



Universität Zürich
Institut für Bildungsevaluation

Institut für Bildungsevaluation
Assoziiertes Institut
der Universität Zürich

Entwicklung und Durchführung der Abschlussprüfungen

Schlussbericht 2008

zuhanden des Rektorats der Weiterbildungsschule Basel-Stadt (WBS)

Florian Keller & Urs Moser
Zürich, August 2008

Inhalt

1	Einleitung	3
2	Das Wichtigste in Kürze	5
3	Fachleistungen der Schülerinnen und Schüler	7
3.1	Mathematikleistungen am Ende der 9. Klasse	7
3.2	Deutschleistungen am Ende der 9. Klasse.....	9
3.3	Leistungsunterschiede zwischen Mädchen und Knaben am Ende der 9. Klasse.....	11
4	Fachleistungen nach Klassen	12
4.1	Mathematik- und Deutschleistungen nach Klassen	12
4.2	Beurteilung der Unterschiede zwischen den Klassen	14
5	Fachleistungen nach Schulen.....	15
6	Leistungsentwicklung an der WBS.....	16
6.1	Durchschnittliche Leistungen der Schülerinnen und Schüler an der WBS....	16
6.2	Fähigkeiten der Schülerinnen und Schüler an der WBS.....	18
6.3	Leistungsentwicklung an den einzelnen Schulen der WBS.....	20
7	Fazit.....	22
	Anhang: Glossar der statistischen Begriffe	23

1 Einleitung

Im Auftrag des Rektorats der Weiterbildungsschule Basel-Stadt (WBS) führte das Institut für Bildungsevaluation nach 2007 zum zweiten Mal die Abschlussprüfungen an der WBS durch.

Der Auftrag umfasst drei Aufgabenbereiche: Erstens die Entwicklung neuer Prüfungsaufgaben und die Produktion der Prüfungshefte sowie die Korrektur der Prüfungen und die Berechnung der Noten für die Schülerinnen und Schüler. Zweitens die Ergebnisrückmeldungen, in denen den Lehrpersonen die Resultate ihrer Schülerinnen und Schüler sowie die Klassenmittelwerte aller WBS-Klassen zurückgemeldet werden. Die Lehrpersonen erhalten damit eine unabhängige Standortbestimmung über den Leistungsstand ihrer Klasse. Drittens wird jedes Jahr ein Schlussbericht zuhanden des Rektorats der Weiterbildungsschule Basel-Stadt erstellt. Im vorliegenden Schlussbericht 2008 werden die Leistungen der Schülerinnen und Schüler, der Klassen und der Schulen der WBS zusammengefasst und in einer Gesamtsicht dargestellt. Zudem werden die Resultate mit den Resultaten der Abschlussprüfungen aus den Jahren 2005 bis 2007 verglichen. Damit dient der jährliche Schlussbericht auch als ein Instrument der Qualitätssicherung beziehungsweise als Teil des Bildungsmonitorings.

2007 wurde zum ersten Mal ein solcher Schlussbericht zu den Abschlussprüfungen der WBS vorgelegt¹. Darin konnte unter anderem aufgezeigt werden, dass sich die durchschnittlichen Leistungen seit 2005 in Deutsch kontinuierlich verbessern, in der Mathematik aber nahezu stagnieren und viele Schülerinnen und Schüler die Ziele des Lehrplans nicht erreichen. Aufgrund dieser Resultate ergriff das Rektorat der WBS unter dem Begriff «Mathepaket» folgende Massnahmen zu einer messbaren Verbesserung der Mathematikleistungen:

1. Erhöhung der Stundendotation von 4 auf 5 Stunden (E-Zug) beziehungsweise von 5 auf 6 Stunden (A-Zug).
2. Obligatorische Durchführung einer Lernstandserhebung zu Beginn der 1. Klasse der WBS.
3. Formulierung klarer Erwartungen für die Schlussprüfung bezüglich Vorbereitung und Resultat durch die Schulleitung.
4. Anforderungen der Schlussprüfungen werden den Lehrpersonen im Rahmen einer speziellen Fachkonferenz transparent gemacht und mit Parallelaufgaben veranschaulicht.
5. Zusammenstellung und Ausarbeitung von zusätzlichem Übungsmaterial und Aufgabensammlungen im Hinblick auf die Schlussprüfung.
6. Orientierung am Kompetenzraster, das im Rahmen des Projekts HarmoS umschrieben wird.

¹ Moser, U. & Keller, F. (2007). *Entwicklung und Durchführung der Abschlussprüfungen. Schlussbericht 2007 zuhanden des Rektorats der Weiterbildungsschule Basel-Stadt (WBS)*. Zürich: Institut für Bildungsevaluation. (www.ibe.uzh.ch/entwicklung/wbsabschluss.html)

Die Punkte 1, 2 und 5 dieses Massnahmepakets werden ab dem kommenden Schuljahr 2008/09 umgesetzt, während der sechste Punkt eher einer längerfristigen Zielformulierung gleichkommt. Bereits realisiert hingegen wurden die Punkte 3 und 4, die Formulierung klarer Erwartungen für die Abschlussprüfung und die Kommunikation der in den Abschlussprüfungen geforderten Fähigkeiten. Mit dem vorliegenden Bericht wird nun unter anderem überprüft, ob und inwieweit die umgesetzten Massnahmen tatsächlich eine Steigerung der Mathematikleistungen bewirkt haben.

Der Bericht ist analog dem Schlussbericht 2007 in sieben Kapitel gegliedert. In Kapitel 2 werden die Durchführung der Abschlussprüfungen sowie die Methoden der Ergebnisanalyse und -darstellung beschrieben. In Kapitel 3 werden einerseits die Leistungen der Schülerinnen und Schüler in der Abschlussprüfung 2008 dargestellt. Andererseits wird inhaltlich beschrieben, über welche Kompetenzen die Schülerinnen und Schüler am Ende der 9. Klasse verfügen. Im vierten Kapitel sind die Ergebnisse der einzelnen Klassen und im fünften Kapitel die Ergebnisse der Schulen dargestellt. In Kapitel 6 wird die Entwicklung der durchschnittlichen Leistungen der Schülerinnen und Schüler sowie der Schulen in den Jahren 2005 bis 2008 beschrieben und es wird auf erste Trends hingewiesen. Im siebten Kapitel werden die wichtigsten Resultate dieses Berichts zusammengefasst.

2 Das Wichtigste in Kürze

Testinstrumente – Die Testinstrumente, die als Abschlussprüfungen am Ende der 2. Klasse der WBS eingesetzt werden, basieren auf Leistungstests, die 2004 im Rahmen der Evaluation der Strukturänderung an der WBS am Institut für Bildungsevaluation entwickelt wurden. Seither werden die Testaufgaben jedes Jahr überarbeitet, erneuert und verbessert.

Der Mathematiktest 2008 umfasst 82 Testaufgaben aus den Lehrplanbereichen «Zahlen und Zahlenoperationen», «Funktionen», «Gleichungen», «Geometrie» und «Stochastik».

Der Deutschtest 2008 umfasst 121 Testaufgaben aus den Lehrplanbereichen «Leseverständnis» und «Sprachbetrachtung». Der Lehrplanbereich «Schreiben» wurde mit einem Schreibenanlass geprüft. Drei Themen standen dabei zur Auswahl: «Ich war Zeuge», «Wie ich mich kleide» und «Das Handy – Fluch oder Segen unserer Zeit?».

Die Tests sind so konzipiert, dass alle Schülerinnen und Schüler unabhängig vom besuchten Klassenzug die gleiche Prüfung lösen können und dass die Prüfungsergebnisse mit dem Referenzrahmen der Evaluation der Strukturänderung an der WBS empirisch verglichen werden können.

Teilnahme – An den Abschlussprüfungen 2008 nahmen 908 Schülerinnen und Schüler aus 52 Klassen teil. 379 Schülerinnen und Schüler (42 Prozent) wurden in einer Regelklasse des A-Zugs, 463 (51 Prozent) in einer Regelklasse des E-Zugs unterrichtet. 66 Schülerinnen und Schüler (rund 7 Prozent) besuchten eine Musikklasse, eine Sportklasse oder die Fremdsprachenklasse.

