



**Universität
Zürich** ^{UZH}

**Institut für Bildungsevaluation
Assoziiertes Institut der Universität Zürich**

Schlussprüfung an der Volksschule des Kantons Basel-Stadt

Schlussbericht 2017

Lukas Giesinger

Zürich, 14. September 2017

Institut für Bildungsevaluation
Assoziiertes Institut der Universität Zürich
Wilfriedstrasse 15
8032 Zürich

Tel: 043 268 39 60
Fax: 043 268 39 67
www.ibe.uzh.ch

lukas.giesinger@ibe.uzh.ch

Inhaltsverzeichnis

1	Einleitung	4
2	Das Wichtigste in Kürze	4
3	Leistungen der Schülerinnen und Schüler	7
3.1	Mathematik: Vergleich zwischen den Leistungszügen	7
3.2	Mathematik: Fähigkeiten der Schülerinnen und Schüler	8
3.3	Deutsch <i>Lesen und Verstehen</i> : Vergleich zwischen den Leistungszügen	9
3.4	Deutsch <i>Lesen und Verstehen</i> : Fähigkeiten der Schülerinnen und Schüler	10
3.5	Leistungsunterschiede zwischen Mädchen und Knaben	10
4	Leistungen nach Klassen	11
4.1	Leistungen der Klassen in Mathematik sowie Deutsch <i>Lesen und Verstehen</i>	11
4.2	Zusammenhang mit der Klassenzugehörigkeit	13
5	Leistungen nach Schulen	13
6	Leistungsentwicklung an der Volksschule des Kantons Basel-Stadt	15
6.1	Entwicklung der durchschnittlichen Leistungen der Schülerinnen und Schüler an der WBS	15
6.2	Fähigkeiten der Schülerinnen und Schüler an der Volksschule des Kantons Basel-Stadt	15
7	Fazit	18
8	Anhang	20
8.1	Glossar der statistischen Begriffe und Abkürzungen	20
8.2	Daten zu den Abbildungen 6 und 7	21

1 Einleitung

Am Ende der obligatorischen Schulzeit wird an der Weiterbildungsschule Basel-Stadt (WBS) eine Schlussprüfung durchgeführt. Mit der Schlussprüfung werden die Fähigkeiten der Schülerinnen und Schüler in den Fächern Mathematik und Deutsch mit einem einheitlichen Leistungstest geprüft und benotet. Die Noten sind Teil des Schlusszeugnisses der Schülerinnen und Schüler.

Seit 2007 werden die Schlussprüfungen vom Institut für Bildungsevaluation (IBE) durchgeführt. Das IBE übernimmt im Auftrag des Erziehungsdepartements und der Volksschulen des Kantons Basel-Stadt die Entwicklung und Erprobung der Prüfungsaufgaben sowie die Korrektur, Auswertung und Rückmeldung der Prüfungsergebnisse. Nach Abschluss der Schlussprüfung werden die Prüfungsergebnisse in einem Bericht dargestellt und mit den bisherigen Ergebnissen verglichen. Die Schlussprüfung ist ein zentrales Instrument der Leistungsmessung und der Qualitätssicherung des kantonalen Schulsystems. Sie fand 2017 zum letzten Mal statt und wird ab dem kommenden Jahr durch den Check S3 des Bildungsraums Nordwestschweiz ersetzt.

Im Jahre 2013 hat das Pädagogische Zentrum Basel Stadt (PZ.BS) neue Prüfungsmaterialien entwickelt. Aus diesem Grund sind Vergleiche der aktuellen Ergebnisse mit den Ergebnissen der Jahre 2012 und früher nur eingeschränkt möglich. Dies gilt insbesondere im Fach Deutsch, wo die Bereiche *Grammatik* und *Rechtschreibung* nicht mehr geprüft werden. Im vorliegenden Bericht werden die diesjährigen Deutschleistungen daher lediglich mit denjenigen der Jahre 2013 bis 2016 verglichen.

Der Bericht ist in sieben Kapitel gegliedert. In Kapitel 2 sind die Durchführung der Schlussprüfung und die Berechnung der Leistungsskalen beschrieben. In Kapitel 3 werden die Leistungen sowie die Fähigkeiten der Schülerinnen und Schüler in den Schlussprüfungen 2017 beschrieben und zueinander in Beziehung gesetzt. In Kapitel 4 sind die durchschnittlichen Leistungen der Klassen und in Kapitel 5 die durchschnittlichen Leistungen der einzelnen Schulen in der Schlussprüfung 2017 dargestellt. In Kapitel 6 richtet sich der Fokus auf die Leistungsentwicklung an der Volksschule des Kantons Basel-Stadt. Der Bericht schliesst mit einer Zusammenfassung der wichtigsten Ergebnisse.

2 Das Wichtigste in Kürze

Teilnahme – An der Schlussprüfung 2017 nahmen 855 Schülerinnen und Schüler aus 57 Klassen teil. 773 Schülerinnen und Schüler (90 Prozent) besuchten Regelklassen der Weiterbildungsschule (WBS), 23 Schülerinnen und Schüler (3 Prozent) waren heilpädagogischen Spezialangeboten (SpA) und 59 Schülerinnen und Schüler (7 Prozent) Fremdsprachenklassen zugeteilt.

Von den Schülerinnen und Schülern der WBS-Regelklassen waren 325 einer Regelklasse des A-Zugs mit allgemeinen Anforderungen und 448 einer Regelklasse des E-Zugs mit erweiterten Anforderungen zugeteilt. Von Letzteren besuchten 42 Schülerinnen und Schüler eine Musik- und 16 eine Sportklasse. Die Gymnasien nehmen seit 2015 nicht mehr an der Schlussprüfung teil, da ihre schulische Situation mit derjenigen von Abschlussklassen nur bedingt vergleichbar ist.

Testinstrumente – Mit der Schlussprüfung wurden die Leistungen in den Fächern Mathematik und Deutsch geprüft. Die Prüfungsaufgaben wurden von Fachexpertinnen und Fachexperten des PZ.BS und erfahrenen Lehrkräften der Oberstufe gemeinsam erarbeitet und einem Testlauf unterzogen. Die Testhefte sind so konzipiert, dass sie von allen Schülerinnen und Schülern des 11. Schuljahres – unabhängig von ihrer Schulform – bearbeitet werden können.

Das Fach Mathematik wurde in zwei Teilen geprüft. Der erste Teil umfasste 45 Aufgaben zu den beiden Bereichen *Arithmetik/Algebra* und *Sachrechnen/Funktionen*, die ohne Taschenrechner gelöst werden mussten. Der zweite Teil enthielt 30 Aufgaben zu den beiden genannten Teilbereichen sowie zu *Geometrie*, die mit Hilfe eines Taschenrechners gelöst werden konnten.

Mit dem Deutschtest wurden zum einen die Fähigkeiten im Bereich *Lesen und Verstehen* und zum anderen die Fähigkeiten im Bereich *Schreiben* geprüft. Für *Lesen und Verstehen* mussten die Schülerinnen und Schüler zwei Zeitungsausschnitte und zwei Kurzgeschichten von unterschiedlicher Länge und Komplexität lesen und Fragen zum Textverständnis beantworten. Im Bereich *Schreiben* mussten die Schülerinnen und Schüler drei Texte schreiben. Im ersten Text mussten die Schülerinnen und Schüler möglichst präzise eine Situation beschreiben, in der sie sich gelangweilt haben. Im zweiten Text ging es darum, zu argumentieren, ob es gut ist, wenn sich Jugendliche langweilen. Die dritte Textaufgabe bestand darin, einen Brief an die Schulleitung zu verfassen, mit der Bitte, eine Klassenreise nach Barcelona zu unterstützen.

Durchführung – Die Schlussprüfung fand Ende Mai 2017 statt und wurde von den Schulen organisiert. Die Prüfungshefte wurden anschliessend an das IBE nach Zürich gebracht, wo sie von einem Team von Fachwissenschaftlerinnen und Fachwissenschaftlern mit Unterrichtserfahrung korrigiert wurden. Nach der Korrektur wurden die Aufgaben elektronisch erfasst und plausibilisiert. Die Testresultate der Schülerinnen und Schüler wurden anhand einer Notenskala, die seit 2007 in nahezu unveränderter Form eingesetzt wird, in Noten umgerechnet.

Zwei Wochen nach der Durchführung der Prüfung wurden die Noten der Schülerinnen und Schüler an die Schulleitungen versandt. Den Lehrpersonen wurden die Prüfungsergebnisse ihrer Schülerinnen und Schüler sowie das durchschnittliche Klassenergebnis zugestellt. Zudem erhielten die Schulleitungen der WBS einen kurzen Bericht zu den durchschnittlichen Leistungen der Klassen an ihrer Schule. Für die Schulkreisleitungen der Volksschule wurde eine Rückmeldung mit den Resultaten ihrer jeweiligen Schulen erstellt.

