



H E Ä T
HERAUSFORDERUNG
HETEROGENITÄT

Heterogenität kognitiver Leistungsdispositionen nach dem Übergang in die Sekundarstufe I

Abschlussbericht über die Sprachstandsdiagnostik 2017

Stefanie Dotzel, Karina Karst & Stefan Münzer

Inhaltsverzeichnis

Zentrale Befunde	2
1. Einleitung	3
1.1 Das Projekt „Herausforderung Heterogenität“	4
2. Methode	5
2.1 Design	5
2.2 Stichprobe	6
2.3 Erhebungsinstrument	7
2.3.1 Testauswertung	8
2.3.2 Normierung	10
2.4 Auswertungsmethode	10
3. Ergebnisse	11
3.1 Heterogenität in den fünften und sechsten Klassen	11
3.2 Gesamtleistung	12
3.3 Hörverstehen	14
3.4 Leseverstehen	15
3.5 Satzbau	16
3.6 Grammatik	18
3.7 Wortschatz	19
3.8 Rechtschreibung	20
4. Fazit	22
Literaturverzeichnis:	24

Zentrale Befunde

Im Projekt Herausforderung Heterogenität – kurz: HEAT – wird erfasst, wie sich die Sprachkompetenz im ersten Schuljahr nach dem Übergang in die Sekundarstufe I in heterogenen und homogenen Klassen entwickelt. In diesem Projektbericht wird vor allem die Beschreibung der Kompetenzen der fünften und sechsten Jahrgangsstufe im Fachbereich Deutsch des Jahrgangs 2017/2018 in den Blick genommen. Die Testergebnisse von insgesamt 741 Schülerinnen und Schülern aus insgesamt wurden 29 Klassen zeigen folgende Ergebnisse:

- Die Ausprägungen im Bereich Heterogenität in Klassen zeigen in der fünften Klasse signifikant höhere Werte als in der sechsten Klasse.
- Sowohl in der fünften, als auch in der sechsten Klasse sind äußerst homogene wie auch äußerst heterogene Klassen vorzufinden.
- Die Gesamt- wie Teilleistungen der Schülerinnen und Schüler der sechsten Klasse sind signifikant höher als die der fünften Klasse.
- Ein Vergleich der Jahrgangsstufe 5 und Jahrgangsstufe 6 zeigt, dass die Gesamtleistung im Test in der sechsten Klasse um durchschnittlich fast zehn Testpunkte höher ist.

1. Einleitung

Heterogenität spiegelt aus bildungswissenschaftlicher Sicht divergente Facetten von Klassenzusammensetzungen wider. So umfasst der Begriff sowohl kulturelle, geschlechtsspezifische, motivationale oder interessen geleitete Aspekte, bezieht jedoch auch leistungsbezogene Unterschiede von Schülerinnen und Schülern mit ein (Altrichter & Hauser, 2007). Letzteres stellt vor allem für Lehrkräfte ein zentrales und relevantes Element der Heterogenität dar – nach Höhmann (2009) vor allem, weil sich Lehrkräfte damit während des Unterrichts stetig konfrontiert sehen.

Kognitiv-leistungsbezogene Heterogenität nimmt das Element des Lernens in den Fokus: so bezieht sie sich auf leistungsrelevante Fähigkeiten und Schwächen von Kindern, wobei diesbezüglich vor allem die unterschiedlichen Voraussetzungen und Bedingungen des Lernens im Zentrum stehen (Heinzel, 2008). Um diesen Verschiedenheiten gerecht zu werden, wird dafür plädiert, Differenzierungsmaßnahmen zu implementieren. So greift die äußere Differenzierungsform vor allem hinsichtlich organisatorischer Kriterien (Angebot von Förderunterricht), während die innere Differenzierung die tatsächliche Umsetzung des Unterrichts in den Blick nimmt (unterschiedliche Ziele für leistungsstärkere und –schwächere Schülerinnen und Schüler).

Die Diskussionen erstrecken sich über die Dreigliedrigkeit vs. Gesamtschule bis hin zu einem Muss an individueller Förderung. Die Debatte darum, welche äußere Differenzierungsform die fruchtbarste sei, blieb bislang jedoch weitgehend erfolglos – vor allem auch aus dem Grund, dass für die individuellen Lernfortschritte der Schülerinnen und Schüler die äußerlichen Umstände relativ wenig Einfluss nehmen und es eine qualitativ hochwertige Unterrichtsqualität ist, die in Lernerfolg mündet (Hattie, 2009).

Die Frage, wie Lerngruppen für ein möglichst effektives Lern-Outcome strukturiert sein sollten, wird also kontrovers diskutiert (Scharenberg, 2012). Vor allem nach dem Übergang in die Sekundarstufe I gewinnt der Zusammenhang zwischen Heterogenität und Lernerfolg an Bedeutung, da die frisch geformten Klassen zu diesem Zeitpunkt von den Lehrkräften wieder neu organisiert werden müssen.

Im Land Baden-Württemberg wurde zum Schuljahr 2012/13 die verbindliche Übergangsempfehlung für die Sekundarstufe I abgeschafft, so dass die Entscheidung der zukünftig besuchten Schulform des Kindes allein von den Eltern getroffen wird (KMK, 2015). Seither kann beobachtet werden, dass die Anzahl der Schülerinnen und Schüler, welche das Gymnasium besuchen, im Land stark angestiegen ist (Statistisches Landesamt Baden-Württemberg, 2018). Fest

steht, dass sich im Schuljahr 2016/2017 mehr als ein Drittel der Schülerinnen und Schüler, welche eine Empfehlung zum Besuch der Werkreal-/Hauptschule oder Gesamtschule erhielten, die Realschule oder das Gymnasium besuchten (ebd.). Des Weiteren wurde jedoch auch festgestellt, dass sich fast ein Sechstel der Kinder, welche eine Gymnasialempfehlung erhielten, für eine andere Schulform entschieden (ebd.).

Aus welchen Leistungsbereichen sich Gymnasialklassen seit dem Vollzug der Übergangsreform zusammensetzen, kann aufgrund des Wegfalls der verbindlichen Empfehlung nicht weiter kontrolliert werden. Anzunehmen ist jedoch, dass es zu einer steigenden Durchmischung kognitiv-leistungsbezogener Konditionen der Schülerschaft kommt.

Damit eine solch ansteigende Durchmischung nicht als Gefahr gesehen, sondern als Chance genutzt werden kann, müssen berufstätige wie auch angehende Lehrkräfte in der Diagnose und im Umgang mit Heterogenität unterstützt werden. Der Umgang mit Heterogenität stellt eine der zentralen schulischen Herausforderungen dar (Bohl, Budde, Rieger-Ladich, 2017) – und hier schließt das Leuchtturmprojekt „Herausforderung Heterogenität“ (HEAT) der Universität Mannheim an.

1.1 Das Projekt „Herausforderung Heterogenität“

Das Projekt „Herausforderung Heterogenität“ (HEAT) startete im Januar 2016 und wird im Jahr 2021 enden. Gefördert wird das fünfjährige Projekt vom Wissenschaftsministerium des Landes Baden-Württemberg und ist Teil eines vielseitigen Programms zur „Förderung von Leuchttürmen in der Lehrerbildung“.

Das zentrale Vorhaben des Programms spiegelt sich in einer Theorie-Praxis-Kooperation zur evidenzbasierten und praxisorientierten Professionalisierung von (angehenden) Lehrkräften an Mannheimer Gymnasien wider.

Die Umsetzung basiert auf drei Elementen: Der *Universität*, welche für die Aus- und Fortbildung der Lehramtsstudierenden zuständig ist – der *Wissenschaft*, welche für die Begleitung und Analyse verantwortlich ist – und den *Kooperationsschulen* der Stadt Mannheim. Das Verbindungsstück wird durch die Bereitstellung von *Service-Learning* Seminaren hergestellt.

Bei dem Begriff Service-Learning handelt es sich um eine didaktische Methode, die universitäre Lehre mit konkreten Praxisprojekten verbindet. Umgesetzt wird dies in der Art, dass den Studierenden zuerst theoretische Begrifflichkeiten und Konzepte nahegelegt werden, welche anschließend in einer in das Seminar implementierten Praxisphase erfahren, angewandt und erprobt werden (Reinders, 2010).

Zentrale Fragestellungen des Projekts nehmen die Kopplung von Leistung an Heterogenität in den Blick: So geht es etwa um die Frage, ob es für den Lernerfolg günstige und weniger günstige Heterogenitätskonstellationen in Schulklassen geben kann. Zudem sollen vor dem Hintergrund heterogener Lernumgebungen die professionellen Kompetenzen der Lehrkräfte und deren diagnostischen Kompetenzen untersucht werden.

Die Analyse dieser Forschungsfragen soll in dem zentralen und langfristigen Ziel münden, den professionellen Umgang mit Heterogenität zu fördern. Lehrkräften wie auch Lehramtsstudierende werden im kompetenten Umgang mit heterogenen Schulklassen unterstützt und geschult. Die hieraus resultierenden Ergebnisse sollen wiederum für die Entwicklung der Lehrerbildung dienen.