Durchführung – Die Durchführung der Abschlussprüfungen fand im Juni 2008 statt und wurde von den sechs Weiterbildungsschulen selbstständig organisiert. Die Prüfungshefte wurden anschliessend ans Institut für Bildungsevaluation nach Zürich gebracht, wo sie von mehreren Fachwissenschaftlern mit Unterrichtserfahrung korrigiert wurden. Die Aufsätze des Deutschtests (Lehrplanbereich «Schreiben») wurden von drei Germanistinnen mit Unterrichtserfahrung anhand eines standardisierten Korrekturrasters nach zehn inhaltlichen und zehn formalen Kriterien bewertet. Nach der Korrektur wurden die Aufgaben elektronisch erfasst und plausibilisiert. Zwei Wochen nach der Durchführung der Abschlussprüfungen standen die Prüfungsergebnisse der Schülerinnen und Schüler in Form von Noten zur Verfügung. Ende Juni wurden zudem die Prüfungsergebnisse der Schülerinnen und Schüler als prozentualer Anteil richtig gelöster Aufgaben sowie die Klassenmittelwerte berechnet und den jeweiligen Lehrpersonen zugestellt.

Skalierung der Leistungsdaten – Die Leistungen der Schülerinnen und Schüler in der Abschlussprüfung werden als Punktzahl auf der WBS-Skala ausgewiesen. Die Punktzahl wurde mittels einer Rasch-Skalierung unter Einbezug der Parameter der Abschlussprüfung 2005 berechnet. Diese Methode ermöglicht es, die Leistungen der Schülerinnen und Schüler in den Jahren 2005 bis 2008 als Punktzahl auf der gleichen Skala zu verorten. Anschliessend wurde die Skala so standardisiert, dass der Mittelwert aller Schülerinnen und Schüler der Abschlussprüfung 2005 in jedem Fach 500 Punkte und die Standardabweichung 100 Punkte beträgt. Veränderungen der durchschnittlichen Schülerleis-

tungen sowie die Leistungsentwicklung an den einzelnen Schulen können somit einfach interpretiert werden: Werte über 500 Punkte entsprechen einer höheren Leistung im Vergleich zur Abschlussprüfung 2005, Werte unter 500 Punkte einer tieferen Leistung. Leistungsunterschiede – beispielsweise zwischen Mädchen und Knaben oder zwischen zwei Prüfungsjahren – sind dann relevant, wenn sie 20 Punkte oder mehr betragen. Unterschiede von 80 und mehr Punkten sind sehr gross.

Ein weiterer wesentlicher Vorteil dieser Skalierung besteht darin, dass neben den Leistungen der Schülerinnen und Schüler auch die Schwierigkeit der Prüfungsaufgaben auf der gleichen Skala mit einer einheitlichen Metrik abgebildet werden kann. 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 62 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 zusammengefasst.

Intervalle – Für die inhaltliche Beschreibung wurden die Leistungen der Schülerinnen und Schüler in Intervalle von jeweils 100 Punkten auf der WBS-Skala eingeteilt. Schülerinnen und Schüler, die ein bestimmtes Intervall erreichen, können mindestens die Hälfte der Aufgaben in diesem Intervall richtig lösen. Erreicht eine Schülerin beispielsweise im Mathematiktest 401 Punkte, dann beträgt die durchschnittliche Lösungswahrscheinlichkeit für die Aufgaben des Intervalls «401 bis 500 Punkte» 50 Prozent. Das heisst, die Schülerin löst vermutlich die Hälfte der Aufgaben dieses Intervalls richtig.

Weiterführende 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 Weiterbildungsschule 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.

Zu beziehen im Internet unter: <http://www.ibe.uzh.ch/downloads/berichte.html>

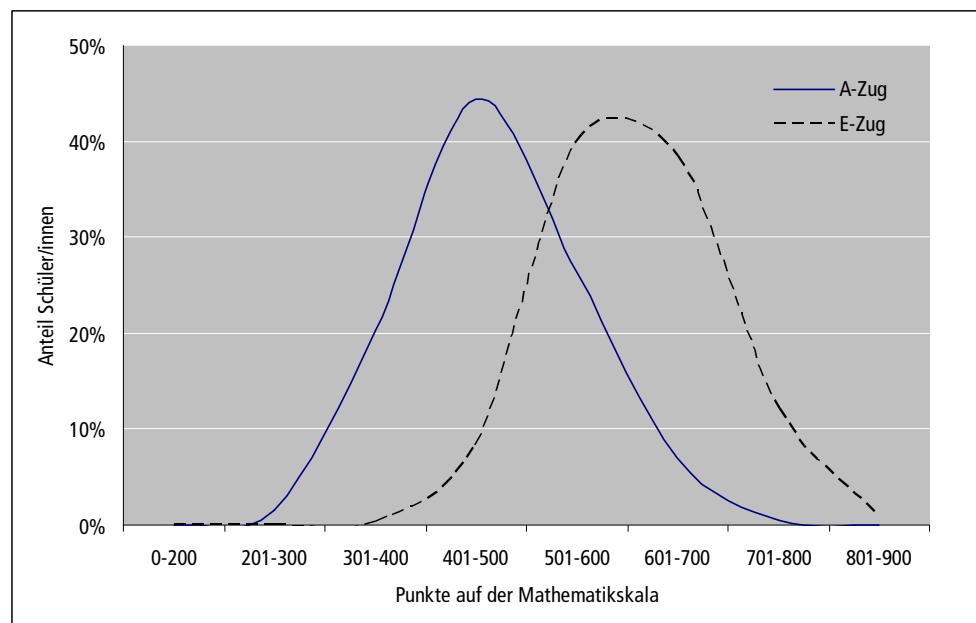
3 Fachleistungen der Schülerinnen und Schüler

3.1 Mathematikleistungen am Ende der 9. Klasse

Abbildung 3.1 zeigt die Verteilung der Mathematikleistungen der Schülerinnen und Schüler in der Abschlussprüfung 2008 nach Leistungszug.

Die Schülerinnen und Schüler des A-Zugs erreichen am Ende der 9. Klasse durchschnittlich 469 Punkte auf der Mathematikskala, die Schülerinnen und Schüler des E-Zugs 606 Punkte. Die Differenz zwischen den durchschnittlichen Mathematikleistungen der Schülerinnen und Schüler des A- und des E-Zugs beträgt 137 Punkte.

Abbildung 3.1 Verteilung der Mathematikleistungen 2008

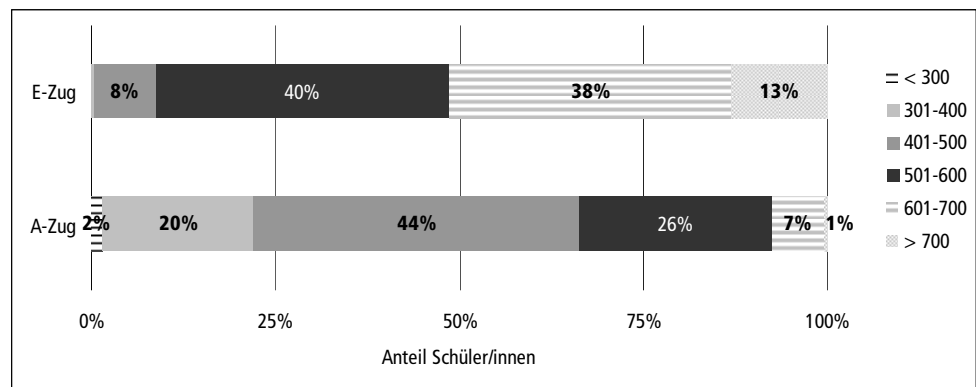


Anmerkungen: ohne Schülerinnen und Schüler der Fremdsprachenklasse
A-Zug: N = 395; M = 469 Punkte; SD = 85 Punkte
E-Zug: N = 502; M = 606 Punkte; SD = 80 Punkte

Trotz der grossen Differenz der Mittelwerte (137 Punkte) überschneiden sich die Leistungsverteilungen der beiden Züge. Rund 6 Prozent der Schülerinnen und Schüler aus dem A-Zug erzielen in der Mathematik eine höhere Leistung als der Durchschnitt der Schülerinnen und Schüler aus dem E-Zug. Umgekehrt erreichen rund 2 Prozent der Schülerinnen und Schüler aus dem E-Zug tiefere Mathematikleistungen als der Durchschnitt der Schülerinnen und Schüler des A-Zugs. Der Überschneidungsbereich der beiden Kurven beträgt rund 41 Prozent. Das heisst, 365 Schülerinnen und Schüler können aufgrund ihrer Mathematikleistungen nicht eindeutig dem A- oder dem E-Zug zugewiesen werden.