Skalierung der Leistungsdaten – In diesem Bericht werden die Leistungen der Schülerinnen und Schüler als Punktzahl auf der Mathematik- beziehungsweise auf der Deutschskala ausgewiesen. Die Punktzahl wurde mittels einer Rasch-Skalierung der Prüfungsergebnisse berechnet. Grundlage für die Skala sind die Ergebnisse des Leistungstests, der 2004 im Rahmen der Evaluation der Strukturänderung an der WBS am IBE durchgeführt wurde. Die Skala wurde so standardisiert, dass der Mittelwert aller Schülerinnen und Schüler der Schlussprüfungen 2005 in jedem Fach 500 Punkte und die Standardabweichung 100 Punkte betrug. Aufgaben, deren Lösungswahrscheinlichkeit sich zwischen den Schulformen oder zwischen den Testjahren stark unterschied (so genanntes Differential Item Functioning; DIF) wurden aus der Skalierung ausgeschlossen.

Die Rasch-Skalierung der Leistungsdaten hat verschiedene Vorteile. Erstens werden mit der Rasch-Skalierung sowohl die Leistungen der Schülerinnen und Schüler als auch die Schwierigkeit der Prüfungsaufgaben auf der gleichen Skala mit einer einheitlichen Metrik abgebildet. Die Schwierigkeit der Aufgaben und die Leistung der Schülerinnen und Schüler stehen dabei in einer definierten Beziehung zueinander. Eine Schülerin beispielsweise, die eine Leistung von 700 Punkten erreicht, kann eine Aufgabe mit der Schwierigkeit von 700 Punkten mit einer Wahrscheinlichkeit von 50 Prozent richtig lösen. Im Wissen um diese Lösungswahrscheinlichkeit können die Schülerleistungen anhand von Testaufgaben inhaltlich beschrieben werden. Damit die Beschreibung der Leistungen übersichtlich wird, wurden ähnlich schwierige Aufgaben zu Intervallen (Kompetenzniveaus) zusammengefasst.

Ein zweiter Vorteil ist, dass die Grösse beziehungsweise die Relevanz von Leistungsunterschieden – beispielsweise zwischen dem WBS-A-Zug und dem WBS-E-Zug oder zwischen Mädchen und Knaben – einfach interpretiert werden kann. Leistungsunterschiede sind dann relevant, wenn sie 20 Punkte oder mehr betragen. Unterschiede von 80 und mehr Punkten sind sehr gross.

Ein dritter Vorteil ist, dass mittels der Rasch-Skalierung der Prüfungsergebnisse auch die Entwicklung der Leistungen während mehrerer Testjahre beschrieben werden kann. Da ausgewählte Aufgaben während mehrerer Jahre identisch eingesetzt werden, bleibt die Metrik der Skala immer konstant. Diese Methode ermöglicht es, die Leistungen der Schülerinnen und Schüler in den Jahren 2005 bis 2017 trotz teilweise neuer Aufgaben auf der gleichen Punkteskala darzustellen und zu vergleichen. Schülerinnen und Schüler, die 2005 500 Punkte erreichten, hatten die gleichen Fähigkeiten wie die Schülerinnen und Schüler, die 2017 500 Punkte erreichten. Auch Leistungsmittelwerte können verglichen werden: Werte über 500 Punkte entsprechen einer höheren Leistung im Vergleich zu den durchschnittlichen Leistungen in der Schlussprüfung im Jahr 2005.

Eine Einschränkung betrifft das Fach Deutsch. Mit der Schlussprüfung wurden die Fähigkeiten in den Bereichen *Lesen und Verstehen* sowie *Schreiben* getestet. Wie frühere Analysen gezeigt haben, sind für die Leistungen im Lesen und im Schreiben unterschiedliche, d.h. nicht direkt vergleichbare Kompetenzen notwendig. Da der Bereich *Schreiben* zudem über die Jahre hinweg nicht identisch korrigiert werden kann, wird für den Schlussbericht nur der Bereich *Lesen und Verstehen* skaliert und analysiert. Zudem werden seit 2013 die Bereiche *Grammatik* und *Rechtschreibung* nicht mehr geprüft. Damit hat sich der fachliche Inhalt der Tests verändert. Ein Vergleich mit den früheren Deutschleistungen ist daher nicht möglich.

Intervalle – Für die inhaltliche Beschreibung werden die Leistungen der Schülerinnen und Schüler in Intervalle von jeweils 100 Punkten eingeteilt. Schülerinnen und Schüler, die ein bestimmtes Intervall erreichen, können mindestens die Hälfte aller Aufgaben in diesem Intervall richtig lösen. Erreicht eine Schülerin 401 Punkte, dann beträgt die durchschnittliche Lösungswahrscheinlichkeit für die Aufgaben des Intervalls *401 bis 500 Punkte* 50 Prozent. Diese Schülerin ist in der Lage, die Hälfte der Aufgaben des Intervalls richtig zu lösen. Mit zunehmender Punktzahl steigt auch die Wahrscheinlichkeit, die Aufgaben eines Intervalls richtig lösen zu können. Erreicht eine Schülerin beispielsweise im Mathematiktest 450 Punkte, dann beträgt die durchschnittliche Lösungswahrscheinlichkeit für die Aufgaben des Intervalls *401 bis 500 Punkte* 62 Prozent. Das heisst, die Schülerin löst die Aufgaben dieses Intervalls mit einer Wahrscheinlichkeit von 62 Prozent richtig.

Weiterführende Informationen – Informationen zum methodischen Vorgehen bei der Skalierung der Prüfungsergebnisse sowie eine ausführliche Beschreibung der Intervalle anhand von Beispielaufgaben finden sich im Schlussbericht zur Evaluation der Strukturänderung an der WBS Basel-Stadt:

Moser, U. & Keller, F. (2006): *Evaluation der Strukturänderung an der Weiterbildungsschule Basel-Stadt. Schlussbericht zuhanden des Erziehungsdepartements des Kantons Basel-Stadt, Ressort Schulen*. Zürich: Kompetenzzentrum für Bildungsevaluation.
<http://www.ibe.uzh.ch/publikationen/Evaluation-WBS.pdf>

Die Berichte mit den Ergebnissen der Schlussprüfungen 2007 bis 2016 können auf der Website des IBE heruntergeladen werden:

<http://www.ibe.uzh.ch/projekte/wbsabschluss.html>

3 Leistungen der Schülerinnen und Schüler

3.1 Mathematik: Vergleich zwischen den Leistungszügen

Abbildung 1 zeigt die Verteilung der Leistungen der Schülerinnen und Schüler in der Mathematikprüfung 2017 für die beiden Leistungszüge der WBS im Kanton Basel-Stadt. Der fett gedruckte Balken zeigt, wo der Mittelwert der Schülerinnen und Schüler liegt. Die Breite der Box gibt den Leistungsbereich an, in welchem die Ergebnisse der mittleren 50 Prozent der Schülerinnen und Schüler liegen. Zählt man noch die Breite der T-förmigen Elemente hinzu, so erhält man den Bereich, in welchem 90 Prozent der Ergebnisse liegen.

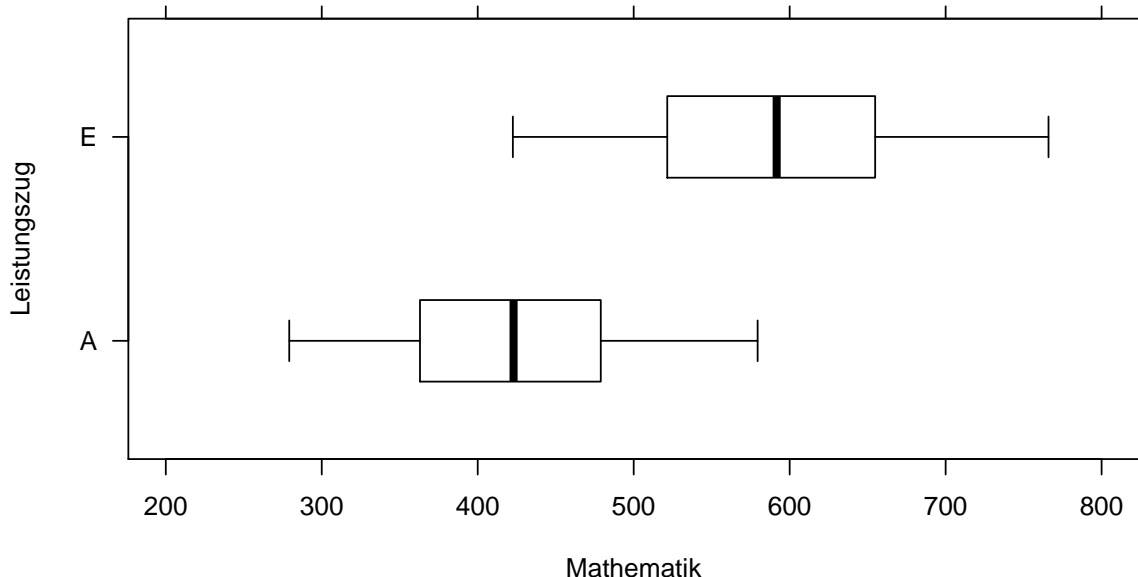


Abbildung 1: Verteilung der Mathematikleistungen nach Leistungszug; nur Regelklassen