2. Methode

2.1 Design

Das Design des Projekts „Herausforderung Heterogenität“ weist ein längsschnittliches Format auf: erhoben wird zu Beginn der fünften und zu Beginn der sechsten Klasse. Im Gegensatz zum Querschnittsdesign können mittels eines Längsschnitts auch zusätzlich Verläufe aufgezeigt werden. Ziel dieses Längsschnittsdesigns ist somit nicht nur die Messung von Heterogenitäts- und Leistungszuständen, sondern auch die Untersuchung und Beschreibung deren Entwicklungen.

Als eine solche leistungs- und heterogenitätsspiegelnde Variable wird die Sprachkompetenz in den fünften und sechsten Klassen genutzt. Analysiert werden also die Leistungen im Fachbereich Deutsch. Begründet werden kann dies, dass bereits Spinner (2004) gezeigt hat, dass die sprachliche Kompetenz als Grundlage für alle weiteren schulischen Leistungen gilt. Zudem konnte nachgewiesen werden, dass die Leistung im Bereich Sprache auch weitere Kompetenzen, wie beispielsweise die der Mathematik, vorhersagen kann (Watson et al., 2003). Eine zusätzliche, speziell dieses Projekt betreffende Begründung geht auf den hohen Migrationsanteil Mannheimer Bürger zurück. So konnte gezeigt werden, dass in der Stadt Mannheim besonders viele Menschen leben, die einen Migrationshintergrund haben (Fernández Amman, Kropp, A. & Müller-Lancé, 2015). Die Relevanz der sprachlichen Kompetenz – vor allem im Bereich schulischer Bildung – kann somit als immanent wichtig gesehen werden.

Um das Ziel der Förderung eines kompetenten Umgangs von Lehrkräften wie auch Lehramtsstudierenden mit Heterogenität in Schulklassen zu erreichen, werden die jährlichen Sprachstandserhebungen im Rahmen des Service Learning Seminars durchgeführt. Die Testung wird durch die Lehramtsstudierenden vollzogen sowie von diesen ausgewertet. Während des Prozesses werden die Studentinnen und Studenten von den Projektverantwortlichen betreut und im Umgang mit den objektiven Messverfahren geschult.

2.2 Stichprobe

Der hier vorgestellte Projektbericht bezieht sich auf die Durchführung im Jahr 2017 und stellt grundlegende Ergebnisse der Erhebung im Querschnitt dar. Die Stichprobe im Jahr 2017 setzt sich aus insgesamt 741 Schülerinnen und Schülern zusammen. 386 von diesen befinden sich in der fünften Klasse, 364 in der sechsten. Die Anzahl der teilnehmenden Schulen beträgt vier. Insgesamt wurden 29 Klassen durch die Sprachstandsdiagnostik untersucht; 15 Klassen hiervon

können der fünften und 14 Klassen der sechsten Jahrgangsstufe zugeordnet werden.

Das Geschlechterverhältnis der Gesamtstichprobe beträgt 347 (53,5%) Jungen zu 301 (46,5%) Mädchen (Abbildung 1). Betrachtet man dieses Verhältnis klassenweise, so zeigt sich in der fünften Klasse ein Anteil von 175 Jungs (51,0%) und 168 (49,0%) Mädchen, während es in der sechsten Klasse 166 Jungen (56,1) und 130 (43,9%) Mädchen sind. Dieses Verhältnis widerspricht dem Normaltypus der Geschlechterverteilung in Gymnasien, in denen Mädchen leicht überrepräsentiert vorkommen (Statistisches Bundesamt, 2016).

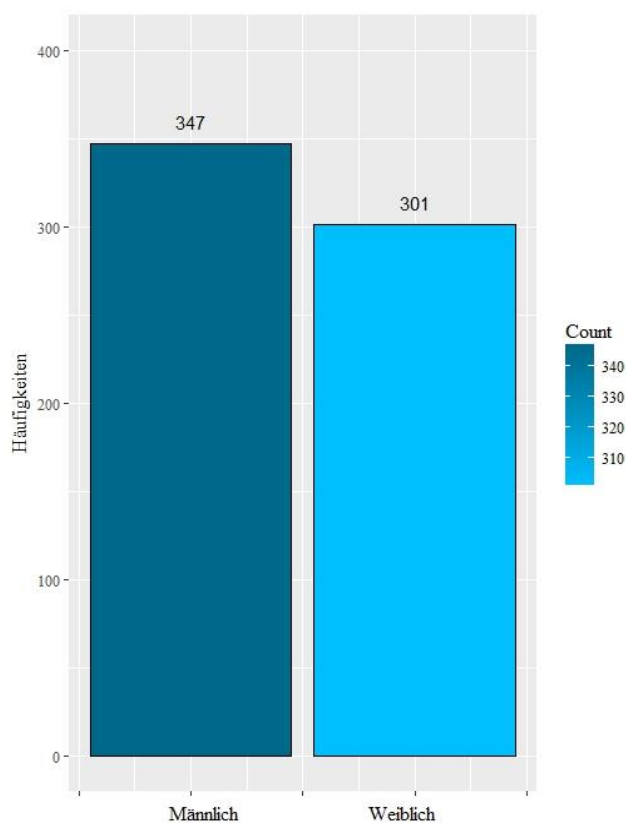


Abbildung 1. Häufigkeiten nach Geschlecht

Fernández Amman et al. (2015) unterstreichen, dass in der Stadt Mannheim besonders viele Menschen mit Migrationshintergrund leben. Dies spiegelt sich auch in den Daten wieder. Bezüglich des Herkunftslandes zeigen sich folgende Verteilungen in den Klassen: In der Gesamtstichprobe haben 300 (46.3%) der Schülerinnen und Schüler einen Migrationshintergrund, 348 (53.7%) hingegen haben keinen (Abbildung 2). In der fünften Klasse beträgt das Verhältnis von Kindern mit und ohne Migrationshintergrund 153 (44.6%) zu 190 (55.4%). In der sechsten Klasse weisen insgesamt 140 (47.3%) Schülerinnen und Schüler einen Migrationshintergrund und 156 (52.7%) Kinder keinen Migrationshintergrund auf.

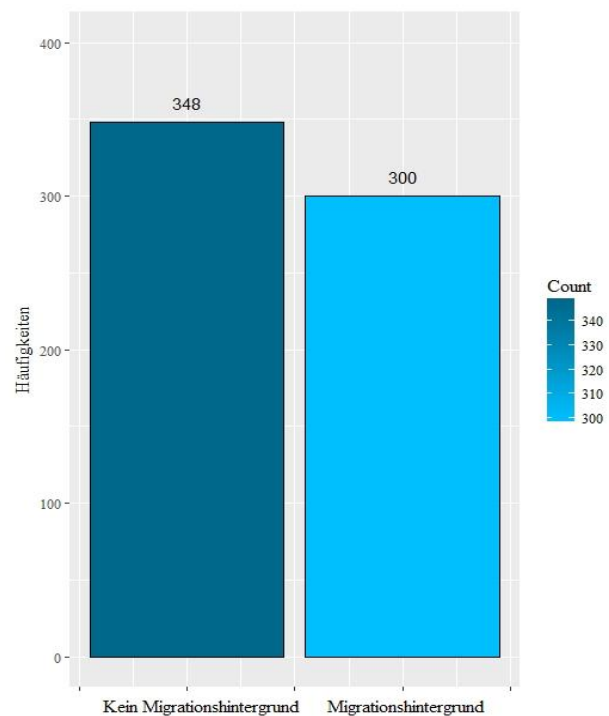


Abbildung 2. Häufigkeiten nach Migrationshintergrund

2.3 Erhebungsinstrument

Die Erhebungsmethode ist quantitativer Art. Analysiert werden die Sprachkompetenzen in den fünften und sechsten Schulklassen. Hierzu wird der Duisburger Sprachstandtest (DST) herangezogen, welcher in fünf Subskalen untergliedert werden kann: Hörverstehen, Leseverstehen, Satzbau, Wortschatz und Grammatik. Zusätzlich bezieht der Test auch die Rechtschreibleistung mit ein. Diese wird bei der Bestimmung der Gesamtpunktzahl jedoch nicht mit in die Berechnung miteinbezogen.

Der Duisburger Sprachstandtest wurde von Blöcker, Endell, Hönig, Kreisherr, Pietsch, Theunissen der Universität Duisburg und von Witzke der Universität Essen entworfen. Der Test benötigt zur Durchführung ungefähr eine Doppelschulstunde. Insgesamt kann ein Ergebnis von 100 Punkten erreicht werden, wobei auf jede der fünf Subskalen 20 Punkte fallen. Wird die Rechtschreibung miteingeschlossen, so können hier zusätzlich 40 Punkte erzielt werden.