Abbildung 3.2 zeigt die Verteilung der Schülerinnen und Schüler am Ende der 9. Klasse auf die Intervalle in der Mathematik nach Leistungszug. 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.

Abbildung 3.2 Verteilung der Schülerinnen und Schüler nach Intervall in der Mathematik (ohne Fremdsprachenklasse)



Rund 22 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS weniger als 400 Punkte auf der Mathematikskaala. Ihnen fehlen in der Mathematik Grundkompetenzen, die zur Lösung von einfachen Grundoperationen notwendig sind.

Alle Schülerinnen und Schüler des E-Zugs sowie rund 78 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 400 Punkte. Sie kennen unter anderem Zahlenarten und Zahlenmengen sowie die wichtigsten Grundoperationen und können sie auch anwenden.

Rund 91 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 34 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 500 Punkte. Sie können unter anderem einfache Prozentrechnungen lösen sowie Grundoperationen mit Dezimalzahlen durchführen.

Rund 51 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 8 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 600 Punkte. Sie können unter anderem Bruchgleichungen lösen und Folgerungen aus grafischen Darstellungen von statistischen Daten ziehen.

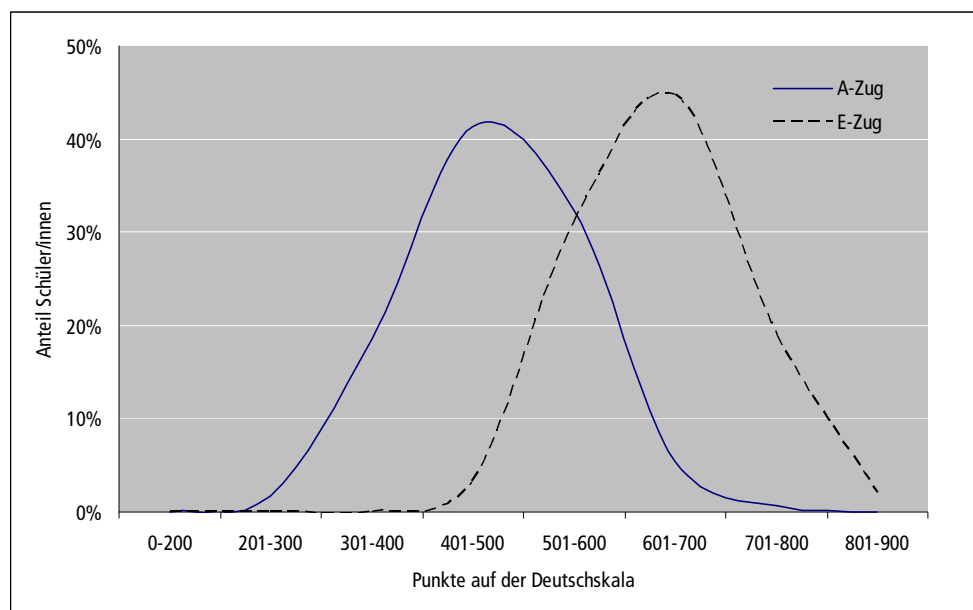
Rund 13 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 1 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 700 Punkte. Sie können unter anderem Potenzen und Quadratwurzeln anwenden sowie Rauminhalte und Flächen von geometrischen Körpern berechnen.

3.2 Deutschleistungen am Ende der 9. Klasse

Abbildung 3.3 zeigt die Verteilung der Deutschleistungen der Schülerinnen und Schüler in der Abschlussprüfung 2008 nach Leistungszug.

Die Schülerinnen und Schüler des A-Zugs erreichen am Ende der 9. Klasse durchschnittlich 472 Punkte auf der Deutschskala, die Schülerinnen und Schüler des E-Zugs 634 Punkte. Die Differenz zwischen den durchschnittlichen Deutschleistungen der Schülerinnen und Schüler des E- und des A-Zugs beträgt 162 Punkte.

Abbildung 3.3 Verteilung der Deutschleistungen 2008

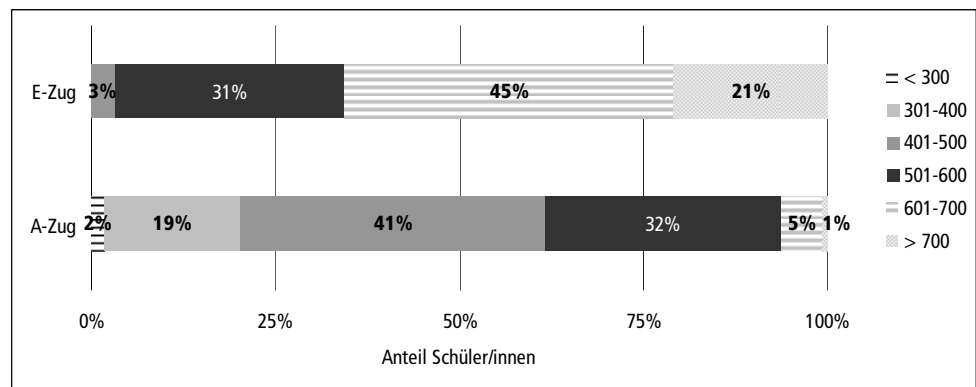


Anmerkungen: ohne Schülerinnen und Schüler der Fremdsprachenklasse
A-Zug: N = 393; M = 472 Punkte; SD = 83 Punkte
E-Zug: N = 502; M = 634 Punkte; SD = 78 Punkte

Die Leistungsunterschiede zwischen den Schülerinnen und Schülern des A- und des E-Zugs sind in Deutsch zwar grösser als in der Mathematik. Trotzdem überschneiden sich die Leistungsverteilungen der beiden Züge auch in Deutsch. 9 Schülerinnen und Schüler aus dem A-Zug (rund 2 Prozent) erzielen in Deutsch eine höhere Leistung als der Durchschnitt der Schülerinnen und Schüler aus dem E-Zug. Umgekehrt erreichen 3 Schülerinnen und Schüler aus dem E-Zug (rund 1 Prozent) tiefere Deutschleistungen als der Durchschnitt der Schülerinnen und Schüler des A-Zugs. Der Überschneidungsbereich der beiden Kurven beträgt rund 32 Prozent. Das heisst, 284 Schülerinnen und Schüler können aufgrund ihrer Deutschleistungen nicht eindeutig dem A- oder dem E-Zug zugewiesen werden.

Abbildung 3.4 zeigt die Verteilung der Schülerinnen und Schüler am Ende der 9. Klasse auf die Intervalle in Deutsch nach Leistungszug. 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.

Abbildung 3.4 Verteilung der Schülerinnen und Schüler nach Intervall in Deutsch (ohne Fremdsprachenklasse)



Alle Schülerinnen und Schüler des E-Zugs sowie rund 98 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 300 Punkte auf der Deutschskala. Sie können den Inhalt eines einfacheren Sachtextes verstehen und ihm die wichtigsten Informationen entnehmen.

Alle Schülerinnen und Schüler des E-Zugs sowie rund 79 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 400 Punkte. Sie kennen unter anderem die wichtigsten Zeitformen und können diese in einem Text erkennen. Zudem sind sie fähig, zwischen den Informationen im Text Beziehungen herzustellen.

Rund 97 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 38 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 500 Punkte. Sie können unter anderem die Zeitformen von Verben erkennen und richtig anwenden sowie die inhaltliche Intention eines Textes erkennen.

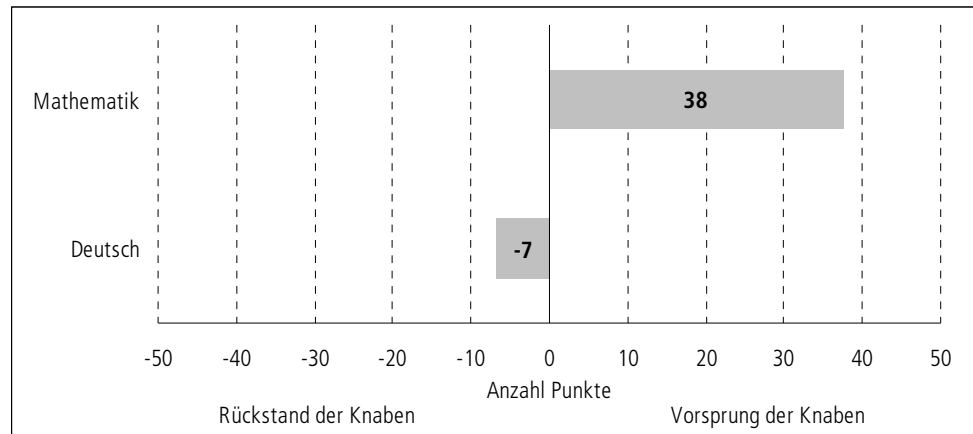
Rund 66 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 6 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 600 Punkte. Sie verstehen unter anderem komplexere Texte und können Zeitformen, Fälle und Pronomen richtig anwenden.

