Tim Fütterer², Colin Cramer²

¹Pädagogische Hochschule Freiburg, ²Eberhard Karls Universität Tübingen

Systematische Übersichtsarbeiten weisen darauf hin, dass die Verwendung von formalisiert eingeholten Unterrichtsrückmeldungen von Schüler:innen (so genanntes Schülerfeedback, SF) und hospitierenden Kolleg:innen (kollegiales Feedback bzw. Hospitation, KF) durch Lehrpersonen positive Effekte auf die Professionalisierung und Unterrichtsqualität haben können (Ridge & Lavigne, 2020; Röhl, 2021). Obwohl das Einholen von Rückmeldungen zum eigenen Unterricht sowohl in den Standards für Lehrerbildung (KMK, 2019) als auch in den Schulqualitätskonzeptionen einiger Bundesländer vorgesehen ist, scheint die Realisierung in Deutschland jedoch stark von der Eigeninitiative von Lehrpersonen und Einzelschulen abhängig zu sein. Die Befunde von Richter und Pant (2016) deuten an, dass in Deutschland 32% der Lehrpersonen eine nicht näher spezifizierte Form von Schülerfeedback zum Unterricht einholen sowie 9% selbst regelmäßig hospitieren und Feedback geben. Allerdings fehlen zur Häufigkeit der Einholung von kollegialem und formalisiertem Schülerfeedback in Deutschland verlässliche Daten. Weiterhin liegen bisher aus Deutschland nur einzelne, meist qualitative Studien zur wahrgenommenen Nützlichkeit vor (z.B. Gärtner & Vogt, 2013; Neuber & Göbel, 2021). Darüber hinaus legen Befunde nahe, dass unterrichtsverbessernde Wirkungen von Feedback von der Durchführung und Ausführlichkeit der Feedbackgespräche mit den Schüler:innen (Gaertner,

2014) bzw. Kolleg:innen (Ridge & Lavigne, 2020) abhängen. Wir gehen davon aus, dass dies auch für die wahrgenommene Nützlichkeit des erhaltenen Feedbacks gilt.

Die vorliegende Studie geht auf Basis einer repräsentativen Befragung den Fragen nach, (1) wie häufig Lehrpersonen in Deutschland Unterrichtsfeedback von Schüler:innen und Kolleg:innen einholen, (2) wie sie im Vergleich die Nützlichkeit der beiden Rückmeldeverfahren (a) insgesamt und (b) in Bezug auf einzelne Entwicklungsbereiche erleben und (3) welchen Effekt die Intensität der Diskussion des Feedbacks mit den Schüler:innen bzw. Kolleg:innen auf dessen Nützlichkeitseinschätzung hat.

Zur Bearbeitung der Fragestellungen wurde im Oktober 2023 eine repräsentative Online-Befragung von N = 615 Lehrkräften in Deutschland hinsichtlich Alter (M=50.9 Jahre, SD=9.5), Geschlecht (63% weiblich) und Schulform (Grundschule: 22.4%, Sekundarstufe: 66.7%, Sonderpädagogik: 9.6%, Sonstige: 1.3%) durch einen Felddienstleister durchgeführt. Die Nutzungshäufigkeit der beiden, im Rahmen der Befragung explizit formal definierten Rückmeldeverfahren wurde jeweils mittels eines Items erhoben („Wie häufig bzw. regelmäßig haben Sie bisher Unterrichtsfeedback von Schüler:innen/Kolleg:innen eingeholt?“, Antwortoptionen „noch nie“, „einmal“, „mehrmals“, „ca. einmal im Jahr“, „mehrmals im Jahr“). Jeweils drei Items erhoben die Nützlichkeitseinschätzung (Beispielitem: „Das Feedback hat mir geholfen, meinen Unterricht zu verbessern“; SF: $\alpha=.90$, KF: $\alpha=.92$).

Die Nützlichkeit in Bezug auf verschiedene Entwicklungsbereiche (angelehnt an Helmke, 2007) wurde jeweils anhand von einer Multiple-Choice-Frage erhoben, in der die Lehrkräfte die Bereiche auswählten, für die das erhaltene Feedback als entwicklungsfördernd wahrgenommen wurde. Jeweils mit einem Item wurde erfragt, ob und wie detailliert das erhaltene Feedback besprochen wurde (Antwortmöglichkeiten „gar nicht“, „kurz“, „ausführlich“, „ausführlich mit gemeinsamer Entwicklung von Verbesserungsmaßnahmen“).

Insgesamt gaben 42.9% der Lehrpersonen an, mindestens einmal jährlich SF einzuholen, wohingegen nur 17.6% regelmäßig KF erhalten. Hinsichtlich der Nützlichkeit der bisherigen Erfahrungen mit den beiden Verfahren zeigten sich leicht niedrigere Bewertungen für SF (M=3.56, SD=0.80) verglichen mit KF (M=3.73, SD=0.78, $t(320)=-3.71$; $p<.001$). Hinsichtlich spezifischer Unterrichts- und Professionalisierungsaspekte schätzten die Lehrkräfte SF als signifikant nützlicher für die Verbesserung des Klassenklimas und der Erklärung von Inhalten ein. KF scheint besser geeignet zu sein, um das Klassenmanagement und den Umgang mit Heterogenität zu verbessern. Weiterhin wiesen multiple latente Regressionsanalysen unter Kontrolle von Alter, Geschlecht und Schulart der Lehrpersonen auf die Relevanz der Diskussion der Rückmeldungen für die wahrgenommene Nützlichkeit hin ($\beta_{SF}=.20$, $p<.001$; $\beta_{PF}=.38$, $p<.001$). Die Befunde tragen aufgrund der Repräsentativität der Stichprobe zum Erkenntnisstand zur Nutzungshäufigkeit von Unterrichtsfeedback in Deutschland bei. Weiterhin ergeben sich Hinweise für einen passgenaueren Einsatz der beiden Unterrichtsfeedbackarten hinsichtlich bestimmter Entwicklungsbereiche und auf die Relevanz des anschließenden Feedbackgesprächs.

Wahrnehmung und Effekte elterlichen Feedbacks auf die Lesemotivation und -leistung von Grundschulkindern

Margarita Knickenberg¹, Fabian Hoya², Frank Hellmich¹

¹Universität Paderborn, ²Pädagogische Hochschule Vorarlberg

Im Rahmen von Family Literacy Konzepten wird die Bedeutung der Familie als informelle Sozialisationsinstanz für den Lesekompetenzerwerb von Kindern betont (Villiger, 2011). Die Beteiligung von Eltern an Lernprozessen ihrer Kinder – beispielsweise in Form von Feedback – wird auch als Parental Involvement bezeichnet (Grolnick & Slowiaczek, 1994). Unter Feedback werden Informationen verstanden, die Kinder während ihrer Lernprozesse oder im Anschluss an eine erbrachte Lernleistung erhalten (Hattie, 2009) – mit dem Ziel, sie zu bestärken, ihre Anstrengungsbereitschaft zu erhöhen oder ihnen alternative Lernwege aufzuzeigen. Allerdings wird Feedback angesichts individueller Lernbiografien unterschiedlich wahrgenommen und verarbeitet, wie das student-feedback interaction model (Lipnevich & Smith, 2022) verdeutlicht. Das Modell berücksichtigt die Relevanz des Feedbackverständnisses der Schüler*innen (kognitive Verarbeitung), ihre emotionale Reaktion darauf (affektive Verarbeitung) und ihre Entscheidung, wie sie darauf reagieren (behaviorale Verarbeitung). Die kognitive Verarbeitung bedeutet, dass Schüler*innen das perzipierte Feedback verstehen und für glaubwürdig halten müssen. Die affektive Verarbeitung betrifft die emotionalen Reaktionen, die auf die Wahrnehmung des Feedbacks folgen und sich beispielsweise in Form von Freude oder Ärger äußern können. Diese affektiven Zustände hängen eng mit der Motivation der Schüler*innen zusammen (Pekrun, 2006). Die behaviorale Verarbeitung umfasst Aktivitäten oder Strategien, die Schüler*innen anwenden, nachdem sie Feedback erhalten haben, um eigene Leistungen zu verbessern (Lipnevich & Smith, 2022). Um Feedback nutzen zu können, stellt sich daher zunächst die Frage, ob die Kinder sich des Feedbacks, das sie erhalten, bewusst sind bzw. ob das gegebene Feedback mit dem wahrgenommenen Feedback übereinstimmt.

In Anlehnung an das student-feedback interaction model (Lipnevich & Smith, 2022) wird daher untersucht, inwieweit das von Eltern vergebene Feedback zu Leseprozessen ihrer Kinder mit dem wahrgenommenen Feedback übereinstimmt (kognitive Verarbeitung). Darüber hinaus wird angenommen, dass elterliches Feedback einen Effekt sowohl auf die Lesemotivation der Kinder (affektive Verarbeitung) als auch auf ihre Leseleistung (als Resultat der behavioralen Verarbeitung) hat.

In einer querschnittlichen Untersuchung nahmen insgesamt N=407 Eltern-Kind-Dyaden teil. Im Detail umfasste die Stichprobe Grundschul*innen mit einem Durchschnittsalter von neun Jahren (M=9.33, SD=0.78) der dritten (n=240) und vierten Klasse (n=167) aus 45 Klassen an 15 Grundschulen in Nordrhein-Westfalen sowie ihre Eltern.

Sowohl die Eltern als auch die Kinder füllten einen Fragebogen aus, in dem sie gebeten wurden, unter anderem Angaben zur Häufigkeit des von den Eltern vergebenen bzw. von den Kindern perzipierten Feedbacks zu machen. Mit jeweils sechs Items, die von Burnett (1999) adaptiert wurden, wurde zum einen positives Feedback („Gut gemacht“, $\alpha=.90$) und zum anderen negatives Feedback („Das war nicht gut“, $\alpha=.81$) erfasst. Darüber hinaus gaben die Kinder Auskunft über ihre intrinsische und extrinsische Lesemotivation („Ich lese, weil es Spaß macht.“, $\alpha=.80$ vs. „Ich lese, weil ich in der Schule besser sein möchte als die anderen in meiner Klasse.“, $\alpha=.85$; Stutz et al., 2014). Zur Messung der Leseleistung wurde der Textverständnistest des ELFE 1–6 (Lenhard & Schneider, 2006) eingesetzt.

Die Befunde eines in Mplus (Muthén & Muthén, 1998–2023) berechneten Strukturgleichungsmodells zeigen, dass elterliches negatives Feedback das von den Kindern wahrgenommene positive ($\beta=-.13^*$) und negative Feedback ($\beta=.25^{***}$) vorhersagt, wohingegen das von den Eltern gegebene positive Feedback keinen signifikanten Prädiktor für das wahrgenommene Feedback darstellt. Gleichzeitig hat das wahrgenommene positive Feedback sowohl einen Effekt auf die intrinsische ($\beta=.33^{***}$) als auch auf die extrinsische Lesemotivation ($\beta=.40^{***}$), während das wahrgenommene negative Feedback ausschließlich mit der extrinsischen Motivation zusammenhängt ($b=.26^{***}$). Die Leseleistung der Kinder kann durch ihre intrinsische ($\beta=.42^{***}$) und extrinsische Lesemotivation ($\beta=-.44^{***}$) vorhergesagt werden (erklärte Gesamtvarianz: 39%).

Die Untersuchungsbefunde unterstützen die Annahmen des student-feedback interaction model (Lipnevich & Smith, 2022) zu kognitiven, affektiven und behavioralen Verarbeitungsprozessen von Feedback bei Kindern im Grundschulalter. Die Bedeutung elterlichen Feedbacks für den Lesekompetenzerwerb von Kindern wird im Rahmen des Vortrags diskutiert.

S-77: Studienabbruch im Lehramt: Ursachen, Prädiktoren und Lösungsansätze

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SO 108

Chair der Sitzung: Stefan Sorge

Chair der Sitzung: Thilo Kleickmann

Chair der Sitzung: Martin Neugebauer

Studienabbruch im Lehramt: Ursachen, Prädiktoren und Lösungsansätze

Chair(s): **Stefan Sorge** (Christian-Albrechts-Universität zu Kiel, Deutschland), **Thilo Kleickmann** (Christian-Albrechts-Universität zu Kiel, Deutschland)

Diskutant*in(nen): **Martin Neugebauer** (Pädagogische Hochschule Karlsruhe, Deutschland)

Studienabbrüche stellen sowohl aus individueller als auch gesellschaftlicher Perspektive eine Herausforderung dar (Wissenschaftsrat, 2015). Besonders im Lehramtsstudium sind Studienabbrüche angesichts des akuten Lehrkräftemangels von hoher Bedeutung (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz, 2023). Der Lehrkräftemangel zeigt sich besonders deutlich in den Fächern und Schulstufen, in denen auch die Abbruchquoten besonders hoch sind. So weisen das Lehramtsstudium Physik beispielsweise Schwundquoten von bis zu 75% auf (Deutsche Physikalische Gesellschaft, 2023) und auch für das Lehramtsstudium der Sekundarstufe I werden hohe Schwundquoten berichtet (Güldener et al., 2020). Dabei existieren zur Beschreibung und Erklärung des Studienabbruchs inzwischen eine Reihe von theoretischen Überlegungen und empirischen Befunden (für einen Überblick siehe Neugebauer et al., 2019). Eine zentrale Theorie stellt dabei das „Student Integration Model“ von Tinto (1975) dar, bei dem davon ausgegangen wird, dass Entscheidungen zum Studienabbruch auf Basis des (Nicht-)Gelingens einer akademischen und sozialen Integration ins Hochschulsystem getroffen werden. Diese Integration hängt dabei sowohl von individuellen Merkmalen als auch von den sozialen und Studienbedingungen ab.

Im Modell des Studienabbruchs von Tinto (1975) finden studiengangs- und berufsspezifische Besonderheiten jedoch nur bedingt Berücksichtigung. Dabei wäre insbesondere für das Lehramtsstudium anzunehmen, dass neben einer sozialen und akademischen Integration ebenso eine Integration in das Berufsfeld Schule bedeutsam ist. So werden universitäre und schulische Lernumgebung häufig als unterschiedliche Lernwelten wahrgenommen und eine unzureichende Verzahnung von Theorie und Praxis wird bemängelt (Feiman-Nemser & Buchmann, 1985; Gröschner & Hascher, 2019). Um dieser Herausforderung zu begegnen, sind im Lehramtsstudium verschiedene Praxisphasen implementiert, die eine frühe Berufsfelderkundung ermöglichen und die Entwicklung professioneller Kompetenz unterstützen sollen. Somit besitzen Praxisphasen im Lehramtsstudium das Potential, einen Einfluss auf Integrationsprozesse zu nehmen. Bisher existieren jedoch nur wenige Studien, die Studienabbruch im Lehramtsstudium und die spezifische Bedeutung der Praxisphasen in den Blick genommen haben (Porsch, 2019).

In diesem interdisziplinären Symposium werden drei Studien zu unterschiedlichen Kohorten von Lehramtsstudierenden vorgestellt. Im ersten Beitrag wird ein einzelner Hochschulstandort und die Bedeutung der dort vorhandenen Praxisphasen sowie die Rolle der Theorie-Praxis-Verzahnung im Praxissemester für die Abbruchintention in den Blick genommen. Der zweite Beitrag identifiziert Prädiktoren für die Abbruchintention spezifisch in mathematisch-naturwissenschaftlichen Fächern unter Berücksichtigung des Professionswissens in einer deutschlandweiten Erhebung. Der dritte Beitrag ergänzt Vergleiche zwischen Studienabbrecher:innen und Absolvent:innen durch eine qualitative Befragung von Studienabbrecher:innen zu Gründen ihres Abbruchs. Abschließend werden die Beiträge aus einer soziologischen Perspektive diskutiert. Aus dieser umfassenden Betrachtung des Studienabbruchs sollen sowohl theoretische Überlegungen zur Erweiterung bestehender Modelle als auch praktische Implikationen für die Gestaltung des Lehramtsstudiums abgeleitet werden.

Beiträge des Symposiums

Zur Rolle von Praxisphasen und deren Vorbereitung für die Abbruchintention angehender Lehrkräfte

Stefan Sorge¹, Andrea Bernholt², Minke Tjardts¹

¹Christian-Albrechts-Universität zu Kiel, Deutschland, ²IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel, Deutschland

Eine der zentralen Herausforderungen im aktuellen Bildungskontext ist der akute Lehrkräftemangel (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz [SWK], 2023). Dieser wird unter anderem dadurch verschärft, dass ein Teil von etwa 10-20% der Studierenden, die sich zunächst für den Lehrberuf entschieden haben, ihr Studium vorzeitig abbrechen (Heublein et al., 2022). Dabei werden für den Studienabbruch sowohl kognitive als auch motivationale Eingangsmerkmale, sowie Merkmale der Studien- und externe Lebensbedingungen diskutiert (Neugebauer et al., 2019). Insbesondere die Wechselwirkung zwischen Person und Studienbedingung bestimmt, inwiefern eine soziale und akademische Integration gelingt, die als maßgeblich für Studienabbruch gesehen wird (Tinto, 1993). Für Studienabbrüche im Lehramtsstudium kann jedoch davon ausgegangen werden, dass ebenso eine Integration in die schulische Berufswelt notwendig wird, da Abwägungsprozesse über Kosten-Nutzen-Relationen die Entscheidung für einen Studienabbruch beeinflussen (Beekhoven et al., 2002). Für die schulische Integration im Lehramtsstudium sind Praxisphasen und deren universitäre Begleitung entscheidend. Studien legen dabei nahe, dass in Praxisphasen nur bedingt eine Integration universitärer und schulischer Lernwelten gelingt (Bleck & Lipowski, 2020) und spätere Praxisphasen Potenzial für eine Berufswahlüberprüfung besitzen (Porsch, 2019). Die vorliegende Studie untersucht daher zwei Fragestellungen. Zunächst wird in Bezug auf drei Pflichtpraktika im Lehramtsstudium untersucht, ob ein Zusammenhang zwischen der Zufriedenheit mit den Praxiserfahrungen im Studium und der Intention, das Studium abzubrechen vorliegt. Darüber hinaus wird geprüft, inwiefern die Zufriedenheit mit dem Praxissemester als ein Mediator zwischen den wahrgenommenen Theorie-Praxis-Bezügen und der Studienabbruchintention wirkt.

Als Datengrundlage dienten zwei Erhebungswellen einer Panelstudie an einer deutschen Universität. Dabei wurden Daten aus Erhebungen im Wintersemester 2021/22 und 2022/23 verwendet. Insgesamt lagen somit Daten von N = 879 Gymnasiallehramtsstudierenden vor, die mindestens ein Pflichtpraktikum durchlaufen haben. Dabei wurde zwischen dem pädagogischen Praktikum nach dem 2. Bachelorsemester, dem fachdidaktische Praktikum nach dem 4. Bachelorsemester und dem Praxissemester im 3. Mastersemester unterschieden. 74.5% der Stichprobe war weiblich, 1% der Teilnehmenden gehörten einer anderen Geschlechtsidentität an. Bei Teilnahme waren die Studierenden durchschnittlich M = 24.9 (SD = 3.6) Jahre alt.

Die Zufriedenheit mit den Praxiserfahrungen wurde über drei Items erhoben (Gruber et al., 2010; Westermann et al., 1996). Die Bewertung wurde jeweils getrennt für die drei Pflichtpraktika erfragt. Die Studienabbruchintention wurde mittels drei Items erfasst (Nagy, 2006). Weiterhin wurden die Theorie-Praxis-Bezüge im Praxissemester (König et al., 2014), sowie weitere Persönlichkeitsmerkmale (z. B. Rammstedt & John, 2005) und Merkmale der Studienbedingung (z. B. Renger et al., 2017) erhoben.

Die Ergebnisse zeigten, dass die Zufriedenheit mit den Praxiserfahrungen mit allen drei Pflichtpraktika signifikant negativ mit der Intention, das Lehramtsstudium abzubrechen, korreliert (Pädagogisches Praktikum: $r = -.16$, $p < .05$, $n = 205$; Fachdidaktisches Praktikum: $r = -.29$, $p < .001$, $n = 349$; Praxissemester: $r = -.21$, $p < .001$, $n = 279$). Multiple Regressionen zeigten zudem, dass die Zufriedenheit mit den Praktika in diesen Analysen auch bei Einschluss von Persönlichkeitsfaktoren und weiteren Merkmalen signifikante Prädiktoren der Abbruchintention bleiben. In Bezug auf die zweite Forschungsfrage konnte ein negativer Zusammenhang zwischen den Theorie-Praxis-Bezügen im Praxissemester und der Abbruchintention identifiziert werden ($\beta = -.18$, $p < .05$), der durch die Zufriedenheit mit dem Praktikum partiell mediert wird (Theorie-Praxis – Zufriedenheit: $\beta = .46$, $p < .001$; Zufriedenheit – Abbruchintention: $\beta = -.16$, $p < .01$).

Die Ergebnisse legen nahe, dass Studierende, die mit ihren Praxiserfahrungen weniger zufrieden sind, eine höhere Abbruchintention bzw. bei einer höheren Zufriedenheit eine geringere Abbruchintention aufweisen. Dies unterstreicht die Bedeutung von Praxisphasen bei Überlegungen zum Studienabbruch und deutet darauf hin, dass die schulische Integration eine Rolle in den Entscheidungsprozessen der Studierenden spielt (Beekhoven et al., 2002). Die Rolle der Theorie-Praxis-Bezügen für die Bewertung des Praxissemesters und die Studienabbruchintention bietet somit einen Ansatzpunkt, um das Lehramtsstudium und seine Praxisphasen weiterzuentwickeln.

Vorhersage der Studienabbruchsintention in mathematisch-naturwissenschaftlichen Lehramtsstudiengängen – Zusammenhänge mit individuellen Merkmalen sowie der Studiums- und Praktikumszufriedenheit

Judith Pollmeier, Thilo Kleickmann

Christian-Albrechts-Universität zu Kiel, Deutschland

Theoretischer Hintergrund

Vor dem Hintergrund des Lehrkräftemangels insbesondere in den MINT-Fächern sind empirische Erkenntnisse über die Ursachen von Studienabbrüchen im Lehramtsstudium von zentraler Bedeutung. In der Literatur zu Studienabbrüchen (nicht beschränkt auf das Lehramtsstudium) werden unter anderem individuelle, personenbezogene Merkmale als relevante Prädiktoren für Studienabbrüche bzw. deren Intention berichtet. Dazu gehören kognitive Voraussetzungen, Persönlichkeitsmerkmale, motivationale Aspekte und Lebensumstände der Studierenden (Bernholt, Zimmermann & Möller, 2023; Fleischer et al., 2019; Tinto, 1975).

Neben personenbezogenen Merkmalen erwiesen sich studienbezogene Merkmale als bedeutsam (Bernholt et al., 2023). So zeigte beispielsweise die Zufriedenheit mit dem Studium über die Berücksichtigung von kognitiven Merkmalen hinaus eine prädiktive Relevanz für Studienabbrüche (Brandstätter, Grillich & Farthofer, 2006). Die Studienzufriedenheit ihrerseits hängt mit der wahrgenommenen Qualität des Studiums bzw. der Lehre zusammen (Blüthmann, 2012). Neben den universitären Lehrveranstaltungen sind schulische Praktika eine zentrale Komponente von Lehramtsstudiengängen. Erste Befunde deuten auf die Relevanz dieser Praktika für die Berufswahlsicherheit sowie den späteren Eintritt und Verbleib in den Beruf als Lehrkraft hin (Horvath et al., 2018; Porsch, 2019). Unklar ist, ob die Praktikumszufriedenheit zur Erklärung von Studienabbrüchen bzw. der Intention zum Studienabbruch beitragen kann und dies über die allgemeine Studienzufriedenheit hinaus.

Fragestellungen

Die vorliegende Studie soll zur Erklärung der Studienabbruchsintention von Lehramtsstudierenden mit mathematisch-naturwissenschaftlichen Fächern beitragen. Dabei berücksichtigen wir neben personenbezogenen Merkmalen (kognitive Grundfertigkeiten, fachspezifisches Vorwissen, Persönlichkeitsmerkmale, fachliches Interesse, Abiturnote) die Studienzufriedenheit sowie die Praktikumszufriedenheit. Zwei Fragestellungen stehen im Mittelpunkt:

1. Welche individuellen Merkmale erklären Varianz in der selbstberichteten Studienabbruchsintention?
2. Welche inkrementellen Beiträge leisten die allgemeine Studienzufriedenheit und die Praktikumszufriedenheit zur Erklärung der Studienabbruchsintention?

Methode

Zur Untersuchung der Fragestellungen wurden die Daten von 330 Lehramtsstudierenden mit mathematisch-naturwissenschaftlichen Unterrichtsfächern von 25 deutschen Hochschulen analysiert. Die Studierenden machten im Sommersemester Angaben zu den von Ihnen absolvierten Praktika und Lehrveranstaltungen der letzten beiden Semester. Zu Beginn des folgenden Wintersemesters bearbeiteten sie Fragebögen zu ihrer Studienabbruchsintention, der Studienzufriedenheit, Persönlichkeits- und motivationalen Merkmalen, dem pädagogisch-psychologischen Wissen, kognitiven Fähigkeiten und demographischen Merkmalen. Von den untersuchten Studierenden gaben 160 Studierende an, im fraglichen Zeitraum ein Praktikum absolviert zu haben. In multiplen Regressionsanalysen wurden zunächst die individuellen Merkmale und dann schrittweise die allgemeine Studienzufriedenheit und die Praktikumszufriedenheit als Prädiktoren der Studienabbruchsintention eingeführt.

Ergebnisse und Folgerungen

Bei der Untersuchung der individuellen Merkmale erwiesen sich das pädagogisch-psychologische Wissen ($\beta = -.24$; $t(306) = -3.99$; $p < .001$), das fachliche Interesse ($\beta = -.19$; $t(306) = -3.57$; $p < .001$) sowie der Neurotizismus ($\beta = .26$; $t(306) = 4.11$; $p < .001$) als signifikante Prädiktoren der Studienabbruchsintention.

Bei zusätzlicher Einführung der Studienzufriedenheit ($\beta = -.45$; $t(305) = -8.51$; $p < .001$) in die Modelle erwies sich diese über die bereits berücksichtigten Variablen hinaus als relevant für die Vorhersage der Studienabbruchsintention. Auch der Neurotizismus und das pädagogisch-psychologische Wissen waren weiter mit der Abbruchsintention assoziiert. Wurde neben den individuellen Merkmalen die Praktikumszufriedenheit berücksichtigt, trug diese ($\beta = -.18$; $t(145) = -2.34$; $p < .05$), der Neurotizismus und das pädagogisch-psychologische Wissen zur Erklärung der Abbruchsintention bei. Bei gleichzeitiger Berücksichtigung von Studiums- ($\beta = -.40$; $t(144) = -5.03$; $p < .001$) und Praktikumszufriedenheit ($\beta = -.14$; $t(144) = -2.04$; $p < .001$) wiesen beide einen signifikanten Effekt neben dem pädagogisch-psychologischen Wissen auf.

Die Ergebnisse stützen frühere Befunde, nach denen sowohl personen- als auch studienbezogene Faktoren mit der Studienabbruchsintention zusammenhängen. Über individuelle Merkmale hinaus erklärten sowohl die Zufriedenheit mit dem Studium insgesamt, wie auch mit den absolvierten Praktika jeweils inkrementelle Teile der Varianz der Abbruchsintention. Die Ergebnisse legen daher die Praktika als Ansatzpunkt für Maßnahmen nahe, um Studienabbrüche in den MINT-Lehramtsfächern zu reduzieren. Im Vortrag wird außerdem darauf eingegangen werden, inwiefern Qualitätsmerkmale der Lehre zur Erklärung der Studiums- und Praktikumszufriedenheit beitragen.

Frühe Prädiktoren und Gründe für den Abbruch eines Lehramtstudiums - Ergebnisse aus dem Studierendenpanel PaLea

Andrea Bernholt¹, Friederike Zimmermann², Jens Möller²

¹IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik Kiel, Deutschland, ²Christian-Albrechts-Universität zu Kiel, Deutschland

Theoretischer Hintergrund

Obwohl die Forschung zum Studienabbruch insgesamt zugenommen hat, lassen sich mit Blick auf die Gruppe der Lehramtsstudierenden nur wenige empirische Studien finden (Herfter, Grüneberg, & Knopf, 2015). Längsschnittliche Daten, mit denen frühe Prädiktoren zum Studienbeginn mit einem späteren Abbruch bzw. Wechsel aus dem Lehramtstudium abgebildet werden können, sind kaum zu finden (Neugebauer, Heublein & Daniel, 2019). Zudem legen bisherige Ergebnisse nahe, dass neben Faktoren wie sie aus der fachübergreifenden Abbruchforschung bekannt sind, auch spezifische, lehramtsspezifische Merkmale wie beispielsweise Studienwahlmotive (z. B. das pädagogische Interesse) und berufsbezogene Selbstkonzepte (unterrichts- und fachbezogenes Selbstkonzept) bedeutsam sind (Suhlmann, Sassenberg, Nagengast & Trautwein, 2018).

Auch der Forschungsstand zu selbst berichteten Gründen zum Abbruch weist auf spezifische Bedingungen und Motivlagen des Lehramtsstudiums hin wie beispielsweise auf eine mangelnde Theorie-Praxis-Verzahnung sowie auf den Mangel an schulpraktischen Anteilen (Blüthmann & Thiel, 2011; Herfter, Grüneberg & Knopf, 2015). Zudem werden interne Ursachen wie eine falsche Berufswahl, aufkommender Widerstand gegen die Schulpraxis oder subjektiv falsche Studienvorstellungen (Blüthmann, Lepa, & Thiel, 2008; Henecka, 1996) genannt.

Fragestellung und Design

Anknüpfend an die skizzierten Forschungslücken mit Blick auf die Spezifitäten eines Lehramtstudiums wurden in der vorliegenden Studie daher die zwei Forschungsfragen adressiert:

o Welche studienbezogenen bzw. individuellen Merkmale zu Studienbeginn sind prädiktiv für einen Abbruch des bzw. Wechsel aus dem Lehramtsstudium?

o Welche Gründe und Ursachen geben Lehramtsstudierende für ihren Studienabbruch an und unterscheiden sich diese zwischen Studierenden, die ihr Studium abbrechen, und Studierenden, die den Studiengang wechseln?

Als Datenbasis für den vorliegenden Beitrag stehen Selbstauskünfte von Lehramtsstudierenden an 13 deutschen Hochschulen des Panels zum Lehramtsstudium (PaLea) zur Verfügung (Kauper, Bernholt, Möller, & Köller, 2023). Neben den Studieneingangserhebungen im Wintersemester 2009 bzw. 2010 wurden Daten aus den Abbruch- und Wechselfragebögen über die Gesamtlaufzeit des Panels (2009 bis 2017) verwendet. Insgesamt liegen Daten von N = 787 Studierenden vor, von denen n = 303 ihr Lehramtsstudium abgebrochen bzw. n = 484 ihr Lehramtsstudium erfolgreich beendet haben (Referenzgruppe).

Die Studierenden Daten in PaLea ermöglichen dabei ein umfangreiches Bild über Studienbedingungen (Lehramtszugang, Semester, Fachwahl, etc.), individuelle Merkmale (Persönlichkeit, soziodemographische Variablen, etc.), allgemeine (Kognitive Fähigkeiten, Abiturnote, etc.) sowie berufsbezogene Kompetenzfacetten wie motivationale Orientierungen (Berufswahlmotivation, berufsbezogene Selbstkonzepte, Interesse, etc.).

Zur Beantwortung der ersten Forschungsfrage wurden schrittweise multinominale logistische Regressionen berechnet, die Analysen der zweiten Forschungsfrage basieren auf deskriptiven Statistiken.

Ergebnisse und Implikationen

Hinsichtlich früher Prädiktoren zeigt sich, dass Studierende, die ein gymnasiales Lehramt studierten, ein höheres Risiko zum Abbruch aufweisen. Zudem lässt sich feststellen, dass sowohl die Fähigkeiten als auch die Motivation entscheidend sind für einen Abbruch des Lehramtsstudiums. So zeigten sich ungünstige Prognosen für Studierende mit schlechteren Leistungen im Abitur, anfangs geringem Fachinteresse, weniger Vorfreude auf das Lehramtsstudium und hoher Unsicherheit, ob das Lehramtsstudium die richtige Wahl ist.

Die berichteten Gründe für den Studienabbruch sind vor allem falsche Erwartungen an das Studium sowie Desinteresse am Lehrerberuf. Ein fehlender Theorie-Praxis-Bezug wurde vor allem von Studierenden als Ursache genannt, die aus dem Lehramtsstudium in ein anderes Studium gewechselt sind.

Zusammenfassend lässt sich festhalten, dass eine studiengangspezifische Erfassung und Untersuchung von Eingangsmerkmalen auch für das Lehramtsstudium sinnvoll erscheinen, da sich diese – neben allgemeinen Merkmalen – empirisch als prädiktiv zeigen. Insbesondere mit Blick auf die Intention, Studienabbrüche im Lehramtsstudium zu verringern, scheint eine spezifische Untersuchung möglicher Prädiktoren, aber auch selbst genannter Ursachen der Studierenden, sinnvoll, um adäquate, adressatengerechte und gezielte Präventionsmaßnahmen entwickeln zu können. Implikationen der Ergebnisse für die Lehrkräftebildung werden in dem Vortrag bzw. im Symposium diskutiert.

S-78: Förderung evidenzinformierter Entscheidungen von Lehrkräften

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 145

Chair der Sitzung: Samuel Merk

Chair der Sitzung: Kirstin Schmidt

Chair der Sitzung: Ulrich Trautwein

Förderung evidenzinformierter Entscheidungen von Lehrkräften

Chair(s): **Samuel Merk** (Pädagogische Hochschule Karlsruhe, Deutschland), **Kirstin Schmidt** (Pädagogische Hochschule Karlsruhe, Deutschland)

Diskutant*in(nen): **Ulrich Trautwein** (Universität Tübingen)

Professionelles Handeln von Lehrkräften ist unter anderem durch deren Rückgriff auf geeignete Wissensbestände bei Entscheidungen und Problemlösungen gekennzeichnet (Dewe et al., 1992). Welche Wissensbestände für die Lösung welcher Probleme als geeignet gelten, hängt jedoch zum einen stark von der Natur des zu lösenden Problems ab, zum anderen von der Transferierbarkeit und Transformierbarkeit des Wissens: Für hochgradig situative Entscheidungen wie z.B. welche*r Schüler*in als nächstes im Unterrichtsgespräch aufgerufen werden soll, kann episodische Evidenz angemessen sein. Konzeptuelle Entscheidungen wie z.B. Zeitpunkt und Form von Feedback hingegen können von der stärkeren Widerspruchsfreiheit, Objektivität und Generalisierbarkeit (bildungs-)wissenschaftlichen Wissens profitieren (Renkl, 2022). Während diese Charakteristika professionellen Wissens abgesehen von Nuancierungen (Stark, 2017) oder vereinzelter Fundamentalkritik (Bellmann & Müller, 2011) in vielen Bildungswissenschaften als weit geteilte Vorstellungen gelten, ist das Bild bei Lehramtsstudierenden und Lehrkräften weniger klar. Während einigen Befunden nach Lehramtsstudierende und Lehrkräfte bildungswissenschaftlichem Wissen durchaus substanziellen Gültigkeitsanspruch und praktische Relevanz beimessen (Schmidt et al., 2022) sind andere deutlich pessimistischer - insbesondere bzgl. der Quellenwahl in Problemlösesituationen und der Persistenz evidenzwidriger Überzeugungen (Menz et al., 2020).

Vor diesem Hintergrund sind im letzten Jahrzehnt einige Rahmenmodelle entstanden, die zum einen den idealtypischen Prozess evidenzinformierten Handelns konzeptuell beschreiben (Groß Ophoff et al., 2023), um daraus zum anderen Barrieren und Katalysatoren abzuleiten. Diese Barrieren und Katalysatoren wiederum lassen sich heuristisch in Charakteristika der Nutzenden (z.B. Lehrkräfte, Studierende) und Anbietenden (z.B. Forschende, Knowledge Brokers) gliedern: Auf Seite der Nutzenden stellt etwa forschungsmethodisches Vorwissen einen Katalysator und eine niedrige Nützlichkeitsüberzeugung eine Barrieren dar. Auf Seite der Anbietenden kann eine verständliche Beschreibung von Untersuchungsdesigns und -ergebnissen evidenzinformierte Entscheidungen katalysieren, aber missverständliche Kommunikation wie die Induktion eines »Practical Significance Bias« (Michal & Shah, 2024) eine Barriere darstellen.

Das Symposium »Förderung evidenzinformierter Entscheidungen von Lehrkräften« umfasst Experimente mit Lehrkräften und Lehramtsstudierenden die jeweils Barrieren oder Katalysatoren auf Angebots- oder Nutzendenseite adressieren: Kühlwein et al. untersuchen etwa in einem Experiment mit Lehrkräften, inwiefern geeignete textliche Beschreibungen oder grafische Darstellungen dazu geeignet sind, eine realistische Vorstellung von Effektstärken bestimmter Interventionen an Lehrkräfte zu kommunizieren (Katalysator auf Angebotsseite), um so den typischen »Practical Significance Bias« klassischer Wissenschaftskommunikation (Barriere auf Angebotsseite) zu verringern. Kollar et al. adressieren hingegen einen Katalysator auf Nutzendenseite, indem sie die Effekte verschiedener Feedbackarten auf den Aufbau von Fertigkeiten zur evidenzinformierten Analyse von Problemsituationen im Unterricht untersuchen. Die Bevorzugung episodischer Evidenz als Barriere der Nutzenden wird von Rochnia et al. in einem Experiment mit Lehramtsstudierenden untersucht. Damit möchte das Symposium einen Beitrag rund um die Diskussion von Chancen und Problemen evidenzinformierter Praxis von Lehrkräften leisten und insbesondere Gelingensbedingungen und Schwierigkeiten bei deren Förderung adressieren.

Beiträge des Symposiums

Effektstärken verständlich an Lehrpersonen kommunizieren

Florian Kühlwein¹, Kirstin Schmidt¹, Jürgen Schneider², Samuel Merk¹

¹Pädagogische Hochschule Karlsruhe, Deutschland, ²Leibniz-Institut für Bildungsforschung und Bildungsinformation

Lehrpersonen können bei ihren beruflichen Entscheidungen auf diverse Informationsquellen zurückgreifen, wie etwa auf Erfahrungswissen oder bildungswissenschaftliche Forschungsergebnisse (Buehl & Fives, 2009).

Die Berücksichtigung von Informationen unter anderem aus wissenschaftlichen Studien bei der Vor- oder Nachbereitung von Unterricht kann als evidenzinformierte Praxis (EIP) bezeichnet werden (Brown et al., 2017; Ferguson, 2021). Ziel von EIP ist es, rationale Entscheidungen im Bildungssystem zu fördern, um beispielsweise die Unterrichts- und Schulqualität zu verbessern (Bauer & Kollar, 2023; EU, 2007).

Damit EIP gelingen kann, müssen Lehrpersonen auch die Kosten und den Nutzen pädagogischer Interventionen einschätzen können. Dies ist ein komplexer Prozess, da sowohl die praktische Relevanz einer Intervention, abgeleitet aus Effektstärken, als auch ihr Umsetzungsaufwand zu berücksichtigen sind. Die akkurate Interpretation von Effektstärken ist jedoch bereits eine notwendige Bedingung für die Entfaltung des Potenzials von EIP: Werden Effektstärken über- oder unterschätzt, besteht die Gefahr, dass aus wissenschaftlichen Ergebnissen dysfunktionales Wissen (Groß-Ophoff et al., 2023) entsteht, welches darauf aufbauende Entscheidungen und Handlungen negativ beeinflusst.

In der Literatur finden sich Hinweise, dass diverse verbale und visuelle Präsentationsformen von Effektstärken z.B. unterschiedlich informativ wahrgenommen werden (Hanel & Mehler, 2019; Schmidt et al., im Druck). Zudem gilt die akkurate Rezeption für Laien zwar als herausfordernd (z.B. Lortie-Forgues et al. 2021; Schmidt et al., 2023; Kim et al., 2022; Zhang et al., 2023), aber auch durchaus erreichbar (Merk et al., 2023). Unklar bleibt, wie akkurat Lehrpersonen Effektstärken rezipieren können und welche Vor- und Nachteile verbale und visuelle Präsentationsmodi aufweisen. Daher leiten wir folgende Forschungsfragen ab:

FF 1: Wie akkurat (abhängige Variable 1/AV1), informativ (AV2) und mit welcher wahrgenommenen Schwierigkeit (AV3) werden verbal und visuell kommunizierte Effektstärken von Lehrkräften rezipiert?

FF 2: Gibt es bezüglich dieser AV Unterschiede zwischen den Präsentationsmodi (verbal vs. visuell)?

In einem randomisierten, 4x6x2 Between-Within-Within-Subject-Experiment erhalten Lehrpersonen jeweils einen Forschungsbericht, welcher zufällig aus vier verschiedenen Themen zu einer pädagogischen Intervention ausgewählt wird

(Between-Subject-Faktor). Dieser kontextualisiert 12 anschließend präsentierte Forschungsergebnisse, die zufällig in der dargestellten Effektstärke (Within-Subject-Faktor) und dem Präsentationsmodus (Within-Subject-Faktor, Beispiele: <https://t1p.de/drK0w>) variieren.

Für den verbalen Präsentationsmodus werden Cohen's U3 Formulierungen verwendet, da diese als intuitiv verständlich und aussagekräftig gelten (Cohen, 1988). Für den visuellen Präsentationsmodus werden Half-Eye-Plots genutzt, da bisherige Ergebnisse andeuten, dass diese als akkurat und z. B. besonders informativ wahrgenommen werden (Schmidt et al., im Druck).

Die primäre abhängige Variable ist die wahrgenommene Effektstärke, welche mittels eines Probability-of-Superiority-Items (vgl. Zhang et al., 2023; Beispiel: <https://t1p.de/drK0w>) erfasst wird und nach Transformation und Abgleich mit der wahren Effektstärke die AV1 bildet. Sekundäre abhängige Variablen sind die wahrgenommene Informativität (AV2) und Schwierigkeit (AV3). Beide abhängigen Variablen werden mittels siebenstufigen Likertitems erfasst.

In einer bereits durchgeführten Pilotstudie mit 240 Effektstärkeneinschätzungen von N = 20 englischsprachige Lehrpersonen (Prolific) wurden diese Messinstrumente zunächst geprüft. Eine Hauptuntersuchung mit N = 60 englischsprachigen Lehrpersonen (Prolific; Poweranalyse: <https://t1p.de/drK0w>) ist geplant.

Die Forschungsfragen werden mit Generalized Linear Mixed Models mithilfe des brms-Pakets (Bürkner, 2017) modelliert. Dabei wird die sogenannte ROPE + HDI Entscheidungsregel angewandt (Kruschke, 2018), um substanzielle gegen vernachlässigbar kleine Effekte zu testen.

Die Analyse der Pilotdaten liefert bereits klare Evidenz dafür, dass die Kommunikation von Effektstärken grundsätzlich mittels beider Präsentationsmodi gelingt. Die Interpretationen visuell kommunizierter Effektstärken sind jedoch weniger verzerrt als diejenigen der verbal kommunizierten Effektstärken ($r = -0.44$, 95% CI = $[-0.56, -0.30]$). Gleichzeitig werden sie aber auch als weniger informativ ($r = -0.27$, 95% CI = $[-0.40, -0.13]$) und schwieriger zu verstehen ($r = -0.31$, 95% CI = $[-0.44, -0.17]$) eingeschätzt. Lediglich die Frage nach einer unterschiedlichen Standardabweichung der Akkuratheit (Noise) blieb inkonklusiv, sollte aber nach der Poweranalyse im Rahmen der Haupterhebung entscheidbar sein. Konsequenzen für Wissenschaftskommunikation und Lehrkräfteprofessionalisierung werden diskutiert.

Förderung evidenzorientierten Denkens von Lehramtsstudierenden mit hybridem KI- und peer-basiertem Feedback

Ingo Kollar, Tugce Özbek, Julia Hornstein, Martin Greisel
Universität Augsburg

Theoretischer Hintergrund

Der Erwerb von Fertigkeiten zur evidenzorientierten Analyse von problematischen Unterrichtssituationen ist ein wichtiges Ziel der universitären Lehramtsausbildung (Bauer & Kollar, 2023). Gleichzeitig fällt es Lehramtsstudierenden häufig schwer, bildungswissenschaftliche Theorien und Befunde systematisch auf unterrichtliche Problemsituationen anzuwenden (Hetmanek et al., 2015).

Eine effektive Möglichkeit, Lehramtsstudierende hierin zu unterstützen, liegt zum einen in Peer-Feedback. Dabei geben sich Studierende gegenseitig Feedback auf ihre Problemanalysen und überarbeiten ihre Analysen entsprechend (Double et al., 2022). Zum anderen ist inzwischen auch Künstliche Intelligenz (z.B. ein Generative Pre-Trained Transformer) dazu in der Lage, auf Basis entsprechenden Prompt Engineerings qualitativ hochwertiges Feedback zu geben (Escalante et al., 2023). Gleichzeitig birgt KI die Gefahr, dass Lernende hochwertige Prozesse wie das Geben von Feedback auf die Analysen von Kommiliton*innen an die KI „auslagern“, sodass deren lernförderliches Potenzial nicht genutzt wird.

Diese Studie untersucht daher die Effekte „reinen“ Peer-Feedbacks, „reinen“ KI-basierten Feedbacks sowie zweier Varianten hybriden Feedbacks („KI-Feedback-plus-Peer-Feedback“ und „KI-Feedback-auf-Peer-Feedback“) auf den Erwerb von Fertigkeiten zum evidenzorientierten Denken von Lehramtsstudierenden. Wir vermuteten, dass die beiden hybriden Feedbackformen günstigere Effekte auf den Erwerb von Fertigkeiten zum evidenzorientierten Denken haben sollten als die Umsetzung reinen Peer-Feedbacks. Zudem nahmen wir an, dass das reine KI-Feedback am wenigsten wirksam ist.

Methode

N = 186 fortgeschrittene Lehramtsstudierende (MFachsemester = 5.12, SD = 1.58; 60.75% weiblich) aus drei Parallelveranstaltungen im Rahmen der psychologischen Lehramtsausbildung wurden zufällig einer von vier experimentellen Bedingungen zugewiesen (reines Peer-Feedback, reines KI-Feedback, KI-Feedback-plus-Peer-Feedback, KI-Feedback-auf-Peer-Feedback).

Nach einem Vorfragebogen analysierten die Teilnehmenden drei schriftlich präsentierte, problematische Unterrichtssituationen innerhalb einer Unterrichtsstunde auf Basis zweier kurzer Theoriezusammenfassungen (ICAP-Modell von Chi & Wylie, 2014; Cognitive-Load-Theorie von Sweller, 2005). Proband*innen der Bedingung „reines Peer-Feedback“ hatten anschließend die Aufgabe, Feedback auf die Problemanalysen (= Erstbearbeitung) zweier Kommiliton*innen zu formulieren. Studierende der Bedingung „reines KI-Feedback“ erhielten stattdessen ausschließlich Feedback von Open AI's GPT-4 auf ihre Erstbearbeitungen; zu diesem Zweck wurden diese von GPT-4 auf Basis eines entsprechenden Prompts analysiert. In der Bedingung „KI-Feedback-plus-Peer-Feedback“ wurden die Proband*innen aufgefordert, Feedback auf die Erstbearbeitungen zweier Kommiliton*innen zu formulieren, erhielten anschließend zusätzlich das GPT-generierte Feedback auf diese Erstbearbeitungen und konnten ihr Feedback auf dieser Basis nochmals überarbeiten, bevor sie es ihren Kommiliton*innen zugänglich machten. In der Bedingung „KI-Feedback-auf-Peer-Feedback“ formulierten die Proband*innen zunächst Peer-Feedback auf zwei Erstbearbeitungen; danach wurde die Qualität des von ihnen formulierten Feedbacks von GPT-4 (auf Basis eines hierfür erstellten Prompts) analysiert und ihnen rückgemeldet, wie sie die Qualität ihres Feedbacks verbessern könnten, woraufhin sie ihr Feedback überarbeiteten und den Kommiliton*innen zur Verfügung stellten. Anschließend überarbeiteten die Proband*innen aller Bedingungen ihre ursprünglichen Erstbearbeitungen.

Im Posttest wurden die Fertigkeiten zum evidenzorientierten Denken mit zwei Instrumenten erfasst. Im ersten wurden unterschiedlich qualitätsvolle evidenzorientierte Analysen zu einem neuen Unterrichtsfall vorgelegt, die ihrer Qualität entsprechend sortiert werden sollten („Erkennen der Qualität von evidenzorientierten Analysen“). Im zweiten Test analysierten die Proband*innen einen weiteren Fall auf Basis zweier neuer bildungswissenschaftlicher Theorien („Durchführung einer evidenzorientierten Analyse“).

Ergebnisse und Diskussion

Hinsichtlich des Erkennens der Qualität von evidenzorientierten Analysen zeigen erste Analysen keine signifikanten Unterschiede zwischen den vier Bedingungen ($F(3,170) = 1.012$, $p = .38$). Die ersten Ergebnisse weisen also nicht darauf hin, dass hybride

Formen von (peer- und KI-basiertem) Feedback mit Blick auf den Erwerb von Fertigkeiten zum evidenzorientierten Denken besser wirken als reines Peer-Feedback. Dass reines KI-basiertes Feedback zu ähnlich guten Leistungen führt, kann als erster Hinweis auf die umfassenden Potenziale von hochwertigem Prompt Engineering zur Erstellung von Feedback auf studentische Fallanalysen interpretiert werden (Escalante et al., 2023). Die Kodierung der Leistungen im Posttest zur Durchführung evidenzorientierter Analysen laufen derzeit und werden bis zur Konferenz abgeschlossen sein. Gleiches gilt für die Analyse der Qualität des (peer- und KI-basierten) Feedbacks.

Die Abwertung episodischer Evidenz bedeutet nicht die Aufwertung wissenschaftlicher Evidenz, oder doch?

Michael Rochnia¹, Anna-Lena Molitor¹, Judith Schellenbach-Zell¹, Marc Bienefeld², Cornelia Gräsel¹

¹Universität Wuppertal, ²Universität Bielefeld

Theoretischer Hintergrund

Trotz unterschiedlicher Nuancierung dominiert in den universitären Anteilen der deutschen Lehrkräftebildung der Konsens, dass wissenschaftliche Theorien und empirische Befunde der Bildungswissenschaften einen Nutzwert für die Praxis innehaben (Diery et al., 2020). Unter Lehramtsstudierenden herrscht dieser Konsens jedoch nicht – sie stehen dem Nutzwert von empirischen Befunden und wissenschaftlichen Theorien stellenweise skeptisch gegenüber (Voss, 2022). Dies ist insofern problematisch, als dass es sich um eine ungünstige motivationale Disposition für die Nutzung dieser Wissensbestände handelt (Greisel et al., 2022). Damit geht die Gefahr einher, dass weniger veridikale Wissensbestände – wie anekdotische Evidenz und subjektive Theorien – in der Praxis genutzt werden (Kiemer & Kollar, 2021). Eine Möglichkeit, dieser Problematik zu begegnen, könnte die Steigerung des Nutzwerts von empirischen Befunden und wissenschaftlichen Theorien mittels einer Intervention sein. In Nutzwertinterventionen entdecken Menschen beispielgestützt die persönliche Relevanz eines Sachverhalts mittels reflexiven Schreibens (Hulleman & Harackiewicz, 2021; Spalding & Wilson, 2002). Rochnia und Gräsel (2022) haben im Rahmen einer Nutzwertintervention teilweise erfolgreich den von Lehramtsstudierenden wahrgenommenen Nutzwert der Bildungswissenschaften gesteigert. Die Lehramtsstudierenden in den Experimentalgruppen erhielten zwei verschiedene Beispiele für den Nutzwert der Bildungswissenschaften und reflektierten anschließend schriftlich den Nutzwert der Bildungswissenschaften. Die Proband*innen in den Kontrollgruppen rezipierten nur die Beispiele für den Nutzwert der Bildungswissenschaften – eine schriftliche Reflexion wurde nicht durchgeführt. Allerdings blieb in der Studie unklar, ob dies auch die Relevanz empirischer Befunde und wissenschaftlicher Theorie aufwertet. In der vorliegenden Sekundäranalyse werden die Daten von Rochnia und Gräsel (2022) unter diesem Gesichtspunkt reanalysiert.

Fragestellung

Bevorzugen Lehramtsstudierende nach einer Nutzwertintervention bei der Suche nach Informationen über guten Unterricht empirische Evidenz und wissenschaftliche Theorien gegenüber anekdotischer Evidenz und subjektiven Theorien?

Methode

In einem experimentellen 2x2 Design wurden 184 Lehramtsstudierende (77% weiblich; $M = 27.74$, $SD = 5.76$) zufällig auf eine von vier Gruppen (EG1, EG2, KG1 und KG2) verteilt. Im Rahmen der Nutzwertintervention erhielt jede Gruppe ein unterschiedliches Treatment: EG1 las empirische Evidenz über die Korrelation von pädagogischem Wissen und Unterrichtsqualität und reflektierte anschließend schriftlich den Nutzwert der Bildungswissenschaften. EG2 reflektierte den Nutzwert der Bildungswissenschaften ebenfalls schriftlich, las davor jedoch wissenschaftliche Theorien über das Reflexionspotenzial der Bildungswissenschaften. KG1 und KG2 lasen nur die empirische Evidenz beziehungsweise die wissenschaftlichen Theorien.

Vor und nach dem Treatment brachten alle Gruppen empirische Befunde, wissenschaftliche Theorien, episodische Evidenz und subjektive Theorien in eine Rangfolge von 1 bis 4. Mit dieser Rangfolge bildeten sie die Relevanz ab, die der Wissensbestand für sie hat, wenn sie Informationen über guten Unterricht suchen. Veränderungen in dieser ordinalen Relevanzreihenfolge über die Zeit und zwischen den Gruppen untersuchen wir mittels Cumulative Link Mixed Models (Tutz & Hennevoogel, 1996).

Ergebnisse

Vor dem Treatment priorisierten die Proband*innen episodische Evidenz am stärksten ($MW = 1.8$, $SD = 0.90$) gegenüber empirischen Befunden ($rRB = -.04$; 95% CI $[-.16, .08]$), wissenschaftlichen Theorien ($rRB = -.43$; 95% CI $[-.52, -.32]$) und subjektiven Theorien ($rRB = -.64$; 95% CI $[-.71, -.57]$). Die Intervention induzierte jedoch eine Veränderung in der Relevanzrangfolge von t1 zu t2: In Cumulative Link Mixed Models zeigte sich eine moderate Senkung des Rankings der episodischen Evidenz über alle Experimentalgruppen hinweg ($d = .40$; 95% CI $[.09, .71]$). Parallel dazu stieg das Ranking der wissenschaftlichen Evidenz an ($d = -.29$; 95% CI $[-.59, .00]$). Die Veränderungen in den Rankings beider Theoriearten waren nicht bedeutsam. Demnach ist zu vermuten, dass bei einer Suche nach Wissen über guten Unterricht nach dem Treatment episodische Evidenz weniger priorisiert wird und stärker mit wissenschaftlicher Evidenz konkurriert. In der Diskussion gehen wir den Schwächen der Nutzwertintervention nach und geben einen Ausblick darauf, wie zukünftige Nutzwertinterventionen gestaltet werden könnten, um größere Effekte zu erzielen. Dies betten wir in die Debatte zur Förderung evidenzinformierter Praxis von Lehrkräften ein (vgl. Bauer & Kollar, 2023).

S-79: Challenges and opportunities for young refugees in Germany: How institutional conditions shape competencies and educational careers

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: SN 169

Chair der Sitzung: Melanie Olczyk

Chair der Sitzung: Julian Seuring

Chair der Sitzung: Gisela Will

Chair der Sitzung: Birgit Becker

English Track

Challenges and opportunities for young refugees in Germany: How institutional conditions shape competencies and educational careers

Chair(s): **Melanie Olczyk** (Martin-Luther-Universität Halle-Wittenberg, Germany), **Julian Seuring** (Leibniz Institute for Educational Trajectories, Germany), **Gisela Will** (Leibniz Institute for Educational Trajectories, Germany)

Diskutant*in(nen): **Birgit Becker** (Goethe-Universität Frankfurt, Germany)

In the mid-2010s, a considerable number of refugees applied for asylum in Germany, many of whom were minors. The integration of these young refugees into the German educational system represents a major challenge for both the refugees and the school system. The initial stages of this process are of particular importance, as they can have a significant impact on the young refugees' future educational careers. In this symposium, we examine the role of institutional conditions in shaping the opportunities and support that young refugees experience when starting their education in Germany. We focus on conditions at different levels and address a variety of factors, including those at the national level (e.g., asylum policies), at the federal state level (e.g., inclusive or separate instruction of young refugees), and within educational institutions (e.g., teaching practices). The four contributions examine the influences of these conditions on the educational attainment, competencies, and wellbeing of young refugees.

The first contribution investigates the conditions of school readiness of refugee children at the start of primary school. School readiness is understood as a set of competencies associated with early development domains (cognitive, linguistic, socio-emotional, and physical development). In addition to individual and familial characteristics (e.g., mental health, language support, parental education), the study focuses on the correlations between institutional support (e.g., preschool attendance, language instruction) and the children's general cognitive abilities and German language proficiency.

The second contribution addresses the enrolment in primary school, focusing on the conditions associated with delayed enrolment and the type of class attended in first grade (i.e., attending a regular class versus a separate class for 'newcomers'). Comparing participation rates across five federal states, the study analyses how regulations of the educational system relate to the two outcomes.

The third contribution explores the relationship between teaching practices in primary school and children's cognitive and psychological outcomes. The study examines the impact of culturally responsive teaching on German language and mathematics competencies, school belongingness, and the well-being of children with and without immigrant background or refugee status.

The fourth contribution examines the relationship between legal regulations and the acquisition of German language proficiency among young refugees attending lower secondary education. The study emphasises institutional conditions at different levels (national, federal states, districts), including the waiting time for school enrolment, attendance in 'newcomer' classes, and residence permit status.

Overall, the symposium synthesises research from educational science, psychology, and sociology addressing a wide range of institutional conditions and academic outcomes, thereby elucidating a comprehensive picture of early educational integration of young refugees.

Beiträge des Symposiums

Refugee children's school readiness: key conditions at the individual, family, and institutional level

Ersoy Erdemir¹, Aileen Edele², Gisela Will³

¹Boğaziçi University, Istanbul, Turkey, ²Humboldt Universität zu Berlin, Germany, ³Leibniz Institute for Educational Trajectories, Germany

School readiness refers to an interdependent set of competencies associated with early development domains that lay the foundation for academic learning in primary school (Busch et al., 2022; Snow, 2006). In educational research, these domains pertain to cognitive, language, socio-emotional, and physical development (Pianta et al., 2020). Decades of research have documented the strong link between school readiness in early childhood and later school achievement in primary school and beyond (Duncan et al., 2007). In line with ecological models of human development (Bronfenbrenner, 1996; Suárez-Orozco et al., 2018), the status of school readiness at primary school entry is shaped by numerous factors at levels of the child, the family, and the institution (Ndijuye, 2023). While robust readiness is vital for all children, socioeconomic and contextual conditions affect the extent to which children enter primary school developmentally ready to thrive academically (Magnuson et al., 2004). Given the exponential rates of refugee migration and the pronounced developmental risks of refugee children, examining school readiness among displaced children is critical (Murphy et al., 2018; UNHCR, 2023). Evidence suggests that displacement in early years disrupts acquisition of developmental foundations required for learning trajectories during primary school (Bouchane et al., 2018). Previous research has documented that refugee children's cognitive, language, and socio-emotional competencies are negatively affected by displacement, hindering school readiness (Bernhardt et al., 2024). Yet, there is a dearth of evidence with respect to how child-, family-, and institution-level conditions influence school readiness among children who first-hand experienced displacement.

Following an ecological approach, this research addresses this gap by elucidating multiple factors associated with levels of school readiness in early childhood among refugee children. The guiding research question is as follows: Which conditions at the child, family, and institutional level are pivotal to refugee children's school readiness at the start of primary school? We expect factors at the levels of the child, the family, and the institution to affect children's school readiness, with a particular strong effect of child mental health, language support, socioeconomic status of the family (SES) along with aspects of institutional quality. The research draws on data of the ReGES study (Refugees in the German Educational System), which followed approximately 2,400 children in Germany after their fleeing from crisis areas (e.g., Syria, Afghanistan and Iraq) across 7 waves of data collection. The current study focuses on the first wave which included young children who were not yet enrolled in school at the time of the measurement.

In first analyses, we focus on cognitive and language skills as dependent variables and determine the role of different conditions at levels of the child (e.g., time since resettlement; asylum status; mental and physical health), the family (e.g., parental education; home learning environment), and at the institutional level (preschool attendance; language instruction).

Linear regression models show that both outcomes are significantly positively related to family support. Institutional conditions are also important: Institutional language instruction appears to be particularly important for German language skills. Refugee-specific conditions, such as an insecure residence status and living in a group accommodation, are significantly negatively correlated with cognitive skills. This research identifying conditions that shape refugee children's school readiness is significant for (1) illuminating whether mechanisms known from previous research on majority and other minority children apply to refugee children, (2) garnering empirical knowledge that may inform early education policies to better respond to needs of young refugee children, (3) and potentially contributing to contents of early intervention programs toward more custom-designs that capitalise on refugee-specific mechanisms leading to improved school readiness outcomes.

Institutional conditions of school enrolment delays and newcomer class attendance among refugee children in Germany

Melanie Olczyk¹, Julian Seuring², Gisela Will²

¹Martin-Luther-Universität Halle-Wittenberg, Germany, ²Leibniz Institute for Educational Trajectories, Germany

For children who fled their country of origin before starting school, enrolling in primary school in the host country marks a crucial step in their educational career. This study examines the institutional conditions that are associated with a delayed primary school enrolment among school-age refugee children and with the type of class they attend upon entering school, i.e., regular classes or 'newcomer' classes (i.e., special classes for refugees and immigrants providing additional support, such as language instruction). Both delayed school entry and attendance in newcomer classes can affect children's future educational paths.

Delayed primary school enrolment results from the interaction between parents, teachers, school leaders, and educators (if preschool was attended). In Germany, prerequisites for enrolment include age and readiness for school, determined by cognitive, social-emotional, and/or physical abilities. However, the enrolment prerequisites vary across federal states. A child's attendance in a newcomer class mainly depends on the availability of such classes and institutionally defined access conditions (e.g., German language skills), which also vary between federal states. Empirical evidence has shown that children from families with lower educational and socioeconomic status (Kratzmann & Schneider, 2008), as well as children from immigrant families (Autorengruppe Bildungsberichterstattung, 2010), are more likely to experience a delayed school entry. Language barriers are a primary reason for this delay among immigrant children (Wehner, 2011), while other school-related skills also play a role (Tuppat & Becker, 2014). In addition, children who are not proficient in the language of instruction are more likely to attend newcomer classes. We know little about the role of the institutional context (Tuppat et al., 2016; for an exception) and the situation of refugee children. Thus, we investigate how institutional regulations affect both delayed school entry and class type attendance among refugee children, emphasising differences between federal states.

We used data from the ReGES study (Refugees in the German Educational System), focusing on a cohort of children not yet enrolled in school at the time of study in 2018. The study was conducted in Bavaria, Hamburg, North Rhine-Westphalia, Rhineland-Palatinate, and Saxony. Our subsample included 1,234 school-age children (48% female; on average 6 years old). We applied linear probability models to predict the likelihood of enrolling in primary school and attending a newcomer class. Our analysis controlled for differences between federal states and considered individual and family characteristics such as German language skills, preschool attendance, residence status, and parents' education level.

Most refugee children in the sample began primary education on time, with only 7.5% experiencing a delayed start. Comparing this to official statistics, refugee children were more likely to start school later. Multivariate analysis indicated that the state in which the children lived mattered, even when controlling for individual and family characteristics. Children in Bavaria, Rhineland-Palatinate, and Saxony were more likely to have a delayed school entry. Regarding our second outcome, most children started in regular classes, with only 7.5% in newcomer classes. The analysis again highlighted the relevance of the institutional context, with Bavaria and Saxony having the highest newcomer class participation rates. Furthermore, there was a notable interplay between delayed entries and newcomer class attendance: as the share of delayed entries decreased, the share of newcomer class participation increased, and vice versa. In general, it is important to understand in detail the interplay between delayed enrolment and class type as well as the underlying mechanisms in order to address potential barriers to educational participation and improve support for refugee children's education. For both outcomes, lacking German language skills was found to be a significant factor. Consequently, enhancing language support for newly arrived children before school starts is crucial.

Valuing cultural diversity in school as a key for the academic and psychological adjustment of immigrant, refugee, and non-immigrant children: results from a multi-informant study

Francesca Ialuna¹, Sauro Civitillo², Maja Schachner³, Philipp Jugert¹

¹University of Duisburg-Essen, Germany, ²Utrecht University, Netherlands, ³Martin-Luther-Universität Halle-Wittenberg, Germany

Ethnic diversity is an integral part of society. Many children in European and German schools have experienced migration or forced displacement. Evidence shows that immigrant, refugee, and ethnic minoritized children often have to face challenges (e.g., discrimination) that may undermine their academic and psychological adjustment (i.e., school achievement, belongingness, wellbeing; see Civitillo et al., 2023; Dimitrova et al., 2016; Stevens et al., 2020). Thus, it is crucial to understand what school factors can support the adjustment of immigrant and refugee as well as non-immigrant children.

Culturally responsive teaching (CRT) is a framework including teaching practices that positively address cultural diversity (Gay, 2018). Teachers can employ CRT practices such as designing culturally relevant lessons, thus making students feel appreciated and included in school. Accordingly, US-based literature highlighted positive effects of CRT for students' adjustment (Aronson & Laughter, 2016; Bottiani et al., 2020). Yet, there is still limited evidence from Europe. Moreover, research examining underlying processes explaining how CRT is linked with students' outcomes is rare. With CRT practices, teachers may contribute to a classroom climate promoting intercultural and heritage learning and equal treatment between students. These two dimensions may explain how CRT is related to students' outcomes. Accordingly, they can enhance school achievement, belongingness and wellbeing of ethnically diverse adolescents (Schachner et al., 2016, 2021). So far, little is known about the associations between CRT, cultural diversity climate and the adjustment of students in primary school, and specifically of immigrant and refugee children.

First, we quantitatively examined the link between teachers' CRT and children's school achievement, belongingness and wellbeing. Second, we investigated to what degree a climate fostering intercultural and heritage learning and equal treatment

between students mediates the association between CRT and children's outcomes. Third, we compared all associations between immigrant and refugee vs. non-immigrant children, in order to isolate potential unique effects related to children's immigrant experience.

We recruited N = 234 fourth-grade children (Mage = 10.45, SD = .55, 38% of first-generation immigrant and refugee) and their classroom teachers (N = 41) in Germany. CRT was assessed through teachers' reports of their CRT self-efficacy, which is currently one of the most used measures of CRT (Lawson et al., 2024) and is linked with CRT implementation (Comstock et al., 2023). We assessed children's school achievement with tests on German vocabulary, reading comprehension and mathematical competence. Children reported their perceptions of heritage and intercultural learning and equal treatment by students, school belongingness and wellbeing with paper-and-pencil questionnaires. We performed multilevel analyses with Mplus 8.6 (Muthén & Muthén, 2017).

Results showed that teachers' CRT self-efficacy was positively linked with children's German vocabulary and mathematical competence. Classroom climate did not significantly mediate these associations. Heritage and intercultural learning was positively associated with children's school belongingness and wellbeing, but negatively with reading comprehension. Equal treatment by students was positively linked with mathematical competence. All associations were similar for immigrant and refugee as well as non-immigrant children.

Results indicate that CRT practices and a climate promoting cultural diversity and equal treatment are beneficial not only for immigrant and refugee, but also for non-immigrant children. Future research including information on classroom ethnic composition may explore the negative link between heritage and intercultural learning and reading comprehension. It could be that in classrooms with more immigrant and refugee children, reading comprehension levels are lower, but teachers may feel important to promote heritage and intercultural learning. Overall, the study suggests that teachers should incorporate discussions on diverse cultural traditions in their lessons, and encourage respectful treatment between children. Therefore, trainings and courses in CRT should become fundamental part of initial teacher preparation and professional development for teachers.

Institutional conditions and acquisition of destination language skills among young refugees: investigating the German context

Oliver Winkler¹, Anne-Kathrin Carwehl²

¹Martin-Luther-Universität Halle-Wittenberg, Germany, ²Federal Office for Migration and Refugees, Germany

For young refugees of school age, the acquisition of language skills in the main language of the host country is key for their educational integration. Opportunities to learn the destination language are structured by legislative and regulatory frameworks established in the host country, especially in the domains of education and asylum policy. In our study, we investigate if 1) shorter waiting times of school enrolment, 2) integration into regular classes instead of separate refugee classes and 3) a secure residence status are positively correlated with German language skills of young refugees. Our research question is: Are different institutional conditions related to individual acquisition of German language skills of young refugees?

Legislations that affect destination language learning opportunities are implemented at different levels of administration (national, federal states, districts). In Germany, policy instruments regarding education are responsibilities of the federal state level. Federal states (and to some degree also schools, i.e. organizations) control the timing of school enrolment and the type of initial schooling (separate preparatory classes vs. regular classes) for newly arrived immigrants (Korntheuer & Damm, 2020; Will et al., 2022). While longer waiting times tend to delay the beginning of destination language learning, preparatory classes have been shown to not successfully reduce inequality in language skills regarding refugee students who have never learnt in a preparatory class (Höckel & Schilling, 2022). National asylum policies, such as those concerning the residence status, can also impact language acquisition opportunities because a secure residency status can increase motivation to learn the host country language (Homuth et al., 2020). Although refugee integration is also a local responsibility, we expect few variations of destination language opportunities at the local level because districts have limited authority in education and asylum legislations (Khalil et al., 2022; Tjaden & Spörlein, 2023).

Empirically, we use secondary data of the ReGES study (Refugees in the German Educational System) which surveys adolescent refugees aged 14-16. We restrict the sample to youths who currently attend regular classes. We use data of the first survey wave, applying cross-sectional analyses. Our dependent variable is the receptive vocabulary of the German language. The independent variables are waiting time until school enrolment in Germany (measured in months), whether the student has ever attended a preparatory class or has always been learning in regular classes (0: no, 1: yes), and legal status (0: secure, 1: insecure). We analyse four area states in the data (Bavaria, North Rhine-Westphalia, Rhineland-Palatinate and Saxony) and apply multi-level analyses (level 1: individuals; level 2: districts). The analysis comprises N = 1,097 students.

Our findings reveal that prolonged waiting times for school enrolment and insecure residence status are significantly associated with fewer destination language skills. Furthermore, young refugees who attended preparatory classes earlier remain to have lower German language proficiency in regular classes compared to those who did never, although the cross-sectional design limits definitive conclusions about the impact of these classes. Overall, the three institutional conditions showed significant coefficients, even after controlling for individual and family resources which emphasises the importance of the analysed institutional conditions to constrain agency of refugee families. Remaining significant differences between federal states indicate that more educational legislations than the ones regarded could contribute to additional differences in destination language skills. Districts were confirmed to play a minor role for second language acquisition. The study concludes that coordinated policy efforts in education and asylum across different administrative levels are necessary to enhance the linguistic integration of refugees.

S-80: Heterogenität der Schülerschaft und herkunftsbedingte Bildungsungleichheiten: Welche Rolle spielen Lehrkräfte, Lehrkrafturteile und Unterrichtsgestaltung?

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: EO 382

Chair der Sitzung: Katja Scharenberg

Chair der Sitzung: Katharina Maag Merki

Heterogenität der Schülerschaft und herkunftsbedingte Bildungsungleichheiten: Welche Rolle spielen Lehrkräfte, Lehrkrafturteile und Unterrichtsgestaltung?

Chair(s): **Katja Scharenberg** (Ludwig-Maximilians-Universität München, Deutschland)

Diskutant*in(nen): **Katharina Maag Merki** (Universität Zürich)

In Bildungseinrichtungen zeigt sich die Heterogenität der Schülerschaft z.B. an einem wachsenden Anteil von Schüler:innen mit Migrationshintergrund oder Fluchterfahrungen oder am steigenden Inklusionsanteil von Kindern und Jugendlichen mit sonderpädagogischem Förderbedarf (SPF). Empirische Studien haben wiederholt gezeigt, dass Schüler:innen aus bestimmten sozialen Gruppen hinsichtlich der Nutzung von Bildungsangeboten und der Erlangung von Bildungsabschlüssen benachteiligt sind: So besuchen Schüler:innen aus Familien mit niedrigerem sozioökonomischen Status oder Lernende aus zugewanderten Familien häufiger eine Förderschule (z.B. Cooc & Kiru, 2018; Kronig, 2003), wiederholen häufiger ein Schuljahr (z.B. Bayard & Schalit, 2016) oder werden auf der Sekundarstufe I häufiger einem Bildungsgang mit einem niedrigeren Anforderungsniveau zugeteilt (z.B. Müller & Hofmann, 2016). Des Weiteren liegen empirische Hinweise auf die soziale Selektivität bei der Vergabe schulischer Fördermaßnahmen vor, z.B. bei der Gestaltung zieldifferenten Unterrichts bei Grundschüler:innen oder der Erteilung von Nachteilsausgleichen (Sahli Lozano et al., 2023). Hierbei zeigten sich Disparitäten in der Bereitstellung entsprechender Bildungsangebote in Abhängigkeit der sozialen Herkunft der Lernenden.

Neben individuellen Herkunftsmerkmalen der Schüler:innen spielen auch Lehrkräfte eine zentrale Rolle, wenn es um die Bildungsbeteiligung geht – sei es bei der Einschätzung schulischer Leistungen oder des Sozialverhaltens, bei Übertrittsempfehlungen, der Einleitung von besonderen Maßnahmen wie z.B. Sonderbeschulung, zieldifferentem Unterricht oder Nachteilsausgleich. Hier zeichnen sich Verzerrungen des Lehrkrafturteils und somit Benachteiligungen von bestimmten Gruppen von Lernenden ab. Südkamp et al. (2018) fanden systematische Verzerrungen der Lehrkräfteeinschätzungen in Abhängigkeit vom SPF. Auch ergaben Studien, dass die Erwartungen von Lehrkräften nach dem Migrationshintergrund der Lernenden verzerrt sind (Lorenz, 2021). Weiter wurde aufgezeigt, dass kognitive Fähigkeiten und die Anstrengungsbereitschaft von Schüler:innen mit Migrationshintergrund oder einer anderen Erstsprache als der Unterrichtssprache auch bei gleichen Leistungen niedriger eingeschätzt werden als bei Schüler:innen ohne Migrationshintergrund (Neuenschwander & Niederbacher, 2021).

Die Gruppe der Lernenden mit Migrationshintergrund, niedrigem SES, geringen Kenntnissen der Unterrichtssprache oder SPF ist jedoch vielfältig und in sich heterogen. Daher stellt sich die Frage, ob bestimmte Subgruppen von Lernenden einer bestimmten sozialen Gruppe (z.B. jene mit Migrationshintergrund oder SPF) besonders von Bildungsbenachteiligung betroffen sind und welche Rolle dabei die Einschätzung der Lehrkräfte und die Unterrichtsgestaltung spielen. Die Beantwortung dieser Fragen ist für die pädagogische Praxis von hoher Relevanz, damit Bildungseinrichtungen der Vielfalt der Schüler:innen gerecht werden und deren individuelle Entwicklung und soziale Partizipation fördern können (Gräsel, Decristan & König, 2017).

Fragestellungen

Unser Symposium adressiert zwei zentrale Fragestellungen: 1.) Sind bestimmte Subgruppen einer jeweiligen sozialen Gruppe besonders von verzerrten Lehrkrafturteilen und in der Folge von Benachteiligungen betroffen? 2.) Werden bestimmte Subgruppen von Schüler:innen bezüglich der Zuweisung zu einer spezifischen Fördermaßnahme zum Ausgleich herkunftsbedingter Disparitäten benachteiligt?

Beiträge des Symposiums

Das Symposium umfasst vier Beiträge. Dabei werden zum einen unterschiedliche Schüler:innengruppen und Altersstufen in den Blick genommen. Zum anderen wird die Rolle der Lehrkräfte beim Ausgleich herkunftsbedingter Bildungsungleichheiten adressiert.

Beitrag 1 untersucht, ob bestimmte Gruppen von Schüler:innen mit Migrationshintergrund in besonderem Maße gefährdet sind, mit verzerrten Leistungserwartungen und Schulniveauserwartungen von Lehrpersonen konfrontiert zu werden.

In Beitrag 2 wird der Frage nach dem Einfluss von sozialer Herkunft und Migrationshintergrund auf die Maßnahmen „zieldifferenten Unterricht“ und „Gewährung eines Nachteilsausgleichs“ in Schulklassen der Sekundarstufe I untersucht.

In Beitrag 3 wird analysiert, wie genau Lehrkräfte die soziale Partizipation der Schüler:innen ihrer Klasse einschätzen können und ob die Lehrkräfteeinschätzungen bei Schüler:innen mit SPF weniger genau ausfallen als bei Kindern ohne SPF.

Beitrag 4 richtet den Blick auf die Lehrpersonen und befasst sich mit der Frage nach der Bedeutung der individuellen Bezugsnormorientierung der Lehrkräfte für das Selbstkonzept von Schüler:innen mit und ohne SPF in inklusiven Schulklassen.

In der abschließenden Gesamtdiskussion werden die Beiträge zusammenfassend gerahmt und die Bedeutung von Lehrkräften und der Unterrichtsgestaltung für den Ausgleich herkunftsbedingter Bildungsungleichheiten diskutiert.

Beiträge des Symposiums

Stereotypen in Klassenzimmer: Schülernationalitäten und verzerrte Lehrpersonenerwartungen

Markus P. Neuenschwander, Ariana Garrote

Fachhochschule Nordwestschweiz

Theoretischer Hintergrund

Forschungsergebnisse belegen moderate, aber konsistente Effekte von Lehrpersonen-Leistungserwartungen (Überzeugungen über die zukünftigen Leistungen der Lernenden) auf die Leistung der Lernenden (Brophy, 1985; Urhahne & Wijnia, 2021; Wang et al., 2020; Zhu et al., 2018). Überdies erklärt die Lehrpersonen-Schulniveauserwartung (Überzeugungen über das zukünftige Schulniveau von Lernenden) die Schulniveauzuweisung in der Sekundarstufe I.

Verzerrte Lehrpersonenerwartungen können zu Diskriminierung führen, weil sie Schüler:innen gewisse Lerngelegenheiten vorenthalten können (Babad et al., 1989). Zahlreiche Studien zeigen, dass Lehrpersonenerwartungen nach dem Migrationshintergrund der Schüler:innen verzerrt sind (Lorenz, 2018; 2021; Glock & Krolak-Schwerdt, 2014; van den Bergh et al., 2010). Obwohl die Gruppe der Lernenden mit Migrationshintergrund vielfältig ist, wissen wir wenig darüber, welche Gruppen von Lernenden mit Migrationshintergrund von Erwartungsverzerrungen besonders betroffen sind. Gemäß dem Stereotype Content Model (Fiske et al., 2002; Cuddy et al., 2009) kann eine Nationalität als kompetent/inkompetent oder warm/kalt gesehen werden. Nationalitäten, die als kalt und kompetent wahrgenommen werden, gelten als besonders bedrohlich. Schüler:innen dieser Nationalitäten werden daher mit negativ verzerrten Lehrpersonenerwartungen konfrontiert.

In einer Studie beurteilten Studierende gemäß diesem Modell die Stereotype über 70 verschiedene Nationalitäten (Neuenschwander et al., 2024). Diese Ratings wurden clusteranalytisch in vier Gruppen mit ähnlichen Stereotypen entlang der Dimensionen Kälte und Kompetenz zusammengefasst. Cluster 1 (Wärme und Kompetenz hoch) umfasst Nationalitäten wie Italien und Deutschland. Cluster 2 (Wärme tief, Kompetenz mittel) umfasst Nationalitäten wie Türkei, Kosovo, Serbien, Nord-Mazedonien. Cluster 3 (Wärme und Kompetenz tief) umfasst Nationalitäten wie Albanien, Eritrea, Afghanistan. Cluster 4 (Wärme hoch, Kompetenz tief) umfasst Nationalitäten wie Sri Lanka, Portugal). Mit dieser Gruppierung können Jugendliche mit Migrationshintergrund identifiziert werden, die ein erhöhtes Risiko haben, diskriminiert zu werden.

Fragestellung

Welche Gruppen von Schüler:innen mit Migrationshintergrund sind gefährdet, mit verzerrten Leistungserwartungen und Schulniveauserwartungen von Lehrpersonen konfrontiert zu werden?

Aufgrund der Forschung zur Intersektionalität sollte der sozio-ökonomische Status (SES) und die Familiensprache der Schüler:innen kontrolliert werden (Cole, 2009).

Methode

Insgesamt nahmen 1036 Schüler:innen aus 74 Klassen der Klassenstufen 4 bis 6 an der Studie teil (50.3% weiblich, 31.3% mit Migrationshintergrund, Durchschnittsalter 10.7 Jahre).

Schüler:innen gaben an, ob sie zu Hause die Unterrichtssprache Deutsch oder nicht Deutsch sprachen. Sie bearbeiteten curricular valide Tests zur Messung der Kompetenzen in Mathematik und Deutsch. Die Nationalitäten der Schüler:innen wurden aufgrund der Typologie von Neuenschwander et al. (2024) in vier Nationalitätengruppen und einer Gruppe ohne Migrationshintergrund gegliedert.

Der SES wurde anhand des Berufs der Eltern erfasst (Elternfragebogen). Es wurde der International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom & Treiman, 2010) berechnet. Pro Schüler:in wurde der höhere ISEI-Wert verwendet (HISEI).

69 Lehrpersonen (69.6% weiblich, 10% mit Migrationshintergrund, Durchschnittsalter 38.2 Jahre) berichteten in einem Papierfragebogen ihre Erwartungen an die Leistungen der Schüler:innen in Deutsch und Mathematik sowie an die Schulniveauzuweisung in der Sekundarstufe I (niedrig, mittel, hoch).

Ergebnisse und ihre Bedeutung

Stufenweise Mehrebenenanalysen mit dummykodierten Nationalitäten-Cluster zeigten, dass unter Kontrolle von Leistungen in Mathematik, HISEI und der Familiensprache die Lehrpersonenerwartungen an die Leistungen in Mathematik für Schüler:innen aus Cluster 2 niedriger waren als an Schüler:innen ohne Migrationshintergrund. Die anderen Nationalitäten-Cluster waren nicht signifikant. Die Ergebnisse für Deutsch waren ähnlich.

Im Vergleich zu Schüler:innen ohne Migrationshintergrund erwarteten die Lehrpersonen, dass Schüler:innen aus Cluster 2 signifikant eher dem niedrigen als dem mittleren oder hohen Schulniveau zugewiesen werden. Die Leistungen in Deutsch und Mathematik, HISEI und Familiensprache wurden kontrolliert. Für die anderen Cluster zeigten sich keine signifikanten Unterschiede.

Lehrpersonenerwartungen sind nur gegenüber gewissen Gruppen von Schüler:innen mit Migrationshintergrund verzerrt (Cluster 2). Die Ergebnisse sind für die Aus- und Weiterbildung von Lehrpersonen zentral, aber auch für die Entwicklung von Strategien gegen Diskriminierung im Schulkontext.

Effekte der sozialen Herkunft auf besondere Maßnahmen in inklusiven Schulen der Sekundarstufe I

Eva-Maria Holzer, Maria Wehren-Müller, Janina Kraft, Mireille Tabin, Elisabeth Moser-Opitz
Universität Zürich

Theoretischer Hintergrund

Studien haben aufgezeigt, dass Schüler:innen aus sozial weniger privilegierten sowie aus zugewanderten Familien im Bildungssystem benachteiligt sind: Kinder aus Familien mit einem niedrigen sozioökonomischen Status (SES) sowie mit Migrationshintergrund besuchen häufiger eine Sonderschule, repetieren häufiger ein Schuljahr, werden auf der Sekundarstufe I häufiger einem Leistungszug mit geringeren Ansprüchen zugeteilt und erhalten häufiger sonderpädagogische Fördermaßnahmen (u.a. Cooc & Kiru, 2018; Hibel et al., 2010; Kölm et al., 2020; Müller & Hofmann, 2016; Sahli Lozano et al., 2023). Im Kontext von inklusiver Beschulung zeigt sich die soziale Selektivität vermehrt auch bei Maßnahmen wie dem Nachteilsausgleich (NTA) oder dem zieldifferenten Unterricht (ZDU), die innerhalb der Regelschule angeboten werden. Die zwei Maßnahmen unterscheiden sich in ihren Zielsetzungen, in der Vorgehensweise der Vergabe und ihrer Wirkung für die weitere Schullaufbahn (z.B. für Deutschland: KMK, 2019; knw, 2018; für die Schweiz: Sahli Lozano et al., 2021). Bei einem NTA werden keine Anpassungen der Lernziele vorgenommen und die Beurteilung erfolgt entsprechend der standardorientierten Kriterien. ZDU beinhaltet eine Reduktion der Lernziele, verbunden mit einer individualnorm-orientierten Beurteilung, die im Zeugnis ausgewiesen ist. Damit kommt ZDU einer vorzeitigen Selektion gleich (Sahli Lozano et al., 2023) und bestimmte Schulabschüsse sind nicht mehr möglich (KMK, 2010). Während für die Vergabe eines NTA in der Regel ein Gutachten einer anerkannten Fachstelle notwendig ist, liegt die Entscheidung für ZDU in den meisten Fällen in der Kompetenz der Lehrpersonen. Dabei scheint der soziale Hintergrund einen Einfluss auf die Zuspache dieser Maßnahmen zu haben: Schüler:innen mit einem niedrigen SES erhielten in der Primarstufe häufiger die Maßnahme ZDU, Schüler:innen aus Familien mit einem hohen SES erhielten häufiger die Maßnahme NTA (Sahli Lozano et al., 2023). Dies weist auf eine soziale Selektivität innerhalb des inklusiven Schulsystems hin. Ob sich dies auch für die Sekundarstufe I zeigt, wird im vorliegenden Beitrag untersucht.

Fragestellung

Inwiefern beeinflussen Merkmale der sozialen Herkunft die Vergabe von NTA und ZDU auf der Sekundarstufe I?

Methode

Stichprobe

Zur Stichprobe gehören 83 Klassen der Sekundarstufe I mit vorwiegend grundlegenden oder mittleren Leistungsansprüchen (N = 1060 Schüler:innen), in denen Lernende mit den Förderschwerpunkten Lernen und emotional-soziale Entwicklung unterrichtet werden (n = 577 Jungen, n = 483 Mädchen).

Erhebungen

ZDU und NTA: Angaben pro Schüler:in (Fragebogen Lehrkräfte)

Migrationshintergrund: Kind im Ausland geboren oder beide Elternteile im Ausland geboren (Angaben der Eltern)

SES: PISA-Skala „Ressourcen zu Hause“ (Mang et al. 2019), Online-Befragung der Schüler:innen

Kontrollvariablen: Geschlecht, Vorhandensein einer Diagnose (ASS, ADHS, LRS, Dyskalkulie; Fragebogen Lehrkräfte); Mathematikleistungen (Eigenentwicklung von Aufgaben basierend auf Moser Opitz et al., 2021; Holzer et al., 2017; Kuhl et al., 2022).

Analysen

Logistische Regressionsmodelle: Analyse des Einflusses des SES und des Migrationshintergrundes unter Kontrolle des Vorhandenseins einer Diagnose, des Geschlechts sowie der Mathematikleistung auf die Vergabe von ZDU sowie von NTA.

Ergebnisse und ihre Bedeutung

Der SES hat einen signifikanten Effekt auf die Festlegung von ZDU: Schüler:innen mit einem niedrigen SES erhalten häufiger ZDU ($OR = 0.685$, $p = 0.008$) als Schüler:innen mit einem höheren SES. Bei der Vergabe von NTA zeigte sich im Unterschied zur Studie von Sahli Lozano et al. (2023) kein Einfluss des SES. Allerdings ergab sich ein Einfluss des Migrationshintergrundes: Schüler:innen ohne Migrationshintergrund erhalten häufiger einen NTA ($OR = 0.445$, $p = 0.003$) als Schüler:innen mit Migrationshintergrund. Damit bestätigen die Ergebnisse den Einfluss der sozialen Herkunft auf die Vergabe der Maßnahmen NTA und ZDU auch für die Sekundarstufe I und die damit verbundene soziale Selektivität im inklusiven Schulsystem. Für zukünftige Studien heißt dies, dass sowohl der SES als auch der Migrationshintergrund berücksichtigt werden müssen.

Wie gerne spielen und arbeiten die Mitschüler:innen mit Matti? Befunde zur Genauigkeit von Lehrkräfteeinschätzungen zur sozialen Partizipation bei Kindern mit und ohne SPF

Sina Schürer, Stefanie van Ophuysen

Universität Münster

Theoretischer Hintergrund

Die Eingebundenheit in positive soziale Beziehungen ist ein grundlegendes menschliches Bedürfnis (Baumeister & Leary, 1995). Die Ermöglichung dieser Eingebundenheit wird als zentrale Aufgabe inklusiver Bildung gesehen (Seitz & Scheidt, 2012). Damit diese Aufgabe gelingt, ist es hilfreich, wenn die Lehrkräfte korrekt einschätzen können, wer in der Pause ein:e (un)beliebte:r Spielpartner:in bzw. im Unterricht Arbeitspartner:in ist. Erste empirische Studien zeigen mittlere Zusammenhänge zwischen Lehrkraft- und Schüler:innenangaben zur sozialen Akzeptanz mit einer großen Streuung zwischen den Klassen (z.B. Südkamp et al., 2018).

Forschungsbefunde zeigen weiterhin, dass insbesondere Kinder mit sonderpädagogischem Förderbedarf (SPF) vergleichsweise weniger akzeptiert bzw. häufiger abgelehnt werden (Schürer, 2020). Hinsichtlich der Akkuratheit der Einschätzung der Partizipation bei Kindern mit SPF ist die Forschungslage jedoch begrenzt und es ist unklar, ob sich Nachteile für Kinder mit SPF ergeben. Wilbert et al. (2018) zeigen, dass Lehrkräfte die soziale Akzeptanz von Kindern mit SPF Lernen akkurater vorhersagen können als die von Kindern mit SPF emotional-soziale Entwicklung oder ohne SPF. Südkamp et al. (2018) finden systematische Verzerrungen der Lehrkräfteeinschätzungen in Abhängigkeit vom SPF.

Inwiefern Unterschiede in der Urteilsakkuratheit mit dem Kontext der Partizipation in Zusammenhang stehen, wurde bislang nicht untersucht. Es ist jedoch plausibel anzunehmen – und erste Studien liefern entsprechende Befunde –, dass die Partizipation in außerunterrichtlichen Kontexten (z. B. in den Pausen) anders ausfällt als im Kontext des Unterrichts (z.B. bei Gruppenarbeiten) (Frederickson & Graham, 1999). Da Lehrkräfte deutlich mehr Zeit mit den Kindern im Unterricht als in außerunterrichtlichen Settings verbringen, haben sie mehr Gelegenheit, die Schüler:innen in Arbeitsphasen zu beobachten als beispielsweise in den Pausen. Entsprechend ist zu erwarten, dass die Lehrkräfteeinschätzung zur Akzeptanz als Arbeitspartner:in genauer ausfallen sollte als die Einschätzung der Akzeptanz als Spielpartner:in.

Fragestellung

Die vorliegende Studie beleuchtet, (1.) wie genau Lehrkräfte die soziale Einbindung der Kinder ihrer Klasse einschätzen können und (2.) ob die Lehrkräfteeinschätzungen bei Schüler:innen mit SPF weniger genau ausfallen als bei Kindern ohne SPF. Darüber hinaus prüfen wir, (3.) ob sich Unterschiede in der Urteilsakkuratheit je nach Kontext der Partizipation (Spiel- vs. Arbeitspartner:in) zeigen.

Methode

Mittels standardisierter Fragebögen wurden Schüler:innen (N=975, Jahrgangsstufen 3 und 4) und ihre Klassenlehrkräfte (N=46) befragt. Die Schüler:innen beantworteten für jedes andere Kind der Klasse auf einer 3-stufigen Ratingskala die Fragen „Wie gerne spielst du in den Pausen mit ...?“ bzw. „Wie gerne arbeitest du im Unterricht mit ...?“. Die Einschätzungen der Mitschüler:innen wurden für beide Kontexte zu je einem Akzeptanzwert gemittelt. Die Lehrkräfte gaben für jedes Kind auf einem 5-stufigen Rating an, wie gut es beim Spielen in der Pause bzw. bei Gruppenarbeiten im Unterricht akzeptiert ist. Zusätzlich liegen Lehrkraftinformationen auf einem 5-stufigen Rating zu jedem Kind für die schulischen Leistungen und die emotional-soziale Kompetenz vor. Diese Werte wurden dichotomisiert: Kinder mit niedrigen Kompetenz-/Leistungswerten wurden als förderbedürftig kategorisiert. Zur Überprüfung der Fragestellungen wurden Korrelations- und Regressionsanalysen berechnet. Die Analysen in Mplus wurden über das Maximum-Likelihood-Verfahren mit einem robusten Schätzer (MLR) durchgeführt. Die hierarchische Struktur der Daten wurde über den Befehl type=complex berücksichtigt. Moderationseffekte wurden über Interaktionsterme der Variablen SPF und soziale Akzeptanz berücksichtigt.

Ergebnisse und ihre Bedeutung

Korrelationsanalysen ergaben mittlere Zusammenhänge zwischen Lehrkraft- und Schüler:inneneinschätzungen mit großen Unterschieden zwischen den Klassen ($r(\text{Spiel}) = .449$, range: $r = .099$ – $r = .833$; $r(\text{Arbeiten}) = .488$, range: $r = .056$ – $r = .836$). Effekte der

schulischen Leistung oder der emotional-sozialen Kompetenz zeigten sich nicht, sodass unsere Befunde nicht auf systematische Nachteile für Kinder mit SPF hindeuten. Es zeigten sich teilweise große Unterschiede in der Akkuratheit in Abhängigkeit vom Kontext der Partizipation. Diese variieren jedoch: Einige Lehrkräfte waren deutlich besser informiert über die Beziehungen in den Pausen, anderen Lehrkräften gelang die Einschätzung der Akzeptanz im unterrichtlichen Kontext hingegen besser.

Die Befunde werden mit Blick auf Implikationen für die schulische Praxis und die Lehrkräfteausbildung diskutiert.

Die Bedeutung der individuellen Bezugsnormorientierung von Lehrkräften für das Selbstkonzept von Schüler:innen in inklusiven Schulklassen

Katja Scharenberg¹, Eva Kouka², Sebastian Röhl², Wolfram Rollett³

¹Ludwig-Maximilians-Universität München, ²Pädagogische Hochschule Freiburg, ³Carl von Ossietzky Universität Oldenburg

Theoretischer Hintergrund

Leistungsbezogene soziale Vergleiche im schulischen Kontext beeinflussen das akademische Selbstkonzept von Schüler:innen. So weisen Schüler:innen mit ähnlichen intellektuellen Fähigkeiten bei einem höheren schulischen Leistungsniveau ein niedrigeres akademisches Selbstkonzept auf als bei einem niedrigeren schulischen Leistungsniveau (Big-Fish-Little-Pond-Effekt; Marsh, 1987, S. 280). Legen Lehrkräfte eine soziale Bezugsnorm bei der Leistungsbeurteilung an, erweist sich dies als hinderlich für die Entwicklung eines positiven Selbstkonzepts von Schüler:innen (Holder & Kessels, 2016; Rossmann et al., 2011). Umgekehrt ergaben Forschungsbefunde zum Big-Fish-Little-Pond-Effekt, dass eine individuelle Bezugsnormorientierung von Lehrkräften das akademische Selbstkonzept, die Motivation und Schulleistung von Schüler:innen verbessern kann (Lüdtke et al., 2005).

Die Gemeinschaftsschulen in Baden-Württemberg stehen als inklusive Bildungskontexte grundsätzlich allen Schüler:innen unabhängig von ihren individuellen oder leistungs- und fähigkeitsbezogenen Voraussetzungen offen (Bohl & Wacker, 2016). Hierbei handelt es sich um eine schulische Organisationsform der Sekundarstufe I, in der Schüler:innen auf individuell unterschiedlichen Anspruchsniveaus innerhalb derselben Schule oder Klasse unterrichtet werden (Edelstein, 2023). Gleichzeitig wird verstärkt nach den Prinzipien des individualisierten und kooperativen Lernens und alternativen Formen der Leistungsbeurteilung gearbeitet. Allerdings ist zu erwarten, dass auch in diesem Kontext individuelle Unterschiede in der Bezugsnormorientierung von Lehrkräften bedeutsame Effekte auf die Selbstkonzeptentwicklung der Schüler:innen haben. Eine in diesem Zusammenhang besonders vulnerable Gruppe sind Schüler:innen mit sonderpädagogischem Förderbedarf.

Ausgehend von diesen Überlegungen wird in unserem Beitrag der Frage nachgegangen, ob eine individuelle Bezugsnormorientierung der Lehrkräfte in inklusiven Schulklassen negative Bezugsgruppeneffekte, die aufgrund sozialer Leistungsvergleiche auftreten können, abmildert, und zwar insbesondere für jene der Schüler:innen mit SPF. Hierzu liegen bislang kaum empirische Erkenntnisse vor (Holder & Kessels, 2018).

Methode

Datengrundlage ist eine schriftliche Befragung von n=821 Schüler:innen aus 52 inklusiven Schulklassen (Jahrgangsstufen 5–7) an 22 Gemeinschaftsschulen in Baden-Württemberg, davon n=111 (13.5%) mit SPF. Die Stichprobe ist, wie ein Vergleich mit Statistiken des Landesinstitutes für Schulentwicklung (2018) zeigt, hinsichtlich der soziokulturellen Schülerzusammensetzung dieser Schulart repräsentativ. Informationen zum diagnostizierten SPF und dem individuellen Kompetenzniveau der Schüler:innen basieren auf Lehrkraftangaben. Das individuelle Kompetenzniveau (grundlegendes, mittleres oder erweitertes Niveau) in Mathematik und Deutsch wurde als Indikator für das Leistungsniveau der Schüler:innen verwendet, auf dem sie unterrichtet wurden. Abhängige Variable war das akademische Selbstkonzept der Schüler:innen (Bos et al., 2005). Angaben zur individuellen Bezugsnormorientierung der Lehrkräfte wurden mit dem Lehrkräftefragebogen zur Erfassung der Bezugsnormorientierung (FEBO; Rheinberg, 1980) erhoben. Fehlende Werte wurden multipel (m=30) geschätzt (Grund et al., 2019). Die Varianzzerlegung der Kriteriumsvariable (ICC=.054, p=.033) verdeutlichte die Notwendigkeit, die genestete Datenstruktur zu berücksichtigen. Daher wurden schrittweise Effekte auf Individual- und Klassenebene geschätzt (R-Paket lme4; Bates et al., 2015).

Ergebnisse und ihre Bedeutung

Bei der Vorhersage des Selbstkonzepts zeigte sich unter Kontrolle des individuellen Kompetenzniveaus der Schüler:innen kein Effekt des mittleren Kompetenzniveaus der Schulklasse ($\beta=.061$, $p=.693$). Der BFLPE ließ sich in unserer Stichprobe für die Gemeinschaftsschulen also nicht bestätigen. In Übereinstimmung mit früheren Untersuchungen hatten Schüler:innen mit SPF zwar ein niedrigeres Selbstkonzept ($\beta=-.510$, $p<.001$) als andere Schüler:innen. Diese Unterschiede wurden jedoch unter Kontrolle des individuellen Kompetenzniveaus insignifikant ($\beta=.070$, $p=.541$). Die individuelle Bezugsnormorientierung der Lehrkräfte erwies sich dagegen als positiver Prädiktor für das Selbstkonzept ($\beta=.096$, $p=.028$), und zwar auch unter Kontrolle des durchschnittlichen Kompetenzniveaus der Schulklassen.

Zusammenfassend zeigt unsere Studie, dass eine individuelle Bezugsnormorientierung von Lehrkräften eine wirksame Strategie bei der Leistungsbeurteilung an Gemeinschaftsschulen sein könnte, um das Selbstkonzept von Schüler:innen mit SPF zu verbessern. Der ausbleibende BFLPE in unserer Stichprobe unterstreicht das Potenzial inklusiver Lernsettings an Gemeinschaftsschulen für die Selbstkonzeptentwicklung.

Da sich das individuelle Kompetenzniveau der Schüler:innen als ein wichtigerer Prädiktor für das Selbstkonzept herausstellte als der SPF, könnte eine lernstandsbezogene schuljahresweise Anpassung des zu unterrichtenden Kompetenzniveaus potenziell negative Auswirkungen von Bezugsgruppeneffekten für die Schüler:innen mit SPF verringern.

S-81: Computational modeling and educational sciences: Synergies and advancements

Zeit: Mittwoch, 29.01.2025: 9:00 - 10:45 · Ort: M 003

Chair der Sitzung: Blazej Baczkowski

Chair der Sitzung: Manuel D.S. Hopp

Chair der Sitzung: Kou Murayama

English Track

Computational modeling and educational sciences: Synergies and advancements

Chair(s): **Blazej M. Baczkowski** (Hector Research Institute for Education Sciences and Psychology, University of Tübingen, Tübingen, Germany), **Manuel D.S. Hopp** (Hector Research Institute for Education Sciences and Psychology, University of Tübingen, Tübingen, Germany)

Diskutant*in(nen): **Kou Murayama** (Hector Research Institute for Education Sciences and Psychology, University of Tübingen, Tübingen, Germany)

Theories on education research are often complex—they describe dynamic relationships of various factors such as learning, skills, metacognition, and emotion. These complicated relationships cannot be easily accommodated by traditional statistical approaches, creating a large gap between theories and empirical testing. In the field of cognitive science, there is growing amount of literature using computational modeling to analyze data, which allows researchers to flexibly apply the exact theoretical model that researchers specified. This computational approach provides educational psychologists with a powerful opportunity to conduct theory-driven (yet flexible) analysis on educational data. The computational approach is different from learning analytics or machine learning approaches (although there are some similarities) in that the approach puts strong emphasis on top-down theoretical modeling, with the goal of elaborating and further developing education theories. In this symposium, we will showcase some computational approaches to analyzing educational data, with the aim of demonstrating their utility and importance for future research on education. The first contribution examines the dynamic interaction of goal setting, performance, and emotions in self-regulated learning using math task data and longitudinal intensive data from an online learning platform. The second contribution uses natural language processing tools to examine how learners retrieve and describe information acquired from reading an expository text. This approach aims to offer insights into the quality of newly acquired knowledge representation that goes beyond what we can typically infer from scoring conventional forced-choice educational quizzes. The third contribution investigates the emergence of domain-specific interest through curiosity-driven learning by employing a computational approach using knowledge network simulations. The fourth contribution simulates learning dynamics in free practice through computational simulations of a Reciprocal-Practice-Success model, integrating learning theories with motivation-achievement reciprocity in educational contexts. The last contribution proposes a performant and scalable knowledge tracing with the interpretability that personalized education needs. With the rise of digital technologies and AI, we now have access to vast amounts of data on learners' behavior. Computational approaches provide researchers with valuable tools to capitalize on this opportunity. These methods not only enable more precise testing of educational theories but also offer the potential to enhance personalized learning strategies. As these methods continue to evolve, they show promise in facilitating the development of more adaptive educational interventions and enhance personalized learning strategies.

Beiträge des Symposiums

Applying computational modeling to understand the effects of emotions on goal setting and performance in dynamic self-regulated learning

Wy Ming Lin¹, Lily FitzGibbon², Maria Theobald³, Jasmin Breitwieser⁴, Garvin Brod⁵, Kou Murayama¹, Michiko Sakaki¹

¹Hector Research Institute for Education Sciences and Psychology, University of Tübingen, Tübingen, Germany, ²Division of Psychology, University of Stirling, Stirling, Scotland, United Kingdom, ³Department of Psychology, University of Trier, Trier, Germany, ⁴DIPF | Leibniz Institute for Research and Information in Education, Frankfurt am Main, Germany, ⁵DIPF | Leibniz Institute for Research and Information in Education and Center for Individual Development and Adaptive Education of Children at Risk (IDeA), Frankfurt am Main, Germany; Department of Psychology, Goethe University, Frankfurt am Main, Germany

Theoretical Background

Self-regulated learning (SRL) is a dynamic process consisting of goal setting, task performance, and emotions (Panadero, 2017). While the interplay of goals and performance in SRL through goal-performance discrepancies (GPD) has been well investigated in control (Carver & Scheier, 1998) and social-cognitive theory (Bandura, 1991), the findings on the effects of emotions on goal setting and performance have been inconsistent. For instance, some theories posit that negative emotions lead to higher performance standards (e.g., Ahrens, 1987), while others suggest that positive emotions should have this effect (e.g., Ilies et al., 2010). Furthermore, positive emotions may encourage coasting and lower performance (Carver, 2003) or enhance effort allocation and higher performance (Rigoli & Pezzulo, 2022). To reconcile these findings, we employed computational modeling to build a formal model that will help us understand the effects of emotions that go beyond the effects of the GPD. We apply our model in two studies: a gamified online math task (Study 1) and in a naturalistic setting where students prepared online for a high-stakes exam (Study 2; Breitwieser et al., 2022).

Research question

What role do emotions play in SRL that go above and beyond those of the GPD? Can a formal model of goal setting, performance, and emotions be used to capture these dynamics?

Method

We applied our computational model on data from two tasks. In Study 1 (n = 357), we used a controlled experimental setting in which participants completed 20 two-minute blocks of math questions. Goals were set before and emotions were assessed after each block. Bonus reward was given relative to goals if participants reached them. In Study 2 (n = 305), students studied for a high-stakes medical exam using an app (at least 20 days). Goals were set before each study session and emotions were assessed before and after each session.

Our model incorporates equations that describe how individuals (or students) set their next goal, how they will perform on the task, and the emotions they will feel afterwards. Each equation has components describing the effect of the GPD (goal setting, performance, and emotions equations) as well as the effect of emotions (goal setting and performance equations).

Results

Our results show that the effect of emotions can be teased apart from the GPD. In Study 1, we found that positive emotions led to high goals while negative emotions led to low goals. We also found evidence of a coasting/pushing effect on performance in which positive emotions led to low performance while negative emotions led to high performance. Study 2 revealed a different pattern of results. While we saw that there were no effects of emotions on goal setting (possibly due to the high-stakes nature of the task), we saw the opposite effect of emotions on performance when compared with Study 1. Here, positive emotions pushed for high performance while negative emotions led to low performance. This finding could be attributed to a feeling of competency that is associated with positive emotions that boost participants' effort allocation (Rigoli & Pezzulo, 2022), leading to high performance.

Significance

Computational modeling offers valuable insights into the role of emotions in SRL, particularly in how they interact with goals and performances across different contexts. While emotions may seem to support performance on a task in one context, emotions may also hinder performance in others (compare the effect of positive emotions on performance in Study 1 vs Study 2). Through computational modeling, this study contributes to our understanding of the dynamics and cyclical interpretations of SRL by providing evidence of a possible formalization of the component phases and their relationships with each other.

Using natural language processing models to recover real-world knowledge acquisition

Blazej M. Baczowski, Sabine Oligschläger, Kou Murayama, Michiko Sakaki

Hector Research Institute for Education Sciences and Psychology, University of Tübingen, Tübingen, Germany

Theoretical background

The acquisition of real-world knowledge and the monitoring of learning progress are crucial areas of interest in educational psychology. Conventional methods of assessing learning outcomes often rely on closed-ended questions (e.g., educational quizzes), which arguably provide limited insights into the student's understanding of the learning material with a univariate measure of performance (e.g., proportion of correct answers). Indeed, to convey the understanding of a complex topic, students must not only retrieve the right memories but also coherently express them by showing how they fluently navigate through the semantic space of the newly acquired knowledge, weaving atomic concepts into a logical whole (Kintch, 1970; Maclellan, 2005). In contrast to traditional scoring of closed-ended questions, using modern tools of natural language processing (NLP) that are capable of analyzing freely generated textual answers to open-ended questions may provide far more detailed insights into the cognitive process of the test-takers.

Research question

Here, we leverage the recent advancements in NLP tools to recover the acquisition of real-world knowledge. Using NLP, we aim to track how individuals traverse the semantic space of the newly acquired knowledge when freely recounting the content of the expository text composed of unique idea units that are logically organized into a coherent narrative.

Method

To illustrate our approach, we present exemplary data collected from participants who read a popular-science text on the physics of snow avalanches that describes what they are, what causes them, and how to prevent destructive ones (Veritasium, 2024). Subsequently, in an immediate recall phase, participants recount the text in a self-paced manner to reproduce its content to the best of their ability. The obtained individual text-recounts of participants are segmented into single sentences. By means of pre-trained large language models (e.g., S-BERT; Reimers & Gurevych; 2019), these sentences are then converted into numerical vectors that capture their contextual meaning in a high-dimensional semantic space (e.g., 768 dimensions). The numerical representation of sentences allows us to train a classifier so that we can identify to what extent these individual sentences match the idea units from the original text. By analyzing these freely generated text-recounts this way, we assess how well participants reproduce the original learning material. Specifically, we quantify not only the overall percentage of content retrieved but also its depth indicating the quality of knowledge representations ranging from gist summaries to detailed description of key points of the text.

Results and their significance

Our findings exemplify the potential of using NLP to capture nuanced aspects of learning that might be obscured by traditional forced-choice formats. Also, this approach can track the formation of new knowledge and its expression in an idiosyncratic way. The implications of using NLP in the analysis of open-ended responses include the potential to scaffold individualized learning experience. For example, educators may offer personalized learning recommendations or retrieval cues to aid knowledge acquisition and retention, respectively. Our approach highlights the promise of NLP in advancing the study of learning and memory by providing deeper insights into the learning process and capturing its naturalistic outcomes.

Curiosity as a catalyst: Simulating the development of domain-specific interest in knowledge networks

Manuel D.S. Hopp, Alexandr Ten, Michiko Sakaki, Kou Murayama

Hector Research Institute for Education Sciences and Psychology, University of Tübingen, Tübingen, Germany

Theoretical background

Interest and curiosity are fundamental drivers of learning, playing crucial roles in educational contexts (Renninger & Hidi, 2016; Loewenstein, 1994). Although these constructs are frequently used interchangeably, they represent discrete psychological phenomena with distinctive characteristics in the learning process (Grossnickle, 2016).

Interest may be defined as a relatively enduring predisposition to reengage with a particular material over time (Hidi & Renninger, 2006). It is characterized by positive affect and increased value for the subject matter. In contrast, curiosity is conceptualized as a "thirst for knowledge" that drives information-seeking behavior (Loewenstein, 1994).

The underlying mechanisms of interest and curiosity have been the subject of extensive study. The development of interest often follows a four-phase model, progressing from situational interest to a more developed individual interest (Hidi & Renninger, 2006). Curiosity is commonly explained by the information-gap theory, which states that individuals are driven to close perceived gaps in their knowledge (Loewenstein, 1994).

Nevertheless, a notable gap remains in our understanding of how domain-specific interest emerges (Renninger & Hidi, 2015). We propose that domain-specific interest represents an inherent consequence of curious learning.

Research question

How does curiosity-driven learning contribute to the emergence and development of domain-specific interest over time?

Method

To investigate this question, we employ a computational approach using knowledge network simulations (Murayama, 2022). In this framework, knowledge is represented as a network where nodes denote entities (e.g., "Earth," "Sun") and edges represent relationships between them (e.g., "orbits"). Learning is conceptualized as the construction of this network representation.

Within this context, we operationalize curiosity as the drive to close knowledge gaps, represented by missing connections in the network. Interest is reflected in the sustained engagement with the network. Domain specificity is captured by the clustering and interconnectedness of nodes within specific subject areas (i.e., modularity).

We compare the behavior of two types of simulated agents:

1. "Curious" learners who actively seek to close knowledge gaps in their previously learned knowledge, and
2. "non-curious" learners who explore the network more randomly.

Agents begin at a central position in a predefined ground truth network and expand their knowledge networks step-by-step based on their respective learning rules. We utilize various knowledge network structures to ensure generalizability of our findings.

To analyze the resulting knowledge networks, we assess a diverse set of topological measures, including clustering coefficients, path lengths, and modularity. These metrics allow us to quantify the development of domain-specific interest and compare the learning outcomes of curious versus non-curious agents.

Results and their significance

Our preliminary findings suggest that curious learners tend to develop more pronounced topic interest compared to non-curious learners. Specifically, the knowledge networks of curious agents show:

1. Higher clustering coefficients within specific domains, indicating the formation of tightly interconnected knowledge structures.
2. Fewer clusters and greater modularity, reflecting the emergence of distinct, domain-specific knowledge clusters.

These effects vary depending on the underlying ground-truth network structure, highlighting the importance of considering domain-specific characteristics in understanding interest development.

These results have significant implications for educational theory and practice. They provide computational support for the idea that curiosity-driven learning facilitates the development of domain-specific interest(s). This suggests that educators should focus on strategies that stimulate curiosity and encourage active exploration of knowledge gaps to foster deep, sustained interest in academic subjects.

Predictive, scalable and interpretable knowledge tracing on structured domains

Hangi Zhou¹, Bamler Robert¹, Charley M. Wu¹, Álvaro Tejero-Cantero²

¹University of Tübingen, Cluster of Excellence Machine Learning, Tübingen AI Center, IMPRS-IS, ²University of Tübingen, Cluster of Excellence Machine Learning

Theoretical background

Computational models play a crucial role in advancing personalized education, particularly in the development of intelligent tutoring systems (Corbett & Anderson, 1994; Sidney, 1926). These systems rely on accurate estimates of learner progress (knowledge tracing; KT) and an understanding of the prerequisite structure of the target domain (Settles & Meeder, 2016; Lynn & Bassett, 2020). While recent deep learning models have achieved high KT accuracy, they often lack the interpretability of psychologically inspired models. This trade-off between performance and interpretability hinders the development of explainable and effective educational

technologies.

Research question

Can we develop a computational model for knowledge tracing that combines high predictive accuracy, scalability for real-world educational applications, and interpretability of both learner-specific traits and knowledge prerequisite structure?

Method

We introduce PSI-KT, a novel computational approach that explicitly models the influence of individual cognitive traits and knowledge prerequisites on learning dynamics. PSI-KT employs a hierarchical generative framework with three key components:

1. A probabilistic state-space model representing learner knowledge states across knowledge components (KCs).
2. Latent variables capturing learner-specific cognitive traits.
3. A shared static graph modeling KC prerequisites.

The model utilizes scalable Bayesian inference techniques, enabling efficient personalization across large numbers of learners and extensive learning histories – a critical feature for real-world educational settings. We evaluated PSI-KT using three datasets from online learning platforms (Assistments 2012, Assistments 2017, and Junyi Academy 2015), comparing our model against eight state-of-the-art computational models in education. Our analyses include a wide range of

assessments of predictive accuracy, scalability in continual learning scenarios, and interpretability of learned representations.

Results and their significance

PSI-KT demonstrated superior multi-step predictive accuracy, particularly with limited training data (100-1000 learners), outperforming baselines that required data from 60,000+ learners for comparable performance. In continual learning scenarios, mirroring real-world educational environments where the model is continually retrained on new learner data, PSI-KT exhibited the most favorable cost-accuracy trade-off. It required minimal retraining time while maintaining superior prediction accuracy compared to other baselines, providing strong advantages for real-world educational applications.

Lastly, the model's learned representations of learner traits showed high specificity, consistency, and disentanglement. Moreover, we relate these inferred learner traits to interpretable educational constructs, such as individualized forgetting rates and transfer ability across topics. PSI-KT's inferred prerequisite graphs also demonstrated superior alignment with human-annotated

curriculum structures and exhibited strong operational interpretability by predicting causal relationships between KCs. This capability offers valuable insights for curriculum design and sequencing in educational settings.

In sum, PSI-KT bridges the gap between high-performance machine learning models and interpretable cognitive modeling approaches, offering a solution to the accuracy-interpretability trade-off in educational data mining.

1. Strong performance with limited data makes it suitable for diverse educational contexts, from small classrooms to massive open

online courses.

2. Scalability in continual learning addresses the need for adaptive educational technologies that evolve with learner interactions.

3. Interpretable representations provide actionable insights for educators and learning scientists, potentially informing evidence based instructional design and personalized learning strategies.

IS-08: Invited Symposium: Die Integration von Geflüchteten in das deutsche Bildungssystem: Herausforderungen und Perspektiven

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 102

Chair der Sitzung: Irena Kogan

Chair der Sitzung: Jörg Dollmann

Die Integration von Geflüchteten in das deutsche Bildungssystem: Herausforderungen und Perspektiven

Chair(s): **Irena Kogan** (Universität Mannheim, Deutschland)

Diskutant*in(nen): **Jörg Dollmann** (Mannheimer Zentrum für Europäische Sozialforschung MZES, Mannheim und Deutsches Zentrum für Integrations- und Migrationsforschung DeZIM, Berlin)

<p>Das Symposium befasst sich mit der dringenden Frage der Integration von geflüchteten Kindern und Jugendlichen in das deutsche Bildungssystem, einem Thema von hoher wissenschaftlicher und gesellschaftlicher Relevanz. Das Symposium bringt Wissenschaftler aus verschiedenen Disziplinen zusammen, um die vielfältigen Herausforderungen zu diskutieren, mit denen diese jungen Menschen konfrontiert sind, darunter sprachliche Hürden, kulturelle Anpassung und Bildungsbrüche.</p>

<p>Rund ein Drittel aller schulpflichtigen Geflüchteten, die im Zuge der Fluchtmigration zwischen 2015 und 2018 nach Deutschland gekommen sind, wurden im deutschen Schulsystem zunächst separat in sogenannten Neuzuwandererklassen (auch: Vorbereitungs-, Integrations- oder Willkommensklassen) beschult (Pagel & Edele 2022, Winkler 2021). Die Bundesländer unterscheiden sich beispielsweise darin, ob Neuzugewanderte direkt in Regelklassen eingeschult werden oder zunächst solche Neuzuwandererklassen besuchen und nach welchen Kriterien der Übergang in Regelklassen erfolgt. Im ersten Beitrag des Symposiums werden diese länderspezifischen Regelungen näher beleuchtet und ihre Bedeutung für die Beschulung junger Geflüchteter untersucht.</p>

<p>Der zweite Beitrag des Symposiums untersucht, ob und inwieweit sich die Deutschkompetenzen und die sozio-emotionale Integration von jungen Geflüchteten in der 9. Jahrgangsstufe unterscheiden. Dabei wird der Vergleich zwischen jenen, die im Grundschulalter zunächst in einer Vorbereitungsklasse beschult wurden, und jenen, die nie oder nur kurzzeitig eine Vorbereitungsklasse besucht haben, angestellt.</p>

<p>Der dritte Beitrag untersucht die Rolle der Lehrkräfte bei der Integration von Geflüchteten, indem er das zielgruppenspezifische Fachwissen der Lehrkräfte in Vorbereitungsklassen beleuchtet. Außerdem werden die Belastungen und Bereicherungen, die sie im Unterricht erleben, sowie der Beitrag, den sie ihrer Meinung nach zur Integration ihrer Schüler*innen leisten können, betrachtet.</p>

<p>Der vierte Beitrag ist eine Fallstudie, die die Bildungsstrategien ukrainischer Flüchtlingseltern für ihre schulpflichtigen Kinder in Deutschland und Polen untersucht. Dabei werden verschiedene Bildungsstrategien betrachtet, einschließlich des Online-Unterrichts im ukrainischen Schulsystem im Ausland mit und ohne Schulbesuch in den jeweiligen Ländern.</p>

<p>In allen Beiträgen werden die strukturellen und politischen Rahmenbedingungen für die Bildung von Geflüchteten in Deutschland untersucht, ihre Wirksamkeit bewertet und Verbesserungsmöglichkeiten aufgezeigt. Empirische Analysen beleuchten die Erfahrungen von Geflüchteten in der Schule und zeigen sowohl Barrieren als auch erfolgreiche Integrationsstrategien auf. Wir erfahren, vor welchen Herausforderungen geflüchtete Eltern im deutschen Bildungssystem stehen und welche Strategien sie anwenden, um sich zurechtzufinden. Darüber hinaus wird die Rolle der Lehrkräfte bei der schulischen und gesellschaftlichen Integration von Geflüchteten beleuchtet.</p>

<p>Durch die Kombination von theoretischen Erkenntnissen und empirischen Studien soll das Symposium ein umfassendes Verständnis der aktuellen Situation vermitteln und evidenzbasierte Lösungsvorschläge zur Verbesserung der Bildungsergebnisse von geflüchteten Kindern und Jugendlichen in Deutschland aufzeigen. Die Diskussionen sollen einen Beitrag zu einem breiteren Diskurs über Bildungspolitik, Chancengleichheit und die Rechte junger Menschen im Migrationskontext leisten.</p>

Beiträge des Symposiums

Die Bedeutung rechtlicher und bildungspolitischer Rahmenbedingungen für die Beschulung neuzugewanderter junger Geflüchteter

Gisela Will¹, Regina Becker¹, Oliver Winkler²

¹Leibniz Institute for Educational Trajectories (LifBi), Bamberg, Germany, ²Martin Luther University Halle-Wittenberg, Halle, Germany

Unter den im letzten Jahrzehnt nach Deutschland gekommenen Geflüchteten befindet sich auch eine große Anzahl an Kindern und Jugendlichen. Die Integration dieser jungen Neuzugewanderten in das Bildungssystem ist eine zentrale Voraussetzung für die Teilhabe in anderen gesellschaftlichen Bereichen, wie etwa für die soziale Einbindung, aber auch für eine spätere erfolgreiche Arbeitsmarktpartizipation.

Bildungsbeteiligung und schulischer Erfolg hängen von einer Vielzahl von Faktoren ab, u.a. auch von institutionellen Kontexten und bildungspolitischen Vorgaben. In Deutschland spielen hierbei bundeslandspezifische Regelungen eine zentrale Rolle. Dies gilt auch für die Bildungsintegration junger Geflüchteter in das allgemeinbildende Schulsystem. So unterscheiden sich die Bundesländer beispielsweise darin, ob kürzlich Zugewanderte direkt in reguläre Klassen eingeschult werden oder zunächst sogenannte Neuzuwandererklassen besuchen sowie bezüglich der Kriterien, die für einen Übergang in reguläre Klassen angewendet werden (z.B. Autor:innengruppe Bildungsberichterstattung, 2024; Massumi et al., 2015).

Im vorliegenden Beitrag werden diese bundeslandspezifischen Regelungen näher betrachtet und deren Bedeutung für die Beschulung von jungen Geflüchteten untersucht. Aufbauend auf bisherigen Forschungen, die bereits auf Zusammenhänge zwischen unterschiedlichen institutionellen Rahmenbedingungen auf Bundeslandebene und Bildungsteilhabe verwiesen haben, diese jedoch nicht systematisch untersucht haben (z.B. Will & Homuth, 2020), betrachten wir Unterschiede in den institutionellen Regelungen hinsichtlich vier abhängiger Variablen: der Dauer bis zur Einschulung in Deutschland, der Art der besuchten Klasse (reguläre Klasse vs. Neuzuwandererkategorie), der besuchten Schulform sowie ob die Schüler:innen altersadäquat in die entsprechende Klassenstufe eingeschult wurden.

Wir stützen uns auf ein theoretisches Modell, das Bildungsentscheidungen als Ergebnis rationaler Kosten-Nutzen-Kalkulationen annimmt (z.B. Erikson & Jonsson, 1996). Die individuellen Bildungsinvestitionen hängen von individuellen Motivationen und Ressourcen innerhalb einer gegebenen Gelegenheitsstruktur ab. Wir integrieren die bundeslandspezifischen Regelungen über die Gelegenheitsstrukturen in das theoretische Modell. Wir erwarten deutliche Bundeslandunterschiede hinsichtlich der Dauer bis zur Einschulung, der besuchten Klassenform und der besuchten Schulform sowie im Hinblick auf die altersadäquate Beschulung, da hier die rechtlichen Rahmenbedingungen stark zwischen den Bundesländern variieren.

Unsere Analysen basieren auf der ersten Welle der Daten der Studie „Refugees in the German Educational System (ReGES)“ (Will et al., 2021). Die ReGES-Studie wurde in fünf Bundesländern durchgeführt, die sich hinsichtlich zentraler Makroindikatoren (z.B. bezüglich ihrer Beschulungsstrategie Neuzugewandelter) unterscheiden. Wir werten Daten der Jugendlichenkohorte aus, die 2.415 Geflüchtete insbesondere aus Syrien, Afghanistan und dem Irak zwischen 14 und 16 Jahren umfasst. Zum Zeitpunkt der ersten Befragung befanden sich die Jugendlichen in der Sekundarstufe I des allgemeinbildenden Schulsystems.

Die Ergebnisse linearer Regressionen und linearer Wahrscheinlichkeitsmodelle zeigen signifikante bundeslandspezifische Unterschiede bezüglich aller untersuchten Indikatoren der Bildungsintegration. Diese sind weitestgehend robust, auch wenn weitere, in der Forschung zu sozialen und ethnischen Bildungsungleichheiten als relevant erachtete Faktoren (z.B. Diehl et al., 2016) kontrolliert werden. Auch zusätzliche Analysen mit Daten der IAB-BAMF-SOEP-Befragung von Geflüchteten, die es erlauben alle 16 Bundesländer zu betrachten, zeigen Zusammenhänge zwischen bundeslandspezifischen Regelungen und der Beschulung neuzugewandelter Jugendlicher. Die Befunde zeigen, dass rechtliche Regelungen die Bildungsbeteiligung Neuzugewandelter stark strukturieren und dass es wenig individuellen Spielraum für Bildungsentscheidungen zu geben scheint. Die Zuweisung zu einem Bundesland bestimmt somit maßgeblich die Dauer bis zur Einschulung, die Einstufung in eine altersgerechte Klassenstufe und den Besuch einer Neuzuwandererklasse. Die bildungspolitischen Regelungen hängen darüber hinaus stark mit den Chancen von Geflüchteten zusammen, ein Gymnasium zu besuchen. Aufgrund der geringen Durchlässigkeit des deutschen Bildungssystems ergeben sich hieraus Pfadabhängigkeiten für den weiteren Bildungs- und Berufsweg der neu zugewanderten Schüler:innen.

Deutschkompetenzen und sozial-emotionale Integration von geflüchteten Jugendlichen: Lassen sich mittelfristig Unterschiede für den Besuch einer separierten vs. Regelklasse feststellen?

Oliver Winkler¹, Aileen Edele², Stefan Schipolowski³

¹Martin-Luther-Universität Halle-Wittenberg, ²Humboldt-Universität zu Berlin, ³Institut zur Qualitätsentwicklung im Bildungswesen an der Humboldt-Universität zu Berlin

Der Unterricht in Neuzuwandererklassen in Deutschland richtet sich vor allem an neuzugewanderte Schüler:innen, deren Deutschkenntnisse nicht ausreichen, um dem Lehrplan in den Regelklassen zu folgen (Massumi et al. 2023). Obgleich Neuzuwandererklassen als zentrale bildungspolitische Maßnahme zur Unterstützung der sprachlichen Integration schulpflichtiger Geflüchteter gelten (Crul et al. 2019), sind sie daher in den Bundesländern und Kommunen unterschiedlich verbreitet (Will et al. 2022). Zudem existieren Neuzuwandererklassen in verschiedenen Organisationsformen (additiv, (teil-)integrativ, parallel), sodass unterschiedliche Grade der separierten Beschulung ausgemacht werden können (Massumi et al. 2015).

Sowohl in der Wissenschaft als auch in der Öffentlichkeit wird die separate Beschulung Neuzugewandelter kontrovers diskutiert. Neuzuwandererklassen werden einerseits als Schutzraum angesehen, der den Erwerb von Deutschkenntnissen erleichtert (Plöger & Barakos 2021). Andererseits werden sie dafür kritisiert, dass das separierte Lernen in Vorbereitungsklassen den Kontakt zu Peers aus nicht-eingewanderten Familien verringert (Karakayalı et al. 2017) und es gibt erste Hinweise darauf, dass die separierten Klassen den Erwerb von Deutschkenntnissen beeinträchtigen (Höckel & Schilling 2022). Die Ständige Wissenschaftliche Kommission der KMK rät von deren Einsatz in Grundschulen ab und empfiehlt sie lediglich für die Sekundarstufe (SWK 2022).

Der Beitrag geht der Frage nach, ob und ggf. wie stark sich die Deutschkompetenzen und die sozial-emotionale Integration von geflüchteten Jugendlichen, die beim Eintritt in das deutsche Schulsystem separat vs. in einer Regelklasse beschult wurden, mittelfristig unterscheiden. Konkret vergleichen wir die Deutschleistungen und die selbstberichtete soziale Eingebundenheit und Schulzufriedenheit von jungen Geflüchteten in der 9. Jahrgangsstufe, die zunächst im Grundschulalter in einer Vorbereitungsklasse gelernt haben, mit jenen junger Geflüchteter, die nie oder nur kurzzeitig eine separate Klasse besucht haben. Zudem stellen wir als Vergleichswert das von zuwanderten Jugendlichen ohne Fluchthintergrund erreichte Niveau der Deutschkompetenz bzw. der sozial-emotionalen Integration dar.

Theoretisch kann zum einen erwartet werden, dass Jugendliche aus geflüchteten Familien, die zunächst eine Klasse für Neuzugewanderte besuchten, auch einige Jahre später in der Sekundarstufe noch geringere Deutschkompetenzen aufweisen als ihre Peers, die nie Neuzuwandererklassen besuchten, weil der Rückstand in den Deutschkenntnissen noch nicht aufgeholt werden konnte. In diesem Fall hätten die Vorbereitungsklassen die sprachlichen Ungleichheiten nicht vollständig beseitigen können. Zum anderen kann aber auch angenommen werden, dass in den Neuzuwandererklassen ein effizientes Lernen ermöglicht wird, sodass allenfalls noch kleine Unterschiede in Deutsch bestehen bleiben. Analog könnte sich der Besuch einer separaten Klasse nachhaltig negativ auf die sozial-emotionale Integration geflüchteter Schüler:innen auswirken und auch in der Sekundarstufe noch zu Beeinträchtigungen führen. Oder aber es könnte sich – analog zu früheren Befunden, die einen Vorteil bei einer teiltintegrativen Beschulung in Regelklassen auf die Schulzugehörigkeit identifizierten (Pagel & Edele, 2022) – ein Vorteil für ein solches Beschulungsmodell auf die sozial-emotionale Integration zeigen.

Empirisch arbeiten wir mit Daten des IQB-Bildungstrends 2022 für Neuntklässler:innen, mit dem sowohl die Lesekompetenz der Jugendlichen als auch Hörverstehensleistungen und orthografische Kompetenzen untersucht wurden sowie die selbst berichtete soziale Eingebundenheit und Schulzufriedenheit. Die Schulleistungsstudie hat einen detaillierten Indikator zum Besuch von Neuzuwandererklassen bei geflüchteten Jugendlichen erhoben („Falls Ihr Kind nicht in Deutschland geboren wurde und bereits 6 Jahre alt oder älter war, als es nach Deutschland gekommen ist: Wie wurde Ihr Kind in den ersten 2 Jahren an einer deutschen Schule unterrichtet? Und für wie lange?“). Das Item differenziert zum einen das schulorganisatorische Modell (additiv, (teil-)integrativ, parallel) und zum anderen die jeweilige Beschulungsdauer im Vorbereitungsklassentypus. In den Analysen verwenden wir zur Untersuchung der Unterschiede in den Deutschkompetenzen und der sozial-emotionalen Integration zwischen den oben genannten Gruppen Propensity Score Matching und Mehrebenenregressionsmodelle unter Berücksichtigung wesentlicher Merkmale der Bildungsbiografie und des familiären Hintergrunds der Jugendlichen.

Abschließend diskutieren wir Limitationen des Erhebungs- und Analyseansatzes sowie die praktischen Implikationen der Befunde.

Belastungs- und Erfolgswahrnehmung von Lehrpersonen in Vorbereitungsklassen für Geflüchtete – Ein Vergleich von Lehrpersonen mit und ohne spezifische Weiterbildung für den Unterricht mit Geflüchteten

Elisabeth Maué, Franziska Spanner

Universität Konstanz

Im Zuge der Zuwanderung aus Krisenregionen kam in den vergangenen Jahren eine erhebliche Zahl schulpflichtiger Kinder und Jugendlicher nach Deutschland (Brücker et al., 2017), was entsprechende Maßnahmen im Bildungsbereich mit Blick auf deren Integration und insbesondere Sprachbildung erforderte. Für Schüler*innen, deren Deutschkenntnisse für den Besuch einer Regelklasse nicht ausreichen, wurden an allgemeinbildenden und beruflichen Schulen sogenannte Vorbereitungsklassen eingerichtet. Diese dienen vor allem der Vermittlung der deutschen Sprache und sollen den Übergang in reguläre Bildungsgänge oder eine Berufsausbildung ermöglichen (Baethge/Seeber, 2016; Emmerich et al., 2017). Durch den täglichen Kontakt mit ihren Schüler*innen gelten die in Vorbereitungsklassen unterrichtenden Lehrpersonen als Vertreter*innen der Aufnahmegesellschaft, die die Integration ihrer Schüler*innen entscheidend prägen (Blossfeld et al., 2016). Um jenen nicht nur sprachliche und berufliche, sondern auch persönliche Entwicklungsmöglichkeiten zu eröffnen, benötigen die Lehrpersonen entsprechende(s) Kompetenzen und Wissen, beispielsweise mit Blick auf die Entwicklung zielgruppenspezifischer Unterrichtsmaterialien, sprachsensiblen Unterricht (Becker-Mrotzek et al., 2021) oder den Umgang mit Traumata (Blossfeld et al., 2016; Weiß et al., 2021).

Die wenigen vorliegenden Untersuchungen zu wahrgenommenen Chancen und Herausforderungen von Lehrpersonen in Vorbereitungsklassen für geflüchtete Schüler*innen sind überwiegend qualitative Studien oder Fallstudien in einzelnen Bundesländern (z.B. Kärner et al., 2016; Karakayali et al., 2017; Reinke/Goller, 2022; Scheiermann/Walter, 2016; Stöbe-Blossey et al., 2019; Weiß et al., 2021).

Im vorliegenden Beitrag wird deshalb untersucht, 1.) über welches zielgruppenspezifische Fachwissen Lehrpersonen in Vorbereitungsklassen verfügen, 2.) welche Belastungen und Bereicherungen sie in Bezug auf den Unterricht in Vorbereitungsklassen wahrnehmen sowie 3.) welchen Beitrag sie ihres Erachtens zur Integration der unterrichteten Schüler*innen leisten können. Dabei soll explizit betrachtet werden, ob sich die Wahrnehmungen von Belastungen, Bereicherungen und des Beitrags zur Integration in Abhängigkeit des zielgruppenspezifischen Fachwissens unterscheiden.

Die Untersuchung war als mixed-methods-Design angelegt und wurde zwischen Mai und September 2017 an beruflichen Schulen in Baden-Württemberg durchgeführt. Der halbstandardisierte Fragebogen enthielt verschiedene Konstrukte (z.B. Belastungen, Bereicherungen, Beitrag zur Integration, Einstellungen zur Integration) mit geschlossenen Fragen (6-stufige Likertskala: 1=gar nicht belastend bis 6=sehr belastend bzw. 1=trifft gar nicht zu bis 6=trifft voll zu) sowie offene Fragen. Absolvierte Weiterbildungen wurden über eine Auswahl mit Mehrfachnennung erfasst: Deutsch als Fremdsprache (DaF), Deutsch als Zweitsprache (DaZ), Weiterbildung (Interkulturalität, Migration, Trauma etc.), Sprachkurs, anderes. Aufgrund des schwierigen Feldzugangs konnten lediglich N=48 Lehrpersonen befragt werden, was multivariate Analysen ausschließt.

Nur etwas mehr als die Hälfte der Lehrpersonen (n=28) verfügen aufgrund von Aus-/Weiterbildungen über zielgruppenspezifisches Wissen, auf 20 Personen trifft dies nicht zu. An 14 Maßnahmen wurde vor Beginn der Tätigkeit in der Vorbereitungsklasse teilgenommen; 20 Maßnahmen wurden erst während oder nach der erstmaligen Tätigkeit in einer Vorbereitungsklasse absolviert. Als mit Abstand größte Belastung empfinden die Lehrpersonen die Leistungsheterogenität (M=4.75, SD=1.14), welche von Lehrpersonen mit spezifischer Aus-/Weiterbildung stärker wahrgenommen wird (t(46)=-2.43, p=.02). Insgesamt fühlen sich die meisten Lehrpersonen dem Stress als Lehrperson einer Vorbereitungsklasse gewachsen (M=5.19, SD=0.17; Gruppenunterschied n.s.). Als negative Aspekte werden in der offenen Frage neben Merkmalen der Schüler*innen (z.B. mangelnde Disziplin, Pünktlichkeit) insbesondere schwierige Rahmenbedingungen (z.B. Vertragssituation, Fluktuation der Schüler*innen) hervorgehoben. Die Lehrpersonen verbinden jedoch auch viele positive Aspekte mit dem Unterricht in Vorbereitungsklassen, wie die Freiheit in Bezug auf Themen- und Methodenauswahl oder die Möglichkeit, Entwicklungen mit Blick auf Lernen und Integration beobachten zu können. Die höchsten Erfolgchancen für die geflüchteten Schüler*innen sehen die Lehrpersonen beim Spracherwerb (M=4.29, SD=1.07) und im Zurechtkommen im Alltag (M=4.13, SD=1.10; Gruppenunterschied n.s.).

Die Ergebnisse verweisen grundsätzlich auf die Bedeutung des Schulbesuchs für die Integration von zugewanderten und geflüchteten Kindern und Jugendlichen sowie auf die in diesem Kontext besondere Funktion der Lehrpersonen. Diese verfügen nur teilweise über zielgruppenspezifisches Fachwissen, was auf einen Bedarf an Aus- und Weiterbildung (Lipowski, 2011) verweist. Zudem werden Verbesserungsmöglichkeiten hinsichtlich der strukturellen Gestaltung von Vorbereitungsklassen deutlich.

Bildungsstrategien ukrainischer Vertriebener in Berlin und Warschau: Die Rolle der transnationalen Chancenstruktur

Céline Teney

Institut für Soziologie, Freie Universität Berlin

Ich untersuche die Bildungsstrategien vertriebener ukrainischer Eltern für ihre schulpflichtigen Kinder, die nach Polen und Deutschland geflohen sind. Zu diesem Zweck führe ich eine Abduktionsanalyse eines einzigartigen qualitativen Paneldatensatzes durch, der aus 88 halbstrukturierten Interviews mit vertriebenen Ukrainer*innen mit schulpflichtigen Kindern in Berlin und Warschau aus dem Sommer 2022 sowie 53 Folgeinterviews aus dem Frühjahr 2023 besteht. In meiner Analyse identifiziere ich drei unterschiedliche Bildungsstrategien: (1) Online-Einschreibung in das ukrainische Schulsystem aus dem Zielland heraus; (2) ausschließliche Integration in das Bildungssystem des Ziellandes; (3) Rückkehr in die Ukraine und Fortsetzung des ukrainischen Schulcurriculums vor Ort. Das Ausmaß, in dem die Befragten eine Strategie vor der anderen bevorzugen, variiert zwischen Warschau und Berlin und zwischen unseren beiden Interviewwellen. Meine Analyse unterstreicht die Bedeutung der "transnationalen Gelegenheitsstruktur" für ein besseres Verständnis der Handlungsfähigkeit und der Handlungsstrategien von Migrant*innen. Ich konzeptualisiere dies als die Gesamtheit der Möglichkeiten und Beschränkungen, die sich aus den Kontexten ergeben, in denen sich Migrant*innen bewegen.

PS-62: Inklusion in der Schule aus Perspektive von Lehrenden & Lernenden

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 129

Chair der Sitzung: David Scheer

Prädiktoren für das schüler*innenspezifische Selbstwirksamkeitserleben von allgemeinen und sonderpädagogischen Lehrkräften

David Scheer¹, Chiara Enderle², Anett Platte², Tatjana Leidig³, Conny Melzer²

¹Pädagogische Hochschule Ludwigsburg, Deutschland; ²Universität Leipzig; ³Universität zu Köln

Bisherige Forschungsvorhaben zum Selbstwirksamkeitserleben von Lehrkräften (SWE-L) im Kontext von Inklusion nehmen eher interindividuelle Unterschiede zwischen Lehrkräften in den Blick. Gerade mit Blick auf unterrichts- und schüler*innen bezogene Aspekte von SWE-L erscheint es jedoch sinnvoll, auch die intraindividuelle Variabilität mit Blick auf Merkmale spezifischer Schüler*innen in den Blick zu nehmen. Hierfür wurde in den Niederlanden auf Basis der Teacher's Sense of Efficacy Scale (Tschannen-Moran & Woolfolk Hoy, 2001) eine schülerspezifische Skala (SS-TSES) entwickelt (Zee et al., 2016). Dieses wurde von Leidig et al. (2021) auf Deutsch übersetzt (SS-TSES-D) und psychometrisch evaluiert (Przibilla et al., in revision; Leidig et al., in Vorb.). Die Struktur des Instruments mit den vier Subskalen Instructional Support, Emotional Support, Student Engagement und Behavior Management konnte im Vergleich mit dem niederländischen Original mittels Multilevel CFA bestätigt werden.

Fragestellung & Methodik

Der Beitrag fokussiert die Fragestellung, inwieweit sowohl Merkmale von Schüler*innen als auch der Lehrkräfte einen Einfluss auf das SWE-L haben.

In einem Querschnittsdesign wurden Schulen in drei Bundesländern postalisch zur Teilnahme aufgerufen. Jede teilnehmende Lehrkraft erhielt einen Satz von acht Fragebögen, beinhaltend den SS-TSES-D sowie den Strength and Difficulties Questionnaire (SDQ; Goodman, 1997). Mittels einer Onlinegenerators wurden den Lehrkräften acht zufällige Nummern zugewiesen, die als Grundlage zur Auswahl der Schüler*innen aus der Klassenliste dienten. Insgesamt haben 90 Lehrkräfte teilgenommen. Dies ergibt eine auswertbare Stichprobe von $N = 717$ Schüler*innen.

In hierarchisch linearen Modellen (HLM) wurden schrittweise Schüler*innenmerkmale als Level-1-Prädiktoren und Lehrkraftmerkmale als Level-2-Prädiktoren aufgenommen: Die Ergebnisse des SDQ wurden als Level-1-Prädiktoren am Clustermittelwert zentriert, als Level-2-Prädiktoren wurden die Clustermittelwerte am Gesamtmittelwert zentriert.

Ergebnisse

In den deskriptiven Statistiken erscheint es zunächst so, dass die sonderpädagogischen Lehrkräfte – entgegen der Forschungslage zu nicht-schüler*innenspezifischer SWE-L im Kontext von Inklusion (Wray et al., 2022) – ihre SWE-L-Werte niedriger ein als die allgemeinen Lehrkräfte. Dieser Befund relativiert sich jedoch bzw. neutralisiert sich bei einer differenzierten Betrachtung der HLM:

Als bestes Modell erwies sich das Fixed-Slopes-Modell mit allen Level-1-Prädiktoren sowie den aggregierten SDQ-Scores als Level-2-Prädiktoren. Eine Hinzunahme der Lehrkraftmerkmale auf Level 2 führte zu weder zu einer signifikant besseren Varianzaufklärung, noch zeigte sich irgendeins der Lehrkraftmerkmale als signifikanter Prädiktor für die schüler*innenspezifische SWE-L.

Über alle SS-TSES-D-Subskalen hinweg zeigten sich die Lehrkräfteeinschätzungen externalisierender, internalisierender und prosozialer Verhaltensweisen der Schüler*innen als stärkste Prädiktoren für die schüler*innenspezifische SWE-L, ohne Unterschiede zwischen sonderpädagogischen und allgemeinen Lehrkräften. Unter Berücksichtigung des tatsächlich wahrgenommenen Schüler*innenverhaltens konnte zudem kein relevanter Einfluss des diagnostizierten sonderpädagogischen Unterstützungsbedarfs festgestellt werden.

Im Vortrag werden die differenzierten Ergebnisse vorgestellt und diskutiert. Implikationen auf die Lehrkräftebildung zur Inklusion werden reflektiert. Zudem werden Anschlussfragen und sich aus den Befunden ergebende Forschungsdesiderate skizziert.

Inklusionsbezogene Einstellungen und Selbstwirksamkeitserwartungen von Lehrpersonen in der Schweiz und die Bedeutung der Schulleitung

Isabella Lussi¹, Julia Schaub¹, Stephan Huber^{1,2}

¹Pädagogische Hochschule Schwyz, Schweiz; ²Johannes Kepler Universität Linz, Österreich

Ausgangslage und theoretischer Hintergrund

Mit der Ratifizierung der UN-BRK im Jahr 2014 verpflichtete sich die Schweiz ein auf allen Ebenen integratives Bildungssystem zu etablieren. Zehn Jahre später wird das Thema der integrativen Schule in der schulischen Praxis und auch im öffentlichen Diskurs in der Schweiz breit diskutiert. In verschiedenen Schweizer Kantonen sind politische Vorstösse zu verzeichnen, die wieder vermehrt separative (i.e. segregierte) Angebote wie z.B. die Wiedereinführung von Kleinklassen fordern. Kantonale Evaluationen zum Thema zeigen auf, dass Schulischen Heilpädagoginnen und Therapeutinnen sowie Schulleitungen gegenüber der integrativen Schule mehrheitlich eher positiv eingestellt sind und sich der Aufgabe auch gewachsen fühlen. Lehrpersonen nehmen jedoch eher eine neutrale Haltung ein und geben häufiger an, mit der integrativen Aufgabe überfordert zu sein (AVS Graubünden, 2022; Mettauer et al. 2021; Moser Opitz & Pool Maag, 2019).

Positive inklusionsbezogene Einstellungen und Selbstwirksamkeitserwartungen von Lehrpersonen gelten als zentrale Voraussetzung für die Gestaltung inklusiver respektive integrativer Schulen (Greiner et al. 2020, McElvany et al., 2018, Bosse et al., 2017, Hösel et al. 2020). Einstellungen und Wirksamkeitserwartungen von Lehrpersonen werden neben Professionswissen und Selbstregulationsfähigkeiten als wichtige Aspekte ihrer professionellen Kompetenz im Umgang mit heterogenen Lerngruppen angesehen (McElvany et al., 2018; Hellmich und Görel, 2014, Hösel et al. 2020). Die Ausprägung inklusionsbezogener Einstellungen und Selbstwirksamkeitsüberzeugungen wird mit unterschiedlichen Faktoren erklärt (Greiner et al. 2020, S: 274). Empirisch gut belegt ist die Bedeutung der Profession (McElvany et al., 2018, Hellmich 2018) und Erfahrungen im Umgang mit heterogenen Lerngruppen (Hellmich und Görel, 2014, Hellmich, 2018). Bisher wenig in den Blick genommen wird die Bedeutung der Schulleitung hinsichtlich der Ausprägung inklusionsbezogener Einstellungen und Selbstwirksamkeitserwartungen. Dies obschon bekannt ist, dass Schulleitungen im Hinblick auf die Gestaltung inklusiver Schulen eine Schlüsselposition einnehmen (Sturm et al., 2015; vgl. Amrhein, 2014; Anderegg, 2019; Criblez et al., 2012; Hattie et al., 2013; Heimlich, 2018; Peperkorn et al., 2021; Schulz & Böttinger, 2022; Wirtz, 2020).

Forschungsfrage und Forschungsdesign

Die vorliegende Studie geht der Frage nach, welche Bedeutung der Schulleitung hinsichtlich der Ausprägung inklusionsbezogener Einstellungen und Selbstwirksamkeitserwartungen von Lehrpersonen und Fachpersonen der integrativen Förderung zukommt. Im Fokus stehen dabei sowohl Führungsmerkmale der Schulleitung als auch deren Einstellungen und Selbstwirksamkeitserwartungen. Zudem wird auch die Einschätzung der kontextuellen Rahmenbedingungen in die Analysen miteinbezogen, da Einstellungen im Unterschied zu Werthaltungen immer auch kontextgebunden sind. Die Studie basiert auf einer Online-Befragung von 219 Schulleitenden, 917 Regellehrpersonen und 268 Fachpersonen der integrativen Förderung aus verschiedenen Kantonen der Deutschschweiz aus dem Schuljahr 2022/23. Die Daten werden mehrbenenanalytisch ausgewertet und die theoretisch hergeleiteten Hypothesen mittel Strukturgleichungsmodellen überprüft.

Bedeutsame Ergebnisse

Die Analysen zeigen auf, dass die befragten schulischen Akteure gegenüber der integrativen Schule insgesamt neutral bis eher positiv eingestellt sind und sich die Umsetzung des integrativen Unterrichts mehrheitlich auch zutrauen. Es gibt diesbezüglich jedoch signifikante Unterschiede zwischen den Professionen und auch nach Schulstufen. Zudem zeigt sich, dass inklusionsbezogene Einstellungen von Lehrpersonen signifikant mit deren Selbstwirksamkeitserwartungen und teilweise mit Merkmalen der Schulleitung zusammenhängen. Schulleitende, die sich eher als Verwalter und weniger als Gestalter wahrnehmen und die Ressourcen zur Umsetzung der integrativen Schule als besonders gering erachten, sind der integrativen Schule gegenüber zudem signifikant kritischer eingestellt.

Die Ergebnisse bestätigen bisherige Befunde, dass Einstellungen gegenüber der integrativen Schule davon abhängen, inwiefern sich Lehrpersonen den Umgang mit ihren Anforderungen zutrauen. Dies verweist auf die Bedeutung von Professionalisierungsmassnahmen im Kontext integrativer Schulen. Darüber macht die Studie deutlich, dass sich Schulleitungen in ihren grundlegenden Haltungen bezüglich ihrer Rolle in Schulentwicklungsprozessen unterscheiden und sich dies auch in Einstellungen gegenüber der integrativen Schule widerspiegelt. Vor diesem Hintergrund ist es bedeutsam, dass auch die Professionalisierung der Schulleitungsfunktion vorangetrieben wird, damit sie die Gestaltung der integrativen Schule auch im Kontext suboptimaler Rahmenbedingungen aktiv angehen kann.

Diagnostisches Handeln im inklusiven Unterricht – ein schulform- und professionsübergreifender Blick auf diagnostische Zugänge Lehrender

Theresa Overbeck, Silvia Greiten

Pädagogische Hochschule Heidelberg, Deutschland

Schulische Lerngruppen sind auf Grund gesellschaftlicher Entwicklungen, wie der Umsetzung der UN-Behindertenkonvention oder weltweiter Migrationsbewegungen, zunehmend vielfältig zusammengesetzt. Die Heterogenität Lernender in Schule und Unterricht zu berücksichtigen gilt als eine der größten aktuellen Herausforderungen des deutschen Bildungssystems (Gräsel, Decristan & König, 2017). Im Umgang mit der Vielfalt wird adaptiver Unterricht als erfolgversprechend erachtet, der sich durch eine möglichst hohe Passung zwischen Lernangeboten und individuellen Voraussetzungen der Schüler:innen auszeichnet (Brühwiler, 2014). Um eine solche Passung herzustellen, benötigen Lehrkräfte Kenntnisse über Lernausgangslagen, Lernstände und Lernmöglichkeiten der Schüler:innen (Fischbach, Mähler & Hasselhorn, 2021). Die Diagnostik von lernrelevanten Merkmalen spielt damit eine zentrale Rolle bei der adaptiven Unterrichtsgestaltung (Beck et al., 2008).

Mit Blick auf Diagnostik im inklusiven Unterricht gelten Formen pädagogischer Diagnostik als besonders bedeutsam (Simon & Geiling, 2016), die formativ ausgerichtet sind und Entscheidungen zur Gestaltung von adaptiven Unterrichts- und Förderangeboten unterstützen (Souvignier, 2022). Das diagnostische Handeln von Lehrkräften kennzeichnend rücken damit u.a. diagnostische Methoden in den Fokus, mit Hilfe derer innerhalb sowie außerhalb des Unterrichts Erkenntnisse über Lernende gewonnen werden können (Greiten, Trumpa & Veber, 2023). Dabei ist in Bezug auf die Schulpraxis weitestgehend ungeklärt, welche diagnostischen Methoden Lehrkräfte unterschiedlicher Professionen und Schulformen im Rahmen von Situationen innerhalb oder außerhalb des inklusiven Unterrichts anwenden und ob diese inklusionspädagogischen Prinzipien entsprechen (Fischbach et al., 2021).

Vor diesem Hintergrund steht (1) die Frage nach den diagnostischen Zugängen, die Lehrkräfte beschreiben, um im inklusiven Unterricht Erkenntnisse über Lernende zu gewinnen und (2) die Einordnung der genutzten diagnostischen Zugänge im Rahmen einer pädagogischen, formativ ausgerichteten Diagnostik nach Zielsetzung und Zeitpunkt der Erkenntnisgewinnung im Mittelpunkt dieses Beitrags (Overbeck & Skribbe, 2024, i.V.). Empirische Grundlage des Beitrags bildet eine im BMBF-geförderten Verbundprojekt „DiaGU“ (Greiten et al., 2023) durchgeführte Interviewstudie, in welcher vier Sonderpädagog:innen und fünf Fachlehrkräfte zu diagnostischen Zugängen im inklusiven Unterricht leitfadengestützt befragt wurden. Die Interviewauswertung erfolgte mit der inhaltlich strukturierenden Qualitativen Inhaltsanalyse (Kuckartz & Rädicker, 2022). Das Kategoriensystem wurde dabei deduktiv entwickelt, induktiv ausdifferenziert und durch konsensuelles Validieren (Gläser-Zikuda, 2013) mit einer Inter-coder-Übereinstimmung zwischen 71% und 85% im Projektteam abgesichert.

Die Befunde der Studie weisen darauf hin, dass Lehrkräfte unterschiedlicher Schulformen und Professionen häufig individualisierenden Unterricht sowie implizit auch lehrkraftgelenkte Phasen zur Diagnostik nutzen. Darüber hinaus erachten die Befragten schulform- und professionsübergreifend informelle Methoden und die Schüler:innen-Lehrkraft-Beziehung als diagnostisch bedeutsam. Bei der Beschreibung diagnostischer Zugänge durch die Interviewten sind formative Aspekte dominant. Insgesamt liefert die vorliegende Studie damit relevante Befunde für die inhaltliche Ausgestaltung einer Lehrer:innenaus- und -fortbildung mit dem Ziel der Bewältigung gegenwärtiger und künftiger inklusionsbedingter gesellschaftlicher Herausforderungen.

Bildungsaspirationen von Schüler:innen mit besonderen Unterstützungsbedarfen: Eine Frage des familiären Hintergrunds und der Noten?

Sebastian Bauer¹, Cornelia Gresch², Ines Reißenweber¹

¹Leibniz-Institut für Bildungsverläufe; ²IQB

Theoretischer Hintergrund Schüler:innen mit besonderen Unterstützungsbedarfen – wie einem sonderpädagogischem Förderbedarf (SPF) oder einer Teilleistungsschwäche (TLS) - erreichen häufig niedrigere Bildungsabschlüsse, was ihren Berufseinstieg erschwert (Blanck, 2020; Shifrer, 2013). Die anvisierten Bildungsabschlüsse hängen unter anderem von den Bildungsaspirationen der Schüler:innen ab (Stocké, 2013). Diese sind nach dem Wisconsin-Modell durch Einflüsse des familiären Hintergrunds und der schulischen Leistung geprägt (Sewell et al., 1957). Da Schüler:innen mit besonderen Unterstützungsbedarfen häufig aus Familien mit niedrigerem sozioökonomischen Status oder Zuwanderungshintergrund

kommen und schwächere schulische Leistungen zeigen (Lenkeit et al., 2022), ist dieser Zusammenhang auch für sie relevant (Lee et al., 2015). In Studien zu Bildungsaspirationen von Schüler:innen mit besonderen Unterstützungsbedarfen wird häufig auf Gottfredsons (1981) Eingrenzungs- und Kompromisstheorie Bezug genommen (z.B. Casey et al., 2006; Rojewski, 1999). Diese besagt, dass Selbstkonzept und wahrgenommene Hindernisse Jugendliche dazu zwingen, Kompromisse in ihren Bestrebungen einzugehen, wenn der Aufwand zu hoch oder das Risiko des Scheiterns zu groß erscheint. Jugendliche mit besonderen Unterstützungsbedarfen könnten daher Selbstselektionsprozessen unterliegen, die zu geringeren Bildungs- und Berufsaspirationen führen, etwa durch Etikettierungen mit einem Label wie „SPF“ oder „TLS“, die zu Identitätsverletzungen führen können (Blanck, 2020; Shifrer, 2013). Auch besondere Formalitäten bei der Benotung auf Grundlage individuell festgelegter Lernziele oder zieldifferenten Unterricht könnten solche Anpassungen der Bildungsziele fördern. In der Regel werden Schüler:innen mit einem SPF im Bereich Lernen oder geistige Entwicklung zieldifferent unterrichtet, Schüler:innen mit anderen Förderbedarfen hingegen zielgleich (KMK, 2017). In Deutschland fehlen jedoch Studien zu den Bildungsaspirationen von Schüler:innen mit unterschiedlichen Unterstützungsbedarfen. Fragestellung Dieser Beitrag untersucht die Bildungsaspirationen von Schüler:innen mit SPF (zielgleich oder zieldifferent) oder TLS mit Blick auf ihre angestrebten Schulabschlüsse am Ende der Sekundarstufe I sowie den Einfluss vom familiären Hintergrund und den Schulnoten auf den gewünschten Schulabschluss. Methode Die Daten stammen aus der INSIDE-Längsschnittstudie und wurden in der siebten Jahrgangsstufe an allgemeinen Schulen in Deutschland erhoben. Als Datenbasis dieser Analysen dienen alle Schüler:innen, die kein Gymnasium besuchen, mit vollständigen Angaben zum gewünschten Schulabschluss, zum besonderen Unterstützungsbedarf und zu den Schulnoten in Deutsch und Mathematik ($N = 1.092$, $n = 77$ Schüler:innen mit SPF, davon 55% zieldifferent, und $n = 89$ mit TLS). Die Schulnoten basieren auf den Endjahresnoten der sechsten Klasse und wurden über die Lehrkräfte erfasst. Zur Erfassung des familiären Hintergrunds wurde der höchste Wert des „International Socio-Economic Index of Occupational Status“ (ISEI) (Ganzeboom et al., 1992) der Eltern genutzt (HISEI) und der Zuwanderungshintergrund der Schüler:innen. Fehlende Informationen zum familiären Hintergrund wurden imputiert (Royston, 2004). Für die Analyse wurden multinomiale logistische Regressionen gerechnet. Ergebnisse und Diskussion Erste Ergebnisse zeigen, dass Schüler:innen mit SPF, die zieldifferent ($OR = 7.90$; $p < .001$) oder zielgleich ($OR = 2.92$; $p = .04$) unterrichtet wurden, signifikant häufiger einen Hauptschulabschluss anstrebten als einen Realschulabschluss. Das Vorliegen von TLS hatte keinen signifikanten Effekt. Bezüglich des Abiturwunsches im Vergleich zum Realschulabschluss gab es keine Unterschiede. Alle Effekte blieben auch nach Kontrolle des familiären Hintergrunds und der Schulnoten stabil. Die Ergebnisse legen nahe, dass Schüler:innen mit SPF geringere Bildungsaspirationen hatten als Schüler:innen ohne besondere Unterstützungsbedarfe, wobei zielgleich und zieldifferent unterrichtete Schüler:innen vergleichbare Bildungsziele hatten. Schüler:innen mit TLS hatten hingegen ähnliche Bildungsaspirationen wie Schüler:innen ohne besondere Unterstützungsbedarfe. Für höhere Bildungsziele waren bei allen Gruppen nicht der Status, d.h. die jeweiligen Unterstützungsbedarfe, sondern der familiäre Hintergrund und die Schulnoten ausschlaggebend. Die Stabilität der Effekte zeigt: Der Status wirkt unabhängig vom familiären Hintergrund und den Schulnoten auf die Herausbildung der Bildungsaspirationen. Die Untersuchung zeigt, dass die Zuweisung von Unterstützungsbedarfen eine Rolle spielt und leistet damit einen Beitrag zur Debatte um mögliche negative Folgen damit verbundener Etikettierungen.

PS-63: Motivation in der Hochschulbildung 2

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 133

Chair der Sitzung: Florian Klapproth

Welche Rolle spielen implizite Intelligenztheorien für Bildungsaspirationen von Studierenden?

Florian Klapproth, Adriana Zörner

Medical School Berlin

Theoretischer Hintergrund Nach Rational-Choice-Theorien (RCT; z. B. Breen & Goldthorpe, 1997) werden Bildungsaspirationen durch das bewusste Abwägen der Bedeutsamkeit des erwarteten Nutzens, der voraussichtlichen Kosten und der antizipierten Erfolgswahrscheinlichkeit in Bezug auf eine Bildungsoption determiniert. Schülerinnen und Schüler oder Studierende wählen danach diejenige Bildungsoption (z. B. einen bestimmten Bildungsabschluss), die mit einer hohen Erfolgswahrscheinlichkeit, geringen Kosten und einem hohen Nutzen verbunden ist (Lewin et al., 1944). Dabei variieren diese drei Parameter zwischen Individuen unterschiedlicher sozialer Klassen (Breen & Goldthorpe, 1997). Beispielsweise werden die Kosten für einen Studienabschluss von weniger privilegierten Individuen als höher wahrgenommen als von stärker privilegierten Individuen (Zimmermann, 2020). Die Determination von Bildungsaspirationen durch die von der RCT postulierten Parameter kann daher auch als Herkunftseffekt aufgefasst werden (Stocké, 2022). Dieser eher ökonomischen Sichtweise auf die Entwicklung von Bildungsaspirationen stehen psychologische Theorien gegenüber, die Bildungsaspirationen auch als Ergebnis unbewusster Prozesse der Selbstregulation auffassen (z. B. Bandura, 1997). Selbstregulative Prozesse können durch implizite Theorien über das Selbst, insbesondere über die eigene Intelligenz, beeinflusst werden (Dweck, 1999). So zeigen Individuen, die ihre Intelligenz als grundsätzlich veränderbar betrachten („growth mindset“), ein höheres Engagement in der Erreichung ihrer Bildungsziele als Individuen, die Intelligenz als stabil betrachten („fixed mindset“) (Martin et al., 2013). Fragestellung Ob implizite Intelligenztheorien einen Effekt auf Bildungsaspirationen ausüben, wurde nach Wissen der Autor:innen bislang nicht untersucht. Es wurde die Hypothese geprüft, dass die in der RCT postulierten Faktoren die Bildungsaspiration von Studierenden (Bachelor- versus Master-Abschluss) vorhersagen. Konkret wurde erwartet, dass die Erfolgserwartung und die Bedeutsamkeit des erwarteten Nutzens positiv, die erwarteten Kosten dagegen negativ mit der Bildungsaspiration zusammenhängen. Ferner wurde geprüft, ob implizite Intelligenztheorien der Studierenden einen Effekt auf die Bildungsaspiration ausüben. Methode Die Stichprobe bestand aus N = 180 Studierenden (MAlter = 22.6 Jahre, SD = 2.6; 57.9 % weiblich) unterschiedlicher Studiengänge an staatlichen (80.3 %) und privaten (19.7 %) Hochschulen. Abhängige Variablen waren die idealistische und die realistische Bildungsaspiration (gewünschter bzw. erwarteter Universitätsabschluss: Bachelor oder Master). Prädiktoren waren die drei RCT-Parameter Erfolgserwartung, Bedeutsamkeit des erwarteten Nutzens und die erwarteten Kosten (jeweils auf einer Skala von 1 bis 5) sowie das Mindset auf einer Skala von 1 (eher „fixed“) bis 5 (eher „growth“), $\alpha = .80$. Zusätzlich erhoben wurden das Geschlecht der Studierenden, ihr Abiturnotendurchschnitt sowie der familiäre Bildungsstatus (operationalisiert über den höchsten Bildungsabschluss in der Familie). Geschätzt wurden unterschiedliche Modelle der multiplen logistischen Regression. Vor der Analyse wurden die Prädiktoren z-standardisiert. Berichtet wird aus Platzgründen nur das vollständige Modell mit 7 Prädiktoren. Ergebnisse und Diskussion Die logistische Regressionsanalyse für die abhängige Variable „idealistische Bildungsaspiration“ ergab einen signifikanten Effekt der Erfolgswahrscheinlichkeit, $B = 0.99$ ($SE = 0.23$), $\text{Exp}(B) = 2.68$, $p < .001$, und einen signifikanten Effekt des Mindsets, $B = 0.49$ ($SE = 0.20$), $\text{Exp}(B) = 1.64$, $p = .015$. Alle anderen Effekte waren nicht signifikant. Für die abhängige Variable „realistische Bildungsaspiration“ zeigte sich ein signifikanter Effekt der Erfolgswahrscheinlichkeit, $B = 1.09$ ($SE = 0.23$), $\text{Exp}(B) = 2.96$, $p < .001$, sowie eine signifikante Kosten x Mindset-Interaktion, $B = 0.43$ ($SE = 0.22$), $\text{Exp}(B) = 1.53$, $p = .049$. Alle anderen Effekte waren nicht signifikant. Die Annahme eines Effekts der RCT-Parameter auf die Bildungsaspiration konnte nur für die erwartete Erfolgswahrscheinlichkeit bestätigt werden. Darüber hinaus konnte gezeigt werden, dass das Mindset einen Effekt auf die idealistische sowie einen moderierenden Effekt auf die realistische Bildungsaspiration ausübt. Die Wechselwirkung zwischen Mindset und erwarteten Kosten bedeutet inhaltlich, dass eine hohe Bildungsaspiration am wahrscheinlichsten ist, wenn Studierende ein Growth-Mindset haben und geringe subjektive Kosten antizipieren.

Eine Analyse von reziproken Effekten zwischen Leistung und dem motivationalen Aspekt des Fachinteresses im Dualen Studium - Ergebnisse eines Random Intercept Cross-Lagged Panel Models (RI-CLPM)

Steffen Wild

Technische Universität Dortmund, Deutschland

Theoretischer Hintergrund

Eine immer noch ungelöste und diskutierte Fragestellung ist, ob Leistung den motivationalen Aspekt des Interesses beeinflusst oder ob die Zusammenhangsrichtung von Interesse auf Leistung verläuft bzw. stärker ist (Scherrer et al., 2020; Vu et al., 2024). Überwiegend wurde diese Fragestellung in der Schulforschung mit keinem eindeutigen und abschließenden Ergebnis aufgearbeitet (Schiefele, 2009; van Maurice et al., 2014). Gründe für die inkonsistenten Resultate werden u.a. in den zeitlich variierenden Messabständen oder bei der unterschiedlichen Leistungsmessungen vermutet (Benden & Lauermann, 2023; Dormann & Griffin, 2015). Folglich drängt sich die Fragestellung von reziproken Effekten zwischen Leistungen und Interesse auch in der Hochschulforschung auf und muss insbesondere vor dem Hintergrund, weil im Hochschulstudium intrinsische Werte zurückgehen und in Zusammenhang mit Bildungsentscheidungen, wie Studienabbruch oder ein Masterstudium aufzunehmen, stehen, aufgearbeitet werden (Benden & Lauermann, 2022; Harackiewicz et al., 2002; Schnettler et al., 2020).

Ausgehend von dem Reciprocal Internal/External Frame of Reference (I/E) Modell (Marsh, 1986; Möller et al., 2011), wo die Effekte vorhergehender Selbstkonzepte auf nachfolgende Leistungen sowie vice versa wirken, wird postuliert dass von diesen ausgelösten dimensionalsten Vergleichen auch der Bereich der Motivation beeinflusst wird. Empirische Effekte belegen in Längsschnittstudien sowohl den ausschließlichen unidimensionalen Zusammenhang von Interesse auf die nachstehende Leistung (Liu & Hou, 2018; Messerer, 2023), als auch den ausschließlichen unidimensionalen Zusammenhang von Leistung auf Interesse (Garon-Carrier et al., 2016; van Maurice et al., 2014; Viljaranta et al., 2014). Weitere Studien geben Hinweise auf eine reziproke Beziehung zwischen Interesse und Leistung, wie etwa bei Höft und Bernholt (2019), Hebbeker et al. (2019) oder Köller et al. (2001).

Fragestellung

Die vorangehenden Ausführungen zeigen, dass das Zusammenspiel von Interesse und Leistung komplex ist. Folglich könnte die Zeitspanne zwischen den Messzeitpunkten, die Fachdisziplin, die Bildungsinstitution oder die Messung der Leistung, einen Einfluss auf das Resultat haben. In diesem Beitrage prüfen wir diese Fragestellung zu reziproken Effekten bei Dual Studierenden

(Graf et al., 2014), deren Zahl sich in den letzten zehn Jahren (zwischen 2012 und 2022) auf 120.000 Studenten in Deutschland fast verdoppelt hat (Hofmann, 2023). Im Einzelnen werden die Hypothesen getestet:

H_1 : Wenn die Dual Studierenden höhere Werte im Fachinteresse besitzen, dann steigen die späteren Leistungen.

H_2 : Wenn Dual Studierende ein höheres Leistungsniveau erreichen, dann entsteht ein zeitlich nachgelagertes höheres Fachinteresse.

Methode

Daten von Dual Studierenden der Startkohorte des Jahres 2016 von der Panelstudie „Studienverlauf – Weichenstellungen, Erfolgskriterien und Hürden im Verlauf des Studiums an der DHBW“ (Deuer & Meyer, 2020) wurden basierend auf jährlichen Messungen und bei mindestens zweimaliger Teilnahme verwendet. Das Fachinteresse wurde anhand fünf Items aus der Skala von Fellenberg und Hannover (2006) mit akzeptablen Reliabilitäten in den drei Messzeitpunkten ($\omega = .85-.89$) gemessen. Die Leistungsdaten (Grade Point Average) des Messzeitpunkts wurden von der Hochschulverwaltung zur Verfügung gestellt. Die Daten wurden anhand eines Random Intercept Cross-Lagged Panel Model (RI-CLPM) ausgewertet (Hamaker et al., 2015; Orth et al., 2021).

Ergebnisse und ihre Bedeutung

Das geschätzte RI-CLPM weist eine annehmbare Güte auf ($\chi^2 = 1014.164$; $df = 156$; $\chi^2 / df = 6.501$; $p \leq .001$; $CFI = .956$; $TLI = .952$; $RMSEA = .056$; $SRMR = .038$). Die Within-Komponente zeigt, dass Leistung im ersten Messzeitpunkt das spätere Fachinteresse im zweiten Messzeitpunkt positiv ($\beta = .060$; $p = .022$) beeinflusst und das Fachinteresse im zweiten Messzeitpunkt beeinflusst die spätere Leistung im dritten Messzeitpunkt ($\beta = .537$; $p = .031$) positiv. Entgegen den Erwartungen zeigt sich ein negativer Einfluss des Fachinteresses im Messzeitpunkt eins auf die Leistung im Messzeitpunkt zwei ($\beta = -.261$; $p = .001$). Offenbar scheinen reziproke Effekte auf einzelne Studienabschnitte begrenzt zu sein. Des Weiteren könnten leistungsstarke Studierende zu Studienbeginn unterfordert zu sein, was möglicherweise mit der Auswahl der Dual Studierenden zusammenhängen könnte.

Individualisierte Interventionen - Motivationsförderung nicht nur im Durchschnitt, sondern für alle?

Liene Brandhuber¹, Lisa Bäumke², Daniel Wiecek², Thorben Jansen³, Samuel Merk¹, Benjamin Nagengast², Cora Parrisus¹

¹Pädagogische Hochschule Karlsruhe; ²Universität Tübingen; ³IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Hintergrund und Fragestellung

Motivationsinterventionen zeigen im Durchschnitt mittlere Effekte (Lazowski & Hulleman, 2016). Allerdings deutet die empirische Befundlage auch darauf hin, dass Motivationsinterventionen zum Teil nicht für alle Lernenden gleichermaßen wirksam sind. Eine mögliche Erklärung für diese Unterschiede lässt sich aus dem *Aptitude-Treatment-Interactions* Ansatz ableiten (Cronbach, 1957). Dieser beruht auf der Annahme, dass ein Treatment dann am effektivsten ist, wenn eine optimale Passung zwischen dem Treatment und den Personenmerkmalen vorliegt (Cronbach & Snow, 1977; Snow, 1989). Demnach sollten Personen, die sich z.B. nicht sozial zugehörig fühlen, stärker von einer *Social-Belonging* Intervention profitieren als Personen, die bereits gut eingebunden sind (Walton & Brady, 2021). Zeitgleich ist davon auszugehen, dass Studierende, die bspw. bereits eine hohe Nützlichkeit in einem Lerngegenstand sehen, weniger von einer Intervention zur Förderung der *Nützlichkeitsüberzeugung* (vgl. Priniski et al., 2018) profitieren. Um den größtmöglichen Erfolg einer Maßnahme zur Motivationsförderung zu erreichen, erscheint es also vielversprechend, Lernenden individuell genau die Intervention darzubieten, die auf ihre eigenen (motivationalen) Bedürfnissen zugeschnitten ist. Anliegen dieser Studie ist es daher, die Wirksamkeit einer individualisierten (d.h. bedarfsgerechten) Zuweisung von verschiedenen Motivationsinterventionen im Vergleich zu einer zufälligen Zuweisung sowohl hinsichtlich proximaler als auch globaler Outcomes zu testen. Hierfür wurden fünf Interventionen eingesetzt, die sich von unterschiedlichen Motivationstheorien ableiteten (Lernzielorientierung, Elliot & McGregor, 2001; Growth-Mindset, Dweck, 2008; Nützlichkeitsüberzeugungen, Eccles & Wigfield, 2020; Social Belonging, Walton & Cohen, 2007; Attributionstheorie, Weiner, 2012).