2.3.1 Testauswertung

Die Punktevergabe hinsichtlich der einzelnen Subkompetenzen ist folgendermaßen strukturiert: Bei dem Teil zum Hörverstehen werden zehn Fragen zu einer Geschichte gestellt, die die Schülerinnen und Schüler zuvor über eine Audio-Disk oder das Vorlesen durch die Lehrkraft gehört haben. Insgesamt sind für den Teil des Hörverstehens zehn Minuten vorgesehen. Zuerst wird eine standardisierte Vorbemerkung durch die Lehrkraft vorgelesen. Hierauf folgt das Abspielen beziehungsweise das Vorlesen der Geschichte (zwei Minuten) und die anschließende Bearbeitungszeit von zirka acht Minuten. Jede der zehn Fragen enthält vier Antwortmöglichkeiten; hiervon ist jedoch nur eine richtig. Auf jede richtige Antwort gibt es zwei Punkte.

Ähnlich ist das Prinzip bezüglich des Leseverstehens. Auch hier besteht der Test aus zehn Fragen, welche sich auf eine Geschichte beziehen. Die Schülerinnen und Schüler bekommen für die Bearbeitung dieses Testteils fünfzehn Minuten Zeit. Bei der Beantwortung der Fragen darf zum Text zurückgeblättert werden. Von den jeweils vier Antwortmöglichkeiten gibt es nur eine Richtige, welche mit zwei Punkten versehen ist.

Der Testbereich zur Grammatik ist in zwei Teile gegliedert – Teil A und Teil B. Teil A ist als Lückentext mit zehn Fragen aufgebaut. Die zu füllenden Worte bekommen die Schülerinnen und Schüler vorgegeben, sie müssen diese jedoch grammatikalisch an den Satz anpassen. Auf jede richtige grammatikalische Antwort wird ein Punkt vergeben.

Beispiel: Mike hat Geburtstag. Er freut sich über _____ (viel/toll/Geschenk).

Er freut sich über **viele tolle Geschenke**.

Teil B stellt eine One-Choice-Zusammenstellung dar – bestehend aus zehn Fragen mit jeweils fünf Antwortmöglichkeiten. Die Fragestellung ist ebenfalls als eine Art Lückentext formuliert, bei welchem die Schülerinnen und Schüler das fehlende, korrekt grammatikalische Wort ankreuzen müssen. Die Antwortmöglichkeiten stellen also unterschiedliche grammatikalische Möglichkeiten dar.

Beispiel: Ich habe meine Stifte Hause vergessen.

- ☐ in
- ☐ nach
- ☐ von
- ☐ **zu**
- ☐ ins

Für den Grammatikteil wird den Schülerinnen und Schülern eine Bearbeitungszeit von zehn Minuten bereitgestellt. Die Bearbeitungsdauer der beiden Teile A und B ist zeitlich nicht festgelegt.

Der Bereich zum Satzbau ist aus 40 Fragen aufgebaut, von denen jeweils vier Fragen einem Fragenblock zugeordnet werden. Insgesamt gibt es also zehn Fragenblöcke. Die Antwortmöglichkeiten stellen unterschiedliche Arten dar, einen Satz zu formulieren. Die Punktvergabe wird hier folgendermaßen vollzogen: Wird ein richtiger Satz angekreuzt, wird ein Punkt vergeben – ansonsten null Punkte. Wird ein falscher Satz nicht angekreuzt, wird ebenfalls ein Punkt vergeben. Wird er nicht angekreuzt, werden null Punkte vergeben. Zur Beantwortung der Fragen bekommen die Schülerinnen und Schüler 15 Minuten Zeit.

Beispiel:

- Weil er rechtzeitig nach Hause will, der Mann möchte losfahren.
- **Der Mann möchte, weil er nach Hause will, rechtzeitig losfahren.**
- **Der Mann möchte losfahren, weil er rechtzeitig nach Hause will.**
- Möchte der Mann losfahren, weil er rechtzeitig nach Hause will.

Um den Wortschatz zu testen, werden insgesamt 20 One-Choice-Fragen herangezogen. Erkannt werden muss ein Synonym für ein Wort (Fragestellung), zu welchem es vier Antwortmöglichkeiten gibt. Auf jede richtige Antwort gibt es einen Punkt, auf jede falsche keinen. Beim Wortschatz wird den Schülerinnen und Schülern ein Bearbeitungszeitraum von fünf Minuten bereitgestellt.

Beispiel: aushecken

- ausgraben
- entlassen
- ausschneiden
- **planen**

Der Teil zur Rechtschreibung ist ebenfalls als Lückentext mit 20 Lücken aufgebaut. Den Schülerinnen und Schülern wurde zuerst der Satz, dann noch einmal das Wort, welches in die Lücke geschrieben werden sollte, laut vorgelesen. Zudem wird im Testmanual ausdrücklich darauf hingewiesen, eine deutliche Aussprache zu verwenden und verlangsames oder überakzentuier-

tes Sprechen zu vermeiden. Für jedes richtig geschriebene Wort wird ein Punkt vergeben. Sobald ein Fehler in dem Wort vorgefunden wird, wird kein Punkt vergeben; d.h. mehrfache Fehler innerhalb eines Wortes wurden nicht gezählt.

Beispiel: Peter _____ ein Glas Limonade.

Peter **trinkt** ein Glas Limonade.

2.3.2 Normierung

Seit dem Jahr 2010 nahmen jährlich mehr als 20.000 Schülerinnen und Schüler an der Erhebung mit dem Duisburger Sprachstandtest in Nordrheinwestfalen teil. Das Testinstrument eignet sich als Test wie auch als Re-Test, weswegen er in der fünften und in der sechsten Klasse durchgeführt werden kann. Dies geht darauf zurück, dass der Test einen relativ großen Headroom aufweist – also viel Verbesserungspotential aufweist – und somit Deckeneffekte als gering einzuschätzen sind. So wird auch noch in der sechsten Klasse ein realitätsgetreues Bild der Schülerleistungen widerspiegelt.

Im Schuljahr 2012/2013 ergab sich im Land Nordrheinwestfalen ein Mittelwert von 72 Punkten in der fünften Klassenstufe, welcher in den Folgejahren 71, 72 und im Schuljahr 2015/2016 68 Punkte betrug. Anhand der teilnehmenden Gymnasien der letzten vier Jahre ergibt sich somit ein Vergleichs- beziehungsweise Erfahrungswert von 71 Testpunkten. Die Re-Testergebnisse aus dem Schuljahr 2015/2016 in Nordrheinwestfalen zeigen, dass sich die Schülerinnen und Schüler der beteiligten Schulen binnen eines Jahres um durchschnittlich acht Testpunkte verbessern. Der Vergleichswert liegt somit bei 79 Leistungspunkten bei den sechsten Klassen.

2.4 Auswertungsmethode

Die Testhefte wurden mithilfe der Auswertungskriterien des DST's bewertet und anschließend in eine Datenmatrix eingegeben. Die deskriptive Auswertung erfolgt mit der Software R. Die hier dargestellten Ergebnisse gelten als eine Beschreibung der Leistung der teilnehmenden Schulen - eine Verallgemeinerung auf weitere Schulen in Deutschland lässt eine solche Auswertung nicht zu.

Ziel der Auswertung der Sprachstandstests ist es, Aussagen über mittlere und klassenspezifische Leistungsniveaus in der fünften und sechsten Stufe treffen zu können.

Die Heterogenität wird durch die jeweilige Standardabweichung der Gesamtleistung einer Klasse gemessen. Die Standardabweichung definiert die durchschnittliche Abweichung vom Klassenmittelwert; d.h. je größer die Standardabweichung, desto stärker unterscheiden sich die Schülerinnen und Schüler einer Klasse in ihrer Gesamtleistung. Dieser Heterogenitätswert wird dann jeder Schülerin und jedem Schüler einer Klasse zugeteilt. Somit wird bedacht, dass jedes Kind derselben Klasse auch denselben Heterogenitätsbedingungen unterliegt.

3. Ergebnisse

Um die Testleistungen der Schülerinnen und Schüler zu beschreiben, werden sowohl die Gesamtleistung, als auch die einzelnen Subkompetenzen betrachtet. Ebenfalls wird zwischen der fünften und sechsten Jahrgangsstufe differenziert.

3.1 Heterogenität in den fünften und sechsten Klassen

Um einen ersten Eindruck über die Heterogenität zu liefern, soll die Ausprägungen der Heterogenität der teilnehmenden Klassen dargestellt werden. Hier wird die Gesamtleistung der Klassen im Duisburger Sprachstandstest in den Blick genommen.

