



Universität Zürich
Institut für Bildungsevaluation



Nach sechs Jahren Primarschule.
Deutsch, Mathematik und motivational-emotionales
Befinden am Ende der 6. Klasse.

Bildungsdirektion des Kantons Zürich (Hrsg.)

Urs Moser, Domenico Angelone
Institut für Bildungsevaluation
Assoziiertes Institut der Universität Zürich

Judith Hollenweger, Alex Buff
Pädagogische Hochschule Zürich

Herausgeber

Bildungsdirektion des Kantons Zürich

Autorinnen und Autoren

Urs Moser, Domenico Angelone
Institut für Bildungsevaluation
Assoziiertes Institut der Universität Zürich

Judith Hollenweger, Alex Buff
Pädagogische Hochschule Zürich

Zitationsvorschlag

Moser, U., Buff, A., Angelone, D. & Hollenweger, J. (2011). *Nach sechs Jahren Primarschule. Deutsch, Mathematik und motivational-emotionales Befinden am Ende der 6. Klasse*. Zürich: Bildungsdirektion Kanton Zürich.

ISBN 978-3-905839-18-0

Bezugsquelle

Bildungsdirektion Kanton Zürich
Bildungsplanung
Walcheturm, Walcheplatz 2
8090 Zürich
www.bi.zh.ch → Veröffentlichungen
(Der Bericht liegt nur in elektronischer Form vor)

Copyright

© Bildungsdirektion Kanton Zürich 2011

Vorwort

Seit dem Jahr 2003 begleitet die Bildungsdirektion einen Teil der damals eingeschulten Schülerinnen und Schüler im Kanton Zürich auf ihrer Schullaufbahn. Die Erhebungen geben wissenschaftlich abgestützte Hinweise zu den Stärken und Schwächen unseres Bildungswesens. Im Sommer 2009 wurde, nach 2003 und 2006, die dritte Lernstandserhebung im Kanton Zürich durchgeführt.

Ziel der hier vorgelegten Erhebung in den sechsten Klassen der Primarschule ist es, den Lernstand in Deutsch und Mathematik festzustellen und zu überprüfen, in welchem Ausmass die Schülerinnen und Schüler die Ziele des Lehrplans erreichen und welche Lernfortschritte sie während der gesamten Primarschulzeit gemacht haben.

Die Studie geht der Frage nach, wie stark Unterschiede im Lernzuwachs von Schülerinnen und Schülern durch ihre familiäre Herkunft, ihren Migrationshintergrund oder ihr Geschlecht erklärt werden können. Geprüft wird auch, ob und wie solche Merkmale die Zuteilung in eine der Sekundarschulabteilungen oder ins Gymnasium beeinflussen. Es wird analysiert, welche Auswirkungen die soziale Zusammensetzung einer Klasse und der Anteil von Schülerinnen und Schüler mit Deutsch als Zweitsprache auf den Lernfortschritt haben. Untersucht wird auch, wie sich Lernmotivation und schulisches Selbstvertrauen über den Zeitraum der ersten sechs Jahre verändern und welche Wechselwirkungen zu den schulischen Leistungen bestehen.

Zur besseren Lesbarkeit gliedert sich der vorliegende Bericht in zwei Teile. Der erste Teil möchte die Resultate in gut verständlicher Form beschreiben. Der zweite Teil beschreibt die verwendeten statistischen Verfahren und präsentiert ausgewählte Resultate detaillierter.

Die Forschenden der Universität Zürich und der Pädagogischen Hochschule Zürich haben mit grossen Fachwissen und hohem Engagement eine sorgfältige Analyse des Lernstands auf der Zürcher Primarschule erstellt. Für alle Personen, die sich mit dem Zürcher Volksschulwesen und dessen Weiterentwicklung beschäftigen, will dieser Bericht eine reichhaltige Quelle an Informationen und Erkenntnissen zur Primarstufe und den Übertritt in die nachfolgende Schulstufe sein.