Methode

Im Rahmen einer Methodenvorlesung für Lehramtsstudierende wurden in einem Prä-Posttest-Design die verschiedenen Motivationsinterventionen implementiert. Mittels Prätest wurde ermittelt, für welche der fünf Motivationsinterventionen ein individuell erhöhter Bedarf bestand. Zur Bestimmung des Bedarfs wurden die z-standardisierten Mittelwerte miteinander verglichen und der intraindividuell niedrigste Wert als Bedarf definiert. Innerhalb der Gruppe der Studierenden desselben Bedarfs wurde jeweils der Hälfte zufällig die bedarfsgerechte Intervention zugewiesen (individualisierte Interventionsbedingung) und der anderen Hälfte eine der vier anderen Interventionen (Kontrollbedingung). Das heißt, dass bspw. die Hälfte derjenigen Studierenden, die besonders niedrige Nützlichkeitsüberzeugungen berichteten, eine Nützlichkeitsintervention erhielt, wohingegen die andere Hälfte zufällig eine der anderen vier Interventionen zugewiesen bekam. Die Motivationsinterventionen bestanden jeweils aus einem Kurzvideo (10-15 Minuten) zu einem der fünf motivationalen Konstrukte und anschließend einem 15-minütigen schriftlichen Arbeitsauftrag. Nach einer Woche folgte der Posttest. Insgesamt liegen von $N = 330$ Bachelorstudierenden (80 % weiblich) vollständige Daten vor.

Ergebnisse und Diskussion

Zunächst wurde überprüft, ob die Studierenden, die eine bedarfsgerechte Motivationsintervention erhielten, eine höhere konstruktsspezifische Motivation nach der Intervention zeigten als die Studierenden, die denselben Bedarf hatten, aber zufällig einer anderen Motivationsintervention zugewiesen wurden. Studierende, die eine bedarfsgerechte Growth Mindset und Nützlichkeitsintervention erhielten, zeigten im Posttest eine höhere konstruktsspezifische Motivation (z.B. höheres Growth Mindset) als Studierende mit demselben Bedarf, die jedoch eine andere Intervention erhielten (Growth Mindset Intervention: $b = -1.30$, $BF_{10} = 5188558$; Nützlichkeitsintervention: $b = -.45$, $BF_{10} = 3.20$). Die Ergebnisse bezüglich der Social-Belonging Intervention, Attributionstraining und Lernzielorientierungs-Intervention sind inkonklusiv.

In einem zweiten Schritt wurde die Gesamtgruppe an Studierenden mit bedarfsgerechter Zuweisung mit der Gesamtgruppe an Studierenden mit zufälliger Zuweisung der Motivationsinterventionen hinsichtlich mehrerer globaler Outcomes verglichen. Studierende, die eine bedarfsgerechte Intervention erhielten (ungeachtet der tatsächlichen Intervention), berichteten niedrigere Wechselintentionen ($b = -.13$, $BF_{10} = 12.79$), niedrigere Abbruchsintentionen ($b = -.13$, $BF_{10} = 8.40$) und niedrigere subjektive Kosten ($b = -.16$, $BF_{10} = 81.35$). Hinsichtlich der Zufriedenheit mit der Vorlesung und des selbstberichteten Engagements sind

die Ergebnisse inkonklusiv. Weitere Analysen mit der Klausurnote und Verhaltensdaten aus einem Lernmanagementsystem als globale Outcomes stehen noch aus.

Die ersten Ergebnisse der vorliegenden Studie deuten auf die Überlegenheit der individualisierten Zuweisung von Motivationsinterventionen gegenüber einem „one-size-fits-all“-Ansatz hin. Dennoch weist die Studie Limitationen auf (bspw. Stichprobengröße auf Ebene der einzelnen Interventionen), die die zum Teil inkonklusiven Ergebnisse möglicherweise erklären können.

Dieses Semester wird alles besser! - Erklärung interindividueller Unterschiede in der Nutzung von Lernaktivitäten und der Intentions-Verhaltenslücke von Lehramtsstudierenden

Vivien Rieder, Eva Bosch, Birgit Spinath

Universität Heidelberg, Deutschland

Die selbstregulierte Nutzung von Lernaktivitäten ist notwendig für ein erfolgreiches Studium (Bernardo et al., 2019). Viele Studierende scheitern allerdings daran, Lernaktivitäten in dem Ausmaß zu nutzen, in dem sie es sich zu Beginn des Semesters vornehmen (Bosch et al., 2021). Dadurch entsteht eine Intentions-Verhaltenslücke im Lernverhalten der Studierenden. Eine mögliche Erklärung dieser Lücke stellt die Abnahme der Intentionen im Verlauf der ersten Semesterhälfte analog zur Abnahme der Motivation in diesem Zeitraum dar (Benden & Lauermann, 2019; Sander et al., 2024).

Interindividuelle Unterschiede in der Größe der Intentions-Verhaltenslücke sollten jedoch auch durch personenseitige Faktoren erklärbar sein. Selbstregulationsfähigkeiten wie beispielsweise das Überwachen der eigenen Zielverfolgung haben sich bereits als ein relevanter personenseitiger Faktor zur Erklärung der Intentions-Verhaltenslücke erwiesen (De Bruin et al., 2012; Sniehotta et al., 2005). Motivationsbezogene Kompetenzen (Schaller & Spinath, 2017; Spinath, 2005) wurden in diesem Zusammenhang noch nicht untersucht. Diese fassen individuelle Fähigkeiten des selbstregulierten Lernens zusammen, die in den unterschiedlichen Stadien motivationaler Phasenmodelle wie des Rubikon-Modells der Handlungsphasen (Heckhausen & Gollwitzer, 1987) zur erfolgreichen Bewältigung notwendig sind. Motivationsbezogene Kompetenzen werden in motivationale Kompetenzen (MoKo), volitionale Kompetenzen der Zielsetzung und Handlungsplanung (VoKo-ZH), volitionale Kompetenzen der Ressourcenaktivierung (VoKo-RA) und bewertende Kompetenzen (BeKo) unterteilt. In der vorliegenden Untersuchung wurde die Annahme überprüft, dass die Ausprägungen dieser motivationsbezogenen Kompetenzen Unterschiede in der Nutzung von Lernaktivitäten und im Ausmaß der Intentions-Verhaltenslücke von Studierenden erklären.

Im Rahmen einer Vorlesung für Lehramtsstudierende untersuchten wir in einer längsschnittlichen Studie mit drei Messzeitpunkten zunächst das Auftreten der Intentions-Verhaltenslücke sowie die Entwicklung der Intentionen in der ersten Semesterhälfte. Anschließend nutzten wir motivationsbezogene Kompetenzen, um zunächst das Lernverhalten der Studierenden und schließlich die Größe der Intentions-Verhaltenslücke vorherzusagen. Die Stichprobe bestand aus 7 Kohorten von Lehramtsstudierenden ($N = 1852$, 68.2% weiblich, $M_{\text{Alter}} = 21.69$ Jahre, $SD_{\text{Alter}} = 3.10$ Jahre) zwischen dem Sommersemester 2018 und dem Sommersemester 2021. Die Intentionen (Anfang und Mitte des Semesters) und deren Umsetzung (Ende des Semesters) wurden durch Selbstauskünfte in Bezug auf sechs unterschiedliche Lernaktivitäten erfasst (Vorlesungsbesuch, Nachbereitung mit Hilfe der Folien, Nachbereitung mit Literatur, Einreichung von Übungsstatements, Selbsttests und Lerngruppen).

Wir fanden eine Intentions-Verhaltenslücke zwischen den Intentionen zu Semesterbeginn und der Nutzung der Lernaktivitäten. Die Intentionen in der Mitte des Semesters waren außerdem signifikant niedriger als zu Beginn. Die hierarchische Regression zur Vorhersage des Lernverhaltens ergab, dass die volitionalen Kompetenzen der Zielsetzung und Handlungsplanung sowie die bewertenden Kompetenzen die Nutzung der Lernaktivitäten über die Intention hinaus vorhersagten. Dabei gingen eine höhere Intention sowie höhere volitionale Kompetenzen der Zielsetzung und Handlungsplanung im finalen Modell mit einer vermehrten Nutzung der Lernaktivitäten einher, während höhere bewertende Kompetenzen mit einer geringeren Nutzung assoziiert waren. Bei der Vorhersage der Intentions-Verhaltenslücke durch motivationsbezogene Kompetenzen wurden lediglich jene Studierende betrachtet, die die Lernaktivitäten in geringerem Maße umsetzten als sie es sich vorgenommen hatten. Dabei gingen höhere volitionale Kompetenzen der Zielsetzung und Handlungsplanung mit einer kleineren Intentions-Verhaltenslücke, höhere bewertende Kompetenzen jedoch mit einer größeren Intentions-Verhaltenslücke zusammen. Insgesamt waren die motivationsbezogenen Kompetenzen nur für eine geringe Varianzaufklärung verantwortlich, weshalb wir in explorativen Analysen weitere Prädiktoren untersuchten. Dabei gingen höhere Intentionen zu Semesterbeginn, ein höheres Alter und eine schlechtere Abiturnote jeweils mit einer größeren Intentions-Verhaltenslücke einher.

Die vorliegende Studie gibt sowohl Hinweise auf die Existenz einer Intentions-Verhaltenslücke im Studienkontext als auch auf eine negative längsschnittliche Entwicklung der Intentionen analog zur Motivationsentwicklung. Motivationsbezogene Kompetenzen konnten zur Erklärung der Intentions-Verhaltenslücke beitragen, weitere Prädiktoren müssen jedoch noch identifiziert werden. Zukünftige Forschung sollte daher sowohl den Verlauf von Intentionen anhand mehrerer Messzeitpunkte als auch weitere mögliche Ursachen für die Intentions-Verhaltenslücke untersuchen.

PS-64: Education and Career Prospects

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 135

Chair der Sitzung: Hanna Beißert

Eine Untersuchung der Auswirkungen des Nichtbestehens einer wichtigen Prüfung auf die Entwicklung von beruflichen Interessen von Studierenden

Gfrörer Thomas

Hector-Institut für Empirische Bildungsforschung, Universität Tübingen, Deutschland

Einleitung

Berufliche Interessen sind assoziiert mit Bildungsentscheidungen (Volodina et al., 2015; Wille et al., 2020), Leistung (Nye et al., 2017, 2018), Zufriedenheit (Hoff, Song, Wee, et al., 2020) sowie dem Verbleib (Allen & Robbins, 2008; Donohue, 2006) in Studium und Beruf. Trotz ihres Einflusses auf Studium, Beruf und andere wichtige Lebensbereiche (Stoll et al., 2017), ist wenig über die Determinanten der Entwicklung von beruflichen Interessen bekannt (Rounds & Su, 2014). In der folgenden Studie wird untersucht, wie sich das Nichtbestehen einer wichtigen Prüfung auf die Entwicklung von beruflichen Interessen von Studierenden auswirkt.

Theoretischer Hintergrund

Auch wenn meta-analytische Befunde darauf hin deuten, dass berufliche Interessen über die Lebensspanne hinweg relativ stabil sind und ihre Stabilität mit dem Alter ansteigt (Hoff et al., 2018; Low et al., 2005), können individuelle Entwicklungsverläufe von diesem Muster abweichen. Beispielsweise weisen Hoff et al. (2020) und Gfrörer et al. (2021, 2024) darauf hin, dass sich die beruflichen Interessen mancher Personen stärker verändern als die von anderen. Theorien zur Entwicklung von Interessen sprechen dem situationalen Interesse hier eine besondere Rolle zu (Hidi & Renninger, 2006; Su et al., 2019). Es wird angenommen, dass das kontinuierliche Erleben von situationalem Interesse, eine Entwicklung dispositionaler Interessen hervorrufen kann (Hidi & Renninger, 2006; Su et al., 2019). Es wird vermutet, dass situationales Interesse, durch positive situationale Auslöser, wie die persönliche Relevanz einer Aktivität, hervorgerufen werden kann (Hidi & Renninger, 2006). Im Rahmen dieser Arbeit wird untersucht, ob ebenfalls ein negativer situationaler Auslöser, wie das Nichtbestehen einer wichtigen Prüfung, die Entwicklung von Interessen beeinflussen kann.

Fragestellung

Hat das Nichtbestehen einer wichtigen Prüfung einen Einfluss auf die Entwicklung von beruflichen Interessen?

Methode

Für die Untersuchung der Fragestellung werden Daten der Studie „Transformation des Sekundarschulsystems und akademische Karrieren“ (TOSCA; Köller et al., 2004) verwendet. Die Analysestichprobe beträgt $N = 412$ Teilnehmende. In die Analysestichprobe wurden Teilnehmende der Wellen zwei (Alter: 21) und drei (Alter: 23) mitaufgenommen, welche sich zu beiden Zeitpunkten im Studium befanden. Die Erfassung des Nichtbestehens einer wichtigen Prüfung erfolgte zur dritten Welle retrospektiv für den Untersuchungszeitraum, anhand einer Skala für Lebensereignisse (Vaidya et al., 2002). Die Studienteilnehmenden sollten angeben, ob sie innerhalb der letzten 2 Jahre von kritischen Lebensereignissen betroffen waren, unter anderem vom Nichtbestehen einer wichtigen Prüfung.

Um die generische Frage auf eine Inhaltsdomäne beziehen zu können, wurden in die Analysestichprobe nur Studienteilnehmende aufgenommen, welche einen MINT-Studiengang belegen. Dadurch soll erreicht werden, dass sich die Antwort der Studienteilnehmenden nur auf die Inhaltsdomäne MINT bezieht und nicht beispielsweise auch auf die Inhaltsdomäne Sozialwissenschaften. Berufliche Interessen im Bereich MINT wurden anhand der Dimensionen Realistic und Investigative des RIASEC-Modells operationalisiert (Holland, 1997) und durch den Allgemeinen Interessen-Struktur-Test sowohl zur zweiten als auch zur dritten Welle der TOSCA Studie gemessen ($.82 < \alpha < .89$; Bergmann & Eder, 2005).

Da sich Teilnehmende, welche eine wichtige Prüfung nicht bestehen, von Teilnehmenden, welche dieses Ereignis nicht erlebt haben vermutlich bereits auf Prätest-Werten unterscheiden, wurde auf ein Verfahren zur Schätzung kausaler Effekte in Beobachtungsstudien zurückgegriffen: Inverse-Probability of Treatment Weighting (Thoemmes & Ong, 2016). Zunächst wurden für die Teilnehmenden Stichprobengewichte berechnet und dann anschließend bei der Schätzung eines multiplen Regressionsmodells mitaufgenommen. Das Gewichtungungsverfahren führt zu einer Balancierung der beiden Gruppen auf wichtigen Hintergrundvariablen, welche den Effekt des Nichtbestehens einer Prüfung auf die Entwicklung beruflicher Interessen konfundieren können.

Ergebnisse

Die vorläufigen Befunde deuten auf einen überwiegend kleinen ($|d| < 0.05$) Einfluss des Nichtbestehens einer wichtigen Prüfung auf die Entwicklung beruflicher Interessen hin. Die Befunde verdeutlichen, dass berufliche Interessen im Erwachsenenalter auch nach Einfluss von starken negativen Auslösern relativ stabil bleiben. Zur Überprüfung der Robustheit der Effekte gegenüber unbeobachteten Konfundierungsvariablen, sind zusätzliche Sensitivitätsanalysen geplant (Ding & VanderWeele, 2016).

Skills to equalize the bills: How skills affect the early gender pay gap of college graduates

Nora Lorenz¹, Frauke Peter², Fabian Trennt³

¹Deutsches Zentrum für Hochschulwissenschaften (DZHW), Deutschland; ²Deutsches Zentrum für Hochschulwissenschaften (DZHW), Deutschland; ³Deutsches Zentrum für Hochschulwissenschaften (DZHW), Deutschland

After its massive expansion in the 21st century, females today outnumber their male counterparts in higher education and have greater educational success. Yet their increased investment in human capital does not translate into earnings. Although there has been a marked decrease in the gender pay gap (GPG) in recent decades, it has declined much more slowly at the top of the wage distribution and is particularly prevalent for highly skilled workers (Blau & Kahn, 2017; Albrecht et al., 2018). Notably, gender differences in earnings are already visible in the first job right after graduation, with the main reason attributed to it being the gender-specific choice of study (Francesconi and Parey, 2018). Campbell et al. (2022) find that high-attaining women systematically undermatch in terms of expected earnings compared to their male counterparts, mostly driven by major choice in higher education. However, there exists a notable gap in knowledge about the extent to which major choice influences occupational sorting. Why, for example, do engineering graduates work as consultants or political scientists as PR professionals?

More recent approaches to understanding the (early) GPG emphasize non-traditional factors such as gender differences in preferences, attitudes and social norms that lead to occupational sorting (Bertrand, 2011). Exley & Kesler (2022) demonstrate, that even school-aged girls assess their skills in male-dominated fields lower compared to boys. Thus, how individuals perceive their own skills affects their life outcomes, including labor market returns and progression (Adamecz-Völgyi and Shure, 2022). Another line of research has found that women and men in fact differ in their skill sets and that changing skill requirements in the labor market due to technological change or other trends affect them differently (Black and Spitz-Oener, 2010; Cortes et al., 2024).

The aim of our research is to explore the link between choice of major, self-assessed skills and labor market returns. We use a unique and rich data set on graduates in Germany, the DZHW Graduate Panel, which includes detailed information on a range of cognitive and non-cognitive perceived skills acquired in higher education, as well as early career choices. Multiple cohorts and waves allow us to assess whether the choice of field of study per se matters when taking into account the perceived skills acquired in higher education. In a first step, we investigate whether certain skills are less rewarding for women than for men by using a Kitagawa-Oaxaca-Blinder decomposition to analyze the early GPG as well as quantile regressions. In a second step we employ instrumental variables approach to validate our results, as well as a range of robustness checks. The aim is to use the unique survey information at our disposal to shed light on the unexplained part of early-career gender differences in labor market outcomes.

Our preliminary results show that perceived skills have a small but significant effect on pay gaps at job entry a share of the EGP that is notably larger than that of socio-economic background. However, different types of skills differ in their relevance for labor market outcomes. In particular, our results suggest that women experience a wage disadvantage because they rate their social skills higher, which reduces earnings *ceteris paribus*, while men rate their analytical skills higher, which increases earnings. Nevertheless, field of study remains the most important explanatory factor for entry wages at the early stages of careers. In the future, we plan to go beyond wages and include a range of labor market outcomes to better understand the gendered career trajectories of highly qualified individuals. In particular, we are interested in occupations graduates go on to pursue, and the role of perceived skills for their career trajectories.

How Employers Screen CVs for Skilled Workers: The Role of Educational Credentials

Michèle Kreienbühl^{1,2}, Miriam Grønning¹, Irene Kriesi¹

¹Eidg. Hochschule für Berufsbildung, EHB, Schweiz; ²Universität Bern

In addressing the social challenge of matching skilled workers and jobs, educational credentials play a crucial role. In countries with a vocationally oriented education system and labor market, such as Germany and Switzerland (Kriesi, Buchmann, and Sacchi 2010; Di Stasio and van de Werfhorst 2016), access to jobs in a given occupational field often requires occupation-specific credentials (Kriesi, Buchmann, and Sacchi 2010; Di Stasio and van de Werfhorst 2016). For most of these fields there are multiple educational credentials at upper-secondary and tertiary level, such as initial VET, tertiary-level professional VET and university diplomas. Research shows that employers' preferences for these credentials depend on the occupation (Rageth, Renold, and Sritharan 2022; McDonald and Korber 2023). A likely reason for this finding is that the task profile of occupations differs (Autor, Levy, and Murnane 2003; Fernández-Macías and Bisello 2022). However, studies on how employers assess educational credentials in the hiring process are scarce, and differences between occupational fields in employers' preferences for these qualifications are underexplored.

Thus, our research questions are: Does the relevance of educational credentials in the CV screening process depend on the task profile of occupational fields? Do preferences for different types of educational credentials depend on the task profile of occupational fields?

We classify occupations using the task approach by Fernández-Macías and Bisello 2022. Tasks are units of work activity that require task-specific skills (Autor 2013). Thus, the task profile determines the skill requirements of an occupation. Drawing upon Human Capital (Becker 1964), Signaling (Spence 1973) and Screening theory (Stiglitz 1975), we assume that employers use jobseeker characteristics on CVs to evaluate whether the jobseeker fulfill the required task-specific skills. Individuals acquire different levels of practical and theoretical as well as occupation-specific and general skills (e.g., problem-solving, and interactive skills) depending on their education (Becker 1964). Thus, jobseekers' credentials are one of the main providers of and signals for task-specific skills.

Our empirical analysis is based on a novel and little explored data base. We use click data of the years 2020 – 2023 from the recruitment platform of the Swiss public employment service, Job-Room.ch. Job-Room.ch provides employers with detailed online CVs of jobseekers which are registered at a regional unemployment office. When searching for suitable candidates, employers specify an occupation and set various filters (e.g., educational credentials, occupation-specific experience). They browse through the list of jobseekers and contact those whom they are interested in. The click data allows the observation of employers' real-life search behavior. Rich information on jobseeker characteristics makes the data well suited to examine the employers' preferences for different types of educational credentials.

We proceed in two steps. First, we analyze the search entries of the employers. Multinomial logistic regression is used to examine the relevance of educational credentials in occupational fields with different task profiles. The dependent variable measures whether the employer screened for different educational credentials using the search filter. Based on the classification by Fernández-Macías and Bisello (2022) we will determine the task profile of the occupation. Second, we run logistic regressions to analyze which types of educational qualifications have an effect on the probability that the employer clicks on the contact button of a jobseeker's profile.

As the data arrived with a delay, we have only just started with analyses and are not yet able to present any empirical results. The empirical analyses will be completed by the time of the conference. Our results will provide valuable insights for young people faced with educational decisions and for education and labor market policy to counteract the shortage of skilled workers.

Langzeitwirkungen von Service-Learning-Seminaren

Marlene Bunte¹, Hanna BeiBert^{1,2,4}, Anne-Sophie Waag³, Julia Derkau³, Manfred Hofer³

¹Goethe Universität Frankfurt; ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation; ³Universität Mannheim;

⁴DeA-Center for Individual Development and Adaptive Education Frankfurt

Service-Learning verbindet zivilgesellschaftliches Engagement (Service) mit hochschulischem Wissenserwerb (Learning) im Kontext von Schulunterricht oder Hochschullehre (Reinders, 2016; Sliwka, 2004). Durch Anwendung in der Praxis wird die erlernte

Theorie reflektierend verankert und es findet eine Auseinandersetzung mit dem eigenen Selbst als Teil der Gesellschaft statt (Reinders, 2016). Fachwissenschaft, Praxis und soziales Engagement werden hierbei verknüpft und es werden prägende Erfahrungen geschaffen, die zur Persönlichkeitsbildung beitragen sollen (Hofer, 2019). Kurzzeitige bis mittelfristige positive Effekte dieser Lehrmethode wurden empirisch bereits umfangreich belegt (Gerholz, 2020; Slepcevic-Zach & Fernandez, 2012). Die vorliegende Studie beantwortet nun erstmals die Frage nach der Nachhaltigkeit und Langfristigkeit der Wirkung von Service-Learning auf Studierende.

Dafür wurden 19 ehemalige Teilnehmende (15 weiblich, 4 männlich, 38-59 Jahre, $M=42.37$, $SD=4.41$) der ersten in Deutschland implementierten Service-Learning-Seminare (durchgeführt zwischen 2003 und 2007 an der Universität Mannheim) zu ihren Erfahrungen im Seminar und den dort erworbenen Fertigkeiten befragt. Zum Befragungszeitpunkt lag das Seminar für die Befragten 15-20 Jahre zurück.

Mithilfe von teilstandardisierten, retrospektiven Leitfadeninterviews wurden langfristige Auswirkungen von Service-Learning auf drei zentrale Zieldimensionen (Selbstwirksamkeit, Theorie-Praxis-Reflexion, Civic Attitudes) untersucht. Der Fragebogen wurde deduktiv-induktiv an der Theorie entwickelt und basierend auf Pilot-Interviews überarbeitet. Er beinhaltete sowohl geschlossene Frageformate (für quantitative Analysen) als auch offene Fragen (für qualitative Analysen). Im vorliegenden Beitrag stehen die Befunde der qualitativen Analysen im Vordergrund. Zuvor werden dennoch kurz die Befunde, die sich aus den geschlossenen Antworten ergeben haben, dargelegt.

Die Teilnehmenden gaben unter anderem an, wie sie rückblickend den Einfluss des Service-Learning-Seminars im Vergleich zu herkömmlichen Seminaren im Hinblick auf Selbstwirksamkeit, Theorie-Praxis-Reflexion und Civic Attitudes einschätzen. Bezüglich der Selbstwirksamkeit bewerteten alle Personen das Service-Learning-Seminar im Vergleich zu anderen Veranstaltungen als förderlicher, wobei mehr als die Hälfte sogar angab, dass das Seminar die berufsbezogene Selbstwirksamkeit im Vergleich zu anderen Veranstaltungen viel mehr gestärkt habe. Bezüglich Theorie-Praxis-Reflexion und Civic Attitudes gab jeweils etwas mehr als die Hälfte der Befragten an, dass das Service-Learning-Seminar förderlicher als andere Seminare gewesen sei. Für alle drei Dimensionen wurden Einstichproben-t-Tests berechnet, die zeigen, dass die Mittelwerte signifikant von den theoretischen Mittelwerten der Skalen abweichen ($p < .001$), was einen förderlicheren Einfluss des Service-Learning-Seminars indiziert.

Für die qualitative Auswertung der Interviews wurde anhand des Interviewmaterials ein Kategoriensystem entwickelt. Die Kodierung wurde mithilfe von MAXQDA durchgeführt. Der Fokus lag auf langfristigen Effekten von Service-Learning auf die berufliche Entwicklung, Selbstwirksamkeit, gesellschaftliche Verantwortungsfähigkeit und Theorie-Praxis-Reflexion. Die Auswertung ergab eindrücklich, dass der Seminarbesuch bei vielen der Befragten nachhaltige Effekte hatte, die sich teilweise bis heute in der Persönlichkeit oder im Berufsleben niedergeschlagen. Insbesondere der praktische Teil des Seminars scheint hier förderlich gewesen zu sein. Viele beschrieben die praktischen Erfahrungen als außergewöhnlich prägend, sei es für spätere berufliche Entscheidungen oder für die Entwicklung der eigenen Persönlichkeit. Die Arbeit mit benachteiligten Kindern stärkte für viele nicht nur die gesellschaftliche Verantwortungsfähigkeit, sondern schärfte grundsätzlich den Blick für eigene Privilegien und gesellschaftliche Probleme und steigerte das Einfühlungsvermögen sowie die gesellschaftliche Perspektivenübernahme in einer Weise, von der die Befragten nach eigenen Aussagen teilweise bis heute profitieren.

Zudem hatte das Seminar auch auf die Entwicklung konkreter Fähigkeiten einen nachhaltigen förderlichen Einfluss und vermittelte Methoden, die die Befragten teilweise noch heute erfolgreich einsetzen, so z.B. die Verknüpfung von theoretischen Inhalten mit praktischer Anwendung, interdisziplinäre Zusammenarbeit oder die Methode der Evaluation. Die angestrebte Theorie-Praxis-Reflexion scheint also zu funktionieren. Darüber hinaus stellte die erlebte Selbstwirksamkeit im Rahmen des praktischen Teils des Seminars für die meisten der Teilnehmenden eine sehr prägende Erfahrung dar. Die Erfolgserlebnisse und das direkte Feedback sorgten vielfach für Selbstwirksamkeitserfahrungen, die wegbereitend für die Berufswahl und weitere berufliche Entscheidungen waren.

Unsere Studie legt nahe, dass Service-Learning als Lehrmethode nicht nur einen unmittelbaren, sondern auch einen erheblichen nachhaltigen Einfluss auf die weitere Entwicklung von Studierenden hat und eine größere Verbreitung der Methode die deutsche Hochschullehre verbessern könnte.

PS-65: Digital Media and Professional Development: Importance for Lifelong Learning and Career Growth

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 256

Chair der Sitzung: Falk Scheidig

Informelles Lernen mit digitalen Medien – Analysen zum Einfluss von Motivation auf die Beteiligungschancen von Erwachsenen

Sara Reiter¹, Julia Gorges²

¹Freie Universität Berlin; ²Philipps-Universität Marburg

Mit Blick auf das Lehren und Lernen in der Weiterbildung halten digitale Medien nicht nur Potenziale für die didaktische Gestaltung formaler und non-formaler Bildungsangebote bereit, sondern eröffnen auch vielfältige Möglichkeiten für informelles Lernen (u. a. Rohs, 2013; Seufert & Meier, 2016). Dem Begriff des informellen Lernens liegt kein einheitliches Verständnis zugrunde (Overwien, 2005; Kaufmann, 2016). In statistischen Erhebungen zur Weiterbildungsteilnahme, wie bspw. dem Adult Education Survey (AES), wird informelles Lernen in Anlehnung an die europäische Lernformtypologie (Eurostat, 2016) als ein intendiertes Lernen außerhalb von institutionalisierten und organisierten Bildungssettings klassifiziert und durch verschiedene Lernaktivitäten (u. a. auch durch das Lernen mit digitalen Medien) erfasst (BMBF, 2024; Bilger, 2016). Folgt man den Befunden des AES, so gaben im Jahr 2022 40 Prozent der 18-69-Jährigen an, durch die Nutzung von Lernangeboten am Tablet, an Smartphones, am Laptop oder Computer informell gelernt zu haben. Damit hat sich die Teilnahme an dieser Form des informellen Lernens im Vergleich zum Jahr 2018 nahezu verdoppelt (23 Prozent) (BMBF, 2024). Allerdings verweisen Studien auch auf soziale Unterschiede. Beispielsweise kommt der „Monitor Digitale Bildung“ zu dem Schluss, dass es häufiger Höherqualifizierte und Berufstätige sind, die mit digitalen Medien lernen (Schmid et al. 2017). Insofern erscheint es lohnenswert, neben der Reichweite des informellen Lernens mit digitalen Medien auch die jeweiligen Voraussetzungen und Bedingungen in den Blick zu nehmen, unter denen ein solches Lernen gelingen kann.

Der vorliegende Beitrag greift die Frage nach relevanten Prädiktoren der Beteiligung am informellen Lernen mit digitalen Medien aus einer motivationstheoretischen Perspektive auf. Rekurrierend auf den wert-erwartungstheoretischen Ansatz nach Eccles (1983) wird davon ausgegangen, dass Bildungsentscheidungen Erwachsener und somit auch die Entscheidung, informell zu lernen, durch (weiter-)bildungsbezogene Wertüberzeugungen beeinflusst werden (Gorges, 2015; Gorges & Hollmann, 2015). Nach Eccles (2005) können bildungsbezogene Wertüberzeugungen sowohl vorteilhafte als auch nachteilige Überzeugungen (Kosten) beinhalten, wobei solche Überzeugungen dem Modell zufolge wiederum durch sozial-kulturelle Faktoren und vorangegangene Erfahrungen geprägt werden (ebenda).

Ausgehend von diesen theoretischen Implikationen wird im Beitrag untersucht, inwiefern Wertüberzeugungen gegenüber dem Lernen mit digitalen Medien in einem Zusammenhang mit der Beteiligung am informellen Lernen mit digitalen Medien Erwachsener stehen und inwiefern die Ausprägung solcher Wertüberzeugungen von den soziodemografischen Merkmalen (Alter, Geschlecht, Bildungshintergrund, Erwerbsstatus) der Befragten und ihren Erfahrungen mit der Nutzung digitaler Medien abhängen.

Für die Untersuchung werden Daten des Adult Education Survey aus dem Jahr 2018 herangezogen, da in diesem Erhebungsjahr Informationen zu den Wertüberzeugungen gegenüber dem Lernen mit digitalen Medien vorliegen. In die Stichprobe der vorliegenden Untersuchung gehen Personen im Alter von 18-69 Jahren ein (n = 5233). Die Überprüfung der theoretisch postulierten Zusammenhänge erfolgt mittels eines Strukturgleichungsmodells, das mit der Software R (lavaan) auf Basis imputierter Daten (m=25) berechnet wurde.

Befunde erster Analysen belegen, dass (vorteilhafte) Wertüberzeugungen gegenüber dem Lernen mit digitalen Medien die Beteiligung Erwachsener am informellen Lernen mit digitalen Medien begünstigen. Entgegen unserer Erwartungen zeigen allerdings auch die wahrgenommenen Kosten gegenüber dem Lernen mit digitalen Medien einen (schwachen) positiven Effekt auf die Beteiligung. Darüber hinaus lassen die Ergebnisse darauf schließen, dass die Effekte der digitalen Nutzungserfahrungen auf die Beteiligung partiell durch die (vorteilhaften) Wertüberzeugungen mediert werden. Für die soziodemografischen Merkmale wie z. B. das Bildungsniveau lassen sich sowohl Effekte auf die motivationalen Faktoren im Modell als auch auf die Beteiligung am informellen Lernen mit digitalen Medien beobachten. Dies lässt darauf schließen, dass für die Erklärung der bildungsbezogenen Disparitäten hinsichtlich der Beteiligung am informellen Lernen mit digitalen Medien neben motivationalen Faktoren weitere Aspekte eine Rolle spielen dürften (z. B. Gelegenheitsstrukturen).

Im Anschluss an die Ergebnispräsentation werden sowohl Limitationen (z. B. Analyse mit Querschnittdaten) als auch weitergehende Forschungsbedarfe aufgezeigt und diskutiert.

Evolving Work, Evolving Skills: How the Use of Digital Technologies Requires Different Forms of Further Training Depending on Task Complexity

Marco Seegers¹, Birgit Zeyer-Glozzo²

¹Bundesinstitut für Berufsbildung (BIBB), Deutschland; ²TU Dortmund, Deutschland

Technological change poses major challenges to society in the 21st century. In particular, the associated changes in job tasks and skill requirements in today's western societies require workers to continuously and comprehensively update their skills. One way to adapt to new and often more complex job and skill requirements is to participate in training. However, it is still unclear what kind of further training is used to adapt to the increasing complexity of tasks and the use of technology in the workplace. This study examines the impact of technology use at work on participation in informal and non-formal further training in Germany, considering the complexity of job tasks. We employ empirical-quantitative analyses to investigate two research questions: 1.) Does technology use influence the likelihood of participating in non-formal and informal further training? 2.) To what extent does the task complexity moderate the impact of technology use on the likelihood of non- and informal further training?

Following the skill-biased technological change thesis (Autor et al., 1998) and the task-based approach (Autor et al., 2003), we argue that technological change affects the demand for different types of skills, disproportionately favouring employees with higher skills and more complex tasks. On the other hand, it is the less complex tasks that are easier to automate and replace, while the use of technology and more complex and abstract problem-solving, communication and management skills, as well as flexibility, creativity and intuition, are more likely to be complementary. Unskilled or low-skilled employees with less complex tasks are therefore under pressure to adapt to technological change and to acquire new knowledge that will enable them to take on more complex tasks that cannot (yet) be automated (Arnold et al., 2016). Based on the specific functionality of different forms of

training (Rogers, 2014), we argue that employees with less complex job tasks are more likely to participate in non-formal training, where new and specific concepts and skills are taught that broaden the knowledge base of employees to take on new tasks in the wake of technological change. In contrast, employees with more complex tasks are more likely to take part in informal training, which is more likely to refine, specify and update existing knowledge.

Using the adult cohort (SC6: 13.0.0) of the German National Education Survey (Blossfeld & Maurice, 2019; NEPS Network, 2022), logistic regression models reveal that increased use of networked digital technologies significantly increases participation in informal training. Conversely, the effect on non-formal training is less pronounced and not statistically significant. With regard to the interaction between technology use and task complexity, it can only be shown for the sample that, as hypothesised, employees with low task complexity participate more often in non-formal training and employees with high task complexity participate more often in informal training in connection with increased technology use. Although the interaction effects cannot be generalised, the data provide a first indication that non-formal and informal training may play different roles in the process of technological change. However, the relationship may be more complex, influenced by the initial level of technology use, and vary by age or sector, as suggested by additional analyses.

The study points to the importance of holistic training strategies, tailored to the different needs of employees based on the complexity of their tasks and the use of technology in the workplace. However, further research is needed to validate the significance of the findings. Furthermore, the context and specific skills imparted by different types of training and their effectiveness in preparing workers for the ongoing digital transformation should be investigated.

ENTFÄLLT: The Relation Between Critical Online Reasoning Skills of Incoming University Students in Germany and Their Epistemological Beliefs

Lisa Martin de los Santos Kleinz¹, Eric Klopp², Anna Horrer³, Jan Zottmann³, Olga Zlatkin-Troitschanskaia¹

¹Johannes Gutenberg-Universität Mainz, Deutschland; ²Universität des Saarlandes, Deutschland; ³Ludwig-Maximilians-Universität München, Deutschland

Theoretical background

The ability to effectively research and evaluate online information is a crucial skill for students and a key factor in academic success [1,2,3,4]. Critical Online Reasoning (COR) encapsulates these essential competencies for effective online learning. It includes the skills needed to search for, evaluate, select, and utilize online information to address a problem [5]. The COR-construct is divided into three skill facets: Acquisition of Online Information, such as the ability to use search engines to find reliable sources purposefully; Critical Information Evaluation, such as the secure differentiation between reliable and unreliable sources, information and arguments from claims, and Reasoning using Evidence, Argumentation, and Synthesis [6].

Epistemological beliefs, as individual assumptions about knowledge and the process of knowledge acquisition, influence how individuals engage with knowledge, including whether they trust, question, reconsider, or reject (online) statements [7,8]. Epistemological beliefs positively affect the appropriate processing of online information and critical thinking skills [9,10,11,12,13,14] and therefore should affect COR-skills and beliefs about Internet-related information sources [15].

Question

This exploratory study investigates the extent to which epistemological beliefs influence basic COR-skills and beliefs regarding subject-related Internet searches for information and whether these differ between the four domains of economics, medicine, sociology, and physics.

Methods

In total, 1939 participants (881 female, $M_{age}=19.91$, $sd_{age}=2.42$) took part in the study, all of whom were in the first semester.

Epistemological beliefs were assessed using the following six dimensions [16] with 3-item-scales: Personal Justification (PJ, $\alpha=.81$), Justification by Authority (JA, $\alpha=.75$), Justification by Multiple Sources (JS, $\alpha=.75$), Justification by the Scientific Community (JC, $\alpha=.60$), Certainty of Knowledge (CK, $\alpha=.75$), and Reflexive Nature of Knowledge (RN, $\alpha=.83$).

The students' self-assessment of the COR-related skills like Acquisition of Online Information (AOI), Critical Information evaluation of Sources (CIS), Critical evaluation of Information Reliability (CIR), Critical Information evaluation of Arguments (CIA), and Critical Information evaluation of Generative systems (CIG) were measured with single-item indicators.

Beliefs about Internet-related information sources [16] were assessed with the following four 3-item-scales: Justification for Internet-based knowing (JIK, $\alpha=.63$), Source of Internet-based knowledge (SOK, $\alpha=.69$), Simplicity of Internet-based knowledge (SIK, $\alpha=.69$), and Certainty of Internet-based knowledge (CIK, $\alpha=.62$).

We employed a structural equation model for multivariate regression analysis for the whole sample using a single-group-model and investigated differences in regression coefficients between the domains using a multiple-group-model and Wald tests.

Results

Concerning COR, the epistemological beliefs of PJ ($\beta=.056^{**}$) and JC ($\beta=-.049^{**}$) predict CIS ($R^2=.026$). The epistemological beliefs PJ ($\beta=.045^{**}$) and JA ($\beta=.082^{***}$) predict CIR ($R^2=.032$), and again PJ ($\beta=.079^{**}$) and JA ($\beta=.079^{**}$) predict CIA ($R^2=.035$). Finally, RN ($\beta=.170^{***}$) predicts CIG ($R^2=.098$). Significant differences exist between the regression coefficients predicting CIG by PJ ($\chi^2=8.768^{***}$, $df=3$). In particular, the regression coefficients in economics, sociology, and medicine differed from those in physics.

Concerning beliefs about internet-related information sources, PJ ($\beta=-0.33^{**}$), JS ($\beta=.118^{***}$), CK ($\beta=.019^{**}$), and RN ($\beta=.040^{**}$) predicted JIK ($R^2=.308$). SOK ($R^2=.094$) was only predicted by RN ($\beta=.070^{***}$). The epistemological beliefs JS ($\beta=.048^{**}$), CK ($\beta=.038^{**}$), and RN ($\beta=.061^{***}$) predicted SIK ($R^2=.098$). Finally, JA ($\beta=.035^{**}$), JS ($\beta=.050^{***}$), CK ($\beta=.037^{***}$), and RN ($\beta=.042^{**}$) predicted CIK ($R^2=.109$). Significant differences exist between the regression coefficients predicting SOK by RN ($\chi^2=9.385^{***}$, $df=3$). Economics differs from physics, sociology from medicine, and medicine from physics. Finally, significant differences exist between the regression coefficients predicting SIK by JS ($\chi^2=9.976^{***}$, $df=3$), where economics and sociology and between sociology and physics.

The results show that epistemological beliefs affect COR-skills and the beliefs about Internet-related information sources despite a low explained variance which results from the measurement. Additionally, the results indicate the need to separate the effects for some domains and epistemological beliefs, which will be critically discussed in the presentation.

Mehr als ChatGPT? Künstliche Intelligenz als Thema der Erwachsenenbildung – eine computergestützte Angebotsanalyse

Falk Scheidig

Ruhr-Universität Bochum, Deutschland

Theoretischer Hintergrund

Im Zusammenhang mit der zunehmenden Durchdringung vieler Lebensbereiche von Digitalisierungsprozessen gewinnt auch der (mögliche) Einsatz Künstlicher Intelligenz (KI) in privaten wie in beruflichen Kontexten an Bedeutung (Seufert & Handschuh, 2024). Für die Nutzung von KI-Tools wie ChatGPT, aber auch für die kritisch-reflektierte Einschätzung von Limitationen und Risiken von KI sind KI-bezogene Kompetenzen erforderlich (Cousseran et al., 2023; de Witt et al., 2023). Im Sinne lebenslangen Lernens können Erwachsenenbildungsangebote einen Beitrag dazu leisten, KI-bezogene Kompetenzen im Erwachsenenalter zu entwickeln bzw. zu vertiefen.

Es ist bislang nicht erforscht, mit welcher inhaltlichen Ausrichtung KI in Angeboten öffentlicher Erwachsenenbildungseinrichtungen als Thema aufgegriffen wird – also welche Facetten und Nutzungsszenarien von KI im Fokus stehen, ob die Angebote informations- oder anwendungsorientiert sind, ob die Auseinandersetzung mit KI von optimistischen Erwartungen oder von skeptischer Distanz getragen sind. Anknüpfend an dieses Desiderat gilt das Erkenntnisinteresse des vorliegenden Beitrags der inhaltlichen Ausrichtung von Veranstaltungen zu KI an Volkshochschulen, dem bekanntesten und reichweitenstärksten Erwachsenenbildungsanbieter in Deutschland (Deutscher Volkshochschul-Verband, 2023a, 2023b).

Fragestellung

Welche inhaltliche Ausrichtung besitzen Volkshochschulveranstaltungen zu Künstlicher Intelligenz?

Methode

Zur Beantwortung der Fragestellung wurden VHS-Veranstaltungen analysiert, die im Juli 2024 im „VHS-Kursfinder“ gelistet waren. Im „VHS-Kursfinder“ werden ca. 266.000 Veranstaltungen von 452 Volkshochschulen geführt. Über eine Datenbankabfrage beim Deutschen Volkshochschul-Verband, der den „VHS-Kursfinder“ hostet, wurden alle Veranstaltungen mit „KI“ oder „Künstlich* Intelligenz“ im Titel exportiert. Nach einer Dedublierung mehrfach gelisteter Veranstaltungen sowie nach dem Ausschluss von Veranstaltungen ohne KI-Bezug konnten 598 aktuelle Veranstaltungen zu KI an Volkshochschulen ermittelt werden.

Bei der anschließenden Analyse der Titel und Ankündigungstexte der 598 Veranstaltungen zu KI wurden zwei Zugänge kombiniert: Erstens wurde – anknüpfend an die in der Erwachsenenbildungsforschung etablierte Methode der Programmanalyse (Fleige et al., 2019) – eine strukturierende qualitative Inhaltsanalyse (close reading) vorgenommen, um die Volkshochschulveranstaltungen zu KI inhaltlich zu kategorisieren (induktive Kategorienbildung). Zweitens wurden für eine vertiefte Analyse mit Topic Modeling und Sentimentanalyse zwei Verfahren der automatischen Textanalyse (distant reading) verwendet, die in der (Erwachsenen-)Bildungsforschung bislang kaum Verwendung fanden: Während Topic Modeling zur Identifizierung latenter Themencenter im Datensatz zum Einsatz kam (Papilloud & Hinneburg, 2018), wurde mit einer Sentimentanalyse ermittelt, ob die Volkshochschulveranstaltungen zu KI eher positiv oder eher negativ konnotiert sind (Schneijderberg et al., 2022).

Ergebnisse

Das Ergebnis der strukturierenden qualitativen Inhaltsanalysen zeigt, dass unter den Volkshochschulangeboten zu KI viele Angebote einer Anwendungsorientierung folgen, wobei der Fokus anwendungsorientierter Veranstaltungen vor allem auf ChatGPT liegt. Mittels Topic Modeling konnten keine (interpretierbaren) latenten Strukturen in den Veranstaltungsausschreibungen identifiziert werden. Die Sentimentanalyse zeigt, dass in Volkshochschulveranstaltungen zu KI deutlich häufiger positiv besetzte Wörter als negativ besetzte Wörter vorkommen, worin sich die Anwendungsorientierung vieler Angebote spiegelt. Gleichwohl legen sowohl die qualitative Inhaltsanalyse als auch die Sentimentanalyse offen, dass Volkshochschulveranstaltungen zu KI durch die Thematisierung von Grenzen und Gefahren von KI potenziell auch zu einer kritisch-reflektierten Haltung Erwachsener beitragen können.

PS-66: Evidenzorientierung und Wahrnehmung in der Lehrkräftebildung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 169

Chair der Sitzung: Linda Schirle

Evidenzorientierung bei Lehrkräften: Trends aus PISA und die Rolle der Zusammenarbeit von Lehrkräften mit Studierenden in verschiedenen Praktikumsformaten

Linda Schirle, Andreas Just, Christine Sälzer

Universität Stuttgart, Deutschland

Theoretischer Hintergrund

Im Zuge der Professionalisierung der Lehrkräftebildung hat sich die Evidenzorientierung im Lehramtsstudium (1. Phase) als fester Bestandteil des Curriculums etabliert. Auch im Referendariat (2. Phase) und im Bereich der Fortbildungen (3. Phase) spielen veröffentlichte, wissenschaftliche Erkenntnisse zu Schule und Unterricht eine wichtige Rolle. Um Evidenzorientierung zu entwickeln, sind möglichst unmittelbare und kontinuierliche Praxissettings nötig, die einen handelnden Umgang mit der im Studium rezipierten Evidenz erlauben. Eine Chance dafür bieten schulpraktische Ausbildungsphasen – nicht nur für Studierende, sondern auch für betreuende Lehrkräfte. Für diese werden zahlreiche Fortbildungen angeboten, um regelmäßig mit aktueller Evidenz in Berührung zu kommen.

Allerdings besuchen Lehrkräfte in Deutschland Fortbildungen deutlich seltener als ihre Kolleg*innen im Ausland (OECD, 2019; Robert Bosch Stiftung, 2024). Dabei schreiben Lehrkräfte in PISA 2015 und 2018 ($N = 125764$) Fortbildungen eine große Bedeutung für den beruflichen Erfolg und die Zufriedenheit zu (Fütterer et al., 2023). Für die Lehrkraft-Kohorte in PISA 2012 wird die Einschätzung der Relevanz von Forschungsergebnissen jedoch unter anderem durch fehlende Zeit beeinflusst (Thomm et al., 2021). Hier zeigt sich eine immerhin moderate Wertschätzung evidenzbasierter Praxis bei Lehrkräften, die mit grundlegenden Konzepten von Forschungsmethoden und Statistik vertraut sind (Thomm et al., 2012). Da sich die betrachtete Stichprobe auf Mathematiklehrkräfte beschränkt, liegt eine Affinität zur Statistik nahe.

Fragestellung:

Dieser Beitrag klärt zwei Fragen: 1) Inwiefern lässt sich die moderate Akzeptanz evidenzbasierter Lehre aus PISA 2012 auch in der PISA-Lehrkraft-Kohorte 2022 finden und unterscheiden sich die Mathematiklehrkräfte hier von Lehrkräften anderer Fächer? 2) Wirkt sich bei Ausbildungslehrkräften die Zusammenarbeit mit Studierenden auf Einstellungen gegenüber evidenzorientierter Lehre aus?

Methode:

Für Fragestellung (1) werden die Angaben von $N = 2516$ Lehrkräften in Deutschland (in Vollzeitbeschäftigung) bei PISA 2022 betrachtet. Diese werden, differenziert nach Mathematik ($N = 832$) und anderen Fächern ($N = 1684$), mithilfe von χ^2 -Tests und t-Tests ausgewertet. Anhand der wöchentlich investierten Zeit im Lehrkräftealltag werden Gruppen mit k-means clustering gebildet, um die Gruppe der Mathematiklehrkräfte genauer zu beschreiben. Für die Forschungsfrage (2) werden Gymnasiallehrkräfte in Baden-Württemberg zu ihren Einstellungen gegenüber evidenzorientierter Lehre befragt. Das Erhebungsinstrument besteht größtenteils aus validierten Skalen (bspw. Georgiou et al., 2020) – unter anderem die von Thomm et al. (2021) diskutierten Skalen aus PISA 2012. Neben den Einstellungen werden Betreuungsaufgaben für schulpraktische Ausbildungsphasen und weitere Personenangaben erhoben. Darauf basierend werden Gruppen gebildet und Einstellungsunterschiede analysiert.

Ergebnisse & ihre Bedeutung:

Im Vergleich mit Lehrkräften anderer Fächer haben Mathematiklehrkräfte in den vergangenen 12 Monaten etwas seltener berufsbezogene Literatur gelesen ($\chi^2(1) = 5.152$, $p < 0.05$; 75% versus 82%).

Laut Thomm et al. (2021) ist Zeitmangel ein Grund dafür, dass Evidenz als irrelevant eingeschätzt wird. Um dies genauer zu verstehen, haben wir die Tätigkeiten der Mathematiklehrkräfte betrachtet. Es lassen sich drei Lehrkrafttypen ableiten, deren Arbeitsschwerpunkt entweder im Unterrichten ($N = 567$), der Korrekturarbeit ($N = 94$) oder sozialen Interaktionen ($N = 171$) liegt, wobei die Zeit für Unterrichtsplanung nicht variiert. Der signifikante Fisher's Exact Test verdeutlicht, dass die hauptsächlich Korrigierenden (67%) seltener als die Unterrichtsfokussierten (72%) und die Netzwerkenden (88%) berufsbezogene Literatur lesen.

Den größten Effekt auf das Lesen berufsbezogener Literatur finden wir im Zusammenhang mit der Fähigkeit, Schüler*innen beim kritischen Denken zu unterstützen ($m_1 = 3$, $m_2 = 2.82$, $t(1780) = 4.3474$, $p < 0.001$, $g = 0.26$), was sich bei Mathematiklehrkräften noch deutlicher zeigt ($m_1 = 2.97$, $m_2 = 2.72$, $t(239) = 2.4108$, $p < 0.05$, $g = 0.36$) als bei der Gesamtstichprobe.

Die Analysen zur Untersuchung der zweiten Forschungsfrage schließen wir nach der Datenerhebung im Herbst 2024 an und werden sie im Januar präsentieren können. Wir erwarten daraus einen Einblick, inwiefern sich Einstellungen der Lehrkräfte, die bereit sind, eine Praxisbetreuungsaufgabe zu übernehmen, von denjenigen anderer Lehrkräfte unterscheiden.

Auseinandersetzung mit Forschungsliteratur als Weiterbildungsaktivität? Eine Latente Klassenanalyse mit Lehrpersonendaten aus PISA 2022

Franziska Locher

St.Gallen University of Teacher Education, Schweiz

Hintergrund

Angestoßen durch die sogenannte «empirische Wende» (Buchhaas-Birkholz, 2009) wird in den letzten fünfundzwanzig Jahren immer mehr gefordert, dass Evidenz in der Bildungspraxis eine stärkere Rolle spielen muss (Bromme, Prenzel, & Jäger, 2014; Hartmann, Decristan, & Klieme, 2016). Wird im Bildungskontext von Evidenz gesprochen, sind damit bildungswissenschaftliche Daten und Befunde gemeint, die durch vielfältige empirisch ausgerichtete Forschungsansätze (Ergebnisse von Large-Scale Assessments, Interventionsstudien, etc.) gewonnen und vor dem Hintergrund spezifischer Fragestellungen interpretiert werden, und dazu dienen, Vermutungen, Hypothesen oder Theorien zu stützen oder zu widerlegen. Die Berücksichtigung von und Orientierung an Evidenz in der Bildungspraxis soll den Erfolg der Handlungsentscheidung erhöhen und die Handlungsgrundlagen verständlicher machen. So sollen auch mögliche fehlerhafte Erklärungen pädagogischer Situationen korrigiert werden können (Bauer & Kollar, 2023; Rochnia, Schellenbach-Zell, Steckel, & Radisch, 2022). Nach dem heutigen Stand ist dies jedoch eher Wunsch als Realität; die Erwartungen hinsichtlich einer tragenden Rolle von Evidenz im Lehrberuf haben sich nicht erfüllt - die Auseinandersetzung mit Forschungsliteratur stellt kaum eine Wissensressource dar, die Lehrpersonen für ihre Arbeit bzw. zur

Weiterbildung nutzen (Besa, Demski, Gesang, & Hinzke, 2023, für einen Überblick). Dabei stellt sich die Frage, ob Lehrpersonen nur die Auseinandersetzung mit Forschungsliteratur nicht als mögliche Weiterbildungsaktivität betrachten, oder ob sie generell wenig Weiterbildungsaktivitäten nutzen, z.B. aufgrund von zu wenig Zeitressourcen. Eine gemeinsame Betrachtung unterschiedlicher Weiterbildungsaktivitäten kann hier gewinnbringend sein.

Fragestellung

Folgende Forschungsfragen werden überprüft: 1) Lassen sich unterschiedliche Muster in den Weiterbildungsaktivitäten von Lehrpersonen erkennen? 2) Lassen sich Unterschiede im selbsteingeschätzten Einfluss der Weiterbildungen auf die eigene pädagogische Praxis und bei der selbsteingeschätzten professionellen Kompetenz der Lehrpersonen durch unterschiedliche Muster der Weiterbildungsaktivitäten vorhersagen?

Methode

Für die Untersuchung wurde die Substichprobe deutscher Lehrpersonen aus PISA 2022 (Programme for International Student Assessment; OECD, 2023) herangezogen ($N = 2673$; Alter: $M = 44.5$; 59% weiblich). Zur Beantwortung der Fragestellung wurde eine Latente Klassenanalyse mit *Mplus* (Muthen & Muthen, 1998-2017) durchgeführt. Ausgangspunkt für die Klassifizierung der Weiterbildungsaktivitäten sind die Antwortmuster der Lehrpersonen auf folgende Frage: «Haben Sie in den letzten 12 Monaten an folgenden Aktivitäten teilgenommen?». Zur Auswahl standen 14 Aktivitäten (Ordinalskala: 1 = "Ja"; 2 = "Nein"), z.B. «Lesen berufsbezogener Literatur (z.B. Fachzeitschriften, evidenzbasierte Artikel oder Abschlussarbeiten)» oder «Hospitieren an anderen Schulen».

Vorläufige Ergebnisse

In einem ersten Schritt wurde eine Latente Klassenanalyse mit den Weiterbildungsaktivitäten durchgeführt, die nach Vergleich verschiedener Klassenlösungen auf Grundlage verschiedener Modelkennwerte (BIC/AIC, LMRT, Gruppengröße, etc.) sowie inhaltlicher Sinnhaftigkeit der Klassen fünf verschiedene Weiterbildungstypen ergab: 1) «Der digital orientierte Typ» ($n = 508$), 2) «Der inaktive Typ» ($n = 235$), 3) «Der tendenziell formale Typ» ($n = 512$), 4) «Der weiterbildungsfreudige Typ» ($n = 994$) und 5) «Der selbstgesteuert-informelle Typ» ($n = 424$). Zu bemerken ist, dass die Auseinandersetzung mit Forschungsliteratur als Weiterbildungsmöglichkeit bei Typ 3 (83%), 4 (93%) und 5 (93%) sehr stark ausgeprägt ist, etwa vergleichbar mit dem informellen Austausch mit Kolleg:innen zur Optimierung des eigenen Unterrichts (82-95%).

Diskussion und Ausblick

Die persönliche Weiterbildung stellt ein Schlüssel dar, um die eigenen professionellen Kompetenzen zu erweitern und berufliche Herausforderungen meistern zu können. Dass das Lesen von Forschungsliteratur im Sinne der gewünschten Evidenzorientierung in der Bildungspraxis für viele Lehrpersonen eine Möglichkeit zur Weiterbildung darstellt, steht im Gegensatz zu bisherigen Befunden. Bei der Interpretation der Ergebnisse sollte jedoch berücksichtigt werden, dass die Lehrpersonen lediglich mittels «ja» und «nein» angaben, ob sie im letzten Jahr Literatur, wie evidenzbasierte Artikel gelesen haben. Diese Methode liefert nur begrenzt Einblick, z.B. was das tatsächliche Lesematerial anbelangt. Globale retrospektive Selbstberichte sind ausserdem anfällig für Verzerrungen (z.B. Erinnerungsfehler, Stereotype; Fahrenberg, Myrtek, Pawlik, & Perez, 2007), was im Beitrag diskutiert wird. Im nächsten Schritt wird untersucht, wie die fünf Weiterbildungstypen mit dem selbsteingeschätzten Einfluss der Weiterbildungsaktivitäten und den selbsteingeschätzten professionellen Kompetenzen zusammenhängt.

Der Einfluss persuasiver Widerlegungstexte auf die Konzeptions- und Einstellungsänderungen von Lehramtsstudierenden hinsichtlich bildungsbezogener Fehlkonzepte

Isabell Tucholka¹, Madeleine Müller², Bernadette Gold¹

¹Technische Universität Dortmund, Deutschland; ²Universität Erfurt, Deutschland

Theoretischer Hintergrund

Es existieren viele Fehlkonzepte über bildungsbezogene Themen, die dem aktuellen Stand der Forschung widersprechen. Sie sind sowohl bei Lehramtsstudierenden als auch bei professionellem Bildungspersonal weit verbreitet (De Bruyckere et al., 2020). Fehlkonzepte beruhen häufig auf einer zu starken Vereinfachung komplexer Zusammenhänge beim Lehren und Lernen und bieten vermeintlich einfache Antworten auf vielschichtige Problemstellungen (Asberger et al., 2020). Problematisch erscheinen Fehlkonzepte besonders dann, wenn sie zur Grundlage professionellen Handelns und Entscheidens werden und zu ineffektiven Lehr- und Lernstrategien oder zu verschwendeten Ressourcen führen, wie z.B. König et al. (2012) zu Fehlkonzepten über Klassenwiederholung zeigen. Darüber hinaus können Fehlkonzepte hinderlich für das Lernen neuer Inhalte sein (Dole & Niederhauser, 1990).

Daher verfolgen sogenannte Widerlegungstexte das Ziel, beim Lesenden einen konzeptuellen Wandel anzuregen, indem sie erst ein Fehlkonzept ansprechen und es anschließend mit dem wissenschaftlich korrekten Konzept kontrastieren (Sinatra & Broughton, 2011). Widerlegungstexte haben sich in dieser Hinsicht als erfolgreich erwiesen (Zengilowski et al., 2021).

Ein konzeptueller Wandel scheint dabei eng mit veränderten Einstellungen einherzugehen, die wiederum in Bezug auf kontroverse Themen den konzeptuellen Wandel befördern oder behindern können (Sinatra & Seyranian, 2016). Persuasive Texte zielen daher spezifisch auf Einstellungsänderungen, indem sie z.B. die Vor- oder Nachteile eines bestimmten Themas hervorheben (Sinatra et al., 2012; Thacker et al., 2020). Die Ergänzung von Widerlegungstexten um einen entsprechenden persuasiven Textabschnitt kann daher die Auswirkungen des Widerlegungstextes auf die konzeptuellen und einstellungsbezogenen Veränderungen verstärken oder verringern (Thacker et al., 2020).

Zielsetzung & Fragestellung

Die vorliegende Studie überträgt diese Erkenntnisse auf den Bildungsbereich, indem sie in Bezug auf bildungsbezogene Fehlkonzepte folgende Fragestellung verfolgt: Wie beeinflussen Widerlegungstexte mit persuasiven Ergänzungen die Konzeptions- und Einstellungsversänderungen von Lehramtsstudierenden?

Methode

In der experimentellen Studie lasen 116 Lehramtsstudierende (Bachelor 74.1%, 85.3% weiblich, $M_{\text{Alter}}=23.10$; $SD=3.69$) Widerlegungstexte zu zwei Themen (direkte Instruktion; Klassengröße), um die Robustheit möglicher Effekte über verschiedene Thematiken zu prüfen. Die Texte (Textlänge: 304-343 Wörter, Flesch-Kincaid-Lesbarkeitsindex: 44-49) enthielten jeweils eine persuasive Ergänzung, die die Widerlegung des Fehlkonzepts entweder bekräftigte oder abschwächte (Faktor: Texttyp). Die Gruppenzuweisung zu Texttyp und auch Textreihenfolge erfolgte randomisiert. Vor und nach dem Lesen bewerteten die Lehramtsstudierenden wie sehr sie verschiedenen Aussagen zu bildungsbezogenen Fehlkonzepten, u.a. zu direkter Instruktion

($\alpha=.90$) und Klassengröße ($\alpha=.75$), zustimmten (QUEBEC, Asberger et al., 2020). Außerdem gaben sie zu beiden Messzeitpunkten ihre Einstellung zu diesen Themen anhand von vier Items an ($.68 < \alpha < .77$; adaptiert nach Heddy et al., 2017).

Die Daten wurden themenspezifisch mittels gemischter ANOVAs analysiert, bei denen die Zustimmung zu Fehlkonzepten und die entsprechende Einstellung zum Thema als abhängige Variablen zu zwei Messzeitpunkten (Innersubjektfaktor) und der Texttyp (begründigende/abschwächende Textergänzung) sowie die Textreihenfolge als unabhängige Variablen (Zwischensubjektfaktoren) verwendet wurden.

Ergebnisse und Diskussion

Die Zustimmung zu bildungsbezogenen Fehlkonzepten verringerte sich in allen Bedingungen (Direkte Instruktion: $F(1,112)=75.74$, $p<.001$, $\eta_p^2=.403$; Klassengröße: $F(1,112)=114.38$, $p<.001$, $\eta_p^2=.505$). Einstellungsänderungen scheinen dagegen themenspezifisch zu sein (Direkte Instruktion: $F(1,112)=0.01$, $p=.942$; Klassengröße: $F(1,112)=45.89$, $p<.001$, $\eta_p^2=.291$, s. auch Trevors, 2022 zu Impfung/Immigration). Signifikante Unterschiede zwischen den zwei Texttypen gab es nicht (Direkte Instruktion: Zustimmung, $F(1,112)=0.09$, $p=.762$; Einstellung, $F(1,112)=0.39$, $p=.534$; Klassengröße: Zustimmung, $F(1,112)=2.26$, $p=.108$; Einstellung, $F(1,112)=0.02$, $p=.907$).

Beide Widerlegungstexte gingen mit einer Verringerung bildungsbezogener Fehlkonzepte einher. Trotz fehlender Kontrollbedingung für eine eindeutige Zuordnung zum Widerlegungstext, wäre dieser Effekt konsistent zu bestehenden Studien (Zengilowski et al., 2021). Zusätzliche Informationen, welche die Widerlegung bekräftigen oder abschwächen, hatten keinen differenzierenden Einfluss auf konzeptuelle oder auch einstellungsbezogene Veränderungen, wie es bei breiter diskutierten gesellschaftspolitischen Themen eher der Fall ist (Berweiger et al., 2023; Thacker et al., 2022). Da sich die meisten Untersuchungen zu Widerlegungstexten auf letztere konzentrieren, gibt die vorliegende Studie einen Hinweis auf Spezifika von Bildungsthemen, die es zu berücksichtigen gilt, um bildungsbezogenen Fehlkonzepten bei Lehramtsstudierenden entgegenzuwirken und damit verknüpften Einstellungen entsprechend begegnen zu können.

Die grafische Darstellung von Lernverläufen beeinflusst die eingenommene Bezugsnorm bei der Interpretation

Erika Lunowa¹, Sarah Bez^{1,2}, Martin J. Tomasik³, Samuel Merk¹

¹PH Karlsruhe; ²Universität Tübingen; ³Universität Zürich

Theoretischer Hintergrund

Formatives Assessment wird auf der Unterrichtsebene als eine lernförderliche Form von datengestützten Entscheidungen (sog. *data-based decision making*) angesehen (Schildkamp, 2019; Souvignier & Hasselhorn, 2018). Die Lernwirksamkeit des formativen Assessments gilt als empirisch hinreichend belegt (Lee et al., 2020; Schütze et al., 2018; Xuan et al., 2022). Für die Lernwirksamkeit scheint die Passung zwischen wiederkehrender Diagnostik und den darauf basierenden Unterrichtsmaßnahmen relevant zu sein (Souvignier & Hasselhorn, 2018). Die dafür erforderliche Leistungsbeurteilung kann jedoch nur durch einen Vergleich mit einem Referenzwert (Heckhausen, 1974) vorgenommen werden. Dabei werden typischerweise drei Bezugsnormen unterschieden: die „kriteriale Bezugsnorm“, die „soziale Bezugsnorm“ sowie die „individuelle Bezugsnorm“ (Rheinberg, 2014). Bei der Leistungsbewertung der Ergebnisse von formativem Assessment gilt die Anwendung der individuellen und kriterialen Bezugsnorm als vorteilhaft (Haupt-Mukrowsky, 2022; Sliwka et al., 2023), denn die kriteriale Bezugsnorm basiert auf im Voraus festgelegten (aufeinander abgestimmten) Kompetenzen, welche im Lerngegenstand selbst begründet liegen (Rheinberg, 2014) und der Vergleich der Leistungen der Lernenden mit dieser Norm legt ihren Lernstand in Bezug zum angestrebten Lernziel offen. Die Anwendung der individuellen Bezugsnorm wird als motivationsförderlich angesehen (Dickhäuser et al., 2017; Janson et al., 2022), da (selbst kleine) individuelle Lernfortschritte sichtbar werden (Dickhäuser & Rheinberg, 2003). Bisherige Studien deuten auf eine vermehrte Anwendung der sozialen Bezugsnorm bei der Leistungsbewertung hin (Bez et al., 2021; Holder & Kessels, 2018; Lintorf & Buch, 2021; Wilbert & Gerdes, 2009; Wilbert & Grünke, 2010), wobei die Bezugsnormorientierung (Rheinberg, 1980) als allgemeine Neigung einer Lehrkraft, eine bestimmte Bezugsnorm anzuwenden, sowohl als recht stabil als auch mit flexiblen Anteilen beschrieben wird (Lintorf & Buch, 2021; Wilbert & Gerdes, 2009). Grundsätzlich besteht bei Datenvisualisierungen die Möglichkeit, unterschiedliche Aspekte hervorzuheben, obwohl die zugrundeliegenden Daten identisch sind (Ratner et al., 2019). Geht man davon aus, dass die Bezugsnormorientierung neben globalen auch situationsspezifische Anteile hat, stellt sich die Frage, ob Lehrkräfte durch gezielte Visualisierungen darin unterstützt werden können, die individuelle und kriteriale Bezugsnorm vermehrt zur Leistungsbeurteilung anzuwenden. Dies soll in der vorliegenden explanativen Studie anhand folgender drei Forschungsfragen untersucht werden:

- (1) Welche Bezugsnormen werden bei der Interpretation von Lernverläufen wie häufig eingenommen?
- (2) Wird die eingenommene Bezugsnorm von der grafischen Darstellung der Lernverläufe beeinflusst?
- (3) Ist die globale Bezugsnormorientierung für die in einer konkreten Bewertungssituation eingenommene Bezugsnorm prädiktiv?

Methode

In einer (within-person) randomisierten Studie mit $N = 223$ Lehramtsstudierenden erhielten diese nacheinander ökologisch valide Lernverlaufsdaten zweier Klassen, die sich im Leistungsniveau und –trend unterschieden.

In einem ersten Schritt wurden die Lernverläufe einer Klasse in einer vergleichsweise neutralen Darstellung präsentiert, bei der keine Bezugsnorm besonders hervorgehoben wurde. Die Studierenden bewerteten die Lernverläufe, wobei ihnen Informationen zu den Kompetenzstufen bzw. zur Vergleichsgruppe in Textform zur Verfügung standen.

Im zweiten Schritt bewerteten die Studierenden die Lernverläufe der zweiten Klasse. Sie sahen eine von drei informationsgleichen Darstellungen, bei denen jeweils eine Bezugsnorm gezielt hervorgehoben wurde.

Schließlich bearbeiteten die Studierenden zur Erfassung der globalen Bezugsnormorientierung die Kleine Beurteilungsaufgabe (Rheinberg, 1980).

Ergebnisse und ihre Bedeutung

Die Ergebnisse der Studie deuten darauf hin, dass Studierende bei einer (vergleichsweise) neutralen Darstellung durchschnittlich mehr als eine Bezugsnorm ($M = 1,24$) einnahmen, wobei vor allem die individuelle und/oder kriteriale Bezugsnorm angewendet wurde (Forschungsfrage 1). Eine bezugsnormhervorhebende Darstellung scheint einen mittleren bis starken Effekt auf die eingenommene Bezugsnorm zu haben (Forschungsfrage 2: $OR = 4.25$ [individuelle Bezugsnorm]; 4.75 [kriteriale Bezugsnorm]; 11.17 [soziale Bezugsnorm]). Die globale Bezugsnormorientierung einer Lehrkraft korreliert nur schwach oder vernachlässigbar mit der eingenommenen Bezugsnorm bei der Interpretation (Forschungsfrage 3: $-.31 \leq r_{RB} \leq .27$). Insgesamt legen die Ergebnisse

nahe, dass die gezielte Hervorhebung der individuellen und kriterialen Bezugsnorm bei der Visualisierung von Lernverläufen Lehrkräfte darin unterstützen könnte, diese zur Leistungsbewertung vermehrt einzunehmen.

PS-67: Emotionale und soziale Dimensionen der (Lehrkräfte-)Professionalisierung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 142

Chair der Sitzung: Anett Wolgast

Ein digitales Training zur Förderung sozial-emotionaler Kompetenzen – Ergebnisse einer Interventionsstudie zu Medizinischen Fachangestellten in der Ausbildung

Julia Raecke, Tanja Tschöpe, Sebastian Konheiser, Annalisa Schnitzler, Markus Peters

Bundesinstitut für Berufsbildung, Deutschland

Theoretischer Hintergrund

Soziale und emotionale Kompetenzen (SEK) spielen in fast allen Bereichen des Lebens eine zentrale Rolle, um einen wertschätzenden Umgang und somit ein friedliches Zusammenleben zu erreichen. Vor allem in der Vielzahl an Berufen, bei denen die Zusammenarbeit in Teams sowie Interaktionen mit Patientinnen und Patienten oder Kundinnen und Kunden im Vordergrund stehen, sind Erwerb und Anwendung von SEK essentiell. Trotzdem wird der Förderung von SEK in der beruflichen Ausbildung noch zu wenig Aufmerksamkeit geschenkt. Dies gilt auch für Medizinische Fachangestellte (MFA), welche täglich anspruchsvolle soziale Situationen bewältigen müssen (z.B. fordernde Patientinnen/Patienten, Konflikte im Team; Vu-Eickmann et al., 2018). Auch ihr Wohlbefinden und ihre Gesundheit hängen von einem positiven Miteinander im Berufsalltag ab.

In einer Anforderungsanalyse identifizierten Tschöpe und Monnier (2016) Emotionsregulation, Perspektivenkoordination und Kommunikationsstrategien als besonders relevante SEK für MFA. Mithilfe eines eigens entwickelten Situational Judgement Tests (SJT) haben sie Entwicklungsbedarfe in allen drei SEK bei MFA in der Ausbildung festgestellt. Um diese Bedarfe zu adressieren, haben wir ein digitales Training zur Förderung von SEK von MFA in der Ausbildung entwickelt.

Forschungsfrage

Kann ein digitales Training einen Beitrag zur Förderung von SEK von MFA in der Ausbildung leisten?

Methode

In einer Interventionsstudie mit Trainings- und Kontrollgruppe und einem Prätest-Posttest-Follow-Up-Design überprüfen wir die Wirksamkeit des Trainings. Das Training umfasst drei 1-stündige digitale Lerneinheiten, die jeweils eine der drei Kompetenzen durch Vermittlung relevanten Wissens, Videobeispiele und Übungen fördern sollen. Die Inhalte des Trainings orientieren sich an psychologischen Theorien und Trainingsansätzen zu den jeweiligen Kompetenzen (z.B. von McRae & Gross (2020), Selman (2003) und Brand-Hörsting (2019), vgl. Peters et al., 2023). Auch hier nutzen wir selbst entwickelte SJT zur Messung der SEK vor und nach dem Training. Von Herbst 2023 bis Frühjahr 2024 nahmen ca. 500 MFA im zweiten Ausbildungsjahr an der Studie teil, ca. 280 davon an allen drei Messzeitpunkten. Wir erwarten eine signifikante Erhöhung aller drei SEK in der Trainingsgruppe (N ≈ 150) im Vergleich zur Kontrollgruppe (N ≈ 130). Die Trainingseffekte werden mithilfe von ANOVAs mit Messwiederholung geprüft.

Ergebnisse und ihre Bedeutung

Die Trainingsgruppe zeigt im Posttest sowie im Follow-Up signifikant höhere Mittelwerte als die Kontrollgruppe. Während die SEK der Trainingsgruppe über die Zeit steigen, bleiben die SEK der Kontrollgruppe gleich bzw. fallen sogar tendenziell. Dies deutet darauf hin, dass SEK mithilfe eines relativ kurzen und flexibel einsetzbaren digitalen Trainings erfolgreich gefördert werden können. Die Effekte erstrecken sich jedoch nicht auf alle drei untersuchten Kompetenzen gleichermaßen und auch in Bezug auf verschiedene Situationstypen im SJT gibt es Effekunterschiede. Detaillierte Ergebnisse und deren Implikationen werden im Vortrag vorgestellt. Zusammengefasst könnten MFA von einer Integration digitaler SEK-Trainings in ihre Ausbildung in Berufsschulen und/oder Betrieben durchaus profitieren. Außerdem sind die grundsätzlichen Inhalte vermutlich auch hilfreich für Auszubildende anderer Berufe und auch übertragbar auf andere Bildungskontexte. Für noch stärkere und langfristige Effekte wäre es vermutlich wichtig, dass Auszubildende das Erlernte ausprobieren und besprechen können. Zukünftige Forschung könnte längere Follow-Up-Phasen einplanen und digitale mit traditionellen Trainingsmethoden vergleichen oder kombinieren, um die besten Ansätze (bzgl. Wirksamkeit und Kosteneffizienz) zur Förderung von SEK zu identifizieren.

Kooperationsformen zwischen Lehrpersonen und ihre Effekte auf die wahrgenommene soziale Unterstützung, wahrgenommene kognitive Empathie der Schulleitung und wahrgenommene eigene Beanspruchung

Anett Wolgast¹, Manuela Keller-Schneider²

¹University of Applied Sciences FHM Campus Hanover, Deutschland; ²Zurich University of Teacher Education, University of Marburg

Theoretischer Hintergrund: Lehrpersonen können von unterschiedlichen Formen der Kooperation mit ihren Kolleginnen und Kollegen profitieren und von ihnen soziale Unterstützung erfahren, um eine hohe Unterrichtsqualität aufrechtzuerhalten (z. B. Steinert et al. 2006; Terhart & Klieme 2006). Diese Erfahrung kann dazu beitragen, berufliche Anforderungen und außerschulische Verpflichtungen zu erfüllen und negative Beanspruchung zu minimieren (z. B. Keller-Schneider, 2018; Klusmann et al., 2008). Fünf empirisch ermittelte Formen der Lehrerverkooperation sind (1) Informationsaustausch, (2) Aufgabenteilung, (3) gemeinsame Planungsaktivitäten, (4) Diskussion über pädagogische Fragen und (5) gemeinsame Verantwortung (z. B. Fussangel & Gräsel, 2012; Gräsel et al., 2006; Keller-Schneider & Albisser, 2013). Ob die Kooperationsformen von Lehrpersonen jedoch später ihre wahrgenommene soziale Unterstützung, wahrgenommene kognitive Empathie der Schulleitung und wahrgenommene Beanspruchung vorhersagen, wurde bisher selten untersucht.

Die wahrgenommene Beanspruchung bei Lehrpersonen bezieht sich theoretisch auf die negativen Folgen, die sich aus der chronischen Belastung durch stressige Anforderungen ergeben können (z. B. Hobfoll, 1989; Rudow, 1999). Belastungsreaktionen können sich in verschiedenen Formen äußern, wie Burnout, Müdigkeit, Angstzustände, Depressionen und Kopf- oder Rückenschmerzen. Das Stress-Belastungs-Ressourcen-Konzept beschreibt, dass die negativen Belastungsreaktionen von Lehrpersonen durch die Erhöhung und optimale Kombination individueller und kollektiver Ressourcen gemildert werden können (Rudow, 1999). Lehrpersonen, die ein hohes Ausmaß an Belastung, aber auch eine kognitive Empathie ihrer Schulleitung mit ihrer Situation wahrnehmen, können sich geschützt fühlen und ihr Grundbedürfnis nach sozialer Eingebundenheit erfüllt sehen (Justice, 2018; Ryan & Deci, 2020). Somit könnte die wahrgenommene kognitive Empathie von der Schulleitung mit einem niedrigen Ausmaß an wahrgenommener negativer Beanspruchung bei Lehrpersonen zusammenhängen.

Fragestellung: In Anlehnung an das Stress-Belastungs-Ressourcen-Konzept (Hobfoll, 1989, Rudow, 1999) war die Frage, ob die fünf Kooperationsformen zwischen Lehrpersonen ein geringes Ausmaß an wahrgenommener negativer Beanspruchung, ein hohes Ausmaß an wahrgenommener sozialer Unterstützung und ein hohes Ausmaß wahrgenommener kognitiver Empathie von der Schulleitung determinieren.

Methode: In dieser Längsschnittstudie füllten 456 Grundschullehrpersonen (68.4% weiblich; Lebensalter: $M = 43.15$, $SD = 11.95$) den gleichen Fragebogen an zwei Messzeitpunkten aus, zu Beginn eines Schuljahres und 19 Monate später am Ende des nächsten Schuljahres. Wir untersuchten mittels Strukturgleichungsmodellierung ($\chi^2[371] = 401.88$, $p = .13$, CFI = .98, RMSEA = .03, 95 % CI [$< .01$, .05], SRMR = .09), ob die Kooperationsformen mindestens eine der folgenden drei Kriterien bei den Lehrpersonen determinierte (unter Kontrolle dieser Konstrukte an Messzeitpunkt 1, der Lehrselbstwirksamkeitserwartung und des Geschlechts): ‚wahrgenommene soziale Unterstützung‘ und ‚wahrgenommene kognitive Empathie von der Schulleitung‘ und ‚wahrgenommene Beanspruchung‘.

Ergebnisse: Der (1) Informationsaustausch als Kooperationsform bei Lehrpersonen determinierte ihre wahrgenommene soziale Unterstützung ($Beta = .57$, $p < .05$) und die wahrgenommene kognitive Empathie von der Schulleitung ($Beta = .31$, $p < .05$) 19 Monate später. Die (4) Diskussion pädagogischer Fragen als Kooperationsform bei Lehrpersonen determinierte die wahrgenommene kognitive Empathie von der Schulleitung ($Beta = .29$, $p < .05$) 19 Monate später. Zudem waren die wahrgenommene soziale Unterstützung (latent $r = -.23$, $p < .05$) und wahrgenommene kognitive Empathie von der Schulleitung (latent $r = -.10$, $p < .05$) statistisch negativ mit der Beanspruchung der Lehrpersonen assoziiert. Die Strukturgleichungsmodellierung ergab für die Lehrpersonen 85 % aufgeklärte Varianz im Kriterium ‚wahrgenommene soziale Unterstützung‘, 53 % aufgeklärte Varianz im Kriterium ‚wahrgenommene kognitive Empathie von der Schulleitung‘ und 94 % aufgeklärte Varianz im Kriterium ‚wahrgenommene Beanspruchung‘. Diese Ergebnisse können genutzt werden, um Schulleiterinnen und Schulleiter dabei zu unterstützen und zu ermutigen, die Zusammenarbeit zwischen Lehrpersonen zu fördern, um negative Beanspruchung abzumildern und eine hohe Unterrichtsqualität zu erhalten.

Unveiling Emotion Regulation Strategies Among Teachers: Insights from Multilevel Latent Profile Analysis

Tanja Bross¹, Anne Frenzel², Ulrike Nett¹

¹Universität Augsburg, Deutschland; ²Ludwig-Maximilians-Universität München, Deutschland

Theoretical Background

Teachers encounter various situations in their daily professional lives, with individual perceptions and interpretations of these situations differing from day to day (Schmidt et al., 2017). Depending on their perceptions of these situations, teachers experience different emotions, with joy and anger being the most frequently experienced (Burić & Frenzel, 2019; see Frenzel, 2014 for an overview). To regulate them, teachers use various strategies (Taxer & Gross, 2018). Present research on emotion regulation shows that individuals are likely to use multiple emotion regulation strategies simultaneously (Ford et al., 2019). Moreover, they adjust their emotion regulation based on the specific situation and the emotions they are experiencing (Rottweiler et al., 2023; Rottweiler et al., 2018). However, little is known about the patterns of strategies that are beneficial for teachers' emotion regulation.

Research Question

The aim of the present study was to exploratively examine the following research questions:

1. What patterns of emotion regulation strategies are identified at the situational level and what profiles are identified at the individual level?
2. How are the different situational patterns adapted to the valence of the situation (positive or negative), the context (teaching or non-teaching), and specific emotions (joy and anger)?

Method

A total of 165 teachers (Level 2) participated in a diary study for sixteen days ($M_{Age} = 43.26$ years, $SD = 11.01$, 83.2% female). Each day, they received a short questionnaire via email between 6 PM and 10 PM. Participants selected one emotionally positive and one negative situation, which determined the valence of the situation (0: negative, 1: positive) resulting in $n = 2136$ measurements (Level 1). All subsequent items were recorded separately for both situations. Joy and anger (Goetz et al., 2013; Rottweiler et al., 2018) as well as eight emotion regulation strategies (Rottweiler et al., 2018) were measured on a five-point Likert scale.

We performed multilevel latent profile analyses (ML-LPA) in Latent GOLD Version 6.0 (Vermunt & Magidson, 2021). After identifying the final ML-LPA model, the covariates situation (positive or negative), context (teaching or non-teaching), and emotions (joy and anger) were included.

Results

A five-pattern solution on Level 1 and a four-profile solution on Level 2 revealed the best model fit. Pattern 1 "Average ER" comprised 38.5% ($n = 823$) of the situations and was characterized by usage of all emotion regulations strategies around the mean. Pattern 2 was labelled "Slight ER + avoidance" (25.1%, $n = 536$), as refocusing and avoidance were used above the mean. Pattern 3 "Low ER" comprised 15.2% ($n = 324$) and was characterized by usage of strategies below the mean. Pattern 4 was characterized by "Suppression & Refocusing" usage (11.0%, $n = 234$) and pattern 5 by "Expression & Social Support" usage (10.2%, $n = 219$).

At the person level (Level 2), the first profile predominantly uses the "Slight ER + Avoidance" pattern (26.8%, $n = 44$). The second profile mainly exhibits "Average ER" (26.5%, $n = 44$). The third profile combines "Average ER" with "Low ER" (23.6%, $n = 39$). The fourth profile is characterized by a diverse use of ER patterns (23.1%, $n = 38$).

Including covariates revealed that teachers used different patterns between positive and negative contexts, that these patterns vary between teaching versus non-teaching situations and in relation to the experience of joy and anger.

Discussion

Our findings enhance the theoretical understanding of teachers' emotion regulation patterns, highlighting that strategies are both situation- and emotion-specific. Effective emotion regulation by teachers is crucial for the well-being of both teachers and students. Therefore, training in emotion regulation should be integrated into university-level educational programs. Further implications will be discussed at the conference.

Coaching von angehenden Lehrkräften: Veränderbarkeit von Kompetenzüberzeugungen und antizipierten Unterrichtsemotionen

Marina Elena Pfeifer, Anne Christiane Frenzel

Universität München (LMU), Deutschland

Coaching erweitert die traditionelle universitäre Fachvermittlung durch eine personenorientierte Begleitung zur Entwicklung von Selbstreflexion, individuellen Kompetenzen und Kompetenzüberzeugungen, welche einen maßgeblichen Einfluss auf die Unterrichtsqualität, das individuelle Wohlbefinden, das Burnout-Risiko und das berufliche Ausscheiden haben (Burić & Frenzel, 2023; Goetz et al., 2023; Schoeps et al., 2019). Entsprechend konnten empirische Studien positive Effekte von Coaching auch für Lehrkräfte zeigen (Wiemer, 2012; Yang et al., 2022). Das Programm „Souverän im Klassenzimmer“ ist auf das Coaching von Lehramtsstudierenden ausgerichtet und umfasst einen zweitägigen Workshop sowie zwei Kleingruppen-Coachings. Die Einheiten wurden nach Erhalt einer zusätzlichen Coaching-Ausbildung von erfahrenen Lehrkräften angeleitet und hatten zum Ziel, Studierende durch Selbsterfahrung, Reflexion und Feedback in ihrer professionsbezogenen Entwicklung zu begleiten und dadurch personale und soziale Kompetenzen aufzubauen (Fiedler & Frenzel, 2022). Die vorliegende Studie untersucht die Wirksamkeit dieses Coaching-Programms in Hinblick auf die Förderung des Selbstkonzepts, der Selbstwirksamkeit sowie der antizipierten Unterrichtsemotionen bei Lehramtsstudierenden.

Die Interventionsstudie folgte einem Prä-Post-Kontrollgruppendesign, für das alle Teilnehmer:innen ($N = 390$, $M_{\text{age}} = 22.76$, $SD_{\text{age}} = 3.96$) im Abstand von etwa sechs Wochen befragt wurden. Die Interventionsgruppe ($N = 176$, $M_{\text{age}} = 23.52$, $SD_{\text{age}} = 3.68$) nahm in der Zwischenzeit am Coaching-Programm teil, während die Kontrollgruppe ($N = 214$, $M_{\text{age}} = 22.14$, $SD_{\text{age}} = 4.07$) entweder monetär entschädigt wurde oder nach dem Posttest am Coaching-Programm teilnahm. Per Selbstbericht wurden Facetten des *beruflichen Selbstkonzepts* (ERBSE-L; Retelsdorf et al., 2014; d.h.: Fach, Diagnostik, Beratung, Erziehung), der *Selbstwirksamkeit* (NTSES; Skaalvik & Skaalvik, 2007; d.h.: Unterrichten, Anpassung an individuelle Bedürfnisse, Klasse motivieren, Disziplin wahren, Kooperation mit Eltern/Kollegium, Veränderungen bewältigen) – ergänzt um *Beziehungsaufbau* (Fiedler & Frenzel, 2022) sowie *antizipierte Unterrichtsemotionen* (Freude, Angst und Ärger; TES; Frenzel et al., 2016) erfasst. Veränderungen zwischen den beiden Messzeitpunkten wurden mittels ANOVA mit Messwiederholung (Interventions- vs. Kontrollbedingung) analysiert.

Die Ergebnisse zeigen erwartungskonform signifikante Interaktionseffekte kleiner bis mittlerer Größe zwischen dem Innersubjektfaktor „Zeit“ und dem Zwischensubjektfaktor „Bedingung“ bei allen Subskalen des Selbstkonzepts, der Selbstwirksamkeit sowie im Freude- und Angsterleben. Post-hoc-Tests bestätigen, dass diese Interaktionseffekte auf günstige Prä-Post-Veränderungen in allen Facetten des Selbstkonzepts (Fach: $F = 7.31$, $p < .01$, $\eta_p^2 = .02$; Diagnostik: $F = 13.76$, $p < .01$, $\eta_p^2 = .03$; Beratung: $F = 5.78$, $p < .05$, $\eta_p^2 = .01$; Erziehung: $F = 16.28$, $p < .01$, $\eta_p^2 = .04$) und der Selbstwirksamkeit (Unterrichten: $F = 14.29$, $p < .01$, $\eta_p^2 = .04$; Anpassen an individuelle Bedürfnisse: $F = 30.44$, $p < .01$, $\eta_p^2 = .08$; Klasse motivieren: $F = 29.92$, $p < .01$, $\eta_p^2 = .07$; Disziplin wahren: $F = 30.44$, $p < .01$, $\eta_p^2 = .07$; Kooperation mit Eltern/Kollegium: $F = 43.67$, $p < .01$, $\eta_p^2 = .10$; Veränderungen bewältigen: $F = 13.34$, $p < .01$, $\eta_p^2 = .03$; Beziehungsaufbau: $F = 22.90$, $p < .01$, $\eta_p^2 = .06$) sowie im Freude- sowie Angsterleben (Freude: $F = 9.18$, $p < .01$, $\eta_p^2 = .02$; Angst: $F = 8.37$, $p < .01$, $\eta_p^2 = .02$; Ärger: $F = 0.51$, $p = .48$, $\eta_p^2 = .00$) in der Interventionsbedingung bei ausbleibenden Veränderungen in der Kontrollbedingung zurückzuführen sind.

Die Ergebnisse unserer Studie dokumentieren die Veränderbarkeit von Kompetenzüberzeugungen und antizipierten Unterrichtsemotionen durch das evaluierte Coaching-Programm. Lehramtsstudierende, die derartige Wahlpflichtprogramme wählen, können durch das Coaching-Angebot in ihren Kompetenzüberzeugungen und antizipierten Unterrichtsemotionen gestärkt werden. Dies bekräftigt bestehende Forschung zur Wirksamkeit von Coaching-Programmen für Lehramtsstudierende (Mok et al., 2023) und unterstreicht trotz des relativ höheren finanziellen Aufwands das Potential solcher Angebote in der Ausbildung von Lehrkräften.

PS-68: Feedback in Lern- und Leistungsprozessen: Wahrnehmung, Wirkung und Gestaltung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: SO 322

Chair der Sitzung: Nikolai Zinke

Wirksamkeit von automatisiertem Feedback: Eine experimentelle Untersuchung der Effekte auf die Schreibleistung

Fabian T.C. Schmidt¹, Jennifer Meyer², Thorben Jansen², Johanna Fleckenstein¹

¹Universität Hildesheim, Germany; ²IPN, Kiel

Theoretischer Hintergrund: Empirische Studien zeigen, dass Feedback ein wirksames Mittel zur Verbesserung der Schreibfähigkeiten ist (Graham et al., 2023; Graham & Harris, 2017; Graham & Perin, 2007). In den letzten Jahren hat die Entwicklung von KI-basierten automatisierten Schreibbewertungssystemen (Automated Writing Evaluation, AWE) es ermöglicht, individualisiertes Feedback effizient zu implementieren. AWE Systeme können den Schreibprozess durch formatives Feedback unterstützen, jedoch sind die Effekte heterogen und insbesondere bei jüngeren Schüler*innen noch unzureichend untersucht (Fleckenstein et al., 2023, 2024; Ngo et al., 2022; Scherer et al., 2024; Zhai & Ma, 2022). Generell ist Feedback am effektivsten, wenn es nicht nur evaluative oder korrigierende Informationen über vergangene Leistungen enthält, sondern auch elaborierte didaktische Hinweise wie Lernende ihre Fähigkeiten verbessern können beinhaltet (Mertens et al., 2022; Narciss & Huth, 2004; Van der Kleij et al., 2015; Wagner et al., 2024). Fragestellung: Diese Experimentalstudie untersucht die Auswirkungen von AWE Feedback mit unterschiedlichem Grad der Individualisierung und Elaboriertheit auf die Überarbeitungsleistung und eine anschließende Transferaufgabe von Schüler*innen der Sekundarstufe I für eine Schreibaufgabe im Fach Englisch. Methode: In einem Experiment wurden N = 388 Schüler*innen der Klassen 7-9 in vier Versuchsgruppen randomisiert eingeteilt. Die Schüler*innen erhielten eine Schreibaufgabe im Fach Englisch, die sie am Computer bearbeiten sollten. Die Aufgaben bestand darin, eine formelle E-Mail zu schreiben, in der Informationen angefordert werden. Die Gruppen erhielten unterschiedliche Formen von AWE Feedback zu ihren Texten: (1) eine Kontrollgruppe ohne Feedback, (2) eine Gruppe, die nur ein Bewertungsrastr mit Schreibzielen erhielt, (3) eine Gruppe mit evaluativem Feedback, die zusätzlich individuelle Leistungsinfos zu den Zielen erhielt, und (4) eine Gruppe mit elaboriertem Feedback, die zusätzlich didaktische Hinweise und Beispiele erhielt. Die Wirksamkeit dieser Interventionen wurde anhand von fünf Beurteilungskriterien bei der Überarbeitungsleistung und bei einer Transferleistung untersucht, wobei Mehrebenen-Regressionsanalysen zur Auswertung verwendet wurden. Ergebnisse: Die Ergebnisse zeigen, dass bei Berücksichtigung der Eingangsleistung, elaboriertes Feedback einen signifikant positiven Effekt auf den Überarbeitungsprozess hatte ($\beta = .24, p < .05$). Im Vergleich dazu hatten allgemeine Ziele ($\beta = .11, p = .17$) und evaluatives Feedback ($\beta = .01, p = .87$) keine signifikanten Auswirkungen auf die Überarbeitungsleistung. Diese Befunde unterstreichen die Bedeutung der Art des Feedbacks: Während ein rein evaluatives Feedback keine Wirkung zeigt, ist ein elaboriertes Feedback mit didaktischen Hinweisen für die Überarbeitung wirkungsvoll. Dennoch bleibt die Frage offen, ob die Nutzung von AWE Feedback langfristig zur Verbesserung der Schreibfähigkeiten beiträgt, da die Effekte auf die Transferleistung in keiner Gruppe signifikant waren. Diskussion: Diese Studie bietet eine systematische Untersuchung der Wirksamkeit von KI-basierter Unterstützung bei jüngeren Schüler*innen mit geringem Vorwissen. Die Befunde zeigen, dass auch diese Zielgruppe von einem elaborierten AWE Feedback profitiert. Ein weiterer zentraler Befund ist, dass elaboriertes AWE Feedback die Überarbeitungsprozesse unterstützt, allerdings nicht die Leistung in einer anschließenden Schreibaufgabe verbessert. Weitere Forschung, die eine wiederholte KI-gestützte Lerngelegenheit bietet, könnte Aufschluss über mögliche verborgene Transferleistungen bieten.

How Feedback Complexity Impacts Its Evaluation: The Role of Performance, Narcissism and Perfectionism

Nikolai Zinke¹, Lenski Sina¹, Mertens Ute², Kuklick Livia², Walpurger Verena³, Hannes Schröter^{1,3}, Martin Merkt¹

¹German Institute for Adult Education, Bonn, Germany; ²Leibniz Institute for Science and Mathematics Education, Kiel Germany; ³FernUniversität in Hagen, Hagen, Germany

Theoretical Background

Recent studies have demonstrated the superiority of complex (elaborated) feedback over simple feedback (Mertens et al., 2022; Hattie & Timperley, 2007). Additionally, preliminary empirical evidence suggests that elaborated feedback may have better affective-motivational impact and was rated as more useful among low performing students after incorrect responses (Kuklick & Lindner, 2023). However, it was primarily shown in low-stake contexts. Our study aims to explore these research questions under higher stake setting. Hence, we assumed that feedback evaluation is elevated after elaborative (vs. simple) feedback (H1a) and, at low performance levels, reduced after simple (vs. elaborated) feedback (H1b). Performance-critical situations are salient for personality traits like narcissism and perfectionism (Pincus et al., 2009). Narcissism is a heterogeneous construct encompassing grandiose narcissism, characterized by arrogance and self-enhancement, and vulnerable narcissism, marked by hypersensitivity and defensiveness (Pincus & Lukowitsky, 2010). Grandiose narcissists respond positively to feedback to maintain self-esteem, whereas vulnerable narcissists may exhibit avoidance behaviours in response to negative feedback (Morf & Rhodewalt, 2001; Pincus et al., 2009). Therefore, we hypothesized that narcissists would exhibit lower feedback evaluation when performance level is low (H2a). Perfectionists view feedback as crucial for improving their performance (Stoeber & Otto, 2006), and thus we postulated a positive relationship between perfectionism and feedback evaluation (H2b).

Method

Undergraduate psychology students (N = 131; 108 female) participated in the survey. They were enrolled in an introductory course and were requested to take a corresponding e-mock exam. The course content was represented by 15 multiple-choice questions, with 5 alternative answers each. Automated feedback was provided after each answer. Different levels of feedback complexity were randomly assigned to three groups: a) only the score achieved in the whole task (KR), b) additionally with a note indicating which answer alternatives were solved correctly (KCR), or c) additionally with a comprehensive explanation (EF). Prior to the mock exam, the students completed questionnaires assessing narcissism (Brailovskaia et al., 2019; Schoenleber et al., 2015), perfectionism (Feher et al., 2020), and control variables. Feedback evaluation (perceived usefulness and positive affective reaction; Attali & van der Kleij, 2017) was measured after each mock exam task and after mock exam.

Results

Elaborative feedback significantly enhanced perceived usefulness: $b = 0.18$, $p < .001$. This effect was also moderated by performance level ($b = -0.03$, $p < .001$), suggesting at low performance levels the perceived usefulness in KR/KCR conditions (vs. EF) was lower. No (interaction) effects for positive affective reaction occurred ($ps \geq .408$). Hence, H1a and H1b were confirmed. Vulnerable narcissism showed a stronger negative correlation only with positive affective reaction at lower performance levels (vs. higher performance levels): $b = 0.02$, $p = .046$, but not with perceived usefulness ($p = .220$). No interactions were found among grandiose narcissists (all $ps \geq .249$). Perfectionism was negatively related only to positive affective reaction ($b = -0.22$, $p = .050$). Hence, only H2a was partly confirmed.

Discussion

Our findings corroborate prior research indicating that elaborative feedback is more beneficial (e.g. Mertens et al., 2022) as it contains informational content, aligning with the formative-corrective character of the learning process, particularly when it refers to students' mistakes (Kuklick & Lindner, 2023; Shute, 2008). Regarding narcissism, our finding supports recent contributions that vulnerable narcissists respond adversely to feedback at low performance levels (Kjærvik & Bushman, 2021). In contrast, neither grandiose narcissism nor perfectionism was found to significantly affect feedback evaluation. However, exploratory analyses revealed that in elaborative conditions at higher performance levels, perceived usefulness was elevated among (grandiose) narcissists and perfectionists. Despite the need for further research, elaborative feedback could address feedback-sensitive personality traits (Suh et al., 2019; Vaillancourt, 2013).

Give me more: Wie Narzissmus und Perfektionismus mit der Wahl und Wahrnehmung von Feedback zusammenhängen

Sina Lenski¹, Nikolai Zinke¹, Livia Kuklick², Ute Mertens², Verena Walpurger³, Hannes Schröter^{1,3}, Martin Merkt¹

¹Deutsches Institut für Erwachsenenbildung e.V., Deutschland; ²Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN), Kiel, Deutschland; ³FernUniversität in Hagen, Hagen, Deutschland

Theoretischer Hintergrund

Insbesondere elaboriertes Feedback kann das Lernen und die Leistung von Studierenden in formativen Prüfungen fördern (Mertens et al., 2022). Motivation und Leistung sind zwei Faktoren, die das Feedback-Abrufverhalten beeinflussen können (Leenknecht & Carless, 2023). Studierende wählen jedoch nicht immer die effektivsten Lernstrategien, selbst wenn sie deren Vorteile kennen (Rea et al., 2022). Bisherige Studien deuten darauf hin, dass Persönlichkeitsmerkmale eine wichtige Rolle bei der Wahl von Feedback spielen können (Ashford et al., 2016). In der vorliegenden Studie wird untersucht, wie neben der Leistung in Testaufgaben auch Persönlichkeitsmerkmale wie vulnerabler Narzissmus und Perfektionismus die Wahl von elaboriertem gegenüber einfacherem Feedback (z. B. *Knowledge of Response* [KR] oder *Knowledge of Correct Response* [KCR]) und dessen wahrgenommene Nützlichkeit beeinflussen. Vulnerabler Narzissmus ist gekennzeichnet durch hohe Empfindlichkeit gegenüber Kritik, Scham und Angst vor Ablehnung (Atlas & Them, 2008). Perfektionismus beschreibt das Streben nach Leistung und hohen persönlichen Standards (Feher et al., 2020).

H1: Die Leistung hängt negativ mit der Wahl von elaboriertem Feedback (**a**) und der wahrgenommenen Nützlichkeit des Feedbacks (**b**) zusammen.

H2: Vulnerabler Narzissmus ist negativ mit der Wahl von elaboriertem Feedback (**a**) und Perfektionismus ist positiv mit der wahrgenommenen Nützlichkeit des Feedbacks (**b**) assoziiert.

Methode

In dieser Online-Studie wurde eine Probeklausur für das Bachelormodul „Einführung in die Psychologie“ (15 Multiple-Choice-Fragen) angeboten. Die Teilnehmenden ($N = 63$ Psychologiestudierende) erhielten nach jeder Antwort ein einfaches Feedback über die erreichte Punktzahl (max. 5 Punkte; KR). Anschließend konnten sie zusätzlich ein Feedback darüber, welche Antwort(en) korrekt waren (KCR) und elaboriertes Feedback (EF) wählen. Die Variablen Narzissmus (Schoenleber et al., 2015), Perfektionismus (Feher et al., 2020) sowie die wahrgenommene Nützlichkeit (Strijbos et al., 2010) des Feedbacks wurden erfasst. Zur Analyse wurden gemischte Modelle verwendet.

Ergebnisse

Die Ergebnisse stützen die Annahme **H1a**, dass Studierende mit höherer Leistung weniger häufig elaboriertes Feedback wählen, $OR=0.36$, 95% $CI=[0.27; 0.48]$, $z=-5.95$, $p<.001$. Wie angenommen (**H1b**), zeigt sich auch, dass Studierende mit höherer Leistung das Feedback als weniger nützlich empfinden, $b=-0.08$, $t(641.33)=-2.35$, $p=.019$. Entgegen der Annahme (**H2a**) war vulnerabler Narzissmus positiv mit der Wahl von elaboriertem Feedback assoziiert, $OR=1.06$, 95% $CI=[1.01; 1.11]$, $p=.035$. Explorative Analysen deuten darauf hin, dass dieser Zusammenhang durch das Leistungsniveau moderiert wird, $OR=0.95$, 95% $CI=[0.92; 0.97]$, $z=-3.75$, $p<.001$. Der positive Zusammenhang trat nur bei niedrigem Leistungsniveau auf, während er sich bei hohem Leistungsniveau ($\geq M+SD$) aufhob ($p=.283$). Anders, als in **H2b** angenommen, hatte Perfektionismus keinen Einfluss auf die wahrgenommene Nützlichkeit von Feedback, $b=0.006$, $t(63.0)=0.58$, $p=.562$. Bei Personen mit ausgeprägtem Perfektionismus war die Präferenz für elaboriertes Feedback zudem unabhängig von ihrer Leistung $b<0.001$, $t(905.41)=0.045$, $p=0.964$.

Diskussion

Die Ergebnisse zeigen, dass Studierende mit höherer Klausurleistung weniger häufig elaboriertes Feedback wählen und dieses als weniger nützlich empfinden. Auch Lipnevich und Smith (2009) fanden heraus, dass leistungsstarke Studierende weniger detailliertes Feedback benötigen. Im Gegensatz zu unserer Annahme zeigte sich, dass vulnerabel narzisstische Studierende ein höheres Bedürfnis nach detailliertem Feedback haben. Pincus et al. (2009) zeigten, dass vulnerabel narzisstische Personen stärker auf externe Validierung angewiesen sind, um Unsicherheiten zu kompensieren. Es ist zu vermuten, dass der Wunsch nach Unsicherheitsreduktion die gegenteilige Tendenz zur Vermeidung von Feedback überlagern könnte (Atlas & Them, 2008). Dieser Prozess könnte sich insbesondere bei niedrigeren Leistungsniveaus verstärken, wie die explorative Moderationsanalyse nahelegt. Perfektionismus zeigte anders als angenommen, keinen signifikanten Einfluss auf die wahrgenommene Nützlichkeit von Feedback, was im Widerspruch zu den Ergebnissen von Besser et al. (2004) steht. Gegenstand zukünftiger Forschung sollte es sein, die zugrunde liegenden Mechanismen und Bedingungen genauer zu untersuchen, die das Feedback-Abrufverhalten und die Präferenzen von Studierenden beeinflussen.

Insgesamt liefert diese Studie in einem authentischen Lernkontext wertvolle Einblicke in das Feedback-Abrufverhalten von Studierenden und unterstreicht die Bedeutung individueller Unterschiede für die Gestaltung effektiver Feedbackstrategien.

Interaktive E-Lernlektionen für Erstsemesterstudierende in der Allgemeinen Chemie - Feedback und Assessment in selbstregulierten Lernphasen

Florian Trauten, Carolin Eitemüller, Maik Walpuski

Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Leistungsprobleme stellen den häufigsten Grund für Studienabbruch dar (Heublein et al. 2017). Die Abbruchquote liegt in der Chemie mittlerweile bei 54 % (Startsemester 2016/17; Fachwechsel oder Abbruch; Heublein, Hutzsch & Schmelzer, 2022). Diese Defizite müssen bereits in der frühen Studieneingangsphase behoben werden, da sich andernfalls das Risiko für Studienabbruch deutlich erhöht (Heublein et al., 2017). Der erhöhte Zeitaufwand zum Aufholen von Wissenslücken (Schwedler, 2017) stellt Studierenden vor Probleme. Rückmeldung zum Lernerfolg oder zu Fehlern muss anders als in der Schule aktiv eingefordert werden. Die Kontaktaufnahme zu anderen Studierenden und zu Lehrenden wird zu einem entscheidenden Faktor für den Verbleib an der Hochschule (Sarcelletti & Müller, 2011). Gerade zu Beginn des Studiums fehlen diese Kontakte aber.

Vor diesem Hintergrund haben sich unter anderem automatisiertes fehlerspezifisches Feedback (Trauten, Eitemüller & Walpuski, 2023) und Assessments (Freasier, Collins & Newitt, 2003) bereits als vielversprechende Möglichkeiten erwiesen, Selbstlernphasen im Studium zu optimieren. In interaktiven E-Lernlektionen lassen sich beide Aspekte realisieren. Zudem lassen sich die Inhalte multimedial aufbereitet und interaktiv darstellen.

Die Art der Einbindung in die Lehrveranstaltung kann dabei die Nutzung und Akzeptanz beeinflussen. Während Karay, Reiss und Schaubert (2020) zeigen konnten, dass selbstbestimmte Nutzung zu einer häufigeren Verwendung und einer höheren Akzeptanz von Assessments führen kann, berichten Trentepohl et al. (2022), dass sich deutlich bessere Lernergebnisse im Rahmen eines Hochschulkurses (Bauingenieurwesen) erzielen lassen, wenn mangelnde Selbstregulation der Studierenden durch regelmäßig von der Lehrperson eingeforderte verbindliche Abgaben kompensiert wird. Denn ein hohes Maß an Selbstbestimmung im Lernprozess stellt gerade für Lernende mit geringem Vorwissen, das häufig mit einer niedrigen Selbstregulationsfähigkeit hinsichtlich des Lernprozesses verbunden ist, ein Problem dar (Winters, Greene & Costich, 2008). Oft unterschätzen sie den Umfang der zu lernenden Inhalte (Thibodeaux, Deutsch, Kitsantas & Winsler, 2017) bei gleichzeitigem Überschätzen der eigenen Fähigkeiten (Bruin, Kok, Lobbestael & Grip, 2017).

Ziel und Fragestellung

Ziel des Projekts ist es daher, zu überprüfen, unter welchen Bedingungen die Einbindung von automatisiertem fehlerspezifischem Feedback und Assessment in Form von interaktiven E-Lernlektionen in die Lehre den Studienerfolg steigern und die Akzeptanz seitens der Studierenden hinsichtlich des Angebots erhöhen kann. Die zugrundeliegende Frage lautet:

FF: Wie wirken sich interaktive E-Lernlektionen mit automatisiertem Feedback in Abhängigkeit von den Implementationsbedingungen (selbstbestimmt vs. verpflichtend) auf (a) die Akzeptanz hinsichtlich der E-Lernlektionen und (b) den Fachwissenserwerb von Erstsemesterstudierenden in der Allgemeinen Chemie aus?

Methode und Design

Diese Fragestellung wurde im Wintersemester 2023/2024 anhand einer Intervention im Zwei-Gruppen Prä-Post-Follow-up-Design an $N = 136$ Studierenden überprüft. Hierzu wurden die Lehramtsstudierenden im Fach Chemie (IG₁: Verpflichtende Bearbeitung) mit Studierenden der Fachwahl Bachelor Chemie und Bachelor Water Science (IG₂: Eigenverantwortliche Bearbeitung) im Rahmen des Moduls „Allgemeine Chemie“ verglichen, das sie – bei gleichen Modulzielen – getrennt voneinander besuchen. Zur Beantwortung der Teilfrage (a) wurde in Anlehnung an Kärchner et al. (2023) ein Fragebogen zur Evaluation digitaler Lernumgebungen mit zwei Subskalen (empfundene Nützlichkeit, Akzeptanz) entwickelt. Beide Subskalen weisen eine gute interne Konsistenz (Cronbachs- $\alpha \geq .82$) auf. Die Teilfrage (b) wird durch den Einsatz eines Fachwissenstest (adaptiert nach Freyer, 2013 und nach Averbek, 2020) vor und nach der verpflichtenden Lernphase sowie am Ende des Wintersemesters untersucht.

Erste Ergebnisse

Die Datenlage zeigt, dass die Lernlektionen fast ausschließlich und sehr umfänglich von IG₂ genutzt wurden. Beide Interventionsgruppen besuchten die reguläre Übung sowie die Vorlesung. Dabei zeigt sich ein grundsätzlicher Vorteil im Klausurerfolg für die Bearbeitung der Lerneinheiten. Im Rahmen einer hierarchischen Regression wurden mehrere Prädiktoren miteingeschlossen. Das Modell mit der besten Varianzaufklärung ($F(4,37) = 11.01, p < .001, R^2 = .494$) weist darauf hin, dass besonders vorwissensschwache Studierende vom Lernen mit den Lerneinheiten profitieren. Die Fragebögen zeigen jeweils ein zufriedenstellendes Ergebnis hinsichtlich Akzeptanz und empfundener Nützlichkeit der Lerneinheiten.

PS-69: Digitale Lehrmethoden und ihre Rolle im Lernprozess: Neue Ansätze für die Fachbildung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: SO 318

Chair der Sitzung: Hanna Meiners

Effekte emotionsanregender Gestaltung in kaufmännischen digitalen Aufgaben – Ergebnisse einer Experience-Sampling Studie

Hanna Meiners

Georg-August-Universität Göttingen, Deutschland

Mit der voranschreitenden Etablierung digitaler Technologien in der Arbeitswelt resultiert ebenfalls eine Verbreitung digitaler Medien in beruflichen Bildungskontexten. Insbesondere bei der Konzeption Aufgabensituationen aus dem kaufmännischen Berufsalltag bietet digitale Medien Potentiale einer authentischen, anspruchsvollen und das Interesse fördernden Aufgabengestaltung (z. B. durch die Einbindung von Anwendungssoftware).

Forschungsseitig zeigt sich, dass über die Gestaltung von Lehr-Lern-Materialien eine positive Wirkung auf die emotionalen und kognitiven Facetten erreicht werden kann. Gestaltungselemente des *Emotional Designs* können dabei gezielt eingesetzt werden, um Emotionen zu induzieren, wobei empirische Befunde zu Effekten noch kein einheitliches Bild zeigen (Brom et al., 2018; Zhou et al., 2019; Wong & Adesope, 2021). Zu differenzieren sind emotionsinduzierende Gestaltungselemente nach Informations- und Interaktionsgestaltung (z. B. Plass & Schwartz, 2014). Die Informationsgestaltung fokussiert visuelle Elemente wie helle Farben. Ergänzend dazu steht die Interaktionsgestaltung durch beispielsweise Storytelling. Ferner ergeben sich daraus für digitale Aufgaben der kaufmännischen Berufsausbildung Potentiale, Emotionen gezielt über Designelemente in digitalen Aufgaben zu steuern, um die Informationsverarbeitung im Aufgabenbearbeitungsprozess zu unterstützen (siehe ICALM, Plass & Kaplan, 2016).

In einer Experience-Sampling-Studie wurden digitale Aufgaben zu kaufmännischen Problemstellungen anlehnend an Situationen des beruflichen Alltags eingesetzt. Zentrale Zielsetzung dieser Studie ist es, die empfundenen Emotionen bei der Aufgabenbearbeitung zu beleuchten und zu untersuchen, welche Effekte innerhalb unterschiedlich gestalteter Aufgaben empirisch nachweisbar sind. Dabei wird der Frage nachgegangen, inwieweit sich die postulierten Gestaltungseffekte auf die emotionale Aktiviertheit sowie die kognitive Belastung im Bearbeitungsprozesses von digitalen kaufmännischen Aufgaben replizieren lassen. Kaufmännische Auszubildende ($N = 518$) wurden dabei zufällig einer Experimentalgruppe (emotionsanregende Aufgaben) ($N_{EG} = 279$) oder Kontrollgruppe (neutral gestaltete Aufgaben) ($N_{KG} = 239$) zugeordnet. Die Aufgaben der Gruppen unterscheiden sich einerseits in der Informationsgestaltung. Statt einer schwarz-grauen Farbgebung sowie Einbindung eckiger Formen in den Aufgaben der Kontrollgruppe, wurden warme helle Farben sowie runde Formen in Aufgaben der Experimentalgruppe eingesetzt. Zusätzlich wurden innerhalb der emotionsanregenden Aufgaben Anwendungssoftware (z. B. Excel) sowie Audio- und Videovignetten als Elemente des Storytellings eingebunden, um Autonomie- und Kontrollerleben zu erhöhen (Astleitner & Leutner, 2000).

Während des Bearbeitungsprozesses dieser digitalen Aufgaben haben die Auszubildenden wiederkehrend jeweils nach drei Aufgabenabschnitten (Aufgabeneinstieg, Aufgabenbearbeitung und Lösungseingabe) ihre emotionale Aktiviertheit mittels der PANAVA-KS (Schallberger, 2005) sowie ihre wahrgenommene kognitive Belastung anhand von Kurzskalen zur kognitiven Belastung (in Anlehnung an Klepsch et al., 2017) eingeschätzt. Erste deskriptive Analysen zeigen, dass durchschnittlich über den gesamten Aufgabenbearbeitungsprozess sowohl das Valenzempfinden ($M_{EG} = 3.50$, $SD_{EG} = 0.95$; $M_{KG} = 3.35$, $SD_{KG} = 0.97$) als auch die positive Aktivierung ($M_{EG} = 2.91$, $SD_{EG} = 0.84$; $M_{KG} = 2.85$, $SD_{KG} = 0.84$) in der Experimentalgruppe höher ausgeprägt ist als in der Kontrollgruppe. Hingegen verhält sich das Empfinden der negativen Aktivierung entgegengesetzt ($M_{EG} = 2.94$, $SD_{EG} = 0.89$; $M_{KG} = 3.04$, $SD_{KG} = 0.90$). Die partielle Betrachtung von Aufgabenabschnitten zeigt einen signifikanten Unterschied zwischen den Gruppen im Valenzempfinden (t-Test: $t(496) = 2.29$, $p = 0.023$) nach der Aufgabenbearbeitung, der auch unter Kontrolle des Empfindens der positiven Aktivierung sowie Valenz im zuvor gegangenen Aufgabeneinstieg existent ist (ANCOVA: $F(1; 494) = 4.49$, $p = 0.035$). Folglich fühlen sich kaufmännische Auszubildende bei der Bearbeitung emotionsanregender Aufgaben glücklicher und zufriedener. Vertiefende aufgabenspezifische statistische Analysen werden durchgeführt, um Erklärungsfaktoren und Mediationseffekte bezüglich der wahrgenommenen kognitiven Belastung und anderen Bedingungsvariablen (z. B. intrinsische Motivation) empirisch nachzuweisen. Die Ergebnisse werden im Hinblick auf mögliche Potentiale für die Gestaltung von digitalen Aufgaben im Lern- und Arbeitskontext der kaufmännischen Berufsausbildung diskutiert.

Wie wirksam sind curriculare Bausteine mit Unterrichtsvideos und Simulationsübungen im Hinblick auf das digitalisierungsbezogene fachdidaktische Wissen von Lehramtsstudierenden und Lernergebnisse ihrer Schülerinnen und Schüler?

Nicoletta Bürger, Tobias Kantorski, Mira Wittenberg, Katrin Hauenschild, Barbara Schmidt-Thieme, Anke Meisert, Jürgen Menthe, Christof Wecker

Universität Hildesheim, Deutschland

Theoretischer Hintergrund

Eine Herausforderung für Lehrkräfte stellt der Einsatz digitaler Technologien zur Unterstützung des Lernens von Schülerinnen und Schülern dar (u. a. Fraillon et al., 2020). Für die Förderung entsprechender Kompetenzen stellt der Ansatz der *practice-based teacher education* in seinem *learning cycle* mit der Modellierung von Kernpraktiken des Unterrichtens oder Simulationsübungen (*rehearsals*) potentiell wirksame Komponenten bereit (Grossman, 2018), die sich auch auf das Unterrichten mit digitalen Technologien anwenden lassen. Bislang liegen jedoch kaum Interventionsstudien zur Wirksamkeit dieses Ansatzes vor.

Fragestellung

Untersucht wurde, 1) inwieweit curriculare Bausteine mit den genannten Komponenten digitalisierungsbezogenes fachdidaktisches Wissen und weitere Merkmale von Lehramtsstudierenden beeinflussen, 2) inwieweit digitalisierungsbezogenes fachdidaktisches Wissen und weitere Merkmale von Lehramtsstudierenden mit Lernergebnissen von ihnen unterrichteter Schülerinnen und Schülern zusammenhängen und 3) inwieweit sich praxisorientierte curriculare Bausteine – vermittelt durch die genannten Merkmale von Lehramtsstudierenden – auf die Lernergebnisse ihrer Schülerinnen und Schülern auswirken.

Methodik

Rahmenbedingungen der Untersuchung. Die Untersuchung fand eingebettet in Vorbereitungs- und Begleitseminare eines Schulpraktikums statt. Der curriculare Baustein umfasste drei Seminarsitzungen, in denen die Studierenden lernen sollten, Schülerinnen und Schülern Kompetenzen in der digitalen Welt zu vermitteln bzw. sie mittels digitaler Technologien beim Erwerb fachlichen Wissens zu unterstützen. Die Studierenden planten eine Unterrichtsstunde mit dem jeweiligen Unterrichtsziel und führten diese anschließend im Praktikum durch.

Design. Die Untersuchung folgte einem Pre-Post-Kontrollgruppenplan mit dem *praxisbezogenen* vs. *traditionellen textbasierten curricularen Baustein* als unabhängiger Variable, die zwischen Seminargruppen manipuliert wurde. Die Studierenden wurden den Bedingungen willkürlich zugewiesen.

Manipulation der unabhängigen Variable. In der Bedingung mit dem *praxisorientierten curricularen Baustein* erarbeiteten sich die Lehramtsstudierenden das jeweilige digitalisierungsbezogene fachdidaktische Wissen mit Videoausschnitten einer modellhaften Unterrichtssequenz und erprobten ihre selbst entwickelte Unterrichtsstunde in einer rollenspielbasierten Simulationsübung. In der Bedingung mit dem *traditionellen textbasierten curricularen Bausteinen* wurde das digitalisierungsbezogene fachdidaktische Wissen anhand verschiedener Grundlagentexte erarbeitet, während sich die Simulationsübung im Seminar auf eine andere, für die vorliegende Untersuchung nicht relevante Unterrichtsstunde, bezog.

Messung von Variablen. Vor und nach der Manipulation im curricularen Baustein bearbeiten die Lehramtsstudierenden Tests und Fragebögen, mit denen das digitalisierungsbezogene fachdidaktische Wissen sowie einstellungs- und motivationsbezogene Lehrkräftemerkmale erfasst wurden. Der im Praktikum durchgeführte Unterrichtsversuch der Lehramtsstudierenden wurde videographiert. Der Lernerfolg der dabei unterrichteten Schülerinnen und Schüler wurde anhand von Tests zum Lernerfolg hinsichtlich des jeweiligen Unterrichtsziels und Fragebögen für Interesse und Selbstkonzept vor und nach der Unterrichtsstunde erfasst. Für den Vergleich über die Themen und Fächer hinweg wurden die Werte zum Lernerfolg anhand der Daten aus den Pretests der jeweiligen Kontrollbedingungen z-standardisiert.

Stichprobe. Von 136 Lehramtsstudierenden liegen Daten aus Tests und Fragebögen vor, sowie bei 90 von ihnen Daten von den von ihnen unterrichteten 923 Schülerinnen und Schülern.

Ergebnisse und Diskussion

(1) In beiden Bedingungen waren das digitalisierungsbezogene fachdidaktische Wissen und verschiedene einstellungs- und motivationsbezogene Merkmale nach dem curricularen Baustein höher ausgeprägt als davor. Während das digitalisierungsbezogene fachdidaktische Wissen nach dem praxisbasierten curricularen Baustein statistisch signifikant höher war als nach dem traditionellen textbasierten curricularen Baustein ($F = 4.946$, $p = .028$, part. $\eta^2 = .040$), gab es bezüglich der übrigen untersuchten Merkmale keine statistisch signifikanten Unterschiede.

(2) Der Enthusiasmus der Lehramtsstudierenden wies statistisch signifikante, allerdings negative Zusammenhänge mit dem Lernerfolg hinsichtlich des jeweiligen Unterrichtsziels ($\beta = -.599$, $SE = .117$, $p < .001$) als auch mit dem Interesse ($\beta = -.733$, $SE = .181$, $p < .001$) und dem Selbstkonzept ($\beta = -.658$, $SE = .223$, $p = .003$) der Schülerinnen und Schüler auf.

(3) Der praxisbezogene curriculare Baustein hatte keinen statistisch signifikanten Effekt auf die Lernergebnisse der Schülerinnen und Schüler; insgesamt konnte kein statistisch signifikanter Mediationseffekt identifiziert werden.

Die Ergebnisse stellen angenommene Wirkungszusammenhänge in Frage und verdeutlichen den dringenden Bedarf an weiterer Forschung sowohl zur Rolle des Wissens von Lehrpersonen für die Lernförderlichkeit ihres Unterrichts als auch zur Effektivität der einzelnen Komponenten des *learning cycle*.

Erklärt erhöhtes Zeitinvestment erhöhten Lernerfolg? Das Lernen mit Videos als Komponente von Flipped Classroom.

Gesa Bintz, Charlotte Kellermeyer, Louis Frye, Stephan Dutke

Universität Münster, Deutschland

Flipped Classroom findet in der universitären Lehre zunehmend Anwendung. Zudem konnte in zahlreichen Überblicks- und Einzeluntersuchungen erhöhter Lernerfolg im Vergleich zu konventioneller Lehre dargestellt werden (Cheng, Ritzhaupt & Antonenko, 2019). Im Rahmen von Flipped Classroom wird häufig selbstgesteuertes Lernen anhand von Lernvideos zur primären Vermittlung neuer Inhalte eingesetzt. Erst danach werden diese Inhalte in Lehrveranstaltungen vertieft. Verglichen mit konventioneller Lehre erhöhen sich mit der Verschiebung des Erstkontakts zu Lernmaterial im Flipped Classroom von in class zu out of class die Möglichkeiten den Fortschritt des Lerninhalts auf die eigene Verarbeitungskapazität anzupassen (Abeysekera & Dawson, 2015). Lernvideos bieten Funktionen wie Pausieren und wiederholtes Ansehen, die im konventionellen Vortrag durch Lehrende nicht zur Verfügung stehen (Noetel et al., 2021). Es fehlen über die überzeugenden Theorien hinaus empirisch belegte Erklärungen für lernförderliche Effekte Flipped Classroom (Kapur et al., 2022). Bereits durch minimale Veränderungen der ersten Phase von Flipped Classroom, der Verschiebung von in zu out of class, kann Lernerfolg gefördert werden (Bintz, Barenberg & Dutke, unveröffentlichtes Manuskript). Demnach erscheint es zentral die einzelnen Komponenten und ihre Wirkung zu verstehen, um Flipped Classroom Lehre optimal einzusetzen. Die vorliegende Untersuchung betrachtet deshalb die selbstgesteuerte Aneignung neuer Lerninhalte mit Lernvideos. In einer Online-Befragung haben Studierende ($N = 180$), ein 13-minütiges Lernvideo zu dem lernpsychologischen Thema evidenzbasierte Lernstrategien bearbeitet. Lernerfolg wurde zu zwei Messzeitpunkten anhand von confidence-weighted-true-false Items erfasst. Eine Teilstichprobe ($N = 51$) wurde zusätzlich zur kognitiven Belastung bei der Bearbeitung des Lernvideos befragt. Angenommen wurde ein Zusammenhang zwischen Intensität der Bearbeitung des Lernvideos und Lernerfolg. Durchschnittlich gaben Studierende an, das Video zwischen zwei und drei Mal pausiert zu haben ($M = 2,59$; $SD = 1,8$). Es wurde eine durchschnittliche Bearbeitungszeit von 20 Minuten ($SD = 13,85$) gemessen. Die Intensität der Bearbeitung des Lernvideos korrelierte mit dem kognitiven Lernerfolg. Für den Einsatz von Pausen bestand ein geringer ($p = .16$, $p < .001$), für die Bearbeitungszeit ein moderater ($p = .38$, $p < .001$) Zusammenhang zum kurzfristigen Lernerfolg. In einer multiplen Regressionsanalyse sagen die Prädiktoren (Konstante, Wiedergabedauer, Anzahl an Pausen) das Kriterium Anteil korrekter Antworten voraus, $F(2, 133) = 6,75$, $p < .01$. Das Modell hat mit einem $R^2 = 0,092$ (korr. $R^2 = 0,078$) eine geringe Anpassungsgüte. Sowohl für kurz- ($p = .32$, $p = .02$, $N = 51$) als auch langfristigen ($p = .49$, $p = .005$, $N = 31$) metakognitiven Lernerfolg wurden moderate Zusammenhänge zur lernförderlichen kognitiven Belastung nachgewiesen. Die Regressionsanalyse ergab eine signifikante Vorhersage des Kriteriums richtig und sichere Antworten zu Messzeitpunkt 2 durch den Prädiktor lernförderliche kognitive Belastung $F(1, 24) = 11,08$, $p < .01$. Das Modell weist mit $R^2 = 0,316$ (Korr. $R^2 = 0,29$) eine hohe Anpassungsgüte auf. Die Ergebnisse eröffnen in Hinblick auf die im Lernvideo gegebenen Kontrollmöglichkeiten sowie die damit einhergehend ermöglichte Verwendung kognitiver Kapazität auf die Erarbeitung der Lerninhalte neue Erkenntnisse. Einerseits geht ein höherer Lernerfolg in der vorliegenden Stichprobe eher mit längerer Wiedergabe des Videos, als mit der selbstberichteten Verwendung

der Pausenfunktion einher. Andererseits scheint die erlebte lernförderliche kognitive Belastung, zwar mit höherer Konfidenz in die eigenen Antworten, dabei aber nicht zwingend mit präziserer Differenzierung zwischen korrekten und nicht-korrekten Antworten einherzugehen.

Analyse der Auswirkungen einer forschungsbasierten Lernumgebung auf die Kompetenzentwicklung von Schüler:innen im Fach Wirtschaft in der Sekundarstufe

Victoria Vochatzer, Malte Ring, Taiga Brahm

Eberhard Karls Universität Tübingen, Deutschland

Einführung und theoretischer Hintergrund

Zahlreiche Ethikskandale haben die Diskussion über die Notwendigkeit von Wirtschaftsethik als integralen Bestandteil der schulischen Ausbildung intensiviert (Heiduk, 2019). Dabei spielt auch die Digitalisierung eine immer größer werdende Rolle (Du & Xie, 2021). Um die Kompetenzen von Lernenden nachhaltig zu fördern, bieten Ansätze des forschenden Lernens, bei denen Schüler:innen sich aktiv an der Entwicklung von Wissen beteiligen, einen guten Ansatzpunkt (Huber, 2014). Für diese konnten bereits positive Effekte auf das Engagement und die Lernergebnisse (Meitriana et al., 2021), die Reflexionsfähigkeiten (Mutlu, 2020) sowie das konzeptionelle Verständnis und Erklärniveau (Almuntasheri et al., 2016) von Schüler:innen gezeigt werden. Dabei wird einer digitalen Umsetzung eine Verbesserung des Erfolgs zugeschrieben (De Jong et al., 2014; Hofhues, 2019), z.B. im Hinblick auf die Analysefähigkeit (Bónus et al., 2024).

Insgesamt liefert die bisherige Forschung erste Hinweise auf die Wirksamkeit von digital gestütztem, forschendem Lernen, allerdings hauptsächlich konzentriert auf naturwissenschaftliche Fächer. (Quasi)-experimentelle Studien, die die Wirksamkeit entsprechender Ansätze im sozialwissenschaftlichen Kontext untersuchen, liegen nach unserem Wissen nicht vor. Offen bleibt, ob sich die Mechanismen, die hinter forschendem Lernen stehen, auf sozialwissenschaftliche, genauer noch, wirtschaftsethische Kontexte übertragen lassen. Die vorliegende Arbeit untersucht daher die Wirkung des Einsatzes einer auf forschendem Lernen basierenden, digitalen Lernplattform im Spannungsfeld Wirtschaft – Ethik – Digitalisierung. Die Lernplattform – entwickelt sowohl für Lehrpersonen als auch für Schüler:innen – enthält verschiedene thematische Module, die wiederum unterschiedliche Medien zur Bearbeitung des Themenfeldes anbieten (z.B. Texte, Grafiken, Videos, Podcasts, Aufgaben etc.).

Im Zentrum der Studie steht die Frage, wie sich die Nutzung der Lernplattform auf die ökonomischen und wirtschaftsethischen Einstellungen, Kompetenzen und das Interesse von Schüler:innen der achten Klasse von allgemeinbildenden Schulen auswirkt.

Methodische Vorgehensweise

Um die Auswirkungen der Implementation der digitalen Lernplattform untersuchen zu können, werden Pre- und Posttests mit Experimental- und Kontrollgruppen auf Klassenebene durchgeführt. Da sich Lehrkräfte und Schulleitungen für das Programm freiwillig melden konnten, ist ein quasi-experimentelles Design notwendig. Die Experimentalklassen nutzten die Plattform im Fach Wirtschaft, Berufs- und Studienorientierung über die Dauer eines Schulhalbjahres im Schuljahr 23/24. Die Kontrollgruppe hatte im gleichen Zeitraum Regelunterricht. Über beide Schulhalbjahre haben n= 588 Schülerinnen und Schüler teilgenommen, davon n = 346 in der Interventions- und n = 242 in der Kontrollgruppe.

Die abhängigen Variablen wurden mit standardisierten Skalen erfasst und umfassen Einstellungen zur Wirtschaft (Oberrauch & Seeber, 2021), ökonomischen Kompetenz (Kaiser et al., 2020), ökonomisches Wissen im Bereich Wirtschaftsethik sowie wirtschaftliches Interesse (Oberrauch & Seeber, 2021) und digitale Kompetenz (Masur et al., 2017). Als Kontrollvariablen wurden neben demografischen Informationen auch kognitive Fähigkeiten (Heller et al., 1998), Einstellungen zu Geld (Barry & Breuer, 2012) und Wettbewerb (Oberrauch & Seeber, 2021), Persönlichkeitsmerkmale (Gerlitz & Schupp, 2005), Präferenzen (Falk et al., 2016), Nachhaltigkeitsbewusstsein (Gericke et al., 2019) und Testmotivation erhoben.

Vorläufige Ergebnisse und Diskussion

Vorläufige Ergebnisse aus dem ersten Schulhalbjahr (n = 125) zeigen keine signifikanten Unterschiede der beiden Gruppen am Ende des ersten Schulhalbjahres im Hinblick auf die untersuchten abhängigen Variablen (Kompetenz, Wissen und Interesse). Eine mögliche methodische Erklärung ist die fehlende Power aufgrund der kleinen Stichprobengröße und die damit verknüpfte unterkomplexe statistische Analyse (Gruppenvergleiche im Posttest). Da durch die Erhebung im zweiten Schulhalbjahr das Sample deutlich größer wird, soll der Effekt der Intervention mithilfe des Differenz-von-Differenzen-Ansatzes und unter Berücksichtigung der Mehrebenen-Struktur genauer untersucht werden. Darüber hinaus werden mit Hilfe einer Befragung von Lehrpersonen noch mögliche Gründe für die möglicherweise fehlende Wirkung erfasst. Zum Tagungstermin liegt die vollständige Datenauswertung vor.

PS-70: Digitale Lernumgebungen als Schlüssel zur Förderung von wissenschaftlichem Denken und Praxis

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 145

Chair der Sitzung: Miriam Lechner

Erlernen der wissenschaftlichen Beobachtung anhand von Worked Example in einer digitalen Lernumgebung

Miriam Lechner, Stephanie Moser, Doris Lewalter

Technische Universität München, Deutschland

Theoretischer Hintergrund: Aktuell beschäftigen sich nur wenige Studien mit den Auswirkungen von heuristischen double-content WE auf den Erwerb grundlegender wissenschaftlicher Kompetenzen (Koenen et al., 2017). WE bestehen dabei aus der Problemstellung, den Lösungsschritten und der Lösung (Renkl, 2014). Speziell digitale WE führen zu einer besseren Kenntnis der Inhalte und können positive Auswirkungen auf die Lernmotivation, insbesondere auf Interesse und Zufriedenheit haben (Van Harsel et al., 2019; Gupta, 2019). Abhängig ist der Erfolg von WE maßgeblich von der kognitiven Belastung während des Lernens, wobei der Einsatz von WE zu einer Verringerung des extraneous CL und gleichzeitig zu einer Erhöhung des germane CL führen kann (Renkl, 2014). Insbesondere faded WE verbessern den Wissenstransfer stärker als non faded WE, wobei faded WE motivierender für erfahrene Lernende sind und non faded WE für Novizen (Paas et al., 2005). Eine Form von WE sind heuristische double content WE, die v.a. grundlegende naturwissenschaftliche Arbeitsmethoden umfassen wie die wissenschaftliche Beobachtung (Koenen et al., 2017). Bei double-content WE wird in der learning domain das zu erlernende Konzept vermittelt, in unserer Studie ist dies die wissen. Beobachtung. Die exemplifying domain beinhaltet dann den spezifischen Inhalt, anhand das Konzept aus der learning domain erlernt werden soll. In unserer Studie ist dies die Gewässerstrukturkartierung. Die wissenschaftliche Beobachtung (learning domain) ist definiert als theoriegeleitete, systematische und selektive Beobachtung konkreter Systeme und Prozesse ohne grundlegende Manipulation (Kohlhauf et al., 2011). Entwickelt werden soll ein Verständnis für die Kriterien einer angemessenen Versuchsplanung. In unserer Studie folgen wir der Literatur von Wellnitz und Mayer (2013), die sechs Schritte der wissenschaftlichen Beobachtung wie folgt definieren: Formulierung der Forschungsfrage(n), Ableitung der Nullhypothese und Alternativhypothese, Planung des Forschungsdesigns, Durchführung, Datenanalyse und Beantwortung der Forschungsfrage(n). Ziel der Studie war die Entwicklung, Analyse und praktische Anwendung einer digitalen Lernumgebung und im Hinblick auf Wissenserwerb, Motivation und kognitive Belastung zu testen. Hierfür wurden unterschiedlich gestaltete double-content WE in Form von Lernvideos mit hohen selbstgesteuerten Lernanteil eingesetzt.

Fragestellung: Die Forschungsfragen lauten: Inwieweit hat der Einbezug bzw. die unterschiedliche Gestaltung von WE (keine vs. non faded vs. faded) einen Einfluss auf 1) den Kompetenzerwerb (Faktenwissen über die wissenschaftliche Beobachtung) und die praktische Anwendung der wissenschaftlichen Beobachtung? 2) die Motivation (situationales Interesse und Basic Needs) 3) den Cognitive Load?

Methode: Die Studie wurde als Feldexperiment mit einfaktoriellen, quasi-experimentelles Pre-/Posttest-Design durchgeführt. Sie begann mit einer Online-Kick-off-Veranstaltung zur Gewässerstrukturkartierung, vorauf sich die Teilnehmenden anschließend über Moodle vorbereiteten. Am Exkursionstag erlernten die Teilnehmenden anhand der digitalen Lernumgebung die wissenschaftliche Beobachtung, während sie fünf Flussabschnitte anhand eines papierbasierten Kartierungsbogens kartierten. Die Stichprobe waren 62 Studierende der Naturwissenschaften und Doktoranden, welche zufällig einer der drei WE-Bedingungen zugewiesen wurden. Die Datenerhebung erfolgte über onlinebasierte Fragebögen sowie über die qualitative Auswertung der Antworten aus der digitalen Lernumgebung anhand eines selbstentwickelten Kodierschemas. Die Fragebögen enthielten Fragen zum Faktenwissen der wissenschaftlichen Beobachtung (Wahser, 2008, adaptiert von Koenen, 2014), Fragen zur Selbstbestimmtheit beim Lernen (Thomas & Müller, 2011), Basic Needs (Willems & Lewalter, 2011), Situationales Interesse (Lewalter & Knogler, 2014), und Cognitive Load (Leppink et al., 2013). Die Datenauswertung erfolgte mittels ANOVA, um die drei WE-Bedingung hinsichtlich der Zunahme des Faktenwissens über die wissenschaftliche Beobachtungsmethode von Pre- zu Posttest zu vergleichen. Eine MANOVA mit den drei Gruppen als festem Faktor und den abhängigen Variablen Anwendungswissen, situationales Interesse, Basic Needs und Cognitive Load wurde gerechnet, sowie Bonferroni- korrigierte Post-hoc-Analysen, um Unterschiede im Anwendungswissen zwischen den drei Gruppen festzustellen.

Ergebnisse: Die Studie zeigt, dass heuristische double-content non-faded und faded WE geeignet sind das Erlernen von Anwendungswissen über die wissenschaftliche Beobachtung signifikant zu steigern und den Aufbau naturwissenschaftlicher Kompetenzen zu fördern. Deskriptiv konnte gezeigt werden, dass die lernrelevante Motivation gefördert werden kann und keine kognitive Überlastung entsteht.

ENTFÄLLT Generative AI in Higher Education: A RAG System to Enhance Student Engagement with Scientific Literature

Dominik Thüs, Sarah Malone, Roland Brünken

Universität des Saarlandes, Deutschland

Theoretical Background

This study investigated the integration and impact of OwlMentor, a self-developed AI-powered learning environment designed to assist university students in comprehending scientific texts. Scientific literacy is crucial for engaging with academic texts, connecting arguments, methods, and data to draw conclusions (Goldman & Bisanz, 2002). Many students struggle due to underdeveloped scientific literacy (Sason et al., 2020). Meta-cognitive strategies such as self-questioning, linking new information to prior knowledge, summarizing, and self-explaining can improve scientific text comprehension (Chi et al., 1989, 1994; Cromley, 2010; Gunn, 2008; Kendeou & van den Broek, 2005, 2007). Generative AI (GAI) and large language models support these methods through interactive engagement, practice questions, and feedback (Kasneci et al., 2023). Studies suggest conversational AI can enhance student performance, motivation, self-efficacy and cognitive engagement (Liang et al., 2023; Wollny et al., 2021; Wu & Yu, 2023). The Technology Acceptance Model (TAM) explains how users come to accept and use technology, focusing on perceived ease of use and perceived usefulness as key determinants of intention to use and actual system use (Venkatesh & Davis, 1996)

Research Questions

Our general interest was to investigate how AI can be effectively integrated into higher education to enhance text comprehension and learning. The specific research questions of this study were:

- How do perceived ease of use and perceived usefulness affect students' intention to use OwlMentor?
- How does general self-efficacy impact system use and perceived usefulness over time?
- What are the relationships between system use and learning outcomes?
- How accurate are OwlMentor's responses and what are common user requests?

Methods

OwlMentor was developed participatorily and integrated into a course, with development and evaluation taking place over two semesters. In the first semester, a pilot study with 17 students tested the technical functionality, usability and quality of AI-generated responses. In the second semester, 16 students used the revised OwlMentor. Based on the TAM, we assessed perceived ease of use, perceived usefulness, intention to use, and actual system use.

Learning outcomes were measured through pre- and post-tests, TAM variables and self-efficacy were assessed at three time points and system use was logged continuously. Expert evaluation of OwlMentor's responses assessed accuracy and relevance, alongside a qualitative analysis of user requests.

Results and Their Meaning

Descriptive analysis indicated higher initial OwlMentor usage (TP1: $M = 11$ active users) which decreased over time (TP2: $M = 6$, TP3: $M = 6$). Students showed significant learning gains from pre- to post-test ($t(13) = -3.56$, $p < .001$, Cohen's $d = 1.93$) with users achieving higher gains ($M = 7.35$, $SD = 3.85$) than non-users ($M = 5.75$, $SD = 2.90$).

Perceived usefulness correlated strongly with intention to use (T1: $r(15) = .94$, $p < .001$; T2: $r(13) = .79$, $p < .001$; T3: $r(10) = .87$, $p < .001$). Multiple regression confirmed perceived usefulness as a significant predictor of intention to use ($\beta = .950$, $t = 9.336$, $p < .001$). High self-efficacy was negatively correlated with perceived usefulness ($r(10) = -.85$, $p = .002$) and intention to use ($r(10) = -.81$, $p = .005$), suggesting students with higher confidence found OwlMentor less necessary.

Qualitative analysis of user requests indicated a need for explanations, summaries, and understanding relationships between concepts. The expert evaluations recognized the high quality of the OwlMentor responses, supporting the RAG approach's effectiveness.

In summary, OwlMentor demonstrates potential in supporting scientific text comprehension. However, its success depends on adapting to user needs, continuously adding value, and integrating with existing learning strategies. Future AI-enhanced educational tools should adopt a flexible, student-centered approach, emphasizing regular assessments and iterative development to remain relevant and effective in higher education.

Auswirkungen digitaler, barrierearmer Lernumgebungen auf das Selbstkonzept im Umgang mit Experimentalanleitungen

Martina Graichen, Linda Hinderer, Silke Mikelskis-Seifert

PH Freiburg, Deutschland

Theoretischer Hintergrund

Voraussetzungen für Teilhabe am Unterricht und am gesellschaftlichen Leben sind naturwissenschaftliche Kompetenz und ein positives Selbstkonzept. Letzteres beeinflusst den Lernzuwachs und kann über Motivation gefördert werden. Lernumgebungen sollten möglichst frei von Lernbarrieren und Stereotypen sein, um allen Schüler*innen den Zugang zum Lernmaterial zu ermöglichen. Dies besonders bezogen auf Mädchen, die zumeist ein geringeres Interesse für Naturwissenschaften aufweisen (Hoffmann, 2002; OECD, 2016).

Das Experimentieren spielt aufgrund seiner Rolle zur Erkenntnisgewinnung und Entwicklung von naturwissenschaftlichem Verständnis und Kompetenz eine wichtige Rolle. Im Sinne der Handlungsfähigkeit stehen die Experimentieranleitungen (Zugang zu Informationen und zum eigentlichen Handeln) im Fokus. Die Gedanken zur Barrierefreiheit (Oettle et al., 2021) und Teilhabe setzt das Konzept „Science for All“ (Stinken-Rösner et al., 2020) um.

Unsere adaptive Lernumgebung basiert auf comicbasiertem Storytelling (Akanca, 2020), welcher, durch die Verknüpfung von Inhalt und Erzählung, die Erinnerung und den Abruf der Inhalte erleichtert (Kromka & Goodboy, 2019). Durch die Kombination aus Text und Bild sind wissenschaftliche Comics dem Lernen aus Text überlegen (Mayer, 2014, 2024).

Fragestellung

Untersucht wurde die Entwicklung des Selbstkonzeptes während des Lernens mithilfe einer digitalen, barrierefreien Lernumgebung. Zunächst überprüfen wir die Barrierefreiheit (H1), anschließend, ob sich das Selbstkonzept im Umgang mit Experimentieranleitungen zwischen den Messzeitpunkten erhöht (H2). Wir gingen davon aus, dass sich das Selbstkonzept von Mädchen in einer stereotypen Lernumgebung stärker entwickelt, (H3). Explorativ untersuchten wir die Entwicklung des Selbstkonzeptes genauer (H4) und prüften, ob die Dimensionen des Selbstkonzeptes einen Einfluss auf das Selbstkonzept in jeder Dimension haben.

Methode

Nach den genannten Konzepten entwickelten wir eine digitale Lernumgebung zum Magnetismus für die Klassenstufe 5/6 ($N=348$).

In einem Prä-Post-Vergleich mit Intervention maßen wir im ersten Termin Selbstkonzepte im Umgang mit Experimentieranleitungen in den drei Dimensionen Handlung, Sprache und Visuell (Graichen et al., in Vorbereitung), (Skala 0-3: 0= sehr gering ausgeprägtes Selbstkonzept; 3= sehr stark ausgeprägtes Selbstkonzept). Eine Woche später bearbeiteten die Schüler*innen die Lernumgebung und führten selbstständig zwei Experimente zum Magnetismus aus. Dabei erhoben wir nach jedem Experiment die wahrgenommene Barrierefreiheit in den Dimensionen (Skala 0-3: 0= keine Barrierefreiheit; 3= volle Barrierefreiheit). Zwei Wochen später wurde das Selbstkonzept in den drei Dimensionen erneut gemessen.

Ergebnisse und Diskussion

Die Ergebnisse zeigen, dass die Schüler*innen die digitalen Anleitungen zum Experimentieren (H1) in allen drei Dimensionen als barrierearm wahrnahmen ($2,44 \geq M \geq 2,57$), da sich die Mittelwerte im oberen, d.h. barrierearmen Bereich der Skala liegen.

Zur Prüfung der Hypothesen 2 und 3 rechneten wir eine ANOVA mit Messwiederholung für die Faktoren Zeit (H2) und Geschlecht (H3) mit den Dimensionen des Selbstkonzeptes. Die Analyse zeigte einen Haupteffekt Zeit (H2: $F(1,295)=10,51$, $p=0,001$,

$\eta^2p=0,034$), der für eine positive Entwicklung des Selbstkonzeptes spricht. Der Haupteffekt für Geschlecht ($H3: F(1,295)=0,02, p=0,902, \eta^2p=0,00$) sowie die Interaktion aus Zeit und Geschlecht ($F(1,295)=0,93, p=0,337, \eta^2p=0,00$) waren beide nicht signifikant, weshalb wir davon ausgehen können, dass es keine geschlechtsspezifischen Entwicklungen zwischen Jungen und Mädchen gab.

Die jeweilige explorative multiple Regression auf jedes der Selbstkonzepte im Umgang mit experimentellen Anleitungen zum zweiten Messzeitpunkt ($H4$) zeigte, dass der jeweilige Selbstkonzeptwert zum ersten Messzeitpunkt einen Einfluss hatte. Außerdem sagte die Barrierefreiheit der Dimension Visuell in allen drei Analysen das Selbstkonzept voraus. Darüber hinaus beeinflusste die Barrierefreiheit der Dimension Sprache die Selbstkonzepte in den Dimensionen Handlung und Sprache.

Die Ergebnisse zeigen, dass die Anleitungen in der digitalen Lernumgebung barrierearm wahrgenommen wurden und bei den Schüler*innen das Selbstkonzept im Umgang mit Experimentieranleitungen gesteigert werden konnte. Die Analysen zeigen zudem, dass es noch offene Fragen gibt: Insbesondere weitere Einflussfaktoren auf das Selbstkonzept und die weitere Auswirkung der barrierefreien auf die Lernwirksamkeit, sollten untersucht werden.

ENTFÄLLT: Vermittlung guter wissenschaftlicher Praxis durch Graphic Novel basierte Lernmodule

André Kolbe, Carolin Eitemüller, Katrin Schübler, Maik Walpuski

Universität Duisburg-Essen, Deutschland

Theoretischer Hintergrund

Gute wissenschaftliche Praxis (GWP) bildet das Fundament der wissenschaftlichen Integrität. Die All European Academies (ALLEA) hat Leitlinien zur Sicherung der GWP veröffentlicht, die kontinuierlich aktualisiert werden. ALLEA hat 2017 den European-Code-of-Conduct-for-Research-Integrity herausgegeben, der als Referenzrahmen für gute wissenschaftliche Praxis in Europa dient (All European Academies, 2017). Diese Leitlinien bieten eine Grundlage für nationale Regelungen. Die Deutsche Forschungsgemeinschaft (DFG) hat 1998 eigene Leitlinien zur Sicherung der GWP veröffentlicht, die ebenfalls kontinuierlich aktualisiert werden (Deutsche Forschungsgemeinschaft, 2022). Ziel der DFG ist es, eine verbindliche Kultur wissenschaftlicher Integrität an Forschungsinstitutionen zu etablieren. Daher ist es unerlässlich, bereits dem wissenschaftlichen Nachwuchs während des Studiums die Grundlagen der GWP zu vermitteln, um eine ehrliche Wissenskultur zu fördern.

Fuerholzer et al. (2020) entwickelten GWP-Kurse für Medizinstudierende und zeigten signifikante Veränderungen im Wissen und den Einstellungen der Studierenden gegenüber den GWP-Leitlinien. Allerdings fehlen flächendeckende GWP-Module für naturwissenschaftliche Fächer sowie Erkenntnisse über lernwirksame Merkmale solcher Module.

Multimediale Lernmaterialien spielen im Selbststudium eine wichtige Rolle, da sie motivationale und selbstregulative Anreize bieten und gegenüber traditionellen Lernmaterialien einen deutlichen Mehrwert hinsichtlich der Lernwirksamkeit bieten (Dannik et al., 2018). Interaktive Lernmaterialien können Lernen effektiver und effizienter gestalten und zu einem tieferen konzeptuellen Verständnis beitragen (Baumgartner & Herber, 2013). Generative Aktivitäten fördern eine aktive Auseinandersetzung mit Lerninhalten (Fiorella, 2021).

Für GWP erscheint neben Kenntnis entsprechender allgemeiner Regeln vor allem die Umsetzung der Regeln, die meist an komplexe und vor allem fachspezifische Situationen gebunden ist, besonders wichtig. Es ist daher anzunehmen, dass Fallbeispiele besonders gute Lerngelegenheiten sein können (Zumbach et al., 2008). Da Sogunro (2014) konnte zeigen, dass textlastige Lerninhalte die Motivation negativ beeinflussen und das Lernen behindern, können die Lernmaterialien durch *Graphic Novels* textlich entlastet werden. *Graphic Novels* können zudem die Lernmotivation steigern und ermöglichen es Lerninhalte interessant zu präsentieren (Fitri Dwi Arini et al., 2016; Sipayung et al., 2020).

Bisher wurde nicht untersucht, ob digitale Selbstlernmaterialien, die mit komplexen Fallbeispielen in Form von *Graphic Novels* angereichert wurden, die Lernwirksamkeit steigern. GWP dient hierbei als exemplarisches Anwendungsgebiet, um die Effektivität solcher Lernmaterialien zu evaluieren.

Fragestellungen

Die Studie zielt darauf ab, die Effektivität digitaler Lernmodule, die mit *Graphic Novels* angereichert sind, zu untersuchen. Dabei werden drei Gruppen gebildet: Kontrollgruppe (KG), Interventionsgruppe 1 (IG1), in der *Graphic Novels* in der Instruktionsphase eingesetzt werden, und Interventionsgruppe 2 (IG2), in der *Graphic Novels* in den Lernaufgaben verwendet werden.

Folgende Forschungsfragen sollen untersucht werden:

Inwiefern unterscheiden sich digitale Lernmodule mit *Graphic Novels* ...

FF1: von Lernmodulen ohne *Graphic Novels* (KG) hinsichtlich Lernwirksamkeit, kognitiver Belastung, Usability und der motivationalen Wirkung.

FF2: in der Instruktionsphase (IG1) von Lernmodulen mit *Graphic Novels* in Lernaufgaben (IG2) hinsichtlich Lernwirksamkeit, kognitiver Belastung, Usability und der motivationalen Wirkung.

Methode

Im Rahmen eines Kooperationsprojekts an drei nordrhein-westfälischen Universitäten wurden GWP-Module für Bachelor- und Masterstudierende der Naturwissenschaften als Selbstlernmaterialien entwickelt und empirisch untersucht. Die Forschungsfragen werden mittels Pre-Post-Kontrollgruppen-Design an einer Stichprobe von 150 Studierenden untersucht.

Im Sommersemester 2024 werden drei digitale GWP-Lerneinheiten hinsichtlich (a) Lernwirksamkeit, (b) kognitiver Belastung, (c) Usability und (d) Motivation evaluiert. Die Lernwirksamkeit wird mittels selbstentwickelten GWP-Fachwissenstest erhoben. Die kognitive Belastung wird mit Items von Kriegstein et al. (2023) und die Usability mit Items nach Schrepp (2023) gemessen. Die Motivation wird mit Items nach Engel (2004), Haugwitz (2009) und Flake et al. (2015) erhoben. Zur Untersuchung von Unterschieden für Lernwirksamkeit, kognitive Belastung, Usability und Motivation zwischen den Gruppen werden einfaktorielle Varianzanalysen mit anschließenden Post-hoc-Tests durchgeführt. Die Ergebnisse dieser Studie werden auf der Tagung präsentiert.

Ausblick

Das Projekt trägt praktisch zur Entwicklung evaluierter GWP-Lerneinheiten bei, die zu GWP-Modulen zusammengestellt und in bestehende Curricula integriert werden können. Theoretisch liefert die Studie Erkenntnisse zur Lernwirksamkeit von *Graphic Novels* in GWP-Lernmodulen unter Berücksichtigung kognitiver Belastung, Usability und Motivation.

PS-71: Pädagogische Überzeugungen und Qualität in der frühkindlichen Bildung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: SO 422

Chair der Sitzung: Franziska Cohen

Zusammenhang zwischen pädagogischen Überzeugungen frühpädagogischer Fachkräfte und ihrer Interaktionsqualität im Guided play: Ergebnisse aus einer deutsch-österreichischen Kitastudie

Anna Bilgeri¹, Katharina Klucznik², Wilfried Smidt³, Franziska Cohen¹

¹Pädagogische Hochschule Freiburg, Deutschland; ²Freie Universität Berlin, Deutschland; ³Leopold-Franzens-Universität Innsbruck, Österreich

In der frühen Bildung gilt das Spiel als bedeutsamer Kontext für kindliche Bildungsprozesse (Leuchter, 2013). Allerdings befindet sich das kindliche Spiel im theoretischen und praktischen Diskurs in einem Spannungsfeld zwischen pädagogischer Einflussnahme und Selbstständigkeit des Kindes (Sauerbrey, 2021). Guided play versucht eine Balance zwischen der Autonomie und Selbststeuerung des Kindes sowie der Begleitung durch die pädagogische Fachkraft (z.B. durch Scaffolding; dialogisch entwickelte, offene Fragen) zu schaffen (Skene et al., 2022; Weisberg et al., 2016).

Das erweiterte Struktur-Prozessmodell pädagogischer Qualität (Anders & Oppermann, 2024) dient als theoretischer Rahmen, um den Zusammenhang zwischen den professionellen Kompetenzen frühpädagogischer Fachkräfte und der beobachteten Interaktionsqualität im Spielkontext zu untersuchen. Die professionellen Kompetenzen der Fachkräfte gelten als entscheidend für den Umgang mit komplexen Anforderungen im pädagogischen Alltag und für eine hohe Interaktionsqualität (Anders & Oppermann, 2024; Fröhlich-Gildhoff et al., 2011). Da pädagogische Überzeugungen als Filter, Leitlinien und Rahmen für das pädagogische Handeln dienen (Fives & Buehl, 2012) stellt sich die Frage, inwiefern sie die Qualität und Quantität von Interaktionsprozessen im Guided play beeinflussen. Internationale Studien zeigen, dass die pädagogischen Überzeugungen das Rollenverständnis sowie die Initiative und Aktivität der Fachkräfte im Spiel beeinflussen (Cheung et al., 2022; Keung & Fung, 2021; McInnes et al., 2011). Allerdings gibt es im deutschsprachigen Raum bisher wenig Forschung zu den pädagogischen Orientierungen im Zusammenhang mit dem Guided Play (Schmitt, Weber, Venitz & Leuchter, 2023). Dieses Forschungsdesiderat wird im Rahmen des vorliegenden Beitrags aufgegriffen. Anhand einer Stichprobe von $N = 502$ Kindern aus 142 Kitas aus Österreich ($n = 261$ Kinder) und Deutschland ($n = 241$) werden Zusammenhänge zwischen den pädagogischen Überzeugungen und der Interaktionsqualität im Guided play analysiert. Guided play wird bei gleichzeitigem Auftreten der Kategorien: Freispiel, Initiative ging vom Kind aus, Teilnahme der Fachkraft und Scaffolding-Verhaltensweisen oder Spielerischer Austausch als gegeben definiert. Die pädagogischen Überzeugungen wurden anhand von Fachkraftbefragungen in Anlehnung an Schmidt & Smidt (2021) mittels 12 Items auf einer 5-stufigen Skala erhoben (Beispielitem: Bei der Förderung von Kindern ist es wichtig, dass ... "die Kinder von der Fachkraft zu eigenen Lösungen ermuntert werden."). Zudem wurde die Interaktionsqualität mit dem Time-Sampling-Verfahren Zielkindbeobachtung (ZIKIB) (Kuger et al., 2006) und dem Beobachtungsinstrument Individualized Classroom Assessment Scoring System (inCLASS) (Downer et al., 2010) standardisiert erhoben. Die Zielkinder wurden in zehnminütigen Zyklen über den Vormittag hinweg mit beiden Instrumenten synchron beobachtet, wobei bis zu vier Beobachtungszyklen durchgeführt wurden. Strukturmerkmale auf Einrichtungsebene (Fachkraft-Kind-Schlüssel) und auf Fachkräfebene (Berufserfahrung, Qualifikation) sowie Kindmerkmale (Geschlecht, Migrationshintergrund und Alter) werden als Kontrollvariablen berücksichtigt. Die hierarchische Datenstruktur wird durch Mehrebenenregressionen und fehlende Werte werden über die FIML-Methode berücksichtigt. Die vorhandenen Daten werden gepoolt analysiert.

Die Zielkinder verbringen im Durchschnitt 23 von insgesamt 40 beobachteten Minuten im Freispiel ($SD = 9,27$). Vergleichsweise selten wurde jedoch Guided play ($M = 1,7$ min, $SD = 3,01$) beobachtet, welches häufiger zu Beginn des Kindergarten tags (32% in Beobachtungszyklus 1: $M = 9:13$ Uhr $SD = 0:22$ und 34 % in Beobachtungszyklus 2: $M = 10:04$ Uhr, $SD = 0:41$) stattfindet. Obwohl Guided play nur selten beobachtet wurde, zeigen erste Ergebnisse einen signifikant positiven Zusammenhang zwischen der pädagogische Orientierung Ko-Konstruktion und Guided play ($r = .246$ $p = .029$). Darüber hinaus zeigt sich ein signifikant positiver Zusammenhang zwischen der Qualität der Fachkraft-Kind-Interaktionen und der Zeit, die Zielkinder mit Guided play verbringen ($r = .435$, $p = .001$). Im Gegensatz dazu besteht ein negativer Zusammenhang zwischen dem Auftreten von Guided play und der Interaktionsqualität mit Peers ($r = -.153$, $p = .017$). Die Ergebnisse werden im Hinblick auf die Bedeutung der Einstellungen für die Qualitätsentwicklung in der frühpädagogischen Forschung und Praxis diskutiert.

Orientierungen pädagogischer Fachkräfte zum Umgang mit „Krisen“ und „Demokratiebildung“ - Empirische Befunde aus dem System der Frühen Bildung

Dr. Leonhard Birnbacher, Judith Durand

Deutsches Jugendinstitut, Deutschland

Vielfältig ausgemachte nationale Herausforderungen sowie multiple globale Krisen führen in politik- und sozialwissenschaftlichen Diskursen zur kritischen Betrachtung des Zustands der demokratisch verfassten Gesellschaft (Selk 2023; Lessenich 2019). Befunde, die auf einen Vertrauensverlust in die formalen Repräsentationsorgane der Demokratie hindeuten (Merkel et al. 2020) oder die Wahrnehmung eines sich zunehmend polarisierenden gesellschaftlichen Klimas (Lux/Gülzau 2022), tragen zur Einschätzung bei, dass die Demokratie vor ernstzunehmenden Herausforderungen steht (Münkler 2022). Auch das System der Frühen Bildung ist hiervon beeinflusst und wird auf unterschiedliche Weise mit Krisenkonstellationen konfrontiert (Meier-Gräwe 2024). Einerseits sollen erziehungswissenschaftliche Subdisziplinen für ihre Gegenstandsbereiche analysieren, wie Erziehungsinstitutionen von Krisen betroffen sind (Otto et al. 2020). Andererseits wird neben Befunden einer Krise im System „Kita“ (u.a. Rauschenbach et al. 2024), der Kita als Akteur eine zentrale Rolle in der Bearbeitung von gesellschaftlichen Krisen zugeschrieben (Wrana et al. 2022). Als Bestandteil der Gesellschaft und erste öffentliche Bildungsinstitution wird von ihr gefordert, gesellschaftliche Veränderungen aufzugreifen und pädagogisch zu bearbeiten (Andresen 2018; Oelkers 2010) sowie zur Entwicklung von Kompetenzen beizutragen, wie Perspektivübernahme, Toleranz, Verantwortungsübernahme, Meinungsartikulation oder Kompromissfähigkeit, um Kinder auf ein Leben in einer Demokratie vorzubereiten. Dabei wird meist wenig beachtet, wie die Akteure der Frühen Bildung selbst Krisen und deren Einfluss auf das tägliche Geschehen in der Institution wahrnehmen und welche Orientierungen zu Demokratie sie zu erkennen geben.

Vor diesem Hintergrund stellen sich *die Fragen*, wie Demokratiebildung in der Kita verstanden und praktiziert wird, wie pädagogische Fach- und Leitungskräfte gesellschaftliche Krisen im Allgemeinen wahrnehmen und welche Einfluss und Relevanz sie diesen im Besonderen für die pädagogische Arbeit mit Kindern und Familien in Kitas zusprechen. Zudem ist von Interesse, welche Orientierungen pädagogische Fachkräfte gegenüber Demokratie als gesellschaftliches Ordnungssystem zeigen, was sie

unter Demokratiebildung verstehen und welche Schlussfolgerungen sie für die pädagogische Praxis daraus ableiten. Grundlage für die Analysen sind Ergebnisse der qualitativen empirischen Studie "Bildung und Demokratie mit den Jüngsten" (Birnbacher et al 2023).

Die Fragestellungen wurden in einem *qualitativen multimethodischen und multiperspektivischen Design* über Dokumentenanalysen (Hoffmann 2018), 23 leitfadengestützten Interviews mit Fach- und Leitungskräften (Helfferich 2014) und Kindern (Weltzien 2012) in 12 Kitas sowie mehrtägige videogestützte Beobachtungen an sechs Standorten (Petschick 2016; Tuma 2013) untersucht. Im Rahmen dieses Vortrags stehen die Auswertungen von 23 leitfadengestützten Interviews mit Fach- und Leitungskräften zu deren Perspektiven auf gesellschaftliche Krisen und deren Auswirkungen auf die Praxis der Frühen Bildung im Fokus, die systematisch anhand der strukturierenden Inhaltsanalyse (Kuckartz 2016) und mit Hilfe der Entwicklung von Kategorien am Material (Schreier 2014) durchgeführt wurde. Deduktiv gesetzte Kategorien, die sich aus einer theoretischen Heuristik (Eberlein et al. 2020) ableiteten, wurden induktiv durch Haupt- und Subkategorien ergänzt. Ein Pretest und die Diskussion von Abweichungen bei der Codierung wurde ein Kategorienhandbuch finalisiert und konsensuelles Codieren möglich (Kuckartz 2016). Die Codierung wurde computergestützt mittels MAXQDA durchgeführt.

Die *Analysen zeigen*, dass unterschiedliche Entwicklungen, Veränderungen und Herausforderungen, die als gesellschaftliche „Krisen“ verhandelt werden, vermittelt und unmittelbar auch in den Kitas aufschlagen und Fachkräfte vor Unsicherheiten und Herausforderungen stellen, die Auswirkungen auf die pädagogische Praxis haben. Diese Verunsicherungen lassen sich auch in den Orientierungen der Befragten zur Demokratie selbst nachvollziehen, die zwischen grundsätzlicher positiver Bezugnahme auf die Demokratie als generelle Ordnungsform und Kritik an deren aktueller konkreter Leistungsfähigkeit und Funktionsweise changieren. Zudem zeigt sich, dass – wie für den Krisentopos generell typisch (Graf 2020) – die Bezugnahme auf eine ausgemachte Problemkonstellation zunächst mit einer Verunsicherungsempfindung des professionellen Selbstverständnisses oder einer Unterbrechung der eigenen Handlungsroutinen verbunden wird. Die Ergebnisse werden im Hinblick darauf diskutiert, inwiefern die Kita als Akteur den Auftrag der Bearbeitung von gesellschaftlichen Krisen wahrnehmen kann, wenn sie selbst von Krisen betroffen ist und Anzeichen der Verunsicherung im professionellen Selbstverständnis der Fachkräfte auszumachen sind.

Metaanalyse zum Einfluss des Personalschlüssels auf die Prozessqualität in Kindertageseinrichtungen

Kristina Hausladen, Katrin Wolf, Lars Burghardt, Yvonne Anders

Otto-Friedrich-Universität Bamberg, Deutschland

Gute strukturelle Rahmenbedingungen in frühkindlichen Bildungseinrichtungen sind die Grundlage für eine effektive und qualitativ hochwertige Erfahrung für alle Kinder (Vandell et al., 2010), da sie die Prozessqualität in Gruppen für das Personal und die Kinder gleichermaßen prägen. Strukturelle Qualitätsaspekte wie der Personalschlüssel können generell leichter durch politische Vorschriften reguliert werden als andere Qualitätsdimensionen und sind daher zentral in der Debatte zur pädagogischen Qualität in Kindertageseinrichtungen. Gemäß der Logik des Struktur-Prozess-Qualitäts-Modells (z. B. Anders & Oppermann, 2024; Pianta et al., 2005) kann davon ausgegangen werden, dass ein geringerer Personalschlüssel in Kindertageseinrichtungen eine wichtige Voraussetzung für regelmäßige und hochwertige Interaktionen mit Kindern ist, was wiederum dazu beitragen kann, frühe Anzeichen von Lernschwierigkeiten vorzubeugen oder besondere Bedürfnisse der Kinder besser zu erkennen bzw. die Entwicklung der Kinder zu fördern (de Schipper et al., 2007; Pianta et al., 2009). Das Ziel dieser Metaanalyse war herauszufinden, ob der Personalschlüssel in Kindertageseinrichtungen das Niveau der Prozessqualität statistisch signifikant vorhersagt. Zusätzlich zu dieser Fragestellung analysieren wir potentielle Moderatoren, die diesen Zusammenhang beeinflussen könnten. Wir betrachten Moderatoren auf struktureller Ebene, sowie methodische Moderatoren. Insgesamt wurden 2.966 Publikationen über verschiedene Quellen ermittelt (Datenbanken, manuelle Suche wichtiger Autor*innen und Arbeiten und Literaturverzeichnis bereits identifizierter Publikationen). Unsere finale Stichprobe besteht aus 52 Publikationen mit Daten aus insgesamt 25.577 Kindertageseinrichtungen. Wir schlossen dabei nur Publikationen ein, die die Prozessqualität mit Beobachtungsinstrumenten erfassen. Wir verwendeten Fisher's z und standardisierte Regressionskoeffizienten als Effektgrößen und führen getrennte Meta-Analysen für beide Effektstärken durch. Die Analysen wurden mit dem metafor-Paket (Viechtbauer, 2010) in R mit einem dreistufigen Random-Effektsmodell durchgeführt, um Heterogenität zwischen den Effektgrößen zu berücksichtigen. Potenzielle Verzerrungen in den Ergebnissen wurden durch Sensitivitätsanalysen überprüft. Die Gesamteffektgröße mit Fisher's z zeigt einen kleinen, aber signifikanten negativen Effekt des Verhältnisses zwischen Kindern und Erzieher*innen auf das Niveau der Prozessqualität auf ($k=142$, $r=-0,117$, 95% CI(-0,178, -0,056), $p<.001$). Die Analysen mit standardisierten Regressionskoeffizienten bestätigen dieses Bild ($k=46$, $\beta=-0,089$, 95% CI(-0,148, -0,029), $p<.01$). Eine Erhöhung der Anzahl der Kinder pro Erzieher*in beeinflusst die Prozessqualität in der Gruppe tatsächlich negativ. Die Moderatorenanalysen zeigen, dass dieser negative Effekt in Gruppen für jüngere Kinder (0-2 Jahre) und in altersgemischten Gruppen größer ist als in Gruppen für ältere Kinder (3-6 Jahre). Auch die Moderatoren zur Messung der Schlüsselvariablen zeigen wichtige Unterschiede zwischen den Publikationen auf. So ist der negative Effekt höher, wenn Autor*innen den Personalschlüssel erfassen, anstatt der direkt beobachtbaren Fachkraft-Kind-Relation. Außerdem war der negative Effekt in Gruppen höher, in denen nur qualifizierte Erzieher*innen bei der Berechnung des Schlüssels berücksichtigt wurden. Bezüglich der Erfassung der Prozessqualität, zeigen unsere Ergebnisse, dass Heterogenität in den Effektstärken durch die Art und Weise wie die Prozessqualität in den Publikationen operationalisiert wurde erklärt werden kann. Am stärksten war der negative Effekt bei Instrumenten zur Erfassung der Prozessqualität, die sich auf die allgemeine Lernumgebung und die Quantität der Prozesse in der Gruppe konzentrieren (ITERS(-R), ECERS(-R), EQOS), gefolgt von Instrumenten, die sich auf die bereichsspezifische Prozessqualität konzentrieren (ECERS-E und Instrumente, z. B. zur Sprachförderung). Der signifikant negative Effekt des Personalschlüssels verschwindet fast bei Instrumenten, die die tatsächlichen Interaktionen zwischen Erziehern und Kindern erfassen (CLASS, CIS, STRS, OCRES). Zusammenfassend zeigen unsere Ergebnisse, dass ein niedrigerer Personalschlüssel, d.h. weniger Kinder pro Fachkraft, generell wichtig für die Lernumgebung in Kindertageseinrichtungen ist. Eine Regulierung des Personalschlüssels ist daher wichtig, um insbesondere jüngeren Kindern eine qualitativ hochwertige Lernumgebung zu ermöglichen. Darüber hinaus konnten wir zeigen, dass die Effektstärken eher heterogen sind und kontextspezifisch interpretiert werden sollten. Sowohl die Erfassung des Personalschlüssels als auch der Prozessqualität beeinflussen die Höhe der Effektstärken. So ist ein niedrigerer Personalschlüssel wichtiger für die Betrachtung der Quantität der Prozesse als für die tatsächlichen Interaktionen zwischen einer Fachkraft und den Kindern.

Der Einfluss von geschlechteratypischen fiktionalen Rollenmodellen auf Geschlechterstereotype und das geschlechtsspezifische Spielverhalten von männlichen Kindergartenkindern

Jan Lenhart, Katharina Nagl, Maria Dürrbeck

Otto-Friedrich-Universität Bamberg, Deutschland

Theoretischer Hintergrund und Fragestellung

Geschlechterstereotype, also Überzeugungen, wie sich Männer und Frauen verhalten oder zu verhalten haben (Eagly, 1987), beeinflussen viele Aspekte des sozialen Lebens, wie die Ausübung von Hobbies (Boiché et al., 2014), das Interesse für bestimmte Schulfächer (Li et al., 2022), und die Entscheidung für oder gegen bestimmte Karrierewege (Janssen & Backes-Gellner, 2016). Obwohl sich Geschlechterstereotype in vielen westlichen Kulturen im Lauf des vergangenen Jahrhunderts abgeschwächt haben (z.B. Eagly et al., 2020), wird geschlechtsnonkonformes Verhalten insbesondere bei Jungen und Männern kritisch gesehen und sozial sanktioniert (Kwan et al., 2020). Da sich Geschlechterstereotype bereits in der frühen Kindheit ausformen (Trautner et al., 2005) und Einfluss auf Entscheidungen und Verhalten nehmen (Seitz et al., 2020), sind Interventionen, die Geschlechterstereotype infrage stellen, ein wichtiger Aufgabenbereich der frühen Bildung.