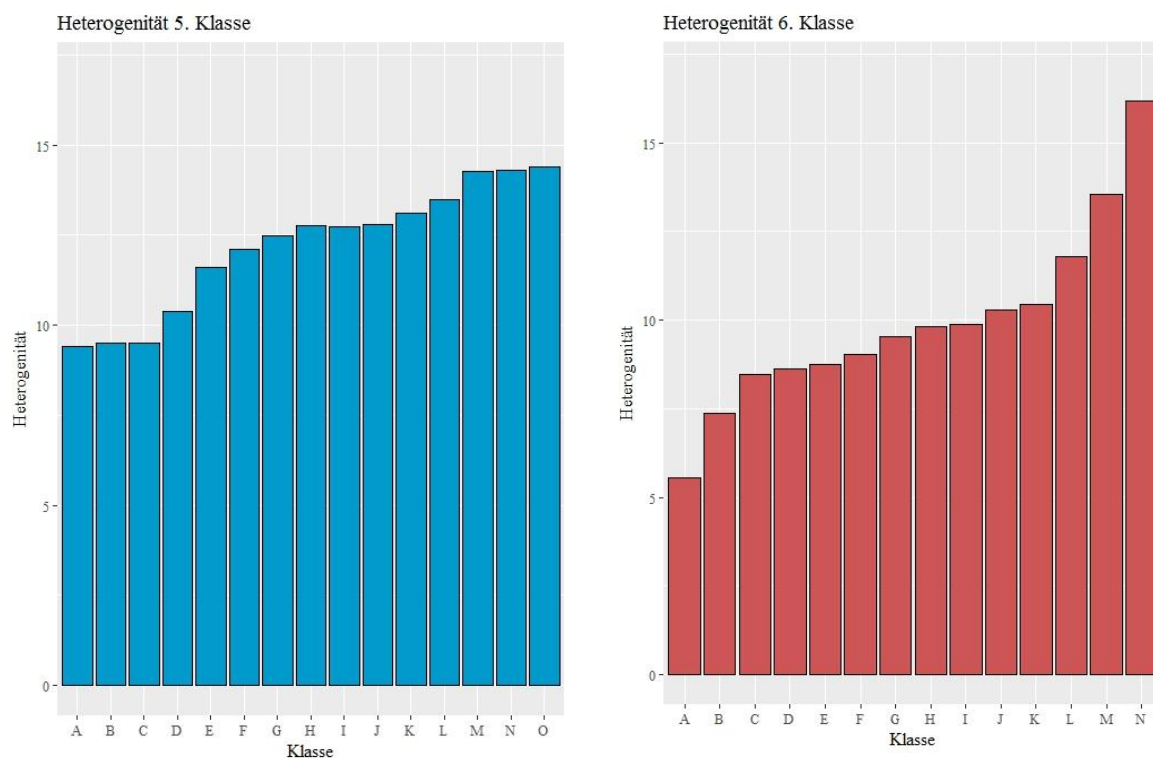


Abbildung 3. Heterogenität der Gesamtleistung in der fünften und sechsten Jahrgangsstufe

Die Gegenüberstellung der Heterogenitäten in Abbildung 3 zeigt deutlich, dass die Streuungen der sechsten Klassen im Durchschnitt sichtbar geringer sind, als die der fünften Klassen. Der Mittelwert der Heterogenität in den fünften Klassen beträgt 12.19 – im Gegensatz hierzu nimmt dieser bei den sechsten Klassen einen durchschnittlichen Wert von 9.95 an.

Ein t-Test zur Testung auf Unterschiede zeigt zudem, dass sich die fünften Klassen in der Heterogenität signifikant von den sechsten Klassen unterscheiden (Tabelle 1).

Tabelle 1: Mittelwerte zur Heterogenität für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Heterogenität	12.19 (1.76)	9.95 (2.60)	2.73**	27

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

3.2 Gesamtleistung

Um die Gesamtleistung tiefer in den Blick zu nehmen, werden die Häufigkeiten der Gesamttestwerte in Abbildung 4 dargestellt.

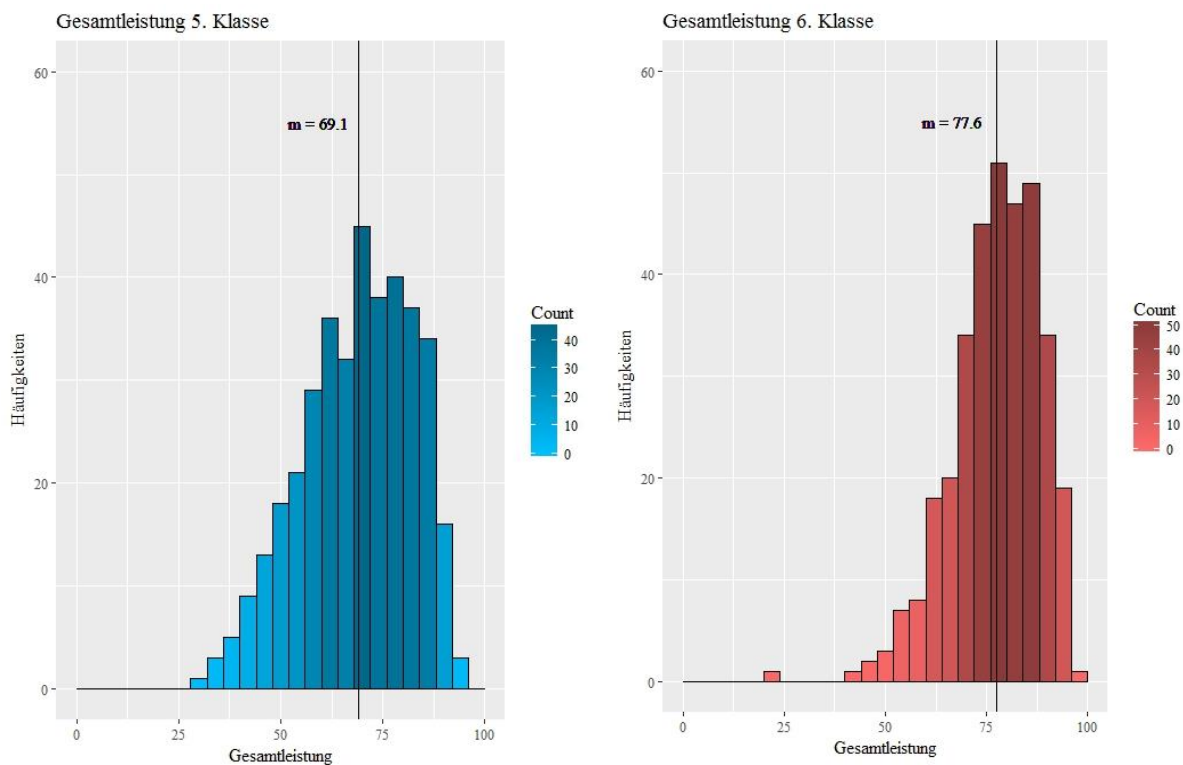


Abbildung 4. Verteilung der Gesamtleistung in der fünften und sechsten Jahrgangsstufe

In der fünften Klasse wurde ein Mittelwert von insgesamt 69.1 Punkten ermittelt. Die Standardabweichung beträgt 13.5, so dass 68% der Schülerinnen und Schüler eine Gesamtleistung erzielt haben, die zwischen 55.6 und 82.6 Punkten liegt. Der maximal erzielte Wert liegt bei 94,5 Testpunkten, hingegen liegt der geringste Wert bei 31,5 Punkten. Die Spannweite zwischen der leistungsstärksten und leistungsschwächsten SchülerIn in der fünften Klasse beträgt somit 63 Testpunkte.

In der sechsten Klasse liegt der Mittelwert bei insgesamt 77.6 Punkten. Zudem ist zu vermerken, dass die Standardabweichung bei einem Wert von 10.9 Punkten liegt. 68% der Leistungswerte streuen also zwischen 66.7 und 88.5 Testpunkten. In der sechsten Klasse sind die durchschnittlichen Abweichungen vom Mittelwert also geringer, als in den fünften Klassen. Die Spannweite zeigt jedoch einen deutlich höheren Wert in den sechsten Klassen. So unterscheidet sich der leistungsstärkste SchülerIn mit 97,5 Punkten vom der leistungsschwächsten SchülerIn mit 23 Punkten mit einer Spannweite von 74,5 Testpunkten.

Die bezüglich der Gesamtleistung heterogenste fünfte Klasse weist eine Standardabweichung von 14.4, die homogenste Klasse eine Standardabweichung von 9.4 Leistungspunkten auf. In der sechsten Jahrgangsstufe beträgt die Standardabweichung der heterogensten Klasse 16.2. Die durchschnittliche Abweichung der homogensten Klasse nimmt einen Wert von 5.6 an.

Die sechsten Klassen unterscheiden sich von den fünften Klassen jedoch nicht nur in der Streuung, sondern auch in der Gesamtpunktzahl. Der Unterschied von fast zehn Leistungspunkten konnte mittels eines t-Tests als signifikant bestätigt werden (Tabelle 2).

Tabelle 2: Mittelwerte der Gesamtleistung für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Gesamtpunktzahl	69.06 (13.49)	77.62 (10.90)	-9.40***	711

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

Im Folgenden werden die einzelnen Subkompetenzen getrennt voneinander betrachtet, um so ein differenzierteres Bild über die Stärken und Schwächen der Schülerinnen und Schüler aufzuzeigen.