Zürich im Juni 2011

Bildungsdirektion, Bildungsplanung

Inhaltsverzeichnis

Teil 1: Ergebnisse	9
1 Einleitung	11
(Urs Moser & Domenico Angelone)	
2 Leistungszuwachs vom Schuleintritt bis zum Ende der Primarstufe	14
(Domenico Angelone & Urs Moser)	
2.1 Einleitung	14
2.2 Leistungszuwachs nach Geschlecht.....	16
2.3 Leistungszuwachs nach Erstsprache.....	18
2.4 Leistungszuwachs nach sozialer Herkunft	21
2.5 Leistungszuwachs nach Einschulungsform	24
2.6 Leistungszuwachs auf der Mittelstufe	27
2.7 Leistungen am Ende der 6. Klasse nach Vorwissen	28
2.8 Fazit	29
3 Fachleistungen am Ende der 6. Klasse	31
(Urs Moser & Domenico Angelone)	
3.1 Ergebnisse nach Anforderungsniveaus	31
3.2 Anforderungsniveaus in Mathematik	32
3.3 Anforderungsniveaus in Deutsch.....	37
3.4 Verteilung der Schülerinnen und Schüler nach den Anforderungsniveaus.....	41
3.5 Leistungen nach individuellen Merkmalen	42
3.6 Vergleich von Ergebnissen in den Jahren 1998 und 2009	46
3.7 Fazit	48
4 Die Bedeutung der Klassenzusammensetzung	50
(Domenico Angelone & Urs Moser)	
4.1 Segregation nach bildungsrelevanten Merkmalen	50
4.2 Zusammensetzung der Klassen	51
4.3 Klassenzusammensetzung und Schulleistungen.....	52
4.4 Die Situation von QUIMS-Schulen	54
4.5 Die Bedeutung der Förderung von Deutsch als Zweitsprache	58
4.6 Leistungsunterschiede zwischen den Klassen.....	59
4.7 Fazit	61

5	Leistungsbeurteilung	63
	(Domenico Angelone & Urs Moser)	
5.1	Leistungsbeurteilung durch Noten	63
5.2	Noten, individuelle Merkmale und Bezugsnorm.....	64
5.3	Fazit	69
6	Übertritt von der Primarschule in die Sekundarstufe I	71
	(Domenico Angelone & Urs Moser)	
6.1	Kritische Schnittstelle	71
6.2	Leistungen nach zugewiesenem Schultyp	71
6.3	Individuelle Merkmale nach zugewiesenem Schultyp	75
6.4	Übertritt ins Langgymnasium	76
6.5	Übertritt in die Abteilung A der Sekundarschule	80
6.6	Fazit	82
7	Motivation und lernbezogene Emotionen nach sechs Jahren Primarstufe in Mathematik und Deutsch	83
	(Alex Buff)	
7.1	Einleitung	83
7.2	Kompetenzüberzeugung in Mathematik und Deutsch.....	85
7.3	Valenzüberzeugung in Mathematik und Deutsch.....	88
7.4	Lernbezogene Emotionen in Mathematik und Deutsch	92
7.5	Fazit	94
7.6	Anmerkungen.....	97
8	Motivation und Leistung in Mathematik und Deutsch	98
	(Alex Buff)	
8.1	Einleitung	98
8.2	Motivation und Leistung in Mathematik	100
8.3	Motivation und Leistung in Deutsch.....	103
8.4	Fazit	106
9	Unterricht, Motivation und Leistung in Mathematik und Deutsch	108
	(Alex Buff)	
9.1	Einleitung	108
9.2	Unterricht, Motivation und Leistung in Mathematik	100
9.3	Unterricht, Motivation und Leistung in Deutsch.....	112
9.4	Fazit	113
9.5	Anmerkungen.....	115