Rund 21 Prozent der Schülerinnen und Schüler des E-Zugs sowie rund 1 Prozent der Schülerinnen und Schüler des A-Zugs erreichen am Ende der WBS mehr als 700 Punkte. Sie sind unter anderem in der Lage, in einen Text eingebettete Informationen zu finden und sprachliche Nuancen zu verstehen. Zudem können sie den Sinn schwieriger Fremdwörter aus dem kontextualen Zusammenhang herleiten.

3.3 Leistungsunterschiede zwischen Mädchen und Knaben am Ende der 9. Klasse

Abbildung 3.5 zeigt die Leistungsunterschiede zwischen Mädchen und Knaben am Ende der 9. Klasse als Punkte auf der WBS-Skala.

Abbildung 3.5 Leistungsunterschiede zwischen Mädchen und Knaben in Mathematik und in Deutsch (ohne Fremdsprachenklasse)



Anmerkungen: ohne Schülerinnen und Schüler der Fremdsprachenklasse
Mathematik:
Mädchen: N = 428; M = 526 Punkte; SD = 104 Punkte
Knaben: N = 469; M = 564 Punkte; SD = 106 Punkte
Deutsch:
Mädchen: N = 427; M = 566 Punkte; SD = 111 Punkte
Knaben: N = 468; M = 559 Punkte; SD = 116 Punkte

In der Mathematik erreichen Knaben durchschnittlich 38 Punkte mehr als Mädchen. Dieser Unterschied ist statistisch signifikant und mit einer mittleren Effektstärke von $d = 0,35$ deutlich.

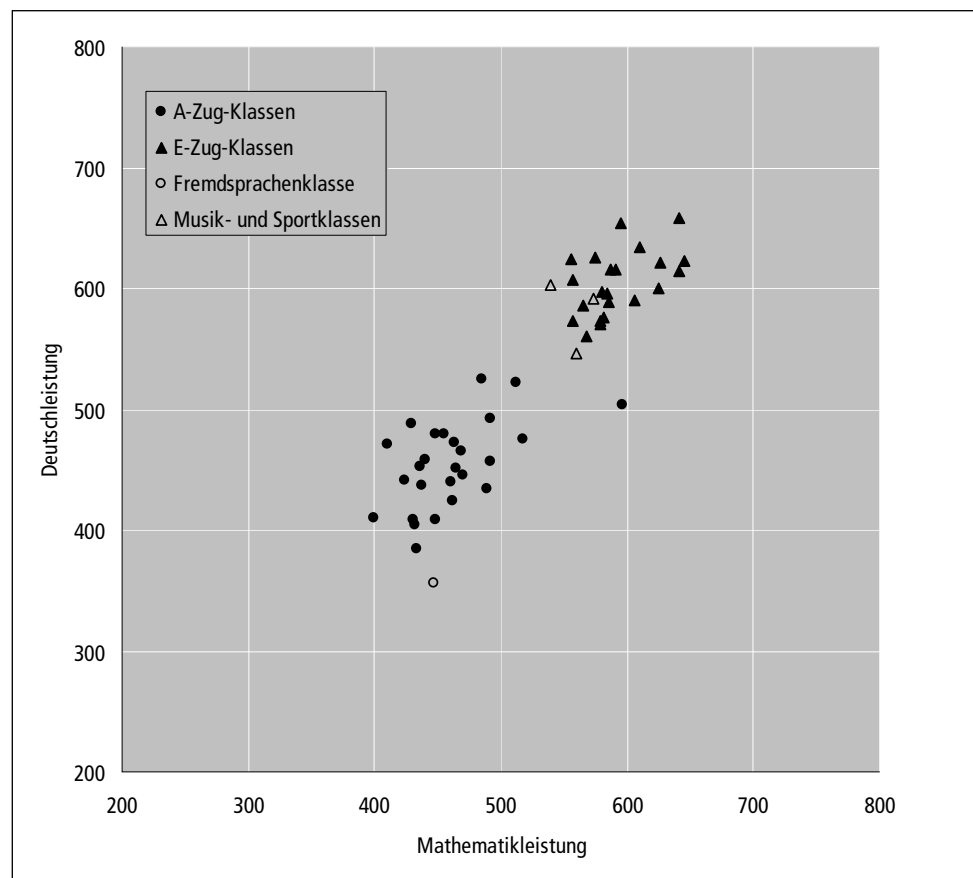
In Deutsch erreichen Knaben durchschnittlich 7 Punkte weniger als Mädchen. Dieser Rückstand ist mit einer Effektstärke von $d = 0,06$ unbedeutend. Das heisst, in Deutsch gibt es keine relevanten Leistungsunterschiede zwischen Mädchen und Knaben.

4 Fachleistungen nach Klassen

4.1 Mathematik- und Deutschleistungen nach Klassen

Abbildung 4.1 zeigt die Ergebnisse der beteiligten Klassen in Mathematik und Deutsch. Die Ergebnisse der Regelklassen des A-Zugs werden als Punkte, die Ergebnisse der Regelklassen des E-Zugs als Dreiecke dargestellt. Jeder Punkt beziehungsweise jedes Dreieck steht für eine Klasse. Die Position der Punkte und Dreiecke ergibt sich aus der durchschnittlichen Punktzahl der Schülerinnen und Schüler einer Klasse.

Abbildung 4.1 Fachleistungen in Deutsch und Mathematik nach Klassen



Anmerkungen: A-Zug-Klassen: Mathematik: N = 26 Klassen; M = 462 Punkte; SD = 40 Punkte
Deutsch: N = 26 Klassen; M = 455 Punkte; SD = 36 Punkte
E-Zug-Klassen: Mathematik: N = 22 Klassen; M = 592 Punkte; SD = 28 Punkte
Deutsch: N = 22 Klassen; M = 605 Punkte; SD = 27 Punkte

Die durchschnittlichen Leistungen in den Klassen des A- und des E-Zugs unterscheiden sich deutlich voneinander. Kein Klassenmittelwert einer E-Klasse liegt in der Mathematik oder in Deutsch unter 550 Punkten, während die meisten Klassen des A-Zugs sowohl in der Mathematik wie auch in Deutsch nicht mehr als 500 Punkte erreichen. Nur jeweils drei A-Klassen erreichen in der Mathematik oder in Deutsch mehr als 500 Punkte.

Die Regelklassen des A-Zugs erreichen in der Mathematik im Durchschnitt 462 Punkte und in Deutsch 455 Punkte. Die Unterschiede zwischen den Klassen sind allerdings sehr gross. In der Mathematik liegt der Klassenmittelwert einer A-Klasse mit 597 Punkten gar über den durchschnittlichen Leistungen der E-Klassen. Die Spannweite zwischen dem höchsten und dem tiefsten Klassenmittelwert beträgt in der Mathematik 197 Punkte und in Deutsch 141 Punkte.

Die Regelklassen des E-Zugs erreichen in der Mathematik im Durchschnitt 592 Punkte und in Deutsch 605 Punkte. Die Streuung zwischen den Klassen des E-Zugs ist deutlich geringer als zwischen den Klassen des A-Zugs. Die Spannweite zwischen dem höchsten und dem tiefsten Klassenmittelwert des E-Zugs beträgt in der Mathematik 89 Punkte und in Deutsch 98 Punkte.

Die Fremdsprachenklasse erbringt in Deutsch deutlich tiefere Leistungen als die Regelklassen des A-Zugs ($M = 357$), unterscheidet sich in der Mathematik aber nicht von den anderen Klassen des A-Zugs ($M = 447$). Die Musik- und Sportklassen unterscheiden sich in ihren Leistungen nicht von den übrigen Regelklassen des E-Zugs.