3.3 Hörverstehen

Beginnend mit dem Bereich Hörverstehen wird in Abbildung 5 die Verteilung der Testwerte in der fünften und sechsten Jahrgangsstufe abgebildet. In der fünften Klasse ergibt sich ein Mittelwert von 13.0 von insgesamt 20 zu erzielbaren Punkten. Die Standardabweichung nimmt einen Wert von 4.1 Punkten an, so dass 68% der SchülerInnen eine Punktzahl zwischen 8.9 und 17.1 erreicht haben. In der fünften Klasse beträgt die Spannweite 20 Punkte.

Betrachtet man die Leistungen im Hörverstehen in der sechsten Klasse, so lässt sich ein Mittelwert von 14.9 Punkten bestimmen. Die durchschnittlichen Abweichungen vom Mittelwert sind im Gegensatz zur fünften Klasse um eine halbe Standardabweichung geringer und betragen 3.3. Somit streuen 68% der Leistungswerte zwischen 11.6 und 18.2 Punkten. Die Spannweite zwischen der leistungstärksten und der leistungsschwächsten SchülerIn betragen 18 Punkte. Dies ergibt sich aus dem Maximalwert von 20 und dem Minimalwert von 2 Punkten.

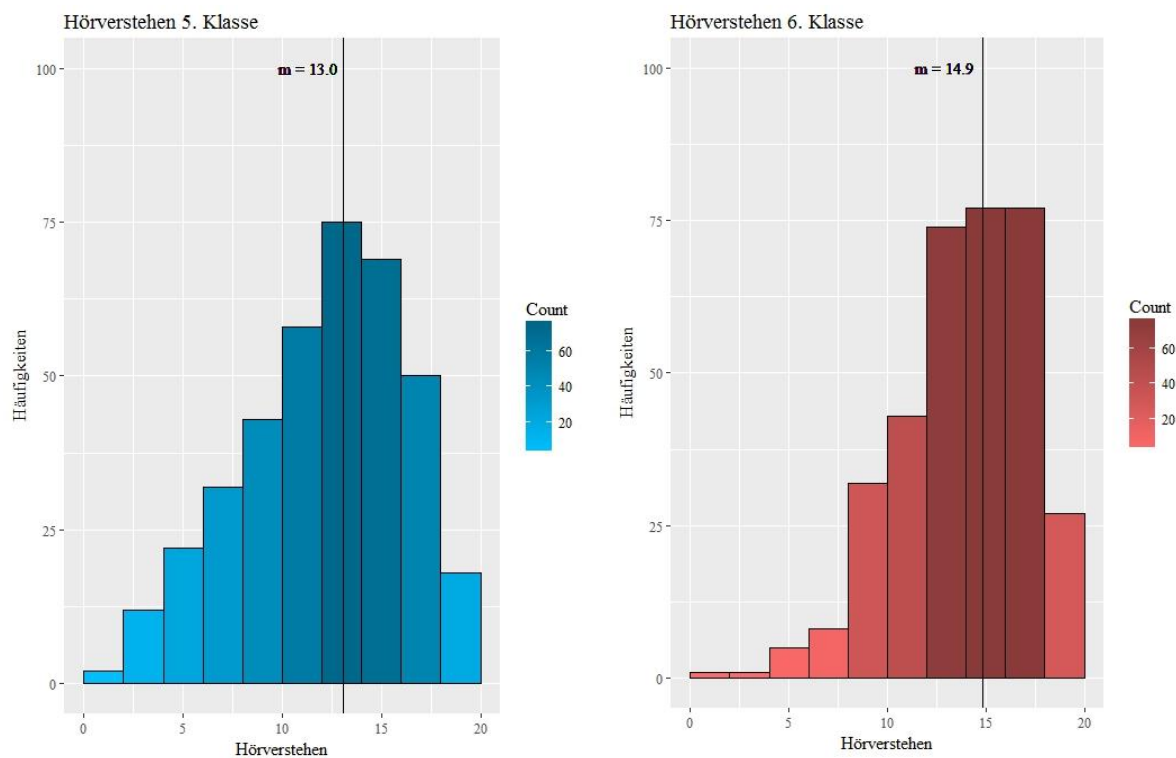


Abbildung 5. Verteilung der Leistung im Bereich Hörverstehen in der fünften und sechsten Jahrgangsstufe

Die in der fünften Jahrgangsstufe heterogenste Klasse weist eine Standardabweichung von 4.4, die heterogenste sechste Klasse von 3.9 Leistungspunkten auf. Im Gegensatz hierzu zeigen die jeweils homogensten Klassen folgende Streuungen: 3.0 für die fünfte und 2.0 für die sechste Klasse.

Die Berechnung eines t-Tests zeigt, dass sich zwischen den fünften und sechsten Klassen ein signifikanter Unterschied in den Testwerten ergibt (Tabelle 3).

Tabelle 3: Mittelwerte im Hörverstehen für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Punktzahl Hörverstehen	13.07 (4.11)	14.87 (3.32)	-6.54***	715

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

3.4 Leseverstehen

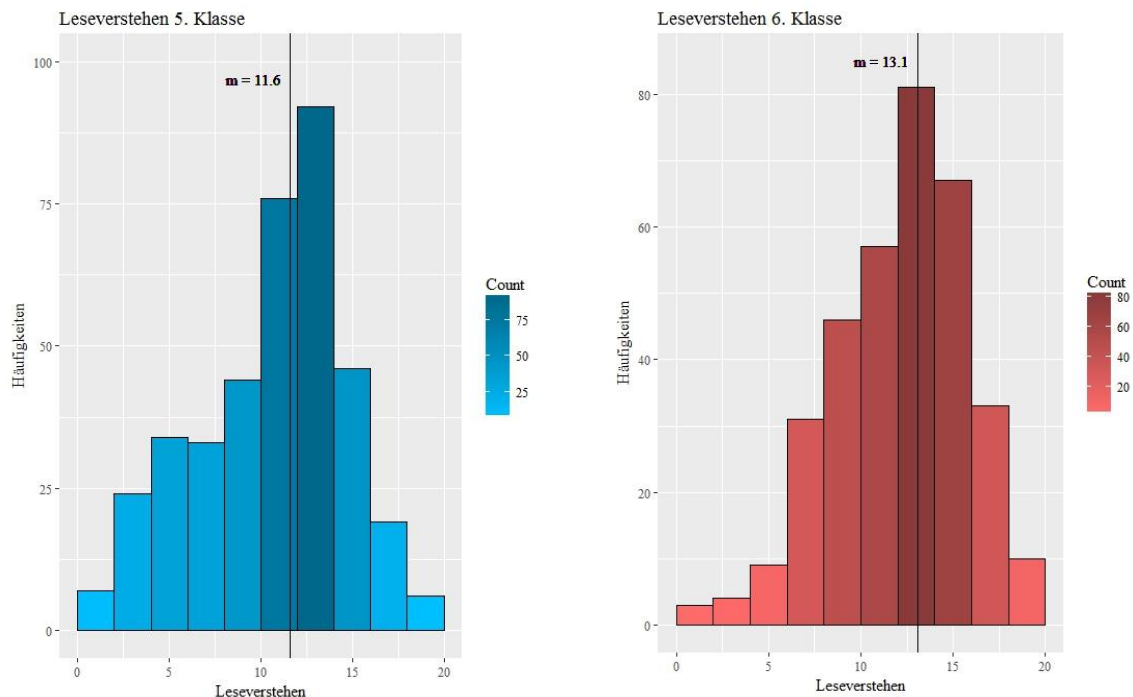


Abbildung 6. Verteilung der Leistung im Bereich Hörverstehen in der fünften und sechsten Jahrgangsstufe

Beim Leseverstehen konnten insgesamt 20 Punkte erreicht werden. In den fünften Klassen zeigt sich ein Mittelwert von 11.6 Punkten. Die Standardabweichung beträgt 4.1 Leistungspunkte. Im Bereich Leseverstehen streuen 68% der Leistungen der SchülerInnen also zwischen 7.5 und 15.7 Testpunkten. In der fünften Klasse beträgt die Spannweite 20 Punkte

In der sechsten Klasse wurde ein Mittelwert von 13.1 Punkten errechnet, welcher mit einer Standardabweichung von 3.6 Leistungspunkten einhergeht. Somit befinden sich 68% der SchülerInnen der sechsten Klasse zwischen einer Leseleistung von 9.5 und 16.7 Punkten. In der sechsten Klasse beträgt die Spannweite 18 Testpunkte, wobei das Maximum bei 20 und das Minimum bei 2 Punkten liegen.

Betrachtet man die Heterogenitäten in den fünften und sechsten Klassen, so zeigen sich folgende Werte: Die fünfte Klasse erstreckt sich von der homogensten Klasse mit einer Standardabweichung von 2.8 bis hin zu heterogensten Klasse mit einer Streuung von 4.9. Die homogenste sechste Klasse weist eine durchschnittliche Abweichung von 2.3 Punkten, die heterogenste Klasse von 4.3 auf.