Insbesondere zu effektiven Interventionen bei Jungen ist die Forschungslage begrenzt, sodass wir in der vorliegenden Studie untersuchten, ob sich der Logik der sozial-kognitiven Theorie (Bussey & Bandura, 1999) folgend eine Geschichte mit antistereotypem Rollenmodell, das für geschlechtsnonkonformes Verhalten durch wichtige Bezugspersonen verstärkt wird, auf die Geschlechterstereotype und das geschlechtsspezifische Spielverhalten von männlichen Kindergartenkindern auswirkt und ob Interventionseffekte vom Alter der Kinder oder der Geschlechterrollenorientierung ihrer Eltern abhängen.

Methode

Im Rahmen eines Pre-Post-Design wurden 82 männliche Kindergartenkinder im Alter von 3 bis 6 Jahren randomisiert einer von zwei Versuchsbedingungen zugewiesen. In der Experimentalgruppe ($n = 44$) hörten die Kinder eine Geschichte von einem Jungen, der mit einem Freund mit zwei „typisch weiblichen“ Spielzeugen spielt, dafür vom Vater gelobt wird, und der von den Eltern erfährt, dass es ganz normal für Jungen ist, mit diesen Spielsachen zu spielen. In der Kontrollgruppe ($n = 38$) hörten die Kinder eine identische Geschichte mit dem Unterschied, dass das „typisch weibliche“ Spielzeug durch „typisch männliches“ Spielzeug ersetzt wurde.

Im Pre- und Posttest durften die Kinder mit acht Spielzeugen (zwei „typisch weibliche“ Targetitems aus der Geschichte, zwei „typisch weibliche“ Transferitems und vier männliche Kontrollitems) in einem Zeitraum von vier Minuten frei spielen. Mithilfe von Videoaufzeichnungen wurde codiert, ob und wie lange die Kinder mit den jeweiligen Spielsachen spielten. Darüber hinaus wurden die Kinder zu jedem Spielzeug gefragt, ob üblicherweise Jungen oder Mädchen mit dem Spielzeug spielen (deskriptive Geschlechterstereotype), ob Jungen oder Mädchen damit spielen dürfen (präskriptive Geschlechterstereotype), und ob sie selbst gerne damit spielen möchten (Spielzeugpräferenz) (jeweils 5-stufige Skala). Zudem wurden das Alter der Kinder in Monaten sowie von Mutter und Vater getrennt die normative Geschlechterrollenorientierung (mit dem Fragebogen von Krampen, 1979) erfragt.

Ergebnisse und ihre Bedeutung

Die Ergebnisse zeigen, dass die Jungen „typisch männliche“ Spielzeuge als deskriptiv und präskriptiv männlich bewerteten, während die „typisch weiblichen“ Spielzeuge deskriptiv und präskriptiv den Mädchen zugeordnet wurden. Die Jungen äußerten zudem deutliche Präferenzen für die „typisch männlichen“ Spielzeuge, spielten häufiger und deutlich länger mit diesen. Während die Intervention keine signifikanten Effekte auf die präskriptive und deskriptive Bewertung der Spielzeuge hatte und keinen Einfluss auf die Spieldauer hatte, erhöhte die Intervention die Präferenz sowie die Wahrscheinlichkeit für das Spielen mit den „typisch weiblichen“ Targetitems. Für die Transferitems und die Kontrollitems zeigten sich hingegen erwartungsgemäß keine Auswirkungen der Intervention. Weder das Alter der Kinder noch die Geschlechterrollenorientierung ihrer Eltern hatten einen Einfluss auf die Effekte der Intervention.

Die Ergebnisse zeigen, dass bereits kurze Interventionen Interesse und das Spielen bzw. Anfassen von „typisch weiblichen“ Objekten anregen können. Geschlechterstereotype selbst sowie die Persistenz von Verhalten scheinen hingegen weniger leicht beeinflussbar zu sein. Allerdings muss berücksichtigt werden, dass in der vorliegenden Studie nur Objekte mit sehr starker Geschlechtsassoziation präsentiert wurden, sodass es möglich ist, dass weniger stark geschlechtsassoziierte Verhaltensweisen auch hier leichter für Veränderungen zugänglich sind.

S-103: Auf dem Weg zu adaptiven Online-Fortbildungen? Voraussetzungen von Lehrkräften und ihr Nutzungsverhalten

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 148

Chair der Sitzung: Tim Fütterer

Chair der Sitzung: Anika Dreher

Chair der Sitzung: Daniela Rzejak

Auf dem Weg zu adaptiven Online-Fortbildungen? Voraussetzungen von Lehrkräften und ihr Nutzungsverhalten

Chair(s): **Tim Fütterer** (Universität Tübingen, Deutschland), **Anika Dreher** (Pädagogische Hochschule Freiburg)

Diskutant*in(nen): **Daniela Rzejak** (Universität Kassel)

Fort- und Weiterbildungen sind zentrale Maßnahmen, um Lehrkräfte auf berufliche Herausforderungen vorzubereiten (Darling-Hammond et al., 2017). In den letzten Jahren haben Online-Fortbildungen zunehmend an Bedeutung gewonnen (Dede et al., 2009; Fütterer et al., 2024a, 2024b; Meyer et al., 2023), da sie sich auf Lehrkräfte- und Schülerebene als wirksam erwiesen haben (Morina et al., 2023), sie potentiell mehr Lehrkräfte erreichen können als Präsenzfortbildungen (Scaling-up) und Lehrkräfte das Fortbildungsangebot häufig nicht nur örtlich, sondern auch zeitlich flexibel nutzen können. Darüber hinaus wird Online-Fortbildungen das Potential zugesprochen, der Heterogenität von Lehrkräften besonders gut gerecht werden zu können (Copur-Gencturk & Orrill, 2023). Im Sinne eines Angebots-Nutzungs-Modells auf Fortbildungsebene (Fütterer et al., 2024a; Lipowsky & Rzejak, 2019) sollen flexible Nutzungsmöglichkeiten eine bessere Passung zwischen den individuellen Voraussetzungen der Lehrkräfte und dem Fortbildungsangebot ermöglichen. Fortbildungen, die nicht einer "one-size-fits-all" Logik folgen, sondern jeweils zu individuellen Voraussetzungen der Lehrkräfte angepasst sind bzw. anpassbar sind, sind adaptive Fortbildungen.

Allerdings liegt bisher kaum Evidenz dazu vor, inwiefern und unter welchen Bedingungen Online-Fortbildungen dieses Potential wirklich ausschöpfen können. An diesem Desiderat setzt das Symposium an. Die drei Beiträge fokussieren dabei unterschiedlich stark auf die relevanten Stellen des Angebots-Nutzungs-Modells:

In Beitrag 1 werden verschiedene individuelle Voraussetzungen von Lehrkräften und schulische Rahmenbedingungen genauer in den Blick genommen und ihre Bedeutung für die Nutzung von Online-Fortbildungen im Vergleich zu Präsenzfortbildungen (als Fortbildungsteilnahme) untersucht. Dazu wurden Zusammenhänge zwischen der Art der Fortbildung und Herausforderungen, die mit der Fortbildungsteilnahme verbunden sind, anhand von Mehrebenen-Pfadmodellen analysiert. Die Befunde deuten darauf hin, dass sich die Bedeutung der untersuchten Voraussetzungen nicht signifikant zwischen Präsenz- und Online-Fortbildungen unterscheidet, d.h. Online-Fortbildungen nicht speziell jene Lehrkräfte erreichen, die eingeschränkte Möglichkeiten haben, an Präsenzfortbildungen teilzunehmen.

Beitrag 2 fokussiert auf die Frage, inwiefern flexible Nutzungsmöglichkeiten verschiedener Angebotstypen einer Online-Fortbildung eine Nutzung im Sinne individueller Lernziele der Lehrkräfte ermöglicht, ohne dass dabei intendierte Lernziele der Fortbildung aus dem Blick geraten. Dazu wurden subjektive und objektive Maße für lernzieltreue Nutzung verschiedener Angebotstypen sowie ihre Zusammenhänge untereinander und mit anderen Nutzungsvariablen (z.B. kollegialer Austausch) untersucht. Die Ergebnisse zeigen positive Zusammenhänge zwischen den beiden Maßen für alle Angebotstypen und liefern Hinweise für die Bedeutung von anderen Nutzungsvariablen für diese Maße.

Beitrag 3 nimmt schließlich die Zusammenhänge zwischen verschiedenen Voraussetzungen von Lehrkräften und deren Nutzung einer Online-Fortbildung in den Blick. Dazu werden Daten zu individuellen motivationalen, kognitiven und metakognitiven Voraussetzungen und auch Nutzungsdaten in Form von digitalen Verhaltensspurdaten erhoben. Zur Untersuchung der Passung von Angebot, Voraussetzungen und Nutzung wird ein logistisches Regressionsmodell spezifiziert und zur Identifikation von Nutzungsprofilen werden latente Profilanalysen durchgeführt.

Zusammengenommen zeichnen die Beiträge ein gemischtes Bild hinsichtlich der Potentiale von Online-Fortbildungen, welches von einer im Bereich Lehrkräftefortbildungen einschlägigen Wissenschaftlerin eingeordnet und diskutiert wird.

Beiträge des Symposiums

Können Onlinelehrkräftefortbildungen andere Zielgruppen erreichen als Fortbildungen in Präsenz?

Nina Mulajmović¹, Eric Richter¹, Rebecca Lazarides², Dirk Richter¹

¹Universität Potsdam, ²Universität Potsdam & Technische Universität Berlin, Berlin, Deutschland

Theoretischer Hintergrund

Fortbildungsveranstaltungen sind essenziell für die berufliche Weiterentwicklung von Lehrkräften (Darling-Hammond et al., 2017; Desimone, 2009). Folglich sollte jede Lehrkraft die Möglichkeit haben, daran teilzunehmen. Allerdings ist bereits bekannt, dass fehlende Angebote in der Region und fehlende zeitliche Ressourcen Barrieren für die Teilnahme an Präsenzfortbildungen darstellen können (Krille et al., 2020). Diese Barrieren können möglicherweise durch Online-Fortbildungen überwunden werden, welche den Vorteil haben, dass sie räumlich und/oder zeitlich unabhängig sind (Bates et al., 2016). Inzwischen besteht auch das Fortbildungsangebot zunehmend aus Online-Fortbildungen (Koscheck et al., 2022). Bislang ist jedoch noch weitgehend unerforscht, ob Online-Fortbildungen tatsächlich dazu beitragen können, dass Lehrkräfte aus peripheren Regionen und mit wenigen zeitlichen Ressourcen stärker an Angeboten der Lehrkräftefortbildung partizipieren.

Forschungsfragen

Die vorliegende Studie adressiert das oben beschriebene Forschungsdesiderat und zielt darauf ab, das grundlegende Verständnis zur Teilnehmer:innenschaft von Präsenz- und Online-Fortbildungen zu erweitern. Dazu stellt sie sich die folgenden zwei Forschungsfragen:

Inwiefern nehmen Lehrkräfte, die eine eingeschränkte Möglichkeit zur Teilnahme an Präsenzfortbildungen haben, unterschiedlich stark an Online- und Präsenzfortbildungen teil?

Unterscheidet sich die Bedeutsamkeit einzelner Bedingungsfaktoren für die Fortbildungsteilnahme zwischen Präsenz- und Online-Fortbildungen?

Methode

Stichprobe: Die zugrundeliegenden Daten stammen aus TALIS 2018. Dementsprechend liegen zur Teilnahme an Präsenz- und Online-Fortbildungen sowie zu Teilnahmebarrieren Angaben von 261.426 Lehrkräften aus 15.672 Schulen aus 48 Ländern vor. Die Mehrheit der Lehrkräfte identifizierte sich als weiblich (68,8 %), arbeitete im Sekundarschulbereich (80,2 %) und verfügte über eine Berufserfahrung von mehr als fünf Jahren (81,5 %).

Instrumente: Zur Erfassung der Fortbildungsteilnahme wurden die Lehrkräfte geschlossen dazu befragt, ob sie in den letzten zwölf Monaten an Präsenz- bzw. Online-Fortbildungen teilgenommen haben. In Hinblick auf die Teilnahmebarrieren wurden sowohl lehrkräfte- als auch schulbezogene Merkmale berücksichtigt. Etwa wurden die Lehrkräfte geschlossen dazu befragt, ob Fortbildungen mit ihrem Arbeitsplan kollidieren oder sie aufgrund familiärer Verpflichtungen keine Zeit haben, um daran teilzunehmen. Zudem wurden Angaben zur Eingebundenheit im Kollegium sowie zu Schulstandort (ländlich vs. städtisch) und -größe (Anzahl der Schüler:innen) eingeholt.

Analyseverfahren: Zur Untersuchung der Forschungsfragen wurden Mehrebenen-Pfadmodelle geschätzt. Für die erste Forschungsfrage wurden unbeschränkte Modelle spezifiziert. Zur Beantwortung der zweiten Forschungsfrage erfolgte ein Vergleich von unbeschränkten und beschränkten Modellen mittels Chi-Quadrat-Differenztests.

Ergebnisse

Erste Ergebnisse zeigen signifikante Zusammenhänge zwischen der mit einer Fortbildungsteilnahme verbundenen Herausforderung „Kollision mit dem Arbeitsplan“ ($b^* = .01, p < .05$), der mit einer Fortbildungsteilnahme verbundenen Herausforderung „familiäre Verpflichtungen“ ($b^* = -.04, p < .01$), der Eingebundenheit im Kollegium ($b^* = .14, p < .01$), der Schulgröße ($b^* = -.09, p < .05$) und der Teilnahme an Präsenzfortbildungen. Jedoch sind die Effekte nach Cohen (1988) als sehr schwach einzuordnen. Weiterhin zeigt sich kein signifikanter Zusammenhang zwischen dem Schulstandort und der Teilnahme an Präsenzfortbildungen ($b^* = .01, p = .67$). Selbiges gilt für die Teilnahme an Online-Fortbildungen. Auch hier zeigen sich signifikante Zusammenhänge für die mit einer Fortbildungsteilnahme verbundene Herausforderung „Kollision mit dem Arbeitsplan“ ($b^* = .02, p < .01$), die mit einer Fortbildungsteilnahme verbundene Herausforderung „familiäre Verpflichtungen“ ($b^* = -.03, p < .01$), die Eingebundenheit im Kollegium ($b^* = .13, p < .01$) und die Schulgröße ($b^* = -.09, p < .01$) mit sehr schwachen Effektstärken sowie kein Zusammenhang für den Schulstandort ($b^* < .01, p = .79$). Ferner zeigt sich, dass sich die Bedeutsamkeit der untersuchten Merkmale nicht signifikant zwischen Präsenz- und Online-Fortbildungen unterscheidet. Diese Befunde deuten darauf hin, dass Online-Fortbildungen nicht, wie bisher angenommen, insbesondere jene Lehrkräfte erreichen, die nur unzureichende Möglichkeiten haben, um an Präsenzfortbildungen teilzunehmen (Krille et al., 2020).

Individuelle Nutzung einer Online-Fortbildung vor dem Hintergrund objektiver und subjektiver Lernziele

Karina Demmler¹, Marita Friesen², Lars Holzäpfel¹, Timo Leuders¹, Anika Dreher¹

¹Pädagogische Hochschule Freiburg, ²Pädagogische Hochschule Heidelberg

Theoretischer Hintergrund

Lehrkräfte, die an Fortbildungen teilnehmen, haben unterschiedliche Voraussetzungen und damit auch unterschiedliche Bedarfe und Ziele (Yoon & Kim, 2022). Daher werden verschiedene Angebotstypen wie Unterrichtsvideos, Materialien zur Unterrichtsplanung oder Erklärvideos zu didaktischen Hintergründen von Fortbildungsteilnehmenden als mehr oder weniger nützlich empfunden (Lipowsky, 2021). Um heterogene Bedürfnisse im Sinne einer Teilnehmendenorientierung bei der Gestaltung von Fortbildungen adressieren zu können, spielen insbesondere Online-Fortbildungen eine zentrale Rolle: Sie können von Lehrkräften dahingehend unterschiedlich genutzt werden, als dass sie eine individuelle Nutzung verschiedener Angebotstypen und mehr Flexibilität bei der zeitlichen Nutzung ermöglichen (Bragg et al., 2021; Friesen et al., 2023). Bei solchen flexiblen Nutzungsmöglichkeiten und der damit einhergehenden hohen Teilnehmendenorientierung besteht jedoch die Gefahr, dass die intendierten Lernziele im Sinne einer hohen Zielorientierung aus dem Blick geraten (Pöhler-Friedrich et al., eingereicht). Für die Entwicklung effektiver adaptiver Online-Fortbildungen ist es daher wichtig, die individuellen Nutzungsweisen von Lehrkräften hinsichtlich ihrer Passung zu subjektiven und objektiven Lernzielen eines Fortbildungsangebots besser zu verstehen.

Fragestellungen

Anhand einer Online-Fortbildung zum Thema Problemlösen im Mathematikunterricht untersuchen wir daher:

(FF1) Inwieweit nutzen Lehrkräfte verschiedene Angebotstypen im Sinne des Fortbildungsziels (objektive Passung) und im Sinne ihrer individuellen Lernziele (subjektive Passung)?

(FF2) Wie hängen diese Maße für lernzieltreue Nutzung untereinander und mit anderen bekannten Nutzungsvariablen (Nutzungsdauer, kollegialer Austausch und Erprobung) für unterschiedliche Angebotstypen zusammen?

Methode

Das Online-Fortbildungsangebot besteht aus drei inhaltlichen Kapiteln zum Problemlösen im Mathematikunterricht, in welchen jeweils die drei Angebotstypen Erklärvideo, Planungsmaterial und Unterrichtsvideo zwei Wochen lang flexibel asynchron genutzt werden konnten (Demmler et al., 2024). Nach jedem Kapitel fand eine synchrone Online-Sitzung zum kollegialen Austausch statt. In einem Fortbildungsjournal wurden Nutzungsdaten von N=125 Mathematiklehrkräften erfasst. Erfragt wurden beispielsweise die Anzahl der Unterrichtserprobungen und die Intensität des kollegialen Austauschs (6-stufige Likertskala). Für jeden Angebotstyp dokumentierten die Lehrkräfte neben der zeitlichen Nutzung ihre Take-Home-Messages und schätzten die wahrgenommene Nützlichkeit auf einer 6-stufigen Likertskala ein (subjektives Maß für lernzieltreue Nutzung). Die Passung der Take-Home-Messages zu den jeweiligen Lernzielen der Fortbildung (objektives Maß für lernzieltreue Nutzung) wurden von zwei Autorinnen auf vier Qualitätsstufen (0 bis 3) bewertet (Krippendorff's $\alpha = 0,84$). Beide Maße für lernzieltreue Nutzung konnten über die Kapitel hinweg zu Skalen mit akzeptabler bis guter Reliabilität zusammengefasst werden. Zur Beantwortung der Forschungsfragen wurden die Daten deskriptiv ausgewertet und Korrelationsanalysen durchgeführt.

Ergebnisse

Das objektive Maß für lernzieltreue Nutzung lag für alle Angebotstypen durchschnittlich zwischen der zweiten und dritten Qualitätsstufe (Erklärvideo: $M=1,72, SD=0,68$; Materialpaket: $M=1,35, SD=0,80$; Unterrichtsvideo: $M=1,65, SD=0,74$). Das subjektive Maß für lernzieltreue Nutzung wurde für alle Angebotstypen durchschnittlich sehr positiv eingeschätzt (Erklärvideo: $M=5,08, SD=0,76$; Materialpaket: $M=5,03, SD=0,79$; Unterrichtsvideo: $M=4,57, SD=1,09$). Für alle Angebotstypen zeigten sich schwache bis moderate Korrelationen zwischen den subjektiven und objektiven Maßen für lernzieltreue Nutzung. Diese fielen am höchsten beim Planungsmaterial ($r=0,30, p=0,001$) und am niedrigsten beim Unterrichtsvideo ($r=0,22, p=0,023$) aus. Ebenso zeigten sich für alle Angebotstypen schwache bis moderate Korrelationen zwischen der Nutzungsdauer und dem subjektiven Maß für lernzieltreue Nutzung, am stärksten für das Unterrichtsvideo ($r=0,31, p<0,001$). Die Nutzungsdauer korrelierte jedoch lediglich für das Erklärvideo ($r=0,22, p=0,020$) und Unterrichtsvideo ($r=0,22, p=0,024$) schwach mit dem objektiven Maß für

lernzieltreue Nutzung. Für das Planungsmaterial zeigte sich dagegen ein schwacher Zusammenhang zwischen der Nutzungsdauer und der Anzahl unterrichtlicher Erprobungen ($r=0,22$, $p=0,016$). Der kollegiale Austausch korrelierte moderat mit dem subjektiven Maß für lernzieltreue Nutzung des Planungsmaterials ($r=0,42$, $p<0,001$) und des Unterrichtsvideos ($r=0,33$, $p=0,006$). Insgesamt deuten die für alle Angebotstypen gefundenen Zusammenhänge zwischen den beiden Maßen für lernzieltreue Nutzung darauf hin, dass auch bei flexibel nutzbaren Online-Fortbildungsangeboten (im Sinne der Teilnehmendenorientierung) eine Zielorientierung aufrechterhalten werden kann. Die berichteten Zusammenhänge mit anderen Nutzungsvariablen liefern Hinweise für ihre Bedeutung hinsichtlich lernzieltreuer Nutzung verschiedener Angebotstypen in Online-Fortbildungen.

Nur Technologien oder auch Didaktik? Untersuchung der Präferenzen von Physiklehrkräften bei der Wahl von technologiebezogenen, adaptiven Online-Fortbildungen

Armin Fabian¹, Salome Flegr², Nina Udvardi-Lakos¹, David Weiler¹, Ulrich Trautwein¹, Sebastian Becker-Genschow³, Tim Fütterer¹

¹Universität Tübingen, ²Ludwig-Maximilians-Universität München, ³Universität zu Köln

Theoretischer Hintergrund

Digitale Medien haben ein großes Potenzial, Schülerinnen und Schüler dabei zu unterstützen, konzeptuelles Physikwissen zu erwerben (z.B. durch Simulationen; Flegr et al., 2023). Um das Potenzial jedoch auszuschöpfen, benötigen Lehrkräfte ein fundiertes technologiebezogenes Professionswissen (Mishra & Koehler, 2006). Dieses Professionswissen wird häufig durch das Technological Pedagogical and Content Knowledge Framework konzeptualisiert (Mishra & Koehler, 2006), welches die Integration verschiedener Wissensfacetten (z.B. technologisches und fachdidaktisches Wissen) zu einer eigenständigen Wissensfacette (d.h. TPCK) postuliert. Für die Förderung von TPCK sind Lehrkräftefortbildungen essentiell (Fabian et al., 2024), welche zunehmend auch als Online-Fortbildungen angeboten werden, um zeitliche und räumliche Ressourcen zu schonen (Dede et al., 2006; Fütterer et al., 2024). Verfügbare Online-Fortbildungen werden allerdings häufig angebotsorientiert konzipiert und bieten damit eine *one-size-fits-all* Lösung (Richter und Richter, 2020), gleichwohl Lehrkräfte unterschiedliche kognitive und motivationale Voraussetzungen haben. Gleichzeitig deuten Studien darauf hin, dass angebotene TPCK-Fortbildungen sehr heterogen sind bezüglich der adressierten Wissensfacetten (Fabian et al., 2024). Insgesamt ist somit weitestgehend ungeklärt, inwieweit Fortbildungsangebote mit den individuellen Voraussetzungen von Lehrkräften übereinstimmen, was vor dem Hintergrund zukünftiger adaptiver Fortbildungen kritisch zu sehen ist. Außerdem basiert bisherige Forschung häufig auf Fragebogendaten, sodass nur eingeschränkte Einsichten in das Nutzungsverhalten und damit den eigentlichen Lernprozess möglich sind.

Forschungsfragen

Um Einsichten in die Passung von Voraussetzungen und Fortbildungen zu TPCK zu erhalten, sowie Prädiktoren für das Nutzungsverhalten zu identifizieren, werden folgende Forschungsfragen (FF) untersucht:

FF 1a) Welche Online-Fortbildungen zum Einsatz digitaler Medien bevorzugen Physiklehrkräfte: Technologie-zentrierte Fortbildungen oder Fortbildungen, die konkret auf TPCK abzielen, also den didaktisch-elaborierten Einsatz digitaler Medien im physikspezifischen Unterrichtskontext schulen?

FF 1b) Wie hängen die individuellen motivationalen, kognitiven und metakognitiven Voraussetzungen der Physiklehrkräfte mit dieser Wahl zusammen?

FF 2a) Wie nutzen Physiklehrkräfte adaptiv gestaltete Online-Fortbildungen zum Einsatz digitaler Medien? Welche Nutzungsprofile gibt es?

FF 2b) Wie hängen die Nutzungsprofile mit den individuellen Voraussetzungen der Physiklehrkräfte zusammen?

Methode

Es handelt sich um eine korrelative Onlinestudie, die im August 2024 in Deutschland durchgeführt wird. Um ausreichende Daten zu sichern, werden neben Physiklehrkräften auch Physiklehramtsstudierende aus ganz Deutschland akquiriert. Wir erheben kognitive und metakognitive Variablen, sowie motivationale Einstellungen und Überzeugungen gegenüber dem Medieneinsatz im Physikunterricht. Außerdem erheben wir digitale Verhaltensspurdaten („log-data“) der Lehrkräfte (z. B. time-on-task oder Häufigkeit verwendeter Hilfestellungen). Zusätzlich erheben wir die Präferenz des Lernmoduls (technologie-fokussiert, pädagogisch-fokussiert oder TPCK-fokussiert). Für FF1 (Passung von Angebot und Nutzung) spezifizieren wir ein logistisches Regressionsmodell mit den Kompetenzfacetten als UV und der Wahl des Moduls als AV. Für FF2 (Identifikation von Nutzungsprofilen) führen wir latente Profilanalysen durch und untersuchen Zusammenhänge mit den Voraussetzungen durch Regressionsmodelle, wobei die Zugehörigkeit zu einem Profil die AV darstellt. Die gesamte Studie wird auf einem eigens entwickelten, KI-basierten Onlinesystem durchgeführt. Zu Beginn beantworten die Teilnehmenden Fragen zu den Kompetenzfacetten und wählen anschließend ihr bevorzugtes Lernmodul (i.e. Präferenz) basierend auf detaillierten Beschreibungen der Wissensfacetten nach Mishra und Koehler (2006). Während der Bearbeitung des gewählten Moduls, mit der Möglichkeit, später auch andere Module zu bearbeiten, werden im Hintergrund digitale Verhaltensspurdaten erfasst. Jedes Lernmodul besteht aus drei Teilen: einem Instruktionsvideo, in dem grundlegende Konzepte des Modulinhalts aufbereitet werden, vertiefende Materialien zum Nachschlagen, sowie Aufgaben, die die Teilnehmenden zur Reflexion anregen. Während der Bearbeitung steht den Teilnehmenden ein integrierter KI-basierter ChatBot zur Seite, den sie bei Nachfragen verwenden können.

Ergebnisse und Relevanz

Die Daten werden im August 2024 erhoben. Um eine ausreichende Stichprobengröße zu gewährleisten, wird die Studie über verschiedene Kanäle (Schulnetzwerke, E-Mailverteiler und Landesinstitute) disseminiert. Auf der GEBF werden zudem die Ergebnisse einer bereits durchgeführten Bedarfsanalyse mit Lehrkräften ($N = 78$) präsentiert, in der 82% angaben, sich adaptive Online-Fortbildungen zu wünschen – was die hohe Relevanz unserer Studie unterstreicht.

S-82: Technologiebasierte Diagnostik und Förderung von Schreibkompetenzen: Herausforderungen und Potenziale

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 151

Chair der Sitzung: Johanna Fleckenstein

Chair der Sitzung: Johannes Hartig

Chair der Sitzung: Vera Busse

Technologiebasierte Diagnostik und Förderung von Schreibkompetenzen: Herausforderungen und Potenziale

Chair(s): **Johanna Fleckenstein** (Universität Hildesheim, Deutschland), **Johannes Hartig** (DIPF Frankfurt)

Diskutant*in(nen): **Vera Busse** (Universität Münster)

Schreiben ist eine fächerübergreifende Kernkompetenz und eine wichtige Grundlage gesellschaftlicher und beruflicher Teilhabe (Gogolin & Duarte, 2016; Graham, 2019). Das Symposium beleuchtet die vielfältigen Aspekte und Dimensionen des Einsatzes digitaler Technologien beim Schreiben in verschiedenen Bildungsstufen von der Grundschule über die Sekundarstufe bis zur Universität. Im Mittelpunkt steht die Frage, wie digitale Technologien zeitgemäß genutzt werden können, um die Schreibkompetenzen von Lernenden zu diagnostizieren und zu fördern, sowie die damit verbundenen Herausforderungen und Potenziale.

Zentrale Herausforderungen für die zielführende Einbindung von technologiebasiertem Schreiben in Lehr-Lern-Settings sind hinreichende Kompetenzen von Schüler:innen beim Schreiben mit Tastatur sowie die Validität und Reliabilität digitaler Schreibaufgaben. In zwei Beiträgen wird daher beleuchtet, inwiefern der digitale Modus bei der Erfassung von Schreibkompetenzen sinnvoll einsetzbar ist und welche Unterschiede und Potenziale es im Vergleich zum analogen Schreiben gibt. So untersucht der erste Beitrag den Umgang von Dritt- und Viertklässler:innen mit Orthografie-Aufgaben in einem Tablet-Papier-Hybridformat wurden spezifische Fehlerquellen, Korrekturstrategien und der Zeitbedarf bei der Bearbeitung der Aufgaben untersucht. Die Studie liefert erste Erkenntnisse darüber, wie Testformate gestaltet werden können, um die Lern- und Leistungsfähigkeit der Schüler:innen zu unterstützen. Der zweite Beitrag befasst sich mit der Messinvarianz zwischen analog und digital geschriebenen Schülertexten in der Sekundarstufe. Diese Ergebnisse deuten darauf hin, dass die Schreibaufgabe grundsätzlich auch im digitalen Modus verwendet werden kann.

Ein großes Potenzial digitalgestützter Lehr-Lern-Settings liegt in der Prozessorientierung beim Schreiben. Zum einen können im Schreibprozess erfasste Daten Aufschluss über Verarbeitungsprozesse geben, zum anderen kann automatisiertes Feedback den Schreibprozess unterstützen. Zwei Beiträge beschäftigen sich mit Fragen rund um den Einsatz von Writing Analytics und (generativer) KI zur Diagnose und Förderung von Schreibkompetenzen. Im dritten Beitrag wird die Verwendung von Keylogs untersucht, um die Effektivität von automatisiertem Feedback beim Schreiben zu verstehen. In einer Studie mit Zehntklässler:innen wurde gezeigt, dass KI-Feedback die Textqualität verbessert und das behaviorale Engagement der Schüler:innen während der Textüberarbeitung erhöht. Der vierte Beitrag untersucht die Wirksamkeit von automatisiertem Feedback, generiert durch eine KI, bei der Überarbeitung wissenschaftlicher Texte von Studierenden. Die Ergebnisse zeigen, dass die Qualität der automatisierten Beurteilung hoch war und das darauf basierende Feedback zu signifikant besseren Schreibleistungen in der in der Feedbackgruppe führte als in der Kontrollgruppe.

Zusammenfassend zeigt das Symposium, dass der Einsatz technologiebasierter Ansätze zur Diagnostik und Förderung von Schreibkompetenzen sowohl erhebliche Potenziale als auch Herausforderungen birgt. Die Forschungsergebnisse bieten wertvolle Einblicke und praxisnahe Empfehlungen für die Integration digitaler Technologien in den Schreibunterricht und die Schreibförderung.

Beiträge des Symposiums

Hybride Testformate in der Grundschule: Eine Videostudie im Kompetenzbereich Orthografie

Seda Yilmaz Wörfel, Kati Pügner, Michael Krelle

TU Chemnitz

Die überarbeiteten Bildungsstandards betonen zunehmend die Bedeutung digitaler Kompetenzen im Schulalltag. Die Integration digitaler Medien in den Unterricht stellt jedoch sowohl Schüler:innen als auch Lehrkräfte vor neue Herausforderungen. Das Schreiben auf Papier und das Tippen auf einer Tastatur stellen unterschiedliche kognitive und motorische Anforderungen (Dürscheid/Brommer 2009). Insbesondere im Bereich der Orthografie können diese Unterschiede zu variierenden Fehlerquellen (Frahm 2012; Jung et al. 2021) und Korrekturstrategien (Jung et al. 2021) führen. Es ist daher wichtig zu untersuchen, i) wie Schüler:innen mit diesen neuen Anforderungen umgehen; ii) welche spezifischen Schwierigkeiten sie dabei haben; iii) wie sich diese Unterschiede auf die Leistung der Schüler:innen auswirken.

Vor diesem Hintergrund wurden zwei Studien geplant. Die erste Studie (HYBRID) untersucht im TBA-Format den Umgang von Dritt- und Viertklässler:innen mit Testaufgaben zur Orthografie in einer „Tablet-Papier-Hybridvariante“: Einige Aufgaben werden digital am Tablet bearbeitet, andere Aufgaben werden (graphomotorisch) auf Papier bearbeitet. Die zweite Studie (TBA) zielt darauf ab, das Verhalten der Schüler:innen bei technologiebasierten Aufgaben am Tablet insgesamt zu analysieren, einschließlich spezifischer Fehlerquellen und des Zeitbedarfs. Im Fokus stehen Lückentext-Aufgaben, die an das digitale Format angepasst sind. Die Daten wurden bei einer Grundschule mit insgesamt 90 Dritt- und Viertklässler:innen erhoben. Bildschirmaufzeichnungen bei digitaler Bearbeitung und Videoaufzeichnungen des Bearbeitungsgegenstands (Papier, Tablet-Bildschirm und Stift Tastatur) wurden von hinten oben gefilmt.

Im Vortrag werden die ersten Ergebnisse vorgestellt und die daraus gewonnenen Erkenntnisse, wie z.B. wie Testformate gestaltet werden sollten, um die Lern- und Leistungsfähigkeit der Schülerinnen zu unterstützen, diskutiert.

Ein Vergleich analoger und digitaler Schreibprodukte. Die Invarianz der MEZ-Schreibaufgabe Jugendliche zwischen den Testmodi „handschriftlich“ und „Tastatur“

Birger Schnoor, Irina Usanova

Universität Hamburg

Der aktuelle Beitrag befasst sich mit der Überprüfung der Messinvarianz zwischen analog und digital geschriebenen Schülertexten in der Sekundarstufe. Breite angelegte Studien zum digitalen Schreiben liegen kaum vor (Döring und Busse 2022).

Vorhandene Befunde liefern Hinweise auf substanzielle Unterschiede zwischen analog und digital produzierten Texten, die sich bei Letzteren in längeren Texten und besserer Textqualität niederschlagen (Goldberg et al. 2003; Morphy und Graham 2012; für einen Überblick siehe Döring und Busse 2022). Studien, die überprüfen, inwiefern sich Instrumente, die sich im analogen Testmodus bewährt haben, auch messinvariant im digitalen Testmodus sind, fehlen bislang.

Im Rahmen des Forschungsprojekts „Mehrsprachigkeitsentwicklung an der Schwelle zum Beruf (MEZ-2)“ (Gogolin et al., 2021) wurde in einer Teilstudie Unterschiede in der Messung von Schreibfähigkeit im Deutschen in der Sekundarstufe zwischen analogem (handschriftlich) und digitalem (Tastatur) Erhebungsmodus untersucht. Grundlage war die MEZ-Schreibaufgabe Jugendliche (Gogolin et al., 2022). Aufgabenstellung war das Schreiben eines Probeartikels für ein Jugendmagazin in Form eines Anleitungstextes, bei dem durch neun Fotos dokumentierte Arbeitsschritte des Baus eines Holzbummerangs zu beschreiben waren.

Für diesen Zweck wurden zu beiden Erhebungswellen von MEZ-2 an je einer zufälligen Teilstichprobe die deutschsprachigen Texte anstatt analog digital erhoben. Insgesamt stehen 196 digitale und 1228 analoge Textprodukte zur Verfügung.

Die Textprodukte wurden anhand eines generischen Modells für Schreibkompetenz ausgewertet, das drei Kompetenzbereiche umfasst: (1) textuell-pragmatische und (2) lexiko-syntaktische Qualifikationen sowie (3) Produktivität (Flüssigkeit) (siehe Abbildung 2). Die Kombination dieser Qualifikationen wird als einheitliches Konstrukt betrachtet (Mehta et al. 2005; Schnoor 2019), das die aufgabenspezifische Gesamtqualität der mit MEZ-Instrumenten erhobenen Texte widerspiegelt (Klinger et al. 2019; Schnoor 2019; Usanova 2019).

Die Überprüfung der Messinvarianz erfolgte anhand konfirmatorischer Faktorenanalysen (Brown, 2015; Millsap, 2011). Hierzu wurde das generische Kompetenzmodell in ein empirisches Messmodell übersetzt und auf faktorielle Invarianz zwischen den Testmodi getestet.

Im Ergebnis erwies sich das Messmodell als partiell skalar invariant zwischen den Testmodi, was bedeutet, dass die Kovarianzstruktur uneingeschränkt und die Mittelwertstruktur eingeschränkt vergleichbar ist. Somit ließ sich nur eine relativ geringe Abhängigkeit der MEZ-Schreibaufgabe Jugendliche vom Erhebungsmodus auffinden, und das Instrument bietet grundsätzlich die Option der digitalen Testung. Inwiefern sich diese Ergebnisse für das Deutsche auch auf die anderen Sprachversionen der MEZ-Schreibaufgabe Jugendliche übertragen lassen, ist Ziel weiterer Forschung.

Effekte Automatisierten Feedbacks verstehen: Schreibprozesse nutzen um die Rolle von Behavioralem Engagement aufzudecken

Ronia Schiller¹, Johanna Fleckenstein², Jennifer Meyer¹

¹IPN Kiel, ²Universität Hildesheim

Feedback kann Lernende in vielen Bereichen unterstützen (Hattie & Timperley, 2007). Jedoch ist Feedback nicht effektiv, wenn es nicht proaktiv genutzt und verarbeitet wird (Lipnevich & Smith, 2022; Winstone et al., 2017). Die aktuelle Feedbackforschung betrachtet Lernende zunehmend als aktiven Bestandteil des Feedbackprozesses, wobei das Lernenden-Engagement ein zentraler Begriff ist (Winstone & Nash, 2023). Bisherige Studien haben Lernenden-Engagement in Verbindung mit Feedback hauptsächlich mittels Selbstberichtfragebögen erfasst. Studien legen jedoch nahe, dass behaviorale Maße Lernenden-Engagement während spezifischer Lernaktivitäten – wie dem Arbeiten mit Feedback – genauer abbilden könnten als Selbstberichte (e.g., Choi et al., 2023; Fuller et al., 2018; Sinatra et al., 2015).

Der Bereich des Schreibens eignet sich besonders gut, um Lernenden-Engagement in Zusammenhang mit Feedback zu untersuchen, da Verhaltensprozesse bei der Textüberarbeitung als Manifestation behavioralen Engagements betrachtet werden können (e.g., Horbach et al., 2022; Nguyen et al., 2017). In einer früheren Studie konnten wir bereits zeigen, dass die Edit Distance und die Überarbeitungszeit als Indikatoren von behavioralem Engagement während der Textüberarbeitung Effekte automatisierten Feedbacks auf die Überarbeitungsleistung von Schüler*innen partiell oder vollständig mediieren. In der aktuellen Studie haben wir Keystroke-Logging genutzt und konnten so noch effizienter behaviorale Maße für Lernenden-Engagement während der Textüberarbeitung gewinnen. Keystroke-Logging-Daten haben den Vorteil, dass sie nicht nur oberflächliche Änderungen am Text sichtbar machen können, sondern Informationen über den Zeitpunkt jedes Tastenschlags sowie die Position des geänderten Zeichens im Text enthalten (Leijten & Van Waes, 2013).

Um anhand dieser Maße die Frage zu beantworten, welche Rolle behaviorales Engagement von Schüler*innen während der Textüberarbeitung für die Effektivität von automatisiertem Feedback spielt, haben wir ein Between-Subjects-Design in einer computer-basierten Studie implementiert. Dabei haben N = 453 Zehntklässler*innen Essays zu vorgegeben Themen auf Englisch geschrieben. Sie wurden randomisiert einer Feedback- oder einer Kein-Feedback-Kontrollbedingung zugewiesen. Abhängig von der Bedingung überarbeiteten alle Schüler*innen ihr Essay entweder mit Hilfe automatisierten Feedbacks oder ohne Feedback. Wir haben ein Large Language Models (d.h., GPT3.5-turbo) genutzt, um Schüler*innen in der Feedback-Bedingung innerhalb weniger Sekunden elaboriertes Feedback zu ihren Essayentwürfen bereitzustellen. Als Indikatoren für behaviorales Engagement haben wir die aktive Schreibzeit sowie die Gesamtzahl der Tastenschläge während der Textüberarbeitung herangezogen. Anschließend an die Textüberarbeitung haben die Schüler*innen ein zweites Essay zu einem anderen vorgegeben Thema geschrieben, um mögliche Lerneffekte zu untersuchen.

Wir haben separate Mediationsanalysen durchgeführt, in denen die Indikatoren für behaviorales Engagement jeweils die Mediatoren des Effekts von Feedback auf die Textqualität der Überarbeitung darstellten. Unter Kontrolle der Textqualität des Erstentwurfs erzielten Schüler*innen in der Feedback-Bedingung bei der Überarbeitung eine bessere Textqualität als Schüler*innen, die kein Feedback erhielten. Zudem zeigten sie höheres behaviorales Engagement bei der Textüberarbeitung. Das heißt, sie haben mehr Zeit genutzt, um ihr Essay aktiv zu überarbeiten und mehr Änderungen am Text vorgenommen, als Schüler*innen, die ohne Feedback überarbeitet haben.

Die Ergebnisse der Mediationsanalysen zeigten, dass sowohl die aktive Schreibzeit als auch die Gesamtzahl der Tastenschläge während der Überarbeitung den positiven Feedbackeffekt auf die Überarbeitungsleistung der Schüler*innen vollständig mediieren. Hinsichtlich der Textqualität des zweiten Essays haben wir positive Zusammenhänge mit dem behavioralen Engagement während der Überarbeitung des ersten Essays, aber keine Effekte des Feedbacks festgestellt.

Unsere Ergebnisse erweitern die aktuelle Studienlage zur Effektivität von automatisiertem Feedback beim Schreiben. Die Bedeutung einer prozessorientierten Herangehensweise bei der Förderung von Schreibfähigkeiten von Schüler*innen unter Verwendung von (automatisiertem) Feedback wird durch die Befunde der Studie hervorgehoben. In Forschung und Schulpraxis sollte nicht nur die Bereitstellung von Feedback und die Leistung, sondern vielmehr auch der Umgang von Schüler*innen mit Feedback im Fokus stehen, um sicherzustellen, dass sie ihre Schreibfähigkeiten weiterentwickeln können. Limitationen der Studie und Potential von Prozessmaßen im Kontext Schreiben werden diskutiert.

Titel: Automatisiertes Feedback und Schreiben wissenschaftlicher Texte: Können Large Language Models helfen?

Robert Köller¹, Johanna Fleckenstein², Fabian Schmidt², Fynn Junge¹, Julian Lohmann¹, Jens Möller¹

¹CAU Kiel, ²Universität Hildesheim

Das Schreiben und Überarbeiten wissenschaftlicher Texte ist ein kognitiv und motivational herausfordernder Prozess. Dieser Prozess kann mit Hilfe von Feedback gefördert werden. Allerdings ist die Bereitstellung von Feedback zu Texten aufwändig. Mit der Einführung von Large Language Models (LLM) stehen Werkzeuge zur Verfügung, die (wissenschaftliche) Texte beurteilen und automatisiertes Feedback generieren können. Die Wirksamkeit von automatisiertem Feedback zum Schreiben im Allgemeinen (Fleckenstein et al., 2023; Scherer et al., 2024) und LLM-generiertem Feedback in der Sekundarstufe im Speziellen (Meyer et al., 2024) konnte bereits gezeigt werden.

In der vorliegenden Studie wurde untersucht, ob LLM-generiertes Feedback die Überarbeitung von wissenschaftlichen Studierendentexten verbessert. Zudem wurde untersucht, wie reliabel das LLM-generierte Feedback ist. Zu diesem Zweck schrieb eine Stichprobe von N = 166 Psychologie- und Lehramtsstudierenden ein Abstract zu einem deutschsprachigen wissenschaftlichen Artikel. In der Feedbackgruppe wurden die Teilnehmenden gebeten, ihre Abstracts zu überarbeiten, nachdem sie mittels GPT-4 generiertes automatisiertes Feedback zu ihrem ersten Entwurf erhalten hatten. Das Feedback von GPT-4 beinhaltete fehlende oder unvollständige Bestandteile der individuellen Abstracts sowie allgemeine Hinweise zur Gestaltung eines Abstracts. Die Kontrollgruppe erhielt ebenfalls die Aufforderung, ihren Text zu überarbeiten, aber lediglich eine Erinnerung an die Aufgabenstellung. Die Textqualität der ersten Fassungen und der Überarbeitungen wurden von zwei geschulten menschlichen Beurteilerinnen kodiert (ICC: 0.84). Die Analyse der Korrektheit des KI-generierten Feedbacks ergab eine Übereinstimmungsquote von 96,13 % mit Expert*innenurteilen. Die Ergebnisse von Regressionsanalysen zeigten signifikante Effekte des Feedbacks auf die Überarbeitungsleistung ($d = 1,19$), unabhängig von Geschlecht, Alter, Studienfach und Abiturnote. Darüber hinaus hatte die Bearbeitungszeit einen signifikanten Einfluss auf die Textqualität.

Die Ergebnisse zeigen die Zuverlässigkeit und die Wirksamkeit von LLM-generiertem Feedback beim wissenschaftlichen Schreiben. Der Einsatz entsprechender Technologien zur Förderung von Schreibkompetenzen im Studium wird diskutiert.

S-83: Digitale Förderung von Kompetenzen im selbstregulierten Lernen – Konzeptionen, Chancen und Herausforderungen

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: M 003

Chair der Sitzung: Carmen Nadja Hirt

Chair der Sitzung: Mathias Mejeß

Chair der Sitzung: Dirk Ifenthaler

Digitale Förderung von Kompetenzen im selbstregulierten Lernen – Konzeptionen, Chancen und Herausforderungen

Chair(s): **Carmen Nadja Hirt** (Universität Zürich, Institut für Erziehungswissenschaft, Schweiz), **Mathias Mejeß** (Pädagogische Hochschule Zürich, Abteilung Professions- und Systemforschung, Schweiz)

Diskutant*in(nen): **Dirk Ifenthaler** (Universität Mannheim, Lehrstuhl für Wirtschaftspädagogik – Learning, Design, & Technology, Deutschland)

Selbstregulierte Lernende setzen sich Ziele für ihr Lernen und überwachen, regulieren und kontrollieren aktiv ihre Kognition, ihre Motivation und ihr Verhalten im Hinblick auf diese Ziele und die gegebenen Kontextbedingungen (Pintrich, 2000). Kompetenzen im Bereich des selbstregulierten Lernens (SRL) wirken sich positiv auf akademische Leistungen sowohl im schulischen als auch im universitären Kontext aus (Dent & Koenka, 2016; Richardson et al., 2012). Lernende sehen sich jedoch nach wie vor mit großen Herausforderungen konfrontiert, wenn es darum geht, ihre Lernprozesse selbst zu regulieren (Blasiman et al., 2017; McCabe, 2011). Deshalb ist es zentral, die SRL-Kompetenzen von Lernenden zu fördern. Eine solche Förderung kann durch direkte (z. B. explizite Instruktion, Feedback), indirekte (z. B. Wahlmöglichkeiten, Lerntagebücher) und transferbezogene Maßnahmen (z. B. Anwendung auf weitere Lernaufgaben/-situationen) erfolgen (Cleary & Zimmerman, 2004; De Smul et al., 2018; Dignath & Veenman, 2021; Fischer et al., 2020; Sweller & Paas, 2017).

Trotz der Möglichkeit, die SRL-Kompetenzen von Lernenden erfolgreich zu fördern, findet eine Förderung in Schulen und Universitäten bislang selten statt (Dignath & Büttner, 2018; Dignath & Veenmann, 2021; Klug et al., 2014). Entsprechend wurden in den vergangenen Jahren unterschiedliche analoge und digitale SRL-Fördermassnahmen und -interventionen konzipiert und empirisch untersucht. Dabei hat sich gezeigt, dass analoge und digitale Förderansätze gleichermaßen effektiv sein können (Dynarski et al., 2007; Sitzmann et al., 2006; van der Beek et al., 2019), wobei jedoch die zunehmende Digitalisierung von Bildungsinstitutionen neue Möglichkeiten zur Förderung von SRL eröffnet (Horz & Schulze-Vorberg, 2017). Offen bleibt die Frage, wie diese neuen digitalen Möglichkeiten aktuell eingesetzt und künftig genutzt werden können, um ihr Potenzial zur Förderung des SRL von Lernenden optimal auszuschöpfen.

Ziel des Symposiums ist es, die Konzeption und Wirksamkeit verschiedener digitaler Fördermassnahmen zum SRL zu präsentieren, deren Chancen und Herausforderungen zu diskutieren und darauf basierend Empfehlungen für Forschung und Praxis abzuleiten. Der erste Beitrag fokussiert den Einfluss der wahrgenommenen Wertschätzung durch die Lehrkraft auf u.a. das SRL von $N = 141$ Schüler*innen der Klassenstufen 5 bis 9, welcher durch den Einsatz digitaler Lerntagebücher untersucht wird. Der zweite Beitrag befasst sich mit der Konzeption und Wirksamkeit eines herausforderungsbasiertes, personalisiertes und gamifiziertes Online-Pilottrainings zur Förderung des SRL von $N = 112$ Studierenden. Der dritte Beitrag beschäftigt sich mit der Frage, wie sich verschiedene Arten von Feedback durch eine adaptive Lerntechnologie auf die verschiedenen Phasen des SRL-Prozesses bei $N = 194$ Studierenden auswirken.

Die drei Beiträge befassen sich auf unterschiedliche Weise mit der direkten, indirekten und transferbezogenen digitalen Förderung von SRL für verschiedene Zielgruppen in unterschiedlichen Kontexten. Aus den jeweiligen Förderkonzepten und Forschungsdesigns ergeben sich vielfältige Chancen und Herausforderungen, die im Rahmen des Symposiums durch Prof. Dr. Dirk Ifenthaler diskutiert werden.

Beiträge des Symposiums

Der Zusammenhang zwischen wahrgenommener Wertschätzung und selbstreguliertem Lernen: Eine Tagebuchstudie mit Schüler*innen der 5.-9. Klasse

Maximilian Resch, Franziska Zmatlik, Marlene Heilmann, Franziska Kunert, Henrik Bellhäuser
Johannes Gutenberg Universität Mainz, Psychologie in den Bildungswissenschaften, Deutschland

Selbstreguliertes Lernen (SRL) gilt als Schlüsselkompetenz, um Bildungsziele zu erreichen und lebenslang zu lernen. Dabei zeigt sich, dass sich die verschiedenen Teilkomponenten von SRL, wie beispielsweise Zielsetzung, Planung, Motivation, Prokrastination, Zielerreichung und Zufriedenheit, von Tag zu Tag gegenseitig beeinflussen (Bellhäuser et al., 2021). Lerntagebücher stellen eine wirksame Methode dar, um die Lernleistung zu verbessern, indem sie die Reflexion und das Monitoring des Lernverhaltens fördern (Dignath et al., 2023).

Ein bisher kaum beachteter Aspekt in der SRL-Forschung ist die Rolle der wahrgenommenen Wertschätzung. Die Stress-as-Offense-to-Self-Theorie (SOS; Semmer et al., 2019) postuliert, dass erfahrene Wertschätzung zur Steigerung eines positiven Selbstwertes und zur Reduktion von Stress beiträgt. Mit dem Ziel der kongruenten und vergleichbaren Definition und Operationalisierung von Wertschätzung, haben Resch & Bellhäuser (2024) das Konstrukt Experienced Appreciation in Social Interactions entwickelt, das unterschiedliche Manifestationen von Wertschätzung in einem integrativen Konstrukt vereint. Innerhalb von zwei Mehr-Studien-Designs konnten die Annahmen der SOS im Arbeitskontext bestätigt werden (Resch & Bellhäuser, 2024; Resch, Nagelmann & Bellhäuser, 2024). Im Schulkontext liegen derzeit noch keine vergleichbaren Studien vor.

Die vorliegende Studie untersucht den Einfluss der täglichen wahrgenommenen Wertschätzung auf das SRL, die Schulzufriedenheit und das Wohlbefinden von Schüler*innen. Es wird vermutet, dass Schüler*innen, die sich wertgeschätzt fühlen, motivierter und zufriedener lernen, sowie höheres Wohlbefinden und einen höheren Selbstwert berichten. Je mehr Wertschätzung durch die Lehrkraft empfunden wird, desto weniger Stress und desto weniger Prokrastination sollten Schüler*innen erleben.

Im vorliegenden Projekt wurden digitale Lerntagebücher eingesetzt, um den Einfluss von täglicher wahrgenommener Wertschätzung auf das SRL zu erfassen. Schüler*innen der Klassenstufen 5 bis 9 dokumentierten dazu täglich in ihren Lerntagebüchern, wie wertgeschätzt sie sich durch ihre Lehrkräfte fühlten und wie ihr Lernverhalten in Bezug auf die genannten SRL-Komponenten war. Die Tagebuchstudie wurde längsschnittlich an mehreren Gymnasien in Rheinland-Pfalz durchgeführt. Während einer vierwöchigen Erhebungsphase füllten die Schüler*innen täglich ein kurzes Lerntagebuch sowie zu Beginn und Ende einen ausführlichen Fragebogen aus.

Erste Analysen mit Multilevel Modellen mit $N = 141$ Schüler*innen zeigen, dass die tägliche wahrgenommene Wertschätzung des Vortags signifikant die intrinsische Motivation ($\beta = 0.21, p < .001$), die Schulzufriedenheit ($\beta = 0.12, p = .002$), das Wohlbefinden ($\beta = 0.24, p < .001$), die Selbstwirksamkeit ($\beta = 0.15, p < .001$) sowie die Volition ($\beta = 0.09, p = .005$) am Folgetag voraussagt. Höhere Werte in der wahrgenommenen Wertschätzung sagten ein allgemein höheres SRL ($\beta = 0.07, p = .001$), sowie geringeres Stressempfinden ($\beta = -0.07, p = .025$) am Folgetag vorher.

Die vorliegende Studie erweitert die bisherige SRL-Forschung, indem sie die Rolle der Wertschätzung als neuen Einflussfaktor untersucht und zeigt, dass die Prinzipien der SOS-Theorie auch auf Schüler*innen angewendet werden können. Die Ergebnisse legen nahe, dass Interventionen zur Steigerung der wahrgenommenen Wertschätzung im schulischen Umfeld das Potential haben, das selbstregulierte Lernen zu fördern und somit zu besseren schulischen Leistungen und höherem Wohlbefinden der Schüler*innen beizutragen.

Ein Online-Training im selbstregulierten Lernen für Hochschulstudierende. Trainingskonzeption und Pilottraining

Carmen Nadja Hirt¹, Kerstin Bäuerlein², Yves Karlen¹

¹Universität Zürich, Institut für Erziehungswissenschaft, Schweiz, ²Pädagogische Hochschule FHNW, Institut Sekundarstufe I und II, Schweiz

Das Studium birgt vielfältige Lernsituationen und Herausforderungen, die hohe Anforderungen an das selbstregulierte Lernen (SRL) stellen (z. B. Prüfungsvorbereitung, Verfassen schriftlicher Arbeiten; Dresel et al., 2015). Die Studierenden müssen ihre Lernprozesse planen, überwachen und regulieren, um ihre Ziele zu erreichen. Allerdings sind viele Studierende unzureichend auf diese Herausforderungen vorbereitet, was sich beispielsweise in der ineffektiven Nutzung von Lernstrategien äussert (Blasiman et al., 2017). Daher benötigen Studierende Unterstützung beim Erwerb und der Weiterentwicklung ihrer SRL-Kompetenzen (z. B. Rinas et al., 2022).

Durch Trainingsmassnahmen lassen sich SRL-Kompetenzen und Lernergebnisse verbessern – in Präsenzkursen ebenso wie in digitalen Trainings. Letztere bieten zusätzliche Personalisierungsmöglichkeiten (Martin & Betrus, 2019) sowie räumliche und zeitliche Flexibilität (van der Beek et al., 2019). Allerdings besteht bei digitalen Trainings ein erhöhtes Risiko des Trainingsabbruchs (z. B. Dubey & Piroška, 2019). Einen Lösungsansatz dafür bietet die Gamification, wobei Game-design Elemente in nicht-spielerischen Kontexten eingesetzt werden (Deterding et al., 2011), um die Motivation zu erhöhen (Sailer et al., 2017).

Bisherige Online-Trainings zum SRL sind oft fach- oder inhaltsabhängig (Foerst et al., 2019) und/oder geben meist einen festen Inhalts- und Anwendungsplan vor (Lobos et al., 2021). Eine Abstimmung auf die individuellen Herausforderungen und Bedürfnisse der Studierenden wäre jedoch für deren Lernerfolg zentral (z. B. Dresel et al., 2015).

Auf diesen Erkenntnissen basierend wurde daher ein Online-Training für Hochschulstudierende unterschiedlicher Fachrichtungen konzipiert. Im Rahmen eines Pilotprojekts (2022-2023) wurde ein Pilottraining umgesetzt, das von der realen Herausforderung der Prüfungsvorbereitung ausgeht und auf die metakognitiven Strategien Ziele setzen und Planen fokussiert. Neben der Online-Plattform wurden Trainings- (Erklärvideos mit virtuellen Lernenden als Modelle und Trainingsbegleiter*innen, Übungen, Quiz etc.) und Evaluationselemente (Wissenstests, Fragebögen etc.) entwickelt. Direkte, indirekte und transferbezogene Trainingsmassnahmen knüpfen an bisherige Kompetenzen der Studierenden an und sind mit Gamification-Elementen angereichert (z. B. Levels, Badges).

Anhand eines experimentellen Prä-Post-Designs wurde untersucht, (1) wie effektiv das Online-Training die metakognitiven Strategien fördert und (2) wie die Studierenden Inhalt, Struktur, Design und Gamification-Elemente beurteilen. Einhundertzwölf angemeldete Studierende wurden randomisiert der Experimental- (EG) oder Wartekontrollgruppe (KG) zugeteilt. Davon nahmen $n = 55$ vollständig an der Studie teil ($n_{EG} = 19, n_{KG} = 36$). Die Studierenden (67,3 % weiblich) waren im Durchschnitt $M = 25,27$ Jahre alt ($SD = 5,15$). 90,9 % studierten in einem Bachelor-Studiengang im Durchschnitt im $M = 3,64$ Semester ($SD = 1,67$). Es gab keine signifikanten Unterschiede zwischen der EG und KG in Bezug auf Geschlecht, Alter, Studiengang und Semesterzahl. Die Experimentalgruppe trainierte durchschnittlich $M = 2.76$ Stunden ($SD = 2.02$). Die Ergebnisse zeigen eine hohe Zufriedenheit mit dem Training (Nützlichkeit für Prüfungsvorbereitung, Inhaltsqualität, Benutzerfreundlichkeit, persönliche Relevanz) sowie eine hohe subjektiv eingeschätzte Nützlichkeit der Trainingselemente (z. B. Erklär-/Modellierungsvideos, Levelstruktur, Standortbestimmung) für den Wissenserwerb und die Motivation. Zudem zeigen mixed ANOVAs signifikant positive Trainingseffekte auf das deklarative und prozedurale metakognitive Strategiewissen der Studierenden und die Strategienutzung im Studium.

Neben der Konzeption des fachunabhängigen Trainings, das auf reale Herausforderungen der Studierenden unterschiedlicher Studienrichtungen anwendbar ist, erwiesen sich hinsichtlich der Evaluation insbesondere das Finden eines geeigneten Erhebungszeitraums und die Stichprobenrekrutierung als herausfordernd, da Studierende verschiedener Hochschulen mit unterschiedlichen Prüfungszeiten Berücksichtigung finden sollten. Trotz umfangreicher Werbemassnahmen wurde nur eine kleine Stichprobe gewonnen, welche sich durch einen hohen Dropout nach dem Prä-Test (vor Trainingsstart) weiter reduzierte. Möglicherweise war der Aufwand für viele Studierende während der Prüfungsvorbereitung zu hoch. Um ökologische Validität zu gewährleisten, erschien es jedoch wichtig, die Studie während der Prüfungsvorbereitungsphase durchzuführen.

Die Studie unterstreicht das grosse Potential eines herausforderungsbasierten, gamifizierten und personalisierten Online-Trainings im SRL für Studierende, denn das Online-Training erwies sich trotz kurzer durchschnittlicher Trainingszeit als zielführend. Zugleich sind jedoch für künftige Projekte Massnahmen zur Rekrutierung sowie zur Erhöhung des Commitments von Teilnehmenden zu diskutieren.

Selbstreguliertes Lernen durch Bildungstechnologien fördern: Eine differenzierte Sicht auf die Rolle von Feedback

Mathias Mejez¹, Yvonne Hemmler²

¹Pädagogische Hochschule Zürich, Abteilung Professions- und Systemforschung, Schweiz, ²Universität Mannheim, Fakultät für Betriebswirtschaftslehre, Deutschland

Die Integration digitaler Technologien in das Bildungswesen hat in den letzten 30 Jahren erheblich zugenommen (Mirata et al., 2020). Dennoch besteht dringender Bedarf, besser zu verstehen, wie Lernende durch Bildungstechnologien, insbesondere adaptive Lerntechnologien (ALT), effektiv in ihrem SRL unterstützt werden können (Broadbent & Poon, 2015). Der vorliegenden Studie liegt das Selbstregulationsprozessmodell des Lernens nach Schmitz & Wiese (2006) zugrunde, das SRL in präaktionale, aktionale und postaktionale Phasen unterteilt. Feedback, das die Aufgaben-, Prozess- und Selbstregulationsebene adressiert,

kann das Lernen der Lernenden verbessern, indem es ihnen hilft, Optimierungspotenziale besser zu erkennen und zu nutzen (Wisniewski et al., 2020). Theobald und Bellhäuser (2022) haben die Auswirkungen von drei verschiedenen Arten von Feedback auf SRL verglichen: (1) direktes, (2) informatives und (3) transformatives Feedback. Wir vermuten, dass diese positiven Effekte von Feedback durch den Einsatz von ALT verstärkt werden können. ALT sind digitale Lernumgebungen, die kontinuierlich Daten über Lernende analysieren, um interaktiv auf deren Aktivitäten zu reagieren und Lernprozesse personalisiert zu unterstützen (Xie et al., 2019). ALT stellen somit eine vielversprechende Möglichkeit für die Bereitstellung von personalisiertem und zeitnahe Feedback dar (Schmid et al., 2022).

Trotz dieser Entwicklungen besteht weiterhin die Frage, wie differenziertes Feedback durch ALT den SRL-Prozess beeinflusst. Die vorliegende Studie zielt darauf ab, die Auswirkungen von Feedback auf den SRL-Prozess in Abhängigkeit von der Art des erhaltenen Feedbacks weiter zu untersuchen (Theobald & Bellhäuser, 2022). An der Studie nahmen $N = 194$ Studierende mit einem Durchschnittsalter von 23,5 Jahren ($SD = 5,6$) teil (74,2% weiblich). Im Rahmen des 12-wöchigen Kurses erhielten die Studierenden die Aufgabe, wöchentliche Hausaufgaben mit Hilfe einer ALT, die zur Förderung von SRL entwickelt wurde, zu bearbeiten (Autor:innen, 2024). Basierend auf den Antworten der Lernenden wurden drei verschiedene Formen personalisierten Feedbacks angeboten: (1) personalisierte Vorschläge für SRL-Strategien (direktives Feedback), (2) ein Dashboard, das die individuelle Entwicklung des SRL der Lernenden über die Zeit anzeigte (feedback) und (3) eine Kombination aus Vorschlägen und Dashboard (transformatives Feedback).

Die Mehrebenen-Modelle zeigen zwischen prä-aktionaler und aktionaler Phase positive Zusammenhänge zwischen Planung ($b = 0.20$, $p < 0.001$), Aufgabenwert ($b = 0.13$, $p = 0.035$), Selbstwirksamkeit ($b = 0.10$, $p = 0.034$), Zielorientierung ($b = 0.12$, $p = 0.007$), Freude ($b = 0.09$, $p = 0.007$) sowie Hoffnung ($b = 0.09$, $p = 0.066$) und Selbstregulation. Ärger ($b = -0.07$, $p = 0.007$) und Angst ($b = -0.11$, $p = 0.015$) hängen negativ mit Selbstregulation zusammen. Freude und transformatives Feedback interagieren negativ ($b = -0.08$, $p < 0.001$). Ärger und direktes Feedback interagieren positiv miteinander ($b = 0.06$, $p < 0.090$). Weiter zeigen die Analysen positive Zusammenhänge zwischen der Zufriedenheit in der postaktionalen Phase und dem Aufgabenwert ($b = 0.10$, $p = 0.012$), der Selbstwirksamkeit ($b = 0.26$, $p < 0.001$), der Freude ($b = 0.07$, $p = 0.019$) und der Hoffnung ($b = 0.27$, $p < 0.001$) in der anschließenden präaktionalen Phase, während Ärger ($b = -0.09$, $p = 0.024$) und Angst ($b = -0.14$, $p = 0.001$) negativ durch die Zufriedenheit beeinflusst werden. Informatives Feedback verstärkt den Zusammenhang zwischen Zufriedenheit und Planung ($b = 0.09$, $p = 0.071$), während transformatives Feedback den positiven Zusammenhang zwischen Zufriedenheit und Aufgabenwert ($b = -0.07$, $p = 0.079$) unerwartet reduziert. Gleichzeitig schwächt direktes Feedback den positiven Zusammenhang zwischen Zufriedenheit und Hoffnung ($b = -0.06$, $p = 0.080$) unerwartet ab.

Die vorliegende Studie zeigt, dass die Art des Feedbacks durch ALT die Motivation, metakognitive Aktivität und die emotionale Lage von Lernenden auf unterschiedliche Weise beeinflussen kann. Während der iterative Zyklus des SRL bestätigt wird, variiert die Wirksamkeit des Feedbacks je nach Komponenten und Phasen des SRL, was den Bedarf an noch individuelleren Feedbackmechanismen unterstreicht.

S-84: Emerging Pathways: Gelingendes Studieren in Fernstudiengängen

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: O 131

Chair der Sitzung: Stefanie Habermann

Chair der Sitzung: Michael Hast

Emerging Pathways: Gelingendes Studieren in Fernstudiengängen

Chair(s): **Stefanie Habermann** (IU Internationale Hochschule), **Michael Hast** (IU Internationale Hochschule, Germany)

Die Hochschullandschaft durchläuft derzeit einen Wandel, da sich immer mehr Studierende für ein Fernstudium entscheiden (Statista Market Insights, 2024). Insbesondere nicht-traditionelle Studierende, die in klassischen Präsenzstudiengängen unterrepräsentiert sind, tendieren zu diesem Studienformat, da es ihnen den Zugang zu höherer Bildung häufig erst ermöglicht (Kerst & Wolter, 2022). Denn Fernstudienangebote richten sich beispielsweise an Personen ohne klassische Hochschulzugangsberechtigung und eignen sich aufgrund ihrer hohen Flexibilität insbesondere für Menschen mit beruflichen und familiären Verpflichtungen (Klinke & Pundt, 2022). Fernstudiengänge bieten somit Bildungschancen, helfen bestehende Bildungsungleichheiten abzubauen und liefern eine „Antwort auf gesellschaftliche, soziale und demografische Veränderungen und Herausforderungen“ (Zawacki-Richter & Stöter, 2020, S. 300). Gleichzeitig sind die Erfolgsquoten von Fernstudierenden geringer als von Präsenzstudierenden (Fojtík, 2018) und sie neigen häufiger dazu, ihr Studium vor Erreichen des Abschlusses abzubrechen (Dahm, 2022). Da das Fernstudium zunehmend an Bedeutung gewinnt, beschäftigen sich derzeit verschiedene Forschungsprojekte mit unterschiedlichen Aspekten dieses Themas. Unter der Fragestellung: „Wie kann Studieren im Fernstudium gelingen?“ werden im Rahmen des Symposiums einige dieser Projekte eingeladen, ihre Forschungsergebnisse zu präsentieren und zu diskutieren. Dabei wird eine interdisziplinäre Perspektive eingenommen, da die in den Projekten tätigen Wissenschaftler:innen aus den Disziplinen Psychologie, Erziehungs-/Bildungswissenschaft, Politikwissenschaft, Soziologie und Sozialer Arbeit stammen.

Der Einzelbeitrag „Übergang ins Fernstudium erfolgreich meistern: Chancen nutzen und Herausforderungen bewältigen“ präsentiert Chancen und Herausforderungen, denen Fernstudienanfänger:innen in der kritischen Phase des Studieneingangs begegnen. Studentische Bewältigungsstrategien, die mittels eines systematischen Literaturreviews und Fokusgruppen identifiziert wurden, werden vorgestellt. Der Einzelbeitrag „Eingangsphase des Fernstudiums: Unterstützungsangebote und Perspektiven“ analysiert die Wahrnehmung von Unterstützungsmaßnahmen durch Fernstudierende zu Beginn ihres Studiums. Er umfasst Ergebnisse eines systematischen Literaturreviews und ersten Fokusgruppendifkussionen, um Einblicke in die Unterstützungssysteme und Bedürfnisse nicht-traditioneller Studierender zu gewinnen und Empfehlungen für die Verbesserung der Studieneingangsphase zu entwickeln. Der Fokus des Einzelbeitrags „Studentische Identität im Fernstudium in Abgrenzung zu traditionellen Studienformaten“ liegt auf der Analyse der Unterschiede in der persönlichen und sozialen Identität sowie der impliziten Identifikation zwischen Fernstudierenden und Präsenzstudierenden und basiert auf Ergebnissen einer Fragebogenuntersuchung sowie eines impliziten Assoziationstests. Zudem wird erforscht, wie diese Identitätsaspekte Abbruchintentionen und Studienerfolg beeinflussen. Der letzte Einzelbeitrag „Lernunterstützung (First-Generation-)Studierender im Fernstudium – Unterstützungspraxen an privaten Hochschulen im Spannungsfeld zwischen Homogenitäts- und Diversitätsorientierung“ untersucht mittels Gruppendiskussionen, teilnehmenden Beobachtungen und Dokumentenanalyse, wie private Hochschulen Unterstützung für First-Generation-Studierende im Fernstudium gestalten. Dabei zeigt sich, dass Studienbetreuende eine diversitätsorientierte Unterstützung bieten, während Lehrende eher homogenitätsorientierte Ansätze verfolgen. Der Beitrag bietet Hinweise für die Anpassung der Unterstützungsstrukturen und deren Bedeutung für Bildungsgerechtigkeit.

Beiträge des Symposiums

Übergang ins Fernstudium erfolgreich meistern: Chancen nutzen und Herausforderungen bewältigen

Michael Hast, Caroline Trautwein, Stefanie Habermann, Roxana Schweighart, Julian Rebien, Anna Maria Pampel
IU Internationale Hochschule

Theoretischer Hintergrund

Die Beliebtheit von Fernstudienangeboten wächst weltweit, mit 11 Millionen Studierenden im Jahr 2017, 31 Millionen im Jahr 2024 sowie einer prognostizierten Zahl von 47 Millionen Studierenden im Jahr 2028 (Statista Market Insights, 2024). Flexible Studienangebote und alternative Hochschulzugangsberechtigungen bauen Barrieren ab und bieten bisher unerreichten Studierendengruppen die Möglichkeit ein Studium zu absolvieren. Gleichzeitig sind die Studienleistungen von Fernstudierenden im Vergleich zu Präsenzstudierenden schlechter (Fojtík, 2018) und die Abbruchquoten höher (Dahm, 2022). Vor allem die Anfangsphase stellt diesbezüglich eine kritische Phase dar, in der Studierende besonders häufig die Entscheidung treffen, ihr Studium zu beenden (Heublein, 2014; Utami et al., 2020). Im Studieneingang begegnen den Studierenden zahlreiche persönliche, organisationale, soziale und inhaltliche Anforderungen (Trautwein & Bosse, 2017). Gelingt es den Studierenden Strategien zu entwickeln, um diese kritischen Anforderungen zu bewältigen, wird gelingendes Studieren wahrscheinlich (Trautwein & Bosse, 2017). Allerdings liegen derzeit nur wenige und vor allem keine systematisierten Ergebnisse zur kritischen Eingangsphase im Fernstudium vor.

Fragestellung

Dieser Beitrag untersucht die Frage: „Welchen Chancen und Herausforderungen begegnen Fernstudierende in der Studieneingangsphase und wie gehen sie mit diesen um?“

Methode

Es werden Ergebnisse aus zwei Einzelstudien – einem Scoping Review und Fokusgruppen – vorgestellt. Das Scoping Review wurde zur Beantwortung der Frage „Was ist über den Übergang ins Fernstudium bekannt?“ durchgeführt. Dazu wurden zwischen November 2023 und Februar 2024 systematisch fünf Datenbanken (ERIC, PubMed, Google Scholar, PsycINFO, and Scopus) anhand von deutschen und englischen Schlüsselbegriffen durchsucht, wobei 1534 Quellen identifiziert wurden (Schweighart et al., 2024). Nach einem mehrstufigen Screeningverfahren durch drei Personen wurden 60 Quellen in das Review eingeschlossen und mittels inhaltlich-strukturierender Inhaltsanalyse (Kuckartz & Rädiker, 2022) ausgewertet. Die Review-Ergebnisse werden mit Fokus auf die Möglichkeiten und Herausforderungen im Fernstudieneingang präsentiert. Basierend auf diesen Ergebnissen werden im Herbst 2024 Fokusgruppen mit Fernstudienanfänger:innen durchgeführt, um mit Blick auf den Studieneingang neben Möglichkeiten, Ressourcen und Herausforderungen insbesondere Bewältigungsstrategien zu erheben. Insgesamt sind acht Fokusgruppen mit jeweils etwa sechs Studierenden an den beiden größten Hochschulen Deutschlands mit

Fernstudienangeboten (IU Internationale Hochschule und FernUniversität in Hagen) geplant. Die Ergebnisse werden im Anschluss an die Transkription mittels Qualitativer Inhaltsanalyse ausgewertet. Neben den bereits jetzt vorliegenden Ergebnissen aus dem Scoping Review werden auf der Tagung erste Ergebnisse der Fokusgruppen vorgestellt.

Ergebnisse und ihre Bedeutung

Laut den Ergebnissen des Scoping-Reviews bietet der Übergang ins Fernstudium für Studienanfänger:innen zahlreiche Chancen, darunter hohe Flexibilität, Unterstützung und Anleitung durch Lehrende, positive Lernerfahrungen, die Stärkung eigener Fähigkeiten sowie wertvolle Interaktionen mit Kommiliton:innen und Lehrenden. Gleichzeitig birgt diese Phase auch Herausforderungen. Fernstudienanfänger:innen berichten häufig von überfordernder Arbeitsbelastung, fehlender Anleitung durch Lehrende, mangelnder Interaktion und Austausch, Schwierigkeiten bei der Navigation der Lernplattformen und mentalen Belastungen wie Einsamkeit oder Ängste. Die Erfahrungen von Anfänger:innen im Fernstudium unterscheiden sich dabei von Präsenzstudierenden, was sich in höheren Abbruchquoten und niedrigeren Erfolgsquoten niederschlägt (Dahm, 2022; Fojtík, 2018; Xavier et al., 2022). Die Untersuchung der spezifischen Möglichkeiten, Herausforderungen und Bewältigungsstrategien im Kontext des Fernstudiums ist von großer Bedeutung und liefert wertvolle Erkenntnisse für Fernhochschulen, deren Mitarbeitende sowie Fernstudierende. Diese Erkenntnisse ermöglichen die Entwicklung gezielter Angebote und Interventionen, die Studierende beim erfolgreichen Übergang ins Fernstudium unterstützen können. Darüber hinaus trägt das Forschungsprojekt dazu bei, Bildungsbarrieren für Menschen – unabhängig von ihrer sozialen Herkunft oder Lebenssituation – abzubauen und die Zugänglichkeit sowie die Qualität von Fernstudienangeboten zu verbessern.

Eingangsphase des Fernstudiums: Unterstützungsangebote und Perspektiven

Sabrina Schaper, Jennifer Lange

FernUniversität in Hagen

Der Beitrag befasst sich aus empirischer Perspektive mit studentischen Perspektiven auf die Eingangsphase des Fernstudiums. Konkret stehen digitale (Lern-) Infrastrukturen, Gestaltungsansätze und konkrete Unterstützungsangebote ebenso im Fokus wie die Frage nach möglichen spezifischen Anforderungen sogenannter nicht-traditioneller Studierender (NTS), die als zentrale Akteursgruppe des Fernstudiums gelten. Damit schließt das Projekt innerhalb des Forschungsdiskurses zur Studieneingangsphase eine große Forschungslücke: Zwar wurde in früheren Forschungsprojekten auf die Vielzahl an- und Herausforderungen des Studieneingangs als generell krisenanfällige Phase des Hochschulstudiums aufmerksam gemacht (z. B. Schaper, 2023; Bosse & Trautwein, 2017), allerdings erweisen sich dabei der Kontext des Fernstudiums sowie die Zielgruppe der nicht-traditionellen Studierenden bislang als unterbeforscht. Um diese Lücke zu schließen, werden Zwischenergebnisse aus zwei Projektphasen vorgestellt:

Im ersten Teil werden Ergebnisse eines systematischen Literaturreviews nach Gough et al. (2017) zu nicht-traditionellen Studierenden im (inter-)nationalen Forschungsdiskurs dargelegt. NTS zeichnen sich durch eine ausgeprägte Heterogenität aus und schließen in einigen Definitionsansätzen Fernstudierende (Isensee & Wolter, 2017) oder ein Fernstudium als alternatives Studienmodell (Schmitz-Feldhaus, 2020) mit ein. Für das Review wurden sechs Datenbanken im Zeitraum von Ende Oktober 2023 bis Anfang Dezember 2023 durchsucht. 193 Publikationen wurden in den Literaturpool aufgenommen und hinsichtlich ihres Verständnisses von NTS untersucht. Die Ergebnisse bestätigen, dass der Begriff auf nicht eindeutigen Definitionen und Gruppenzuordnungen beruht und bestimmte Facetten studentischer Heterogenität stark vereinfacht oder bündelt (Chung et al., 2014; Lübken et al., 2015; Wolter, 2013). Entsprechend lassen sich sowohl Implikationen, als auch kritische Aspekte und Anforderungen hinsichtlich empirischer Forschung zu NTS ableiten, die im eigenen qualitativen Vorgehen aufgegriffen und reflektiert werden. Dazu werden 2) erste Ergebnisse aus Gruppendiskussionen mit Fernstudierenden der zwei größten Fernhochschulen Deutschlands (FernUniversität in Hagen und IU Internationale Hochschule) vorgestellt. Die Durchführung der insgesamt acht Fokusgruppen mit etwa sechs Teilnehmenden ist innerhalb des letzten Jahresquartals 2024 geplant. Die Auswertung erfolgt mithilfe der qualitativen Inhaltsanalyse. Die Datenerhebung soll offenlegen, wie bestimmte Unterstützungsstrukturen, -ansätze und -maßnahmen von Studierenden – unter Berücksichtigung ihrer Heterogenität – wahr- und angenommen werden. Ebenso können sich Leerstellen abzeichnen und darauf aufbauend bedarfs- und zukunftsorientierte Implikationen für den Übergang in ein Fernstudium abgeleitet werden. Dadurch ermöglichen die Befunde des Forschungsprojekts einen differenzierteren Blick auf den Forschungsdiskurs zu NTS im Kontext des Fernstudiums.

Studentische Identität im Fernstudium in Abgrenzung zu traditionellen Studienformaten

Clemens Klink, Annica Paul-Poudevigne, Marcus Eckert

APOLLON Hochschule der Gesundheitswirtschaft

Theoretischer Hintergrund

Im Kontext der Bildungssozialisation gewinnt die Entwicklung einer ausgeprägten studentischen Identität besondere Bedeutung (Matschke et al., 2023). Diese Identität ist nicht nur ein Ausdruck der individuellen Selbstwahrnehmung, sondern auch ein Spiegelbild der sozialen Zugehörigkeit und des Engagements im akademischen Umfeld. Vergangene Forschungsvorhaben identifizierten zudem verschiedene Formen der akademischen Identität als einflussreichen Faktor für Studienerfolg an Präsenzhochschulen (Meens et al., 2018; Talaifar et al., 2021). Vor allem die implizite Identifikation mit dem eigenen Studienfach wurde als ein guter Abbruchprädiktor identifiziert (Roland et al., 2018). Zudem wurde herausgefunden, dass die akademische Identität bei nicht-traditionellen Studierendengruppen, welche stärker abbruchgefährdet sind, niedriger ausgeprägt ist (Achtelik, 2023; Janke et al., 2017) und über den Studienverlauf abnimmt (Majer, 2018; Matschke et al., 2023). Sie könnte deshalb ein wichtiger Zugangsweg für mehr Chancengerechtigkeit sein.

Bisher gibt es kaum Forschungsansätze, welche die studentische Identität im Kontext des Fernstudiums untersuchen, obwohl die Studierendenschaft hier von nicht-traditionellen Studierenden geprägt ist und höhere Abbruchraten vorweist als an Präsenzhochschulen (Simpson, 2013; Klink et al., 2024). Das Fernstudium stellt gerade deshalb einen besonders spannenden Kontext hinsichtlich der Untersuchung der studentischen Identität dar.

Fragestellung

Wie unterscheidet sich die studentische Identität von traditionellen und nicht-traditionellen Studierenden im Präsenz- und Fernstudium und inwiefern wirkt sich diese Identitätsbildung auf die Abbruchintention beziehungsweise den akademischen Erfolg der Studierenden aus?

Methode

Im Rahmen des BMBF-geförderten Forschungsprojekts „Studienerfolge und -abbrüche im Fernstudium“ (SaFe) wurde die studentische Identität in einer querschnittlichen Studie an Stichproben aus zwei deutschen Fernhochschulen, der APOLLON

Hochschule der Gesundheitswirtschaft (N = 734) sowie der Europäischen Fernhochschule Hamburg (N = 300), untersucht. Hierzu wurden explizite Fragebögen angepasst, welche mit Likert-Skalen die Ausprägung persönlicher (Moreira et al., 2019) und sozialer (Janke et al., 2017) studentischer Identität erfassten. Um darüber hinaus die implizite studentische Identität zu erheben, wurde in Anlehnung an Roland et al. (2018) bei der Hälfte der Teilnehmenden (N = 492) ein reaktionszeitbasierter impliziter Assoziationstest angewendet (Greenwald et al., 2021). Diese Identitätsmaße wurden schließlich mit Bildungsbiografien sowie dem sozialen Umfeld der Studierenden in Beziehung gesetzt. Im Rahmen eines Regressionsmodells wurde zudem die Abbruchintention, die durch den MISAND-Fragebogen erfasst wird (Deuer & Wild, 2019), vorhergesagt und die akademische Motivation (Schnettler et al., 2022) kontrolliert (Klinke et al., under review).

Ergebnisse und ihre Bedeutung

Die Ergebnisse belegen, dass Studierende im Fernstudium explizit eine starke studentische Identität aufweisen. Die persönliche Identität (M = 4.70, SD = .98) liegt hier höher als die soziale Identität (M = 4.55, SD = 1.79). Im Vergleich zu Ergebnissen aus früherer Forschung stellte sich heraus, dass Letztere zudem vergleichbar ist zu traditionellen Studienformaten (Janke et al., 2017), während die implizite studentische Identität schwächer zu sein scheint (M = .28, SD = .56; Roland et al., 2018). Da die studentische Identität die Abbruchintention der Studierenden im Regressionsmodell auch nach dem statistischen Kontrollieren der akademischen Motivation vorhersagte ($\beta = -.10$, $p < .05$), untermauern die Ergebnisse zudem deren Bedeutung.

Ein bemerkenswertes Ergebnis zeigte sich bei den beruflich qualifizierten Studierenden. Im Gegensatz zu früheren Erkenntnissen an Präsenzhochschulen (Achtelik, 2023; Janke et al., 2017; Majer, 2018; Matschke et al., 2023) hatten beruflich qualifizierte und somit nicht-traditionelle Studierende im Vergleich zu traditionell qualifizierten Studierenden eine stärker ausgeprägte soziale und persönliche studentische Identität (p-Werte $< .01$; $d_{\text{sozial}} = .24$, $d_{\text{persönlich}} = .23$), während auf impliziter Ebene das Gegenteil der Fall war: Hier wiesen sie eine schwächere implizite studentische Identität auf als traditionell Qualifizierte (p-Wert $< .01$; $d = -.26$). Die Anonymität, die das Fernstudium zum Teil bieten kann, könnte als eine sichere Umgebung angesehen werden, in der individuelle Herkunft und Verhaltensweisen irrelevant werden und der wichtigere Aspekt die Übereinstimmung der eigenen Identität mit dem gewählten akademischen Bestreben sein könnte.

Lernunterstützung (First-Generation-)Studierender im Fernstudium –Unterstützungspraxen an privaten Hochschulen im Spannungsfeld zwischen Homogenitäts- und Diversitätsorientierung

Aline Fuß¹, Stefanie Hoffmann², Stefanie Kessler¹, Olaf Dörner²

¹IU Internationale Hochschule, ²Otto von Guericke Universität Magdeburg

Private Hochschulen heben sich durch Profilunterschiede von öffentlichen Hochschulen ab, indem sie nicht nur verstärkt weiterbildend agieren (und damit eine große Zielgruppe erschließen), sondern auch gezielt auf weitere Bildungsbedarfe reagieren, z.B. von beruflich bereits qualifizierten Studierenden, Studierenden des Dritten Bildungsweges oder älteren Studierenden, die sich bereits in einem Berufs- und Familienleben wiederfinden, das den Anspruch einer höheren Flexibilität an die Bildungsteilnahme stellt (vgl. Autor:innengruppe Bildungsberichterstattung 2024, S. 205-207). Dementsprechend unterscheidet sich die Zusammensetzung der Studierendenschaft an privaten und öffentlichen Hochschulen (Frank et al. 2020).

Insbesondere ein Fernstudium wird von der oben angeführten Zielgruppen ob einer Kompatibilität mit ihren Ansprüchen gewählt (Pundt/Schöpke, 2024, S. 9). Während ein Fernstudium, das insbesondere auf ihre Anforderung reagiert, an öffentlichen Hochschulen lediglich zu 9 % als Studienformat geführt wird, sind es an privaten Hochschulen 28% (vgl. ebd., S. 206). Rund 58 % der Fernstudierenden sind an einer privaten Hochschule immatrikuliert (Diermeier/Geis-Thöne 2023, S. 11).

Eine besonders signifikante Diversitätsdimension an der IU Internationalen Hochschule (die in diesem Beitrag exemplarisch für die private Hochschullandschaft herangezogen wird) ist die der Bildungsherkunft. Laut den Erstsemesterbefragungen stammen hier 70% der Neuimmatrikulierten aus einem nicht-akademischen Elternhaus, wohingegen es an öffentlichen Hochschulen lediglich circa 45% sind (vgl. Kracke et al. 2024, S. 3).

Diese Gruppe steht im Fokus des BMBF-geförderten Verbundprojektes „First-Generation-Studierende als Zielgruppe privater Hochschulen - Formen und Bedingungen organisationaler Unterstützung“ (FiPHo) der IU Internationalen Hochschule und der Otto-von-Guericke-Universität Magdeburg. Im Projektkontext werden die Forschungsfragen beantwortet, 1) welche kollektiven Orientierungen in Bezug auf FGS an privaten Hochschulen rekonstruiert werden können und 2) inwiefern diese Orientierungen die Unterstützungsstrukturen an privaten Hochschulen prägen. Ziel ist herauszuarbeiten, inwiefern sich in formalen Unterstützungsstrukturen im Studium sowie in informellen Unterstützungspraktiken kollektiv geteilte Orientierungen dokumentieren. Dafür werden in drei Teilprojekten Gruppendiskussionen (Studierende, Lehrende, Studienberater:innen & Studienbetreuende) und ergänzend teilnehmende Beobachtungen (Beratungs- sowie Lehr-Lernsettings) durchgeführt, die mit der Dokumentarischen Methode der Interpretation (vgl. Bohnsack 1999) ausgewertet werden.

Für den Beitrag werden Daten aus den zwei Teilprojekten zur Aneignungs- und Lernunterstützung im Fernstudium (Lehrende und Studienbetreuende) herangezogen, die zur Kontrastierung erhoben wurden. Die Frage, der im Beitrag nachgegangen wird, lautet: Welche Personen lassen sich in Bezug auf die Lernunterstützung (First-Generation-)Studierender im Fernstudium an privaten Hochschulen rekonstruieren? Ziel ist es abzubilden, inwiefern sich die Adressierung diverser Zielgruppen auch in der Gestaltung der Lehre und der Begleitung durch Studierendebetreuende wiederfindet und wie die Praxis der Lernunterstützung – von der eine höhere Verbleib- und Bildungsabschlusswahrscheinlichkeit erwartet werden darf – hergestellt wird. Schließlich stellt eine Diversität auch entsprechende Anforderungen an verschiedene Akteur*innen der Hochschule, wenn aus Zielgruppen Teilnehmende werden (zur Problematisierung siehe Ahmed 2012). Der Forschungsfrage entsprechend werden Ergebnisse aus dem Teilprojekt Studienbetreuende (insgesamt 5 Gruppendiskussionen und 6 Beobachtungen) und dem Teilprojekt Lehrende (insgesamt 5 Gruppendiskussionen und 2 Teilnehmende Beobachtungen) des Verbundvorhabens herangezogen und Perspektiven hinsichtlich ihrer Lernunterstützung verschränkt.

Die statistischen Daten zur diversen Zusammensetzung an der IU Internationalen Hochschule verweisen auf eine hochschulische Praxis, die das Herstellen einer kulturellen Passung (Bourdieu/Passeron 1973) in Bezug auf First-Generation-Studierende vermuten lässt. Erste Ergebnisse verweisen auf eine Diversitätssensibilität und Habitussensibilität (vgl. Schmitt 2019) vor allem in der Gruppe der Studierendenbetreuenden. Bei der Gruppe der Lehrenden lässt sich konträr dazu eher eine homogenitätsorientierte Orientierung (vgl. Hanft 2015) auf Lernunterstützung rekonstruieren.

Der Beitrag und die empirische Betrachtung der Lernunterstützungspraktiken im Fernstudium an einer privaten Hochschule ermöglicht die Untersuchung von Öffnungsprozessen des Sozialraums Hochschule. Darin liegt das Potential, aus der rekonstruierten Organisationspraxis heraus wichtige Transferhinweise für die Organisationsentwicklung öffentlicher Hochschulen zu erhalten. Dies ist unter dem Gesichtspunkt von Bildungsgerechtigkeit – und somit der Öffnung nicht nur vom Sozialraum Hochschule, sondern auch der Erweiterung von Handlungsmöglichkeiten im gesamtgesellschaftlichen sozialen Raum (Diermeier/Geis-Thöne 2023) – von hoher Relevanz.

S-85: Herausforderungen einer indikatorengestützten Bildungsberichterstattung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: SN 169

Chair der Sitzung: Susanne Lochner

Chair der Sitzung: Bettina Arnoldt

Chair der Sitzung: Josefine Lühse

Herausforderungen einer indikatorengestützten Bildungsberichterstattung

Chair(s): **Susanne Lochner** (Deutsches Jugendinstitut (DJI)), **Bettina Arnoldt** (Deutsches Jugendinstitut (DJI)), **Lühse Josefine** (Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF))

Eine zentrale Bestandsaufnahme des deutschen Bildungswesens wird seit 2006 durch die nationale Bildungsberichterstattung gewährleistet. Der nationale Bericht "Bildung in Deutschland" erscheint alle zwei Jahre im Auftrag der Ständigen Konferenz der Kultusminister der Länder (KMK) und des BMBF und stellt in bildungsbereichsspezifischen und -übergreifenden Kapiteln umfassend und empirisch die aktuelle Lage und Entwicklungen im ihrem Gesamtzusammenhang dar. Der im Grundgesetz als Gemeinschaftsaufgabe von Bund und Ländern verankerten Bildungsberichterstattung kommt als Grundlage für evidenzbasierte Bildungspolitik eine Schlüsselfunktion zu. Im Gegensatz zu bildungswissenschaftlichen Studien, die beispielsweise spezifische Ursache-Wirkungs-Strukturen untersuchen, zielt die Bildungsberichterstattung darauf ab, national repräsentative und fortlaufend aktualisierbare Daten zu den Bildungsvoraussetzungen, -wegen und -ergebnissen von Individuen zu sammeln. Diese Daten werden aus einer interdisziplinär zusammengesetzten Autor:innen- und Co-Autor:innengruppe mit Informationen zu Qualitätsmerkmalen von Bildungseinrichtungen und weiteren Kontextfaktoren kombiniert, um das Zusammenspiel von bildungsorganisatorischen und lebensweltlichen Aspekten mithilfe wissenschaftlicher Methoden transparent zu machen. Das analytische Potenzial liegt in der systematischen Zusammenstellung, Einordnung und Reflexion relevanter Daten unter Berücksichtigung aktueller wissenschaftlicher, fachlicher sowie bildungspolitischer und -rechtlicher Rahmenbedingungen. Bildungsberichterstattung ist daher als ein umfassender Prozess zu verstehen, der die Gewinnung, wissenschaftliche Aufbereitung, Analyse sowie die Darstellung und Bereitstellung zentraler Daten umfasst. Sie geht über die bloße Aufbereitung (hauptsächlich amtlicher) Bildungsstatistiken hinaus, indem sie kontinuierlich Informationen zu politischen, rechtlichen und fachlichen Entwicklungen im Bildungswesen recherchiert, systematisiert und zusammenstellt. Dieses, auch interdisziplinäre Vorgehen ermöglicht eine analytische Einordnung der Ergebnisse in gesamtgesellschaftliche Zusammenhänge sowie die Teilnahme an (fach-)öffentlichen Debatten.

Das Symposium will daher der Frage nachgehen, wie mit einer indikatorengestützten Bildungsberichterstattung sowohl eine Lebensverlaufs Perspektive als auch ein breiter Bildungsbegriff, der explizit auch non-formale Bildungsprozesse umfasst, datenbasiert erschlossen werden kann und welche Herausforderungen sich hierbei stellen. Eine Antwort könnte in der Einführung einer Individualstatistik liegen, die den Bildungsverlauf personenbezogen analysierbar macht. In einem ersten Beitrag soll zu Beginn des Symposiums auf umfassende Weiterentwicklungen in der Berichterstattung eingegangen werden, die sich besonders in der Neukonzipierung des Kapitel I zu Bildungsverläufen, Kompetenzentwicklung und Bildungserträgen niederschlagen. Erst mit längsschnittlichen Personensurveys wie dem Nationalen Bildungspanel (NEPS) ist es möglich, Bildungs- und Erwerbsverläufe sowie individuelle Bildungserträge und -biografien zu rekonstruieren und zu analysieren. Anschließend soll in einem zweiten Beitrag ein Fokus auf die non-formale Bildung im Lebensverlauf gelegt werden. Aufgrund einer geringen wissenschaftlichen Fundierung und mangelhafter Datenbasis stellt dieser Bereich eine Herausforderung für die Bildungsberichterstattung dar. Der Beitrag will Weiterentwicklungsmöglichkeiten im Sinne einer indikatorengestützten Berichterstattung aufzeigen. In einem abschließenden Beitrag sollen die Zielsetzungen und bisherigen Aktivitäten zur Einführung eines Bildungsverlaufsregisters dargestellt werden. Die angestrebte bildungsbereichsübergreifende statistische Datenbasis würde erstmals Bildungsverlaufsanalysen mit amtlichen Daten ermöglichen. Welchen Mehrwert die Einführung einer Personen-ID für die nationale Bildungsberichterstattung bereithält und mit welchen Herausforderungen im weiteren Prozess zu rechnen sind, soll als Diskussionsgrundlage aufgegriffen werden.

Beiträge des Symposiums

Zur Darstellung von Übergängen und Verläufen in der nationalen Bildungsberichterstattung – Nutzung von administrativen und Surveydaten

Maximilian Bach¹, Madlain Hoffmann², Christian Kerst³, Maria Richter⁴

¹Statistische Ämter des Bundes und der Länder (Destatis, StLA), ²Leibniz-Institut für Bildungsverläufe (LifBi), ³Deutsches Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), ⁴Soziologisches Forschungsinstitut Göttingen (SOFI)

Das Ziel der nationalen Bildungsberichterstattung besteht darin, datengestützte Informationen über Bedingungen, Prozesse, Ergebnisse und Erträge von Bildung im gesamten Lebensverlauf aufzuzeigen (Maaz & Kühne, 2018). Dabei steht die kontinuierliche Beobachtung des gesamten Bildungswesens im Vordergrund, die durch eine systematische Bestandsaufnahme alle zwei Jahre erfolgt und somit langfristige Entwicklungen und Trends in den Fokus nimmt. Hierfür sind repräsentative Datengrundlagen und fortschreibbare Indikatoren erforderlich, die es ermöglichen, den individuellen Lebensverlauf als Abfolge sukzessiver Bildungsstufen und Bildungsübergänge zu analysieren.

Von Beginn an stand der nationale Bildungsbericht vor der Herausforderung, dass nicht zu allen Themen ausreichend Daten vorliegen, die eine regelmäßige Fortschreibung, eine Differenzierung nach soziodemographischen und sozioökonomischen Merkmalen, motivationalen Faktoren oder nach Land und Region sowie eine längsschnittliche Betrachtung von Bildungs- und Erwerbsverläufen und Kompetenzentwicklungen ermöglichen. Dabei stellt insbesondere die Analyse von Übergängen und Verläufen ein zentrales Element in der Bildungsberichterstattung dar. So wird in den bereichsspezifischen Kapiteln der Perspektive von „Bildung im Lebenslauf“ häufig dadurch entsprochen, dass bei der Darstellung von Übergängen jeweils auf den vorhergehenden bzw. nachgelagerten Bildungsbereich Bezug genommen wird. In den letzten Jahren haben neue Datenzugänge, insbesondere die längsschnittlichen Daten des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019), aber auch thematische Erweiterungen in bereits vorliegenden querschnittlichen Datenbeständen, wie bspw. dem Mikrozensus, zu inhaltlichen und methodischen Weiterentwicklungen in der Bildungsberichterstattung geführt. Neben der Nutzung administrativer Daten werden daher zunehmend quer- und insbesondere auch längsschnittliche Surveydaten herangezogen, die die Analyse und somit stärkere Fokussierung auf individuelle Verläufe ermöglichen. Dies erfolgt im nationalen Bildungsbericht einerseits in den bereichsspezifischen Kapiteln mit der Betrachtung individueller beruflicher Ausbildungs- oder Hochschulverläufe, etwa durch die verstärkte Nutzung von Sequenzmusteranalysen oder durch Querschnittsdaten, mit denen im Kohortenvergleich

Unterschiede im Bildungserwerb zu bestimmten Zeitpunkten analysiert werden. Andererseits wird mit der Neukonzipierung des Kapitels I zu Bildungsverläufen, Kompetenzentwicklung und Erträgen nunmehr eine bereichs- und altersübergreifende Perspektive auf Bildung eingenommen, die die jeweiligen bildungsbereichsspezifischen Perspektiven um eine längerfristige Verlaufs- und Entwicklungsperspektive ergänzen. So ermöglichen längsschnittliche Datengrundlagen die Darstellung individueller Bildungs- und Lebensverläufe sowie individueller Kompetenzentwicklungen.

Die konzeptionellen und thematischen Weiterentwicklungen bringen jedoch auch neue Herausforderungen für die nationale Bildungsberichterstattung mit sich. Im Beitrag werden diese an bildungsbereichsspezifischen und bildungsbereichsübergreifenden Beispielen anhand folgender Fragen aufgegriffen und diskutiert: Welche Möglichkeiten gibt es, Übergänge und Verläufe indikatorisiert darzustellen? Welche Daten und Methoden sind hierfür heranzuziehen? Welche Vor- und Nachteile gehen mit den verschiedenen Datengrundlagen einher? Welche Konsequenzen ergeben sich hieraus für die adressatengerechte Kommunikation und Darstellung der Ergebnisse?

Aus einer bildungsbereichsspezifischen Perspektive werden erstens für die berufliche Ausbildung die Möglichkeiten und Grenzen einer indikatorengestützten Analyse des Übergangs bzw. individuellen Verlaufs in eine berufliche Ausbildung sowie des Ausbildungsverlaufs anhand von verfügbaren administrativen Daten sowie längsschnittlichen Befragungsdaten diskutiert. Als zweites Beispiel werden Übergänge zwischen beruflicher Ausbildung und Hochschulstudium thematisiert und aus einer stärker bildungsbereichsübergreifenden Perspektive werden drittens, schulische und berufliche Bildungs- und Erwerbsverläufe beginnend von der 9. Jahrgangsstufe bis ins junge Erwachsenenalter vorgestellt. An diesen Beispielen werden neue Entwicklungen und Herausforderungen der adressatengerechten Darstellung diskutiert.

Non-formale Bildung im Lebensverlauf – Herausforderungen im Rahmen der indikatorengestützten Bildungsberichterstattung

Bettina Arnoldt¹, Jonathan Kohl², Susanne Lochner¹

¹Deutsches Jugendinstitut (DJI), ²Deutsches Institut für Erwachsenenbildung – Leibniz-Zentrum für Lebenslanges Lernen (DIE)

Der nationalen Bildungsberichterstattung liegt das Konzept des lebenslangen Lernens sowie ein breiter Bildungsbegriff (Bock, 2008) zugrunde, da aus einer individuenzentrierten und Lebensverlaufs-perspektive Bildung nicht erst mit dem Schuleintritt beginnt und nicht bereits mit dem ersten berufsqualifizierenden Abschluss endet. Somit sind nicht nur formale Bildungskontexte wie die Schule, sondern auch Bildungserfahrungen an non-formalen oder informellen Lernorten außerhalb der Schule bzw. formaler Kontexte zu betrachten. Der nationale Bildungsbericht trägt diesem Bedarf von Beginn an Rechnung: das Kindes- und Jugendalter wird mit einem eigenständigen Kapitel zu Früher Bildung (Kapitel C) sowie Indikatoren zu non-formalen Lernwelten im Schulalter (Kapitel D) adressiert, das Erwachsenenalter mit einem eigenständigen Kapitel zur Weiterbildung (Kapitel G) (Konsortium Bildungsberichterstattung, 2005). Non-formale Lernkontexte gewinnen zudem an Bedeutung, etwa aufgrund des Strukturwandels oder demografischer Veränderungen. Ein aktuelles Beispiel stellt das auf europäischer Ebene im Rahmen des European Pillar of Social Rights Action Plan betriebene Benchmarking zur Beteiligung an Weiterbildungsaktivitäten im Erwachsenenalter dar, für das sich Deutschland bis 2030 einen Zielwert gesetzt hat. Definitorisch orientiert sich die nationale Bildungsberichterstattung an internationalen Klassifikationsschemata, wie der Classification of Learning Activities (European Union, 2016). Gleichwohl bleiben Unklarheiten, die im Rahmen der Berichterstattung nur bedingt gelöst werden können. Dies beginnt bei der Verortung Früher Bildung zur formalen oder non-formalen Bildung in internationalen vs. nationalen Klassifikationen. Bildungsbereichsübergreifend können nicht nur konzeptionelle Unschärfen in der Abgrenzung von non-formalen, formalen und informellen Lernaktivitäten, sondern auch weitreichende methodologische Unterschiede in der Art der Erfassung non-formaler Bildungsaktivitäten sowie eine breite, aber zugleich lückenhafte und häufig nicht harmonisierte Datenlage beobachtet werden.

Vor diesem Hintergrund will der Vortrag folgenden Fragestellungen nachgehen: Kann über die verschiedenen Bildungsbereiche ein einheitliches Verständnis von non-formaler Bildung erreicht werden? Wie lässt sich non-formale Bildung datengestützt darstellen? Wo zeigen sich Datenlücken und welche Optimierungsmöglichkeiten bieten sich für die indikatorengestützte Bildungsberichterstattung, die sodann auch die Forschungsaktivitäten in diesem Bereich stärken könnten?

Die bislang genutzten Datenquellen zur Beschreibung non-formaler Bildung im Lebensverlauf sollen anhand der zentralen Differenzierungslinien der nationalen Bildungsberichterstattung „Input-Prozess-Output“ aufgezeigt und entsprechende Datendesiderate abgeleitet werden. Hierbei kann differenziert werden zwischen Kennzahlen, die vorrangig der Deskription des Feldes dienen und Daten aus Studien, die Aussagen über Prozesse und Wirkungen erlauben. Während die Datenlage bezogen auf die Systemebene der Frühen Bildung durch die Kinder- und Jugendhilfestatistik und Monitoringprojekte als umfassend und valide bezeichnet werden kann, sieht es im Bereich der non-formalen Bildung im Schulalter schlechter aus (Kuger & Lochner, 2024), was unter anderem auf heterogene Anbieterstrukturen in der Zivilgesellschaft und der im Vergleich zu formaler Bildung begrenzten Ressourcen zurückgeführt werden kann. Die 2015 erstmalig erhobene Kinder- und Jugendarbeitsstatistik kann als Beispiel für die Datengenerierung angeführt werden, um das Angebot der Kinder- und Jugendarbeit systematisch zu erfassen. Jedoch befindet sich diese Statistik weiterhin im Aufbau und der kontinuierlichen Weiterentwicklung (Pothmann, 2020). Mit Hilfe von repräsentativen sozialwissenschaftlichen Erhebungen (NEPS, AID:A) sowie Teilnehmendenstatistiken von zivilgesellschaftlichen Akteuren wird versucht, das Feld der außerschulischen non-formalen Bildung in Ansätzen abzubilden. Für die Weiterbildung haben sich historisch gewachsen verschiedene Monitoringstudien etabliert, die non-formale Bildungskontexte insbesondere auf Input-Ebene u.a. aus Sicht von Teilnehmenden oder Betrieben erfassen (Adult Education Survey, Continuing Vocational Training Survey, PIAAC). Sie werden ergänzt von Kennzahlen aus der amtlichen Statistik (Mikrozensus) oder sozialwissenschaftlichen Erhebungen (NEPS, SOEP, Deutscher Alterssurvey). Sie eint jedoch, dass sie über kein harmonisiertes Erhebungsinstrument verfügen und bei Subgruppenanalysen mitunter zu uneinheitlichen Ergebnissen führen (Widany et al., 2019). Mit Blick auf vulnerable Personengruppen lassen sich eklatante Datenlücken aufzeigen. Der Vortrag möchte Entwicklungsmöglichkeiten wie Datenharmonisierungen aufzeigen, von der das Monitoring non-formaler Bildung im Lebensverlauf profitieren kann.

Bildungsverlaufsregister

Katharina Giar, Franziska Hohlstein
Statistisches Bundesamt

Das Statistische Bundesamt stellt die Zielsetzung und den aktuellen Sachstand des Projekts „Bildungsverlaufsregister im Verbund“ vor. Mit dem Aufbau eines Bildungsverlaufsregisters in Deutschland soll eine bildungsbereichsübergreifende statistische Datenbasis geschaffen werden, die zum einen Bildungsverlaufsanalysen ermöglicht und zum anderen Bildungsangaben für den registerbasierten Zensus sowie für europäische Bildungsstatistiken bereitstellen kann. Die Statistischen Ämter des Bundes und der Länder konzipieren hierzu gemeinsam mit den Kultusministerien und dem BMBF den Aufbau eines

auf amtlichen Bildungsstatistiken basierenden „Bildungsverlaufsregisters im Verbund“ (BVR-V). Hierfür wurde im Statistischen Verbund bereits eine Machbarkeitsstudie und ein gemeinsames Zielbild erarbeitet und beschlossen, welche die Grundlagen für die fachliche Konzeption eines Bildungsverlaufsregisters im Verbund festhalten.

Der föderale Aufbau des BVR-V entspricht der Hoheit der Organe über die Statistiken. Bundesrechtlich geregelte Statistiken wie die Hochschulstatistik, die Berufsbildungsstatistik oder die Pflegeausbildungsstatistik werden in einem Bundesbildungsverlaufsregister (BVR-B) verarbeitet, länderrechtlich geregelte Statistiken wie die Schulstatistiken oder die Anerkennungsstatistik der Länder werden wiederum in Bildungsverlaufsregistern der Länder (BVR-L) verarbeitet. Eine Zusammenführung der getrennten Statistikregister soll perspektivisch dann über das Bildungsverlaufsregister im Verbund stattfinden.

Durch das Zuspielen einer Bildungs-ID, die mithilfe statistikübergreifender personenidentifizierender Merkmale in einer separaten Stelle, einer Trusted Third Party, zugeordnet wird, können pseudonymisierte Bildungsverläufe erzeugt und ausgewertet werden. Hiermit wird – analog zum möglichen Forschungsdatengesetz – eine bereits lange geäußerte Forderung aus Wissenschaft und Forschung nach der Verknüpfbarkeit von registerbasierten Bildungsdaten erfüllt. Gegenüber stichprobenbasierten Paneldaten bietet das BVR-V den Vorteil, durch die Nutzung amtlicher Daten perspektivisch und im Sinne einer Vollerhebung alle formalen Bildungsteilnahmen und -abschlüsse zu erfassen und im Verlauf darstellen zu können. Dadurch entfallen typische methodische Probleme von Panelerhebungen, wie beispielsweise Panelmortalität, Selbstselektivität bzw. selektive Ausfälle und Gewöhnungseffekte. Ein weiterer Vorteil der amtlichen Daten ist der direkte Bezug der Informationen aus der Verwaltung, also beispielsweise Prüfungsämter oder Industrie- und Handelskammern. Diese Informationen werden anschließend im Rahmen der statistischen Aufbereitung einer intensiven Plausibilisierung unterzogen. Dieser Prozess gewährleistet die Validität, Objektivität und die Reliabilität der gelieferten Angaben.

Grundlage eines BVR-B ist der Erlass eines Gesetzes zur Einführung eines Bundesbildungsverlaufsregisters durch den Gesetzgeber auf Bundesebene. Die BVR-L können im jeweiligen Land durch Landesrecht begründet werden. Für die mögliche Schaffung eines gemeinsamen BVR-V ist ein Staatsvertrag zwischen Bund und Ländern notwendig. Weiterhin sind vielschichtige Arbeiten zum Aufbau der technischen und organisatorischen Infrastruktur des BVR-B sowie des BVR-V erforderlich.

Der Bedarf eines Bildungsverlaufsregisters ist bei den Stakeholdern der Bildungspolitik und -forschung unstrittig. Diese Bedarfe fließen neben den zentralen datenschutzrechtlichen Anforderungen bereits in die Konzeption ein. In Verbindung mit dem geplanten Forschungsdatengesetz wird der Zugang der Bildungsverlaufsdaten für die Wissenschaft deutlich verbessert werden.

S-86: Wie können soziale Bildungsungleichheiten reduziert werden? Eine interdisziplinäre Perspektive auf formale und non-formale Bildung

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: SO 108

Chair der Sitzung: Jan Scharf

Chair der Sitzung: Hannah Kleen

Chair der Sitzung: Hanna Dumont

Wie können soziale Bildungsungleichheiten reduziert werden? Eine interdisziplinäre Perspektive auf formale und non-formale Bildung

Chair(s): **Jan Scharf** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland), **Hannah Kleen** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland)

Diskutant*in(nen): **Hanna Dumont** (Universität Potsdam)

Kindern und Jugendlichen unabhängig von ihrer familiären Herkunft und elterlichen Unterstützung gleiche Bildungschancen zu bieten, gilt als eine der wichtigsten Voraussetzungen für gesellschaftliche Partizipation, zugleich jedoch auch als große bildungspolitische Herausforderung. Während Bildungsbeteiligung und Bildungsangebote in den vergangenen Jahren gestiegen sind (Autor:innengruppe Bildungsberichterstattung 2024: 6), stellen soziale Ungleichheiten im Bildungserwerb – sowohl hinsichtlich Kompetenzen (vgl. Lewalter et al. 2023; Stanat et al. 2022, 2023) als auch mit Blick auf Bildungsverläufe und -abschlüsse (vgl. etwa Kurz et al. 2019; Scharf et al. 2020; Schindler & Bittmann 2021) – aber nach wie vor einen robusten Befund der Bildungsforschung dar (im Überblick: Autor:innengruppe Bildungsberichterstattung 2024: 21f). Nicht erst seit den öffentlichen Diskursen um soziale Benachteiligungen und Lernrückstände im Zuge der COVID-19-Pandemie (Dietrich et al. 2021) rücken Fragen der gezielten Kompensation und Reduzierung von Bildungsungleichheiten vermehrt in den Fokus, zuletzt durch die Ergebnisse der jüngsten PISA-Studie nochmals – und international vergleichend – verdeutlicht: „Um dieser Situation zu begegnen, bedarf es entschiedener gemeinsamer Anstrengungen [...], um flächendeckende Maßnahmen auf den Weg zu bringen, die eine Reduktion von Ungleichheiten gezielt in den Fokus nehmen“ (Lewalter et al. 2023: 320f).

Die Bildungsforschung hat Ursachen von Bildungsungleichheiten nach sozialer Herkunft, die in familiären Interaktionsformen, Ressourcenausstattungen und Bildungsentscheidungen zu suchen sind sowie in institutionellen Entstehungsbedingungen, unter anderem in Form von Diskriminierungen, für verschiedene Altersstufen in einem breiten Forschungsstand umfänglich nachgewiesen (vgl. etwa Dumont et al. 2019; Pit-ten Cate & Glock 2023; Scharf et al. 2020) und systematisiert (Bachsleitner et al. 2022). Sie bieten einen wichtigen Ausgangspunkt für Forschungsarbeiten zur Reduzierung von Ungleichheit: Dieser Forschungsgegenstand wiederum erfordert es, zum einen interdisziplinäre Perspektiven einzunehmen, zum anderen neben Maßnahmen in formaler Bildung auch den Beitrag non-formaler Lernumwelten und deren Wechselwirkungen mit formalen Bildungsangeboten in den Blick zu nehmen (Hadjari et al., 2022; Autor:innengruppe Bildungsberichterstattung, 2024). Das Symposium bringt drei Studien zusammen, die auf verschiedenen Analyseebenen angelegt sind und sich auch in ihren methodischen Herangehensweisen und Zielen unterscheiden, zugleich aber abschließend übergreifend und interdisziplinär diskutiert werden können. Während sie die Frage nach wirksamen Maßnahmen zur Erhöhung von Bildungschancen bzw. Reduzierung von Ungleichheit eint, werden erstens die Effekte einer Intervention zur Reduzierung von stereotypen Urteilen, zweitens die Effekte einer Schulstrukturreform und drittens die Breite und Evidenz des bisherigen empirischen Forschungsstands zum Thema fokussiert.

Aus psychologischer Perspektive wird im ersten Vortrag untersucht, ob eine bias awareness Intervention Stereotype von angehenden Lehrkräften reduzieren kann. Im deutschen Kontext ist bisher wenig über die bias awareness von Lehrkräften gegenüber Schüler*innen mit niedrigem sozio-ökonomischen Status bekannt. Angenommen wird, dass eine höhere Bias Awareness bei Lehrkräften den Einfluss von Stereotypen auf Urteile verringern kann. Anhand der vorläufigen Befunde zeigt sich eine Tendenz, dass die Bias Awareness der Teilnehmenden steigt und somit der eigene implizite Bias stärker reflektiert wird.

Der zweite Vortrag behandelt die soziologische Perspektive auf Effekte von Schulreformen. Basierend auf den quasiexperimentellen Daten der BERLIN-Studie untersucht der Beitrag, wie sich die Durchlässigkeit zum Abitur und die soziale Ungleichheit im Abiturserwerb durch die Umstellung auf ein zweigliedriges Schulsystem verändert haben. Dabei wird der Frage nachgegangen, ob sich durch einen Anstieg der Abiturient*innenquote und den höheren Anteil an Schüler*innen nicht privilegierter Herkunft an den nichtgymnasialen Schulformen nach der Reform soziale Ungleichheiten im Abiturserwerb verringert haben. Die differentiellen Befunde zeigen unter anderem, dass sich die sozialen Ungleichheiten im Abiturserwerb nach der Reform nicht signifikant verändert haben.

Abschließend bringt eine aktuelle Forschungssynthese interdisziplinäre Perspektiven ein, indem ausgehend von einem weiten Bildungsbegriff Konzepte in formaler, non-formaler und informeller Bildung gebündelt und systematisiert werden. So können Befunde vergleichend betrachtet und generalisierbare Erkenntnisse sowie Forschungslücken aufgezeigt werden. Der Beitrag beschreibt, wie Bildungsbarrieren in der Forschung konzeptualisiert und welche Maßnahmen evaluiert werden.

Beiträge des Symposiums

Bias Awareness als Kurzintervention: Reflexion von Bias gegenüber Schüler*innen mit niedrigem sozio-ökonomischen Status

Noa Gadov, Meike Bonefeld

Institut für Erziehungswissenschaft, Erziehung und Sozialisation, Universität Freiburg

Schüler*innen erhalten abhängig von ihrem sozio-ökonomischen Status bei gleicher Leistung schlechtere Leistungsbeurteilungen (Tobisch & Dresel, 2017). Außerdem nehmen (angehende) Lehrkräfte Schüler*innen mit hohem sozio-ökonomischen Status als signifikant ehrgeiziger, fleißiger, intelligenter, konzentrierter und motivierter wahr als Schüler*innen mit niedrigem sozio-ökonomischen Status (Tobisch & Dresel, 2020). Des Weiteren assoziieren angehende Lehrkräfte Schüler*innen mit niedrigem sozio-ökonomischen Status weniger mit Sprachkompetenz, Fähigkeiten und guten Arbeitsgewohnheiten. Zudem haben sie mehr negative implizite Einstellungen gegenüber Schüler*innen mit niedrigem sozio-ökonomischen Status als gegenüber Schüler*innen hohem sozio-ökonomischen Status und (Glock & Kleen, 2020). Zahlreiche Studien belegen, dass implizite Einstellungen (angehender) Lehrkräfte eine Ursache dafür sein können, dass (angehende) Lehrkräfte Schüler*innen abhängig von leistungsfremden Merkmalen unterschiedlich beurteilen (Bonefeld & Dickhäuser, 2018; Glock 2016 et al.; Glock 2018; Nürnberger et al. 2015; Quinn 2020).

Um den Einfluss von impliziten Einstellungen auf die Schüler*innenbeurteilung zu reduzieren, ist es erforderlich, dass (angehende) Lehrkräfte zunächst ein Bewusstsein für ihre eigenen impliziten Einstellungen entwickeln (bias awareness, Perry et al., 2015). Bislang existieren insbesondere im deutschen Kontext nur wenige Studien zur bias awareness im Allgemeinen sowie insbesondere zur bias awareness im schulischen Kontext. Die wenigen bestehenden Studien zeigen auf, dass die bias awareness angehender Lehrkräfte eher moderat ausgeprägt ist (Bonefeld 2022). Es konnte ebenfalls nachgewiesen werden, dass sich die bias awareness angehender Lehrkräfte und deren Bewusstsein für die eigenen impliziten Einstellungen gegenüber Menschen mit Migrationshintergrund sowie Frauen und Männern durch eine Kurzintervention in Form eines 90-minütigen Workshops signifikant erhöhen lässt (Bonefeld, 2022). Noch nicht untersucht wurde die bias awareness (angehender) Lehrkräfte in Bezug auf Menschen mit niedrigem sozio-ökonomischen Status. Im Rahmen einer Interventionsstudie wird die Bias Awareness (angehender) Lehrkräfte in Bezug auf Menschen mit niedrigem sozio-ökonomischen Status sowie die Veränderbarkeit dieser durch den Einsatz einer adaptierten Version der Kurzintervention von Bonefeld (2022) untersucht.

Es wurde eine a-priori Poweranalyse mit G*Power Version 3.1.9.7 (Mayr et al., 2007) durchgeführt, um die Mindeststichprobengröße zu bestimmen, die zur Prüfung der Wirksamkeit der Intervention erforderlich ist. Die Ergebnisse der Analyse zeigen, dass eine Stichprobengröße von $N = 45$ erforderlich ist, um eine Power von 95% für den Nachweis eines mittleren Effekts bei einem Signifikanzkriterium von $\alpha = .05$ zu erreichen. Die deskriptiven Daten einer vorläufigen Datenanalyse ($N = 16$) weisen darauf hin, dass die Intervention die Bias Awareness angehender Lehrkräfte in Bezug auf Menschen mit niedrigem sozio-ökonomischen Status erhöht (t_1 : $M = 3.83$, $SD = 1.42$; t_2 : $M = 4.27$, $SD = 1.47$; t_3 : $M = 4.06$, $SD = 1.46$). Eine Messwiederholungs-ANOVA ergab keine signifikanten Effekte der Intervention auf die bias awareness der Teilnehmenden. Dies ist vermutlich auf die zu geringe Stichprobengröße zurückzuführen. Es ist zu erwarten, dass bei Erreichen der Mindeststichprobe (Abschluss der Erhebung 08/2024) die Wirksamkeit der Intervention nachgewiesen werden kann. Die Ergebnisse werden vor dem Hintergrund der bisherigen Befunde zu impliziten Einstellungen bezüglich des sozio-ökonomischen Status diskutiert.

Soziale Ungleichheiten im Abiturerwerb: Wie durchlässig ist das zweigliedrige Schulsystem und für wen?

Anna Bachsleitner¹, Markus Lörz¹, Marko Neumann¹, Michael Becker²

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland, ²TU Dortmund

Die in Deutschland frühe Aufteilung der Schüler*innen in verschiedene Schulformen und die relativ geringe Durchlässigkeit des Bildungssystems werden als ein Erklärungsansatz für die international vergleichsweise hohen sozialen Bildungsungleichheiten diskutiert, auch wenn sich die Kopplung zwischen Schulform und Schulabschluss in den letzten Jahrzehnten deutlich gelockert hat. Zudem erfolgte in einer Reihe von Bundesländern eine Umstellung der Schulsysteme auf Zweigliedrigkeit, sodass neben dem Gymnasium nur noch eine zentrale nichtgymnasiale Schulform besteht. Diese ermöglicht entweder durch eine eigene gymnasiale Oberstufe oder durch Kooperationen mit anderen Schulen auch den Erwerb eines Abiturs. Eine der zentralen Zielsetzungen der Schulstrukturreformen war eine Reduzierung des Einflusses der sozialen Herkunft auf den Bildungserfolg. Empirische Befunde zur Umstellung der Schulstruktur auf Zweigliedrigkeit in Berlin zeigen einen Anstieg der Oberstufenzugangsberechtigungen sowie der realistischen Abituraspirationen an den nichtgymnasialen Schulformen nach der Reform, von der alle sozialen Herkunftsgruppen in ähnlicher Weise profitieren konnten (Maaz et al., 2017). Bisher wurde jedoch weder für Berlin noch für andere Bundesländer mit zweigliedrigen Schulsystemen analysiert, inwieweit sich soziale Ungleichheiten im tatsächlichen Erwerb des Abiturs durch die Reform verändert haben. Daher untersucht der vorliegende Beitrag die Forschungsfrage, ob die Umstellung auf Zweigliedrigkeit den Zugang zum Abitur erleichtert und wenn ja, wie diese verbesserte Durchlässigkeit die sozialen Ungleichheiten beim Zugang zum Abitur verändert.

Es wird vermutet, dass sich mit den verbesserten Möglichkeiten des Abiturerwerbs der Anteil an Abiturient*innen, insbesondere an den nichtgymnasialen Schulformen, grundsätzlich erhöht. Hinsichtlich der neuen Entscheidungspunkte für den Abschlusserwerb, die sich durch die geöffneten Bildungswege an den nichtgymnasialen Schulformen ergeben, sind mit Bezug auf Rational-Choice-Ansätze (u.a. Erikson und Jonsson, 1996) zwei herkunftsspezifisch gegenläufige Effekte denkbar: Einerseits kann aus Gründen des Statuserhalts (Breen & Goldthorpe, 1997) vermutet werden, dass Jugendliche aus privilegierten Familien stärker die Möglichkeiten nutzen, ihren bisherigen Bildungspfad zu korrigieren und doch noch ein Abitur zu erwerben. Andererseits kann vermutet werden, dass Schüler*innen aus weniger privilegierten Familien von den späteren Entscheidungspunkten profitieren, da sie ihre Erfolgswahrscheinlichkeit für den Erwerb eines Abiturs zu einem späteren Zeitpunkt im Bildungsverlauf besser einschätzen können. Bezogen auf die Veränderung der Ungleichheit im Abiturerwerb insgesamt wird angenommen, dass es durch einen Anstieg der Abiturient*innenquote und einem hohen Anteil an Schüler*innen nicht privilegierter Herkunft an den nichtgymnasialen Schulformen nach der Reform auf Ebene der Gesamtkohorte zu einer Verringerung sozialer Ungleichheit kommen sollte. Schüler*innen mit nichtakademischen Elternteilen würden dann ihren Anteil an Abiturient*innen nach der Reform stärker steigern können als Schüler*innen aus Akademikerfamilien.

Die theoretischen Annahmen werden basierend auf den Daten der BERLIN-Studie analysiert, mit der die Umstellung des Sekundarschulsystems auf Zweigliedrigkeit in Berlin untersucht wurde. Die BERLIN-Studie hat ein quasi-experimentelles Studiendesign, mit dem Schüler*innen der 9. Jahrgangsstufe vor der Reform (im Jahr 2011) und nach der Reform (im Jahr 2014) längsschnittlich verfolgt wurden. Die Schüler*innenkohorte, die nach der Schulstrukturreform untersucht wurde, bildet die Reformkohorte der Studie, die Kohorte vor der Schulstrukturreform die Kontrollkohorte. Es werden lineare Wahrscheinlichkeitsmodelle zum Abiturerwerb mit Interaktionen zwischen der sozialen Herkunft und der Kohortenzugehörigkeit sowohl für die nichtgymnasialen Schulformen als auch über alle Schulformen gerechnet.

Die empirischen Befunde zeigen, dass der Anteil an Abiturient*innen, insbesondere an den nichtgymnasialen Schulformen, nach der Reform deutlich gestiegen ist. An den nichtgymnasialen Schulformen konnten alle sozialen Herkunftsgruppen ihre Abiturquote steigern, insbesondere jedoch Schüler*innen aus privilegierten Elternhäusern. Dieses Ergebnis steht im Einklang zu den theoretischen Annahmen der Statuserhaltstheorie – Jugendliche aus privilegierten Familien nutzen stärker die Korrekturoptionen an den nichtgymnasialen Schulformen, da ein Abitur notwendig zur Reproduktion des elterlichen Status ist. Insgesamt, über alle Schulformen betrachtet, haben sich die sozialen Ungleichheiten im Abiturerwerb nach der Reform nicht signifikant verändert. Die Befunde und ihre Implikationen werden im Vortrag diskutiert und eingeordnet.

Bildungsbarrieren in formaler und non-formaler Bildung identifizieren, Evidenz zu deren Abbau systematisieren: Erste Ergebnisse einer Forschungssynthese

Christina Möller¹, Selina Kirschey², Monika Lindauer¹, Ingeborg Jäger-Dengler-Harles², Jan Scharf², Andreas Herz¹, Susanne Kuger¹, Kai Maaz²

¹Deutsches Jugendinstitut e. V., ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Sozial benachteiligte Kinder, Jugendliche und junge Erwachsene haben geringere Bildungschancen und -erfolge. Dieser Zusammenhang von Bildungserwerb und sozialer Herkunft ist in Deutschland besonders stark und persistent (vgl. Lewalter et al.

2023; Autor:innengruppe Bildungsberichterstattung 2024; Simon 2023), mit Folgen auch für die Teilhabe und Partizipation am gesellschaftlichen Leben: an Kultur, Sport, Medien, an Erwerbsarbeit und weitere Bereiche. Daher ist die Frage, wie Bildungsbenachteiligungen kompensiert beziehungsweise möglichst früh verhindert werden können, umso bedeutsamer für die Gesellschaft.

Die Ursachen sozialer Bildungsungleichheit in Deutschland sind inzwischen umfangreich erforscht und systematisiert (Bachsleitner, Lämmchen & Maaz 2022). Empirische Studien untersuchen zudem, wie Bildungsungleichheiten abgebaut oder deren Entstehen verhindert werden kann (vgl. Blatter et al. 2020; Solga & Weiß 2015), konzentrieren sich dabei jedoch auf einzelne Bildungssettings oder Lebensphasen (Autor:innengruppe Bildungsberichterstattung 2020) und erfassen damit nicht, dass Kinder und Jugendliche sich in einem Geflecht aus formalen, non-formalen und informellen Lernorten, neben der Schule etwa in Vereinen oder Jugendhilfeeinrichtungen und im Familien- und Freundeskreis, bewegen (Zdun 2021). Damit fehlt es für den Untersuchungsraum Deutschland bisher an einem systematischen Überblick zu bestehenden Forschungsarbeiten, die die Möglichkeiten zum Abbau von Bildungsbarrieren in verschiedenen Lernumwelten eruieren und diese insbesondere hinsichtlich einer empirisch geprüften Wirksamkeit untersuchen.

Die im Rahmen des Symposiums vorgestellte Studie knüpft an diesem Forschungsdesiderat in Form einer systematischen Literaturrecherche (SLR) bzw. Forschungssynthese an. Die ursprünglich im medizinischen Bereich entwickelte Methode dient der systematischen Bündelung vorhandener wissenschaftlicher Erkenntnisse und der Offenlegung noch bestehenden Forschungsbedarfs zu einer ausgewählten Fragestellung (Newmann & Gough 2020) und findet zunehmend auch Anwendung im sozialwissenschaftlichen Bereich (Petticrew & Roberts 2006; Wetterich & Plänitz 2021). Konkret werden für diese SLR relevante qualitative und quantitative Längs- und Querschnittstudien der Bildungsforschung und ihrer Lehnsdisziplinen recherchiert, die die Wirksamkeit konkreter Maßnahmen zum Abbau von Bildungsbarrieren untersuchen. Dabei werden formale, non-formale und informelle Bildungssettings für Zielgruppen der frühen Kindheit bis zum jungen Erwachsenenalter (0-27 Jahre) berücksichtigt, die ab dem Publikationsjahr 1965 für den Untersuchungsraum Deutschland evaluiert wurden. Hinsichtlich der Frage, was unter einer Bildungsbarriere zu verstehen ist, zielen die Einschlusskriterien auf Untersuchungen, die mit Konzepten von Bildungsbarrieren auf verschiedenen Ebenen arbeiten (Schmidt-Hertha 2018, S. 831; Wenzel 2008, S. 430): auf der Mikro-Ebene (individuelle wissens-, kompetenz- oder dispositionsbezogene Barrieren), auf Meso-Ebene (z. B. Lernumwelten) und auf der Makro-Ebene (z. B. politische Diskurse).

Eine Recherche in der Datenbank FIS Bildung und eine anschließende kriteriengeleitete Filterung unter Zuhilfenahme der Software R ergab eine Treffermenge von 61.366 Artikeln, die potenziell Antworten auf die leitende Forschungsfrage geben können. Den weiteren methodischen Schritten einer SLR folgend (vgl. Newman & Gough 2020; Wetterich & Plänitz 2021), schließt ein umfassender Screening Prozess an, der auch die Nutzung der Software Rayyan und eine damit verbundene künstliche Intelligenz zur Vorsortierung der noch zu screenenden Artikel nach Passungswahrscheinlichkeit einbezieht (vgl. Ouzzani 2016). Insgesamt 3.343 Treffer wurden als potenziell relevant eingestuft und in das Volltext-Screening übernommen. Die an diesem zweiten Screening-Prozess anschließende Kodierung der Studien erfolgt entlang eines deduktiv-induktiv entwickelten Kategoriensystems, worüber die relevanten Informationen der ausgewählten Arbeiten zum theoretischen Zugang, zum Forschungsdesign, zu den untersuchten Maßnahmen sowie zur Wirksamkeit festgehalten und inhaltsanalytisch aufbereitet werden (Kuckartz & Rädiker 2022). Dies ermöglicht es im Sinne eines cumulative reviews (Templier & Paré 2015), eine erste Orientierung über den Stand der Forschungsliteratur zu schaffen: Wie wird für verschiedene Zielgruppen und Bildungssettings zum Abbau von Bildungsbarrieren geforscht, welche Maßnahmen kommen zum Einsatz und wie wirken diese?

Ziel des Beitrags ist es, die Ergebnisse des Screening- und Kodierprozesses, insbesondere mit dem Fokus auf den Abbau von Bildungsbarrieren in formalen und non-formalen Bildungssettings, darzustellen. Ausblickend wird das Potenzial dieser Forschungssynthese für die Identifizierung von Forschungslücken und weiteren Potenzialen zur Chancenförderung für Kinder, Jugendliche und junge Erwachsene beleuchtet.

S-87: Transitionen und Karriereverläufe in der Wissenschaft

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 145

Chair der Sitzung: Martin Daumiller

Chair der Sitzung: Johanna Ott

Chair der Sitzung: Ronja Steinhauser

Transitionen und Karriereverläufe in der Wissenschaft

Chair(s): **Martin Daumiller** (Universität Augsburg, Deutschland), **Johanna Ott** (Universität Augsburg, Deutschland), **Ronja Steinhauser** (Universität Mannheim)

In der Gesellschaft wird das Streben nach evidenzbasierten Lösungen und einer stärkeren Beachtung wissenschaftlicher Erkenntnisse unverändert größer. Wissenschaftler:innen spielen hierbei eine zentrale Rolle. Trotz der großen gesellschaftlichen Bedeutung von Wissenschaftler:innen sind deren Arbeits- und Ausbildungsbedingungen häufig von Herausforderungen (z. B. prekäre Arbeitsbedingungen, Wissenschaftszeitgesetz, Wettbewerbsdruck) gekennzeichnet, was zu Wohlbefindenseinbußen, einer Abnahme von Motivation oder sogar dem Ausscheiden aus dem Wissenschaftssystem führen kann (hierzu beispielsweise Gonzalez et al. 2021; Roach et al. 2017). Dabei könnten insbesondere Transitionsphasen eine besondere Bedeutung zukommen.

Diese sind bisher in der Forschung nur wenig adressiert und es wird immer dringlicher, genauer zu verstehen, wer wieso eine Karriere in der Wissenschaft wählt und welche Faktoren Karriereverläufe begünstigen.

In diesem Symposium fokussieren wir entsprechend Transitionen in der wissenschaftlichen Karriere – von der Entscheidung für eine Promotion, über den Verbleib in der Wissenschaft nach Abschluss der Promotion bis zu alternativen wissenschaftlichen Karrierewegen.

Der erste Beitrag berichtet von einer längsschnittlichen Befragung, die über 200 Teilnehmende vom Studium in die Promotion begleitete. Die Ergebnisse zeigen, dass die erwartete Passung zur Promotion am Ende des Studiums prädiktiv für Wohlbefinden, individuell wahrgenommenen Fortschritt sowie Abbruchintentionen in der Promotion ist. Die Befunde stellen somit erste Ansatzpunkte für eine Unterstützung von Studierenden bei ihrer Entscheidung für oder gegen eine Promotion dar.

Im zweiten Beitrag werden der Zusammenhang zwischen soziokulturellem Hintergrund und Selbstkonzept mit Karrierewünschen betrachtet. Zusätzlich wird untersucht, welche Rolle die Verinnerlichung des Narrativs akademischer Exzellenz dabei spielt. Ziel ist es, eine mögliche Selbstselektion des wissenschaftlichen Korpus durch dieses Exzellenznarrativ zu betrachten, um einen möglichen Mangel an Diversität in der Wissenschaft zu erkennen. Die Datengrundlage bildet eine quantitative Befragung von Forschenden an drei Berliner Universitäten und der Charité Berlin zu Forschungskarrieren.

Im dritten Beitrag wird die Veränderung von Lern- und Erscheinungszielen bei Nachwuchswissenschaftler:innen im Übergang in den Postdoc betrachtet. Ergebnisse von zwei Stichproben (Doktorand:innen am Ende ihrer Promotion und Postdoktorand:innen zu Beginn ihrer Postdoc-Phase), die längsschnittlich untersucht wurden, zeigen intraindividuelle Veränderungen sowie Unterschiede der Ziele zwischen den Stichproben.

Im vierten Beitrag wird das Wissenschaftsmanagement als alternativer Karrierewege für Wissenschaftler:innen zur Professur im universitären Umfeld anhand von qualitativen und quantitativen Ergebnisse einer großen Mixed-Methods-Studie abgebildet. Dabei werden verschiedene Wege ins und im Wissenschaftsmanagement, Umweltfaktoren, sowie individuelle Personenmerkmale betrachtet. Die Analyse der Daten zeigt, dass Wissenschaftsmanager*innen eine hochqualifizierte, interdisziplinär aufgestellte Berufsgruppe sind, die komplexe Aufgaben im Wissenschaftssystem erfüllt und deren Qualifikationen oft aus eigenen Forschungserfahrungen und vielfältigen Weiterbildungen resultieren. Zudem wird deutlich, dass Wissenschaftsmanagement, früher als „Plan B“ zu einer akademischen Karriere angesehen, heute zunehmend bewusst gewählt wird und mit hoher Zufriedenheit verbunden ist.

Damit werden die zentralen Transitionswege wissenschaftlicher Karrieren mit dem Beginn der Transition und der Entscheidung für oder gegen eine Laufbahn in der Wissenschaft, dem Übergang in die Postdoc-Phase und möglichen Alternativen zur wissenschaftlichen Karrieren in diesem Symposium vereint. Dazu präsentieren wir Ergebnisse aus mehreren, teils qualitativen und quantitativen, querschnittlichen und längsschnittlichen, national-repräsentativen Untersuchungen. Zudem bringt das Symposium Wissenschaftler:innen aus der Psychologie, der Persönlichkeits- und Organisationsentwicklung sowie der Soziologie zusammen und liefert neben relevanten Impulsen für zukünftige Forschung zu diesem bedeutsamen Themengebiet auch wesentliche praktische Anknüpfungspunkte.

Im Anschluss an die Einzelbeiträge werden wir mit dem interdisziplinären Expert:innenteam und dem Auditorium eine digital angeleitete Diskussion führen. Dabei erwarten wir, dass insbesondere auch die Diskussionsbeiträge der im Plenum anwesenden Wissenschaftler:innen von besonderem Wert für das Symposium sein werden, da diese mit ihren individuellen Erfahrungen in unterschiedlichen Transitionsphasen und dem Wissenschaftssystem einen wichtigen Beitrag leisten.

Beiträge des Symposiums

Promovieren oder nicht? Zur Bedeutung von Fit-Prognosen für Wohlbefinden, wahrgenommenen Fortschritt und Abbruchintention bei Promovierenden

Ronja Steinhauser¹, Belinda Merkle¹, Johanna Ott², Oliver Dickhäuser¹, Markus Dresel³, Martin Daumiller², Stefan Janke¹

¹Universität Mannheim, ²Universität Augsburg, Deutschland, ³Universität Augsburg

Die Ausbildung von Promovierenden ist eine zentrale Aufgabe an Hochschulen und ein wichtiger Grundstein für das Sicherstellen eines qualitativ-hochwertigen Wissenschaftssystems. In den letzten Jahren entscheiden sich immer mehr Absolvent:innen für eine Promotion (OECD, 2021). Gleichzeitig berichten Promovierende von eingeschränktem Wohlbefinden (Levecque et al., 2017) und mehr als ein Sechstel der Promovierenden schließen ihre Promotion nicht erfolgreich ab (Fabian et al., 2013). Ein möglicher Ansatzpunkt für das bessere Verständnis dieser Phänomene ist die Person-Environment-Fit-Theorie, die postuliert, dass eine fehlende Passung zwischen den Interessen (Interessenspassung) beziehungsweise den Fähigkeiten (Fähigkeitspassung) einer Person und ihrem (Arbeits-)Umfeld (z. B. Promotionskontext) zu geringerem Wohlbefinden, eingeschränktem Arbeitserfolg und frühzeitigem Drop-Out führen kann (Etzel & Nagy, 2016; Le et al., 2014).

Es stellt sich die Frage, ob eine solche fehlende Passung vor Beginn der Promotion erkannt werden kann, um Personen dabei zu unterstützen, die für sie ideale Entscheidung zu treffen. Studien im Kontext der Studienfachwahl zeigen, dass Personen bereits vor dem Beginn ihres Studiums Prognosen über ihre Passung für ein Studienfach abgeben können und dass diese Prognosen

mit ihrem späteren Wohlbefinden in Zusammenhang stehen (Merkle et al., 2023). Bislang unklar ist jedoch, ob sich diese Befunde auf Promovierende übertragen lassen und ob die Passungsprognosen (=erwartete Passung) prädiktiv für Erfolgsfaktoren, wie Wohlbefinden, den individuell wahrgenommenen Fortschritt sowie Abbruchintentionen in der Promotion sind.

Um diese Forschungslücke zu schließen, führten wir eine Längsschnittstudie durch, in der wir die Teilnehmenden der Studie vom Ende ihres Studiums über die ersten sechs Monate ihrer Promotion begleiteten. In einer ersten Befragung wurden deutschlandweit 648 Studierende mit Promotionsinteresse zu der von ihnen erwarteten Interessens-Promotions-Passung und Fähigkeits-Promotions-Passung befragt (t1). Von diesen begannen 209 Personen (65,6% weiblich; MAlter = 25,38, 48,1% in MINT-Fächern) tatsächlich eine Promotion. Sechs Monate nach Promotionsbeginn wurden diese Promovierenden zu ihrer Arbeitszufriedenheit, ihrem affektiven Erleben (Freude, Ärger, Angst bei der Forschungsarbeit), ihrem wahrgenommenen Fortschritt sowie ihrer Abbruchintention befragt (t2). Zur Beantwortung der Forschungsfrage berechneten wir ein Pfadmodell unter Verwendung einer Maximum-Likelihood-Schätzung mit robusten Standardfehlern.

Die vor Beginn einer Promotion erwartete Interessens-Promotions-Passung war prädiktiv für Arbeitszufriedenheit ($\beta = .24$, $p = .021$), erlebte Freude bei der Forschungsarbeit ($\beta = .28$, $p = .007$), wahrgenommenen Fortschritt ($\beta = .24$, $p = .015$) und Abbruchintentionen ($\beta = -.30$, $p = .003$) sechs Monate nach Beginn der Promotion. Erlebte Angst und erlebter Ärger bei der Forschungsarbeit während der Promotion waren nicht mit der erwarteten Interessens-Promotions-Passung vor der Promotion assoziiert ($p > .391$). Die erwartete Fähigkeits-Promotions-Passung sagte - abgesehen von erlebter Angst bei der Forschungsarbeit ($\beta = -.23$, $p = .024$) - keines der Outcomes signifikant vorher ($p > .085$).

Unsere Ergebnisse zeigen, dass die erwartete Interessens-Promotions-Passung eine Vielzahl bedeutender Outcomevariablen vorhersagt und daher ein hilfreiches Instrument für die Beratung von Studierenden bei Ihrer Entscheidung für oder gegen eine Promotion sein könnte. So könnten beispielsweise Self-Assessment-Tests nach Vorbild von Studienorientierungstests vor Aufnahme einer Promotion Studierende dazu einladen, ihre eigenen Interessen mit den Tätigkeiten und Arbeitsinhalten einer Promotion abzugleichen. Hinsichtlich der erwarteten Fähigkeits-Promotions-Passung ist unklar, ob diese eine geringere Rolle für Promotionserleben und Promotionserfolg spielt oder ob die eingeschränkte prädiktive Validität darauf zurückzuführen ist, dass es Studierenden schwerer fällt, einzuschätzen, ob ihre Fähigkeiten zu einer Promotion passen.

ENTFÄLLT Exzellenz im Wissenschaftssystem: sozio-strukturelle Unterschiede in Karriereaspirationen von Forschenden in frühen und mittleren Karrierephasen

Miriam Fischer

Humboldt-Universität zu Berlin, Goethe-Universität Frankfurt

Die Wissenschaft ist, vielleicht mehr als jedes andere Berufsfeld, von Narrativen über akademische Exzellenz basierend auf Annahmen von Meritokratie und Objektivität durchdrungen, welche die Rolle systemischer und struktureller Faktoren in akademischen Laufbahnen verschleiern (Merzle & Winkelmayer, 2022; Hark & Hofbauer, 2023). Diese institutionelle Betonung von Exzellenz und scheinbar objektiven Leistungsindikatoren vernachlässigt die damit verbundene Subjektivität und die daraus resultierenden Pfadabhängigkeiten (Bol, Vaan & van de Rijt, 2018, 2022; Casper; van der Molen & Bol, 2024). Es ist daher kaum verwunderlich, dass die Gruppe der Personen, die eine akademische Forschungslaufbahn beginnen, in Bezug auf verschiedene sozialstrukturelle Merkmale deutlich weniger vielfältig ist als die allgemeine Bevölkerung. So sind Frauen, rassifizierte Minderheiten, Angehörige der Arbeiterklasse und Menschen mit Behinderungen unter in der Wissenschaft nach wie vor zahlenmäßig unterrepräsentiert, was sich mit zunehmendem Karrierestufen noch verschärft (Casad et al., 2021; Zapfel et al., 2021).

Da das akademische Selbstkonzept ein bekannter Prädiktor für den Bildungserfolg in schulischen Kontexten ist (Möller et al., 2020), konzentriert sich dieser Artikel auf das akademische Selbstkonzept von Wissenschaftler:innen in der frühen und mittleren Karrierephase (Doktorand:innen, Postdocs und Forscher:innen mit befristeten Verträgen) und die Beziehung zur ihrer Karriereplanung. Darüber hinaus werden verinnerlichte Narrative über akademische Exzellenz als ein Mechanismus hinter gruppenspezifischen Unterschieden im akademischen Selbstverständnis und in der Karriereplanung betrachtet. Individuen, die von diesen Erzählungen beeinflusst werden, können solche Botschaften verinnerlichen, was dazu führt, dass sie sich selbst akademische Schwierigkeiten zu schreiben, anstatt strukturelle Barrieren zu erkennen. Jüngste Erkenntnisse aus Deutschland haben gezeigt, dass die Umgestaltung der Sozialstruktur nur verstanden werden kann, wenn sie aus einer mehrdimensionalen Perspektive untersucht wird (Grunow & Khoudja, 2024). Dementsprechend wird die soziale Positionierung durch eine Kombination von sozialstrukturellen Merkmalen entlang der sozialen Trennlinien Geschlecht, Rassifizierung/Ethnizität, Klassenherkunft und Behinderung gemessen. Die zentralen Forschungsfragen lauten: Inwieweit unterscheiden sich Nachwuchswissenschaftler:innen und Nachwuchswissenschaftler mit unterschiedlichem soziostrukturellem Hintergrund (Geschlecht, Rassifizierung/Ethnizität, Klassenzugehörigkeit, Behinderung) in ihren Karriereplanung und akademischen Selbstkonzepten? Welche Rolle spielt die Verinnerlichung des akademischen Exzellenznarrativs bei diesen potenziellen Unterschieden?

Anhand von Daten der Attitudes towards Diversity, Objectivity and Meritocracy in Science (ADOMiS)-Befragung wird in dieser Studie das Zusammenspiel zwischen soziostruktureller Position, akademischem Selbstkonzept und Karriereverläufen untersucht. Die ADOMiS-Befragung (Feldphase Herbst 2024) umfasst Informationen zu retrospektiven und prospektiven Laufbahnen, sowie Einstellungen und Informationen zum Universitätsalltag von Forschenden aller Karrierestufen (Prädoc, Mittelbau, Professuren) an drei Berliner Universitäten und der Charité Berlin. Aufgrund der expliziten Unterstützung der Präsident:innen der jeweiligen Einrichtungen ist eine gute Rücklaufquote zu erwarten. Die Rekrutierung findet über institutsinterne Kanäle statt (Mailinglisten, Intranet, Poster). Etwaige Verzerrungen der Stichprobe werden soweit möglich mit administrativen Daten korrigiert. Mittels Regressionsanalysen wird der Zusammenhang zwischen sozialer Positionierung und Karriereaspirationen für Prädoc- und Postdoc-Forschende getrennt betrachtet und dabei Exzellenznarrative und akademisches Selbstbild als mögliche Mechanismen geprüft. Im zweiten Schritt werden Prädoc- und Postdoc-Forschende verglichen. Die Ergebnisse können als Grundlage für Strategien dienen, die die Entwicklung langfristiger wissenschaftlicher Karrieren von Personen mit marginalisierten soziostrukturellen Merkmalen unterstützen. Vorläufige Ergebnisse und ihre Implikationen werden in diesem Beitrag diskutiert.

Veränderungen der Zielorientierung in der Transition von der Promotion in die Postdoc-Phase

Johanna Ott¹, Ronja Steinhauser², Oliver Dickhäuser², Stefan Janke², Markus Dresel¹, Martin Daumiller¹

¹Universität Augsburg, ²Universität Mannheim

Wissenschaftler:innen haben eine Schlüsselrolle bei der Generierung und Verbreitung von Wissen (Weinberg et al., 2014), wobei Postdoktorand:innen wegen ihrer Forschungserfahrung essenziell sind. Allerdings bilden wenig Studien Postdoktorand:innen und ihre Karriereverläufe und Motivation ab.

Da die Wissenschaft ein markanter Lern- und Leistungskontext ist, bewährte sich die Achievement Goal Theorie zur Beschreibung der Motivation von Wissenschaftler:innen (Daumiller et al., 2020). Wir fokussierten uns auf Lern- und Erscheinungsziele mit ihrer jeweiligen Annäherungs- und Vermeidungsausprägung. Diese erscheinen hinsichtlich der Übergangsphase besonders relevant, da sie Lern- und Selbstdarstellungsaspekte umfassen, die für eine gelingende Transition mit ihren Konkurrenzswängen und Lerngelegenheiten bzw. -erfordernissen bedeutsam sein dürften. Querschnittlich wurden Zusammenhänge von Annäherungslernzielen mit Erfolgsvariablen wissenschaftlicher Karrieren assoziiert- Wohlbefinden, Lernzuwachs und Forschungsleistung (Daumiller et al., 2023; Rinas et al. 2022). Die Befundmuster zu Vermeidungslernzielen sind hingegen spärlich, verweisen aber auf Zusammenhänge mit Ärger und Nutzung von Lerngelegenheiten (Daumiller et al., 2020a; Rinas et al., 2020). Annäherungserscheinungsziele sind mit extrinsischen Lebensaspirationen und dem Einsatz fragwürdiger Forschungspraktiken assoziiert (Janke et al., 2019; Janke et al., 2019). Vermeidungserscheinungsziele hängen indes negativ mit beruflichem Lernen und Wohlbefinden zusammen (z.B. Daumiller et al., 2020).

Längsschnittliche Befunde bei Schüler:innen deuten auf einen Zusammenhang zwischen Veränderungen der Umwelt und Veränderung der Lernzielorientierung hin (Meier et al., 2013). Unklar ist, ob diese Befunde auf Transitionen im Erwachsenenalter übertragbar sind.

Speziell nehmen wir an, dass Annäherungserscheinungsziele am Ende der Promotion und zum Eintritt in die Postdoc-Phase zunehmen und in der Postdoc-Phase weiter steigen, wohingegen Vermeidungserscheinungsziele beim Eintritt in den Postdoc abnehmen und in der Postdoc-Phase weiter sinken sollten. Bezüglich Annäherungslernzielen und Vermeidungslernzielen ist anzunehmen, dass diese am Ende der Promotion sinken, beim Eintritt in den Postdoc ansteigen und im Laufe der Postdoc-Phase erneut abnehmen.

Diese Hypothesen prüften wir anhand einer Längsschnittstudie mit vier Messzeitpunkten in dreimonatigen Abständen (letzter Messzeitpunkt September 2024), an der 538 Wissenschaftler:innen am Ende ihrer Promotion (315 weiblich, 211 männlich, 5 divers) und 131 zu Beginn ihrer Postdoc-Phase (76 weiblich, 52 männlich, 1 divers) von deutschen Universitäten teilnahmen. Die genannten Ziele wurden mit jeweils 4 Items von Daumiller und Dresel (2020a) gemessen (gute Konsistenz: $\alpha > .83$). Es lag kein systematischer Drop-Out.

Die Verteilung der Daten berücksichtigend, berechneten wir Wilcoxon-Rangsummentests mit den ersten zwei Messzeitpunkten und bestimmten die Effektgrößen durch Rang-Biserial-Korrelationen. Zur Konferenz liegen alle Messzeitpunkte vor und werden mittels latenter Wachstumskurvenmodelle ausgewertet.

Für Annäherungserscheinungsziele zeigten sich weder inter- noch intraindividuelle Veränderungen.

Vermeidungserscheinungsziele waren bei Doktorand:innen statistisch signifikant stärker ausgeprägt als bei Postdoktorand:innen ($Mdn=7$ vs $Mdn=6.5$; $p=.045$; kleiner Effekt Cohen's $d=.19$). In der Analyse der individuellen Veränderungen zeigten Doktorand:innen eine statistisch signifikante Abnahme ihrer Vermeidungserscheinungsziele ($Mdn=7$ vs $Mdn=6.75$; $p=.015$; kleiner Effekt Cohen's $d=.12$). Für die Postdoktorand:innen zeigte sich keine signifikante Veränderung.

Im Vergleich der beiden Gruppen zeigten Postdoktorand:innen statistisch signifikant höhere Annäherungslernziele als Doktorand:innen ($Med=7.75$ vs $Mdn=7.50$; $p=.04$; kleiner Effekt Cohen's $d=.21$). Annäherungslernziele nahmen bei Postdoktorand:innen statistisch signifikant ab ($Med=7.75$ vs $Med=7.25$; $p=.025$, kleiner Effekt Cohen's $d=.33$). Dasselbe zeigte sich für Promovierende am Ende ihrer Promotion sowohl für ihre Annäherungslernziele ($Mdn=7.50$ vs $Mdn=7.00$; $p<.001$; mittlerer Effekt Cohen's $d=.55$), sowie ihre Vermeidungslernziele ($Mdn=7.00$ vs. $Mdn=6.625$; $p<.001$, kleiner Effekt Cohen's $d=.24$). Vermeidungslernziele zeigten indes keine signifikanten Veränderungen für Postdoktorand:innen, sowie keine signifikanten Gruppenunterschiede.

Unsere Hypothesen zu Annäherungserscheinungszielen wurden nicht bestätigt, woraus eine relative Stabilität dieser geschlussfolgert werden kann (vgl. Fryer & Elliot, 2007). Die Abnahme der Vermeidungserscheinungszielen könnte durch den abnehmenden Druck durch die beendete Promotion und die Sicherheit des anschließenden Postdoc-Arbeitsverhältnisses erklärt werden. Analysen mit allen Messzeitpunkten werden diese Tendenzen spezifizieren.

Die jeweils kleinen Effekt zur Veränderung der Lernzielorientierung decken sich mit der Literatur zu Transitionsphasen bei Schüler:innen (s.o.). In Anbetracht bisheriger querschnittlicher Befunde könnte eine weitere Abnahme Anlass für Interventionen sein.

Karrierewege ins und im Wissenschaftsmanagement

Julia Rathke¹, René Krempkow², Kerstin Janson²

¹PH Ludwigsburg, ²IU Internationale Hochschule

Mit der Einführung neuer Steuerungsinstrumente haben die Hochschulen mehr Autonomie erhalten und müssen im Gegenzug mehr Managementaufgaben selbst übernehmen (Schimank 2009). Sowohl in der universitären Selbstverwaltung als auch in der Verwaltungsbürokratie (Hüther 2010: 129f.) haben sich daher Managementmodelle etabliert, in denen akademische Leistungsprozesse mit Ressourcenmanagement verknüpft sind (Bogumil et al. 2009). In der Folge haben sich an den Schnittstellen zwischen Wissenschaft und Verwaltung neue Berufsbilder im Wissenschaftsmanagement etabliert (Banscherus et al. 2017: 126f; Windfuhr et al. 2019; Anonymus), die insbesondere für Wissenschaftler*innen eine interessante Karriereoption darstellen und deren Arbeitsaufgaben sich im Übergangsbereich zwischen Wissenschaft und Wissenschaftsverwaltung befinden.

Um der Frage empirisch nachzugehen, welche Karrierewege im Wissenschaftsmanagement existieren und inwiefern sich diese von der Wissenschaft sowie der Verwaltung unterscheiden, nutzen wir umfangreiche Daten, die im Rahmen des Projektes Anonymus erhoben wurden. Zum einen sind dies zwei Online-Surveys mit jeweils mehr als 1.000 Teilnehmenden, zum anderen 111 qualitative Interviews mit Wissenschaftsmanager*innen, Hochschulleitungen und Weiterbildungstrainer*innen. Die Online-Surveys wurden hauptsächlich auf Basis der Mitglieder beruflicher Netzwerke und der Absolvent*innen von Wissenschaftsmanagement-Weiterbildungsstudiengängen mittels personalisierter Einladung jeweils zum Jahreswechsel 2019/20 und 2021/22 durchgeführt. Die qualitativen Interviews erfolgten im Verlaufe der Jahre 2020/21.

Die Daten zeigen, dass Wissenschaftsmanager*innen sich häufig für dieses Berufsfeld entscheiden, weil sie weder „in die reine Wissenschaft“ (Interviewzitat) noch in die Verwaltung wollen und sie es überdies als ein spannendes Berufsfeld mit großen Entscheidungsspielräumen wahrnehmen. Mehrheitlich auf umfangreichen wissenschaftlichen Erfahrungen aufbauend, wollen sie der Wissenschaft verbunden bleiben und ihre erworbenen Kompetenzen weiterhin für den Wissenschaftsbereich einsetzen. Dabei sind die Arbeitsbedingungen im Wissenschaftsmanagement zwar deutlich stabiler (d.h. ca. drei Viertel der Stellen sind unbefristet) als Stellen für Nachwuchswissenschaftler*innen. Die Kehrseite jedoch ist, dass Karrierewege deutlich weniger gut strukturiert sind als in der Wissenschaft oder der Verwaltung. Da Stellen im Wissenschaftsmanagement häufig nach individuellem Hochschulbedarf und oft projektförmig entstehen, ist ungewiss, auf welche zukünftige Stellen Kandidat*innen sich vorbereiten

können. Somit sind Einstiegsmöglichkeiten relativ intransparent, sowie Weiterentwicklungs- und Aufstiegsmöglichkeiten sowie Einkommenszugewinne unklar und recht begrenzt. Zugleich bietet die damit bislang bestehende relative Offenheit des Berufsfeldes Potenziale für soziale Aufsteiger und andere im deutschen Wissenschaftssystem unterrepräsentierte Personen.

In solchen Berufsfeldern spielen – so unsere erste These – sogenannte Schlüsselkompetenzen (sowohl digitale als auch nichtdigitale, vgl. Stifterverband 2018; Anonymus) eine besonders wichtige Rolle. Sie ermöglichen es – so vermuteten wir zweitens – ergänzend zu fachlichen (bzw. im Hochschulbereich: disziplinspezifischen) Kompetenzen, mit neu auftretenden Problemen besser umzugehen und Lösungsvorschläge zu erarbeiten. Dies dürfte – so unsere dritte These – für die betreffenden Personen tendenziell von Vorteil sein und sollte sich auch in deren Karriereweg bzw. beruflichem Erfolg als positiv erweisen.

In unserem Symposium-Beitrag planen wir, empirische Belege für diese Thesen mit den aus dem KaWuM-Projekt vorliegenden quantitativen und qualitativen Daten vorzustellen. Hierfür möchten wir sowohl Ergebnisse aus unseren beiden Surveys als auch aus qualitativen Interviews zeigen und zu ausgewählten Aspekten im Sinne einer Methodentriangulation komplementär zusammenführen.

Unsere Ergebnisse könnten möglicherweise, und dies möchten wir im Anschluss an die Präsentation unserer Ergebnisse zur Diskussion stellen, als symptomatisch für künftig voraussichtlich noch häufiger entstehende neue Berufsfelder gelten: Auch für andere neue Berufsfelder ist ungewiss, auf welche zukünftige Stellen Kandidat*innen sich vorbereiten können. Daher könnte auch hierfür eine umfassendere Bildung unter stärkerer Einbeziehung digitaler und nichtdigitaler Kompetenzen ein Schlüssel sein, um mit künftigen gesellschaftlichen und beruflichen Herausforderungen angemessen(er) umgehen zu können.

S-88: Non-formale berufliche Weiterbildung als Schlüssel zur Anpassung an eine sich wandelnde Arbeitswelt

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 157

Chair der Sitzung: Benjamin Schimke

Non-formale berufliche Weiterbildung als Schlüssel zur Anpassung an eine sich wandelnde Arbeitswelt

Chair(s): **Benjamin Schimke** (Bergische Universität Wuppertal, Deutschland)

Der technologiebedingte strukturelle Wandel der Arbeitswelt ist eine zentrale und aktuelle gesellschaftliche Herausforderung (Bundesministerium für Arbeit und Soziales 2022). Mit zunehmendem Einsatz digitaler Technologien verändern sich sowohl Arbeitsinhalte als auch -prozesse, wodurch Arbeitsmarktteilnehmende einem stetigen Anpassungsdruck ausgesetzt sind, um ihre Marktposition zu erhalten (Autor et al. 2003; Ioannidou und Parma 2022). Ein Großteil der sich daraus ergebenden Qualifizierungsbedarfe von Beschäftigten wird über sogenannte non-formale Weiterbildung gedeckt, d.h. über berufsbezogene Kurse von relativ kurzer Dauer und ohne allgemein anerkannten Abschluss (Eurostat 2016).

Obwohl Arbeitsanbietende durch fortschreitende technische Entwicklungen eine Entwertung ihrer Fachkenntnisse und Kompetenzen erfahren, sind es insbesondere Arbeitsnachfragende, die über die Teilnahme an diesen Weiterbildungsmaßnahmen entscheiden (Cedefop 2015). Neben der Bereitschaft zu lebenslangem Lernen von Seiten der Arbeitskräfte, haben demnach die betrieblichen Rahmenbedingungen eine zentrale Position für die Bewältigung der skizzierten strukturellen Herausforderungen.

Das Symposium wird mithilfe von quantitativen Beiträgen – auf Basis unterschiedlicher Datenquellen – Erkenntnisse zu Determinanten, Moderatoren und Erträgen non-formaler Weiterbildung vor dem Hintergrund eines sich wandelnden Arbeitsmarktes vorstellen und diskutieren. Dabei werden sowohl beide Perspektiven (Arbeitnehmende und Arbeitsnachfragende) in den Blick genommen als auch unterschiedliche Zugänge für die Bewältigung der Anpassungsleistungen und deren Erträge beleuchtet.

Der erste Beitrag widmet sich hierzu einer offenen Frage zur Ungleichverteilung von Weiterbildungschancen in frühen Karrierephasen. Er fokussiert dabei die Interaktion zwischen betrieblichen Rahmenbedingungen für Weiterbildung und den individuellen Bildungserfahrungen junger Arbeitskräfte, die sich aus der früheren schulischen Bildungsepisode speisen. Die vorläufigen Analysen auf Basis der Startkohorte 4 des NEPS zeigen, dass sowohl ein weiterbildungsförderlicher betrieblicher Kontext als auch positive individuelle Bildungserfahrungen mit Weiterbildungsaspirationen in Verbindung stehen. Es findet sich ebenfalls empirische Unterstützung für den angenommenen Moderationseffekt: junge Arbeitskräfte mit positiven Vorerfahrungen profitieren am stärksten von finanzieller und arbeitszeitbefreiender Unterstützung des Betriebs.

Der zweite Beitrag schaut ebenfalls auf Determinanten non-formaler Weiterbildung, hat dabei jedoch einen dezidiert betrieblichen Fokus. Es wird mithilfe des BIBB-Qualifizierungspanels der Frage nachgegangen, welche Auswirkungen die Einführung neuer Informations- und Kommunikationstechnologien (IKT) in Betrieben auf die Beteiligung von Beschäftigten an non-formaler betrieblicher Weiterbildung hat. Die Befunde weisen darauf hin, dass mit IKT-Einführung die Weiterbildungsbeteiligung der Beschäftigten zunimmt. Noch ausstehende Analysen sollen einer potenziellen Heterogenität dieses Zusammenhangs nach dem betrieblichen Digitalisierungsgrad nachgehen, um möglicherweise passgenauere Förderbedarfe für bestimmte Betriebsstrukturen aufzudecken.

Der dritte Beitrag geht mithilfe groß angelegter ‚Informationsbereitstellungsexperimente‘ (Haaland et al. 2023) u.a. der Frage nach, wie Arbeitskräfte die Kosten und den Nutzen von Investitionen in neue Fähigkeiten einschätzen, um damit ihre Beschäftigungschancen in ihrem und anderen Berufen zu verbessern. Dabei erhalten die Befragten auch Informationen darüber, inwieweit ihr derzeitiger Beruf vom technologischen Wandel betroffen sein könnte. Der Beitrag liefert mit seinem innovativen methodischen Ansatz darüber hinaus Hinweise wie Arbeitskräfte, die einem hohen Automatisierungsrisiko ausgesetzt sind, potenzielle Berufswechsel und die damit verbundenen Kosten und Nutzen einschätzen. Die Befragung läuft zurzeit und im Rahmen des Symposiums sollen erste Ergebnisse präsentiert werden.

Der letzte Symposiumsbeitrag adressiert eine offene Frage des Weiterbildungsnutzens und schaut dabei erneut auf betriebsseitige Mechanismen. Im Zentrum steht die Frage, welchen Einfluss berufliche Weiterbildung in Berufen mit unterschiedlichem Automatisierungsrisiko auf die Beschäftigungschancen von Arbeitskräften hat. Auf Basis eines faktoriellen Surveys mit Personalrekrutierenden kann gezeigt werden, dass der Nutzen von Weiterbildungsnachweisen in Personalauswahlverfahren mit steigendem Automatisierungsrisiko des Zielberufs abnimmt. Dieser Effekt gilt sowohl für Kurse und darin erworbene spezifische Fertigkeiten, die mit den Stellenanforderungen korrespondieren, als auch für Kurse, die generell einsetzbare Qualifikationen vermitteln.

Beiträge des Symposiums

Continuing Training Aspirations of Young Workers in the Context of Previous Educational Experiences and Companies' Training Opportunities

Mortimer Schlieker¹, Matthias Siembab¹, Martin Ehlert²

¹Federal Institute for Vocational Education and Training, BIBB, ²WZB Berlin Social Science Center and FU Berlin

Due to rapid technological advances, continuous training (CT) has become increasingly important for workers' employability and career development. At the same time, employers are forced to continuously update the skills and knowledge of their employees in order to remain competitive (World Economic Forum 2023). However, participation in CT is unevenly distributed: highly educated workers are more likely to participate. This accumulation of advantages, often referred to as the 'Matthew effect' in reference to Merton's (1968) research, is mainly due to placement in the labor market. Job and firm characteristics explain most of the relationship between education and training (Hornberg et al. 2024). As a result, workers with privileged educational backgrounds are in advantageous positions in the labor market, allowing them to adapt to new technologies more quickly.

To explore the roots of inequality in CT, we look more closely at the link between schooling and initial job placement in the early stages of the life course. It is during this formative period that fundamental career decisions are made and aspirations for participation in CT are established, laying the foundation for future career development and adaptation. Previous research has shown that low qualifications and negative prior educational experiences can hinder CT participation throughout the career (Sandau 2022). Our study adds to this by examining how young workers' prior educational experiences interact with training opportunities available in their firms to form training aspirations. Accordingly, we address two interrelated research questions:

1. To what extent do prior educational experiences influence the aspirations of young workers to participate in CT?
2. Do firm characteristics moderate the effect of prior educational experience on CT aspiration?

Theoretically, we combine individual-level and firm-level theories embedded in a life-course framework to develop our hypotheses. At the individual level, value expectancy theory (Eccles et al. 1983) suggests that positive educational experiences lead to higher expectations of success and higher value placed on education. This should lead to higher motivation to participate in CT (Gorges und Kandler 2012). At the firm level, we build on transaction cost theory (Williamson 1985), which suggests that firms will invest in employee training only if they can expect a payoff. This is more likely for highly educated workers. However, this can be overcome by institutions such as company policies and works councils (Wotschack und Solga 2014). Therefore, we expect that fostering firm CT structures as well as CT practices will mitigate the effect of negative prior training experiences.

We use data from the Starting Cohort 4 of the National Educational Panel Study (NEPS), which follows a cohort of 9th graders through their first years of employment. Using linear probability models, we examine the relationship between prior educational experiences (operationalized by school satisfaction), supportive company training contexts (operationalized by financial and time support as well as the existence of CT structures), and young workers' aspirations for non-formal training.

Preliminary findings indicate that positive educational experiences and promoting company contexts are positively associated with young workers' aspirations for non-formal training. Moreover, company support seems to moderate (strengthen) the association between school satisfaction and training aspirations: young workers who were (moderately) highly satisfied benefit most from offered time off and financial support for training. Thus, firm policies seem to amplify inequalities due to previous educational experiences. Further analyses will examine the mechanisms behind this more closely to find out why early educational experiences shape the training trajectories and hence the employability on a labor market influenced by rapid technological change.

Digitale Transformation von Betrieben und die Teilnahme von Beschäftigten an betrieblicher Weiterbildung

Felix Lukowski, David Samray

Bundesinstitut für Berufsbildung, BIBB

Gesellschaftliche Phänomene wie die Digitalisierung, Dekarbonisierung, der demografische Wandel und die De-Globalisierung stellen die Wirtschaft zunehmend vor neue Herausforderungen. Infolgedessen unterliegt die Arbeits- und Berufswelt einem fortwährenden Wandel. Besonders im Kontext der Technologienutzung lassen sich Veränderungen in Arbeitsprozessen beobachten. So verändern sich infolge technologischer Neuerungen nicht nur Arbeitsinhalte, sondern auch die Anforderungen an die Arbeitnehmenden. Um auf Veränderungen von Arbeitsinhalten infolge der digitalen Transformation zu reagieren, wird vor allem auf die Bedeutung beruflicher Weiterbildung als Anpassungsmechanismus verwiesen (Anbuhl 2019; Bundesministerium für Arbeit und Soziales 2022).

Der Beitrag schaut aus betrieblicher Perspektive auf einen sich verändernden Arbeitsmarkt und nimmt dabei Auswirkungen auf betriebliche Weiterbildung in den Blick. Im Speziellen soll untersucht werden, wie sich die Einführung neuer Informations- und Kommunikationstechnologien (IKT) in Betrieben auf die Beteiligung von Beschäftigten an non-formaler, betrieblicher Weiterbildung auswirken. Hierzu zählen etwa betrieblich finanzierte Kurse, Lehrgänge und Seminare für Beschäftigte, die nicht zu einem staatlich anerkannten Berufsabschluss führen. Besonderes Augenmerk wird in den Analysen auf Heterogenitäten zwischen Betrieben mit einem unterschiedlichen Digitalisierungsgrad gelegt.

Der Fokus auf betrieblich geförderte non-formale Weiterbildung wird dadurch begründet, dass Betriebe als die größte Finanzierungsquelle beruflicher Weiterbildung in Deutschland gelten (Schönfeld und Behringer 2017, S. 73), während non-formale berufliche Weiterbildung die am häufigsten praktizierte Weiterbildungsform darstellt (Ebner und Ehlert 2018, S. 213).

Als theoretische Grundlage werden die Humankapitaltheorie und der Kontingenzansatz herangezogen. Auf Grundlage der Humankapitaltheorie (Becker 1993) kann argumentiert werden, dass Investitionen in das Humankapital der Beschäftigten einen ökonomischen Nutzen für Unternehmer:innen darstellen. Beispielsweise kann Weiterbildung dazu beitragen, dass technische Arbeitsmittel von Beschäftigten produktiv genutzt werden können. Insbesondere in Folge der Neueinführung von Technologien sollte es darum zu Investitionen in betriebliche Weiterbildung kommen. Ebenfalls tragen situative organisationstheoretische Ansätze wie der Kontingenzansatz (Fiedler 1967) zur Erklärung betrieblicher Weiterbildung bei. Demzufolge kann angenommen werden, dass betriebliche Weiterbildung einen Mechanismus darstellt, mit dem sich Betriebe an eine sich wandelnde Umwelt anpassen (Käpplinger 2016, 97f.). Beispielsweise erzeugt technologischer Wandel einen Qualifizierungsbedarf, der dann zu einer Reaktion des Unternehmens in Form von Weiterbildung seiner Mitarbeitenden führt.

Datenbasis der Untersuchung sind Längsschnittdaten des BIBB-Betriebspanel zu Qualifizierung und Kompetenzentwicklung (kurz: BIBB-Qualifizierungspanel), mit denen Auswirkungen der Einführung von IKT auf die Beteiligung von Beschäftigten an non-formaler Weiterbildung mittels Fixed-Effects-Modellen berechnet werden. Das BIBB-Qualifizierungspanel ist eine jährliche Befragung von Betrieben in Deutschland zur Personalsituation und betrieblichem Qualifizierungsgeschehen. Für die Analysen werden die Erhebungswellen 2016 bis 2023 herangezogen. Die Stichprobengröße der Befragung lag in diesen Zeitraum jeweils bei rund 4.000 Betrieben pro Befragungswelle.