4.2 Beurteilung der Unterschiede zwischen den Klassen

Die unterschiedlichen Leistungen der Klassen lassen sich auf Merkmale der Schülerinnen und Schüler und auf Merkmale der Klassen zurückführen. Je stärker die Unterschiede zwischen den Klassen variieren, desto eher können sie durch Merkmale der Klasse beziehungsweise durch Merkmale des Unterrichts erklärt werden.

Tabelle 4.1 Varianzen zwischen und innerhalb der Klassen: Mathematik und Deutsch, 2007 und 2008

	Varianz zwischen den Klassen		Varianz zwischen den Schülerinnen und Schülern	
	2007	2008	2007	2008
E-Zug Mathematik	9%	8%	91%	92%
E-Zug Deutsch	9%	7%	91%	93%
A-Zug Mathematik ²	5%	10%	95%	90%
A-Zug Deutsch	3%	15%	97%	85%

Tabelle 4.1 zeigt für Mathematik und für Deutsch die Prozentanteile, die entweder durch individuelle Merkmale der Schülerinnen und Schüler oder durch Unterrichts- und Klassenmerkmale erklärt werden können. Als Vergleichswerte sind zudem die Prozentanteile aus den Abschlussprüfungen 2007 angegeben.

Im E-Zug können 2008 in der Mathematik 8 Prozent der Leistungsunterschiede auf Merkmale des Unterrichts oder der Klasse zurückgeführt werden. 92 Prozent können mit individuellen Unterschieden der Schülerinnen und Schüler erklärt werden. In Deutsch können im E-Zug 7 Prozent der Leistungsunterschiede mit Merkmalen des Unterrichts und 93 Prozent mit Merkmalen der Schülerinnen und Schüler erklärt werden. Im E-Zug gibt es zwischen den Jahren 2007 und 2008 nur marginale Unterschiede.

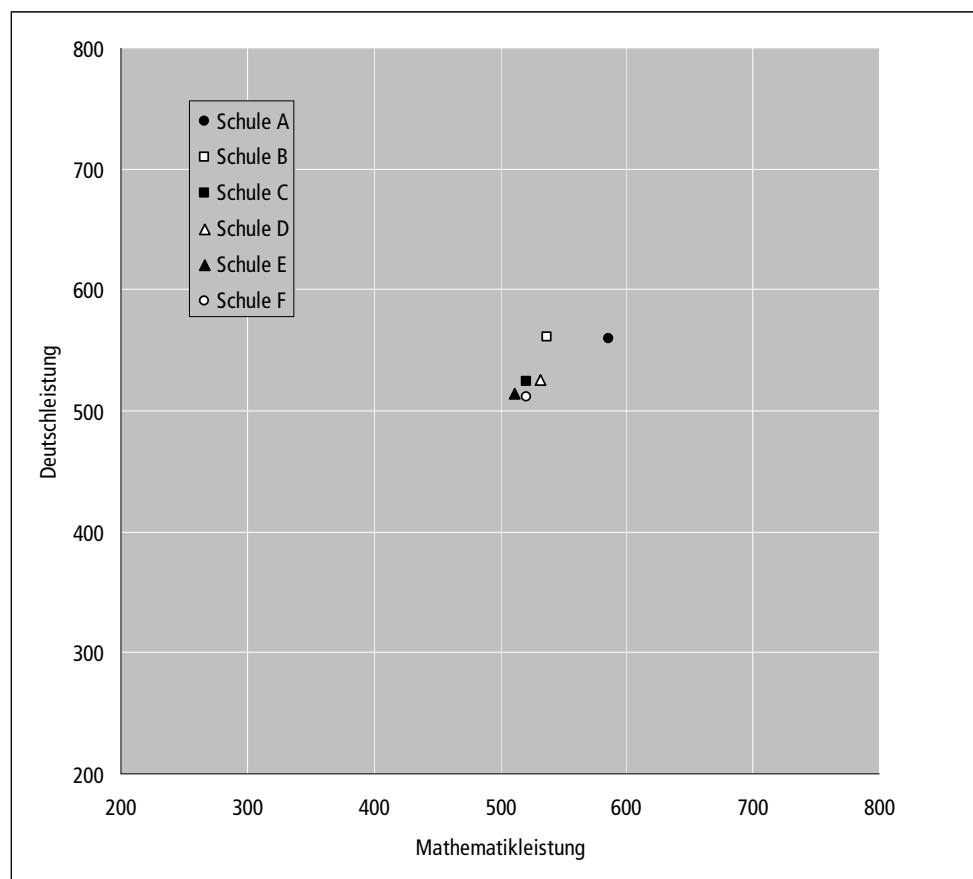
Im A-Zug hingegen können die Leistungsunterschiede viel stärker auf Unterschiede zwischen den Klassen zurückgeführt werden. In der Mathematik können 10 Prozent und in Deutsch 15 Prozent der Leistungsunterschiede mit Merkmalen des Unterrichts erklärt werden. Im Vergleich zu 2007 ist die Varianz zwischen den A-Zug-Klassen sprunghaft angestiegen. Das heisst, ein Teil der A-Klassen erreicht deutlich höhere Leistungen als die übrigen A-Klassen. Das heisst aber auch, dass es für den Erfolg in den Abschlussprüfungen relevant ist, welche Klasse die Schülerinnen und Schüler des A-Zugs besuchen.

² Die Prozentanteile in Tabelle 4.1 wurden für den A-Zug ohne die Klasse berechnet, deren durchschnittliche Mathematikleistung über dem Mittelwert der E-Klassen liegt. Vergleicht man die Varianzen aller A-Klassen miteinander, so können in der Mathematik 23 Prozent und in Deutsch 16 Prozent der Leistungsunterschiede auf Merkmale der Klasse oder des Unterrichts zurückgeführt werden.

5 Fachleistungen nach Schulen

Abbildung 5.1 zeigt die Ergebnisse der Schülerinnen und Schüler der Regelklassen pro Schule. Die Position der Schulmittelwerte ergibt sich aus der durchschnittlichen Punktzahl, die an einer Schule in Mathematik und Deutsch erreicht wurden. Um den Einfluss der unterschiedlichen Schülerschaft auf das Ergebnis einer Schule auszugleichen, wurde das Leistungsniveau der Schülerinnen und Schüler statistisch kontrolliert. Das heißt, die Ergebnisse von Schulen mit einem hohen Anteil Schülerinnen und Schüler des A-Zugs wurden angehoben, die Ergebnisse von Schulen mit einem tiefen Anteil Schülerinnen und Schüler des A-Zugs wurden entsprechend reduziert.

Abbildung 5.1 Fachleistungen in Deutsch und Mathematik nach Schulen
(nach statistischer Kontrolle der Leistungsniveaus)



Wie Abbildung 5.1 zeigt, unterscheiden sich die Leistungen an den Schulen C, D, E und F nur geringfügig. Die durchschnittlichen Leistungen der Schülerinnen und Schüler dieser Schulen befinden sich sowohl in der Mathematik als auch in Deutsch zwischen rund 510 und 530 Punkten.

Statistisch signifikant bessere Leistungen werden an den Schulen A und B erbracht. Die Schülerinnen und Schüler aus dem Schulhaus A erreichen in der Mathematik durch-

schnittlich 585 Punkte. Das sind rund 50 Punkte mehr als der Mittelwert aller Schulen und 73 Punkte mehr als die schwächste Schule. Die Differenz zwischen dem besten und dem schwächsten Schulmittelwert ist statistisch signifikant und mit einer mittleren Effektgrösse von $d = 0,68$ relevant. Auch in Deutsch werden an der Schule A Leistungen erbracht, die rund 30 Punkte über dem Mittelwert aller Schulen liegen.