Bezüglich der Subkompetenz Leseverstehen zeigt sich ein signifikanter Unterschied in den Mittelwerten der fünften und sechsten Klassen (Tabelle 4).

Tabelle 4: Mittelwerte im Leseverstehen für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Punktzahl Leseverstehen	11.59 (4.10)	13.10 (3.58)	-5.32***	720

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

3.5 Satzbau

Im Teil zur Subkompetenz Satzbau konnten ebenfalls 20 Punkte erzielt werden. Die Verteilung der Leistungen ist in Abbildung 6 dargestellt.

In den fünften Klassen zeigt sich ein Mittelwert von 16.3 Leistungspunkten. Die Standardabweichung beträgt 1.9. 68% der Schülerinnen und Schüler streuen somit in ihren Leistungen zwischen 14.4 und 18.2 Testpunkten. Wie auch im Hörverstehen beträgt die Spannweite 14 Punkte. Der Maximalwert beträgt 20, der Minimalwert 6 Punkte.

Im Vergleich dazu zeigt die sechste Klasse einen durchschnittlich erzielten Wert von 17.2 Leistungspunkten im Bereich Satzbau. Die Standardabweichung nimmt einen Wert von 1.4 Punkten an, so dass sich 68% der erreichten Testwerte zwischen 15.8 und 18.6 Punkten befinden. In der

sechsten Klasse zeigt die Spannweite mit 9,5 Punkten einen deutlich geringeren Wert als in den fünften Klassen. Der Maximalwert beträgt 20, der Minimalwert 10,5 Punkte.

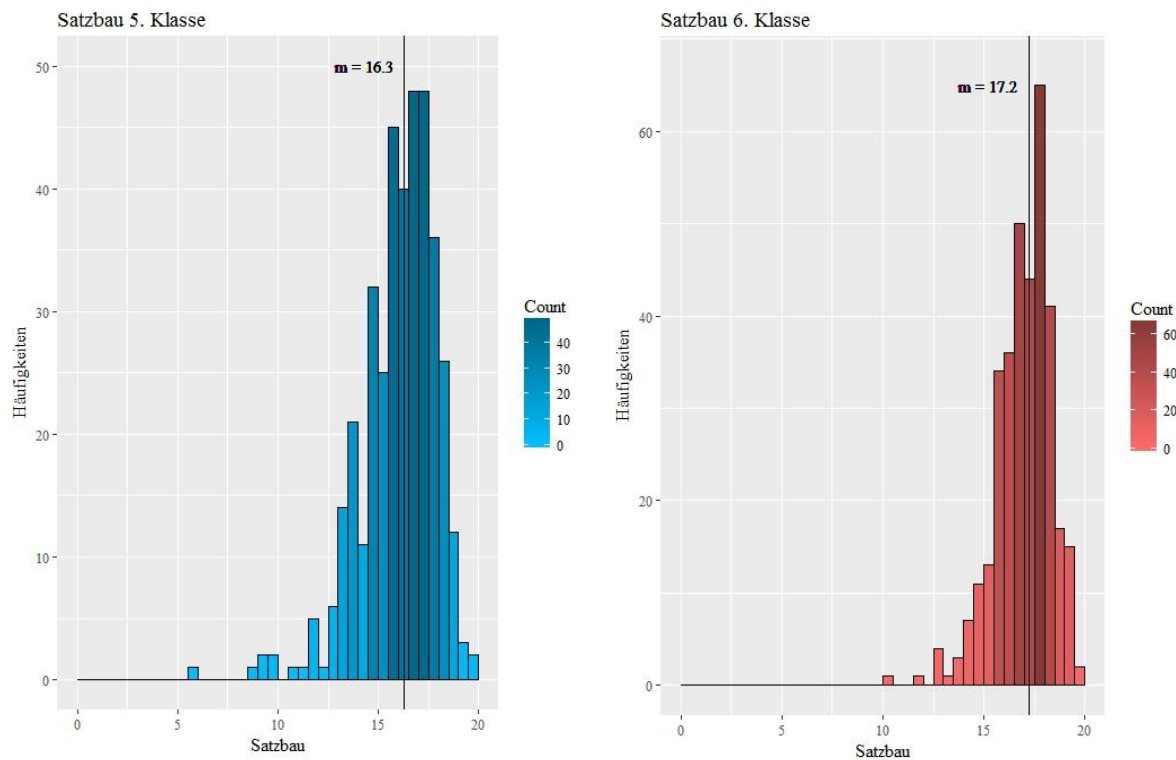


Abbildung 7. Verteilung der Leistung im Bereich Satzbau in der fünften und sechsten Jahrgangsstufe

Hinsichtlich des Satzbaus zeigt sich, dass die heterogenste fünfte Klasse eine durchschnittliche Abweichung von 2.5 Punkten aufweist, welche in der sechsten Klasse 1.9 beträgt. Hingegen lässt sich für die homogenste fünfte Klasse ein Wert von 1.1 und für die sechste Klasse ein Wert von 0.9 berechnen.

Der t-Test bezüglich der Leistungsunterschiede im Bereich Satzbau zeigt sich auch hier als signifikant (Tabelle 5). Die Schülerinnen und Schüler der fünften und sechsten Klasse weichen somit mit einer Mittelwertsdifferenz von 0.96 Punkten signifikant voneinander ab.

Tabelle 5: Mittelwerte im Bereich Satzbau für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Punktzahl Satz- bau	16.26 (1.88)	17.22 (1.39)	-7.84***	699

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

3.6 Grammatik

In der fünften Klasse nimmt der Mittelwert der Subkompetenz Grammatik den Wert 15.2 an. Auch hier konnten insgesamt maximal 20 Leistungspunkte erreicht werden. Die Standardabweichung beträgt 3.7. Somit können 68% der Schülerinnen und Schüler ein Wert zwischen 11.5 und 18.9 zugewiesen werden. Die mit 3 Punkten leistungsschwächste SchülerIn ist um 17 Punkte zur mit 20 Punkten leistungsstärksten SchülerIn entfernt.

In der sechsten Klasse zeigt sich ein in Relation zu den weiteren Subkompetenzen hoher Mittelwert mit insgesamt 17.2 Punkten. Die Standardabweichung beträgt 2.8, so dass 68% der Werte zwischen 14.4 und 20.0 Punkten streuen.

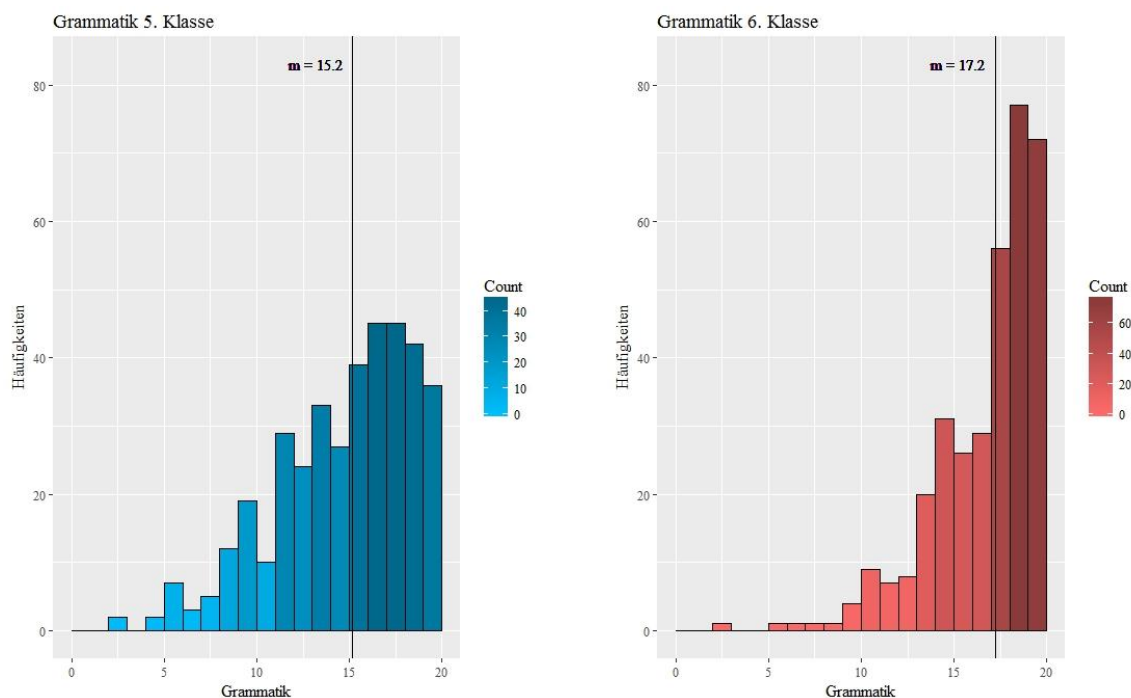


Abbildung 8. Verteilung der Leistung im Bereich Grammatik in der fünften und sechsten Jahrgangsstufe

Abbildung 8 zeigt, dass im Bereich Grammatik keine Normalverteilung vorliegt – in dem Sinne, dass die Randgruppen im Gegensatz zum Durchschnitt gering ausgeprägt sind – sondern die Ausprägung einer hohen Punktzahl das tatsächlich am häufigsten erreichte Ergebnis widerspiegelt. Die Spannweite beträgt 17 Punkte, welcher den Abstand vom maximal höchsten Wert mit 20 Punkten zum geringsten Wert mit 3 Punkten beschreibt.