Der Beitrag baut auf Untersuchungen auf, in denen bislang noch nicht zwischen Betrieben mit unterschiedlichem Digitalisierungsgrad differenziert wurde. Hier zeigte sich ein positiver Zusammenhang zwischen IKT-Einführung und der Teilnahme der Beschäftigten an betrieblichen Weiterbildungsmaßnahmen (Baum und Lukowski 2022). Unter Betrieben in Deutschland herrscht jedoch eine große Heterogenität, was den Stand der Technik anbelangt (Baum und Fournier). Der Einfluss von IKT-Einführung auf (non-formale) betriebliche Weiterbildung soll deshalb differenziert nach betrieblichem Digitalisierungsgrad untersucht werden. Die Befunde sollen anschließend dahingehend diskutiert werden, ob sich aus ihnen Förderbedarfe für bestimmte Betriebsgruppen ergeben (z.B. KMU mit geringem Digitalisierungsstand). Ebenso soll diskutiert werden, welche weiteren Forschungsdesiderate sich aus der Empirie ergeben.

Understanding On-the-Job Search

Pascal Heß¹, Armando Miano²

¹Institute for Employment Research, IAB, ²University of Naples Federico II, CESF

The goal of this project is to improve our understanding of how workers decide about job mobility, focusing especially on workers' beliefs about the costs of searching for a new job within their current or in another occupation and their perceptions of their opportunities in the labor market. From a theoretical perspective, job mobility – moving to a new job within the same or in another occupation – is a crucial driver of labor market dynamism and an important factor shaping workers' personal development and

earnings growth. Job mobility improves matches between employers and employees, as it pushes workers to search for jobs that better suit their skills and helps employers to find workers with the skills they need more easily.

Yet, we still know relatively little about the search behavior of employed workers and about barriers preventing workers to move to other (better-suited) jobs. People who are already in the job market often are not well informed about other career opportunities – other jobs where they would be paid more, but also opportunities for personal development and skills enhancement (Jäger et al. 2024; Miano 2023). We know even less about the costs and the barriers that workers face to move to a new job and how workers perceive these costs.

At the same time, occupation mobility has gained increased relevance in times of profound transformations in the labor market. Certain occupations are particularly affected by technological change and the green transition and, thus, are at risk of becoming obsolete (Heß et al. 2023). However, the literature does not yet provide clues as how well employees are informed about such potential job-change induced career changes. How do workers think about changing jobs? How do they assess the costs and benefits of looking for new job opportunities? Do they consider moving to a different occupation and how do they perceive the benefits and costs of that? How informed are they about their opportunities in theirs and other occupations? How do workers perceive the costs and benefits of investing in new skills to take advantage of opportunities in theirs and other occupations?

To answer these questions, we design a new comprehensive online survey. We plan to administer this survey to a sample of 5,000 German employees, focusing especially on workers in occupations that are more at risk of being displaced by structural changes in the labor market – automation, green transition. In the survey we will elicit workers' beliefs about how costly it would be for them to find a new job within their current occupation and, crucially, in other occupations, as well as their beliefs about what occupations are 'adjacent' to theirs. By linking the survey records with detailed administrative data – conditional on respondents' consent – we will be able to investigate how workers' past experiences shape their beliefs and how (possibly distorted) perceptions affect workers' employment choices and career trajectories – 'allocative effect' of (mis-)perceptions.

The survey will feature multiple information provision experiments (Haaland et al. 2023) where randomly selected groups of respondents will receive information on which occupations are closest to them in terms of task similarity paired with additional information on the benefits or costs (need to re-retrain or acquire a license) of moving to such occupations. Some respondents will also receive information about the extent to which their current occupation could be affected by technological change or the green transition. The panel dimension of the survey will allow us to study beliefs updating, persistence of the treatment effects, and, especially, to compare expectations with realizations for respondents who will change jobs in between the two surveys.

Signalwert von Weiterbildungsnachweisen in Berufen mit unterschiedlichem Automatisierungsrisiko: Experimentelle Evidenz aus einem faktoriellen Survey

Benjamin Schimke

Bergische Universität Wuppertal

Der zunehmende Einsatz von Technologien hat tiefgreifende Auswirkungen auf Arbeitsmärkte weltweit. Insbesondere die Automatisierung von Routineaufgaben und die technologische Komplementierung von Nicht-Routineaufgaben führt zu einer Polarisierung der Arbeitsmärkte. Berufe, die weniger automatisierungsgefährdet sind, erleben eine positive Entwicklung der (Wieder-)Beschäftigungschancen – in einigen Ländern auch Lohnzuwächse – während in Berufen mit hohem Automatisierungsrisiko schlechtere Arbeitsmarktchancen bestehen (Acemoglu et al. 2022; Goos et al. 2014; Autor und Dorn 2013; Spitz-Oener 2006). Im Kontext dieses 'skill-biased technological change' stellt sich die Frage, wie betroffene Arbeitskräfte ihre Beschäftigungsfähigkeit erhalten (Autor et al. 2003).

Ein Wandel der Qualifikationsanforderungen und die damit „schrumpfende Haltbarkeit von Fachkenntnissen“ (Ra et al. 2019) dürften zu einer Entwertung von Kompetenzen führen und damit zu einer verstärkten Notwendigkeit von Weiterbildungsmaßnahmen, um das eigene Humankapital zu erhalten (Becker 1993). Allerdings werden rund 90% aller non-formalen Weiterbildungen durch Arbeitgebende finanziert (Cedefop 2015). Es sind folglich Arbeitsnachfragende, die über Weiterbildungsinvestitionen entscheiden und das Training ihrer Arbeitskräfte mit Routinetätigkeiten ist keine rationale Strategie, wenn Technologien im Vergleich zu menschlicher Arbeitskraft einen Wettbewerbsvorteil bieten. Während empirische Evidenz zur Frage, wie Automatisierungsrisiken die Weiterbildungsbeteiligung bestimmen in den letzten Jahren zugenommen hat, sind Erkenntnisse zur möglichen Heterogenität des Weiterbildungsnutzens rar (Heß et al. 2023; Lukowski et al. 2021; Ehlert 2020). Die vorliegende Studie untersucht vor diesem Hintergrund Auswirkungen beruflicher Weiterbildung auf Einstellungschancen von Arbeitskräften in Berufen mit unterschiedlichem Automatisierungsrisiko. Dabei wird mithilfe von filter- sowie passungstheoretischen Überlegungen (Arrow 1973; Jovanovic 1979) geprüft, ob Weiterbildungsmaßnahmen, die auf Nicht-Routineaufgaben abzielen, die Chancen auf eine neue Beschäftigung erhöhen.

Zur Beantwortung dieser Frage wurde ein Vignettenexperiment mit Personalrekrutierenden (N=717) durchgeführt, die in ihrem Arbeitsalltag Bewerbende in einem von 15 Berufen rekrutieren (Auspurg und Hinz 2015). Zu Beginn des Experiments wurde jedem Probanden eine fiktive Stellenausschreibung aus jenem Beruf präsentiert, in dem sie selbst Einstellungsexpertise besitzen. Im Anschluss sollten sie nacheinander insgesamt acht tabellarische Lebensläufe auf einer 11-stufigen Skala (von 0 bis 100%) hinsichtlich der Wahrscheinlichkeit einer Vorstellungsgesprächseinladung bewerten. Die experimentell variierten Faktoren umfassten u.a. das Geschlecht, Abschlussnoten sowie ein Faktor mit Weiterbildungsnachweisen. Das Design umfasst insgesamt 10 Faktoren, ist strukturell identisch für alle 15 Berufe und basiert auf 192 Lebensläufen, die experimentell aus dem Vignettenuniversum (2634) (D-Effizienz: 92,8%) ausgewählt und anschließend experimentell auf 24 Vignettensets verteilt wurden. Die Analyse der bewerteten Vignetten (V=5.443) erfolgt mittels hierarchisch-linearer Mehrebenenregression. Die Automatisierungsrisiken der Berufe wurden über den Indikator von Dengler und Matthes (2018) abgebildet und zwei Robustheitsprüfungen aller Analysen mit alternativen Messungen nach Somers und Fouarge (2022) und (Webb 2019) durchgeführt.

Die Ergebnisse zeigen, dass Weiterbildungsmaßnahmen, die auf die Anforderungen der Stellenausschreibung abgestimmt sind, die Chancen auf eine Einladung zum Vorstellungsgespräch signifikant erhöhen. Ein entsprechender Nachweis erhöht die durchschnittliche Wahrscheinlichkeit einer Einladung um 3,4 Prozentpunkte im Vergleich zu Bewerbungen ohne Weiterbildung. Maßnahmen, die nicht direkt mit den Stellenanforderungen in Zusammenhang stehen, aber von den Rekrutierenden als grundsätzlich sinnvoll für die Ausübung des Berufs angesehen werden, bieten im Mittel einen geringeren Vorteil von 2,1 Prozentpunkte. Der Beschäftigungsnutzen dieser Nachweise nimmt mit steigendem Automatisierungsrisiko jedoch ab. In Berufen mit hohem Automatisierungsrisiko ist der Vorteil von Weiterbildungszertifikaten, die eine Passung mit den Stellenanforderungen aufweisen signifikant geringer, aber bleibt auch in Berufen mit maximaler Automatisierungsgefährdung signifikant. Der positive Effekt für Weiterbildungsnachweise, die zu Fertigkeiten führen, welche in der Stellenausschreibung nicht direkt nachgefragt wurden, ist bei einem Automatisierungsrisiko von über 75% hingegen nicht mehr von Null differenzierbar ist.

Zusammenfassend zeigt die Studie, dass non-formale berufliche Weiterbildung ein wirksames Mittel zur Erhöhung der Beschäftigungsfähigkeit darstellt, wobei die Wirksamkeit durch das Automatisierungsrisiko des Berufs moderiert wird. Die Erkenntnisse bieten wertvolle Hinweise für Unternehmen und politische Entscheidungsträger, um Weiterbildungssysteme an die Herausforderungen der digitalen Transformation anzupassen und die Wettbewerbsfähigkeit von Arbeitskräften zu sichern.

S-89: Reproduzierbarkeit und Datenqualität bei Forschungssynthesen: Status Quo und Lösungsansätze für eine transparente und nachhaltige Methode

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 159

Chair der Sitzung: Nadeshda Jung

Chair der Sitzung: Tamara Heck

Chair der Sitzung: Tanja Burgard

Reproduzierbarkeit und Datenqualität bei Forschungssynthesen: Status Quo und Lösungsansätze für eine transparente und nachhaltige Methode

Chair(s): **Nadeshda Jung** (DIPF, Deutschland), **Tamara Heck** (DIPF, Deutschland)

Diskutant*in(nen): **Tamara Heck** (DIPF, Deutschland), **Tanja Burgard** (ZPID, Deutschland)

Reproduzierbarkeit und Datenqualität sind zentrale Prinzipien in der wissenschaftlichen Forschung, insbesondere bei Forschungssynthesen wie systematischen Reviews und Meta-Analysen. Diese Prinzipien sind essenziell, um die Verlässlichkeit, Transparenz, Interoperabilität und Nachnutzung wissenschaftlicher Erkenntnisse sicherzustellen.

Reproduzierbarkeit bedeutet, dass Forschungsergebnisse durch Anwendung derselben Methoden und unter denselben Voraussetzungen zu denselben Befunden kommen (Rahal et al., 2022). Dies bedeutet, dass Methoden und Daten gemäß bestimmten Kriterien zur Transparenz und Nachvollziehbarkeit aufbereitet und verfügbar gemacht werden müssen, damit Forschende die Erkenntnisse der Primärstudien nachvollziehen und überprüfen können, ob sie bei der Anwendung derselben Schritte die gleichen Ergebnisse erzielen.

Bei Forschungssynthesen, die auf der höchsten Evidenzstufe angesiedelt sind (Burns et al., 2011) und die in der Bildungsforschung sowohl für die Wissenschaft als auch für die Bildungspraxis und Politik als Entscheidungsbasis dienen sollen, kommt Reproduzierbarkeit und Datenqualität eine besondere Rolle zu. Diese sind nicht nur von Relevanz, um die Qualität der Forschungssynthesen und damit ihre Aussagekraft zu bewerten, sondern sie sind auch essentiell, wenn es um nachhaltige Verwertung und Fortschreibung der Forschungssynthesen geht. Wissenschaftliche Erkenntnisse sind einem ständigen Wandel unterworfen, da neue Studien frühere Ergebnisse von Forschungssynthesen bestätigen, erweitern oder widerlegen können. Die Erstellung wiederholbarer, sogenannter „Living Reviews“ ermöglicht es, transparente und reproduzierbare Methoden sowie Daten aus bestehenden Forschungssynthesen zu nutzen. Dadurch können Evidenzen effektiv geprüft und aktualisiert werden (Elliot et al., 2014). Voraussetzung dafür ist jedoch die konsequente Anwendung von Regeln zur Reproduzierbarkeit und Sicherstellung der Datenqualität.

Beide Kriterien stellen jedoch auch erhebliche Herausforderungen für die durchführenden Forschenden dar. Zum einen können bei Forschungssynthesen hochinferente Beurteilungen über die Relevanz und Qualität von Studien notwendig sein, was die Dokumentation und Nachvollziehbarkeit der Prozesse und Daten erschwert. Zum anderen erfordern Tools zur Automatisierung von Forschungssynthesen (Campos et al., 2024; Marshall et al., 2022), dass Prozesse angepasst und Daten in verschiedene Formate transferiert werden müssen. Zudem werden mit diesen Tools neue Daten und Metadaten generiert, die transparent dokumentiert werden müssen.

In diesem Symposium werden Forschungsarbeiten vorgestellt, die sich mit der Reproduzierbarkeit, Datenqualität und der Dokumentation von Prozessen, Ergebnissen und Daten bei Forschungssynthesen beschäftigen und den Fragen nachgehen, wie die Methode der Forschungssynthese transparenter und nachhaltiger werden kann und welche Standards benötigt werden. Beitrag 1 stellt Ergebnisse einer Untersuchung der Reproduzierbarkeit von Forschungssynthesen in der Bildungsforschung und Psychologie sowie Qualitätskriterien zur Überprüfung dieser vor. Beitrag 2 diskutiert aus sozialwissenschaftlicher Sicht Ansätze zur qualitätsgetriebenen und reproduzierbaren Datenselektion, die bei verschiedenen Prozessen während einer Forschungssynthese zum Tragen kommen. Beitrag 3 diskutiert die Relation zwischen dem Publikations-Bias verschiedener Disziplinen und den Auswirkungen bei der Literaturrecherche und -selektion.

Beiträge des Symposiums

Analyse der Reproduzierbarkeit bei Forschungssynthesen in der Bildungsforschung und Psychologie: Ein Beitrag zu Standards bei der (Daten-)Dokumentation

Nadeshda Jung¹, Claudiu Petrulic², Tanja Burgard², Tamara Heck¹

¹DIPF, Deutschland, ²ZPID, Deutschland

Reproduzierbarkeit und Datenqualität sind grundlegende Prinzipien in der wissenschaftlichen Forschung, insbesondere bei Forschungssynthesen wie systematischen Reviews und Meta-Analysen. Sie gewährleisten die Verlässlichkeit, Transparenz und Nachnutzung wissenschaftlicher Erkenntnisse. Reproduzierbarkeit bedeutet, dass ein Forschungsergebnis bei Anwendung derselben Methoden und Voraussetzungen reproduziert werden kann (Rahal et al., 2022). Das setzt voraus, dass Methoden und Daten transparent aufbereitet und zugänglich gemacht werden, damit andere Forschende die Studien nachvollziehen und überprüfen können.

Jedoch bestätigen Studien zur Transparenz und Reproduzierbarkeit von publizierten Forschungssynthesen diesbezüglich Unzulänglichkeiten, wenn sich auch gleichsam die Dokumentation der Methode insgesamt über die Jahre verbessert zu haben scheint (Polanin et al., 2020). Wong und Bouchard (2023) untersuchten 33 Metaanalysen aus der Kriminologie und konnten bei keiner eine Vollständigkeit der 22 angelegten Kriterien feststellen. Ähnliche Ergebnisse sind auch aus der Medizin (Sun, et al., 2019; Maticic et al., 2019) und der Psychologie bekannt (Polanin et al., 2020; Lakens et al., 2017). In einer aktuellen Untersuchung aus der Bildungsforschung waren basierend auf 12 ausgewählten Kriterien von 361 Forschungssynthesen nur 16 voll reproduzierbar (Buntins et al., 2023).

Um diese Problematik anzugehen, sollen in diesem Beitrag die Entwicklung von Standards zur Prozess- und Datendokumentation von Forschungssynthesen sowie deren Beschreibung als Metadaten thematisiert werden. Diese können sowohl von Forschenden genutzt werden, um ihre Arbeit transparent und reproduzierbar zu machen, als auch von Forschungsdatenzentren zur Qualitätskontrolle bei der Aufnahme von Forschungsdaten von Forschungssynthesen.

Dabei werden folgende Forschungsfragen untersucht: a) Wie sollte ein Kriterienkatalog für die Dokumentation unterschiedlicher Typen von Forschungssynthesen gestaltet sein? b) Wie ist dieser Kriterienkatalog als Entscheidungshilfe für Forschende anzuwenden? c) Ist dieser Kriterienkatalog geeignet, um die Reproduzierbarkeit von Forschungssynthesen in der Bildungsforschung und Psychologie zu prüfen?

Dafür wurden spezifische Kriterien zur Bewertung der Reproduzierbarkeit entwickelt, angelehnt an vorhandene PRISMA-Leitlinien (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Page et al., 2021; Rethlefsen et al., 2021) und den FAIR-Prinzipien (Wilkinson et al., 2016). Zudem wurde sich an der Untersuchung verschiedener Typen von Forschungssynthesen orientiert (Sutton et al., 2019) und Vergleiche angestellt, um Ähnlichkeiten und Unterschiede in den Metadatenattributen zu identifizieren. Ein besonderer Fokus bei der Kriterienentwicklung liegt dabei auf der vollständigen und transparenten Dokumentation aller Prozessschritte, einschließlich der Suchstrategien, Auswahlkriterien und Datenanalyseverfahren. Für die Untersuchung der Anwendung des Kriterienkatalogs für die Reproduzierbarkeit von Forschungssynthesen wurde eine Stichprobe aus etwa 100 Forschungssynthesen aus der Psychologie und den Bildungswissenschaften gezogen. In zentralen Datenbanken (PSYINDEX, FIS Bildung, ERIC; SSCI, SCI-EXPANDED, AHCI, ESCI) wurde dafür mit zehn geprüften Suchanfragen (Jäger-Dengler-Harles et al., 2022) recherchiert. Insgesamt wurden 11896 Ressourcen in den drei Datenbanken gefunden, die meisten davon mit den Begriffen "systematic review" und Variationen von "systematic literature review". Für eine gemischte Stichprobe sind die Daten für jede Disziplin entsprechend den deutsch- und englischsprachigen Veröffentlichungen ausbalanciert.

Der derzeitige Kriterienkatalog umfasst eine Vielzahl verschiedener Kriterien, während die Entscheidungsmatrix mehrere Familien bzw. Typen von Forschungssynthesen unterscheidet. Diese Matrix hat eine anpassbare Struktur, die eine Anwendung auf unterschiedliche Forschungsergebnisse in verschiedenen Disziplinen ermöglicht. Die Kriterien wurden kartiert, den Prozessschritten zugeordnet und Redundanzen entfernt.

Die Ergebnisse dieser Arbeit dienen zur Entwicklung von Standards zur Prozess- und Datendokumentation von Forschungssynthesen sowie deren Beschreibung mit Hilfe von Metadaten. Diese können sowohl von Forschenden genutzt werden, um ihre Arbeit transparent und reproduzierbar zu machen, als auch von Forschungsdatenzentren zur Qualitätskontrolle bei der Aufnahme von Daten von Forschungssynthesen. Die beiden betrachteten Disziplinen decken dabei ein breites Spektrum an methodischem Vorgehen bei Forschungssynthesen ab, sodass die Ergebnisse auch für andere Fächer adaptierbar sind. Zur Unterstützung von Forschenden bei der Anwendung des Kriterienkatalogs wurde eine Entscheidungsmatrix entwickelt, die Forschende mit Leitfragen zu ihrem Typ der Forschungssynthese und dem damit verbundenen Kriterienkatalog führt.

Qualitätsgetriebene und reproduzierbare Datenselektion in Forschungssynthesen

Jessica Daikeler
GESIS, Deutschland

In der sozialwissenschaftlichen Forschung sind Reproduzierbarkeit und Datenqualität von zentraler Bedeutung, insbesondere bei Forschungssynthesen wie systematischen Reviews und Meta-Analysen. Beitrag 2 unseres Symposiums widmet sich der qualitätsgetriebenen Datenselektion.

Datenqualität ist ein zentraler Begriff in der wissenschaftlichen Forschung und kann in intrinsische und extrinsische Qualität unterteilt werden (Fröhling et al., 2023; Herzog et al., 2007). Diese Qualitätsdimensionen sind essentiell, um die Verlässlichkeit und Reproduzierbarkeit wissenschaftlicher Ergebnisse sicherzustellen und die Zweitnutzung von Daten voranzutreiben. Besonders in der sozialwissenschaftlichen Forschung, wo Datenquellen durch die Erweiterung traditioneller Datenquellen wie Umfragedaten durch digitale Verhaltensdaten und verlinkte Daten immer vielfältiger und komplexer werden, spielt die Bewertung der Datenqualität eine entscheidende Rolle (Fröhling et al., 2023; Sen et al., 2021).

Das systematische Review von Daikeler et al. (2024) gibt eine Übersicht zu Konzepten und zur Bewertung der Datenqualität im Zeitalter der digitalen Sozialforschung. Die Studie zeigt, dass traditionelle Qualitätskriterien und -Indikatoren erweitert und an die spezifischen Eigenschaften digitaler Daten angepasst werden müssen, um deren Potenzial voll auszuschöpfen. Traditionelle intrinsische Qualitätsdimensionen wie die Repräsentations- und Messverzerrungen von Datenquellen eignen sich nach wie vor als Ansatzpunkte zur Datenqualitätsbeschreibung. Eine häufige Kritik bei der Synthese von Evidenz ist die Perspektive des "Waste Management Systems", also der Vorwurf, dass nur qualitativ schlechte Studien in eine Evidenzsynthese integriert werden, wodurch durch deren Akkumulation nur ein "Waste Management" betrieben wird und keine hochwertige Evidenz subsumiert werden kann (Borenstein et al., 2009). Gemäß Viechtbauer (2010) ist das erste Ziel dieses Beitrags die Diskussion der Rolle und Adressierung von Qualitätsdimensionen bei der Synthese von Evidenz.

Das zweite Ziel dieses Beitrags ist die Diskussion eines Konzeptes zur didaktischen Vermittlung sozialwissenschaftlicher Datenqualität, wie es im neu geschaffenen Kompetenzzentrum Datenqualität in den Sozialwissenschaften (KODAQs) verfolgt wird (Dehne et al., im Erscheinen). KODAQs fokussiert auf die Bewertung und Verbesserung von Datenqualität in den beiden didaktischen Angeboten der Data Quality Academy und der Data Quality Toolbox.

Zusammenfassend bietet Beitrag 2 einen umfassenden Überblick über die Bedeutung und Definition der Datenqualität in der sozialwissenschaftlichen Forschung und stellt praxisnahe Ansätze zur Sicherstellung und didaktischen Vermittlung von Datenqualität vor. Ferner wird die Brücke zur Adressierung von Qualitätsaspekten bei der Akkumulation von Evidenz näher beleuchtet.

Prävalenz und akteursabhängige Risikofaktoren von Publikationsverzerrungen in den Sozial-, Wirtschafts- und Verhaltenswissenschaften

Désirée Nießen¹, Carolin Poppa², Jessica Daikeler¹, Henning Silber¹, Bernd Weiß¹, David Richter²

¹GESIS, Deutschland, ²SHARE BERLIN Institute GmbH

Seit über einem Jahrzehnt wird über die (Nicht-)Replizierbarkeit empirischer Ergebnisse debattiert (z. B. Pashler & Harris, 2012). Im Zuge dieser Replikationskrise wurden einige potenzielle Ursachen für die geringe Replizierbarkeit identifiziert und durch die Förderung von z. B. Präregistrierungen, offenen Daten und Materialien angegangen (z. B. Pennington, 2023). Dennoch kommt es weiterhin vor, dass wissenschaftliche Ergebnisse nicht repliziert werden können. Publikationsverzerrungen wurden hierbei als ein möglicher Faktor identifiziert (z. B. Berinsky et al., 2021; Gollwitzer et al., 2020). Eine Publikationsverzerrung kann als die (Nicht-)Veröffentlichung von Studienergebnissen angesehen werden, die auf deren Richtung oder Stärke zurückzuführen ist (z. B. Dickersin, 2005). Das heißt, dass z. B. statistisch signifikante Ergebnisse mit größerer Wahrscheinlichkeit veröffentlicht werden als statistisch nichtsignifikante (z. B. Auspurg & Hinz, 2011; Franco et al., 2014). Dadurch wird die Bedeutsamkeit von Befunden überschätzt. Infolgedessen erschweren Publikationsverzerrungen die Beurteilung des wahren Wissensstandes zu einer bestimmten Forschungsfrage, was zu wissenschaftlichen Debatten und politischen Entscheidungen führen kann, die auf unzureichenden und unvollständigen Ergebnissen beruhen, so dass falsche Empfehlungen gegeben werden. Mehrere Faktoren beeinflussen das Entstehen von Publikationsverzerrungen, vor allem die direkt beteiligten Akteure (d. h. Autoren, Herausgeber und Gutachter). Darüber hinaus wirkt sich auch der disziplinäre und kulturelle Hintergrund auf die Wahrscheinlichkeit und das Ausmaß von Publikationsverzerrungen aus (z. B. Auspurg et al., 2014).

In der vorliegenden Studie konzentrieren wir uns auf die verschiedenen Schritte des Entscheidungsprozesses von Forschern, die zur selektiven Veröffentlichung von Ergebnissen führen können, während andere Ergebnisse in der „Aktenschublade“ verbleiben. Zu diesem Zweck betrachten wir die erfolgreichen Studieneinreichungen zwischen 2012 und 2021 (N = 178) bei zwei im deutschen Kontext einzigartigen probabilistischen Panels (dem GESIS Panel und der Innovationsstichprobe des Sozio-ökonomischen Panels [SOEP-IS]), bei denen externe Forscher aus allen sozial-, wirtschafts- und verhaltenswissenschaftlichen Disziplinen eigene Studien einreichen können. Diese müssen allerdings eine ausführliche Beschreibung u. a. des theoretischen Hintergrunds, der Forschungsfragen und des Forschungsdesigns enthalten, die dann einem Peer Review unterzogen werden.

Mithilfe eines eigens entwickelten Codierschemas extrahieren wir ausführliche Informationen zu drei Elementen (d. h. Studieneinreichung, Veröffentlichung und Autor). Darüber hinaus führen wir eine Autorenbefragung durch, um weitere akteursbezogene Daten zu Persönlichkeitsmerkmalen sowie Motiven, Einstellungen und Werten in Bezug auf den Forschungsprozess zu erheben. Auf diese Weise kann der gesamte Forschungszyklus beobachtet werden: von der Studieneinreichung, über die Durchführung, zur (idealerweise) anschließenden Veröffentlichung.

Konkret verfolgen wir zwei Ziele: Erstens wollen wir die Prävalenz und das Ausmaß von Publikationsverzerrungen in den Sozial-, Wirtschafts- und Verhaltenswissenschaften in Deutschland messen. Dabei untersuchen wir die Diskrepanz zwischen u. a. den vorgeschlagenen Forschungsfragen und Hypothesen, die in den GESIS-Panel- und SOEP-IS-Studieneinreichungen formuliert wurden, und der Darstellung in der anschließenden tatsächlichen Veröffentlichung. Jede Abweichung von der ursprünglichen Einreichung kann als Hinweis auf eine Publikationsverzerrung interpretiert werden. Zum Beispiel kann eine Studieneinreichung eine oder mehrere Hypothesen enthalten, die dann in einer Publikation berichtet werden (keine Publikationsverzerrung). Es ist aber auch möglich, dass nicht alle Hypothesen aus einer Studieneinreichung in einer Publikation berücksichtigt werden (Publikationsverzerrung auf der Hypothesenebene) oder dass überhaupt keine Ergebnisse zu einer bestimmten Studie veröffentlicht werden (Disseminationsverzerrung auf der Publikationsebene). Zweitens wollen wir untersuchen, welche Studien- und Autorenmerkmale mit einer höheren Wahrscheinlichkeit mit einer Veröffentlichung verbunden sind.

Dieser Beitrag fokussiert sich hierbei zum einen auf die Gegenüberstellung der panelspezifischen Unterschiede (Verteilung der Disziplinen, Anzahl der Publikationen, Art der Hypothesendarstellung etc.). Zum anderen werden die Art der Veröffentlichung, die Art des Forschungsdesigns sowie die Vorarbeiten und Gestaltung der Studieneinreichung als mögliche Ursachen und Risikofaktoren von Publikationsverzerrungen unter die Lupe genommen.

S-90: Forschungswissen für die Bildungspraxis: Forschungsbasierte Gestaltung von Unterstützungsangeboten für (angehende) Lehrkräfte

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 162

Chair der Sitzung: Pia Sander

Chair der Sitzung: Steven Beyer

Chair der Sitzung: Isabell van Ackeren-Mindl

Forschungswissen für die Bildungspraxis: Forschungsbasierte Gestaltung von Unterstützungsangeboten für (angehende) Lehrkräfte

Chair(s): **Pia Sander** (Universität Duisburg-Essen, Deutschland), **Steven Beyer** (Universität Duisburg-Essen, Deutschland)

Diskutant*in(nen): **Isabell van Ackeren-Mindl** (Universität Duisburg-Essen, Deutschland)

Bildungsforschung zielt u.a. darauf ab, das Bildungssystem zu verbessern, indem sie aktuelle gesellschaftliche Herausforderungen und damit verbundene Bildungsanliegen adressiert. Sie liefert beispielsweise Erkenntnisse und Evidenz darüber, wie Lehrkräfte die Potenziale der Digitalisierung nutzen und Chancengerechtigkeit fördern können (z.B. Hauck-Thum, Heinz & Hoiß, 2023). Dennoch greifen (angehende) Lehrkräfte Forschungswissen selten in ihrem Denken und Handeln auf (Bauer & Koller, 2023). Gründe dafür reichen von individuellen Merkmalen der Lehrkräfte (Greisel et al., 2023) über eine unzureichende praktische Relevanz (Brühwiler & Leutwyler, 2020) bis hin zu strukturellen Rahmenbedingungen in Forschung und Praxis (Holtappels, 2019). Obwohl damit in den letzten Jahren wichtige Erklärungsansätze für den unzureichenden Transfer zwischen Forschung und Praxis aufgezeigt wurden, wurden diese Erkenntnisse bisher kaum genutzt, um Unterstützungsangebote für Lehrkräfte zu gestalten, die die Nutzung von Forschungswissen in der Bildungspraxis fördern (Gorard et al., 2022; Leuders et al., 2023).

Ein wichtige Quelle für Forschungswissen sind Forschungssynthesen. Forschungssynthesen haben im letzten Jahrzehnt in der Bildungsforschung deutlich an Popularität zugenommen, so dass mittlerweile Synthesen zu vielfältigen Themen mit Relevanz für die Schulpraxis vorliegen (z.B. Bai, Hew, & Huang, 2020; Graham, Kihara, & MacKay, 2020; Leber et al., 2023). Sie fassen aktuelle Forschung systematisch und pointiert zusammen und können zu fundierten und evidenzorientierten Entscheidungen von Lehrkräften beitragen (Grant & Booth, 2009). Zur Nutzung von Forschungswissen aus bildungswissenschaftlichen Synthesen liegen bisher jedoch kaum Erkenntnisse vor (Gorard et al., 2022). Erste Studien zu Clearing-House Angeboten, d.h. Online-Portalen mit zielgruppengerechter Aufbereitung von Forschungssynthesen, deuten darauf hin, dass diese von Lehrerbildenden als nützlich wahrgenommen werden (Diery et al., 2020).

Dieses Symposium widmet sich den Fragestellungen, ob und wie (angehende) Lehrkräfte Forschungswissen für ihre pädagogische Praxis nutzen, um basierend auf diesen Erkenntnissen Gestaltungsmerkmale für forschungssynthesenbasierte Transferangebote zu identifizieren. Diese sollen das Auffinden, Rezipieren und Nutzen von Forschungswissen verbessern. Die Beiträge reflektieren dabei die Annahme, dass ein Transferangebot insbesondere dann erfolgreich sein kann, wenn es gemeinsam und ko-konstruktiv von Forschenden und Praktiker:innen gestaltet wird (Brühwiler & Leutwyler, 2020; Kerres, Sander, Waffner, 2022).

Der erste Beitrag stellt eine Bedarfsanalyse von Lehrkräften und Lehrkräftebildenden vor. Die Untersuchung beleuchtet, welche Informationsquellen die Zielgruppen nutzen (und nicht nutzen), um Forschungswissen für ihre pädagogische Praxis zu finden und einzusetzen, und welche Formate zur Aufbereitung von Forschungswissen bevorzugt werden. Ziel ist es, die Inhalte eines neuen Clearing-House-Angebots zielgruppengerecht und ko-konstruktiv zu gestalten. Der zweite Beitrag befasst sich ebenfalls mit der Gestaltung von digitalen Transferangeboten, wie etwa Clearing-House-Webseiten. In einer mixed-method Studie wird untersucht, ob die Nutzungsintentionen von Grundschullehrkräften für Forschungswissen aus Clearinghouse-Artikeln von deren themenabhängigen Einstellungen und Selbstwirksamkeitserwartungen beeinflusst werden. Der dritte Beitrag untersucht mit einer qualitativ-strukturierenden Inhaltsanalyse, wie Lehramtsstudierende eine digitalisierungsbezogene Forschungssynthese bezüglich ihrer Nutzungsrelevanz bewerten, welche Nutzungspraktiken sie im Kontext einer Anwendungsaufgabe zeigen und welche Gestaltungsmerkmale von Unterstützungsangeboten sich hieraus für sie ergeben und als förderlich wahrgenommen werden.

Zusammenfassend zielen die Beiträge darauf ab, die vielfältigen Aspekte der Unterstützung von Lehrkräften bei der evidenzorientierten Praxis zu beleuchten. Durch die Einbindung von (angehenden) Lehrkräften werden wichtige Impulse für die Gestaltung effektiver Unterstützungs- und Transferangebote gegeben. Das Symposium schließt mit einer Diskussion ab, die den Blick weitet und kooperative Ansätze von Transfer diskutiert, die zum Ziel haben, die Brücke zwischen Bildungsforschung und schulischer Praxis zu stärken.

Beiträge des Symposiums

Sources of information used by teachers and educators to prepare and deliver instruction: choice, use and suitability for evidence-based practice

Martí Quixal¹, Franziska Tschönnens¹, Salome Wagner², Andreas Lachner³

¹Tübingen Center for Digital Education, Universität Tübingen, ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ³Institut für Erziehungswissenschaft & Tübingen Center for Digital Education, Universität Tübingen

Evidence-based education is the best strategy to improve school quality and learners' success (Bauer et al., 2015; Davies, 1999; Leuders et al., 2023). Evidence-based education requires the transfer of research knowledge, i.e., making (pre-service) teachers and teacher educators (henceforth educators) acquainted with the knowledge gained from educational science to improve their teaching. However, teachers rarely rely on empirical evidence to improve their pedagogical strategies (Greisel et al., 2023; Hetmanek et al., 2015). Additionally, little is known about the resources they use to prepare and deliver instruction (Buntins et al., 2024).

We conducted a study to find out if and how often (pre-service) teachers and educators use resources compatible with evidence-based practice.

Our research questions were:

1. What sources of information do (pre-service) teachers and educators use to prepare and deliver instruction?
2. What kind of evidence can be linked to these sources of information?

3. Why do (pre-service) teachers and educators use (or not use) certain sources?

Based on the Theory of Planned Behavior (Ajzen, 1991) –excluding the behavioral intention component– we assume the frequency of use of different sources of information can be predicted by the attitudes towards the epistemological nature of these resources, the purpose of use and the perceived ease of use. Understanding the choices made by teachers and educators will allow for the design of target group-specific transfer products.

Using an online questionnaire, we investigate how target audiences select and use their information sources. To apprehend the frequency and reasons of use of such information sources, we developed a scale based on previously existing inventories (Diery et al., 2020; Jette et al., 2003; Johnston et al., 2003; Young & Ward, 2001). We used well-established scales based on empirical findings to measure attitudes towards resources (Aarons et al., 2012; Diery et al., 2020; Rosman & Kerwer, 2022), the purpose of use (Diery et al., 2020; Greisel et al., 2023; Lysenko et al., 2014) and the perceived ease of use (Diery et al., 2020).

Recruitment was conducted through university and professional distribution lists. The study was approved by the university's ethics committee and the corresponding government administration. N = 616 participants took part: n = 262 pre-service teachers, n = 137 in-service teachers, n = 72 educators, n = 137 participants active as in-service teachers and educators, and n = 8 did not reveal their roles (collection goes on until July 31, 2024 to reach 300 complete questionnaires).

Following Franke (2021) we grouped sources of information into potentially evidence-based (Group 1), potentially theory-based (Group 2), and potentially based on everyday experiences (Group 3). A full analysis pending, preliminary data show that all target groups are unaware of or less likely to use information sources in Group 1 than in Groups 2 and 3. The data also show that information sources in Group 3 are used more frequently than those in Group 2. For example, 48% of pre-service teachers and 58% of teachers and educators reported being unaware of clearinghouses. In contrast, only 17% of pre-service teachers and 6% of teachers and educators reported being unaware of public portals offering professional development materials (though 63% of pre-service teachers and 47% of teachers and educators do not use them). Additionally, 95% of pre-service teachers and 92% of teachers and educators reported relying on colleagues as a source of information.

A multigroup path analysis to elucidate the influence of the independent variables and the differences between target groups is underway. Results and analysis will be complete in time for the conference. Implications for a target-group centered design of a clearinghouse –e.g., adapting transfer products and reach-out strategies– will be discussed.

Wann nutzen Grundschullehrkräfte digitale Transferangebote zu Leistungsheterogenität – themenspezifische Bedarfe, Einstellungen und Selbstwirksamkeit

Anna Maria Leitz, Hannah Kleen, Mareike Kunter

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Wissenstransfer erfolgt dann, wenn ein Bedarf an Informationen zum Umgang mit Herausforderungen und Problemen entsteht (Krause & Möller, 2008; Case & Given, 2016). Trotz evidenzfordernder Standards (European Commission, 2007) präferieren Lehrkräfte statt Forschungswissen eher Anekdoten und Erfahrungswissen (Bråten & Ferguson, 2015; Kiemer & Kollar, 2021). Begrenzte Ressourcen, Beschaffbarkeit und Verständlichkeit wissenschaftlicher Literatur sind einige Gründe dafür (u.a. Williams & Coles, 2007; Bromme et al., 2016). Um den Wissenschaft-Praxis-Gap zu überbrücken (Farley-Ripple et al., 2018), stellen Clearinghouses (Informationswebseiten) Forschungssynthesen niedrigschwellig, komprimiert und zielgruppenangepasst bereit (Seidel et al., 2017; Schneider et al., 2021).

Wichtig erscheint auch, dass nicht nur das Format der Informationen an Bedarfe der Lehrkraft angepasst wird, sondern auch die Inhalte (Farley-Ripple et al., 2018). Das vorgestellte Projekt widmet sich der Lernförderung im Kontext von Heterogenität – ein Bereich, zu dem Lehrkräfte vielfach Fortbildungsbedarf äußern (Weber et al., 2024). Es mangle an Material zu Leistungsheterogenität und insbesondere Grundschullehrkräfte empfinden Belastungen durch Leistungsheterogenität (Eduversum & Universitätsmedizin Mainz, 2024).

Neben einer als vertrauenswürdig und nützlich wahrgenommenen Quelle (Schoor & Schütz, 2021) legen Verhaltens- und Kompetenzmodelle sowie Befunde zur Technologienutzung von Lehrkräften nahe, dass Einstellungen und Selbstwirksamkeit bezüglich eines Mediums das darauf bezogene (Nutzungs-)Verhalten beeinflussen (Yildiz Durak, 2021; Ajzen & Fishbein, 2005; Kunter et al., 2011). Welche Rolle spielen nun Überzeugungen zum inhaltlichen Themenbereich eines Mediums? Bei Inklusionsthemen zeigten sich Einstellungen und Selbstwirksamkeitserwartungen, spezifisch zu Inklusion, als verhaltensrelevant (Opoku et al., 2021; Mentel et al., 2024). Die Nutzung digitaler Clearinghouse-Informationen geht vermutlich über die Abhängigkeit vom Medium oder der Wissenschaft als Quelle hinaus: Wer sich den Umgang mit dem Themenbereich Leistungsheterogenität nicht zutraut oder ablehnende Einstellungen dazu hat, habe weniger Intentionen, sich darüber zu informieren. Gleichzeitig ist der jeweilige Kontext der Lehrperson zu beachten.

Fragestellung

Die Studie untersucht Informationsbedarfe von Lehrkräften bezüglich der Lernförderung im Kontext von Leistungsheterogenität und prüft, welche Faktoren Unterschiede der Nutzungsbereitschaft zweier digitaler Informationsangebote erklären.

1. Zu welchen Inhalten und Herausforderungen beim Umgang mit dem unteren Leistungsbereich haben Grundschullehrkräfte Informationsbedarf?

2. Inwiefern beeinflussen Überzeugungen zu wissenschaftlicher Evidenz, leistungsheterogenitätsbezogene Einstellungen sowie Selbstwirksamkeitserwartungen und Kontextvariablen die Nutzungsintentionen digitaler Informationsangebote?

Es ergeben sich drei Hypothesen: Positivere Einstellungen zu Leistungsheterogenität (H1), höhere Selbstwirksamkeitserwartungen im Umgang mit dem unteren Leistungsbereich (H2) und eine höhere Nützlichkeitswahrnehmung wissenschaftlicher Quellen (H3) führen zu einer erhöhten Nutzungsintention themenspezifischer a) Clearinghouse-Webseiten b) Newsletter-Abonnements.

Methode

Wir erheben qualitative sowie quantitative Online-Fragebogendaten in einer Studie mit Grundschullehrkräften, die präregistriert ist und für die eine Poweranalyse eine notwendige Stichprobengröße von N = 130 ergab. Sie wird über den Verteiler eines Berufsverbands rekrutiert.

Ein offenes Antwortformat erfragt inhaltliche Informationsbedarfe. Selbstwirksamkeit messen wir durch eine Skala von Hachfeld et al. (2012; 4 Items), adaptiert für den Umgang mit dem unteren Leistungsbereich. Fünf Items von Dignath et al. (2020) erfassen themenspezifische Einstellungen, 14 Items von Kiemer und Kollar (2021) die Nützlichkeit wissenschaftlicher vs. erfahrungsbasierter Quellen. Nutzungsintentionen der Clearinghouse-Webseite erfassen wir über eine selbstberichtete

Wahrscheinlichkeit der Nutzung. Um mehr als eine hypothetische Verhaltensintention zu testen, können Lehrkräfte am Fragebogenende freiwillig – als tatsächliches Verhalten – einen themenspezifischen Newsletter abonnieren. Als Kontrollvariablen werden Kontextbedingungen, u.a. Alter, Berufserfahrung und der Anteil leistungsschwacher Kinder, mitbedacht.

Fragestellung 1 beantworten wir durch die zusammenfassende Inhaltsanalyse (Mayring, 2022), die Kategorien identifiziert, deren Häufigkeiten wir interpretieren. Die Hypothesen zu Fragestellung 2 werden anhand multipler linearer Regressionen geprüft.

Ergebnisse/Bedeutung

Da die Datenerhebung noch läuft, berichten wir die Ergebnisse zum Konferenzzeitpunkt und diskutieren sie als Erklärungspotentiale des Wissenschaft-Praxis-Gaps. Sie erweitern das Verständnis darüber, welche Informationen sich Lehrkräfte wünschen und welche lehrkraftseitigen Voraussetzungen einer themenspezifischen Evidenznutzung vorausgehen.

So können wir unser wissenschaftsseitiges, digitales Clearinghouse-Informationsangebot für Lehrkräfte noch näher an den identifizierten Bedarfen ausrichten und damit auf die dringenden Herausforderungen von Lehrkräften im Kontext unterer Leistungsbereich reagieren.

Unterstützung der Rezeption und Nutzung von Forschungswissen aus digitalisierungsbezogenen Forschungssynthesen

Pia Sander, Steven Beyer, Tanja Mayer
Universität Duisburg-Essen

Theoretischer Hintergrund

Evidenzorientierte Bildungspraxis

Evidenzorientierte Bildungspraxis ist relevant, da evidenzbasierte Entscheidungen die Schulpraxis und Schülerleistungen verbessern können (z.B. Bauer & Kollar, 2023). Evidenz bietet Lehrkräften wertvolle Hinweise zur Unterrichtsgestaltung, jedoch gibt es Herausforderungen bei der Rezeption und Nutzung, die zu seltener Implementation führen (Gorard et al., 2022; Thomm et al., 2021). Die professionelle Entwicklung und Unterstützung der (angehenden) Lehrkräfte ist daher unerlässlich, obwohl wenig Einigkeit darüber besteht, welche Kompetenzfacetten wie gefördert werden sollten (Bauer & Kollar, 2023; Voss et al., 2020). Übergreifendes Ziel ist der Auf-/Ausbau eines Verständnisses von Evidenz als Ressource zur Reflexion und Begründung von Praktiken (Bauer & Kollar, 2023). Der Beitrag unterscheidet hinsichtlich der Forschungsnutzung in Anlehnung an Lysenko et al. (2016) zwischen Instrumental Use (transmissive Praxisanwendung), Conceptual Use (individueller Erkenntnisgewinn) und Symbolic Use (Handlungsbegründung).

Unterstützungsansätze

Es besteht ein Defizit an belastbarer Evidenz bezüglich der Gestaltung von Unterstützungsansätzen zur Evidenznutzung in der Bildungspraxis (Gorard et al., 2022; Leuders et al. 2023). Gorard et al. (2022) zeigen auf, dass sowohl die passive Verbreitung von Forschungsergebnissen als auch viele Interventionen weitgehend ineffektiv hinsichtlich der Einstellungs- und Verhaltensänderung sind. Zudem bleibt unklar, ob die bloße Darstellungsvereinfachung von Forschungsergebnissen zu besseren Ergebnissen führt. Die bestehenden Forschungsergebnisse deuten darauf hin, dass zusätzliche Maßnahmen über die Bereitstellung und Vereinfachung hinaus erforderlich sind, um die tatsächliche Rezeption und Nutzung von Evidenz in der Praxis zu fördern (Gorard et al., 2022).

Fragestellung

Der Beitrag untersucht die Unterstützung von Lehramtsstudierenden und die Verbesserung der Forschungskommunikation durch Unterstützungsansätze unter besonderer Berücksichtigung der Potentiale digitalisierungsbezogener Forschungssynthesen bezüglich evidenzorientierter Bildungspraxis (Rycroft-Smith & Stylianides, 2022). Aufgrund des umfassenden Evidenzdefizits stehen folgende Forschungsfragen im Zentrum:

1. Wie schätzen Lehramtsstudierende die Relevanz einer exemplarischen digitalisierungsbezogenen Forschungssynthese hinsichtlich des Instrumental Use, Conceptual Use und Symbolic Use ein?
2. Welche Praktiken der Nutzung von Forschungswissen aus einer exemplarischen Forschungssynthesen zeigen Lehramtsstudierende?
3. Welche Merkmale eines Unterstützungsangebotes fördern die Rezeption und Nutzung von digitalisierungsbezogenem Forschungswissen?

Methode

Lehrveranstaltungsbezogen lasen Masterstudierende im Grundschullehramt (n = 16) eine digitalisierungsbezogene Forschungssynthese und bearbeiteten hierzu schriftlich offene Items und Aufgaben. Die offenen Items zur Nutzungsrelevanz (FF1) wurden adaptiert von Afdal & Spernes (2018). Die Anwendungsaufgabe (FF2) bestand darin, anhand der Forschungssynthese zu begründen, wie eine analoge Unterrichtsaufgabe von Herget & Scholz (2018) digital weiterentwickelt werden kann. Die offenen Items zur Erfassung förderlicher Merkmale eines Unterstützungsangebotes (FF3) wurden von gestaltungsorientierten Ansätzen adaptiert (z.B. Doorley et al., 2018).

Alle erhobenen Daten wurden mittels qualitativ-strukturierender Inhaltsanalyse sowie Fallvergleichen, z.B. anhand absolvierter digitalisierungsbezogener Lerngelegenheiten, Unterrichtserfahrung, Selbsteinschätzung studentischer Forschungskompetenz nach Böttcher & Thiel (2018), ausgewertet (Schreier, 2012).

Ergebnisse und ihre Bedeutung

Vorläufige Ergebnisse zeigen eine unterschiedliche Bewertung der Relevanz im Zusammenhang mit den drei Nutzungstypen. Für den Conceptual Use sowie Symbolic Use wird der Forschungssynthese eine hohe Relevanz zugesprochen. Die Synthese wird als hilfreiche Quelle zur Orientierung anhand des aktuellen Forschungsstandes gesehen. Herausforderungen wurden im Zusammenhang mit dem Aufwand der Lektüre im Rahmen eines potentiellen Schulalltags berichtet. Darüber hinaus wurden insbesondere Schwierigkeiten beim Verständnis der untersuchten digitalen Medien berichtet. Dieser Befund ist im Kontext der Lehrkräftebildung vor dem Hintergrund der unsystematischen digitalisierungsbezogenen Lerngelegenheiten (siehe Eickelmann & Drossel, 2020) erwartungskonform, weil das mediendidaktische/-pädagogische Vorwissen ggf. nicht ausreichend für das Verständnis ist. In Folge der Lektüreschwierigkeiten wird die Relevanz der Forschungssynthese für den Instrumental Use als gering eingeschätzt, weil die Studierenden keine unmittelbares Nutzungspotential der digitalen Medien im Unterrichtskontext sahen.

Der Vortrag stellt die ausführlichen Ergebnisse der Inhaltsanalyse sowie Fallvergleiche zu weiteren Aspekten der Relevanz (FF1), Nutzungspraktiken im Kontext der Anwendungsaufgabe (FF2) und förderlichen Merkmalen von Unterstützungsansätzen (FF3) vor. Ziel ist es, erste konzeptuelle Annahmen zu Gestaltungsmerkmalen eines Unterstützungsangebotes zu entwickeln, das evidenzorientierte Praxis auf Grundlage von digitalisierungsbezogenen Forschungssynthesen unterstützt und in Folgestudien getestet werden soll.

S-91: ENTFÄLLT: Persönlichkeitsmerkmale und ihre Relevanz auf Beurteilungen im Bildungsbereich

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 242

Chair der Sitzung: Valerie Gitzi

Chair der Sitzung: Vanda Capon-Sieber

Chair der Sitzung: Richard Göllner

ENTFÄLLT: Persönlichkeitsmerkmale und ihre Relevanz auf Beurteilungen im Bildungsbereich

Chair(s): **Valerie V. Gitzi** (Universität Zürich, Schweiz), **Vanda Capon-Sieber** (Universität Zürich, Schweiz)

Diskutant*in(nen): **Richard Göllner** (Eberhard Karls Universität Tübingen, Tübingen)

Theoretischer Hintergrund:

Die Wahrnehmung von Situationen wird massgeblich durch Persönlichkeitsmerkmale beeinflusst (Pawlik, 1973; Roberts, 2018). Persönlichkeitsmerkmale sind stabile, überdauernde Eigenschaften, die individuelle Unterschiede im Verhalten und Erleben erklären, gängig operationalisiert durch das Big-Five-Modell (Hagemann et al., 2022). Dies gilt auch im Kontext der Bildung, wo individuelle Persönlichkeitsmerkmale der Akteur*innen insbesondere relevant für Beurteilungen sind. Während vielfältige Untersuchungen zur Relevanz von Persönlichkeitsmerkmalen für das Erleben und Verhalten von Personen durchgeführt wurden (z.B. Hagemann et al., 2022), gibt es bislang vergleichsweise wenig Studien im Bildungsbereich (z.B. Meyer et al., 2023) und speziell für Beurteilungen. In diesem Symposium wird die Bedeutsamkeit von Persönlichkeitsmerkmalen unterschiedlicher Akteur*innen im Bildungsbereich – Schüler*innen, Lehrpersonen und externe Beurteilende des Unterrichts – für verschiedene Arten von Beurteilungen untersucht.

Beiträge:

Der erste Beitrag widmet sich der Frage, inwiefern die Beurteilung von Unterrichtsqualität mit Persönlichkeitsmerkmalen von Schüler*innen korrelieren. Im Sinne des Person-Environment-Fit-Modell (Chartrand, 1991) wird geprüft, ob Unterrichtsqualitätsdimensionen (*kognitive Aktivierung, konstruktive Unterstützung, effektives Klassenmanagement und Individualisierung*) und Persönlichkeitsmerkmale von Schüler*innen korrespondieren und daraus Zusammenhänge festgestellt werden können. Mittels Mehrebenenregressionsanalysen an Daten von $N = 700$ Schüler*innen können substanzielle Zusammenhänge zwischen Persönlichkeitsmerkmalen und Facetten der Unterrichtsqualitätsdimensionen festgestellt werden.

Im zweiten Beitrag wird die Persönlichkeit von externen Beobachtenden beleuchtet. Dabei wird untersucht, ob Persönlichkeitsmerkmale (*Gewissenhaftigkeit, Verträglichkeit und Offenheit für Neues*) bei externen Beobachtenden einen Einfluss auf die Validität (*Übereinstimmung mit Masterratings*) ihrer Unterrichtsqualitätseinschätzungen haben. Es wird sowohl der Einfluss auf untrainierte als auch trainierte Beobachtende und drittens auf den Lernzuwachs im Training analysiert. Die Ergebnisse zeigen, dass die untersuchten Persönlichkeitsmerkmale der externen Beobachtenden in der vorliegenden Stichprobe keine signifikante Erklärung der Unterschiede in der Validität von Unterrichtsqualitätseinschätzungen bieten.

Die Bedeutung von Persönlichkeitsmerkmalen von Lehrpersonen auf die Beurteilungsakkuratesse von Textanalysen wird im dritten Beitrag adressiert und der Zusammenhang zwischen dem Persönlichkeitsmerkmal *Gewissenhaftigkeit*, weiteren Faktoren, wie *Motivation* sowie *aufgewendeter Zeit von Lehrpersonen*, und deren Einfluss auf die Urteilsakkuratesse von Schüler*innen-Texten untersucht. Die Analysen wurden mit Hilfe eines neuentwickelten hierarchischen Modells (Lohmann et al., under review) durchgeführt, das die simultane Schätzung von *Rang-, Niveau- und Differenzierungskomponenten* in einem Modell ermöglicht. Hier zeigt sich, dass *Gewissenhaftigkeit* allein keine signifikante Begründung der Urteilsakkuratesse bietet, ebenso wenig die *Motivation*. Lediglich bei der *aufgewendeten Zeit* zeigte sich ein signifikanter positiver Zusammenhang.

Diskussion:

Zusammengefasst zeigen die drei Beiträge, dass sich Zusammenhänge zwischen Persönlichkeitsmerkmalen und Beurteilungen im Bildungskontext nur zum Teil identifizieren lassen und entsprechend weitere Merkmale in den Blick genommen werden müssen. Die Implikationen dieses Symposiums werden im Hinblick auf die Schulpraxis und -forschung diskutiert.

Beiträge des Symposiums

ENTFÄLLT: Unterrichtsqualität und Persönlichkeitsmerkmalen von Schüler*innen

Miriam Marleen Gebauer

Otto-Friedrich-Universität Bamberg

Theoretischer Hintergrund:

Persönlichkeitsmerkmale beschreiben relativ stabile Muster von Gedanken, Gefühlen und Verhaltensweisen (Hampson & Goldberg, 2006; Roberts, 2018). Eine der bekanntesten Taxonomien von Persönlichkeitsmerkmalen ist das Big-Five-Inventary, welches die Merkmale Neurotizismus, Extraversion, Offenheit, Gewissenhaftigkeit und Verträglichkeit umfasst (Rammstedt & John, 2005). Forschungsbefunde haben wiederholt die Bedeutung von Persönlichkeitsmerkmalen für schulischen Erfolg gezeigt: So werden Zusammenhänge zwischen Leistungsmerkmalen und Persönlichkeitsmerkmalen berichtet (Meyer et al., 2023), aber auch zwischen leistungsrelevanten Merkmalen wie Beteiligung am Unterricht oder Erledigung der Hausaufgaben (Gray & Watson, 2002).

Die Dimensionen von Unterrichtsqualität (*kognitive Aktivierung, konstruktive Unterstützung, effektives Klassenmanagement, Individualisierung*) sind von wesentlicher Bedeutung für das Lernen und den Lernerfolg im Unterricht (Hattie & Yates, 2013; Praetorius & Gräsel, 2021). Die Bedeutung von Unterrichtsqualität ist in Bezug auf eine Vielzahl an Merkmale untersucht worden und es werden Zusammenhänge zu Leistungsergebnisse von Schüler*innen (Kunter et al., 2013), aber auch zu individuellen Merkmalen wie zum Beispiel der motivationalen Orientierung berichtet (Sakiz et al., 2012).

Persönlichkeitsmerkmale einer Person beeinflussen die Art und Weise, wie diese Person auf Situationen oder Verhaltensweisen anderer reagiert (Roberts, 2018). Im Sinne eines Person-environment fit (Chartrand, 1991) kann ein Zusammenhang zwischen Persönlichkeitsmerkmalen und der Wahrnehmung der Unterrichtsqualität als eine Korrespondenz zwischen diesen Merkmalen interpretiert werden (Pawlowska et al., 2014). Aktuell vorliegende Befunde berichten Zusammenhänge zwischen Neurotizismus und der Teilnahme an Gruppendiskussionen (Chamorro-Premuzic et al., 2007) oder Verträglichkeit und Gewissenhaftigkeit und der wahrgenommenen Zufriedenheit im Unterricht (Pawlowska et al., 2014).

Fragestellung:

Korrespondieren die Dimensionen der Unterrichtsqualität mit Persönlichkeitsmerkmalen von Schüler*innen? Können Zusammenhänge zwischen den Big-Five Persönlichkeitsmerkmalen (Neurotizismus, Extraversion, Offenheit, Gewissenhaftigkeit und Verträglichkeit) und den Dimensionen von Unterrichtsqualität (kognitive Aktivierung, konstruktive Unterstützung, effektives Klassenmanagement und Individualisierung) festgestellt werden?

Methode:

In diese Analysen einbezogen wurden Daten von N = 700 Schüler*innen an Gymnasien, die zum Zeitpunkt der Befragung in der 7. Jahrgangsstufe waren. Das Durchschnittsalter betrug $M = 12.2$ Jahre ($SD = .07$). Zur Beantwortung der Forschungsfragen wurden ein Doubly-manifest Regressionsmodell zur Prüfung der Zusammenhänge geschätzt. Aufgrund der begrenzten Stichprobengröße mussten einzelne Manifest-Indikatoren anstatt latenter Skalen verwendet werden. Auf der Klassenebene wurde eine manifeste Aggregation einzelner Ratings verwendet, um manifeste Indikatoren zu erzeugen. ICCs für alle Unterrichtsqualitätsvariablen waren .10 oder höher, was auf einen erheblichen Unterschied zwischen den Klassen hinweist. Korrelationen zwischen den manifesten Unterrichtsqualitätsvariablen reichten von $r = .35$ bis $r = .95$ mit einem Median von $med = .71$.

Ergebnisse und ihre Bedeutung:

Die Analysen der Mehrebenenregression zeigen statistisch signifikante Zusammenhänge zwischen den Dimensionen der Unterrichtsqualität und den Persönlichkeitsmerkmalen der Schüler*innen. Auf within-Ebene können statistisch signifikante Zusammenhänge zwischen Facetten des effektiven Klassenmanagements und den Persönlichkeitsmerkmalen Extraversion, Verträglichkeit und Gewissenhaftigkeit festgestellt werden sowie zwischen Individualisierung und Neurotizismus und Offenheit und kognitiver Aktivierung. Auf der between-Ebene sind Facetten der kognitiven Aktivierung bedeutsam für Extraversion, Gewissenhaftigkeit und Offenheit. Facetten des effektiven Klassenmanagements zeigen statistisch signifikante Zusammenhänge zu den Persönlichkeitseigenschaften Extraversion, Verträglichkeit, Offenheit und Neurotizismus.

Es zeigt sich insgesamt, dass Schüler*innen mit unterschiedlichen Ausprägungen von Persönlichkeitseigenschaften offenbar auf unterschiedliche Art und Weise auf das Unterrichtsangebot reagieren. So scheinen beispielsweise eher ängstliche Schüler*innen weniger von herausfordernden Aufgaben und mehr von Strukturen zum Umgang mit Unterrichtsstörung zu profitieren. Eher extrovertierte Schüler*innen hingegen scheinen sich von kognitiven Aktivierungen angesprochen zu fühlen. Die Ergebnisse werden vor dem Hintergrund ihrer Bedeutung für zukünftige Forschungen und schulpraktischen Implikationen diskutiert.

ENTFÄLLT: Einfluss von Persönlichkeitsmerkmalen externer Beobachtender auf die Validität von Unterrichtsqualitätseinschätzungen und auf den Lernzuwachs in Trainings

Valerie V. Gitzl¹, Vanda Capon-Sieber¹, Thilo Kleickmann², Johannes Hartig³, Mirjam Steffensky⁴, Aiso Heinze⁵, Olga K. Lichtner², Anna-Katharina Praetorius¹

¹Universität Zürich, Schweiz, ²Christian-Albrechts-Universität zu Kiel, Deutschland, ³Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland, ⁴Universität Hamburg, Deutschland, ⁵Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, Deutschland

Theoretischer Hintergrund:

Zur Unterrichtsqualitätseinschätzung werden häufig externe Beobachtende eingesetzt, die für valide Bewertungen trainiert werden müssen (Helmke, 2017). Diese Trainings sollen ein gemeinsames Verständnis theoretischer Grundlagen (Seidel et al., 2005) und die Notwendigkeit des genauen Befolgens von Anweisungen vermitteln, auch als Adhärenz bezeichnet. Individuelle Unterschiede in der Adhärenz im medizinischen Kontext konnten durch Persönlichkeitsmerkmale wie Gewissenhaftigkeit, Verträglichkeit oder Offenheit für Neues erklärt werden (siehe Meta-Analyse von Molloy et al., 2013). Für den Bildungskontext vermuteten Cash et al. (2012), dass diese Persönlichkeitsmerkmale auch Unterrichtsqualitätseinschätzungen und den Lernzuwachs in Trainings beeinflussen können. Unsere Studie untersucht diese Vermutung empirisch.

Fragestellung:

1. Wie hängen Persönlichkeitsmerkmale untrainierter Beobachtender mit der Validität ihrer Unterrichtsqualitätseinschätzungen zusammen?
2. Wie hängen Persönlichkeitsmerkmale trainierter Beobachtender mit der Validität von Unterrichtsqualitätseinschätzungen zusammen?
3. Unterscheiden sich Beobachtende mit unterschiedlichen Persönlichkeitsmerkmalen darin, inwieweit sie von kürzeren oder längeren Trainings profitieren?

Hypothesen:

Basierend auf der Meta-Analyse von Harari et al. (2015) erwarteten wir, dass Gewissenhaftigkeit und Offenheit bei untrainierten Beobachtenden positiv mit der Validität ihrer Unterrichtseinschätzungen assoziiert sind, während Verträglichkeit negativ korreliert (FF1). In Bezug auf trainierte Beobachtende nehmen wir an, dass alle drei Merkmale positiv mit der Validität verbunden sind (FF2) (Müller & Pietsch, 2011). Explorative Analysen wurden im Hinblick auf den potenziellen Einfluss der Persönlichkeitsmerkmale auf den Lernzuwachs in unterschiedlich langen Trainings zur Unterrichtsqualitätseinschätzung durchgeführt (FF3).

Methode:

Wir analysierten die Daten eines Prä-Post-Experiments, das den Einfluss der Halbierung der Trainingsdauer auf die Validität von Unterrichtseinschätzungen untersuchte. Da die Zuteilung zur Trainingsdauer randomisiert durchgeführt wurde und die Kürzung der Inhalte lediglich den Detailgrad der Erarbeitung betraf, konnten die Daten für die vorliegende Fragestellung uneingeschränkt verwendet werden.

Insgesamt nahmen N = 75 Personen (91% weiblich) mit Hintergrund in Psychologie oder Pädagogik am Ursprungsexperiment teil. Etwa zwei Drittel waren Studierende. Das Alter lag zwischen 20 und 53 Jahren ($M = 26.3 \pm 6.7$). Vor dem Training wurden drei der fünf Big-Five Persönlichkeitsmerkmale mittels Big-Five-Inventory (BFI-2; Danner et al., 2016) erfasst. Die Teilnehmenden hatten vor und nach dem Training zwei Wochen Zeit für die Unterrichtsqualitätseinschätzungen von unter anderem fünf videographierten Naturwissenschaftslektionen anhand von vier Unterrichtsqualitätsdimensionen (Klassenführung, konstruktive Unterstützung, kognitive Aktivierung sowie Auswahl und Thematisierung des Inhalts). Die Teilnehmenden erhielten eine Aufwandsentschädigung.

Unsere Hypothesen testeten wir mittels linearer Regressionen. Als abhängige Variable diente jeweils die Validität (Übereinstimmung mit Masterratings) einer der vier Dimensionen von Unterrichtsqualität. Für FF1 wurden jene Abweichungen vom

Masterrating vor dem Training verwendet, für FF2 wurden jene Abweichungen vom Masterrating nach dem Training verwendet und für FF3 die Differenz dieser Abweichungen vor und nach dem Training, wobei für FF2 und FF3 ein Haupteffekt und eine Interaktion mit der Trainingsbedingung hinzugenommen wurde. Die Persönlichkeitsmerkmale dienten jeweils als unabhängige Variablen.

Ergebnisse:

FF1 und FF2: Entgegen unseren Annahmen gab es keine statistisch signifikante Assoziation zwischen den Persönlichkeitsmerkmalen und der Validität der Unterrichtsqualitätseinschätzungen bei untrainierten und trainierten Beobachtenden.

FF3: Es liessen sich einzelne Interaktionseffekte zwischen den Persönlichkeitsmerkmalen und der Trainingsbedingung auf den Zuwachs an Validität (Lernzuwachs) feststellen. Personen im langen Training mit hohen Verträglichkeits- und hohen Offenheitswerten zeigten den grössten Lernzuwachs in der Dimension kognitive Aktivierung ($b = -.91$, $p = .01$) bzw. Personen im kurzen Training mit hohen Verträglichkeitswerten zeigten den kleinsten Lernzuwachs in der Dimension konstruktive Unterstützung ($b = .72$, $p = .03$). Diese signifikanten Werte verschwanden jedoch nach der Alpha-Fehler-Korrektur für multiples Testen.

Diskussion:

Unsere Ergebnisse zeigen, dass die Ausprägung der Persönlichkeitsmerkmale der Beobachtenden keine ausreichende Erklärung für Unterschiede in der Validität von Unterrichtseinschätzungen bieten können und unterstreichen somit die Notwendigkeit, andere Beobachtenden Merkmale wie bspw. die fachliche Expertise bzgl. eines möglichen Einflusses auf die Validität der Unterrichtsqualitätseinschätzungen zu untersuchen.

ENTFÄLLT: Wie hängen Beurteilungsdauer, Gewissenhaftigkeit und Motivation von Lehrkräften mit der Beurteilungsakkuratheit zusammen?

Julian F. Lohmann¹, Flavio Lötscher², Stefan Keller², Jens Möller¹

¹Christian-Albrechts-Universität zu Kiel, ²Pädagogische Hochschule Zürich, Schweiz

Theoretischer Hintergrund:

Akkurate Leistungseinschätzungen von Lehrkräften gelten als integraler Bestandteil von gutem Unterricht (DeLuca et al., 2016; Elliot et al., 2001) und sind positiv mit wichtigen Bildungsergebnissen assoziiert (z.B. Kolovou et al., 2024; Trautwein et al., 2006). Schriftsprachliche Leistungen wie Aufsätze, die im sekundären wie tertiären Bildungsbereich essentielle Bestandteile von Prüfungsformaten sind, sind in den letzten Jahren verstärkt in den Fokus der Forschung zu Lehrkräfturteilen gerückt (z.B. Möller et al., 2022; Jansen et al., 2021, Vögelin et al., 2021).

Ein zentraler Aspekt von theoretischen Modellen zu Lehrkräfturteilen sind Moderatoren der Urteilsakkuratheit (Loibl et al., 2020; Südkamp et al., 2012). Hierbei lassen sich verschiedene Typen von Moderatoren unterscheiden, wie beispielsweise Testcharakteristika, Schüler*innen- oder Lehrkräfteigenschaften. Obwohl viele Studien bedeutsame Unterschiede in der Urteilsakkuratheit zwischen Lehrkräften nahelegen, gibt es bisher noch immer wenige Studien zu Lehrkraftvariablen als Moderatoren der Urteilsakkuratheit (z.B. Urhahne & Winija, 2021). Die meisten existierenden empirischen Studien zu Lehrkraftvariablen haben sich vor allem mit der Bedeutung der (Lehr-)Erfahrung für die Beurteilungsakkuratheit befasst (e.g., Fleckenstein et al., 2018; Jabůrek et al., 2022; McElvany et al., 2009; Möller et al., 2022). Weitestgehend unerforscht sind hingegen Persönlichkeitseigenschaften, motivationale Merkmale, oder Beurteilungsstrategien von Lehrkräften im Zusammenhang mit Urteilsakkuratheit.

Basierend auf der Arbeit von Schrader und Helmke (1989) werden in der Forschung zur Urteilsakkuratheit die Rang-, Niveau- und Differenzierungskomponente unterschieden (Urhahne & Winija, 2021). Die Rangkomponente ist als Korrelation zwischen den Lehrkräfturteilen und der tatsächlichen Leistung operationalisiert. Die Niveauelemente reflektiert, ob eine Lehrkraft im Mittel zu Über- oder Unterschätzung tendiert. Die Differenzierungskomponente drückt aus, wie gut die Lehrkräfturteile die Leistungsvarianz der bewerteten Schüler*innenleistungen repräsentieren.

Fragestellung:

1. Wie hängt die Gewissenhaftigkeit von Lehrkräften mit verschiedenen Komponenten von Urteilsakkuratheit zusammen?
2. Wie hängen die Motivation und das Selbstkonzept in Bezug auf Bewertungsaufgaben mit verschiedenen Komponenten der Urteilsakkuratheit zusammen?
3. Wie hängt die von Lehrkräften aufgewendete Zeit pro Text mit verschiedenen Komponenten der Urteilsakkuratheit zusammen?

Methode:

Zur Beantwortung der Forschungsfrage wurden Daten von $N = 325$ Lehramtsstudierenden des Fachs Englisch aus Deutschland und der Schweiz ausgewertet. Alle Teilnehmenden bewerteten fünf zufällig ausgewählte Texte aus einem grossen Korpus von authentischen Schüler*innentexten auf analytischen und holistischen Skalen. Diese Texte wurden zuvor in einem aufwendigen Ratingverfahren von geschulten Expert*innen der International Association for the Evaluation of Educational Achievement (IEA) eingeschätzt (Keller et al., 2024) und mit einem IRT-basierten Rater*innenmodell um systematische Rater*inneneffekte bereinigt (Keller et al., 2024; und z.B. Eckes, 2023; Wu, 2017). Insgesamt lagen der Analyse über 6000 Einzelurteile zugrunde.

In der vorliegenden Arbeit verwenden wir ein neues hierarchisches Modell (Lohmann et al., under review), das es erlaubt, die drei Komponenten der Urteilsakkuratheit simultan und unter Berücksichtigung multipler Moderatoren zu untersuchen.

Ergebnisse und ihre Bedeutung:

In Bezug auf Forschungsfrage eins und zwei zeigten sich keine signifikanten Effekte. Weder die Rang-, noch die Level- oder Differenzierungskomponente wurden durch die Gewissenhaftigkeit, die Motivation oder das Selbstkonzept der Lehrkräfte vorhergesagt. Die Überprüfung von Forschungsfrage drei zeigte zwei signifikante Moderationseffekte des Beurteilungsverhaltens: Je mehr Zeit sich die Lehrkräfte für die Bewertung der Texte nahmen, desto höher waren ihre Rang- und Differenzierungskomponente.

In Anlehnung an Forschung zur Verwendung von Hinweisreizen (Cues) bei der Beurteilung von Schüler*innenleistungen (e.g., Loibl et al., 2020), kann vermutet werden, dass Lehrkräfte, die sich mehr Zeit für die Bewertung nahmen, mehr leistungsrelevante Aspekte der Schüler*innentexte in ihr Urteil einfliessen liessen. Durch diese elaboriertere Verarbeitung der relevanten Textqualitätsaspekte erreichten diese Personen eine höhere Beurteilungsakkuratheit hinsichtlich der Rangkomponente. Die höhere Differenzierungskomponente könnte in ähnlicher Weise interpretiert werden: Personen, die sich mehr Zeit für die

Bewertung der Texte nahmen, stellten auch mehr (qualitative) Unterschiede zwischen den Texten fest. Dies könnte sich in einer höheren wahrgenommenen Varianz widerspiegeln, die die Leistungsvarianz allerdings etwas überschätze.

S-92: Anstrengung als Schlüssel zur Bewältigung von Lernanforderungen

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 382

Chair der Sitzung: Laura Kehle

Chair der Sitzung: Detlef Urhahne

Chair der Sitzung: Diana Raufelder

Anstrengung als Schlüssel zur Bewältigung von Lernanforderungen

Chair(s): **Laura Kehle** (Universität Passau, Deutschland), **Detlef Urhahne** (Universität Passau)

Diskutant*in(nen): **Diana Raufelder** (Universität Greifswald)

Als Hull (1943) das Gesetz von der geringsten Anstrengung formulierte, nach der von zwei Möglichkeiten, die mit dem geringsten Aufwand zur Zielerreichung gewählt wird, hatte er im Sinn, dass Organismen dazu tendieren, Energie zu sparen, um ihr Überleben zu sichern. Da Anstrengung – ganz gleich, ob physischer oder mentaler Natur – von unangenehmen, sich steigenden Gefühlserscheinungen bis hin zur völligen Erschöpfung begleitet sein kann, galt es unnötige Anstrengung unbedingt zu vermeiden.

Heutzutage wissen wir, dass die intentionale Mobilisierung von Anstrengung jedoch sehr positive Auswirkungen auf das Lernen und den Wissenserwerb besitzt. Das langt von wünschenswerten Erschwernissen bei Lernaufgaben (Schalk & Ziegler, 2023), die zusätzliche Anstrengung erfordern, über günstige Folgen von Kausalattributionen auf Anstrengung bei Erfolg wie Misserfolg (Badewitz et al., 2021) bis hin zum Erwerb einer gelernten Betriebsamkeit (Eisenberger, 1992), so dass bereits der mit Anstrengung bewältigte Lernweg und nicht nur das reine Ergebnis einen belohnenden Charakter erhält.

Das Symposium setzt sich aus drei Beiträgen zusammen, die alle die Funktion von Anstrengung als Schlüssel zur Bewältigung von Lernanforderungen aufzeigen. Im ersten Beitrag werden kognitions- und motivationspsychologische Erklärungen für das Lernen mit Anstrengung zusammengeführt. Basierend auf der Theorie der kognitiven Belastung (Sweller et al., 2019), der Selbstbestimmungstheorie (Ryan & Deci, 2020) und der situierten Erwartungs-Wert-Theorie (Eccles & Wigfield, 2020) werden drei Arten von Anstrengung unterschieden. Anstrengung als Komplexität, um die Verarbeitungsanforderungen auf Objektebene zu bewältigen. Anstrengung durch Bedürfnisversagung, um mit den negativ bewerteten affektiv-emotionalen Verarbeitungsanforderungen des Lernens richtig umzugehen, sowie Anstrengung als Zuteilung von Ressourcen zur Bewältigung eines Lernproblems.

Im zweiten Beitrag werden diese drei Arten von Anstrengung empirisch messbar gemacht und ihr Zusammenhang mit der Leseleistung von Schülerinnen und Schülern der siebten und achten Jahrgangsstufe untersucht. Explorative Faktorenanalysen ergaben für die drei Arten der Anstrengung eine dreifaktorielle Lösung, was auf die Trennbarkeit der hypothetischen Konstrukte hindeutet. Als signifikanter Prädiktor der Lesegenauigkeit erwies sich die Anstrengung als Zuteilung von Ressourcen, die es erlaubt, die Intensität des Lernprozesses selektiv zu verstärken.

Im dritten Beitrag wird diese Anstrengung als Zuteilung von Ressourcen als zusätzliche Lernanstrengung gedeutet, um Fehler beim Mathematiklernen in der Grundschule selbstgesteuert zu berichtigen. Grundschülerinnen und -schüler arbeiteten mit Tablets und hatten während des Lernens aufgrund von Feedback zu falsch gelösten Aufgaben die Möglichkeit, ihre Lernleistung zu verbessern. Anstrengung wurde als Verhaltensmaß sowie als Selbstbericht im Lernprozess und zum Abschluss der Lerneinheit ermittelt. Alle drei Indikatoren der adaptiven Lernanstrengung nach Fehlern sind in der Lage, den Wissenserwerb der Grundschülerinnen und -schüler zu erklären.

Beiträge des Symposiums

Was macht das Lernen anstrengend? Erklärungen aus kognitions- und motivationspsychologischer Perspektive

Matthias Nückles¹, Axel Grund², Stefan Fries³, Alexander Renkl¹, Julian Roelle⁴

¹Universität Freiburg, ²Universität Luxemburg, ³Universität Bielefeld, ⁴Universität Bochum

Anstrengung wird in der empirischen Bildungsforschung als Konstrukt mit vielen Facetten dargestellt. Sich anzustrengen ist wünschenswert, aber auch unangenehm. Anstrengung kann lange Zeit auf gleichbleibendem Niveau ausgeübt werden, sich aber auch schnell verändern, wenn die Aufgabenstellung es erforderlich macht. Anstrengung kann durch Ziele ebenso angeregt werden wie durch die Menge an zu verarbeitenden Informationen. Anstrengung kann erschöpfend, aber auch sehr befriedigend wirken, wenn man beim Lernen sein Bestes gegeben hat.

In diesem Beitrag wollen wir eine Brücke zwischen einer kognitions- und einer motivationspsychologischen Perspektive auf Anstrengung schlagen, um zu einer vollständigeren Betrachtung und Erklärung der Ursachen von Anstrengung zu kommen. In der kognitionspsychologischen Forschung wird mentale Anstrengung zumeist mit der Aufteilung und Verwendung kognitiver Ressourcen zur Bewältigung von Lernaufgaben in Verbindung gebracht (Sweller, van Merriënboer & Paas, 2019). In der motivationspsychologischen Forschung dient Anstrengung typischerweise der Intensivierung und Aufrechterhaltung von Lernprozessen bis zur Zielerreichung (Vollmeyer & Rheinberg, 2000). Wir plädieren in Verknüpfung dieser Sichtweisen für eine dreiteilige Konzeption von Anstrengung im Kontext des Lernens in formalen Settings wie Schule und Hochschule und machen in diesem Zuge verschiedene Quellen für das Erleben und Investieren von Anstrengung verantwortlich (vgl. Grund et al., 2024).

In einer ersten Bedeutung reflektiert Anstrengung die Schwierigkeit oder Komplexität einer Lernaufgabe (Anstrengung durch Komplexität). Hier geht es um die Verarbeitungsanforderungen auf Objektebene und wie viel Anstrengung erforderlich ist, um das Problem zu lösen. Diese Art von Anstrengung ist durch die zu bewältigende Datenmenge verursacht und lässt sich messen, indem man fragt: „Wie schwierig ist diese Aufgabe für dich?“

In einer zweiten Bedeutung spiegelt Anstrengung die Belastung oder Aversion, die eine Person bei der Aufgabenbearbeitung empfindet (Anstrengung durch Bedürfnisversagung). Anstrengung resultiert hierbei aus den negativ bewerteten affektiv-emotionalen Verarbeitungsanforderungen des Lernens selbst. Diese Art von Anstrengung ist durch die Aufgabenerledigung verursacht und lässt sich messen, indem man fragt: „Wie anstrengend ist es für dich an dieser Aufgabe zu arbeiten?“

In einer dritten Bedeutung bezieht sich Anstrengung auf die Initiierung, Intensität und Ausdauer von Lernverhalten (Anstrengung durch Zuteilung). Hier geht um die selbstgewählte Zuteilung von Ressourcen zur Bewältigung eines Lernproblems. Diese Art von Anstrengung ist durch die Bewertung einer Aufgabe hinsichtlich Erwartung und Wert verursacht. Um Anstrengung durch Zuteilung von Ressourcen zu messen, lässt sich fragen: „Wie sicher bist du dir/wichtig ist es dir, diese Aufgabe zu lösen?“

Jede der drei Bedeutungen von Anstrengung ist mit einer bedeutsamen theoretischen Perspektive verbunden. Die Theorie der kognitiven Belastung (Sweller et al., 2019) spiegelt sich im Konstrukt Anstrengung durch Komplexität. Die Selbstbestimmungstheorie (Ryan & Deci, 2020) findet sich in der Anstrengung durch Bedürfnisversagung wieder und die situierte Erwartungs-Wert-Theorie (Eccles & Wigfield, 2020) ist an die Anstrengung durch Zuteilung geknüpft.

In der Praxis verbinden sich jedoch alle drei Bedeutungen von Anstrengung miteinander. Wenn eine Schülerin am Abend noch ihre Hausaufgaben erledigen soll, weil die Eltern es so wünschen, kann durch die Einschränkung ihrer Freizeit das Bedürfnis nach Selbstbestimmung frustriert sein. Stellen sich die Aufgaben als objektiv anspruchsvoll heraus, ist auch ihr Empfinden von Anstrengung durch Komplexität hoch. Ihre Erwartung, die Aufgaben zu so später Stunde noch gut zu bewältigen, ist gering und der Wert, den sie der Erledigung der Hausaufgaben beimisst, vielleicht ebenso.

Die differenzierte Betrachtung von Anstrengung erlaubt es, unser Verständnis von Anstrengung zu erweitern. Wenn mentale Anstrengung in der Theorie der kognitiven Belastung als motivationale Kosten dargestellt werden (Feldon et al., 2019), ist es wichtig, zwischen den objektiven und den affektiv-emotionalen Aufgabenanforderungen zu unterscheiden. Wenn Anstrengung als frustrierend und befriedigend zugleich erlebt wird, kann es daran liegen, dass die Aufgabe zwar als bedürfnisversagend, aber die Motivation zur Aufgabenerledigung durch hohe Erwartungs-Wert-Überzeugungen charakterisiert ist. Im Vortrag werden weitere Vorteile und Implikationen dieses Anstrengungsverständnisses vorgestellt.

Drei Arten der Lernanstrengung: Wie wichtig sind sie beim Textlesen?

Laura Kehle, Detlef Urhahne

Universität Passau

Anstrengung als motivationaler Prädiktor der Lernleistung wird in der Literatur unterschiedlich definiert. Grund, Fries, Nückles, Renkl und Roelle (2024) schlagen daher eine Unterscheidung von drei verschiedenen Arten der Anstrengung vor: (1) Anstrengung als die Wahrnehmung von Komplexität bzw. Schwierigkeit einer Lernaufgabe, (2) Anstrengung als Frustration von Bedürfnissen bei der Bearbeitung einer Lernaufgabe und (3) die persönliche Anstrengung als Mobilisierung und Einsatz von Ressourcen zur Bewältigung einer Lernaufgabe.

Die vorliegende Studie untersucht diese drei Arten von Anstrengung und deren Zusammenhang mit der Leseleistung. Der Theorie entsprechend wird angenommen, dass Anstrengung als Komplexität und Anstrengung als Frustration negativ mit der Leseleistung zusammenhängen, während die persönliche Anstrengung als Mobilisierung von Ressourcen positiv mit der Leseleistung korreliert.

An der Studie nahmen 143 Studierende (57.3% weiblich) im Alter von $M = 13.47$ Jahren ($SD = .79$) teil. Es wurde der Lesegeschwindigkeits- und -verständnistest für die Klassen 5–12 (LGVT 5-12+; Schneider et al., 2017) als Leseaufgabe eingesetzt. Dabei wurde die Leseleistung als Lesegenauigkeit und -geschwindigkeit gemessen. Die drei Anstrengungsarten wurden zu Beginn, während der Testbearbeitung sowie nach der Testbearbeitung mittels reliabler Fragebogenskalen (nach Gendolla & Wright, 2009 und Grund et al., 2024) erfasst.

Mittels explorativer Faktorenanalysen, Korrelationsberechnungen und multipler Regressionsanalysen wurden die vermuteten Zusammenhänge empirisch geprüft.

Explorative Faktorenanalysen zu den verschiedenen Messzeitpunkten ergaben für die drei Arten der Anstrengung jeweils eine dreifaktorielle Lösung, was auf die Trennbarkeit der hypothetischen Konstrukte hindeutet.

Multiple Regressionen zur Vorhersage der Lesegenauigkeit anhand der drei Anstrengungsarten zu den drei Messzeitpunkten ergaben signifikante Modelle (MZP1: $F(3, 139) = 9.09$, $p < .001$, $R^2 = .15$; MZP2: $F(3, 139) = 5.02$, $p < .01$, $R^2 = .08$; MZP3: $F(3, 137) = 10.57$, $p < .001$, $R^2 = .17$). Persönliche Anstrengung als Mobilisierung von Ressourcen war zu jedem Messzeitpunkt ein signifikanter Prädiktor (tMZP1 = 5.06, $p < .001$; tMZP2 = 3.41, $p < .001$; tMZP3 = 4.58, $p < .001$), Anstrengung als Komplexität (tMZP1 = -.45, n.s.; tMZP2 = .38, n.s.; tMZP3 = -.12, n.s.) oder Frustration (tMZP1 = 1.08, n.s.; tMZP2 = .24, n.s.; tMZP3 = -1.82, n.s.) beeinflussten die Lesegenauigkeit jedoch nicht. Multiple Regressionen zur Vorhersage der Lesegeschwindigkeit anhand der drei Anstrengungsarten zu den drei Messzeitpunkten führten zu keinen signifikanten Modellen (MZP1: $F(3, 139) = 2.07$, n.s.; MZP2: $F(3, 139) = 1.09$, n.s.; MZP3: $F(3, 137) = .86$, n.s.).

Die Ergebnisse lassen vermuten, dass tatsächlich zwischen drei verschiedenen Arten der Anstrengung beim Lernen unterschieden werden kann. Dabei sagt insbesondere die persönliche Anstrengung die Leseleistung vorher. Der Einfluss der persönlichen Anstrengung auf das Lesen bezieht sich nicht auf die Quantität des gelesenen Textes, sondern beeinflusst vielmehr die Intensität der Lesetätigkeit. Die Wahrnehmung von Anstrengung als Schwierigkeit bzw. Komplexität sowie die Aversion gegenüber einer Aufgabe (Anstrengung als die Frustration von Bedürfnissen) stellen eine Vermeidungshaltung dar. Diese ruft zwar Unbehagen gegenüber einer Aufgabe hervor, hat jedoch keinen Einfluss auf die Leseleistung. Es könnte jedoch sein, dass diese Arten der Anstrengung bei anderen, unangenehmeren und damit mühevolleren Lernaufgaben eine prädiktive Rolle einnehmen.

Adaptive Lernanstrengungen nach Fehlern und deren Wissensseffekt bei Grundschulkindern: Ergebnisse einer prozessorientierten Studie im authentischen Unterrichtsetting

Jana Spear¹, Donna Bryce¹, Carolin Burmeister², Robert Grassinger², Markus Dresel¹

¹Universität Augsburg, ²Pädagogische Hochschule Weingarten

Fehler sind ein wichtiger Bestandteil der Unterrichtspraxis, da sie viel Lernpotenzial bieten (Metcalfe, 2017). So können Fehler beispielsweise helfen, negatives Wissen aufzubauen (Oser & Spychiger, 2005) oder den Lernprozess zu steuern, indem sie den Lernenden anzeigen, worauf das Lernverhalten fokussiert werden sollte. Eine solche Anpassung des Lernverhaltens wird im Konzept der auf die Lernhandlung bezogenen Reaktionen (sog. handlungsadaptive Reaktionen) auf Fehler beschrieben und im Sinne einer zusätzlich aufgebrachten Lernanstrengung über (meta-)kognitive Prozesse und Lernhandlungen, die auf die gezielte Überwindung des Fehlers ausgerichtet sind, definiert (Dresel et al., 2013). Für diese adaptive Lernanstrengung nach Fehlern wird postuliert, dass sie den Wissenserwerb bedeutsam unterstützt (Grassinger et al., 2018). Die aktuelle Forschungslage zeigt bisher aber kaum Befunde für den Altersbereich von Grundschulkindern und beschränkt sich außerdem weitgehend auf die Berücksichtigung von Selbstberichtsdaten. Obwohl dem Selbstbericht von Kindern zwar zugeschrieben wird, eine wichtige interne Perspektive zu liefern und dementsprechend berücksichtigt werden sollte (Sturges et al., 2002), ist ein reliables und valides Antwortverhalten nicht immer gegeben (Conijn et al., 2020). Folglich ist die Notwendigkeit einer multimodalen Erfassung des Lernverhaltens bei Kindern im Grundschulalter besonders hoch (Winne, 2010). Dazu wird in der vorliegenden Feldstudie ein Untersuchungsparadigma erstmals an Grundschulkindern angewendet, mit dem der Effekt der adaptiven Steuerung der

Lernanstrengung direkt innerhalb einer authentischen Lernsituation untersucht und mit einer multimodalen Erfassung des Lernverhaltens nach Fehlern verbunden wird.

Dafür bearbeiteten 271 Grundschulkinder der dritten Jahrgangsstufe (Alter: $M = 9.99$ Jahre, $SD = 0.74$) einen Vorwissenstest, eine digitale Lernumgebung auf Tablets sowie einen Posttest zur Erfassung des Lernzuwachses im Fach Mathematik. Das Vorwissen wurde durch Subskalen des DEMAT 2+ (Krajewski et al., 2020) erhoben, die die Inhalte der darauffolgenden digitalen Lernumgebung abdeckten. Die Lernumgebung konnte in der vorgegebenen Zeit im eigenen Lerntempo bearbeitet werden und beinhaltete vier Kapitel mit auf den Lehrplan abgestimmten Inhalten. Nach einer Lernphase mit Erklärungen und Beispielen sollten die Kinder in jedem Kapitel fünf Aufgaben selbst berechnen, worauf jeweils ein Richtig/Falsch Feedback folgte. Am Ende jedes Kapitels erhielten die Kinder zusätzlich ein zusammenfassendes Feedback darüber, wie viele Fehler gemacht wurden. Daraufhin konnten die Kinder entscheiden, ob sie zum nächsten Kapitel weitergehen oder ob sie das vorherige Kapitel nochmal wiederholen möchten und damit die Möglichkeit haben, die Inhalte zu wiederholen und falsche Aufgaben erneut zu bearbeiten. Basierend auf dieser Steuerung des Lernverhaltens nach Fehlern wurde ein behavioraler Indikator zur Erfassung der adaptiven Lernanstrengung nach Fehlern abgeleitet. In einem Posttest wurde der Lernzuwachs mit 12 Items erfasst, die die Lerninhalte der digitalen Lernumgebung abdeckten mit jeweils 3 Items pro Kapitel. Die Scores wurden als Anteil an richtig gelösten Items berechnet. Im Posttest war zusätzlich eine adaptierte Fassung der Selbstberichtsskalen von Dresel et al. (2013) zur Erfassung der adaptiven Lernanstrengung nach Fehlern enthalten. Ein Item dieser Skala wurde als zusätzliche State-Messung der Lernanstrengung nach Fehlern direkt am Ende jedes Kapitels eingesetzt.

Alle drei Indikatoren zeigten übereinstimmend einen Wissensseffekt von adaptiven Lernanstrengungen nach Fehlern. In einem gemeinsamen Modell aller Indikatoren konnte ein Wissensseffekt von $\beta = .23$ gezeigt werden. Allerdings waren die Korrelationen zwischen den Selbstberichten und der behavioralen Messung gering. Diese Ergebnisse zeigen einerseits die kritisierte Differenz zwischen Selbstberichtsdaten von Kindern und ihrem tatsächlichen Verhalten auf und unterstützen damit erneut die Notwendigkeit der multimodalen Erfassung des Lernverhaltens (Spörer & Brunstein, 2006). Andererseits wird aber auch die Bedeutung eines adaptiven Umgangs mit Fehlern sowie die Steuerung der Lernanstrengung für den erfolgreichen Wissenserwerb hervorgehoben – bereits im Alter von Grundschulkindern.

S-93: Erwachsenenbildung und gesellschaftlicher Wandel. Methodische Herausforderungen und Innovationen in der Erfassung von (Weiter-) Bildungsinhalten.

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 150

Chair der Sitzung: Jana Costa

Chair der Sitzung: Lena Nusser

Chair der Sitzung: Sonja Hahn

Erwachsenenbildung und gesellschaftlicher Wandel. Methodische Herausforderungen und Innovationen in der Erfassung von (Weiter-) Bildungsinhalten.

Chair(s): **Jana Costa** (Leibniz-Institut für Bildungsverläufe, Deutschland), **Lena Nusser** (Leibniz-Institut für Bildungsverläufe, Deutschland)

Diskutant*in(nen): **Sonja Hahn** (Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Im Horizont multipler Herausforderungen und beschleunigter (global-) gesellschaftlicher Entwicklungen wird dem lebenslangen Lernen von Erwachsenen eine grundlegende Bedeutung zugeschrieben (z.B. Bagnall & Hodge, 2022; Field, 2006). Gesellschaftliche Herausforderungen können dabei zugleich als Anlass für die Lern- und Bildungsaktivitäten Erwachsener betrachtet werden wie auch als Mittel zu ihrer Bewältigung (z.B. Schreiber-Barsch & Mauch, 2019; Singer-Brodowski, 2023).

Themen und Inhalte von Weiterbildung sowie deren Funktionen für Individuen, Organisationen und die Gesellschaft sind bislang allerdings unzureichend erfasst und wir wissen wenig darüber, mit welchen Themen sich Erwachsene angesichts gesellschaftlicher Herausforderungen in Fort- und Weiterbildungen, aber auch eigeninitiativ und selbstgesteuert, auseinandersetzen (z.B. Kühn & Costa, 2024, in press). Dies begrenzt die Leistungsfähigkeit von Bildungsberichterstattung und Bildungsforschung, das Steuerungspotential politischer und zivilgesellschaftlicher Akteure sowie die Lern- und Entwicklungsmöglichkeiten der Weiterbildungspraxis.

Methodisch basiert die empirische Weiterbildungsforschung zur Erfassung von Themen und Inhalten der Erwachsenenbildung zumeist auf inhaltsanalytischen Verfahren. Diese Methoden (grundlegend Berelson (1952), Krippendorff (2022)) werden beispielsweise auf veröffentlichte Weiterbildungsprogramme und Angebote angewendet. Ebenso kommen sie bei der Codierung offener Angaben in Survey-Studien wie dem Nationalen Bildungspanel (NEPS), dem Adult Education Survey (AES) und dem Mikrozensus zum Einsatz.

Der Mangel an belastbaren Informationen über die Themen und Inhalte von Weiterbildungen sowie deren Konjunktur ist auf zwei Hauptgründe zurückzuführen: Zum einen ist die Vielfalt der Weiterbildungsangebote empirisch schwer erfassbar, zum anderen ist der Codieraufwand angesichts der Datenfülle sehr hoch. Obwohl für die Klassifikation textlicher Angaben bereits seit längerem maschinelle Verfahren wie Natural Language Processing (NLP) (z.B. Goldberg, 2017) und zunehmend auch Large Language Modelle (Li et al. 2022) eingesetzt werden, finden diese Methoden in der Weiterbildungsforschung bislang keine Anwendung.

Das vorgeschlagene Symposium greift dieses Desiderat auf und widmet sich methodischen Herausforderungen und Innovationen in der Erfassung von Themen und Inhalten der Lernaktivitäten von Erwachsenen. In drei Beiträgen wird das Potenzial offener (Inhalts-)Angaben herausgearbeitet, indem Möglichkeiten der Reanalyse bestehender Datensätze (NEPS) vorgestellt und methodische Innovationen, insbesondere im Hinblick auf die Unterstützung durch Large-Language-Modellen, diskutiert werden.

Beitrag 1: Informelles Lernen im Horizont gesellschaftlicher Herausforderungen: Trends in der Beschäftigung mit nachhaltigkeitsbezogenen Themen im Erwachsenenalter.

Im ersten Beitrag werden die Potenziale offener Angaben zur Erfassung von Lerninhalten im Feld der Erwachsenenbildung unter einer spezifischen thematischen Perspektive herausgearbeitet. In Rückgriff auf die Daten des Nationalen Bildungspanels werden mit einem klassischen inhaltsanalytischen Zugang offene Angaben hinsichtlich ihres Bezugs zum Themenfeld Nachhaltigkeit identifiziert und in ihrer zeitlichen Dynamik abgebildet. Im Beitrag werden insbesondere methodische Herausforderungen thematisiert, die in den folgenden Beiträgen aufgegriffen werden.

Beitrag 2: Ableitung eines initialen Kategoriensystems zur Klassifikation von Weiterbildung unter Verwendung eines Large – Language-Models

Im zweiten Beitrag wird anhand eines Vorhabens zur Klassifikation von Weiterbildungsangeboten untersucht, inwiefern spezifische Komponenten und Arbeitsschritte der für dieses Feld typischen Inhaltsanalysen von LLM übernommen werden können. Im Zentrum steht dabei die initiale Ableitung eines Kategorienschemata aus bestehenden Klassifikationen und Kodierleitfäden zur weiteren Anwendung auf Ankündigungstexte. Der Beitrag demonstriert dies anhand einer ersten Stichprobe.

Beitrag 3: Zur Entwicklung und Erprobung einer Klassifikation von Inhalten und Themen der Erwachsenen- und Weiterbildung auf der Basis von Kursdatenbanken

Der dritte Beitrag schließt an den zweiten Vortrag an. Von einem initialen Kategoriensystem ausgehend wird untersucht, inwiefern ein Verfahren entwickelt werden kann, welches es in einem iterativen Prozess ermöglicht, das Kategoriensystem empirisch zu validieren und zu erweitern. Dabei werden die für das inhaltsanalytische Vorgehen zentralen Arbeitsschritte der Strukturierung an ein LLM übertragen. Das Verfahren wird in dem Beitrag veranschaulicht und erste Ergebnisse präsentiert.

Beiträge des Symposiums

Informelles Lernen im Horizont gesellschaftlicher Herausforderungen: Trends in der Beschäftigung mit nachhaltigkeitsbezogenen Themen im Erwachsenenalter

Jana Costa

Leibniz-Institut für Bildungsverläufe, Deutschland

Theoretische Hintergrund

Im Horizont globalgesellschaftlicher Herausforderungen gewinnt die Auseinandersetzung mit Fragen des lebenslangen Lernens im Kontext von Nachhaltigkeit zunehmend an Bedeutung. Es ist dabei unstrittig, dass eine transformative Wende in Richtung Nachhaltigkeit Erwachsene als wichtige Akteure mit einbeziehen muss und lebenslanges Lernen erfordert (z.B. Singer-Brodowski 2023; Adoßment 2015). Nachhaltigkeitsbezogene Themen werden in der Erwachsenenbildung bislang traditionell in den Bereichen Umweltbildung (z.B. Apel, 2016; Kandler & Tippelt, 2011), Politische Bildung (z.B. Klemm, 2019) oder Globales Lernen

(z.B. Scheunpflug & Franz, 2012) bearbeitet und theoretisch-konzeptionell sowie didaktisch weiterentwickelt (z.B. Franz et al., 2009; Singer-Brodowski, 2023; Schreiber-Barsch, 2020). Empirische Studien zum Themenfeld fokussieren bislang vorwiegend den non-formalen Bildungsbereich. So werden z.B. außerschulische Bildungsangebote untersucht (z.B. Michelsen et al., 2013) und Maßnahmen der Qualitätssicherung diskutiert (z.B. Boden, 2020). Darüber, welche Rolle Nachhaltigkeit für das Lernen Erwachsener jenseits vorstrukturierter Angebote spielt, ist wenig bekannt (z.B. Van Poeck et al., 2020).

Fragestellung

Anknüpfend an dieses Desiderat steht in diesem Beitrag die Beschäftigung mit nachhaltigkeitsbezogenen Lernthemen in verschiedenen Kontexten (Fachmessen und -kongresse, Fachvorträge & Internet) im Mittelpunkt. Besonderer Fokus liegt dabei auf der Frage, inwiefern sich auf aggregierter Ebene thematische Trends im Zeitverlauf skizzieren lassen. Dazu wird u.a. danach gefragt, ob sich zu bestimmten gesellschaftlichen Ereignissen (wie z.B. dem Aufkommen der "Fridays for Future"-Bewegung) eine Häufung von spezifischen Themen in den Daten abzeichnet.

Methodisches Design

Den Ausgangspunkt der Studie bildet ein interdisziplinäres Forschungsprojekt, in welchem die Untersuchung des informellen Lernens Erwachsener im Kontext von Nachhaltigkeit unter Rückbezug auf vorliegende Datensätze des Nationalen Bildungspanels (NEPS) im Mittelpunkt steht. Die Daten der Erwachsenenkohorte des NEPS werden dabei reanalysiert, indem die erfassten offenen Angaben zu Lernthemen in verschiedenen informellen Kontexten in einem ersten Schritt inhaltsanalytisch ausgewertet werden. Auf Basis der offenen Angaben zu Lernthemen der Erwachsenen im Jahr 2020/21 (n=6.674; ca. 3.000 offene Angaben) in verschiedenen informellen Kontexten (Fachmessen und -kongresse, Fachvorträge & Internet) wurde mithilfe einer induktiv-deduktiven Inhaltsanalyse ein Kategoriensystem entwickelt, das Aufschluss über explizite nachhaltigkeitsbezogene Lernthemen gibt. Dieses Kategoriensystem wurde wieder an den Datensatz angewendet: Bislang wurden 17.395 offene Aussagen aus vier Erhebungswellen (2017/18; 2019/20; 2020/21; 2021/22) von drei studentischen Hilfskräften kodiert (Gwet's agreement coefficient: 0,92) und den Autor:innen validiert, wobei bis zum Beitrag noch weitere Wellen kodiert sein werden. Mit den vier Erhebungswellen liegen in der Analytestichprobe bislang Informationen von insgesamt 6.793 Personen vor.

Ergebnisse

Erste Analysen verweisen darauf, dass über die Kontexte hinweg lediglich bei 6,61% der Befragten ein expliziter Bezug zum Thema Nachhaltigkeit kodiert wurde. Aufgrund der häufig abstrakt formulierten Angaben können allerdings unter dem Deckmantel der Abstraktheit weitere Lernthemen verschleiert bleiben, die zwar einen Bezug zu Nachhaltigkeit aufweisen, aufgrund der Formulierung allerdings nicht in die Auswertung aufgenommen wurden. Die Befunde deuten weiterhin darauf hin, dass die Wahrscheinlichkeit, sich mit nachhaltigkeitsbezogenen Lernthemen zu beschäftigen, mit höherer formaler Bildung, höherem Einkommen sowie steigendem beruflichen Status zunimmt. Auffällig ist, dass 32,21% der Befragten, bei denen ein Nachhaltigkeitsbezug kodiert wurde, sich auch mehrfach in verschiedenen Erhebungswellen und/oder Lernkontexten mit entsprechenden Themen beschäftigen.

Für den eingereichten Beitrag werden gezielt thematische Trends auf Aggregatebene herausgearbeitet und vor dem Hintergrund gesellschaftlicher Entwicklungen diskutiert. Hierfür bildet die Kodierung weiterer Erhebungswellen die Grundlage. Da die offenen Angaben zum informellen Lernen jährlich erfasst werden, lassen sich thematische Trends dann potenziell bis in das Jahr 2009/10 zurückverfolgen.

Diskussion

In dem Beitrag wird das Potential bereits vorliegender offener Angaben zu Lerninhalten aufgegriffen, indem verschiedene Forschungsparadigmen sowie deren Stärken und Potenziale zusammengebracht werden. Dies eröffnet sowohl inhaltliche als auch methodische Diskussionsperspektiven, die im Beitrag aufgegriffen und insbesondere mit Blick auf methodische Innovationen in der Perspektive auf die Anwendung von Large-Language-Models in der Erfassung von Lernthemen im Erwachsenenalter diskutiert werden.

Ableitung eines initialen Kategoriensystems zur Klassifikation von Weiterbildung unter Verwendung eines Large –Language-Models

Andreas Martin, Josef Schrader

Deutsches Institut für Erwachsenenbildung

Theoretischer Hintergrund

Weiterbildung ist in Deutschland in Folge ihrer historischen Genese mehr als alle anderen Bildungsbereiche durch Vielfalt und Heterogenität ihrer Angebote geprägt. Eine Folge dieser Heterogenität ist, dass Weiterbildungsangebote zumeist keiner curricularen Systematik folgen und weitgehend unstandardisiert gestaltet werden (Schrader, 2019). Dies stellt die empirische Forschung zu den Ursachen und Erträgen von Weiterbildung vor große Herausforderungen, denn die zu erwartenden Effekte sind in hohem Maße durch die jeweiligen Weiterbildungsinhalte vermittelt. Bisher fehlt jedoch eine umfassende, die Vielfalt der Weiterbildungsangebote abdeckende Klassifikation von Weiterbildungsangeboten, die es ermöglicht, auf der Grundlage unterschiedlicher Datenzugänge eine inhaltsvalide und reliable Erfassung von Weiterbildungsaktivitäten zu gewährleisten. Der vorgeschlagene Beitrag widmet sich der Frage, ob und wie mit Hilfe von Large-Language-Modellen ein umfassendes Klassifikationssystem entwickelt werden kann.

Die Entwicklung von Klassifikationen ist in der Weiterbildungsforschung bisher ein Schwerpunkt der inhaltsanalytischen Programmforschung, folgt vor allem in Anlehnung an Mayring (2015) einem induktiv – deduktiven oder abduktiven Prozess (Gieseke & Robak, 2022) und basiert aufgrund des hohen Aufwandes zumeist auf geringen Fallzahlen, insofern die Erhebungen thematisch, regional und zeitlich begrenzt durchgeführt werden. In der Weiterbildungsforschung gibt es infolgedessen eine Vielzahl von Kategoriensystemen mit begrenzter Aussagekraft. Ziel des hier vorgestellten Vorhabens ist es, einen entsprechenden Prozess der Kategorienbildung unter Verwendung von Large-Language-Modellen zu erproben, der zu einem umfassenden Kategoriensystem führt, das generalisierende Aussagen erlaubt. Ausgangspunkt ist dabei die Ableitung eines initialen Klassifikationssystems aus bestehenden Klassifikationen. Dieses soll anschließend in einem iterativen Prozess an einer großen Zahl von Ankündigungstexten zu Weiterbildungsangeboten empirisch gesättigt werden (s. dazu Beitrag 3 dieses Symposiums). Eine erste Herausforderung ergibt sich aus der dimensionalen Unschärfe der in den bestehenden Klassifikationssystemen der Erwachsenen- und Weiterbildung genutzten Begriffe. Die Einordnung und Klassifikation von Lerngelegenheiten orientiert sich sehr häufig an einer praxisbezogenen Terminologie, in der Lerngelegenheiten durch die spezifische Kombination von Inhalt, Methode, Lernort, Lernziel, Verwendungszweck und Zielgruppe (Schlutz, 2006) definiert werden. Dies zeigt sich z.B. in Kategorien wie „Integrations- und Orientierungskurse“, „finanzielle Grundbildung“, „Nachgeholt Schulabschlüsse“, „wissenschaftliche Weiterbildung“ oder „Bildung für Ältere“. Die Entwicklung eines alle bestehenden

Klassifikationen zusammenfassenden umfassenden und disjunkten Kategoriensystems setzt die dimensionale Differenzierung der bislang verwendeten Kategorien und deren Synthese in ein integriertes, mehrdimensionales Klassifikationssystem voraus.

Fragestellung

In vorbereitenden Arbeitsschritten wurde der Forschungsstand zur Programm- und Angebotsforschung (Fleige et al., 2019) aufgearbeitet und ein Scoping Review vorbereitet. Im Zuge dieser Recherchen wurden bislang ca. 60, zumeist hierarchisch strukturierte Kategorienschemata recherchiert, die den Ausgangspunkt für die Entwicklung eines umfassenden Kategoriensystems bilden. Aufgrund des Umfangs des Ausgangsmaterials nutzen wir die Potentiale von Large-Language-Modellen. Im Zentrum des Beitrages steht darum die Frage, inwiefern der Einsatz von LLMs dieses Vorgehen ermöglicht und die Ergebnisse dem Urteil von Expert:innen standhalten.

Methodik

Im hier zunächst vorgestellten Prozess der Entwicklung eines umfassenden und diskunkten Kategoriensystems soll ein Large Language Model genutzt werden. Um den Prozess der Verarbeitung durch GPT 4o nachvollziehbar zu gestalten, wird hier nur eine Teilmenge dieser Klassifikationen sequenziell über eine API-Schnittstelle verarbeitet. Dazu wurde ein Prompting entlang der PASSD (Persona, Aufgabe, Situation, Strategie, Dialog) - Systematik entwickelt, um zunächst ein Aggregat der in unterschiedlichen Klassifikationssystemen verwendeten Kategorien zu bilden und dieses anhand der spezifischen Kombination von Inhalt, Methode, Lernort, Lernziel, Verwendungszweck und Zielgruppe aufzuschlüsseln. Anschließend sollen Vorschläge zur Integration der Teildimensionen in ein mehrdimensionales Klassifikationssystem synthetisiert werden.

Ergebnisse

Erste Ergebnisse zeigen, dass die Gestaltung der Dialog-Sequenzen durch die API-Schnittstelle eine stetige Rekontextualisierung der vorhergehenden Dialoge erfordert und so zu einer Kumulation der Kommunikation führt. Da die kommunizierbare Tokenzahl über die API-Schnittstellen begrenzt ist, muss der Dialog rationalisiert werden.

Zur Entwicklung und Erprobung einer Klassifikation von Inhalten und Themen der Erwachsenen- und Weiterbildung auf der Basis von Kursdatenbanken

Niklas Kallinger¹, Andreas Martin², Josef Schrader², Augustin Kelava¹

¹Universität Tübingen, ²Deutsches Institut für Erwachsenenbildung

Hintergrund

Die Forschung zur Programm- und Angebotsentwicklung in der Erwachsenen- und Weiterbildung stützt sich typischerweise auf deduktiv-induktive Verfahren der kategorialen Erfassung von Ankündigungstexten (Gieseke & Robak 2022; Nolda, 2015), die jedoch auf ausgewählte Themenbereiche oder wenige Anbieter begrenzt sind und häufig keine Aussagen zur Validität und Reliabilität der Codierungen, die zumeist von studentischen Hilfskräften vorgenommen werden, treffen. In diesem Beitrag wird ein weiterentwickeltes Vorgehen vorgestellt, das eine breitere empirische (Daten-)Grundlage auf der Basis von Weiterbildungsdatenbanken und Large-Language-Modellen (LLM) nutzt.

Fragestellung

Die Ausgangslage ist ein bestehendes, initial deduktiv erarbeitetes Kategoriensystem, welches eine umfassende Klassifikation von Erwachsenenbildungs- und Weiterbildungsangeboten erlaubt. Es stellt sich die Frage, wie ein Verfahren entwickelt werden kann, welches es in einem iterativen Prozess ermöglicht, das Kategoriensystem empirisch zu validieren und zu erweitern. In diesem Sinne schließt das Vorgehen an den in Inhaltsanalysen wesentlichen Arbeitsschritt der Strukturierung an (Mayring, 2015, 2019) und überträgt diesen an ein Large-Language-Model. Dieses Verfahren wird im Beitrag vorgestellt und eine erste Implementation eines Ausschnitts des Verfahrens veranschaulicht.

Methode

Das allgemeine Verfahren sieht vor, dass eine große Datengrundlage von Ankündigungstexten aus 77 regionalen und überregionalen Weiterbildungsdatenbanken, die in der Metadatenbank „InfoWeb Weiterbildung“ zusammengeführt werden auf das Kategoriensystem angewendet werden. Diese Metadatenbank umfasst 3.498.398 aktuelle Ankündigungstexte, ergänzt durch Kontextinformationen zu Anbietern und Regionen sowie Kurstitel und Schlagworte.

Das grundsätzliche iterative Verfahren sieht mehrere Schritte vor, die im Beitrag vorgestellt werden, wobei aus Anschaulichkeitsgründen der Fokus auf der konkreten Darstellung des 4. Punktes des Verfahrens (nämlich Prompting und LLM-Anwendung) liegt:

1. Ergänzung des initialen Kategoriensystems: Zunächst wird eine Residualkategorie für nicht zuweisbare Inhalte dem initialen Kategoriensystem als Erweiterung zugeordnet.
2. Datensatzaufteilung: Ferner erfolgt die Aufteilung (Splitten) des Datensatzes in einen Test- und einen Trainingsdatensatz.
3. Zuordnung idealtypischer Ankündigungstexte: Experten ordnen idealtypische Ankündigungstexte dem initialen Kategoriensystem zu.
4. Prompting und LLM-Anwendung: Danach erfolgt die Übergabe des angereicherten Kategoriensystems an ein vortrainiertes Transformer-basiertes LLM z.B. GPT4, OpenAI, 2024a, um weitere Ankündigungstexte zu klassifizieren. Das konkrete „Wie“ (Prompting und Vorstellung des Application Programming Interfaces, API; OpenAI, 2024b) wird in dem Beitrag exemplarisch elaboriert. Dabei sollen auch die praktischen Unwägbarkeiten beleuchtet werden.
5. Interrater-Reliabilität: Danach erfolgt die Überprüfung der Übereinstimmung zwischen LLM und Experten.
6. Iterative Verfeinerung: Wiederholung des Prozesses, bis die „Interrater-Reliabilität“ zwischen Experten und LLM auf einem hohen Niveau konvergiert.
7. Abschließend werden Ankündigungstexte der Residualkategorie durch Expertinnen und Experten neu zugeordnet und das Kategoriensystem entsprechend final erweitert.

Ergebnisse

Das Ergebnis dieses Verfahrens ist ein neues, optimiertes Kategoriensystem, das den Anforderungen der empirischen Bildungsforschung und Bildungsberichterstattung gerecht werden soll. Es wird erwartet, dass durch die dynamische Anpassung und ständige Verbesserung des Kategoriensystems die Zahl der nicht zuordenbaren Fälle sukzessive reduziert wird, auch wenn gelegentlich neue Residualfälle auftreten.

Die erste Implementation an einem Teildatensatz, insbesondere des 4. Schrittes (Prompting und LLM-Anwendung), zeigt, dass Ankündigungstexte inhaltsvalide und reliabel klassifiziert werden können, wenn die Instruktionen des LLMs elaboriert sind.

S-94: Eine ganzheitliche Perspektive auf Bildung wagen: Lehrkräfte und Schüler*innen an Schulen in herausfordernden Lagen stärken

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: SN 163

Chair der Sitzung: Patrick Hawlitschek

Chair der Sitzung: Simone Jambor-Fahlen

Chair der Sitzung: Hans Pant

Eine ganzheitliche Perspektive auf Bildung wagen: Lehrkräfte und Schüler*innen an Schulen in herausfordernden Lagen stärken

Chair(s): **Patrick Hawlitschek** (Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin e.V., Deutschland), **Simone Jambor-Fahlen** (Universität zu Köln)

Diskutant*in(nen): **Hans Anand Pant** (Humboldt-Universität zu Berlin & IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik)

Auch in 2024 verweist die nationale Bildungsberichterstattung erneut auf den engen Zusammenhang zwischen sozialer Herkunft und dem Bildungserfolg von Schüler*innen in Deutschland (Autor*innengruppe Bildungsberichterstattung, 2024). Dieser spiegelt sich nicht nur in den mathematischen und (schrift-)sprachlichen Kompetenzen von Schüler*innen, sondern z.B. auch in ihrer Bildungsbeteiligung wider. Um den Abbau sozialer Bildungsungleichheiten in Deutschland zu unterstützen, kommt daher der gezielten Förderung von Schulen in herausfordernden Lagen eine wachsende Bedeutung zu (Marx & Maaz, 2023). Im Allgemeinen zeichnen sich Schulen in herausfordernden Lagen durch eine Vielzahl an externen Merkmalen aufseiten ihrer Schüler*innen aus, die für das Lernen ungünstig sind (z. B. geringes Vorwissen, schwierige Lebensverhältnisse, hoher Anteil mit nichtdeutscher Erstsprache, aggressives Verhalten). Andererseits weisen diese Schulen interne Merkmale auf, die auf eine niedrige Prozessqualität hindeuten (z. B. ungünstige Organisations- und Unterrichtskultur, verringerte Kooperationsfähigkeit sowie hohes Belastungserleben unter Lehrkräften; Holtappels & Brücher, 2021; van Ackeren et al., 2021). Um die Prozessqualität an diesen Schulen zu stärken, bedarf es daher ganzheitlicher Unterstützungsangebote, die sowohl die Schüler*innen, das pädagogisch tätige Personal als auch den schulischen Sozialraum in den Blick nehmen (Dietrichson et al., 2017).

Vor diesem Hintergrund geht das Symposium der Frage nach, wie Schulen in herausfordernden Lagen unterstützt werden können, um innerschulische Prozessqualitäten weiterzuentwickeln. Dazu werden Befunde aus der wissenschaftlichen Begleitung der Bund-Länder-Initiative „Schule macht stark“ (SchuMaS) vorgestellt.

Im ersten Beitrag beleuchten Woerfel et al. die Wirksamkeit eines Leseflüssigkeitstrainings, das mit der Einbindung eines Audiostiftes darauf abzielt, neben fachlichen Kompetenzen auch die Lesemotivation von Schüler*innen zu stärken. Der zweite Beitrag von Hawlitschek et al. beschäftigt sich mit datengestützter Unterrichtsentwicklung als eine mögliche Strategie, um Unterricht an die Lernbedarfe der Schüler*innen auszurichten. Es werden erste Hinweise zur Weiterentwicklung der professionellen Kompetenz von Lehrkräften (Wissen, Selbstwirksamkeit) als auch Zusammenhänge zur intendierten Verhaltensänderung untersucht. Im dritten Beitrag stellen Klusmann et al. erste Ergebnisse zur Wirksamkeit einer Professionalisierungsmaßnahme vor, die Lehrkräfte beim Abbau von beruflichem Belastungserleben unterstützen soll. Dabei liegt der Fokus darauf, mit welchen Ausgangsvoraussetzungen (z.B. emotionale Erschöpfung) sich Lehrkräfte überhaupt für eine Teilnahme entscheiden. Im vierten Beitrag untersuchen Karst et al. Fürsorge-Praktiken von Lehrkräften gegenüber armutsbetroffenen Schüler*innen. Der Beitrag leitet daraus Implikationen für die Lehrkraft-Schüler*innen-Interaktionen an Schulen in herausfordernden Lagen ab.

Insgesamt können die Beiträge des Symposiums somit nicht nur Erkenntnisse darüber liefern, über welche einzelnen Ansatzpunkte Schüler*innen und Lehrkräfte an Schulen in herausfordernden Lagen unterstützt werden können, sondern auch, welchen Beitrag ein Schulentwicklungsprojekt zu einer ganzheitlichen Schul- und Unterrichtsentwicklung leisten kann.

Beiträge des Symposiums

Leseförderung als Ziel der Unterrichtsentwicklung: Effekte eines Leseflüssigkeitstrainings für Schüler*innen an Schulen in sozial schwierigen Lagen

Till Woerfel¹, Anne Griepentrog², Seda Yilmaz Woerfel¹, Simone Jambor-Fahlen¹, Alexandra Zepter¹

¹Universität zu Köln, ²Universität Chemnitz

Theoretischer Hintergrund:

Ziel der Bund-Länder-Initiative „Schule macht stark – SchuMaS“ ist es, v.a. Schulen in schwierigen sozialen Lagen und ihre Lehrkräfte über Qualifizierungsangebote darin zu unterstützen, Bildungschancen auch sozial benachteiligten Lernenden zu ermöglichen und dafür die sprachlichen Basiskompetenzen zu stärken. Mit Blick auf die Sprachförderung stehen insbesondere Schüler*innen im Vordergrund, die zu Hause eine andere Sprache als Deutsch sprechen. Diese Schüler*innen besuchen überproportional häufig Schulen in sozial schwierigen Lagen und erzielen im Durchschnitt, auch nach Kontrolle sozioökonomischer Faktoren, am Ende der vierten Klasse schlechtere Ergebnisse in der Lesekompetenz als einsprachig deutschsprachige Schüler*innen (Stanat et al., 2022). Diese Ergebnisse legen nahe, dass besonders hierarchieniedrige, phonologische Fähigkeiten gefördert werden sollten, da diese als wichtige Prädiktoren für die Lesekompetenz gelten. Empirisch hat sich gezeigt, dass die Förderung der Leseflüssigkeit besonders wirksam ist (Philipp & Schilcher, 2012). Dies wurde im Inhaltscluster „Unterrichtsentwicklung“ aufgegriffen und auf der Grundlage von Wirksamkeitsergebnissen der Leseforschung ein trainingsförmiges Leseflüssigkeitskonzept entwickelt, das in Blended-Learning-Qualifizierungsmodulen an die an SchuMaS beteiligten Deutschlehrkräfte vermittelt wird und Gegenstand des Vortrags ist.

Methode:

Das Training wurde von Lehrkräften, die die Blended-Learning-Qualifizierung durchlaufen haben, selbstständig innerhalb von 7–8 Wochen umgesetzt. Um die Wirksamkeit des Trainings hinsichtlich der Förderung der Leseflüssigkeit zu bestimmen, wurde an sechs an SchuMaS beteiligten Schulen in zwei Bundesländern eine quasi-experimentelle Interventionsstudie im Prä-Post-Kontrollgruppenvergleich (N=416; EG=276; KG=140) in 3. (N=191) und 4. Klassen (N=225) durchgeführt. Auf der Grundlage der Schüler*innenangaben sprechen 25,5 % der Teilnehmenden zu Hause nur Deutsch, 58,45% Deutsch und eine weitere Sprache und 16,05% nur eine andere Sprache außer Deutsch.

Als Outcomemessinstrumente wurden der ELFE II (Lenhard et al., 2017), der die Leseleistung mit drei Untertests (Wort-, Lese- und Textverständnis) erfasst, ein Fragebogen zur Erfassung der Lesemotivation, sowie ein C-Test zur Diagnose der allgemeinen Sprachkompetenz eingesetzt.

Um die Effekte des Trainings über die Zeit im Gruppenvergleich zu bestimmen, wurde ein lineares gemischtes Modell (Pinheiro et al., 2024) berechnet, das Individuen als zufällige Effekte sowie Zeit, Gruppe, Sprachstand und Familiensprache als feste Effekte berücksichtigte. Zusätzlich wurden die Interaktionen zwischen Zeit, Gruppe und Klasse im Modell berücksichtigt.

Ergebnisse und ihre Bedeutung:

Die Ergebnisse zeigen, dass sich die Gesamtleseleistung zwar verbessert, jedoch keine signifikanten Unterschiede zwischen der Experimental- und Kontrollgruppe festzustellen sind ($F(1, 286) = 1.016$, $p = 0.3143$). Es zeigt sich zwar ein Haupteffekt der Klassenzugehörigkeit zugunsten der dritten Klasse mit signifikanten Effekten auf die ELFE-Gesamtwerte ($F(1, 287) = 72.69$, $p < .0001$), die Klasseneffekte variieren jedoch nicht zwischen den Gruppen und Messzeitpunkten ($F(1, 254) = 0.058$, $p = 0.81$). Auch die Auswertung der ELFE-Untertests zeigt, dass sich die Schüler*innen der Experimentalgruppe im Zeitverlauf nicht signifikant von der Kontrollgruppe unterscheiden. Die Ergebnisse der festen Effekte zeigen, dass eine höhere Rechtschreibleistung auch zu höherer Leseleistung führt und dass die zu Hause gesprochene Sprache keinen signifikanten Einfluss auf die Veränderung der Leseleistung hat.

Um zu klären, welche Faktoren dazu geführt haben, dass das Training nicht die intendierte Wirkung erzielt hat, sind weitere Analysen vorgesehen. Dabei werden insbesondere die inter-individuelle Intensität der Trainings sowie die Lesemotivation der Schüler*innen berücksichtigt und die Veränderung der Leseleistung bei besonders schwachen Schüler*innen gesondert analysiert. Diese Analysen werden ebenfalls im Vortrag berichtet und mögliche Anpassungen des Trainings und zusätzliche unterstützende Maßnahmen für die zukünftige Umsetzung diskutiert.

Konstruktvalidität von Wissen und Selbstwirksamkeit zur datengestützten Unterrichtsentwicklung: Dimensionalität, Veränderungssensitivität und Zusammenhänge mit der intendierten Datennutzung bei Lehrkräften

Patrick Hawlitschek, Sofie Henschel

Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin e.V.

Theoretischer Hintergrund:

Datengestützte Unterrichtsentwicklung (DGUE) umfasst den systematischen Einbezug verschiedener Datenquellen (z. B. aus Kompetenztests, Unterrichtsbeobachtungen, Schulstatistik) in die Unterrichtsgestaltung (Mandinach & Schildkamp, 2021), wovon insbesondere Schüler*innen aus Schulen herausfordernden Lagen profitieren können (Dietrichson et al., 2017). Das idealtypische Vorgehen folgt einem zyklischen Entwicklungskreislauf aus Analyse und Interpretation von schulischen Daten, der darauf bezogenen Maßnahmenplanung und -umsetzung sowie der Erfolgsüberprüfung. Dabei handelt es sich um einen voraussetzungsreichen Prozess, der gezielt unterstützt werden muss. Ansatzpunkte zur Unterstützung von Lehrkräften sind Merkmalsbereiche, die sich als prädiktiv für eine erfolgreiche Datennutzung erwiesen haben (Schildkamp et al., 2017). Hierzu zählen neben Merkmalen der Daten und der Schule (z. B. zeitnahe Rückmeldung, Schulleitungshandeln) vor allem individuelle Lehrkraftmerkmale (z. B. Wissen, Überzeugungen wie Selbstwirksamkeit), die im Fokus dieser Studie stehen. Unter letzteren gelten einerseits eine gut vernetzte kognitive Wissensbasis sowie andererseits positiv ausgeprägte Selbstwirksamkeitsüberzeugungen als wesentlich für die Umsetzung datengestützter Unterrichtsentwicklungsprozesse im schulischen Alltag (Hooglund et al., 2016). Auf die Stärkung beider Aspekte sollten deshalb auch Fortbildungen abzielen. In bisherigen Studien wurde häufig das selbstberichtete Wissen erfasst, wobei ähnliche Operationalisierungen entweder als kognitiver Indikator für professionelles Wissen (z. B. Michos et al., 2022) oder als motivationaler Indikator für Selbstwirksamkeitsüberzeugungen (z.B. Dunn et al., 2013) genutzt wurden. Bislang liegen keine Studien vor, in denen sowohl selbstberichtetes Wissen als auch ein stärker objektives (testbasiertes) Wissen und die Selbstwirksamkeit bezogen auf DGUE gemeinsam erfasst wurden. Folglich ist unklar, inwieweit die Konstrukte empirisch voneinander abgrenzbar sind. Zudem ist unklar, inwieweit die Konstrukte einerseits differenziell veränderungssensitiv und andererseits mit dem intendierten Datennutzungsverhalten als einem transfer-relevanten Verhaltensmerkmal differenziell zusammenhängen. Daher untersuchen wir die Konstruktvalidität des Wissens und der Selbstwirksamkeit:

Fragestellungen:

1. Lässt sich das selbstberichtete Wissen über Datennutzung von Lehrkräften empirisch vom testbasiert erfassten Wissen und der Selbstwirksamkeit unterscheiden?
2. Inwieweit wirkt sich eine Blended-Learning-Fortbildung zur Datennutzung gleichermaßen positiv auf alle drei Konstrukte im Vergleich zu einer Kontrollgruppe aus?
3. Hängen die drei Konstrukte differenziell mit der intendierten Datennutzung nach Abschluss der Fortbildung zusammen?

Methode:

Den Analysen liegen Daten einer quasi-experimentellen Feldstudie mit Prä-Post-Kontrollgruppendesign zugrunde, bei der Lehrkräfte der Interventionsgruppe (IG; $n_{IG}=78$, Berufserfahrung: $M=13.48$, $SD=8.97$, 69% weiblich, 37% Primarstufe) für die DGUE mit den bundesweiten Vergleichsarbeiten (VERA) in einem Blended-Learning Format (insg. 8 Stunden mit 3 E-Learning Einheiten im Selbststudium, 2 Online-Workshops) fortgebildet wurden, während Lehrkräfte der (Warte-)Kontrollgruppe (KG; $n_{KG}=38$, Berufserfahrung: $M=16.69$, $SD=8.90$, 74% weiblich, 50% Primarstufe) erst nach Abschluss des Posttests an der Fortbildung teilnahmen. Alle Lehrkräfte bearbeiteten im Prätest und Posttest einen Wissenstest (13 Items; $\omega_{Prä/Post}=.72/.67$) und Skalen zum selbstberichteten Wissen (4 Items; $\omega_{Prä/Post}=.84/.81$), zur Selbstwirksamkeit (5 Items; $\omega_{Prä/Post}=.80/.82$) sowie zur intendierten Datennutzung (4 Items; $\omega_{Post}=.89$; nur IG). Fehlende Werte (8-35% pro Variable) wurden unter Berücksichtigung von Hintergrundmerkmalen multipl imputiert ($k=35$ Datensätze). Zur Analyse der ersten Fragestellung wurden latente Dimensionsanalysen und für die Fragestellungen 2 und 3 Pfadanalysen mit R durchgeführt.

Ergebnisse und ihre Bedeutung:

Vergleichende konfirmatorische Faktoranalysen zeigen, dass ein dreidimensionales Modell ($.42 \leq r \leq .47$), welches zwischen testbasiertem und selbstberichtetem Wissen sowie Selbstwirksamkeit differenziert, die beste Modellanpassung liefert (Fragestellung 1). In separaten Pfadanalysen zeigten sich unter Kontrolle der Prätestwerte signifikante Interventionseffekte zugunsten der IG für alle drei Outcomes, wobei in einem dreidimensionalen Modell mit korrelierten Outcomes nur das selbsteingeschätzte Wissen in der IG im Vergleich zur KG höher ausfiel (Fragestellung 2). Gleichzeitig hing in der IG nur die Selbstwirksamkeit unter Kontrolle der Wissensfacetten mit der intendierten Datennutzung im Posttest zusammen (Fragestellung

3). Die Ergebnisse zeigen, dass sich die drei Konstrukte empirisch trennen lassen und unter gegenseitiger Kontrolle differenziell durch die Fortbildung verändert werden sowie mit dem intendierten Datennutzungsverhalten zusammenhängen. Daher scheint es angebracht, die dreidimensionale Struktur in der Entwicklung von Unterstützungsangeboten zur DGUE zu berücksichtigen.

Lehrkräfte an Schule in herausfordernden Lagen: Wer nimmt an Online-Angeboten zur Stressreduktion teil und für wen sind sie wirksam?

Uta Klusmann¹, Dirk Richter², Lisa Benckwitz³, Anna Hilz³, Tabea Schulze-Hagenest³

¹Humboldt-Universität zu Berlin & IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik,

²Universität Potsdam, ³IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Theoretischer Hintergrund:

Das berufliche Wohlbefinden von Lehrkräften, welches Arbeitszufriedenheit, positive Emotionen und reduzierte negative Emotionen sowie Stress oder Erschöpfung umfasst, ist ein Thema, das in wissenschaftlichen Kreisen, in der Praxis und in der Bildungsadministration intensiv diskutiert wird (Hascher & Waber, 2021). Geringes berufliches Wohlbefinden wird mit erhöhten Fehlzeiten, einer höheren Fluktuationsrate und einer verminderten Qualität der Lehrer-Schüler-Interaktion in Verbindung gebracht (z. B. Wartenberg et al., 2023). Infolgedessen wächst die Besorgnis über das hohe Risiko von Burnout und Stress bei Lehrkräften (Karakus et al., 2024), insbesondere bei Lehrkräften, die in Schulen in sozioökonomisch benachteiligten Lagen arbeiten, in denen Stress- und Burnout-Raten tendenziell höher sind. Aufgrund des Mangels an repräsentativen Daten gibt es jedoch nur wenige empirische Belege zu diesem Thema. Dennoch ist es plausibel anzunehmen, dass Lehrkräfte in sozioökonomisch benachteiligten Gebieten mit zahlreichen arbeitsbezogenen Stressfaktoren konfrontiert sind, wie beispielsweise einer hohen Prävalenz von Schüler*innen mit sozio-emotionalen und motivationalen Problemen, Störungen des Unterrichts und aggressivem Verhalten (Bradley & Corwyn, 2002). Angesichts der entscheidenden Rolle, die Lehrkräfte bei der Entwicklung sozioökonomisch benachteiligter Schüler*innen spielen, könnte ein vermindertes Wohlbefinden bei Lehrkräften daher eine unvorteilhafte Abwärtsspirale der Interaktion zwischen Lehrkräften und Schüler*innen begünstigen.

Frühere Forschung konnte jedoch gezeigt, dass es wirksame Interventionen zur Verringerung von Stress und Burnout-Symptomen bei Lehrkräften gibt (Iancu et al., 2018). Die erfolgreichsten Interventionen beinhalten typischerweise kognitiv-behaviorale oder achtsamkeitsorientierte Ansätze (Iancu et al., 2018). Die Teilnahme an solchen Interventionsprogrammen erfordert jedoch einen erheblichen Zeitaufwand, der von Personen mit geringem Wohlbefinden als nicht verfügbar empfunden werden kann, wodurch die Teilnahme möglicherweise auf Lehrkräfte mit einem bestimmten Grundniveau an Wohlbefinden beschränkt und andere ausgeschlossen werden.

In diesem Beitrag sollen die folgenden Forschungsfragen beantwortet werden:

Fragestellungen:

1) Auf welchem Niveau des beruflichen Wohlbefindens entscheiden sich Lehrkräfte an Schulen für die Teilnahme an einem Online-Training zur Stressbewältigung?

2) Für wen ist die Teilnahme an einem Online-Stress-Interventionsprogramm am wirksamsten?

Methode:

Im Sommer/Herbst 2021 nahmen 2923 Lehrkräfte aus 196 Schulen in sozioökonomisch benachteiligten Lagen an einer Online-Umfrage teil, wobei Angaben zur emotionalen Erschöpfung und Zynismus, zur Arbeitszufriedenheit sowie zum positiven und negativen Affekt gemacht wurden.

Von den befragten Lehrkräften meldeten sich 229 für eine Online-Intervention zur Stressbewältigung (4 Sitzungen) an. Die Intervention bot ein kognitiv-behaviorales Stressmanagement-Training, angereichert mit Achtsamkeitsansätzen, in Gruppen von 5-15 Teilnehmer*innen mit geschulten Trainer*innen. Bei der Studie handelt es sich um ein Prä-Post-Kontrollgruppen-Design mit 175 Lehrkräften in der Interventionsgruppe und 54 Lehrkräften in der Wartekontrollgruppe. Für die Mittelwertvergleiche wurden außerdem auch Daten aus einer repräsentativen Stichprobe von Lehrkräften in Deutschland mit einbezogen.

Ergebnisse und ihre Bedeutung:

Zunächst sollte beantwortet werden, auf welchem Niveau des beruflichen Wohlbefindens sich Lehrkräfte für die Teilnahme an einer Intervention entscheiden. Die Ergebnisse der Mittelwertvergleiche zeigten, dass die Lehrkräfte, die sich für die Intervention anmeldeten, ein ähnliches Maß an emotionaler Erschöpfung und Arbeitszufriedenheit angaben, wie die Gesamtstichprobe der Lehrkräfte im Programm und auch die repräsentative deutsche Stichprobe. Allerdings waren die Lehrkräfte in der angemeldeten Intervention deutlich weniger zynisch als die Gesamt- und die repräsentative Stichprobe.

Zweitens zeigten sich statistisch signifikante positive Auswirkungen der Intervention auf den positiven Affekt. Für die emotionale Erschöpfung, die Arbeitszufriedenheit und den negativen Affekt konnte jedoch kein signifikanter Effekt der Intervention festgestellt werden. Für den Zynismus stellten wir fest, dass Lehrkräfte mit höheren Ausgangswerten signifikant von der Intervention profitieren.

Die Ergebnisse zeigen, dass Interventionsprogramme nicht unbedingt die besonders erschöpften oder unzufriedenen Lehrer erreichen. Ein hohes Maß an Zynismus stellte jedoch eher ein Hindernis für die Teilnahme dar. Die Intervention selbst war nur bei der Steigerung des positiven Affekts und bei Lehrkräften mit hohem Zynismus wirksam. Dies deutet darauf hin, dass das Interventionsprogramm Lehrkräften helfen könnte, ihren Zynismus zu überwinden und ihren positiven Affekt zu verbessern.

Wie gestaltet sich Fürsorge in SchuMaS-Schulen? Einblicke in armutsbezogene Fürsorge-Interaktionen zwischen Lehrer*innen und Schüler*innen

Karina Karst¹, Oscar Yendell¹, Carolina Claus²

¹Universität Mannheim, ²Humboldt-Universität zu Berlin

Theoretischer Hintergrund:

In Deutschland empfangen knapp zwei Millionen Kinder Transferleistungen (Bürgergeld), deren Lebenssituation folglich als „arm“ bezeichnet werden kann (Bundesagentur für Arbeit, 2023; Goebel & Krause, 2018). Zur Finanzierung schulischer Bedarfe (bspw. Schulmaterialien oder Mittagessen) können Schüler*innen im Transferleistungsbezug zusätzlich „Leistungen für Bildung und Teilhabe“ beantragen (Kaps & Marquardsen, 2017). Einem relationalen Verständnis von Armut folgend, führt der Empfang entsprechender Leistungen zu einem „empfangenden“ Verhältnis gegenüber einer „gebenden“ Gesellschaft und einer entsprechend öffentlichen Zuschreibung als „arm“ (Fritsch & Verwiebe, 2018). Praktisch zeigt sich dieses Verhältnis in Schule, indem Lehrer*innen armutsbetroffene Schüler*innen und ihre Familien in der Finanzierung beispielsweise für Mittagsverpflegung

oder Schulmaterialien unterstützen, wodurch die Armut der Schüler*innen in die Interaktionen zwischen Lehrer*innen und Schüler*innen hineingetragen wird (Kaluza & Schimnek, 2023; Steiner et al., 2024). Die innerschulische Prävention von Armutsfolgen ist jedoch nicht alleinige Aufgabe von Lehrer*innen, vielmehr werden auch weitere Professionen adressiert (bspw. Schulsozialarbeit), wodurch ein (idealtypisch) multiprofessionelles Vorgehen entsteht (Gigerl & Breser, 2024; Steiner et al., 2024). Theoretisch können Interaktionen zur Prävention von Armutsfolgen als individuelle fürsorgliche Praktiken gefasst werden, indem Lehrer*innen auf die Angewiesenheit entsprechender Schüler*innen reagieren (Dietrich, 2024). Diese Praktiken umfassen vier Phasen: 1. Die Wahrnehmung eines Fürsorgebedarfs, 2. Die Übernahme von Verantwortung für Fürsorge, 3. Konkrete Fürsorge-Interaktionen, 4. Die Reaktion der Fürsorge empfangenden Person (Tronto, nach Dietrich, 2024). Bisher ist jedoch unbeleuchtet, wie sich konkrete Fürsorge-Interaktionen zwischen Lehrer*innen und Schüler*innen in Bezug auf den Bürgergeldbezug von Schüler*innen gestalten und somit zu den klassisch unterrichtsbezogenen und pädagogischen Tätigkeiten hinzustoßen. Gleichzeitig ist unklar, ob und wie Lehrer*innen sich für entsprechende Tätigkeiten überhaupt in der Verantwortung sehen.

Fragestellung:

In einer ersten Fragestellung wird daher untersucht, wie sich aus der Perspektive von Lehrer*innen Bürgergeld-bezogene Fürsorge-Interaktionen zwischen Lehrer*innen und Schüler*innen gestalten. Damit verbunden wird der Frage nachgegangen, wie Lehrer*innen sich für diese Fürsorge in der Verantwortung sehen.

Methode:

Zur Beantwortung der Forschungsfragen wurden von Juni bis November 2023 acht Fokusgruppen mit jeweils 3-4 Lehrer*innen derselben Schule geführt (Barbour, 2018). Die Fokusgruppen wurden an drei weiterführenden SchuMaS-Schulen (ein Gymnasium, zwei Gemeinschaftsschulen) in verschiedenen Großstädten geführt, an denen der Anteil an Schüler*innen mit Transferleistungsbezug zwischen 30-85% variierte. Die Fokusgruppen wurden nach der Grounded Theory ausgewertet, indem beginnend offen in den einzelnen Fokusgruppen, anschließend axial und fokusgruppenübergreifend und abschließend selektiv kodiert wurde (Strauss & Corbin, 1998).

Ergebnisse und ihre Bedeutung:

Es wird gezeigt, dass Lehrer*innen bürokratische Fürsorge-Interaktionen zwischen Schüler*innen, Familien und Behörden (Jobcenter) gestalten, indem sie Schüler*innen und deren Familien im Bürgergeldbezug in der Beantragung von Mitteln für „Bildung und Teilhabe“ unterstützen (bspw. zur Finanzierung von Nachhilfe oder Mittagsverpflegung). Diese Interaktionen variieren zwischen den Schulen, indem es in einigen Schulen konkrete Lehrer*innen als Ansprechpartner*innen oder im Kollegium abgesprochene (digitalisierte) Abläufe gibt, während Abläufe und Zuständigkeiten in anderen Schulen ungeklärt bleiben. Wenngleich alle Lehrer*innen einen entsprechenden Fürsorgebedarf wahrnehmen, nehmen sie eine unterschiedliche Verantwortung für eben jene wahr. Einige Lehrer*innen dienen für Schüler*innen und ihre Familien als erste Ansprechpartner*innen für bildungsbezogene behördliche Fragestellungen und unterstützen sie in der Finanzierung schulischer Bedarfe, die ungenügend über „Bildung und Teilhabe“ abgedeckt werden. Andere Lehrer*innen weisen diese Verantwortung von sich und argumentieren, dass eine entsprechende Unterstützung belastend sowie Aufgabe von Behörden oder Schulsozialarbeit sei. Durch die unterschiedlichen Verantwortungsübernahmen gestalten sich nicht nur zwischen den Schulen, sondern auch innerschulisch Fürsorge-Interaktionen unterschiedlich.

Die Ergebnisse werden vor dem Hintergrund diskutiert, wie sich die unterschiedliche Verantwortungsübernahme der Lehrer*innen auf die vier Phasen von Fürsorge-Interaktionen auswirkt (Dietrich, 2024). Zudem wird diskutiert, inwiefern ein multiprofessionelles Vorgehen zu einer kohärenten innerschulischen Prävention von Armutsfolgen beitragen kann (Gigerl & Breser, 2024; Steiner et al., 2024).

S-95: Educational Neuroscience: Bridging Brain Mechanisms and Learning Practices

Zeit: Mittwoch, 29.01.2025: 13:15 - 15:00 · Ort: EO 154

Chair der Sitzung: Sabine Oligschläger

Chair der Sitzung: Michiko Murayama

Chair der Sitzung: Kou Murayama

English Track

Educational Neuroscience: Bridging Brain Mechanisms and Learning Practices

Chair(s): **Sabine Oligschläger** (Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany), **Michiko Sakaki** (Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany; Research Institute, Kochi University of Technology, Kochi, Japan)

Diskutant*in(nen): **Kou Murayama** (Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany)

Over the past decade, a lively debate has persisted about the extent to which neuroscience can contribute to educational research. Learning in the classroom is a complex process affected by the school system, teachers, peers, parents, and learning materials and it seems impossible to investigate such complex and multifaceted phenomena in a controlled laboratory setting with neuroimaging techniques. At the same time, there is no doubt that learning is at least partly affected by brain mechanisms. Thus, neuroscience research has the potential to provide novel insights into mechanisms underlying learning that psychological methods do not easily provide.

The purpose of this symposium is to revisit this conversation by showcasing a series of exemplary studies from the field of educational neuroscience. We will feature four contributions, each illustrating the use of different neuroimaging techniques to investigate a variety of cognitive and affective domains relevant to learning. We aim to provide the education science community with an exemplary overview of the types of research being conducted in the field of educational neuroscience and to illustrate the potential of these techniques to complement and enhance traditional educational research.

The first presentation will feature functional magnetic resonance imaging (fMRI) to investigate the role of the reward network in the brain (traditionally associated with extrinsic reward/motivation) for intrinsic motivation to engage in difficult, even intractable tasks. This presentation provides a unique perspective on how similarly intrinsic and extrinsic motivations are supported in the brain, despite their different implications for learning. The second contribution focuses on metacognition. The presentation will discuss brain mechanisms of metacognition and their relationship to arithmetic ability using electroencephalography (EEG), a portable neuroimaging technique. This research provides insights into the dynamics of metacognitive processes that existing methods, such as self-report or think-aloud techniques, would not be able to easily examine. The third presentation focuses on individual differences in mathematical abilities. A series of studies will be presented that have shed light on how children and adults solve difficult math problems differently. The presentation will also highlight another method, functional near-infrared spectroscopy (fNIRS), a portable technique for measuring brain activity from hemodynamic activity using near-infrared light which is particularly suitable for children. Finally, the fourth contribution uses fMRI to investigate how individual differences in the spatial layout of cortical connections are related to general cognitive ability. This research will highlight the importance of individual differences in brain network architecture for understanding the mechanisms underlying general cognitive abilities.

Taken together, this symposium seeks to continue the dialogue on the role of neuroscience in education, and how interdisciplinary approaches can complement our understanding of learning processes.

Beiträge des Symposiums

Role of the reward network in the brain for learners' motivation

Michiko Sakaki

Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany; Research Institute, Kochi University of Technology, Kochi, Japan

Previous research in neuroscience suggests that the reward network in the brain is important for motivation for primary and secondary rewards, such as food and money. Such extrinsic rewards are known to undermine learning and motivation in educational psychology (Hidi, 2016 for a review). However, research in the past decade suggests that the same reward network in the brain can play a critical role even in other types of motivations that are crucial for learning (e.g., intrinsic motivation). This presentation highlights recent research on the role of the reward network in the brain in human motivation.

The first part of the presentation focuses on motivation for challenges (Sakaki et al., 2023). Research in educational psychology suggests that it is important for students to approach challenges, become resilient against setbacks and failures, and sustain effort in an achievement setting over time. To examine the conditions under which people express motivation for difficult and nearly impossible challenges, we asked participants to play a game-like task (i.e., a stop-watch task) in a magnetic resonance imaging (MRI) scanner with and without performance-contingent monetary rewards. Participants (irrespective of the conditions) were asked to play three versions of the task. The first one was an extremely easy task, where participants were able to almost always succeed. The second version was an intermediate level of difficulty, where participants were able to succeed on approx. 50% of trials. The final version was an extremely difficult task, in which they were never able to succeed. When participants played the task without extrinsic incentives, their task enjoyment increased as the chance of success decreased, even if the task was almost impossible to succeed. In contrast, such challenge-based motivation was not observed when participants received monetary rewards for the task. The activation in the reward network also tracked the pattern of task enjoyment.

The second part of the presentation focuses on approach goals (i.e., goals to obtain desirable outcomes) vs. avoidance goals (i.e., goals to avoid the loss of desirable outcomes; Sakaki et al., 2024). Previous field studies repeatedly suggest that approach goals facilitate learners' intrinsic motivation, while avoidance goals induce anxiety. To examine the role of the reward network of the brain, we asked participants to play the stop-watch task with an approach goal frame (i.e., "you would gain points when you succeed") vs. an avoidance goal frame (i.e., "you would lose points when you fail") in an MRI scanner. Participants reported stronger positive emotions when they focused on approach goals, but stronger anxiety and disappointment when they focused on avoidance goals, replicating previous findings from field studies. In addition, the activation pattern in the striatum was different depending on the goal frame, suggesting that the striatum is differently involved in the same task depending on the goal frames.

Furthermore, the distinct patterns of activation in the striatum were associated with distinct subjective experiences participants reported between the approach and the avoidance conditions.

Across these studies, our results suggest that people are intrinsically motivated to challenge a nearly impossible task and enjoy the same task under the approach goal frame more than the avoidance goal frame. The results also suggest that the reward network in the brain is important for these seemingly higher-order motivations for humans, such as intrinsic motivations for challenges and enjoyment for approach over avoidance goals.

Metacognitive monitoring and its link to mathematical task performance

Stephan E. Vogel

Educational Neuroscience, Institute of Psychology, University of Graz, Austria

A growing body of evidence indicates that there is a significant relationship between mathematical task performance and metacognitive regulation (Beelon et al., 2019). Metacognitive regulation is often defined as a collection of self-reflective, higher-order cognitive resources that monitor (i.e., detect) and control (i.e., adapt) mental activities during task performance (Nelson, 1990). While monitoring relates to the subjective self-assessment of how well a task will be/is/has been performed, control refers to activities that enable the adaptation of different cognitive operations (e.g., change in strategy use) to increase task performance and subsequent learning (Roebers et al., 2014). In this talk, I will present two studies, one behavioral and one using electroencephalography (EEG), that aimed to better understand the association of metacognitive monitoring in the context of conceptual knowledge and arithmetic task performance.

Metacognitive monitoring is important in tasks in which naïve concepts or calculations interfere with scientific knowledge (Stricker et al., 2022). Awareness of this conflict is crucial for inhibiting interfering concepts and selecting the appropriate answer. The main aim of the first study was to behaviorally test whether metacognitive monitoring is linked to task performance and how it is moderated by mathematical abilities. We collected behavioral data from 82 adults who verified the correctness of mathematical statements on a computer. Half of the statements were incongruent (naïve and scientific concepts were in conflict), the other half were congruent (concepts were nonconflicting). After verifying each statement, participants rated their response confidence on a continuous scale. Mathematical competence was measured with the “Mathematics test for selection of personnel (M-PA)”. Results demonstrated that participants were slower and more error prone when naïve concepts conflicted with scientific concepts (i.e., incongruent condition). Participants also showed significantly lower metacognitive calibration for incongruent compared to congruent statements. In addition, the accuracy of metacognitive calibration was moderated by mathematical abilities such that individuals with higher mathematical competence demonstrated more accurate calibration. These findings indicate a significant role of metacognitive monitoring during conflict resolution in mathematics.

Despite its importance, little is currently known about the neurocognitive basis of metacognitive monitoring in mathematical tasks. Previous electrophysiological (EEG) research has demonstrated that error-related potentials, i.e., error-related negativity (Ne) and error-related positivity (Pe), are modulated by metacognitive monitoring (Boldt et al., 2015). The main aim of the second study was to examine whether the Ne and Pe are also indicative of metacognitive monitoring during a mathematical task and whether it is modulated by mathematical abilities. We recorded EEG data (64 active electrodes/10-20 system) from 39 adults (aged 18-35 years) who performed an “arithmetic lure task” (Heidekum et al., 2019). Participants verified the correctness of 540 multiplications on a computer screen and rated their decision confidence on a 6-point Likert scale. Participants also completed a paper-pencil arithmetic fluency task to measure arithmetic abilities. EEG analyses focused on frontal-to-parietal midline electrodes where the Ne and Pe typically occur. The results showed that the Pe amplitude correlated significantly with subjects’ confidence ratings. An additional correlation analysis indicates a significant, but small, association between the Pe and arithmetic fluency. These findings imply a strong association between the Pe, often associated with explicit error detection, and metacognitive monitoring during arithmetic problem solving.

Overall, this work demonstrates a significant link between metacognitive monitoring and mathematical task performance and shows that neural signatures of metacognitive monitoring appear 200-400 ms after a response. Further investigation of these signatures in the context of education and mathematical learning is essential.

Educational Neuroscience of Mathematics

Christina Artemenko

Department of Psychology, University of Tuebingen, Germany; LEAD Graduate School & Research Network, University of Tuebingen, Germany

Educational Neuroscience builds a bridge from neuroscience over cognitive, developmental, and educational psychology to educational science and practice. Thereby, such research has the potential to not only identify that there are problems but also to understand why these problems occur. In the domain of mathematics, my research addresses the questions (1) what makes mathematics difficult, (2) why some people have more problems in math than others, and (3) how children learn mathematics. As a neuroimaging method, I use functional near-infrared spectroscopy (fNIRS) in addition to behavioral measures, because it is appropriate for children due to motion tolerance and for studying mathematics in natural settings due to its ecological validity.

(1) The difficulty of mathematics increases, when calculating with larger numbers (problem size effect) or when carry or borrow operations in arithmetic are needed. The carry operation in addition (carrying a decade when the sum of the units exceeds 9) and the borrow operation in subtraction (borrowing a decade when the unit of the minuend is smaller than the unit of the subtrahend) reflect major obstacles in math education. Behaviorally, higher problem sizes as well as carry and borrow operations increase the difficulty of arithmetic, as they increase solution times and are more error prone (Artemenko et al., 2018a, 2019, 2021).

(2) Interestingly, the carry and borrow effects as well as problem size effects are even larger in individuals with low math abilities, suggesting that they have particular problems in more complex math tasks (Artemenko et al., 2018a, 2019). Moreover, individuals with high math ability performed complex arithmetic by using resources in the brain, which individuals with low math ability needed even for simple arithmetic (Artemenko et al., 2018a, 2019). These findings point at interindividual differences in how individuals deal with arithmetic complexity.

(3) During mathematical development, a functional specialization of mathematics takes place in the brain, while domain-general resources are needed less for arithmetic (Artemenko et al., 2018b, 2021). During arithmetic learning, typically developing children get better and also use less brain resources for processing complex arithmetic (Soltanlou et al., 2017, 2018). This shows that children learn differently than adults. Moreover, even children with mathematical disabilities learn differently than typically developing children (Soltanlou et al., 2022). Finally, during initial learning of mathematics, the systematic use of fingers leaves traces in the brain, supporting the approach of embodied cognition (Artemenko et al., 2022).

Taken together, my research focuses on the carry and borrow effects as well as the problem size effect that increase mathematical difficulty, and on the inter- and intraindividual differences in these effects during development. These findings contribute to educational neuroscience and provide unique insights that might inform educational practice.

Cognitive abilities are associated with shorter connectivity in the cerebral cortex of the brain

Sabine Oligschläger

Hector Research Institute of Education Sciences and Psychology, University of Tübingen, Tübingen, Germany

Cognitive abilities carry profound implications for life trajectories. For example, measures of general intelligence, commonly expressed as the "g" factor, reflect the psychological trait with the greatest predictive validity for real-life outcomes (Gottfredson, 1997), including educational achievement in school (Johnson et al., 2006), academia, and career (Jensen, 1993). Understanding how individual differences in cognitive abilities arise from differences in brain organization remains poorly understood, yet holds a potential for education science. By investigating the mechanisms by which general cognitive abilities affect learning-related processing, we may better identify opportunities for cognitive growth and tailor educational materials to individual processing styles. This is particularly important for specific groups, such as gifted children with exceptionally high intelligence, where tailored learning approaches may optimize their educational experiences and outcomes.

Recent trends in neuroimaging have highlighted interindividual differences in the spatial organization of the cerebral cortex of the brain, revealing differences in the way functions and connections are spatially arranged within the cortical layout (Feilong et al., 2023; Haxby et al., 2020). This line of research has produced early evidence that differences in this spatial organization may play a role in cognitive abilities (Cui et al., 2020; Feilong et al., 2021; Keller et al., 2022).

From a spatial perspective, most cortical regions are strongly connected to their neighbors, with fewer and fewer connections to regions at greater distances (Ercsey-Ravasz et al., 2013). It has been proposed that these long-range connections, despite being costly, provide benefits to brain function (Bullmore & Sporns, 2012). Based on this proposition, an intuitive assumption would be that individuals with more long-range connections would have better cognitive abilities. However, just the opposite pattern was found in a study that compared highly intelligent adolescents with adolescents and adults of average intelligence. Overall fewer rather than longer connections were associated with high intelligence (Ma et al., 2017).

Given the counterintuitiveness of these findings, as well as some methodological limitations (e.g., small sample size), such findings warrant closer examination. The research question of the present work is therefore: How do individual differences in the spatial extent of cortical connectivity relate to individual differences in cognitive ability?

Here, we used functional magnetic resonance imaging (fMRI) to measure time-course correlations of activity between cortical regions, a measure referred to as functional connectivity. We chose this measure because of its higher spatial resolution to resolve connectivity between nearby regions. We used anatomical reconstructions of the cortical surface to measure the average distance of a region from the areas to which it is connected. For each cortical location, we assessed how the variance in connectivity distance across individuals relates to variances in performance in executive control and general cognitive ability.

Throughout the cortex, we found clusters for which interindividual differences in connectivity distance were associated with those in cognitive test performance. For most clusters, the association was negative, meaning that shorter connectivity distances predicted better performance. Thus, for specific areas, the shorter the connections in an individual, the better the cognitive performance, suggesting that better cognitive abilities do not rely on a greater number of long but on short connections.

These results suggest that individual differences in cognitive abilities may be associated with individual cortical organizations that favor shorter rather than longer connections. Such findings help us to better understand the neural origins of individual differences in cognitive abilities and learning-related processes.

IS-09: Invited Symposium: Digitale Souveränität als Schlüssel zur Bewältigung gesellschaftlicher Herausforderungen - Beiträge der politischen Bildung

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 102

Chair der Sitzung: Monika Oberle

Chair der Sitzung: Johannes Hartig

Digitale Souveränität als Schlüssel zur Bewältigung gesellschaftlicher Herausforderungen - Beiträge der politischen Bildung

Chair(s): **Monika Oberle** (Goethe-Universität Frankfurt, Deutschland), **Johannes Hartig** (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation)

Angesichts eines digitalen Strukturwandels der demokratischen Öffentlichkeit ist die digitale Souveränität der Bürger:innen für freiheitliche Demokratien heute eine zentrale Funktionsvoraussetzung. So hat das Internet zwar die Zugänglichkeit von Informationen enorm gesteigert, zugleich aber auch zu einer wachsenden Flut an ungeordneter Information geführt. Nachrichten können nun von allen Bürger:innen permanent produziert, reproduziert, verbreitet und kommentiert werden, und auch Politiker:innen und Regierende kommunizieren via sozialer Medien direkt mit ihren Followern. Die klassischen journalistischen Massenmedien haben ihre gatekeeper-Funktion in weiten Teilen eingebüßt, in der sie früher die Verifizierung und Einordnung von Fakten und Ereignissen übernahmen. Dies stellt Bürger*innen vor große Herausforderungen hinsichtlich der Selektion, Verifikation und Einordnung von Informationen – benötigt werden entsprechende Orientierungskompetenzen. Zugleich bieten sich in der e-democracy neue gesellschaftliche und politische Beteiligungsmöglichkeiten, deren Nutzung allerdings oftmals von sozialen Ungleichheiten gekennzeichnet ist. Fake News, hate speech, Verlust an Datensouveränität und algorithmische Diskriminierung sind Stichworte der großen gesellschaftlichen Herausforderungen, für deren Bewältigung im privaten und politischen Handeln Bürger:innen auch durch politische Bildung befähigt werden müssen. Angesichts sozialer Ungleichheiten sind dabei auch kompensatorische Effekte zu anderen Instanzen der politischen Sozialisation relevant. (Gapski et al. 2017; Hubacher/Waldis 2021; Oberle 2022).

Von der politischen Bildung wird erwartet, einen Beitrag zur Bewältigung dieser gesellschaftlichen Herausforderungen zu leisten. Dem steht ein eklatanter Mangel an empirischen Erkenntnissen über Schulunterricht und Lehrkräftebildung in diesem Feld gegenüber. Es fehlt an empirischen Befunden dazu, wie politische Bildung effektiv zur Förderung einer digitalen Souveränität beitragen kann und wie es um die Praxis einer „politischen Bildung in der digitalen Welt“ (Oberle/Heldt 2022) in Schule und Lehrkräftebildung heute bestellt ist. Dieses Symposium möchte anhand von Ergebnissen aktueller empirischer Studien die Ausgangsbedingungen auf Seiten der Lernenden und Lehrenden, die aktuelle Unterrichtspraxis sowie die Bedarfe an und Potenziale von Lehrkräftefortbildungen aus fachdidaktischen sowie bildungs- und kulturwissenschaftlichen Perspektiven in den Blick nehmen.

Vorgestellt werden zwei qualitative und zwei quantitative Studien, die auf Schüler:innen, Lehramtsstudierende und Fachlehrkräfte fokussieren. Die Beiträge untersuchen digitale Partizipationserfahrungen und Bildungserwartungen von Jugendlichen in der Schweiz (Waldis/Balsiger), analysieren die fachlichen und fachdidaktischen Vorstellungen von Digitalisierung bei angehenden Politiklehrkräften in Rheinland-Pfalz (Heldt/Theophil), geben Einblick in die aktuelle schulische Praxis politischer Bildung zur Förderung einer digital literacy in Nordrhein-Westfalen (Abs/Deimel/Welsand) und beleuchten die Potenziale von Online-Fortbildungen für Politiklehrkräfte zur politischen Medienbildung (Oberle). Dabei werden theoretische Modelle von digital citizenship, digitaler Souveränität und politischer Medienkompetenz von Lehrkräften und Schüler:innen mit empirischen Befunden verbunden, um sie empirisch zu fundieren, aber auch kritisch weiterzuentwickeln.

Beiträge des Symposiums

Digital citizenship bei Jugendlichen – Online-Praktiken einer vernetzten Bürgerschaftlichkeit

Monika Waldis¹, Mattia Balsiger²

¹PH FHNW, ²Zentrum für Demokratie Aarau

Hintergrund

Digitale Medien schaffen eine erweiterte soziale und politische Realität. Heranwachsende können durch die Teilnahme an digitalen Prozessen aktiv mitwirken oder passiv-beobachtend gesellschaftliche Diskurse mitverfolgen. In demokratietheoretischer Sicht zielt Partizipation auf die Beteiligung möglichst vieler zur Teilhabe am Gemeinwesen und seiner Ausgestaltung ab. In der Folge umfasst Onlinepartizipation netzbasierte Aktivitäten zur Meinungsbildung und Mitsprache, wie etwa die Suche nach politischer Information, das Followen von politisch engagierten Personen oder Politiker*innen, das Liken, Teilen, Posten oder Kommentieren von Beiträgen in den Sozialen Medien oder die Nutzung digitaler Civic Tech Tools (Theocharis & van Deth, 2018). Ob Jugendliche im digitalen Raum eher passiv-konsumierend unterwegs sind, Formen der Mikro-partizipation praktizieren oder aktiv politische Ziele verfolgen: Heranwachsende gestalten mit ihrem Netzhandeln digitale Inhalte, Aktivitäten und Strukturen mit. Auf diese Weise sind sie an den aktuellen sozio-technischen Entwicklungen und der Entstehung digitaler Öffentlichkeiten beteiligt, die das Verständnis von und die Anforderungen an Bürgerschaft transformieren.

In der Literatur zu «Digital Citizenship» werden eine Vielzahl von Kompetenzen erwähnt, über die Bürger*innen verfügen sollten: Informations- und Medienkompetenz, digitale Kommunikations- und Partizipationsfähigkeiten, Online-Ethik und Selbstsorge sowie die Fähigkeit zum Widerstand, verstanden als kritisch-reflexive Analysefähigkeit des digitalen Raums und daraus abgeleitete Handlungsstrategien (Choi, 2016). Fest steht, dass die neuen Plattformtechnologien die Grundlagen öffentlicher Kommunikation verändern und zu multidimensionalen Machtverschiebungen führen. Inwiefern netzimmanente Machtbeziehungen, die sogenannte «network power» (Castells, 2009; Hubacher, 2021), von jugendlichen Produzent*innen wahrgenommen und reflektiert werden, und welche Schlüsse sie für ihr eigenes Handeln daraus ziehen, ist indes kaum bekannt.

Fragestellung und Methode

Im Rahmen dieses Beitrags werden die folgenden Forschungsfragen bearbeitet:

- Welche alltäglichen Praktiken im Umgang mit dem Politischen im Netz pflegen Jugendliche und über welche Formen der Onlinepartizipation berichten und reflektieren sie? In welcher Hinsicht fühlen sie sich kompetent und wo bestehen Handlungsunsicherheiten?
- Wie nehmen Jugendliche die politische Dimension des digitalen Raumes und die darin wirksamen Machtstrukturen wahr? Welche Strategien leiten diesbezüglich ihr Netzhandeln an?

Im Rahmen einer Pilotstudie hat ein Forschungsteam der PH FHNW zwischen November 2022 und April 2023 mit 20 Jugendlichen im Alter zwischen 13 und 19 Jahren themenzentrierte Interviews geführt. Der Interviewleitfaden war in drei Teile gegliedert und beinhaltete Fragen zum politischen Interesse und zur politischen Sozialisation, Praktiken der Onlinepartizipation sowie Inputs zur Reflexion netzinhärenter Machtstrukturen.

Die Stichprobengewinnung orientierte sich an der Zusammenstellung eines kontrastreichen Samples. In einem ersten Schritt wurden Jugendliche im Kontext von Bewegungen und in aktivistischen Kreisen gesucht ($n = 7$), in einem zweiten Schritt wurde der Zugang zu Schüler*innen gesucht, die nach eigenen Aussagen sich als politisch eher passiv bezeichneten ($n = 9$). Weitere Kontakte führten zu politisch interessierten Jugendlichen ohne direkte Bezüge zu aktivistischen Kreisen ($n = 4$). Die Interviews wurden transkribiert, in MaxQDA eingelesen und inhaltsanalytisch-interpretativ ausgewertet.

Ergebnisse und ihre Bedeutung

Obschon sich die Hälfte der Befragten als noch zu jung für Politik einschätzten, berichteten alle über sozialisierende Effekte und Anknüpfungspunkte zum Politischen, wie etwa die Klimakrise, Corona, der Ukrainekrieg, Black lives matter, die Flüchtlingskrise oder der familiäre Migrationshintergrund. Humor, insbesondere in der Form von Memes, ist äußerst beliebt; zugleich beschäftigt die Grenze von Satire. Die Jugendlichen nutzen einen Medienmix, wobei die politisch weniger Interessierten klassischen Medien mehr Vertrauen schenken; Aktivist*innen hingegen jeglicher Information kritisch begegnen. Betreffend Glaubwürdigkeitsprüfung von Informationen verlassen sich die Befragten einerseits auf die eigene Urteilskraft und andererseits auf die Eltern respektive das Klimastreiknetzwerk. Das Verständnis für Machtstrukturen im digitalen Raum ist erwartungsgemäß eher implizit angelegt, jedoch setzen sich viele Jugendliche mit diesen recht intensiv auseinander. Erwähnt werden u.a. der Normalisierungsdruck, Schweigespiralen, Hate-Speech, Shadow-Banning und Doxing. Insgesamt verweisen die empirischen Ergebnisse auf Aufgaben und Lücken einer «Digital Citizenship Education», die nach Ansicht der Befragten möglichst früh beginnen sollte.

Voraussetzungen für die Vermittlung digitaler politischer Kompetenzen. Eine Analyse zu Lehrkräften in ICCS 2022

Hermann Josef Abs, Daniel Deimel, Nina Welsandt
Universität Duisburg-Essen

Hintergrund

Aufgrund der gesellschaftlichen Bedeutung von Digitalisierung für den Bildungsbereich entstanden übergreifende Systematisierungen zu digitalen Kompetenzen (z.B. Council of Europe, 2022; Vuorikari et al., 2022). Diese repräsentieren einen weitgehenden Konsens über Ziele digitaler Bildung - wie technische Fähigkeiten, Online-Sicherheit und Partizipation. Fraglich ist jedoch, ob diese Kompetenzrahmen schon hinreichend für die politische Bildung spezifiziert sind (Huschle et al., 2024; Heldt & Krösche, 2023; Cortesi et al., 2020; Waldis, 2020).

Die International Civic and Citizenship Education Study, ICCS 2022, legt einen Themenschwerpunkt auf Digitalisierung (Schulz et al., 2023). Ein Ergebnis ist, dass Lehrkräfte in NRW im europäischen Vergleich seltener digitale Bildungsaktivitäten in ihren Unterricht integrieren (Welsandt et al. 2024). Dies ist problematisch, weil Lehrkräfte nachweislich digitalisierte demokratische Teilhabe fördern können (Oberle et al., 2020).

Fragestellung und Methode

Vor diesem Hintergrund verfolgt der Beitrag zwei Ziele: Zuerst wird betrachtet, inwieweit ICCS digitalisierungsbezogene Aspekte aufgreift, die sich (nicht) in internationalen Kompetenzrahmen wiederfinden und inwiefern diese zu einem spezifischen Profil digitalisierungsbezogener politischer Bildung beitragen. Danach greifen wir Ergebnisse zu digitalen Bildungsaktivitäten in ICCS 2022 auf (Schulz et al., 2024, 250f). Gefragt wird, welche Bedingungen im Schulkontext und bei einzelnen Lehrpersonen die Häufigkeit der Durchführung von digitalen Bildungsaktivitäten mit Schüler:innen vorhersagen können.

Die Grundlage der Sekundäranalyse bildet repräsentative Lehrkraftdatensatz in ICCS 2022 in Nordrhein-Westfalen. Die Stichprobe umfasst $N = 2.896$ Lehrkräfte aller Schulformen, die im Schuljahr 2021/22 eine achte Klasse unterrichten haben. Hiervon unterrichteten $N = 1.102$ Lehrkräfte in Jahrgang 8 ein Fach, das "mit politischer Bildung und Demokratieerziehung zu tun hat". Von diesen berichteten wiederum $N = 196$ Lehrkräfte, eine staatliche Lehrbefähigung in "Politik, Sozialwissenschaft, Soziologie".

Die Lehrkräfte wurden gefragt, wie häufig sie mit Schüler:innen der 8. Jahrgangsstufe verschiedene Aktivitäten zur Nutzung des digitalen Raums durchgeführt haben („Nie“, „Einmal“, „Einmal im Monat“, oder „Öfter als einmal im Monat“). Insgesamt wurden vier Aktivitäten erfasst, etwa zur Informationssuche und Beurteilung von Informationen, zur Online-Sicherheit und zur Partizipation in Online-Diskussionen und. Die Fragen bilden eine Skala, die international auf einen Mittelwert von $M = 50$ mit einer Standardabweichung von $SD = 10$ standardisiert wurde. Werte über 50 bedeuten also, dass Lehrkräfte auf diesem Merkmal über dem internationalen Durchschnitt von ICCS 2022 liegen.

Ergebnisse und ihre Bedeutung

Erste Ergebnisse zeigen wenig systematischen Einfluss der Schulebene: Weniger als 2% der Varianz des Skalenwertes werden durch die Zugehörigkeit zu einer bestimmten Schule erklärt. Deutlich wird jedoch, dass eine fachliche Einschlägigkeit individueller Lehrkräfte die Häufigkeit entsprechender Aktivitäten erhöht.

Lehrkräfte, die kein sozialwissenschaftliches Fach unterrichten, erreichen in NRW den niedrigsten Skalenwert von $M = 43,9$ Skalenpunkten. Lehrkräfte, die ein entsprechendes Fach in Jahrgang 8 unterrichten, aber über keine sozialwissenschaftliche Fakultät verfügen, bewegen sich mit $M = 49,9$ Skalenpunkten nahe am internationalen Durchschnitt. Lehrkräfte mit einschlägiger Fakultät übertreffen mit $M = 53,0$ Skalenpunkten den internationalen Durchschnitt.

Während fachlich qualifizierte Lehrkräfte in NRW zwar überdurchschnittlich oft Aktivitäten zu digitaler Demokratiebildung durchführen, liegt das Niveau dieser Gruppe nicht höher als das Niveau, das ein Viertel der europäischen Vergleichsländer mit den jeweiligen Gesamtstichproben ihrer Lehrkräfte erreicht; beispielsweise liegt der Durchschnitt aller befragten Lehrkräfte in Dänemark bei $M = 54,8$ Skalenpunkten. Regressionsanalysen vertiefen die Frage, welche individuellen Bedingungen und Merkmale von Schulsystemen dazu beitragen, dass Lehrkräfte Aktivitäten zur digitalen politischen Bildung durchführen.

In Bezug auf Deutschland zeigt sich, dass der Kompetenzaufbau im Studium eine höhere Bedeutung für die Entwicklung einschlägiger Bildungsaktivitäten in der Schule hat als die schulische Sozialisation durch das Unterrichten eines einschlägigen Faches. Daraus folgert eine hohe Bedeutung, entsprechende Bildungsaktivitäten systematisch (im Studium) anzuleiten. Da Fächer der politischen Bildung überwiegend fachfremd unterrichtet werden, sollte die Verankerung entsprechender Themen im bildungswissenschaftlichen Studienanteil gestärkt werden.