E-Zug: $n = 442$; $M = 592$ Punkte; $SD = 104$ Punkte

A-Zug: $n = 316$; $M = 423$ Punkte; $SD = 94$ Punkte

Die Schülerinnen und Schüler des E-Zugs erreichten am Ende des 11. Schuljahres durchschnittlich 592 Punkte auf der Mathematikskala, die Schülerinnen und Schüler des A-Zugs 423 Punkte. Die Differenz zwischen den durchschnittlichen Mathematikleistungen der Schülerinnen und Schüler des A- und des E-Zugs betrug also 169 Punkte. Dieser Unterschied ist sehr gross (Effektgrösse $d = 1.69$).

Betrachtet man in Abbildung 1 die Verteilungen der Leistungen der verschiedenen Züge, so fällt eine deutliche Trennung zwischen den Leistungen der Schülerinnen und Schüler des A- und E-Zugs auf. Die Leistungsüberschneidungen zwischen den beiden Zügen fielen gering aus: 3 Prozent der Schülerinnen und Schüler des A-Zugs erzielten in Mathematik eine höhere Leistung als der Durchschnitt der Schülerinnen und Schüler des E-Zugs, und 5 Prozent der Schülerinnen und Schüler des E-Zugs erreichten tiefere Mathematikleistungen als der Durchschnitt der Schülerinnen und Schüler des A-Zugs.

3.2 Mathematik: Fähigkeiten der Schülerinnen und Schüler

Damit sich die Mathematikleistungen inhaltlich interpretieren lassen, werden sie in sogenannte Intervalle (Kompetenzniveaus) eingeteilt. Die Intervalle erlauben eine Beschreibung, was die Schülerinnen und Schüler innerhalb eines Leistungsbereiches wissen und können. Schülerinnen und Schüler, die aufgrund ihrer Leistungen einem bestimmten Intervall zugeteilt werden, können mindestens die Hälfte der Aufgaben eines Intervalls richtig lösen. Aufgaben eines tieferen Intervalls werden mit einer grösseren Wahrscheinlichkeit richtig gelöst, Aufgaben eines höheren Intervalls werden mit geringerer Wahrscheinlichkeit richtig gelöst. Tabelle 1 zeigt die Verteilung der Schülerinnen und Schüler am Ende des 11. Schuljahres auf die Intervalle in Mathematik nach Leistungszug.

Tabelle 1: Verteilung der Schülerinnen und Schüler auf die Intervalle in Mathematik; ohne Schülerinnen und Schüler der Fremdsprachen-, Musik- und Sportklassen sowie der SpA¹

Leistungszug	Intervall				
	≤ 400	401–500	501–600	601–700	> 700
WBS E	2%	17%	36%	29%	16%
WBS A	43%	38%	16%	3%	–

2 Prozent der Schülerinnen und Schüler des E-Zugs sowie 43 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS weniger als 400 Punkte auf der Mathematikskala. Diesen Schülerinnen und Schülern fehlten in Mathematik jene Grundkompetenzen, die zur Lösung von einfachen Grundoperationen notwendig sind.

98 Prozent der Schülerinnen und Schüler des E-Zugs sowie 57 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 400 Punkte. Sie kannten unter anderem Zahlenarten und Zahlenmengen sowie die wichtigsten Grundoperationen und konnten sie anwenden.

81 Prozent der Schülerinnen und Schüler des E-Zugs sowie 19 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 500 Punkte. Sie konnten beispielsweise einfache Prozentrechnungen lösen, Umfang und Fläche von Vierecken sowie das Volumen eines Quaders berechnen.

45 Prozent der Schülerinnen und Schüler des E-Zugs sowie 3 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 600 Punkte. Sie konnten unter anderem Grundoperationen mit Dezimalzahlen durchführen und lineare Gleichungen lösen und die Fläche eines Dreiecks berechnen.

16 Prozent der Schülerinnen und Schüler des E-Zugs erreichten am Ende der WBS mehr als 700 Punkte. Sie konnten unter anderem mit Potenzen und Quadratwurzeln umgehen sowie Rauminhalte und Flächen von geometrischen Körpern berechnen. Keine Schülerin, kein Schüler des A-Zugs erreichte dieses Kompetenzniveau.

¹ Aufgrund von Rundungsfehlern kann es vorkommen, dass sich die Zahlen einer Tabellenzeile nicht auf 100 Prozent addieren. Die ungerundeten Anteile der Schülerinnen und Schüler addieren sich jedoch auf 100 Prozent.

3.3 Deutsch *Lesen und Verstehen*: Vergleich zwischen den Leistungszügen

Abbildung 2 zeigt die Verteilung der Leistungen der Schülerinnen und Schüler im Bereich *Lesen und Verstehen* der Deutschprüfung 2017. Der fett gedruckte Balken zeigt für die beiden Leistungszüge der WBS, wo der Mittelwert der Schülerinnen und Schüler liegt. Die Breite der Box gibt den Leistungsbe-
reich an, in dem die mittleren 50 Prozent der Schülerinnen und Schüler liegen. Zählt man noch die Breite der T-förmigen Elemente hinzu, so erhält man den Bereich, in dem 90 Prozent der Schülerinnen und Schüler liegen.

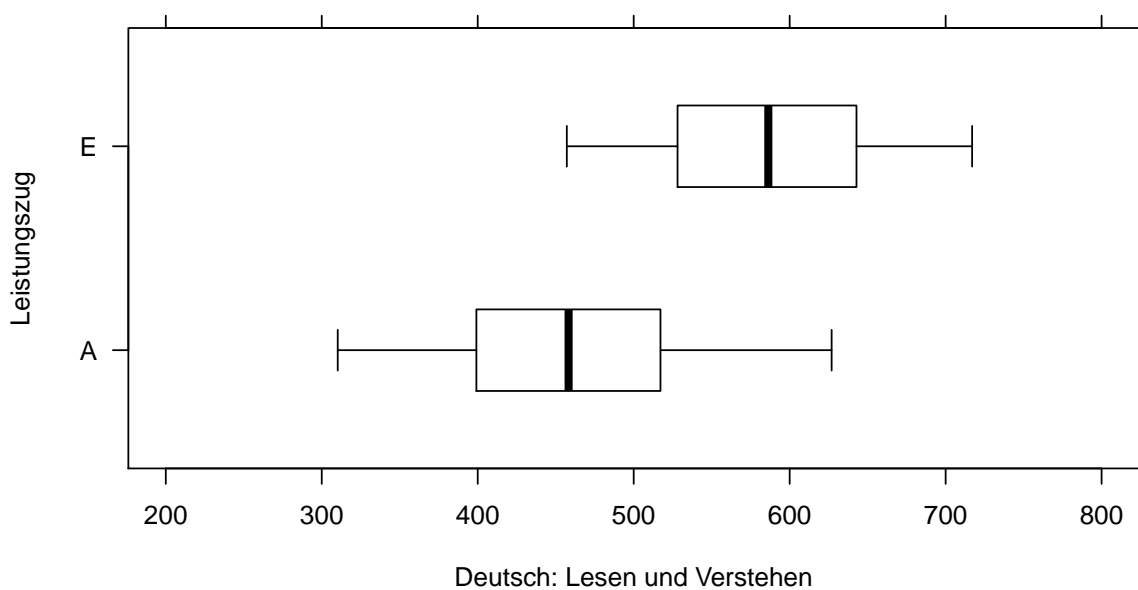


Abbildung 2: Verteilung der Leistungen im in Deutsch *Lesen und Verstehen* nach Leistungszug; nur Regelklassen

E-Zug: $n = 442$; $M = 586$ Punkte; $SD = 84$ Punkte

A-Zug: $n = 318$; $M = 458$ Punkte; $SD = 97$ Punkte

Die Schülerinnen und Schüler im A-Zug der WBS erreichten am Ende des 11. Schuljahres durchschnittlich 458 Punkte im Bereich *Lesen und Verstehen*, die Schülerinnen und Schüler des WBS E-Zugs 586 Punkte. Die Differenz zwischen den durchschnittlichen Leistungen der Schülerinnen und Schüler des E- und des A-Zugs betrug also 128 Punkte. Dieser Unterschied ist mit einer Effektgrösse von $d = 1.43$ sehr gross.