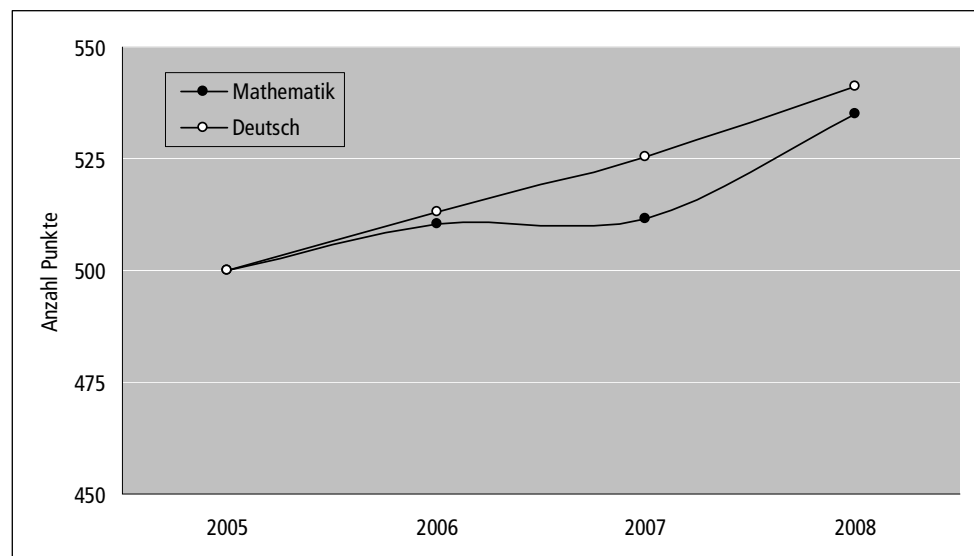
An der Schule B werden mit 561 Punkten die besten durchschnittlichen Deutschleistungen erbracht. Diese Leistung liegt rund 30 Punkte über dem Mittelwert aller Schulen und 50 Punkte über dem schwächsten Schulmittelwert. Dieser Unterschied ist statistisch signifikant (Effektgrösse $d = 0,45$).

6 Leistungsentwicklung an der WBS

6.1 Durchschnittliche Leistungen der Schülerinnen und Schüler an der WBS

Abbildung 6.1 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in den Fächern Mathematik und Deutsch in den Jahren 2005 bis 2008. Die Mittelwerte in Mathematik sind als schwarze Punkte, die Mittelwerte in Deutsch als weisse Punkte dargestellt.

Abbildung 6.1 Leistungsentwicklung in den Fächern Mathematik und Deutsch



Anmerkungen: Mathematik: 2005: $N = 759$; $M = 500$; $SD = 93$;
 2006: $N = 770$; $M = 511$; $SD = 96$;
 2007: $N = 910$; $M = 511$; $SD = 103$;
 2008: $N = 897$; $M = 535$; $SD = 107$;

Deutsch: 2005: $N = 759$; $M = 500$; $SD = 96$;
 2006: $N = 770$; $M = 513$; $SD = 101$;
 2007: $N = 908$; $M = 525$; $SD = 124$;
 2008: $N = 895$; $M = 541$; $SD = 114$;

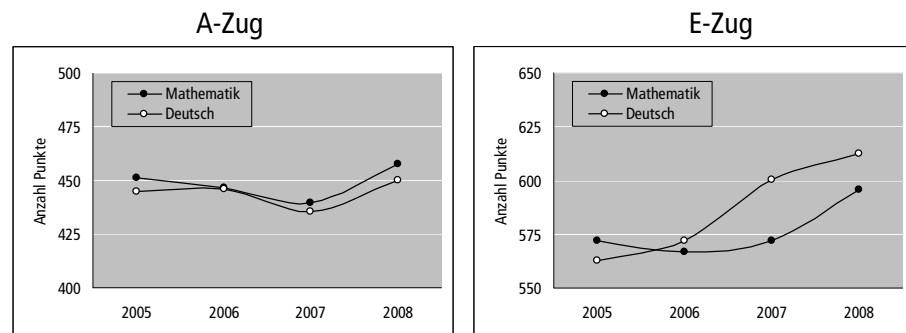
Im Jahr 2005 wurde die Skala für Mathematik und Deutsch auf 500 Punkte festgesetzt. Seither haben sich die Leistungen an der WBS sowohl in Deutsch als auch in der Mathematik leicht erhöht.

In Deutsch erhöhten sich die durchschnittlichen Leistungen seit 2005 von 500 Punkten auf 541 Punkte. Diese Leistungssteigerung ist mit einer mittleren Effektgrösse von $d = 0,41$ statistisch signifikant. Im Vergleich zum Vorjahr konnten 2008 die Deutschleistungen noch einmal leicht um 16 Punkte gesteigert werden.

In der Mathematik konnten die durchschnittlichen Leistungen seit 2005 um 35 Punkte auf 535 Punkte erhöht werden (Effektgrösse $d = 0,35$). Seit 2007 konnten die Mathematikleistungen um 24 Punkte statistisch signifikant gesteigert werden.

Die Abbildung 6.2 zeigt die durchschnittlichen Leistungen der Schülerinnen und Schüler in Mathematik und Deutsch in den Jahren 2005 bis 2008. In der linken Grafik ist die Leistungsentwicklung der Schülerinnen und Schüler des A-Zugs und in der rechten Grafik die Leistungsentwicklung der Schülerinnen und Schüler des E-Zugs dargestellt.

Abbildung 6.2 Leistungsentwicklung in den Fächern Mathematik und Deutsch nach Leistungszug



Anmerkungen: Anteil Schülerinnen und Schüler im E-Zug:
 2005: Mathematik: 40%; Deutsch: 47%
 2006: Mathematik: 53%; Deutsch: 53%
 2007: Mathematik: 54%; Deutsch: 54%
 2008: Mathematik: 56%; Deutsch: 56%

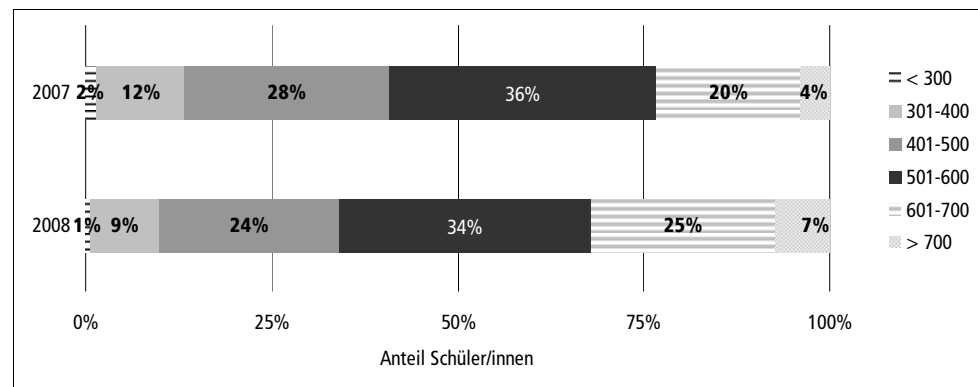
Im A-Zug verlaufen die Leistungskurven in den Fächern Mathematik und Deutsch weitgehend parallel: Nachdem die Leistungen der Schülerinnen und Schüler zwischen 2005 und 2007 kontinuierlich gesunken sind, konnte 2008 eine Trendwende erreicht werden. Die durchschnittlichen Ergebnisse der Schülerinnen und Schüler des A-Zugs sind 2008 so gut wie noch nie seit 2005.

Auch im E-Zug konnte 2008 eine deutliche Leistungssteigerung erzielt werden. In der Mathematik ist der Leistungszuwachs zwischen 2007 und 2008 mit 24 Punkten etwas grösser als im A-Zug (18 Punkte). In Deutsch ist die Leistungssteigerung weniger deutlich und beträgt im E-Zug 12 Punkte (A-Zug: 14 Punkte).

6.2 Fähigkeiten der Schülerinnen und Schüler an der WBS

Neben der Entwicklung der durchschnittlichen Leistungen ist auch von Interesse, ob alle Schülerinnen und Schüler gleichermaßen von der beschriebenen Leistungssteigerung profitieren können. Konkret stellt sich die Frage, ob der Anteil Schülerinnen und Schüler reduziert werden konnte, der in der Mathematik oder in Deutsch nicht über die notwendigen Grundkompetenzen verfügt. Um diese Frage zu beantworten, wurden die Schülerinnen und Schüler aufgrund ihrer Leistungen in Intervalle von jeweils 100 Punkten eingeteilt. Anhand dieser Einteilung können die Fähigkeiten der Schülerinnen und Schüler inhaltlich beschrieben werden. Abbildung 6.3 zeigt die Verteilung der Schülerinnen und Schüler in der Mathematik nach Intervall in den Jahren 2007 und 2008.