„Digitalisierung“ in den Vorstellungswelten angehender Lehrkräfte der Politischen Bildung

Inken Heldt¹, Manuel Theophil²

¹Universität Passau, ²(Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau

Hintergrund

Die Förderung von ‚digitaler Bildung‘ und ‚digitaler Kompetenz‘ ist das Thema der Stunde. In den akademischen, medialen und (bildungs)politischen Diskursen wird unter diesen Begriffen eine große Bandbreite unterschiedlicher Ansprüche und Inhalte verfolgt. Der vorliegende Beitrag liefert einen Diskussionsimpuls, indem er empirisch erhobene Perspektiven angehender Lehrer*innen der politischen Bildung präsentiert. Entlang von drei Fragen – der Frage nach dem übergreifenden Unterrichtsziel, der Frage nach den wichtigsten Unterrichtsthemen und der Frage nach den größten Gefahren der Digitalisierung – werden Innovationspotentiale der ersten Phase der Lehrkräftebildung in Bezug auf den Lerngegenstand „Digitalisierung“ im politischen Fachunterricht skizziert. Der Beitrag will Anstoß geben zu einer breiteren innerfachlichen und öffentlichen Diskussion über die Frage, wie die politische Bildung ihr Aufgabenverständnis mit Blick auf den digitalen Wandel bestimmen kann.

Die hier ausschnitthaft vorzustellende Studie fokussiert die Art und Weise, wie das Phänomen der Digitalisierung in den Vorstellungen von Lehramtsstudierenden der politischen Bildung in Erscheinung tritt und welche Charakteristika und Bezogenheit dabei zu beobachten sind. Auf welche Begriffe, Relevanzsetzungen und Problembeschreibungen, auf welche Lernziele und -inhalte nehmen angehende Politiklehrkräfte Bezug, wenn es um ‚Digitalisierung‘ geht?

Das Offenlegen der themenspezifischen Deutungsmuster angehende Lehrkräfte ist für die fachdidaktische Lehrkräfteausbildung relevant, weil auf dieser Grundlage identifiziert werden kann, welche Thematisierungen als sinnerschließende Impulse an die bestehenden Vorstellungen der Lehramtsstudierenden anschlussfähig sind. Damit soll ein Beitrag zur Konzeption universitärer Lehrangebote geleistet werden, die an den lebensweltlichen Deutungsmustern der Beteiligten ansetzen. Die Kontrastierung der subjektiven Vorstellungselemente mit wissenschaftlichen Perspektiven bewirkt ein Spannungsverhältnis, das Eigenheiten und charakteristische Elemente der wissenschaftlichen und lebensweltlichen Zugänge zum Phänomenbereich sichtbar macht und aus dem theoretisch-konzeptionelle Lernzugänge, -inhalte und -ziele abgeleitet werden können.

Fragestellung und Methode

Die Studie ist explorativ angelegt und liefert erste Einblicke in ein Feld, über das bisher kaum empirisch gestützte Erkenntnisse vorliegen. Die Studie beruht auf einer qualitativen Erhebung bestehend aus n = 20 problemzentrierten Interviews. Das Sample besteht aus Bachelor- und Master-Studierenden des Lehramts Sozialkunde (für Gymnasium, Berufsschulen sowie Realschule Plus) an der TU Kaiserslautern. Die Interviews wurden zwischen Dezember 2021 und März 2022 geführt. Die Auswertung der Daten erfolgt als strukturierende qualitative Inhaltsanalyse (Kuckartz & Rädiker, 2022, S. 129).

Ergebnisse und ihre Bedeutung

Unter den von uns befragten angehenden Lehrkräften ist die Vorstellung präsent, dass ‚die Digitalisierung‘ mit einem neuen, spezifischen Set von Lerninhalten und -zielen einhergeht. Diese Forderung lässt eine Vorstellung des Digitalen erkennen, die in ihren Implikationen hinterfragt werden muss. Gemeint ist die Annahme, ‚Digitalisierung‘ beziehe sich auf den vermeintlich abgrenzbaren Gegenstandsbereich der „digitalen Welt“ und die unterrichtliche Auseinandersetzung befähige zu einem verantwortungsvollen, vorsichtigen Umgang mit dieser. „Digitale Bildung“ bzw. „Medienbildung“ wird damit als etwas ‚Zusätzliches‘ verstanden, also als ein spezifisches Lernfeld, das additiv zu anderen Lernfeldern des Politikunterrichts und ihren jeweiligen Zielen aufzugreifen ist (Kerres, 2020).

Die These unseres Beitrages ist, dass die in der Datenanalyse offengelegte implizite Gegenüberstellung von ‚digitaler‘ und ‚analoger‘ Welt einer Denkfigur anhängt, die den Wirkungsbereich des Digitalen unterschätzt – und daher irreführend ist. Die politische Bildung steht vor der Herausforderung, Digitalität künftig nicht allein als Bereichsphänomen von Medien im Sinne einer Reduktion auf ‚mediale‘, kommunikations- und informationszentrierte Phänomene der Digitalisierung zu thematisieren, sondern als einen Prozess, der die mit Digitalisierung verbundenen – und häufig über Medien und digitale Technologien vermittelten – individuellen und gesellschaftlichen Veränderungen stärker reflektiert und zu normativen Fragen in Beziehung setzt. Nicht der Umgang von Lernenden ‚mit‘ dem Digitalen steht damit im Fokus des Politikunterrichts, sondern die Frage danach, wie Subjekte in ihrer Welt- und Selbstsicht fortlaufend durch die Digitalität geprägt und beeinflusst werden – wofür die Interessen der Akteure ‚hinter‘ den Nutzeroberflächen als solche (und damit als politisch) erkennbar zu machen sind.

Politiklehrkräfte zur politischen Medienbildung befähigen – empirische Befunde zu Wirkungen digitaler Lehrkräftefortbildungen

Monika Oberle, Natalie Grobshaeuser

Goethe-Universität Frankfurt am Main

Hintergrund

Angesichts der gewandelten Bedingungen der „digitalen Demokratie“ (Borucki et al., 2020) bzw. der „e-democracy“ (Kneuer, 2016) bedarf es für eine lebendige und funktionsfähige Demokratie neuer Orientierungsfähigkeiten seitens der Bürger*innen. Politische Bildung hat die Aufgabe, zur Förderung einer politischen Medienkompetenz (Oberle, 2022) der Lernenden als Bestandteil ihrer digitalen Souveränität (Blossfeld et al., 2018) beizutragen. Für eine mündige politische Urteilsbildung müssen Schüler*innen u.a. dazu befähigt werden, Fake News und politische Desinformation (Guess & Lyons, 2020) zu erkennen und deren Verbreitungsmotive und -strategien zu verstehen. Lehrkräfte müssen für diese Aufgabe sensibilisiert werden und zugleich didaktische Wege kennen, die politische Medienkompetenz ihrer Schüler*innen zu fördern. Dies kann grundsätzlich mit digitalen Tools (bspw. serious games) erfolgen, aber auch mit analogen didaktischen Ansätzen. Da es oftmals an einer Ausbildung von Politiklehrkräften für die Förderung fachbezogener digitaler Schülerkompetenzen fehlt, auch bei jüngeren Lehrer*innen (vgl. Welsandt et al., 2024), sind Fortbildungen zur politischen Medienbildung für praktizierende Politiklehrkräfte sehr relevant. Empirische Befunde zu Wirkungen solcher Fortbildungen liegen bisher kaum vor.

Im Frankfurter Teilprojekt des BMBF-geförderten Verbundprojekts „Digitale Souveränität als Ziel wegweisender Lehrer:innenbildung für Sprachen, Gesellschafts- und Wirtschaftswissenschaften“ (DiSo-SGW) werden Lehrkräftefortbildungen entwickelt, implementiert, begleitend beforscht und evidenzbasiert optimiert, in denen Politiklehrkräfte der Sekundarstufen lernen, serious games und andere digitale Tools zielgruppenorientiert einzusetzen und dabei fachliche Kompetenzen digital und digitale Kompetenzen fachlich zu fördern. Ein besonderer Schwerpunkt liegt auf der Befähigung der Lehrkräfte, die Souveränität von Lernenden im Umgang mit dem Phänomen Fake News bzw. politischer Desinformation zu stärken.

Fragestellung und Methode

Der Beitrag stellt erste Ergebnisse der Begleitforschung zu fünf eintägigen digitalen Lehrkräftefortbildungen vor, an denen 2024 ca. 100 Lehrkräfte von Schulen der Sekundarstufen aus ganz Deutschland teilgenommen haben. Die Fortbildungen fokussierten die Themen Fake News als Herausforderung für Demokratie und politische Bildung sowie Ansätze des digital game-based learning zur Förderung der politischen Medienkompetenz der Schüler*innen.

Präsentiert werden ausgewählte Ergebnisse einer längsschnittlichen Teilnehmendenbefragung zu zwei Messzeitpunkten, die direkt vor und direkt nach den Veranstaltungen mit Hilfe erfolgte. Der teilstandardisierte Online-Fragebogen enthielt Item-Batterien zu fachlichen und fachdidaktischen Dispositionen der Lehrkräfte, die vor und nach Fortbildungsteilnahme erhoben wurden, sowie Items zur Bewertung der Veranstaltung inklusive Durchführungsqualität und subjektiv wahrgenommenen Lerneffekten. Der Fokus der Präsentation liegt auf den fachdidaktischen Selbstwirksamkeitsüberzeugungen der Lehrkräfte (Oberle et al., 2015) bezüglich politischer Medienbildung im Unterricht, ihren Ausprägungen, Einflussfaktoren und Veränderungen.

Ergebnisse und ihre Bedeutung

Insgesamt weisen die Daten auf eine hohe Zufriedenheit der Lehrkräfte mit den Fortbildungen, ihrer Prozessqualität und ihren Ergebnissen hin. Zuwächse zeigen sich in den fachlichen Kenntnissen und den fachdidaktischen Selbstwirksamkeitsüberzeugungen der Lehrkräfte bezüglich politischer Medienbildung im Unterricht. Die Lehrkräfte trauen sich die zielgruppengerechte Thematisierung von Fake News, auch mit Hilfe digitaler Tools, und die Sensibilisierung von Schüler*innen für politische Desinformation nach Fortbildungsteilnahme eher zu. Die Wirkungen hängen systematisch mit den Vorkenntnissen der Lehrkräfte zusammen. Die Ausgangskompetenzen der Lehrkräfte sind wiederum abhängig von ihrer Ausbildung.

Die Befunde unterstreichen einerseits den Bedarf an Lehrkräftefortbildungen zur Förderung der digitalen Souveränität im Politikunterricht, womit teilweise eine fehlende Ausbildung der Lehrkräfte kompensiert werden muss. Andererseits weisen sie auf die Potenziale solcher Fortbildungen hin, die auch in relativ kurzen, eintägigen und digitalen Formaten in digitaler Form signifikante Effekte erzielen können.

Bedarf besteht an einer Untersuchung längerfristiger Wirkungen. Ein solches Forschungsdesign mit Follow-Up-Erhebung ist in der zweiten Projektphase in 2025 geplant. Dabei soll untersucht werden, inwiefern mehrmodulare Fortbildungen, die auch eine Praxisphase sowie eine Begleitung bzw. gemeinsame Reflexion der gesammelten Erfahrungen beinhalten, zusätzliche nachhaltige Wirkungen auf die Kompetenzen und Unterrichtspraxis der Lehrkräfte entfalten können.

PS-72: Prüfungen und Innovationen: Fairness, Emotionen und neue Ansätze im Hochschulkontext

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: SO 418

Chair der Sitzung: Hernán González Cruz

Diversitätsgerechte Optimierung der E-Prüfungssoftware Dynexite: Eine Mixed-Methods-Studie zur Verbesserung der User-Experience

Kirsten Gropengießer¹, Nikolai Zinke², Sina Lenski², Natalia Reich-Stiebert¹, Sohail Bayzaei¹, Stefan Stürmer¹

¹FernUniversität Hagen, Deutschland; ²Deutsches Institut für Erwachsenenbildung e.V., Bonn, Deutschland

Theoretischer Hintergrund

Insbesondere Fernhochschulen erfahren eine zunehmende Digitalisierung der Bildungsprozesse (Blumentritt et al., 2019) und weisen eine diverse Studierendenschaft auf (Stürmer et al., 2018). Zu den Diversitätsmerkmalen zählen unter anderem Alter und Sprachkenntnisse, aber auch Vorerfahrungen mit E-Prüfungssystemen sind von Interesse (Froehlich et al., 2023).

Studierende mit geringerer Technikaffinität und Vorerfahrung im Umgang mit digitalen Medien stoßen häufig auf Barrieren im Studium (Senkbeil et al., 2019). Ein Modell zur Erklärung der Akzeptanz neuer Technologien ist das Technology-Acceptance-Model (TAM). Die Einflussvariable Nutzerfreundlichkeit des TAM, die als wichtigste Komponente der User Experience gilt (Brüggemann & Wiepcke, 2023), wird durch die Vorerfahrung moderiert (Davis, 1989; Venkatesh et al., 2003). Ziel der Studie ist es, die User Experience des an der RWTH Aachen entwickelten E-Prüfungssystems Dynexite (vgl. Corves et al., 2019) mittels des TAM zu evaluieren und das System bedarfsgerecht weiterzuentwickeln, um die Akzeptanz bei den Studierenden zu erhöhen.

Fragestellungen

RQ1: Welche Verbesserungsbedarfe zur Steigerung der Akzeptanz von E-Prüfungssystemen lassen sich mithilfe der Concurrent-Think-Aloud-Methode identifizieren?

RQ2: Welche Zusammenhänge bestehen zwischen den Diversitätsmerkmalen und der Akzeptanz von E-Prüfungssystemen?

RQ3: Zeigt sich eine Steigerung der Akzeptanz nach Weiterentwicklung der E-Prüfungssysteme und Optimierung der Aufgabengestaltung?

Methode

Mittels eines Mixed-Methods-Ansatzes wurde das Verbesserungspotenzial von Dynexite anhand eines Prä-Post-Designs untersucht. Es wurden Diversitätsvariablen wie Alter, Geschlecht, Erstsprachen- und Behindertenstatus, Prüfungsangst, Computerangst sowie als zusätzliches Merkmal die Einstellung/Akzeptanz zu Dynexite berücksichtigt. Insgesamt nahmen $N = 27$ Psychologiestudierende an der Studie teil. Im Prätest äußerten diese ihre Gedanken während der Bearbeitung unterschiedlicher E-Prüfungs-Testaufgaben mittels der Concurrent-Think-Aloud-Methode (CTA; Nielsen et al., 2021) und füllten anschließend einen Online-Usability-Fragebogen aus. Auf Basis des Feedbacks der Studierenden konnten mittels Inhaltsanalyse entsprechende Verbesserungsbedarfe, sogenannte Feature Requests (FR), abgeleitet werden. Die daraufhin modifizierte Dynexite-Version wurde in einem Posttest erneut mittels CTA und Online-Usability-Fragebogen evaluiert. Das daraus resultierende Feedback wurde auf Passung mit den bereits im Prätest generierten FR überprüft und entsprechend zugeordnet.

Ergebnisse

Die Ergebnisse zu **RQ2** zeigten einen negativen Zusammenhang zwischen Nutzerfreundlichkeit mit unstrukturiertem Fließtext ($r = -.40, p = .041$). Zudem bestand ein negativer Zusammenhang zwischen Prüfungsangst und Scrollen nach einem Aufgabenwechsel ($r = -.41, p = .039$). Ein positiver Zusammenhang wurde zwischen dem Behindertenstatus und einem durch die Aufgabe erforderlichen Wechsel zwischen einem PDF-Dokument und der eigentlichen Aufgabenstellung festgestellt ($r = .43, p = .027$). Es gab einen positiven Trend von der Anzahl absolvierter Prüfungen mit Dynexite zur Akzeptanz gegenüber Dynexite ($r = .35, p = .079$). Auch zeigte sich, dass höhere Computerangst den Zusammenhang zwischen Nutzerfreundlichkeit und der Einstellung zu Dynexite verstärkte ($b_{PEOU \times CA} = 0.19, p = .005$). Im Posttest (**RQ3**) wurden Verbesserungen bei zwei FR nach Modifikation und bei zwei FR ohne Modifikation festgestellt. Dies zeigte sich in einer Reduktion der Anzahl der Nennungen im Posttest.

Diskussion

Die Ergebnisse zeigen, dass gezielte Modifikationen von Dynexite zu einer Verbesserung der User Experience beitragen können, insbesondere hinsichtlich der Strukturierung von Texten und durch Vermeidung unnötigen Scrollens. Die Anzahl absolvierter E-Prüfungen war tendenziell positiv mit Akzeptanz assoziiert. Dieses Ergebnis steht im Einklang mit Befunden von Froehlich et al. (2023), wonach die Vorerfahrung zu einer Steigerung der Akzeptanz von E-Prüfungen führt. Studierende mit geringeren Ressourcen, z. B. weniger Vertrautheit mit dem System, scheinen stärker von der wahrgenommenen Nutzerfreundlichkeit des Systems abhängig zu sein. Dies entspricht der Empfehlung von Klammer (2019), das System in seinen Funktionen gezielt weiterzuentwickeln und durch den Abbau von Barrieren Chancengleichheit zu gewährleisten. Um zukünftig standardisierte Veränderungsmessungen zu realisieren, sollten iterative Nutzerfeedbackschleifen mit größeren Stichproben im laufenden Betrieb durchgeführt werden. Insgesamt liefert die Studie Erkenntnisse für die nutzerzentrierte Weiterentwicklung und Verbesserung der Gebrauchstauglichkeit und damit der Akzeptanz von E-Prüfungssystemen, insbesondere durch die Identifizierung von Verbesserungsbedarfen für das E-Prüfungssystem Dynexite.

Cheating behavior in online exams can be reduced by considering learner-centered aspects of university students

Marco Rütth, Kai Kaspar

Universität zu Köln, Deutschland

Theoretical

Online assessments such as online tests are among the most frequently used evaluation methods in universities (e.g., Flores et al., 2015; Gonçalves et al., 2020). However, students reported to cheat in online exams significantly more often during the Covid-19 pandemic than before (Newton & Essex, 2023). Cheating has been a serious and prevalent problem regarding traditional face-to-face exams (e.g., Rettinger & Gallant, 2022; Whitley, 1998). Regarding online exams, common measures against cheating such as proctoring imply ethical questions regarding academic integrity and should be carefully weighed with alternatives (cf. Coghlán et al., 2021). So, rather than focusing on countermeasures such as proctoring to fix the symptom (cheating), we scrutinize

background

whether a lack of learner-centered aspects is a possible cause for cheating in online exams. Question

We formulated three research questions on the role of learner-centered aspects in cheating in online exams, which has yet been underexplored. First, with reference to self-determination theory (Anderman et al., 2022; Ryan & Deci, 2020), we investigated how effective addressing university students' needs is to reduce their cheating intention in online exams (RQ1). Second, based on previous research on the role of students' self-regulation in higher education (Brown, 2021; Flores et al., 2020), we examined how university students' conceptions of online exams, age, and gender relate to their future and past cheating behavior (RQ2). Third, to better understand students' motives regarding cheating, we investigated university students' most important reasons for and against cheating in past online exams (RQ3).

Method

By means of an online study, we collected data from 339 students from universities in Germany (MAge = 25.01, SDAge = 4.89, 276 female, 61 male, 2 diverse) who had written at least one online exam since 2020. The students were enrolled in a bachelor's degree program (55.46%), a master's degree program (37.46%), or a state examination program (7.08%). Most participants did not specify their course of studies (38.64%), while the other participants studied psychology (29.50%), teacher education (5.60%), or business administration (2.65%). Regarding students' needs (RQ1), we experimentally tested whether students' cheating intention is lower in four scenarios in which their basic psychological need for autonomy or competence is satisfied (vs. frustrated) and conducted t-tests for independent samples. Regarding the role of students' conceptions, age, and gender in cheating intention and behavior (RQ2), we calculated bivariate correlations and multiple linear regression models. Test requirements were checked, and the analyses were based on instruments with good psychometric properties (all Cronbach's $\alpha \geq 0.72$). Regarding students' reasons (RQ3), an inductive content analysis of responses to open-ended questions was performed, resulting in good interrater agreement (Cohen's Kappa = 0.84).

Results and their significance

Our results support that learner-centered aspects play an important role in university students' cheating behavior in online exams. Students' cheating intention was lower in three out of four scenarios in which their needs were addressed. In specific, students may cheat less if teachers illustrate throughout the course why students need to master the content of the examination for their future professional life, if they include problem-solving tasks that are as authentic as possible to the demands of students' future profession, and if they provide performance feedback and detailed feedback to exam results if possible. Moreover, all seven conceptions of online exams were related to students' cheating intention and behavior. These results corroborate our theoretical approach towards motivation (self-determination theory) and self-regulation (conceptions of online exams). We also identified students' most important reasons for and against cheating, including the quality of teaching and assessment as well as moral standards, respectively. Overall, our insights allow to formulate several practical ways to address and reduce cheating behavior.

Je dichter die Prüfung, desto höher die Angst? – Leistungsangst im Semesterverlauf vor einer Klausur am Semesterende

Jörn R. Sparfeldt, Lea Charleen Schwindt

Universität des Saarlandes, Deutschland

Viele Studierende erleben vor bedeutsamen Prüfungen Leistungsangst; die Leistungsangst nimmt mit zeitlicher Nähe zur Prüfung häufig deutlich zu (z.B. Lotz & Sparfeldt, 2017; Rafetty et al., 1997). Außerdem scheint die Leistungsangstfacette Besorgtheit bereits weiter vor der Prüfung anzusteigen und hoch zu bleiben, während Aufgeregtheit erst unmittelbar vor oder während der Prüfung zuzunehmen scheint (vgl. Zeidner, 1998). Zumeist wurden jedoch nur wenige Messzeitpunkte oder ein kurzer Zeitraum vor der Prüfung betrachtet. Auch wurden beide Leistungsangstfacetten selten getrennt untersucht. Daher sollte der zeitliche Verlauf von Besorgtheit und Aufgeregtheit über ein Semester (mit multiplen Messzeitpunkten) bis zu einer bedeutsamen Prüfung erfasst werden. Üblicherweise korrelieren auf eine Lehrveranstaltung und Prüfung bezogene allgemeine Leistungsangst (trait) und Zustandsangst (state) mindestens mittelhoch (z.B. Lotz & Sparfeldt, 2017).

Erwartungs-Wert-Theorien (z.B. Eccles, 1983) und die Control-Value-Theorie (Pekrun, 2006) postulieren, dass die Erwartung möglicher Ergebnisse und ihrer Konsequenzen sowie deren subjektiver Wert die Angstaussprägung beeinflussen. Daher sollte die Angstaussprägung aus dem ebenfalls mehrfach erfassten akademischen Selbstkonzept (für „Erwartung“) und intrinsischen Wert (für „Wert“) im Semesterverlauf analysiert werden.

Wir erwarteten (1) eine Zunahme der klausurbezogenen Leistungsangst im Semesterverlauf mit zunehmender Nähe zur Klausur, (2) positive und mindestens mittelhohe Korrelationen zwischen trait- und state- Besorgtheit und -Aufgeregtheit sowie zwischen trait- Besorgtheit und -Aufgeregtheit und (3) signifikante Vorhersagen der zu späteren Messzeitpunkten erfassten Leistungsangst aus dem Selbstkonzept und dem intrinsischen Wert des jeweils vorangegangenen Messzeitpunkts.

Methoden

Die Daten wurden in einer einsemestrigen bildungswissenschaftlichen Lehramtsvorlesung an acht (zumeist) zwei Wochen auseinander liegenden Messzeitpunkten erhoben (t1: erste Sitzung, t8 unmittelbar vor der für den weiteren Studienverlauf bedeutsamen Abschlussklausur am Semesterende). Studierende, die nicht zu t1 und t8 teilnahmen, wurden von den Berechnungen ausgeschlossen. Die Analysestichprobe bestand aus 73 Lehramtsstudierenden (Frauen/Männer: $n = 54/18$; $M = 22.07$ Jahre, $S = 3.29$).

Besorgtheit und Aufgeregtheit (Hodapp et al., 2011) wurden als trait (t1, Besorgtheit/Aufgeregtheit: Cronbachs $\alpha = .78/.90$) und als state ($.93 \leq \alpha \leq .95/.80 \leq \alpha \leq .93$) in Bezug auf die Abschlussklausur (t1 – t8) mit fünfstufiger Antwortskala erfasst (1: nie, 5: sehr häufig). Das auf die Vorlesung bezogene akademische Selbstkonzept (Items aus Rost et al., 2007) und der intrinsische Wert (Steinmayr & Spinmayr, 2010) wurden ebenfalls als trait (Selbstkonzept/Wert: t1, $\alpha = .76/.52$) und state (t2 – t8, $.58 \leq \alpha \leq .77/.80 \leq \alpha \leq .93$) erhoben.

Ergebnisse

Zur Angstaussprägung im Semesterverlauf zeigten Modellvergleiche (Wachstumskurvenmodelle; LGCM) eine statistisch signifikant bessere Anpassung eines Modells mit quadratischem Term als Modelle mit linearem Verlauf oder keiner Veränderung über die Zeit. Für Besorgtheit zeigte sich deskriptiv zunächst eine Abnahme (t1: 3.64 bis t4: 3.38), gefolgt von einer langsamen Zunahme (t8: 3.61). Aufgeregtheit nahm ebenfalls zunächst ab (t1: 2.70, t3: 2.30), gefolgt von einer relativen Stabilität bis zu t6 (-2.3) und dann einer stärkeren Zunahme (t8: 2.80).

Trait- und State-Kennwerte hingen für Besorgtheit ($.45 \leq r \leq .74$) und Aufgeregtheit ($.56 \leq r \leq .82$), bedeutsam zusammen. Trait-Besorgtheit und -Aufgeregtheit korrelierten stark ($r = .56$). Die um Selbstkonzept und Intrinsischen Wert erweiterten Strukturgleichungsmodelle konvergierten nicht.

Diskussion

Über den Angstanstieg mit zunehmender Nähe zur Klausur am Semesterende hinaus erlaubt die Analyse eines Semesters Erkenntnisse zur Dynamik auch am Semesterbeginn. Eine Erklärung für den Abfall am Semesteranfang könnte in der Reduktion von angststeigernder Unsicherheit (über z.B. Inhalte und Lehrperson) liegen. Trotz eher geringer Veränderung um die Semestermitte verweist der bereits bei zeitlicher Distanz zu beobachtende Anstieg von Besorgtheit (bei Aufgeregtheit erst kurz vor der Klausur) auf die Relevanz einer getrennten Betrachtung beider Leistungsangstfacetten. Auch wenn der Verlauf durch die Modellerweiterung um Selbstkonzept und intrinsischen Wert nicht aufgeklärt werden konnte, geben die Ergebnisse zur Dynamik im Semesterverlauf vor einer Klausur Hinweise für Ansatzpunkte und Interventionsideen, die im Vortrag diskutiert werden.

Investigating task-specific self-efficacy as a moderator of the relationship between performance goals and cheating

Hernán González Cruz¹, Tanja Fritz², Martin Daumiller², Stefan Janke¹

¹University of Mannheim, Germany; ²University of Augsburg, Germany

Academic dishonesty among students is a widespread problem affecting various aspects of higher education, including the validity of assessments and competence development (McCabe et al., 2001). To address this issue effectively, it is crucial to understand how known factors influencing cheating interact. Current motivational frameworks suggest that performance goals are positively associated with cheating. This association stems from the perceived utility of cheating in performance-oriented contexts (Anderman & Koenka, 2017). Additionally, theoretical models and meta-analyses highlight task-specific self-efficacy, as an important predictor of cheating (Lee et al., 2020; Murdock & Anderman, 2006). An interaction between these factors is theoretically plausible. The perceived usefulness of cheating as a performance strategy may be especially strong when task-specific self-efficacy is low, and the performance goal cannot be achieved through one's own competence (Murdock & Anderman, 2006). Despite its plausibility, this interaction remains unstudied.

This study examined if task-specific self-efficacy moderates the relationship between performance goals and academic dishonesty. We hypothesized that low self-efficacy would strengthen the positive association between performance goals and cheating. We aim to understand how self-efficacy affects dishonesty as a strategy, contributing to knowledge on student cheating motivations. Furthermore, we employ a novel application of paradata methodology (Diedenhofen & Musch, 2017) to assess academic dishonesty in online environments, potentially informing future research and integrity strategies.

Our study employed a 2x2 factorial design, manipulating self-efficacy (high vs. low) and performance goals (induced vs. not induced) in an online experimental setting. Participants (N = 203) completed an aptitude test framed for academic positions, consisting of six numerical puzzles somewhat like Sudoku, with three designed to be especially challenging. To manipulate task-specific self-efficacy, participants in the high self-efficacy condition received training on useful puzzle-solving strategies, while the low self-efficacy group completed difficult puzzles without guidance. During the practice phase, all participants were presented with an online puzzle solver to check their answers. They were then explicitly instructed to close the solver and not use any unauthorized aids, including the solver, during the actual test. To induce performance goals, participants were told that their goal was to demonstrate high performance as this was linked to their compensation. Cheating behavior was assessed using paradata methodology, tracking tab-switching events during the test (Diedenhofen & Musch, 2017). The average duration of the experiment was 34 minutes.

Manipulation checks showed that participants in the performance conditions scored higher on normative-approach performance goals ($t(119.04) = 2.55, p = .012, d = .38$), supporting the effectiveness of the performance manipulation. Surprisingly, participants in the high task-specific self-efficacy conditions did not report higher self-efficacy compared to the active control conditions ($t(120.65) = 1.46, p = .148, d = .21$). Regarding paradata, we found no significant interaction for tab-switching events during the test (i.e., unallowed access to the online solver). Specifically, we tested the interaction effect of the performance condition with a continuous measure of task-specific self-efficacy, given the failed manipulation check ($b = 0.10, p = .716$).

These preliminary findings highlight the challenge of robust manipulation of task-specific self-efficacy for the study of dishonest behavior in experimental settings, even in the context of a task that was novel to our sample. The failed manipulation check for self-efficacy limits our ability to draw firm conclusions about its moderating role, underscoring the complexity of factors influencing academic dishonesty. Future research should explore other robust methods for manipulating self-efficacy and studying the interplay between self-efficacy, performance goals, and cheating behavior. This work emphasizes a general need for innovative experimental methodologies in academic integrity research, which may ultimately shed more light on how individual and contextual determinants give rise to this phenomenon in higher education.

EduTopics – Eine interaktive App zur Exploration von Forschungsfeldern mit Textmining am Beispiel der ECER von 1998 bis 2024

Alexander Christ, Jens Röschlein, Christoph Schindler

DIPF - Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Konferenzen als „Marketplaces for Ideas“ (Hansen & Budtz Pedersen, 2018) stellen einen der zentralen Dreh- und Angelpunkte des (interdisziplinären) wissenschaftlichen Austauschs dar. Auf ihnen werden Einblicke in aktuelle Forschungsprojekte gegeben als auch ihre Ergebnisse und Implikationen diskutiert. Dabei werden zugleich neue Forschungs(sub)felder und Paradigmen verhandelt (Gross & Fleming, 2011). Eine Analyse und Kartierung (Böhm et al., 2002) der Schwerpunkte, Trends und Desiderate wie auch der beteiligten Forschenden und deren Kooperationen ermöglicht daher einen Überblick über vergangene und aktuelle Entwicklungen der internationalen Forschung. Wie auch bei der Erarbeitung des Forschungsstandes aus der Literatur, verhindert die schiere Anzahl an Forschungsarbeiten, beteiligten (Sub-)Disziplinen und die daraus resultierende Heterogenität bei periodischen großen Konferenzen einen einfachen Überblick. Die Nutzung von Methoden aus der Computerlinguistik und dem maschinellen Lernen für die Analyse und Clustering der Forschungsarbeiten ist daher eine grundlegende Voraussetzung für die Schaffung eines Überblicks über Schwerpunkte, Trends und Desiderate.

Dieses Ziel wird in diesem Beitrag verfolgt, indem auf Basis der über $n = 30\,000$ Einreichungen der European Conference on Educational Research (ECER) von 1998 bis 2024 ein frei zugängliches Open Access und Open Source Web-Tool (<https://dipf-lis.shinyapps.io/EduTopicsECER/>) entwickelt wurde. Das Tool ermöglicht Nutzer:innen die Exploration der Beiträge nach geclusterten inhaltlichen Themen und weiteren Kovariaten wie Publikationsjahr, Person, ECER-Netzwerk oder Affiliation. Zu diesem Zweck wurden die Dokumente mittels Topic Modelling nach ihren inhaltlichen Schwerpunkten analysiert (Blei & Lafferty, 2009). Topic Modelling ist ein exploratives Cluster-Verfahren mit welchem die latenten inhaltlichen Themen (Topics) eines Korpus

anhand häufig (gemeinsam) vorkommender Wörter in Dokumenten bestimmt werden können. Die Ergebnisse dieses Verfahrens wurden anschließend differenziert anhand der oben erwähnten Kovariaten, um den Nutzer:innen die Suche an Schnittstellen der Kovariaten mit den latenten Themen zu ermöglichen (Madzik et al., 2022; Xie et al., 2022). Alle Schnittstellen zwischen den bibliographischen Variablen und latenten Themen wurden per Shiny (Chang et al., 2024) interaktiv visualisiert, wodurch nicht nur die Suche in diversen Tabellen ermöglicht wird, sondern auch die Exploration des Korpus anhand graphischer Darstellungen wie Heatmaps, Streudiagrammen oder Graphenvisualisierungen vorgenommen werden kann.

Für die über $n = 30\,000$ Dokumente zeigte sich anhand von Simulationen wie Gibbs-Sampling (Griffiths & Steyvers, 2004) und der qualitativen Beurteilung der Topics, dass $k = 50$ Topics die inhaltlichen Themen des Korpus zufriedenstellend darstellen. Ein Großteil von ihnen entsprach Forschungsthemen wie beispielsweise „Sustainability Education“, „Intercultural Education and Immigration“ oder „E-Learning in (Higher) Education“. Hinzu kamen auch eher methodisch geprägte Topics wie „Qualitative Interview Studies“, „Reviews and Bibliometry“ oder „Ethnographic Studies“. Für die zeitliche Entwicklung der Topics über die insgesamt 25 Konferenzen zeigten sich eindeutige Trends, ebenso wie für die Zusammenhänge der Verortung innerhalb eines ECER-Netzwerkes. Bei der Graphenvisualisierung der Themen zeigte sich, dass Querschnittsthemen wie „Attitudes towards Learning“ oder „School System and (Student) Performance“ zentral für den ganzen Korpus sind, während Topics wie z.B. „Labour Market and Neoliberalisation“ oder „Gender and Gender Differences“ tendenziell distinkt von anderen Themen behandelt wurden. Die Nutzung der Shiny-App mitsamt der Exportfunktion als auch einer Funktion zur Verortung eigener Texte anhand der identifizierten Topics wird im Beitrag exemplarisch in Laufzeit dargestellt.

Dieser Beitrag verfolgt zwei zentrale Ziele: (1) Die Vorstellung des Tools für Nutzer:innen anhand exemplarischer explorativer Auswertungen der ECER-Beiträge inklusive eines Einblicks in eine mögliche zukünftige Ausweitung der Methode für den Einsatz bei Literaturrecherchen und (2) eine Diskussion der Potenziale, die Verfahren aus der Computerlinguistik und maschinellen Lernen und daraus entstehender Analyse- und Visualisierungstools für die „klassische“ empirische Forschung mit sich bringen. Es werden anschließend zukünftige Entwicklungsschritte wie die Integration weiterer Verfahren aus dem Bereich des maschinellen Lernens und die Übertragbarkeit des Tools auf weitere Korpora diskutiert.

PS-73: Mathematiklernen: Didaktik, Aufgaben und Kompetenzen

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 154

Chair der Sitzung: Ann-Katrin van den Ham

Visualisierungen von Brüchen in Schulbüchern: Diskret, diskretisiert oder kontinuierlich?

Sabrina Schwarzmeier, Andreas Obersteiner

Technische Universität München, Deutschland

Theoretischer Hintergrund

Das Erlernen von Brüchen ist für Schülerinnen und Schüler oft schwer. Eine Ursache ist, dass Kinder Vorwissen über natürliche Zahlen auf Brüche übertragen, obwohl dies nicht immer mathematisch korrekt ist (Lortie-Forgues et al., 2015; Ni & Zhou, 2005).

Der Größenvergleich von Brüchen ist ein typisches Beispiel dafür. Lernende betrachten die Bruchkomponenten manchmal getrennt voneinander und ziehen falsche Schlussfolgerungen (z. B. "3/9 ist größer als 2/5, weil $3 > 2$ und $9 > 5$ "). Visualisierungen wie Bruchstreifen (längliche, rechteckige Darstellungen) könnten beim Aufbau von Größenvorstellungen (Fähigkeit, die numerische Größe eines Bruchzahlsymbols schnell einzuschätzen) und somit bei der Überwindung von Schwierigkeiten beim Größenvergleich hilfreich sein, weil visuell schnell erkennbar ist, wie groß ein Bruch ist (Atagi et al., 2016). Dabei werden frühe kindliche Fähigkeiten beim Einschätzen visuell dargestellter Anteile genutzt (Jeong et al., 2007).

Begolli et al. (2020) zeigten, dass Kinder visuell dargestellte Anteile besser einschätzen können, wenn *kontinuierliche* Visualisierungen (ohne sichtbare Unterteilungen) verwendet werden, im Vergleich zu *diskretisierten* (mit sichtbaren Unterteilungen) oder *diskreten* Visualisierungen (räumlich getrennte Objekte). Ein Grund dafür ist, dass Kinder bei diskretisierten und diskreten Visualisierungen häufiger fehleranfällige und ineffiziente Zählstrategien verwenden, die auf ihrem Vorwissen über natürliche Zahlen basieren, während sie bei kontinuierlichen Visualisierungen eher effiziente Strategien anwenden, die auf dem Einschätzen von Bruchgrößen basieren (Schwarzmeier et al., 2024).

Fragestellung

Diese Erkenntnisse legen nahe, dass kontinuierliche Visualisierungen Vorteile bieten könnten, die im Unterricht genutzt werden sollten. Schulbücher spielen eine entscheidende Rolle im Unterricht, da sie die in Lehrplänen abstrakt formulierten Lernziele in konkrete Instruktionsmöglichkeiten übersetzen und Lehrkräften bei der Unterrichtsplanung eine Orientierung bieten (Valverde et al., 2002). Dennoch existieren bisher nur wenige systematische Erkenntnisse darüber, ob kontinuierliche Visualisierungen von Brüchen in Schulbüchern angemessen berücksichtigt werden und ob sie Lerngelegenheiten zum Aufbau von Größenvorstellungen (Aufgaben zum Einschätzen oder Überschlagen) beinhalten. Daher analysiert die vorliegende Studie a) die Verwendung unterschiedlicher Visualisierungsarten (kontinuierlich, diskretisiert und diskret) von Brüchen und b) das Auftreten von Aufgaben mit Visualisierungen zum expliziten Einschätzen von Bruchgrößen in Schulbüchern.

Methode

Analysiert wurden acht deutsche Schulbuchreihen aus verschiedenen Schulformen (gymnasial, nicht-gymnasial), Bundesländern (aus jedem Bundesland mindestens eine zugelassene Schulbuchreihe) und von unterschiedlichen Verlagen (C.C.Buchner, Cornelsen, Klett, Westermann). Ein Kodierschema diente der Einordnung in die Kategorien Visualisierungstyp (*kontinuierlich*, *diskretisiert* und *diskret*) bzw. Aufgabentyp (*Schätzaufgabe* bzw. *keine Schätzaufgabe*) durch zwei unabhängige Rater. Die Interrater-Reliabilität der Kodierung war jeweils hoch (Cohens Kappa $\kappa = .85$ bzw. $\kappa = .82$).

Ergebnisse und Diskussion

Insgesamt wurden $N=2066$ Bruchvisualisierungen betrachtet. Davon waren lediglich 5,4% *kontinuierlich*, während 91,7% *diskretisiert* und 2,9% *diskret* waren. Von $N=1564$ Visualisierungen, die im Rahmen von Aufgaben verwendet wurden, wurden 1,2% in Aufgaben zum *Schätzen* verwendet, während 98,8% *kein Schätzen* erforderten. Die Anteile variierten je nach Themenbereich innerhalb der Bruchrechnung. Insgesamt zeigt sich, dass sich die Schulbücher stark auf diskretisierte Visualisierungen fokussieren und kaum Schätzaufgaben mit Visualisierungen zur Förderung von Größenvorstellungen nutzen.

Der Einsatz von diskretisierten Visualisierungen im Unterricht ist durchaus sinnvoll, weil sie – anders als kontinuierliche Visualisierungen – die exakte Übersetzung zwischen symbolischer und visueller Darstellung ermöglichen. Jedoch kann der fast ausschließliche Einsatz von diskretisierten Visualisierungen als problematisch angesehen werden, da das Potenzial kontinuierlicher Visualisierungen zum Aufbau von Größenvorstellungen (Begolli et al, 2020; Schwarzmeier et al., 2024) nicht ausgeschöpft wird. Zudem werden fast keine expliziten Schätzaufgaben mit Visualisierungen eingesetzt, was zu den oft niedrigen Leistungen von Lernenden in diesem Bereich beitragen könnte.

Insgesamt scheinen die in Schulbüchern verwendeten Visualisierungen und damit verbundenen Aufgabenformate eher wenig geeignet, um den zentralen Bereich der Größenvorstellung von Brüchen zu fördern. Daraus ergeben sich besondere Herausforderungen für Lehrkräfte, das Potenzial vorhandener Aufgaben im Unterricht zu nutzen. Sie sollten gezielt auch Schätzaufgaben in Verbindung mit kontinuierlichen Visualisierungen einsetzen.

Mathematikschulbücher als potenziell implementiertes Curriculum: ein systematisches Review zum Forschungsstand

Constanze Koschwitz, Ann-Katrin van den Ham

Universität Hamburg, Deutschland

Theoretischer Hintergrund

Schulbücher sind eine zentrale Lernressource für den Mathematikunterricht und Grundlage für die Instruktion vieler Lehrkräfte (Mullis et al., 2012; van den Ham & Heinze, 2018). Valverde et al. (2002) konzeptualisierten das Schulbuch als potenziell implementiertes Curriculum, das abstrakte Inhalte des intendierten Curriculums (z.B. Lehrpläne) in das implementierte Curriculum (konkreter Unterricht) übersetzt. Dieses beeinflusst das erreichte Curriculum (z.B. Leistung).

Modell: intendiertes Curriculum-->potenziell implementiertes Curriculum-->implementiertes Curriculum-->erreichtes Curriculum

Dieses Modell wird oft zur Begründung in der Schulbuchforschung herangezogen (vgl. Bayazit, 2013; Bingolbali & Bingolbali, 2019). Fan et al. (2013) zeigen jedoch in einem systematischen Review, dass sich ein Großteil dieser Forschung auf Schulbuchanalysen und -vergleiche konzentriert (vgl. potenziell implementiertes Curriculum), während die Forschung in anderen Bereichen (vgl. implementiertes und erreichtes Curriculum) fragmentiert und inkonsistent ist. Dabei besteht ein großer Forschungsbedarf zum Zusammenhang zwischen Schulbüchern und Schüler*innenleistungen.

Ziel dieses Beitrages ist es, zu analysieren, (1) inwieweit das Modell von Valverde et al. (2002) durch aktuelle Forschung ab 2013 gestützt wird, und (2) die Forschung zur Schulbuchnutzung im Unterricht sowie (3) den Zusammenhang zwischen Schulbüchern und Schüler*innenleistungen zu synthetisieren.

Methode

Es wurde ein systematisches Review gemäß PRISMA-Guidelines durchgeführt (Page et al., 2021). In den Datenbanken ERIC, Web of Science und PsycInfo wurde nach englischsprachigen peer-reviewed Studien im Bereich „Education“ ab 2013 gesucht. Die Suchbegriffe umfassten „mathematics“ oder mathematikspezifische Begriffe (z.B. „algebra“) sowie „textbook“. Nach Ausschluss von Duplikaten wurden 543 Artikel identifiziert. Weitere 259 Studien wurden ausgeschlossen, da das Schulbuch nicht zentraler Untersuchungsgegenstand war, die Bücher nicht für Schüler*innen konzipiert waren oder es sich nicht um empirische Studien handelte. Basierend auf dem Modell von Valverde et al. (2002) wurden Kategorien entwickelt und die 284 Studien von zwei Rater*innen unabhängig kodiert. Anschließend erfolgte die inhaltliche Auswertung der Studien zur Schulbuchnutzung im Unterricht und Schulbucheffekten auf Leistungen.

Ergebnisse und Diskussion

Zur Beantwortung der Forschungsfrage 1 wurden die 284 Artikel nach dem Modell von Valverde et al. (2002) eingeordnet. 26 Artikel analysieren die Übereinstimmung zwischen Lehrplan und Schulbuch (intendiertes Curriculum --> potenziell implementiertes Curriculum), 197 Artikel führen reine Schulbuchanalysen durch (potenziell implementiertes Curriculum), 38 Artikel untersuchen die Nutzung von Schulbüchern durch Lehrkräfte und Schüler*innen (potenziell implementiertes Curriculum -> implementiertes Curriculum) und 23 Artikel Schulbucheffekte auf Leistung, wobei das implementierte Curriculum ausgeblendet wird (potenziell implementiertes Curriculum --> erreichtes Curriculum). Dies zeigt, dass der Forschungsschwerpunkt wie bei Fan et al. (2013) weiterhin auf Schulbuchanalysen liegt, während das Modell zur Schulbuchwirkung weiter kaum empirisch untersucht ist.

Für die Forschungsfrage 2 wurden Studien zur Schulbuchnutzung analysiert. Die Ergebnisse zeigen, dass Schulbücher in vielen Ländern eine wesentliche Ressource für den Mathematikunterricht darstellen. Es gibt jedoch teilweise Abweichungen von Schulbuchinhalten und Unterschiede in der Implementation zwischen Ländern und Lehrkräften. Die meisten Studien beschäftigen sich mit Lehrkräften (n=29) und beruhen auf kleinen Stichproben aus Sekundarschulen. Größere Stichproben basieren ausschließlich auf Selbstberichten. Die Ergebnisse sind daher nur bedingt verallgemeinerbar.

In Bezug auf die Forschungsfrage 3 wurden Studien zu Zusammenhängen zwischen Schulbüchern und Schüler*innenleistung analysiert. Die meisten Studien zeigen Zusammenhänge zwischen den verwendeten Schulbüchern und der Schüler*innenleistung (n=18), während einige Studien keine (n=3) oder nur teilweise (n=2) Zusammenhänge feststellen. Die divergierenden Befunde könnten auf unterschiedliche Methoden und Designs zurückzuführen sein, welche im Beitrag näher diskutiert werden sollen. Hervorzuheben ist, dass der Unterricht als implementiertes Curriculum in keiner der gefundenen Studien berücksichtigt wurde.

Zusammenfassend lässt sich feststellen, dass das Valverde-Modell nicht ausreichend empirisch untersucht ist. Erste Untersuchungen zeigen zwar die Plausibilität der angenommenen Effekte, eine belastbare Überprüfung fehlt jedoch. Die Wirkung des Schulbuchs vom implementierten zum erreichten Curriculum wurde bisher nicht untersucht, obwohl diese Annahme im Modell wesentlich ist. Es ist an der Zeit den Schwerpunkt der Schulbuchforschung von deskriptiven Schulbuchanalysen auf die gesamte Wirkungskette zu erweitern.

Textverständnis als Mediator beim Lösen realitätsnaher mathematischer Aufgaben: Der Einfluss von sprachlicher Komplexität, kognitiven Faktoren und sozialem Hintergrund

Eileen Klotz, Prof. Dr. Timo Ehmke, Prof. Dr. Dominik Leiss

Leuphana Universität Lüneburg, Deutschland

Realitätsnahe Aufgaben haben in den vergangenen Jahren an Bedeutung gewonnen und können als ein primäres Ziel des Mathematikunterrichts angesehen werden (Niss, 2016). Eine Möglichkeit zum Lösen dieser Aufgaben stellt der Modellierungskreislauf von Blum und Leiss (2007) dar. Beim Durchlaufen müssen Schüler:innen eine textbasierte Realsituation verstehen, in Mathematik übersetzen, dadurch die Aufgabe lösen und die Lösung mit der Ausgangssituation abgleichen (Verschaffel et al., 2000). Dies zeigt, dass neben mathematischer Kompetenz Textverständnis benötigt wird (Abedi et al., 2003). Innerhalb des Lösungsprozesses kommt es insbesondere in der für den weiteren Verlauf zentralen Phase des Aufstellens eines Situationsmodells, welches die kognitive Struktur darstellt, auf die der Verstehensprozess ausgerichtet ist (Reusser, 1986), zu Schwierigkeiten (Leiss et al., 2010). Untersuchungen ergaben, dass Aufgaben mit textuellem Input eine niedrigere Lösungsrate aufweisen als solche, die kein Textverständnis erfordern (Duarte et al., 2011). Daher liegt die Vermutung nahe, dass neben Effekten soziodemografischer Merkmale (OECD, 2013) und kognitiver Faktoren (Fuchs et al., 2006) auf die Lösung fehlendes Textverständnis zu Problemen führt.

Es wurde bereits untersucht, ob Merkmale sprachlicher Komplexität die Leistung von Schüler:innen beeinflussen. Aktuelle Ergebnisse sind jedoch uneindeutig. Bislang ist zudem nicht bekannt, ob der Einfluss des situationsbezogenen Textverstehens durch seine bedeutsame Rolle im Lösungsprozess als Mediator für soziodemografische Merkmale und kognitive Faktoren im Hinblick auf die mathematische Lösung dient und welche Rolle systematische Variation sprachlicher Komplexität in diesem Zusammenhang spielt. Ziel war herausfinden, ob eine Fokussierung auf situationsbezogenes Textverständnis und sprachliche Umformulierung von Texten als Unterstützungsoption für Schüler:innen angesehen werden kann, um den Umgang mit realitätsnahen Aufgaben im Mathematikunterricht voranzubringen. Folgende Forschungsfragen wurden beantwortet:

1. Wie unterscheidet sich der Lösungsdurchschnitt situationsbezogener Textverständnisaufgaben und mathematischer Aufgaben in Abhängigkeit von der Variation der sprachlichen Komplexität des Textes?
2. Inwieweit kann die Varianz in der mathematischen Lösung durch soziodemografische Merkmale und kognitive Faktoren erklärt werden?
3. Inwieweit spielt das situationsbezogenen Textverständnis auf verschiedenen Sprachniveaus eine vermittelnde Rolle in diesem Kontext?

Methode

Es wurde eine experimentelle Studie in Form eines Papier-Bleistift-Tests durchgeführt, an welcher 428 Schüler:innen aus 26 Klassen der Jahrgänge 7 bis 10 einer kooperativen Gesamtschule teilnahmen. Drei Aufgaben zu linearen Funktionen wurden in verschiedenen komplexen Realkontexten entwickelt und jeweils auf drei Sprachanforderungsniveaus (SPA1, SPA2, SPA3),

gekennzeichnet durch deren sprachlichen Komplexität, variiert. Um Klasseneffekte auszuschließen, wurden drei verschiedene Versionen der Testhefte nach dem Zufallsprinzip ausgeteilt. Auf jeden Text folgten Fragen zum situationsbezogenen Textverständnis sowie eine mathematische Frage. Zusätzlich wurden soziodemografische Merkmale (Geschlecht, kulturelle Ressourcen, Zuwanderungshintergrund, elterlicher Bildungsabschluss) sowie sprachliche Fähigkeiten mittels eines Sprachtests und innermathematische Fähigkeiten zu linearen Funktionen mittels eines Mathematiktests erhoben.

Ergebnisse

Generell kann gezeigt werden, dass situationsbezogenes Textverständnis und mathematische Lösung bei Texten mit geringer linguistischer Komplexität signifikant besser sind. Der Zusammenhang zwischen soziodemografischen Merkmalen, kognitiven Fähigkeiten und der mathematischen Lösung wurde durch Regressionsanalysen untersucht. Ergebnisse zeigen, dass innermathematische Fähigkeiten ($\beta=0,598$, $p<0,001$) und generelle Sprachkompetenz ($\beta=0,303$, $p<0,001$) die stärksten Prädiktoren für die mathematische Lösung sind. In Mediatoranalysen wurde zudem geprüft, inwieweit situationsbezogenes Textverständnis von Texten verschiedener sprachlicher Komplexität diese Zusammenhänge mediiert. Ergebnisse zeigen, dass situationsbezogenes Textverständnis (SPA1: $\beta=0,253$, $p<0,001$, SPA2: $\beta=0,292$, $p<0,001$, SPA3: $\beta=0,168$, $p<0,01$) nach innermathematischen Fähigkeiten ($\beta=0,401$, $p<0,001$) der bedeutsamste Prädiktor für die mathematische Lösung ist und die Beziehung zwischen dem Einfluss kultureller Ressourcen, kognitiver Faktoren und der Lösungskompetenz der Aufgaben mediiert. Unterschiede zwischen Sprachniveaus können nicht festgestellt werden. Insgesamt können durch den Einbezug des situationsbezogenen Textverständnisses 84,7% der Varianz in der mathematischen Lösung aufgeklärt werden.

Die Ergebnisse weisen die Relevanz sprachlicher Komplexität mathematischer Texte für die Lösung auf. Zudem wird durch die vermittelnde Rolle deutlich, dass auf Grund des starken Einflusses auf die mathematische Lösung ein verstärkter Fokus im Unterricht auf fachbezogene Textverständnisübungen gelegt werden sollte.

Kognitive, domänenspezifische Funktionen und ihre Bedeutung für mathematische Kompetenzen in der Grundschule

Vroni Hischa¹, Frank Niklas²

¹Katholische Universität Eichstätt-Ingolstadt; ²Ludwig-Maximilians-Universität München

Die Entwicklung mathematischer Kompetenzen wird durch eine Vielzahl an kognitiven Funktionen geprägt (z.B. Ashkenazi & Adi, 2023; Purpura et al., 2017; Xenidou-Dervou et al., 2018). Bisher wurden dabei häufig das Approximate Number System (ANS) und Mapping als domänenspezifische Funktionen untersucht. Das ANS ist ein angeborenes, kognitives System, welches unpräzise Repräsentationen von großen, nicht-symbolischen Mengen ermöglicht (Halberda & Odic, 2015; Starr & Brannon, 2015). Auf Basis dessen können solche Mengen beispielsweise bezüglich ihrer Mächtigkeit miteinander verglichen werden. Unter Mapping wird die Zuordnung zwischen nicht-symbolischen und symbolischen Repräsentationen von Mengen verstanden (Mundy & Gilmore, 2009). Aufgrund einer heterogenen Befundlage wird die Bedeutung des ANS für mathematische Kompetenzen im Vor- und Grundschulalter diskutiert (z.B. Chen & Li, 2014; Chu et al., 2015; Coolen et al., 2022; Gimbert et al., 2019). Da das Mapping nicht-symbolische Mengenrepräsentationen nutzt, wird eine medierende Rolle des Mapping im Zusammenhang zwischen ANS und mathematischen Kompetenzen vermutet, jedoch ist auch hier die Befundlage heterogen (z.B. Fazio et al., 2014; Gimbert et al., 2019; Jang & Cho, 2018). Ziel der Studie ist es daher, die Bedeutung von ANS und Mapping für mathematische Kompetenzen in der Grundschule näher zu beleuchten.

Insgesamt nahmen an der Studie 180 Dritt- und Viertklässler*innen (97 weiblich) im Alter von 8 Jahren und 2 Monaten bis 11 Jahren und 11 Monaten ($M = 9$ Jahre und 5 Monate; $SD = 8$ Monate) teil. Zunächst bearbeiteten alle Grundschüler*innen in Gruppensitzungen (ca. 90 min) insgesamt 9 Subtests der tabletgestützten Applikation EI-MAG (Oesterlen et al., 2016) zur Erfassung der Kapazität der drei Arbeitsgedächtnis-Komponenten nach dem Modell von Baddeley (1986, 2003), zentrale Exekutive, visuell-räumlicher Skizzenblock und phonologische Schleife. Nach einer kurzen Pause bearbeiteten die Drittklässler*innen den DEMAT 2+ (Krajewski et al., 2020) und die Viertklässler*innen den DEMAT 4 (Gölitz et al., 2006). In Einzelsitzungen (ca. 35 min) bearbeiteten alle Grundschüler*innen den Subtest Nonverbale Intelligenz des BUEGA-II (Esser et al., 2021), anschließend nicht-symbolische Mengenvergleichsaufgaben unter Verwendung der Software Panamath (Halberda et al., 2008) an einem Laptop und Zahlenstrahlzuordnungsaufgaben am Tablet (ähnlich zu Siegler & Opfer, 2003). Bei den nicht-symbolischen Mengenvergleichsaufgaben mussten jeweils zwei Punktwolken verglichen werden und per Tastendruck angegeben werden, welche von beiden mehr Punkte enthält. Bei den Zahlenstrahlzuordnungsaufgaben wurden Zahlen präsentiert, welche auf einem unskalierten Zahlenstrahl platziert werden mussten.

Es zeigt sich, dass das ANS und Mapping in der dritten und vierten Klasse unter Berücksichtigung von Alter, Geschlecht, Migrationshintergrund und der Kapazität des Arbeitsgedächtnisses signifikante Prädiktoren für mathematische Kompetenzen sind. Ähnliche Ergebnisse berichten Fazio et al. (2014) für das ANS in der fünften Klasse und Schneider et al. (2018) in ihrer Meta-Analyse zum Mapping zwischen 4 und 14 Jahren. Mediationsanalysen zeigten keine Mediation durch Mapping im Zusammenhang zwischen ANS und mathematischen Kompetenzen bei den untersuchten Dritt- und Viertklässler*innen. Dies steht im Einklang mit den Ergebnissen von Gimbert et al. (2019), welche eine solche Mediation durch Mapping zwar im Kindergartenalter berichten, aber in der zweiten Klasse nicht mehr. Vor dem Hintergrund des heterogenen Forschungsstandes werden die Ergebnisse unter Berücksichtigung methodischer Unterschiede in den verschiedenen Studien weiter diskutiert und eingeordnet. Die vorliegende Studie trägt mit ihren Befunden zum aktuellen Forschungsstand zu kognitiven, domänenspezifischen Funktionen und deren Bedeutung für mathematische Kompetenzen bei. Auf Basis empirischer Studien wie dieser kann in künftigen Studien untersucht werden, ob und inwiefern die Erfassung von ANS und Mapping zum Beispiel im Rahmen von Screenings für mathematische Kompetenzen eingesetzt werden kann.

PS-74: Social Background, Educational Success, and Integration: The Impact of Education on Inequality and Aspiration

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 150
Chair der Sitzung: Ann-Marie Sommerfeld

English Track

ENTFÄLLT Soziale Herkunft und Bildungserfolg: Zur empirischen Gültigkeit ökonomischer Axiome.

Benjamin Gröschl, Katja Scharenberg

Ludwig-Maximilians-Universität München, Deutschland

Die soziale Herkunft eines Kindes ist in der empirischen Bildungsforschung ein zentraler Prädiktor für den Bildungserfolg (Breen et al., 2012; Marks & O'Connell, 2021; Müller & Gangl, 2006; Sirin, 2005). Herkunftsbedingte Disparitäten in den Schulleistungen von Grundschüler:innen (Herrmann et al., 2022) und die Einflüsse der elterlichen Bildung auf den Übergang in die Sekundarstufe (Dumont et al., 2019) verdeutlichen die direkte Bedeutung sozialer Herkunft. Zusätzlich wirkt sich die soziale Herkunft auf die allgemeine und psychosoziale Gesundheit (Fliesser et al., 2019; KiGGS Study Group et al., 2014; Lampert et al., 2018) oder das Verletzungsrisiko (Yuma-Guerrero et al., 2017) aus, was indirekt auch die schulische Kompetenzentwicklung beeinflusst. Mit Blick auf diese multikriterialen Outcomes verdeutlicht die soziale Herkunft so wie kaum ein anderes Konzept den interdisziplinären Charakter empirischer Bildungsforschung.

Der Bildungsstand der Eltern und deren sozioökonomischer Status (SES) gelten, auch aufgrund einschlägiger Empfehlungen (APA, 2007; OECD, 2018; U.S. Department of Education, 2012; U.S. Department of Justice, 2014), als die meistverwendeten Kontrollvariablen der empirischen Bildungsforschung (Sirin, 2005). Die „Big Three“ der sozialen Herkunft (Rofe, 2021; Sirin, 2005; White, 1982) umfassen Einkommen, Bildungsabschluss und Beruf der Eltern. Das Einkommen spiegelt wirtschaftliche Ressourcen und sozialen Status wider, der Bildungsabschluss das kulturelle Kapital und der Beruf das soziale Ansehen und die Verdienstmöglichkeiten (Becker, 1964; Ditton et al., 2019; Eccles, 2005; Lareau, 2000; Marks, 2021; Minzer, 1958; Sirin, 2005).

Unser Beitrag untersucht die empirische Tragfähigkeit ressourcenorientierter Handlungsmodelle in der empirischen Bildungsforschung. Hierfür werden zwei grundlegende ökonomische Theoreme mithilfe semiparametrischer generalisierter Regressionsmodelle überprüft. Datengrundlage (n=4830) ist das Nationale Bildungspanel (NEPS; Startkohorte 2; Artelt & Sixt, 2023; Blossfeld & Roßbach, 2019).

Im ersten Analyseschritt wird die Elastizität von Bildungsinvestitionen gegenüber existenziellen Bedürfnissen und subjektiven Präferenzsystemen vor dem Hintergrund ökonomischer Theorieansätze betrachtet. Bei knappen oder unzureichenden Ressourcen haben existenzielle Bedürfnisse (z.B. Essen, Wohnen und Sicherheit) Vorrang vor sekundären Bedürfnissen (z.B. kulturelle Teilhabe und individuellem Bildungserfolg) (Becker, 1964; Maslow, 1943). Zudem wird der hieraus abgeleitete "Threshold"-Effekt identifiziert, bei dem bis zu einer Mindestausstattung an ökonomischen Ressourcen der Zusammenhang zwischen SES und der erzielten Mathematikkompetenz in Stärke und Erklärungskraft eine sehr hohe Streuung aufweist, während er darüber hinaus homogener und in geringerem Maße variiert und so parametrische Schätzungen ermöglicht.

Der Schwellenwert wird durch verschiedene parametrische und parameterfreie iterative Maximierungsverfahren empirisch abgeleitet und dient als Grundlage weiterer partiell-linearer Modellierungen. Unterhalb des Schwellenwertes ($24 < ISEI < 32$, $p < 0.001$) zeigt sich so in den meisten Bildungsgruppen keine signifikante parametrische Abhängigkeit zwischen dem SES und der Mathematikkompetenz von Schüler:innen der 4. Klasse. Dagegen bestehen aber deutliche kategorielle Unterschiede in der Mathematikkompetenz in Abhängigkeit vom Bildungsniveau der Eltern ($-0.38 < \beta_0 < 0.14$; $p < 0.001$). Fehlende finanzielle Ressourcen scheinen, abhängig vom Bildungsniveau, unterschiedlich erfolgreich substituiert zu werden. Da in der untersten Bildungsgruppe keine Substitution unzureichender ökonomischer Ressourcen durch das kulturelle Kapital der Eltern zu erfolgen scheint, sind die nichtlinearen Marginaleffekte des SES auf die Mathematikkompetenz hier besonders ausgeprägt ($R^2 = 0.33$; $0.01 [ISEI = 60] < \partial(\beta_1) < 0.03 [ISEI = 20]$; $p < 0.001$).

Im zweiten Analyseschritt wird die Gültigkeit des ökonomischen Gesetzes vom abnehmenden Grenznutzen (Gossen, 1854) in der Bildungsforschung untersucht. Der zusätzliche Nutzen von Bildungsinvestition (z.B. Nachhilfeunterricht, mediale und technische Ausstattung) nimmt mit dem Konsum jeder weiteren Einheit ab. Dieser nichtlineare Sättigungseffekt lässt sich außer bei akademischen Eltern in allen Bildungsgruppen nachweisen ($0.006 [ISEI = 60] < \partial(\beta_1) < 0.16 [ISEI = 30]$; $p < 0.001$).

Die Ergebnisse verdeutlichen die empirische Relevanz mikroökonomischer Theorien in der empirischen Bildungsforschung. Sie betonen zum einen die Notwendigkeit interdisziplinärer Perspektiven, zum anderen aber auch die Notwendigkeit, diese nichtlinearen Effekte statistisch zu berücksichtigen. Aus makroökonomischer und gesellschaftspolitischer Perspektive sind die aufgezeigten empirischen Zusammenhänge ebenfalls von großer Relevanz, weisen sie doch deutlich auf besonders hohe nichtlineare "Renditen" von Bildungsinvestitionen vor allem in den niedrig gebildeten und zusätzlich ökonomisch schwachen Familien hin.

Schooling Exposure and Achievement Inequality – Revisiting the ‘Great Equalizer’ Hypothesis with the Differential Exposure Approach

Giampiero Passaretta¹, Jan Skopek², Joseph Workman³

¹Pompeu Fabra University; ²Trinity College Dublin; ³University of Missouri Kansas City, United States of America

Background

Understanding the role of schools for social inequality in educational outcomes is a central concern in sociology. Critical perspectives view schools as institutions of social reproduction that reinforce inequality by allocating students to unequal learning environments. On the other hand, positive perspectives view schools as ‘equalizers’ because they provide more equal learning opportunities than do non-school environments. The equalizer perspective has received support from seasonal comparison design (SCD) studies in the United States that found SES achievement gaps grow during non-school periods (pre-school and summer breaks) and decline/are stable during school years (Alexander, et al. 2007; Downey et al., 2004; von Hippel & Hamrock, 2019; von Hippel et al., 2018).

However, evidence on school equalization in the United States is more mixed than ever. Recent studies suggest that SES equalization occurs in early childhood rather than in later school grades (von Hippel et al. 2018; Quinn et al. 2016). Additionally, findings on seasonality in learning inequality exhibit a great deal of instability across datasets (Workman et al. 2023).

Exposure approaches, which separate effects of exposure to school from the effects of aging, have been proposed as complementary strategies to identify the role of schooling for inequality in educational outcomes. Adopting a differential exposure approach (DEA) that identifies exposure effects of schooling through random variation in test dates, Passaretta & Skopek (2021) demonstrated strong effects of schooling on learning of first graders in Germany. In contrast to US findings, they did not find evidence for schooling to equalize inequality in learning by socioeconomic status.

In our study, we apply the DEA to data on students from the United States. In doing so, we extend the study of school effects in the US context and assess whether the DEA and SCD methods yield similar findings.

Research question

How does exposure to schooling influence SES disparities in mathematics and reading achievement?

Data and methodology

We apply the DEA to data from the Early Childhood Longitudinal Study Kindergarten Cohort of 2011 (ECLS-K 2011), the latest nationally-representative data used to reconstruct seasonal patterns of inequality (von Hippel et al. 2018). We exploit fall and spring measures of reading and mathematics achievement for kindergarten and first grade (four waves in total). Our exposure approach estimates the effects of school exposure on learning in contrast to the effects of being older by comparing, within a school year, students who were the same age at the time of testing but were tested at (randomly) different times in the school year (in fall or spring). Student learning was assessed in mathematics and reading with test scores scaled to be vertically integrated across waves. To ensure exogeneity of birth dates, a requirement of the DEA, we restricted the sample to children with compliant school enrollment. Crucially, we tested if schooling effects were moderated by students' socioeconomic background (SES) as the equalization hypothesis would translate to low SES children benefiting more from school exposure.

Results

Preliminary findings from our DEA models show considerable positive effects of schooling for mathematics and reading achievement which are numerically consistent with Passaretta & Skopek's (2021) results for Germany. Moreover, preliminary findings for kindergarten indicate support for the equalization perspective as evident in a negative interaction between SES and school exposure. Hence, exposure to kindergarten (regardless of aging) effectively reduced SES gaps in achievement. However, we found no evidence for school equalization in grade one.

Based on those preliminary findings we conclude that school exposure fosters learning and equalizes gaps in achievement by SES background. Those equalization effects, however, seem to concentrate on early phases of schooling and fade at later stages of schooling.

The Effect of Schooling on Parental Integration

Ann-Marie Sommerfeld

Friedrich-Schiller-Universität Jena, Deutschland

Exploiting the age-at-enrollment policies in 16 German states as exogenous source of variation, I examine whether the schooling of the oldest child in a migrant household affects parents' integration. My analysis links administrative records on primary school enrollment cutoff dates with micro data from the German Socioeconomic Panel (SOEP). Using a regression discontinuity design around the school enrollment cutoff and an instrumental variable approach I show that children's schooling improves the integration of parents along several dimensions, such as labor market outcomes, financial worries, and German language skills. Labor market outcomes are most positively affected for mothers. Additional analysis of underlying mechanisms suggests that results are driven by gains in disposable time and exposure to the German language and culture.

PS-75: Mitbestimmung und Interaktion im Unterricht

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 157

Chair der Sitzung: Victoria Wiederseiner

Mitbestimmung in der Grundschule für alle Kinder? Wie Grundschulkinder ihre Mitbestimmungsmöglichkeiten im schulischen Kontext wahrnehmen

Victoria Wiederseiner, Selma Ceivan, Vanessa Jandl

Institut für Grundschulforschung, Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Ausgehend von der UN-Kinderrechtskonvention (2019) werden Kinder als Rechtssubjekte betrachtet, denen Kinderrechte zugesprochen werden – so auch das Recht auf Mitbestimmung (Art. 12, UN-KRK) (BMFSJ 2019). Nehmen aber alle Kinder Mitbestimmung gleich wahr? In der vorliegenden Studie wird Mitbestimmung in Anlehnung an Ertl, Martschinke und Grüning (2022) als kollektiver Aushandlungsprozess (hier vorrangig zwischen Lehrkräften und Schüler*innen) entlang von vier theoretisch konzeptualisierten Facetten mit einem unterschiedlich hohen Maß an Mitbestimmung operationalisiert: (1) *Informiert- und Gehört-werden*, (2) *Mitplanen* und *Mitberaten*, (3) *Mitwirken* und *Mitgestalten* sowie (4) *Mitentscheiden*. Nationale und internationale Studien zeigen, dass die Wahrnehmung der Mitbestimmung zwischen den schulischen Akteur*innen unterschiedlich ausfällt (Grundmann et al. 2000; Forde et al. 2018). So nehmen Kinder eher in geringem Umfang und in geringer Qualität Mitbestimmungsmöglichkeiten in der Schule wahr – im Vergleich zu ihren Lehrkräften, die die eröffneten Mitbestimmungsmöglichkeiten höher einschätzen. Die derzeitige Forschungslage weist jedoch deutliche Desiderata auf, ob die theoretisch gebildeten Facetten sich auch empirisch bestätigen lassen und inwiefern Unterschiede in der individuellen Wahrnehmung der Kinder auf ihre Mitbestimmungsmöglichkeiten im schulischen Kontext bestehen bzw. wie sie erklärt werden können.

Der Beitrag geht daher drei zentralen Forschungsfragen (FF) nach:

1. Lassen sich die theoretisch definierten Mitbestimmungsfacetten in einer Drei-Faktorenstruktur abbilden?
2. In welchem Ausmaß nehmen Grundschulkinder ihre Mitbestimmungsmöglichkeiten entlang der Facetten wahr?
3. Lassen sich hinsichtlich der wahrgenommenen Mitbestimmungsmöglichkeiten der Kinder Cluster identifizieren und unterscheiden sich diese in Abhängigkeit von Personen- und Klassenkompositionsmerkmalen?

Im Rahmen von Gruppendiskussionen wurden 173 Kinder aus den Bundesländern Bayern, Sachsen und Hessen (46% weiblich, 10% Klasse 1, 38% Klasse 2, 21% Klasse 3, 31% Klasse 4) neben ihrer kollektiven Wahrnehmung, auch nach ihren individuell wahrgenommenen Mitbestimmungsmöglichkeiten im schulischen Kontext mittels Fragebogen befragt. Die wahrgenommenen Mitbestimmungsmöglichkeiten der Kinder wurden anhand der theoretisch definierten Facetten kindgerecht auf einer vierstufigen Skala (1=selten bis 4=häufig) erfasst. Nach Ceivan und Martschinke (2024, i. D.) werden Facette 2 und 3 in der Datenauswertung aufgrund mangelnder inhaltlicher Trennschärfe gebündelt. Die faktorielle Struktur der Daten wurde daher mit drei Facetten mittels einer konfirmatorischen Faktorenanalyse (CFA) geprüft (FF1). Nach den deskriptiven Analysen (FF2) wurden hierarchische Clusteranalysen (ward method) auf den drei Facetten der Mitbestimmungsmöglichkeiten berechnet (FF3).

Die Mitbestimmungsfacetten lassen sich nach den Ergebnissen der CFA in einer Drei-Faktorenstruktur abbilden (FF1). Das Drei-Faktorenmodell weist eine akzeptable bis gute Modellpassung der empirischen Daten ($\chi^2(10) = 163.75$, $p = .095$, $RMSEA = 0.08$, $CFI = 0.98$, $TLI = 0.91$, $SRMR = 0.03$) nach den Kriterien von Hu und Bentler (1999) auf. Die befragten Schüler*innen ($N = 173$) geben an, häufig Mitbestimmung hinsichtlich des „Informiert- und Gehört-werdens“ (Facette 1) ($M = 3.47$; $SD = .49$) wahrzunehmen, während sie beim „Mitplanen/Mitberaten & Mitwirken/Mitgestalten“ (Facette 2) ($M = 2.85$; $SD = .73$) und „Mitentscheiden“ (Facette 3) ($M = 2.69$; $SD = 1.05$) weniger wahrgenommene Mitbestimmungsmöglichkeiten angeben (FF2). In den hierarchischen Clusteranalysen zu den Mitbestimmungsfacetten konnten zwei Cluster identifiziert werden: Schüler*innen des Clusters 1 ($n = 70$) kennzeichnen sich durch weniger wahrgenommene Mitbestimmungsmöglichkeiten hinsichtlich Facette 1 ($M = 3.32$), Facette 2 ($M = 2.40$) und Facette 3 ($M = 1.59$), während Schüler*innen des Clusters 2 ($n = 97$) mehr Mitbestimmungsmöglichkeiten (Facette 1: $M = 3.58$; Facette 2: $M = 3.17$; Facette 3: $M = 3.45$) wahrnehmen. Die Cluster unterscheiden sich hinsichtlich der wahrgenommenen Mitbestimmungsmöglichkeiten signifikant (Facette 1: $F(1,165) = 12.40$, $p < .001$; Facette 2: $F(1,165) = 61.93$, $p < .001$; Facette 3: $F(1,165) = 570.52$, $p < .001$). Bezüglich der Personenmerkmale Geschlecht und der Klassenkompositionsmerkmale (Schule, Bundesland und Klasse) zeigen sich hingegen keine signifikanten Unterschiede zwischen den Clustern (FF3).

Die Ergebnisse zeigen, dass die Grundschulkinder am häufigsten in Facette 1 Mitbestimmungsmöglichkeiten wahrnehmen. Auch zwischen den beiden Clustern werden deutliche Unterschiede in den Mitbestimmungsmöglichkeiten hinsichtlich Facette 2 und 3 erkennbar. Besonders in Cluster 1 zeichnet sich ab, dass Facette 3 in der Wahrnehmung der Kinder noch unterrepräsentiert ist. Das verweist auf die Notwendigkeit, alle Mitbestimmungsfacetten in den Blick zu nehmen und weitere Personen- und Klassenkompositionsmerkmale zur Erklärung der Unterschiede hinzuzuziehen.

Zusammenhang zwischen der wahrgenommenen Mitbestimmung und Merkmalen der Unterrichtsqualität in der Grundschule unter Berücksichtigung weiterer Prädiktoren

Leonora Gerbeshi¹, Sonja Ertl²

¹Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland; ²Universität Augsburg, Deutschland

Mitbestimmung in der (Grund)Schule ist insbesondere aus demokratiepädagogischer (Baumgardt, 2022) und kinderrechtlicher Perspektive (CRC, 2009) von Bedeutung. Dabei wird Mitbestimmung als kollektiver Prozess verstanden, der von der individuellen Selbstbestimmung abzugrenzen ist (Müller-Kuhn & Häbig, 2022), wenngleich ein Zusammenhang zwischen Mit- und Selbstbestimmung besteht (Pupeter & Schneekloth, 2018). Mitbestimmung wird im Folgenden in Anlehnung an Ertl, Martschinke & Grüning (2022) über drei Facetten operationalisiert: Informiert- und Gehört-werden, Meinung und Abstimmung sowie Diskussion und Austausch, die zwischen der Organisations- (z. B. Klassenregeln) und Unterrichtsebene (z. B. Unterrichtsthemen) im Klassenkontext unterschieden werden.

Es ist zu erwarten, dass die Mitbestimmung in Zusammenhang mit der Unterrichtsqualität steht, insbesondere mit den Dimensionen *Konstruktive Unterstützung* und *Klassenführung* (Klieme, 2006). Eine konstruktive Unterstützung impliziert z. B. das Einbinden von Interessen der Kinder (Decristan et al., 2020). Zudem könnte die Mitbestimmung in einem Zusammenhang mit der Klassenführung stehen, da sich eine gute Klassenführung beispielsweise dadurch auszeichnet, dass alle beim Lernen beteiligt sind, eine klare Kommunikation herrscht sowie Regeln und Routinen begründet werden (ebd.).

In (intern-)nationalen Studien zeichnet sich ab, dass es in der (Grund)Schule und auf unterrichtlicher Ebene an Mitbestimmungsmöglichkeiten mangelt (Forde et al., 2018; Müthing et al., 2018). In bisherigen Analysen (anonymisiert) zeigt sich, dass Grundschulkinder signifikant mehr Mitbestimmung auf Organisations- als auf Unterrichtsebene wahrnehmen und sie sich informiert und gehört fühlen, Austausch- und Abstimmungsprozesse jedoch deutlich weniger stark ausgeprägt wahrnehmen. Der Effekt personenbezogener Prädiktoren, z.B. Geschlecht, auf die wahrgenommene Mitbestimmung ist nicht ausreichend erforscht (Beck & Meusburger, 2023). Bezüglich der Unterrichtsqualität konnte für die Sekundarstufe gezeigt werden, dass eine hohe wahrgenommene Mitbestimmung der Schüler*innen mit einer hohen Qualität der Unterrichtsgestaltung einhergeht (Holtappels, 2004).

Es sind keine Studien bekannt, die den Zusammenhang zwischen der wahrgenommenen Mitbestimmung von Grundschulkindern und der Unterrichtsqualität untersuchen.

Der Vortrag geht den Fragen nach, ob es (1) einen Zusammenhang zwischen der wahrgenommenen Mitbestimmung und der Unterrichtsqualität (konstruktive Unterstützung und Klassenführung) gibt und (2) welche Rolle dabei personenbezogenen Merkmale auf Seiten der Schüler*innen (Geschlecht, Migrationsgeschichte), die wahrgenommene Selbstbestimmung sowie Klassengröße spielen.

Im Rahmen eines quantitativen Querschnittsdesigns wurden 472 Grundschulkinder der 3. Jahrgangsstufe aus 28 Klassen mittels eines Fragebogens befragt (50.1 % weiblich; 32.6 % Migrationsgeschichte). Die Reliabilitäten der sechs selbst konstruierten Skalen zu den Mitbestimmungsfacetten auf Organisations- und Unterrichtsebene liegen im akzeptablen bis guten Bereich ($\alpha = .60$ bis $.86$). Die Erfassung der Unterrichtsqualität erfolgte über etablierte Skalen zur wahrgenommenen konstruktiven Unterstützung (Fauth, 2014) sowie Klassenführung (Fauth, 2014). Die wahrgenommene Selbstbestimmung wurde über eine Skala von Röder & Kleine (2009) erhoben.

Zur Beantwortung der Forschungsfrage wurden Mehrebenenanalysen berechnet, indem das Geschlecht, die Migrationsgeschichte und die Selbstbestimmung als Prädiktoren auf Individualebene modelliert wurden, wohingegen die wahrgenommene konstruktive Unterstützung und Klassenführung als aggregierte Prädiktoren auf Level 2 in das Modell integriert wurden. Alle metrischen Variablen wurden z-standarisiert.

Auf Organisationsebene erweist sich die konstruktive Unterstützung als ein signifikanter Prädiktor für alle drei Facetten der wahrgenommenen Mitbestimmung, wobei auf Unterrichtsebene die konstruktive Unterstützung lediglich für die Skala Informiert- und Gehört- werden ein relevanter positiver Prädiktor ist. Die Klassenführung stellt für die Mitbestimmungsfacetten Diskussion und Austausch auf Organisationsebene einen negativen signifikanten Prädiktor dar. Auf Unterrichtsebene erweist sich die Klassenführung als nicht relevant für die wahrgenommene Mitbestimmung. Bezüglich weiterer Prädiktoren handelt es sich bei der Migrationsgeschichte und Selbstbestimmung um signifikante Effekte für alle Mitbestimmungsfacetten sowohl auf Organisations- als auch Unterrichtsebene.

Kinder nehmen in einem Unterricht, in dem sie mehr Mitbestimmung wahrnehmen, die Klassenführung trotz veränderter Struktur- und Kontrollmechanismen (Martschinke & Grüning, 2022) insgesamt weder als effektiver noch ineffektiver wahr. Mitbestimmung scheint daher nicht mit mehr Störungen und weniger Aufmerksamkeit (Bleck & Lipowsky, 2019) seitens der Kinder einherzugehen. Die Ergebnisse werden im Vortrag differenziert diskutiert und weitere Forschungsdesiderata aufgezeigt werden.

Viel Lob für Esma? Lernenden- und Lehrendenmerkmale als Prädiktoren der Interaktionsqualität in Grundschulklassen

Josephine Gatzweiler, Yvonne Erhardt, Sina Schürer, Stefanie van Ophuysen
Universität Münster, Deutschland

Theoretischer Hintergrund

Die Beziehung zwischen Lehrkräften und Schüler*innen ist entscheidend für Lernerfolg, Lernmotivation, Sozialverhalten und auch emotionales Wohlbefinden der Lernenden (Schübel & Winkelhofer, 2021; Hamre et al. 2013). Die Beziehungsqualität ergibt sich u.a. aus der Interaktionsgestaltung durch die Lehrkraft, wobei Hamre et al. (2013) die drei Dimensionen Emotional Support, Classroom Organisation und Instructional Support unterschieden. Unser Fokus liegt auf der Dimension Emotional Support, welche sowohl positive als auch negative Interaktionen zwischen Lehrkraft und Schüler*innen umfasst und die Berücksichtigung der Lernendenmerkmale und ihrer Bedürfnisse betont (Hamre et al., 2013). Diese Interaktionen lassen sich weiterhin in anerkennende (z.B. loben, Trost) und verletzende Interaktionsformen (z.B. schimpfen, ignorieren) unterteilen (Winkelhofer & Prengel, 2018). Doch nicht alle Schüler*innen erleben die gleiche Lehrkraftinteraktion. So zeigen erste empirische Befunde, dass Lernenden mit Förderbedarf im Verhalten, Jungen und leistungsschwache Schüler*innen weniger positive Interaktionen erleben (Huber, 2021; McGrath & Van Bergen, 2014), während Schüler*innen mit Migrationshintergrund vergleichsweise häufig positive Lehrkraftrückmeldungen zu ihrem Verhalten erhalten (Schwab et al., 2022).

Ein Lehrkraftmerkmal, welches pädagogische Beziehungen beeinflussen kann, so unsere Annahme, ist die inklusive Haltung, verstanden als Wertschätzung von Individualität, Vielfalt und Gemeinschaft als Leitlinien des professionellen Handelns. Mit Fokus auf Emotional Support, also die Beachtung der Merkmale und Bedürfnisse der individuellen Schüler*innen bei der Interaktionsgestaltung, erscheint uns besonders der Wert der Individualität bedeutsam. Dieser umfasst die bedingungslose Akzeptanz und Wertschätzung der Individualität jedes einzelnen Kindes unabhängig von Stärken und Schwächen (Kullmann et al. 2014). Bisher fehlen empirische Befund zum Zusammenhang zwischen inklusiver Haltung und Interaktionsqualität.

Fragestellungen

- 1) Stehen Individualmerkmale der Schüler*innen (Migration, Geschlecht, Leistung und Emotional-soziale Fähigkeiten) im Zusammenhang mit Lehrkraft-Schüler*in-Interaktionen?
- 2) Ist die inklusive Haltung der Lehrkraft ein Prädiktor für die Qualität der Interaktion mit den Schüler*innen?
- 3) Fällt der Zusammenhang zwischen Individualmerkmalen der Schüler*innen und der Lehrkraft-Schüler*in-Interaktion durch die inklusive Haltung schwächer aus, weil alle Lernenden gleich wertschätzend und positiv behandelt werden?

Methode

Zur Erfassung der positiven und negativen Lehrkraft-Schüler*innen-Interaktion als Indikator der Dimension Emotional Support gaben für jedes Kind je sieben Mitschüler*innen auf einer fünfstufigen Ratingskala (sehr wenig bis sehr viel) an, wie häufig es von der Klassenlehrkraft gelobt bzw. mit diesem Kind geschimpft wird. Diese Werte wurden für jedes Kind zu einem Loben- und Schimpfen-Wert gemittelt. Der Migrationshintergrund wurde über die Familiensprache erfasst. Für jedes Kind liegen weiterhin Lehrkraftangaben zu den Schulleistungen und den emotional-sozialen Kompetenzen vor. Die inklusive Haltung (Subdimension

Individualität) wurde über ein selbstentwickeltes Instrument (7 Items, z.B. „Jede*r Schüler*in ist ein gleichermaßen wertvoller Teil der Klassengemeinschaft“, Cronbach's $\alpha = .81$) erhoben.

Mit Daten von ca. 800 Schüler*innen aus 39 Klassen werden hierarchische lineare Modelle zur Vorhersage der Interaktionsqualität (Loben, Schimpfen) angepasst. Neben den Individualmerkmalen als Level-1-Prädiktoren und der inklusiven Haltung als Level-2-Prädiktor wird im Falle von signifikanten Variationen der Steigungsparameter der Level-1-Prädiktoren mittels Cross-Level-Interaktionen überprüft, inwiefern die inklusive Haltung der Lehrkräfte diese erklären kann.

Ergebnisse

Erste bivariate Analysen zeigen signifikante Zusammenhänge zwischen Loben und den Prädiktoren Leistung der Lernenden ($r = .345$, $p < .001$, $N = 784$) und emotional-sozialen Fähigkeiten ($r = .475$, $p < .001$, $N = 785$). Weiterhin unterscheidet sich die wahrgenommene Lobhäufigkeit in Abhängigkeit vom Geschlecht (MJunge: 3.01, MMädchen: 3.34, $t(786) = -6.096$; $p < .001$; $d = .435$) und dem Migrationshintergrund (MMig: 3.04, MDt: 3.32, $t(692) = 4.856$; $p < .001$; $d = .370$). Vergleichbare Ergebnisse zeigen sich für die negative Interaktion. So korreliert Schimpfen signifikant mit der Leistung ($r = -.253$, $p < .001$, $N = 784$) und den emotional-sozialen Fähigkeiten ($r = -.474$, $p < .001$, $N = 785$) und es liegen signifikante Mittelwertunterschiede in Abhängigkeit vom Geschlecht (MJunge: 2.58, MMädchen: 2.05, $t(786) = 8.777$; $p < .001$; $d = .626$) und dem Migrationshintergrund (MMig: 2.59, MDt: 2.11, $t(692) = -7.407$; $p < .001$; $d = .564$) vor. Korrelationen der Interaktionsqualität mit der inklusiven Grundhaltung auf Klassenebene sind nicht signifikant.

Im Vortrag werden Ergebnisse der Zwei-Ebenen-Modelle vorgestellt und Implikationen für die schulische Praxis und Anschlussforschung diskutiert.

Wie Schüler:innen die Bedürfnisunterstützung im Unterricht wahrnehmen – Eine latente Profilanalyse zur Rolle personen-, herkunfts- und leistungsbezogener Merkmale

Cornelia Gar¹, Katja Scharenberg¹, Katrin Lohrmann¹, Stefan Markus², Michaela Gläser-Zikuda³

¹Ludwig-Maximilians-Universität München; ²Bergische Universität Wuppertal; ³Friedrich-Alexander-Universität Erlangen-Nürnberg

Theoretischer Hintergrund und Fragestellungen

Motivationalen Aspekten kommt eine zentrale Rolle bei der Erklärung und Vorhersage von Schulleistungen zu (den Brok et al., 2010; Hornstra et al., 2015). Der Selbstbestimmungstheorie (Ryan & Deci, 2017) zufolge hängen Motivation und Leistung von Schüler:innen mit ihrem Erleben von Autonomie, Kompetenz und sozialer Eingebundenheit im Unterricht zusammen. Die Unterrichtsgestaltung durch die Lehrkraft ist daher relevant (Stroet et al., 2013). Enge Zusammenhänge zwischen Bildungserfolg, schulischen Leistungen sowie der sozialen Herkunft der Schüler:innen (Stanat et al., 2022) wurden in Studien ebenso aufgezeigt wie Geschlechtsunterschiede (Spinath et al., 2014). Des Weiteren hängen Schüler:innenleistungen mit dem Verhalten der Lehrkraft im Unterricht zusammen (Good et al., 2018; Domen et al., 2020; Reeve, 2009). Bislang existieren jedoch keine Studien, die eine ganzheitliche und gleichzeitig differenzierte Betrachtung der Indikatoren eines bedürfnisunterstützenden Lehrkraftverhaltens mit personen-, herkunfts- sowie leistungsbezogenen Merkmalen von Schüler:innen vornehmen. Daher wurde ein personenzentrierter Ansatz gewählt, um diese Forschungslücke zu adressieren und folgende Forschungsfragen zu beantworten:

- (1) Welche Schüler:innen-Profile können hinsichtlich der schüler:innenperzipierten Bedürfnisunterstützung identifiziert werden?
- (2) Inwiefern unterscheiden sich Schüler:innen-Profile hinsichtlich personen-, herkunfts- sowie leistungsbezogener Merkmale von Schüler:innen?

Methodik

Zur Beantwortung der Fragestellungen werden Querschnittsdaten des Forschungsprojekts BASiS (Beziehungsgestaltung, Autonomie und Soziale Eingebundenheit in der Schule) genutzt. Die Stichprobe setzt sich aus $n = 1781$ Schüler:innen (Jahrgangsstufe 6–11) aus 153 Schulklassen an 15 Schulen zusammen. Die durch die Lehrkraft gebotene Unterstützung von Autonomie ($\omega = .74-.83$), Kompetenz ($\omega = .73-.88$) und sozialer Eingebundenheit ($\omega = .81-.92$) in Mathematik wurde aus Perspektive der Schüler:innen erhoben (Admiraal et al., 2019; Thiel et al., 2013; Torsheim et al., 2012). Die Mathematikkompetenz wurde über die Zeugnisnote und einen standardisierten Kompetenztest (IQB, 2013; ISB, 2019) erfasst. Der sozioökonomische Status wurde über die angegebenen Wohn- und Eigentumsverhältnisse sowie Berufsklassen der Eltern erhoben, der Migrationshintergrund der Schüler:innen über die zu Hause gesprochene Sprache (Wendt et al., 2016). Mittels latenter Profilanalysen wurden Schüler:innenprofile hinsichtlich der Bedürfnisunterstützung identifiziert und unter Verwendung des BCH-Ansatzes in Mplus Version 8.10 analysiert (Muthén & Muthén, 2017). Hierfür wurden die Daten unter Berücksichtigung ihrer genesteten Struktur sowie getrennt nach Schulstufen schulformspezifisch analysiert.

Ergebnisse und Diskussion

Unter Berücksichtigung statistischer Fit-Indizes, der Wahrscheinlichkeit der Profilzuordnungen und der Profilgröße erwiesen sich 3- sowie 4-Profil-Modelle als optimal, wobei sich schulstufen- sowie schulformspezifische Unterschiede hinsichtlich der schüler:innenperzipierten Bedürfnisunterstützung zeigen (FF1). Hinsichtlich FF2 zeigt der BCH-Differenztest zwischen den Profilen der Mittel- und Oberstufe des Gymnasiums signifikante Haupteffekte ($\chi^2_{\text{Mittelstufe}}(7) = 26.76$, $p < .001$, $\chi^2_{\text{Oberstufe}}(7) = 11.14$, $p = .011$) für das Geschlecht. Paarweise Differenztests weisen auf signifikante Mittelwertsunterschiede zwischen Profilen mit niedriger und moderater bis hoher Bedürfnisunterstützung, überwiegend zuungunsten der Mädchen, hin. Für die Mathematikkompetenz zeigen sich in der Mittelstufe signifikante paarweise Differenzen zugunsten des Profils, das sich durch eine höhere Bedürfnisunterstützung auszeichnet ($\chi^2_{1/2}(7) = 4.69$, $p = .030$; $\chi^2_{2/4}(7) = 5.29$, $p = .021$). Für die nicht-gymnasialen Schulformen zeigen sich zwischen den Profilen der Mittel- und Oberstufe signifikante Haupteffekte ($\chi^2_{\text{Mittelstufe}}(7) = 6.56$, $p = .038$; $\chi^2_{\text{Oberstufe}}(7) = 8.88$, $p = .012$) hinsichtlich des Migrationshintergrunds. Zwischen den Profilen mit niedriger, moderater und hoher Bedürfnisunterstützung zeigen sich ebenfalls signifikante paarweise Differenzen, und zwar sowohl zugunsten als auch zuungunsten der Schüler:innen-Profile mit höherem Migrationsanteil. Für die Indikatoren der Mathematikkompetenz bestehen in allen drei Schulstufen signifikante paarweise Differenzen zugunsten der Schüler:innen-Profile mit der höchsten Bedürfnisunterstützung sowie ein signifikanter Haupteffekt für die Oberstufe ($\chi^2(7) = 9.42$, $p = .009$).

Die Ergebnisse zeigen, dass vorrangig am Gymnasium die wahrgenommene Bedürfnisunterstützung im Mathematikunterricht hinsichtlich des Geschlechts variiert und herkunfts- sowie leistungsbezogene Merkmale dagegen insbesondere in den nicht-gymnasialen Schulformen für ein unterschiedliches Ausmaß an Bedürfnisunterstützung im Mathematikunterricht von Bedeutung zu sein scheinen. Größtenteils reihen sich diese sowie weitere Ergebnisse in den bisherigen Forschungsstand ein (z.B.

Vasconcellos et al., 2020). Daraus abgeleitete Implikationen sollen für das schüler:innenperzipierte bedürfnisunterstützende Verhalten der Lehrkraft und dessen Variation hinsichtlich zentraler Schüler:innenmerkmale sensibilisieren und zudem Hinweise auf mögliche Effekte von Schulformen und -stufen geben.

PS-76: School Development in Transition: Current Challenges and Changes

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 382

Chair der Sitzung: Marko Neumann

English Track

Do Research-Practice Partnerships address 21st Century Teaching Challenges? – Results of a Systematic Literature Review

Désirée Theis, Raphaela Schlicht-Schmälzle, Jonas Visé, Ulrike Hartmann

DIPF, Deutschland

Introduction

Societies consistently strive for progress and innovation. As a result, they face rapid and far-reaching processes such as globalization, increasing diversity, and technological and digital transformations. Navigating this ever-changing world requires individuals to be equipped with essential skills, including critical thinking, problem-solving, and creativity (OECD, 2023, 2018, 2019). In this context, schools and teachers play a crucial role. They must respond to global changes and transformations by adapting curricula and teaching practices accordingly (OECD, 2023, 2018). Collaborating in Research-Practice Partnerships (RPPs) could assist teachers and schools in using research findings to deal with these changes. RPPs are long-term, mutualistic efforts characterized by shared responsibility among participants that aim to address practical problems and to produce usable and accessible research findings for educators (Coburn and Penuel, 2016). RPPs have gained increasing attention over the past decade (Cooper et al., 2020). Nonetheless, knowledge about these partnerships remains incomplete (Coburn & Penuel, 2016). Expanding knowledge about the issues RPPs focus on and their relationship with structural characteristics of the partnerships — such as geographical location, collaboration duration, and the stakeholders involved — could help to identify specific fields of interest in RPP work and determine whether the issues addressed depend on the context of the collaborations. Therefore, we examine whether RPPs address 21st century challenges, structural characteristics of RPPs (geographic context, involved stakeholders, duration of partnerships), and the relationship between the issues discussed by RPPs and their structural characteristics.

Method

The present study builds on and utilizes data from a systematic literature review conducted by Authors A. For the literature review, searches in the EBSCO Education Research Complete and the Social Science Citation Index (SSCI) by Web of Science databases were performed, searching for articles that used the term "Research-Practice Partnership*" in titles, abstracts, or keywords. After a series of exclusion and inclusion processes the search resulted in a sample of 141 empirical RPP studies published in English from 2002 to the end of 2022. We employed content analysis in MAXQDA to examine whether the thematic foci, defined according to the OECD framework of 21st century teaching challenges (effective use of technology, teaching strategies for 21st century skills, migration and mobility, well-being, equity and inclusion; OECD, 2018, 2023), were mentioned in the articles. Moreover, we looked for information on the geographic location of the RPPs, the duration of the collaboration, and the stakeholders involved in the collaboration. Finally, we analyzed whether the RPPs that focused on 21st century challenges differed in terms of the structural characteristics we assessed.

Results and discussion

This study represents an initial step toward enhancing our understanding of RPPs and their role in addressing 21st century challenges in education. Moreover, our findings suggest preliminary hypotheses guiding future research. Our findings indicate that RPPs address the challenges faced by educators in their daily practice. Over half of the RPPs examined in the literature focus on 21st century challenges. Key among these challenges are the effective teaching of 21st century skills and promoting equity and inclusion. Teaching 21st century skills emerges as a critical focus of RPPs across multiple continents, including Europe, North America, South America, and Australia but less often in Africa and Asia. However, our analysis also highlights significant gaps in the reporting of structural details such as partnership durations within the RPP literature. While Coburn and Penuel (2016) have established criteria for RPPs, our review indicates limited adherence to these criteria in the literature. Furthermore, we find geographical differences in the prevalence of RPPs. The vast majority are concentrated in North America, while countries on other continents host far fewer partnerships.

Bewertung von Reformen in der gymnasialen Oberstufe - eine Frage schulkontextueller Merkmale?

Marko Neumann

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Theoretischer Hintergrund und Fragestellungen:

Das deutsche Schulsystem durchläuft fortlaufende Reformprozesse, die sich auf alle Schulstufen von der Grundschule bis zur gymnasialen Oberstufe erstrecken. Als eine wichtige Gelingensbedingung für die erfolgreiche Implementation und Wirkung von Reformmaßnahmen wird dabei die Akzeptanz der Neuerungen bei den beteiligten Akteuren gesehen (vgl. Gräsel & Parchmann, 2004; Goldenbaum 2013; Holtappels, 2014). Insbesondere Schulleitungen kommt bei der Implementation von schulischen Reformmaßnahmen eine Schlüsselrolle zu, da sie als treibende Kraft bzw. Unterstützer („agent of support“, Hallinger & Heck, 2011, S. 4) für die reformbezogenen schulischen Veränderungsprozesse vor Ort fungieren. Welche Rolle Faktoren des schulischen Kontextes wie Schulgröße, Schulform, Zusammensetzung der Schülerschaft, personelle und räumliche Ressourcen der Reformbewertung von Schulleitungen zukommt, wurde bislang nur vereinzelt untersucht (Lee, 2020). An diesem Desiderat setzt die vorliegende Untersuchung an.

Den Kontext bildet eine im Land Schleswig-Holstein umgesetzte Reform der gymnasialen Oberstufe, in deren Zentrum die Einführung getrennter Anspruchsniveaus (grundlegendes und erweitertes Anspruchsniveau) in den drei Kernfächern Deutsch, Mathematik und fortgeführte Fremdsprache, die im bisherigen System von allen Schülerinnen und Schülern im gemeinsamen Klassenverband auf erhöhtem Anspruchsniveau besucht wurden, steht. Zudem wurde mit der Einführung des so genannten „Profilseminars“ Raum für das fächerübergreifende und projektorientierte Lernen geschaffen (vgl. MBWH, 2021). Beide Reformbestandteile bringen bedeutende organisatorische und inhaltlich-konzeptionelle Anpassungsbedarfe mit sich, etwa mit Blick auf die Kursplanung (sowohl hinsichtlich Stundentafel als auch personeller Planung und Bedarfe) oder die inhaltliche Ausgestaltung der beiden Anspruchsniveaus und des Profilseminars.

Vor diesem Hintergrund sollen im eingereichten Beitrag folgende Forschungsfragen adressiert werden:

1. Wie bewerten die Schulleitungen die Neuerungen in der Oberstufe und ihren Umsetzungsstand und inwieweit variieren die diesbezüglichen Einschätzungen zwischen den Schulen?
2. In welchem Maße weisen die reformbezogenen Bewertungseinschätzungen der Schulleitungen Zusammenhänge mit Merkmalen des schulischen Kontextes auf?

Methodisches Vorgehen:

Die Untersuchung erfolgt im Rahmen einer wissenschaftlichen Begleitstudie zur neuen Oberstufe in Schleswig-Holstein, deren primäre Datengrundlage Online-Befragungen von Schulleitungen (Vollerhebung, Teilnahmequote 76 %) und Lehrkräften (an rund einem Drittel der Schulen) darstellen. Die Befragung wurde im Jahr 2023 rund ein halbes Jahr nach Eintritt des ersten Schüler*innenjahrgangs in das neu gestaltete Kurssystem durchgeführt. Für die vorliegende Untersuchung werden die Angaben von insgesamt $N = 108$ Schulleitungen aus allgemeinbildenden Gymnasien und Gemeinschaftsschulen mit gymnasialer Oberstufe herangezogen, die zu verschiedenen Bewertungs- und Umsetzungsaspekten bezüglich der Neuerungen sowie ausgewählten Merkmalen des schulischen Kontextes befragt wurden. Dabei kamen überwiegend Befragungsinstrumente aus anderen Reformstudien zum Einsatz, die soweit erforderlich an den spezifischen Reformkontext angepasst und um eigenentwickelte Fragen ergänzt wurden. Als Merkmale des schulischen Kontextes wurden die Schüler*innenzahl in der Oberstufe, die soziale und leistungsbezogene Schüler*innenkomposition, die Einwohnerzahl am Schulstandort sowie die Einschätzung der räumlichen und personellen Ausstattung (einschließlich Unterrichtsversorgung) berücksichtigt. Die Auswertungen erfolgten sowohl auf Skalenebene als auch auf Ebene von Einzelitems. Für Fragestellung 1 werden Mittelwerte und Streuungen (Häufigkeitsverteilungen und Standardabweichung) berechnet. Für die Zusammenhangsanalysen in Fragestellung 2 werden Partialkorrelationen unter Kontrolle der Schulform ermittelt und um ausgewählte multivariate lineare Regressionsanalysen zur Vorhersage der Reformbewertung durch schulkontextuelle Merkmale ergänzt.

Ergebnisse:

Die Bewertung der beiden Reformbestandteile durch die Schulleitungen fiel im Mittel positiv aus (vierstufige Antwortskala von 1-4; Skala Bewertung Einführung Niveaudifferenzierung: $M = 3,35$; $SD = 0,52$; Skala Bewertung Einführung Profilseminar: $M = 2,78$; $SD = 0,66$). Auch bezüglich der Umsetzung an den Schulen fanden sich im Mittel positive Einschätzungen (Skala Umsetzung Neuerungen insgesamt: $M = 3,19$; $SD = 0,56$; Skala Umsetzung Profilseminar: $M = 2,81$; $SD = 0,58$). Die bivariaten Zusammenhangsanalysen erbrachten vereinzelte statistisch signifikante Zusammenhänge der Bewertungseinschätzungen mit schulkontextuellen Merkmalen, die in ihrer Größenordnung jedoch moderat ausfielen ($r_{\max} = -.35$). Gleichwohl liefern die bivariaten Analysen erste Hinweise auf die Relevanz schulkontextueller Merkmale für die Bewertung von Reformmaßnahmen, die im Beitrag um ausgewählte multivariate Analysen ergänzt werden sollen.

Schulautonomie in Deutschland – wie Schulleitungen sie wahrnehmen und wie sie sich seit 2000 entwickelt hat

Dominik Endemann¹, Jannis Burkhard²

¹Freie Universität Berlin; ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Der Paradigmenwechsel von der Input- zur Outputsteuerung, bzw. hin zu einer evidenzbasierten Bildungspolitik begann in Deutschland ab Mitte der 1990er Jahre, angestoßen durch das mittelmäßige Abschneiden deutscher Schüler*innen in der ersten TIMSS-Studie 1995 (Kuhn, 2014; Wendt et al., 2016). Im Zuge des Paradigmenwechsels kam es zwischen 1999 und 2004 auch zu einem deutlichen Anstieg an implementierten rechtlichen Regelungen, die die Eigenständigkeit von Einzelschulen erweiterten (Rürup & Heinrich, 2007). Trotzdem erfuhr das Thema Schulautonomie seit Ende der 2000er Jahre in der Forschung eine vergleichsweise geringe Beachtung. Schulautonomie wird als eine grundlegende Voraussetzung für effektive pädagogische Arbeit an Einzelschulen angesehen, da sie zu einer effizienteren Ressourcenzuweisung und einem höheren Engagement der Mitarbeitenden führen soll (Nikolai & Helbig, 2013; Dou et al., 2017). Inwieweit und in welchen Bereichen Einzelschulen in Deutschland auch tatsächlich selbstständig handeln können, gilt jedoch weiterhin als Forschungsdesiderat. Die Arbeit von Rürup und Heinrich (2007) und ein „Expertenrating“ des Aktionsrats Bildung (vbw – Vereinigung der bayerischen Wirtschaft e.V., 2010) untersuchen die rechtliche Verbreitung der „Idee“ Schulautonomie, daneben gibt es Erkenntnisse aus der Begleitforschung zu Modellprojekten (Pfeiffer, 2008). Mithilfe von Daten aus Large-Scale Assessments wurde außerdem untersucht, wie Schüler*innenleistung und Schulautonomie zusammenhängen (Wößmann et al., 2007; Nikolai & Helbig, 2013). Die Wahrnehmung der Handlungsspielräume von Schulleitungen in verschiedenen Bereichen über mehrere Jahre hinweg wurde bislang nicht untersucht. Der vorliegende Beitrag widmet sich daher der Fragestellung, wie Schulleiter*innen die Autonomie ihrer eigenen Schule bewerten. Ziel ist es, die Entwicklung dieser Einschätzungen im Zeitverlauf von 2000 bis 2022 nachzuzeichnen.

Zur Untersuchung der Fragestellungen wurden Daten der PISA-Studien von 2000 bis 2022 herangezogen. Die Schulleitungen beantworteten in den Kontextfragebögen zwölf Fragen dazu, wer die Hauptverantwortung für die Bereiche Personal, Finanzen, Schülerangelegenheiten und Unterrichtsgestaltung hat (Maslowski et al., 2007). Je nach Jahr liegt die Stichprobengröße der Befragung zwischen 216 und 257 Schulleiter*innen. Um die Angaben der verschiedenen Jahre miteinander vergleichen zu können, wurden die Antwortoptionen von 2000 – 2015 im ersten Schritt harmonisiert. Dafür wurde eine Variable erstellt, die abbildet, ob die Schulleiter*innen die Verantwortung für den jeweiligen Autonomiebereich ausschließlich bei der Schule sehen. Anschließend wurde die Entwicklung der verschiedenen Schulautonomiebereiche von 2000 bis 2015 deskriptiv untersucht. Aufgrund unterschiedlicher Antwortformate konnten die Daten der PISA-Studie 2022 nicht mit in die längsschnittliche Untersuchung einfließen. Deshalb wurde im zweiten Schritt die Situation im Jahr 2022 anhand der aktuellsten verfügbaren Daten analysiert und es wurde mittels Clusteranalyse explorativ der Frage nachgegangen, ob sich Schulen bestimmten Typen von Schulautonomie zuordnen lassen.

Ein zentrales Ergebnis dieser Analysen ist, dass die Schulautonomie in der Mehrzahl der untersuchten Bereiche (Personell, Finanziell, Schülerangelegenheiten, Curricular) in der Wahrnehmung der Schulleiter*innen zwischen 2000 und 2015 tendenziell abgenommen hat. Für 2022 kann zusammenfassend festgestellt werden, dass viele pädagogische Entscheidungen zu einem großen Teil auf Schulebene getroffen werden. Dagegen gaben Schulleitungen an, dass bei administrativen Entscheidungen, also z.B. zur Festlegung des Schulbudgets, meist übergeordnete Instanzen entscheiden. Die anschließende Clusteranalyse ergab zwei Cluster: „zentralisierte Steuerung“ (161 Schulen) und „dezentralisierte Steuerung“ (39 Schulen). Besonders auffällig sind die Unterschiede zwischen den beiden Clustern im Bereich Personal. Im Cluster „zentralisierte Steuerung“ ist hier der Anteil an Schulleiter*innen, die die Ministerien als hauptverantwortlich betrachten, je nach Item bis zu 88 Prozentpunkte höher.

Dass sich die wahrgenommene Schulautonomie nur wenig verändert hat stützt die These Rürups (2019), der resümiert, dass die Schulautonomie derzeit als ein Instrument zur Umsetzung schulpolitischer Anliegen dient, da top-down-Implementationen als ineffektiv angesehen werden.

Eine Limitation der Studie sind Hinweise darauf, dass die (wahrgenommenen) Handlungsspielräume von Einzelschulen je nach Bundesland und Schulgröße deutlich variieren (Warwas, 2012; vbw – Vereinigung der bayerischen Wirtschaft e.V., 2010).

Input Stratification? School Resources and Track Type. Empirical evidence from early tracking countries using PISA data.

Maximilian Brinkmann, Janna Teltemann

Universität Hildesheim, Deutschland

The early external differentiation (i.e. tracking) is a polarizing issue. Amidst numerous empirical studies on this topic, two central findings are beginning to emerge: there is little evidence supporting the higher efficiency of these education systems, but substantial evidence indicating that such systems lead to greater social inequality (Terrin & Triventi, 2022, meta-analysis). Despite these findings, the topic remains highly polarizing (Esser & Seuring, 2020; Heisig & Matthewes, 2022). In this context, we find it crucial to describe mechanisms within tracked systems that can explain the empirical evidence. Our contribution focuses on one such mechanism: the stratification of school resources.

In their influential study, Brunello & Checchi (2007, p. 795) describe, among other things, a possible mechanism for the higher inequality in tracked systems: unequally distributed resources between school tracks. They provide (descriptive) figures from 2004 for five countries suggesting that higher school tracks exhibit a lower (i.e., better) student-teacher ratio. This seemingly led to resource stratification being described either as a conceivable or even as an established mechanism in the literature (Betts, 2011, p. 33; Bol, 2014, p. 1548; Brinkmann et al., 2024, p. 181; Pomianowicz, 2023, pp. 299, 304; Terrin & Triventi, 2022, p. 7; Werfhorst, 2018, p. 24).

The data provided by Brunello & Checchi (2007; 795) thus appear to be the best evidence available. In this context, our contribution seeks to provide more comprehensive evidence using PISA data pursuing two main goals. First to replicate and extend the figures (i.e. student-teacher ratios) presented by Brunello & Checchi (2007; 795). Second, to gauge the distribution of other school resource indicators covered in PISA. We focus on tangible indicators of school resources, such as student-teacher ratios or student-computer ratios, but also consider qualitative indicators like student behavior that disrupts instruction.

We use PISA data from 2000 to 2022. Utilizing multiple waves allows us to confirm the robustness of findings for various indicators (e.g., student-teacher ratios) and to include indicators present in only a few waves. Our analysis proceeds in two steps. First, we examine whether the distribution of school resources systematically differs between tracked and comprehensive systems by computing measures of unevenness, such as the Gini coefficient, for school indicators within a country. Then, we compare these indicators between tracked and untracked systems using a simple regression framework. Next, we analyze resource distribution within tracked systems. That is, we predict resource endowment based on track type using regression.

In order to distinguish tracked and untracked/comprehensive systems, we draw on previous research (Brinkmann et al., 2024; Strello et al., 2022). To indicate track type within tracked systems we use two approaches. First, we use PISA's "program" variable to code track type. Second, we follow Wößmann (2009, p. 26) who characterizes tracked systems as containing "school types, hierarchically structured by performance" and use the school's average achievement.

Our preliminary findings show that while track systems tend to exhibit a more uneven distribution of resources compared to untracked systems, the differences are rather small. For our analysis within tracked systems we find robust evidence contradicting the figures presented by Brunello & Checchi (2007; 795). That is, higher tracks exhibit higher student to teacher ratios. In general we find that, if anything, lower tracks seem to exhibit more favorable resource endowments when it comes to "tangible" indicators. Only for qualitative indicators such as student behavior, we find the opposite pattern.

Our next step is to use multiverse-analysis in order to make our findings and the numerous data-analytic steps involved in this process more transparent. We aim to discuss the implications of our findings for the tracking literature and elaborate on mechanisms (word limit).

PS-77: Perspektiven auf Heterogenität im Unterricht

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 129

Chair der Sitzung: Andrea Vorderobermeier

Einfluss lehr-, lerntheoretischer sowie motivationaler Überzeugungen auf die Wahrnehmung von Unterrichtsmerkmalen in heterogenen Klassengefügen

Andrea Vorderobermeier, Maximilian Sailer

Universität Passau, Deutschland

Theoretischer Hintergrund & Fragestellungen

Vielfach wurde gezeigt, dass lehr-, lerntheoretische und motivationale Überzeugungen von Lehrpersonen u.a. die Wahrnehmung der Unterrichtsmerkmale sowie Lernleistungen der Schüler:innen beeinflussen (z.B. Helmke, 2015; Löwen et al., 2011; Zee & Koomen, 2016). Ziel dieses Beitrags ist es zu untersuchen, ob diese Erkenntnisse auch für heterogene Klassenstrukturen gelten. Die zweiwöchige Sommerschule in Österreich dient hierbei als Unterrichtsetting (Stichprobe umfasst Lehrende $N = 532$ und Schüler:innen $N = 3.517$), in dem Schüler:innen in Kleingruppen (max. 15 Schüler:innen) klassen- und schulstufenübergreifend unterrichtet werden. Da Überzeugungen durch spezifische Kontextbedingungen aktiviert werden können (Fives & Buehl, 2012), soll geklärt werden, welche Überzeugungen (*Konstruktion, Transmission, Schüler:innenorientierung, Selbstwirksamkeit*) bei den Lehrpersonen festzustellen sind. Zudem wird untersucht, ob diese Überzeugungen im gegebenen Setting die Wahrnehmung der Unterrichtsmerkmale *Instruktionsstrategien, Klassenmanagement, Schüler:innenmotivierung, Klarheit und Strukturiertheit, anspruchsvolles Üben, Individualisierung und Differenzierung* beeinflussen. Es ergeben sich folgende Forschungsfragen:

FF1: Welche Struktur weisen die lehr-, lerntheoretischer Überzeugungen in den Dimensionen *Transmission, Konstruktion* und *Schüler:innenorientierung* sowie motivationaler Überzeugungen (*Selbstwirksamkeit*) bei den Lehrpersonen der Sommerschule auf?

FF2: Haben Überzeugungen der Lehrpersonen einen Einfluss auf die Wahrnehmung von Unterrichtsmerkmalen durch Schüler:innen in heterogenen Gruppengefügen?

Methodik

Daten von 532 Lehrpersonen (im Dienst befindliche Lehrkräfte $N = 377$ und Lehramtsstudierende $N = 155$) und deren 3.517 Schüler:innen ($M_{\text{Alter}} = 12,75$; $SD = 1,58$; $M_{\text{Klassengröße}} = 10,39$) wurden über einen Online-Fragebogen im Rahmen einer Evaluierung der Sommerschulmaßnahme erhoben. Aufgrund der Datenstruktur wurde unter Berücksichtigung theoretischer Überlegungen, Intraklassenkorrelationen (Individual- und Klassenebene) sowie Mehrebenen-Konfirmatorische Faktorenanalysen geprüft, ob eine Mehrebenenanalyse durchzuführen ist. Es zeigte sich, dass darauf verzichtet werden kann (schwache Hierarchisierung und geringer Unterschiede zwischen den Gruppen). Zur Beantwortung der FF1 wurden t-Tests und ANOVAs durchgeführt. Da die Stichprobe der Lehrpersonen aus zwei Personengruppen besteht, wurde sichergestellt, dass Messinvarianz gegeben ist. Alle latenten Konstrukte weisen mindestens metrische Messinvarianz auf. Um die interessierenden latenten Konstrukte sowohl auf Schüler:innen- als auch auf Lehrkräfte-Ebene und deren Zusammenhang zu repräsentieren (FF2), werden, getrennt für Studierende und im Dienst befindliche Lehrkräfte, Strukturgleichungsmodelle herangezogen.

Ergebnisse & Schlussfolgerungen

Alle verwendeten Skalen, die Überzeugungen abbilden, weisen in der Stichprobe hohe Zustimmungswerte ($M = 3.92$ bis 4.46 ; max. 5) auf. Bei Lehramtsstudierende sind leicht höhere Werte in der *Konstruktion* sowie *Schüler:innenorientierung* festzustellen. Die Berufserfahrung der Lehrkräfte hat auf die Zustimmung zur *Transmission* sowie zur *Selbstwirksamkeit* einen kleinen signifikanten Einfluss. Es unterscheiden sich vor allem Lehrpersonen mit geringer Berufserfahrung von solchen, die bereits länger als 10 Jahre im Dienst sind (signifikant höhere Werte).

Das Strukturgleichungsmodell der Lehrkräfte zeigt, dass sich die *Selbstwirksamkeit* auf die Unterrichtsmerkmale *Klassenmanagement* ($\beta = .12$), *Schüler:innenmotivierung* ($\beta = .12$), *Individualisierung* ($\beta = .15$) und *Differenzierung* ($\beta = .22$) jeweils schwach positiv auswirkt. *Konstruktion* hat einen Einfluss auf *Instruktionsstrategien* ($\beta = .17$), *Klassenmanagement* ($\beta = .14$) *Schüler:innenmotivierung* ($\beta = .14$) und *Individualisierung* ($\beta = .16$). Transmissive Überzeugungen haben einen kleinen negativen Effekt auf *Individualisierung* ($\beta = -.14$) und *Differenzierung* ($\beta = -.28$). Ebenfalls negativ wirkt sich die *Schüler:innenorientierung* auf die Merkmale *Instruktionsstrategie* ($\beta = -.16$), *Klassenmanagement* ($\beta = -.17$), *Schüler:innenmotivierung* ($\beta = -.14$), *Klarheit und Strukturiertheit* ($\beta = -.12$) und *Individualisierung* ($\beta = -.16$) aus.

Bei den Studierenden werden lediglich transmissive Überzeugungen signifikant, mit schwach positiven Auswirkungen auf die Merkmale *Anspruchsvolles Üben* ($\beta = .24$), *Individualisierung* ($\beta = .20$) und *Differenzierung* ($\beta = .25$).

Es konnte gezeigt werden, dass selbst in einem heterogenen Umfeld, wie es die Sommerschule darstellt, Überzeugungen von Lehrpersonen einen geringen, aber signifikanten Einfluss auf die Wahrnehmung einiger Unterrichtsmerkmale haben. Sie spielen damit eine Rolle in der Interaktion zwischen Lehrenden und Lernenden. Erklärungsgründe für die Unterschiede zwischen Lehrkräften und Lehramtsstudierenden können die fehlende Unterrichtserfahrung sowie individuelle und kontextuelle Gegebenheiten und damit die Nicht-Übereinstimmung von Überzeugung und tatsächlicher Unterrichtspraxis bei Studierenden sein.

Heterogene Schüler:innen - heterogene Unterstützungsbedarfe? Einblicke in die professionelle Wahrnehmung von Schulleitungen

Susanne Enssen

Universität Duisburg-Essen, Deutschland

Die Verbesserung der Bildungschancen benachteiligter Schüler:innen, die Schulen in sozialräumlich deprivierten Lagen besuchen, rückt zunehmend in den Fokus von Wissenschaft und Politik, wie verschiedene Projekte von Bund und Ländern zeigen (BMBF, 2021; 2024). In diesem Kontext erscheint es sinnvoll, die Komposition benachteiligter Schüler:innen in den Blick zu nehmen, um eine bedarfsorientierte Unterstützung zu ermöglichen. Jurczok & Lauterbach (2014), die sich mit der „Geografie von Bildungschancen in benachteiligten städtischen Bildungsräumen“ (ebd.) auseinandergesetzt haben, kommen zu dem Ergebnis, dass eine Homogenisierung von Schüler:innen an diesen Schulen stattfindet.

Dieser Befund stützen weitere Studien, nach denen Schulen in sozialräumlich deprivierten Lagen zumeist ein Abbild städtischer Quartiere bzw. Sozialräume darstellen, die von „sozioökonomischen Polarisierungs- und Entmischungsprozessen“ (Bremm et al., 2016, S. 325) geprägt sind. Andere Studien hingegen zeigen, dass auch an Schulen in sozialräumlich deprivierten Lagen

durchaus eine heterogene Schüler:innenschaft vorzufinden ist und Schulleitungen bzw. Lehrkräfte einen differenzierten Blick auf die Vielfalt ihrer Schüler:innen und die damit einhergehenden Herausforderungen im Schulalltag haben (Enssen et al., 2024, i.V.; Böttcher et al., 2017). Folglich stellen sich die Fragen, wie sich die Heterogenität von Schüler:innen an Schulen in sozialräumlich deprivierten Lagen äußert und welche Auswirkungen dies auf den Schul- und Unterrichtsalltag hat. Um sich diesen Fragen anzunähern, beschäftigt sich der Beitrag mit folgenden Fragestellungen:

- 1) Wie nehmen Schulleitungen die Heterogenität ihrer Schüler:innen wahr?
- 2) Welche Unterstützungsbedarfe gehen mit der Heterogenität der Schüler:innen einher?

Die Datenbasis stammt aus der Bund-Länder-Initiative *Schule macht stark*. Die Grundlage für den Beitrag bilden 76 geführte Interviews mit Schulleitungen bzw. Schulleitungsteams von Schulen mit Sekundarstufe I, Sekundarstufe I & II sowie Förderschulen. Die Auswertung erfolgte in Anlehnung an die dokumentarische Methode (Nohl, 2017; Kleemann et al., 2013).

Erste Ergebnisse zeigen, dass die Schulleitungen einen differenzierten Blick auf die Heterogenität ihrer Schüler:innen haben. Dies gilt sowohl für Schulen, deren Schüler:innen ausgeprägte Möglichkeiten zur sozial-räumlichen Mobilität haben, als auch für Schulen, deren Schüler:innen darin eingeschränkt sind. Darüber hinaus können die Schüler:innen auf Basis der Erkenntnisse in unterschiedliche Typen unterteilt werden. Auf Grundlage der Befunde zur Heterogenität der Schüler:innen konnten verschiedene Unterstützungsbedarfe bei diesen abgeleitet werden. Die Ergebnisse zeigen, dass die Unterstützungsbedarfe der Schüler:innen den Schul- und Unterrichtsalltag nachhaltig prägen. Einige dieser Bedarfe, wie die Stärkung der Kompetenzen in Deutsch und Mathematik, liegen im Kernbereich der Schule und werden im Rahmen des Unterrichts sowie durch ergänzende außerunterrichtliche Angebote adressiert. Andere Bedarfe gehen deutlich über den unterrichtlichen Kontext hinaus und lassen einen hohen Bedarf an bedarfsorientierter emotional-sozialer Unterstützung erkennen, die insbesondere auf die Persönlichkeitsentwicklung der Schüler:innen abzielt.

Die Ergebnisse werfen weitere Fragen auf, die im Anschluss an den Beitrag diskutiert werden können: Inwieweit werden angehende Lehrkräfte im Studium auf ihre Tätigkeit an Schulen in sozialräumlich deprivierten Lagen vorbereitet? Inwieweit sollten Schul- und Unterrichtsalltag flexibilisiert werden, um eine bedarfsorientierte Unterstützung der Schüler:innen zu ermöglichen? Welche Potenziale und Grenzen gehen mit einer bedarfsorientierten Unterstützung von Schüler:innen an Schulen in sozialräumlich deprivierten Lagen einher?

Auswirkungen von Heterogenität in Schulklassen auf psychosoziale Merkmale von Schülerinnen und Schülern: Ein Systematic Review

Claudia Neuendorf, Svenja Hascher, Camilla Rjosk

Universität Potsdam, Deutschland

Hintergrund

Die Schulklassse als Ökosystem (Classroom ecosystem, Gest & Rodkin, 2011), ist einer der wichtigsten Entwicklungskontexte für Kinder und Jugendliche und hat damit potenziell einen beträchtlichen Einfluss auf ihre kognitive, psycho-soziale und emotionale Entwicklung (Bronfenbrenner, 1994). Neben Eltern und Lehrkräften sind die Mitschüler:innen bestimmende Faktoren im täglichen Erleben von Kindern und Jugendlichen (Horoz et al., 2022; Kilday & Ryan, 2022). Dabei unterscheiden sich Klassengemeinschaften darin, wie heterogen die Schülerschaft bezüglich unterschiedlicher Merkmale ist. Insgesamt zeigt sich in der Gesellschaft ein Trend hin zu größerer Heterogenität: Der Anteil geflüchteter Kinder und Jugendlicher steigt aufgrund humanitärer Krisen in der Welt; mehr Schüler:innen mit sonderpädagogischem Förderbedarf (SPF) werden inklusiv beschult; zunehmend entstehen Schulen mit mehreren Bildungsgängen, an welchen eine heterogenere Schülerschaft beschult wird (Autor:innengruppe Bildungsberichterstattung, 2024). Vermutlich ist ein großer Teil dieser Heterogenität innerhalb der Schulen und Klassen, also im direkten Umfeld der Schüler:innen verortet.

Fragestellung

Der Einfluss der Klassenzusammensetzung auf individuelle Schüler:innenoutcomes wurde bislang vor allem im Hinblick auf die mittlere Ausprägung der Mitschüler:innen untersucht und bezog sich meist auf lern- und leistungsbezogene Outcomes, wie Schulleistungen und motivationale Merkmale (z.B. Dumont et al., 2013; Hattie, 2002; Marsh et al., 2015; Marsh et al., 2007). Forschung zum Einfluss von Heterogenität in Schulklassen auf Schüler:innenoutcomes bezieht sich überwiegend auf leistungsbezogene und ethnische Heterogenität und ihre Effekte auf Schüler:innenleistung (z.B. Dotzel et al., 2021; Rjosk, 2022). Forschung zu Heterogenitätseffekten auf psycho-soziale Merkmale, wie Sozialkompetenzen oder Zugehörigkeitsgefühl, ist hingegen rar. Theorien, auf die sich Studien zur Heterogenität in Schulklassen beziehen, verdeutlichen, dass deren Wurzeln in die Soziologie (Allport, 1954; Blau, 1977), Sozialpsychologie (Turner et al., 1979), Entwicklungspsychologie (Piaget, 1932; Vygotsky, 1978), Kriminologie (Sampson & Groves, 1989; Shaw & McKay, 1942) und die Netzwerkforschung (McPherson et al., 2001) reichen. Dementsprechend handelt es sich um ein fragmentiertes Forschungsfeld, welches von einer Integration profitieren würde. Aus diesem Grund wird im Beitrag ein systematisches Review durchgeführt. Folgende Fragen sollen beantwortet werden: 1. Wie umfangreich ist die Evidenz zum Einfluss von Heterogenität auf psychosoziale Merkmale von Schüler:innen? 2. Welche Heterogenitätsdimensionen und Outcomes wurden untersucht? 3. In welche Richtung deuten die bisherigen Befunde? 4. Welche Forschungslücken bestehen? 5. Welche Forschungsdesiderate ergeben sich?

Methode

Es wurden drei Datenbanken (PsychInfo, ERIC, Scopus) nach englisch- und deutschsprachigen Studien in der Primar- und Sekundarstufe durchsucht, welche in Titel, Abstract oder Schlüsselwörtern sowohl einen Begriff zu Heterogenität (z.B. heterogeneity, mix, diversity), als auch aus dem Bereich psychosozialer Outcomes enthielten (z. B. Sozialverhalten: victimization, violence, social support, cooperation; sozio-emotionale Merkmale: attachment, belonging, classroom climate; Sozialkontakte: friendship, peer relations, segregation, inter-ethnic contact; Sozialkompetenzen: social competences, empathy, soft skills, tolerance). Eine erste Suche lieferte 6.999 Ergebnisse, welche nach mehreren Screeningschritten auf 55 Studien reduziert wurden. Aktuell wird ein Update der Suche durchgeführt.

Ergebnisse und ihre Bedeutung

Zunächst zeigte sich, dass ein Großteil der Studien, welche vorgaben, Heterogenität zu untersuchen, kein Heterogenitätsmaß verwendeten, sondern Klassenmittelwerte verschiedener Merkmale. Diese Studien wurden nicht berücksichtigt. Die Mehrheit der inkludierten Studien untersucht ethnische bzw. sprachliche Heterogenität von Schulklassen, ein kleinerer Teil sozio-ökonomische, nur einzelne Studien Leistungsheterogenität, Geschlechterbalance oder Heterogenität im SPF. Es wurden weiterhin eine Vielzahl an Outcomes - sozio-emotionale Merkmale, soziale Kompetenzen, Sozialverhalten und Sozialbeziehungen – untersucht. Die Ergebnislage ist uneindeutig: Etwa ein Drittel der Studien zu ethnischer Heterogenität zeigt

positive, ein weiteres Drittel negative und der Rest keine Effekte. Die anderen Dimensionen von Heterogenität zeigen negative bzw. Null-Effekte. Lediglich die Studie zu Heterogenität im SPF findet positive Effekte auf soziale Beziehungen. Zukünftige Studien sollten die Unterschiede in den Effekten erklären, indem beispielsweise unterschiedliche Datensätze herangezogen und Moderatoren untersucht werden. Weiterhin sollten mehr unterschiedliche Heterogenitätsdimensionen einbezogen werden. Studien sollten explizitere Bezüge zu den zitierten Theorien herstellen und diese rigoroser testen. Hierfür wären auch experimentelle Designs geeignet.

Umgang mit Heterogenität im adaptiven Unterricht: Bedeutsame Schüler*innenmerkmale aus Sicht von Lehrkräften

Thora Schwarze, Simon Ohl, Hanna Dumont

Universität Potsdam, Deutschland

Theoretischer Hintergrund

Zur optimalen Förderung von Schüler*innen wird zunehmend ein produktiver Umgang mit Heterogenität im Unterricht gefordert (Budde, 2012; Prengel, 2019). Eine Herausforderung dabei ist jedoch, dass es kein geteiltes Verständnis des Heterogenitätsbegriffs gibt und selten konkretisiert wird, welche Heterogenitätsdimensionen berücksichtigt werden sollen (Arnold & Lindner-Müller, 2017; Trautmann & Wischer, 2022; Walgenbach, 2021). Das Konzept des adaptiven Unterrichts stellt einen vielversprechenden Ansatz zur Realisierung eines produktiven Umgangs mit Heterogenität dar und kann herangezogen werden, um Dimensionen des Heterogenitätsbegriffs zu konkretisieren. Im adaptiven Unterricht wird das Lernangebot kontinuierlich an die individuellen Lernvoraussetzungen der Schüler*innen angepasst (Corno, 2008; Dumont, 2019; Hardy et al., 2019; Randi, 2022). So benötigen Schüler*innen mit weniger günstigen Lernvoraussetzungen einen höheren Grad an Unterstützung durch die Lehrkraft, während Schüler*innen mit günstigeren Lernvoraussetzungen von einem niedrigeren Unterstützungsgrad profitieren, da sie bereits selbstreguliert lernen können (Corno, 2008; Kalyuga, 2007). Als besonders zentral für Lehr-Lernprozesse im adaptiven Unterricht werden kognitive und psychosoziale Schüler*innenmerkmale verstanden (Corno & Snow, 1986). Ungeklärt ist jedoch, ob die im Konzept des adaptiven Unterrichts genannten Heterogenitätsmerkmale in der Praxis tatsächlich von Lehrkräften als bedeutsam erachtet werden.

Fragestellung

Vor diesem Hintergrund geht der vorliegende Beitrag der Frage nach, welche Schüler*innenmerkmale Lehrkräfte als bedeutsam für adaptives Unterrichten erachten. Dies erfolgt anhand einer gezielt ausgewählten Stichprobe von Lehrkräften, die an Preisträger-Grundschulen des Deutschen Schulpreises unterrichten, die u.a. aufgrund ihres produktiven Umgangs mit Heterogenität ausgezeichnet wurden.

Methode

Die Datengrundlage stellen leitfadengestützte Interviews mit 32 Klassenlehrkräften aus sieben Grundschulen dar, die mit dem Deutschen Schulpreis ausgezeichnet wurden. Der Interviewleitfaden gliederte sich in fünf Themenkomplexe: Charakteristika des Unterrichts, Umgang mit Heterogenität, Werte und Einstellungen zu Heterogenität, Rolle der Klassengemeinschaft und interne Kooperation im Lehrkollegium. Anhand der qualitativen Inhaltsanalyse nach Kuckartz und Rädiker (2022) wurden bedeutsame Schüler*innenmerkmale aus Sicht der Lehrkräfte identifiziert.

Ergebnisse und ihre Bedeutung

Die Ergebnisse deuten darauf hin, dass die in der Literatur zum adaptiven Unterricht genannten Heterogenitätsmerkmale in einem hohen Maße mit denen übereinstimmen, die von den befragten Lehrkräften als bedeutsam erachtet werden. Insgesamt nannten die Lehrkräfte die Merkmalsdimension der kognitiven Merkmale und hierbei das Merkmal des bereichsspezifischen Vorwissens am häufigsten. Dies deckt sich mit der Annahme, dass adaptiver Unterricht vor allem durch die Anpassung des Unterrichts an die kognitiven Lernvoraussetzungen gekennzeichnet ist (Corno, 2008; Corno & Snow, 1986; Hardy et al., 2019). Daneben waren psychosoziale Schüler*innenmerkmale und in diesem Zusammenhang selbstregulative Kompetenzen für die interviewten Lehrkräfte relevant. Auch für Corno (2008) stellt selbstreguliertes Lernen einen integralen Bestandteil adaptiven Unterrichts dar. Das Merkmal sonderpädagogischer Förderbedarf wurde von einem Großteil der Lehrkräfte als lehr-lernrelevant benannt. In der Literatur zum adaptiven Unterricht werden (Lern-)Behinderungen zwar thematisiert, jedoch in einem geringeren Umfang als kognitive und psychosoziale Merkmale (Corno, 2008). Am seltensten sprachen die Lehrkräfte – trotz ihrer nachgewiesenen Bedeutung für schulisches Lernen und den Kompetenzerwerb (Kunter & Ewald, 2016; Maaz & Dumont, 2019) – von soziodemographischen Merkmalen wie der familiären Herkunft. Daraus kann jedoch nicht geschlossen werden, dass soziodemographische Merkmale für Lehr-Lernprozesse im adaptiven Unterricht nicht relevant seien. Vielmehr sind diese vermittelt über andere Merkmale wirksam (Cortina & Pant, 2018).

Der vorliegende Beitrag konkretisiert relevante Heterogenitätsdimensionen im adaptiven Unterricht. Davon ausgehend werden Implikationen für Unterrichtsforschung und -praxis abgeleitet. Mit Blick auf die quantitative Heterogenitätsforschung ist die Erfassung vielfältiger Schüler*innenmerkmale auf der Klassenebene notwendig, um lehr-lernrelevante Heterogenitätsdimensionen umfassend zu analysieren. Für die qualitative Untersuchung von Heterogenität bieten sich rekonstruktive Verfahren an, um Schüler*innenmerkmale zu untersuchen, die Lehrkräfte möglicherweise unbewusst berücksichtigen. Im Hinblick auf die schulische Praxis legen die Ergebnisse nahe, dass die Konkretisierung von lehr-lernrelevanten Schüler*innenmerkmalen einen wichtigen Baustein darstellt, um einen produktiven und adaptiven Umgang mit Heterogenität im Unterricht zu realisieren.

PS-78: schulisches Engagement und Entwicklung: Individuelle, soziale und familiäre Faktoren

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 242

Chair der Sitzung: Felix Bittmann

Entwicklung mentaler Fähigkeiten bei jungen Lernenden: Deutschkompetenzen sagen Anpassungsfähigkeit, Emotionsregulation und Coping vorher

Inga Benz, Kristin Kersten

Universität Mannheim, Deutschland

Kinder mit Migrationshintergrund schneiden in großen Bildungsstudien oft schlechter ab als ihre Altersgenossen, was jedoch nicht auf den Migrationshintergrund selbst, sondern auf häufig damit korrelierende soziale Unterschiede zurückzuführen ist (OECD, 2018). Zudem steht Schulerfolg in Verbindung mit (bildungs-)sprachlichen Kompetenzen in der Unterrichtssprache Deutsch (Gogolin & Lange, 2011). Andererseits wird in der jüngsten Debatte zur *Bilingual Advantage (BA) Hypothesis* angenommen, dass häufige Sprachwechsel, bedingt durch migrationsbedingte Mehrsprachigkeit, zu kognitiven Vorteilen führen können (Bialystok, 2017). Entsprechende Befunde existieren in Bezug auf herkunftssprachliche und schulisch bedingte Mehrsprachigkeit (Nicolay & Poncelet 2015, Trebits et al. 2021). Sie können auf ein bisher übersehenes Potenzial für Schüler*innen mit mehrsprachigem Hintergrund hinweisen, auf das es im Bildungskontext verstärkt einzugehen gilt (s.z.B. Ansätze zum sprachsensiblen Unterricht, Gogolin et al., 2020; Rossner & Bolitho, 2022).

BA-Ergebnisse bezüglich exekutiver Funktionen sind jedoch uneinheitlich und werden kontrovers diskutiert (z.B. Lehtonen et al., 2018; Donnelly et al., 2019; Übersicht: Festman et al., 2023). Im Rahmen dieser Debatte unterscheiden Greve et al. (2021) und Nicoladis (2016) zwischen Effekten auf grundlegende exekutive Funktionen und höhere mentale Fähigkeiten wie divergentes Denken (Kharkhurin, 2009), Emotionsregulation und Anpassungsfähigkeit (Koch, 2023). Studien mit jungen Erwachsenen legen eine Verbindung zwischen Mehrsprachigkeit und höheren mentalen Fähigkeiten nahe (Greve et al., 2021; Koch et al., 2023). Während die BA den Sprachwechsel als zugrundeliegenden Mechanismus für das Training exekutiver Funktionen annimmt (Bialystok, 2017), vermuten Koch et al. (2023) eine Stimulierung der Kognition durch sprachlich und kulturell heterogene Lebenskontexte von Mehrsprachigen.

Um differenzielle Effekte von Mehrsprachigkeit und bildungssprachlichen Kompetenzen auf höhere mentale Fähigkeiten für jüngere Lerner zu untersuchen, wurde eine querschnittliche Fragebogenstudie mit 144 Kindern (5.-7. Klasse, $M=12,23$ Jahre, $SD=0,81$) durchgeführt.

Fragestellungen:

1. Sagt der Grad der Mehrsprachigkeit (GM) von jungen Lernenden ihre mentalen Fähigkeiten höherer Ordnung vorher?
2. Sagen die Deutschkompetenzen (DK) von jungen Lernenden ihre mentalen Fähigkeiten höherer Ordnung vorher?

Zur Erfassung der abhängigen Variablen wurden standardisierte Skalen zur Emotionsregulation (FREE-Skala, Burton & Bonanno, 2016), Anpassungsfähigkeit (FGA-Skala, Brandtstädter, 2009) und Coping (CFS-Skala, Kato, 2012) verwendet (deutsche Übersetzungen in kindgerechter Sprache). Die DK wurden als Summe selbsteingeschätzter Kompetenzen im Sprechen, Hören, Lesen und Schreiben, der GM als Summe der Sprachkompetenzen im Deutschen, der besten weiteren Sprache und der Häufigkeit des Sprachgebrauchs operationalisiert (LSBQ; Anderson et al., 2018).

Separate Regressionsanalysen mit GM und DK als Prädiktoren und den kognitiven Variablen zeigen keine Effekte für den Grad der Mehrsprachigkeit; die Deutschkompetenzen sagen die mentalen Fähigkeiten signifikant vorher (FREE: $R^2=.040$, $\beta=.218$, $p=.012$; FGA: $R^2=.045$, $\beta=.212$, $p=.018$, Coping : $R^2=.032$, $\beta=.179$, $p=.037$. Diese Effekte bleiben unter Kontrolle des Alters bestehen.

Während diese Befunde die Ergebnisse von Greve et al. (2021) und Koch et al. (2023) bei jungen Erwachsenen nicht bestätigen, können sie darauf hinweisen, dass sich Anpassungs- und Regulationsfähigkeiten anhand sprachlich-kulturell vielfältiger Situationen erst während des Erwachsenwerdens ausbilden (Koch, 2023). Sie zeigen zudem die Bedeutung der bildungssprachlichen Entwicklung in der weiterführenden Schule. Sie implizieren die Notwendigkeit von weiteren Studien zur Rolle von Sprachkompetenzen für die kognitive Entwicklung und legen nahe, geeignete Fördermaßnahmen für den Aufbau hoher Sprachkompetenzen zu implementieren (s. Ansätze zum sprachförderlichen/sprachsensiblen Unterricht, Gogolin et al., 2020; Sahlender & Gantefort, 2022).

Eine aktuelle Bewertung von Ablenkungseffekten in der Sekundarstufe

Felix Bittmann

Lifbi Bamberg, Deutschland

In den vergangenen Jahrzehnten wurden zahlreiche Reformen des deutschen Bildungssystems mit der Absicht durchgeführt, das System insgesamt transparenter und durchlässiger zu gestalten (Hurrelmann 2013). Besonders der Sackgassencharakter einiger Schulformen wurde dabei möglichst reduziert, beispielsweise durch die Möglichkeit, Bildungsabschlüsse kontinuierlich erwerben zu können und einen sequentiellen Bildungsweg durch das System zu wählen. So sollte heutzutage auch der Übergang auf eine nicht-gymnasiale Schulform nach der Primarstufe nicht ausschließen, letztlich das Abitur erwerben zu können. Empirisch zeigt sich zudem tatsächlich, dass ein nicht zu vernachlässigender Anteil von SchülerInnen auf nicht-gymnasialen Schulformen Aspirationen für das Abitur berichtet, dieser Abschluss also als finales Bildungsziel betrachtet wird (Bittmann 2023). Wenn nun SchülerInnen mit anfangs hohen Aspirationen auf nicht-gymnasiale Schulformen übergehen, aber diese hohen Aspirationen im Zeitverlauf verlieren, so bezeichnen wir dies als Ablenkungsprozess (Müller und Pollak 2004). Der vorliegende Beitrag überprüft empirisch mit Längsschnittdaten, wie stark diese Ablenkungsprozesse in der Sekundarstufe tatsächlich ausfallen.

Der aktuelle Forschungsstand ist ausbaufähig. Zwar verweisen einige ältere Studien auf potentielle Ablenkungsprozesse, jedoch werden diese nicht quantifiziert und näher untersucht (Becker und Hecken 2008). Eine neuere Studie untersucht mit Retrospektivdaten für Personen, die zwischen 1960 und 1980 geboren wurden, ob die Einführung neuer Schulformen insgesamt Bildungsungleichheit reduzieren konnte, kann dafür jedoch keine Belege finden (Schindler und Bittmann 2021). Eine nachfolgende Untersuchung mit Prospektivdaten für Kinder, die ab 2010 in die Sekundarstufe übergegangen sind, zeigt auf, dass sich deutliche Anzeichen für Ablenkungsprozesse finden (Bittmann und Schindler 2021). So sinken selbst für ein positiv selektiertes Sample von SchülerInnen in den nicht-gymnasialen Schulformen die Aspirationen für das Abitur von Anfangs 100%

bis zur Klassenstufe 9 auf knapp 80%. Da der Beobachtungszeitraum der Untersuchung damit endet, kann die letztlich realisierte Ablenkung nicht genau beurteilt werden. Eine aktuelle Studie, die sich mit der Frage befasst, ob die Abschaffung bzw. Zusammenlegung von Schulformen (De-tracking) die soziale Ungleichheit reduziert, kommt zu einem negativen Fazit und berichtet eine tendenziell eher steigende Ungleichheit durch derartige Reformen (Holtmann u. a. 2024). Die Frage bleibt daher, welchen Einfluss Schulformen insgesamt auf das Erreichen hoher Bildungstitel ausüben und wie stark Ablenkungsprozesse ausfallen.

Mit aktuellen Daten der NEPS Startkohorte 3 können nun erstmals auch finale Bildungstitel untersucht werden, da alle SchülerInnen die Sekundarstufe inzwischen abgeschlossen haben. Gemessen wurde dabei, ob letztlich ein Abitur erreicht wurde oder nicht. Zur Beantwortung unserer Fragestellung wählen wir ein Sample von SchülerInnen, die zu Beginn der Sekundarstufe hohe Aspirationen (für das Abitur) aufweisen. Wir berichten rein deskriptive Ergebnisse, möchten aber auch den Schulzweigeffekt abschätzen. Dazu erzeugen wir mittels Entropy Balancing ein Sample von SchülerInnen, die in Klassenstufe 5 verschiedene Schulformen besuchen (Gymnasium vs. andere Schulform), jedoch in ihren sonstigen Charakteristika sehr ähnlich sind. Wir matchen dabei auf Kompetenzen in den Fächern Mathematik und Deutsch, soziale Herkunft der Familie, Geschlecht, Bundesland und zahlreiche weitere Faktoren, um einen statistisch validen Vergleich zu ermöglichen.

Unsere Ergebnisse zeigen auf, dass ein rein deskriptiver Unterschied von ca. 27 Prozentpunkten besteht ($N = 1.938$). Dies bedeutet, dass SchülerInnen mit anfangs hohen Aspirationen, die nicht auf das Gymnasium übergegangen sind, eine ca. 27 Prozentpunkte geringere Wahrscheinlichkeit für den Erwerb des Abiturs haben als SchülerInnen auf dem Gymnasium. Nach der Durchführung des Matchings reduziert sich diese Differenz auf ca. 13 Prozentpunkte. Robustheitsanalysen mit Regressionen kommen zu sehr ähnlichen Werten. Dies zeigt auf, dass für eine vergleichbare Schülerpopulation der „Effekt“ der Schulform deutlich reduziert ist, wenn man Selektion in die Schulformen berücksichtigt. Wir bewerten diesen Unterschied zwischen den Schulformen zwar als problematisch, jedoch als weniger gravierend als auf Basis früherer Studien angenommen. Offenbar bestehen im deutschen System weiterhin Ablenkungseffekte, die noch weiter reduziert werden sollten, jedoch scheinen diese nicht übermäßig stark.

Bedeutung individueller und familiärer bildungsbezogener Faktoren für Schulentfremdung

Ann-Kathrin Ouarda, Isabell Martin, Michaela Gläser-Zikuda, Berit Breins

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Schulentfremdung kann sowohl auf individueller als auch auf gesellschaftlicher Ebene weitreichende Konsequenzen haben (Hadjar et al., 2019; Hascher & Hagenauer, 2010; Quenzel & Hurrelmann, 2019). Unter Schulentfremdung werden negative Kognitionen und Affekte gegenüber sozialen und akademischen Aspekten der Schule verstanden und es wird zwischen der Entfremdung vom Lernen sowie der Entfremdung von Mitschüler:innen und Lehrkräften differenziert (Gläser-Zikuda et al., 2024; Hascher & Hadjar, 2018; Marcin et al., 2020). Individuell kann Schulentfremdung zu geringerer Partizipation am Unterricht bis hin zum Schulabbruch sowie eingeschränkter gesellschaftlicher Teilhabe führen (Hadjar et al., 2019; Hascher & Hagenauer, 2010). Für die Gesellschaft fehlen Fachkräfte und es entstehen erhebliche Kosten für Sozialhilfearbeit (Allmendinger et al., 2011; Quenzel & Hurrelmann, 2019). Die Bedeutung verschiedener schulbezogener Faktoren für Schulentfremdung, wie die Lehrer:innen-Schüler:innen-Beziehung, Schüler:innen-Schüler:innen-Beziehung und Unterrichtsqualität wurde bereits beleuchtet (Gläser-Zikuda et al., 2024; Mahmoudi et al., 2018; Morinaj et al., 2023; Schmid et al., 2021). Über die Relevanz individueller und familiärer lern- bzw. bildungsbezogener Faktoren für Schulentfremdung, wie z.B. die Bildungsaspiration oder die Lernunterstützung durch Eltern, ist hingegen weniger bekannt. Aber es gibt Hinweise darauf, dass z.B. autonomieförderndes instruktionales Elternverhalten mit Blick auf die Lern- und Leistungsentwicklung von Heranwachsenden günstig ist (Teuber et al., 2022). Diese Studie untersucht daher, welche Bedeutung individuelle (Bildungsaspiration, Resilienz, Leistung) und familiäre Faktoren (elterliches Instruktionsverhalten) auf die Schulentfremdung haben. Insgesamt $n = 3172$ Schüler:innen (46.2% weiblich; Alter: $M = 12.97$, $SD = 1.66$) aus 330 Klassen der Jahrgangsstufen 5 bis 9 eines deutschen Bundeslandes nahmen an einer Befragung teil. Die drei Dimensionen von Schulentfremdung (Hascher & Hadjar, 2018), die eigene Bildungsaspiration der Schüler:innen (Ditton et al., 2002), die Resilienz (Leppert et al., 2008) und das kindperzipierte Instruktionsverhalten der Eltern (mit den Subskalen autonomieunterstützende Instruktion, kontrollierende Instruktion, Interesse an Schule sowie Hausaufgabenhilfe) (Wild et al., 2001) wurden gemessen. Die Skalen zeigen alle zufriedenstellende bis gute Reliabilitäten (von $\alpha = .71$ bis $\alpha = .88$). Die Ergebnisse von Mehrebenenregressionen zeigen, dass sich die kontrollierende Instruktion ($B = -.13$), das Interesse der Eltern an der Schule ($B = -.10$) und die Resilienz ($B = -.22$) negativ auf die Entfremdung vom Lernen auswirken. Zudem zeigt sich, dass Jungen entfremdeter vom Lernen sind als Mädchen ($B = -.05$). Für die Entfremdung von den Lehrkräften erwiesen sich die eigene Bildungsaspiration ($B = .29$) und das Alter ($B = .07$) als signifikant positive Prädiktoren, während die kontrollierende Instruktion ($B = -.22$) und die Resilienz ($B = -.20$) signifikant negative Prädiktoren darstellen. Mädchen scheinen entfremdeter von ihren Lehrkräften ($B = .10$) sowie von ihren Mitschüler:innen ($B = .13$) zu sein. Die Bildungsaspiration kann als positiver Prädiktor für die Entfremdung von den Mitschüler:innen identifiziert werden ($B = .12$). Die Resilienz ($B = -.25$), das Interesse der Eltern an der Schule ($B = -.10$), die autonomieunterstützende Instruktion ($B = -.06$) und die kontrollierende Instruktion ($B = -.07$) sind negative Prädiktoren der Entfremdung von den Mitschüler:innen. Die Ergebnisse zeigen, dass sich resilientere Schüler:innen, die kontrollierendes elterliches Instruktionsverhalten wahrnehmen, in allen drei Dimensionen weniger entfremdet fühlen. Das Interesse der Eltern an der Schule kann als präventiv für die Entfremdung vom Lernen und von den Mitschüler:innen betrachtet werden, während eine hohe Bildungsaspiration mit einer höheren Entfremdung von den Lehrkräften und Mitschüler:innen einherzugehen scheint. Einschränkend ist allerdings mit Blick auf die Itemformulierungen anzunehmen, dass eher erlebter Leistungsdruck erfasst wurde. Für die Entfremdung von den Mitschüler:innen erweist sich das autonomieunterstützende Instruktionsverhalten der Eltern als schützender Faktor. Zusammenfassend legen die Ergebnisse nahe, dass sowohl Resilienz als auch die elterliche Instruktion mit Blick auf Schulentfremdung als präventiv eingeschätzt werden können. Entsprechende Implikationen ergeben sich für die Förderung der Resilienz von Lernenden sowie für die Zusammenarbeit von Schule und Eltern.

PS-79: Berufliche Anforderungen und Bildungswege

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 133

Chair der Sitzung: Helena Geisler

Bildung als Schlüssel zum Arbeitsmarkt? Wie Rekrutierende den Bildungsabschluss als Auswahlkriterium begründen

Helena Geisler, Miriam Grønning, Irene Kriesi

Eidgenössische Hochschule für Berufsbildung EHB, Schweiz

In Ländern mit hochgradig stratifizierten Bildungssystemen wie Deutschland, den Niederlanden und der Schweiz sind jobspezifische Bildungsabschlüsse meist Zugangsvoraussetzung für Fachkräftestellen (Imdorf et al. 2023) – und damit Schlüssel für gesellschaftliche und ökonomische Teilhabe. Um diese Schlüsselrolle von Bildung im Zugang zum Arbeitsmarkt besser zu verstehen, bedarf es einen Fokus auf Personalauswahlprozesse. In Studien zur Bewertung einzelner Abschlüsse bzw. Bildungswege bei Auswahlverfahren zur Stellenbesetzung werden meist akademische und berufsbildende Abschlüsse gegenübergestellt (Hippach-Schneider und Weigel 2013; Maier et al. 2020). So zeigen z. B. Hippach-Schneider und Weigel (2013), dass Personalverantwortliche in der Schweiz und Deutschland Bildungsabschlüsse präferierten, denen sowohl praktische als auch theoretische Erfahrung zugesprochen wird.

Das Auswahlkriterium «Bildungsabschluss» wird mit dem Verweis auf tatsächlich vorhandene Fähigkeiten (Becker 1962), auf das Signal für Lernfähigkeit (Spence 1973) und auf externe Vorgaben (Collins 1979; Mackert 2004) begründet (vgl. Tholen 2023). Bildungsabschlüsse haben im Personalauswahlprozess weitere Funktionen: Sie werden z. B. dazu genutzt, die Anzahl der Bewerbungen reduzieren, oder als Zeichen für Überqualifikation oder zu hohe Ansprüche gewertet (vgl. Imdorf 2009). Es bleibt jedoch unklar, wie Funktion und Begründung des Auswahlkriteriums «Bildungsabschluss» mit weiteren Merkmalen der Bewerber*innen wie z. B. Berufserfahrung, Alter und Nationalität sowie Kontextmerkmalen der Stelle, Firma und Branche zusammenhängen.

Um sowohl die Bewertung unterschiedlicher Abschlüsse und ihre Funktion als Auswahlkriterium als auch den Zusammenhang mit weiteren Bewerber*innen- und Kontext-Merkmalen zu verstehen, untersuche ich die Begründungen der Rekrutierenden: Wie legitimieren sie ihre Auswahlverfahren, wie verwenden, bewerten und begründen sie Bildungsabschlüsse als Auswahlkriterium? Ich arbeite heraus, welche Muster sich in der Begründung bzw. Nicht-Begründung von spezifischen Bildungsabschlüssen als Auswahlkriterium zeigen und wie sie sich durch weitere Bewerber*innen- und Kontext-Merkmale erklären lassen.

Theoretischer Rahmen meiner Analyse sind die Konventionensoziologie (Boltanski und Thévenot 2018; Diaz-Bone 2018) und die Soziologie der Regimes of Engagement (Thévenot 2001; Hansen 2023): Rekrutierende verweisen mit ihren Begründungen bzw. dem Unbegründet-Lassen auf die verschiedenen Arten, einzelnen Aspekten einer bewerbenden Person, ihrer Bewerbungsunterlagen und insbesondere ihrer Abschlüsse Wert zuzuweisen, ausserdem auf ihre persönlichen Rekrutierungsziele und auf gewohnte Auswahlabläufe. Mit den zugrundeliegenden Konventionen, Zielen und Gewohnheiten legitimieren die Rekrutierenden in unseren Interviews die Funktion und Bewertung von Bildungsabschlüssen.

Datengrundlage sind 34 qualitative, semistrukturierte Interviews mit Rekrutierenden, die als Kontaktpersonen auf aktuellen schweizerischen Stelleninseraten genannt wurden. Die Stelleninserate wurden in sechs Berufsbereichen gesucht, die folgende Kriterien erfüllen: Sie sind sowohl über den berufspraktischen als auch über akademischen Ausbildungsweg zugänglich, unterscheiden sich in ihrem Tätigkeitsspektrum, haben unterschiedliche Geschlechterverteilung und gehören zu den Berufsgruppen mit den meisten ausgeschriebenen Stellen. Um Muster im gesamten Datensatz zu analysieren, eignet sich Reflexive Thematic Analysis (Braun und Clarke 2022). Dazu wurden die für die Forschungsfrage relevanten Interviewpassagen codiert, die Codes anschliessend geclustert und Themes konstruiert, welche die den Codes zugrundeliegenden Konzepte fassen und die Forschungsfrage adressieren.

Erste Ergebnisse zeigen unterschiedliche Rekrutierungsziele und Legitimationen je nach Stelle, Unternehmen und Branche. Ein jobspezifischer Bildungsabschluss wird für Fachkräftestellen vorausgesetzt, da dieser als Nachweis für jobspezifische Qualifikationen gilt. Wie ein Bildungsabschluss bewertet wird, ist von Zuschreibungen zu anderen Bewerber*innen-Merkmalen wie Nationalität und Alter abhängig: Ausländische Abschlüsse werden meist schon in der Vorauswahl ausgeschlossen, unabhängig von Stufe und Art. Ausserdem signalisieren Abschlüsse junger Menschen, dass die Bewerbenden nicht lange auf der Stelle bleiben, da sie sich weiterqualifizieren wollen oder sollen: Insbesondere Kandidat*innen mit einer 3–4-jährigen Berufsausbildung und anschliessender Berufsmatur (Berufsabitur), die für ein Studium berechtigt, wird abgesprochen, dass sie lange auf der Stelle bleiben. Aber auch jungen Menschen mit einer Berufsausbildung ohne Berufsmatur wird abgesprochen, dass sie lange auf der Stelle bleiben: Hier steht die Konvention im Hintergrund, dass sich junge Leute weiterentwickeln sollen und daher Stellen wechseln werden. Tertiär-Abschlüssen ohne Berufserfahrung, also insbesondere Uni/ETH-Abschlüssen dagegen wird eine zu hohe Lohnforderung bei fehlender direkter Einsetzbarkeit zugesprochen.

Berufliche Aspirationen und kognitives Anforderungsniveau von Ausbildungsberufen

Anett Friedrich, Annalisa Schnitzler

Bundesinstitut für Berufsbildung (BIBB), Deutschland

Ein entscheidender Schritt im Bildungsprozess von Jugendlichen ist die Wahl eines Berufes, welcher für einen großen Teil der Jugendlichen ein Beruf aus dem dualen oder schulischen Ausbildungssystem ist (Bundesinstitut für Berufsbildung 2023: 76). Die Ausbildungsberufe in Deutschland sind dabei heterogen, etwa in ihren Zugangschancen (Protsch/Solga 2016), ihren beruflichen Inhalten (Paulus/Matthes 2013) und ihrem kognitiven Anforderungsniveau (Friedrich et al. 2023). Berufe mit höherem kognitiven Anforderungsniveau sollten in stärkerem Maße die Möglichkeit bieten zusätzliche Kompetenzen sowie vermehrt transferierbare Fähigkeiten zu erlernen, die auf dem Arbeitsmarkt entsprechend entlohnt werden sollten (Meyer/Sacchi 2020). Der Zugang zu diesen Berufen hängt stark mit dem Schulabschluss zusammen, je höher dieser ist desto eher gelingt der Übergang (Friedrich et al. in Begutachtung). Gerade Jugendlichen mit maximal Hauptschulabschluss gelingt der Übergang in einen kognitiv anspruchsvolleren Ausbildungsberuf eher selten (ebd.), obwohl sie zum Teil kognitive Fähigkeiten besitzen, die mit denen von Absolventen und Absolventinnen mit Realschulabschluss oder Abitur vergleichbar sind (Holtmann et al. 2018). Allerdings bleibt bisher unklar, ob Jugendliche mit maximal Hauptschulabschluss eher an betrieblichen und institutionellen Barrieren scheitern oder ob sie anspruchsvollere Berufe gar nicht aspirieren. An dieser Stelle setzt unser Beitrag an, wir untersuchen, wer Berufe mit höherem kognitiven Anforderungsniveau aspiriert und welche individuellen Merkmale mit der Aspiration zusammenhängen. Wir betrachten damit die kognitiven Anforderungen, in Ergänzung zu bspw. Interessen und dem Motiv des Statuserhalts, erstmalig als eine gesonderte Dimension von beruflicher Aspiration.

Theoretisch stützen wir uns dabei auf das *Expectancy-Value-Model* (Eccles/Wigfield 2002). Kern der Theorie ist zum einen die subjektive Erfolgserwartung, zum anderen der *subjective task value*, welcher vier Elemente umfasst: (1) den *interest-enjoyment value* (Interesse und Spaß an Tätigkeit); (2) den *attainment value* (Relevanz für Bestätigung der eigenen Identität); (3) den *utility value* (Nützlichkeit für eigene Ziele) und schließlich (4) die *relative costs* (Kosten in Form von z.B. benötigter Anstrengung, Versagensangst und Opportunitätskosten). Gute schulische Leistungen und Lernmotivation sollten mit einem Bedürfnis nach Kompetenzerleben und Spaß an kognitiven Aktivitäten einhergehen (Deci/Ryan 1985; Cacioppo et al. 1996). Darüber hinaus sollte unter Kostengesichtspunkten die rationale Bildungswahl zu leistungsspezifischen Entscheidungen bzgl. der Ausbildungsberufswahl führen (Breen/Goldthorpe 1997; Erikson/Jonsson 1996). Schulabschlüsse stellen häufig die erste Hürde dar, um überhaupt im Auswahlprozess zu bleiben (Protsch 2017; Protsch/Solga 2016), was wiederum von den Jugendlichen antizipiert wird und ihre berufliche Aspiration beeinflussen kann, um übermäßige Kosten zu vermeiden und trotzdem das Ziel eines Ausbildungsabschlusses zu erreichen. Weiterhin nehmen wir an, dass die Risikobereitschaft und der bisherige Lernstress der Jugendlichen damit zusammenhängen sollte, wie die Jugendlichen ihre Kosten in Form von Anstrengung und möglichem Versagen in Bezug auf eine Ausbildung in einem bestimmten Ausbildungsberuf einschätzen.

Als Datengrundlage für unsere Analysen nutzen wir die Startkohorte 3 des Nationalen Bildungspanels (NEPS-Netzwerk 2024). Unsere Grundgesamtheit umfasst Jugendliche, die sich jemals für eine Ausbildung interessiert haben. Die zu erklärende Variable der Untersuchung ist das kognitive Anforderungsniveau des realistisch aspirierten Ausbildungsberufs. Wir verwenden hierfür einen Indikator des Berufspsychologischen Service (BPS) der Bundesagentur für Arbeit (Friedrich et al. 2023). Für unsere Analysen nutzen wir *linear probability models* (Wooldridge 2010).

Die vorläufigen Ergebnisse zeigen, dass Ausbildungsinteressierte mit besseren Schulnoten eher Berufe mit höheren anstelle von niedrigeren kognitiven Anforderungen aspirieren. Jugendliche, die einen Hauptschulabschluss anstreben, aspirieren seltener Berufe mit höheren anstelle von niedrigeren kognitiven Anforderungen, was dafürsprechen könnte, dass sie ihre schlechteren Chancen antizipieren. Die Ergebnisse für die übrigen Merkmale stehen noch aus, werden aber bis zur Konferenz vorliegen.

Gerade in Zeiten des Fachkräftemangels ist es gesellschaftlich relevant zu ergründen, wie man allen Jugendlichen den Weg auch in kognitiv anspruchsvolle Berufe ebnen kann. Dabei hilft es zu verstehen, ob Jugendliche, auch je nach Schulabschluss, dies Berufe gar nicht erst aspirieren oder ob ihnen der Weg durch institutionell und betriebliche Barrieren versperrt wird.

More demanding, more money? Returns to varying levels of cognitive requirements and specificity in vocational education and training

Filiz Koneberg^{1,2}, Paula Protsch^{1,2}, Daniela Rohrbach-Schmidt¹

¹Bundesinstitut für Berufsbildung (BIBB); ²Universität zu Köln

Research on educational returns often focuses on the comparison of secondary educational tracks such as vocational education vs. further general schooling or academic education (Shavit and Müller, 2006; Lavrijsen and Nicaise, 2017; McNally et al., 2020; Silliman and Virtanen, 2022). However, vocational education is highly heterogeneous and stratified along occupations with different levels and types of attainable skills and different labor market returns (Menze, 2019; Grønning and Kriesi, 2022). Recently, skill differences in training occupations of vocational education have received increasing academic attention in explaining differences in returns to training. Eggenberger et al. (2018) examine whether the degree of occupational specificity of skills influences the wage prospects of VET graduates. Their results suggest that graduates with specific skill bundles receive higher monetary returns when working in their trained occupation. In addition, Grønning et al. (2020) show that the impact of occupational specificity on earnings varies by gender and gender type of training occupation (female, neutral or male) and can partly explain the gender wage gap in Switzerland. Research on the importance of cognitive (or general) skills for those with vocational qualifications shows that very broad skills lead to a better match between training occupations and first jobs (Menze, 2017) and that a higher proportions of cognitive skills in training curricula (Langer and Wiederhold, 2023) as well as used cognitive skills increase earnings opportunities of VET graduates. This positive effect of cognitive skills is most often attributed to an increase in flexibility in the labour market (Lavrijsen and Nicaise, 2017) and/or a better basis for lifelong learning (Tobback et al., 2024).

In sum, we know from previous research that occupational specificity and cognitive skills play a role in explaining differences in VET returns. However, we know little about how the interplay of occupational specificity and cognitive level might shape the earnings prospects of graduates. To address this research gap, we focus on the case of Germany and pose the following research questions:

1. Whether and to what extent do monetary returns to vocational education differ by the cognitive requirement level and the specificity of the training occupation?
2. Do monetary returns to cognitive level and specificity vary by the gender-type of the training occupation?

Theoretical assumptions on how different levels of cognitive requirements and occupational specificity might affect the wages are derived from human capital theory (Becker 1964), labour market signaling (Spence, 1973; Arrow, 1973; Stiglitz, 1975), skill complementarity approaches (Lazear, 2004; Allinson and Hayes, 1996) and cultural devaluation theories (Acker, 1991; England, 1992).

We run multilevel models with individuals nested in occupations utilizing the BIBB/BAuA Employment Survey ("ETB") of 2012, 2018 and 2024. The ETB is a cross sectional survey of "core workers" and is therefore representative for German employees who are at least 15 years of age and work a minimum of 10 hours per week. First results show, that higher cognitive requirements positively affect wages of graduates regardless of job changes, while higher specificity only pays off in jobs that are related to the training occupation. Although theoretically expected by complementarity approaches, we do not find any significant complementary effects of high cognitive level and high specificity increasing earnings of VET graduates. Overall we see, that higher cognitive demands seem to be the main driver in explaining differences in wages attributable to the training occupation. This could be due to the fact, that most of the VET occupations are by design quite specific and therefore higher degrees of specificity are less relevant for earnings than higher cognitive demands.

Neudefinition und Bewertung von Schwierigkeitsfaktoren in kaufmännischen Abschlussprüfungen.

Sebastian Schlunkes, Esther Winther

Universität Duisburg-Essen, Deutschland

Abstract: Dieses Paper zielt darauf ab, drei Schwierigkeitsfaktoren für Prüfungsaufgaben neu zu definieren und ihre Wirksamkeit zu bewerten, basierend auf der umfangreichen Forschung von Frau Professor Esther Winther aus dem Jahr 2010. Die Studie

analysiert die Ergebnisse von drei schriftlichen Prüfungen, die über drei Jahre hinweg mit insgesamt 458 Auszubildenden durchgeführt wurden. Diese Auszubildenden befinden sich in der Ausbildung zum Industriekaufmann in Deutschland.

Die Schwierigkeitsfaktoren wurden wie folgt kategorisiert: 1) Spezifität, die widerspiegelt, wie spezifisch die Wissensinhalte sind, die in einer Aufgabe abgefragt werden. 2) Kognition, die das erforderliche Verständnisniveau zur erfolgreichen Lösung einer Aufgabe beschreibt, basierend auf der Taxonomie von Bloom; und 3) Modellierung, die die Komplexität einer Aufgabe sowie die benötigte Arbeitsgedächtniskapazität und kognitive Belastung widerspiegelt, basierend auf den Forschungen von Sweller (1988).

Fragestellung: Kann eine strukturierte Bewertung von Aufgaben anhand dieser drei definierten Schwierigkeitsfaktoren die Vorhersage der Aufgabenschwierigkeit verbessern und eine verlässliche Anleitung für ehrenamtliche Prüfungsaufgabenersteller/innen bieten, um ausgewogenere Prüfungsaufgaben zu erstellen?

Methode: Unsere Studie konstruierte eine kaufmännische Abschlussprüfung für Industriekaufleute mit 25 Aufgaben. Diese Testprüfung wurde über drei Jahre hinweg mit 458 Auszubildenden durchgeführt. Die Lösungsquote jeder Aufgabe wurde aufgezeichnet, und die Aufgaben wurden anhand der drei Schwierigkeitsfaktoren bewertet. Jeder Faktor wurde in drei Stufen unterteilt.

Wir führten eine univariate ANOVA und eine Regressionsanalyse durch, um die Bedeutung der Schwierigkeitsfaktoren und ihrer Stufen auf die Lösungsraten zu bewerten. Die ANOVA prüfte die Haupteffekte und Interaktionen zwischen den Faktoren, während die Regressionsanalyse den Einfluss jedes Faktors auf die Aufgabenleistung untersuchte.

Ergebnisse: Die Ergebnisse der ANOVA zeigten, dass alle drei Faktoren—Spezifität ($F(2,14)=75,478$, $p<0,001$), Kognition ($F(2,14)=54,449$, $p<0,001$) und Komplexität ($F(2,14)=27,685$, $p<0,001$)—signifikante Effekte auf die Lösungsraten hatten, wobei keine signifikanten Interaktionseffekte zwischen den Faktoren festgestellt wurden. Post-hoc-Tests ergaben signifikante Unterschiede zwischen allen Stufen innerhalb jedes Faktors ($p<0,001$). Die Regressionsanalyse unterstützte diese Ergebnisse, wobei alle drei Faktoren signifikante negative Koeffizienten aufwiesen (Spezifität: $\beta=-0,547$, $p<0,001$; Kognition: $\beta=-0,490$, $p<0,001$; Modellierung: $\beta=-0,346$, $p<0,001$), was darauf hinweist, dass höhere Stufen dieser Faktoren mit niedrigeren Lösungsraten korrelieren.

Schlussfolgerung: Die Studie zeigt, dass die Verwendung von Spezifität, Kognition und Modellierung als strukturierte Schwierigkeitsfaktoren eine verlässliche Methode zur Vorhersage der Aufgabenschwierigkeit bietet. Dieser Ansatz kann als wertvolle Anleitung für ehrenamtliche Prüfungsaufgabenersteller dienen, um ausgewogenere Prüfungen mit einer genaueren Darstellung von leichten, mittleren und schweren Aufgaben zu erstellen. Die Implementierung dieses Rahmens kann die Qualität und Zuverlässigkeit der Kompetenzbewertungen in der deutschen Berufsausbildung verbessern und letztlich die Ausbildungsergebnisse der Auszubildenden optimieren.

PS-80: Bildungsentscheidungen und ihre Auswirkungen auf Persönlichkeit und Kompetenzen

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 135

Chair der Sitzung: Kamila Cygan-Rehm

DIPALOG - Förderung überfachlicher Lebens- und Schlüsselkompetenzen (ÜLS) von Kindern und Jugendlichen in der Ostschweiz

Dölf Looser

Pädagogische Hochschule St. Gallen, Schweiz

Theoretischer Hintergrund

Stanford (2010) forderte schon 1977, dass in neuen Schulklassen von Beginn weg durch soziales Lernen die Förderung personal-sozialer Kernkompetenzen geschaffen werden sollte. Die Wirksamkeit solcher Interventionen konnte durch empirische Studien teilweise nachgewiesen werden (Moch, 2002; Zick, 2002; Boeger, Dörfler & Schut-Ansteeg, 2006; Looser, 2017). Es kann bilanziert werden, dass (Schulentwicklungs-)Projekte mit erlebnispädagogischen und gruppendynamischen Elementen zu Beginn einer neuen Schulklasse durchaus Wirkung auf die ganzheitliche Entwicklung der Persönlichkeit erzielen können.

Diesem Ansatz pflichtet auch Roth (2011) bei: Der Persönlichkeitsbildung der Schüler sollte eine ebenso grosse Bedeutung wie der Wissensvermittlung zukommen. Sie sollte einerseits in stärkerem Masse und andererseits professioneller in den Schulen stattfinden (S. 291). Er verweist auf die grossen Herausforderungen an Lehrpersonen, welche sich mit sehr unterschiedlichen Sozialisationsbedingungen ihrer Schüler konfrontiert sehen. Auch Evertson und Weinstein (2006) fordern eine erweiterte Definition der Klassenführung, bei der nicht nur die Aufrechterhaltung von Ruhe und Ordnung, sondern auch die Förderung von personal-sozialen Kompetenzen bei Schülern vermehrt berücksichtigt werden. Sie definieren "Classroom Management" als "the actions teachers take to create an environment that supports and facilitates both academic and social-emotional learning" (S. 4).

Zusammenfassend lässt sich ableiten, dass sich Investitionen in die Förderung überfachlicher Kompetenzen gerade in Zeiten von diversen gesellschaftlichen Herausforderungen und unterschiedlichen Sozialisationsbedingungen lohnen dürften.