Die Leistungsüberschneidungen zwischen dem A- und dem E-Zug waren im Bereich *Lesen und Verstehen*, ähnlich wie in Mathematik, vergleichsweise gering: Rund 8 Prozent der Schülerinnen und Schüler des A-Zugs erzielten in *Lesen und Verstehen* eine höhere Leistung als der Durchschnitt der Schülerinnen und Schüler des E-Zugs. Umgekehrt erreichten rund 6 Prozent der Schülerinnen und Schüler des E-Zugs tiefere Leistungen als der Durchschnitt der Schülerinnen und Schüler des A-Zugs.

3.4 Deutsch *Lesen und Verstehen*: Fähigkeiten der Schülerinnen und Schüler

Tabelle 2 zeigt die Verteilung der Schülerinnen und Schüler am Ende des 11. Schuljahres auf die Intervalle im Bereich *Lesen und Verstehen*. Schülerinnen und Schüler, die einem bestimmten Intervall zugeordnet werden, können die Aufgaben in diesem Intervall sowie alle Aufgaben der tieferen Intervalle mit einer Wahrscheinlichkeit von mindestens 50 Prozent richtig lösen.

Tabelle 2: Verteilung der Schülerinnen und Schüler auf die Intervalle in Deutsch *Lesen und Verstehen*; ohne Schülerinnen und Schüler der Fremdsprachen-, Musik- und Sportklassen sowie der SpA (vgl. Fussnote 1)

Leistungszug	Intervall				
	≤ 400	401–500	501–600	601–700	> 700
WBS E	1%	14%	40%	38%	7%
WBS A	26%	42%	24%	7%	1%

1 Prozent der Schülerinnen und Schüler des E-Zugs und 26 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende des 11. Schuljahres weniger als 400 Punkte im Bereich *Lesen und Verstehen*. Sie verstanden die Hauptaussagen eines einfachen, kontinuierlichen Textes.

99 Prozent der Schülerinnen und Schüler des E-Zugs sowie 74 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 400 Punkte. Sie verstanden einen einfachen Text so weit, dass sie zwischen den Informationen im Text Beziehungen herstellen konnten.

85 Prozent der Schülerinnen und Schüler des E-Zugs sowie 32 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 500 Punkte. Sie erkannten in einem Text verschiedene argumentative Zusammenhänge und konnten diese in eigenen Worten zusammenfassen.

45 Prozent der Schülerinnen und Schüler des E-Zugs sowie 8 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 600 Punkte. Sie verstanden komplexere Texte und konnten Folgerungen aus einem Text ableiten und diese anhand von Textstellen begründen. Sie konnten zudem über die gestalterischen Merkmale eines fiktionalen Texts Auskunft geben.

7 Prozent der Schülerinnen und Schüler des E-Zugs sowie 1 Prozent der Schülerinnen und Schüler des A-Zugs erreichten am Ende der WBS mehr als 700 Punkte. Sie waren unter anderem in der Lage, in einen Text eingebettete Informationen zu finden und sprachliche Nuancen zu verstehen.

3.5 Leistungsunterschiede zwischen Mädchen und Knaben

Abbildung 3 zeigt die Leistungsunterschiede zwischen Mädchen und Knaben am Ende des 11. Schuljahres. In Mathematik erreichten die Knaben im Durchschnitt 58 Punkte mehr als die Mädchen. Dieser Unterschied ist statistisch signifikant und mit einer Effektstärke von $d = 0.45$ gering ausgeprägt. In *Lesen und Verstehen* erreichten die Knaben im Durchschnitt 10 Punkte mehr als die Mädchen. Dieser Unterschied ist statistisch nicht signifikant und mit $d = 0.06$ sehr gering ausgeprägt. Die zur Abbildung gehörige deskriptive Statistik ist in Tabelle 3 angegeben.

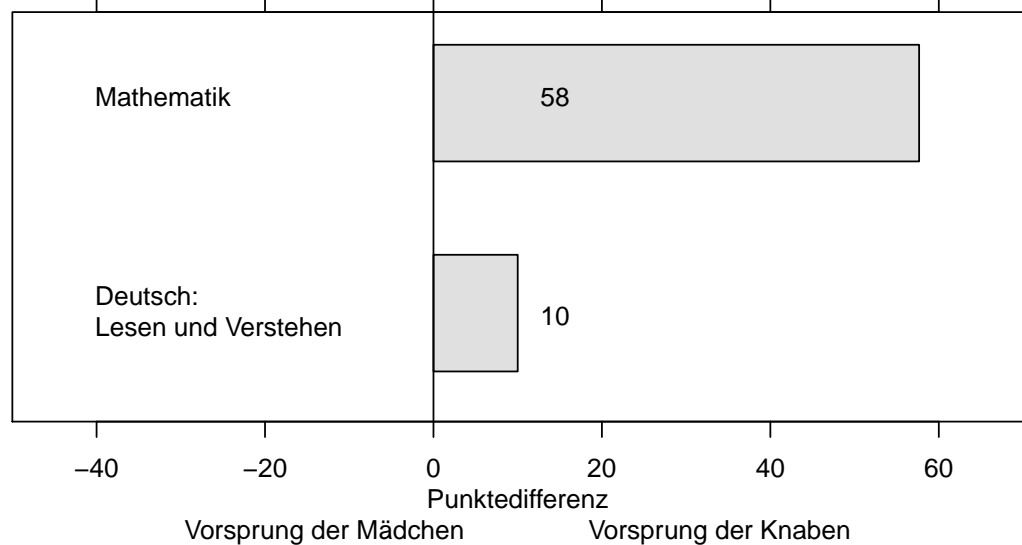


Abbildung 3: Leistungsunterschiede zwischen Mädchen und Knaben. Positive Werte entsprechen einem Leistungsvorteil der Knaben.

Tabelle 3: Leistungen in Mathematik sowie Deutsch *Lesen und Verstehen* nach Geschlecht: Stichprobengrösse n , Mittelwert M , Standardabweichung SD

Geschlecht	Mathematik			Deutsch <i>Lesen und Verstehen</i>		
	n	M	SD	n	M	SD
Knaben	461	538	134	460	521	122
Mädchen	362	480	121	362	514	118

4 Leistungen nach Klassen

4.1 Leistungen der Klassen in Mathematik sowie Deutsch *Lesen und Verstehen*

Die Klassen des E-Zugs erreichten in Mathematik im Durchschnitt 592 Punkte und in *Lesen und Verstehen* 589 Punkte. Die Spannweite zwischen dem höchsten und dem tiefsten Klassenmittelwert des E-Zugs betrug in Mathematik 132 Punkte und in *Lesen und Verstehen* 138 Punkte.

Die Klassen des A-Zugs erreichten in Mathematik durchschnittlich 420 Punkte und in *Lesen und Verstehen* 458 Punkte. Die Spannweite zwischen dem höchsten und dem tiefsten Klassenmittelwert des A-Zugs betrug in Mathematik 138 Punkte und in *Lesen und Verstehen* 95 Punkte.

Wie Abbildung 4 zeigt, unterschieden sich die Klassenleistungen nicht nur innerhalb des A- bzw. des E-Zugs erheblich. Auch zwischen den beiden Zügen bestanden deutliche Leistungsunterschiede. So erbrachten sämtliche WBS-E-Klassen bessere Leistungen in den Bereichen Mathematik und *Lesen und Verstehen* als die WBS-A-Klassen.