Teil 2: Analysen zu den Ergebnissen.....	117
1 Die Zürcher Lernstandserhebung im Überblick	119
(Domenico Angelone & Urs Moser)	
1.1 Grundgesamtheit und Stichprobe.....	119
1.2 Leistungsskalen.....	124
1.3 Soziale Herkunft, Erstsprache und kognitive Grundfähigkeiten	126
2 Analysen zum Leistungszuwachs vom Schuleintritt bis zum Ende der Primarstufe	129
(Domenico Angelone)	
2.1 Einleitung.....	129
2.2 Methode.....	129
2.3 Ergebnisse zum Leistungszuwachs von der 1. bis 6. Klasse.....	130
2.4 Ergebnisse zu den Leistungen am Ende der 6. Klasse nach Vorwissen.....	134
3 Methodisches Vorgehen und Analysen zu den Fachleistungen am Ende der 6. Klasse	137
(Urs Moser & Domenico Angelone)	
3.1 Bildung von Anforderungsniveaus	137
3.2 Leistungstests für den Vergleich zwischen 1998 und 2009	138
3.3 Ergebnisse zum Leistungsvergleich zwischen 1998 und 2009.....	140
3.4 Ergebnisse zum Vergleich der Selbsteinschätzungen von Schul- und Unterrichtsmerkmalen zwischen 1998 und 2009	142
4 Analysen zur Bedeutung der Klassenzusammensetzung für die Leistungen	145
(Domenico Angelone)	
4.1 Einleitung.....	145
4.2 Methode.....	145
4.3 Ergebnisse zur Bedeutung der Klassenzusammensetzung für die Leistungen am Ende der 6. Klasse.....	149
4.4 Ergebnisse zum DaZ-Unterricht.....	157
5 Analysen zur Leistungsbeurteilung.....	163
(Domenico Angelone)	
5.1 Einleitung.....	163
5.2 Methode.....	163
5.3 Ergebnisse zur Erklärung der Zeugnisnoten in Deutsch.....	163
5.4 Ergebnisse zur Erklärung der Zeugnisnoten in Mathematik	166

6	Analysen zum Übertritt von der Primarschule in die Sekundarstufe I	170
	(Domenico Angelone)	
6.1	Einleitung	170
6.2	Methode.....	170
6.3	Ergebnisse zum Übertritt in das Langgymnasium.....	171
6.4	Ergebnisse zum Übertritt in die Abteilung A der Sekundarschule.....	175
7	Analysen zu Motivation und lernbezogene Emotionen nach sechs Jahren Primarstufe in Mathematik und Deutsch	178
	(Alex Buff)	
7.1	Einleitung	178
7.2	Methode.....	178
7.3	Ergebnisse	180
8	Analysen zu Motivation und Leistung in Mathematik und Deutsch	187
	(Alex Buff)	
8.1	Einleitung	187
8.2	Methode.....	187
8.3	Ergebnisse	189
9	Analysen zu Unterricht, Motivation und Leistung in Mathematik und Deutsch	193
	(Alex Buff)	
9.1	Einleitung	193
9.2	Methode.....	193
9.3	Ergebnisse	196
	Literatur	199

Teil 1

Ergebnisse

1 Einleitung

Urs Moser & Domenico Angelone

Zürcher Längsschnittstudie

Mit der Zürcher Längsschnittstudie werden im Abstand von drei Jahren fachliche und fachübergreifende Kompetenzen von Schülerinnen und Schülern erfasst, die im Sommer 2003 in die 1. Klasse der Primarschule eingetreten sind. Die Ergebnisse der dritten Erhebung im Sommer 2009 sind im vorliegenden Bericht dargestellt.

Ziel der Zürcher Längsschnittstudie ist es, den Lernstand nach drei, sechs und neun Jahren mit den Zielen des Lehrplans zu vergleichen und den Lernzuwachs in Abhängigkeit von individuellen Merkmalen, wie der sozialen Herkunft, zu beschreiben. Dargestellt wird auch die Entwicklung der motivationalen und emotionalen Lernvoraussetzungen.