Forschungsfrage

Das interkantonale Schulentwicklungsprojekt nimmt Bezug auf die überfachlichen Kompetenzen im Lehrplan der 21 Deutschschweizer Kantone (Lehrplan 21), erweitert diesen mit weiteren Lebens- und Schlüsselkompetenzen (21st century skills). Ziel des Forschungsbeitrages ist es, die Wirkung eines dialogischen Standortgesprächs zwischen Eltern, Kind und Lehrperson und nachfolgender Förderung anhand von DIPALOG-Lehrmaterialien mittels eines Quasi-Experimentaldesigns mit zwei Messzeitpunkten bei Schülern der Volksschule inklusive Kontrollgruppenvergleich in den ÜLS (überfachliche Lebens- und Schlüsselkompetenzen) nachzuweisen.

Methode

An der Untersuchung nahmen im Schuljahr 2023-24 insgesamt 10 Schulen (88 Schulklassen der Versuchsgruppe) und 3 Schulen (13 Klassen der Kontrollgruppe) der Volksschule in der Ostschweiz teil. Dabei wurden mit einem klassischen quasi-experimentellen Design (zwei Messpunkte, Versuchs- und Kontrollgruppe) nach dem ersten Quartal und am Schuljahresende 2024 per Fragebogen überfachliche Variablen wie Sorgfalt, Zuverlässigkeit, Konfliktfähigkeit, Einfühlungsvermögen, Lern- und Leistungsmotivation sowie z.B. Umgang mit Stress erhoben. Die im Projekt erarbeiteten Instrumente und Skalen werden ebenfalls vorgestellt. Als zentrale Intervention wurde zwischen t1 und t2 eine mehrwöchige Trainingsphase eingeschaltet, in welcher die Schüler individuell oder klassenbezogen in einer ÜLS schwerpunktmässig arbeiteten. Die jeweiligen Lehrpersonen-Teams der Versuchsgruppe durchliefen im Vorhinein zwei halbtägige Weiterbildungen. Mit Hilfe von t-Test und Varianzanalysen mit Messwiederholung mit einem dreifaktoriellen Design (experimentelle Bedingung, Geschlecht, Messzeitpunkte) werden Wirkungs- und Interaktionseffekte untersucht.

Ergebnisse

und

Diskussion

Erste Ergebnisse weisen eine positive Bilanz der DIPALOG-Trainings aus: Während sich die Mittelwerte zur Sorgfalt, Zuverlässigkeit, Konfliktfähigkeit, zum Einfühlungsvermögen, Lern- und Leistungsmotivation sowie zum Umgang mit Stress bei der Versuchsgruppe signifikant verbesserten ($720 < n < 612$), zeigt sich bei diesen Themen in der jeweiligen Varianzanalyse kein signifikantes Ergebnis. Bei über einem Drittel der Kompetenzen sind die Mittelwerte der Kontrollstichprobe im ersten Messzeitpunkt signifikant höher als die der Versuchsgruppe. Signifikante Effekte können bei einzelnen Kompetenzen wie Konzentration und Zeitmanagement in der Varianzanalyse nachgewiesen werden. Obwohl viele der Ergebnisse sich wie erwartet theoriekonform zeigen, können signifikante längsschnittliche Effekte mit Kontrollgruppenvergleich nicht durchwegs nachgewiesen werden.

Bei der Interpretation dieser nicht-randomisierten Studie muss mit Bezug zum quasi-experimentellen Design auf den Zeitdruck und das manchmal mangelnde Commitment der Lehrpersonen in den Innovationsschulen und insofern auf mögliche Stichprobenverzerrungen hingewiesen werden. Die Erkenntnisse sollen praktizierende Klassenlehrpersonen und Schulen motivieren, sich den zwar schwierig zu messenden, aber sehr bedeutsamen Konstrukten der überfachlichen Kompetenzen anzunehmen und diese zu fördern.

Timing of school entry and personality traits in adulthood

Anton Barabasch^{1,3}, Kamila Cygan-Rehm^{1,2}, Andreas Leibing^{1,4}

¹Technische Universität Dresden (TUD); ²Leibniz-Institut für Bildungsverläufe (IfBi Bamberg); ³Friedrich-Alexander Universität Erlangen-Nuremberg (FAU); ⁴Deutsches Institut für Wirtschaftsforschung (DIW Berlin)

Research question: This paper examines the long-term consequences of later school entry on personality traits. Entering school is an important life event, and school enrollment laws are among the most common public policies worldwide. They typically define a cut-off date such that all children born before this date are enrolled in a given school year, and those born after this date are enrolled one year later. The cut-off date largely determines the timing of a child's schooling, and extensive research has examined its consequences for human capital development.

Research to date has focused primarily on cognitive skills and typically found that older school entrants have a significant advantage in terms of early student achievement (e.g., Bedard and Dhuey, 2006; Puhani and Weber, 2008). However, these differences disappear over time (e.g., Smith, 2009; Crawford et al., 2014). Few studies have documented desirable effects on students' behavioral and socioemotional development in the early years of school (e.g., Mühlenweg et al., 2012; Crawford et al., 2014). Much less is known about the consequences of later school entry for the development and persistence of personality traits.

Theoretical background: Long-run effects on personality traits would be consistent with the dynamic model of skill formation (Cunha and Heckman, 2007, 2008), i.e., even if some differences initially arise solely because of chronological age advantage, self-productivity and dynamic complementarity of skills create mechanisms that could sustain the effects into the teenage years and beyond. Moreover, starting school can be a stressful life event (Landersø et al., 2020), and it occurs at a time in a child's life that is critical for human capital formation (Currie and Almond, 2011). Once formed, personality traits do not change rapidly in later life and follow relatively stable trends (e.g., Cobb-Clark and Schurer, 2013; Elkins et al., 2017). Thus, the unique combination of the importance of school enrollment and its timing within the life cycle is another reason why we might expect persistent effects on personality traits.

Method: We provide a comprehensive analysis of the long-term effects on the Big Five Inventory of personality traits. To establish causality, we exploit plausibly exogenous, institutionally induced variation in the timing of individuals' school entry. Specifically, we exploit statutory cut-off rules for school enrollment in Germany within a regression discontinuity design (RDD). The German context is particularly suitable for our research design because, unlike the US or the UK, there is no automatic relationship between the timing of school entry and compulsory schooling requirements. We apply this method to large-scale survey data from the National Educational Panel Study (NEPS-SC6) and the German Socio-Economic Panel (SOEP).

Results and their significance: We find that older school entrants enjoy a long-term reduction in levels of neuroticism. This trait refers to emotional volatility and is typically associated with poor well-being and worse socioeconomic outcomes (Gensowski et al., 2021). The effect is driven by women, which is consistent with recent evidence for South Korea of positive effects on higher self-esteem during middle and high school for girls but not for boys (Shin, 2023). However, we also show that the effect on neuroticism persists into adulthood. This has important implications for gender gaps in the labor market, as women typically score significantly higher on neuroticism at all stages of life, which puts them at a disadvantage. Moreover, while previous research suggests that personality differences are mainly driven by socioeconomic inequalities and parenting styles (e.g., Elkins and Schurer, 2020; Falk et al., 2021), our results are more consistent with the literature showing that educational policies may also matter (e.g., Chetty et al., 2011; Heckman et al., 2013; Algan et al., 2022).

Zwischen Schule und Universität: Wie unterscheidet sich das sozioemotionale Erleben von Frühstudierenden in den beiden Kontexten und wovon hängt es ab?

Lorena Fleischmann, Lilly Knoll

Ruprecht-Karls-Universität Heidelberg, Deutschland

Theoretischer Hintergrund: Deutschlandweit absolvieren aktuell mehr als 1,384 Schüler(innen) ein sog. Frühstudium, d.h. sie besuchen bereits während ihrer Schulzeit reguläre Lehrveranstaltungen an Hochschulen, nehmen an Prüfungen teil und erwerben ggf. Scheine, die auf ein späteres Regelstudium angerechnet werden können (Deutsche Telekom Stiftung, 2011, 2018). Über die demografischen Merkmale der Frühstudierenden (vgl. Deutsche Telekom Stiftung, 2006, 2013, 2018; Solzbacher, 2006–2007, 2011) sowie über deren Bewertung der Fördermaßnahme im Nachhinein (vgl. Fleischmann et al., 2022; Stumpf & Gabert, 2016) ist bereits einiges bekannt. Allerdings gibt es bisher kaum Studien, die untersucht haben, wie die Frühstudierenden ihre Teilnahme an den zwei unterschiedlichen Bildungskontexten sozial und emotional erleben (z.B. Pannell et al., 2021). Dies ist vor allem vor dem Hintergrund erstaunlich, dass das Frühstudium teilweise hohe zeitliche und fachliche Anforderungen an die teilnehmenden Schüler(innen) stellt (z.B. Kaden, 2016; Solzbacher, 2006–2007) und dass möglicherweise Kontrasteffekte zwischen den beiden Kontexten auftreten können (z.B. Möller & Marsh, 2013).

Fragestellung und Methode: Wir gingen in unserer Studie der Frage nach, wie sich das sozioemotionale Erleben von Frühstudierenden – konzeptualisiert als Motivation (Selbstkonzept und Werte), Zufriedenheit, Wohlbefinden, Belastungserleben und wahrgenommene Peerakzeptanz – in den Kontexten „Schule“ und „Universität“ unterscheidet und welche Rahmenbedingungen das Erleben beeinflussen. Dazu befragten wir im Sommersemester 2024 insgesamt 347 Frühstudierende (156 weiblich, 183 männlich, 8 divers), die an unterschiedlichen Universitäten in ganz Deutschland eingeschrieben waren. Die Frühstudierenden waren zum Befragungszeitpunkt im Schnitt 16.39 Jahre alt ($SD = 1.35$) und besuchten etwa das 2. Fachsemester ($M = 2.19$; $SD = 1.41$). Zur Erhebung des sozioemotionalen Erlebens griffen wir auf etablierte Messinstrumente zurück (z.B. *Skalen zur Erfassung des schulischen Selbstkonzepts* von Schöne et al., 2012, und *Skala zur Erfassung subjektiver schulischer Werte* von Steinmayr & Spinath, 2010).

Ergebnisse: Die Ergebnisse zeigten zunächst, dass sich das sozioemotionale Erleben der Frühstudierenden über beide Kontexte hinweg weitgehend positiv darstellt (Mittelwerte zwischen $M = 1.97$ [Belastungserleben] und $M = 3.13$ [Wohlbefinden] auf einer 4-stufigen Skala von *trifft gar nicht zu* [1] bis *trifft voll zu* [4]). Allerdings ließen sich klare Unterschiede zwischen den beiden Kontexten feststellen: So war insbesondere die wahrgenommene Peerakzeptanz der Frühstudierenden an der Universität deutlich niedriger ausgeprägt als in der Schule, $t(277) = 14.02$, $p < .001$, $d = 0.84$, während sich für die studienbezogenen Werte, $t(301) = -12.90$, $p < .001$, $d = -0.74$, das Belastungserleben, $t(290) = -3.67$, $p < .001$, $d = -0.22$, und die Zufriedenheit, $t(289) = -22.07$, $p < .001$, $d = -1.30$, ein umgekehrtes Muster ergab. Weiterhin berichteten die Frühstudierenden ein höheres Selbstkonzept an der Schule im Vergleich zur Universität, $t(299) = 17.62$, $p < .001$, $d = 1.02$; das Wohlbefinden der Frühstudierenden war in beiden Kontexten ähnlich ausgeprägt, $t(289) = 1.82$, $p = .071$, $d = 0.11$. Als mögliche Ursachen für die Unterschiede ließen sich in (Regressions-)Analysen in erster Linie (nicht) vorhandene Maßnahmen zur Unterstützung der Frühstudierenden in der Schule bzw. an der Universität ausmachen.

Diskussion und praktische Implikationen: Die Befunde der vorliegenden Studie legen den Schluss nahe, dass das sozioemotionale Erleben der Frühstudierenden in den Kontexten „Schule“ und „Universität“ trotz erheblicher Belastungsfaktoren bzw. möglicher Kontrasteffekte insgesamt recht positiv ausfällt. Gleichzeitig wird jedoch auch deutlich, dass besonders im Hinblick auf die soziale Einbindung der Frühstudierenden an der Universität noch Handlungsbedarf besteht. Diese könnte z.B. durch mehr Unterstützungsangebote wie eine gezielte Betreuung der Schüler(innen) während des Frühstudiums durch Dozent(inn)en oder Tutor(inn)en verbessert werden. Auch an den Schulen könnte sich die Zufriedenheit der Schüler(innen) ggf. durch Maßnahmen wie mehr individuelle Förderung durch die Lehrer(innen) noch steigern lassen (vgl. auch die Befunde von Solzbacher, 2011).

Kleine Fische, große Haie? Wie sich eine Klassenwiederholung auf multikriteriale Ziele von Schule und Unterricht auswirkt.

Paul Fabian, Katja Scharenberg

LMU München, Deutschland

Theoretischer Hintergrund

Eine Klassenwiederholung hat weitreichende Konsequenzen für die betroffenen Lernenden. Die Veränderungen im Leistungsniveau dieser Lernenden während des Wiederholungsjahres sind gut erforscht, wobei überwiegend neutrale Effekte festgestellt wurden (Goos et al., 2021). Die Forschung zum Wohlbefinden der betroffenen Schüler:innen ist uneinheitlich: Einige Studien berichteten positive Effekte auf sozio-emotionale Outcomes (Ehmke et al., 2017; Marsh et al., 2017), während andere Studien keine signifikanten Unterschiede zwischen wiederholenden und versetzten Lernenden fanden (Pipa et al., 2024). Zudem deuteten einige Forschungsarbeiten darauf hin, dass es keine langfristig negativen Auswirkungen auf das Wohlbefinden der Wiederholenden gibt (z.B. Rathmann et al., 2020).

Die wechselseitige Beziehung zwischen schulischer Leistung und subjektivem Wohlbefinden (SWB; Kleinkorres et al., 2020) unterstreicht die Notwendigkeit, auch die Auswirkungen der Klassenwiederholung auf das Wohlbefinden zu untersuchen. Die Theorien von Bronfenbrenner (1979, 1986; Ökologie der menschlichen Entwicklung), Ryan und Deci (2000; Selbstbestimmungstheorie) sowie beispielsweise der Etikettierungsansatz von Becker (1963) und der Ansatz des Big-Fish-Little-Pond-Effekts (BFLPE) nach Marsh (1987) bieten dabei eine nützliche Rahmung zum Verständnis der Auswirkungen der Klassenwiederholung auf das Wohlbefinden und die schulische Leistung.

Fragestellung

Unsere Forschungsfrage lautet demnach: Welchen Einfluss hat eine Klassenwiederholung auf das SWB und weitere sozioemotionale Faktoren in den Jahren nach der Wiederholung?

Methode

Unsere Studie erweitert den bisherigen Forschungsstand, indem sie sich auf verschiedene Dimensionen des SWB konzentriert, einschließlich physischem, kognitivem, emotionalem und sozialem Wohlbefinden. Hierbei wurden weitere sozio-emotionale Outcomes (Selbstkonzept in Deutsch, Mathematik und Lesen, Lesemotivation und Anstrengungsbereitschaft) im schulischen Kontext analysiert. Datengrundlage war das Nationale Bildungspanel (NEPS; Startkohorte 3; NEPS-Network, 2024) mit $n=4.029$ Schüler:innen, von denen $n=94$ die 7. Klasse wiederholten. 50,6% der gesamten Stichprobe waren weiblich und 3,7% hatten einen Migrationshintergrund.

Fehlende Daten wurden mit MICE (van Buuren & Groothuis-Oudshoorn, 2011) imputiert. Messungen von Variablen vor und nach der Klassenwiederholung (t0 respektive t1-t4) wurden mit dem lavaan-Paket (Rosseel, 2012) extrahiert. Mittels Propensity Score Matching (PSM) im Paket MatchThem (Pishgar et al., 2021) wurden vergleichbare Lernende identifiziert, die sich nur im Treatment der Klassenwiederholung unterschieden. Dabei wurden die Messungen der abhängigen Variablen vor der Klassenwiederholung berücksichtigt. Difference-in-Difference-Modelle (DiD) wurden mit fixest (Berge & Iverson, 2020) berechnet, wobei das Geschlecht und der besuchte Schultyp kontrolliert wurden. Dieser Ansatz eliminiert den Einfluss unbeobachteter Störvariablen, wobei PSM die Annahme paralleler Trends zwischen Treatment- und Kontrollgruppen in den DiD-Modellen stärkt. Alle Analysen wurden in R 4.4.0 (R Core Team, 2024) durchgeführt.

Ergebnisse und Diskussion

Die Ergebnisse zeigten keine Unterschiede in der Schulfreude (emotionale Facette), der sozialen Integration in der Klasse (soziale Facette), dem körperlichen Wohlbefinden (physische Facette) und der allgemeinen und schulbezogenen Zufriedenheit (kognitive Facette). Auch in den meisten weiteren sozio-emotionalen Outcomes (Selbstkonzept in Deutsch und Mathematik, Lesemotivation, Anstrengungsbereitschaft) waren keine Effekte zu verzeichnen. Allerdings zeigte sich, dass ein positiver Trend des Leseselbstkonzeptes ($\beta=.29$ (.16), $df=1,974$, $p=.075$) ein Jahr nach dem Wiederholungsjahr (t2) nachweisbar ist.

Zusammenfassend leistet die vorliegende Studie einen bedeutenden Beitrag zum bislang vorliegenden empirischen Kenntnisstand zum Einfluss von Klassenwiederholungen auf SWB und damit verbundene sozio-emotionale Outcomes. Die Effekte der Klassenwiederholung auf verschiedene Facetten des SWB waren insignifikant, mit Ausnahme des Selbstkonzepts im Lesen. Unser gewählter theoretischer Referenzrahmen ließ die Erwartung zu, dass sowohl positive als auch negative Effekte der Klassenwiederholung auf das SWB existieren können. Daher ist es plausibel, dass sich solche zu erwartenden Effekte potenziell gegenseitig aufheben könnten. Beispielsweise könnte ein Anstieg des Selbstkonzepts durch den BFLPE, aber auch ein Rückgang des Selbstkonzepts durch das Erleben von Versagen oder Spott dazu führen, dass sich Schüler:innen insgesamt weder besser noch schlechter fühlen als vor der Klassenwiederholung. Basierend auf unseren Forschungsergebnissen erscheint die Klassenwiederholung als eine ineffektive Intervention – sowohl in Bezug auf die schulischen Leistungen als auch auf das SWB.

PS-81: Weiterbildung: Motive, Gelingensbedingungen, Effekte

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: SO 318

Chair der Sitzung: Luca Farina Hollricher

Warum an Weiterbildung teilnehmen? Dimensionen der Weiterbildungsmotivation erwachsener Lerner:innen an deutschen Volkshochschulen

Luca Farina Hollricher, Julia Gorges

Philipps-Universität Marburg, Deutschland

Studien zum Weiterbildungsverhalten Erwachsener belegen immer wieder den engen Zusammenhang zwischen Bildungsniveau und Weiterbildungsverhalten im Erwachsenenalter (Bilger & Strauß, 2019; 2022). Vor allem bei Erwachsenen mit niedrigem Bildungsniveau ist die Weiterbildungsmotivation der Adressat:innen ein zentraler Prädiktor der Weiterbildungsbeteiligung, während Weiterbildungsbeteiligung bei Personen mit höherem Bildungsniveau (d.h., Akademiker:innen) vor allem auf günstige Gelegenheitsstrukturen zurückgeführt werden kann und Weiterbildungsmotivation an prädiktiver Validität verliert. Trotz dieser wichtigen Rolle von Weiterbildungsmotivation für die Bildungsbeteiligung insbesondere benachteiligter Zielgruppen von Weiterbildung ist bislang wenig über ihre Konzeptualisierung und ihre Struktur bekannt.

Aus der Perspektive der Erwartungs-Wert-Theorie (EWT) werden bildungsbezogene Wahlentscheidungen auf zwei zentrale motivationale Determinanten zurückgeführt: Erfolgserwartung („kann ich das?“) und subjektiver Wert der Handlungsalternativen („will ich das?“) (Eccles, 2005; Gorges, 2015). Laut Eccles (2005) umfasst der subjektive Wert eine utilitaristische, intrinsische und persönliche Komponente, sowie in Kosten. Der subjektive Wert lässt sich laut Vorarbeiten zu dessen Struktur aus dem Weiterbildungskontext zudem untergliedern in inhaltsbezogene, statusbezogene und abschlussbezogene Wertüberzeugungen (Gorges et al., 2017). Laut Gorges et al. (2017) können die inhaltsbezogenen, abschlussbezogenen und statusbezogenen Wertkomponenten und die Kostenkomponente von einem Bilanzwert abgegrenzt werden, der ein Gesamturteil hinsichtlich der Teilnahme an einem Weiterbildungskurs reflektiert (i.S. einer cost-benefit Analyse; Eccles, 1983).

Der vorliegende Beitrag untersucht die konzeptuelle Binnenstruktur der Weiterbildungsmotivation und prüft systematische Zusammenhänge zwischen Wertüberzeugungen, Erfolgserwartung und Bilanzwert.

Der Analyse liegen Daten von Teilnehmer:innen an Volkshochschulkursen im Bereich EDV und Englisch zugrunde (N = 232, 69 % weiblich, Alter: M = 51,7 Jahre, SD = 14,1), die Angaben zum subjektiven Wert des besuchten Kurses, zur Erfolgserwartung und zum Bilanzwert gemacht haben. Der subjektive Wert wurde mit insgesamt neun Items mit inhaltsbezogener (3 Items), abschlussbezogener (3 Items) und statusbezogener (3 Items) Ausrichtung auf einer 6-Stufigen Likert-Skala erfasst, wobei jeweils ein utilitaristisches (z.B. „Die Kursinhalte sind für die Erreichung meiner Ziele sehr nützlich“), ein intrinsisches (z.B. „Es wird mir Spaß machen mich mit den Kursinhalten zu beschäftigen“) und ein persönliches (z.B. „Ich erwarte von mir, dass ich mich mit den Kursinhalten auskenne“) Motiv angesprochen wurde, sowie vier Items zu den Kosten (z.B. „Die Teilnahme an dem Kurs ist sehr teuer“). Die Erfolgserwartung wurde anhand von drei Items (z.B. „Ich werde den Kurs sehr wahrscheinlich erfolgreich besuchen“). Der Bilanzwert wurde mit vier Items erfasst (z.B. „Alles in allem lohnt es sich, den Kurs zu besuchen“).

Konfirmatorische Faktorenanalysen stützen die Annahme der 4-Faktoren-Struktur des subjektiven Wertes mit teilnahmeförderlichen Wertüberzeugungen, die sich nach den Bezugspunkten Kursinhalt, Status als Teilnehmer:in und Abschluss des Kurses untergliedern, sowie einer Kostenkomponente mit monetären und nicht-monetären Anteilen. Der subjektive Wert lässt sich weiterhin von einem Bilanzwert und einem Faktor Erfolgserwartung abgrenzen (CFI = .934; TLI = .912; RMSEA = .068; SRMR = .066). Der Bilanzwert wird primär vom Faktor der Erfolgserwartung vorhergesagt ($\beta = .73, p < .05$). Die ausdifferenzierte Struktur zeigt eine bessere Passung zu den vorliegenden Daten als ein Modell, in dem die Wertkomponenten nicht unterschieden werden (CFI = .649; TLI = .601; RMSEA = .139; SRMR = .176). Die Differenzierung nach Bezugspunkten zeigt außerdem einen besseren Modellfit als eine Differenzierung nach intrinsischem, utilitaristischem und persönlichem Wert (CFI = .908; TLI = .877; RMSEA = .080; SRMR = .097).

Die Ergebnisse der Faktorenanalysen bestätigen die Annahme zur Binnenstruktur des subjektiven Wertes nach bestimmten Aspekten eines Kurses im Anschluss an Gorges et al. (2017) für die vorliegende Stichprobe. Befunde von Gorges et al. (2017) erweiternd, konnte der Faktor der Erfolgserwartung als primärer (und einziger) Prädiktor des Bilanzwertes identifiziert werden. Möglicherweise stellt der Bilanzwert einen der Wertkomponenten übergeordneten Faktor dar. Zukünftige Studien könnten die Annahme prüfen, ob systematische Unterschiede in der Binnenstruktur und der prädiktiven Validität der Faktoren zwischen den Fachbereichen (EDV und Englisch) bestehen.

Professionalisierung des Weiterbildungspersonals: Befunde einer Interventionsstudie mit Lehrenden aus der berufsbezogenen Weiterbildung

Alisha Koch^{1,2}, Susanne Wißhak¹

¹Rheinland-Pfälzische Technische Universität Kaiserslautern Landau (RPTU); ²Universität Konstanz

Theoretischer Hintergrund und Fragestellung

Weiterbildung und lebenslanges Lernen gewinnen u.a. durch die Digitalisierung und den Fachkräftemangel zunehmend an Bedeutung (Autor:innengruppe Bildungsberichterstattung, 2024). In berufsbezogenen Weiterbildungsmaßnahmen sollen Erwachsene neues Wissen, Fähigkeiten und Einstellungen erwerben, um auf aktuelle und zukünftige Herausforderungen am Arbeitsplatz reagieren zu können (Cerasoli et al., 2018). Ziel solcher Maßnahmen ist der Transfer, also die anschließende Anwendung des Gelernten am Arbeitsplatz (Baldwin & Ford, 1988). Aufgrund der Relevanz von Transfer für Lernende, Organisationen und ganze Gesellschaften (Aguinis & Kraiger, 2009) existiert bereits eine Vielzahl an Studien, die sich mit den Faktoren des Gelingens und Misslingens von Transfer beschäftigen (Wißhak, 2022). Die Befunde werden häufig den drei Einflussfaktoren nach Baldwin und Ford (1988) zugeordnet: Teilnehmendenmerkmale, Weiterbildungsdesign und Arbeitsumfeld.

Analog zu Modellen und Befunden aus der Unterrichtsforschung (z.B. Baumert & Kunter, 2006; Kunter et al., 2013; Voss et al., 2014) wird davon ausgegangen, dass das professionelle Wissen der Lehrenden die Bildungserfolge und Weiterbildungsqualität beeinflusst und somit gefördert werden muss. Mit Blick auf das Ziel des Weiterbildungstransfers, sollten die Lehrenden, welche die Weiterbildungsmaßnahmen planen und durchführen, die Befunde der Transferforschung kennen und wissen, wie sie den Transfer fördern können. In diesem Beitrag wird Transferwissen definiert „als das evidenzbasierte, also empirisch fundierte Wissen über die Determinanten und Fördermöglichkeiten von Transfer“ (Koch et al., 2022, S. 92). Dabei wird eine dreidimensionale Struktur für das Konstrukt Transferwissen angenommen, welches sich aus Wissen zu den Teilnehmendenmerkmalen, zum Weiterbildungsdesign und zum Arbeitsumfeld der Lernenden zusammensetzt. Über das

Transferwissen der Lehrenden ist bislang wenig bekannt. Forschungsergebnisse zeigen aber, dass Lehrende die Transferförderung zwar als wichtig erachten (Wißhak & Hochholding, 2018), gleichzeitig aber den Forschungsbefunden wenig Bedeutung beimessen (Hutchins et al., 2010).

Zur Weiterbildung und individuellen Professionalisierung nutzen Lehrende häufig non-formale Weiterbildungsangebote (Autor:innengruppe Bildungsberichterstattung, 2024), welche selbst lern- und transferförderlich gestaltet sein sollten (Schrader, 2010). Im Rahmen der vorliegenden Arbeit wurde ein digitales Weiterbildungsangebot (Intervention) für das berufsbezogene Weiterbildungspersonal entwickelt, welches auf die Stärkung des Transferwissens abzielt und selbst auf Basis der Transferforschung konzipiert wurde. Dabei soll das Weiterbildungsangebot die individuelle Professionalisierung der Lehrenden i.S. des Erwerbs von professionellem Handlungswissen fördern. Ziel war es zu untersuchen, inwieweit das Transferwissen von Lehrenden aus der berufsbezogenen Weiterbildung im Rahmen einer Intervention erhöht werden kann.

Methode

Die quasi-experimentelle Interventionsstudie (Prä-Post-Design) mit Kontrollgruppe wurde mit $N = 81$ Lehrenden aus der berufsbezogenen Weiterbildung durchgeführt ($n_{EG} = 47$; $n_{KG} = 34$). Die Experimentalgruppe absolvierte das 5-wöchige digitale Weiterbildungsangebot (insgesamt 17-20 Stunden) mittels Selbstlernphasen und gemeinsamen Austauschsituationen. Inhaltlich lernten die Lehrenden die Befunde der Transferforschung kennen und setzten die Inhalte direkt an eigenen Weiterbildungskonzepten um. Die Kontrollgruppe erhielt keine Intervention. Die Erfassung des Transferwissens erfolgte mittels eines Instruments mit 31 Items (Koch et al., 2022), welches das Konstrukt Transferwissen entlang der drei Dimensionen (Teilnehmermerkmale, Weiterbildungsdesign und Arbeitsumfeld) abbildet. Anhand von mixed ANOVAs wurden mögliche Interaktionseffekte (Zeit x Gruppe) untersucht.

Ergebnisse und ihre Bedeutung

Die Ergebnisse der mixed ANOVAs zeigen signifikante *Interaktionseffekte* (Zeit x Gruppe) für das *gesamte Konstrukt Transferwissen* ($F_{(1, 79)} = 49.74$; $p < 0.001$; partielles $\eta^2 = .386$) sowie für alle drei Dimensionen *Teilnehmendenmerkmale* ($F_{(1, 79)} = 19.28$; $p < 0.001$; partielles $\eta^2 = .196$), *Weiterbildungsdesign* ($F_{(1, 79)} = 25.81$; $p < 0.001$; partielles $\eta^2 = .246$) und *Arbeitsumfeld* ($F_{(1, 79)} = 30.58$; $p < 0.001$; partielles $\eta^2 = .279$). Damit weist die Experimentalgruppe einen signifikant höheren Zuwachs an Transferwissen über die Zeit auf als die Kontrollgruppe.

Die Ergebnisse sprechen für die Wirksamkeit der Intervention zur Stärkung des Transferwissens. Demnach lässt sich Transferwissen im Rahmen eines Weiterbildungsangebots vermitteln und erlernen. Weitere Studien sind notwendig, um den Einfluss eines ausgeprägten Transferwissens der Lehrenden auf den Transfererfolg der Lernenden zu untersuchen.

Weiterbildungsevaluation aus Sicht der Verantwortlichen

Selina Ebert¹, Caroline Bonnes², Sabine Hochholding¹

¹Universität Konstanz; ²Deutsches Institut für Erwachsenenbildung

Damit Unternehmen in einer sich permanent verändernden Welt erfolgreich bestehen können, müssen ihre Beschäftigten kontinuierlich dazulernen und sich weiterbilden (Schäfer & Ebersbach, 2021). Effektive und hochwertige Weiterbildungs- bzw. Personalentwicklungsmaßnahmen beeinflussen die Wettbewerbs- und Innovationsfähigkeit von Unternehmen und deren Beschäftigten maßgeblich (Aguinis & Kraiger, 2009). Die betrieblichen Ausgaben für Weiterbildung steigen kontinuierlich (Seyda & Placke, 2020). Umso wichtiger ist es, dass Unternehmen die Qualität und die Wirksamkeit ihrer betrieblichen Weiterbildungsangebote kennen, um diese strategisch zu steuern und nachhaltig zu handeln.

Der Fokus der Studie liegt auf der Evaluation betrieblicher Weiterbildung, welche einen wesentlichen Teil des Bildungscontrollings darstellt (Hoffmann, 2018). Ein zentrales und seit langem etabliertes wissenschaftliches Konzept zur Weiterbildungs- oder Trainingsevaluation ist das Vier-Ebenen-Modell von Kirkpatrick (Kirkpatrick & Kirkpatrick, 2016). Dabei stellt sich die Frage, wie Weiterbildungen in Unternehmen konkret evaluiert werden, worüber im deutschsprachigen Raum bisher nur wenig bekannt ist. Das Ziel der explorativen Interviewstudie besteht deshalb darin, vertiefte Einblicke in die Praxis betrieblicher Weiterbildungsevaluation anhand der folgenden Forschungsfragen zu erhalten und zu diskutieren:

- F1: Welche Evaluationsebenen im Modell von Kirkpatrick decken Unternehmen ab?
- F2: Welche Erhebungsmethoden werden für die Weiterbildungsevaluation genutzt?
- F3: Was macht für die Zuständigen die Qualität einer Evaluation aus?
- F4: Welche Auswirkungen haben die Evaluationsergebnisse im Unternehmen?
- F5: Welche Faktoren beeinflussen aus Sicht der Verantwortlichen die betriebliche Weiterbildungsevaluation?

Die Interviews wurden leitfadengestützt mit 18 Evaluationsverantwortlichen aus Unternehmen geführt und mit einer qualitativen Inhaltsanalyse (Mayring, 2022) ausgewertet. Das entwickelte Kategoriensystem besteht aus 105 Kategorien, welche sich in 20 Ober-, 76 Unterkategorien sowie 9 Unter-Unterkategorien gliedern und sich in fünf Themenschwerpunkte unterteilen lässt. Der zur Bestimmung der Interraterreliabilität ermittelte Kappa-Koeffizient (Brennan & Prediger, 1981) beträgt 0,89. Diese hohe Übereinstimmung dokumentiert eine sehr gute Zuverlässigkeit des Kategoriensystems.

Die Ergebnisse zeigen, dass in allen befragten Unternehmen betriebliche Weiterbildungen evaluiert werden. Jedoch erfolgt dieser Prozess nur bei einer Minderheit der Unternehmen modellgestützt und wissenschaftsorientiert. De facto wird aber das Vier-Ebenen-Modell bei vielen Unternehmen umgesetzt. Methodisch setzen alle Unternehmen auf schriftliche Befragungen. Zudem werden mündliche Befragungen oder Beobachtungen genutzt. Die Mehrheit der Unternehmen führt die Weiterbildungsevaluation digitalisiert durch. Die Evaluation basiert auf Selbsteinschätzung der Teilnehmenden oder Fremdeinschätzung z.B. durch die Führungskraft. Kein Unternehmen berichtet, dass experimentelle Untersuchungsdesigns genutzt werden. Bei den Eigenschaften eines guten Evaluationsprozesses werden die Verwertbarkeit und Aussagekräftigkeit der Daten sowie die Individualität des Prozesses von den Befragten am Häufigsten genannt. Verbesserungspotenzial am eigenen Evaluationsprozess sehen die Studienteilnehmenden insbesondere bei der Transferüberprüfung. Die durch die Evaluation gewonnenen Daten werden über eine reine Dokumentation hinaus nur beschränkt weiterverarbeitet und nicht systematisch für die Weiterbildungsgestaltung genutzt. Vorhandene Ressourcen, die Wichtigkeit der Personalentwicklung im Unternehmen sowie die Akzeptanz der Evaluation durch beteiligte Personen werden unter anderem als evaluationsbeeinflussende Faktoren identifiziert.

Insgesamt wird deutlich, dass die Unternehmen der Studie durchaus Evaluationsdaten erheben, aber ihre Evaluationspraxis noch weiter professionalisieren könnten. Ein möglicher Ansatzpunkt wäre, die Evaluationsdaten mit verschiedenen Methoden und durch unterschiedliche Quellen zu erheben, um deren Validität zu erhöhen (Rossi et al., 2019). Weiterhin sollte die Evaluation

auf Basis wissenschaftlicher Konzepte entwickelt werden. Um kausale Effekte von Weiterbildungen festzustellen, sind experimentelle Designs nötig. Digitalisierung und Automatisierung könnten die Qualität der Evaluation erhöhen bzw. den Arbeitsaufwand verringern. Auch die Nutzung von künstlicher Intelligenz z.B. bei der Auswertung freier Kommentarfelder wäre denkbar. Großes, wenig genutztes Potenzial liegt zudem in der systematischen Nutzung und Kommunikation der Evaluationsergebnisse, was zu einem vollständigen Evaluationsprozess gehört. Letztendlich benötigt es für all diese Punkte ausreichende zeitliche, finanzielle und personelle Ressourcen, um eine qualitativ hochwertige und zielgerichtete Weiterbildungs-evaluation durchzuführen.

Digitale Medien im Unterricht: Der Einfluss von Erfahrungen und beruflicher Weiterbildung auf die Selbstwirksamkeitserwartung von Lehrkräften

Claudia Menge, Christiane Annemann, Len Heinatzky, Luisa Grützenmacher, Julia Gerick
TU Braunschweig, Deutschland

Die Digitalisierung hat als zentraler Treiber des gesellschaftlichen Wandels in der gesamten Lebens- und Arbeitswelt zu Transformationsprozessen geführt (Lenk und Pohle 2021). Auch das Schulsystem steht, nicht zuletzt aufgrund veränderter politischer Zielsetzungen in Bezug auf das Lehren und Lernen mit digitalen Medien, unter Anpassungs- und Handlungsdruck (KMK 2017, 2021). Die Integration digitaler Medien in Schule und Unterricht wird von Lehrkräften häufig als zusätzliche Belastung im Berufsalltag wahrgenommen (Maus 2020; Mußmann et al. 2021; Schmechtig et al. 2020). Sie erfordert nicht nur geeignete Rahmenbedingungen, sondern auch spezifische Kompetenzen, wie eine hohe Selbstwirksamkeitserwartung (Schulze-Vorberg et al. 2021).

Selbstwirksamkeitserwartungen sind im Modell der professionellen Kompetenz von Lehrkräften (Baumert und Kunter 2006) Teil der nicht-kognitiven Kompetenzfacetten „motivationale Orientierungen“. Ausgehend davon, dass sie keine globalen, kontextunabhängigen Dispositionen darstellen, sondern durch Erfahrungen und Interventionen beeinflussbar sind (Seifert und Schaper 2018), besteht das Ziel des vorliegenden Beitrags in der Untersuchung von Prädiktoren der Selbstwirksamkeitserwartung von Lehrkräften in Bezug auf den Einsatz digitaler Medien im Unterricht, im Folgenden als digitalisierungsbezogene Selbstwirksamkeitserwartung bezeichnet.

Theoretische Grundlage des Beitrags ist die sozial-kognitive Lerntheorie von Bandura (1977, 1997). Bandura definiert Selbstwirksamkeit als den Glauben einer Person an die eigene Fähigkeit, Problemsituationen erfolgreich zu bewältigen. Digitalisierungsbezogene Selbstwirksamkeit beschreibt demnach, inwieweit Lehrkräfte sich den Einsatz digitaler Medien im Unterricht zutrauen und annehmen, ihre Lehrziele dabei erreichen zu können. Bandura (1977) unterscheidet vier Quellen der Selbstwirksamkeit: Am stärksten wird sie durch eigene Erfahrungen mit der erfolgreichen Bewältigung von Anforderungen sowie auch durch stellvertretende Erfahrungen positiv beeinflusst. Von verbaler Ermutigung geht ebenfalls ein geringer positiver Einfluss aus, während negative Emotionen die Selbstwirksamkeit verringern.

Vor diesem Hintergrund wird im vorliegenden Beitrag der Frage nachgegangen, inwiefern sich die digitalisierungsbezogene Selbstwirksamkeit durch berufliche Erfahrungen im Umgang mit digitalen Medien im Unterricht, berufliche Weiterbildungen sowie die emotionale Erschöpfung von Lehrkräften vorhersagen lässt. Es wird davon ausgegangen, dass Weiterbildungen direkte und stellvertretende Erfahrungen ermöglichen sowie verbale Ermutigung bieten können und folglich mit der digitalisierungsbezogenen Selbstwirksamkeit zusammenhängen (Schulze-Vorberg et al. 2021). Zudem wird erwartet, dass erschöpfte Lehrkräfte aufgrund von Vermeidungsverhalten weniger positive Erfahrungen sammeln können (Bandura 1977) und folglich eine niedrigere Selbstwirksamkeit aufweisen (Lazarides und Warner 2020).

Die Analysen basieren auf Daten der Lehramtsstichprobe der Startkohorte 5 des Nationalen Bildungspanels (NEPS-Netzwerk 2022), die von 2014 bis 2022 ein zusätzliches lehramtspezifisches Befragungsprogramm erhielt (Ortenburger et al. 2023; Schaeper et al. 2023). Für den vorliegenden Beitrag wurden Angaben von $n=433$ Lehrkräften ausgewertet, die im Frühjahr 2020 zu beruflichen Weiterbildungen (1 Item) und emotionaler Erschöpfung (4 Items, $\alpha=0,76$) sowie im Herbst 2020 zu Erfahrungen (1 Item) und Selbstwirksamkeitserwartungen bezogen auf den Einsatz digitaler Medien im Unterricht (5 Items, $\alpha=0,78$) befragt wurden.

Durchschnittlich berichten die Lehrkräfte, die sich durch ihren Studienbeginn im Wintersemester 2010/2011 noch am Anfang ihrer beruflichen Laufbahn befinden, eher hohe digitalisierungsbezogene Selbstwirksamkeitserwartungen ($M=4,32$; $\min=1$, $\max=6$). Ein Strukturgleichungsmodell, gerechnet mit der Statistiksoftware Mplus (Muthén und Muthén 1998-2017), bestätigt signifikante Effekte der Weiterbildungsaktivitäten in Bezug auf den Einsatz digitaler Medien im Unterricht ($\beta=0,09$), der Erfahrungen im Einsatz digitaler Medien ($\beta=0,60$) sowie der emotionalen Erschöpfung ($\beta=-0,12$) auf die digitalisierungsbezogene Selbstwirksamkeitserwartung. Die berücksichtigten Kontrollvariablen – Alter, Geschlecht und Fächergruppe (MINT) – haben erwartungsgemäß (siehe Šabić et al. 2022) keine signifikanten Effekte. Das Modell erklärt 37,6 % der Varianz der Selbstwirksamkeitserwartung. Es verfügt über eine gute Passung auf die Daten (RMSEA=.043, CFI=.959, TLI=.949, SRMR=.039). Mit der anstehenden Veröffentlichung des nächsten Scientific Use Files liegen die beschriebenen Daten auch für das Jahr 2021 vor. Es ist geplant, die vorliegende Untersuchung, vorbehaltlich ausreichender Fallzahl, im Längsschnitt in Form autoregressiver cross-lagged Panel Modelle (Selig und Little 2011) zu wiederholen. Die Ergebnisse der Analysen werden im Beitrag präsentiert und diskutiert.

PS-82: Bildungsverläufe: Bedeutung kontextueller Faktoren

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 142

Chair der Sitzung: Lukas Lazzara

ENTFÄLLT Wie unterscheiden sich die Berufsaspirationen von Schüler*innen mit sonderpädagogischem Förderbedarf in verschiedenen institutionellen Rahmenbedingungen?

Christoph Homuth, Monja Schmitt, Sebastian Bauer

Leibniz-Institut für Bildungsverläufe (LIfBi), Deutschland

Der Anteil von Schüler*innen mit sonderpädagogischen Förderbedarfen (SPF) nimmt seit Jahren zu (vgl. Bpb, 2024). Wenngleich die Quote der inklusiv lernenden Schüler*innen mit SPF anwächst, bleibt die Förderschulquote konstant. Mehr als 50% aller Schüler*innen mit SPF besucht nach wie vor eine Förderschule (ebd.; Klemm et al., 2023). Der Förderschulbesuch kann als Risiko angesehen werden, da mehr als 70% der Förderschüler*innen keinen anerkannten Schulabschluss erreichen (vgl. Deutsches Institut für Menschenrechte, 2023), während Schüler*innen mit vergleichbaren Voraussetzungen, die eine Hauptschule besuchen, deutlich häufiger einen Schulabschluss erlangen (Blanck, 2020). Meist erfolgt nach dem Förderschulbesuch kein direkter Übergang in eine berufliche Ausbildung (Ginnold, 2009; Jochmaring, 2019; Basendowski & Werner, 2010; Gaupp & Geier, 2010) und Förderschulabgänger*innen haben große Schwierigkeiten, stabile Erwerbskarrieren zu erreichen (Menze et al., 2021).

Berufsaspirationen sind ein zentraler Aspekt zur Erklärung erfolgreicher Übergänge in die Berufsausbildung (Schoon & Parsons, 2002; Schoon & Lyons-Amos, 2016; Solga, 2005). Die bisherige Forschung zum Zusammenhang von SPF und Berufsaspirationen ist noch lückenhaft. Sie besteht meist aus qualitativen Studien und fokussiert auf Förderschüler*innen (Van Essen, 2013). Förderschüler*innen passen ihre Aspirationen häufig an niedrigere Übergangsziele bis zum Übergang in das berufliche Bildungssystem an („Cooling-Out“) (Blanck, 2020; Pfahl, 2011; Solga, 2005). Welche Rolle es dabei spielt, dass Schüler*innen mit SPF mit mehr Bildungsressourcen häufiger inklusiv lernen und seltener eine Förderschule besuchen, bleibt dabei unberücksichtigt (Powell, 2006). Internationale Studien liefern allerdings Hinweise darauf, dass die Art der Beschulung eine Rolle spielt und inklusiv lernende Schüler*innen mit SPF höhere Berufsaspirationen haben als exklusiv beschulte (Casey et al., 2013).

Der vorliegende Beitrag untersucht, wie sich Berufsaspirationen von Schüler*innen mit SPF im Verlauf der Sekundarstufe I entwickeln und vergleicht dabei inklusiv an Regelschulen lernende Schüler*innen mit exklusiv an Förderschulen lernenden.

Nach dem Wisconsin-Modell zur Erklärung von schichtspezifischer Herausbildung unterschiedlicher Aspirationsniveaus sind neben der Sozialisation im Elternhaus maßgeblich signifikante Andere (Freunde, Peers, Referenzgruppen) wichtig für die Entstehung von Aspirationen (Sewell et al. 1970). Wir erwarten daher, dass Schüler*innen mit SPF in Regelschulen deutlich höhere Berufsaspirationen aufweisen als Förderschüler*innen, u.a. aufgrund des höheren Aspirationsniveaus an Regelschulen im Vergleich zu Förderschulen. Dies sollte auch unter Kontrolle der sozialen Herkunft der SuS der Fall sein, wenngleich davon ausgegangen werden kann (Boudon, 1974; Bourdieu & Passeron, 1971), dass insbesondere Kinder mit SPF aus sozial bessergestellten Familien eher eine Regel- als eine Förderschule besuchen.

Zur Beantwortung der Forschungsfragen vergleichen wir regressionsanalytisch die sozialstrukturelle Höhe der idealistischen Berufsaspirationen (ISEI des Berufswunsches) am Ende der Sekundarstufe I für die Gruppen Förderschüler*innen als exklusiv lernende Schüler*innen mit SPF, inklusiv lernende Regelschüler*innen mit SPF an Hauptschulen, an Hauptschulzweigen bzw. berufsorientierten Schulformen und an nicht-gymnasialen Schulen in Bundesländern mit zweigliedrigen Sekundarschulsystemen, sowie Regelschüler*innen ohne SPF an denselben Schulformen.

Unserer Datenbasis sind Startkohorten (SC) 3 und 4 des NEPS (NEPS Netzwerk, 2024a, 2024b). Mit der SC3 betrachten wir die Höhe der Berufsaspirationen in Klassenstufe 8 und 9. Die SC4-Stichprobe dient zur ergänzenden Analyse der Berufsaspirationen in K9.

Unsere Ergebnisse zeigen, dass inklusiv lernende Schüler*innen mit SPF in Klassenstufe 8 unter Kontrolle der sozialen Herkunft ähnlich hohe Berufsaspirationen aufweisen wie die Vergleichsgruppe der Regelschüler*innen ohne SPF, während Förderschüler*innen statistisch signifikant deutlich niedrigere Berufsaspirationen zeigen. Der Vergleich von Klassenstufe 8 zu 9 zeigt, dass Schüler*innen mit SPF generell niedrigere Berufsaspirationen aufweisen je näher der erste Schulabschluss rückt, während dies nicht für Regelschüler*innen ohne SPF gilt.

Die Ergebnisse passen zu den Befunden eines Cooling-Out-Effekts und der „schulischen Behinderungsthese“ (Blanck, 2020), wonach der Besuch einer Förderschule den weiteren Bildungsverlauf negativ beeinflusst. Angesichts der höheren Wahrscheinlichkeit von sozial benachteiligten Schüler*innen mit SPF eine Förderschule, statt eine Regelschule zu besuchen, deuten die Ergebnisse darauf hin, dass eine höhere Inklusionsquote auch zur Reduktion von sozial ungleichen Bildungschancen beitragen könnte.

Veränderungen überfachlicher Kompetenzen bei Schüler*innen mit und ohne sonderpädagogischen Unterstützungsbedarf in Abhängigkeit von inklusiven Rahmenbedingungen

Lena Nusser¹, Cornelia Gresch²

¹LIfBi, Deutschland; ²Humboldt-Universität zu Berlin

Theoretischer Hintergrund

Neben akademischen Fachkompetenzen wird bei Jugendlichen mit Blick auf Übergänge und die spätere berufliche Ausbildung immer wieder die Bedeutung überfachlicher Kompetenzen (auch „berufsrelevante Kompetenzen“ oder „Schlüsselqualifikationen“) betont (vgl. z. B. Eckert & Schmitt 2007; Hofmann-Lun 2011; Niehaus 2009; Stein & Stummbaum 2010). Hierbei handelt es sich nach Pohlmann und Heckt (2011) um „Fähigkeiten, Fertigkeiten und Einstellungen, die zur Teilhabe am gesellschaftlichen Leben und zur Aneignung von fachlichem Wissen notwendig sind.“ (Pohlmann & Heckt, 2011, S. 36). Dem Lernen und dem Lernkontext in der Schule wird dabei eine wesentliche Rolle bei der Entwicklung überfachlicher Kompetenzen zugesprochen (vgl. Fend 2006, S. 116; Horstkemper & Tillmann 2018).

Dabei gibt es Hinweise, dass Schüler*innen mit sonderpädagogischen Unterstützungsbedarfen (SPU) im Mittel über geringere überfachliche Kompetenzen verfügen als Schüler*innen ohne SPU und dass dies auch mit dem Schulkontext zusammenhängen kann (Schuck et al., 2018). Gresch und Nusser (2023) untersuchten die durch Lehrkräfte eingeschätzten überfachlichen Kompetenzen „Selbstkompetenzen“, „Sozial-kommunikative Kompetenzen“ und „Lernmethodische Kompetenzen“ bei

Schüler*innen mit und ohne SPU. Die Autorinnen zeigten, dass schulische Lernbedingungen, die gezielt auf inklusive Bildung ausgerichtet sind, für alle Schüler*innen mit höheren überfachlichen Kompetenzen einhergehen. Hierbei waren inklusive Rahmenbedingungen für die überfachlichen Kompetenzen von Schüler*innen ohne SPU in stärkerem Maße bedeutsam als für diejenigen mit SPU. Ob sich die überfachlichen Kompetenzen zwischen den Gruppen darüber hinaus über die Zeit differentiell verändern, also beispielsweise bei Schüler*innen mit SPU weniger stark zunehmen, und welche Rolle hierbei inklusive Rahmenbedingungen spielen, wurde bisher nicht untersucht.

Fragestellung

Der Beitrag untersucht, wie Lehrkräfte die Veränderung der überfachlichen Kompetenzen von der sechsten zur siebten Jahrgangsstufe bei Schüler*innen mit und ohne SPU einschätzen und welche Rolle inklusive, schulische Rahmenbedingungen und die Ausgangsniveaus der jeweiligen überfachlichen Kompetenzen spielen.

Methode

Datengrundlage bilden die ersten beiden Erhebungszeitpunkt der ersten Kohorte des Projekts „Inklusion in der Sekundarstufe I in Deutschland – INSIDE“ mit $N = 1.883$ teilnehmenden Schüler*innen an 192 Schulen. In der zweiten Erhebung wurden die Lehrkräfte gebeten, für die einzelnen Schüler*innen jeweils für drei Items auf einer 5-stufigen Likert-Skala einzuschätzen, wie sich die drei Dimensionen „Selbstkompetenzen“, „Sozial-kommunikative Kompetenzen“ und „Lernmethodische Kompetenzen“ der überfachlichen Kompetenzen im Laufe des Schuljahres verändert haben. In der ersten Welle wurde zudem das Ausgangsniveau der überfachlichen Kompetenzen über die Lehrkräfte erfasst (Schuck et al., 2018). Informationen über den SPU basieren auf Angaben der Schulen. Darüber hinaus wurden über die Schulleitungen inklusive schulische Rahmenbedingungen erhoben, basierend auf der Qualitätsskala zur inklusiven Schulentwicklung (QU!S, Heimlich et al. 2018). Hintergrundmerkmale der Schüler*innen (Geschlecht, sozioökonomischer Status, kognitive Grundfähigkeiten) wurden für die Analysen kontrolliert. Die Auswertung erfolgte anhand von Regressionsanalysen mit dem Package lavaan (Rosseel, 2012) in der Software R (R Core Team, 2021). Fehlende Werte wurden mittels FIML-Ansatzes berücksichtigt.

Ergebnisse

Die Faktorstruktur der drei theoretisch angenommenen Dimensionen der überfachlichen Kompetenzen und ihren Veränderungen bildet sich in den Daten ab ($CFI = .982$; $TLI = .971$; $RMSEA = .073$; $SRMS = .026$). Regressionsanalysen zeigen einen Zusammenhang zwischen der eingeschätzten Veränderung der überfachlichen Kompetenzen mit individuellen Merkmalen der einzelnen Schüler*innen. Schüler*innen mit SPU wird eine geringere Entwicklung überfachlicher Kompetenzen durch die Lehrkräfte zugeschrieben als ihren Peers ohne SPU. Als besonders prädiktiv stellt sich die frühere Einschätzung der Lehrkräfte der überfachlichen Kompetenzen dar. Der positive Zusammenhang deutet darauf hin, dass für Schüler*innen mit ohnehin hohen Ausgangsniveaus auch ein stärkerer Zuwachs an überfachlichen Kompetenzen beobachtet wird. Die inklusiven Rahmenbedingungen an der Schule zeigen keine signifikanten Zusammenhänge mit der Veränderung überfachlicher Kompetenzen der Schüler*innen.

Die Befunde deuten auf einen Bedarf der Förderung überfachlicher Kompetenzen von Schüler*innen mit SPU über bisherige inklusive Angebote in den Schulen hinaus hin, um sowohl die Aneignung fachlicher Kompetenzen als auch nachschulische Übergänge in eine Ausbildung und den Arbeitsmarkt zu unterstützen.

Zusammenhänge struktureller und personaler Merkmale auf die Qualität der Umsetzung naturwissenschaftlicher Bildungsangebote in der Kita

Lukas Lazzara¹, Laura Venitz², Anika Bürgermeister³, Ada Haen⁴, Ilonca Hardy⁵, Mirjam Steffensky⁴, Sina-Kristin Koschick⁵, Henrik Saalbach³, Miriam Leuchter¹

¹RPTU Kaiserslautern-Landau; ²Hochschule für Technik und Wirtschaft Saar, Saarbrücken; ³Universität Leipzig; ⁴Universität Hamburg; ⁵Goethe-Universität Frankfurt am Main

Theorie

Die Förderung von Kindern in verschiedenen Entwicklungsbereichen wird als ein zentrales Merkmal der Prozessqualität in Kindertagesstätten angesehen [2]. In diesem Zusammenhang sind pädagogische Fachkräfte (PFK) verantwortlich für die Bereitstellung hochwertiger Bildungsangebote in der Interaktion mit Kindern, zum Beispiel in der frühen naturwissenschaftlichen Bildung. Die Prozessqualität in Kindertagesstätten wird auch durch strukturelle Merkmale (z. B. Fachkraft-Kind-Schlüssel), durch personale Merkmale wie Orientierungsqualität (pädagogische Überzeugungen) sowie professioneller Kompetenz (z. B. fachdidaktisches Wissen) (mit-)bestimmt. [1]. Im Kontext der naturwissenschaftlichen Bildung ist das Zusammenspiel von fachübergreifenden und fachspezifischen Struktur-, Orientierungs-, Kompetenz- und Prozessqualitätsmerkmalen bislang wenig untersucht. Die vorliegende Studie analysiert deshalb die Interdependenz zwischen diesen Aspekten bei der Umsetzung eines naturwissenschaftlichen Bildungsangebots.

Fragestellung

Welchen Zusammenhang gibt es zwischen Merkmalen der Strukturqualität, Orientierungsqualität, Kompetenzmerkmalen und der Prozessqualität im Rahmen eines naturwissenschaftlichen Bildungsangebots in der Kita?

Methode

Im Rahmen des Projekts Profink wurden Daten von 76 PFK erhoben (Durchschnittsalter 40.14 Jahre, $SD = 11.23$). Die Umsetzung eines naturwissenschaftlichen Bildungsangebots zum Thema *Materialien und ihre Eigenschaften* wurde videografiert. Zusätzlich wurden die Fachkräfte schriftlich befragt.

Indikatoren der Strukturqualität

Der *Fachkraft-Kind-Schlüssel* wurde durch die Anzahl der Kinder und der Anzahl der Fachkräfte in der Gruppe errechnet.

Die *Rolle von Naturwissenschaften* in der Kita wurde mit 11 Items auf einer vierstufigen Skala (1= selten, 4= regelmäßig) gemessen ([6]; $\alpha = .82$).

Indikatoren der Orientierungsqualität

Fachübergreifende pädagogische Überzeugungen wurden mit 12 Items von [5] auf einer vierstufigen Skala erfasst (Ko-Konstruktion, $\alpha = .74$; Selbstbildung, $\alpha = .68$; Instruktion, $\alpha = .76$).

Kompetenzmerkmale

Die *Selbstwirksamkeit pädagogischer Fachkräfte in der naturwissenschaftlichen Bildung* wurde mit sechs Items auf einer vierstufigen Skala (s.o.) erfasst ([6]; $\alpha = .87$).

Das *naturwissenschaftliche fachdidaktische Wissen* wurde über 14 Items erfasst ([6]; $\alpha = .75$).

Indikatoren der Prozessqualität

Die Prozessqualität dieses Bildungsangebots wurde mittels des *Classroom Assessment Scoring System* (CLASS, [3]) in den drei Domains *emotionale Unterstützung*, *Gruppenführung* und *Anregungsqualität* auf einer siebenstufigen Skala (1/2: niedrige Qualität, 3-5: mittlere Qualität, 6/7: hohe Qualität; ICC = .89) bewertet.

Zur Beantwortung der Forschungsfrage wurden Deskriptiva, Korrelationen und eine CFA berechnet.

Ergebnisse und Diskussion

Deskriptiv zeigte sich, dass *Anregungsqualität* gering ausgeprägt war, die *emotionale Unterstützung* und *Gruppenführung* eher hoch.

Aus den signifikanten Korrelationen zwischen den Prozessqualitätsmerkmalen und strukturellen, Orientierungs- sowie Kompetenzmerkmalen wurde ein Modell abgeleitet (Abbildung 1) und mittels CFA geprüft. Der Devianztest sowie die Fitindizes deuteten auf eine gute Passung des Modells hin ($\chi^2(14) = 13.00$; $p = .526$; CFI = 1.00; TLI = 1.03; RMSEA = 0.00; SRMR = 0.07). Der *Fachkraft-Kind-Schlüssel* klärte 16% der Varianz in der Domain *emotionale Unterstützung* auf, der Prädiktor *fachübergreifende instruktive Überzeugungen* hatte in diesem Zusammenhang keinen Einfluss. Die *Berufserfahrung* und die *Rolle von Naturwissenschaften* klärten 22.2% der Varianz in der Domain *Gruppenführung* auf, wobei *Berufserfahrung* hier keinen Einfluss hatte. Für die Domain *Anregungsqualität* waren *fachübergreifende ko-konstruktive Überzeugungen* (13% Varianzaufklärung) prädiktiv, jedoch war der Zusammenhang negativ.

Im Kontext der naturwissenschaftlichen Bildung in der Kita erwiesen sich aus fachspezifischer Perspektive das Strukturmerkmal *Rolle von Naturwissenschaften*; aus fachübergreifender Perspektive das Strukturmerkmal *Fachkraft-Kind-Schlüssel* und das Orientierungsmerkmal *ko-konstruktive Überzeugungen* als Prädiktoren von Prozessqualitätsmerkmalen. Die Studie liefert erste Hinweise, wie naturwissenschaftsspezifische und fachunabhängige Struktur- und Orientierungsmerkmale die Prozessqualität im Kontext der frühen naturwissenschaftlichen Bildung beeinflussen. Besonders auffällig ist der starke Zusammenhang des fachspezifischen Strukturmerkmals *Rolle von Naturwissenschaften* mit dem Prozessqualitätsmerkmal *Gruppenführung*. Dieser könnte auf die relative breite Konzeption der Domain *Gruppenführung* in CLASS zurückgeführt werden, die als Antizipation von Problemen, dem Geben kontingenter Hinweisreize sowie klares Erwartungsmanagement seitens der Fachkraft verstanden wird [3]. Der negative Zusammenhang fachübergreifender *ko-konstruktiver Überzeugungen* mit der Domain *Anregungsqualität* könnte auf die hohen fachlichen Anforderungen hinweisen, ko-konstruktive Bildungsangebote in einem naturwissenschaftlichen Kontext umzusetzen. Dennoch scheint es PFK bei kleineren Gruppen eher zu gelingen, *emotionale Unterstützung* zu leisten.

Instruktionssensitivität als Indiz für eine differenzielle Kompetenzentwicklung in unterschiedlichen Schulformen: DIF-Analysen am Beispiel der ökonomischen Bildung in der Jahrgangsstufe 8 in Nordrhein-Westfalen

Fabio Fortunati, Esther Winther

Universität Duisburg-Essen, Deutschland

In der ökonomischen Bildung in Deutschland treffen Schüler*innen auf höchst heterogene Lerngelegenheiten, was das Stundenkontingent, die Fächerstruktur sowie die Lehrkräfteausbildung anbelangt (Fortunati & Winther, 2024b; Macha, 2019; Siegfried & Ackermann, 2020; Siegfried & Hagen, 2020). Der Assessmentforschung gelingt es zwar, interindividuelle Unterschiede zuverlässig zu erheben, jedoch bleiben instruktionssensitive Merkmale meistens unberücksichtigt. Es bedarf daher methodischer Zugänge, um empirische Befunde zu Unterrichtseffekten präziser detektieren zu können. Das deutsche Bildungssystem selektiert die Schülerschaft früh in verschiedene Schulformen mit unterschiedlichen normativen Bildungszielen und Unterrichtstraditionen. Dies wirft unmittelbar die Passung zwischen Curriculum und Assessment, aber auch zwischen erfahrener Instruktion und Assessment auf. So können Inhaltsbereiche eines Assessments sich instruktionssensitiv auf spezifische Gruppen auswirken, da diesen mehr Lernmöglichkeiten oder kognitiv aktivierender Unterricht zu Verfügung standen (Naumann, Musow et al., 2019; Naumann, Rieser et al., 2019).

Der Fokus dieses Beitrages liegt darin, über das Konstrukt der Instruktionssensitivität schulformspezifische Unterschiede eines Testinstruments exemplarisch für die ökonomische Kompetenz von Schülern der Sekundarstufe I im Bundesland Nordrhein-Westfalen zu untersuchen. Die Schulform kann hier als Proxy für unterschiedliche unterrichtsbezogene sowie andere institutionelle Einflussfaktoren betrachtet werden. Gleichwohl ist anzumerken, dass für eine tiefer gehende Analyse von Unterrichtseffekten der Einbezug relevanter unterrichtsbezogener Merkmale unerlässlich ist (Naumann, Rieser et al., 2019). Auf Grundlage einer repräsentativen Stichprobe soll mithilfe eines Testinstruments zur Messung ökonomischer Kompetenz geprüft werden, ob dieses (1) instruktionssensitiv, bezogen auf die Schulform ist, und (2), welche inhalts- und kognitionsbezogenen Merkmale diese Items aufweisen und (3), ob sich diese Unterschiede durch unterschiedliche curriculare Vorgaben erklären lassen.

Generell kann unter dem Konzept der Instruktionssensitivität verstanden werden, ob Lernerfolge unterrichtsbezogen attribuiert werden können (Polikoff, 2010). In diesem Beitrag werden der Judgmental- und der empirische Ansatz miteinander verbunden, in dem ex ante durch eine Analyse der Curricula der ökonomischen Domäne (Fortunati & Winther, 2024a) Hypothesen über mögliche instruktionssensitive Inhalte in den verschiedenen Schulformen formuliert und im Anschluss empirisch über IRT-basiertes Differential-Item-Functioning (DIF) geprüft werden (Deutscher & Winther, 2018). Kernelement DIF-basierter Studien ist es, dass bei Schüler*innen mit gleicher Personenfähigkeit aufgrund eines gruppenbezogenen Merkmals unterschiedliche Lösungswahrscheinlichkeiten für das Assessment als Ganzes und auf Itemebene auftreten (Masters, 1988; Naumann et al., 2016).

Die Datenerhebung der Hauptstudie wurde im Frühjahr 2022 computergestützt als Querschnittsdesign durchgeführt. Es handelt sich um eine stratifizierte zweistufige Cluster-Stichprobe ($N=2.852$), in der durch explizite Stratifizierung Schulen mit ähnlichen Merkmalen zu sinnvollen Einheiten (Strata) zusammengefasst wurden. Die ökonomische Kompetenz wurde mithilfe eines authentisch konstruierten und computergestützten Assessments erhoben (Winther & Abs, 2024). Das psychometrische Messmodell basiert auf der Item-Response-Theory. Für die Analyse der Daten wird ein polytomes 1PL-IRT-Modell, Multidimensional-Random-Coefficients-Multinomial-Logit-Modell (MCMMLM) gewählt (Adams et al., 1997). Zur Berücksichtigung der Mehrebenenstruktur (Individual- und Klassen-/Schulformebene) wurde eine hierarchische logistische Regression zur Detektion von DIF-Unterschieden durchgeführt (Li et al., 2017). Auf Individualebene werden zur präziseren Attribuierung des

DIF-Effekts weitere Variablen wie das Geschlecht, der sozioökonomische Hintergrund usw. in die Analyse mit einbezogen. Hinsichtlich der Beurteilung des DIF-Werts wird auf das Klassifizierungsschema von Paek und Wilson (2011) Bezug genommen.

Zum jetzigen Zeitpunkt liegen die Ergebnisse der Hauptstudie noch nicht vollständig vor. Die Befunde der Pilotstudie lassen sowohl auf Test- als auch auf Itemebene signifikante Schulformunterschiede erkennen. Die instruktionssensitiven Items weisen schulformspezifische Unterschiede hinsichtlich inhaltlicher und kognitiver Charakteristika auf, die sich jedoch nur teilweise auf unterschiedliche curriculare Standards zurückführen lassen.

Die Ergebnisse können zum einen dabei helfen, instruktionssensitive Items für Assessments in der ökonomischen Bildung zu konstruieren, und zum anderen für die Fort- und Weiterentwicklung intendierter wie implementierter Curricula im Sinne instruktionaler Kohärenz genutzt werden. Die gefundenen inhalts- und kognitionsspezifischen Ergebnisse ermöglichen ein präziseres Verständnis der erwartbaren Fähigkeiten der Schüler und können zu einer besseren Binnendifferenzierung im Unterricht beitragen.

PS-83: Einstellungen: Bedeutung für Bildungsentscheidungen, -verläufe und -outcomes

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 145

Chair der Sitzung: Dennis Hauk

Politiklehrlpersonen als Agenten des digitalen Wandels? Eine empirische Analyse des digitalisierungsbezogenen Berufswahlmotivs im Lehramtsstudium

Dennis Hauk¹, Johannes Schuster¹, Michael May², Gudrun Heinrich³

¹Universität Leipzig; ²Friedrich-Schiller-Universität Jena; ³Universität Rostock

Theoretischer Hintergrund

Politiklehrkräfte spielen weltweit eine entscheidende Rolle bei der digitalen Transformation, um Kindern und Jugendlichen politische Teilhabe, Selbstständigkeit und Mündigkeit in einer digitalen Gesellschaft zu ermöglichen (Stoddard, 2014; Waldis, 2020). Studien zeigen, dass Politiklehrkräfte davon überzeugt sind, einen bedeutsamen gesellschaftlichen Beitrag zur digitalen Bürgerschaftsbildung zu leisten (Vajen et al., 2020). Die Forschungslage offenbart jedoch auch, dass unter angehenden Lehrkräften das Berufswahlmotiv, einen Beitrag zur digitalen Transformation der Gesellschaft leisten zu wollen, nur mäßig ausgeprägt ist (Martínez-Moreno und Petko, 2023). Unklar ist zudem, inwieweit sich dieser allgemeine Befund mit der fachspezifischen Berufswahlmotivation von Politiklehrlpersonen deckt. Das Beheben dieses Forschungsdesiderats ist für die politische Bildung von Bedeutung, da hier die Digitalisierung nicht nur eine fachübergreifende Querschnittsaufgabe darstellt, sondern zum Kernbereich der Domäne gehört (Autoren, 2021). Das Wissen über digitalisierungsbezogene Berufswahlmotive besitzt zudem eine Relevanz für die Ausbildung von Politiklehrlpersonen, um adaptive und zielgerichtete Entscheidungen hinsichtlich der curricularen Anforderungen und professionsbezogenen Ausbildungsstandards treffen zu können. Die vorliegende Studie hilft zu verstehen, inwieweit Politiklehrlpersonen am Beginn ihrer Berufslaufbahn den digitalisierungsbezogenen Gestaltungsauftrag als Motiv ihrer Berufswahlentscheidung berücksichtigen. Gleichzeitig werden Prädiktoren der digitalisierungsbezogenen Berufswahlmotivation untersucht und hinsichtlich ihrer Gültigkeit für die Politiklehrlerbildung überprüft.

Forschungsfragen:

1. Wie stark ist bei Politiklehramtsstudierenden das Berufswahlmotiv, einen Beitrag zur digitalen Transformation leisten zu wollen, im Vergleich zu anderen Berufswahlmotiven ausgeprägt?
2. Inwieweit beeinflussen die selbsteingeschätzte digitale Kompetenz, positive digitale Vorerfahrung, Alter und Geschlecht das Berufswahlmotiv einen Beitrag zur digitalen Transformation leisten zu wollen?

Methode

An der Onlinebefragung nahmen $N = 224$ Politiklehramtsstudierende an drei Hochschulstandorten (Jena, Leipzig & Rostock) im Wintersemester 2023/24 teil. Die Studierenden sind $M = 21.4$ ($SD = 2.9$) Jahre alt, überwiegend weiblich ($n_w=117$, $n_m=103$, $n_{div}=4$) und studieren mehrheitlich für das gymnasiale Lehramt ($n_{gym}=154$, $n_{rsos}=54$). Für die Befragung wurden personenbezogene Merkmale (Alter, Geschlecht und Fachrichtung), die digitalisierungsbezogene Vorerfahrung und Kompetenz sowie Studienwahlmotive (Beitrag zur digitalen Transformation, Nützlichkeit, Pädagogisches Interesse, Fähigkeitsüberzeugung, Soziale Einflüsse, Schwierigkeitsgrad, Fachinteresse; 6-Punkt-Likert Skalen, Reliabilität $\alpha = .76-.90$; Martínez-Moreno und Petko, 2023) erhoben. Die Auswertung der Daten erfolgt mittels einfaktorieller ANOVA (Forschungsfrage 1) und multipler linearer Regression (Forschungsfrage 2) in SPSS 29.

Ergebnisse

Mit der Forschungsfrage 1 wurde die Ausprägung des digitalisierungsbezogenen Berufswahlmotivs untersucht. Die Befunde zeigen, dass dieses Berufswahlmotiv mittelstark ausgeprägt ist ($M = 4.64$, $SD = .93$) und im Vergleich ($F(1,178) = 328.493$, $p < .001$) signifikant bedeutsamer eingeschätzt wird als die Nützlichkeit ($M = 4.10$, $SD = .91$), Soziale Einflüsse ($M = 3.66$, $SD = 1.11$) oder der erwartete geringe Schwierigkeitsgrad des Studiums ($M = 2.11$, $SD = .98$) (alle Unterschiede $p < .001$). Signifikant höher eingeschätzt werden das Pädagogische Interesse ($M = 5.17$, $SD = .66$), die Fähigkeitsüberzeugung ($M = 4.85$, $SD = .59$) und das Fachinteresse ($M = 5.11$, $SD = .62$) (alle Unterschiede $p < .001$).

Forschungsfrage 2 fragt nach Prädiktoren für das digitalisierungsbezogene Berufswahlmotiv. Das berechnete Modell ist signifikant ($F(4,199) = 19.855$, $p < .001$) und hat mit einem $R^2 = .29$ (korrigiertes $R^2 = .28$) eine hohe Varianzaufklärung. Die Ergebnisse zeigen, dass die selbsteingeschätzte digitale Kompetenz ($B = .464$, $SE = .073$, $t = 6.339$, $p < .001$) und das Geschlecht ($B = -.218$, $SE = .105$, $t = -2.083$, $p = .039$) einen statistisch signifikanten Einfluss auf das digitalisierungsbezogene Berufswahlmotiv haben.

Die Studie zeigt, dass Politiklehramtsstudierende ein hohes Interesse an der pädagogischen Arbeit mit Kindern und Jugendlichen haben, das Motiv mit ihrer beruflichen Tätigkeit einen Beitrag zur digitalen Transformation leisten zu wollen, steht dabei jedoch weniger im Fokus. Bedeutsame Prädiktoren, die das digitalisierungsbezogene Berufswahlmotiv beeinflussen, sind die digitale Kompetenz und das Geschlecht. Im Unterschied zu allgemeinen Befunden zum Lehramt besitzt die positive digitale Vorerfahrung keine prädiktive Wirkung (Martínez-Moreno und Petko, 2023). Im Vortrag werden die Befunde hinsichtlich bestehender Limitationen und der Bedeutung für die fachinhaltliche Ausbildungscurricula und fachdidaktischen Professionalisierungsangebote diskutiert.

„...irgendwann ist auch mal Schluss mit dem Ganzen!“ Einstellungen und Praktiken von Lehrer*innen und Eltern zu Mehrsprachigkeit im Bildungskontext

Martin Schastak¹, Melanie David-Erb², Sabrina Geyer³

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland; ²Goethe-Universität Frankfurt am Main;

³Pädagogische Hochschule Heidelberg

Mehrsprachigkeit wird in Deutschland allgemein wertgeschätzt (Schastak, in review), allerdings spielt die Mehrsprachigkeit von Schüler*innen in institutionalisierten Bildungsprozessen kaum eine Rolle (Gogolin et al., 2019). Mehrsprachigkeit weist aus der Perspektive des Translanguaging (García & Wei, 2014) Potentiale für das Lernen auf, für deren Entfaltung es aber positive Einstellungen zu Mehrsprachigkeit sowie lernförderliche Praktiken mehrsprachigen Handelns im Handlungsfeld Schule benötigt. In bisherigen Studien berichten Lehrkräfte von einer Orientierung an einer monolingualen Norm sowie einer ungenügenden Ausbildung im Umgang mit Mehrsprachigkeit (Bien 2012; Wischmeier 2012), Bedenken hinsichtlich Kontrollverlust und negativem Klassenklima (Bredthauer & Engfer, 2018) sowie einer Defizitperspektive auf bestimmte Herkunftssprachen und einem

Konkurrenzverhältnis zwischen Sprachen um Lernzeit (Wlossek & Rost-Roth, 2016). Allerdings weisen Lehrkräfte mit Erfahrung in der Beschulung von Seiteneinsteiger*innen in Intensivkursen positivere Einstellungen zu Mehrsprachigkeit als Ressource für Bildungsprozesse auf und implementieren mehrsprachiges Handeln (vgl. ebd.). Die bisherige Befundlage erlaubt keine kontrastierenden Aussagen über Einstellungen zu Mehrsprachigkeit in verschiedenen Handlungsfeldern (z. B. Schule vs. Familie). Zudem könnten sich Lehrkräfte und Eltern aufgrund Erfahrungen mit und Praktiken von Mehrsprachigkeit in den jeweiligen Handlungsfeldern hinsichtlich ihrer Einstellungen zu Mehrsprachigkeit unterscheiden.

In unserer Studie untersuchen wir Einstellungen und Praktiken von Lehrkräften und Eltern zu Mehrsprachigkeit in verschiedenen Handlungsfeldern. Es wird davon ausgegangen, dass die Einstellungen zu Mehrsprachigkeit im Handlungsfeld Schule negativer ausfallen als zu Mehrsprachigkeit als individuelle und gesellschaftliche Ressource und zu mehrsprachigem Aufwachsen in der Familie (H1). Zudem wird angenommen, dass Grundschullehrer*innen und Eltern von Grundschüler*innen sich hinsichtlich ihrer Einstellungen zu Mehrsprachigkeit in verschiedenen Handlungsfeldern unterscheiden (H2). Weiterhin werden subjektive Sinnkonstruktionen und Praktiken mehrsprachigen Handelns im Handlungsfeld schulischer Bildung von Lehrer*innen und Eltern eruiert (F1) und als mögliche Erklärungen für die oben angeführten, vermuteten Unterschiede betrachtet.

Die Studie ist als Mixed-Methods-Vertiefungsdesign (Kuckartz, 2014) angelegt. N=65 Lehrer*innen und n=97 Eltern füllten einen Online-Fragebogen zu Einstellungen zu Mehrsprachigkeit in verschiedenen Handlungsfeldern aus, bevor auf Basis ihrer Einstellungen ein heterogenes Sample von n=7 Lehrer*innen und n=7 Eltern von Grundschüler*innen für leitfadengestützte, problemzentrierte Interviews gezogen wurde. Die in H1 und H2 postulierten within- und between-subjects Unterschiede werden mit einer ANOVA mit Messwiederholung respektive einer einfaktoriellen MANOVA sowie einer Mixed ANOVA geprüft, während F1 mithilfe der Qualitativen Inhaltsanalyse nach Kuckartz & Rädiker (2022) untersucht wird.

Eine explorative Faktorenanalyse etablierter, adaptierter und selbsterstellter Items zu Einstellungen von Mehrsprachigkeit in verschiedenen Handlungsfeldern (KMO=.81) bestätigt mithilfe einer Parallelanalyse die theoretisch angenommene Drei-Faktoren-Struktur mit Eigenwerten der drei Faktoren > 1 und eine Hauptachsenanalyse mit drei Faktoren und Promax-Rotation klärt 40% der Varianz auf (RMSEA = 0.077; 90% Konfidenzintervall = 0.065-0.09). Damit lassen sich drei Faktoren identifizieren: Einstellungen zu „Mehrsprachigkeit als individuelle und gesellschaftliche Ressource“ (MS-IGR), „Mehrsprachigkeit für schulische Bildungsprozesse“ (MS-SB) und „Mehrsprachiges Aufwachsen in der Familie“ (MS-MAF).

Eine ANOVA mit Messwiederholung weist einen Haupteffekt der Handlungsfelder auf (Huynh-Feldt $F(1,940, 300.70) = 197.40, p < .001, \eta^2 = .560$). Bonferroni-korrigierte paarweise Vergleiche zeigen, dass Einstellungen zu MS-SB signifikant niedriger als zu MS-IGR sowie zu MS-MAF ausfallen und die Einstellungen zu MS-IGR signifikant positiver als zu MS-MAF sind.

Eine einfaktorielle MANOVA identifiziert einen Haupteffekt für Unterschiede in den Einstellungen zu Mehrsprachigkeit in verschiedenen Handlungsfeldern der beiden Gruppen ($F(3, 152) = 7.63, p < .001, \eta^2 = .131$, Wilk's $\Lambda = .869$). Post-hoc einfaktorielle ANOVAs weisen signifikante Unterschiede zwischen der Gruppe der Lehrer*innen und der Gruppe der Eltern für MS-IGR und MS-MAF, aber nicht für MS-SB ($F(1, 154) = 0.62, p = .434, \eta^2 = .004$) auf.

Eine Mixed ANOVA bestätigt einen Interaktionseffekt der Einstellungen und Gruppen ($F(2, 308) = 7.48, p < .001, \eta^2 = .046$).

Zur Tagung werden neben den hier vorgestellten Ergebnissen auch erste Ergebnisse zu F1 berichtet und diese zusammen mit den quantitativen Ergebnissen interpretiert.

Passungswahrnehmung versus berufliches (Fehl-)Konzept? Determinanten der Berufswahlmotivation für den Lehrberuf

Lara Dahlke¹, Fani Lauermann¹, Sylvia Rahn^{2,3}, Birgit Ziegler⁴, Nico Dietrich⁴

¹Universität Bonn; ²Bergische Universität Wuppertal; ³Deutsches Institut für Erwachsenenbildung; ⁴Technische Universität Darmstadt

Theoretischer Hintergrund/ Fragestellung

Vor dem Hintergrund eines sich weiter verschärfenden Lehrkräftemangels (insbesondere im Bereich beruflicher Schulen; OECD, 2023), der durch hohe Abbruchzahlen noch verstärkt wird (Gundlach et al., 2024), gewinnt die Frage der Rekrutierung geeigneter Personen für den Lehrberuf zunehmend an Bedeutung. Dementsprechend untersuchen viele Studien welche Motivationen eine besondere Rolle bei der Berufswahlentscheidung für oder gegen ein Lehramt spielen.

Diese oft auf Eccles' Erwartungs-Wert-Theorie (Eccles et al., 1983) beruhende Forschung zeigt, dass (angehende) Lehrpersonen, die sich für den Lehrberuf aus eher intrinsischen (z.B. Freude am Unterrichten) statt extrinsischen Gründen (z.B. Jobsicherheit) entscheiden, später eine höhere Berufswahlzufriedenheit berichten und langfristig im Lehrberuf verbleiben möchten (Watt & Richardson, 2007; 2012). Zudem zeigen erste Studien, dass intrinsische Berufswahlmotive das Interesse von Schüler:innen am Lehrberuf vorhersagen können (z.B. Renger et al., 2022).

Bisherige Erkenntnisse beruhen allerdings vorwiegend auf retrospektiven Untersuchungen von Berufswahlmotiven (z.B. als Lehrkraft unterrichten wollen) oder Berufsvorstellungen (z.B. Lehrberuf als sicherer Job) von (angehenden) Lehrkräften. Einschlägige nicht retrospektive Untersuchungen erfassen zwar Selbstvorstellungen, jedoch keine lehramtsspezifischen Berufsvorstellungen. Insgesamt ist daher wenig darüber bekannt, inwieweit (mangelndes) Interesse am Lehrberuf mit (möglicherweise fehlerhaften) Vorstellungen über den Beruf vor der erfolgten Berufswahl zusammenhängen. Insbesondere bleibt hierbei unberücksichtigt, inwieweit diese Vorstellungen zu den angegebenen Berufswahlmotiven passen.

In diesem Beitrag werden daher aus einer Person-Environment-Fit-Perspektive nicht nur die Berufswahlmotivation im Rahmen der Selbstvorstellungen, sondern auch die jeweils zugehörigen Vorstellungen über den Lehrberuf als Berufskonzept noch vor der Berufswahlentscheidung untersucht und gegenübergestellt.

Methode

Im Rahmen regulärer Unterrichtsstunden bearbeiteten 1452 Berufs- und Gymnasialschüler:innen einen Online-Fragebogen. Wie wichtig ihnen bestimmte Aspekte in ihrer Berufswahl sind (Selbstvorstellungen) und für wie typisch sie ebendiese für den Lehrberuf halten (Berufskonzept), wurde mittels parallel formulierter Items erfasst. Berufswahlmotivationen wurden basierend auf der Erwartungs-Wert-Theorie mithilfe validierter Skalen erfasst, einschließlich intrinsischer und extrinsischer Motive sowie lehramtsspezifischer Berufsvorstellungen.

Zunächst wurden interindividuelle Unterschiede im Interesse am Lehrberuf in Abhängigkeit von Schüler:innenmerkmalen untersucht (z.B. Geschlecht). Unter Kontrolle dieser Unterschiede wurden anschließend 8 Modelle getestet, um die inkrementellen prädiktiven Effekte von zugehörigen Selbst- und Berufseinschätzungen mit parallel formulierten Items zu erfassen (z.B. Jobsicherheit: Eine sichere Berufslaufbahn ist mir wichtig / ist typisch für den Lehrberuf). Schließlich wurden im

Gesamtmodell die prädiktiven Effekte aller 16 Konstrukte getestet. Messvarianzannahmen wurden für die jeweiligen Selbst- und Berufseinschätzungen überprüft.

Ergebnisse

In Einklang mit bestehenden Forschungsergebnissen unterstreichen die separaten Analysen auch für Schüler:innen die primäre Bedeutung intrinsischer Berufswahlmotive gegenüber extrinsischen für das Interesse am Lehrberuf. Während das unterschiedliche Interesse zu 17% auf die Merkmale der Schüler:innen zurückzuführen ist, erklären die intrinsischen (21%-35%) und extrinsischen Motive (18%-21%) der Lernenden bis zu 29% zusätzliche Varianz.

Die Analysen deuten zudem darauf hin, dass intrinsische Berufswahlmotive stärker auf eine mangelnde Passung zum Lehrberuf hinweisen als auf Fehlvorstellungen. Während die Vorstellungen der Schüler:innen über ihr Interesse am Unterrichten, fachliches Interesse, andere beurteilen und ein niedrigeres Belastungsempfinden in der Arbeit mit Menschen prädiktiv für ihr Interesse am Lehrberuf sind, zeigen ihre Vorstellungen diesbezüglich keine oder deutlich geringere prädiktive Effekte.

Unterschiedliche Berufsvorstellungen im Entfaltungsbedürfnis und extrinsischen Nutzwert weisen hingegen auf mögliche Fehlvorstellungen hin; Eine höhere Einschätzung über Entfaltungsmöglichkeiten, Freizeit und Jobsicherheit im Lehrberuf hängen positiv mit dem Interesse zusammen.

Insgesamt deuten die Ergebnisse darauf hin, dass die Förderung der Bildung eines Berufskonzepts insbesondere in Bezug auf extrinsische Berufswahlmotive (bspw. Arbeitszeit) sowie die Herausstellung oder Schaffung von Entfaltungsmöglichkeiten im Lehrberuf mögliche Ansatzpunkte für nachhaltige Rekrutierungsmaßnahmen darstellen.

Die Messinvarianzanalysen zeigen, dass sich die Konstrukte über fachliches Interesse, Beurteilen, Freizeit sowie die Arbeit mit Menschen bezogen auf Selbstvorstellungen von denen über das Berufskonzept unterscheiden. Hierauf basierende Rückschlüsse auf eine berufliche Passung sind somit nicht ohne Einschränkung möglich, sofern sie sich nicht auf den Lehrberuf beziehen.

Vertrauen von Schüler:innen in Teamteacher:innen als Puffer für die Abnahme schulischen Engagements: Längsschnitt-Ergebnisse aus Schulen in herausfordernden Lagen

Annalena Schröder¹, Dorothea Gläßer¹, Christopher Holl¹, Laura McCullagh², Kristin Mitte¹, Nicole Harth¹

¹Ernst-Abbe-Hochschule Jena; ²Universität Münster

Nach der Selbstbestimmungstheorie von Deci und Ryan (1985) wird Motivation durch verschiedene Faktoren gefördert, wodurch wiederum Engagement gesteigert werden kann. Im schulischen Kontext steht hierbei insbesondere die Beziehung zwischen Lehrkraft und Schüler:in mit Motivation und schulischem Engagement in Verbindung. Schulisches Engagement umfasst verhaltensbezogene, affektive und kognitive Aspekte und ist ein Schlüsselfaktor, um den Schulbesuch aufrecht zu erhalten und Schulabbrüche zu verringern (Pascarella & Terenzini, 2009; Reeve, 2012). In unserer Studie wiesen alle 17 teilnehmenden, weiterführenden Schulen im Bundesland Thüringen zum ersten Messzeitpunkt eine Abbruchquote von über 20 Prozent auf. Um dieser Herausforderung zu begegnen, wurde auf Grundlage des Europäischen Sozialfonds ein sozialpädagogisches Teamteaching in spezifischen Klassen implementiert und deren Einflussfaktoren auf den Schulerfolg untersucht. Bestehende Forschungsergebnisse zeigen, dass das Vertrauen von Schüler:innen in Lehrkräfte positives Verhalten und Leistungen von Schüler:innen vorhersagt (Adams, 2014). Die Analyse der qualitativen Daten aus unserem Projekt zeigte außerdem, wie wichtig eine vertrauensvolle Beziehung zwischen den Schüler:innen und ihren Teamteacher:innen ist. In dieser Studie wurde angestrebt, diesen Zusammenhang in 23 ethnisch und sozioökonomisch diversen Klassen zu replizieren. 474 Schülerinnen und Schüler nahmen über drei Messzeitpunkte (2020-2022) an einem Selbstbeurteilungsfragebogen teil. Die Längsschnittdaten wurden mithilfe von Mehrebenenmodellen analysiert. Die Ergebnisse zeigen einen insgesamt negativen Effekt der Zeit auf das schulische Engagement. Dieser Effekt ist jedoch bei Schüler:innen, die mehr Vertrauen in ihre Teamteacher:innen hatten, abgeschwächt. Vertrauensvolle Beziehungen scheinen demnach eine stabilisierende Wirkung zu haben. Unsere Ergebnisse unterstreichen die Bedeutung des Vertrauens in pädagogisches Personal für den künftigen schulischen Erfolg von Schüler:innen mit unterschiedlichen Bildungsvoraussetzungen an Schulen mit hohen Abbruchsquoten. Es scheint wichtig zu sein, sozioemotionale Aspekte bei Kindern und Jugendlichen in ihrer schulischen Entwicklung aufzugreifen und zu fördern, um ihnen gleiche Bildungschancen bei gesellschaftlich ungleichen Ausgangspositionen zu ermöglichen.

S-100: Affektives Lernen mit XR-Technologien in der Umweltbildung

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 162

Chair der Sitzung: Jule Krüger

Chair der Sitzung: Josef Buchner

Affektives Lernen mit XR-Technologien in der Umweltbildung

Chair(s): **Jule M. Krüger** (Universität Potsdam, Deutschland)

Diskutant*in(nen): **Josef Buchner** (PH St. Gallen, Schweiz)

In der Forschung zur Umweltbildung und zur Bildung für nachhaltige Entwicklung (BNE) spielen neue technologische Entwicklungen eine immer größere Rolle. Die Möglichkeiten, die Virtual Reality (VR), Augmented Reality (AR) und ähnliche immersive Technologien (zusammengefasst kurz „XR“) bieten, werden hierbei aktuell schon vielseitig erprobt.

XR eröffnet die Möglichkeit, in andere Welten abzutauchen. Hier kann man in immersiver VR (iVR) simulierte, vom Klimawandel beeinflusste Wälder der Zukunft besuchen (Huang et al., 2021), in 360° Videos an abgelegenen Orten der Welt wie tropischen Gletschern die Auswirkungen des Klimawandels ganz nah erleben (Adjanin & Brooks, 2023), oder unter der Meeresoberfläche die Verschmutzung der Ozeane in iVR betrachten (Straßmann et al., 2020). Aber auch inter- und pro-aktive Erfahrungen, wie ein iVR Spiel zum Aufräumen des Ozeans (Lisle et al., 2022), eine iVR-Simulation wichtiger Handlungsschritte während einer Hochwasserkatastrophe (Araujo-Junior et al., 2024), oder ein Mobile Game in dem Entscheidungen der Lernenden die Landschaft verändern (Hernández-Pérez et al., 2023), werden durch XR-Technologien ermöglicht.

Um effektive Umweltbildung mit XR-Technologien zu gestalten, müssen die relevanten Faktoren genauer untersucht werden. Neben den kognitiven Faktoren, wie Wissen über die Auswirkungen des Klimawandels oder über umweltfreundliches Verhalten, spielen auch Emotionen und Einstellungen eine große Rolle. Erlebnisse in iVR-Technologie können emotionale Reaktionen auslösen, bei denen Personen virtuelle Erfahrungen als reale Ereignisse verarbeiten (Susindar et al., 2019). In einer Meta-Analyse konnte ein Einfluss von VR-Umgebungen mit natürlichen Inhalten, aber nicht von VR-Umgebungen mit erbauten Inhalten auf Emotionen gefunden werden (Bolouki, 2024). Im Bereich der Umweltbildung beschäftigen sich jedoch erst wenige Studien mit affektiven Faktoren in XR. Zum Beispiel wurde gezeigt, dass Unterwasser-iVR emotionale Reaktionen hervorruft (Fauville et al., 2024), Präsenzgefühl einen Einfluss auf subjektiv erfahrenes affektives Lernen in Desktop-VR (Bodzin et al., 2024) und auf Einstellungs- und Verhaltensänderungen in iVR (Buchner, 2024) hat, und AR-gestützte Ausflüge zum lokalen Teich affektive Variablen erhöhen können (Kamarainen et al., 2013). Darüber hinaus konnte Brambilla et al. (2022) in einem systematischen Review zeigen, dass virtuelle Naturerfahrungen emotionale Prozesse wie Naturverbundenheit fördern können (Brambilla et al., 2022).

In dem hier beschriebenen Symposium soll das Thema der affektiven Reaktionen in der Umweltbildung, die durch XR-Technologien hervorgerufen werden können, näher beleuchtet werden. Dazu widmen sich die Beiträge allgemein der Fragestellung „Welche Rolle spielen affektive Faktoren und Einstellungen bei der Umweltbildung mit XR-Technologien?“. Hierzu wird im ersten Beitrag ein Systematic Literature Review beschrieben, in dem der aktuelle Stand der Forschung in der Umweltbildung beschrieben wird. Es zeigt sich, dass nur wenige Studien die Einstellung der Lernenden einbeziehen und fast keine der begutachteten Studien emotionsbezogene bzw. affektive Variablen betrachteten. Im zweiten Beitrag wird die Perspektive von Lehrkräften betrachtet, die eine XR-basierte Anwendung zum Thema Artenvielfalt im Unterricht eingesetzt haben. Die Lehrkräfte wurden hierzu befragt, ob sie bei den Schüler:innen Änderungen in den Einstellungen oder dem Verhalten durch die Nutzung der Anwendung festgestellt haben. Im dritten Beitrag wird eine Studie beschrieben, die sich auf die emotionsbezogenen bzw. affektiven Auswirkungen der Nutzung einer immersiven VR-Umgebung bezieht. Insbesondere wird hierbei der Einfluss von Mitgefühl („Compassion“) auf Naturverbundenheit untersucht.

Die beschriebenen Beiträge sollen einen Überblick über erste Effekte beim affektiven Lernen mit XR-Technologien in der Umweltbildung geben. In der Diskussion des Symposiums werden die Ergebnisse der Studien zusammengeführt und weitere Forschungsstränge diskutiert. Hierbei wird, zusätzlich zum Beitrag des Diskutanten, genügend Zeit für Anregungen und Ideen aus dem Publikum eingeplant.

Beiträge des Symposiums

Der Einfluss von XR-Technologien in der Umweltbildung – Ein Systematic Literature Review

Jule M. Krüger¹, Kieren "Ren" Rende Mendoza², Noah Glaser³, Mohan Yang⁴, Kimberly Moeller³

¹Universität Potsdam, Deutschland, ²University of Nebraska Omaha, United States, ³University of Missouri, United States, ⁴Old Dominion University, United States

Umweltbildung ist ein wesentlicher Bestandteil der Entwicklung einer informierten und engagierten Gesellschaft, die nachhaltige Entscheidungen treffen und Umweltprobleme bewältigen kann (Core Writing Team et al., 2014). Umweltwissen reicht hierbei nicht aus, um umweltfreundliches Verhalten zu inspirieren (Otto & Pensini, 2017; Roczen et al., 2014). Umwelteinstellungen, also umweltbezogene Werte, Überzeugungen und Wahrnehmungen, sind wichtige Vermittler, die die Wahrscheinlichkeit erhöhen, dass Wissen in Verhalten umgesetzt wird (Milfont & Duckitt, 2010). Eine zentrale Herausforderung in der Umweltbildung besteht in der Identifikation effektiver Strategien zum Wissensaufbau und der Förderung umweltfreundlicher Einstellungen für langfristige nachhaltiges Handeln. In neueren Entwicklungen werden Virtual Reality (VR) und anderen XR-Technologien zur Darstellung alternativer Erfahrungen eingesetzt, um Wissen und Naturverbundenheit zu erhöhen. Studien zeigten das Potential von VR Umweltwissen, umweltfreundliche Einstellungen, Empathie, Naturverbundenheit oder die Bereitschaft zu ökologisch nachhaltigem Verhalten zu erhöhen (Ahn et al., 2016; Markowitz et al., 2018; Martingano et al., 2021; Meidenbauer et al., 2020; Zelenski et al., 2015). Es ist jedoch weitere Forschung nötig, um die genauen Faktoren zu bestimmen, durch die XR-Technologien Umweltbildung unterstützen können.

Das vorliegende Systematic Literature Review beschäftigt sich mit der Nutzung von XR-Technologien in der Umweltbildung, die im Feld implementiert und evaluiert wurden. Forschungsfragen behandeln hierbei den Einsatz von XR-Systemen in der Umweltbildung und ihre einzigartigen technologischen Merkmale und Möglichkeiten, und die eingesetzten Forschungsmethoden und -ergebnisse in Bezug auf Einstellung, Verhalten und Wissen. Für die Struktur des Reviews wurden die PRISMA-Richtlinien genutzt (Moher et al., 2010). In einer ersten Suche in sechs Datenbanken im Juni 2023, wurden Suchbegriffe zu verschiedenen XR-Technologien (z.B. Virtual Reality - VR, Augmented Reality - AR, Mixed Reality - MR) und zu Umweltbildung (z.B. Klimawandel, Biodiversität, Nachhaltigkeit, Umwelt, Ökosystem) kombiniert. Es wurden nach Entfernung von Duplikaten 834 englischsprachige Artikel, die seit Januar 2015 (kommerziellen Verfügbarkeit von VR-Brillen) veröffentlicht wurden, gefunden. Eingeschlossen wurden Studien, die eine Implementierung eines XR-Systems, eine Forschungskomponente mit gemessenen

Ergebnissen, einen Schwerpunkt auf Umweltbildung und ein pädagogisches Setting beinhalten. Nach Begutachtung der Titel und Abstracts wurden 53 Artikel zum Volltext-Screening und schlussendlich 29 Artikel für Einschluss in das Review ausgewählt.

Die Ergebnisse zeigen, dass die Themen Naturschutz (n=8), Klimawandel (n=7), Nachhaltigkeit (n=6), Ökosysteme (n=5) sowie Umweltverschmutzung und Abfall (n=3) behandelt werden. Die XR-Erfahrungen enthielten VR (n=14) und AR (n=12) Anwendungen, die als Simulationen (n=11), virtuelle Exkursionen (n=9), Spiele (n=4), ortsbezogene AR-Anwendungen (n=3), eigenständige Lerneinheiten (n=3) und digitale Bilderbücher (n=2), umgesetzt wurden. Die meisten Studien (n=14) wurden in Primar- und Sekundarschulen durchgeführt, n=9 an Universitäten und n=4 in informellen Lernumgebungen. Die Methoden von n=33 Studien beschrieben Zwischensubjektvergleiche (n=8), Innersubjektvergleiche (n=4), gemischte Designs (n=11), Posttest (n=8) und Querschnittskorrelationen (n=2). Hierbei wurden in den Zwischensubjektvergleichen n=6 Medienvergleiche und n=10 Innermedienvergleiche vorgenommen.

In n=22 Studien wurden wissensbezogene Variablen erfasst. Die meisten Innersubjektvergleiche zeigten nach Anwendungsnutzung einen Anstieg von Wissen und Umweltbewusstsein. Medienvergleiche zeigten gemischte Ergebnisse zum Lernen mit verschiedenen Medien. Die Nutzung von gezielt eingesetzten pädagogischen Strategien führte zu erhöhtem Lernerfolg. Gemessene Variablen zur Einstellung, Motivation und Affekt enthielten Einstellung zur Umwelt (n=8), Selbstwirksamkeit (n=6), Lernmotivation (n=5) und affektive Reaktionen wie Angst und Empathie (n=3). Ergebnisse zu kognitiven Einstellungen zeigten positive Effekte in Bezug auf Selbstwirksamkeit, Überzeugungen, Motivation, und umweltbezogene Ansichten und Werte. Ergebnisse zu affektiven Einstellungen zeigten erhöhte Umweltempathie, Sorge um Umweltgesundheit und positive Einstellungsänderung in Bezug auf Ökosysteme. Verhaltensintentionen und tatsächliches Verhalten wurden in n=11 der Studien gemessen und auch hier wurden positive Effekte festgestellt.

Die beschriebene Studienlandschaft zeigt, dass sich noch wenige der Studien mit Einstellungen und affektiven Reaktionen auseinandersetzen. Diese sind jedoch relevant für die Verhaltensänderung, da Wissen alleine nicht ausreicht um umweltfreundliches Verhalten hervorzurufen (Otto & Pensini, 2017; Roczen et al., 2014). Weitere Studien sollten auf die Erfassung von affektiven Variablen und Einstellungs- und Verhaltensänderungen zielen.

Lernen mit der greenpeace App – Ergebnisse einer qualitativen Untersuchung von Lehrkräften zum Einsatz von XR im Schulunterricht

Miriam Mulders, Lara Kirner

Universität Duisburg-Essen, Deutschland

Im Beitrag wird die von greenpeace entwickelte XR-Lernapplikation Der Artenvielfalt auf der Spur mit Fokus auf der Umgebung des Amazonas-Regenwaldes untersucht. Die Aussagen von fünf Lehrkräften, die Schüler:innen bei der Nutzung begleiteten, wurden qualitativ ausgewertet. Die Lehrkräfte verweisen darauf, dass die Anwendung durch ihre Charakteristika (z.B. Interaktivität) die Initialmotivation der Schüler:innen erhöhe, sich mit Umweltthemen auseinanderzusetzen. Für eine Einstellungs- und Verhaltensänderung müsse jedoch eine längerfristige Auseinandersetzung stattfinden.

Theoretischer Hintergrund: Mit der Einführung der 17 Ziele für nachhaltige Entwicklung (United Nations, 2015) rückte auch in Deutschland das Augenmerk der Schulbildung auf die Verbesserung der Lebensbedingungen auf unserem Planeten (Weselek et al., 2022). Parallel dazu hat sich das Lernen vieler Schüler:innen seit der Pandemie zunehmend in den digitalen Raum verlagert. Eine aktuelle Herausforderung besteht darin, digitale und physische Lernumgebungen zu kombinieren und gleichzeitig die Ziele für nachhaltige Entwicklung zu vermitteln. Technologien wie Virtual und Augmented Reality (VR/AR) werden dabei als vielversprechend angesehen (Gadelha, 2018).

Fragestellung: Dieser Beitrag untersucht eine XR-Lernanwendung zum Thema Artenvielfalt. Die Anwendung zielt darauf ab, bei Schüler:innen das Interesse an einer nachhaltigen Entwicklung unserer Umwelt zu wecken und Veränderungen in Einstellung und Verhalten anzustoßen. Die webbasierte Applikation, die über mobile Endgeräte erkundet wird, ermöglicht es Schüler:innen, verschiedene Lebensräume (z.B. Amazonas-Regenwald) virtuell zu explorieren. Die virtuelle Umgebung ist durch auditive (z.B. Regenwaldgeräusche) und visuelle Inhalte (z.B. einheimische Frau) gekennzeichnet.

Die Applikation besteht sowohl aus virtuellen (VR) als auch aus Elementen, bei denen die Realität durch virtuelle Artefakte ergänzt wird (AR). Wir folgen daher dem Vorschlag von Rauschnabel et al. (2022), den Überbegriff XR zu nutzen, wobei das X als Platzhalter verstanden wird.

XR-Technologien versprechen neue Lernwelten zu schaffen, die in der Realität nicht oder kaum erfahrbar sind. Sie ermöglichen authentische und selbstgesteuerte Lernerfahrungen (Makransky & Petersen, 2021; Slater et al., 2022). Um die Wirksamkeit von XR für den Schulunterricht zu beurteilen, wurde 2023 ein Feldexperiment mit 274 Schüler:innen der Mittelstufe durchgeführt. In einer Doppelstunde testete die Hälfte der Schüler:innen die XR-Applikation (blinded for review). Dieser Artikel legt den Fokus auf Aussagen von Lehrkräften, welche die Schüler:innen begleiteten. Dabei wurde diesen Forschungsfragen nachgegangen:

Hat die Anwendung aus Sicht der Lehrkräfte Änderungen auf Einstellungs- und Verhaltensebene bei den Schüler:innen angestoßen?

Welche Prozesse beobachteten die Lehrkräfte bei den Schüler:innen?

Methode: Nach der Hauptuntersuchung wurden die beteiligten 12 Lehrkräfte kontaktiert. Drei Lehrkräfte nahmen an einem Online-Interview teil und zwei Lehrkräfte füllten einen Online-Fragebogen aus. Beide Methoden beinhalteten denselben Fragenkatalog (<https://shorturl.at/vtdjl>). Die Daten wurden zusammengeführt und anhand einer fokussierten Inhaltsanalyse (Kuckartz & Rädiker, 2020) ausgewertet.

Ergebnisse:

1. Forschungsfrage: Als wichtigen Faktor für Änderungen auf Einstellungs- und Verhaltensebene benennen die Lehrkräfte, dass Betroffenheit bei den Schüler:innen entsteht (N=4). Gemäß den Lehrkräften gelingt es der App, das Interesse zu steigern und Impulse zur Reflektion des eigenen Denkens und Handelns zu geben (N=5). Jedoch betonen sie, dass eine Einstellungs- und Verhaltensänderung nur über eine längerfristige Auseinandersetzung erfolgen kann: „Dazu hat die Zeit nicht gereicht. Ich glaube, dass dafür mehr Beschäftigung und Betroffenheit hergestellt werden muss.“ Sie fordern konkrete Bezüge zur Lebenswelt der Schüler:innen sowie praktische Umsetzungstipps für eine nachhaltigere Lebensweise (N=4).

2. Forschungsfrage: Die Lehrkräfte beobachteten qualitativ andere Lernprozesse bei den Schüler:innen (im Vergleich zu traditionellen Unterrichtssettings). Das Erleben eines Gefühls von „Vor-Ort-Sein ändert die Betroffenheit“ (N=2) und auch das interaktive Erleben einer fremden Lebenswirklichkeit motiviere die Schüler:innen (N=2).

Die Ergebnisse der Befragungen der Schüler:innen (blinded for review) und der Lehrkräfte verweisen auf das Potential von XR-Technologien, Schüler:innen zur Auseinandersetzung mit einer nachhaltigen Entwicklung unserer Umwelt zu motivieren und

letztlich positive Effekte auf affektiver und behavioraler Ebene zu erzeugen. Größere und repräsentativere Befragungen von Lehrkräften sind jedoch nötig.

Becoming a Tree - Die Verkörperung der Natur in iVR beeinflusst Mitgefühl und Naturverbundenheit

Pia Spangenberg¹, Jule M. Krüger¹, Sonja Maria Geiger², Georg Felix Reuth¹, Lena Baumann¹, Steve Nebel¹

¹Universität Potsdam, Deutschland, ²Murdoch University, Australia

Sowohl in der Forschung zur Bildung für Nachhaltige Entwicklung (BNE) als auch in der Umweltpsychologie, wird ein stärkerer Fokus auf affektive Prozesse gefordert, die im Zusammenhang mit transformativen Lernprozessen und umweltfreundlichem Verhalten diskutiert werden (Brosch & Steg, 2021; Grund et al., 2024). Während als relevante affektive Prozesse im Kontext transformativer Lernprozesse sogenannte edge-emotions (Grund et al.) diskutiert werden, gilt in der Umweltpsychologie die Naturverbundenheit als relevanter Einflussfaktor auf umweltfreundliches Verhalten (Barragan-Jason et al., 2022; Mayer & Frantz, 2004). Gleichzeitig wird immersive Virtual-Reality (iVR)-Technologie als vielversprechender Forschungsansatz für die Erreichung dieses Ziels diskutiert (Spangenberg et al., 2024), da iVR zugeschrieben wird, besonders geeignet zu sein, um Emotionen zu induzieren (Susindar et al., 2019). Allerdings steckt die Forschung zu den zugrundeliegenden emotionalen Prozessen wie Angst, Wut, Traurigkeit, Freude oder auch Mitgefühl, welche die starke Wirkung von immersiver VR auf die Naturverbundenheit (Brambilla et al., 2022; Spangenberg et al., 2022, 2024) erklären können, noch am Anfang (Spangenberg et al., 2024).

Vor diesem Hintergrund wurden in der Studie in einem randomisierten experimentellen Forschungsdesign drei Forschungsfragen untersucht: a) Inwiefern die Verkörperung eines Baumes in iVR die Naturverbundenheit fördert b) wie affektive Prozesse die Wirkung der Verkörperung auf die Naturverbundenheit beeinflussen und c) inwiefern die Art des erlebten Szenarios (positiv vs. negativ) die Wirkung der Verkörperung auf affektive Prozesse sowie die Naturverbundenheit moderiert.

In der präregistrierten Studie testeten 85 Teilnehmende die iVR-Anwendung „Tree“ (Liu, 2017). Nach dem Zufallsprinzip erlebten die Proband*innen entweder eine vollständige Version mit dem negativen Ende eines Waldbrandes ($n = 42$) oder einer verkürzten Version ohne das negative Ende ($n = 43$). Die Teilnehmenden waren Studierende im Alter von 19 bis 49 Jahren ($M = 25,30$, $SD = 5,49$; männlich = 36, weiblich = 48). Die abhängigen Variablen waren Embodiment, erhoben als virtual body ownership mit einer angepassten Skala von Ahn et al. (2016), Emotionen (Wut, Angst, Traurigkeit, Freude), gemessen mit der Kurzversion der Positive and Negative Affect Scale (PANAS; Mackinnon et al., 1999), und Mitgefühl, erhoben mit einer angepassten Skala von Pfattheicher et al. (2016). Naturverbundenheit wurde mit der Skala „Inclusion of Nature in Self“ (Schultz, 2002) in der grafischen Version von Klepsies et al. (2021) vor und nach der iVR-Erfahrung erfasst. Zusätzlich wurde als objektives Maß die körperliche Anspannung mit elektrodermaler Hautleitfähigkeit gemessen.

Hinsichtlich unserer ersten Forschungsfrage fanden wir einen starken Anstieg der Naturverbundenheit von pre ($M = 51,80$, $SD = 18,70$) zu post ($M = 73,60$, $SD = 24,00$) der iVR, mit einem signifikanten Unterschied und einem großen Effekt in einem t-Test für gepaarte Stichproben, $t(84) = 10,6$, $p < .001$, $d = 1,15$. In einer linearen Regression zum Einfluss der Verkörperung auf die Veränderung der Naturverbundenheit fanden wir jedoch keinen signifikanten Effekt, $R^2 = .003$, $F(1, 83) = 0,259$, $p = .612$. Für unsere zweite Forschungsfrage führten wir fünf Mediationsanalysen mit den vier Emotionen (d.h. Wut, Angst, Traurigkeit, Freude) und Mitgefühl als Mediatoren, virtual body ownership als Prädiktor und Veränderung der Naturverbundenheit als Ergebnis durch. Dieser Vermittlungseffekt wurde nur für Mitgefühl gefunden, mit einem signifikanten indirekten Effekt, $\beta = .12$, $z = 2,21$, $p = .027$. Der direkte Effekt war deskriptiv negativ, aber nicht signifikant, $\beta = -.06$, $z = -0,54$, $p = .589$. Der vollständige Mediationseffekt zeigt, dass die Verkörperung das Mitgefühl erhöht, was wiederum die Naturverbundenheit erhöht. Betrachtet man die dritte Forschungsfrage und damit die Moderation der Effekte durch Framing, so finden wir Interaktionseffekte bezüglich Wut ($F(1, 81) = 6,58$, $p = .012$, $\eta^2p = 0,08$), Angst ($F(1, 81) = 4,81$, $p = .031$, $\eta^2p = 0,06$) und Mitgefühl ($F(1, 81) = 4,44$, $p = .038$, $\eta^2p = 0,05$), die alle zeigen, dass Embodiment einen positiven Effekt beim negativen, aber keinen Effekt beim positiven Szenario hatte. Für Traurigkeit, Freude und Naturverbundenheit wurden keine Interaktionseffekte gefunden.

S-101: Understanding and Fostering Giftedness: Approaches to Talent Development from Multiple Disciplines

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: SN 163

Chair der Sitzung: Ulrich Trautwein

Chair der Sitzung: Katrin Kunz

English Track

Understanding and Fostering Giftedness: Approaches to Talent Development from Multiple Disciplines

Chair(s): **Ulrich Trautwein** (Hector Research Institute of Education Sciences and Psychology, University of Tübingen), **Katrin Kunz** (Hector Research Institute of Education Sciences and Psychology, University of Tübingen)

Talent development is a central educational goal in Germany and is emphasized by policy frameworks established by the Kultusministerkonferenz (KMK) in 2009. In a collective position of all federal states, the KMK stated that all students, including talented children, should receive special support in education, where the diverse needs of students are acknowledged and addressed, focusing on a developmental and ability-oriented approach. However, there is still no systematic approach for a comprehensive or nationwide supported promotion of gifted children, which is also the result of insufficient knowledge from research on optimal talent development for gifted students.

Whereas traditional views on giftedness often focused on above average general cognitive functioning as the sole criterion, more recent models emphasize a multidimensional perspective, considering not only cognitive abilities, but also factors like creativity, motivation and personality traits (Preckel et al., 2020). This broader approach in the attempt of understanding giftedness recognizes that gifted individuals exhibit diverse characteristics and that their development is influenced by a variety of environmental, cognitive, and emotional factors.

In this symposium, we present current research on and explore various approaches to talent development from multiple disciplines, such as education science, STEM didactics and developmental psychology. The symposium aims at furthering our understanding about the antecedents of talent and giftedness, and how the unique potentials of talented and gifted children and adolescents can be fostered.

The first contribution presents an individual participants data meta-analysis regarding the overall effectiveness of extracurricular STEM-related courses for gifted children from primary school and explores differential effects of these courses for boys and girls by considering achievement- as well as motivation-related outcomes.

The second contribution investigates the effectiveness of two specifically designed interventions to enhance computational thinking and mathematical structuring abilities in talented elementary students.

The third contribution aims at reconceptualizing gifted underachievement by comparing multidimensional and intelligence-only models, pointing out the importance of a holistic approach to identifying and supporting underachieving students from various age groups.

The fourth contribution considers creative achievements as a crucial outcome of gifted students. Through an integration of personality, motivation, and intelligence as predictors of creative achievements in gifted adolescents, this contribution examines the complex interplay of these constructs in understanding giftedness.

Together, these contributions (a) shed light on state-of-the-art ways to foster talent and giftedness across various topics and (b) offer empirical examinations of competing approaches of understanding talent and giftedness.

Beiträge des Symposiums

Integrating the Effects of a STEM Enrichment Program for Primary School Children

Benjamin Goecke, Jessika Golle, Benjamin Nagengast, Ulrich Trautwein

Hector Research Institute of Education Sciences and Psychology, University of Tübingen

Extracurricular enrichment programs are one important possibility for nurturing talented or gifted students. This importance is backed up by previous meta-analytic evidence that showed substantial positive effects of such programs on academic achievement and socioemotional development across various age groups (Kim et al., 2016). In the state of Baden-Württemberg, the Hector Children's Academy Program (HCAP) offers a comprehensive framework of extracurricular enrichment (Trautwein et al., 2023). The Hector Course Courses (HCCs), which focus on STEM-related topics, are a unique element of HCAP. Importantly, HCCs are developed and evaluated following a rigorous empirical step-wise approach consisting of a series of randomized controlled trials (RCTs) (efficacy, effectiveness, scaling-up) in which the courses are tested regarding their effectiveness towards central achievement- and motivation-related outcomes. Although several single studies have already shown the effectiveness of single courses in different stages of their evaluation (e.g., Herbein et al., 2018; Schiefer et al., 2021), it remains to be determined whether (a) the HCCs as a specific element of the HCAP are effective across evaluation stages and across the different STEM-related topics they cover (e.g., chemistry) and (b) whether differential effects for boys and girls that could be observed in single studies are generalizable across evaluation stages and courses.

To address these open questions, we re-analyzed existing data from several efficacy, effectiveness, scaling-up studies of N = 9 STEM-related HCCs with a total N = 2.356 participating primary school children. Specifically, we implemented an individual participants data two-step meta-analytical approach (Burke et al., 2016). In the first step, we specified the same regression models for each course and outcome (we distinguished between achievement- and motivation-related outcomes) to make the regressions coefficients comparable to each other (i.e., use the same control variables in each dataset). In the second step, we meta-analyzed the standardized regression coefficients () of the specified regression models to examine the overall effect of the single main effects (e.g., gender, treatment group, and pretest score) and interactions (gender * treatment group). To account for the nested data structure due to the experimental designs of the single studies (e.g., cluster-randomizations), we corrected the standard errors of our models with a twofold strategy. First, the standard errors of the regression models were corrected based on a design-based approach (e.g., DiStefano & Zhang, 2022). Second, the standard errors of the three-level random effects meta-analytical restricted maximum likelihood (REML) models were estimated using a biased-reduced linearization adjustment (Pustejovsky & Tipton, 2022). In total, 86 achievement-related and 52 motivation-related effect sizes were considered.

Our results first revealed that the main effect of treatment was substantial for achievement-related outcomes, but zero for motivation-related outcomes. Second, there was no overall main effect of gender for either achievement- or motivation-related

outcomes. Third, the overall effect of pretest scores was substantial across both achievement- and motivation-related outcomes. Lastly, the overall interaction effects for both outcome types were non-significant indicating no advantages or disadvantages for either gender in relation to the treatment received.

Our analyses replicate findings of previous single studies within the HCAP and extend research on the overall effectiveness of extracurricular enrichment programs for talented and gifted students. Considering possible differential gender effects surpasses previous meta-analytic studies. It demonstrates that whereas both boys and girls benefited from the HCCs, especially with respect to achievement-related outcomes, there were no significant differences in the magnitude of these benefits across gender. Concluding, our study provides robust evidence supporting the effectiveness of the HCCs and emphasizes the importance of considering gender differences in the evaluation of enrichment programs. These findings have significant implications for the design and implementation of extracurricular programs aimed at fostering talented students.

Promoting Computational Thinking and Mathematical Structuring Abilities in Talented Children

Katrin Kunz¹, Judith Havemann², Ann-Kathrin Jaggy¹, Katerina Tsarava¹, Walther Paravicini², Ulrich Trautwein¹

¹Hector Research Institute of Education Sciences and Psychology, University of Tübingen, ²Department of Mathematics, University of Tübingen

In recent years, there has been an increased emphasis on integrating computational thinking (CT) into educational curricula to prepare students for the demands of the digital age (Bocconi et al., 2022). CT is defined as a general problem-solving ability based on thinking processes needed for programming and can be applied in other fields of computer science and beyond (e.g. Shute, 2017). As such, it has been coined a crucial 21st-century skill and suggested to be acquired early on in education (e.g., Wing, 2006).

Mathematical structuring competence (MSC) refers to the ability to structure mathematical content by recognizing, applying and building patterns and structures (Ehrlich, 2015). MSC is integral to mathematical thinking and essential for identifying relationships, forming generalizations, and constructing mathematical models. It is considered a crucial component in fostering mathematical talent (e.g. Kapnick, 1998; Krutetskii, 1976; Sijts, 2017).

As there is a general lack of systematic approaches and empirical studies in both the fields of fostering CT and MSC in talented students, two interventions were developed specifically for talented elementary school students to promote CT and MSC competences respectively.

The present study evaluated the effectiveness of these 1-week long online interventions and investigated the interventions' impact on students' CT and MSC abilities respectively, as well as on their motivational dispositions towards programming and mathematics. Data were collected from a sample of 104 students (32 female) from grades 3 to 6, who were part of an extracurricular STEM enrichment program.

A randomized controlled block design with repeated measures was used to assess the effectiveness of the interventions. Students were randomly assigned to either the CT intervention, which involved learning the programming language Python while transitioning from block-based to text-based programming, or the MSC intervention, which focused on mathematically rich pattern tasks using inquiry-based learning. Multiple linear regression analyses were conducted to evaluate the intervention effects on CT, MSC and motivational dispositions.

Results indicated that students assigned to the CT intervention showed no significant improvement in CT performance compared to the MSC group ($\beta = .11, p = .577$). This suggests that the 1-week duration may not have been sufficient to significantly impact CT abilities to a measurable extent. However, the CT intervention was effective in promoting a positive self-concept ($\beta = .75, p < .001$) and intrinsic value ($\beta = .72, p < .001$) for programming. Additionally, the CT intervention improved students' self-concept ($\beta = .24, p = .002$) regarding mathematics. This suggests potential cross-domain benefits of integrating CT activities in educational curricula. Children assigned to the MSC intervention developed significantly more sophisticated mathematical structuring competence compared to those in the CT group ($\beta = .47, p < .001$). No significant positive effects of the MSC intervention were found regarding self-concept ($\beta = -.24, p = .021$), and intrinsic value ($\beta = -.13, p = .408$) towards mathematics.

Overall, the study highlights the importance of designing and implementing targeted interventions to develop essential skills and positive attitudes in talented students. The findings emphasize the potential of short-term interventions to positively influence elementary school students' abilities and motivational dispositions towards mathematics and programming.

Reconceptualizing Gifted Underachievement: Comparing Multidimensional and Intelligence-Only Approaches

Mingjing Zhu, Jessika Golle, Benjamin Nagengast, Ulrich Trautwein

Hector Research Institute of Education Sciences and Psychology, University of Tübingen

Although contemporary models of giftedness emphasize a multidimensional perspective (e.g., Preckel et al., 2020; Renzulli, 2005; Subotnik et al., 2011), most empirical research on gifted underachievement has primarily relied on intelligence as the main criterion for giftedness (e.g., Agaliotis & Kalyva, 2019; Desmet et al., 2020; Mofield & Parker Peters, 2019; Obergrösser & Stoeger, 2015; Rubenstein et al., 2012; Siegle et al., 2020). This study aims to explore the similarities and differences in identifying gifted underachieving students (GUAs) using theoretical frameworks distilled from previous models of giftedness and talent development. We categorized these frameworks into three representative models: the Intelligence model, which considers intelligence the sole measure of giftedness; the Three-Ring model, based on Renzulli's Three-Ring Conception of Giftedness, which incorporates ability, creativity, and motivation; and the Comprehensive model, which builds upon the others by adding a fourth aspect: personality.

We analyzed data from primary (2,785 third graders), middle (14,126 ninth graders), and high school (4,912 thirteenth graders) students to address three research questions: 1) Would different numbers of gifted underachieving students be identified if a single indicator of intelligence vs. multiple indicators were used to define giftedness? 2) To what extent would there be an overlap between gifted underachieving students identified by multiple indicators of giftedness and those identified by intelligence alone? 3) How would gifted underachieving students identified by single and multiple indicators of giftedness differ in terms of both personal and environmental factors? The findings show that the Intelligence model consistently identified more GUAs than the Three-Ring and Comprehensive models. Little overlap occurred in identified GUAs between the single- and multiple-indicator models. GUAs identified by the Intelligence model had higher intelligence, lower creativity, lower motivation, and perceived less teacher support than those identified by the Three-Ring and Comprehensive models. Their socioeconomic status was higher, and they were less open and conscientious than GUAs identified by the Comprehensive model. Exploring these nuances sends a clear message that decisions about underlying concepts make a difference in identifying GUAs. It signals the need to take the

multidimensionality of giftedness into account when conceptualizing gifted underachievement for future empirical studies and developing strategies to address gifted underachievement in educational settings.

Predicting Creative Achievements – A Comprehensive Approach Integrating Personality, Motivation, and Intelligence of Gifted Children

Ann-Kathrin Jaggy¹, Brent Roberts²

¹Hector Research Institute of Education Sciences and Psychology, University of Tübingen, ²University of Illinois at Urbana-Champaign

Creativity is key to innovation and thus an important 21st-century skill; essential to addressing complex global challenges such as rapid digitalization, climate change, economic instability, and social inequalities. Understanding the antecedents of creativity is critical for understanding how societies can better support individuals in developing these capacities from an early age, thus ensuring more creative achievement later in life. However, historically, much of the research on children's talent development and achievements has largely focused on traditionally defined cognitive ability (e.g., IQ) and academic achievements (Gajda et al., 2017), neglecting the importance of creative achievements as developmental milestones earlier in life. Moreover, studies often focus on only single personality domains, that are related to creativity and creative achievements, such as cognitive ability (Benedek, 2024). Thus, there is a lack of multivariate studies focusing on interindividual differences related to creative achievements in adolescents.

Taking a comprehensive approach, the study aims to contribute to a better understanding of the nomological network of creativity by investigating creative achievements in a sample of gifted secondary school children. We combine a confirmatory and exploratory approach integrating a wide range of theoretically derived variables from different research areas and examine their importance for predicting creative achievements. To organize the variety of factors that contribute to creative achievements, the Neo-Socioanalytic model of personality (Roberts et al., 2024) is used, which considers children's personality traits, motivation, and abilities.

In this multivariate and multi-modal study, we collected self-reports from N = 431 gifted children on the big five personality traits, creative self-efficacy, and vocational interest, as well as creative activities and achievements. Children's cognitive abilities, including fluid intelligence, spatial thinking, and divergent thinking skills, were assessed using standardized tests. Additionally, parents provided informant reports on children's personality traits. A machine learning approach using random forests (Breiman, 2001) was employed to identify the most important predictors of creative achievements, accounting for the high number and multicollinearity of included variables.

Preliminary results show that the most important predictors of creative achievements were (in descending sequence) creative activities, artistic vocational interest, assertiveness, openness, creative self-efficacy, and divergent thinking. A more in-depth analysis of these variables using a structural equation model will provide further insights into the interplay between these factors in predicting creative achievement.

These findings demonstrate that various aspects of personality traits, motivation, and cognitive abilities collectively contribute to creative achievements in adolescents, highlighting the importance of considering all these domains of personal characteristics in research and practice. We discuss a comprehensive model that illustrates the complex interplay between personal characteristics and creativity in real life. The study contributes to a better understanding of the relationship between creative potential and performance in talent development.

S-102: Förderung der Bewertungskompetenzen von Informationen im postfaktischen Zeitalter

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: SN 169

Chair der Sitzung: Philipp L. Marten

Chair der Sitzung: Ingo Kollar

Förderung der Bewertungskompetenzen von Informationen im postfaktischen Zeitalter

Chair(s): **Philipp L. Marten** (Ruhr-Universität Bochum)

Diskutant*in(nen): **Ingo Kollar** (Universität Augsburg)

In der heutigen, sich schnell entwickelnden Welt sind Individuen daran gewöhnt, sich jederzeit selbstständig Wissen aneignen zu können. Dabei müssen sie komplexe, von Algorithmen kuratierte Online-Umgebungen navigieren, wissenschaftliche Informationen interpretieren und die Qualität dargebotener Informationen eigenständig beurteilen, die teilweise auch irreführende oder falsche Informationen enthalten können (Ecker et al., 2024). Diese allgegenwärtige Herausforderung befördert ein postfaktisches Zeitalter, das durch Unsicherheiten in Bezug auf vorhandenes Wissen, die Mittel zum Wissenserwerb und die Glaubwürdigkeitsbewertung von Informationsquellen geprägt ist (Barzilai & Chinn, 2020).

Ein Ausweg aus diesem Dilemma könnte in der Ausbildung von Individuen zu „kompetenten Außenseitern“ (Osborne & Pimentel, 2022) liegen. Um dieses Ziel zu erreichen, müssen diese dazu in der Lage zu sein, begründete Plausibilitäts- und Vertrauensurteile zu fällen, die auf dem Wissen darüber basieren, was Wissenschaftler und die Wissenschaft als System zu glaubwürdigen Wissensquellen macht („informiertes Vertrauen“; Bromme, 2020).

Eine solche Ausbildung muss jedoch durch weitere Kenntnisse flankiert werden. Beispielsweise bedarf es eines grundlegenden Verständnisses über die Verwendung und die Funktionsweise von Algorithmen, um digitale Umwelten kritisch reflektieren und kompetent navigieren zu können (Dogruel et al., 2021). Der versierte Einsatz von Evaluationsstrategien muss eingeübt sein, um informierte Urteile über Quellen zu fällen und Fakten von Falschinformationen zu unterscheiden, ohne über tiefgreifendes Fachwissen zu verfügen (Duncan et al., 2018). Und Wissenschaftler müssen dazu in der Lage sein, ihre Erkenntnisse verständlich zu kommunizieren, ohne die notwendige Tiefe der wissenschaftlichen Ergebnisse zu verlieren (Scharrer et al., 2017).

Um zukünftig erfolgreiche Interventionen zu gestalten, die das Ziel verfolgen, „kompetente Außenseiter“ auszubilden, ist es deshalb von entscheidender Bedeutung, die Effekte von potenziell erfolgsversprechenden Ansätzen empirisch zu überprüfen. Dieses Symposium trägt zur Forschung über die Förderung individueller Bewertungskompetenzen von Informationen im postfaktischen Zeitalter bei, indem es drei empirische Studien zusammenbringt.

Im ersten Symposiumsbeitrag wurde mit einem Prä-Post-Kontrollgruppendesign ($N = 492$) eine videobasierte Kurzintervention zur Steigerung von Awareness und Wissen über Algorithmen ausgewertet und in Zusammenhang mit der Wahl vertrauenswürdiger Quellen analysiert. Die Ergebnisse der hierarchischen Regressionsanalysen stützen die Annahme, dass insbesondere Wissen über die Funktionsweise von Algorithmen von zentraler Bedeutung ist, um algorithmenkuratierte Umwelten zu navigieren und vertrauenswürdige Quellen zu identifizieren.

Der zweite Symposiumsbeitrag untersuchte die Effekte von Fachsprache auf die Glaubwürdigkeit von Informationen und das einhergehende Vertrauen in die kommunizierende Person. In drei Gruppen eingeteilte Teilnehmende ($n = 1192$) erhielten in einer randomisierten Online-Umfrage einen wissenschaftlichen Text über die Effekte eines fiktiven Medikaments, bei dem jeweils die verwendete Menge an Fachjargon unterschiedlich stark ausgeprägt war (niedrig vs. mittel vs. hoch). Die Ergebnisse mehrerer ANOVAs legen nahe, dass niedrige Fachsprache nicht unbedingt eine Überschätzung der eigenen Urteilsfähigkeit über wissenschaftliche Geltungsbehauptungen fördert und sogar positive Effekte auf die Urteile zur Vertrauenswürdigkeitsbewertung von Experten haben kann.

Der dritte Symposiumsbeitrag evaluierte in einem gemischten between-within-Design die Effekte eines Strategietrainings für die Mittelstufe ($N = 514$) zur Förderung von Evaluationsstrategien, insbesondere Sourcing und Corroboration. Die individuellen Fähigkeiten wurden mit einer aktiven Kontrollgruppe zu drei Messzeitpunkten verglichen. Kontrastanalysen legen nahe, dass Jugendliche des Strategietrainings sich in der Beurteilung von Quellen sowie in der Bewertung von irreführenden Social-Media-Behauptungen signifikant verbesserten. Die selbstsichere Identifikation von Deepfake-Videos verbesserte sich jedoch nicht.

Ein renommierter Diskutant wird die Studienergebnisse zusammenbringen, Potentiale identifizieren und Herausforderungen adressieren. Zu den erwarteten wissenschaftlichen und pädagogischen Erkenntnissen gehören ein besseres Verständnis über die Bedingungen für eine erfolgreiche Algorithmen-Awareness-Schulung, Einschätzungen zu adäquaten Formulierungen von wissenschaftlichen Erkenntnissen sowie ein Ausblick auf die Möglichkeiten zur Förderung von Evaluationsstrategien bei Jugendlichen.

Beiträge des Symposiums

Awareness und Wissen über Algorithmen: Überprüfung einer Kurzintervention und der Zusammenhang mit der Wahl vertrauenswürdiger Quellen

Eva Thomm, Johannes Bauer, Leyla Dogruel, Sven Jöckel
Universität Erfurt

Datenbasierte Technologien durchdringen zunehmend unseren Alltag. Aufgrund dessen ist es wichtig, Wissen und Fähigkeiten im Umgang mit solchen algorithmenkuratierten Online-Umwelten zu vermitteln (Frau-Meigs, 2024). Studien deuten bereits an, dass fehlendes Wissen und mangelnde Kompetenzen im Umgang mit Algorithmen zu einem neuen Digital Divide führen können – einer Kluft, die sich u.a. im Zugang zu und der Nutzung solcher Technologien zeigt (z.B. Gran et al., 2021). Besonders die Awareness über die Verwendung von Algorithmen und das Wissen über deren Funktionsweise sind entscheidend, um digitale Umwelten kritisch reflektieren und kompetent navigieren zu können. Beides bedarf jedoch einer gezielten Vermittlung (Dogruel et al., 2021).

Auch kompetentes Sourcing setzt ein Verständnis von Algorithmen voraus. Da traditionelle Gatekeeper als Garanten verlässlicher Informationen fehlen, ist die systematische Berücksichtigung und Bewertung von Quelleninformationen von zentraler Bedeutung (Bråten et al., 2018). Hierfür ist es wichtig zu verstehen, dass Suchergebnisse im Internet von Daten über individuelle Präferenzen und Einstellungen beeinflusst werden (Caulfield & Wineburg, 2023). Bisher ist jedoch wenig über den Zusammenhang zwischen Sourcing und dem Verständnis von Algorithmen und ihrer Funktionsweise bekannt.

Diese Studie verfolgte zwei Ziele. Erstens wurde untersucht, ob ein kurzes Erklärvideo über die Funktionsweise von Algorithmen die Awareness und das Wissen über deren Verwendung verbessert. Zweitens gingen wir der Frage nach, inwieweit diese beiden Aspekte mit der Auswahl vertrauenswürdiger Quellen im Internet zusammenhängen. Wir nahmen an, dass eine höhere Awareness bzw. mehr Wissen mit einer stärkeren Präferenz für vertrauenswürdige Quellen einhergeht.

An der Studie nahmen Personen einer quotenrepräsentativen Stichprobe eines Online Access Panels der deutschen Allgemeinbevölkerung teil (N = 492, M = 48 Jahre (SD = 15), 51% Frauen, 57% Abitur). Die Studie folgte einem Prä-Post-Kontrollgruppendesign. Die Teilnehmenden sahen ein Erklärvideo entweder zur Funktionsweise von Algorithmen (Experimentalgruppe) oder zur Barrierefreiheit im Internet (Kontrollgruppe). Vor der Videopräsentation wurden die Awareness über die Verwendung von und das Wissen über die Funktionsweise von Algorithmen erhoben (je 11 Items; Dogruel et al., 2021). Unmittelbar nach der Präsentation wurde eine Anwendungsaufgabe bearbeitet, die das Verständnis von Algorithmen anhand konkreter Alltagsszenarien erfasste (Dogruel et al., 2021). Zwei Wochen später wurden Algorithmen-Awareness und Algorithmen-Wissen analog zu T1 erneut erhoben. Zudem bewerteten die Teilnehmenden eine simulierte Trefferliste einer Suchmaschine zur Frage, ob koffeinhaltige Energy-Drinks die Leistungsfähigkeit verbessern. Die Trefferliste enthielt sehr, moderat und wenig vertrauenswürdige Quellen. Sie wählten aus neun Treffern die drei besten aus. Aus dieser Auswahl wurde ein Präferenzscore für vertrauenswürdige Quellen berechnet. Alter, Geschlecht und Schulbildung wurden kontrolliert.

Mittels hierarchischer Regressionsanalysen wurde der Effekt der videobasierten Intervention auf (a) die Bearbeitung der Anwendungsaufgabe (T1), (b) die Algorithmen-Awareness und (c) das Algorithmen-Wissen zu T2 untersucht. Dabei wurden Algorithmen-Awareness und Algorithmen-Wissen zu T1 sowie Alter, Geschlecht und Schulbildung berücksichtigt. Es zeigte sich kein signifikanter Effekt der Intervention, jedoch erwies sich das Wissen über Algorithmen (T1) als signifikanter Prädiktor: Je mehr Wissen die Teilnehmenden bereits zu Beginn der Befragung aufwiesen, desto besser schnitten sie in der Anwendungsaufgabe ab und zeigten auch zu T2 eine höhere Awareness für die Verwendung von Algorithmen. Das Algorithmen-Wissen zu T1 sagte zudem die Präferenz für vertrauenswürdige Quellen positiv vorher. Dieser Zusammenhang blieb auch nach Kontrolle eines möglichen Einflusses der Intervention sowie der Kovariaten statistisch signifikant. Entgegen unseren Erwartungen zeigte sich kein Effekt der Algorithmen-Awareness.

Die Ergebnisse stützen die Annahme, dass insbesondere Wissen über die Funktionsweise von Algorithmen von zentraler Bedeutung ist, um algorithmenkuratierte Umwelten zu navigieren und vertrauenswürdige Quellen zu identifizieren. Dies unterstreicht die Notwendigkeit einer gezielten Vermittlung, um eine informierte Nutzung digitaler Technologien zu fördern. Hierzu scheint es jedoch mehr zu bedürfen als der Präsentation eines kurzen Erklärvideos. Die Ergebnisse werden hinsichtlich der Limitationen der Studie sowie im Kontext aktueller Forschung zum Verständnis von Algorithmen und Sourcing diskutiert.

Die Effekte von Fachsprache auf die Glaubwürdigkeit von Informationen und das Vertrauen in die kommunizierende Person

Julian Fick, Luca Rudolph, Friederike Hendriks
Technische Universität Braunschweig

Wissenschaftliches Wissen ist wichtig zur Lösung alltäglicher Probleme (z.B. die Entscheidung, ein Medikament einzunehmen), aber auch Grundlage politischer Teilhabe (z.B. Diskurs um Energiesicherheit). Daher ist es wichtig, wissenschaftliche Informationen bewerten zu können, was aber durch die hohe Komplexität und Polarisierung in heutigen Informationsumwelten und – insbesondere im Internet – den Verlust von Gatekeepern, die über die Qualität von Informationen wachen, herausfordernd ist (McGrew et al., 2018; Osborne & Pimentel, 2023). Zudem führen oft Oberflächenmerkmale wissenschaftlicher Informationen zu vorschnellen Urteilen über ihre Glaubwürdigkeit. Es wurde beispielsweise herausgefunden, dass eine Vermeidung von Fachsprache in wissenschaftlichen Informationen bei den Rezipierenden zu einer Überschätzung der eigenen Fähigkeiten im Themenbereich des Textes führen kann (easiness effect of science popularization; Scharrer et al., 2012, 2017, 2019). Der vermutete Mechanismus liegt darin, dass die Vermeidung von Fachsprache zwar die Worte im Text (d.h. die Textoberfläche; van Dijk & Kintsch, 1983), nicht aber das zugrundeliegende Thema verständlich macht, sich aber das Gefühl der Einfachheit bei den Rezipierenden vom Text auf das Thema überträgt und dadurch nicht nur der Text, sondern fälschlicher Weise auch das zugrundeliegende Thema als einfach empfunden wird. Außerdem bewerteten Lesende die Expertise des Autors bzw. der Autorin als niedriger, wenn er bzw. sie weniger Fachsprache verwendete (Zimmermann & Jucks, 2018). In anderen Studien zeigten sich diese negativen Effekte der Vermeidung von Fachsprache nicht oder nur teilweise (Jonas et al., 2024; Thon & Jucks, 2017).

Ziel der Studie ist es nicht nur die genannten Befunde, insbesondere den easiness effect, konzeptuell zu replizieren, sondern auch zu erweitern, indem wir untersuchen, ob es einen Mittelweg gibt, bei dem die Vorteile der Reduzierung von Fachjargon – insbesondere die Erhöhung der Textverständlichkeit – genutzt werden können, ohne die berichteten unerwünschten Effekte zu verursachen. Daher führten wir eine dritte Bedingung (zusätzlich zu maximalem und minimalem Fachjargon) mit einer mittleren Menge an Fachjargon ein. Außerdem untersuchten wir empirisch, ob die Verarbeitungsflüssigkeit sowie metakognitive Urteile den easiness effect erklären können.

Wir führten eine Online-Umfrage durch (n = 1192), bei der wir die Teilnehmenden zufällig in drei Gruppen einteilten und sie mit einem wissenschaftlichen Text über die Effekte eines fiktiven Medikaments konfrontierten, der sich in der verwendeten Menge an Fachjargon unterschied (niedrig vs mittel vs hoch). Anschließend wurden sie gefragt, ob sie der zentralen Behauptung des Textes zustimmen, wie sicher sie sich bei dieser Entscheidung waren, ob sie einen Experten konsultieren wollten und wie sie die Fachkompetenz, Integrität und Wohlwollen des Autors wahrgenommen haben. Insgesamt konnten wir die negativen Effekte der Vermeidung von Fachjargon sehr bedingt replizieren. Nur für die Zustimmung zur Behauptung zeigte sich ein sehr kleiner signifikanter Effekt ($F(2, 1189) = 3.39, p = .034, \eta^2 = .006$). Stattdessen führte ein geringeres Ausmaß an Fachjargon zu einer höheren Wahrnehmung des Wohlwollens ($F(2, 1189) = 4.74, p = .009, \eta^2 = .008$) und der Integrität des Autors ($F(2, 1189) = 6.14, p = .002, \eta^2 = .010$). Außerdem medierte die Verarbeitungsflüssigkeit, nicht aber die metakognitiven Urteile, die Beziehung zwischen dem Gebrauch von Fachjargon und der Zustimmung zu der zentralen Behauptung des Textes, oder zur Sicherheit der eigenen Urteilsfähigkeit bzw. dem Wunsch weitere Experten zu konsultieren.

Aus den Ergebnissen können wir schlussfolgern, dass das Vermeiden von Fachsprache – wenn es angemessen ist – nicht unbedingt eine Überschätzung der eigenen Urteilsfähigkeit über wissenschaftliche Geltungsbehauptungen fördert und sogar positive Effekte auf die Urteile zur Vertrauenswürdigkeit wissenschaftlicher Expert*innen haben kann. Wir werden das spezifische Thema (Gesundheit), das Studiendesign (between-subjects), die absolute Verständlichkeit des verwendeten Textes in unserer Studie, sowie die Robustheit des easiness effects als Gründe für das Ausbleiben der Replikation diskutieren und Implikationen für den Gebrauch von Fachjargon in der Wissenschaftskommunikation und in inzidentellen Lernkontexten ableiten.

Kann ich dieser Information vertrauen? Effekte eines Strategietrainings für die Mittelstufe zur Förderung von Evaluationsstrategien

Philipp L. Marten, Marc Stadtler

Ruhr-Universität Bochum

In den vergangenen Jahren wurden in diversen Interventionsstudien verschiedene Ansätze zur Verbesserung des individuellen Umgangs mit Online-Informationen untersucht (Kozyreva et al., 2024; Roozenbeek et al., 2023). Die Förderung von Evaluationsstrategien hat sich als besonders vielversprechender Ansatz für verschiedene Altersgruppen herausgestellt (Brante & Strømsø, 2018; McGrew, 2024). Zentrale Evaluationsstrategien sind Sourcing, also die kritische Bewertung und Nutzung verfügbarer Informationen über den Urheber einer Behauptung (Bråten et al., 2018), sowie Corroboration, also der Abgleich einer Behauptung mit weiteren vertrauenswürdigen Quellen (Wineburg, 1991). Um informierte Urteile über Quellen zu fällen und Fakten von Falschinformationen zu unterscheiden, ohne über tiefgreifendes Fachwissen zu verfügen, müssen Evaluationsstrategien versiert eingesetzt werden (Duncan et al., 2018). Leider werden junge Menschen bislang nicht ausreichend im Einsatz von Evaluationsstrategien gefördert (Kheredmand, 2022).

In dieser Studie wurde deshalb die Lernwirksamkeit eines theoretisch fundierten, von Lehrkräften geleiteten Strategietrainings für die 7. und 8. Jahrgangsstufe untersucht. Wir nahmen an, dass Teilnehmende des Strategietrainings im Vergleich zu einer aktiven Kontrollgruppe nach der Intervention besser H1) zwischen vertrauenswürdigen und nicht vertrauenswürdigen Quellen differenzieren, H2) falsche Behauptungen in sozialen Medien erkennen, und H3) Deepfakes mit größerer Sicherheit identifizieren würden.

514 Jugendliche aus 21 intakten Klassen nahmen an den Trainings teil. Diese wurden zufällig entweder dem Strategietraining (11 Klassen, $M = 12.83$, $SD = .70$, 50% weiblich) oder dem Kontrolltraining (10 Klassen, $M = 12.93$, $SD = .86$, 52% weiblich) zugeteilt. Faktoren wie Schulform oder Schul-Sozialindexstufe wurden ausbalanciert. Beide Trainings bestanden aus einer Sequenz von vier Unterrichtseinheiten à 90 Minuten. Während Jugendliche im Strategietraining Sourcing- und Corroboration-Strategien kennenlernten und an authentischen Anwendungsbeispielen einübten, vermittelte das Kontrolltraining ein umfassendes Verständnis von Falschinformationen und deren Auswirkungen, ohne jedoch Evaluationsstrategien praktisch einzusetzen. Die Datenerhebung umfasste drei Messzeitpunkte: Pretest, Posttest und um einen Monat verzögerter Follow-up-Test.

Zu jedem Messzeitpunkt a) gaben die Jugendlichen bei zehn Quellen mit variierender Vertrauenswürdigkeit an, ob sie diese für ein Schulreferat verwenden würden; b) beantworteten sie vier Multiple-Choice-Fragen zu einem irreführenden Social-Media-Beitrag; und c) entschieden bei zwei Videos (ein Deepfake, ein genuines Video), ob es sich um Deepfakes handelte (ja/nein) mit Angabe ihres Sicherheitsgrads. Lesekompetenz, Sozioökonomischer Status und Interesse am Thema wurden kontrolliert.

Kontrastanalysen zeigten eine signifikante Interaktion zwischen Trainingsgruppe und Messzeitpunkt in Bezug auf Quellenbewertungen (H1), $F(1, 222) = 9.10$, $p = .003$, $\eta^2 = .04$. Die Strategietrainingsgruppe zeigte beim Posttest eine bessere Differenzierung der unterschiedlich vertrauenswürdigen Quellen als die Kontrollgruppe, $t(361.56) = 4.90$, $p < .001$, $d = .51$, und behielt ihr Leistungsniveau beim Follow-up-Test bei, $t(346.74) = 5.85$, $p < .001$, $d = .61$.

Hinsichtlich der richtigen Antworten zu einem irreführenden Social-Media-Post ergab die Kontrastanalyse ebenfalls eine signifikante Interaktion (H2), $F(1, 338) = 8.68$, $p = .003$, $\eta^2 = .03$. Die Strategietrainingsgruppe zeigte im Vergleich zur Kontrollgruppe bessere Corroboration-Fähigkeiten zum Posttest, $t(274.04) = 3.43$, $p < .001$, $d = .37$, sowie zum Follow-up-Test, $t(316.85) = 2.86$, $p < .001$, $d = .34$.

Bei der Fähigkeit zwischen Deepfake und genuinen Videos zu unterscheiden (H3), ergab die Kontrastanalyse keine statistisch signifikante Interaktion, $F(1, 352) = 3.02$, $p = .083$, $\eta^2 = .01$.

Die Ergebnisse liefern Belege dafür, dass Jugendliche der 7. und 8. Jahrgangsstufe ihre Fähigkeiten zur Identifizierung vertrauenswürdiger Quellen und zuverlässiger Informationen mit Hilfe einer gezielten Förderung der Evaluationsstrategien Sourcing und Corroboration verbessern können. Wir führen diese Verbesserung auf die praktische Einübung anhand authentischer Beispiele in einer realen Internetumgebung zurück (Wineburg et al., 2022). Es zeigten sich jedoch keine verbesserten Fähigkeiten, KI-generierte von genuinen Videos mit größerer Sicherheit auseinanderzuhalten. Die Ergebnisse deuten insgesamt darauf hin, dass erfahrene Lehrkräfte unsere Lehrmaterialien mit moderater Vorbereitung wirksam einsetzen konnten. Die als Open-Educational-Ressource veröffentlichten Materialien könnten somit schon jetzt flächendeckend eingesetzt werden und Jugendlichen einen großen Nutzen bringen.

S-104: Berufsorientierung und Übergänge in berufliche Ausbildung von geflüchteten Jugendlichen und jungen Erwachsenen

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · *Ort:* SO 108

Chair der Sitzung: Regina Becker

Chair der Sitzung: Gisela Will

Chair der Sitzung: Mona Granato

Berufsorientierung und Übergänge in berufliche Ausbildung von geflüchteten Jugendlichen und jungen Erwachsenen

Chair(s): **Regina Becker** (Leibniz-Institut für Bildungsverläufe, Bamberg), **Gisela Will** (Leibniz-Institut für Bildungsverläufe, Bamberg)

*Diskutant*in(nen):* **Mona Granato** (Bundesinstitut für Berufsbildung (BIBB), Bonn)

Mitte der 2010er Jahre kam es zu einem starken Anstieg an Fluchtzuwanderung in Deutschland. Ein großer Anteil der damals Angekommenen gehört der Gruppe der 16-30-Jährigen an (Statistisches Bundesamt, 2019; BAMF, 2017). Für diese Jugendlichen und jungen Erwachsenen ist einerseits die Entwicklung beruflicher Aspirationen und berufliche Orientierung (Rojewski, 2005), andererseits der Übergang in berufliche Ausbildung von zentraler Bedeutung für den späteren beruflichen Werdegang (Granato et al., 2006; Becker, 2021). Dabei sind die Jugendlichen bzw. jungen Erwachsenen als relativ späte Seiteneinsteiger:innen in das deutsche Bildungs- und Ausbildungssystem mit besonderen Herausforderungen konfrontiert. Zu diesen können bspw. eine geringere Vertrautheit mit dem deutschen Bildungs- und Ausbildungssystem, geringere Deutschsprachkenntnisse, fehlende Kenntnisse über Voraussetzungen für spätere Arbeitsmarktpartizipation und Erträge von Berufsausbildungen sowie mögliche Barrieren beim Zugang zu betrieblicher Ausbildung etwa aufgrund von Diskriminierung (Eberhard/Schulz, 2021) zählen.

Im Symposium wird der Frage nachgegangen, wie geflüchtete Neuzugewanderte, die als Jugendliche oder junge Erwachsene in Deutschland angekommen sind, unter diesen Voraussetzungen berufliche Orientierungen entwickeln und wie ihr Übergang in berufliche Ausbildung gelingen kann. Die vier Beiträge fokussieren dabei unterschiedliche Bildungsetappen und reichen thematisch von der Entwicklung von Berufsaspirationen in der Sekundarstufe I des allgemeinbildenden Schulsystems über den Übergang in berufsvorbereitenden Maßnahmen bis hin zum Zugang zu beruflicher Ausbildung. Ein Diskussionsbeitrag reflektiert die Einzelbeiträge und setzt diese zueinander in Beziehung.

Die ersten beiden Beiträge beschäftigen sich mit geflüchteten Schüler:innen als Seiteneinsteiger:innen ins deutsche Bildungssystem und verwenden Daten der Studie „ReGES – Refugees in the German Educational System“ (Will et al., 2021). Der erste Beitrag thematisiert die Entwicklung von beruflichen Aspirationen von geflüchteten Schüler:innen. Dies ist insbesondere dahingehend relevant, da vorhandene Opportunitätsstrukturen des Bildungs- und Ausbildungssystem möglicherweise noch wenig bekannt sind. Untersucht wird der Zusammenhang der Entwicklung der Berufsaspirationen der Jugendlichen und elterlichen Berufsaspirationen, dem Besuch eines akademischen Schulzweigs und antizipierter Diskriminierung auf dem Ausbildungsmarkt. Der zweite Beitrag betont die Bedeutung des berufsvorbereitenden Sektors als erste Station nach dem Besuch einer allgemeinbildenden Schule. Es wird näher beleuchtet, welche Arten von Berufsvorbereitungen besucht werden und von welchen Faktoren es abhängt, ob diese beispielsweise zum Nachholen von Bildungsabschlüssen genutzt werden oder dem Deutschspracherwerb und Kennenlernen des berufsbildenden Systems dienen.

Beitrag drei und vier widmen sich der Frage, welche Faktoren eine erfolgreiche Einmündung geflüchteter Jugendlicher und junger Erwachsener in berufliche Ausbildung beeinflussen. Der dritte Beitrag beschäftigt sich mit der Bedeutung von berufsbezogenen Fördermaßnahmen, die die Integration von Geflüchteten in den Ausbildungs- und Arbeitsmarkt erleichtern sollen. Hierin wird, auf Basis von Daten der IAB-BAMF-SOEP-Befragung von Geflüchteten (Brücker et al., 2016) analysiert, ob der Besuch einer berufsbezogenen Fördermaßnahme mit einer höheren Übergangschancen in eine Berufsausbildung einhergeht und ob dies in unterschiedlichem Maße für verschiedene Altersgruppen gilt. Der vierte Beitrag adressiert die Rolle von betrieblichen Rahmenbedingungen für die Einstellungschancen in Ausbildung für Geflüchtete. Mit den Daten des BIBB-Betriebspanel wird untersucht, welche betrieblichen Ausbildungsmotive und welche Arten von Tätigkeiten in den Betrieben eine Einstellung Geflüchteter begünstigen. Eine Zeitverlaufsperspektive erlaubt Unterschiede aufgrund gesellschaftlicher Veränderungen (z.B. der Coronapandemie) zu beleuchten.

Beiträge des Symposiums

Entwicklung beruflicher Aspirationen von geflüchteten Schüler:innen nach Klasse 7 und 8

Robin Busse¹, Oliver Winkler²

¹Technische Universität Darmstadt, ²Martin-Luther-Universität Halle-Wittenberg

In Deutschland hat die berufliche Ausbildung eine große Bedeutung für eine gelingende Integration Geflüchteter (Jeon, 2019). Im Zuge der Berufswahl kommt der Herausbildung von Berufsaspirationen eine besondere Rolle zu (Rojewski, 2005). Heranwachsende müssen dabei ihre Berufsaspirationen möglicherweise an eine weitgehend kompetitive Angebotsstruktur anpassen (Tomasik et al., 2009), da sie ihre beruflichen Entscheidungen innerhalb der Möglichkeiten des segmentierten und überwiegend marktorientierten Berufsbildungssystems treffen müssen (Protsch/Solga, 2016). Junge Geflüchtete stehen zusätzlich vor der Herausforderung, in einem neuen gesellschaftlichen Kontext berufliche Ziele zu entwickeln, von dem ihnen die vorhandenen Opportunitätsstrukturen vermutlich noch wenig bekannt sind (Bergseng et al., 2019). Ziel dieser Studie ist es, die Entwicklung der beruflichen Aspirationen bei jungen Geflüchteten, die zwischen 2014 und 2017 nach Deutschland migriert sind, zu untersuchen.

Wir definieren Berufsaspirationen als die Ziele einer Person in Bezug auf ihre zukünftige Karriere (Rojewski, 2005). Berufsaspirationen werden schon in der Kindheit gebildet, bleiben bis zur frühen Adoleszenz jedoch noch relativ vage (Gottfredson, 1981). Sie konkretisieren sich, wenn Bildungs- und Berufsentscheidungen näher rücken (Heckhausen/Tomasik, 2002), da sich Heranwachsende in dieser Phase zunehmend mit wahrgenommenen Chancenstrukturen auseinandersetzen (Gottfredson, 1981). Der Entwicklungsprozess von Berufsaspirationen wird durch verschiedene Faktoren beeinflusst. Anlehnend an Status-Attainment-Ansätze (Sewell et al., 1969) erwarten wir, dass höhere elterliche Berufserwartungen zu höheren Berufsaspirationen von jungen Geflüchteten führen (H1). Außerdem erwarten wir, dass junge Geflüchtete in akademischen Schulzweigen höhere Berufsaspirationen entwickeln als jene in nicht akademischen Schulzweigen (H2) (Basler/Kriesi, 2019).

Zudem erwarten wir, dass antizipierte Diskriminierungen auf dem Ausbildungsmarkt die Entwicklung der Berufsaspirationen negativ beeinflussen (H3) (Busse/Winkler, 2024).

Für die Analysen wird die Refugee Cohort 2 (RC2) der Studie „Refugees in the German Educational System (ReGES)“ herangezogen (Will et al., 2021). Die geflüchteten Jugendlichen waren zum Zeitpunkt der Stichprobenziehung (Oktober 2017) zwischen 14 und 16 Jahre alt und besuchten zum ersten Messzeitpunkt noch die Sekundarstufe I. Unsere Analysen beziehen sich auf die Erhebungswelle 1 (2017/2018, T1) und 4 (2019, T2) sowie auf junge Geflüchtete, die sich in Welle 1 in Klasse 7 oder 8 befanden ($n=1.050$). Die abhängige Variable ist die realistische Berufsaspiration der Befragten (Rojewski, 2005). Für die Analysen wird der International Socio-Economic Index of Occupational Status (ISEI) (Ganzeboom, 2010) der Berufsaspiration verwendet. Der ISEI ist ein auf Basis von Bildung und Einkommen generierter Index, dessen Wertebereich von 16 (z. B. Reinigungspersonal) bis 90 (z. B. Richter:innen) reicht. Wir verwenden ein Cross-Lagged-Strukturgleichungsmodell mit imputierten Daten (FIML).

Unsere Ergebnisse zeigen, dass die jungen Geflüchteten in Klasse 7 oder 8 (T1) erwarten, Berufe mit einem vergleichsweise (z. B. Basler/Kriesi, 2019 für Daten zu Schüler:innen ohne Fluchthintergrund) hohen ISEI aufzunehmen ($M=69,9$; $SD=20,2$). Im betrachteten Zeitverlauf (im Durchschnitt 10 Monate; $SD=1,6$ Monate) verringern sie ihre Berufsaspirationen im Durchschnitt um 7 ISEI-Punkte nach unten ($\Delta=-7,51$). Das Cross-Lagged-Modell ($\chi^2=3,07$, $df=3$, $RMSEA=0,02$; $CFI=0,96$) verweist auf systematische Unterschiede in den Berufsaspirationen der jungen Geflüchteten zwischen beiden Untersuchungszeitpunkten. So besteht in T1 ein signifikanter Zusammenhang zwischen den wahrgenommenen elterlichen Berufserwartungen und den Berufsaspirationen der geflüchteten Schüler:innen ($\beta=0,72$, $p<0,001$). Der Einfluss der wahrgenommenen elterlichen Berufserwartung hat sich zwischen T1 und T2 in Bezug auf die Berufsaspirationen der Schüler:innen reduziert und ist statistisch nicht mehr bedeutsam ($\beta=0,11$, $p=0,077$). Zudem finden wir Hinweise für den Einfluss des Schulzweigs auf das Ausgangsniveau der Berufsaspirationen in T1 ($\beta=4,00$, $p<0,001$). Hinweise für den Einfluss antizipierter Diskriminierungen auf dem Ausbildungsmarkt liegen dagegen nicht vor.

Unsere Befunde verdeutlichen die hohe Plastizität und Dynamik beruflicher Aspirationen junger Geflüchteter. Sie könnten nahelegen, dass sich die Orientierungen junger Geflüchteter von ihren eigenen Wahrnehmungen der Berufserwartungen der Eltern entkoppeln und sich dagegen stärker an den Opportunitätsstrukturen des Berufsbildungssystems orientieren. Schulzweige und daran anknüpfende Opportunitäten für Ausbildungswege kanalisieren die Orientierungen der Heranwachsenden in bedeutendem Umfang.

Geflüchtete Jugendliche im Übergangssektor

Regina Becker

Leibniz-Institut für Bildungsverläufe, Bamberg

Der Bildungsteilhabe kommt für geflüchtete Jugendliche eine besondere Bedeutung zu. Einerseits für das Ankommen in Deutschland und für soziale Teilhabe, andererseits, weil dadurch wichtige Weichen für den weiteren Berufsweg gestellt werden. Jugendliche, die als Quereinsteiger:innen ins Bildungssystem eingegliedert werden, sind jedoch mit besonderen Herausforderungen konfrontiert. Dazu zählen bspw. geringere Deutschsprachkenntnisse, eine geringere Vertrautheit mit dem Schulsystem und daran anschließenden Bildungsoportunitäten, insbesondere dem Ausbildungssystem. Ein späterer Einstieg ins Bildungssystem erschwert oftmals den Erwerb höherer Schulabschlüsse oder den direkten Übergang in berufliche Ausbildung. Für Ausbildungsvorbereitung, dem Nachholen von Deutschsprachkenntnissen sowie dem Erwerb von allgemeinbildenden Schulabschlüssen stellen berufsvorbereitenden Maßnahmen daher für geflüchtete Jugendliche eine wichtige Bildungsmöglichkeit dar (Braun/Lex, 2016). Insgesamt gibt es eine Vielzahl an Maßnahmen und Unterstützungsangeboten (Wehking, 2020).

Erste Forschungen in diesem Bereich zeigen, dass geflüchtete Jugendliche im Übergangssektor überrepräsentiert sind (Christ, 2023) und insbesondere ältere Jugendlichen aus dem allgemeinbildenden Schulsystem in berufsvorbereitende Maßnahmen einmünden (Becker, 2024). Bisher wissen wir wenig darüber, wie bzw. welche Arten von berufsvorbereitenden Maßnahmen geflüchtete Jugendliche nutzen. Dieses Wissen ist jedoch wichtig, um die Bedeutung und Möglichkeiten des berufsvorbereitenden Sektors für die Teilhabe am Bildungs- und Ausbildungssystem besser zu verstehen. Der vorliegende Beitrag nimmt sich dieser Thematik an und fragt: Welche Arten von Berufsvorbereitungsmaßnahmen besuchen geflüchtete Jugendlichen und von welchen Faktoren hängt der Besuch dieser Maßnahmen ab?

Angelehnt an bisherigen Studien (Menze/Holtmann, 2019), unterscheide ich vier Arten von Maßnahmen: (a) Maßnahmen, die dem Nachholen eines Schulabschlusses dienen, (b) Maßnahmen, mit stärkerer schulischer Komponente, die als erstes Ausbildungsjahr anerkannt werden können, (c) Maßnahmen, die vornehmlich praktische Erfahrungen in Betrieben ermöglichen, sowie (d) Maßnahmen, die sich explizit an Neuzugewanderte und Geflüchtete richten und oftmals insbesondere dem Deutschspracherwerb dienen. Aufbauend auf Menze und Holtmann (2019) kann angenommen werden, dass die Wahrscheinlichkeit eines Besuchs einer bestimmten Maßnahme von Merkmalen der Schulbiografie, kognitiven und sozialen Fähigkeiten und dem sozialen Hintergrund der Jugendlichen abhängt. So kann vermutet werden, dass insbesondere Jugendliche mit besseren Grundvoraussetzungen (z. B. schulische Vorerfahrungen im Herkunftsland, höhere elterliche Bildung, längere Erfahrungen mit dem Schulsystem in Deutschland) eine höhere Wahrscheinlichkeit haben, eine Maßnahme zu besuchen, die das Nachholen eines Schulabschlusses ermöglicht. Jugendliche mit ähnlichen Voraussetzungen, die bereits einen Schulabschluss erworben haben, besuchen mutmaßlich eher Maßnahmen, die als erstes Ausbildungsjahr anerkannt werden können. Jugendliche mit schlechteren schulischen Leistungen, weniger Bildungserfahrung und mit niedrigerer elterlicher Bildung besuchen vermutlich eher andere Maßnahmen. Jugendlichen, die darüber hinaus eher kurze Zeit in Deutschland beschult wurden und geringere Deutschsprachkenntnisse haben, nutzen mit höherer Wahrscheinlichkeit Maßnahmen, die insbesondere Zugewanderte adressieren. Darüber hinaus soll untersucht werden, welche Rolle das Einschulungsalter in Deutschland in Kombination mit der verbleibenden Dauer der Schulpflicht sowie einer nicht alters-adäquaten Einschulung spielt.

Die empirischen Analysen beruhen auf Daten der Panel-Studie „ReGES – Refugees in the German Educational System“ (Will et al., 2021). Die Jugendlichen im Sample sind zwischen 2014 und 2018 in Deutschland angekommen und besuchten zum Zeitpunkt der ersten Befragung die Sekundarstufe I des allgemeinbildenden Schulsystems. Verwendet werden die Daten der Erhebungszeitpunkte 2018, 2019, 2020 und 2022. Die multivariaten Analysen basieren auf gepoolten linearen Wahrscheinlichkeitsmodellen, wobei nur die erste besuchte Berufsvorbereitung berücksichtigt wird. Fehlende Werte wurden multiple imputiert (White et al., 2011).

Vorläufige Analysen weisen nicht darauf hin, dass elterliche Bildung, bessere schulische Leistungen (im Herkunftsland) oder mehr Erfahrungen mit dem deutschen Schulsystem und der Besuch einer Maßnahme, die zu einem Schulabschluss führt, zusammenhängen. Der Besuch einer Maßnahme, die als erstes Ausbildungsjahr anerkannt werden könnte, scheint in signifikant

positiven Zusammenhang mit der bisherigen Beschuldungsdauer und einem höheren Alter zu stehen, wohingegen ein signifikant negativer Zusammenhang mit einem bereits erworbenen Abschluss besteht. Jugendliche, die kürzer in Deutschland leben und im Herkunftsland keine Schule besucht haben, besuchen mit einer höheren Wahrscheinlichkeit eine Maßnahme für Neuzugewanderte.

Integration junger Geflüchteter in die berufliche Ausbildung: Wie zielführend sind berufsbezogene Fördermaßnahmen in Deutschland?

Franziska Meyer

Martin-Luther-Universität Halle-Wittenberg

Fast die Hälfte der Personen, die im Zuge der Flucht migrationsbewegung 2015/16 nach Deutschland kamen, ist zwischen 16 und 30 Jahre alt und somit im ausbildungsfähigen Alter (44%, BAMF, 2017). Um ihnen die Integration in den Ausbildungs- und Arbeitsmarkt zu erleichtern, wurden in den vergangenen Jahren mehrere berufsbezogene Fördermaßnahmen initiiert und ausgebaut, allen voran das vom Europäischen Sozialfonds (ESF) und dem Bundesamt für Migration und Flüchtlinge (BAMF) finanzierte ESF-BAMF-Programm sowie verschiedene Maßnahmen der Bundesagentur für Arbeit (BA). Mittels dieser werden die Teilnehmenden unter anderem über die Relevanz eines Berufsbildungsabschlusses für die Arbeitsmarktintegration in Deutschland informiert und beim Erwerb deutschsprachlicher sowie berufsbezogener Fachkompetenzen unterstützt. Bisherige Forschungsarbeiten haben die Wirksamkeit dieser Maßnahmen überwiegend für Einstiegsprozesse in den Arbeitsmarkt, nicht jedoch in das Berufsbildungssystem untersucht (vgl. Bonin et al., 2021; Brücker et al., 2019, 2020). Ziel des Beitrags ist es, dieses Forschungsdesiderat zu adressieren und der Frage nachzugehen: Ist die Teilnahme an einer berufsbezogenen Fördermaßnahme mit größeren Übergangschancen junger Geflüchteter in eine Berufsausbildung verbunden, und gilt dies in unterschiedlichem Maße für Geflüchtete verschiedener Altersgruppen?

Die Aufnahme einer Berufsausbildung kann als das Ergebnis eines wechselseitigen Entscheidungsprozesses zwischen den potenziellen Ausbildungsplatzbewerber:innen und den Ausbilder:innen gesehen werden (vgl. Sørensen/Kalleberg, 1981; Logan, 1996). Die Geflüchteten erwerben durch die Teilnahme an einer berufsbezogenen Fördermaßnahme Kenntnisse über die Vorteile einer Berufsausbildung gegenüber einem direkten Arbeitsmarkteinstieg, was ihre Entscheidung, sich für einen Ausbildungsplatz zu bewerben, positiv beeinflussen kann. Auf der Ausbilder:innenseite kann die Einstellung von Absolvent:innen der Fördermaßnahmen insofern attraktiv sein, als dass den Teilnehmenden marktfähige Sprach- und Fachkenntnisse vermittelt wurden, was mit einer Stärkung ihrer jeweiligen Wettbewerbsposition einhergeht (Fossati/Liechti, 2020).

Gleichzeitig haben Geflüchtete mit Anfang 20 gegenüber jenen mit Ende 20, ungeachtet einer Teilnahme an einer berufsbezogenen Fördermaßnahme, vermutlich aussichtsreichere Chancen auf eine Ausbildungsstelle: Zum einen wurden viele von ihnen im Rahmen der Berufsschulpflicht für Minderjährige bereits mit dem deutschen Berufsbildungssystem vertraut gemacht. Zum anderen profitieren sie davon, dass Ausbilder:innen bevorzugt jüngere Auszubildende einstellen, da vermutet wird, dass diese leichter zu führen und in das betriebliche Sozialgefüge zu integrieren seien (Imdorf, 2012). Da jüngere Geflüchtete eine vergleichsweise günstigere Wettbewerbsposition einnehmen, kann angenommen werden, dass ihr zusätzlicher Nutzen aus der Teilnahme an einer berufsbezogenen Fördermaßnahme tendenziell geringer ist als für ältere Geflüchtete (deadweight effect, vgl. Brown/Koettl, 2015). Die Teilnahme an einer beruflichen Fördermaßnahme sollte daher zwar mit einer generellen Erhöhung der Übergangschancen von Geflüchteten in die Berufsausbildung assoziiert sein (H1), dies sollte jedoch in einem stärkeren Ausmaß für ältere Geflüchtete gelten (H2).

Die Untersuchung dieser Annahmen erfolgt auf Basis der IAB-BAMF-SOEP-Befragung von Geflüchteten mithilfe eines diskreten Ereigniszeitmodells. Die Stichprobe umfasst Geflüchtete, die zwischen 2013 und 2019 einen Asylantrag gestellt haben, zum Beobachtungszeitpunkt keine Schule besuchen und sich mit 18-30 Jahren im ausbildungsfähigen Alter befinden (N=6.004). Ihre Zeit bis zum Übergang in die Berufsausbildung (duration time) wird für den Zeitraum von 2016 bis 2020 analysiert. Personen, die im Beobachtungszeitraum ein Studium aufgenommen haben, wurden in einem Robustnesscheck aus dem Sample ausgeschlossen.

Deskriptiv wird ersichtlich, dass 14% des Samples im Untersuchungszeitraum eine Berufsausbildung aufgenommen haben. Die multivariaten Analysen zeigen, dass ältere gegenüber jüngeren Geflüchteten tatsächlich signifikant schlechtere Chancen haben, eine Berufsausbildung aufzunehmen. Fernerhin ist sowohl die Teilnahme am ESF-BAMF-Programm als auch an den BA-Maßnahmen signifikant positiv mit den Übergangschancen der 25-30-jährigen Geflüchteten in die Berufsausbildung assoziiert, nicht jedoch mit jenen der 18-24-jährigen.

Angesichts der Relevanz eines Berufsbildungsabschlusses für eine günstige Arbeitsmarktpositionierung in Deutschland (Hausner et al., 2015) sowie der vergleichsweise ungünstigen Chancen älterer Geflüchteter auf einen Ausbildungsplatz, kann dieses Ergebnis als klarer Erfolg der untersuchten Fördermaßnahmen erachtet werden. Hiervon ausgehend kann somit ausdrücklich für einen Ausbau der bestehenden berufsbezogenen Förderstrukturen für Geflüchtete plädiert werden, insbesondere für die Gruppe der älteren jungen Erwachsenen.

Betriebliche Ausbildung Geflüchteter: Analyse der Chancen auf Ausbildung auf Basis des BIBB-Betriebspanels

Christian Gerhards

Bundesinstitut für Berufsbildung (BIBB), Bonn

Ein Ausbildungsabschluss ist der zentrale Schlüssel für eine gelingende Integration in Deutschland und in den Arbeitsmarkt (Granato et al., 2006; Becker, 2011). Bislang gibt es kaum Daten, die es erlauben, die Rolle von betrieblichen Rahmenbedingungen für Einstellungschancen von Geflüchteten in Ausbildung zu untersuchen. Ziel des Beitrags ist es, durch eine multivariate Panelanalyse des BIBB-Betriebspanels Einflüsse verschiedener betrieblicher Faktoren für die Ausbildungschancen Geflüchteter zu identifizieren und besser zu verstehen.

Die bisherige Forschung zeigt seit 2015 einen deutlichen Anstieg von Betrieben mit geflüchteten Auszubildenden, insbesondere aufgrund von Asylsuchenden aus Syrien und der Ukraine (Gerhards, 2022; Bundesamt für Migration und Flüchtlinge, 2021; Kosyakova/Brücker, 2024). Analysen zur betrieblichen Ausbildung Geflüchteter zeigen geringere Einstellungschancen für Geflüchtete in bestimmten Bereichen, etwa bei unternehmensnahen Dienstleistungen oder in nicht IHK- oder Handwerkskammer-gebundenen Betrieben (Gerhards, 2022). Unklar bleiben aber bisher Gründe für solche Unterschiede.

Im Beitrag werden Tätigkeiten in und Ausbildungsmotive von Betrieben in den Fokus gerückt. Gemäß des Task-based approach (Autor et al., 2003) werden betriebliche Tätigkeiten und deren Wandel als entscheidend für die Deckung des Qualifikationsbedarfs betrachtet, wobei zwischen analytisch-kommunikativen, manuellen und Routinetätigkeiten unterschieden wird. Hierbei wird

erwartet, dass Sprachkenntnisse bei manuellen und Routinetätigkeiten weniger bedeutsam sind, was Geflüchteten bessere Einstiegschancen bietet (Acemoglu/Autor, 2011). Im Rahmen des soziologischen Neoinstitutionalismus werden Ausbildungsmotive wie das Produktionsmotiv hervorgehoben, welches auf kurzfristigen Nutzen abzielt, im Gegensatz zum Investitionsmotiv, das einen langfristigen Verbleib der Auszubildenden anstrebt (Pfeifer et al., 2009; Merrilees, 1983; Muehlemann et al., 2007). Geflüchtete könnten aufgrund rechtlicher Unsicherheiten besonders für Betriebe interessant sein, die nach dem Produktionsmotiv handeln.