Alle drei Fremdsprachenklassen zeigten geringere Leistungen in *Lesen und Verstehen* als sämtliche

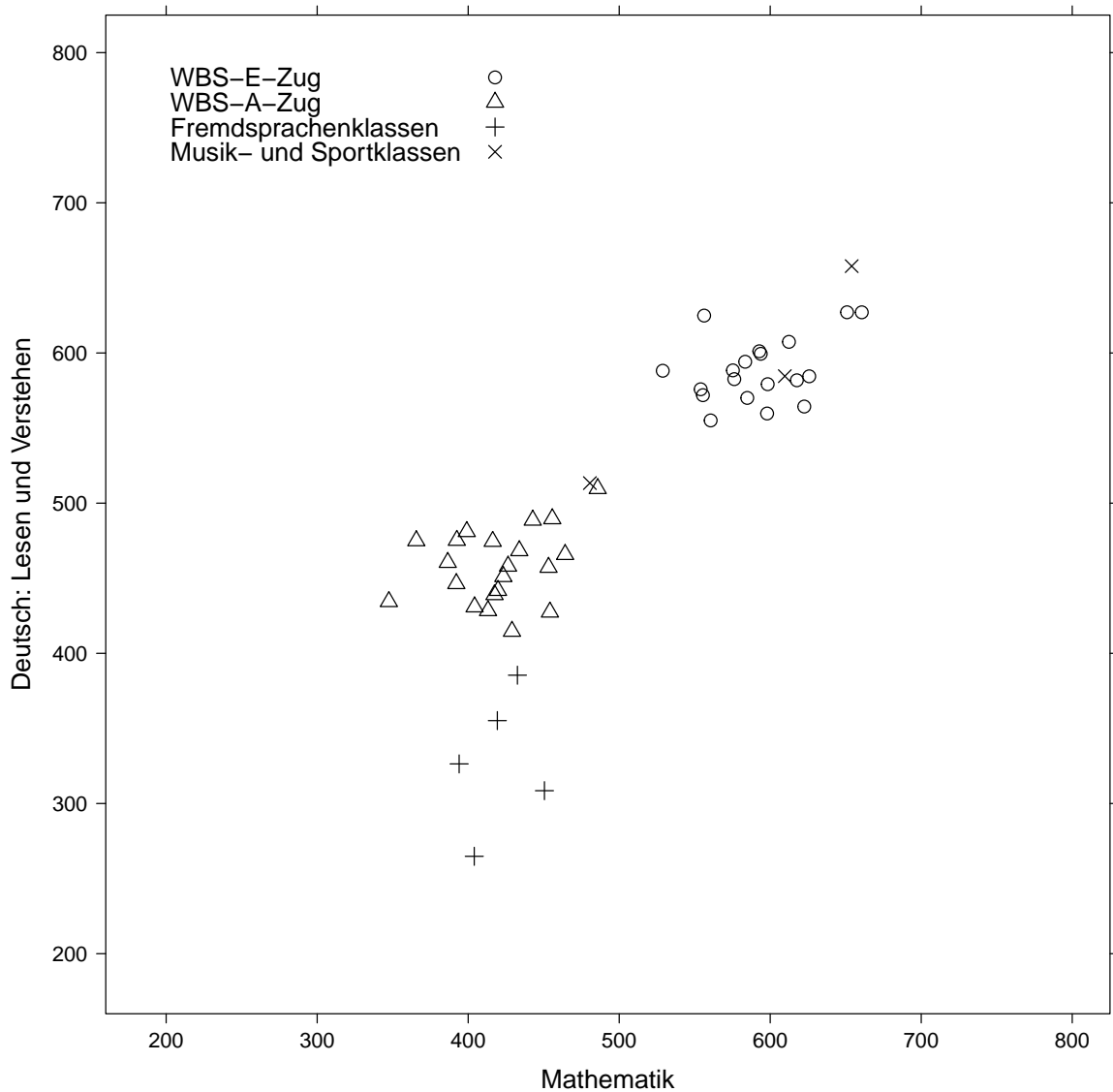


Abbildung 4: Klassenmittelwerte in Mathematik sowie Deutsch *Lesen und Verstehen*; Klassen mit weniger als fünf Schülerinnen und Schülern sowie die Klassen der Spezialangebote (SpA) sind nicht dargestellt.

übrigen Klassen. In Mathematik hingegen lagen die Leistungen der Fremdsprachenklassen in einem ähnlichen Bereich wie diejenigen des A-Zugs.

Die Musik- und Sportklassen sind vom Leistungsniveau her dem E-Zug zuzurechnen. Entsprechend zeigten sie Leistungen, die denjenigen von Klassen des E-Zugs beziehungsweise der leistungsstärksten Klasse des A-Zugs entsprechen.

Tabelle 4 zeigt die mittleren Klassenleistungen nach Leistungszug.

Tabelle 4: Leistungen der Klassen in Mathematik sowie Deutsch *Lesen und Verstehen* nach Leistungszug: Stichprobengrösse n , Mittelwert M , Standardabweichung SD

Leistungszug	Mathematik			Deutsch <i>Lesen und Verstehen</i>		
	n	M	SD	n	M	SD
WBS E	19	592	34	19	589	22
WBS A	21	420	33	21	458	24

4.2 Zusammenhang mit der Klassenzugehörigkeit

Die Unterschiedlichkeit der Leistungen der Schülerinnen und Schüler kann einerseits durch die Klassenzugehörigkeit, andererseits durch individuelle Merkmale der Schülerin oder des Schülers erklärt werden. Je stärker die Leistungen zwischen den Klassen variieren, desto eher können sie durch Merkmale der Klasse oder des Unterrichts erklärt werden.

Tabelle 5 zeigt für Mathematik sowie Deutsch *Lesen und Verstehen* die Prozentanteile an der Schülerleistung, die – innerhalb eines Leistungszugs – durch die Klassenzugehörigkeit (obere Zeile) beziehungsweise durch individuelle Merkmale (untere Zeile) erklärt werden können.

Tabelle 5: Einfluss von Klassen- versus individuellen Merkmalen auf die Schülerleistung in Mathematik sowie Deutsch *Lesen und Verstehen*; ohne Sport-, Musik- und Fremdsprachenklassen und SpA

Schülerleistung abhängig von ...	Mathematik		Deutsch <i>L. u. V.</i>	
	WBS E	WBS A	WBS E	WBS A
... Klassenmerkmalen	7%	5%	4%	0%
... individuellen Merkmalen	93%	95%	96%	100%

Der Anteil der Leistungsunterschiede, welcher durch die Klassenzugehörigkeit erklärt werden kann, unterscheidet sich zwischen E- und A-Zug. In Mathematik können 7 bzw. 5 Prozent durch die Klassenzugehörigkeit erklärt werden. In *Lesen und Verstehen* können im E-Zug 4 Prozent durch die Klassenzugehörigkeit erklärt werden, während die Klassenzugehörigkeit für die Leistungsunterschiede im A-Zug keine Rolle spielt.

Die Klassen- und Unterrichtsmerkmale haben insgesamt nur einen begrenzten Einfluss auf Unterschiede in der Schülerleistung. Zu beachten ist, dass Klassen innerhalb des E- bzw. A-Zugs ähnliche mittlere Schülerleistungen zeigen, was die vergleichsweise geringen Einfluss von Klassenmerkmalen erklärt.

Im E-Zug kann ein grösserer Anteil der Varianz durch Klassenmerkmale erklärt werden als im A-Zug. Der familiäre Hintergrund und die genetische Ausstattung der Schülerin oder des Schülers erklären die Unterschiedlichkeit der Schülerinnen und Schüler des A-Zugs daher in geringerem Masse als diejenige der Schülerinnen und Schüler des E-Zugs.

5 Leistungen nach Schulen

Abbildung 5 zeigt die Ergebnisse der Schülerinnen und Schüler nach Schulen im Kanton Basel-Stadt. Die Position einer Schule ergibt sich aus der durchschnittlichen Punktzahl, die an einer Schule in Ma-

thematik und in Deutsch *Lesen und Verstehen* erreicht wurde. Die Ergebnisse der WBS-Standorte sind als Dreiecke, die Ergebnisse der Spezialangebote (SpA) als Viereck dargestellt.