Die homogensten und heterogensten Klassen in der jeweils fünften und sechsten Jahrgangsstufe sind folgendermaßen ausgeprägt: In der fünften Klasse nimmt die kleinste Standardabweichung einen Wert von 2.6 Leistungspunkten, die größte von 4.1 an. In der sechsten Klasse zeigt sich

ein deutlich geringerer Unterschied zwischen den Klassen mit dem kleinsten Wert von 1.8 und einem größten von 3.8 Testpunkten.

Auch bezüglich der Subkompetenz Grammatik zeigen sich signifikant unterschiedliche Leistungen zwischen der fünften und sechsten Jahrgangsstufe (Tabelle 6).

Tabelle 6: Mittelwerte im Bereich Grammatik für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Punktzahl	15.19	17.22	-8.42***	703
Grammatik	(3.68)	(2.81)		

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

3.7 Wortschatz

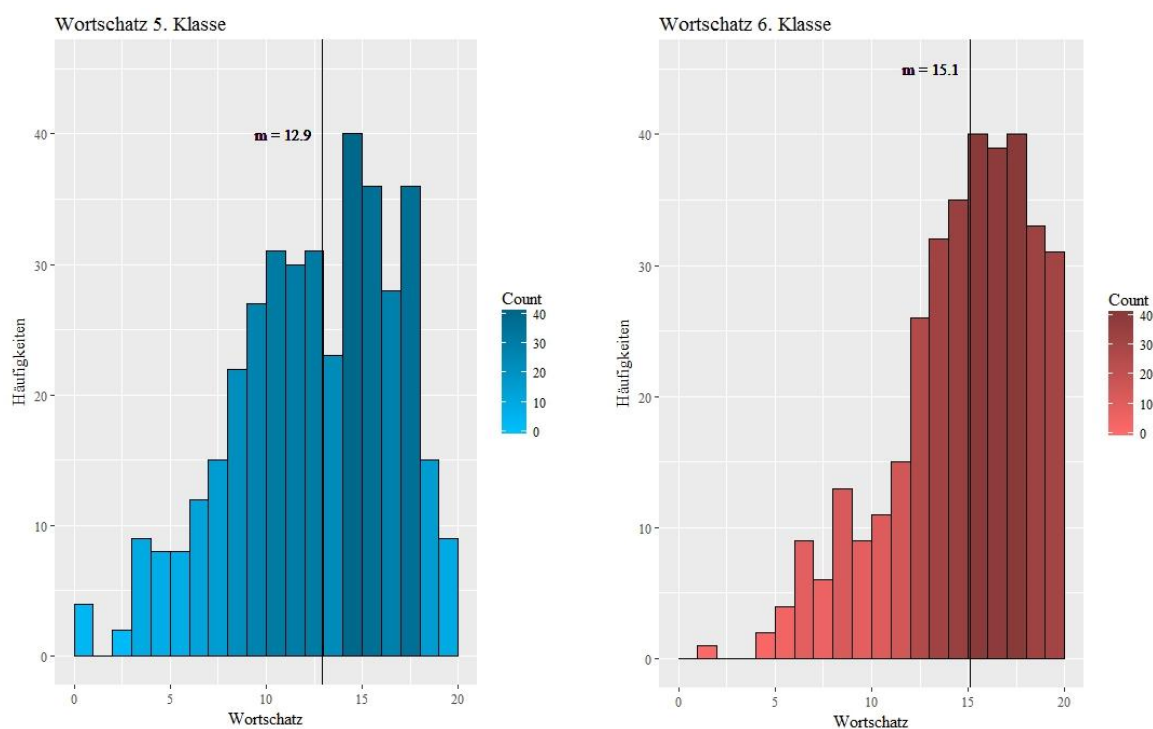


Abbildung 9. Verteilung der Leistung im Bereich Wortschatz in der fünften und sechsten Jahrgangsstufe

Der Mittelwert im Wortschatz, bei welchem insgesamt 20 Punkte erreichbar sind, liegt in der fünften Klasse bei 12.9 Punkten. Da die Standardabweichung 4.2 beträgt, streuen 68% der Testwerte zwischen 8.4 und 17.1 Punkten. Die Spannweite beträgt 20 Testpunkte.

Hinsichtlich der Werte der sechsten Klasse ergibt sich ein durchschnittlicher Wert von 15.1 Punkten im Sprachstandstest. Die Standardabweichung beträgt 3.6 Punkte. Betrachtet man also den Bereich einer Standardabweichung nach oben und nach unten vom Mittelwert, so erstrecken sich die Werte von 11.5 und 18.7 Punkten. Auch in der sechsten Klasse wurden Maximalwerte von 20 Punkten erreicht. Der Minimalwert beträgt hingegen 2 – so ergibt sich eine Spannweite von 18 Punkten im Bereich Wortschatz in den sechsten Klassen.

Hinsichtlich der Heterogenität kann gezeigt werden, dass sich die fünften Klassen von einer Streuung von 3.0 bis hin zu 5.0 erstrecken. Der in den sechsten Klassen geringste Wert beträgt 2.1 und der höchste Wert 5.0.

Der t-Test im Bereich Wortschatz (Tabelle 7) zeigt, dass sich die Leistungswerte der fünften Jahrgangsstufe signifikant von denen der sechsten Jahrgangsstufe unterscheiden.

Tabelle 7: Mittelwerte im Bereich Wortschatz für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Punktzahl Wortschatz	12.94 (4.22)	15.13 (3.63)	-7.55***	729

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

3.8 Rechtschreibung

Im Gegensatz zu den anderen Subskalen können im Bereich Rechtschreibung 20 Punkte erzielt werden. Abbildung 10 veranschaulicht, wie häufig eine Punktzahl im Teil Rechtschreibung des DST's erreicht worden ist.

Der Mittelwert beträgt in den fünften Klassen 14.3 Leistungspunkte. Mit einer Standardabweichung von 3.4 erreichten 68% der Schülerinnen und Schüler einen Testwert zwischen 10.9 und 17.7 Punkten. Die Spannweite in den fünften Klassen beträgt 17 Punkte. Dies ergibt sich aus der Differenz der leistungsschwächsten SchülerIn mit 3 und der leistungstärksten SchülerIn mit 20 Punkten.

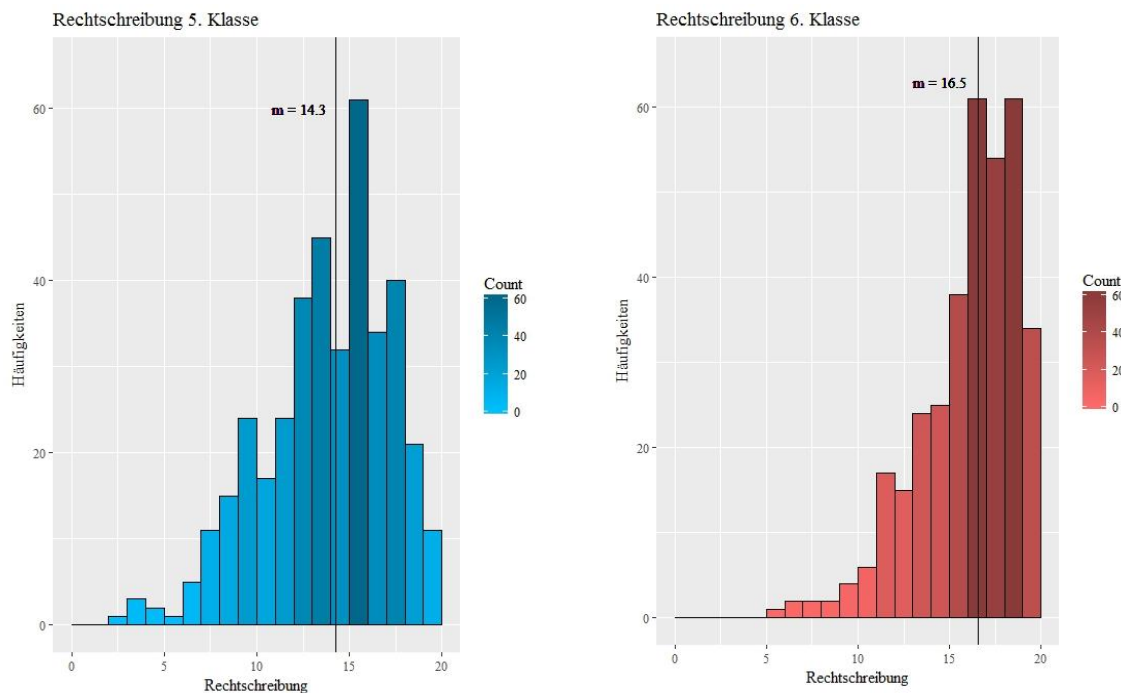


Abbildung 10. Verteilung der Leistung im Bereich Rechtschreibung in der fünften und sechsten Jahrgangsstufe

Im Vergleich dazu beträgt der Mittelwert in den sechsten Klassen 16.5 Punkte. Da die Standardabweichung 2.7 Punkte beträgt, streuen 68% der Leistungswerte zwischen 13.8 und 19.2 Punkten. In der sechsten Klasse zeigt sich in Relation zur fünften Klasse eine deutlich kleinere Spannweite von 14 Testpunkten. Der erzielte Maximalwert beträgt 20 und der Minimalwert 6 Punkte.