Neben diesen organisationssoziologischen Ansätzen bleiben ökonomische Überlegungen relevant. Betriebe könnten die Ausbildung Geflüchteter aufgrund von Sprachbarrieren oder weniger vergleichbaren Abschlüssen als riskanter betrachten. Es wird erwartet, dass Betriebe mit Stellenbesetzungsproblemen eher geneigt sind, Geflüchtete auszubilden (Sung, 2010). Das geforderte Qualifikationsniveau spielt ebenfalls eine Rolle, wobei höhere Anforderungen Betriebe dazu veranlassen könnten, (qualifiziertere) Bewerbende ohne Fluchthintergrund zu bevorzugen (Troltsch/Walden, 2012; Troltsch et al., 2014).

Die Analyse basiert auf Daten des BIBB-Betriebspanels, das seit 2011 jährlich in ca. 3.500 deutschen Betrieben durchgeführt wird, und untersucht als abhängige Variable die Einstellung von Geflüchteten in Ausbildung von 2015 bis 2023.

Für den Fokus auf Ausbildungsmotive der Betriebe und benötigte Tätigkeiten, werden diese zunächst faktoranalytisch zusammengefasst. Die in den Betrieben benötigten Tätigkeiten werden – entsprechend des oben skizzierten Ansatzes – in analytisch-kommunikative Tätigkeiten, solche, die Fingerfertigkeit und Geschick erfordern, sowie Routinetätigkeiten eingeteilt. Bei den Ausbildungsmotiven wird zwischen einem Produktionsmotiv (Einsatz bereits während der Ausbildung) und einem Investitionsmotiv unterschieden (langfristige Entwicklung eigener Fachkräfte).

Zudem werden die Chancen auf dem Ausbildungsstellenmarkt durch die Betrachtung von Neueinstellungen, Bewerbungen von geflüchteten und nicht-geflüchteten Auszubildenden sowie Stellenbesetzungsproblemen analysiert. Kontrollvariablen umfassen das Tätigkeitsniveau im Betrieb, Betriebsgröße, Branchenzugehörigkeit und Standort unterteilt nach Ost- und West-Deutschland. Die statistische Analyse erfolgt mittels einer Hybrid-Panelregression (Perales/Schunck, 2016).

Die Studie zeigt, dass Betriebe mit einem Anstieg manueller Tätigkeiten eher bereit sind, Geflüchtete auszubilden. Interpretation wäre, dass diese Tätigkeiten weniger Kundenkontakt und geringere Deutschsprachkenntnisse erfordern. Bei den Ausbildungsmotiven ergibt sich ein schwacher positiver Effekt beim Produktionsmotiv; möglicherweise fallen bei diesen kurzfristiger orientierten Betrieben mögliche asylrechtliche Unsicherheiten weniger ins Gewicht. Die Untersuchung der Ausbildungschancen von Geflüchteten offenbart, dass diese in Betrieben mit vielen Bewerbungen schlechtere Chancen haben und – überraschenderweise – bei zunehmenden Stellenbesetzungsproblemen seltener eingestellt werden.

Die Kontrolle über die Jahre hinweg zeigt bessere Einstellungschancen in den Jahren 2018 und 2019, also vor der Corona-Pandemie und dem Anstieg der Geflüchtetenzahlen seit dem Ukraine-Krieg. Die Hybridmodelle verdeutlichen, dass größere Betriebe und der Sektor der personennahen Dienstleistungen (z. B. Verkehr, Beherbergung & Gastronomie, Erholung) höhere Einstellungschancen bieten.

Insgesamt verdeutlichen die Ergebnisse, dass neben strukturell-marktwirtschaftlichen Gründen auch Tätigkeiten und Ausbildungsmotive der Betriebe entscheidend für die betriebliche Entscheidung zur Integration von Geflüchteten in den Ausbildungsmarkt sind.

S-105: Prozessuale Perspektiven auf Lehrkraftmotivation und Unterricht

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 145

Chair der Sitzung: Markus Dresel

Chair der Sitzung: Rebecca Lazarides

Prozessuale Perspektiven auf Lehrkraftmotivation und Unterricht

Chair(s): **Markus Dresel** (Universität Augsburg Deutschland), **Rebecca Lazarides** (Universität Potsdam)

Der Motivation von Lehrkräften gilt als zentrale Voraussetzung für effektives Unterrichten. Während dies durch die Befundlage mittlerweile gut bestätigt ist, ist allerdings noch weitgehend ungeklärt, über welche Prozesse sie sich auf den Unterricht auswirkt. Das Symposium greift dieses Desiderat auf und bündelt Beiträge aus der Pädagogischen Psychologie, Erwachsenenbildung und Unterrichtsforschung. Gemeinsam verfolgen sie das Ziel, aus einer prozessbezogenen Perspektive die instruktionalen Konsequenzen multipler Dimensionen der Lehrkraftmotivation aufzuklären.

In Beitrag 1 wird der emotionale Gehalt in der Sprache von Lehrkräften untersucht und gefragt, wie dieser mit selbstberichtetem Enthusiasmus im Zusammenhang steht. Berichtet werden Befunde einer Befragungs- und Beobachtungsstudie mit Mathematiklehrkräften und ihren Schüler*innen. Die Befunde indizieren, dass der Enthusiasmus von Lehrkräften auf ihren Sprachgebrauch messbaren Einfluss hat.

In Beitrag 2 wird der Frage nachgegangen, ob Lehrkräfte Enthusiasmus spezifisch für einzelne Lernende erleben und ausdrücken. DaZ-Lehrkräfte schätzten dazu ihr Erleben in Bezug auf einzelne Lernenden ein und wurden in der Interaktion mit ihnen videografiert. Die Ergebnisse verweisen auf nennenswerte intraindividuelle, dyadenspezifische Variation im Enthusiasmus.

In Beitrag 3 werden attributionale Prozesse von Lehrkräften adressiert und untersucht, wie diese die emotionalen Reaktionen sowie Feedback- und Hilfeverhalten von Lehrkräften beeinflussen. Zwei Vignettenstudien mit Lehrkräften und Lehramtsstudierenden hierzu verweisen darauf, dass hierbei insbesondere die Zuschreibung eigener Kontrolle und Verantwortlichkeit durch die Lehrpersonen relevant ist.

Schließlich nähert sich Beitrag 4 der Frage an, wie Ziele von Lehrenden deren Aufmerksamkeit leiten. Die Ergebnisse einer Eye-Tracking-Studie, bei der Lehramtsstudierende ein Video eines Unterrichtsgesprächs im Fach Mathematik betrachteten, indizieren, dass sich Lehrende zahlreiche Ziele spezifisch für die einzelnen Schüler*innen setzen und diese die Aufmerksamkeitsallokation steuern.

Die Beiträge des Symposiums werden anhand von Leitfragen zur Konvergenz der Ergebnisse, zu den Implikationen für die theoretische Konzeption von Lehrkraftmotivation und deren Wirkungen auf der Prozessebene des Unterrichts sowie zu den methodischen Innovationen strukturiert mit den Beitragenden und Teilnehmenden diskutiert.

Beiträge des Symposiums

Die Macht der Worte: Der Enthusiasmus von Lehrkräften und ihre natürliche Sprache im Unterricht – eine Sentimentanalyse

Hannah Kleen¹, Anne Frenzel², Anton K. G. Marx², David F. Sachs², Franziska Baier-Mosch³, Mareike Kunter⁴

¹DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, ²Department of Psychology, Ludwig-Maximilians-Universität München, ³DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation & Universität Siegen, ⁴DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation & Goethe Universität Frankfurt am Main

Der Enthusiasmus von Lehrkräften wirkt sich nachweislich positiv auf die Motivation und die Leistungen von Schüler*innen aus (Keller et al., 2016; Moè et al., 2021). Enthusiastischere Lehrkräfte zeigen eine bessere Unterrichtsqualität (Fauth et al., 2019; Kunter et al., 2008), was sich nachfolgend auch auf die Ergebnisse der Schüler*innen auswirkt (Feng et al., 2023; Kunter et al., 2013). Obwohl die Annahme, dass begeisterte Lehrkräfte besser unterrichten, naheliegend ist, gibt es kaum theoretisch fundierte Erklärungen und empirische Belege dafür, wie sich der Enthusiasmus der Lehrkräfte in ihrem tatsächlichen Verhalten im Klassenzimmer manifestiert (Keller et al., 2016).

Im vorliegenden Beitrag beleuchten wir einen potenziellen Mechanismus, durch den die Begeisterung der Lehrkräfte zu einem qualitativ hochwertigen Unterricht führen kann: Der emotionale Gehalt (Sentiment) in der Sprache der Lehrkräfte. Sentiment, das auch als affektive Bedeutung eines Textes bezeichnet wird, wird in Bezug auf die Valenz (positiv vs. negativ) und die Intensität (mehr oder weniger positiv/negativ, z.B. auf einer Skala von -1 [sehr negativ] bis 1 [sehr positiv]; Taboada et al., 2011) quantifiziert. Bisherige Forschung zu Sentimentanalysen im Bildungsbereich wurde zumeist in Massive Open Online Courses und nicht in tatsächlichen Klassenzimmern durchgeführt (Zhou & Ye, 2020). Bezüglich des Sentiments der Lehrkräftesprache und des Enthusiasmus der Lehrkräfte, postulieren wir, dass enthusiastischere Lehrkräfte in ihrer natürlichen Sprache, wie sie im Klassenzimmer verwendet wird, ein positiveres Sentiment ausdrücken. Darüber hinaus verwendeten wir einen explorativen qualitativen Ansatz, um zu untersuchen, welche positiv konnotierten und Begeisterung vermittelnden Lehrkräfte verwenden, um ihre Schüler*innen zu begeistern.

Die Datengrundlage der vorliegenden Studie sind transkribierte Videoaufzeichnungen sowie Fragebogendaten von Unterrichtsstunden von N = 19 Mathematiklehrkräften und deren N = 393 Schüler*innen der Sekundarstufe. Es wurden die Selbstschätzungen der Lehrkräfte zu ihrem trait Enthusiasmus (5 Items, z.B. "In dieser Klasse unterrichte ich mit Begeisterung"; Frenzel et al., 2016; Taxer & Frenzel, 2018) sowie dem in-situ Enthusiasmus in der videografierten Unterrichtsstunde (2 Items, z. B. "In dieser Schulstunde hat mir das Unterrichten Spaß gemacht") erfasst. Außerdem bewerteten die Schüler*innen den wahrgenommenen Enthusiasmus ihrer Lehrkräfte ebenfalls generell (5 Items, z.B. "Meine Mathelehrkraft scheint großen Spaß am Unterrichten zu haben"; Taxer & Frenzel, 2018) als auch in der beobachteten Stunde (2 Items, z.B. „In dieser Schulstunde hat meine Lehrkraft mit Begeisterung unterrichtet“). Das Sentiment – Valenz und Intensität – der Lehrkräftesprache wurde mittels Sentimentanalysen ermittelt, wofür wir das SentiWS-Lexikon (Remus et al., 2010) verwendeten.

Für die Analyse berechneten wir Korrelationen zwischen den Enthusiasmuseinschätzungen der Lehrkräfte und Schüler*innen und dem Sentiment der Lehrkräftesprache. In Übereinstimmung mit unserer Hypothese gab es mittelgroße positive Zusammenhänge zwischen dem Sentiment während der aufgezeichneten Unterrichtsstunde und dem in-situ-Enthusiasmus nach der Unterrichtsstunde (Kendalls tau = .44, p = .011) als auch den Selbstberichten zum trait Enthusiasmus beim Unterrichten (Kendalls tau = .38, p = .029). Lehrkräfte, die sich selbst als enthusiastischer beschrieben, verwendeten mehr positiv konnotierte Wörter in ihrem Unterrichtsgespräch. Im Gegensatz zur Hypothese gab es jedoch keine konsistenten Zusammenhänge zwischen dem Sentiment der Lehrkräfte und dem von den Schüler*innen eingeschätzten Enthusiasmus der Lehrkräfte.

Zusätzliche explorative Analysen zeigten, dass von allen bewerteten Wörtern 18,2 % eine negative und 81,8 % eine positive Valenz aufwiesen. Darüber hinaus wurden von den Lehrkräften häufig positive Wörter wie "gut", "genau", "richtig", "nett" oder "einfach" verwendet. Von den Wörtern mit einer besonders hohen SentiWS-Valenz wurden "perfekt" und "super" am häufigsten von den Lehrkräften verwendet, allerdings weniger häufig als die zuvor genannten Wörter. Nicht jede Lehrkraft benutzte diese Wörter, stattdessen sagten diejenigen, die Wörter wie "super" verwendeten, diese häufiger.

Unsere Ergebnisse zeigen insgesamt, dass der Enthusiasmus der Lehrkräfte in ihrem Sprachgebrauch gegenüber Schüler*innen einen messbaren Einfluss hat. Die vorliegenden Ergebnisse müssten mit weiteren Datenquellen repliziert und idealerweise auf verschiedene Unterrichtskontexten ausgeweitet werden.

Sieht man meine Begeisterung für den Unterricht mit dir? Wie Lehrkräfte Enthusiasmus für jeden Lernenden erleben und in 1-zu-1-Situationen ausdrücken

Anne Kosubek¹, Hanna Gaspard¹, Hannes Schröter², Fani Lauermann³

¹Institut für Schulentwicklungsforschung, TU Dortmund, ²Deutsches Institut für Erwachsenenbildung, ³Bonner Zentrum für Lehrerbildung, Universität Bonn

Theoretischer Hintergrund:

Im Lern- und Bildungskontext wird ein positiver Zusammenhang der Motivation von Lernenden sowohl mit dem erlebten Unterrichtsenthusiasmus von Lehrkräften (z.B. Keller et al., 2018; Lazarides & Schiefele, 2021) als auch mit dem Ausdruck dieses Lehrkraftenthusiasmus im Unterricht angenommen (z.B. Frenzel et al., 2009; Patrick et al., 2000). Allerdings wird der Unterrichtsenthusiasmus von Lehrkräften nicht gleichermaßen erlebt und ausgedrückt, sondern variiert beispielsweise zwischen Unterrichtsstunden (Gaspard & Lauermann, 2021) und Lerngruppen (Frenzel et al., 2020). Ebenso sind Variationen des erlebten Lehrkraftenthusiasmus für das Unterrichten jedes einzelnen Lernenden insbesondere in heterogenen Lerngruppen höchstwahrscheinlich (ten Hagen, 2022). Allerdings sind Untersuchungen des Lehrkraftenthusiasmus in sprachlich- und kulturell-heterogenen Lerngruppen generell rar (Hachfeld et al., 2012; Keller et al., 2016), liegen lediglich auf Klassenebene vor (Kosubek et al., im Druck) oder betrachten die Varianz des Unterrichtsenthusiasmus von Lehrkräften für jeden Lernenden und dessen Zusammenhang mit der Lernendenmotivation fokussiert auf eine Komponente des Unterrichtsenthusiasmus (ten Hagen, 2022).

Fragestellungen:

Daher soll in der vorliegenden Untersuchung den folgenden Fragestellungen nachgegangen werden:

1. Inwiefern besteht ein Zusammenhang zwischen dem intraindividuell-variiierenden Erleben und dem intraindividuell-variiierenden Ausdruck des Unterrichtsenthusiasmus von Lehrkräften in heterogenen Lerngruppen?
2. Inwiefern besteht ein Zusammenhang zwischen dem intraindividuell-variiierenden Erleben, dem intraindividuell-variiierenden Ausdruck des Unterrichtsenthusiasmus von Lehrkräften und der Motivation von Lernenden in heterogenen Lerngruppen?

Methode:

Die Untersuchung basiert auf Querschnittdaten von 33 Lehrkräften und 281 Lernenden in Vorbereitungsklassen (VK, Sekundarstufe I in Regelschulen) mit Deutsch als Zweitsprache (DaZ) sowie von 35 Lehrkräften und 329 Lernenden im DaZ-Unterricht von Integrationskursen (IK, Erwachsenenbildung) in Deutschland. Die Lehrkräfte schätzten ihr intraindividuell-variiierendes Erleben des Unterrichtsenthusiasmus für jeden Lernenden ein. Zudem machten die Lernenden Angaben zu ihrer intrinsischen Motivation für den Deutschunterricht. Die Erfassung des intraindividuell-variiierenden Ausdrucks des Unterrichtsenthusiasmus der Lehrkräfte erfolgt durch standardisierte continuous moment-by-moment Beobachtungen anhand von Videosequenzen einer Unterrichtsstunde. Externe Beobachtungspersonen schätzten den Ausdruck mit Hilfe von acht Indikatoren ein: (1) Stimme, (2) Augen/Blickkontakt, (3) Gestik, (4) Körperbewegungen im Raum, (5) Gesichtsausdruck, (6) Wortwahl, (7) Akzeptanz der Ideen und Gefühle der Lernenden, (8) allgemeines Energielevel der Lehrkräfte (Collins, 1978). Diese Einschätzungen basieren auf Beobachtungen von direkten 1-zu-1-Kommunikationssituationen der Lehrkraft mit einem Lernenden, welche aus den vollständigen Unterrichtssequenzen extrahiert wurden. Aktuell liegen Kodierungen von 8 Klassen vor. Die Zusammenhänge wurden mittels hierarchisch-linearer Regressionsanalysen (Raudenbush & Bryk, 2010) untersucht.

Ergebnisse und ihre Bedeutung:

Es fand sich in den IK ein positiver linearer Zusammenhang zwischen dem intraindividuell-variiierenden Erleben des Unterrichtsenthusiasmus der Lehrkräfte für jeden Lernenden und der intrinsischen Motivation dieses Lernenden für den Deutschunterricht ($\beta = .140$, $p = .049$). Somit scheinen Lernende von Lehrkräften, die ein höheres Maß an Enthusiasmus für den Unterricht mit ihnen empfinden, stärker intrinsisch motiviert für den Deutschunterricht. In den Regelschulen zeigten sich ähnliche Tendenzen ($\beta = .097$, $p = .074$).

Erste Ergebnisse zum intraindividuell-variiierenden Ausdruck des Unterrichtsenthusiasmus von Lehrkräften anhand von 8 Klassen (VK: 3 Lehrkräfte & 33 Lernende; IK: 5 Lehrkräfte & 48 Lernende) basieren auf Beobachtungen von 1-zu-1-Kommunikationssituationen mit mindestens 5 Sekunden Dauer, die in Summe etwa 42% einer 48-minütigen Unterrichtsaufzeichnung ausmachten. Davon verbrachte jede Lehrkraft im Durchschnitt etwa 2 Minuten aufgeteilt auf etwa 8 Sequenzen im Laufe der Unterrichtsstunde mit jedem Lernenden der Lerngruppe in einer 1-zu-1-Situation. Bisherige Analysen deuten darauf hin, dass Lehrkräfte im Vergleich zum selbstberichteten Erleben (VK: $M = 5.98$, $SD = 1.18$; IK: $M = 6.16$, $SD = 0.97$) eher wenig Unterrichtsenthusiasmus in 1-zu-1-Kommunikationssituationen ausdrückten (VK: $M = 2.40$, $SD = 0.81$; IK: $M = 2.31$, $SD = 0.83$).

Bis zur Tagung werden vollständige Analysen des intraindividuell-variiierenden Ausdrucks des Unterrichtsenthusiasmus der Lehrkräfte auf Basis der gesamten Stichprobe, die Einschätzung einer weiteren Beobachtungsperson sowie hierarchisch-lineare Regressionsmodelle zur Berechnung der Zusammenhänge mit der intrinsischen Motivation der Lernenden für den Deutschunterricht vorliegen.

Teachers' Self-Directed Attributions, Professional Responsibility, and Classroom Interventions: Connecting Inter- and Intrapersonal Perspectives in Attribution Theory

Fani Lauermann, Daria K. Benden

Bonner Zentrum für Lehrerbildung, Universität Bonn

Theoretical Background:

Teachers' attributions for classroom outcomes like student achievement (i.e., their subjective explanations for the underlying causes of these outcomes) and their ascriptions of responsibility influence their help-giving intentions and classroom intervention

strategies (Reyna & Weiner, 2001; Twardawski et al., 2020; Weiner, 2000). According to Weiner's attribution theory (2000), teachers—when confronted with unexpected or adverse events such as student failure—conduct an attributional search focusing on why the student failed. If students fail due to internal and controllable causes (e.g., lack of effort), they are considered personally responsible, leading teachers to feel anger and adopt punitive measures (e.g., harsh grading). If students fail due to internal but uncontrollable factors (e.g., lack of aptitude despite effort), teachers tend not to hold students responsible, leading to feelings of sympathy and increased willingness to provide supportive feedback. Thus, teachers' causal attributions and responsibility judgments can affect their emotional reactions and classroom intervention strategies.

Although much research has focused on teachers' attributions and ascriptions of responsibility to students, the role of teachers' self-directed attributions and responsibility for classroom outcomes is relatively unexplored. Drawing on Weiner's attribution theory, we conducted two vignette studies on teachers' student- and self-directed attributions and responsibility judgments, emotional reactions, and classroom intervention strategies. We expected that teachers' self-directed responsibility will predict feelings of guilt and willingness to help, beyond student-directed beliefs.

Method:

Study 1 included secondary school preservice (n=224) and in-service (n=134) teachers. In a within-person design, they responded to two vignettes adapted from Reyna and Weiner (2001) describing a student who failed an exam either due to (1) lack of effort despite high ability (i.e., internal/controllable cause) or (2) lack of ability despite effort (i.e., internal/uncontrollable cause). The vignettes also described the teachers' behavior, who prepared the whole class for the exam, but did not undertake any specific measures to support the failing student.

Study 2 was an experimental study with 179 preservice teachers who responded to Vignette (1) from Study 1 after being randomly assigned to one of three conditions. Participants in the two experimental conditions generated counterfactual should statements that focused on (a) teacher responsibility or (b) student responsibility for the failure, whereas participants in the control group described their opinion of the study's topic.

Participants in both studies rated student/teacher responsibility, emotional reactions (anger, sympathy, guilt), feedback goals, and willingness to help. Participants in Study 2 described up to three strategies they could use to help the failing student. These strategies were categorized by two independent coders: teaching adaptations, outsourcing the problem to the student or others, and controlling practices.

Results & Discussion:

Path analyses for both vignettes in Study 1 indicated that student responsibility positively predicted anger, which positively predicted endorsement of punitive feedback and negatively willingness to help. This pattern is consistent with attribution theory. Expanding upon Reyna and Weiner (2001), we found that teacher responsibility predicted feelings of guilt in both vignettes and was the strongest predictor of willingness to help. Accordingly, normative expectations (teachers are responsible for providing help) emerged as a key factor shaping teachers' willingness to help.

ANOVAs in Study 2 showed that preservice teachers' emotional reactions and willingness to help were comparable across conditions. However, participants developed vastly different help-giving strategies. Focusing on teacher responsibility led to strategies involving teaching adaptations, whereas focusing on student responsibility led to strategies that outsourced the problem to others (e.g., the student, parents).

Thus, even when teachers' attributions lead to a higher willingness to help, their personal versus student-directed responsibility can lead them to qualitatively different classroom intervention strategies.

Wie Ziele von Lehrpersonen die Aufmerksamkeitsverteilung im Klassenzimmer lenken: Eine Eye-Tracking Studie zur Bedeutung von schüler*innenbezogenen Zielen für die Wahrnehmung von Unterrichtsinteraktionen

Ricardo Böheim¹, Martin Daumiller², Aldin Alijagic², Andreas Gegenfurtner², Doris Lewalter¹, Tina Seidel¹, Markus Dresel²

¹Technische Universität München, ²Universität Augsburg

Hintergrund:

Zielen von Lehrpersonen wird eine bedeutende Rolle für Unterrichtsqualität zugeschrieben. Befunde der Lehrkraftmotivationsforschung indizieren, dass schüler*innenbezogene Ziele (d.h. Ziele, die sich Lehrpersonen für ihre Schüler*innen setzen) professionelles Handeln von Lehrpersonen beeinflussen (Daumiller et al., 2022). Aus kognitionspsychologischer Perspektive wird angenommen, dass Ziele Wahrnehmung und Informationsverarbeitung durch eine selektive Ausrichtung der visuellen Aufmerksamkeit steuern (Seidel et al., in press). Das bewusste Lenken der Aufmerksamkeit auf bestimmte, lernrelevante Interaktionen wird als Noticing bezeichnet (Seidel & Stürmer, 2014), wobei die Erhebung der Blickbewegungen von Lehrpersonen mittels Eye-Tracking zunehmend an Bedeutung gewinnt (Keskin et al., 2024). Unterschiede im Noticing werden bislang häufig auf Wissensunterschiede zurückgeführt, während die Rolle der Lehrkräftemotivation bislang kaum berücksichtigt ist (Kosel et al., 2023).

Ziele sind ein zentraler Aspekt der Lehrkräftemotivation und lenken den Fokus auf persönlich bedeutsame und relevante Ereignisse – hier: im Unterricht (Dijksterhuis & Aarts, 2010). Hierfür sind endogene (top-down) Aufmerksamkeitsprozesse relevant: diese leiten, abhängig von ihrer Relevanz für die Zielerreichung, die Menge und Dauer der Aufmerksamkeit, die eingehenden Informationen gewidmet wird (Corbetta & Shulman, 2002). Der Inhalt der Aufmerksamkeit repräsentiert die Ziele, die zu einem bestimmten Zeitpunkt aktiv sind, und Aufmerksamkeit übersetzt Ziele in Verhalten (Monsell & Driver, 2000). Ziele von Lehrpersonen können sich dabei auf einzelne Schüler*innen oder die gesamte Klasse beziehen und unterschiedliche Inhalte haben (z.B. Schüler*innen sollen Wissen erweitern, sozial eingebunden werden oder gute Leistungen erzielen). Überraschenderweise wurden die Auswirkungen von Zielen auf Aufmerksamkeitsprozesse in pädagogischen Kontexten bislang kaum analysiert (Ausnahme siehe Kücherer et al., 2021); eine Lücke, die wir mit der vorliegenden Studie schließen, in der wir Noticing als prozessuale und kognitive Konsequenz von Lehrkraftmotivation im Unterricht betrachten.

Fragestellungen:

Die vorliegende Eye-Tracking Studie, bei der Lehramtsstudierende eine Videovignette betrachten, hatte zum Ziel, Hinweise dazu zu liefern, inwiefern schüler*innen- und klassenbezogene Ziele von Lehrpersonen mit deren Aufmerksamkeitsverteilung beim Beobachten einer Unterrichtssituation zusammenhängen. Konkret:

1. Welche Ziele verfolgen Lehrpersonen für ihre Schüler*innen in einer spezifischen, beobachteten Unterrichtssituation?

2. Inwiefern zeigen sich Zusammenhänge dieser Ziele mit der visuellen Aufmerksamkeit von Lehrpersonen (Anzahl und Dauer der Fixationen auf einzelne Schüler*innen)?

Methode:

An der Studie nahmen 51 Lehramtsstudierende (73 % weiblich) teil. Zunächst wurde ihnen ein Ausschnitt (1.5 min) eines videografierten Unterrichtsgesprächs einer 11. Klasse im Mathematikunterricht gezeigt. Anschließend wurden sie zu den Zielen befragt, die sie dort als Klassenlehrkraft verfolgen würden. Danach beobachteten die Teilnehmenden den weiteren Verlauf der Unterrichtsszene während ihre Blickbewegungen mittels Eye-Tracking aufgezeichnet wurden. Zur Analyse der Blickbewegungsparameter wurden für alle Schüler*innen im Video separate Areas of Interest (AOIs) definiert. Die Kodierung der transkribierten Ziele erfolgte anhand eines adaptierten Kodierschemas zu Zielorientierungen von Daumiller und Dresel (2020), mit vier Rater*innen (Beurteilerübereinstimmung: kappa=.78; Nicht-Übereinstimmungen konsensuell geklärt). Für die Analyse der Zusammenhänge von Zielen mit Dauer und Häufigkeit ganzer Fixation auf AOIs wurden lineare Mehrebenenmodelle geschätzt.

Ergebnisse und Bedeutung:

Die Ergebnisse erbrachten, dass Lehramtsstudierende trotz des kurzen Einblicks in das Unterrichtsvideo bereits eine Reihe von Zielen für einzelne Schüler*innen ($M=2.1$, $SD=2.1$) als auch die gesamte Klasse ($M=2.7$, $SD=1.2$) formulierten. Hinsichtlich der Zielinhalte waren Ziele zu den Lernaufgaben, zur Beziehungsqualität sowie zu Verhaltensregeln und Disziplin der Schüler*innen dominant. Diese Ziele waren von Bedeutung für die Aufmerksamkeitsverteilung: Schüler*innenbezogene Ziele prädizierten Anzahl ($\beta=.17$) und Dauer ($\beta=.21$) der Fixationen auf den/die jeweilige/n Schüler/in. Für klassenbezogene Ziele lag bei beiden Metriken eine negative Ergebnistendenz vor, was für eine hierbei stärker bottom-up gesteuerte Wahrnehmung spricht.

Zusammenfassend sprechen die Ergebnisse (1) für die Nützlichkeit und Brauchbarkeit des Paradigmas zur Untersuchung von Motivationseffekten auf die der Unterrichtswahrnehmung sowie (2) die Bedeutung der Ziele von Lehrpersonen für die situative Steuerung der visuellen Aufmerksamkeit in unterrichtlichen Settings. Die Studie leistet somit einen Beitrag zur Aufklärung aufmerksamkeitsbezogener Prozesse als der Transmission von Lehrkräftemotivation in Unterrichtsqualität.

S-106: Unterrichtsqualität zwischen Interdisziplinarität, Intersubjektivität und interaktivem Wissensaustausch

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 151

Chair der Sitzung: Erin Gerlach

Chair der Sitzung: Alfred Lindl

Chair der Sitzung: Wiebke Langer

Chair der Sitzung: Mirjam Steffensky

Unterrichtsqualität zwischen Interdisziplinarität, Intersubjektivität und interaktivem Wissensaustausch

Chair(s): **Erin Gerlach** (Universität Hamburg), **Alfred Lindl** (Universität Regensburg), **Wiebke Langer** (Universität Hamburg)

Diskutant*in(nen): **Mirjam Steffensky** (Universität Hamburg)

Die Qualität von Unterricht ist eines der am lebendigsten diskutierten Themen der aktuellen Bildungsforschung (z. B. Reusser, 2024; Rothland, 2024; Kubsch et al., 2023). Vor dem Hintergrund fachspezifischer Belange und von Übersichtsarbeiten vornehmlich aus der Mathematikdidaktik wurden in der jüngeren Vergangenheit bestehende theoretische Konzeptualisierungen wie beispielsweise die drei (vermeintlich) generischen Basisdimensionen (Praetorius et al., 2018) reflektiert und teilweise restrukturiert oder ergänzt. In diesem Zug wurden sowohl generische Merkmale, die in diesem Forschungskontext bislang noch kaum Beachtung fanden (z. B. Auswahl und Präsentation von Inhalten und Fachmethoden) als auch solche Facetten in Erwägung gezogen, die eine Fachspezifik aufweisen oder zumindest durch eine fachliche Akzentuierung charakterisiert sind (z. B. Potenzial zur kognitiven, ästhetischen oder bewegungsbezogenen Aktivierung; Lindl et al., 2024; Praetorius & Nehring, 2020). Diese wiederum fanden Eingang in reformierte Rahmenmodelle zur Unterrichtsqualität (z. B. Praetorius & Charalambous, 2020 oder das MAIN-TEACH-Modell von Charalambous & Praetorius, 2020).

Insbesondere unter einer fachdidaktischen Perspektive wurde die Forderung deutlich, dass sich Unterrichtsqualität nur dann angemessen modellieren lässt, wenn (1) bei der Konzeptualisierung und Einschätzung fachlich gefärbter Merkmale fachdidaktische Expertise einbezogen wird und die fachspezifischen Aspekte die zugrunde liegenden Ziele, Inhalte und Methoden adressieren (Begrich et al., 2023; Praetorius et al., 2020; Reusser, 2024). Zudem ist (2) weiterhin unklar, in welcher Form Unterrichtsqualität von verschiedenen in den Unterrichtsprozess involvierten Akteur:innen konzeptualisiert wird (z. B. Lehrkräfte, Schüler:innen, Fachdidaktiker:innen) und wer diese anhand welcher Kriterien angemessen zu beurteilen vermag (z. B. Lehrkräfte, Schüler:innen, Beobachter:innen). Das Anwendungspotenzial derartiger Konzeptualisierungen ist (3) schließlich vielfältig und reicht von der empirischen Messung, dem systematischen Monitoring sowie der Sicherung eines qualitativ hochwertigen Unterrichtsangebots im Sinne eines weiten Forschungstransfers bis zur evidenzbasierten Reflexion und Professionalisierung von Lehrkräften. Diese drei Aspekte finden sich in den auch in den Leitfragen des Symposiums wieder:

1. Inwieweit können Merkmale von Unterrichtsqualität als fachspezifisch oder generisch gelten (*Interdisziplinarität*)?
2. Wie können Unterrichtsprozesse perspektivenspezifisch valide konzeptualisiert und beurteilt werden (*Intersubjektivität*)?
3. Wie lassen sich empirisch fundierte Aspekte gewinnbringend in einen wechselseitigen Transfer mit der pädagogischen Praxis einbringen (*interaktiver Wissensaustausch*)?

Im vorliegenden Symposium vergleichen die zwei Forschungsgruppen FALKO-PV (Fachspezifische Lehrkraftkompetenzen - Prädiktive Validierung, Universität Regensburg) und QualiTePE (Quality of Teaching in Physical Education, Pädagogische Hochschule Zürich, Universität Hamburg & Universität Luxemburg) in einem gemeinsamen und zwei projektspezifischen Vorträgen ihre unterschiedlichen Zugänge, Antworten auf diese Fragestellungen zu finden. Hierbei berücksichtigen sie einerseits sieben verschiedene Fachdidaktiken, andererseits durch Daten aus zehn europäischen Ländern auch internationale Blickwinkel auf Unterrichtsqualität. Die jeweiligen Ansätze in den drei Vorträgen werden in einem abschließenden Diskussionsbeitrag von Mirjam Steffensky (Universität Hamburg) kritisch reflektiert.

Beiträge des Symposiums

Identifikation und Genese fachspezifischer Merkmale und Modelle von Unterrichtsqualität – zwei Forschungsansätze im Vergleich

Alfred Lindl¹, Wiebke Langer²

¹Universität Regensburg, ²Universität Hamburg

Theoretischer Hintergrund

Bei der Untersuchung und Beurteilung von Unterrichtsqualität wird auf die große Bedeutung fachspezifischer Aspekte hingewiesen, da diese als besonders lernwirksam gelten (Begrich et al., 2023; Seidel & Shavelson, 2007). Daher rücken diese oder noch feiner granulierte, inhalts- und fertigungsbezogene Aspekte verstärkt in den Fokus bildungswissenschaftlicher und fachdidaktischer Forschung (z. B. Begrich et al., 2023; Lindl et al., 2024; Reusser, 2024; Rothland, 2024). Während die diesbezügliche Studienlage in mathematisch-naturwissenschaftlichen Domänen reichhaltig ist (z. B. Pauli et al., 2024; Praetorius et al., 2018; Schlesinger et al., 2018), sind Ansätze, welche Unterrichtsqualität in anderen Schulfächern unter generischen oder fachspezifischen Gesichtspunkten betrachten, vergleichsweise selten (z. B. Deutsch: Hesse, 2024; Wiprächtiger-Geppert et al., 2021; Ev. Religion: Rothgangel, 2021; Schweitzer, 2024; Fremdsprachen Englisch und Latein: Guttke, 2023; Lindl et al., 2024; Wilden, 2021; Musik: Kranefeld, 2021; Gutmiedl et al., 2024; Sport: Herrmann & Gerlach, 2020; Richartz & Kohake, 2021).

Fragestellung

Diese Forschungslücke wurde in den Projektgruppen FALKO-PV (Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung) und QualiTePE (Quality of Teaching in Physical Education) aufgegriffen, deren identische konzeptionelle Ansatzpunkte und ähnliche methodische Umsetzung, aber differierender fachlicher Fokus eine Synthese der Ergebnisse nahelegen (FALKO-PV: Deutsch, Englisch, Ev. Religion, Latein, Musik; QualiTePE: Sport). Zentrale Fragestellungen beider Projekte sind, inwieweit sich bestehende Modelle von Unterrichtsqualität (z. B. Charalambous & Praetorius, 2020) auf die genannten Fächer übertragen lassen, welche Adjustierungen oder Erweiterungen um spezifische Aspekte notwendig sind, und schließlich, wie sich potenzielle Qualitätsmerkmale – generisch oder fachspezifisch – konzeptualisieren, operationalisieren und validieren lassen.

Methode

Theoretische Ausgangspunkte beider Projekte waren zunächst die aktuellen Forschungsstände in den jeweiligen Fachdidaktiken und der Bildungswissenschaft bzw. insbesondere das erweiterte Syntheseframework respektive MAIN-TEACH-Modell (Praetorius et al., 2018; Charalambous & Praetorius, 2020). Darauf basierend wurden in beiden Projekten Delphi-Studien als Onlinebefragungen durchgeführt, die zur fachbezogenen Modellbildung sowie Identifikation und Validierung fachspezifischer Merkmale in den genannten Fächern beitragen sollten. Darin sollten von Teilnehmenden in FALKO-PV zunächst drei in Hinblick auf ein bestimmtes Fach absolut spezifische Qualitätskriterien in Freitextfeldern beschrieben werden. Anschließend galt es, die fachbezogene Relevanz der 21 Subdimensionen des Syntheseframeworks sowie elf bis 19 Kriterien aus dem bildungswissenschaftlichen und fachdidaktischen Diskurs zu gutem Unterricht auf einer fünfstufigen Likertskala zu bewerten (Hofrichter et al., 2024). In QualiTePE diente ein qualitativer Kommunikationsprozess mit den Projektpartner:innen in der Null-Runde zur Exploration fachbezogener Aspekte und diskursiven Adjustierung des theoretischen Rahmenmodells, bevor in der ersten Folgerunde die sportpädagogische Relevanz von 28 extrahierten Qualitätsmerkmalen auf einer siebenstufigen Ratingskala beurteilt wurde.

Ergebnisse und ihre Bedeutung

An der FALKO-PV-Onlinebefragung nahmen 348 Personen aus Deutschland und Österreich, an der QualiTePE-Delphi-Studie 324 Personen aus 17 europäischen Ländern teil. Diese Gelegenheitsstichproben umfassen jeweils Vertreter:innen der universitären Didaktiken und der Schulpraxis aller beteiligten Fächer. Die Freitextantworten in FALKO-PV und in QualiTePE wurden von (mind.) zwei fachkundigen Kodierer:innen inhaltsanalytisch (Kuckartz & Rädiker, 2022) und unabhängig voneinander (mittlere Proz. Übereinstimmung: 83 % bzw. 81 %) ausgewertet. Bei Betrachtung der offenen Antworten zeigt sich in beiden Projekten unter anderem eine interfachlich variierende Schwerpunktsetzung auf verschiedenen schülerseitigen Aktivierungsformen (z. B. psychomotorisch, ästhetisch), die über kognitive Aspekte hinausgehen. Die Relevanzeinschätzungen wurden – soweit inhaltlich vergleichbar – zur projektübergreifenden Auswertung einheitlich skaliert und zusammengefasst.

In Varianzanalysen zu den mittleren Relevanzeinschätzungen pro Item bzw. Dimension werden sowohl zwischen den sechs Fächern (z. B. hohe Werte für kognitive Aktivierung in Deutsch und niedrige für sozio-emotionale Unterstützung in Latein) als auch zwischen den Ländern (z. B. niedrige Werte für sozio-emotionale Unterstützung in Serbien und für kognitive sowie motorische Aktivierung in der Slowakei) Unterschiede sichtbar. Unter Einbezug fachspezifischer Besonderheiten liefern diese Resultate wertvolle Impulse für lerngegenstandsspezifische Ausdifferenzierungen generischer Rahmenmodelle und Dimensionen. Sie bilden zudem die Grundlage für die Entwicklung digitaler Instrumente zur Unterrichtsevaluation, die in den weiteren Vorträgen des Symposiums diskutiert werden.

Formative Evaluation der Unterrichtsqualität im Sportunterricht – Entwicklung des digitalen QualiTePE-Evaluationsinstruments

Christian Herrmann¹, Wiebke Langer², Erin Gerlach², Angelo Crapa¹

¹Pädagogische Hochschule Zürich, ²Universität Hamburg

Theoretischer Hintergrund

In der Diskussion um die fachspezifische Fundierung der Unterrichtsqualitätsforschung wird deutlich, dass es nicht nur einer gezielten Auseinandersetzung mit den spezifischen Unterrichtsmerkmalen und Lerngegenständen der Fächer bedarf, sondern dass diese Erkenntnisse auch die Schulpraxis und Unterrichtsentwicklung erreichen müssen. Eine Herausforderung der zukünftigen Unterrichtsqualitätsforschung kann in der Neu- und Weiterentwicklung von fachspezifischen Feedback- und Evaluationsinstrumenten gesehen werden, um damit einen Beitrag zur Steigerung der pädagogischen Handlungswirksamkeit zu leisten (u. a. Begrich et al., 2023; Reusser, 2024).

Um dieses Desiderat für das Fach Sport zu untersuchen, wurde im Rahmen des Erasmus+ geförderten Projekts „Quality of Teaching in Physical Education“ (QualiTePE) in Kooperation mit zehn europäischen Partnerländern das Ziel verfolgt, ein fachspezifisches Rahmenmodell zur Unterrichtsqualität im Fach Sport zu entwickeln und zu operationalisieren. In einem strukturierten Delphi-Verfahren wurden sechs Dimensionen identifiziert, die europaweit konsensfähig sind und Grundlage des fachspezifischen QualiTePE-Frameworks bilden (vgl. Symposiumsbeitrag 1). Dieses wurde im QualiTePE-Fragebogeninstrument operationalisiert, in das digitale QualiTePE-Evaluationstool integriert und in der QualiTePE-Pilotstudie in die Unterrichtspraxis implementiert und evaluiert.

Fragestellung

Aktuell fehlen validierte Evaluationsinstrumente, die eine fachspezifische und praxiswirksame Evaluation der Unterrichtsqualität im Fach Sport ermöglichen und dabei die unterschiedlichen fachdidaktischen Diskurse innerhalb Europas berücksichtigen (Herrmann & Gerlach, 2020). Daher stellten sich drei Fragen: (1) Wie können die Subdimensionen des QualiTePE-Frameworks in einem QualiTePE-Fragebogeninstrument operationalisiert werden? (2) Wie muss ein digitales QualiTePE-Evaluationstool konstruiert sein, das eine multiperspektivische (Schüler:in, Lehrperson, Beobachtende), mehrsprachige (u. a. Deutsch, Englisch) und adaptive Evaluation und Auswertung ermöglicht? (3) Wie beurteilen die Anwender:innen (Dozierende, Lehrpersonen) in der QualiTePE-Pilotstudie das QualiTePE-Evaluationstool hinsichtlich der Umsetzbarkeit, Verständlichkeit und Nützlichkeit in verschiedenen Lehr-Lernszenarien?

Methode

Zur Operationalisierung des QualiTePE-Frameworks wurde auf bereits validierte, sportspezifische Fragebogeninstrumente (Herrmann et al., 2015, 2023) zurückgegriffen. Um alle Subdimensionen inhaltlich valide abbilden zu können, wurden einzelne Skalen und Items ergänzt. Das entwickelte QualiTePE-Fragebogeninstrument wurde systematisch in die zehn jeweiligen Landessprachen der Projektländer übersetzt. Das QualiTePE-Fragebogeninstrument wurde sukzessive in das digitale QualiTePE-Evaluationstool überführt. In der QualiTePE-Pilotstudie wurde das QualiTePE-Evaluationstool von allen Partnerländern in die Lehrpersonenausbildung sowie in der Unterrichtspraxis implementiert. Die Implementationen wurden durch qualitativen und quantitativen Erhebung der Anwender:innen-Erfahrungen begleitet (u. a. Schrepp et al., 2017). Die Rückmeldungen zur Anwender:innen-Freudlichkeit und inhaltlicher Verständlichkeit wurden direkt für die Überarbeitung des Evaluationstools genutzt.

Ergebnisse

Auf Basis des im QualiTePE-Framework definierten sechs Dimensionen (u.a. kognitive und motorische Aktivierung) und den untergeordneten 20 Subdimensionen (u.a. Anspruchsniveau, kognitive Verarbeitung) des QualiTePE-Frameworks wurden in 23 Skalen (jeweils 3-5 Items) operationalisiert. Das QualiTePE-Fragebogeninstrument enthält 78 Items, wovon 58 Items aus bereits existierenden Studien adaptiert und 20 Items neu entwickelt wurden. Das Fragebogeninstrument liegt in jeweils drei Perspektiven

(Lernende, Lehrperson, Beobachtende) und zehn Sprachversionen (u.a. Englisch, Deutsch, Französisch, Spanisch, Italienisch) vor (Herrmann et al., 2024).

Das digitale QualiTePE-Evaluationstool ermöglicht eine dynamische Anpassung der zu erhebenden Subdimensionen und führt die verschiedenen QualiTePE-Fragebögen in ihren unterschiedlichen Sprachversionen und Perspektiven zusammen. Die Durchführung der Erhebung erfolgt mittels klassenbezogener QR-Codes oder Weblinks. Die Erhebungsdaten werden automatisch ausgewertet, indem die Perspektiven u.a. in einem Spider-Diagramm gegenübergestellt werden (Crapa et al., 2024).

Im quantitativen Teil der QualiTePE-Pilotierung konnte gezeigt werden, dass das QualiTePE-Evaluationstool in den Bereichen Effizienz, Verständlichkeit und Zuverlässigkeit moderate bis hohe Werte erzielte. Das Tool wurde von Lehrpersonen, Forschenden in der Lehrpersonenbildung und Dozierenden an Hochschulen und Universitäten als potenziell nützliches Instrument zur Unterstützung der Professionalisierung von (angehenden) Lehrpersonen eingeschätzt (Borghouts & Slingerland, 2024).

Bedeutung

Das entwickelte digitale QualiTePE-Evaluationstool ermöglicht eine adaptive, mehrperspektivische und mehrsprachige Erfassung und Auswertung der Unterrichtsqualität im Fach Sport. Durch den minimierten Arbeitsaufwand für die Anwender:innen sowie die hohe Benutzerfreundlichkeit werden die Hürden für eine formative Evaluation in der Unterrichtspraxis gesenkt. Vor dem Hintergrund der Betrachtung einiger generischer Qualitätsdimensionen des Unterrichts, kann es darüber hinaus für andere Fächer von Bedeutung sein.

Ästhetisch-emotionales Lernen als generisches Merkmal von Unterrichtsqualität? Erste Einblicke in die interdisziplinäre Vergleichsstudie FALKO-PV

Laura Simböck, Alfred Lindl, Patrick Ehrich, Mara Rader, Maria Hofrichter, Maximilian Gutsmedl, Stefan Boehringer
Universität Regensburg

Theoretischer Hintergrund

Im Forschungsdiskurs über Unterrichtsqualität werden generische Modelle wie das der drei Basisdimensionen (z. B. Klieme et al., 2006; Praetorius et al., 2018) zunehmend unter fachdidaktischer Perspektive reflektiert und fachspezifische Anpassungen sowie Ausdifferenzierungen vorgeschlagen (z. B. Begrich et al., 2023; Reusser, 2024). Maßgebend ist dabei unter anderem die Frage nach der transdisziplinären Passung und Gültigkeit von vornehmlich im mathematisch-naturwissenschaftlichen Kontext entwickelten theoretischen Konzepten und darin gewonnenen Erkenntnissen (z. B. Lipowsky & Bleck, 2019; Praetorius et al., 2020a). Insbesondere ist unklar, inwiefern zur Beschreibung von Unterrichtsqualität anerkannte Merkmale in anderen schulischen Fachbereichen – gerade auch hinsichtlich ihrer spezifischen Lerngegenstände und zu vermittelnden Fertigkeiten – anders zu akzentuieren bzw. zusammenzustellen sowie um weitere Facetten zu ergänzen sind (z. B. Lindl et al., 2024b; Praetorius et al., 2020b; Praetorius & Gräsel, 2021; Reusser, 2024). Beispielsweise wird in der Sportpädagogik eine motorische Aktivierung (Herrmann & Gerlach, 2020) oder in der Musikpädagogik eine ästhetische Aktivierung als Teilfacette der oder zusätzlich zur kognitiven Aktivierung in Erwägung gezogen (z. B. Kranefeld, 2021).

Fragestellung

Auch in der Forschungsgruppe FALKO-PV (Fachspezifische Lehrkraftkompetenzen – Prädiktive Validierung), deren übergeordneter Forschungsfokus auf der prognostischen Validität fachspezifischer Lehrkraftkompetenzen bezüglich der Qualität von Unterricht sowie dessen kognitiver und affektiver Zielkriterien liegt, wird eine Form ästhetischer Aktivierung als Merkmal von Unterrichtsqualität reflektiert (Lindl et al., 2024a). Mit Blick auf die Fächer Deutsch, Englisch, Evangelische Religion, Latein, Mathematik und Musik wird untersucht, inwiefern sich dieser Aspekt fachübergreifend oder -spezifisch modellieren und wie er sich aus der Perspektive von Schüler:innen messbar machen lässt. Von Interesse ist zudem, wie sensitiv die diesbezüglich entwickelten Skalen für Unterschiede zwischen Klassen und Fächern sind, ob die theoretisch implizierte Faktorstruktur empirisch sichtbar wird und welche Zusammenhänge mit Merkmalen von Lehrkräften und Schüler:innen bestehen.

Methode

Als Ergänzung der sieben generischen Dimensionen des MAIN-TEACH-Modells (Charalambous & Praetorius, 2020) wurde unter Rekurs auf die Forschungsstände in den Fachdidaktiken und eine Onlinebefragung von Lehrkräften der genannten Fächer (vgl. Symposiumsbeitrag 1) das Angebot, Potenzial und die Unterstützung zu ästhetisch-emotionalem Lernen als Merkmal von Unterrichtsqualität mit fünf Subdimensionen konzeptualisiert (I: ästh.-emot. Er- und Vorleben, 3 Items, II: Bedingungen und Freiräume für ästh.-emot. Erfahrungen, 4 Items, III: Schaffung ästh.-emot. Handlungsanlässe, 7 Items, IV: Potenzial zu Veränderungen infolge ästh.-emot. Erfahrungen, 5 Items; V: Sensibilität für die Beeinflussungsmöglichkeiten und Gefahren bei ästh.-emot. Erfahrungen, 4 Items; Lindl et al., 2024a). Von den konstruierten 23 Items lauten zehn Items in allen Fächern gleich, bei 13 Items ist die Formulierung der Forderung nach inhaltspezifischeren Operationalisierungen entsprechend (z. B. Reusser, 2024) lerngegenstandsbezogen konkretisiert. Alle wurden nach zwei Pilotierungsrunden in einen digitalen Fragebogen zu Merkmalen von Unterrichtsqualität, der mittels einer projekteigenen Webapp für Schüler:innen administriert wird, implementiert und sind auf einer fünfstufigen Likert-Skala zu beantworten.

Ergebnisse und ihre Bedeutung

In einer Studie, die im Schuljahr 2023/24 an bayerischen Gymnasien in den genannten Fächern durchgeführt wurde, nahmen 58 Lehrkräfte (62 Klassen) mit ihren Schüler:innen teil (N = 1487). Darin wurden die operationalisierten Merkmale von Unterrichtsqualität in einem Rotationsdesign mit 3x2 Durchgängen erfasst. Bezüglich der Skala ästhetisch-emotionalen Lernens zeigen mehrfaktorielle Varianzanalysen für generische sowie fachlich konkretisierte Items signifikante Unterschiede mit teils großen Spannweiten zwischen Lehrkräften, aber nicht zwischen Fächern. Mittels des Reliabilitätsindikators Cronbachs Alpha ermittelte Reliabilitäten sind für die Subdimensionen und die Gesamtskala angemessen. In einer konfirmatorischen Faktorenanalyse lassen sich vier der fünf Subdimensionen wie angenommen dem ästhetisch-emotionalen Lernen zuordnen und ein Modellvergleich weist auf eine empirische Trennbarkeit zwischen diesem und dem Merkmal der kognitiven Aktivierung hin. Darüber hinaus finden sich signifikante bivariate Korrelationen beispielsweise mit dem Enthusiasmus von Lehrkräften oder der Lernfreude von Schüler:innen. Diese Befunde regen zur Weiterentwicklung der Fragebogenskala und Diskussion darüber an, welchen Mehrwert fachspezifische Ausdifferenzierungen von Unterrichtsqualitätsmerkmalen besitzen.

S-96: Adaptive Unterrichtsplanung: Konzepte, Praxis und Förderung schüler*innenbezogener Entscheidungsprozesse bei Lehrkräften

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 148

Chair der Sitzung: Daniel Scholl

Chair der Sitzung: Simon Küth

Chair der Sitzung: Katharina Schnitzler

Adaptive Unterrichtsplanung: Konzepte, Praxis und Förderung schüler*innenbezogener Entscheidungsprozesse bei Lehrkräften

Chair(s): **Daniel Scholl, Prof. Dr.** (Universität Siegen, Deutschland), **Simon Küth** (Universität Siegen, Deutschland)

Diskutant*in(nen): **Katharina Schnitzler, Dr.** (Universität Potsdam, Deutschland)

Ein adaptiver Unterricht gilt als entscheidend für den Umgang mit schulischer Heterogenität. Einen solchen Unterricht zu verwirklichen setzt voraus, dass Lehrkräfte früh lernen, ihn bereits ab seiner Planung flexibel an die individuellen Bedürfnisse der Lernenden anzupassen (Jäger et al., 2019; KMK, 2022; Corno & Snow, 1986; Parsons et al., 2018). Dadurch können Unterrichtsmaßnahmen einerseits differenziert auf die Determinanten von Schulleistungen (Brühwiler et al., 2017) in ihren heterogenen Ausprägungen (Trautmann & Wischer, 2011) bezogen werden, um dem Recht aller Schüler*innen auf „effective education“ (United Nations, 2006, Artikel 24, 2d) zu entsprechen (Dumont et al., 2024). Andererseits kann die Diversität der Schüler*innen als Vielfalt ihrer Persönlichkeitsressourcen (Prenzel, 2018) berücksichtigt und zur Befriedigung ihres Bedürfnisses nach sozialer Teilhabe, Freundschaft, unbedingter Wertschätzung und individueller Kompetenzentwicklung beigetragen werden (Deutsche UNESCO-Kommission e.V., 2014; Nilhom & Göransson, 2017).

Trotz dieses Konsenses über die Bedeutung eines adaptiven Unterrichts und einer entsprechenden Planung fordern grundlegende Fragen die Forschung heraus: Wie können unterschiedliche disziplinäre Perspektiven zum adaptiven Unterricht und seiner Planung produktiv integriert werden? Welche adaptiven Planungsentscheidungen treffen Expert*innen im Vergleich zu Noviz*innen genau? Wie lassen sich makro- und mikroadaptive Strategien in der Unterrichtsplanung (konzeptuell) verbinden? Und inwieweit können Interventionsprogramme die Entwicklung adaptiver Planungskompetenzen bei angehenden Lehrkräften fördern?

In diesem interdisziplinären Symposium werden qualitative und quantitative Methoden vereint, darunter literaturbasierte Übersichtsarbeiten, vignettenbasierte Fähigkeitstests, qualitative Interviews und Interventionsstudien, um Antworten auf diese Fragen zu suchen und zur Weiterentwicklung der empirischen Unterrichtsplanungsforschung und zur Professionalisierung von Lehrkräften beizutragen.

In Beitrag 1 werden die theoretischen Grundlagen unterrichtlicher Adaptivität aus verschiedenen disziplinären Perspektiven untersucht. Im Ergebnis wird eine multiparadigmatische Annäherung an das Konzept gefordert, um die Komplexität des Adaptivitätskonstrukts mit den spannungsreichen Dynamiken unterrichtlicher Interaktionen zu relationieren.

In Beitrag 2 wird die Fähigkeit zur adaptiven Unterrichtsplanung als Teilaspekt des interdependenten Planungsentscheidens von erfahrenen Lehrkräften und Lehramtsstudierenden mittels vignettenbasierter Fähigkeitstests und netzwerkanalytischer Verfahren triangulativ verglichen. Die Ergebnisse zeigen, dass erfahrene Lehrkräfte eine signifikant höhere adaptive Planungskompetenz aufweisen, und legen offen, welche Planungsentscheidungen auf die Lernvoraussetzungen der Schüler*innen bezogen werden.

In Beitrag 3 wird der Zusammenhang zwischen diagnostischem Wissen und der Fähigkeit von Lehrkräften beleuchtet, adaptive Unterrichtsentscheidungen auf Makro- und Mikroebene zu treffen. Die Ergebnisse weisen darauf hin, dass die prospektive Mikroadaptation eine Brücke zwischen langfristiger Planung und situativer Anpassung darstellt.

In Beitrag 4 wird eine Interventionsstudie zur Förderung der adaptiven Planungskompetenz von Lehramtsstudierenden präsentiert. Durch ein Seminarkonzept und einen vignettenbasierten Test wird untersucht, wie angehende Lehrkräfte besser auf die Herausforderungen heterogener Lerngruppen vorbereitet werden können.

Beiträge des Symposiums

Multiparadigmatische Annäherungen an unterrichtliche Adaptivität – Eine theoriebasierte Relationierung unterschiedlicher Perspektiven im Adaptivitätsdiskurs

Julia Frohn, Dr.¹, Johannes Ludwig², Vivien Heller, Prof. Dr.³, Vera Moser, Prof. Dr.²

¹Humboldt-Universität zu Berlin, Deutschland, ²Goethe-Universität Frankfurt, Deutschland, ³Bergische Universität Wuppertal, Deutschland

Die nationale und internationale Forschung zu unterrichtlicher Adaptivität hat besonders in den letzten zwei Jahrzehnten den Diskurs um den Umgang mit schulischer Heterogenität maßgeblich mitbestimmt (Parsons et al., 2018). Dabei haben sich die Perspektiven auf das – vor allem vom pädagogisch-psychologischen Paradigma geprägten – Thema im Zeitverlauf durchaus gewandelt: Waren ursprünglich „die Begriffe Adaptivität und Instruktion häufig verbunden“ (Bohl et al., 2012, S. 48; zur Kritik siehe auch Wischer & Trautmann, 2012), sodass besonderer Fokus auf das lehrkräftezentrierte instruktive Handeln bestand, attestiert die aktuelle Forschung dem Konstrukt der Adaptivität „einen hohen Grad an Flexibilität in der Organisation von Unterricht“ (Dumont et al., 2024, o. S.). Um diesen offenen Unterrichtsformen adäquat zu begegnen, wird die Notwendigkeit der Vielfalt theoretischer und empirischer Forschungsansätze zu unterrichtlicher Adaptivität zunehmend in den Vordergrund gerückt (Dumont et al., 2024). Während aus Sicht der empirischen Lehr-Lern-Forschung ein theoriebezogenes Desiderat v.a. in der Verhältnisbestimmung von Adaptivität zu den sogenannten Basisdimensionen von Unterricht besteht (Dumont et al., 2024.), erscheint es aus erziehungswissenschaftlicher und fachdidaktischer Sicht bedeutsam, eine theoretische Verortung des Adaptivitätskonstrukts in den spannungsreichen Strukturlogiken (fach)unterrichtlicher Interaktion vorzunehmen, um dem Konstrukt in seiner Komplexität gerecht zu werden. Auch aus inklusionspädagogischer Perspektive bedarf es entsprechender (Ziel)Diskussionen, etwa mit Blick auf die Persönlichkeitsentwicklung der Lernenden oder auf mögliche normative Setzungen (Martschinke, 2015; Hertel et al., 2016).

Dieser Beitrag zielt damit als integrierende Übersichtsarbeit auf eine literaturbasierte Entwicklung eines erkenntnistheoretischen Grundgerüsts zur Beschreibung unterrichtlicher Adaptivität aus unterschiedlichen disziplinären Perspektiven. Im Sinne eines multiparadigmatischen Ansatzes geht es dabei darum, „unnötigen und erkenntnisverengenden Vereinseitigungen“ (Heinrich et al., 2019) mit dem Ziel entgegenzuwirken, adaptives Unterrichten anhand verschiedener Zugänge theoretisch zu beschreiben und Verbindungen und Unterschiede in den Forschungsansätzen produktiv für die weiterführende Forschung nutzbar zu machen.

Aufbauend auf Arbeiten zur Abgrenzung von Adaptivität zu anderen unterrichtlichen Konzepten im Umgang mit heterogenen Lerngruppen, etwa Individualisierung, Binnendifferenzierung oder Personalisiertes Lernen (siehe z. B. Dumont, 2019; Klieme & Warwas, 2011; Stebler et al., 2021), sucht dieser Beitrag damit die teils als „fremde Schwestern“ (vgl. Gebauer, 2019) beschriebenen Konzepte unter der Annahme von Adaptivität als Herstellung von Passungen zwischen Angebot und Nutzung bzw. zwischen institutionellen Normen und situativer sozialer Praxis sinnstiftend zu relationieren.

Dabei wird adaptives Unterrichten auf den Ebenen der Planung, des Prozesses und der Reflexion unterrichtlicher Interaktion aus den Perspektiven der empirischen Lehr-Lern-Forschung, der rekonstruktiv-erziehungswissenschaftlichen Unterrichtsforschung, der Fachdidaktik sowie der Inklusionspädagogik analysiert. Im Gegensatz zu oft eher allgemein ausgerichteten Forschungsansätzen zu unterrichtlicher Adaptivität soll dabei ein Fokus auf Fragen der – auch ko-konstruktiven und partizipativen – Gegenstandskonstitution und der damit verbundenen Fachlichkeit adaptiven Unterrichts liegen. Ferner soll die „Vorstellung der Steuerbarkeit schulischen Geschehens“ (Wiesemann, 2015, S. 38) durch Überlegungen zu unterrichtlichen Spannungsverhältnissen im adaptiven Unterrichten hinterfragt und im Sinne einer angestrebten Passung zwischen unterrichtlichen Determinanten produktiv geöffnet werden.

Wer trifft welche adaptiven Planungsentscheidungen? Triangulation von Studienergebnissen kontrastiver Expertisevergleiche

Daniel Scholl, Prof. Dr.¹, Simon KÜth¹, Jana Meier, Dr.², Christoph Vogelsang, Dr.²

¹Universität Siegen, Deutschland, ²Universität Paderborn, Deutschland

Theoretischer Hintergrund: In den meisten (fach)didaktischen Unterrichtsplanungsmodellen (Scholl et al., 2020) wird von Lehrkräften gefordert, alle unterrichtsrelevanten Entscheidungen auf die Schüler*innenvoraussetzungen zu beziehen und durch didaktische Adaptivität eine „fachunspezifische Passung der kognitiven Voraussetzungen der Lerngruppe und der Aufgaben für die geplante Unterrichtsstunde“ (König et al., 2015, S. 375) herzustellen. Als spezieller Aspekt des Interdependenzprinzips, einem der wichtigsten fachübergreifenden Planungskriterien (Vogelsang et al., 2017; Scholl et al., 2022), ist diese Forderung eingebettet in die noch grundsätzlichere, eine Passung zwischen Planungsentscheidungen zu Lehr-Lernzielen, -inhalten, -methoden und -medien sowie den situativen Bedingungen und den individuellen Voraussetzungen der Lernenden herzustellen.

Während klassische präskriptive (fach)didaktische Planungsmodelle keine expliziten Aussagen über die notwendigen Planungsfähigkeiten zur Umsetzung interdependenter Planungsüberlegungen treffen, betonen kompetenztheoretische Ansätze, dass eine interdependente Unterrichtsplanung nur auf Basis einer adaptiver Planungskompetenz möglich sei (z. B. Beck et al., 2008). Erst durch diese Kompetenz könnten Lehrkräfte potenziell adaptiv bedeutsame Entscheidungsmomente wahrnehmen (Bräuer, 2015) und angemessen berücksichtigen (Brühwiler, 2014).

In der Planungspraxis ist die adaptive Planungskompetenz allerdings hochgradig unterschiedlich ausgeprägt: Während zumindest Expert*innen die Lernvoraussetzungen und -bedürfnisse ihrer Schüler*innen (teilweise) in ihre Planung einbeziehen und ihre Lehrstrategien entsprechend anpassen können, fällt es Noviz*innen schwer, diagnostische Informationen zu integrieren (z. B. König et al., 2020): In ihren eher linearen Unterrichtsplanungen (Westermann, 1991) beachten Noviz*innen nur wenige Planungsbereiche – Unterrichtsinhalte oder -ziele besonders oft nicht (z. B. Karlström et al., 2021) –, die sie zudem unzureichend miteinander verschränken (z. B. Seifried, 2009) und ungenügend auf die Lernvoraussetzungen beziehen (Gassmann, 2013), sodass die Pläne kaum adaptiv sind (Black et al., 2019).

Fragestellung: Obwohl die Diskrepanz zwischen der eingeforderten und realisierten adaptiven Unterrichtsplanung gut belegt ist, gibt es zu wenige Studien, die die konkrete Interdependenz der Planungsentscheidungen von Lehrkräften kontrastierend untersuchen, um Grundlagen für evidenzbasierte Planungsstrategien zu schaffen (König et al., 2020). Deshalb wurde gefragt, welche Planungsentscheidungen genau erfahrene Lehrkräfte und Lehramtsstudierende auf die Lernvoraussetzungen beziehen und wie sich beide Gruppen unterscheiden.

Methode: In einer Gelegenheitsstichprobe von 126 Lehramtsstudierenden und 172 Lehrkräften wurde ein zweidimensionaler vignettenbasierter Fähigkeitstest zur Unterrichtsplanung eingesetzt (Scholl et al., 2022), in dem Planungsentscheidungen aufeinander und auf situative und individuelle Lernvoraussetzungen bezogen werden müssen, um eine möglichst hohe Interdependenz herzustellen. Die Testergebnisse wurden auf signifikante Gruppenunterschiede untersucht. Zusätzlich wurden in einer qualitativen Studie mit 22 Teilnehmenden (6 Bachelorstudierende, 6 Masterstudierende, 4 Referendar*innen und 6 Lehrkräfte), leitfadengestützte Interviews zu zwei Planungsvignetten geführt, die in der Form des lauten Denkens (Ericsson & Simon, 1993) bearbeitet wurden. Die Auswertung durch geschulte Rater*innen folgt der Qualitativen Inhalts- (Mayring, 2022) und der Netzwerkanalyse (u. a. Hollstein & Straus, 2006).

Ergebnisse und ihre Bedeutung: Lehrkräfte erreichen signifikant höhere Scores im Planungsfähigkeitstest, $F(1, 296) = 4,58$, $p = 0,033$, $\eta^2 = 0,015$: Ihre Planungsentscheidungen ($M = 5,92$, $SD = 1,92$) sind interdependenter als die der Lehramtsstudierenden ($M = 5,43$, $SD = 2,05$). Insbesondere intentionale Entscheidungen werden häufiger adaptiv auf einen Planungshintergrund bezogen. Die netzwerkanalytische Auswertung der transkribierten Interviews bestätigt, dass die Anzahl der getroffenen Planungsentscheidungen und ihrer Verknüpfungen mit zunehmendem Expertisegrad steigt (BA-Studierende $M = 7,67$ Planungselemente und $M = 4$ Verknüpfungen vs. Lehrkräfte $M = 15,17$ Elemente und $M = 6,33$ Verknüpfungen). Unabhängig von der Expertise werden Lernvoraussetzungen insbesondere bei Methoden- und Inhaltsentscheidungen mitgedacht, selten bei Ziel- oder Medienentscheidungen.

Ohne Anspruch auf Verallgemeinerbarkeit zeigen die Ergebnisse genauer, wie sich Ansätze zur Adaptivität in den Planungsentscheidungen im Vergleich zwischen Expert*innen und Noviz*innen darstellen. Werden diese empirisch nachweisbaren Ansätze in Folgestudien auf die Qualität des anschließenden adaptiven Unterrichts bezogen, könnten praxisnahe Prinzipien einer adaptiven Unterrichtsplanung empirisch untermauert und das Stigma der (adaptiven) Planungskompetenz als reinem Phänomen der Lehrer*innenbildung relativiert werden (Rothland, 2022). Je nach Schwerpunktsetzung: Probleme der Ausprägung von adaptiver Planungskompetenz, die einerseits nach gründlicherer Aufklärung, andererseits nach spezifischen Fördermöglichkeiten verlangen.

Prospektive Mikroadaptation als Komponente adaptiver Unterrichtsplanung

Silvia Greiten, Prof. Dr., Theresa Overbeck, Samira Skribbe

Pädagogische Hochschule Heidelberg, Deutschland

Theoretischer Hintergrund: Bildung und adaptiver Unterricht sind untrennbar miteinander verbunden, da adaptiver Unterricht darauf abzielt, den Bildungsprozess auf die individuellen Bedürfnisse der Schüler*innen abzustimmen (Beck et al. 2008). Dazu sind Lehrkräfte herausgefordert, differenzierende Lernsituationen auf Adaptivität hin zu planen (Decristan, 2020) mit dem Ziel,

eine effektive und gerechte Bildung für alle Schüler*innen zu gewährleisten. Die in diesem Beitrag vorgestellte Studie untersucht den Zusammenhang von förderbezogener Diagnostik und Folgen für die Unterrichtsplanung und das Handeln von Lehrkräften im Unterricht. Konzepte des adaptiven Unterrichts, konkretisiert mit Makro- und Mikroadaptation (Corno, 2008; Corno & Snow, 1986) bilden den theoretischen Hintergrund. Die Makroadaptation umfasst die langfristige Planung von Unterrichtseinheiten mit Planungen von Material, Methoden, Lernsituationen usw., die Mikroadaptation beschreibt die kurzfristigen Anpassungen in spezifischen Unterrichtssituationen. Lehrkräfte müssen dazu flexibel reagieren und im laufenden Unterricht Entscheidungen treffen. Dabei handelt es sich um ein interaktives, kommunikatives Geschehen mit Feedback, Fragen, Materialanpassungen (Parsons et al., 2018) in konkreten Lehr-Lern-Interaktion, verknüpft mit tiefenstrukturellen Qualitätsdimensionen des Unterrichts (Praetorius et al., 2020). Aktuell finden sich in Forschung und Schulpraxis eine Vielzahl an Konzeptualisierungen adaptiven Unterrichts (Hardy et al., 2019), basierend auf Annahmen zur Förderung von Schüler*innen. Forschung zur Wirksamkeit adaptiven Unterrichts (Dumont et al., 2024), zu Grenzen der Schüler*innenperspektive in Planungsprozessen (Fischer, 2019) und auch bspw. zum Zusammenhang von Diagnostik und Unterrichtsplanung (Scribner et al., 2024; Overbeck et al. 2024) markieren aktuelle Kontexte.

In diesem Beitrag wird eine Studie aus dem vom BMBF geförderten DiaGU-Projekt (Greiten et al. 2023) vorgestellt. In der Studie wurden Interviews, Planungs- und Reflexionsgespräche von Lehrkräften dahingehend untersucht, inwiefern Lehrkräfte während makroadaptiver Prozesse auch Optionen für mikroadaptives Handeln adaptieren und daraufhin gezielt entsprechende Lernsituationen planen. Diese Prozesse wurden unter dem Begriff der prospektiven Mikroadaptation gefasst (Publikation i.V.). Im Beitrag werden Ergebnisse aus der Interviewstudie vorgestellt.

Fragestellung: Welche Zusammenhänge zwischen diagnostischem, didaktischem Wissen und Prozessen der Unterrichtsplanung in Bezug auf Makro- und Mikroadaptation lassen sich aus Interviews mit Lehrkräften beschreiben?

Methode: Das Fallsampling bestand aus 8 Interviews mit im inklusiven Unterricht kooperierenden Lehrkräften. Zur Datenerhebung wurden Leitfadeninterviews (Helfferich 2022) genutzt, zur Datenauswertung die inhaltlich-strukturierende Qualitative Inhaltsanalyse (Gläser-Zikuda et al., 2020; Kuckartz & Rädiker 2022). Die Kategorienbildung erfolgte deduktiv und induktiv (Waldherr et al. 2019).

Ergebnisse und ihre Bedeutung: Die Ergebnisse dokumentieren, dass, wenn Lehrkräfte über mikroadaptive Strategien verfügen und ihnen die Zusammenhänge zwischen Lernsituationen und mikroadaptiven Strategien bewusst sind, sie dann auch förderbezogene, binnendifferenzierte Lernsituationen planen, um mikroadaptiv handeln zu können. Bereits im Planungsprozess stellen sie sich mögliche mikroadaptive Handlungen und Interaktionen prospektiv vor. An Interviewausschnitten und Kommentierungen von Lehrkräften und mit Bezug auf Dokumentationen zu Planungsprozessen werden diese Zusammenhänge aufgezeigt. Die prospektive Mikroadaptation kann somit als Brücke zwischen Makro- und Mikroadaptation bezeichnet werden und ist für die Wirksamkeit adaptiver Unterrichtsplanung relevant.

Intervention zur Förderung adaptiver Planungskompetenz von angehenden Lehrpersonen (InterADAPT): Ein binationales Lehr-Projekt zum adäquaten Umgang mit Heterogenität

Thomas Rey, Dr., Valentin Unger, Dr.
Pädagogische Hochschule St.Gallen, Schweiz

Theoretischer Hintergrund: Gesellschaftliche Entwicklungen wie Globalisierung, Migration, Pluralisierung von Lebensformen und demografischer Wandel tragen zu einer wachsenden Heterogenität in unserer Gesellschaft bei. Der adäquate Umgang mit Heterogenität in Schule und Unterricht gilt dabei als eine der größten Herausforderungen des Lehrberufs (z. B. Herzmann & König, 2016; OECD, 2021), die (angehende) Lehrer*innen länderübergreifend zu bewältigen haben. Um dieser Notwendigkeit zu entsprechen sowie der unterrichtlichen Herausforderung produktiv zu begegnen, gehen aktuelle Ansätze zur professionellen Lehrpersonenkompetenz davon aus, dass die Fähigkeit guten Unterricht zu halten das Ergebnis eines lebensbiographischen Entwicklungsprozesses ist (vgl. Baumert & Kunze, 2006). Somit benötigen Lehrer*innen für die Bewältigung zentraler Anforderungen ein berufsbezogenes Wissen und Können, das sie im Laufe der Lehrpersonenbildung aufbauen. Das Konzept der adaptiven Lehrkompetenz (Beck et al., 2008; Rogalla & Vogt, 2008) nimmt sich dieser Notwendigkeit an, indem es situationsspezifische Fähig- und Fertigkeiten einer Lehrperson im adaptiven Unterricht beschreibt, der möglichst günstige Bedingungen für das Erreichen individueller Lernziele schafft. Hierbei konnte sowohl das Unterrichtshandeln als auch die Unterrichtsplanungskompetenz als Facetten adaptiver Lehrkompetenz identifiziert werden.

Im Rahmen der Unterrichtsplanung stellt adaptive Planungskompetenz die „Expertise zur Passung von fachunspezifischen Voraussetzungen und Bedürfnissen der Lernenden mit planerischen Unterrichtsentscheidungen als generische Anforderung“ (Rey et al., 2018, S.127) dar. Um Lehrpersonen bei den planerischen Tätigkeiten in heterogenen Lerngruppen zu unterstützen, wurde von Rey (2022) Allgemein didaktische Kriterien zur Erfassung einer adaptiven Planungskompetenz im Umgang mit Heterogenität (AIPako) entwickelt und validiert. Auf diesem evidenzbasierten Testinstrument und einem Strategie-Kategorien-Modell adaptiver Planungskompetenz im Umgang mit Heterogenität (Rey et al., 2023) aufbauend, wurde ein Hochschulseminar (CH und D) konzipiert, um künftige Lehrer*innen auf den Unterrichtsalltag in heterogenen Lerngruppen vorzubereiten und ihnen möglichst praxisnahe und effektive Handlungs- und Unterstützungsmöglichkeiten bei der Unterrichtsplanung aufzuzeigen.

Fragestellungen: Im Beitrag werden die Erfassung und Förderung adaptiver Planungskompetenz angehender Lehrpersonen fokussiert und nachfolgende Forschungsfragen verfolgt:

- (1) Lässt sich adaptive Planungskompetenz über entwickelte Textvignetten valide und reliabel erfassen?
- (2) Kann mittels des konzipierten Seminars adaptive Planungskompetenz gefördert werden?

Methode: Ein Ziel des Seminars ist es, ein nachhaltiges Konzept zu entwickeln, dessen Wirksamkeit mithilfe einer Interventionsstudie im Prä-Post-Vergleichsgruppendesign überprüft wird. Neben der Vorstellung des Seminarkonzepts mit besonderem Praxisbezug wird im Vortrag ein Vignettestest zur Erfassung adaptiver Planungskompetenz vorgestellt. Im Sinne des advokatorischen Ansatzes (Oser et al., 2007) bearbeiten die Lehramtsstudierenden im Prä-Post-Design die Unterrichtsvignetten, identifizieren unterrichtsplanungsbezogene Widrigkeiten und beschreiben adäquate Handlungsalternativen. Von März bis Juni 2024 haben Studierende der Pädagogischen Hochschule St.Gallen (Interventionsgruppe CH; MZP_1: n = 56, MZP_2: n = 49) und der Universität Konstanz (Interventionsgruppe D; MZP_1: n = 32, MZP_2: n = 28) teilgenommen. Zudem wird der Test in regulären Seminaren eingesetzt (Kontrollgruppen; derzeit n = 35).

Die entwickelten Vignetten zur Erfassung adaptiver Planungskompetenz wurden in einem zweistufigen Expert*innenrating validiert und werden psychometrisch geprüft. Zur Auswertung der Interventionsstudie werden regressionsanalytische Verfahren herangezogen, mittels derer der Interventionseffekt geprüft werden kann. Als abhängige Variable wird der Score MZP2 aufgenommen, erste unabhängige Variable ist der Score zu MZP1. Weitere Prädiktoren können als unabhängige Variablen mit

aufgenommen werden. Neben den Vignetten wurden die Studierenden offen zu ihren Seminarerfahrungen befragt. Die Antworten werden in einer qualitativen Inhaltsanalyse kategorisiert.

Ergebnisse und Bedeutung: Erste Befunde der zweistufige Expertinnen*validierung zeigen, dass das Testinstrument zur Erfassung adaptiver Planungskompetenz signifikant verbessert werden konnte ($t(19) = 3.759$, $p = .001$). Die Ergebnisse der zweiten Evaluationsstufe (1 = trifft überhaupt nicht zu bis 6 = trifft voll und ganz zu) weisen auf eine hohe Expert*innenvalidität ($M = 5.72$, $SD = .30$) hin.

Gleichwohl noch keine Ergebnisse hinsichtlich der Intervention aufgrund der jüngst abgeschlossenen Erhebung vorliegen (liegen zum Vortrag vor), gründet der Innovationsgehalt des Seminars in der Verzahnung von empirisch validierten Forschungsergebnissen mit starkem unterrichtlichem Praxisbezug einer länderübergreifenden Problematik und zeigt dadurch seine Bedeutsamkeit auf.

S-97: (Post-)Sekundäre Sozialisationskontexte und ihr Einfluss auf Bildungs- und Berufsaspirationen

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: O 131

Chair der Sitzung: Annabell Daniel

Chair der Sitzung: Frederick de Moll

Chair der Sitzung: Steffen Schindler

(Post-)Sekundäre Sozialisationskontexte und ihr Einfluss auf Bildungs- und Berufsaspirationen

Chair(s): **Annabell Daniel** (LMU München, Deutschland), **Frederick de Moll** (Universität Bielefeld)

Diskutant*in(nen): **Steffen Schindler** (Universität Bamberg)

Schule, Ausbildung und Hochschule fungieren als zentrale Sozialisationskontexte während der Adoleszenz und im frühen Erwachsenenalter. Sie vermitteln in systematischer Weise Kompetenzen und Qualifikationen, geben Werte, Normen und Orientierungen weiter und bereiten Jugendliche bzw. junge Erwachsene auf ihre spätere Position und Rolle in der Gesellschaft vor. Damit kommt ihnen eine wesentliche Bedeutung für die Entwicklung der sozialen und beruflichen Identität zu. (Post-)sekundäre Bildungsinstitutionen beeinflussen aber nicht nur individuelle Bildungsprozesse, sondern werden auch durch die jugendlichen und jungen Erwachsenen selbst geprägt. Diese dynamischen Wechselwirkungen zwischen Individuum und Institution bilden die Sozialisationsbedingungen, die theoretisch und empirisch aus unterschiedlichen Perspektiven beleuchtet werden können.

Unter Rückgriff auf die Person-Environment Fit Theorie (Edwards et al. 1998) zeigt sich über die verschiedenen Bildungsinstitutionen hinweg, dass eine hohe Passung zwischen persönlichen Bedürfnissen und der Bedürfnisbefriedigung durch die Umwelt einhergeht mit schulischer Motivation (Eccles et al., 1993), guten Schulleistungen (Fraser & Fisher 1983), Anstrengungen in der Berufsausbildung (Neuenschwander et al. 2024) und dem Studienerfolg (Etzel & Nagy 2016). Untersuchungen im Kontext der Theorie der sozialen Identität (Tajfel 1978) legen zudem nahe, dass das Gefühl, sozial eingebunden und ein geschätzter Teil der Gemeinschaft zu sein, sich positiv auf Leistung und Wohlbefinden (Walton & Cohen 2011) sowie den Verbleib in Ausbildung (Kemmler 2012) und Studium (Walton et al. 2023) auswirkt. Einen weiteren Ansatz, um erfolgreiche Sozialisationsprozesse im postsekundären Bildungsbereich zu untersuchen, bietet das Integrationsmodell von Tinto (1975). Studien zeigen, dass eine hohe akademische und Integration (i.S. einer gelungenen Bewältigung von Leistungsanforderungen) und eine hohe soziale Integration (i.S. der Kontakthäufigkeit zu Lehrenden und Mitlernenden) die Abbruchneigung sowohl in der beruflichen Ausbildung (Holtmann & Solga 2023) als auch im Hochschulstudium (Klein 2019) verringern können.

Die Literatur identifiziert dabei auch vulnerable Gruppen wie Lernende mit Zuwanderungshintergrund, Erstakademiker:innen oder nicht-traditionelle Studierende, für die sich in den Sozialisationskontexten unterschiedliche Herausforderungen stellen (u.a. Williams et al. 2020; Janke et al. 2024).

Das interdisziplinär ausgerichtete Symposium greift diese verschiedenen Perspektiven auf, indem es nach der Bedeutung von (post-)sekundären Sozialisationsprozessen für Bildungs- und Berufsaspirationen fragt und dabei unterschiedliche Akteursgruppen in den Blick nimmt.

Der erste Beitrag untersucht soziokulturelle, geschlechts- und migrationsbezogene Ungleichheiten im berufswahlbezogenen Informationsverhalten, das wiederum zentral für die Entwicklung beruflicher Aspirationen ist. Analysen auf Basis der PISA 2022-Daten zeigen, dass die soziale Integration – hier abgebildet über die Lehrer-Schüler-Beziehung und die soziale Eingebundenheit – den Zugang zu berufsrelevanten Informationen begünstigen und damit Ungleichheiten tendenziell kompensieren kann.

Der zweite Beitrag beschäftigt sich mit dem gymnasialen zweiten Bildungsweg. Er analysiert, wie Lehrkräfte mit dem Spannungsverhältnis zwischen den persönlichen Bedürfnissen der erwachsenen Schüler:innen und den Leistungsanforderungen eines gymnasialen Bildungsgangs umgehen und inwieweit es zu einer Passung zwischen der Praxis der Lehrkräfte und den Erwartungen der Schüler:innen kommt. Die Ergebnisse der Latent-Class-Analysen verweisen auf substanzielle Passungsprobleme, die auch mit einer höheren Abbruchneigung auf dem zweiten Bildungsweg einhergehen dürften.

Basierend auf den Daten des Schweizer DAB-Panels untersucht der dritte Beitrag die Entwicklung beruflicher Aspirationen am Übergang von der Sekundarstufe I in den Beruf. Die Befunde der Wachstumskurvenmodelle verweisen darauf, dass nachobligatorische Bildungswege, die neue Sozialisationsräume eröffnen, Informationen bereitstellen und bei gegebener Durchlässigkeit des Systems neue Karriereoptionen eröffnen, einen Anstieg beruflicher Aspirationen begünstigen können.

Im vierten Beitrag wird der Frage nachgegangen, inwieweit hochschulische Sozialisationserfahrungen die Absicht, ein Promotionsstudium aufzunehmen, beeinflussen. Studierende mit einer gesundheitlichen Studienschwerenden Beeinträchtigung unterscheiden sich in ihrer Promotionsabsicht zwar kaum von Studierenden ohne Beeinträchtigungen, die analysierten Daten der Studierendenbefragung in Deutschland zeigen aber, dass die Promotionsabsicht nach der Beeinträchtigungsart und dem Ausmaß der akademischen und sozialen Integration variiert.

Beiträge des Symposiums

ENTFÄLLT Unraveling How Social Integration in Educational Institutions Shapes Youth Career Information-Seeking Practices

Sercan Erer, Frederick de Moll
Universität Bielefeld

Theoretical Background

We explore how social integration in educational contexts influences disparities in students' future career information-seeking practices (FCISP). Utilizing Social Cognitive Career Theory (SCCT) and data from PISA 2022, we investigate the role of personal attributes and contextual factors in shaping these practices. SCCT, as developed by Lent et al. (1994) and building on Bandura's theories, suggests that career development results from the dynamic interplay between personal and environmental characteristics. Personal attributes include self-efficacy, goals, and outcome expectations, as well as demographic factors like gender, socioeconomic status (SES), and ethnicity (Wang et al., 2022). Contextual factors involve school environments and family involvement, which shape students' learning experiences and their career choices.

Research indicates that disparities in career aspirations and FCISP often align with gender, SES, and ethnicity. Al-Bahrani et al. (2020), Conlon et al. (2023), Howard (2011), Stoet et al. (2022), and Tang et al. (2008) highlight how these factors create varying levels of access and opportunity for students. Social integration within educational contexts is significantly affected by the sense of belonging and quality of teacher-student relationships, which are positively correlated with educational aspirations and self-efficacy (Chinwé et al., 2008; Wong et al., 2019). Moreover, the influence of peers is crucial, as supportive peer environments can lead to more positive educational outcomes and career aspirations (Ruschhoff et al., 2021). Parental involvement further enhances social integration by aligning home and school expectations, fostering higher aspirations and engagement (Hill & Tyson, 2009).

Research Questions

This study specifically seeks to answer two research questions:

- (1) What are the inequality profiles of students concerning FCISP based on gender, economic-social-cultural status (ESCS), and migration status?
- (2) How do contextual variables like school type, parental involvement, and school student gender composition affect students' engagement in FCISP?

Data and Methods

Data from 221,828 students in 11,525 schools across 50 countries were analyzed using multilevel modeling. This approach considered individual factors like school belonging, teacher-student relationships, educational expectations, and curiosity. It also examined how school characteristics, such as public versus private, gender composition, and parental involvement, shape students' engagement in seeking career information.

Results

The findings reveal that social integration within schools plays a significant role in influencing FCISP. Male students and those from higher ESCS backgrounds tend to engage more in FCISP, highlighting a persistent socioeconomic divide. After accounting for the gender composition of schools, ESCS remained a significant factor, suggesting that socioeconomic disparities strongly influence students' social integration and access to career information. Furthermore, public school students and those with higher parental involvement showed greater engagement in FCISP, indicating that institutional and familial support structures are crucial for fostering social integration and enabling access to career-related information. The presence of a higher proportion of male students in schools also correlated with increased FCISP, suggesting gender dynamics within educational settings can influence social integration and access to career resources.

Passungsverhältnisse zwischen Lehrkräften und Studierenden im zweiten Bildungsweg

Claudia Schuchart, Doris Bühler-Niederberger

Bergische Universität Wuppertal

Einleitung: In Deutschland können junge Erwachsene, die die allgemeinbildende Schule mit ohne Abitur verlassen haben, über den zweiten Bildungsweg (ZB) die Berechtigung zum Studium erlangen. Der gymnasiale ZB weist jedoch hohe Abbruchzahlen auf, was im Zusammenhang mit einem für diesen Bildungsweg charakteristischen Spannungsverhältnis gesehen werden muss: Es gelten die Leistungsanforderungen des allgemeinbildenden Gymnasiums, deren Erreichung in standardisierten Abschlussprüfungen überprüft wird. Ein hoher Anteil der Studierenden weist jedoch gebrochene Bildungsbiografien sowie Lern- und Persönlichkeitsprobleme auf (Bellenberg et al. 2019). Wir untersuchen, wie Lehrkräfte in ihrer Praxis auf diese Spannung reagieren und inwieweit es zu einer Passung zwischen der Praxis der Lehrkräfte mit den Erwartungen der Studierenden kommt.

Theorie: Soziologisch lässt sich dieses Spannungsverhältnis als ein Konflikt zwischen „Institution“ und „Person“ beschreiben (Dreeben 1968), da insbesondere vulnerable Lernende Erwartungen an die Anerkennung ihrer Bedürfnisse, z. B. die Anerkennung ihrer persönlichen Situation, haben, die den Anforderungen eines gymnasialen Bildungsgangs widersprechen können. Lehrkräfte können dieses Spannungsverhältnis nicht auflösen, sondern es nur bearbeiten (Helsper 2004; 2010). Sie entwickeln Praktiken dahingehend, ob und wie sie den Bedürfnissen ihrer Lernenden in konkreten Situationen nachgeben (Collins & Pratt 2011). Wir untersuchen drei Aspekte dieser Praxis, die eher „personen-“ oder „institutionenorientiert“ sein können: a) Unterstützung schwacher Lernender vs. Erwartung selbstständigen Lernens; b) Berücksichtigung persönlicher Umstände bei Benotung und Verhaltensregeln vs. Gleichbehandlung aller; c) Interesse an der ganzen „Person“ vs. Interesse am „Studierenden“ als Ausschnitt der Person. Wir fragen: 1) Wie gehen Lehrkräfte in ihrer Praxis mit Person-Institution-Konflikten um, und lässt sich eine dominante Praxis des gymnasialen ZB erkennen? Die Entscheidung für eine bestimmte Praxis kann nicht für alle Lernenden gelten, da diese sich in Abhängigkeit von individuellen Merkmalen (Schuchart & Bühler-Niederberger 2020) darin unterscheiden, wie stark sie die Berücksichtigung ihrer persönlichen Bedürfnisse in der Schule erwarten. Im zweiten Schritt untersuchen wir: Inwieweit und für welche Schülergruppen kommt es am gymnasialen ZB zu einer Passung zwischen der Praxis der Lehrkräfte den Erwartungen der Studierenden?

Methoden: Wir verwenden einen Datensatz von 176 Lehrkräfte und 600 Studierenden in 9 Schulen des gymnasialen ZB in Nordrhein-Westfalen. Für Frage 1 verwenden wir 11 Items zur Lehrpraxis (Seidel 2003), die sich auf die oben beschriebenen Aspekte beziehen und den Lehrenden vorgelegt wurden. Für Frage 2 verwenden wir 5 selbst entwickelte bipolare Items, die sich auf diese Aspekte beziehen und sowohl von Lehrkräften als auch von Studierenden beantwortet wurden. Zur Ermittlung der Gruppenzugehörigkeit verwenden wir Latent-Class-Analysen, die wir für Frage 1 im Sample der Lehrenden und für Frage 2 im gemeinsamen Sample der Lehrenden und Lernenden durchführen.

Ergebnisse: Zwei Drittel der Lehrkräfte gibt an, dass sie schwache Studierende besonders unterstützen und sich um Studierende als „Personen“ kümmern, während sie auf die Einhaltung allgemeiner Regeln für Verhalten und Benotung bestehen (Frage 1). Die gemeinsame Analyse der Lehrenden und Studierenden (Frage 2) ergibt jedoch, dass nur knapp 40% der Studierenden Erwartungen an die Lehrkräfte hat, die diese mit ihrer Praxis bedienen können. Fast die Hälfte der Studierenden – und damit deutlich mehr, als Lehrkräfte annehmen – wünscht sich mehr Anerkennung ihres Erwachsenenstatus, vor allem durch die Gewährung von Autonomie, professioneller Beziehungsdistanz und Anerkennung ihrer persönlichen Situation hinsichtlich der Verhaltensregeln.

Bedeutung: Die Ergebnisse deuten auf ein Dilemma hin: Wenn Lehrkräfte auf den signifikanten Anteil der Lernenden eingehen, die Anleitung, Unterstützung und Betreuung benötigen, weisen sie möglicherweise diejenigen Lernenden zurück, die erwarten, dass ihre Rolle als Erwachsene ernst(er) genommen wird (Ross & Gray 2015). Diese Art von Passungsproblemen könnte mit einer höheren Abbruchneigung im Zusammenhang stehen (Schuchart & Schimke 2022). Es stellt sich die Frage, ob dieses Dilemma durch eine stärkere Orientierung an Strukturen, Materialien und Methoden der Erwachsenenbildung sowie begleitenden Unterstützungssystemen für diejenigen, die diese benötigen, besser bearbeitet werden kann.

Entwicklung von Berufsaspirationen im nachobligatorischen Bildungsweg

Sara Möser
Universität Bern

Persönlichkeit und Identität gelten als die wichtigsten Einflussfaktoren für die Entwicklung von Berufsaspirationen in jungen Jahren (Gottfredson, 2002). Während jedoch grundlegende Persönlichkeitsmerkmale im Jugendalter als relativ stabil gelten (Hirschi, 2010; Low et al., 2005), ist davon auszugehen, dass im nachobligatorischen Bildungserwerb sekundäre Sozialisationsprozesse stattfinden, die Lernende in ihrer neuen Rolle prägen und ihre Identität weiter formen. Bildungsinstitutionen wie Schulen und Ausbildungsbetriebe stellen somit zentrale soziale Kontexte und Sozialisationsräume dar, in denen Individuen ihre beruflichen Ambitionen ausbilden und weiterentwickeln (Heckhausen & Shane, 2015). Darüber hinaus erhalten die Lernenden im Verlauf der nachobligatorischen Bildungslaufbahn weitere Informationen über die Zugänglichkeit von Karriereoptionen und lernen neue Möglichkeiten kennen. Aufgrund des engen Zusammenhangs zwischen Bildungsabschlüssen und dem beruflich segmentierten Arbeitsmarkt sind die Berufschancen in der Schweiz eng mit frühen Entscheidungen in der Bildungs- und Berufslaufbahn verknüpft (Buchmann & Sacchi, 1997). Institutionelle Regelungen definieren dabei sowohl die Möglichkeiten als auch die Einschränkungen auf dem Arbeitsmarkt.

Dieser Beitrag untersucht, wie sich die Berufswünsche junger Erwachsene am Übergang von der Schule in den Beruf entwickeln, mit einem besonderen Fokus auf die strukturierende Wirkung von nachobligatorischen Bildungsverläufen.

Zu diesem Zweck werden die Daten der DAB-Panelstudie verwendet (Becker et al., 2020). Das DAB-Panel erhebt seit 2012 Längsschnittdaten zur beruflichen und schulischen Situation junger Erwachsener in der Deutschschweiz. Seit dem achten Schuljahr wird eine Bruttostichprobe von knapp 4'000 jungen Erwachsenen begleitet, die in bisher elf Erhebungen zu ihrer aktuellen Ausbildungssituation sowie zu ihren Ausbildungs- und Berufswünschen befragt wurden. Die Berufswünsche wurden im Rahmen des DAB-Panels wiederholt abgefragt. In den ersten drei Wellen wurden die Schülerinnen und Schüler, die zu diesem Zeitpunkt die Sekundarstufe I besuchten, gefragt, welchen Beruf sie später ausüben möchten. Nach Abschluss der obligatorischen Schule wurden die Befragten in den späteren Wellen gebeten, anzugeben, welchen Beruf sie im Alter von 30 Jahren ausüben werden. Für die Analyse wird das Bildungsniveau des angestrebten Berufs anhand der Daten der Schweizer Strukturerhebung ermittelt. Der Wert gibt an, wie viel Prozent der Erwerbstätigen in diesem Beruf einen Tertiärabschluss haben. Die Skala reicht von 0.99% für Anwalt/-innen zu 0.05% für Lastwagenschaffner/-in.

Zur Modellierung der Entwicklung von Berufswünschen werden Wachstumskurvenmodelle berechnet. Latente Wachstumsmodelle schätzen die interindividuelle Varianz intraindividuellen Entwicklungsmuster über die Zeit. Durch die Erweiterung des Wachstumsmodells um ein Regressionsmodell werden die latenten Verlaufsparemeter als Ergebnisse unter dem Einfluss von Kovariaten behandelt. Dies ermöglicht es, den Einfluss individueller Ressourcen und struktureller Möglichkeiten auf individuelle Entwicklungsmuster zu untersuchen. Dabei richtet sich der Fokus insbesondere auf die Unterschiede im Aspirationsniveau- und der Aspirationsentwicklung nach Bildungsverlauf. Darüber hinaus werden das Geschlecht, der elterliche Bildungshintergrund sowie der Migrationshintergrund als Kontrollvariablen berücksichtigt.

Im Durchschnitt erhöhen die Befragten ihre Berufsaspirationen und streben mit Mitte zwanzig einen Beruf mit einer höheren Tertiärquote an (55%) als während der obligatorischen Schulzeit (29%). Dies kann auf die Durchlässigkeit des Schweizer Bildungssystems zurückgeführt werden. Sowohl auf dem beruflichen als auch auf dem akademischen Weg stehen den Schülerinnen und Schülern verschiedene Weiterbildungsmöglichkeiten zur Verfügung, und selbst Bildungsgänge mit geringen kognitiven Anforderungen verfügen über Anschlusslösungen, die eine Weiterqualifizierung ermöglichen. Diese Durchlässigkeit ist vermutlich ausschlaggebend, dass Jugendliche in der Schweiz ihre Berufsaspirationen stetig nach oben korrigieren, während die internationale Literatur eher eine Stagnation oder einen Abwärtstrend feststellt (Clark, 1960; Lee & Rojewski, 2009; Tomasik et al., 2009). Dafür spricht auch, dass es insbesondere die Jugendlichen in Ausbildungen mit Berufs- und Fachmaturitäten sind, welche den Aufwärtstrend tragen. Während die Berufsaspirationen von Gymnasiast:innen auf hohem Niveau stagnieren, steigen die Aspirationen von Auszubildenden in dualen Berufslehren mit niedrigen Anforderungen nur schwach an.

Promovieren mit einer gesundheitlichen Beeinträchtigung – Welche Rollen spielen rationale Abwägungen und die soziale Integration für die Promotionsintention?

Ulrike Schwabe¹, Jannis Burkhard², Charlotte Bröder³, Benjamin Gröschl⁴

¹DZHW | Deutsches Zentrum für Hochschul- und Wissenschaftsforschung, ²DIPF | Leibniz Institut für Bildungsforschung und Bildungsinformation, ³Freie Universität Berlin, ⁴Ludwig-Maximilians-Universität München

Unter allen Studierenden sind aktuell fast 24 % gesundheitlich beeinträchtigt, bei 16 % wirkt sich diese Beeinträchtigung erschwerend auf das Studium aus (Steinkühler et al. 2023; Gerdes et al. 2024). Aus der vorhandenen Literatur wissen wir, dass Studierende mit gesundheitlichen studienerschwerenden Beeinträchtigungen ein höheres Risiko für einen Studienabbruch haben (Rußmann et al. 2023, 2024), seltener nach dem Bachelorabschluss ein Masterstudium anstreben (Steinkühler et al. 2023), häufiger von Diskriminierungserfahrungen im Studium berichten (Gerdes et al. 2024) sowie in geringerem Maß im Studium sozial integriert sind (Rußmann et al. 2024).

Insbesondere für die Promotionsphase ist die empirische Befundlage noch überschaubarer (beispielsweise Rhein et al. 2024), was nicht zuletzt auf die Datenverfügbarkeit zurückzuführen ist. Folglich mangelt es bisher an gesicherten Erkenntnissen darüber, ob und warum sich Studierende mit und ohne Beeinträchtigung hinsichtlich der Promotionsaufnahme unterscheiden.

Deshalb untersucht dieser Beitrag, welche Faktoren für Studierende mit einer gesundheitlichen studienerschwerenden Beeinträchtigung besonders relevant sind nach dem Master- oder Staatsexamensabschluss eine Promotion anzuschließen. Wir untersuchen insbesondere, ob die Bedeutung der Einflussfaktoren für die Promotionsintention nach Art der Beeinträchtigung und Schwere der Beeinträchtigung variiert, wie dies beispielsweise für die Abbruchintentionen gezeigt werden konnte (Rußmann et al. 2024).

Der Beginn einer Promotion ist ein sehr später Übergang im Bildungssystem, an dem leistungsabhängige und leistungsunabhängige Faktoren eine Rolle spielen (Boudon 1974). Unabhängig von den objektiven Studienleistungen kann die Promotionsaufnahme über rationales Entscheidungsverhalten (Bergann et al. 2019; Breen & Goldthorpe 1997) oder die soziale Integration in die Hochschule (Tinto, 1975) erklärt werden. Bei beiden theoretischen Mechanismen gehen wir von Unterschieden in der relativen Stärke zwischen Studierenden mit und ohne Beeinträchtigung sowie unterschiedlicher Beeinträchtigungsarten und der Schwere der Beeinträchtigung aus (Moderatorhypothesen).

Basierend auf den Daten von „Die Studierendenbefragung in Deutschland (2021)“ vergleichen wir die Promotionsintentionen von Studierenden mit (n = 9.405) und ohne (n = 57.809) eine gesundheitliche studienerschwerende Beeinträchtigung sowie

unterschiedlicher Beeinträchtigungsarten und Schweregrade innerhalb der Gruppe der Studierenden mit einer gesundheitlichen studienerschwerenden Beeinträchtigung miteinander (Becker et al., 2023; Beuße et al., 2022). Den differentiellen Einfluss verschiedener Erklärungsfaktoren überprüfen wir mittels Interaktionen in logistischen Regressionsmodellen. Die Ergebnisse für den Inter- und Intragruppenvergleich werden als mittelwertzentrierte Marginaleffekte präsentiert.

Es zeigt sich zunächst, dass sich die Promotionsintentionen von Studierenden mit und ohne Beeinträchtigung nicht signifikant voneinander unterscheiden ($n = 46.631$; $\beta = -0.05$; $p = 0.07$): In beiden Gruppen erwägen etwa 30 % der Studierenden nach dem Master- oder Staatsexamensabschluss eine Promotion aufzunehmen. Jedoch zeigen Studierende mit einer psychischen Beeinträchtigung (28,5 %), einer gleichschweren Mehrfachbeeinträchtigung (26,8 %) und einer anderen Beeinträchtigung (26,5 %) besonders niedrige Promotionsintentionen. Sofern Studierende aus einem akademischen Elternhaus kommen, planen sie signifikant häufiger zu promovieren ($n = 45.662$; $\beta = 0.45^{***}$). Sowohl Kosten-Nutzen-Abwägungen als auch die soziale Integration in die Hochschule sind für verschiedene Gruppen unterschiedlich relevant. Unsere Analysen zeigen aber auch, dass die Promotionsabsicht insbesondere auch von vorherigen Bildungsentscheidungen und -erfolgen abhängt, beispielsweise davon, ob an einer Universität oder Hochschule für angewandte Wissenschaften studiert wird sowie von den vorherigen Abschlussnoten.

Zusammengefasst verdeutlichen unsere Befunde, dass die Gruppe der Studierenden mit einer gesundheitlichen studienerschwerenden Beeinträchtigung heterogen ist und sich hinsichtlich der Beeinträchtigungsart und der Schwere der Beeinträchtigung teilweise deutlich unterscheidet. Umfassendes empirisches Wissen über Studierende mit einer gesundheitlichen studienerschwerenden Beeinträchtigung ist wichtig für Entscheidungsträger an Hochschulen, um die Hochschulbildung in Deutschland – und insbesondere den Zugang zur Doktorand:innenausbildung – inklusiver zu gestalten.

S-98: Beobachtungsbasierte Erfassung von Unterrichtsqualität: Auf zu neuen Wegen?

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: EO 159

Chair der Sitzung: Anna-Katharina Praetorius

Chair der Sitzung: Thilo Kleickmann

Chair der Sitzung: Richard Göllner

Beobachtungsbasierte Erfassung von Unterrichtsqualität: Auf zu neuen Wegen?

Chair(s): **Anna-Katharina Praetorius** (Universität Zürich), **Thilo Kleickmann** (Universität Kiel)

Diskutant*in(nen): **Richard Göllner** (Universität Tübingen)

In den letzten Jahren haben sich bei der Erfassung von Unterrichtsqualität mittels Beobachter*innen-Einschätzungen im deutschsprachigen Raum bestimmte Vorgehensweisen etabliert. Dazu gehört beispielsweise die Auswahl einer Gelegenheitsstichprobe von Studierenden als Rater*innen, das standardmässige Durchlaufen von Rater*innen-Trainings und das anschliessende Rating gesamter Unterrichtsstunden mittels eines hoch-inferenten Zugangs. Wenngleich mit dem Etablieren solcher Standards viele Vorteile einhergehen, stellt sich die Frage, inwiefern diese Setzungen mit den oft inkonsistenten Befunden der Unterrichtsqualitätsforschung in Zusammenhang stehen. Entsprechend scheint es an der Zeit zu sein, diese Setzungen zu überprüfen. Das Symposium beleuchtet aus fachdidaktischen und pädagogisch-psychologischen Perspektiven drei zentrale Setzungen, die in der Forschung zu Unterrichtsqualität mittels Beobachter*innen-Ratings üblicherweise vorgenommen werden.

In Vortrag 1 wird die Annahme in den Blick genommen, dass eine Gelegenheitsstichprobe von Rater*innen eine gute Ausgangsbasis für durchzuführende Unterrichtseinschätzungen darstellt. Dazu wird untersucht, inwiefern sich Ratings und zugehörige Begründungen von fachspezifischen versus fachfremden Rater*innen in Bezug auf die Unterrichtsqualitäts-Dimensionen kognitive Aktivierung und konstruktive Unterstützung in typischen Situationen des Mathematikunterrichts unterscheiden. Es zeigt sich, dass sich die Ratings der generischen Beobachtungskriterien zwar nicht zwischen den beiden Gruppen unterscheiden, wohl aber z.T. die Begründungen der Ratings.

In Vortrag 2 wird die übliche Vorgehensweise von Rater*innen-Trainings hinterfragt, bei der neben einem theoretischen Input an Beispielvideos geübt und deren Einschätzungen in Gruppendiskussionen besprochen werden. Konkret wird in der Studie die mittels Begründungen der Rater*innen erfasste Validität der Ratings für zwei Durchführungsarten kontrastiert: die übliche und vergleichsweise aufwändige Vorgehensweise mit gemeinsamen Arbeits- und Diskussionsphasen sowie ein reiner Selbstlernkurs. Es zeigen sich keine eindeutigen Unterschiede zwischen den beiden Ansätzen.

In Vortrag 3 wird die Annahme hinterfragt, Unterrichtsqualität global in Bezug auf gesamte Unterrichtsstunden einzuschätzen. Stattdessen wird die Einschätzung der Unterrichtsqualität von Unterrichtsphasen vergleichend in den Blick genommen. Dabei wird untersucht, inwiefern das eingeschätzte Niveau einzelner Qualitätsdimensionen sowie die Reliabilität und prädiktive Validität sich zwischen klassenöffentlichen Phasen und Schüler*innen-Arbeitsphasen unterscheiden. Die (bislang vorliegenden) Befunde zeigen, dass je nach Unterrichtsqualitätsdimensionen sich sowohl das Niveau als auch die Reliabilität der Einschätzungen voneinander unterscheiden.

Beiträge des Symposiums

Beobachtungsratings von Mathematikunterricht – Welche Rolle spielt der fachliche Hintergrund von Rater:innen?

Linn Hansen¹, Marita Friesen², Anika Dreher¹

¹Pädagogische Hochschule Freiburg, ²Pädagogische Hochschule Heidelberg

Theoretischer Hintergrund und Fragestellung

Beobachtungsratings werden häufig zur Erfassung von Unterrichtsqualität eingesetzt und teilweise sogar als Königsweg bezeichnet (vgl. Helmke, 2021). In verschiedenen Studien zeigte sich allerdings, dass Beobachtungsratings Validitätsproblemen unterliegen. Beispielsweise variieren Ratings von verschiedenen Beobachtenden häufig erheblich (Bell et al., 2015). Ein naheliegender Grund hierfür sind unterschiedliche Wahrnehmungen und Interpretationen von Unterrichtsereignissen, die als unterschiedliche Noticingprozesse der Rater:innen gedeutet werden können (Dreher & Leuders, 2021; Praetorius et al., 2012). Ein Blick in die Noticingforschung zeigt, dass Noticingprozesse professionelles Wissen erfordern (Stahnke et al., 2016), insbesondere fachliches und fachdidaktisches Wissen (Dreher & Kuntze, 2015). Dies wirft die Frage auf, ob fachfremde Rater:innen die Qualität von Fachunterricht mit fachspezifischen Lernzielen und Lernprozessen adäquat einschätzen können (Lindmeier & Heinze, 2020). Diese Frage stellt sich insbesondere für Qualitätsaspekte, denen eine hohe Fachspezifität zugeschrieben wird (Aspekte bezüglich kognitiver Aktivierung und konstruktiver Unterstützung) (Schlesinger et al., 2018). Bislang wurde die Rolle des fachlichen Hintergrunds der Rater:innen in der Diskussion um Beobachtungsratings jedoch nicht ausreichend in den Blick genommen. Unsere Studie setzt hier an, indem folgende Forschungsfrage untersucht wird: Inwiefern unterscheiden sich Ratings und zugehörige Begründungen von fachspezifischen versus fachfremden Rater:innen in Bezug auf Unterrichtsqualitätsmerkmale der Dimensionen kognitive Aktivierung und konstruktive Unterstützung in typischen Situationen des Mathematikunterrichts?

Methode

Hierfür wurden im Rahmen des UFB-Videoprojekts Mathematik in Kooperation mit dem Institut für Bildungsanalysen Baden-Württemberg und der Universität Tübingen sechs geskriptete Videos von typischen Situationen im Mathematikunterricht entwickelt. Jedes Video dauert ca. 10 Minuten und fokussiert einen Unterrichtsqualitätsaspekt der Dimensionen kognitive Aktivierung und konstruktive Unterstützung nach dem Unterrichtsfeedbackbogen (Fauth et al., 2021). Die Videos wurden so erstellt, dass sie jeweils ein typisches kritisches Ereignis enthalten, das aus fachdidaktischer Sicht bezüglich des fokussierten Qualitätsaspekts negativ einzuschätzen ist. Beispielsweise liegt das kritische Ereignis für den Aspekt „Qualität des Feedbacks“ darin, dass die Lehrkraft einen typischen Fehler einer Schülerin zwar aufgreift und im Unterrichtsgespräch thematisiert, den Fehler aber nur oberflächlich korrigiert. Eine Auseinandersetzung mit der dahinterliegenden Fehlvorstellung und ihrer Überwindung findet nicht statt.

Zur Datenerhebung wurden die Unterrichtsvideos Master-Lehramtsstudierenden mit unterschiedlichem professionellem Hintergrund (fachspezifisch vs. fachfremd) in einer Onlinebefragung vorgelegt. Die Rater:innen waren nicht spezifisch geschult, hatten aber bereits mindestens zwei Fachdidaktikveranstaltungen (fachspezifische Rater:innen) oder zwei Veranstaltungen zu Themen der Unterrichtsforschung besucht (insbesondere fachfremde Rater:innen). Nach jedem Video sollten die Teilnehmenden

das entsprechende generische Beobachtungskriterium aus dem Unterrichtsfeedbackbogen auf einer 4-Punkt-Likertskala raten und ihre Einschätzung begründen (z.B.: „Das Feedback, das die Lehrkraft den Schülerinnen und Schülern gibt, ist zum Weiterlernen hilfreich“ für den Aspekt „Qualität des Feedbacks“). Die Begründungen der Ratings wurden mit Hilfe qualitativer Inhaltsanalyse danach codiert, ob das kritische Ereignis erkannt wurde oder nicht. Anschließend wurden die Ratings und die Begründungen der beiden Gruppen (fachspezifisch vs. fachfremd) verglichen.

Vorläufige Ergebnisse und Diskussion

Erste Daten von N= 50 Lehramtsstudierenden (30 fachspezifisch; 20 fachfremd) führen zu folgenden Ergebnissen: Die Ratings der generischen Beobachtungskriterien unterscheiden sich nicht signifikant zwischen den beiden Gruppen. Bezüglich der Begründungen der Ratings zeigen sich jedoch teilweise Unterschiede: Die fachspezifischen Rater:innen führten in zwei der sechs Videos in ihren Begründungen signifikant häufiger das jeweilige kritische Ereignis an als die fachfremden Rater:innen (Fisher's Exact Test, Video 1: $p = .007$, $\phi = .436$; Video 6: $p = .016$, $\phi = .381$).

Diese vorläufigen Ergebnisse zeigen zwar gewisse Unterschiede in den Begründungen, allerdings scheint sich das Erkennen der kritischen Ereignisse nicht maßgeblich in den Ratings widerzuspiegeln. Auch wenn diese Ergebnisse nur mit Vorsicht zu interpretieren sind, könnten mögliche Gründe in einer mangelnden spezifischen Schulung der Teilnehmenden oder in einer fehlenden Fachspezifität der Beobachtungskriterien vermutet werden. In jedem Fall sind weiterführende Untersuchungen zur Klärung der Rolle des fachlichen Hintergrunds von Rater:innen hinsichtlich der Validitätsprobleme von Beobachtungsratings von Unterrichtsqualität notwendig.

Einfluss eines Ratertrainings auf die Reliabilität und Validität von Unterrichtsqualitätsratings

Julia Larissa Blank¹, Vanessa Weißhaupt², Tosca Daltoè³, Evelin Ruth-Herbein¹, Benjamin Fauth³

¹Institut für Bildungsanalysen Baden-Württemberg, ²Universität Tübingen, ³Institut für Bildungsanalysen Baden-Württemberg, Hector-Institut für Empirische Bildungsforschung, Universität Tübingen

Theoretischer Hintergrund

Als Zugang zur Unterrichtsqualität wird in Forschung oder Schul- und Unterrichtsentwicklungsprozessen häufig die externe Beobachtung gewählt. Auch wenn externe Beobachtende als Königsweg zur Erfassung von Unterrichtsqualität bezeichnet wurden (Helmke, 2009), gehen objektive, reliable und valide Einschätzungen mit gewissen Herausforderungen einher. Beispielsweise ist für einzelne Unterrichtsqualitätsaspekte eine gewisse fachliche Expertise notwendig (Dreher & Leuders, 2021; Fauth & Leuders, 2021). Zudem lassen sich Unterrichtsqualitätsaspekte unterschiedlich gut beobachten (Kane & Staiger, 2012). Darüber hinaus ist jede Art der pädagogischen Diagnostik durch Urteilsfehler behaftet (Whiting et al., 2004).

Um trotz dieser Herausforderungen objektive, reliable und valide Unterrichtsqualitätsratings zu erreichen, sollten Beobachtende ausreichend geschult werden (Praetorius et al., 2013). Doch obwohl solche Schulungen vielfach eingesetzt werden, ist bislang wenig darüber bekannt, wie diese gestaltet sein müssen, um Unterrichtsbeobachtungen bestmöglich zu unterstützen (Gitzi et al., 2023).

Hier setzt die vorliegende Studie an und untersucht Unterrichtsqualitätsratings im Kontext einer Schulung für den Unterrichtsfeedbackbogen Tiefenstrukturen (UFB; Fauth et al., 2022), einem empirisch untersuchten Unterrichtsbeobachtungsbogen aus Baden-Württemberg (Maier et al., 2021, 2022). In zwei systematisch variierten Qualifizierungskursen wurden Lehrkräfte im Umgang mit dem Instrumentarium geschult. Folgende Forschungsfragen stehen im Zentrum dieser randomisiert-kontrollierten Studie:

1. Führt das Durchlaufen des Qualifizierungskurses zu reliablen und validen Einschätzungen der Unterrichtsqualität?
2. Unterscheiden sich Teilnehmende unterschiedlicher Durchführungsarten des Qualifizierungskurses systematisch voneinander?

Methode

An der Studie nahmen 92 Lehrkräfte unterschiedlicher Fachbereiche teil (N weiblich = 61; N männlich = 23; Berufserfahrung: M = 19.08 Jahre, SD = 7.5). Die Teilnehmenden wurden randomisiert einer von zwei Durchführungsarten eines Qualifizierungskurses zugewiesen und durchliefen diesen entweder als reinen Selbstlernkurs oder als Selbstlernkurs ergänzt um gemeinsame Arbeits- und Diskussionsphasen. Die Teilnehmenden schätzten vor und nach dem Qualifizierungskurs 10-minütige Videosequenzen der TALIS-Videostudie Deutschland (Grünkorn et al., 2020) hinsichtlich ihrer Unterrichtsqualität ein. Die Einschätzung der 11 UFB-Items erfolgte anhand eines geschlossenen Ratings auf einer vierstufigen Likert-Skala. Zur Auswertung der Beobachtungsübereinstimmung wird der Average Deviation Index (ADM; Burke et al., 1999) herangezogen.

Zudem wurden stichwortartige offene Begründungen für das vergebene Rating erfragt. Die Auswertung der offenen Begründungen erfolgte anhand eines Kodierleitfadens (angelehnt an Irion, 2023). Anhand einer Haupt- sowie einer Unterkategorie wurde kodiert, ob es sich bei der gegebenen offenen Begründung um eine valide Antwort im Sinne des Instrumentariums handelt: 1. Greift die Antwort einen der Indikatoren zum jeweiligen Item aus dem UFB sinngemäß auf? (Hauptkategorie); 2. Welche anderen, nicht validen Antworten werden in den schriftlichen Begründungen genannt? (Unterkategorie). Die Kodierung erfolgte durch eine geschulte Raterin. 20% der offenen Begründungen wurden von einer zweiten Raterin kodiert. Die prozentuale Übereinstimmung der Hauptkategorie liegt bei 92%, zusätzlich ist eine gute Interrater-Reliabilität von $\kappa = .795$ gegeben. Bei abweichenden Kodierungen schloss sich eine Konsensbildung an.

Ergebnisse und ihre Bedeutung

Vorläufige Ergebnisse zeigen keine eindeutigen Unterschiede in der Beobachtungsübereinstimmung zwischen den Durchführungsarten (ADM Posttest Selbstlerngruppe: Min = 0.28, Max = 0.74; ADM Posttest Austauschgruppe: Min = 0.33, Max = 0.75). In beiden Durchführungsarten zeigen sich tendenziell positive Entwicklungen vom Prätest zum Posttest. Die Beobachtungsübereinstimmung unterscheidet sich stark zwischen den Items: Manche Aspekte wie z.B. Störungen (ADM = 0.31) oder Wertschätzung/Respekt (ADM = 0.46) wurden übereinstimmender beobachtet als andere Aspekte wie z.B. herausfordernde Aufgaben (ADM = 0.71) oder individuelle Unterstützung (ADM = 0.74). Erste Auswertungen der offenen Rückmeldungen legen nahe, dass sich bei Items mit geringer Beobachtungsübereinstimmung (ADM) auch weniger valide Antworten zeigen. Weitere Auswertungen stehen noch aus und sollen bei der Tagung im Januar 2025 vorgestellt werden.

Die Befunde liefern Hinweise, welche Unterrichtsqualitätsaspekte besonderen Schulungsbedarf aufweisen und wie die Herausforderungen, reliable und valide Einschätzungen der Unterrichtsqualität zu erhalten, im Rahmen von Raterschulungen adressiert werden können. Limitationen sowie Implikationen für Forschung und Praxis werden diskutiert.

Variabilität von Unterrichtsqualität zwischen klassenöffentlichen Phasen und Schüler:innen-Arbeitsphasen

Annelie Thom¹, Anna-Katharina Praetorius², Thilo Kleickmann¹

¹Universität Kiel, ²Universität Zürich

Theoretischer Hintergrund

Unterrichtsqualität wird oft als ein für Lehrkräfte relativ stabiles Merkmal angenommen. Allerdings zeigt sich, dass Unterrichtsqualität nicht nur zwischen Lehrkräften variiert, sondern auch bei derselben Lehrkraft über die Zeit, und zwar zwischen wie auch innerhalb von Unterrichtsstunden (Gabriel-Busse & Lipowsky, 2021; Jentsch et al., 2020; Kane & Staiger, 2012; Praetorius et al., 2014). Hess und Lipowsky (2020) fanden beispielsweise, dass das kognitive Frageniveau von Lehrkräften zwischen klassenöffentlichen Phasen (KP) und Schüler:innen-Arbeitsphasen (SAP) variierte. Unterschiede in der Unterrichtsqualität zwischen Unterrichtsphasen könnten in den unterschiedlichen unterrichtlichen Anforderungen an die Lehrkraft begründet liegen. Viele parallele Ereignisse in einer SAP könnten eine Lehrkraft beispielsweise im Hinblick auf Monitoring oder individuelle Lernunterstützung anders beanspruchen als Prozesse im gemeinsamen Unterrichtsgespräch (Breidenstein, 2014; Doyle, 2006). Entsprechend könnte auch die Qualität der Unterrichtseinschätzungen zwischen Unterrichtsphasen variieren (Dreher & Leuders, 2021). Zum Beispiel dürften Interaktionen zwischen Lehrkraft und Schüler:innen in KP besser verständlich sein als in SAP. Diese Annahmen wurden allerdings noch nicht systematisch empirisch für verschiedene Basisdimensionen der Unterrichtsqualität geprüft.

Vor diesem Hintergrund untersucht die vorliegende Studie die Variabilität von Basisdimensionen der Unterrichtsqualität innerhalb einer Unterrichtsstunde in Abhängigkeit von Unterrichtsphasen (KP und SAP). Dabei wird erstens die Frage nach der Ausprägung der Unterrichtsqualität und zweitens die Frage nach der Reliabilität und Validität der Beurteilungen durch Beobachter:innen im Vergleich von KP und SAP untersucht.

Methode

Zunächst wurden N = 60 Unterrichtsvideos (Sachunterricht zum Thema natürlicher Wasserkreislauf, 4. Klasse) in KP und SAP sequenziert (Hugener, 2006). Die Gesamtzahl der Phasen variierte zwischen Min = 2 und Max = 11, wobei mit 39,7% der Großteil der Unterrichtsstunden aus drei Phasen bestand. Die Phasen wurden von zwei geschulten Rater:innen hinsichtlich der Dimensionen Klassenführung, motivational-emotionale Unterstützung, kognitive Aktivierung sowie Auswahl und Thematisierung des Inhalts beurteilt (Rogh et al., 2020). Bisher sind Daten von N = 30 Videos vorhanden, zur Tagung wird der vollständige Datensatz vorliegen. Die Interraterreliabilitäten (relative Generalisierbarkeitskoeffizienten) über alle Phasen der Unterrichtsstunden lagen zwischen $G = .69$ bei kognitiver Aktivierung und $G = .84$ bei Auswahl und Thematisierung des Inhalts. Für die Analysen auf Ebene der Unterrichtsphasen wurden die Beurteilungen separat über KP und SAP aggregiert.

Ergebnisse und ihre Bedeutung

Deskriptive Analysen ergaben geringe Unterschiede im Niveau der Unterrichtsqualität zwischen KP und SAP für Klassenführung (Cohen's $d = .10$) und kognitive Aktivierung ($d = .06$). Größere Unterschiede lagen bei motivational-emotionaler Unterstützung ($d = .60$) und Auswahl und Thematisierung des Inhalts ($d = .88$) vor. Klassenführung in KP und SAP korrelierte hoch ($r = .70^{***}$), motivational-emotionale Unterstützung ($r = .40^*$) sowie Auswahl und Thematisierung des Inhalts ($r = .53^{**}$) korrelierten moderat. Bei kognitiver Aktivierung war die Korrelation nicht signifikant ($r = .02$). Die Interraterreliabilitäten variierten zwischen KP und SAP am stärksten bei motivational-emotionaler Unterstützung ($G = .67$ in KP vs. $G = .42$ in SAP) sowie Auswahl und Thematisierung des Inhalts ($G = .82$ in KP vs. $G = .64$ in SAP), weniger jedoch bei Klassenführung ($G = .56$ in KP vs. $G = .58$ in SAP) und kognitiver Aktivierung ($G = .66$ in KP vs. $G = .72$ in SAP). Ergebnisse zur prädiktiven Validität liegen zur Tagung vor.

Die vorläufigen Ergebnisse deuten darauf hin, dass vor allem motivational-emotionale Unterstützung sowie Auswahl und Thematisierung des Inhalts zwischen KP und SAP variieren, was sich auch in unterschiedlichen Interraterreliabilitäten widerspiegelt. Klassenführung zeigte sich relativ stabil in Bezug auf Niveau- und Rangstabilität. Die Ergebnisse deuten somit auf Abhängigkeiten zwischen Oberflächen- (Unterrichtsphasen) und bestimmten Tiefenmerkmalen (Basisdimensionen) hin (Decristan et al., 2015; Hess & Lipowsky, 2020) und legen eine phasenbezogene Beurteilung nahe. Weitere inhaltliche wie methodische Implikationen werden im Vortrag diskutiert.

S-99: Herausforderungen in der Entwicklung von neuen Lehr-Lernformaten am Beispiel der Universitätsschule Dresden

Zeit: Mittwoch, 29.01.2025: 15:15 - 17:00 · Ort: M 003

Chair der Sitzung: Anke Langner

Chair der Sitzung: Birgit Lütje-Klose

Herausforderungen in der Entwicklung von neuen Lehr-Lernformaten am Beispiel der Universitätsschule Dresden

Chair(s): **Anke Langner** (TU Dresden, Deutschland)

Diskutant*in(nen): **Birgit Lütje-Klose** (Universität Bielefeld, Deutschland)

Veränderte schulische Bildungsprozesse könnten eine Lösung für die bestehenden gesellschaftlichen Herausforderungen darstellen. Selbstbestimmung und auch Solidaritätsfähigkeit im Sinne von Klafki stärker in den Mittelpunkt der schulischen Bildung zu stellen, wird als ein möglicher Weg gesehen, vor allem dann, wenn die damit verbundenen schulischen Veränderungsprozesse evidenzbasiert gestaltet werden. Am Beispiel der Universitätsschule Dresden sollen die Herausforderungen eines solchen Gestaltungsprozesses bezogen auf die Entwicklung von innovativen Lehr-Lernformaten anhand von drei Beiträgen nachgezeichnet und diskutiert werden.

In der Universitätsschule werden über die letzten fünf Jahre vor allem Lernformate entwickelt, die ein stärker selbstreguliertes Lernen der Schüler:innen unterstützen (Langner & Pesch, 2024) wie sie aber auch die Kooperation zwischen den Schüler:innen fördern (Projektgruppe der Universitätsschule, Stand: 2022). Das zentrale Unterrichtsformat ist die Projektarbeit (Langner, Heß & Wiechmann, 2021). Der auf 15 Jahre angelegte Schulversuch wurde durch die TU Dresden beantragt und wird in einem ko-konstruktiven Prozess von Schule und Wissenschaft gestaltet. Wissenschaftlich begleitet impliziert dabei auch, dass die TU Dresden die Möglichkeit hat, alle im Kontext des Schulversuches erzeugten Daten zu nutzen. Sowohl mit qualitativen als auch quantitativen Daten sollen die Herausforderungen wie auch bestehende Befürchtungen, die mit der Gestaltung neuer Lehr-Lernformate einhergehen, im Rahmen des Symposiums nachgezeichnet und diskutiert werden. Angelehnt an das Modell von Schulentwicklung nach Rolff (2019) wird der Annahme gefolgt, dass Unterrichtsentwicklung – die Etablierung von neuen Lehr-Lernformaten zur Kompetenzentwicklung der Schüler:innen – in Wechselwirkung mit Personal- und Organisationsentwicklung stehen. Dementsprechend werden in einem ersten Beitrag die empfundenen Herausforderungen aus Sicht der Schüler:innen rekonstruiert, womit auch Fragen der Professionalisierung aufgeworfen werden. Die mit der Entwicklung von neuen Lehr-Lernformaten für eine gelingende Kompetenzentwicklung von Schüler:innen notwendigen organisatorischen Veränderungen werden auf der Basis von Chatdaten zwischen Schulleitung und wissenschaftlicher Leitung rekonstruiert (Beitrag zwei). Mit der Perspektive der Akteure der Projektarbeit Schüler:innen und Lehrer:innen und dem Blick auf den Steuerungsprozess, kann die Umsetzung von Veränderungen exemplarisch charakterisiert werden und Risiken für den Umsetzungsprozess eingeschätzt werden. Um die Risiken des Verlusts an Grundfertigkeiten abwägen zu können, werden abschließend im Rahmen des Symposiums die über drei Jahre erhobenen standardisierten Leistungstest in Mathematik (DEMAT, bspw. Roick et al. 2018) und den orthographischen Kompetenzen (Hamburger Schreibprobe, May et al. 2019) auch zu Vergleichsschulen in Dresden diskutiert.

Beiträge des Symposiums

„manchmal ist die kommunikation sehr schlecht“ - Welche Herausforderungen erleben Schüler:innen in der kooperativen Projektarbeit?

Marlis Pesch, Anke Langner

TU Dresden, Forschungsstelle Universitätsschule

Theoretischer Hintergrund:

Die Relevanz selbstregulierten Lernens in schulischen Lernsettings ist ebenso anerkannt wie kooperative Lernformate (Konrad & Traub, 2019; Perels et al., 2020) und die positiven Effekte auf Lernerfolge und Kompetenzentwicklung sind für beide empirisch nachgewiesen (Dignath & Büttner, 2008; Kyndt et al., 2013). Beide werden als zentrale Fähigkeiten verstanden, die Schüler:innen erwerben sollen (OECD, 2019) und in projektorientierten Lernformaten werden sie zusammengebracht (Traub, 2012).

Deshalb sind in den letzten Jahren Prozesse des Socially Shared Regulated Learning (SSRL) in den Fokus gerückt. SSRL befasst sich mit sozialen Regulationsprozessen in kooperativen Lernsettings (Panadero & Järvelä, 2015). Allerdings ist das Forschungsfeld vergleichsweise klein. Insbesondere für schulische Kontexte gibt es bisher kaum Forschung, möglicherweise auch deshalb, da das Lernen in kooperativen, projektorientierten Lernsettings bisher (noch) nicht kontinuierlich Anwendung in schulischen Kontexten findet.

Für eine gelingende Gestaltung derartiger Lernformate scheint es relevant, aus Perspektive der Akteur:innen zu fragen, welche Herausforderungen auftreten, um daran anknüpfend geeignete Formate und/oder Unterstützungsangebote zu konzipieren (Hadwin et al., 2018). Denn nur wenn Schüler:innen entsprechend angeleitet werden, stellen sie sich herausfordernden Aufgaben (Perry, 1998).

In bisherigen Studien zu Herausforderungen beim kooperativen Lernen wurden u.a. folgende Bereiche identifiziert: motivationale, sozial-emotionale, kognitive und metakognitive Herausforderungen sowie Herausforderungen, die sich auf Aspekte der Lernumgebung beziehen (Bakhtiar, 2019; Hadwin et al., 2018; Koivuniemi et al., 2017). Allerdings waren die Befragten Erwachsene und lernten lediglich zeitweise in kooperativen Lernsettings.

Fragestellung:

In der vorliegenden Studie soll daher untersucht werden, welche Herausforderungen Schüler:innen berichten, die täglich kollaborativ in Projektarbeit lernen (Langner et al., 2021).

Methode:

Die Datenerhebung erfolgte mittels Online-Fragebogen (N=404 Schüler:innen, Jahrgänge 4 bis 9). In Anlehnung an Koivuniemi et al. (2017) wurden ein offenes Antwortformat verwendet und die Schüler:innen gebeten, in einem Freitextfeld zu schildern, welche Herausforderungen sie erlebt haben.

Entsprechend der Literatur (Früh, 2015; Mayring, 2022) wurde das Datenmaterial strukturierend inhaltsanalytisch mit MAXQDA kategorisiert. Hierfür wurde auf Basis bisheriger Forschung (Hadwin et al., 2018; Järvenoja et al., 2013; Koivuniemi et al., 2017)

ein theoriegeleitetes Kategoriensystem entwickelt und nach einem ersten Materialdurchlauf induktiv erweitert und angepasst (Mayring, 2015). Nach einem finalen qualitativen Kodierzyklus wurden die Daten quantitativ analysiert.

Ergebnisse:

Insgesamt machten lediglich 13 % (N=51) der Schüler:innen keine Angabe, 6% (N=24) gaben an, keine Herausforderungen erlebt zu haben. 81% (N=329) der Schüler:innen beschrieben konkrete Herausforderungen; ein erheblicher Anteil (30%) benannte mehr als eine Herausforderung; insgesamt wurden N=514 Herausforderungen benannt.

Das Kategoriensystem umfasst sechs übergeordnete Herausforderungsbereiche: Ein erheblicher Teil der Schüler:innen (33%) berichtet Situationen, die der Sozialen Interaktion & Kooperation zuzuordnen sind (Auseinandersetzungen in der Projektgruppe, Planungen & Absprachen). 15% der genannten Herausforderungen bezogen sich auf Situationen, in denen es um den Wissenserwerb/Lernzuwachs ging. Das Umsetzen eines Projektprodukts (6%) war für viele Schüler:innen eine Herausforderung. Methoden/Tools/Materialien (10%) und das Zeitmanagement (10%) stellen u.a. Herausforderungen im übergeordneten Bereich der Ressourcen (24%) dar. Als emotionale & motivationale Herausforderung (13%) wurde z.B. das konzentrierte Arbeiten von 8% der Schüler:innen als besonders fordernd benannt. Auch mit kritischen Ereignissen (4%, z.B. Krankheit, Verlust von Material) sind die Schüler:innen konfrontiert; 5% waren Sonstige.

Fazit:

Die sehr hohe Antwortrate bei einem offenen Antwortformat unterstreicht die Relevanz der Herausforderungen für die Schüler:innen. Während Zeitmanagement, Kooperation, Planung sowie kognitive Anforderungen und Konzentration auch in anderen Studien häufig benannt werden (Hadwin et al., 2018; Koivuniemi et al., 2017), gibt es dort mitunter weniger Angaben zu emotionalen Herausforderungen, u.a. auch, da diese in geschlossenen Antwortformaten teils nicht aufgeführt werden.

Die Ergebnisse geben Hinweise auf Aspekte, die besonders fokussiert werden sollten, um die sozialen Regulationsprozesse in der kooperativen Projektarbeit zu unterstützen. Ergänzende Daten zum Umgang der Schüler:innen mit den Herausforderungen sowie von Lehrkräften wahrgenommenen Herausforderungen in der kollaborativen Projektarbeit können weitere Hinweise für eine gelinge Gestaltung kooperativer Projektarbeit geben.

Ist die Entwicklung von Lehr-Lernformaten eine Frage der Steuerung durch Schulleitung?

Matthias Ritter, Anna Weigel

TU Dresden, Forschungsstelle Universitätsschule

Theoretischer Hintergrund:

Nicht selten wird Unterrichtsentwicklung zwar in Kontext von Organisationsentwicklung (Rolff 2016) gesetzt, doch in der Regel wird in den Untersuchungen zur Unterrichtsentwicklung der Einfluss der Organisationsentwicklung und auch dem Leitungshandeln nicht thematisiert, sicher nicht zuletzt aus Mangel an Daten. Wenn Unterrichtsentwicklung jedoch nachhaltig in Schule vollzogen werden soll, ist die Verzahnung von Organisations- und Unterrichtsentwicklung unumgänglich (Göhlich et al. 2018). Eine enge Verzahnung von Wissenschaft und schulischer Praxis wird im Diskurs um evidenzbasierte Schulentwicklung immer wieder gewünscht (Schreiner et al. 2019), kann jedoch kaum als Praxis verstanden werden (van Ackeren et al. 2013). An der Universitätsschule Dresden ist eine enge Kooperation von Schulleitung und wissenschaftlicher Leitung seit dem Start der Schule 2019 etabliert. Ein Ausschnitt der Kommunikation beider Akteur:innen bilden Chat-Daten eines Messenger-Dienstes, welche einen besonderen Zugang darstellen, um die strategischen und konzeptionellen Überlegungen zwischen Schulleitung und wissenschaftlicher Leitung wie dessen abgeleiteten Entscheidungen für die Schulorganisation zu rekonstruieren.

Fragestellung:

Im Beitrag wird untersucht, welche Themenfelder der Unterrichtsentwicklung in der Chat-Kommunikation zwischen schulischer und wissenschaftlicher Leitung der Universitätsschule Dresden besprochen werden, so dass die konkreten Herausforderungen in der Zusammenarbeit zwischen Forschung und Praxis exemplarisch dargelegt werden.

Methode:

Für die Analyse der Chat-Daten wird die strukturierende Inhaltsanalyse nach Mayring (2022) angewendet. Das Untersuchungsmaterial umfasst Text- und vor allem Audionachrichten via Smartphone der beiden Akteur:innen im schulischen Kontext, die wöchentlich durchschnittlich rund 15 Transkriptseiten bzw. etwa 1,5 Stunden lang sind. Als Untersuchungszeitraum wurden jeweils vier Monate zu Beginn des Schuljahres im Jahr 2020 und 2023 festgelegt. Die Kategorien wurden deduktiv – basierend auf Rolff (2016) erstellt und zusätzlich um induktiv abgeleitete Codes aus dem Textkorpus erweitert.

Ergebnisse:

Unterrichtsentwicklung wird in den Chats von 2020 zwischen wissenschaftlicher und schulischer Leitung – im Vergleich zu Organisations- und Personalentwicklung – in einem geringeren Umfang thematisiert. Lediglich 11% der Codes sind der Unterrichtsentwicklung zuzuordnen, während die Maincodes Personalentwicklung (20%) und dominierend die Organisationsentwicklung (69%) der Schule besprochen werden. Im Jahr 2023 steigert sich der Umfang der Kommunikation zu Unterrichtsentwicklung auf 20%, während die Relevanz der Organisationsentwicklung um neun Prozentpunkte abnimmt und der Personalentwicklung gleichbleibt. Ein Grund für die höhere Relevanz der Unterrichtsentwicklung könnte sein, dass im Zeitraum zwischen 2020 und 2023 der organisatorische Rahmen für die Schule ausgehandelt und weiter definiert wurde, sodass sich nun dem Kerngeschäft, dem Unterricht gewidmet und intensiver auf inhaltlicher Ebene gearbeitet werden kann. Dies zeigt sich auch an den Subcodes der Kategorie Unterrichtsentwicklung: So ist im Jahr 2020 besonders die „Vision für Unterricht“ der Universitätsschule relevant und wird drei Mal häufiger als später in 2023 thematisiert. Hier findet eine Schärfung des Schulkonzepts sowie des pädagogischen Leitbilds statt, welche konkret über die Subcodes „Projektarbeit“, „Atelierarbeit“ und „Ganztag“ ausdifferenziert werden.

Es wird in den Daten deutlich, dass zwischen dem pädagogischen Anspruch, prozessorientiert und nah an der*dem Schüler:in arbeiten zu wollen und notwendiger Regularien von Schule vermittelt werden muss. So werden in 2023 Gespräche zum Subcode „Bewertungssystem“ relevant. Es werden unterschiedliche Interessensparteien wie Schüler:innen, Lernbegleitung, Eltern, Schulleitung und das Landesamt für Schule genannt, die jeweils andere Erfordernisse aufweisen und den Diskurs dieses Subcodes prägen.

Fazit:

Das Spannungsfeld zwischen Wissenschaft und Praxis in der Schulentwicklung stellt ein bedeutendes Forschungsdesiderat dar (Steffens et al. 2019). Eine vertiefte Untersuchung dieses Bereichs ist notwendig, um fundierte Orientierungen für den Schulaufbau und die organisatorische Transformation zu entwickeln und wie diese im Sinne einer research-practice partnership

(Zala-Mezö & Klein 2020) gestaltet werden kann. Durch weiterführende Forschung können Erkenntnisse gewonnen werden, die nicht nur zur theoretischen Wissensbildung beitragen, sondern auch praktische Handlungsempfehlungen für die schulische Praxis liefern.

Ergebnisse standardisierter Lernstandserhebungen in Mathematik und Rechtschreibung der Schüler:innen der Universitätsschule

Christoph Hassler, Anke Langner
Universitätsschule

Theoretischer Hintergrund:

Lehr-Lernprozesse die stärker von der Sinnhaftigkeit des Schülers/ der Schülerin ausgehen, stehen nicht selten im Verdacht, dass damit verbunden ist, dass Schüler:innen sich die Grundfertigkeiten der Kulturtechniken nicht ausreichend aneignen konnten. Um diese These zu prüfen, wurden inzwischen über 3 Jahre die Schüler:innen der USD über standardisierte Lernstandserhebungen, die idealerweise auch einen Lernprozessverlauf (Schneider et al. 2021) für die einzelnen Schüler:innen dokumentieren, getestet.

Fragestellung:

Nachgegangen wird der Frage, wie die mathematischen und orthographischen Kompetenzen der Schüler:innen der Universitätsschule Dresden im Vergleich zu Schülerinnen anderer Schulen ausgeprägt sind?

Methode:

Es wurde dafür seit dem Schuljahr 2021/22 die Hamburger Schreibprobe (HSP, May et al. 2019) und der Deutsche Mathematiktest (DEMAT, Roick et al. 2018) ab dem Jahrgang 2 bei allen Schüler:innen durchgeführt. Für die letzten zwei Jahre wurden zudem Daten von Vergleichsschulen (Grundschule, Oberschule, Gymnasium und Gemeinschaftsschule) aus dem Ballungsraum Dresden erhoben, wobei (aktuell) Auswertungsergebnisse für die Daten des Schuljahres 2022/2023 vorliegen. Die Auswertung umfasst dabei einen Querschnitt für die Jahrgänge 3 und 5 des Gymnasiums (HSP N=157, DEMAT N=85), der Oberschule (HSP N= 46, DEMAT N=47), der Grundschule (HSP N=74, DEMAT N=76) und der Gemeinschaftsschule (HSP N=85, DEMAT N=93), wobei die Daten in den folgenden Monaten mit den Ergebnissen des Schuljahres 2023/2024 und für die Jahrgänge 3 bis 6 angereichert werden. Von den aus dem Schuljahr 2022/2023 maximal 494 Schüler:innen der USD liegen für die HSP 482 (96%) und für den DEMAT 349 von 374 (93% der Anzahl der Schüler*innen der Jahrgänge 2 bis 6) Testergebnisse vor. An alle Testungen war ein standardisiertes Instrument zur Erfassung der sozialen Herkunft angebunden. Die Erhebung der Kompetenztests wurde entsprechend der Vorgaben an allen Schulen jeweils Ende des Schuljahres und während der regulären Unterrichtszeit durchgeführt. Die erhobenen Daten wurden anschließend deskriptiv ausgewertet, wobei die Vergleiche der Schulen und die Entwicklungen der Jahrgangsstufen der USD über multivariate Varianzanalysen sowie Post-Hoc T-Tests analysiert wurden.

Ergebnisse:

Die Untersuchung zeigt, dass die auf Jahrgangsebene aggregierten Ergebnisse der Schüler:innen der USD des Schuljahres 2022/2023 im Durchschnittsbereich der HSP und des DEMAT anzusiedeln sind. Einzige Ausnahme hiervon stellt der Subtest Arithmetik des 4. Jahrgangs dar, der knapp unterdurchschnittlich liegt. Im Entwicklungsvergleich der Jahrgänge an der USD zeigt sich in den meisten Subtests zwar eine deskriptiv positive Tendenz, jedoch erweisen sich die Unterschiede zu einem Großteil als statistisch nicht signifikant. Im Schulvergleich beider Kompetenztests zeigt sich, dass sich die USD nicht signifikant von der Gemeinschaftsschule unterscheidet. Das Gymnasium unterscheidet sich hingegen signifikant von allen anderen Schulen. Zwischen der USD und der Oberschule sowie der Grundschule zeigen sich Unterschiede vor allem hinsichtlich der Testbereiche der HSP, vereinzelte Unterschiede liegen jedoch auch für Subtests des DEMAT vor.

Fazit:

In der Universitätsschule lernen alle Schüler:innen binnendifferenziert miteinander ohne dass ein Ausleseprozess stattgefunden hat und sie lernen nach ihrem Tempo nicht im Gleichschritt miteinander und dennoch – so das allgemeine Resümee der bisherigen Daten – erreichen die Schüler:innen vergleichbare Kompetenzen. Die Schüler:innen können sich also mit Schüler:innen der Regelschule vergleichen, ob sich die Schüler:innen der Regelschule hinsichtlich der Kompetenzen des selbstregulierten Lernens und der Kooperativität mit Schüler:innen der Universitätsschule messen können, wäre ein nächster wichtiger zu vollziehender Vergleich.

Nachwuchstagung: Postersession "Auswertungsstrategien"

Zeit: Donnerstag, 30.01.2025: 9:15 - 10:45 · Ort: O 102

Bildungsaspirationen von Schüler:innen im Übergang zur gymnasialen Oberstufe und ihren Eltern im Schulformvergleich für Baden-Württemberg

Melina Klein, Evelyn Schnauffer, Wolfgang Wagner, Lisa-Marie König
Universität Tübingen, Deutschland

1.

Theoretischer

Hintergrund

Astleithner, Vogl & Parzer (2021) konnten zeigen, dass soziale Herkunft sich auf die Bildungsaspirationen auswirkt. Aspirationen von jugendlichen Schüler:innen scheinen maßgeblich von signifikanten Anderen (wie den Eltern) beeinflusst zu werden (Sewell, Haller & Portes 1969, zitiert nach Bittmann 2021). Im Vergleich von TOSCA-2002 und TOSCA-2006 hat sich zudem gezeigt, dass sich die gymnasialen Oberstufen von allgemeinbildenden Gymnasien und beruflichen Gymnasien mit verschiedenen Profilen mit Blick auf die soziokulturelle Zusammensetzung (bestehend u.a. aus sozioökonomischem Status, Bildungshintergrund der Eltern, Migrationshintergrund) zugunsten des allgemeinbildenden Gymnasiums unterscheiden (Becker, Maaz & Neumann 2010). Untersuchungswürdig wäre daher, ob sich neben der unterschiedlichen soziokulturellen Zusammensetzung an den Schulformen auch die Bildungsaspirationen der Schüler:innen unterscheiden. Diese Analyse fragt folglich insbesondere nach schulformbezogenen Unterschieden in dem Berufswunsch von Schüler:innen der Eingangsstufe der gymnasialen Oberstufe und ihrer Eltern.

2.

Methode

Daten der Schüler:innen- und Elternbefragung des ersten Erhebungszeitraums 2023 der Längsschnittstudie TOSCAneo des Hector-Instituts für Empirische Bildungsforschung an der Universität Tübingen bildeten die Grundlage der statistischen Analysen. Die Nettostichprobe in dieser Analyse umfasste abzüglich fehlender Angaben $N = 2733$ Schüler:innen und $N = 1657$ Eltern. Die Datenaufbereitung fand in Rstudio (Posit PBC, Version 2024.09.1+394) statt. Die Analysen erfolgten in Mplus (Muthén & Muthén, Version 8.6).

Zentrale Fragestellungen sind: (F1) Gibt es eine relative Übereinstimmung zwischen dem Berufswunsch von Schüler:innen der Eingangsstufe der gymnasialen Oberstufe und dem Berufswunsch der Eltern für ihre Kinder? (F2) Variiert der Berufswunsch von Schüler:innen der Eingangsstufe der gymnasialen Oberstufe beziehungsweise ihrer Eltern in Abhängigkeit von der Schulform?

Als Variablen dienten der Berufswunsch Schüler:innen, der Berufswunsch Eltern und die Schulform. Der Berufswunsch wurde in den Fragebögen durch eine offene Frage erhoben und anhand der Skala des International Socio-Economic Index of Occupational Status (ISEI, Ganzeboom et al. 1992) als Wert zwischen 11 und 90 codiert. Als Schulformen wurden allgemeinbildendes Gymnasium (AG, Referenzgruppe), Gemeinschaftsschule (GMS) und berufliches Gymnasium (BG) mit den Profilen BG-LHP (Lebens- und Humanwissenschaftliches Profil), BG-TEP (Technisches Profil) und BG-WWP (Wirtschaftswissenschaftliches Profil) unterschieden.

Zur Beantwortung der ersten Fragestellung wurde die bivariate Korrelation zwischen den Variablen Berufswunsch Schüler:innen und Berufswunsch Eltern ermittelt. Anschließend wurden für die zweite Fragestellung die Differenzen in dem Berufswunsch in Abhängigkeit von der Schulform durch eine multiple Regressionsanalyse untersucht, wobei die Schulformen als Dummies codiert waren. Es wurden clusterrobuste Standardfehler berücksichtigt sowie Survey-Gewichte einbezogen und für die Schulebene kontrolliert. Die Variablen wurden z-standardisiert, sodass die gruppenbezogenen Unterschiede in standardisierten Einheiten zu interpretieren sind.

3.

Ergebnisse

Die Variablen Berufswunsch Schüler:innen und Berufswunsch Eltern wiesen eine mittlere positive Korrelation von $r = .45$ auf. Es war eine statistisch signifikante relative Übereinstimmung ($p < .001$) der Bildungsaspirationen festzustellen.

Zur Beantwortung der zweiten Fragestellung wurde der Zusammenhang der Variablen Schulform, Berufswunsch Schüler:innen und Berufswunsch Eltern ermittelt. Statistisch signifikante Unterschiede im Berufswunsch der Schüler:innen bestanden zwischen den Schulformen BG-WWP und AG ($b = -2.44$, $SEb = 1.09$, $p = .025$), GMS und BG-TEP ($\Delta b = 2.67$, $SEb = 1.22$, $p = .029$), GMS und BG-WWP ($\Delta b = 3.47$, $SEb = 1.13$, $p = .002$), BG-LHP und BG-TEP ($\Delta b = 3.20$, $SEb = 1.46$, $p = .028$) sowie zwischen BG-LHP und BG-WWP ($\Delta b = 4.01$, $SEb = 1.38$, $p = .004$). Bezüglich des Berufswunsches der Eltern für ihre Kinder wurden Unterschiede zwischen den Schulformen nur zwischen BG-TEP und AG signifikant ($b = -3.04$, $SEb = 1.52$, $p = .046$).

4.

Diskussion

Festzuhalten ist, dass die Erhebung in der Eingangsstufe der gymnasialen Oberstufe stattfand. Ob der Berufswunsch der Schüler:innen sich im Laufe der gymnasialen Oberstufe verändert oder erst konkreter ausbildet – es sei auf die relativ hohe Anzahl an fehlenden Angaben mit $N = 1330$ verwiesen – und somit ggf. größere bzw. geringere Differenzen zwischen den Schulformen entstehen, müsste in nachfolgenden Untersuchungen erforscht werden.

Co-Orchestrierung von Lehrkraftaufgaben im Unterricht: Szenario-basierte Präferenzanalyse von Automatisierungsgraden bezüglich Funktionen Tutorieller Lernsoftware.

Ann-Christin Falhs

Ruhr-Universität Bochum/ Center for Advanced Internet Studies, Deutschland

Ann-Christin Falhs^{1,2} ; Prof. Dr. Nikol Rummel^{1,2}; Prof. Dr. Frank Reinhold³
¹Ruhr-Universität Bochum, Deutschland; ²Center for Advanced Internet Studies (CAIS), Bochum, Deutschland; ³PH Freiburg, Deutschland

Die zunehmende Heterogenität von Schulklassen sowie die damit einhergehenden Herausforderungen im Monitoring und Controlling von Lernprozessen stellen Lehrkräfte vor komplexe Anforderungen (Hoth et al., 2016). Adaptive Lernumgebungen sind in der Lage, Parameter wie die Lernzeit, den Umfang der Instruktionen sowie die Reihenfolge oder den Schwierigkeitsgrad der Aufgaben an die individuellen Lernvoraussetzungen und -fortschritte der Schüler:innen anzupassen (Leutner, 2002). Intelligente Tutoring-Systeme (ITS) haben sich als wirksame Instrumente zur Förderung des individuellen Lernens erwiesen, indem sie die Lernprozesse der Schüler:innen verfolgen und basierend auf deren Lernpfaden adaptive Aufgabenauswahl sowie personalisierte Rückmeldungen bereitstellen (Kulik & Fletcher, 2016; Ma et al., 2014; Steenbergen-Hu & Cooper, 2013; Xu et al., 2019). Durch das Sammeln, Analysieren und Visualisieren von Lernprozessdaten haben Intelligente Tutoring-Systeme und Teacher Dashboards somit das Potenzial, Lehrkräfte beim Monitoring und Controlling von Lernprozessen zu entlasten (Van

Leeuwen & Rummel, 2019; Xhakaj et al., 2017). Dabei können Teacher Dashboards Lehrkräfte bei der Entscheidungsfindung unterstützen, etwa durch die Bereitstellung relevanter Informationen oder Handlungsvorschläge (Van Leeuwen & Rummel, 2019).

Obwohl viele Lehrkräfte den Einsatz adaptiver Lern-Apps befürworten (Robert Bosch Stiftung, 2023), ist bislang wenig erforscht, welche Unterstützung Lehrkräfte durch ITS bevorzugen und welche Automatisierungsgrade sie für spezifische Tätigkeiten als hilfreich erachten. Diese Studie untersucht die Präferenzen von Lehrkräften in Bezug auf die Automatisierung zentraler Tätigkeitsfelder wie die Auswahl von Lernmaterialien, die Verfolgung des Lernfortschritts sowie die Bereitstellung von Feedback und Hinweisen sowie inwiefern diese durch individuelle Faktoren beeinflusst werden. Zur Konzeptualisierung der Automatisierungsgrade werden das 6-Layer-Modell (Molenaar, 2022) sowie die Dashboard-Funktionen Mirroring, Advising, Alerting und Intervening berücksichtigt (Van Leeuwen & Rummel, 2019) herangezogen.

Die Datenerhebung erfolgt durch ein querschnittliches Studiendesign mit zwei Kohorten in Nordrhein-Westfalen: Mathematiklehrkräften im Dienst sowie Lehramtsstudierenden mit Lehrerfahrung. Zunächst werden die bestehenden Unterrichtspraktiken im Hinblick auf die Häufigkeit (Single-Choice Items) sowie die Art der Tätigkeiten (Multiple-Choice Items zzgl. Ranking) erfragt. Anschließend werden mittels sechs storybasierter Szenarien Präferenzen für verschiedene Automatisierungsgrade im Hinblick auf die zentralen Tätigkeitsbereiche erfasst. Durch darauf folgende Single-Choice Items wird erhoben, inwiefern Lehrkräfte über potenziell automatisiert ausgeführte Aufgaben informiert werden möchten, unter welchen Bedingungen und aus welchen Gründen. Das technologisch-pädagogische Wissen wird mittels eines Performanztests (Franke et al., in prep.) erfasst, während die kontextbezogene Einstellung zu Künstlicher Intelligenz (KI) auf Basis der wahrgenommenen Nützlichkeit, der wahrgenommenen Benutzerfreundlichkeit sowie der Verhaltensintention (Mesenhöller & Böhme, 2024) ermittelt wird.

Die zu diskutierenden Auswertungsstrategien schlagen eine Kombination aus multivariaten und inferenzstatistischen Verfahren vor. Um die Präferenzen für unterschiedliche Automatisierungsgrade zu modellieren, werden multiple lineare Regressionen verwendet. Bivariate Korrelationsanalysen erfassen Zusammenhänge zwischen den Präferenzen und den Informationsbedürfnissen der Lehrkräfte. Darüber hinaus werden mithilfe von Moderations- und Mediationsanalysen geprüft, inwiefern das technologisch-pädagogische Wissen (TPK) sowie die Einstellung zu Künstlicher Intelligenz (KI) die Präferenzen beeinflussen. Clusteranalysen dienen schließlich der Identifikation von Gruppenunterschieden, insbesondere im Hinblick auf die Lehrerfahrung.

Die Ergebnisse dieser Studie tragen dazu bei, ein fundiertes Verständnis der Anforderungen und Präferenzen von Lehrkräften im Umgang mit intelligenten Tutoriellen-Systemen zu entwickeln. Im Fokus stehen insbesondere die Einsatz- und Unterstützungsmöglichkeiten, aber auch mögliche Implikationen für die Gestaltung adaptiver Technologien.

Datenchaos oder Erfolgsmuster? Wie Lehrkräfte formative Assessments für datenbasierte Entscheidungen nutzen

Thorsten Fahrbach^{1,2}, Maren Specker², Stephan Blank², Benjamin Fauth²

¹Universität Tübingen, Deutschland; ²Institut für Bildungsanalysen Baden-Württemberg

Theorie

Lernbegleitende Evaluationen, sogenannte Formative Assessments, haben durch die einfache Handhabung mittels digitaler Möglichkeiten zunehmend Einzug in die Klassenzimmer gehalten (Blumenthal et al., 2022). Die Ergebnisse der formativen Assessments ermöglichen es den Lehrkräften, ihren Unterricht an die individuellen Kompetenzen der Schülerinnen und Schüler anzupassen und somit das Lernen zu verbessern (Schütze et al., 2018). Eine solche datengestützte Entscheidungsfindung (eng. Databased Decision-Making, DBDM) umfasst nach Mandinach (2012) vier Schritte: Daten sammeln, Daten analysieren und interpretieren, Entscheidungen auf Basis der Daten treffen, die sich dann im vierten Schritt positiv auf die Lernergebnisse der Schülerinnen und Schüler auswirken.

Der erste Schritt, das Sammeln von Daten, wird durch das jeweilige Tool umgesetzt. Ob sich Lehrkräfte jedoch für den Einsatz eines formativen Assessments entscheiden, hängt von verschiedenen Faktoren ab (Yan et al., 2021). In welchen Settings und mit welcher Intensität computerbasierte Tools außerhalb wissenschaftlicher Studien eingesetzt werden, ist bisher kaum bekannt.

Im Schritt der Datenanalyse und -interpretation zeigt sich für Lehrkräfte die Herausforderung, über das reine Lesen der Daten hinauszugehen (Van Den Bosch et al., 2017; Zeuch et al., 2017). Zudem bleibt die Frage offen, wie viel Zeit Lehrkräfte in die Auswertung der Daten investieren und wie intensiv sie sich damit auseinandersetzen. Bisherige Studien zeigen, dass sich eine intensivere Auseinandersetzung der Lehrkraft mit den Daten positiv auf didaktische Entscheidungen und den Lernverlauf der Schülerinnen und Schüler auswirkt. (Hebbecker et al., 2022).

Damit sich formative Assessments positiv auf die Lernleistungen aller Schülerinnen und Schüler auswirken, gilt die individuelle Förderung als entscheidend. Bisherige Studien deuten jedoch darauf hin, dass diese didaktische Entscheidung von Lehrkräften nur in geringem Maße als dritter Schritt umgesetzt wird (Gelderblom et al., 2016; Van Geel et al., 2016). Insgesamt ist in der Forschung bisher wenig darüber bekannt, wie Lehrkräfte die Bereitstellung von Daten aus formativen Assessments für die Gestaltung ihres Unterrichts nutzen (Farley-Ripple et al., 2018).

Forschungsfragen

Der Zusammenhang zwischen allen vier Schritten wurde bisher kaum untersucht (Hebbecker et al., 2022). Zudem basieren viele Befunde zu datenbasierter Entscheidungsfindung und formativen Assessments auf Daten aus wissenschaftlichen Studien. Untersuchungen, die auf Daten aus der Praxis basieren, sind uns bisher nicht bekannt. Daher möchten wir anhand von realen Nutzungsdaten eines computergestützten formativen Assessments folgende zwei Forschungsfragen untersuchen:

1. Welche Muster lassen sich jeweils in den vier Schritten der datengestützten Entscheidungsfindung identifizieren?
2. In welchem Zusammenhang stehen die identifizierten Muster im jeweiligen Schritt mit Mustern aus den anderen Schritten?

Methodik

Seit dem Schuljahr 2022/23 steht allen Grundschulen in Baden-Württemberg das digitale formative Evaluationstool *quop* (Förster & Souvignier, 2015) zur Verfügung. Mit *quop* können Lehrkräfte die Lese- und Mathematikkompetenzen ihrer Schülerinnen und Schüler in acht Kurztests über das Schuljahr verteilt erfassen. Die Daten aller *quop* nutzenden Lehrkräfte und ihrer Klassen sollen genutzt werden, um Muster in den Schritten der Datenerhebung, der Datenanalyse und des Lernzuwachses zu identifizieren. Zu diesem Zweck werden die Häufigkeit, Regelmäßigkeit und der Zeitpunkt der Tests (Datenerhebung), die Daten aus den Logfiles

über die Betrachtung der Ergebnisse durch die Lehrkraft (Analyse und Interpretation) und die Testergebnisse (Lernzuwachs) verwendet.

Um den dritten Schritt, die datenbasierte Entscheidung, zu untersuchen, wurden die Lehrkräfte in den Schuljahren 2022/23 und 2023/24 gebeten, einen Fragebogen zu ihrer Nutzung von quop auszufüllen. Insgesamt nahmen $n = 1537$ Lehrkräfte teil.

Zur Auswertung der Daten soll zunächst eine Clusteranalyse durchgeführt werden, um mögliche unterschiedliche Muster zu identifizieren (Forschungsfrage 1). Im nächsten Schritt sollen korrelative Zusammenhänge zwischen den Mustern in den jeweiligen Schritten der datengestützten Entscheidungsfindung mit Mustern in anderen Schritten untersucht werden (Forschungsfrage 2).

Digitale Leseförderung: Eine Studie zur Effektivität bei leseschwachen Schüler:innen mit multiplen Risikofaktoren

Christina Ingelfinger, Tobias Dörfler, Cornelia Glaser

Pädagogische Hochschule Heidelberg, Deutschland

Schulleistungsstudien belegen seit mehreren Jahren, dass zahlreiche Schüler:innen gravierende Defizite im Lesen aufweisen (Heine et al., 2023; Hußmann et al., 2017; McElvany et al., 2023; Sälzer, 2021). Besonders betroffen sind hierbei Schüler:innen mit Migrationshintergrund oder niedrigem sozioökonomischem Status (McElvany et al., 2023, S. 18–19; Sachse et al., 2022, S. 156; Stubbe et al., 2023, S. 171). Internationale Studien zeigen, dass digitale Förderprogramme die Leseleistung deutlich verbessern können (z.B. Alqahtani, 2020; Jamshidifarsani et al., 2019). Allerdings liegen im deutschsprachigen Raum, trotz der Vielzahl an digitalen Leseförderprogrammen, bislang nur wenige Studien vor, die deren Wirksamkeit sowie die zugrunde liegenden Bedingungen systematisch untersuchen.

Diese Studie geht der Frage nach, inwieweit eine digitale Leseförderung die Lesekompetenz von Schüler:innen mit Leseschwierigkeiten und multiplen Risikofaktoren effektiv fördern kann. Hierfür wurde im Rahmen eines Pre-Posttest-Designs die App *Meister Cody – Namagi* (Meister Cody GmbH, 2022) eingesetzt. Die Stichprobe umfasste 207 Schüler:innen der vierten bis sechsten Klasse eines Sonderpädagogischen Bildungs- und Beratungszentrums mit dem Förderschwerpunkt Lernen.

Neben der Wirksamkeit hinsichtlich der Leseflüssigkeit (SLRT-II; Moll & Landerl, 2014) wurden mögliche Transfereffekte auf das Leseverständnis (ELFE-II; Lenhard et al., 2020), die Lesemotivation, das Leseselbstkonzept und die Leseaktivität (Fragebogen; vgl. McElvany et al., 2023) sowie die Langfristigkeit der Intervention analysiert. Zusätzlich wurden der rezeptive Wortschatz (GraWo; Seifert et al., 2017) sowie emotionale und soziale Kompetenzen (SDQ; Goodman, 1997) als potenzielle Einflussfaktoren erhoben. Als Vergleichsgruppen kamen eine Alternativtrainingsgruppe, die den Reading-While-Listening-Ansatz nutzte, und eine Unterrichtskontrollgruppe zum Einsatz. Während die Alternativtrainingsgruppe eine etablierte Methode des Lautleseverfahrens einsetzte, diente die Unterrichtskontrollgruppe als Vergleichsbedingung der regulären Unterrichtspraxis. Ergänzend wurden Prozessmaße nach jeder Interventionssitzung erfasst, um die Motivation, den Kompetenzerwerb und die Passung der Aufgabe einzuschätzen.

Obwohl die Analyse der Ergebnisse noch aussteht, deuten erste Befunden auf eine große Heterogenität der Ausgangsleistungen sowohl in der Leseflüssigkeit als auch im Leseverständnis hin. Alle Schüler:innen nahmen fokussiert an den 30-minütigen Interventionssitzungen teil, wobei die Mehrheit ein hohes Maß an Motivation aufwies. Gleichzeitig wurden Einschränkungen der digitalen Förderung sichtbar, insbesondere durch begrenzte Möglichkeiten zur manuellen Anpassung, wie beispielsweise die Erhöhung der Darbietungszeit oder die Variation des Schwierigkeitsniveaus, was die Effektivität der Intervention bei heterogenen Lerngruppen beeinträchtigen könnte.

Ziel der Studie ist es, den Forschungsstand zur Effektivität digitaler Leseinterventionen im deutschsprachigen Raum zu erweitern und praxisorientierte Empfehlungen für den schulischen Einsatz abzuleiten.

Effects of Computer-Based Feedback on Motivational Outcomes

Arya-Ronahi Dogan^{1,2}, Ute Mertens¹, Livia Kuklick¹, Marlit Annalena Lindner^{1,2}

¹Europa-Universität Flensburg; ²IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften

Abstract folgt bis zum 31.12.24

Einflussfaktoren der Karriereentscheidung am Übergang Schule-Beruf bzw. Schule-Studium - Eine systematische Literaturübersicht

Julia Haas

Universität Konstanz, Deutschland

Wird zusammen mit dem Poster bis 31. Dezember 2024 eingereicht.

Erfolgsfaktoren von digitalen educational Escape Games im MINT-Studium - eine kausalanalytische Studie

Melissa Gruber¹, Stefanie Faßbender²

¹Pädagogische Hochschule Karlsruhe, Deutschland; ²Hochschule Karlsruhe, Deutschland

Relevanz: Bedingt durch die digitale Transformation stehen Hochschullehrende der Herausforderung, die Lern-/Lehrangebote an die Erwartungen und Interessen der Studierenden anzupassen (vgl. Harris-Huermann/Pohlenz/Mitterauer, 2018, S. 41). Digitale Lernformate und -methoden, die sowohl ergänzend zu als auch in synchronen Lehrveranstaltungen eingesetzt werden, weisen positive Effekte auf die Kompetenzen der Lernenden auf. Neben der Digitalkompetenz werden beispielsweise Selbstorganisation und Medienkompetenz gefordert und gefördert (vgl. Fthenakis/Walbiner, 2018, S. 99). Ein Beispiel für innovative digitale Lehrformate ist das digitale educational Escape Game, das als Serious Game im Rahmen von Game-Based Learning eingesetzt wird (vgl. Cheng u. a., 2017, S. 261). Dabei befinden sich Studierende in einem virtuellen Raum, der durch eine Fragestellung oder ein Rätsel verschlossen ist. Durch die Interaktion mit im Raum platzierten virtuellen Objekten sind Hinweise zu finden, die zur Beantwortung der Fragestellung oder Rätsellösung dienen und somit den Schlüssel aus dem verschlossenen Raum liefern (vgl. Bilbao-Quintana u. a., 2021, S. 4).

Theoretischer Rahmen: Für den Bildungskontext sind digitale educational Escape Games besonders interessant, da sie basierend auf der konstruktivistischen Lerntheorie sowie Selbstbestimmungstheorie die Aktivierung der Lernenden im Spiel fördern. Durch die intensive die Auseinandersetzung mit dem Lernmaterial kann in digitalen educational Escape Games eine aktive Konstruktion von Wissen stattfinden, während der individuelle Lernprozess von Lernenden unterstützt wird (vgl. Buchner, 2022, S. 59). So gibt es in der relevanten Forschungsliteratur Hinweise, dass sich digitale educational Escape Games positiv auf die Motivation von Studierenden und die Entwicklung ihrer Team- und Kommunikationsfähigkeit auswirken (vgl. Rosillo/Montes, 2021, S. 14; Vergne/Smith/Bowen, 2020, S. 2846). In der Literatur wird kontrovers diskutiert, ob diese Lernspiele einen signifikanten positiven Effekt auf den Lernerfolg von Studierenden haben (vgl. Kavanaugh u. a., 2020, S. 709; Manzano-León u. a., 2021, S. 1; Rosillo/Montes, 2021, S. 14). Zum aktuellen Zeitpunkt fehlen Untersuchungen, wie digitale educational Escape Games konzipiert sein müssen, um dem Bildungskontext am besten gerecht zu werden (vgl. Bilbao-Quintana u. a., 2021, S. 15; Wei Jie Ang/Ni Annie Ng/Shen Liew, 2020, S. 2849; Yow, 2022, S. 1351).

Fragestellung: Aus diesem Grund werden in diesem Promotionsprojekt Determinanten des Lernerfolgs der Studierenden unter dem Einsatz moderner Lehr-/Lernmethoden in der Hochschullehre untersucht, welche Aspekte den Lernerfolg von Studierenden beim Spielen von digitalen educational Escape Games beeinflussen.

Methode: Im Rahmen einer leitfadengestützten Befragung von Fachkundigen wurde zunächst qualitativ exploriert, welche möglichen Determinanten den Lernerfolg der Studierenden beim Spielen von digital educational Escape Games beeinflussen könnten. Gemäß den Empfehlungen der befragten Fachkundigen wurde ein digital educational Escape Game Computerspiel zum Thema Kryptografie für den MINT-Bereich entwickelt. Das Spiel wird im Wintersemester 2024/2025 experimentell untersucht, wobei ein Strukturgleichungsmodell die Grundlage der Untersuchung bildet. Das Modell mit PLS Verfahren wurde aufgrund seiner Strukturen-prüfenden Eigenschaft ausgewählt, um kausalanalytische Zusammenhänge zwischen einzelnen Elementen von digitalen educational Escape Games und dem Lernerfolg von Studierenden zu untersuchen (vgl. Backhaus/Erichson/Weiber, 2013, S. 11). Diese Posterpräsentation stellt die Auswertung der mittels Pretest/Posttest gesammelten Daten sowie die aus dem Computerspiel gewonnenen Daten zum Spielverhalten der Studierenden vor. Erste Ergebnisse zu wichtigen didaktischen sowie spielerischen Faktoren sollen auf mögliche Determinanten des Lernerfolgs schließen.

Fashion DeSIrE "How to Develop a Sustainable Design by Implementing Education for Sustainable Development as a Guiding Principle in Fashion Design Education"

Hanieh Choopani^{1,2}, Martina Glomb², Anne-Marie Grundmeier¹

¹Pädagogische Hochschule Freiburg; ²Hochschule Hannover

This doctoral research aims to explore the implementation of Education for Sustainable Development (ESD) in fashion design higher education, considering the importance highlighted by UNESCO's 2030 Agenda (UNESCO, 2020) and the EU Strategy for Sustainable and Circular Textiles, including the key initiative of the Ecodesign for Sustainable Products Regulation (European Commission, 2023). It emphasizes on how education in fashion design can address environmental, economic, and social challenges by equipping students with the knowledge and skills needed to create sustainable practices within the industry. The study employs a Design-Based Research (DBR) methodology (Oppl et al., 2022), encompassing six phases and multiple iterative processes and addresses this question: how to implement ESD as a guiding principle in fashion design higher education?

The systematic literature review (first phase) follows Paul and Criado's (2020) guidelines to assess ESD implementation in higher education fashion and apparel design B.A. and M.A. programs in Germany and Iran. Publications in German, Persian, and English from January to September 2023 were screened using specific inclusion and exclusion criteria. The three-phase screening identified a significant research gap in Germany and a lack of analysis in Iran (Choopani et al., 2023). For the document analysis (second phase), curricula from 14 German universities and the Iranian Ministry of Science, Research, and Technology were collected online. German universities have 23 distinct curricula, while Iranian universities follow a standardized curriculum. Documents were translated into English and analyzed using MAXQDA software. Applying Mayring's (2022) content analysis method, categories were selected based on theoretical frameworks. Through deductive-inductive analysis, two categories were eliminated, and three new categories emerged. This phase reveals that nine German universities have integrated sustainability aspects into their curricula, whereas only three modules in Iranian curricula address sustainability. Additionally, five German universities have yet to incorporate any sustainability aspects. Overall, sustainability integration varies widely across both countries (Choopani et al., 2023).

Problem-centered expert interviews (Witzel & Reiter, 2022) with fashion design lecturers and professors in Germany and Iran, and fashion design students and graduates at Hochschule Hannover University of Applied Sciences and Arts (HSH), and participant observations for fashion design courses at HSH in Hannover, Germany, were also conducted. To understand the adoption of ESD in fashion and apparel design education, twenty professors and lecturers in Germany and eleven in Iran were contacted, resulting in interviews with twelve in Germany and six in Iran. Participant observations (Kawulich, 2005) were conducted during two winter semesters in 2022 and 2023, recording classroom interactions and course content. These activities identify and evaluate ESD implementation practices in fashion design courses and the data will be transcribed and analyzed using Mayring's (2022) content analysis method on MAXQDA.