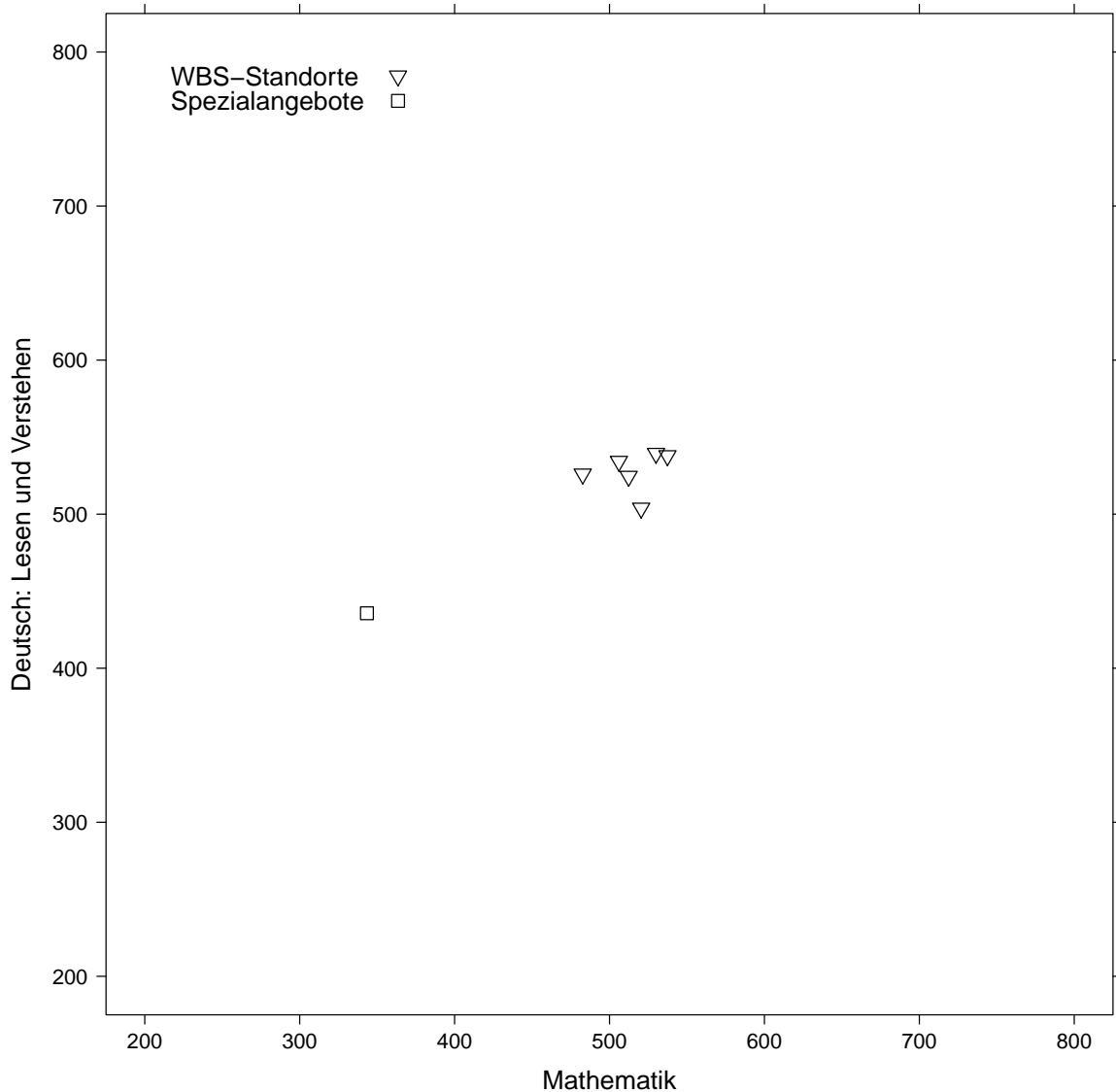


Abbildung 5: Fachleistungen in Mathematik sowie Deutsch *Lesen und Verstehen* nach Schulen. Ohne Sport-, Musik- und Fremdsprachenklassen. Die Leistungen der Regelklassen wurden für den Einfluss des Leistungszugs korrigiert.

Um an den WBS-Standorten den Einfluss der unterschiedlich grossen Anteile an A- und E-Klassen auf das Ergebnis eines Standorts auszugleichen, wurde für den Leistungszug der Schülerinnen und Schüler statistisch kontrolliert. Das heisst, die Ergebnisse von WBS-Standorten mit einem hohen Anteil von Schülerinnen und Schülern im A-Zug wurden nach oben korrigiert, die Ergebnisse von Standorten mit einem tiefen Anteil Schülerinnen und Schüler im A-Zug wurden entsprechend nach unten korrigiert. Die durchschnittlichen Leistungen an den WBS-Schulen unterschieden sich nicht signifikant. Innerhalb

der WBS lag die mittlere Mathematikleistung der leistungsstärksten Schule (537 Punkte) 55 Punkte über derjenigen der leistungsschwächsten Schule. In *Lesen und Verstehen* lag das mittlere Ergebnis der leistungsstärksten Schule (539 Punkte) 36 Punkte über demjenigen der leistungsschwächsten Schule.

Deutlich geringere Leistungen zeigten die Schülerinnen und Schüler in den Spezialangeboten. Sie erreichten in Mathematik durchschnittlich 343 Punkte und in *Lesen und Verstehen* 436 Punkte. Ihre Leistungen lagen in Mathematik 139 Punkte und in *Lesen und Verstehen* 68 Punkte unter den durchschnittlichen Leistungen am schwächsten WBS-Standort.

6 Leistungsentwicklung an der Volksschule des Kantons Basel-Stadt

6.1 Entwicklung der durchschnittlichen Leistungen der Schülerinnen und Schüler an der WBS

Abbildung 6 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Regelklassen des E-Zugs in den Bereichen Mathematik und Deutsch *Lesen und Verstehen* seit 2012. Die Mittelwerte in Mathematik sind als schwarze Punkte, die Mittelwerte in *Lesen und Verstehen* als weisse Punkte dargestellt.

2017 erreichten die Schülerinnen und Schüler des E-Zugs 581 Punkte auf der Mathematikskala. Das sind 15 Punkte weniger als 2016. Diese Leistungsveränderung ist statistisch nicht signifikant und mit einer Effektgrösse von $d = 0.15$ sehr gering. In Deutsch *Lesen und Verstehen* erreichten die Schülerinnen und Schüler des E-Zugs durchschnittlich 567 Punkte, 9 Punkte weniger als im Vorjahr. Dieser Unterschied ist mit $d = 0.13$ ebenfalls sehr gering und statistisch nicht signifikant.

Abbildung 7 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Regelklassen des A-Zugs in den Bereichen Mathematik und *Lesen und Verstehen*. Im Fach Mathematik erreichten die Schülerinnen und Schüler des A-Zugs im Mittel 409 Punkte, 9 Punkte weniger als im Vorjahr. Dieser Unterschied ist statistisch nicht signifikant und mit $d = 0.06$ sehr gering. In *Lesen und Verstehen* sank die Leistung in den WBS-A-Klassen von 445 Punkten im Jahr 2016 auf 437 Punkte im Jahr 2017. Dieser Unterschied von 8 Punkten ist mit $d = 0.08$ sehr gering und statistisch nicht signifikant.

6.2 Fähigkeiten der Schülerinnen und Schüler an der Volksschule des Kantons Basel-Stadt

Neben den durchschnittlichen Leistungen ist der Anteil der Schülerinnen und Schüler, der in Mathematik oder in Deutsch *Lesen und Verstehen* nicht über die notwendigen Grundkompetenzen verfügt, von Interesse. Dazu wurden die Schülerinnen und Schüler aufgrund ihrer Leistungen in Intervalle von jeweils 100 Punkten eingeteilt.

Da die Gymnasien seit 2015 nicht mehr an der Schlussprüfung teilnahmen, sind die Fähigkeitseinstufungen des aktuellen Jahres nicht mit denjenigen seit 2012 vergleichbar. Daher werden die Leistungen des Jahres 2016 hier nur mit denjenigen seit 2015 verglichen.

Tabelle 6 zeigt die Verteilung der Schülerinnen und Schüler am Ende des 11. Schuljahres auf die Intervalle in Mathematik. Schülerinnen und Schüler, die einem bestimmten Intervall zugeordnet werden, konnten die Aufgaben in diesem Intervall sowie alle Aufgaben der tieferen Intervalle mit einer Wahrscheinlichkeit von mindestens 50 Prozent richtig lösen.