Abbildung 6.3 Verteilung der Schülerinnen und Schüler nach Intervall in der Mathematik 2007 und 2008 (ohne Fremdsprachenklasse)



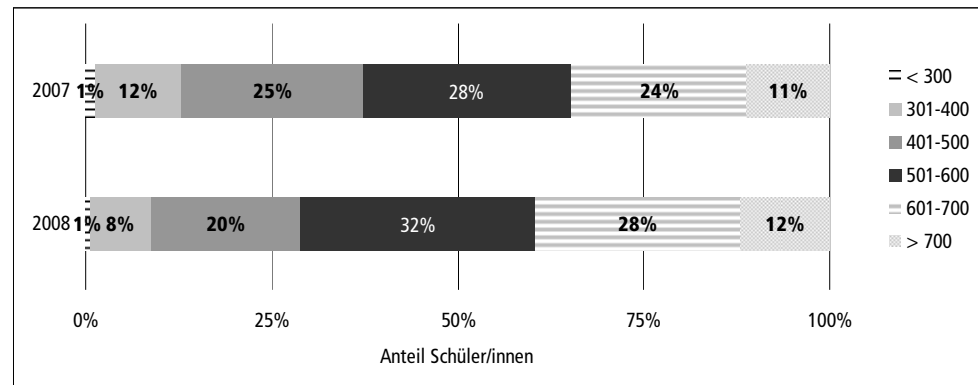
2007 erreichten in der Mathematik 14 Prozent der Schülerinnen und Schüler weniger als 400 Punkte. Diese Schülerinnen und Schüler verfügen nicht über die notwendigen Grundkompetenzen, um einfache Grundoperationen zu lösen. Dieser Anteil konnte im Jahr 2008 auf 10 Prozent reduziert werden.

Auch der Anteil jener Schülerinnen und Schüler, die zwar einfache Grundoperationen lösen können, nicht aber einfache Prozentrechnungen oder Grundoperationen mit Dezimalzahlen (Intervall 401–500), reduzierte sich von 28 Prozent im Jahr 2007 auf 24 Prozent im Jahr 2008.

Gleichzeitig erhöhte sich der Anteil Schülerinnen und Schüler, die mehr als 600 Punkte erreichen, von 24 Prozent im Jahr 2007 auf 32 Prozent im Jahr 2008. Schülerinnen und Schüler mit mehr als 600 Punkten sind in der Lage, Bruchgleichungen zu lösen und Folgerungen aus grafischen Darstellungen statistischer Daten zu ziehen.

Abbildung 6.4 zeigt die Verteilung der Schülerinnen und Schüler in Deutsch nach Intervall in den Jahren 2007 und 2008.

Abbildung 6.4 Verteilung der Schülerinnen und Schüler nach Intervall in Deutsch 2007 und 2008 (ohne Fremdsprachenklasse)



2007 erreichten insgesamt 13 Prozent der Schülerinnen und Schüler weniger als 400 Punkte. 2008 konnte dieser Anteil auf 9 Prozent reduziert werden. Diese Schülerinnen und Schüler sind fähig, einen einfachen Sachtext zu verstehen und Fälle in einfachen Sätzen richtig anwenden. Sie kennen aber beispielsweise die wichtigsten Zeitformen noch nicht.

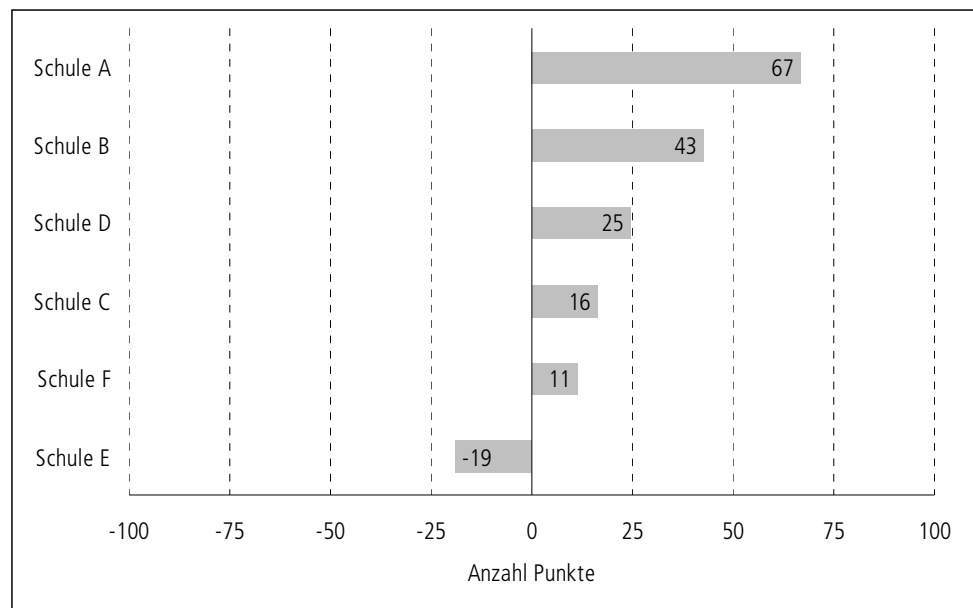
Auch der Anteil jener Schülerinnen und Schüler, die zwar die wichtigsten Fälle, Zeitformen und Kommaregeln kennen, diese aber noch nicht in Satzkonstruktionen anwenden können (Intervall 401–500), reduzierte sich von 25 Prozent im Jahr 2007 auf 20 Prozent im Jahr 2008.

Gleichzeitig erhöhte sich der Anteil Schülerinnen und Schüler, die mehr als 600 Punkte erreichen, von 35 Prozent im Jahr 2007 auf 40 Prozent im Jahr 2008. Schülerinnen und Schüler mit mehr als 600 Punkten verstehen komplexere Texte und können Zeitformen, Fälle und Pronomen richtig anwenden.

6.3 Leistungsentwicklung an den einzelnen Schulen der WBS

Abbildung 6.5 zeigt die Veränderungen der durchschnittlichen Leistungen der Schülerinnen und Schüler in der Mathematik zwischen den Abschlussprüfungen 2007 und 2008, dargestellt nach den einzelnen Schulen der WBS.

Abbildung 6.5 Leistungsentwicklung in der Mathematik zwischen 2007 und 2008 nach Schulen (nur Regelklassen)



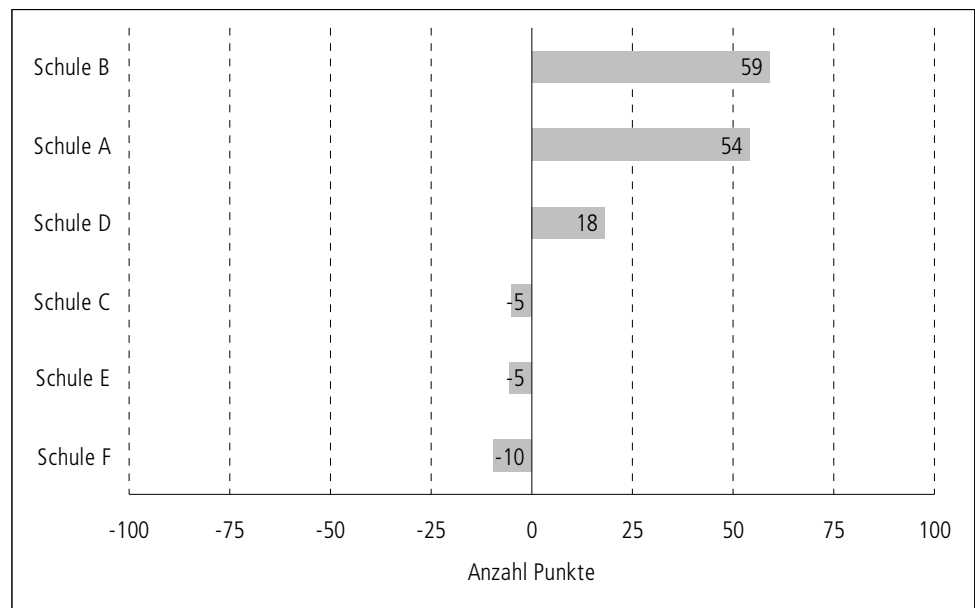
Die grösste Leistungsveränderung in der Mathematik kann im Schulhaus A festgestellt werden. An den Abschlussprüfungen 2008 erzielen die Schülerinnen und Schüler eine durchschnittliche Mathematikleistung, die um 67 Punkte über den Resultaten der Abschlussprüfungen 2007 liegt. Diese markante Verbesserung in den Abschlussprüfungen 2008 mit einer mittleren Effektgrösse von $d = 0,66$ kann nicht allein auf den geringeren Anteil Schülerinnen und Schüler des A-Zugs oder auf einen besonders starken «Jahrgang» zurückgeführt werden. Möglich ist, dass im Schulhaus A die getroffenen Massnahmen zur Leistungssteigerung konsequenter umgesetzt wurden als in anderen Schulen.