Mit der dritten Untersuchung nach sechs Jahren werden insbesondere Fragen zum Übertritt von der Primarschule in die Schultypen der Sekundarstufe I beantwortet: Wie gut stimmen die Ergebnisse in den Leistungstests mit den Noten der Lehrpersonen überein? Welche Rolle spielt die soziale Herkunft beim Übertritt in die Schultypen der Sekundarstufe I? Welche Bedeutung haben private Vorbereitungskurse für den Erfolg beim Übertritt ins Langgymnasium?

Ein weiterer Schwerpunkt der Analysen betrifft die Klassenzusammensetzung: Welche Rolle spielt die soziale Zusammensetzung der Klasse einerseits für die Leistungen der Schülerinnen und Schüler und andererseits für die Beurteilung mit Noten? Weil bereits im Jahr 1998 eine Lernstandserhebung am Ende der 6. Klasse durchgeführt wurde, können zudem die Ergebnisse der beiden Erhebungen miteinander verglichen werden.

Längsschnittstichprobe und Klassenstichprobe

Von den anfänglich hundert Schulklassen, auf die sich die rund 2000 Schülerinnen und Schüler verteilten, bestanden am Ende der 6. Klasse nur noch wenige in ihrer ursprünglichen Form. Der Grund liegt darin, dass die Schulklassen im Kanton Zürich in der Regel nach drei Jahren Primarschule neu zusammengesetzt werden. Die Zürcher Längsschnittstichprobe wurde deshalb für die Erhebung im Sommer 2009 mit einer repräsentativen Stichprobe von hundert 6. Klassen ergänzt. Dadurch wurde erreicht, dass die Leistungen am Ende der 6. Klasse für den Kanton Zürich beschrieben und zugleich die Bedeutung der Schulklasse für die schulischen Leistungen überprüft werden konnten.

Die *Längsschnittstichprobe* umfasste ursprünglich 1941 Schülerinnen und Schüler, die im Sommer 2003 in die 1. Klasse der Primarschule eintraten, sowie 105 Schülerinnen und Schüler, die in die Kleinklasse A eintraten. Von diesen insgesamt 2046 Schülerinnen und Schülern sind bis im Sommer 2009 rund 12 Prozent nicht mehr Teil der Stichprobe, weil sie den Kanton Zürich verlassen oder an eine Privatschule gewechselt haben. Weitere 6 Prozent nahmen aus anderen Gründen nicht an der Erhebung teil, weshalb die vorliegenden Ergebnisse auf den Informationen von 1679 Schülerinnen und Schülern beruhen, was rund 82 Prozent der Ausgangsstichprobe entspricht.

Rund 0.5 Prozent der Schülerinnen und Schüler besuchten nach sechs Jahren bereits eine Schule auf der Sekundarstufe I. Rund 8 Prozent hatten hingegen eine Klasse repetiert und besuchten nach sechs Jahren die 5. Klasse. Rund 2 Prozent waren in der Regelklasse gestartet und besuchten nach sechs Jahren eine Kleinklasse.

Die *Klassenstichprobe* umfasste ursprünglich 2093 Schülerinnen und Schüler aus hundert Klassen. Zwei Klassen beziehungsweise 112 Schülerinnen und Schüler nahmen aus verschiedenen Gründen an der Lernstandserhebung im Sommer 2009 nicht teil. Die vorliegenden Ergebnisse beruhen auf den Informationen von 1981 Schülerinnen und Schülern aus 98 Klassen. Dies entspricht einer Rücklaufquote von knapp 95 Prozent.

Durchführung

Die dritte Erhebung der Zürcher Längsschnittstudie wurde im Sommer 2009 von den beteiligten Schulen nach einer standardisierten Anleitung selbstständig durchgeführt. Dieses Vorgehen wurde deshalb gewählt, weil sich die beteiligten Schülerinnen und Schüler auf 509 Klassen aus 251 Schulen des Kantons Zürich verteilten und dadurch eine externe Durchführung sehr aufwendig geworden wäre.