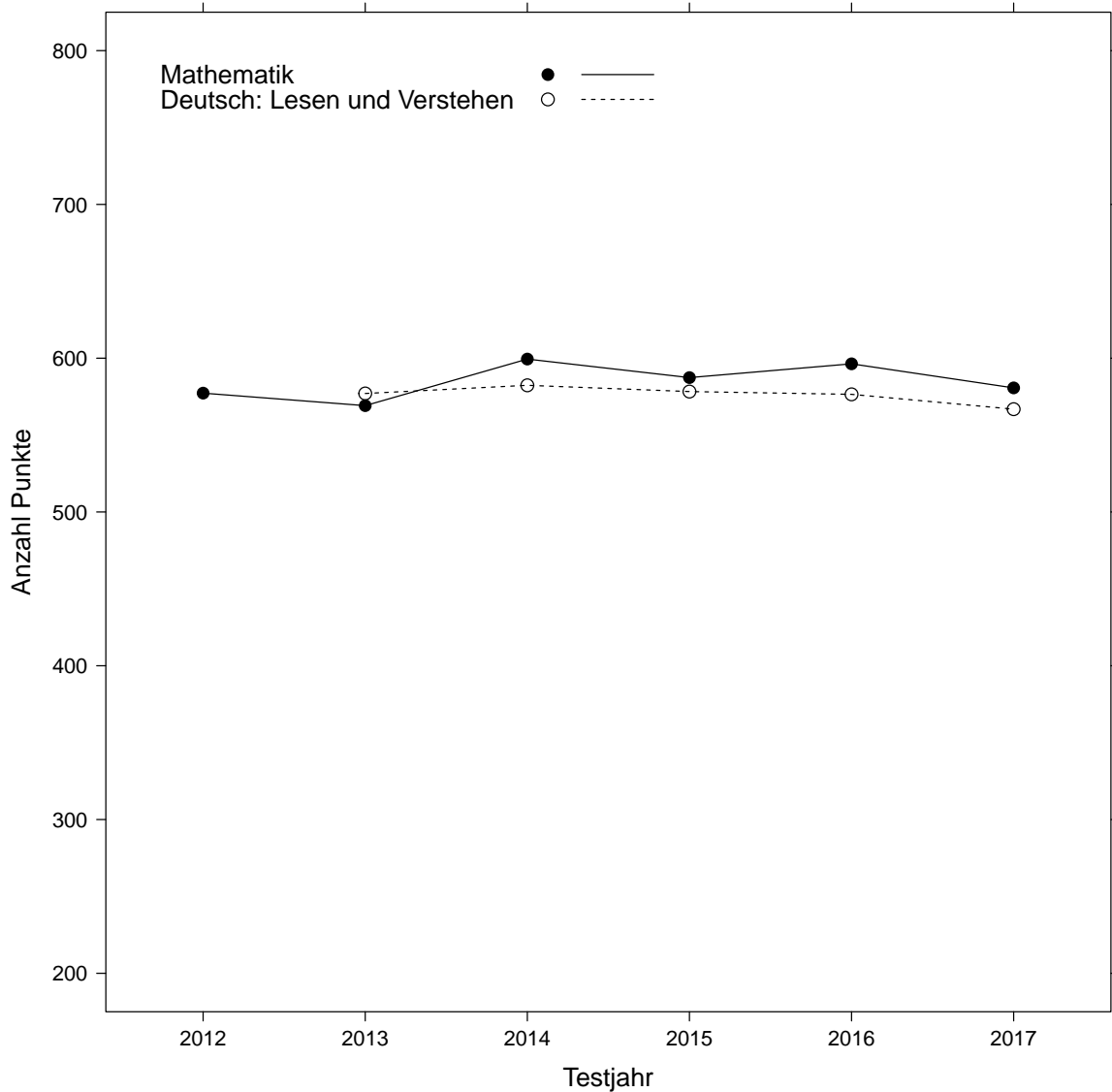


Abbildung 6: Leistungsentwicklung in den Bereichen Mathematik sowie Deutsch *Lesen und Verstehen* im E-Zug (nur Regelklassen); Daten zur Abbildung siehe Anhang

Tabelle 6: Verteilung der Schülerinnen und Schüler auf die Intervalle in Mathematik; alle Schülerinnen und Schüler des 11. Schuljahres, inklusive SpA

Jahr	Intervall				
	≤ 400	401–500	501–600	601–700	> 700
2015	17%	22%	27%	22%	12%
2016	16%	24%	32%	21%	7%
2017	22%	26%	27%	17%	9%

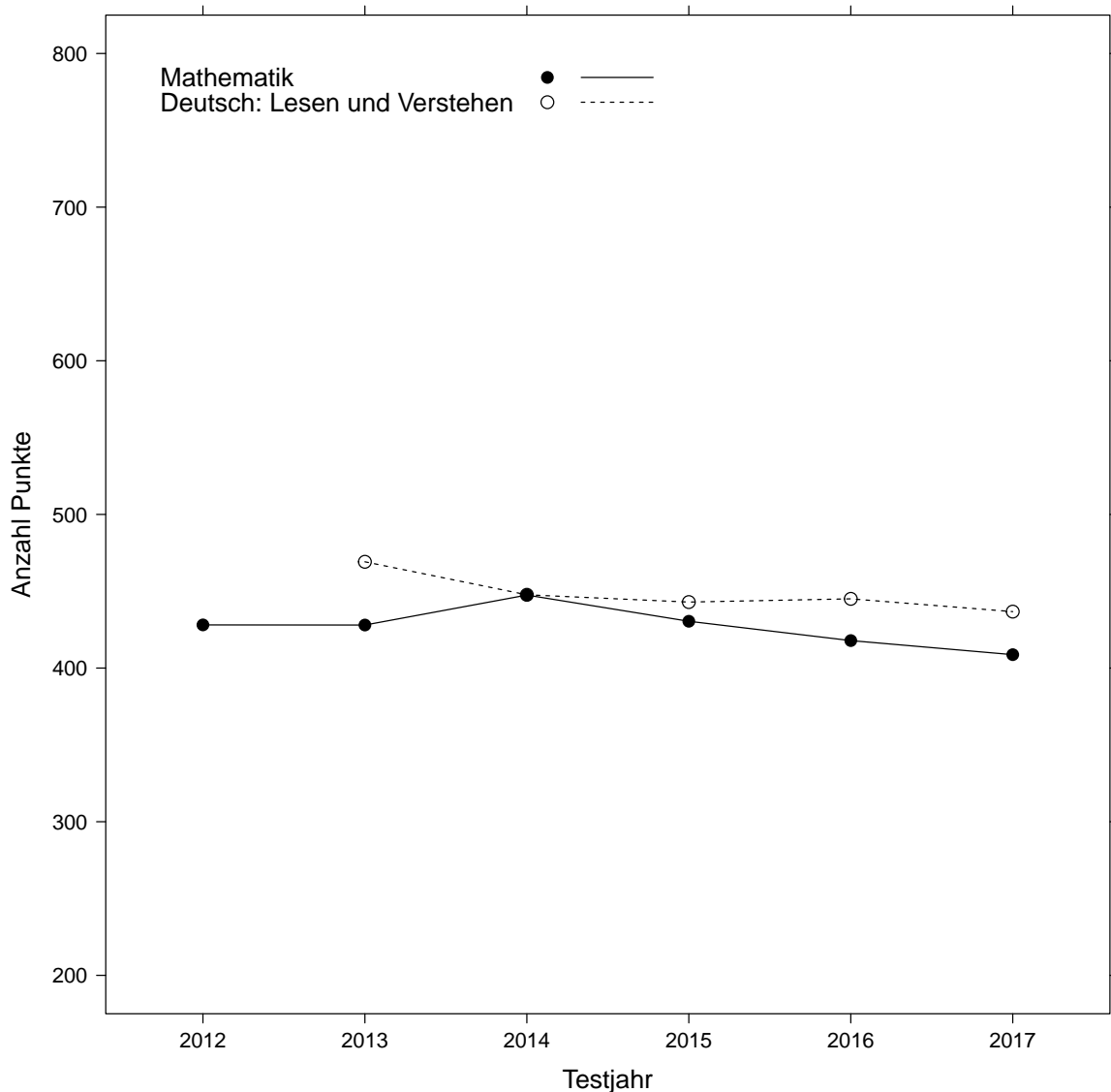


Abbildung 7: Leistungsentwicklung in den Bereichen Mathematik sowie Deutsch *Lesen und Verstehen* im A-Zug; nur Regelklassen; Daten zur Abbildung siehe Anhang

Schülerinnen und Schüler, die in Mathematik 400 oder weniger Punkte erreichen, verfügen noch nicht über die notwendigen mathematischen Grundkompetenzen. Diese Schülerinnen und Schüler sind noch nicht in der Lage, einfache Grundoperationen zu lösen. 2017 erreichten 22 Prozent der Schülerinnen und Schüler weniger als 400 Punkte.

Die besten Schülerinnen und Schüler erreichten in Mathematik mehr als 600 Punkte. Sie waren in der Lage, Bruchgleichungen zu lösen und Folgerungen aus grafischen Darstellungen statistischer Daten zu ziehen. Der Anteil Schülerinnen und Schüler, die mehr als 600 Punkte erreichten, lag 2017 bei 26 Prozent.

Tabelle 7 zeigt die Verteilung aller Schülerinnen und Schüler der Volksschule Basel-Stadt in Deutsch

Lesen und Verstehen nach Intervall.

Tabelle 7: Verteilung der Schülerinnen und Schüler auf die Intervalle in Deutsch *Lesen und Verstehen*; alle Schülerinnen und Schüler des 11. Schuljahres, inklusive SpA

Jahr	Intervall				
	≤ 400	401–500	501–600	601–700	> 700
2015	13%	22%	34%	24%	7%
2016	14%	25%	27%	26%	8%
2017	17%	26%	30%	23%	4%

Schülerinnen und Schüler mit 400 oder weniger Punkten verfügen noch nicht über die notwendigen Grundkompetenzen in *Lesen und Verstehen*. Sie sind nicht in der Lage, einen längeren Text zu verstehen und verschiedene argumentative Zusammenhänge zu erfassen. 2017 erreichten 17 Prozent der Schülerinnen und Schüler in *Lesen und Verstehen* weniger als 400 Punkte.