Ebenfalls eine deutlich positive Veränderung der Mathematikleistungen wird an der Schule B erzielt. Die Ergebnisse in der Abschlussprüfung 2008 liegen durchschnittlich 43 Punkte über den Ergebnissen der Abschlussprüfung 2007.

Ein schlechteres Ergebnis in den Mathematikprüfungen wird nur an der Schule E erzielt (minus 19 Punkte), obwohl der Anteil Schülerinnen und Schüler des A-Zugs im Vergleich zu 2007 leicht gesunken ist.

Abbildung 6.6 zeigt die Veränderungen der durchschnittlichen Leistungen der Schülerinnen und Schüler in Deutsch zwischen den Abschlussprüfungen 2007 und 2008, dargestellt nach den einzelnen Schulen der WBS.

Abbildung 6.6 Leistungsentwicklung in Deutsch zwischen 2007 und 2008 nach Schulen (nur Regelklassen)



Wie in der Mathematik können auch in Deutsch die grössten Leistungsveränderungen an den Schulen A (54 Punkte) und B (59 Punkte) festgestellt werden. An diesen beiden Schulen werden in den Abschlussprüfungen 2008 deutlich bessere Deutschleistungen erzielt als in den Abschlussprüfungen 2007 und auch deutlich bessere Deutschleistungen als an den übrigen Schulen.

Die Leistungen an den Schulen C, D, E und F verändern sich in Deutsch nur wenig. Im Schulhaus F, das bereits in den Abschlussprüfungen 2007 zu den schwächsten Schulen gehörte, haben sich die durchschnittlichen Deutschleistungen weiter reduziert.

7 Fazit

Aufgrund der stagnierenden Leistungen an den Abschlussprüfungen der WBS ergriff das Rektorat der WBS zusammen mit den Lehrpersonen unter dem Namen «Mathe-Paket» mehrere Massnahmen, um die Mathematikleistungen der Schülerinnen und Schüler zu verbessern. Bereits im laufenden Schuljahr konnten einige kleinere Massnahmen des «Mathe-Pakets» realisiert werden.

Wie die diesjährigen Prüfungsergebnisse zeigen, haben sich die umgesetzten Massnahmen bereits positiv ausgewirkt. Die Ergebnisse der Schülerinnen und Schüler sind statistisch signifikant besser als 2007. Dabei haben sich nicht nur die durchschnittlichen Mathematikleistungen verbessert, auch der Anteil Schülerinnen und Schüler mit fehlenden Grundkompetenzen konnte in der Mathematik gegenüber 2007 um 4 Prozent reduziert werden. Gleichzeitig konnte der Anteil Schülerinnen und Schüler erhöht werden, die fähig sind, schwierigere mathematische Aufgaben wie beispielsweise Bruchgleichungen zu lösen. Trotzdem erreichen sowohl in der Mathematik als auch in Deutsch noch immer rund 10 Prozent aller Schülerinnen und Schüler die Lernziele der WBS nicht. Sie sind nicht fähig, einfache Grundoperationen durchzuführen, Prozentrechnungen zu lösen oder mit Dezimalzahlen und Brüchen umzugehen. In Deutsch sind sie nicht in der Lage, Fälle und Zeitformen in einfachen Sätzen korrekt anzuwenden.

Die Ergebnisse in den Abschlussprüfungen offenbaren zudem grosse Leistungsunterschiede zwischen den A-Klassen. Im Gegensatz zu 2007 gelingt es einigen A-Klassen, leistungsmässig weitgehend zu den E-Klassen aufzuschliessen. In der Mathematik ist eine A-Klasse gar besser als der Durchschnitt der E-Klassen. Andere A-Klassen wiederum liegen deutlich unter den durchschnittlichen Leistungen aller WBS-Klassen. Die Gründe für die Leistungsunterschiede zwischen den Klassen sind sicher vielfältig. So unterscheiden sich die A-Klassen in der sozialen Zusammensetzung der Schülerschaft und im Anteil fremdsprachiger Kinder. Wie allerdings die Varianzanalysen zeigen, kann ein beträchtlicher Teil der Leistungsunterschiede auch auf Merkmale des Unterrichts zurückgeführt werden. Die Vermutung liegt daher nahe, dass die eingeleiteten qualitätssteigernden Massnahmen in einigen Klassen besser akzeptiert und konsequenter umgesetzt wurden und so zu einem grossen Teil für den Erfolg einzelner A-Klassen verantwortlich sind.

Eine weitere Veränderung zu den Abschlussprüfungen 2007 sind die grösseren Unterschiede zwischen den Schulen. An einigen Schulen werden zum einen deutlich bessere Leistungen erbracht als an den anderen Schulen. Zum anderen verzeichnen diese Schulen zwischen den Ergebnissen der Abschlussprüfungen 2007 und 2008 auch eine deutlich stärkere Leistungssteigerung als die übrigen Schulen. Ob diese Unterschiede ebenfalls auf die getroffenen Massnahmen zur Leistungssteigerung zurückzuführen sind, die je nach Schulhaus mit unterschiedlichem Fokus oder in unterschiedlicher Konsequenz umgesetzt wurden, lässt sich nur von den Schulen selbst abschliessend beurteilen.

Deutlich wird jedoch, dass eine jährliche Evaluation anhand der Abschlussprüfungen wenig bringt, wenn aus den Ergebnissen nicht qualitätssteigernde Massnahmen für den Unterricht abgeleitet und diese auch umgesetzt werden. Mit dem bisherigen Vorgehen scheint die WBS auf gutem Weg, die Leistungen der Schülerinnen und Schüler an der WBS nachhaltig zu erhöhen.

Anhang: Glossar der statistischen Begriffe

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

Signifikantes Ergebnis – Ein Ergebnis (Unterschied oder Zusammenhang) ist statistisch signifikant, wenn es durch ein statistisches Testverfahren überprüft und für gültig befunden wurde. Es kann mit einer bekannten, im Voraus festgelegten Irrtumswahrscheinlichkeit α (üblicherweise $\alpha = 0.05$) von der Stichprobe auf die Population geschlossen werden. In diesem Bericht erfüllen signifikante Ergebnisse die Bedingung $p < 0.05$. Die Signifikanz eines Ergebnisses sagt aber nichts aus über dessen Relevanz. Um signifikante Ergebnisse zu beurteilen, wird daher die Effektgrösse berechnet.

Effektgrösse – Zur Interpretation von statistisch signifikanten Unterschieden wird üblicherweise die Effektgrösse «d» berechnet, indem die Differenz der Mittelwerte durch die Standardabweichungen dividiert wird. Unterschiede, die aufgrund von verschiedenen Skalen zustande gekommen sind, werden so standardisiert und vergleichbar. Eine Effektgrösse von $d = 0.2$ weist auf einen schwachen Effekt hin, eine Effektgrösse von $d = 0.5$ auf einen mittleren Effekt und eine Effektgrösse von $d = 0.8$ auf einen starken Effekt. Auf der standardisierten WBS-Skala können dementsprechend Differenzen von 20 Punkten als klein, von 50 Punkten als mittelstark (oder deutlich) und von 80 Punkten als sehr gross bezeichnet werden.

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 zwischen dem Mittelwert \pm eine Standardabweichung $[M \pm SD]$ liegen. Wird der Bereich um je eine Standardabweichung erweitert – Mittelwert \pm zwei Standardabweichungen $[M \pm 2 SD]$ –, dann befinden sich darin rund 95 Prozent der Einzelwerte. Für die Leistungsdaten auf der standardisierten WBS-Skala ($SD = 100$ Punkte) heisst das, dass die Ergebnisse von rund zwei Drittel der Schülerinnen und Schüler zwischen 400 und 600 Punkten liegen und rund 95 Prozent der Ergebnisse zwischen 300 und 700 Punkten liegen. Die Standardabweichung wird im Bericht mit «SD» abgekürzt.

Überschneidungsbereiche – Als Überschneidungsbereich wird jener Teil zweier Verteilungen bezeichnet, in dem sich sowohl Ergebnisse der einen als auch der anderen Verteilung befinden. Vorausgesetzt, die Verteilungen sind annähernd normalverteilt, so lässt sich der Überschneidungsbereich in Form einer Prozentzahl berechnen.