Die in der Rechtschreibung heterogensten Klassen zeigen in der fünften Klasse eine Streuung von 4.1 und in der sechsten Klasse von 3.3. Für die homogenste fünfte Klasse lässt sich ein Wert von 2.6 errechnen, welcher in der sechsten Jahrgangsstufe 2.0 beträgt.

Auch bei dieser Subkompetenz zeigt der t-Test einen signifikanten Wert. Die Rechtschreibleistungen der fünften Klasse weichen somit signifikant von denen der sechsten Klasse ab (Tabelle 8).

Tabelle 8: Mittelwerte im Bereich Wortschatz für die fünfte und sechste Klassenstufe

	Klassenstufe		<i>t</i>	<i>df</i>
	5.	6.		
Punktzahl Rechtschreibung	14.29 (3.40)	16.55 (2.73)	-9.95***	720

* = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$. Standardabweichung in Klammern unter den Mittelwerten.

4. Fazit

Der deutsche Bildungsauftrag plädiert für die simultane Umsetzung von Leistungshomogenisierung und höchstmöglicher Qualifizierung in Schulklassen (Helmke, 1988). Dieser Anspruch kann vor allem auf das starke Festhalten am dreigliedrigen Schulsystem zurückgeführt werden. So soll nach Abschluss der Grundschule eine Einteilung der Schülerschaft in unterschiedliche Leistungskategorien erfolgen, um somit eine bestmögliche individuelle Förderung zu gewährleisten.

Um Leistungshomogenisierung beziehungsweise –heterogenisierung zu formen, werden in der Theorie zwei Ansätze unterschieden: Die Egalisierungsstrategie und die Elitestrategie. Kern der Egalisierungsstrategie ist die Förderung von leistungsschwachen Schülerinnen und Schülern und somit die Zusammenstauchung des Leistungsniveaus einer Klasse (Opp & Speck-Hamdan, 2001). Die Elitestrategie grenzt sich von der egalisierenden Methode insofern ab, dass hier vor allem leistungsstarke Schülerinnen und Schüler zusätzlich gefördert werden und die Klasse somit in ihrem Können zusätzlich auseinanderfällt (ebd.).

Auch wenn der Zweck der Dreigliedrigkeit vor allem darin besteht, Leistungshomogenisierung zu vereinfachen, wird nicht erst seit kurzem die Kritik laut, dass das deutsche Schulsystem an einer fiktiven Homogenität in Klassen festhalte (Tillmann, 2004).

Betrachtet man die deskriptiven Daten des Projekts „Herausforderung Heterogenität“, so fällt auf, dass die Heterogenitätswerte der sechsten Klasse deutlich geringer sind, als in der fünften Jahrgangsstufe. Auf Basis dessen kann vermutet werden, dass die einzelnen Klassen über die Zeit in der Leistungsstreuung sinken. Auch der signifikante Unterschied in der Heterogenitätsausprägung zwischen fünften und sechsten Klassen legt die Vermutung nahe, dass Lehrkräfte den Unterricht vorherrschend egalisierend strukturieren. Begründet werden können diese Annahmen jedoch nicht, da der verwendete Datensatz verschiedene fünfte und sechste Klassen beinhaltet und somit keine Entwicklungsverläufe aufgezeigt werden können. Hinsichtlich des Anspruchs einer größtmöglichen Qualifizierung zeigen die Daten, dass die sechsten Klassen in allen Leistungsbereichen signifikant über den fünften Klassen liegen. Jedoch ist auch hier kein Schluss auf eine individuelle Leistungssteigerung möglich, da auch hier keine Verlaufsdaten vorliegen.

Vor diesem Hintergrund stellt sich jedoch die Frage, wie sich die Leistungshomogenisierung binnen eines Jahres derart ausbauen kann. Deswegen untersucht das Projekt HEAT (Herausforderung Heterogenität), wie sich Schülerinnen und Schüler in heterogenen versus homogenen Klassen in der Sprachkompetenz entwickeln und wie Lehrkräfte im Umgang mit Heterogenität

unterstützt werden können. Die in den Daten beobachtete Kombination aus abnehmender Heterogenität über die Zeit und durchschnittlich höherer Leistung in den sechsten Klassen lässt den Apell laut werden, die einzelnen Leistungsgruppen genauer zu untersuchen. So soll zukünftig geklärt werden, für welche Leistungsgruppen eine eher heterogene oder homogene Klassenkonstellation günstig ist.

Literaturverzeichnis:

Altrichter, H. & Hauser, B. (2007). Umgang mit Heterogenität lernen. *Journal für Lehrerinnen und Lehrerbildung*, 1, 4-11.

Fernández Amman, E.M., Kropp, A. & Müller-Lancé, J. (2015). Foreign Language Teaching and Learning in Multilingual Contexts: L3 Teacher Professionalization in Baden-Wuerttemberg. In M.P. Sanfont-Jordá & L. Portolés Falomir (Eds.), *Learning and Using Multiple Languages. Current Findings from Research on Multilingualism* (p. 189 – 216). Cambridge: Cambridge Scholars Publishing.

Hattie, J. (2009): *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Heinzel, F. (2008). Umgang mit Heterogenität in der Grundschule. In J. Ramseger & M. Wagener (Hrsg.), *Chancengleichheit in der Grundschule. Ursachen und Wege aus der Krise* (S. 133 – 138). Wiesbaden: VS Verlag für Sozialwissenschaften.

Helmke, A. (1988). Leistungssteigerung und Ausgleich von Leistungsunterschieden in Schulklassen: Unvereinbare Ziele? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 20 (1), 45-76.

Höhmnn, K. (2009). Heterogenität: Eine begriffliche Klärung. In K. Höhmnn, R. Kopp, H. Schäfers & M. Demmer (Hrsg.), *Lernen über Grenzen. Auf dem Weg zu einer Lernkultur, die vom Individuum ausgeht* (S.27 – 35). Opladen: Barbara Budrich.

KMK (2015). Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (19.02.2015). *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe)*. Zugriff am 08.11.2016 https://www.kmk.org/fileadmin/Daten/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf

Opp, G. & Speck-Hamdan, A. (2001). Heterogenität der Schulanfänger – Herausforderung für die Schule. In Faust-Siehl, G. & Speck-Hamdan, A. (Hrsg.), *Beiträge zur Reform der Grundschule: Schulanfang ohne Umwege. Mehr Flexibilität im Bildungswesen* (Band 111, S. 175-189) Frankfurt a. M.: Grundschulverband – Arbeitskreis Grundschule.

Reinders, H. (2010). Lernprozesse durch Service Learning an Universitäten. *Zeitschrift für Pädagogik*, 56 (4), 531-547.

Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie* (Empirische Erziehungswissenschaft, Bd. 36). Münster: Waxmann.

Spinner, K.H. (2004). Lesekompetenz in der Schule. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Verteilende Analysen im Rahmen von PISA 2000* (S. 125 – 138). Wiesbaden: VS Verlag für Sozialwissenschaften.

Statistisches Bundesamt (2016). *Schulen auf einen Blick*. Zugriff am 25.05.2017 https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/BroschuereSchulenBlick0110018169004.pdf?__blob=publicationFile

Statistisches Landesamt Baden-Württemberg (2018). *Gemeinschaftsschulen erstmals mit rückläufigen Übergangsquoten. Baden-Württemberg: Zahl der Übergänge auf weiterführende Schulen leicht steigend*. Zugriff am 20.02.2018 <https://www.statistik-bw.de/Presse/Pressemitteilungen/2018013>

Tillmann, K.-J. (2004). System jagt Fiktion. Die homogene Lerngruppe. *Friedrich Jahresheft*, 23, 6–9.

Watson, C. S., Kidd, G. R., Horner, D. G., Connell, P. J., Lowther, A. Eddins, D. A., Krueger, G., Goss, D. A., Rainey, B. B., Gospel, M. D. & Watson, B. U. (2003). Sensory, Cognitive, and Linguistic Factors on the Early Academic Performance of Elementary School Children: The Benton-IU Project. *Journal of Learning Disabilities*, 36 (2), 165- 197.