The fourth phase focuses on developing, piloting, and evaluating a sustainable fashion design course using the DBR approach. The course incorporates ESD competencies and the "USE-LESS Product Lifecycle" (USE-LESS Centre for Sustainable Design Strategies, n.d.) to enhance students' skills and understanding of sustainable fashion design. From December 2023 to May 2024, it was developed through monthly meetings. The course integrates sustainable design strategies, aiming to impart knowledge on sustainability's environmental, social, and economic aspects (Claxton & Kent, 2020; European Commission, 2023; Gimenez et al., 2012; Amed et al., 2023). Piloted from March to July 2024 for 28 B.A. Fashion Design students at HSH, the assessment included e-portfolios, session-specific questions, and a final group interview, analyzed using deductive-inductive qualitative content analysis (Mayring, 2022) with MAXQDA software. The course will be further implemented at a vocational school in Hannover and a university in Tehran, Iran, in the fifth phase. Recommendations will be compiled from all phases on ESD implementation in fashion design higher education.

Feedback im Online-Lerntool für den Unterricht Deutsch als Zweitsprache: Eine Untersuchung der Perspektiven von Lehrenden und Lernenden

Betül Mungan

Universität Tübingen, Deutschland

In diesem Poster werden die Ergebnisse von Befragungen zur Nutzung eines digitalen Lerntools im Bereich Deutsch als Zweitsprache ausgewertet und dargestellt. Der Schwerpunkt liegt hierbei bei der Bewertung des Feedbacks in dem digitalen Lerntool.

Förderung der Intention und des Wertes von Making-Aktivitäten durch Autonomie und Nutzen: Eine experimentelle Video-Vignetten-Studie

Sophie Uhing

Friedrich-Alexander-Universität Erlangen-Nürnberg, Deutschland

Ab dem Alter von 12 Jahren nimmt das Interesse an MINT (Mathe, Informatik, Naturwissenschaften, Technik) stetig ab (Sadler et al., 2012). Sogenannte Making-Aktivitäten (z.B. 3D-Druck, Lasercuttern), werden oft als besonders geeignet angesehen, um Jugendlichen MINT-Bildungserfahrungen zu vermitteln (Smolarczyk & Kröner, 2023). Insbesondere bei Making-Aktivitäten sind Mädchen und benachteiligte Jugendliche mit niedrigem SES bisher unterrepräsentiert. Die Theorie des geplanten Verhaltens (TPB; Ajzen, 1991) bietet einen Ansatzpunkt, um die Teilnahmeintention von Jugendlichen an Making-Aktivitäten zu erklären: Verhaltensbezogene Überzeugungen (z.B. Intrinsischer Wert und persönliche Wichtigkeit), normative Überzeugungen und Kontrollüberzeugungen (personenbezogen und umweltbezogen) erklären demnach die Intention. Diesbezüglich bieten verhaltensbezogene Überzeugungen einen guten Ansatzpunkt für Interventionen. Intrinsischer Wert lässt sich insbesondere über einen hohen Grad an Autonomie fördern (Rosenzweig et al., 2022). Die persönliche Wichtigkeit lässt sich erhöhen, indem Aufgaben z.B. mit der eigenen Identität verbunden werden und der Nutzen dargestellt wird (Rosenzweig et al., 2022). Dies wurde auch in Studien zu Making exploriert, indem die Bedeutung von Making für einen selbst oder andere aufgezeigt wurde (Peppler et al., 2021). Sowohl Autonomie als auch Nutzen sind wichtige Anknüpfungspunkte für die Gestaltung von Making-Aktivitäten. Es fehlen bisher vor allem Hinweise, welche Überzeugungen für die Gewinnung von (unterrepräsentierten) Jugendlichen geeignet sind. Deshalb widmete sich diese Studie der Fragestellung, wie sich verhaltensbezogene Überzeugungen sowie die Teilnahmeintention durch eine Videovignetten-Intervention mit den Faktoren Autonomie und Nutzen beeinflussen lassen. Die Studie wurde als 2x2 faktorielles Zwischensubjekt-Design mit den Faktoren Autonomie (hoch-niedrig) und Nutzen (hoch-niedrig) konzipiert. Wir erwarteten, dass Teilnehmer, die ein Video mit einem hohen Nutzen-Faktor ansehen, Zuwächse in den abhängigen Variablen intrinsischer Wert, persönliche Wichtigkeit und Teilnahmeabsicht vom Pretest zum Posttest zeigen.

An der Studie nahmen $N = 223$ Schüler*innen aus drei bayrischen Mittelschulen teil. Das genutzte Messinstrument erfasste alle Überzeugungen der TPB und wurde in einer Vorstudie (Uhing et al., in Arbeit) pilotiert. In dieser Vorstudie zeigte sich, dass persönliche Wichtigkeit und intrinsischer Wert einen Faktor höherer Ordnung bilden, was auch für die vorliegenden Analysen berücksichtigt wurde. Der Manipulationscheck wurde zuvor mittels Regressionen durchgeführt und zeigte, dass dieser nur für verschiedene Facetten des Faktors Nutzen (Nutzen f. Zukunft: $t(216) = 2.98, p = .003$; Nutzen d. Gegenstände: $t(216) = 1.98, p = .047$), jedoch nicht den Faktor Autonomie ($t(216) = 1.42, p = .158$) erfolgreich war.

Für die Hauptanalysen wurde auf ein Latent Change Score Model (Steyer et al., 2000) zurückgegriffen, um differenzierte Aussagen über die Veränderungen von t_1 zu t_2 zu erhalten. Hier wurde für die Intention sowie den Faktor höherer Ordnung für verhaltensbezogene Überzeugungen ein Change Score und ein Grundwert gebildet: Während alle Items des Prä- und Posttests auf den Grundwert t_1 laden, laden auf den Change Score nur die Items von t_2 . Der Change Score stellt somit den latenten Veränderungswert dar. Die Intercepts und Faktorladungen wurden für t_1 und t_2 gleichgesetzt. Die Regressionspfade der Faktoren wurden sowohl auf den Change Score verhaltensbezogener Überzeugungen als auch den Change Score der Intention modelliert. In Anlehnung an die TPB wurden die Regressionspfade des Change Scores der verhaltensbezogenen Überzeugungen, die verhaltensbezogenen Überzeugungen, die subjektive Norm, die Selbstwirksamkeit und umweltbezogene Kontrolle modelliert.

Die Ergebnisse zeigen, dass der Faktor Nutzen den Change Score verhaltensbezogener Überzeugungen beeinflusst, während Autonomie keinen Effekt hatte. Die Intention wurde hauptsächlich durch den Change Score verhaltensbezogener Überzeugungen sowie den Grundwert erklärt. Die Hypothese, mit der ein Einfluss der Faktoren Autonomie und Nutzen auf die Veränderung der verhaltensbezogenen Überzeugungen und der Intention postuliert wurde, konnte somit nur teilweise bestätigt werden. In weiteren Untersuchungen sollen noch Geschlecht und Bildungshintergrund der Eltern berücksichtigt werden.

Interdisziplinäre Interventionsstudie zu formativem Peer- und Selfassessment im Mathematikunterricht der Primarschule

Karolin Maskos¹, Andreas Schulz¹, Ladina Feucht¹, Selin S. Oeksuez², Sandra Zulliger², Alois Buholzer²

¹Pädagogische Hochschule Zürich, Schweiz; ²Pädagogische Hochschule Luzern, Schweiz

Das prozessbegleitende, formative Assessment (FA) trägt eine entscheidende Rolle zum Lernen in der Schule bei (Black & Wiliam, 2009; Black et al., 2011; Hattie, 2009). Es dient im deutschschweizerischen Lehrplan als eine von drei Beurteilungssäulen kompetenzorientierten Unterrichts (Bildungsdirektion des Kantons Zürich, 2017). Von transparenten Kriterien mit engem fachlichen Bezug ausgehend, werden im FA Lernaufgaben eingesetzt, die dem Kind und der Lehrperson den aktuellen Lernstand aufzeigen. Dieser wird zur Erreichung neuer Lernziele in lernförderlichem Feedback rückgemeldet und der Unterricht entsprechend adaptiert. Sowohl im Peerassessment (PA), dem kooperativen Beurteilen, als auch im Selfassessment (SA), der Selbstreflexion und -beurteilung, werden die Schüler:innen als Ressource füreinander und als Agenten ihres eigenen Lernens aktiviert (Black & Wiliam, 2009; Heritage, 2007). Eine wichtige Rolle bei der Umsetzung von PA und SA kommt dem fachdidaktischen und pädagogischen Professionswissen der Lehrpersonen zu (Ball et al., 2008; Harris & Brown, 2013; Heritage, 2007; Schmidt, 2020).

Der Unterrichtseinsatz von PA und SA im Mathematikunterricht wird als deutlich entwicklungsfähig beschrieben (Buholzer et al., 2020). Als erfolgsversprechend gelten zielgerichtete Weiterbildungen, die das Professionswissen von Lehrpersonen zu PA und SA kombiniert mit einem engen fachlichen Fokus vermitteln und konkrete unterrichtspraktische Anregungen geben (Oeksuez et al., 2024). Es fehlen jedoch entsprechend evaluierte Weiterbildungskonzepte insbesondere im deutschsprachigen Raum (Hebbecker & Souvignier, 2018). Hier setzt die folgend vorgestellte, vom Schweizerischen Nationalfond finanzierte, interdisziplinäre Interventionsstudie an.

Im Rahmen der Studie wurde eine dreiteilige, hybride Weiterbildung (Intervention) von insgesamt 8,5 Stunden (inklusive Vor-/Nachbereitung) für Primarlehrpersonen der Klassenstufen 4 und 5 durchgeführt. Die Intervention wurde auf Grundlage von pädagogisch-psychologischen und mathematikdidaktischen Theorien konzipiert (Oeksuez et al., 2024). Sie befasst sich mit der lernförderlichen Umsetzung von PA bzw. SA am Beispiel des Rechnens auf eigenen Wegen in der Multiplikation und Division (Greiler-Zauchner, 2022; Schulz, 2018). Das variable Anwenden von Rechenstrategien und die Freiräume in deren Notation (Schütte, 2004) eignen sich für die Beurteilung im Rahmen von PA oder SA in Verbindung mit dem Vergleichen von Rechenwegen

(Rittle-Johnson & Star, 2011; Smith & Stein, 2018). Vorangegangene Studien (Schulz & Leuders, 2018) zeigten grosse Effekte des Vergleichens und Beurteilens von Rechenwegen, das eine vertiefte, lernförderliche Durchdringung von Rechenstrategien ermöglicht.

Das Poster fokussiert die mathematikdidaktischen Forschungsziele der Studie: (1) Hat eine inhaltspezifisch ausgerichtete Weiterbildung zu PA/SA eine Wirkung auf das fachliche Lernen der Schüler:innen?, (2) Hat diese Weiterbildung auch eine Wirkung auf das fachdidaktische Professionswissen der Lehrperson?, (3) Beeinflusst das Professionswissen der Lehrperson das Lernen der Schüler:innen? und (4) Vermittelt das Professionswissen der Lehrperson die Wirksamkeit der Weiterbildung auf das Lernen der Schüler:innen? (Mediatoranalyse).

Die Stichprobe ($N_{\text{Lehrpersonen}} = 96$, $N_{\text{Schüler:innen}} = 1731$, Stand 27.06.24) unterteilt sich in 50 vierte, 43 fünfte und drei jahrgangsgemischte Klassen der deutschsprachigen Schweiz. Die randomisierte Zuteilung erfolgte in Interventionsgruppen PA ($n=28$ Klassen) bzw. SA ($n=28$) und Wartekontrollgruppe ($n=39$).

Die Daten wurden in einem randomisierten Prä-Post-Follow-up-Wartekontrollgruppendesign im Schuljahr 2023/24 erhoben. Die Lehrpersonen beantworteten vor der Weiterbildung und nach der Unterrichtsdurchführung Fragebögen. Diese umfassten im mathematikdidaktischen Teil das Professionswissen zu Multiplikations- und Divisionsstrategien sowie möglichen Fehlern. Ausserdem gaben die Lehrpersonen Prognosen an, wie oft sie die Strategien bzw. Fehler in der eigenen Klasse erwarten. Die Schüler:innen beantworteten zu drei Messzeitpunkten Textaufgaben (nur t1) sowie Multiplikations- und Divisionsaufgaben. Die Daten sollen mittels mehrbenenanalytischen Regressionsmodellen ausgewertet werden.

Zum Zeitpunkt der 12. Tagung der GEBF ist die Datenerhebung und Kodierung der mathematikdidaktischen Lehrpersonen und Schüler:innen-Fragebögen abgeschlossen. Das Poster berichtet erste Befunde zur Wirkung der Weiterbildung auf das mathematikdidaktische Professionswissen der Lehrpersonen sowie auf das mathematische Lernen der Schüler:innen. Angedacht sind hierfür mehrbenenanalytische Auswertungen. Zudem wird auch der Abgleich der Prognosen der Lehrpersonen zusammen mit den tatsächlichen Leistungsdaten ihrer Klassen thematisiert.

Konstitutionsprozesse immersiver virtueller Lernräume

Fiona Schmidbauer^{1,2}

¹Pädagogische Hochschule Karlsruhe, Deutschland; ²Hochschule Karlsruhe, Deutschland

Hintergrund

und

Forschungsziel

Mit der zunehmenden Digitalisierung und Verlagerung von Bildung in virtuelle Lernräume, beschleunigt durch die Corona-Pandemie, gewinnt immersive virtuelle Realität (VR) im Bildungssektor stetig an Bedeutung. Es fehlt jedoch an theoretischen und empirischen Ansätzen, die Lernen in VR über quantifizierbare Aspekte hinaus betrachten. Das Forschungsprojekt adressiert diese Forschungslücke, indem die Konstitutionsprozesse immersiver virtueller Lernräume untersucht werden. Theoretisch stützt sich die Forschung auf die Raumsoziologie von Martina Löw (2017) und erweitert diesen Ansatz durch die Akteur-Netzwerk-Theorie (ANT) (Latour 2007). Die Erweiterung von Löws Raumsoziologie durch die ANT ist relevant für immersive virtuelle Räume, da die ANT die Rolle nicht-menschlicher Akteure in der Raumkonstitution berücksichtigt und eine ganzheitliche Betrachtung der Raumkonstitutionsprozesse in immersiver VR ermöglicht.

Forschungsfragen

Die Forschungsfragen lauten: (1) Durch welche Praktiken der Raumkonstitution werden immersive virtuelle Lernräume hergestellt? (2) Welche menschlichen und nicht-menschlichen Akteure sind an der Konstitution immersiver virtueller Lernräume beteiligt? (3) Inwiefern beeinflussen Praktiken der Raumkonstitution die Erkenntnisprozesse in immersiver VR?

Methode

Die Erhebung und Analyse der Daten erfolgt mittels Videointeraktionsanalyse (Tuma et al. 2013). Die Videointeraktionsanalyse ist eine iterative Methode, in der auf das ethnographische Sampling ein internes Sampling und gegebenenfalls die Erhebung weiterer Daten erfolgt, bis eine theoretische Sättigung zur Beantwortung der Forschungsfragen erreicht ist (Tuma et al. 2013). Das ethnographische Sampling erfolgte durch Videoaufzeichnungen von Vorlesungen im „VR-Campus“, einer immersiven VR-Anwendung der Hochschule Karlsruhe. Die Daten wurden videographisch erhoben, indem Forschende an VR-Vorlesungssitzungen teilnahmen und durch das Screenrecording ihrer VR-Brille audiovisuelle Videodaten erzeugten. Bisher wurde rund 298 Minuten Videomaterial aus Vorlesungssitzungen von 2022 bis 2024 aufgezeichnet, zusätzlich rund 62 Minuten Videomaterial des fokussierten Interviews in VR und Zoom mit dem VR-Campus-Entwickler Jonas Deuchler. Das erhobene Material wurde bereits grob nach für die Forschungsfragen relevanten Sequenzen durchsucht, die in Datensitzungen mit anderen Forschenden analysiert wurden. Der Fokus der Analysen lag auf der Identifikation von Handlungssequenzen, in denen die Interaktion, Kommunikation und performativen Handlungen der Akteure den immersiven virtuellen Lernraum konstituieren. Die Ergebnisse aus den Datensitzungen werden für das interne Sampling verwendet, um in den vorhandenen Daten Sequenzen für die Feinanalyse zu identifizieren. Gegebenenfalls müssen weitere Daten im Feld erhoben werden, um eine theoretische Sättigung zu erreichen.

Erste

Ergebnisse

In einer Datensitzung im September 2024 wurden Materialsequenzen analysiert, die als relevant für die Forschungsfragen identifiziert wurden. Dabei wurden drei zentrale Phänomene identifiziert, die für die Raumkonstitution immersiver virtueller Lernräume relevant sind: (1) die Dozent:innen-Studierenden-Interaktion, (2) Störungen durch Studierende und (3) die Positionierung der Akteure im virtuellen Raum. Die Positionierung wurde als zentral für die Forschungsfragen eingeschätzt und daher in einer Datensitzung im Oktober 2024 Suchfragen zur räumlichen Positionierung entwickelt, darunter: Wie organisieren die Nutzer:innen ihre räumliche Verortung und Abstände zueinander? Existiert ein individuelles „Territorium“ im virtuellen Raum, ähnlich eines eigenen Stuhls im physischen Seminarraum, und inwiefern kehren Nutzer:innen nach einer Unterbrechung an denselben virtuellen Ort zurück? Die Daten werfen zudem die Frage auf, ob es spezifische Konventionen oder Handlungsoptionen wie virtuelle Sitzplätze gibt, die raumkonstituierende Handlungen im immersiven virtuellen Raum prägen, oder ob die Nutzer:innen Erfahrungen aus dem physischen Raum übernehmen, um Prozesse der Orientierung und „Plazierung“ (Löw 2017, 158f.) zu gestalten.

Ausblick

Auf Grundlage der Suchfragen wird ein internes Sampling des restlichen Materials aus dem VR-Campus durchgeführt. Wenn durch das vorhandene Material keine theoretische Sättigung erreicht wird, müssen erneut Daten im Feld erhoben werden, die voraussichtlich aus weiteren immersiven multi-user VR-Lernanwendungen stammen. Ziel ist es, zu verstehen, inwiefern Nutzer:innen nach Raumgewohnheiten physischer Lernräume agieren und durch welche Regeln sich immersive virtuelle Lernräume von diesen unterscheiden. Ein besseres Verständnis darüber, wie immersive virtuelle Lernräume konstituiert werden

und welche menschlichen wie nicht-menschlichen Akteure daran beteiligt sind, kann wertvolle Erkenntnisse für Lernprozesse in immersiven virtuellen Umgebungen liefern.

Lehrkraft-handlungen zur Unterstützung selbstregulierten Lernens

Kirsten Roth-Kuppler

RPTU Kaiserslautern Landau, Deutschland

Selbstreguliertes Lernen ist eine Grundvoraussetzung für lebenslanges Lernen. Ein wesentlicher Bestandteil dessen ist die Fähigkeit, Lernstrategien zielführend auszuwählen und korrekt anzuwenden. Diese lernstrategische Kompetenz ist bereichsspezifisch, weshalb es sinnvoll ist, Strategieranwendung in einem inhaltlichen Kontext zu trainieren (Schrader et al., 2008). Die hier vorgestellte Studie untersucht Unterstützungshandlungen von Sekundarstufenlehrkräften im Englischunterricht.

Um Verstehensprobleme beim Lesen fremdsprachlicher Texte selbstständig zu lösen, müssen SchülerInnen in der Lage sein, geeignete (adäquate) Strategien anzuwenden. Viele Strategien werden jedoch häufig, aber erfolglos eingesetzt (Finkbeiner & Ludwig, 2008). Daher müssen SchülerInnen darin unterstützt werden, adäquate Strategien zu wählen und sie korrekt anzuwenden.

In der hier vorgestellten Studie werden Videografien von Unterrichtsstunden analysiert, die im Rahmen der ADEQUA-Studie (ebd.) erstellt wurden und Interaktionen zwischen SchülerInnen und Lehrpersonen handlungsnah erfassten. Es zeigte sich höherer Lernerfolg bei adäquater Strategieranwendung, jedoch bleibt die Frage offen, wie genau Lehrpersonen ihre SchülerInnen darin unterstützen können, Strategien adäquat anzuwenden. Es ist zu vermuten, dass Lehrpersonen, die sich bei Unterstützungshandlungen an Elementen selbstregulierten Lernens orientieren, ihre SchülerInnen besser unterstützen und somit selbstreguliertes Lernen gezielter fördern.

Methode:

Ein Zusammenhang zwischen Strategieranwendung und Lernerfolg zeigte sich bislang in Analysen auf mikroanalytischer Ebene, wohingegen Fragebogenstudien diesen Zusammenhang nicht immer bestätigten (z. B. Artelt, 1999).

Die Datengrundlage der hier vorgestellten Studie bilden 24 Videografien von Unterrichtsstunden, die im Rahmen von ADEQUA (Finkbeiner & Ludwig, 2008) in zwölf neunten Klassen acht verschiedener Schulen aller Bildungsgänge entstanden. Elf Lehrpersonen und 267 SchülerInnen bilden die Analysegrundlage. Während der Unterrichtsstunden arbeiteten die SchülerInnen in Partnerarbeit an englischsprachigen Texten.

Vor dem Hintergrund, dass Lernprozesse interne Prozesse sind, das Unterrichtsgeschehen jedoch ebenso Sequenzen externer Steuerung beinhaltet, wird das Verhalten der Lehrpersonen in drei Unterrichtsphasen mit Hilfe eines Kategoriensystems bewertet. Dieses beinhaltet 21 hypothetische Handlungen, die aus einem erweiterten Prozessmodell selbstregulierten Lernens abgeleitet wurden. Das Modell wurde für diese Studie entwickelt und orientiert sich am Prozessmodell von Schmitz und Schmidt (2007) und einem erweiterten Gesamtmodell motivierten selbstregulierten Lernens (Otto, 2007). Die Phase des selbstregulierten Lernens der SchülerInnen ist durch die Einstiegsphase zu Beginn und die Ergebnissicherungsphase am Ende des Unterrichts umrahmt. Die Erarbeitungsphase besteht in der eigentlichen Textarbeit und beinhaltet wiederum drei Phasen des selbstregulierten Lernprozesses der SchülerInnen. Somit werden sowohl interne als auch externe Faktoren berücksichtigt.

Ergebnisse:

Die Videografien wurden transkribiert und kodiert. Im nächsten Schritt werden die Lehrpersonen nach Art ihrer vorwiegenden Unterstützungshandlungen gruppiert. Ebenso werden die SchülerInnen nach Ergebnissen im Textverstehenstest gruppiert. Dieser dient als Indikator für den Lernerfolg. Überdies werden SchülerInnen gruppiert nach ihren Ergebnissen im Selbstregulierungstest. Geplant ist die Überprüfung von Zusammenhängen zwischen der Art der Unterstützungshandlung der Lehrkräfte und dem Lernerfolg sowie der Selbstregulierungskompetenz der SchülerInnen auf Gruppenebene. Zudem soll auf einer individuellen Ebene geprüft werden, ob SchülerInnen mit bestimmten Lernvoraussetzungen besser oder weniger gut unterstützt werden.

Relevanz:

Mit Hilfe der zu gewinnenden Erkenntnisse wird ein Moderationsmodell entwickelt, auf dessen Basis Handlungsempfehlungen für Lehrkräfte formuliert werden. Diese sollen Lehrkräfte einerseits in der Diagnostik selbstregulatorischer Kompetenzen der SchülerInnen unterstützen, andererseits Hinweise geben, wie SchülerInnen gezielter in adäquater Strategieranwendung unterstützt und so in selbstreguliertem Lernen gefördert werden können. Zudem kann das Modell eine Grundlage für die Entwicklung weiterer Unterrichtsmaterialien darstellen. Das Moderationsmodell konkretisiert Prozessmodelle selbstregulierten Lernens für die Unterrichtspraxis.

Aufgabenformate, die auf der Grundlage des neu zu entwickelnden Modells erstellt werden, spielen eine wichtige Rolle für SchülerInnen, deren Erstsprache nicht die Unterrichtssprache ist. Über den Fremdsprachenunterricht hinaus können diese Formate auf andere Unterrichtsfächer übertragen werden. Arbeitsmaterialien, die selbstreguliertes Lernen fördern und fordern, gehören vor dem Hintergrund des Lehrkräftemangels und Homeschooling-Szenarien zum notwendigen Portfolio praktizierender und angehender Lehrkräfte.

Mentale Modelle von Programmierkonzepten in semi-strukturierten Interviews

Katrin Kunz, Katerina Tsarava, Ulrich Trautwein

Hector-Institut für Empirische Bildungsforschung, Universität Tübingen, Deutschland

Relevanz:

Programmieren ist eine Fähigkeit, die nicht nur für das Verständnis und die Interaktion mit Computern und anderen Geräten erforderlich ist. Das Erlernen von Programmierfähigkeiten ist auch mit kognitiven Fähigkeiten wie Kreativität, logischem Denken, Metakognition, allgemeiner Problemlösefähigkeit und Mathematischen Kompetenzen verbunden und für diese von Vorteil (Román-González et al., 2017; Scherer et al., 2019). Code literacy, die Fähigkeit die Sprache von Computern zu lesen und zu schreiben, kann den Unterschied zwischen „Programmieren oder programmiert werden“ (Rushkoff, 2010), „App-fähig oder App-abhängig“ (Gardner & Davis, 2013) und „Technologie schaffen oder konsumieren“ (Resnick et al., 2009) ausmachen. Die Entwicklung valider mentaler Modelle für Programmierkonzepte gilt als wichtiger Schritt beim Erlernen von Programmieren (Bower & Falkner, 2015; Du Boulay, 1986; Sorva, 2013). Die Forschung zu mentalen Modellen von Programmierkonzepten hat sich dabei bisher auf Studierende (Götschi et al., 2003; Ma et al., 2011) konzentriert. Mentale Modelle bauen jedoch auf

Erfahrungen auf und die meisten Programmieranfänger*innen beginnen mit dem Programmieren mittlerweile vor dem Eintritt in die Universität. Das bedeutet, dass die mentalen Modelle jüngerer Schüler*innen erforscht werden sollten, um fehlerhaften Mentalen Modellen im Unterricht frühzeitig zum Beispiel mithilfe von conceptual-change-Ansätzen begegnen zu können (Chi, 2009; DiSessa et al., 2004).

Theoretischer

Rahmen:

Gentner und Stevens (1983) definieren mentale Modelle als Vorstellung von mechanischen Prozessen und Zuständen. Mentale Modelle hängen von der Interaktion einer Person mit ihrer Umwelt ab und liefern Vorhersagen und Erklärungen für die beobachteten Phänomene. Norman (1983) betonte, dass die folgenden Dinge zu berücksichtigen sind, wenn es um mentale Modelle von Programmierkonzepten geht: Das Zielsystem, also die Computerhardware und -software; das konzeptionelle Modell, das von den Lehrkräften als Repräsentation der Programmierkonzepte bereitgestellt wird; das Systembild, also die Aspekte des Zielsystems, mit denen die Schüler*innen auf der Grundlage des implementierten konzeptionellen Modells interagieren; das interne mentale Modell der Schüler*innen und die Konzeptualisierung dieses mentalen Modells durch die Wissenschaftler*innen. Sorva (2013) erläuterte, dass mentale Modelle beobachtet werden können, wenn man die Schüler*innen die Ergebnisse von Computerprogrammen vorhersagen lässt, da bei einem solchen Tracing das mentale Modell zur mentalen Simulation der Programmausführung verwendet wird. Erkannte Fehlvorstellungen lassen Rückschlüsse zu, ob sie inkohärent, teilweise kohärent oder kohärent mit einem fehlerhaften oder validen mentalen Modell verbunden sind (Chi, 2009).

Fragestellung:

1. Welche Mentale Modelle von Programmierkonzepten haben Grundschulkinder?
2. Wie hängen Mentale Modelle von Programmierkonzepten bei Grundschulkindern mit der Programmiersprache und Programmierungsumgebung zusammen?
3. Wie entwickeln sich Mentale Modelle von Programmierkonzepten bei Grundschulkindern im Verlauf von Informatik-Kursen?

Methode:

Im Rahmen der vorliegenden Arbeit wurden 46 Schüler*innen der zweiten bis vierten Klasse in Baden-Württemberg, die einen von drei außerschulischen Informatikkursen der Hector Kinderakademien (z.B. Trautwein et al., 2023) besuchten, befragt. Der erste Kurs des Hector Informatikcurriculums zeigt Zweitklässler*innen Programmierkonzepte „unplugged“, also ohne Computer, im zweiten Kurs wechseln Drittklässler*innen zum Programmieren in Scratch, während Viertklässler*innen im dritten Kurs von block- zu textbasiertem Programmieren in Python übergehen (Tsarava et al., 2024). In vier semi-strukturierten Interviews vor ihrem Kursbesuch, zu zwei Zeitpunkten während des Kurses und nach dem Kursbesuch wurden die Kinder zu ihren Vorstellungen von den Programmierkonzepten „Schleife“ und „bedingte Verzweigung“ befragt. Zusätzlich sollten sie in Think-Aloud Protokollen ihr Tracing von Programmieraufgaben verbalisieren. Die Interviews wurden transkribiert und zum Teil bereits in einer qualitativen Inhaltsanalyse nach Mayring und Fenzl (2019) bezüglich vorkommender Fehlvorstellungen ausgewertet.

Ergebnisse:

Erste Ergebnisse zeigen Fehlvorstellungen in Scratch, die bereits von Sorva (2017), Swidan et al. (2018) und Hartmann et al. (2024) kategorisiert werden konnten. Die Konzeptualisierung von Mentalen Modellen wird jedoch im Allgemeinen als schwer angesehen (Nitz & Fechner, 2018). Weitere Auswertungsstrategien wie etwa die SOLO-Taxonomie (Jimoyiannis, 2011), Grounded Theory (Charmaz, 2015) oder Phänomenologie (Mühling & Große-Bölting, 2023) könnten helfen tiefere Strukturen in den Aussagen der Kinder zu erkennen.

Schulische Einflüsse auf populistische Einstellungen im Jugendalter

Elise Grunwald, Katharina Eckstein, Astrid Körner, Peter Noack

Friedrich-Schiller-Universität Jena, Deutschland

Angeichts der zunehmenden globalen Krisen und der wachsenden Unterstützung rechtspopulistischer Parteien wird die Bildung demokratiebewusster Bürgerinnen und Bürger immer dringlicher. Selbst in etablierten Demokratien verbreiten sich populistische und teils antidemokratische Narrative verstärkt. Die Analyse der Entwicklung populistischer Einstellungen aus jugendpsychologischer Perspektive ist entscheidend, um junge Menschen trotz dieser Herausforderungen als kritische Unterstützer der Demokratie zu stärken. Dies ist besonders relevant, da das Jugendalter eine Schlüsselphase in der politischen Einstellungsbildung darstellt (*Impressionable Years Hypothesis*, Sears & Levy, 2003). Schul- und Bildungskontexte sind nicht nur rechtlich zur Erziehung zur Demokratie verpflichtet, sie bieten auch wichtige Sozialisationskontexte, die durch formale (z. B. curriculare Inhalte) und informelle Lernprozesse (z. B. Klassenklima) junge Menschen beeinflussen können (Noack & Eckstein, 2023). Während Wahlergebnisse einen populistischen Trend unter Jugendlichen vermuten lassen, gibt es bislang nur wenige Befunde zu den schulischen Einflüssen auf populistische Einstellungen im Jugendalter (Noack & Eckstein, 2023; Jungkunz & Weiss, 2024).

Unsere Studie erfasst mithilfe eines für Jugendliche und junge Erwachsene adaptierten Messinstruments (basierend auf Castanho Silva et al., 2018) populistische Einstellungen auf drei Dimensionen: *Volkszentrismus*, *Anti-Elitismus* und *Manichäische Weltanschauung*. Die Datengrundlage bilden zwei Erhebungswellen (Anfang 2024 und Anfang 2025) von t_1 $N = 478$ Schüler*innen im Alter von 14 bis 21 Jahren aus Niedersachsen und Thüringen. Mithilfe von Multilevel Structural Equation Modelling (MSEM) werden die Stabilität populistischer Einstellungen sowie die Wechselwirkungen zwischen diesen Einstellungen und zwei schulbezogenen Variablen – *demokratisches Klassenklima* und *curriculare Demokratiebildung* – untersucht. Zentrale Forschungsfrage ist, ob schulische Erfahrungen die Entwicklung populistischer Einstellungen kompensieren können (Noack & Eckstein, 2023). Dabei kontrollieren wir für die Kovariaten *Migrationshintergrund*, *Geschlecht* und *Schultyp*. Das MSEM ermöglicht es, Effekte sowohl auf Individual- als auch auf Klassenebene zu modellieren.

Wir erwarten erstens, dass sich populistische Einstellungen über die Zeit stabilisieren. Zweitens gehen wir davon aus, dass das Erleben eines demokratischen Klassenklimas sowie verstärkte Demokratiebildung zu Zeitpunkt t_1 mit niedrigeren populistischen Einstellungen zu Zeitpunkt t_2 assoziiert sind. Drittens vermuten wir einen moderierenden Effekt des Alters, in dem Sinne, dass demokratische Erfahrungen, sowohl informell als auch curricular, mit zunehmendem Alter stärkere Effekte zeigen. Der Posterbeitrag diskutiert das Auswertungsverfahren im Hinblick auf die anstehende Datenerhebung der zweiten Welle Anfang 2025. Ziel ist es, mit den Studienergebnissen einen empirischen Ausblick auf schulische Interventionsmöglichkeiten zur Prävention populistischer Einstellungsmuster im Jugendalter zu bieten.

Soziodemografische Unterschiede in der Leseleistung von Schüler:innen - Eine intersektionale Perspektive

Mia Herkommer

Eberhard Karls Universität Tübingen, Deutschland

Bildungsungleichheit und Lesekompetenz

Lesekompetenz ist entscheidend für schulischen Erfolg und gesellschaftliche Teilhabe (Heine et al., 2023; Paetsch et al., 2016). Seit PISA 2000 ist bekannt, dass die Lesekompetenzen in Deutschland stark mit Geschlecht, sozioökonomischem Status (SES) und Migrationshintergrund korrelieren. Während die Ungleichheiten bis 2015 abnahmen, verschärfte die Corona-Pandemie bestehende Disparitäten (Anger & Plünnecke, 2021; Bethhäuser et al., 2023; Ludewig et al., 2022). Frühere Studien betrachteten diese Merkmale oft isoliert. Dadurch wurde vernachlässigt, dass Überschneidungen dieser Kategorien spezifische Benachteiligungen erzeugen können, die über additive Effekte hinausgehen. Um diese komplexen Ungleichheiten besser zu verstehen, ist ein intersektionaler Ansatz unerlässlich (Else-Quest & Hyde, 2016). Dabei hat sich die *Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy* (MAIHDA) als besonders geeignet für die quantitative Forschung erwiesen (Van Dusen et al., 2024).

Zielsetzung und methodischer Ansatz

In dieser Untersuchung wird MAIHDA erstmals für einen Vergleich der Lesekompetenz von 2018 und 2022 für Deutschland herangezogen, um die Komplexität der Ungleichheiten besser zu erfassen. Additive und intersektionale Interaktionseffekte können mithilfe folgender Kennwerte differenziert quantifiziert werden (Evans et al., 2024; Keller et al., 2023):

- Der VPC (Variation Partition Coefficient) beschreibt die Gesamtvarianz, die durch die Gruppenzugehörigkeit erklärt wird.
- Der VPCadj. (adjustierter VPC) gibt Auskunft über die verbleibende Varianz unter Kontrolle der additiven Haupteffekte.
- Der PCV (Proportional Change in Variance) misst zusätzlich den Anteil der Varianz auf Gruppenebene, welcher durch additive Effekte erklärt werden kann, bzw. welcher auf intersektionale Interaktionseffekte hindeutet (1-PCV).

Für die Analyse wurden zunächst die relevanten Hintergrundmerkmale der Lesekompetenz identifiziert: Geschlecht, Migrationshintergrund, Bildungsabschluss (HISCED) und Berufsstatus der Eltern (HISEI). Nach der Imputation fehlender Werte umfasst die Stichprobe schließlich 5.451 Schüler:innen im Jahr 2018 und 6.116 Schüler:innen im Jahr 2022.

In den MAIHDA-Modellen wird die Lesekompetenz auf Ebene 1 modelliert, während die Hintergrundvariablen auf Ebene 2 zu 40 verschiedenen Gruppen kombiniert werden. Ziel der Analyse ist die Beantwortung der folgenden Forschungsfragen:

1. Wie unterscheiden sich die Lesekompetenzen zwischen den intersektionalen Gruppen?
2. Inwieweit tragen komplexe Interaktionen der Hintergrundmerkmale zusätzlich zur Erklärung von Ungleichheiten in der Lesekompetenz bei?
3. Wie unterscheiden sich die zentralen Kennwerte zwischen den Jahren 2018 und 2022?

Ergebnisse

Die Analyse zeigt deutliche Unterschiede zwischen den intersektionalen Gruppen: Die leistungsstärkste Gruppe („weiblich, ohne Migrationshintergrund, mind. ein Elternteil mit Abitur und höchstem Berufsstatus“) übertrifft die leistungsschwächste Gruppe („männlich, mit Migrationshintergrund, kein Elternteil mit Abitur, und niedrigstem Berufsstatus“) um 169 Punkte (2018) bzw. 159 Punkte (2022). Dies entspricht einem erheblichen Kompetenzunterschied und einem Lernzuwachs von etwa 4 Jahren (OECD, 2010).

16.6% (2018) bzw. 13.9% (2022) der Gesamtvarianz in der Lesekompetenz können allein durch die Gruppenzugehörigkeit erklärt werden (VPC). Komplexe Interaktionen haben nur geringen Einfluss auf die Rangfolge der Gruppen in beiden Jahren. Einzig in der leistungsstärksten Gruppe ist die Lesekompetenz sowohl 2018 als auch 2022 signifikant höher, als allein durch additive Haupteffekte erklärbar wäre. Hier deutet dies auf Interaktionseffekte hin, die positiv auf die Lesekompetenz wirken.

Der niedrige VPCadj von 1,5 % (2018) bzw. 0,9 % (2022) zeigt, dass unter Kontrolle der additiven Effekte auf Gruppenebene nur eine geringe Residualvarianz verbleibt. Dies legt nahe, dass die additiven Haupteffekte den Großteil der Varianz erklären. Dennoch bleiben 7.8% (2018) bzw. 5.4% (2022) der Varianz ungeklärt, was auf das Vorhandensein von Interaktionseffekten hinweist (1-PCV). Die Unterschiede der zentralen Kennwerte zwischen 2018 und 2022 sind jedoch nicht signifikant, was darauf hindeutet, dass sich die beobachteten Muster der Ungleichheit zwischen den beiden Zeitpunkten kaum verändert haben.

Insgesamt liefern die Ergebnisse 2018 wie auch 2022 Hinweise auf relevante intersektionale Ungleichheiten im Bildungssystem, auch wenn diese der Analyse folgend vorrangig auf die additiven Haupteffekte der Hintergrundmerkmale zurückzuführen sind.

Sprachformale Textrevisionen mit digitalen Medien in der Primarstufe fördern Erste Ergebnisse einer Interventionsstudie in der Primarstufe

Tina Neff

Pädagogische Hochschule Karlsruhe, Deutschland

In dem Promotionsprojekt „Sprachformale Textrevisionen mit digitalen Medien in der Primarstufe fördern“ geht es um die Nutzung digitaler Rechtschreibhilfen während der Rechtschreibkorrektur von eigenen Texten.

Angeichts der Tatsache, dass Medienbildung kein eigenständiges Schulfach darstellt, ist es zu einer zentralen Bildungsaufgabe geworden, Medien- und Fachdidaktik sinnvoll miteinander zu verbinden (vgl. Anskeit, 2020) und neue Erkenntnisse für die Forschung und Schulpraxis zu generieren. Digitale Rechtschreibhilfen können die eigene Textproduktion entlasten, indem sie im Schreibprozess - eingesetzt während der sprachformalen Textrevision - Hinweise auf konkrete Rechtschreibfehler geben. Zentral dabei ist, dass Lernende mit digitalen Rechtschreibhilfen interagieren müssen – sie also nicht nur mechanisch anwenden, sondern lernen, orthographische Strukturen sowie Korrekturvorschläge zu hinterfragen und diese auf die Sinnhaftigkeit in Bezug zu ihrem eigenen Text zu überprüfen (vgl. Berndt, 2001). In den Bildungsstandards des Faches Deutsch in der Primar- und Sekundarstufe 1 wird im Bereich Schreiben bereits explizit auf die Arbeit mit digitalen Rechtschreibhilfen wie Korrekturprogrammen und Wörterbüchern hingewiesen (vgl. KMK, 2022a, 14; KMK, 2022b, 23). Wie eine didaktische Implementierung digitaler Rechtschreibhilfen in den Unterricht aussehen kann, zeigen jedoch lediglich erste Best-Practice-Beispiele (vgl. Neff, 2021). Voraussetzung für den nachhaltigen Einsatz digitaler Rechtschreibhilfen in den mediengestützten Deutschunterricht ist eine didaktische Implementierung, die sich auf evidenzbasierte Erkenntnisse stützt.

Diesen beschriebenen Desideraten wird im Rahmen der empirischen Interventionsstudie begegnet. Ziel des Projekts ist es, geeignete Fördermaßnahmen zur Interaktion mit digitalen Rechtschreibhilfen zu entwickeln und in Hinblick auf ihre Wirksamkeit zu evaluieren. Diese Fördermaßnahmen werden für die Primarstufe konzipiert und in Form von drei unterschiedlichen digitalen, interaktiven Lernpfaden realisiert. Die drei Lernpfade unterscheiden sich in ihrer didaktischen Konzeption hinsichtlich der Art des Rechtschreibtrainings: Der erste Lernpfad umfasst ein Training im Umgang mit den digitalen Rechtschreibhilfen (Tool-Training). Der zweite Lernpfad beinhaltet neben dem Tool-Training ein konkretes Training zu ausgewählten orthographischen Teilbereichen (Tool-Training + Orthographie-Training). Im dritten Lernpfad erhalten die Kinder lediglich ein Orthographie-Training ohne konkreten Bezug zu den digitalen Rechtschreibhilfen. Zur Evaluation der Wirksamkeit der drei Lernpfade verfassen und revidieren die Schülerinnen und Schüler eigene Texte zu einer profilierten Schreibaufgabe (vgl. Bachmann & Becker-Mrotzek, 2010) in einem Pre-Posttest-Design – unter Verwendung der digitalen Rechtschreibhilfen. Die Daten für die Studie werden im Feld - also im Deutschunterricht - erhoben, wobei drei Klassen pro Interventionsgruppe jeweils in Doppelstunden an der Erhebung teilnehmen.

Um die Wirksamkeit der Fördermaßnahmen zu evaluieren, werden methodisch vorwiegend quantitative Erhebungs- und Auswertungsverfahren eingesetzt, vereinzelte qualitative Elemente sollen diese Erkenntnisse ergänzen (Mixed-Method-Design). In einem Prä- und Posttest schreiben die Lernenden jeweils vor und nach dem Einsatz der digitalen Lernpfade einen eigenen, digitalen Text auf dem Tablet und revidieren ihre individuellen Rechtschreibfehler mithilfe der digitalen Rechtschreibhilfen. Die eigene, digitale Textproduktion im Prä- und Posttest samt sprachformaler Textrevision liefert alle relevanten Informationen für die quantitative Auswertung. Für diese Auswertung der Erst- und Zweitfassung der eigenen Texte wird jeweils die Rechtschreibleistung und die Textqualität analysiert. Zusätzlich wird qualitativ bei einer ausgewählten Stichprobe die digital gestützte sprachformale Textrevision analysiert. Im Rahmen dessen wird geschaut, inwieweit die Revisionen in Bezug zu einer speziellen Reaktion der digitalen Rechtschreibhilfe glücken bzw. missglücken. Für diese Beurteilung wird ein Kriterienraster entwickelt, das die Revisionshandlung in Bezug zur Reaktion der digitalen Rechtschreibhilfe setzt. Zusätzlich ergänzen Screencasts aus Prä- und Posttest von digitaler sprachformaler Textrevision sowie Fragebögen von einzelnen Schüler:innen, in Bezug zu Rechtschreibregeln, -strategien und der Bewusstheit über individuelle Stärken und Schwächen, die quantitativen Erkenntnisse.

Um die externe Validität der Studie zu erhöhen, werden mögliche (personenbezogene) Störvariablen erhoben und kontrolliert. Als Kovariate werden standardisierte Rechtschreibtests, analoge Textproduktionsprozesse, Lesetests, Aufgaben zur Computernutzung und Informationen zur Familiensprache in die Datenauswertung einbezogen.

Im Fokus der Posterpräsentation stehen neben dem Forschungsdesign und ersten empirischen Erkenntnissen vor allem die Auswertungsstrategien der quantitativen und qualitativen Ergebnisse.

Storytelling in multimodalen Schreibarrangements: Förderung der narrativen Textproduktionskompetenz durch multimodales Erzählen

Maximilian Stoller

Pädagogische Hochschule Karlsruhe, Deutschland

Digital Storytelling in multimodalen Schreibarrangements: Förderung der narrativen Textproduktionskompetenz durch multimodales Erzählen

Der vermehrte Einsatz von Tablets im Deutschunterricht ermöglicht Lehransätze, die individuell, authentisch und kollaborativ sind und sich an den gesellschaftlichen, pädagogisch-didaktischen und fachlichen Veränderungen orientieren (Aufenanger, 2020; Döring & Mohseni, 2020; Frederking et al., 2018; Kearney et al., 2020; KMK, 2021; Möbius et al., 2015). Digital Storytelling (DS) bietet in diesem Kontext eine Methode, die digitale Werkzeuge nutzt, um Geschichten zu erzählen, die verschiedene Medien wie Text, Bilder, Audio und Video kombinieren. Diese Form des multimodalen Erzählens unterstützt nicht nur sprachliche und narrative Fähigkeiten, sondern auch digitale Kompetenzen (Robin/McNeil 2012). „DS combines the ancient tradition of telling stories with the use of digital tools and artefacts to compose and share stories“ (Rubio-Hurtado et al., 2022).

Die zentrale Forschungsfrage lautet: Fördert Digital Storytelling in Form von multimodalem Erzählen die narrative Textproduktionskompetenz von Schüler*innen der Orientierungsstufe?

Das Forschungsprojekt basiert auf einem Pre-Post-Test-Design mit zwei Messzeitpunkten. Zu Beginn der Studie verfassen die Schüler*innen eine Erstfassung ihrer Texte (Pretest), nach der Intervention folgt die Überarbeitung zu einer Zweitfassung (Posttest). Der Vergleich dieser beiden Textversionen soll Aufschluss über die Entwicklung der narrativen Textproduktionskompetenz geben.

Nach einem einführenden Lernpfad zur Textsorte „Lügengeschichten“ erhalten die Schüler*innen eine der folgenden drei Fördermaßnahmen:

1. Interventionsgruppe I: Digital Storytelling (DS) – Die Schüler*innen nutzen multimodale Werkzeuge wie Bild, Ton und Video, um ihre Geschichten weiterzuentwickeln. Ziel dieser Intervention ist es, durch die Nutzung digitaler Medien ihre narrative Textproduktionskompetenz zu fördern und ihre Fähigkeit zur Selbstanalyse ihrer Geschichten zu stärken. Diese Intervention basiert auf den „Adaptable Books“ (Hauck-Thum, 2018).

2. Interventionsgruppe II: Textanalyse – Die Schüler*innen analysieren Mustertexte und ihre eigenen Lügengeschichten mit einem Fokus auf „Ereignis“, „Situation“ und „Figur“. Durch diese Analyse sollen Schwächen in den Geschichten aufgedeckt und gezielte Überarbeitungen ermöglicht werden. Diese Intervention orientiert sich am Regensburger Schreibmodul „Burg Adlerstein“, das im Rahmen der RESTLESS-Studie entwickelt wurde (Schilcher et al., 2020; Wild, 2020).

3. Kontrollgruppe: Diese Gruppe erhält keine spezifische Unterstützung zur Überarbeitung ihrer Texte.

Das quasi-experimentelle Design gewährleistet die Generalisierbarkeit der Ergebnisse auf reale Schulsituationen. Die Durchführung in einem schulischen Umfeld stellt sicher, dass die erhobenen Daten eine hohe externe Validität aufweisen. Um die interne Validität zu verbessern, wurden die Schulklassen gleichmäßig auf die Interventionsgruppen verteilt. Die Durchführung der Interventionen erfolgt durch den Versuchsleiter, um Konsistenz und Vergleichbarkeit sicherzustellen (Steinhoff & Marx, 2017).

Im Rahmen der Datenerhebung werden verschiedene Kovariate erfasst, darunter der Frankfurter Leseverstehenstest (FLVT) (Souvignier et al., 2008), der SMiSK-Fragebogen zur Schreibmotivation (Hennes et al., 2023), SES mittels Family Affluence Scale (Currie et al., 2024), Tastaturschreibkompetenz, Literalitätserfahrung, Mediensozialisation, Familiensprache, Alter und Geschlecht. Auf Basis einer Poweranalyse werden derzeit (Oktober bis Dezember 2024) in neun Realschulklassen der 6. Jahrgangsstufe in Karlsruhe rund 200 Erst- und Zweitfassungen erhoben.

Die Wirksamkeit der Interventionen wird sowohl holistisch mit der NAEP-Skala (Gorman, 2010) als auch analytisch mit der RANT-Skala (Wild, 2020) gemessen. Die Texte werden von vier Ratern (2x NAEP, 2x RANT) bewertet, um eine umfassende Analyse der Textqualität zu ermöglichen und die Entwicklung der narrativen Textproduktionskompetenz der Schüler*innen differenziert zu erfassen.

Dieses Projekt ist Teil des Forschungs- und Nachwuchskollegs „Aufgabenqualität im digital gestützten Unterricht“ (AQUA-d) der Pädagogischen Hochschule Karlsruhe und der Universität Tübingen, gefördert durch das MWK Baden-Württemberg. Ziel des Kollegs ist es, die Qualität digitaler Aufgaben im Schulunterricht zu steigern und die Potenziale digitaler Werkzeuge umfassend zu nutzen.

Strategiebasierte Förderung literarischer Kompetenzen. Eine Interventionsstudie in der Sekundarstufe I

Magdalena Schlint

Universität Passau, Deutschland

Diese Dissertationsstudie beschäftigt sich mit der Förderung literarischer Kompetenz von Schüler:innen der 6./7. Jahrgangsstufe. Untersuchungen zeigen, dass literarische Kompetenz an medial unterschiedlichen Texten gefördert werden kann: Literarische Kompetenz, erworben an populären audiovisuellen Texten, kann auf kanonische Printliteratur übertragen werden (Scheubeck, 2023), narrative Computerspiele initiieren literarisches Lernen (Boelmann, 2015; Emmersberger, 2022) und der Einsatz künstlerischer Bilder trägt zur Bildung von Deutungshypothesen bei (Heese, 2021). Ausgehend vom Kompetenzmodell Literarischen Lernens auf semiotischer Grundlage (Schilcher/Pissarek, 2018) wurden Strategien für die Analyse und Interpretation von (multimedialen) Texten entwickelt. Für die Lernbereiche Lesen und Schreiben sind Strategien bereits etabliert (Philipp, 2012, 2016; Leubner/Saupe, 2014; Bay, 2022) und ihre Vermittlung durch das Self-Regulated Strategy Development (SRSD) nach Graham/Harris stellt einen der effektivsten Förderansätze im Bereich Schreiben dar (Philipp, 2012, 2014; Graham/Perin, 2007), bei dem auch die Selbstwirksamkeitserwartung (Bandura, 1997) gestärkt wird. Im Literaturunterricht hingegen sind Strategien bisweilen noch nicht explizit Gegenstand der Förderung. Diese Interventionsstudie untersucht

Mit diesem Desiderat befasst sich diese Studie, der drei Forschungsfragen zu Grunde liegen:

1. Ist der Zuwachs an lit. Kompetenz von Schüler:innen durch einen Literaturunterricht mit der Methode der Modellierung von Strategien größer als in einem Unterricht mit der Methode des Interpretationsgesprächs?
2. Gibt es einen Zusammenhang zwischen der Ausprägung lit. Kompetenz an printliterarischen Texten und der an audiovisuellen Texten?
3. Verändert sich die Selbstwirksamkeitserwartung (SWE) von Schüler:innen bei der Bearbeitung literarischer Aufgaben durch die Verwendung von Strategien?

In einem Zwei-Gruppen-Design wird die Methode der Strategievermittlung mit der des Interpretationsgesprächs (vgl. dazu z.B. Zabka et al., 2022, S. 134-164; Spinner, 2019; Kammer, 2004) verglichen. Dazu führen beide Gruppen Interventionen im Umfang von acht Unterrichtsstunden durch, die auf denselben Lernzielen und Textgrundlagen basieren, sich aber in der Vermittlung unterscheiden: Während in der Experimentalgruppe Strategien nach der SRSD-Methode vermittelt werden, wird in der Kontrollgruppe nach der Methode des Interpretationsgesprächs unterrichtet.

In einer Prä-Post-Messung wird die Veränderung der lit. Kompetenz bei printliterarischen und audiovisuellen Narrationen sowie die Entwicklung der Selbstwirksamkeitserwartung erfasst. Die Erhebung fand im Schuljahr 2023/24 an 16 Klassen mit über 400 Schüler:innen der 6./7. Jahrgangsstufe statt.

Die Auswertung der geschlossenen Fragen zeigt folgende erste Ergebnisse:

- Der Zuwachs der literarischen Kompetenz ist bei Schüler:innen, die durch die Methode der Modellierung lernen, signifikant größer als bei denen, die mit der Methode des Gesprächs lernen.
- Insgesamt zeigt das Training in Bezug auf die Zuwächse der literarischen Kompetenz einen mittleren Effekt nach Cohen.
- Es gibt einen signifikanten, aber geringen Zusammenhang zwischen den Zuwächsen der literarischen Kompetenz bei Printtexten und Filmen.
- Die SWE nimmt in allen Gruppen signifikant zu. In der Gruppe Modellierung gibt es einen kleinen Effekt nach Cohen.

Teacher Training Competences (TeTraCom+) - Validierungsstudie

Franziska Ohl

Johannes Gutenberg Universität Mainz, Deutschland

Self-reflection on educational skills plays a role in developing student teachers by enabling them to evolve their teaching skills critically. However, the question arises is whether students recognize different teacher training competences, distinguish between them, and assess them for themselves. This study aims to develop a competence reflection tool for student teachers based on established standards (e.g., Kultusministerkonferenz, 2022; Kunter & Baumert, 2006; Metsapelto et al., 2022). This tool serves as a foundation for autonomous and repeatable self-reflection, enabling students to cultivate a habit of reviewing their learning success, identifying further learning needs, and monitoring their progress (TeTraCom+). A seven-factor structure was found in a multi-study analysis (N1 = 508; N2 = 426), revealing seven competency areas. This study's findings, particularly the practical applicability of competencies for student teachers, have direct and significant implications for teacher training and education, making it a valuable resource for educators and professionals. These data are to be replicated and validated in a new study (planned N = 350).

The Effect of Valence and Text Difficulty on Attention in Reading with Visual Distractors

Maré Kondruß, Klaus-Peter Wild

Universität Regensburg, Deutschland

In digital learning environments, students often face multimodal content, with educational texts competing against pop-up notifications or advertisements for the reader's attention. Because recent research suggests disruptive effects of auditive and

visual distractors on text reception and learning success (e.g., Cauchard et al., 2012; Lan 2023; Sörqvist et al., 2010; Sung & Mayer, 2015; Zhang et al., 2018), the present eye tracking study wants to investigate the effects of different visual distractor valences on attention and test performance using eye movement and performance measures. Additionally, text difficulty and the individual ability of attention control and text interest are investigated. In the experimental study, participants read six modified texts from textbooks on human anatomy with either low or high text difficulty on screen with a screen-based eye tracker. Text difficulty is determined through the number of anatomical terms and connections between paragraphs within texts. We use commercials as distractors that were rated on attractiveness in a precedent study. While participants read, a distractor with positive valence, neutral valence, or no distractor is presented on the right side of the screen. Following each text, participants answer three multiple-choice questions on the anatomical terms and relationships explained in the text. Results are analysed using linear and negative binomial mixed models. For eye movement measures (e.g., Cauchard et al., 2012; Lan et al., 2023), higher (relative) fixation times, number of fixations, saccade length, regression probability, and shifts between text and distractor in the positive in contrast to the neutral or no distractor condition are expected, as well as in the high in comparison to the low text difficulty condition. Additionally, an increased distractor effect with high text difficulty is expected. Considering the ongoing digital transformation in higher education, the study aims to retrieve practical implications for (digital) learning environments.

Validierung eines Instruments zur Erfassung von Einstellungen zu Inklusion von Lehramtsstudierenden im Hinblick auf Schüler:innen mit Verhaltensauffälligkeiten

Felix Piegsda

Universität Erfurt, Deutschland

Abstract

Einstellungen gelten als wichtige Grundlage kompetenten Lehrer:innenhandelns im Unterricht und gleichzeitig als Voraussetzung einer erfolgreichen Umsetzung von Inklusion an Schulen (Reusser & Pauli, 2014). Bisherige Messinstrumente zur Erhebung von Einstellungen in Bezug auf inklusiven Unterricht fokussieren meist ausschließlich kognitive Einstellungen im Hinblick auf allgemeine inklusive Settings (Junker et al., 2020). Dabei zeigen Untersuchungen, dass die Ausprägung der Einstellungen von der Art der Behinderung der zu unterrichtenden Kinder abhängig ist (Bosse & Spörer, 2014). Zudem werden in der Literatur drei Komponenten von Einstellungen beschrieben: affektiv, konativ und kognitiv (Rosenberg & Hovland, 1960). Die kognitive Komponente wiederum beinhaltet eine auf Wissen und Erfahrung basierende Objektbewertung bspw. hinsichtlich des Werts oder aufzubringenden Kosten (Eagly & Chaiken, 1993; Ajzen, 2001). Aus der Literatur lassen sich daher vier Komponenten von Einstellungen herausarbeiten: Affektion, Konation sowie Kognition-Wert und Kognition-Kosten.

Daraus ergibt sich für die in der Posterpräsentation vorzustellende Studie das Forschungsdesiderat, ein Erhebungsinstrument zu entwickeln und zu validieren, das vier Komponenten von Einstellungen hinsichtlich Verhaltensauffälligkeit als Förderbedarf sowie Behinderung allgemein valide erfassen kann.

Mittels eines Online-Fragebogens wurden 330 angehende Lehramtsstudierende verschiedener Universitäten in Deutschland befragt. Die Einstellungen zu beiden Heterogenitätsdimensionen wurden durch jeweils 17 parallel formulierte Items abgefragt. Fünf Items fragen nach den Einstellungen Kognition-Wert. Die anderen drei Einstellungskomponenten wurden durch jeweils vier Items abgefragt. Die Auswertung erfolgte mit R durch explorative (EFA) und konfirmatorische (CFA) Faktorenanalysen.

In der EFA als auch der CFA wurden die beiden Dimension jeweils getrennt ausgewertet. Dabei zeigen die Ergebnisse, dass die Ausgangsmodelle der Dimensionen Verhaltensauffälligkeiten und Behinderung allgemein keinen zufriedenstellenden Modellfit ergeben. Eindeutig abbilden lässt sich in beiden Dimensionen lediglich die konative Einstellungskomponente. Erst durch Testung weiterer Modelle und dem Ausschluss von vier Items mit Trennschärfen $< .4$ lassen sich die vier Komponenten in beiden Dimensionen darstellen. Die ähnlichen Ergebnisse der beiden Dimensionen finden sich in der konfirmatorischen Faktorenanalyse nicht wieder. So erreicht die Dimension Verhaltensauffälligkeiten auf Basis des durch die EFA optimierten Modells und Ausschluss eines weiteren Items einen zufriedenstellenden Modellfit ($\chi^2(58)=75.712$, $p=.077$, $RMSEA=.044$, $SRMR=.053$, $CFI=.973$). Hingegen zeigt die Dimension Behinderung allgemein in einem ähnlichen Modell keinen zufriedenstellenden Modellfit ($\chi^2(59)=109.798$, $p=.000$, $RMSEA=.078$, $SRMR=.071$, $CFI=.919$). Zufriedenstellend ist der Modellfit für die Dimension Behinderung allgemein lediglich, wenn die Komponente Kognition-Kosten gestrichen wird ($\chi^2(40)=66.010$, $p=.006$, $RMSEA=.067$, $SRMR=.053$, $CFI=.950$).

Die Ergebnisse deuten darauf hin, dass bei der Erhebung von Einstellungen verschiedenen Komponenten differenziert werden können. So finden sich nach der Auswertung der EFA die vier in der Theorie beschriebenen Komponenten wieder. Doch lässt sich durch die CFA in der Dimension Behinderung allgemein keine Komponente Kognition-Kosten darstellen. Dies kann als Hinweis darauf verstanden werden, dass nicht nur Einstellungen hinsichtlich von Behinderung allgemein beachtet, sondern auch einzelne Förderbedarfe abgefragt werden sollten. Gleichzeitig zeigen die Unterschiede zwischen den Dimensionen, dass eine parallele Formulierung der Items nicht zwangsläufig zielführend ist. Gerade im Hinblick auf die Komponente zu kognitiven Kosten ist es den Studierenden womöglich leichter gefallen, sich die Konsequenzen der inklusiven Beschulung von Kindern mit Verhaltensauffälligkeiten für den Unterricht vorzustellen als ganz allgemein für Kinder mit Behinderung. Eine auf die Einstellungsdimension angepasste Formulierung führt möglicherweise bei den Studierenden zu einem klareren Bild über die alltäglichen Konsequenzen inklusiven Unterrichts.

Wie nachhaltig sind die Effekte der Covid-19-Pandemie für das berufliche Wohlbefinden von Lehrern und Lehrerinnen? Ergebnisse der COACTIV-ExpeRt-Längsschnittstudie

Meret Bunge¹, Thamar Voss², Dirk Richter³, Mareike Kunter⁴, Gyde Wartenberg⁴, Uta Klusmann¹

¹Humboldt-Universität Berlin; ²Universität Freiburg; ³Universität Potsdam; ⁴DIPF - Leibniz-Institut für Bildungsforschung und Bildungsinformation, Frankfurt

Das berufliche Wohlbefinden und die Motivation von Lehrern und Lehrerinnen spielt im Schulalltag eine wichtige Rolle für die Unterrichtsgestaltung sowie für die Motivation und Leistung der Schüler und Schülerinnen. Verschiedene Studien haben einen Anstieg an Stress und Erschöpfung während der Covid-19-Pandemie gezeigt. Theorien wie die Set-Point-Theorie und die Conservation-of-Resources-Theorie treffen unterschiedliche Aussagen darüber, wie sich das subjektive Wohlbefinden nach einem einschneidenden Lebensereignis entwickelt, d. h. ob eine Rückkehr des Wohlbefindens zum vorpandemischen Niveau oder längerfristige Veränderungen zu erwarten sind.

Bisher sind die langfristigen Auswirkungen der Pandemie für das berufliche Wohlbefinden und die Motivation im Lehrberuf noch weitestgehend unbekannt. Daher soll im aktuellen Beitrag die Frage untersucht werden, wie sich das berufliche Wohlbefinden

und die Motivation nach der Pandemie entwickelt haben. Merkmale wie die Selbstregulation sowie Unterstützung und Zusammenarbeit im Schulkontext wurden als Prädiktoren herangezogen. Im Rahmen der COACTIV-ExpeRt-Längsschnittstudie wurden 432 Lehrer und Lehrerinnen zu insgesamt vier Messzeitpunkten vor, während und nach der Pandemie zu ihrer emotionalen Erschöpfung, ihrem Unterrichtsenthusiasmus und ihrer Berufszufriedenheit befragt.

Latente Differenzmodelle ergaben, dass die durch die Pandemie erhöhte emotionale Erschöpfung im Laufe der Pandemie und danach signifikant gesunken und der Unterrichtsenthusiasmus signifikant wieder gestiegen ist. Die emotionale Erschöpfung unterscheidet sich jedoch weiterhin signifikant von ihrer vorpandemischen Ausprägung. Die Berufszufriedenheit veränderte sich im Verlauf der Pandemie und danach nicht signifikant. Mithilfe von Regressionsanalysen wurde festgestellt, dass der G-Typ mit einer höheren Berufszufriedenheit zusammenhängt. Eine Ressource für die emotionale Erschöpfung nach der Pandemie stellte zudem die Unterstützung durch die Schulleitung dar.

Die Ergebnisse zeigen den Förderungsbedarf des beruflichen Wohlbefindens insbesondere nach herausfordernden Ereignissen wie der Pandemie auf.

Wirksamkeit unterschiedlicher Fortbildungsformate auf das fachdidaktische Wissen und Handeln pädagogischer Fachkräfte im Bereich der Sprachbildung

Stefanie Horner

Rheinland Pfälzische Technische Universität Kaiserslautern Landau, Deutschland

In den letzten Jahren sind alltagsintegrierte Ansätze zur Unterstützung der Sprachentwicklung von Kindern zunehmend in den Fokus gerückt (Schieler, 2024). Die Nutzung spontaner Sprechansätze stellt jedoch hohe Anforderungen an die Kompetenzen der pädagogischen Fachkräfte, da pädagogische Alltagssituationen in hohem Maße offen, komplex und unvorhersehbar sind (Fröhlich-Gildhoff, Nentwig-Gesemann, Pietsch, Köhler & Koch, 2014). Empirische Studien aus dem nationalen und internationalen Raum zeigen, dass die für die Sprachentwicklung von Kindern bedeutsame Anregungsqualität in Kindertageseinrichtungen eher gering ausgeprägt ist (z.B. Eckhardt & Egert, 2017; Peisner-Feinberg, Schaaf, Hildebrandt & Pan, 2015). Vor diesem Hintergrund sind Fortbildungsmaßnahmen zur Weiterentwicklung der Anregungsqualität pädagogischer Fachkräfte vielversprechend. Auf der Grundlage von Logikmodellen wird davon ausgegangen, dass Fortbildungen zur Veränderung der Kompetenzen von pädagogischen Fachkräften führen und dadurch die Anregungsqualität weiterentwickelt werden kann (Egert, Fukkink & Eckhardt, 2018). Diese Wirkmechanismen wurden bislang kaum nachgewiesen (Markussen-Brown et al., 2017). Damit bleibt unklar, inwiefern die Veränderung spezifischer Kompetenzfacetten, wie beispielsweise fachdidaktisches Wissen, die Anregungsqualität nach der Fortbildung beeinflusst. Des Weiteren wurde noch nicht hinreichend geklärt, welche Bedeutung unterschiedliche Fortbildungsformate für die Weiterentwicklung der Anregungsqualität haben. Es liegen sowohl Ergebnisse von Metaanalysen vor, die zeigen, dass individualisierte Formate wie Coaching besonders effektiv sind (Egert, Dederer & Fukkink, 2020), als auch solche, die die Kombination verschiedener Formate (z. B. Kurs und Coaching) als wirksam hervorheben (Werner, Linting, Vermeer & Van IJzendoorn, 2016).

Vor dem Hintergrund dieser Forschungsdesiderate sollen folgende Forschungsfragen untersucht werden:

(1) Wie wirken sich Coaching, Fortbildung und die Kombination beider Formate auf das fachdidaktische Wissen und die Anregungsqualität von pädagogischen Fachkräften aus?

(2) Werden Effekte der Fortbildungsformate auf die Anregungsqualität durch das fachdidaktische Wissen mediert?

Zur Beantwortung dieser Fragestellungen wurde eine experimentelle Prä-Post-Studie mit randomisierter Zuordnung der pädagogischen Fachkräfte ($N = 120$) zu einem der drei Fortbildungsformate oder einer Wartekontrollgruppe durchgeführt, wobei der Umfang der Fortbildungen vergleichbar gehalten wurde. Inhaltlich basieren die Fortbildungsformate auf dem Ansatz „Mit Kindern im Gespräch“ zur alltagsintegrierten Sprachbildung, für den bereits eine langfristige Erhöhung der Anregungsqualität durch ein kombiniertes Format nachgewiesen werden konnte (Kammermeyer et al., 2019). Die Anregungsqualität wurde mit dem hoch-inferenten Verfahren *Classroom Assessment Scoring System 2nd Edition* (CLASS, Pianta, La Paro & Hamre, 2008; vgl. auch Teachstone, 2022) erfasst. Die Daten werden mittels Strukturgleichungsmodellen mit der Software lavaan (Rosseel, 2012) in RStudio ausgewertet.

Zeit - Eine qualitative Untersuchung des Erlebens von Grundschulkindern

Elisabeth Dötter-Burnickl

Universität Regensburg, Deutschland

Die Grundschule zielt darauf, eine differenzierte Perspektive auf *Zeit* auszubilden und damit die Identitätsentwicklung der Schüler*innen zu unterstützen. Wenn Kinder lernen sich mit und in der *Zeit* zu orientieren, erschließen sie sich ihre Lebenswelt. Dies trägt auch dazu bei, die eigenen Sinn- und Wertvorstellungen diesbezüglich zu reflektieren (BayEUG: Art. 1 Bildungs- und Erziehungsauftrag; Köhnlein 2022, 100 ff.; Tänzer, Schomaker & Heran-Dörr 2014, 25 ff.). Dazu bedarf es Kenntnisse über Zeitvorstellungen im Grundschulalter und wie Kinder die Frage - Was ist *Zeit*? - beantworten. Denn für Didaktik und Unterricht ist es bedeutsam, dass Schüler*innenvorstellungen sowohl bekannt sind als auch Lehrkräfte diese berücksichtigen (Möller 2018, 44 ff.; Adamina, Kübler, Kalcsics, Bietenhard & Engeli 2018, 11). Deshalb ist es „[e]ine Aufgabe von Lehrer*innen (...) [die] Zeitvorstellungen ihrer Schüler*innen kennen zu lernen (...)“ (Kühberger 2012, 52). „[Doch] der Bedarf scheint hoch, Methoden für eine qualitative (...) Diagnostik im Bereich des Zeitkonzepts zu entwickeln (...)“ (Hofmann-Reiter 2015, 331). Zusammenfassend „[gibt] es (...) massive Forderungen (...), die Thematik ‚Zeit‘ in all ihren Facetten stärker in die Bildung und den Schulunterricht zu integrieren, [dennoch] wird diesem Wunsch nach wie vor wenig Rechnung getragen. Bei Fragen rund um das Thema ‚Zeit‘ handelt es sich um eine (...) vernachlässigte Komponente der Bildungslandschaft (...)“ (Schmidt 2016, 204). Daran knüpft die Dissertation an: Die qualitativ-rekonstruktiv angelegte Studie untersucht Erleben/-weisen (Marton & Booth 2014, 174, 177) von Grundschulkindern zum Phänomen *Zeit*. Da zu diesem Forschungsgegenstand bislang kaum empirisch belegte Erkenntnisse vorliegen, wird dieser durch einen phänomenographischen Zugang (Kallweit, Lüschen, Murmann, Pech & Schomaker 2019, 46 f.) konkretisiert und untersucht. Die phänomenographische Untersuchung kennzeichnet eine *Perspektive zweiter Ordnung* (Marton & Booth 2014, 209), entsprechend erfolgt eine Beschreibung des Phänomens aus der Perspektive der Lernenden. Um sich dieser anzunähern, wurde ein zweistufiges Design entwickelt: Das Sample setzt sich dabei aus 26 Kindern der zweiten Jahrgangsstufe der bayerischen Grundschule zusammen. Für die Untersuchung wurden thematische Erzählbilder (Gläser 2014) zu einer Rahmengeschichte erhoben und diese als Gesprächsanlass für ein anschließendes narratives Einzelinterview mit problemzentrierten Komponenten (Vogl 2015) genutzt. Die Bilder werden mittels qualitativer Inhaltsanalyse nach Holzwarth (2006) und Kuhn (2007) deskriptiv analysiert und sachbezogen interpretiert. Daran knüpft eine

phänomenographische Auswertung der Einzelinterviews nach Marton und Booth (1997, 2014) sowie Dahlgren und Fallsberg (1991) an. Intention ist, die gewonnenen Erkenntnisse abschließend für methodisch-didaktische Impulse fruchtbar zu machen. Eine Aufarbeitung der gewonnenen Forschungsergebnisse ist in verschiedenen Formen denkbar: Das kann beispielsweise ein *Concept Cartoon* sein, um mit Schüler*innen zum Phänomen ins Gespräch zu kommen und Erlebensweisen verbalisieren zu lassen.

Das Poster stellt Auswertungsstrategien und erste Ergebnisse, die das Phänomen in seiner Variation aus Lerner*innenperspektive nachzeichnen, vor.

Zweitberuf Religionslehrer:in. Eine Berufsinteressensforschung von Quer- und Seiteneinsteiger:innen in den Lehrer:innenberuf für Katholische Religion

Maria Hochwarter

KPH Edith Stein, Österreich

Das Berufsfeld der Pädagog:innen wird in verschiedenen gesellschaftlichen Kontext, einschließlich des Alltags, der Politik sowie den Medien, intensiv diskutiert und mit einer Vielzahl von Aspekten in Beziehung gesetzt. Häufig werden mögliche Vorteile dieses Berufsfeldes als Vorurteile formuliert und als primäre Motivationen für die Berufswahl als Lehrkraft dargestellt. In bildungspolitischer und medialer Betrachtung rückt jedoch in den jüngsten Jahren ein übergeordnetes Thema in den Fokus, das treffend unter dem Terminus "Lehrkräftemangel" subsumiert werden kann.

Die EU-Kommission hat 2020 in der Mitteilung zum Bildungsraum bis 2025 den Mangel an Lehrpersonal thematisiert. Es braucht Maßnahmen den Lehrkräftemangel entgegen zu wirken vor allem in den Bereichen der MINT-Fächer (Mathematik, Informatik, Naturwissenschaften und Technik) sowie bei der Betreuung von Schüler:innen mit sonderpädagogischem Förderbedarf.¹ Auf den Umstand des fehlenden Lehrpersonals hat die Republik Österreich die Bildungsoffensive "Klasse Job" ins Leben gerufen, welche darauf abzielt, das Lehrpersonal aufzustocken. Dabei sollen einerseits junge Menschen für den Lehrer:innenberuf begeistert werden und andererseits Quereinsteiger:innen über die Möglichkeiten informiert werden, Lehrkraft zu werden. „Um dieser steigenden Nachfrage an Religionslehrer:innen gerecht zu werden, wurde dieser Hochschullehrgang konzipiert. Der Abschluss des Hochschullehrgangs ist die Grundlage für die außerordentliche Befähigung zur Erteilung von konfessionellem Religionsunterricht in der Primarstufe. Diese kann nach Zertifizierung des Lehrganges entsprechend den Vorschriften der jeweiligen Kirchen und Religionsgemeinschaften von diesen ausgestellt werden.“

Die Forschungsfrage dieses Projekts liegt im Bereich der Berufsmotivationsforschung. Sind es pragmatische Gründe wie beispielsweise jene bereits formulierten Vorteile des Lehrer:innenberufs die Menschen zum Quer- und Seiteneinstieg in den Beruf des:r katholischen Lehrperson motivieren oder sind es andere Beweggründe. Daher liegt der Schwerpunkt auf dem Verständnis und der Identifikation der Gründe für den Berufswechsel bzw. -einstieg hin zum Beruf (Religions)Lehrer:in. Ein besonderes Augenmerk liegt auf dem Prozess der beruflichen Entscheidungsfindung, der Selbsteinschätzung der Eignung sowie der Rolle sozialer Einflüsse bei der Festlegung auf den Beruf. Zusätzlich wird untersucht, ob sich die Motive von Quer- und Seiteneinsteigern signifikant von denen von Lehramtsstudierenden im herkömmlichen Bildungsweg unterscheiden.

Bei der Gruppe von Quer- und Seiteneinsteiger:innen handelt es sich keinesfalls um eine homogene Gruppe, sie unterscheiden sich unter anderem in ihren individuellen Biografien signifikant von Lehramtsstudierenden des ersten Ausbildungsweges. Im weiteren Verlauf wird die Frage nach den unterschiedlichen Motiven von Theologiestudierenden im Lehramt und der untersuchten Personengruppe aufgeworfen. Eine Spezifizierung des Forschungsumfelds erfolgt durch die Festlegung, dass die Teilnehmer:innen Lehrkräfte für katholische Religion in der Primar- und Sekundarstufe 1 sein werden. Eine Erweiterung der Stichprobe ist nicht vorgesehen.

Im Rahmen dieser Forschung liegt der Fokus auf der Untersuchung der Motive von Quer- und Seiteneinsteiger:innen für ihren Berufswechsel in den Lehrer:innenberuf (Primar- und Sekundarstufe 1). Dabei sollen entscheidende Faktoren und Aspekte für diesen Berufswechsel identifiziert werden. Ein spezielles Augenmerk liegt auf der Unterscheidung zwischen den Motiven von Personen im zweiten Bildungsweg und Studierenden, die sich im ersten Bildungsweg für den Lehrer:innenberuf entschieden haben, sowie Theologiestudierenden im Lehramtsstudium.

Die Ergebnisse dieser Forschung sollen dazu beitragen, die Motive von Quer- und Seiteneinsteiger:innen für ihren Berufswechsel zu verstehen. Zudem sollen mögliche Unterschiede in den Motiven zwischen dieser Gruppe und regulären Lehramtsstudierenden untersucht werden, um die Auswirkungen von Geschlecht, Alter, sozioökonomischem Hintergrund, Glauben und anderen Faktoren auf die berufliche Entscheidung besser zu verstehen. Dadurch soll ein tieferes Verständnis für die Beweggründe zum Berufswechsel erlangt werden. Die Ergebnisse dieser Forschungsarbeit können nicht nur als Ausgangspunkt für ein besseres Verständnis von Quer- und Seiteneinsteiger:innen dienen, sondern auch als Grundlage für zukünftige Interventionen in der Forschung, um gezielte Programme zur Erhaltung und Steigerung der Motivation im Berufsalltag zu entwickeln. Des Weiteren können die Ergebnisse Erkenntnisse für die Fort- und Weiterbildung liefern, um die individuellen Karrieremöglichkeiten von Lehrer:innen im zweiten Bildungsweg zu erweitern. .

Nachwuchstagung: Postersession "Erhebungsdesign"

Zeit: Donnerstag, 30.01.2025: 9:15 - 10:45 · Ort: O 001

Adaptive Teaching in the Digital Classroom: A Systematic Meta-Review

Marie Pauline Godow^{1,2}, Marlit Annalena Lindner^{1,2}

¹EUF - Europa-Universität Flensburg; ²IPN - Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik

Theoretical Background

Students' differences in cognitive abilities, prior knowledge, interest, or motivation create a complex teaching environment, where traditional one-size-fits-all teaching approaches do not suffice to ensure effective learning for all students (Vock & Gronostaj, 2017). To address student heterogeneity, adaptive teaching has emerged as a promising teaching approach for decades, where teachers continuously modify the learning content and instruction to meet the individual learning needs of each student (Corno, 2008). However, adaptive teaching necessitates teacher resources, for instance, the continuous assessment of students' current performance and progress (i.e., formative assessment; Decristan et al., 2017). The digitalization of education and teaching has introduced new opportunities for designing and implementing learning and teaching processes. Digital tools, such as Intelligent Tutoring Systems, are capable of analyzing student performance and providing individualized feedback in real-time. The implementation of educational technologies offers the potential to facilitate adaptive teaching by automating processes that would otherwise require significant teacher effort and resources (Lachner et al., 2020).

Relevance and Research Questions

Yet, despite promising indications of adaptive teaching, especially in combination with digital tools, empirical research is limited and still developing (Dumont et al., 2024). The field is further marked by diverse terminologies, leading to separated research strands which impedes a comprehensive understanding of adaptive teaching. Based on existing reviews and meta-analyses, our systematic meta-review aims to evaluate the current research trends and to examine what is known about the implementation of adaptive teaching and its effects on learning – especially with the use of digital media. The review addresses the following research questions:

1. How is adaptation in learning and teaching conceptualized and implemented in empirical research and teaching practice according to systematic reviews and meta-analyses?
2. Which approaches of implementing technology to support adaptive learning and teaching have been identified and described as particularly beneficial?
3. What are the cognitive and affective-motivational effects of adaptive learning and teaching practice, as reported in the synthesized literature?
4. Which factors moderate the effectiveness of adaptive learning and teaching approaches?

Methods

In our investigation of the literature, we employ a PRISMA-guided process to systematically search academic databases (PsycINFO, Web of Science, ERIC) with a pre-defined search string. Given the diverse terminology in the field of adaptive teaching, we anticipate that the search string alone may not capture all relevant studies. Therefore, reference lists in identified articles are screened for additional studies of interest (i.e., "snowballing" technique; Greenhalgh & Peacock, 2005).

In a detailed step-by-step screening process, titles, keywords, and abstracts are read in the initial stage, followed by full-text screening of the remaining articles. The meta-review includes those articles that (1) involve an adaptive teaching or learning approach, (2) are written in English or German, (3) are a review or meta-analysis, (4) refer to elementary, secondary, or higher education, and (5) are available in full-text.

Data extraction is carried out systematically in a predefined coding table that captures key information, such as authors, year, title, country, journal, terminology, technology use, research methods, outcomes, or moderator variables.

Outlook

The present work is in an early stage, currently in the systematic literature search. Once all articles have been identified, coded, and analyzed, our meta-review aims to provide a comprehensive synthesis of the current state of research on adaptive learning and teaching. For teaching practice, it can offer an overview connecting the findings regarding different adaptive teaching concepts with their practical implementation, impact on learning, and the effective use of digital media. For research, we aim to identify both well-studied and underexplored research areas and connect empirical findings to direct future research on adaptive learning and teaching.

Beliefs und Selbstwirksamkeitserwartungen zur Integration gesellschaftlicher Fragen in den Informatikunterricht: Entwicklung eines Erhebungsinstruments

Juliane Sperling, Nils Dyck

Friedrich-Schiller-Universität Jena, Deutschland

Beliefs und Selbstwirksamkeitserwartungen (SWE) von Lehrpersonen gelten neben professionellem Wissen als zentrale Einflussfaktoren für die Unterrichtsqualität und werden zunehmend im Kontext des Informatikunterrichts untersucht (z. B. Kartal & Başarmak, 2022; Ronan et al., 2023). Parallel richten sich Diskurse der Fachdidaktik Informatik vermehrt auf soziotechnische Fragestellungen im Schulunterricht (Rücker, 2023; Morales-Navarro & Kafai, 2023). Trotz der curricularen Verankerung der fokussierten Kompetenzen (z. B. im Themenfeld *Informatik, Mensch, Gesellschaft*; IMG (GI, 2016)) ist die unterrichtspraktische Umsetzung, und damit einhergehend Beliefs und SWE von Informatiklehrpersonen im Bereich IMG, bisher kaum erforscht. Das in diesem Beitrag vorgestellte Erhebungsinstrument gliedert sich daher an eine entsprechende Untersuchung an und soll zur systematischen Beantwortung folgender Forschungsfrage beitragen: Welche Beliefs und Selbstwirksamkeitserwartungen haben Informatiklehrpersonen hinsichtlich der Integration des Themenbereichs *Informatik, Mensch, Gesellschaft* in den Informatikunterricht?

Als Erhebungsinstrument dient ein Online-Fragebogen, der sich neben einem Teil zu demographischen Angaben in einen Teil zu Beliefs und SWE hinsichtlich des Unterrichtsfachs Informatik und des Themenbereichs IMG sowie in einen Teil, in dem die Teilnehmenden Items zu Beliefs und SWE bezüglich spezifischer Kompetenzen bearbeiten, gliedert.

Die neun Items des allgemeinen Belief-Teils basieren auf einem validierten Instrument zu Sichtweisen angehender Biologielehrpersonen zu gesellschaftswissenschaftlichen Unterrichtsthemen (Kara, 2012) und wurden für den Informatikunterricht adaptiert (z. B. „Ich denke, dass es sinnvoller ist, das gesellschaftliche, soziale und ethische Aspekte in anderen Fächern als im Informatikunterricht behandelt werden“). Die 18 Items zur Erfassung der SWE basieren sowohl auf Kara (2012) als auch auf Items von Ronan et al. (2023) zur Erhebung von Einstellungen und Überzeugungen von Lehrpersonen in Bezug auf Informatik. Da ein erheblicher Teil praktizierender Informatiklehrpersonen kein grundständiges Lehramtsstudium absolviert hat, wird die SWE für das Unterrichtsfach im Allgemeinen als Kontrollgröße für die SWE im Bereich IMG erhoben (jeweils neun Items, z. B. „Ich traue mir zu, Lehr- und Lernmaterialien für den Informatikunterricht/für IMG zu erstellen“).

Der dritte Fragebogenteil erfasst Beliefs und SWE zur Vermittlung von 14 konkreten IMG-Kompetenzen, die in einem mehrstufigen Verfahren anhand von Lehrplänen aller Bundesländer für den Informatikunterricht in den Sekundarstufen I und II ermittelt wurden: Nach der Identifikation von 124 Kompetenzen, die eindeutig dem Themenbereich IMG zuzuordnen sind, ergaben sich auf Basis der Kompetenzformulierungen fünf Kategorien: (1) Daten- und Informationssicherheit, (2) Gesellschaftliche und individuelle Wechselwirkungen, (3) Rechtliche und ethische Rahmenbedingungen, (4) Historie und Zukunftsperspektiven, (5) Technologische Grundlagen. Die Kodierung der Kompetenzen durch zwei unabhängige BewerterInnen erzielte eine Interrater-Reliabilität von $\kappa = 0,78$ (substanzial bis nahezu perfekt). Unstimmigkeiten wurden durch Konsenskodierung gelöst. Aus den kategorisierten Kompetenzen wählten drei BewerterInnen 14 repräsentative Kompetenzen aus, die die Inhalte und den Umfang der fünf Kategorien möglichst präzise abbilden. Beispiele hierfür sind: „Die Lernenden diskutieren den Einfluss von Datenfilterung auf Wahrnehmung und Meinungsbildung.“; „Die Lernenden beschreiben und verwenden Verfahren zur Sicherung von Vertraulichkeit, Authentizität und Integrität von Daten.“; „Die Lernenden benennen ökologische, ökonomische und soziale Interessen bei der Gestaltung von Informatiksystemen.“. Im Fragebogen wird erhoben, ob diese Kompetenzen nach Ansicht der Lehrpersonen eher im Informatikunterricht, in anderen Unterrichtsfächern oder nicht in der Schule vermittelt werden sollen. Zudem wird erfragt, inwiefern sich Lehrpersonen in der Lage sehen, diese Kompetenzen zu vermitteln, und ob sie dies tatsächlich in ihrem Informatikunterricht tun.

Außer dem Item zum präferierten Ort der Kompetenzvermittlung werden alle Items auf einer fünfstufigen Likert-Skala (1 = „stimme nicht zu“ ... 5 = „stimme zu“) beantwortet. Neben deskriptiven Analysen (z. B. Welche SWE haben Informatiklehrpersonen abhängig von ihrer Berufserfahrung?) bieten sich unter anderem Regressionsanalysen (z. B. Lässt sich aus der allgemeinen SWE zu IMG die SWE bezüglich konkreter IMG-Kompetenzen vorhersagen?) für die Auswertung an.

Der Fragebogen wird derzeit mit Informatiklehramtsstudierenden pilotiert. Die Hauptstudie mit Informatiklehrpersonen der Sekundarstufen I und II ist im Frühjahr 2025 geplant.

Cultural Mismatch neu betrachtet: Die Rolle kultureller Peer-Normen auf die akademische Leistung von Erstakademiker*innen

Paula Schmelzer

Universität Mannheim, Deutschland

Erstakademiker*innen (Studierende, bei denen kein Elternteil einen Hochschulabschluss hat) begegnen an Universitäten häufig besonderen Herausforderungen, die sich beispielsweise in erhöhtem Stresserleben (House et al., 2020), schlechteren Leistungen (Phillips et al., 2020) und höheren Abbruchquoten (Müller & Klein, 2023) äußern. Eine Theorie, die versucht diese Phänomene zu erklären, ist die Cultural Mismatch-Theorie (Stephens et al., 2012). Diese geht davon aus, dass sich die vorherrschenden kulturellen Normen in Elternhäusern von Erstakademiker*innen von den vorherrschenden kulturellen Normen an Universitäten unterscheidet. Hierbei spielt insbesondere eine Diskrepanz zwischen interdependenten kulturellen Normen, die häufig in nicht-akademischen Kontexten vorherrschen, und den independenten kulturellen Normen, die an Universitäten dominieren eine Rolle. Einer interdependenten kulturellen Norm folgend, sollte man sich an den vorherrschenden Kontext und an Bedürfnisse anderer anpassen und mit anderen verbunden sein, während eine independente kulturelle Norm verlangt, dass man sich unabhängig vom Kontext und den Bedürfnissen anderer verhält und unabhängig von anderen ist (Markus & Kitayama, 2010). Der Theorie zufolge ist diese Diskrepanz ursächlich für schlechtere universitäre Leistungen von Erstakademiker*innen. In bisherigen Studien konnten diese Grundannahmen der Theorie bereits gezeigt werden (Phillips et al., 2020; Stark et al., 2024; Stephens et al., 2012).

In bisherigen Untersuchungen zum Cultural Mismatch von Erstakademiker*innen wurde die an der Universität vorherrschenden kulturellen Normen durch die von der Universitätsverwaltung und -leitung vorgegebene Kultur operationalisiert (Stephens et al., 2012). Die wahrgenommenen Normen an einer Institution werden allerdings nicht nur „von oben“ geformt, sondern auch „von unten“ (Morris & Liu, 2015; Stephens & Townsend, 2015). Die kulturellen Normen, denen Studierende an Universitäten ausgesetzt sind, werden also auch von anderen Studierenden geformt und mitgestaltet. In einem experimentellen Design soll nun die Rolle dieser, durch andere Studierende geprägten, kulturellen Normen für einen empfundenen Mismatch mit der Universitätskultur sowie die Auswirkung auf akademische Leistung untersucht werden.

Für dieses Vorhaben soll ein Laborexperiment mit einer studentischen Stichprobe durchgeführt werden. Für ein möglichst ausgeglichenes Verhältnis von Erstakademiker*innen und Folgeakademiker*innen soll das Experiment verstärkt in Lehramtsstudiengängen erhoben werden (Heine, 2010). Zu Beginn des Experiments sollen die kulturellen Normen von Studierenden durch einen Brief der Fachschaft, der eine interdependente oder eine independente Studierendenkultur nahelegt, manipuliert werden. Im Anschluss soll die Leistung in einer Anagrammaufgabe, sowie demografische Merkmale, einschließlich des Bildungshintergrunds, erhoben werden. Für zusätzliche weitergehende Analysen soll außerdem der individuelle kulturelle Hintergrund (independent und interdependent) der Versuchspersonen erhoben werden.

Es wird erwartet, dass in der Bedingung, in der eine independente Studierendenkultur induziert wird, bei Erstakademiker*innen ein Mismatch entsteht und diese folglich schlechtere Leistungen im anschließenden Test zeigen als Folgeakademiker*innen. Bei einer interdependenten Studierendenkultur sollte sich diese Leistungsdiskrepanz verringern.

Die aus der Studie gewonnen Erkenntnisse könnten die Cultural Mismatch-Theorie sinnvoll ergänzen. Außerdem können sie zur Diskussion beitragen, wie die universitäre Kultur und ihre Wahrnehmung besser an die Bedürfnisse von Erstakademiker*innen angepasst werden könnte. Diskutiert wird der optimale Erhebungszeitpunkt im Semester- bzw. Studienverlauf sowie die Effektivität verschiedener Manipulationsansätze der Studierendenkultur.

Einstellungen der Allgemeinbevölkerung zu schulischer Inklusion

Lea Rüger, David Scheer

Pädagogische Hochschule Ludwigsburg, Deutschland

Einstellungen zu schulischer Inklusion werden national wie international viel beforscht. Die Studien fokussieren meist direkt involvierte Akteur*innen der schulischen Inklusion, nämlich (angehende) Lehrkräfte, Eltern und Schüler*innen, und bewegen sich damit stark innerhalb der pädagogischen Disziplin(en). Es gibt bislang nur wenige Studien, die anstelle dieser Zielgruppe die Einstellungen in der Allgemeinbevölkerung betrachten. Einstellungen in der Allgemeinbevölkerung können Aufschluss darüber geben, inwieweit politische Entscheidungen zur Förderung oder Einschränkung inklusiver Praktiken akzeptiert werden und ob sich möglicherweise gesellschaftliche Veränderungen abzeichnen (Burge et al., 2008; Jury et al., 2021; Tempel, 2022). Es kann angenommen werden, dass positive Einstellungen in der Bevölkerung die Entwicklung von einem separierenden hin zu einem inklusiven Schulsystem bzw. die Akzeptanz und damit den Erfolg solcher Entwicklungen unterstützen. Deshalb sind nicht nur die Einstellungen selbst von Interesse, sondern auch Faktoren, die Varianz in den Einstellungen der Allgemeinbevölkerung aufklären. Insbesondere gilt es, Einflüsse zu identifizieren, die nicht im schulischen Kontext, sondern in der Gesellschaft zu verorten sind und den Weg in ein inklusives Bildungssystem behindern.

Ausgehend vom nationalen und internationalen Forschungsstand und den Erkenntnissen einer Pilotstudie wird das Erhebungsdesign einer Hauptstudie präsentiert, welche Einstellungen der Allgemeinbevölkerung untersucht. In einem methodisch interdisziplinär verorteten Vorhaben werden bewährte Erhebungsinstrumente der sonder- und inklusionspädagogischen Einstellungsforschung mit Items aus der politik- und sozialwissenschaftlichen Forschung kombiniert, um Wechselwirkungen zwischen gesellschaftlichen Bedingungen und schulbezogenen Einstellungen analysieren zu können. Durch ein systematisches Literaturreview und einer vorgelagerten Pilotstudie konnten bereits potenzielle Prädiktoren identifiziert werden, die in die Entwicklung des Fragebogens für die Hauptuntersuchung eingeflossen sind. Ziel ist es, ein umfassenderes Verständnis dafür zu entwickeln, wie die Allgemeinbevölkerung gegenüber schulischer Inklusion eingestellt ist und welche Faktoren die Akzeptanz oder Ablehnung inklusiver Bildung beeinflussen.

Emotionale Arbeit und Lehrperson-Schüler:innen-Beziehungen in kulturell diversen Klassen

Lara Fuchs

Universität Salzburg, Österreich

Im Schuljahr 2023/24 gaben 26% der Schüler:innen in Österreich an, im Alltag eine andere Sprache als Deutsch zu sprechen (STATISTIK AUSTRIA, 2024). Das zeigt, dass kulturelle Diversität mittlerweile fester Bestandteil der österreichischen Schullandschaft ist. Trotzdem fühlen sich viele Lehrpersonen unzureichend auf den Umgang mit kulturell diversen Klassen (KDK) vorbereitet, wodurch negative Emotionen auftreten können (Moldenhawer, 2023; Semião et al., 2023). Emotionen spielen eine zentrale Rolle im Lehrberuf (Hargreaves, 2000), da sie u.a. mit dem Unterrichtsverhalten, der Qualität der Lehrperson-Schüler:innen-Beziehung (LSB) und auch dem Wohlbefinden in Beziehung stehen (Frenzel et al., 2021). Trotz dieser Bedeutung von Lehrer:innenemotionen ist die Forschung zu Emotionen von Lehrpersonen im Kontext von Diversität begrenzt (Hry-Beihammer et al., 2019). Auch LSB in KDK sind wenig erforscht (Ialuna et al., 2024).

Für professionelles Lehrpersonenhandeln ist der adäquate Ausdruck von Emotionen bedeutsam. Das Konzept der emotionalen Arbeit bezieht sich auf die Regulierung und den Ausdruck von Emotionen, um beruflichen Erwartungen zu entsprechen (de Ruiter et al., 2021). Dabei wird unter Surface Acting, Deep Acting (Hochschild, 1983) und Modulated Acting (Horner et al., 2020) unterschieden. Je nach Qualität der LSB werden andere Formen des Actings verwendet (de Ruiter et al., 2021). Lehrpersonen unterliegen zudem Display Rules und Feeling Rules (Horner et al., 2020). Positive LSB wirken sich positiv auf das Wohlbefinden von Lehrpersonen (Spilt et al., 2011) und den schulischen Erfolg der Schüler:innen (Bergin & Bergin, 2009; Cornelius-White, 2007), insbesondere aus sozioökonomisch benachteiligten Verhältnissen (Roorda et al., 2011), aus. LSB in KDK können durch Lehrpersonen, die sich im Umgang mit kultureller Diversität als wirksam empfinden, gestärkt werden (Choi & Lee, 2020).

In der Studie wird untersucht, (1) wie sich emotionale Arbeit in kulturell diversen Klassen entfaltet und welche Rolle Feeling Rules in diesen Kontexten spielen. Es wird zudem untersucht, (2) wie Lehrpersonen LSB in KDK fördern und wie emotionale Arbeit die Qualität dieser Beziehungen mitbeeinflusst.

Die vorliegende Studie verwendet ein Mixed-Methods-Design. Im qualitativen Teil werden Leitfadenterviews mit Lehrpersonen in KDK durchgeführt, die anschließend nach qualitativer Inhaltsanalyse ausgewertet werden. Die quantitative Komponente umfasst einen Fragebogen, der auf den qualitativen Ergebnissen basiert, um die Beziehung zwischen emotionaler Arbeit (inkl. der Feeling Rules), Maßnahmen zur Beziehungsförderung und der Qualität der LSB in KDK zu testen. Dieser Teil der Studie wird als Längsschnittuntersuchung mit zwei Messzeitpunkten durchgeführt. Die Daten werden mehrbenenanalytisch ausgewertet (Schüler:innen getestet in Klassen).

Entwicklung fachdidaktischer Kompetenzen durch VR gestützte Reflexionsprozesse

Robert Mühlendorfer

Universität Mannheim, Deutschland

Theoretischer

Die Fähigkeit, verständlich und adressatengerecht zu erklären, gehört zu den zentralen Aufgaben von Lehrpersonen (Ball et al., 2005). Diese Kompetenz umfasst das sachgerechte Aufbereiten und Darstellen von Fachinhalten, sodass Lerninhalte verständlich vermittelt werden.

Angesichts der hohen Bedeutung der Erklärkompetenz ist es wichtig, diese bereits im Studium gezielt zu fördern (Findeisen et al., 2021). Der Core-Practice-Ansatz in einer Virtual-Reality (VR) – Umgebung bietet dafür eine vielversprechende Möglichkeit. Der Core-Practice-Ansatz beschreibt zentrale Unterrichtspraktiken (Grossman, Hammerness, et al., 2009). Es handelt sich um regelmäßig eingesetzte und lernwirksame Unterrichtsbausteine, die für den Erwerb von Wissen und Kompetenzen der Lernenden von Bedeutung sind (ebd.). Diese Kernpraktiken sind erlern- und anpassbar (ebd.) und bestehen aus kompakten Handlungseinheiten, die regelmäßig geübt und im Schulalltag intuitiv abrufbar sind (Fraefel, 2022). Eine VR-Umgebung ermöglicht das realitätsnahe Üben dieser Handlungseinheiten und bietet Studierenden einen sicheren Raum für praxisnahes Training mit beliebig oft wiederholbaren Bedingungen (Hamilton et al., 2021). So können Fähigkeiten gezielt und erfolgreich in realen Problemsituationen angewendet werden (ebd.).

Fragestellung

Dieser Beitrag überprüft, welche Auswirkungen eine VR-Simulation auf die Entwicklung der Erklärkompetenz im Längsschnitt hat und inwiefern ein Praxistransfer ermöglicht wird. Die Trainingsumgebung bildet das VR-Klassenzimmer Teach-R (Huang et al., 2021). Zum Ende ihres ersten Mastersemesters werden Studierende vor ihrem ersten vierwöchigen Schulpraktikum einer Experimental- (EG) und einer Wartekontrollgruppe (KG) zugeordnet. Beide Gruppen haben die Aufgabe, eine Erklärung zu einem gängigen betriebswirtschaftlichen Thema (Rechnungswesen - Abschreibung) zu gestalten. Die Unterrichtserklärungen werden

Hintergrund

und

methodischer

Zugang

während des Masterstudiums dreimal durchgeführt, vor dem ersten Schulpraktikum (t_0 Ende ersten Semester), vor dem zweiten Schulpraktikum t_1 (Ende drittes Semester) und t_2 (nach dem zweiten Schulpraktikum, zwischen drittem und vierten Semester). Die Gruppen unterscheiden sich wie folgt:

- Experimentalgruppe (EG) Die Erklärungen bei t_0 bis t_2 in der VR-Umgebung statt. Die Studierenden interagieren mit 30 Avataren, die typische Verständnisprobleme zeigen. Nach t_0 wird die Unterrichtserklärung durch zwei unabhängige Rater*innen eingestuft.
- Kontrollgruppe (KG) Eine Studierende übernimmt die Rolle der Lehrperson, während drei unabhängige Proband*innen als Lernende agieren. Diese verhalten sich entsprechend vorgegebener Rollenkarten mit spezifischen Verständnisproblemen. Auch hier erfolgt durch zwei unabhängige Rater*innen eine Einordnung als Beginner oder Intermediate.

Videografien dokumentieren das unterrichtliche Handeln zu jedem Zeitpunkt. Beide Gruppen reflektieren nach jeder Aufnahme ihre Unterrichtsqualität. Die Erklärqualität wird anhand der fünf Kriterien zur Analyse von Erklärprozessen bewertet: Fachlicher Gehalt, Lernendenzentrierung, Prozessstruktur, Repräsentation und Sprache. (Findeisen et al., 2021) Zur Vermeidung von Überforderung werden bei den Beginnern nur die wesentlichen Unterpunkte näher betrachtet, hierarchisiert und nach den Charakteristika für geeignete Kernpraktiken ausgewählt. (Grossman, Compton, et al., 2009). Die Überprüfung des Praxistransfers erfolgt durch Videografien unter realen Bedingungen am Ende des zweiten Schulpraktikums sowie durch eine authentische VR-Simulation. Dies erlaubt eine genaue Analyse der Performanz der Proband*innen im realen Umfeld, wobei die VR-Simulation gezielt Situation schafft, in denen alle Kernpraktiken beobachtet werden können.

Ergebnisse und ihre Bedeutung
Es wird erwartet, dass die adaptiv angepasste VR-Methode im Vergleich zum traditionellen Rollenspiel zu einer stärkeren und nachhaltigeren Verbesserung der praktischen Fähigkeiten sowie des Wissenstransfers führt.

Erfassung des cognitive loads bei Praxisanleitenden im Kontext des Einsatzes von xMOOCs. Ein quasi-experimenteller Forschungsansatz mittels Eye-Tracking

Katharina Kettler, Ulrike Weyland

Universität Münster, Deutschland

Fort- und Weiterbildungen stellen Unternehmen vor die Herausforderung, das Personal hierfür freizustellen und dennoch den laufenden Betrieb aufrecht zu erhalten. MOOCs werden als eine Möglichkeit für die flexible Gestaltung von Fort- und Weiterbildungen diskutiert (Ebner, 2019). Aus den ursprünglichen Ansätzen entwickelten sich diverse MOOC-Varianten, die eine flexible Nutzung ermöglichen. xMOOCs weisen hohe Bezüge zu klassischen Web-Based-Trainings auf, während sich cMOOCs durch konnektivistische und konstruktivistische Elemente hervorheben (Ebner et al., 2019). xMOOCs gehen mit einer besonders hohen Flexibilität einher, da sie vollständig orts- und zeitflexibel genutzt werden können. Jedoch wird häufig die didaktische Qualität von xMOOCs sowie die hohe Dropout-Rate kritisiert (Eickhoff, 2023; Ebner et al., 2019). Daher gibt es erste Untersuchungen, die Blended-Learning-Ansätze und MOOCs vereinen, um der Kritik der unzureichenden didaktischen Qualität und einer damit einhergehenden fehlenden Vernetzung und Austauschmöglichkeit der Teilnehmenden entgegenzuwirken (Ebner, Schön & Käfmüller, 2015).

Eine häufig verwendete Möglichkeit, um xMOOCs zu evaluieren, ist eine Befragung der Teilnehmenden. Auch gibt es Handlungsempfehlungen von MOOC-Entwickler:innen für die formale Gestaltung (Schön et al., 2022). Darüber hinaus liegen wenige empirische Befunde bzgl. der Bewertung zur (e-)didaktischen Gestaltung von xMOOCs vor (Schön et al., 2022; Deimann & Friedl, 2020). In Anlehnung an den Einsatz von Erklärvideos lässt sich aufgrund der bisherigen Befundlage annehmen, dass die Art der Gestaltung Einfluss auf die Motivation und die Verarbeitungstiefe nimmt (Persike, 2020). In diesem Zusammenhang ist auch eine Hinwendung zur Messung des cognitive loads erkennbar.

Ziel der Cognitive Load Theory ist es zu erklären, wie die Informationsverarbeitung beeinflusst wird. Grundlegend wird angenommen, dass die kognitive Verarbeitung durch das begrenzte menschliche Arbeitsgedächtnis limitiert ist. Durch unnötige Anforderungen an das kognitive System kann es zu einem cognitive overload kommen, was den eigentlichen Lernprozess limitiert (Sweller, van Merriënboer & Paas, 2019). Sevchenko et al. (2021) resümieren, dass es nicht eine umfassende Methode zur Messung des cognitive loads gibt, sondern eine Vielzahl an Methoden mit diversen Vor- und Nachteilen in der Forschung vertreten sind. Jedoch haben sich insbesondere die physiologischen Evaluationsformate, wie z.B. das Eye-Tracking, enorm weiterentwickelt (Sweller, van Merriënboer & Paas, 2019). Die Überprüfung von Fixationszeiten, Sakkaden, Blinzelraten und Pupillenerweiterungen können Hinweise auf einen möglichen cognitive overload liefern (Sevchenko et al., 2023). So weisen Befunde darauf, dass eine Häufung von Blinzelbewegungen vor und nach Phasen der intensiven Informationsverarbeitung auftreten oder dass eine höhere Aufgabenkomplexität mit weniger, aber längeren Fixationen einhergeht (Clifton et al., 2016; Siegle, G., Ichikawa, N. & Steinhauer, S. 2008). Auch wenn Eye-Tracking bereits eine etablierte Methode zur Erfassung des cognitive loads darstellt, sind anatomische und physiologische Parameter immer auch anfällig für weitere Einflüsse (Sevchenko et al., 2021). Daher verwenden viele Eye-Tracking-Studien komplementäre Instrumente zur Selbsteinschätzung (Sevchenko et al., 2023; Yulianandra et al., 2023). Hier bietet sich die Kombination aus Eye-Tracking und einer Selbsteinschätzung der Teilnehmenden an. Klepsch et al. (2017) validierten ein deutschsprachiges Instrument zur Erhebung des cognitive loads, welches neben der Erfassung der Blickdaten eingesetzt werden soll.

In diesem Dissertationsvorhaben, welches in einem BMBF geförderten Projekt zur MOOC-gestützten Fortbildung von Praxisanleitenden in der Pflege eingebunden ist, wird der übergeordneten Fragestellung nachgegangen, inwiefern der cognitive load der Teilnehmenden bei der Nutzung von xMOOCs durch konnektivistische analoge Elemente beeinflusst wird. Die Gewinnung von Eye-Tracking-Daten während einer xMOOC-Schulung können Aufschluss darüber geben, inwiefern sich während der Schulung ein cognitive overload einstellt (Yulianandra et al., 2023). Im Rahmen eines quasi-experimentellen Studiendesigns wird dieser während der xMOOC-Nutzung mittels Eye-Tracking und Selbsteinschätzung erfasst. In einem Drei-Gruppen-Plan wird eruiert, inwiefern sich der cognitive load verändert, wenn Lernende mit c-Elementen konfrontiert werden. EG1 und EG2 erhalten im Vergleich zur KG eine voran- bzw. eine nachgestellte analoge, konnektivistisch orientierte Schulung. Es wird überprüft, inwieweit eine Verbindung von c-Elementen die xMOOC-Nutzung positiv beeinflusst.

Evidenz(en) für den Unterricht? Wahrnehmung, Verarbeitung und Transfer wissenschaftlicher und nicht-wissenschaftlicher Evidenzen bei Lehrkräften der 1. Ausbildungsphase

Fabian Reinwarth, Maximilian Knogler, Meg Farrell, Tina Seidel

Das folgende Studiendesign untersucht, wie Lehrkräfte in der 1. Ausbildungsphase wissenschaftliche Evidenzen verständnisvoll rezipieren sowie bewerten können. Hierbei werden Metaanalysen im Rahmen des "Clearinghouse Unterricht" (TU München) ausgewählt, bewertet, als Plain Language Summaries zusammengefasst und disseminiert. Hierdurch soll ein Beitrag geleistet werden, den Lehrerberuf hin zu einer forschungsbasierten Profession zu entwickeln, indem wissenschaftliche Studien zielgruppengerecht kommuniziert werden.

Feedback und Beziehung in der Schule: Die Rolle von Oxytocin und Cortisol

Jana Duckwitz

Justus Liebig Universität Gießen, Deutschland

Das Ziel der geplanten Untersuchung ist es, die Wirkung verschiedener Arten von Feedback (lösungsprozessbezogenes vs. sozialvergleichendes Feedback) auf psychischer (wahrgenommene Nützlichkeit des Feedbacks, wahrgenommene Beziehungsqualität zw. Lehrkraft und Schülerinnen und Schülern, Selbstwirksamkeitserwartung) und physiologischer (Hormone Oxytocin und Cortisol) Ebene bei Schülerinnen und Schülern der 9. Klasse experimentell zu untersuchen. Ein besonderes Interesse gilt der Frage, inwieweit die Beziehungsqualität zwischen Lehrkräften und Schülerinnen und Schülern (LSB) bzw. Bindungsstile von Lehrkräften und Schülerinnen und Schülern die Wirkung und Wirkungsweise der verschiedenen Feedbackarten beeinflussen. Die Ergebnisse der Bindungsforschung zeigen, dass Kinder je nach ihren Bindungserfahrungen mit primären Bezugspersonen unterschiedliche Bindungsstile entwickeln (Bowlby, 1969), welche auch auf zukünftige Kontexte, wie den schulischen Bereich, übertragen werden und dort soziale Interaktionen prägen können (Julius et al., 2008). Unterschiedliche Bindungserfahrungen können sich somit auf die Gestaltung der LSB auswirken (Langer, 2019). Vor dem Hintergrund des bildungswissenschaftlichen Angebots-Nutzungsmodells können die individuellen Bindungserfahrungen und die Beziehungsqualität zur Lehrkraft als wichtige individuelle Lernvoraussetzungen der Lernenden betrachtet werden, die die Wirkung von Feedback beeinflussen können (Vieluf et al., 2020). Lösungsprozessorientiertes Feedback, welches auf einem Stärken-Schwächen-Profil mit Strategien zur Arbeit an den Schwächen beruht, erwies sich bereits unter bestimmten individuellen Voraussetzungen und Verarbeitungsmodalitäten, als lern- und motivationsförderlicher gegenüber sozialvergleichendem Feedback, welches lediglich die Information einer Note mit Notenspiegel einer Vergleichsgruppe beinhaltet (Rakoczy et al., 2013). Im Rahmen der geplanten Studie sollen Bindungsstile und die Beziehungsqualität der LSB als weitere potenziell wichtige Einflussfaktoren untersucht werden. Auf physiologischer Ebene spielen die Hormone Oxytocin und Cortisol eine zentrale Rolle bei der Ausprägung von Bindungsstilen, in sozialen Interaktionen und insbesondere bei der Verarbeitung von sozialem Feedback (Cameron et al., 2017; Wang et al., 2022). Eine dynamische Wechselwirkung zwischen diesen Hormonen kann soziale Bindungen fördern oder stressbedingte Reaktionen modulieren (Takayanagi & Onaka, 2021), was wiederum die Wahrnehmung und Reaktion auf Feedback beeinflussen kann. Die Ergebnisse vorherrschender Bindungs- und psychobiologischer Forschung verdeutlichen, dass frühe Bindungserfahrungen die sozial-emotionale Entwicklung eines Kindes nachhaltig prägen, was sich auf physiologischer Ebene des Oxytocin-Systems widerspiegelt (Feldman, 2015). Untersuchungen dieser physiologischen Reaktionen im Forschungsfeld Schule, und insbesondere im Rahmen einer Feedbacksituation stehen noch aus. Das Forschungsvorhaben zielt darauf ab zu untersuchen, wie zwei verschiedene Arten von Feedback im schulischen Kontext die subjektive Wahrnehmung sowie die Oxytocin- und Cortisol-Reaktionen bei Schülerinnen und Schülern beeinflussen und inwiefern diese Wirkung von Bindungsstilen und der Beziehungsqualität der LSB beeinflusst werden. Die Hauptfragestellungen des geplanten Forschungsprojekts lauten:

1. Wie wirken sich zwei verschiedene Arten von Feedback auf die subjektive Wahrnehmung, die Oxytocin- und die Cortisol-Reaktionen bei Schülerinnen und Schülern aus?
2. Hängt die Wirkung der beiden Feedbackarten von Beziehungsaspekten bzw. Bindungsstilen von Lehrkräften und Schülerinnen und Schülern ab?
3. Wie hängen Oxytocin- und Cortisol-Reaktionen während der Feedbacksituation zusammen?

Für die Untersuchung der Fragestellungen wird mit einem Stichprobenumfang von ca. 140 Schülerinnen und Schülern sowie deren Mathematiklehrkräfte geplant.

Das Forschungsvorhaben strebt an, durch die Verknüpfung von Erkenntnissen aus der Unterrichtsqualitätsforschung, der empirischen Bildungsforschung und der psychobiologischen Forschung ein tieferes Verständnis der zugrundeliegenden Mechanismen zu schaffen, um fundierte Gestaltungsmöglichkeiten für qualitativ hochwertiges und individuelles Feedback zu entwickeln.

Forschungsbasierte Entwicklung und Evaluation eines Freizeitbildungsangebotes unter der Einbindung von Lernendenvorstellungen

Eileen Reckmann, Katrin Temmen

Universität Paderborn, Deutschland

Auf dem Poster wird mein Promotionsvorhaben vorgestellt. Die Zusammenfassung sowie das Poster folgen im Dezember.

Information literacy in secondary students

Deirdre Fels¹, Franziska Baier-Mosch², Carolin Hahnel¹

¹Ruhr-Universität Bochum, Germany; ²DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation, Deutschland

Secondary and tertiary students often do not possess sufficient Information Literacy (IL) skills (Argelagós et al., 2022; Eickelmann et al., 2024; Fraillon, 2024; McGrew et al., 2019). They have difficulties searching for information on the internet and evaluating the trustworthiness of the encountered websites. Simultaneously, access to the internet plays an important role in their lives: around 99% of secondary students in Germany have access to their own smartphone or a family laptop, and up to 83% of secondary students use the internet multiple times a week or daily during school lessons (Medienpädagogischer Forschungsverbund Südwest, 2024). Accordingly, there seems to be a discrepancy between the need to use IL in everyday life and the actual skill levels of students. This discrepancy creates the need to foster IL, which is broadly defined as a "set of competences needed to *find, evaluate and use* information" (Nierenberg & Dahl, 2023, p. 58), in the online context.

Fostering IL requires interventions that are effective. To develop effective interventions, research syntheses can be helpful in identifying particularly effective intervention characteristics. However, currently, there are no research syntheses focusing on general IL interventions in the online context. Previous literature reviews in relation to IL interventions have investigated the effectiveness of only one specific method, such as library instruction (Grabowsky & Weisbrod, 2020), or have focused on differences between face-to-face, online, and hybrid formats (e.g., Weightman et al., 2017). Additionally, they have focused on specific aspects of IL, such as dealing with misinformation (Kiili et al., 2024) and media literacy (Lu et al., 2024), since IL is often defined in various ways, with different definitions focusing on different aspects of the construct (Mbandje et al., 2023; Nierenberg et al., 2021; Nierenberg & Dahl, 2023). Thus, a general overview of IL interventions in the online context does not exist yet.

Consequently, to examine the availability and effectiveness of IL interventions in the online context, we plan to conduct a systematic review that focuses on the availability and effectiveness of different IL interventions targeting high school and university students, using the broad IL definition of Nierenberg and Dahl (2023). Since different studies used different methods to evaluate the intervention effectiveness (e.g., no control group (e.g., Allison et al., 2016) or no follow-up assessment (e.g., Şendurur & Yildirim, 2019)), we decided to also include the methods used to assess the intervention effectiveness. Overall, we have formulated the following research questions:

RQ 1: What interventions are available to promote information literacy in secondary school and university students?

- RQ 1.1: Which aspects of information literacy are covered by the interventions?
- RQ 1.2: Who are the target groups of available interventions?
- RQ 1.3: Have interventions been evaluated? In which way have they been evaluated?

RQ 2: How effective are interventions in promoting information literacy?

- RQ 2.1: Which instructional methods are particularly effective?
- RQ 2.2: Under what conditions are IL interventions particularly effective? As moderating contextual factors, we will investigate the IL aspect covered and the age of the participants.

To find relevant literature, the search will be conducted in several databases (PsyInfo, PsyArticles, Psynindex, ERIC, FIS Bildung, peDOCS, Web of Science). To be included, literature will have to meet the following inclusion and exclusion criteria: studies must be in German or English, should be primary studies, need to provide information about an IL intervention, should target secondary and/or tertiary students, and should have been published in or after 2007. The year 2007 was chosen due to the launch of the first iPhone (Grossman, 2007) which marks the start of the availability of the smartphone for the general public (Sarwar & Soomro, 2013).

Investigating AI-Driven Conversation-Based Assessment

Lauritz Schewior^{1,2}, Marlit Annalena Lindner^{1,2}

¹Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN Kiel), Deutschland; ²Europa-Universität Flensburg (EUF), Deutschland

Theoretical Background

Formative assessment refers to informal testing of students' understanding that supports their learning goals through timely feedback and adaptation of instructional strategies (Black & Wiliam, 1998). Research shows that formative assessment has the potential to enhance both teachers' instruction (Furtak et al., 2016) and students' achievement (Faber et al., 2017). In addition, the so-called testing effect highlights that actively recalling information through regular testing leads to better long-term retention and understanding compared to simply reviewing or restudying material (Rowland, 2014).

However, students often show low engagement during low-stakes assessments, a category that includes formative assessments. This underscores the need for innovative approaches to motivate students to participate in testing activities (Alamri et al., 2021). To increase engagement, assessments should integrate seamlessly into students' learning processes, sustain their interest (Hidi & Renninger, 2006), evoke positive emotional reactions, and foster motivation.

A New Approach: Conversation-Based (Formative) Assessment

Conversation-Based Assessment (CBA) is a promising approach to enhance motivation by leveraging natural, interactive conversations to engage students more deeply in the assessment process (Yildirim-Erbasli & Bulut, 2023). Engagement is expected to be enhanced through cognitively stimulating conversations and parasocial interactions that create a more immersive and personalized testing environment. The assumed benefits of CBA can be understood through two dimensions of high-quality teaching (Praetorius et al., 2018): cognitive activation and social-emotional support. However, experimental research on CBA remains limited (Yildirim-Erbasli & Bulut, 2023), mainly because it has historically required considerable effort to build sufficiently good CBA tools.

Recent advancements in artificial intelligence, particularly in Large Language Models (LLMs; Vaswani et al., 2017), open new possibilities for the development of innovative CBA tools. We are currently developing a domain-flexible LLM-driven CBA tool to provide a solid foundation for robust experimental research on CBA. The LLM-driven conversational agent will offer students questions and immediate, individualized feedback within a guided, para-social chat-based dialogue, while we retain sufficient control to facilitate controlled experimental studies.

A Pilotstudy with the First Prototype

A preliminary pilot study with $N = 25$ students revealed promising results. Participants provided positive feedback on the tool's usability, reporting enjoyment and motivation while using it. They also found the feedback from the agent helpful and expressed interest in using the tool for learning purposes. Individual comments highlighted appreciation for "the flow" of the assessment and noted that the agent encouraged them to reflect further on their answers. Participants also offered constructive suggestions for improvement. For instance, a student reported partially answering a question correctly but receiving feedback indicating it was entirely incorrect. In response to this feedback, echoed by multiple participants, we introduced a feature allowing the agent to deliver more nuanced responses to partially correct answers.

Future Studies

In upcoming studies, we aim to gather more feedback on the prototype using a larger dataset to further refine and improve the tool. Additionally, we aim to investigate the effects of CBA on performance compared to traditional formative assessment methods, for example, by conducting post-tests on the material studied. A key focus will also be exploring individual learner

factors, such as motivation or prior knowledge, that may influence the effectiveness of CBAs. Advances in artificial intelligence will also make it possible to conduct conversational assessments via audio in the near future. This raises an important research question: does the modality of interaction - text-based versus audio-based - affect the tool's effectiveness? Addressing these questions will provide further insights into the potential of CBA tools and their application in educational settings.

Kognitive Aktivierung im Kontext der Differenzialrechnung (SEK II)

Rauno Friedlin

Universität Duisburg-Essen, Deutschland

Mathematische Kompetenzen sind ein entscheidender Prädiktor für den Studienerfolg in vielen Studiengängen (Müller et al., 2018). Hohe Abbruchquoten, gerade bei Studienanfängern, verdeutlichen den schulischen Handlungsbedarf und lassen die Ursachen in der vorausgehenden **Sekundarstufe II** vermuten. Die **Differenzialrechnung** spielt dort eine tragende Rolle.

Das Projekt **QuaMath** beschäftigt sich in diesem Zusammenhang mit der Verbesserung des Mathematikunterrichts in der Sekundarstufe II. **Kognitive Aktivierung** ist dabei ein zentrales Unterrichtsprinzip und eine Voraussetzung für den erfolgreichen Erwerb mathematischer Kompetenzen (Prediger et al., 2022, Holzapfel et al., 2024). In der Sekundarstufe II ist bisher wenig über kognitive Aktivierung bekannt. Wie dies gelingen kann, ist daher von besonderem Interesse. Forschungen aus der Sekundarstufe I zeigen, dass kognitive Aktivierung spezifisch für die Fächer zu denken ist (Bruder, R., 2018, Praetorius, A.-K., Gräsel, C., 2021), entsprechende Befunde für die Sekundarstufe II jedoch bislang nicht existieren. Auf dem **Syntheseframework** (Praetorius et al., 2018, 2020) aufbauend, wird in Lehrkräftefortbildungen des QuaMath-Moduls für Differenzialrechnung die Forschung mit der Schulpraxis verknüpft und vertieft.

Als **Ziel dieser Arbeit** sollen **fachspezifische Maßnahmen** für die kognitive Aktivierung in der SEK II modelliert und durch einen passenden **Fragebogen** überprüfbar bzw. empirisch „greifbar“ gemacht werden. Anhand verschiedener Facetten des Konstrukts werden sowohl fachspezifische als auch **generische Aspekte** berücksichtigt. Zur Validierung des Testinstruments werden unterschiedliche Ansätze genutzt. Geplant ist eine Kreuzvalidierung mit Trainings- und Teststichprobe. Nach einer Expertenvalidierung und einer Itemselektion, findet der erste Durchlauf statt. Daran anschließend wird durch Videographie des Unterrichts und externe Ratings eine Prüfung auf Konsistenz mit den Antworten in den Fragebögen ermöglicht. **Leistungstests** im Prä-/Post-Design erfassen auf Seiten der **SchülerInnen** den Lernfortschritt. Daneben können Selbsteinschätzungen sowohl auf Schüler- als auch auf Lehrerebene durchgeführt werden. Die Erhebungen erfolgen in verschiedenen **Kohorten**.

Kriterien und Gelingensbedingungen für den Besuch außerschulischer Lernorte - ein kumulatives Promotionsvorhaben

Ricardo Puppe, Eva Blumberg

Universität Paderborn, Deutschland

Mit diesem Poster soll das Promotionsvorhaben von Herrn Ricardo Puppe zum Thema "Kriterien und Gelingensbedingungen für den Besuch außerschulischer Lernorte" vorgestellt werden. Bei dem Promotionsvorhaben handelt es sich um eine kumulative Dissertation. Hierzu sollen in dem noch frühen Stadium des Promotionsvorhabens erste geplante bzw. angedachte Erhebungsdesigns des kumulativen Promotionsvorhabens auf der Nachwuchstagung der GEBF präsentiert werden. Zudem werden zentrale Fragen aufgeworfen, um gewinnbringende Gespräche zu initiieren.

Ausführlichere Infos folgen mit der Einreichung bzw. dem Abstract zum 31.12.

Lehrkräftefortbildungen im Kontext digitaler Transformation und Mediendidaktik an beruflichen Schulen

Francy Alisch, Susanne Miesera, Claudia Nerdel

Technische Universität München, Deutschland

Die digitale Transformation stellt eine gesamtgesellschaftliche Herausforderung dar. Die Dynamik der Digitalisierung konfrontiert Schulen mit Weiterentwicklungsaufgaben im Bereich von Personal- und Unterrichtsentwicklung. Das Kompetenzzentrum KoKon des Kompetenzverbunds lernen:digital hat das Ziel, die digitale Schulentwicklung an allgemeinbildenden und beruflichen Schulen ko-konstruktiv über alle Ebenen der Schule zu fördern. Das dem Kompetenzzentrum KoKon zugeordnete Teilprojekt der Technischen Universität München (TUM) befasst sich mit der Unterrichts- und Personalentwicklung in der beruflichen Fachrichtung Ernährungs- und Hauswirtschaftswissenschaften. Ziel ist die Entwicklung und Bereitstellung umfangreicher Fortbildungsmodule für die Förderung professioneller Kompetenzentwicklung der Lehrkräfte und digitalgestützten Unterricht, insbesondere im Bereich der digitalen Transformation und der digitalen Mediendidaktik. Die Fortbildungsmodule sind ein Schlüsselement zur Verknüpfung dieser beiden Schwerpunkte und basieren auf den Kompetenzbereichen des DigCompEdu (Redecker, 2017). Im TUM-DigiLLab, dem digitalen Lehr-Lern-Labor, der TUM, werden Fortbildungsangebote entwickelt, die in Präsenz, hybrid und in Online-Modulen angeboten werden. Das Teilprojekt soll auf diese Weise zur Sensibilisierung der Lehrkräfte für die Anforderungen des digitalen Wandels im Berufsfeld beitragen und ihre Kompetenzen im lernwirksamen Einsatz digitaler Medien stärken. Durch berufsfeldspezifische und berufsfeld-übergreifende Kooperation soll die Unterrichts- und Personalentwicklung an beruflichen Schulen nachhaltig gefördert werden.

Die Fortbildungsmodule umfassen (fach-)didaktisch aufbereitete Elemente, die auf der Plattform UnterrichtOnline.org in Form von Videovignetten, Dokumenten und Begleitmaterial bereitgestellt werden. Die Module zeigen Unterrichtsplanungsgespräche, Unterrichtsplanungsmaterialien, Unterrichtssequenzen und Lehrplananalysen jeweils mit konkreten Aufgaben. Grundlage dafür sind die Ordnungskategorien für die Berufsbildung in der digitalen Transformation nach Wittmann & Weyland (2020) sowie die Leitfragen einer mediendidaktischen Analyse nach Kerres (2000). Durch die Bereitstellung der Fortbildungselemente auf der Plattform UnterrichtOnline.org eröffnen sich individuelle Nutzungsszenarien zu Wiederholungs- und Übungszwecken für die Lehrkräfte, auch über den Projektzeitraum hinaus.

Die Lehrkräfte sollen durch die Fortbildungsmodule ein tiefes Verständnis von digitaler Transformation und digitaler Mediendidaktik im Berufsfeld erlangen und in ihrer Fähigkeit gestärkt werden, den Unterricht auf curricularer Ebene weiterzuentwickeln und kritisch zu reflektieren. Die Fortbildungsmodule unterstützen die professionelle Entwicklung der Lehrkräfte, indem sie kollaborative, feedbackgestützte Problemlösestrategien fördern.

Das Forschungsdesign des Teilprojekts wird als Mixed-Method-Studie angesetzt. Die Erhebung erfolgt zu drei Messzeitpunkten (Prä-, Post sowie Follow-Up-Erhebung 6 Monate nach der Fortbildung). Als Intervention finden Lehrerfortbildungen in Form einer Präsenzveranstaltung oder als synchrone Online-Veranstaltung und/oder auf der Plattform UnterrichtOnline.org statt. Die

Fortbildung stellt die unabhängige Variable dar. Die Selbsteinschätzung der digitalen Kompetenzen der Lehrkräfte vor und nach der Intervention wird mittels Skalen erhoben. Das Instrument zur Selbstwirksamkeit von Lehrerinnen und Lehrern in Hinblick auf die unterrichtliche Integration digitaler Technologie umfasst zwölf Items mit einer sechsstufigen Likert-Skala (Doll & Meyer, 2021); die Skala Utility Value umfasst vier Items mit einer vierstufigen Likert-Skala (Fütterer et al., 2023) und das adaptierte TPACK Modell im Berufsfeld Ernährung & Hauswirtschaft (Miesera et al., 2021) umfasst acht Items mit einer vierstufigen Likert-Skala. Die qualitative Erhebung erfolgt mittels Leitfaden-Interviews ca. 6 Monate nach der Lehrerfortbildung. Erfasst werden vier Kategorien: Digitale Kompetenzen und Unterrichtsplanung der Lehrkräfte mit digitalen Elementen, Implementation von Fortbildungselementen und Kooperation im Rahmen der Unterrichtsplanung als abhängige Variablen.

Die Pilotierung der Fortbildung erfolgte im Juni 2024. Im Herbst 2024 folgen erste Durchführungen der Fortbildung inklusive Prä- und Posttests. Erste Ergebnisse der quantitativen Forschung werden auf der Tagung vorgestellt und diskutiert.

Politische Bildung und aufsuchende Jugendarbeit

Paulina Miliczek

Pädagogische Hochschule Karlsruhe, Deutschland

Spätestens seit den Landtagswahlen in Brandenburg, Sachsen und Thüringen wird sichtbar, dass die AfD einen großen Zuspruch von Wähler*inne*n erhält (Heinrich, 2024). Dieser wird nicht zuletzt vor allem Jungwähler*inne*n zugeschrieben (Olanigan, 2023; Mujagić, 2024). Gleichzeitig ergab die zuletzt veröffentlichte Mitte-Studie, dass menschen- sowie demokratiefeindliche Tendenzen innerhalb der deutschen Bevölkerung zunehmen (Friedrich-Ebert-Stiftung, 2023). Das Erstarken rechtsextremer Kräfte, demokratiefeindlicher Einstellungen und die Diskursverschiebung ins rechte Spektrum erfordern einen Handlungsbedarf, welcher in Hinblick auf junge Menschen in einem lauter werdenden Ruf nach politischer Bildung mündet (Rauschenbach, 2028, S. 3). In meinem Promotionsprojekt gehe ich davon aus, dass die Etablierung einer politischen Haltung und Meinung sowie die Entwicklung von Kompetenzen, welche als politisch bewertet werden können, als ein lebenslanger Prozess angesehen werden müssen, der von unterschiedlichen Personen sowie Institutionen beeinflusst wird. Neben öffentlichen Diskursen wirken ebenfalls Eltern, Peers, Institutionen wie die Schule sowie außerschulische Kontexte auf politische Subjektpositionierungen ein (Fischer & Partetzke, 2015, S. 5; Eckstein & Noack, 2018, S. 372). Politische Bildung findet jedoch - die verschiedenen Schulformen betrachtend - ungleich statt, weshalb in erster Linie Schüler*innen, die eine als gesellschaftlich höher eingestufte Schulform besuchen, von politischer Bildung profitieren. Demnach werden durch die Art und Weise, wie und in welchem Umfang politische Bildung an Schulen stattfindet, Ungleichheiten reproduziert oder sogar verstärkt (Wohnig, 2021, S. 76-80). Neben der Schule findet politische Bildung im privaten und im außerschulischen Kontext statt. Da sich der private Kontext der Öffentlichkeit und somit der gesellschaftlichen Einflussnahme weitgehend entzieht, widme ich mich in meinem Projekt der politischen Bildung in außerschulischen Feldern. Der Fokus meines Promotionsprojekts liegt auf jungen Menschen, die in gesellschaftlich prekären Verhältnissen aufwachsen. Der Zugang soll hierbei über die aufsuchende Jugendarbeit stattfinden, die mit jungen Menschen im Sozialraum arbeitet. Ich möchte bewusst die politische Haltung sowie Subjektpositionierung junger Menschen erforschen, die nicht freiwillig an politischen Angeboten teilnehmen bzw. sich aktiv daran beteiligen. Hier gilt es für mich, die Komm-Struktur bestehender Angebote zu überwinden und im Rahmen der aufsuchenden Jugendarbeit zu den Jugendlichen zu gehen und mit ihnen zu sprechen. Demnach stellen Jugendliche als Adressat*inn*en der aufsuchenden Jugendarbeit das Hauptaugenmerk des Promotionsprojekts dar. Es soll der übergeordneten Frage nachgegangen werden, inwiefern Momente des Politischen innerhalb der aufsuchenden Jugendarbeit sichtbar werden und welche Bedeutung die Jugendlichen diesen im Hinblick auf ihre eigene politische Subjektpositionierung beimessen.

In meinem Beitrag im Rahmen der Nachwuchstagung möchte ich mein triangulatives Forschungsdesign vorstellen.

Im ersten Schritt möchte ich zunächst ein Informationsgespräch mit einer Führungskraft der aufsuchenden Jugendarbeit führen, um auf diese Weise einen Zugang und einen ersten Überblick über das Tätigkeitsfeld zu erhalten. Im zweiten Schritt wird eine Gruppendiskussion mit den Fachkräften der aufsuchenden bzw. mobilen Jugendarbeit stattfinden, um auf diese Weise zu erfahren, wie politische Bildung in ihre Arbeit miteinfließt, welche Vorerfahrungen und Kenntnisse sie bezüglich politischer Bildung mitbringen und wie sie ihre Adressat*inn*en in Hinblick auf ihre politische Subjektposition einschätzen.

Der dritte Schritt besteht aus einer teilnehmenden Beobachtung, innerhalb welcher ich die Fachkräfte in ihrer alltäglichen Arbeit mit den Jugendlichen begleiten werde. Fokussiert soll hier der Frage nachgegangen werden, welche Momente des Politischen in der Arbeit sichtbar werden und wie sich sowohl die Fachkräfte als auch die Adressat*inn*en innerhalb dieser positionieren. In einem vierten Schritt sollen problemzentrierte Interviews mit den Adressat*inn*en der aufsuchenden Jugendarbeit geführt werden. Hier steht die Herausbildung bzw. (Weiter-)Entwicklung ihrer eigenen politischen Subjektposition im Mittelpunkt sowie die Bedeutung der Jugendarbeit für diese. Abschließend werden die Ergebnisse den Fachkräften zurückgespiegelt, um daraus mögliche Veränderungspotentiale für die aufsuchende Jugendarbeit ableiten zu können.

Preisbildung als systemische Interaktion verstehen mit digitalen Simulationen

Jonathan Heitzler

PH Freiburg, Deutschland

Problemstellung und Zielsetzung

Ein wichtiges Ziel ökonomischer Bildung ist es, die Lernenden zu mündigen Akteurinnen und Akteuren unserer Gesellschaft auszubilden. Dies setzt ein Verständnis darüber voraus, wie der Preismechanismus als eines der zentralen Steuerungsinstrumente unserer Wirtschaft funktioniert und welche Informationen über den Preis transportiert werden. Der Preisbildungsprozess ist daher ein wichtiger Kernpunkt des ökonomischen Curriculums. Dabei ist zu berücksichtigen, dass das Verständnis des Preisbildungsprozesses eine Vielzahl von zum Teil sehr komplexen prozeduralen, deklarativen und konzeptuellen Wissenskomponenten umfasst. Eine besondere Lernhürde ist das Verständnis der Preisbildung als systemischer Prozess. Der Einsatz von Klassenzimmerexperimenten („classroom experiments“) kann Lernende bei der Überwindung der Lernhürde unterstützen, wenn die Ergebnisse des Experiments (in Form von darauf aufbauendem intermediärem Wissen) in einer anschließenden instruktionalen Unterrichtsphase zur Entwicklung des Zielwissens eingesetzt werden.

In der geplanten Studie soll ein mehrphasiges Unterrichtssetting zu Förderung konzeptueller Wissensfacetten des Themas Preisbildung untersucht werden. In der ersten, divergenten Unterrichtsphase soll ein Klassenzimmerexperiment („Apfelmarktexperiment“) durchgeführt werden, in dem Lernende in einer Marktsituation miteinander interagieren. In einer anschließenden konvergenten Unterrichtsphase sollen verschiedene Versionen eines digitalen Tools eingesetzt werden, um Lernprozesse zu initiieren, die konzeptuellen Wandel hin zu einem stabilen Verständnis des Preisbildungsprozesses anregen. Ziel der Studie ist es, zu untersuchen, ob ausgearbeitete Lösungsbeispiele in Verbindung mit Selbsterklärungsprompts und ob eine offene oder geschlossen gestaltete Lernumgebung bei Lernenden kognitive Aktivitäten auslösen, die den Erwerb des Zielwissens fördern.

Methoden

Die Studie wird als Experiment im 2x2-Design mit Pretest und Posttest durchgeführt. Eine Gruppe der Probandinnen und Probanden erhält eine Toolversion mit Selbsterklärungsprompts. Die andere Gruppe erhält eine Toolversion ohne Selbsterklärungsprompts. Versuchspersonen beider Gruppen werden außerdem zufällig einer offenen bzw. geschlossenen Lernumgebung zugewiesen.

Um den Prozess des Wissenserwerbs zu erfassen wird ein zu entwickelnder Leistungstest durchgeführt. Der Test wird zu drei Zeitpunkten – vor dem Unterricht, nach der ersten Unterrichtsphase und nach dem Unterricht – durchgeführt, um die fortschreitende Wissensentwicklung der Lernenden zu analysieren.

Zur Erfassung der Lernprozesse werden die individuellen Lernaktivitäten bei der Nutzung des digitalen Tools über die mündlichen Selbsterklärungen und Logfiles erfasst. Es wird geprüft, ob die Elaborationsqualität der Selbsterklärungen und die Qualität und Dauer der Toolnutzung den Lernerfolg medieren.

Forschungsfragen

1. Fördern ausgearbeitete Lösungsbeispiele in Verbindung mit Selbsterklärungsprompts in einer konvergenten Unterrichtsphase eines mehrphasigen Unterrichtssettings den Erwerb des Verständnisses des Preisbildungsprozesses als systemische Interaktion?

Es wird davon ausgegangen, dass durch ausgearbeitete Lösungsbeispiele in Verbindung mit Selbsterklärungsprompts eine signifikante Lernwirkung erreicht wird. Eine medierende Wirkung der Elaborationsqualität der Selbsterklärungen auf den Lernerfolg wird angenommen.

1. Ist eine stärkere oder weniger starke Strukturierung der Lernumgebung in der konvergenten Unterrichtsphase besonders geeignet zur Erreichung eines konzeptuellen Verständnisses des Preisbildungsmechanismus?

Es wird angenommen, dass eine stärkere Strukturierung der Lernumgebung aufgrund der Komplexität des Zielwissens eine signifikante Lernwirkung hat. Eine moderierende Wirkung des intermediären Wissens wird angenommen.

Die Studie befindet sich derzeit in der Planung. Im Rahmen der GEBF-Nachwuchstagung soll der Schwerpunkt auf das Erhebungsdesign gelegt werden.

Soziotechnische Urteilskompetenz im Informatikunterricht und die diagnostische Kompetenz von Lehrkräften

Nils Dyck

Friedrich-Schiller-Universität Jena, Deutschland

Betrachtet man exemplarisch das Programm dieser Tagung, sticht sofort der Begriff "Künstliche Intelligenz" (KI) ins Auge. Spätestens seit der freien Verfügbarkeit generativer vortrainierter Transformer wie ChatGPT-3 ist der Einsatz von KI in Bildungskontexten ein großes Thema der Forschung geworden. Die Frage nach dem Einsatz von KI in verschiedenen Kontexten stellt nach der Definition von Zeidler und Nichols (2009) ein soziotechnisches Problem dar, denn hier erfordert ein wissenschaftliches Thema von kontroverser Natur moralische Argumentationen, deren ethische Auswirkungen nicht vernachlässigt werden können.

Während die technische Realisierung von KI zweifelsfrei Teil der Informatik ist, sollen auch die gesellschaftlichen Auswirkungen im Rahmen des Inhaltsbereichs *Informatik, Mensch und Gesellschaft* (GI, 2016) eine Rolle spielen. Dies ist nicht nur in Deutschland der Fall, sondern ein „Bewusstsein für soziale, ethische, rechtliche und datenschutzrechtliche Fragen und Auswirkungen von Informatiksystemen“ ist auch Teil des K-12-Informatikunterrichts in Frankreich, Schweden, Russland, Südkorea und Neuseeland sowie im Vereinigten Königreich (Hubwieser et al., 2015, eigene Übersetzung).

Auch in der Fachdidaktik richten sich die Diskurse vermehrt auf soziotechnische Fragestellungen im Schulunterricht (z. B. Morales-Navarro und Kafai, 2024; Rücker, 2023).

Trotzdem ist Rücker (2024) zufolge die unterrichtliche Umsetzung eine noch bestehende Herausforderung. Zwar behandeln verschiedene Ansätze wie *Computer Science in Context* (Knobelsdorf & Schulte, 2007) oder *Critical Computing* (Ko et al., 2020) die Thematik, doch ist die Zieldimension bisher unklar. Deshalb stellt sich die Frage, welches konkrete Wissen, welche Fähigkeiten und welches Verhalten zu den sozialen und ethischen Aspekten von Informatiksystemen Schüler*innen erlernen sollen (Rücker, 2024).

Außerdem ist unklar, ob Informatiklehrkräfte diesbezüglich ausreichend qualifiziert sind, denn Inhalte im Bereich „Informatik und Gesellschaft“ sind nur für sehr wenige angehende Informatiklehrkräfte während ihrer Ausbildung verpflichtend (Pancratz & Grave, 2023). Also stellt sich die Frage, ob Informatiklehrkräfte in der Lage sind, soziotechnische Urteilskompetenz zu diagnostizieren und zu fördern bzw. wie sie in die Lage versetzt werden können.

Daraus ergeben sich die zwei Forschungsfragen: Welche Niveaustufen zu den Teilkompetenzen des Kompetenzstrukturmodells zur soziotechnischen Urteilskompetenz existieren, und wie kann die diagnostische Kompetenz von Lehrkräften in Bezug auf soziotechnische Urteilskompetenz gefördert werden?

Die erste Frage soll im Rahmen des von der DFG geförderten Projekts „Kompetenzen und Urteilsprozesse zu den Auswirkungen informatischer Systeme“ beantwortet werden. Hierfür sollen in einer Unterrichtsstudie anhand von schriftlich zu bearbeitenden Aufgaben und Diskussionen zu soziotechnischen Entscheidungen verschiedene Niveaustufen qualitativ identifiziert werden, um ein Kompetenzstufenmodell zu entwickeln. Die Studie soll Ende März im Rahmen einer sechsstündigen Unterrichtsreihe im Informatikunterricht der 9. und 11. Klasse beginnen. Später soll die Studie in der 7. und 8. Klasse einer Gemeinschaftsschule wiederholt werden, um die Datengrundlage variabler und größer zu gestalten.

Im zweiten Teil soll der Fokus auf Lehrpersonen gelegt werden, um zu untersuchen, inwiefern diese in der Lage sind, soziotechnische Urteilskompetenz zu diagnostizieren und zu fördern. Hierfür wird im Rahmen des Kompetenzmodells nach Leuders et al. (2018) die diagnostische Leistungsfähigkeit von Lehrkräften in einer Interventionsstudie analysiert. Dazu wird ein Fortbildungskonzept erarbeitet, das die Schwerpunkte auf die Vermittlung des Kompetenzstrukturmodells (Rücker, 2024) und des zu entwickelnden Kompetenzstufenmodells legt sowie auf die Diagnose und die Ableitung von Fördermaßnahmen anhand dieser. Hierfür kann das COSIMA-Modell zur Förderung der diagnostischen Kompetenz als Anregung dienen (Heitzmann et al., 2019). Darüber hinaus stellen sich jedoch weitere methodische Fragen. Bisher gibt es kein quantitatives Messinstrument, um die diagnostische Kompetenz zu soziotechnischen Urteilsprozessen von Lehrkräften zu erfassen. Zwar ist es denkbar, mithilfe der gewonnenen Daten aus der Unterrichtsstudie Vignetten zur entsprechenden Diagnose zu erstellen. Diese würden dann jedoch lediglich auf dem qualitativen Kompetenzstufenmodell beruhen. Die genaue Fortbildungsstruktur im Rahmen der Interventionsstudie ist ebenfalls noch zu erarbeiten.

Sprachdiagnostik am Übergang in die Grundschule: VER-ES 2.0 Schulanmeldung - eine multiperspektivische Überprüfung

Wynona Kühn

Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau, Deutschland

Der Übergang von der Kita in die Grundschule stellt eine entscheidende Entwicklungsaufgabe dar, insbesondere für die Ausbildung basaler Sprachkompetenzen. Für Kinder, die keine Kita besuchen, wurde in Rheinland-Pfalz das Verfahren zur Einschätzung der Sprachkompetenzen bei der Schulanmeldung (*VER-ES 2.0 Schulanmeldung*) Viereinhalbjähriger entwickelt, um einen Sprachförderbedarf frühzeitig zu erkennen. Das Verfahren basiert auf aktuellen Erkenntnissen der Sprachentwicklungsforschung und berücksichtigt sowohl das Sprachverstehen als auch die Sprachproduktion (Wildemann, Kammermeyer & Kühn, 2024). Ziel ist eine Einschätzung der Sprachkompetenzen (u.a. Benennen von Gegenständen, Verbbeugung, Verwendung von Präpositionen, Äußerungslänge), um auf dieser Basis eine Empfehlung für eine alltagsintegrierte Sprachförderung innerhalb einer Kindertagesstätte nach dem rheinland-pfälzischem Landescurriculum „Mit Kindern im Gespräch“ (Kammermeyer et al. 2023) auszusprechen.

Die Entwicklung von Sprachkompetenzen ist für einen erfolgreichen Einstieg ins Bildungssystem von zentraler Bedeutung. Dabei zeigt die Forschung, dass sowohl monolinguale Kinder als auch Kinder mit Deutsch als Zweitsprache spezifische Entwicklungsverläufe durchlaufen, die bei der Einschätzung berücksichtigt werden müssen (Skintey, 2020, Ruberg & Rothweiler, 2012). Insbesondere Kinder ohne Kita-Besuch können von einer frühzeitigen alltagsintegrierten Sprachförderung profitieren, die ihnen vielfältige und gezielte Lerngelegenheiten bietet (Wildemann et al., 2024). Mit *VER-ES 2.0 Schulanmeldung* liegt ein Verfahren vor, welches in Kooperation von Kita und Schule und in Orientierung am Kind durchgeführt wird. Da das Verfahren bisher als Entwurfsfassung vorliegt und eine Evaluation noch aussteht, gilt es u.a. Fragen zur Akzeptanz, Wiedergabetreue und Gütekriterien auf Seiten der Expert:innen aus Praxis und Wissenschaft nachzugehen. In der Entwicklung von Verfahren noch wenig Berücksichtigung findet zudem das Wohlbefinden des Kindes in der Prüfsituation (Ramirez, Rahmann, Klug & Kuger, 2023).

Aus den Erfordernissen abgeleitet ergeben sich Fragestellungen:

1. *Wie wird die Durchführung und Beobachtung des Verfahrens von Praktiker:innen und Wissenschaftler:innen bewertet (u.a. Akzeptanz und Wiedergabetreue)?*
2. *Schätzen Expert:innen aus Praxis und Wissenschaft das Verfahren aufgrund ihrer professionellen Expertise als geeignet ein?*
3. *Erfüllt VER-ES 2.0 Schulanmeldung die wissenschaftlichen Gütekriterien der Objektivität, Reliabilität und Validität?*
4. *Wie erleben Kinder die Durchführung des Verfahrens in Bezug auf Stresserleben und Wohlbefinden?*

Die Eignung des Verfahrens wird aus drei Perspektiven – Praxis, Wissenschaft und Kind – untersucht, um eine umfassende Bewertung zu ermöglichen. Dabei soll das Verfahren gleichzeitig als Maßnahme zur Sprachdiagnostik erprobt werden, um eine Grundlage für eine entsprechende Weiterentwicklung und Evaluation zu schaffen – welche als zentral für die Entwicklung eines diagnostischen Verfahrens gilt (Henschel, Stanat, Becker-Mrotzek, Hasselhorn & Roth, 2014).

Perspektive Praxis:

Die Einschätzung durch Praktiker:innen erfolgt über eine Fragebogenerhebung ($N \geq 100$) und (optionale) leitfadengestützte Experteninterviews. Ziel ist es, die Akzeptanz, Wiedergabetreue und den Nutzen des Verfahrens zu erfassen.

Perspektive Wissenschaft:

Eine Expertenbewertung erfolgt durch Wissenschaftler:innen mithilfe einer Videoanalyse. Die Videoaufnahmen zeigen reale Durchführungen des Verfahrens, anhand derer Fachpersonen den Sprachförderbedarf einschätzen und ihre Ergebnisse in leitfadengestützten Interviews reflektieren ($N = 6$). Zusätzlich werden die Ergebnisse aus der Praxis mit wissenschaftlichen Standards verglichen, um Erkenntnisse zur Validität zu gewinnen.

Zur Überprüfung der Gütekriterien ist eine Stichprobe von $N = 200$ Kindern vorgesehen. Ihre Sprachkompetenzen werden von zwei geschulten Personen eingeschätzt und mit einem weiteren Diagnoseverfahren verglichen. Nach vier Wochen erfolgt eine wiederholte Überprüfung, um die Reliabilität zu bewerten. Das dichotome Antwortformat (Ja/Nein) des Beobachtungsbogens stellt sicher, dass die Sprachhandlungen klar und objektiv erfasst werden (Eid & Schmidt, 2014).

Perspektive Kind:

Das Wohlbefinden der Kinder während der Beobachtung wird durch Smileyerhebungen, Elternfragebögen und die Analyse des Stresserlebens mithilfe eines Sensorarmbandes, welches physiologische Reaktionen misst (Sasse & Leuchter, 2021), erfasst. Diese Dimension ist entscheidend, da das emotionale Erleben der Kinder die Ergebnisse maßgeblich beeinflussen kann (Ramirez et al., 2023).

Die Auswertung folgt einem Mixed-Methods-Ansatz: Quantitative Daten werden deskriptiv sowie mittels Korrelations-, Regressions- und Rasch-Analysen sowie Reliabilitätsprüfungen ausgewertet. Qualitative Daten werden mithilfe der qualitativen Inhaltsanalyse nach Mayring analysiert.

The "BesserLesen"-project. Reading promotion in primary schools using educational technology

Manuel Bruns

The reading competencies of German students have been declining during the last 25 years. In addition, the distribution of reading competencies has become more unequal. As a result of these two developments, around a quarter of all students in Germany now have reading competencies that are too low for successful participation in school and society. Reading competencies are closely linked to reading motivation and reading habits. 37 percent of students in Germany read for less than 30 minutes per day outside of school, and the proportion of students with a high reading motivation decreased as well during the last two decades. The "BesserLesen" project asks if an educational technology in the form of an AI-based reading buddy can help address these problems. We conduct an intervention study with pre-, post-, and follow-up assessments of reading competencies and reading motivation in Bavarian primary schools. The classes are divided into a control group and two intervention groups. Both intervention groups will use the "BesserLesen"-app during school lessons. Intervention Group 2 will, in addition, adopt a teaching concept that revolves around the book that is integrated into the app. We hypothesize that the reading competence, reading motivation, and reading self-concept of all children will improve in both intervention groups and surpass those of the children in the control group. In addition, we hypothesize that children in Intervention Group 2, where the app is part of a didactic reading concept, will perform better than children in Intervention Group 1. We also hypothesize that the children's socioeconomic background influences the effects we observe. In order to test this, we will administer parent questionnaires that tell us about the children's home literacy environment.

Umsetzungsqualität von kooperativem Lernen im inklusiven Unterricht – ein kumulatives Promotionsvorhaben

Linda Hinderer

Pädagogische Hochschule Freiburg, Deutschland

Theoretischer Hintergrund

Inklusive Bildung und Fragen zu Teilhabe und Gleichstellung gewinnen zunehmend an Bedeutung. Seit den 1990er Jahren fordert die UNESCO eine „Bildung für Alle“ und die Ausweitung inklusiver Bildungsangebote (Deutsche UNESCO-Kommission, 1991, 1994). Das vierte Ziel der UN-Nachhaltigkeitsziele strebt bis 2030 eine inklusive und chancengerechte Bildung für alle an (Deutsche UNESCO-Kommission, 2016). In Schulen liegt der Fokus zunehmend auf Inklusion. Kooperative Lernformen spielen dabei eine wichtige Rolle, da sie sich besonders in heterogenen Lernsettings als wirksam erwiesen haben (Johnson & Johnson, 2009). Die geplante Dissertation untersucht die Umsetzungsqualität von kooperativem Lernen im inklusiven Unterricht. Dabei arbeiten Schüler:innen in Kleingruppen, wobei kognitive Aktivierung und konstruktive Unterstützung im Vordergrund stehen, um die Prozessqualität des Unterrichts zu fördern (Adl-Amini & Völlinger, 2021; Hattie, 2009; Seidel & Shavelson, 2007).

Kooperatives Lernen basiert auf fünf Elementen: positive gegenseitige Abhängigkeit, individuelle Verantwortlichkeit, unterstützende Interaktion, kooperative Kompetenzen und Reflexion (Adl-Amini & Völlinger, 2021). Lehrkräfte spielen trotz der starken Schüler:innenzentrierung eine wichtige Rolle beim kooperativen Lernen. Sie gestalten die Gruppenzusammensetzung und agieren als Lerncoach sowie Expert:in (Adl-Amini & Völlinger, 2021). Auch die Schüler:innen übernehmen spezifische Rollen, die die Lernwirksamkeit positiv beeinflussen (Adl-Amini & Völlinger, 2021). Konstruktionsprozesse, kreative Lösungswege, die Auswahl relevanter Materialien, Regulationsprozesse und individuelle Aufgabenbearbeitung sind zentrale Einflussfaktoren (Adl-Amini & Völlinger, 2021). Die Basiselemente sind nicht nur kennzeichnend für kooperatives Lernen (Johnson & Johnson, 1989), sondern auch entscheidend für seine Wirksamkeit (Adl-Amini & Völlinger, 2021; Veenman et al., 2002).

In der Praxis zeigt sich, dass diese Elemente oft unzureichend umgesetzt werden, was die positiven Effekte deutlich abschwächt (Supanc et al., 2017; Veenman et al., 2000, 2002). Lehrkräfte nennen Herausforderungen wie hoher Vorbereitungsaufwand, fehlende soziale Kompetenzen und Heterogenität der Klassen als Gründe für eine mangelhafte Umsetzung (Abramczyk & Jurkowski, 2020; Adl-Amini et al., 2023; Gillies & Boyle, 2010; Völlinger et al., 2018), obwohl kooperatives Lernen konzeptionell inklusiv ist (Adl-Amini & Völlinger, 2021).

Fragestellung

Die Dissertation untersucht, die Umsetzungsqualität von kooperativem Lernen im inklusiven Unterricht und geht dabei auf die Validität bisheriger Forschungsergebnisse sowie zusätzliche Einflussfaktoren ein. Ein zentrales Anliegen ist die Identifizierung verschiedener Lehrkräftetypen hinsichtlich der Umsetzungsqualität von kooperativem Lernen im inklusiven Unterricht.

Es werden folgende Fragen gestellt:

- Welche Befunde liegen zur Umsetzungsqualität von kooperativem Lernen im inklusiven Unterricht vor?
- Wie reproduzierbar ist der in empirischen Untersuchungen festgestellte Zusammenhang zwischen der Umsetzungsqualität kooperativen Lernens und den Überzeugungen der Lehrkräfte im inklusiven Kontext? Welche zusätzlichen Faktoren beeinflussen diesen Zusammenhang?
- Welche Faktoren beeinflussen die Umsetzungsqualität von kooperativem Lernen im inklusiven Grundschulunterricht?
- Welche Lehrkräftetypen lassen sich hinsichtlich der Umsetzung von kooperativem Lernen im inklusiven Grundschulunterricht identifizieren?

Methode

Zur Beantwortung der ersten Forschungsfrage werde ich ein systematisches Literaturreview durchführen. Aufbauend auf der Studie von Adl-Amini et al. (2023) möchte ich in einer empirischen Studie die Gestaltung inklusiver Unterrichtssettings sowie die Analyse von Schüler:inneninteraktionen mittels Strukturlegetechniken (Scheele & Groeben, 2010) untersuchen. Eine weitere Studie widme ich dem Einfluss der Überzeugungen von Lehrkräften auf die Umsetzungsqualität. In einer nicht-experimentellen Lehrkräftebefragung im Mixed-Method-Design möchte ich Skalen zu Überzeugungen (Bönke et al., 2024), Herausforderungen (Abramczyk & Jurkowski, 2020), Umsetzungsqualität (Adl-Amini et al., 2023; Völlinger et al., 2018) und Selbstwirksamkeit (Meschede & Hardy, 2020; Völlinger et al., 2018) und die Nutzung von kooperativem Lernen (Abramczyk & Jurkowski, 2020) im inklusiven Unterricht erheben. Die Daten werden mit statistischen Verfahren sowie fallorientierter Inhaltsanalyse ausgewertet (Keller-Schneider, 2020; Kuckartz & Rädiker, 2022).

(Erwartete) Ergebnisse

Das systematische Literaturreview wird den aktuellen Forschungsstand zu kooperativem Lernen im inklusiven Unterricht beleuchten. Die empirischen Studien überprüfen die Validität bisheriger Ergebnisse und identifizieren zusätzliche

Einflussfaktoren. Die Lehrkräftebefragung wird zur Identifizierung unterschiedlicher Lehrkräftetypen beitragen, die das kooperative Lernen im inklusiven Unterricht unterschiedlich umsetzen.

Verbesserung der Anregungsqualität von pädagogischen Fachkräften in Ausbildung: Ein Vergleich von Vollzeit-Ausbildung und berufsbegleitender Ausbildung in der Fachschule für Sozialpädagogik

Sven Uhle, Gisela Kammermeyer

RPTU Kaiserslautern-Landau, Deutschland

Die Qualität der sprachlichen Anregung durch pädagogische Fachkräfte ist entscheidend für die Sprachentwicklung von Kindern im frühkindlichen Bereich. So konnte nachgewiesen werden, dass Fortbildungen, die auf die Interaktionsqualität von Fachkräften ausgelegt sind, nicht nur die Anregungsqualität von Fachkräften langfristig verbessern (Kammermeyer, Leber et al. 2019), sondern auch das Sprachverständnis von Kindern gefördert werden kann (Horner et al. 2024). Jedoch gibt es zugleich wenig empirisch gesichertes Wissen über die Wirkung entsprechender Fortbildungen für pädagogische Fachkräfte in Ausbildung.

Fortbildungen zur alltagsintegrierten Sprachbildung, wie zum Beispiel mit dem Fortbildungskonzept *'Mit Kindern im Gespräch'*, haben das Potenzial, die Interaktionsqualität von Fachkräften, die nach ihrer Berufsausbildung an einer Weiterbildung teilnehmen, signifikant zu steigern (vgl. u.a. Egert et al. 2020; Kammermeyer, Metz et al., 2019; Hamre et al., 2018).

Neben der Bedeutung der Fortbildung für Berufserfahrene ist es ebenso wichtig, dass angehende Fachkräfte bereits während ihrer Ausbildung gezielt darin geschult werden, eine hohe sprachliche Anregungsqualität zu entwickeln. Zwar kann nach Thole & Croos (2006, zitiert nach Mischo, 2017) festgestellt werden, dass Fachkräfte mit zunehmender Berufserfahrung wieder verstärkt zu ihren subjektiven Theorien zurückkehren, was „die Grenzen der bestehenden Form der Professionalisierung [...] aufzeigt“. Darüber hinaus sind Wissensstrukturen nur schwer veränderbar (Fussangel, Rürup & Gräsel 2010). Um dem entgegenzuwirken, ist es notwendig, bereits in der Ausbildung gezielt praxisnahe Kompetenzen zur Verbesserung der Anregungsqualität zu vermitteln, damit diese langfristig erhalten bleiben und nicht allein von späteren Fortbildungen abhängig sind. Bislang konnten viele Studien zeigen, dass sich Weiter- und Fortbildungen positiv auf die Interaktionsqualität auswirken, jedoch gibt es bislang kaum Studien, die aufzeigen, wie sich Konzepte der alltagsintegrierten Sprachbildung auf die Anregungsqualität von päd. Fachkräften in Ausbildung auswirken (vgl. u. a. Fukkink, Jilink et al. 2019; Fukkink, Helmerhorst et al. 2019). Für die Ausbildung von Fachkräften in Deutschland gibt es bislang keine vergleichbar angelegten Studien, die die Interaktionsqualität von Fachkräften während der Ausbildung belegen können.

Weshalb das hier dargelegte Forschungsvorhaben darauf abzielt, die Frage zu beantworten, ob die Teilnahme an einer 60-stündigen Unterrichtsreihe zur alltagsintegrierten Sprachbildung die Anregungsqualität von Fachschülerinnen im Rahmen von dialogischen Bilderbuchbetrachtungen verbessert. Und, ob Unterschiede in der Verbesserung der Anregungsqualität zwischen Fachschüler*innen in der Vollzeit-Ausbildung und in der berufsbegleitenden Ausbildung existieren.

In dieser Untersuchung wird der Effekt einer 60-stündigen Fortbildung zur alltagsintegrierten Sprachbildung mit dem Konzept *'Mit Kindern im Gespräch'* (MiKiG) auf die Anregungsqualität von Fachschüler*innen in der Sozialpädagogik untersucht. Im Fokus steht der Vergleich zwischen Fachschüler*innen, die entweder eine Vollzeit-Ausbildung oder eine berufsbegleitende Ausbildung absolvieren. Die Fachschüler*innen nehmen sowohl vor als auch nach der Intervention eine 20-minütige Videoaufnahme einer dialogischen Bilderbuchbetrachtung mit drei Kindern auf. Diese Videos werden von CLASS-zertifizierten Rater*innen mit dem „Classroom Assessment Scoring System“ (CLASS; Teachstone 2023) ausgewertet, um die Interaktionsqualität zu messen.

Es handelt sich hierbei um eine quasiexperimentelle prä-post-Interventionsstudie mit dem Ziel, die Auswirkungen des MiKiG-Ansatzes auf die Anregungsqualität von Fachschüler*innen zu messen und mögliche Unterschiede zwischen den beiden Ausbildungsformen zu identifizieren. Die Ergebnisse sollen einen Beitrag zur Optimierung von Ausbildungs- und Unterrichtskonzepten an Fachschulen für Sozialpädagogik leisten, um die Qualität der frühkindlichen Bildung weiter zu verbessern.

Vorbilder als Wegweiser? Konzeption einer Interviewstudie zur (Be-)Deutung von "vorbildlichen" Lehrkräften für die Lehrer*innenbildung

Marieke Schaper

Universität Tübingen, Deutschland

Die Frage nach „guten“ Lehrer*innen wird seit Jahrzehnten innerhalb der deutschen (Haag & Streber, 2020; Harder, 2014; Pirner, 2012; Weinert & Helmke, 1996) und internationalen Erziehungswissenschaft (Ida, 2017; Devine, Fahie & McGillicuddy, 2013; Hamachek, 1969) diskutiert. Ein Aspekt der in diesem Diskurs bisher wenig beleuchtet wurde, ist die „Vorbildlichkeit“, der insbesondere durch den Begriff der „exemplarity“ (Korsgaard, 2020; Dupeyron, 2020) bzw. „pedagogical exemplarity“ (Korsgaard, 2019) im anglo-amerikanischen Sprachraum Einzug erhalten hat. Basierend auf Linda Zagzebski's Konzept des „moral exemplarism“ (Zagzebski, 2017, 2013) beschreibt Korsgaard die Dimension der „pedagogical exemplarity“ als „moments when someone serves an exemplary function in educational processes and relations“ (Korsgaard, 2019, S. 280 f.). Diese Vorbildfunktion wird als vielversprechende Handlungsorientierung für die pädagogische Praxis diskutiert, da sie insbesondere für die Ausbildung der pädagogischen Urteilskraft (Arendt, 2003, S. 145 f.) als relevant erachtet wird (Korsgaard, 2020, S. 1364). Obwohl bereits viele theoretische Auseinandersetzungen zu dem Konzept der „exemplarity“ in unterschiedliche Debatten eingeflossen sind, so mangelt es dennoch an empirischen Forschungsarbeiten zu diesem Themenfeld.

Die konzipierte Interviewstudie fokussiert eben diese qualitativ-empirische Forschungslücke zum Aspekt der „exemplarity“ und zielt darauf ab, Lehramtsstudierende zu ihren ehemaligen Lehrkräften in ihrer eigenen Schullaufbahn zu befragen. Dabei sollen konkret die folgenden Forschungsfragen beantwortet werden: 1) Von welchen Erfahrungen berichten Lehramtsstudierende aus ihrer Schulzeit, in denen ehemalige Lehrkräfte eine Vorbildfunktion eingenommen haben? 2) Wie deuten sie diese Erfahrungen im Hinblick auf ihr Verständnis einer „vorbildlichen“ Lehrkraft? 3) Welche Bedeutung schreiben sie diesen Erfahrungen für ihr (berufliches) Leben zu?

Um diese Forschungsfragen zu beantworten, wird das episodische Interview als Interviewform herangezogen, was insbesondere durch die Unterscheidung von semantischem und episodischem Wissen konstituiert ist. Dies passt zu der Forschungsfrage, da einerseits die Erfahrungen von Lehramtsstudierenden mit ehemaligen Lehrkräften erfragt (episodisch) und andererseits auf ihre Deutungen einer vorbildlichen Lehrkraft im Lichte ihres Fachwissens abzielt (semantisch). Insgesamt wurde sich für Einzelinterviews – und nicht bspw. Gruppeninterviews – entschieden, da das zugrundeliegende Thema das Teilen von persönlichen Erfahrungen und Deutungen umfasst, die möglicherweise in Gruppensituationen entweder zurückgehalten oder durch sie beeinflusst und ggf. in ihrer Darstellung verzerrt werden könnten.

Es sollen insgesamt ca. 10 Interviews mit Lehramtsstudierenden geführt werden. Um geeignete Interviewpartner*innen zu finden und auszuwählen, soll das Schneeballsystem sowie die Selbstaktivierung als sekundäres Sampling (durch Aushänge und E-Mailverteiler) genutzt werden.

Um Qualitätsstandards bezüglich der Daten einzuhalten, soll eine Datentriangulation vorgenommen werden, sodass die Interviewpartner*innen sich hinsichtlich des Bundeslands, der Hochschule (Universität und Fachhochschule/Pädagogische Hochschule), der studierten Schulform sowie der studierten Fächer unterscheiden. Eine Methodentriangulation ist bei der Interviewform des episodischen Interviews bereits als "methodeninterne Triangulation" (Flick, 2011, S. 273) mitinbegriffen, da durch diese Interviewform unterschiedliche Wissensformen in den Blick genommen werden. Zusätzlich soll den Befragten mittels der Kommunikativen Validierung die Möglichkeit zur Rückmeldung gewährleistet werden. Letztlich sollen der Forschungsprozess und erste Ideen zur Interpretation regelmäßig in einem Forschungskolloquium diskutiert werden.

Das episodische Interview zählt zu den halb-strukturierten Interviewformen, in denen ein Leitfaden erstellt wird, mit dem jedoch flexibel umgegangen werden kann. Zur Leitfadenerstellung wird das „SPSS-Prinzip“ verwendet (Helfferich, 2011, S. 182). Die Interviews sollen in Präsenz und per Zoom durchgeführt werden und eine Dauer von einer halben Stunde bis Stunde betragen. Zusätzlich zu dem Interview soll ein demografischer Fragebogen herausgegeben werden, der zentrale Eckdaten festhält.

Die geführten Interviews sollen im weiteren Verlauf mit der qualitativen Inhaltsanalyse nach Kuckartz (Kuckartz & Rädiker, 2022) ausgewertet werden. Die qualitative Inhaltsanalyse ist deshalb geeignet, da sie sowohl deduktive wie induktive Zugänge beachtet und somit einerseits das theoretische Fundament zur "exemplarity" als Bezugspunkt miteinbezogen werden kann, jedoch ebenso eine Offenheit für Abweichungen und neuartige Aspekte durch die Befragten besteht. Hinsichtlich der gewählten Interviewform soll in der Auswertung explizit zwischen semantischen und episodischen Sequenzen differenziert werden.

Woher sie kommen, wohin sie gehen: Ein Studiendesign zu Bildungswegen und Karrierezielen von Promovierenden in der Sozialen Arbeit

Eileen Wittenberger

Frankfurt University of Applied Sciences, Deutschland

Immer mehr Sozialarbeiter*innen mit Fachhochschulabschluss zeigen Interesse an einer Promotion und Forschungskarriere (Ehlert et al. 2017). Diejenigen, die eine Promotion anstreben, stehen dann vor einer großen Frage: Wie wollen sie promovieren? Während es für Universitätsabsolvent*innen zumeist klar ist, dass diese an einer Universität promovieren werden, ist dies bei Hochschulen für angewandte Wissenschaften (HAW) komplexer. Sie können in einer Nachbardisziplin, meist in den Erziehungswissenschaften oder der Soziologie, an Universitäten promovieren, mit oder ohne Einbindung einer HAW, oder aber an einem Promotionszentrum oder -kolleg in Sozialer Arbeit. Letztere Option ist eine recht neue Form der Promotion, die immer weiter ausgebaut wird. Obwohl inzwischen die ersten Jahrgänge nach dem neuen Promotionsmodell abgeschlossen haben, existieren kaum Forschungen zu diesen. Es ist nicht erfasst, welche Hintergründe die Promovierenden mitbringen, aus welchen Gründen sie sich für die Promotion entschieden haben und welche berufliche Laufbahn sie damit anstreben. Es stellt sich daher die Frage, ob sich die Promovierenden der Sozialen Arbeit, die an Promotionszentren oder -kollegs promovieren, von anderen Promovierendengruppen unterscheiden. Sozialarbeiter*innen, die an oder kooperativ mit Universitäten promovieren, sind selten explizit erforscht, sie verschwinden zumeist unter den Promovierenden der Nachbardisziplinen. Daher sind auch hier die sozioökonomischen Hintergründe und Karriereaspirationen weitestgehend unbekannt. Promovierende, die an Universitäten promovieren haben und dort promovieren, sind meist gut erforscht. Die Forschungsfragen lauten daher: Unterscheiden sich Promovierende der Sozialen Arbeit je nach Promotionskontext voneinander und unterscheiden sie sich von Promovierenden universitärer Nachbardisziplinen hinsichtlich ihres sozioökonomischen Hintergrundes und ihrer Karriereziele?

In den Blick genommen werden sollen daher drei Gruppen: Sozialarbeiter*innen an Promotionszentren und -kollegs, Sozialarbeiter*innen, die an oder mit Universitäten promovieren und Promovierende der nahen Bezugswissenschaften. Hinsichtlich des sozioökonomischen Hintergrundes soll der Fokus vor allem auf der Bildungsherkunft, dem Alter, Geschlechterunterschieden und Care-Verpflichtungen liegen. Der zweite Schwerpunkt der Forschung liegt bei der Messung der Karriereaspirationen. Ziel ist, einen Überblick über die Zusammensetzung und die Karriereziele der Promovierenden der Sozialen Arbeit generell, aber insbesondere derer die in einer der neuen Promotionsformen promovieren, zu erfassen. Hierfür wird methodisch auf die standardisierte Befragung zurückgegriffen.

Aufgrund der fehlenden Datengrundlage zu HAW-Promovierenden hinsichtlich ihrer Zusammensetzung, Karriereziele und Erfahrungen widmet sich das Vorhaben einer innovativen Fragestellung. Insbesondere die Promotionszentren und -kollegs wurden bisher nicht in den Blick genommen. Somit sind auch die Karriereziele promovierender Sozialarbeiter*innen bisher nicht bekannt. Studien zeigen jedoch, dass in den universitären Bezugswissenschaften der Wunsch nach einer wissenschaftlichen Karriere deutlich höher als in den MINT-Fächern (Briedis et al. 2014, Kerst/Wolter 2024). Hinsichtlich der sozioökonomischen Hintergründe zeigt sich im bisherigen Forschungsstand, dass Frauen gerade in Studienfächern, in denen sie gut vertreten sind (wozu auch die Soziale Arbeit sowie die Erziehungswissenschaften zählen), seltener als Männer in eine Promotion übergehen (Kerst/Wolter 2024, Daniel et al. 2024). Ein Grund hierfür ist die mangelnde Vereinbarkeit von Wissenschaft und Familie, die für Frauen von größerer Relevanz ist als für Männer (Jaksztat et al. 2012, BuWiN 2021). Laut der Längsschnittstudie Nacaps sind 15% der Promovierenden Eltern, in dem Fächerbereich dem die Soziale Arbeit angehört sind es mit 19% etwas mehr (Daniel et al. 2024). Explizite Daten für Sozialarbeitspromovierende gibt es bisher nicht. Hinsichtlich der Bildungsherkunft existieren wenige Studien zu Promovierenden, gerade HAW-Promovierende werden dahingehend selten betrachtet. Jedoch studieren Erstakademiker*innen besonders häufig an HAWs, schließen jedoch sehr selten eine Promotion ab (Stifterverband für die Deutsche Wissenschaft 2014, Flöther 2021, Stifterverband für die Deutsche Wissenschaft 2022).

Im Rahmen meiner Masterthesis wird das Studiendesign sowie das Erhebungsinstrument erarbeitet. Anschließend soll idealerweise in einem Promotionsvorhaben die Erhebung und Auswertung erfolgen.