Die besten Schülerinnen und Schüler erreichten in *Lesen und Verstehen* mehr als 600 Punkte. Sie verstanden komplexere Texte, konnten Folgerungen aus einem Text ableiten und diese anhand von Textstellen begründen. 2017 erreichten 27 Prozent der Schülerinnen und Schüler mehr als 600 Punkte.

7 Fazit

Insgesamt 855 Schülerinnen und Schüler aus 57 Klassen nahmen 2017 an der Schlussprüfung teil. Die meisten dieser Schülerinnen und Schüler ($n = 773$) gehörten einer Regelklasse der Weiterbildungsschule an, die übrigen verteilten sich auf Fremdsprachenklassen und Spezialangebote.

Wie die Ergebnisse der Schlussprüfung zeigen, verfügten rund 27 Prozent der Schülerinnen und Schüler über die Fähigkeit, in Deutsch *Lesen und Verstehen* einen vergleichsweise komplexen Text zu verstehen, im Text argumentative Zusammenhänge zu erfassen und daraus inhaltlich begründbare Schlüsse abzuleiten. Rund 17 Prozent der Schülerinnen und Schüler waren am Ende des 11. Schuljahres hingegen nicht in der Lage, einen einfacheren Text so weit zu verstehen, dass sie ihn beispielsweise in eigenen Worten hätten zusammenfassen können.

In Mathematik waren 26 Prozent der Schülerinnen und Schüler in der Lage, unter anderem lineare Gleichungen zu lösen oder die Fläche von Drei- und Vierecken zu berechnen. 22 Prozent der Schülerinnen und Schüler verfügten am Ende des 11. Schuljahres hingegen noch nicht über die notwendigen mathematischen Fähigkeiten, um einfache Rechnungen mit Grundoperationen zu lösen.

Vergleicht man die individuellen Leistungen der Schülerinnen und Schüler in beiden Leistungszügen, so bestätigen sich die Ergebnisse der Schlussprüfung 2016. Sowohl in *Lesen und Verstehen* als auch in Mathematik bestanden klare Differenzen zwischen den Leistungen im A-Zug und den Leistungen im E-Zug.

Im aktuellen Jahr zeigten Schülerinnen und Schüler in beiden geprüften Fächern eine gefingfügige Tendenz zu einer verminderten Leistung im Vergleich zum Vorjahr (siehe z.B. Tabellen 6, 7). Diese Tendenz lässt sich teilweise mit der veränderten Zusammensetzung der Leistungsniveaus erklären. Während der Anteil der Schülerinnen und Schüler im E-Zug 2017 im Vergleich zum Vorjahr um 7 Prozent sank (2016: 59 %, 2017: 52 %), stieg der Anteil der Schülerinnen und Schüler im A-Zug (+ 2 %), in Fremdsprachenklassen (+ 4 %) und in den Spezialangeboten (+ 1 %). Dies könnte auf eine veränderte

Zusammensetzung der Schülerpopulation hindeuten.

Zugleich ist diese Tendenz auch innerhalb des A- beziehungsweise E-Zugs erkennbar. Die geringfügig verminderte Leistung beruht direkt auf einer geringeren Anzahl korrekt gelöster Aufgaben. Bezogen auf die 51 Mathematik-Aufgaben beispielsweise, welche zwischen 2016 und 2017 identisch geblieben sind, lösten die Schülerinnen und Schüler des A-Zugs durchschnittlich 1.2 Aufgaben weniger als im Vorjahr. Die Schülerinnen und Schüler des E-Zugs lösten durchschnittlich 1.5 Aufgaben weniger als im Vorjahr. Während diese Tendenz im E-Zug und in Deutsch über die Jahre hinweg sehr schwach ausgeprägt ist, lässt sie sich in den Leistungen des A-Zugs im Fach Mathematik bis ins Jahr 2014 zurückverfolgen.

Der vorliegende Bericht schliesst das Projekt Schlussprüfung Basel-Stadt ab. Zusammen mit den Kantonen Aargau, Basel-Landschaft und Solothurn führt der Kanton Basel-Stadt im Rahmen des Bildungsraums Nordwestschweiz den Check S3 flächendeckend ein. Der Check S3 ersetzt die Schlussprüfung und enthält auch Rückmeldungen zu Deutsch Sprache im Fokus sowie zu Fremdsprachen.

8 Anhang

8.1 Glossar der statistischen Begriffe und Abkürzungen

Effektgrösse – Zur Interpretation Unterschieden wird die Effektgrösse d berechnet, indem die Differenz der Mittelwerte durch die Standardabweichung dividiert wird. Die Effektgrösse macht die Grösse von Unterschieden unabhängig von den verwendeten Massen vergleichbar. Eine Effektgrösse von $d = 0.2$ weist auf einen schwachen Unterschied hin, eine Effektgrösse von $d = 0.5$ auf einen mittleren Unterschied und eine Effektgrösse von $d = 0.8$ auf einen starken Unterschied. Auf der standardisierten WBS-Skala können dementsprechend Differenzen von 20 Punkten als klein, solche von 50 Punkten als mittelstark (oder deutlich) und solche von 80 Punkten als sehr gross bezeichnet werden.

Mittelwert – Der Mittelwert entspricht dem arithmetischen Mittel aller Einzelwerte. Der Mittelwert wird im Bericht mit M abgekürzt.

N, n – Die Anzahl Schülerinnen und Schüler der Gesamtstichprobe wird mit N , die Anzahl Schülerinnen und Schüler einer Teilstichprobe mit n abgekürzt.

Signifikantes Ergebnis – Ein Ergebnis (Unterschied oder Zusammenhang) ist statistisch signifikant, wenn es es sehr unwahrscheinlich ist, dass ein Ergebnis dieser Grösse zufällig zustande gekommen sein könnte. Die Irrtumswahrscheinlichkeit α bzw. p gibt die Wahrscheinlichkeit an, dass in einer Stichprobe ein Unterschied festgestellt wird, obwohl in der Grundgesamtheit kein solcher Unterschied vorhanden ist. In diesem Bericht erfüllen signifikante Ergebnisse die Bedingung $p < 0.05$. Die Signifikanz eines Ergebnisses sagt nichts über dessen Relevanz aus. Um die Relevanz signifikanter Ergebnisse zu beurteilen, wird daher die Effektgrösse d berechnet (siehe oben).

Standardabweichung – Die Standardabweichung ist ein quantitatives Mass für die Streuung der Einzelwerte um den Mittelwert. Entspricht die Verteilung der Einzelwerte einer Normalverteilung, dann besitzt die Standardabweichung die Eigenschaft, dass rund zwei Drittel (68 Prozent) der Einzelwerte höchstens eine Standardabweichung über oder unter dem Mittelwert [$M \pm 1 SD$] liegen. Wird der Bereich um je eine Standardabweichung erweitert [$M \pm 2 SD$], dann befinden sich darin rund 95 Prozent der Einzelwerte.

8.2 Daten zu den Abbildungen 6 und 7

Zu Abbildung 6: Leistungsentwicklung in den Bereichen Mathematik und Deutsch *Lesen und Verstehen* im E-Zug

Mathematik				
Testjahr	<i>n</i>	<i>M</i>	<i>SE</i>	<i>SD</i>
2012	411	577	4.6	94
2013	409	569	4.8	96
2014	410	599	4.6	93
2015	427	587	4.9	101
2016	447	596	4.7	100
2017	384	581	5.1	100

Deutsch <i>Lesen und Verstehen</i>				
Testjahr	<i>n</i>	<i>M</i>	<i>SE</i>	<i>SD</i>
2013	409	577	4.2	84
2014	410	582	4.7	96
2015	427	578	4.3	89
2016	454	576	4.0	86
2017	384	567	4.2	82

Zu Abbildung 7: Leistungsentwicklung in Mathematik sowie Deutsch *Lesen und Verstehen* im A-Zug

Mathematik				
Testjahr	<i>n</i>	<i>M</i>	<i>SE</i>	<i>SD</i>
2012	296	428	4.4	77
2013	343	428	5.7	106
2014	367	448	5.3	102
2015	313	430	5.1	91
2016	305	418	5.6	97
2017	316	409	5.4	96

Deutsch <i>Lesen und Verstehen</i>				
Testjahr	<i>n</i>	<i>M</i>	<i>SE</i>	<i>SD</i>
2013	341	469	5.2	95
2014	368	448	5.9	112
2015	314	443	5.1	90
2016	318	445	5.2	93
2017	318	437	5.3	95