Darstellung der Ergebnisse

Der vorliegende Bericht enthält zwei Teile: einen ersten Teil mit der Beschreibung und Interpretation der Ergebnisse und einen zweiten Teil mit der Darstellung der methodischen und statistischen Grundlagen der Ergebnisse. Während im ersten Teil nahezu vollständig auf die Angaben zu den Analysen verzichtet wird, werden im zweiten Teil die Ergebnisse der statistischen Analysen ausführlich dargestellt und diskutiert.

Für das Verständnis der Ergebnisse sind einige Angaben über die Untersuchung hilfreich. Der Leistungszuwachs von der 1. Klasse bis zur 6. Klasse wurde aufgrund von *Leistungstests* erfasst, die zu allen drei Erhebungszeitpunkten eingesetzt wurden. Die Leistungstests wurden mithilfe von Lehrpersonen des Kantons Zürich entwickelt.

Während die Lernstandserhebungen zu Beginn der 1. und am Ende der 3. Klasse aufgrund der grossen Lernfortschritte der Kinder in den ersten drei Schuljahren mit voneinander unabhängigen Leistungstests durchgeführt wurden, konnten die Lernstandserhebungen am Ende der 3. und am Ende der 6. Klasse miteinander verbunden werden. Dadurch konnte der Leistungszuwachs auf einer einheitlichen Kompetenzskala entsprechend den Zielen des Lehrplans des Kantons Zürich ausgewiesen werden.

Für die Erfassung motivationaler Aspekte oder Angaben zum Unterricht füllten die Schülerinnen und Schüler einen *Fragebogen* aus. Es handelt sich bei diesen Angaben um Selbsteinschätzungen der Schülerinnen und Schüler.

Zur Bestimmung der *sozialen Herkunft* wurde ein Index aus verschiedenen bildungsrelevanten Angaben, wie der Ausbildung der Eltern oder der Anzahl Bücher zu Hause, gebildet. Für die Darstellung der Ergebnisse nach der sozialen Herkunft wurden aufgrund des Index vier gleich grosse Gruppen gebildet: Schülerinnen und Schüler (1) mit benachteiligter sozialer Herkunft, (2) mit eher benachteiligter sozialer Herkunft, (3) mit eher privilegierter sozialer Herkunft und (4) mit privilegierter sozialer Herkunft.

Zur Unterscheidung von Schülerinnen und Schülern mit Deutsch als *Erstsprache* und solchen mit Deutsch als *Zweitsprache* wurde erfasst, welche Sprache die Kinder am häufigsten mit ihren Eltern sprechen.

Die allgemeinen kognitiven Fähigkeiten wurden zum ersten Erhebungszeitpunkt beim Schuleintritt mit einem Intelligenztest erfasst.

2 Leistungszuwachs vom Schuleintritt bis zum Ende der Primarstufe

Domenico Angelone & Urs Moser

2.1 Einleitung

Die Erhebung beim Schuleintritt zeigte in den Bereichen Lesen, Wortschatz und Mathematik grosse Unterschiede im Lern- und Entwicklungsstand auf. Während ein Drittel der Schülerinnen und Schüler zu Beginn der 1. Klasse bereits einfache Sätze und Wörter lesen konnte, kannte ein Drittel der Kinder erst wenige Buchstaben, die sie aber noch nicht zu Wörtern zusammenfassen konnten. Während rund ein Fünftel der Schülerinnen und Schüler sich bereits im Zahlenraum bis 100 zurechtfinden und einfache Additionen und Subtraktionen im Zahlenraum bis 20 löste, beschränkten sich bei einem anderen Fünftel die mathematischen Kompetenzen auf die Kenntnis der Zahlen, das Zählen bis 20 sowie das Erkennen kleiner Mengen.

Die Leistungsunterschiede beim Schulstart sind gross. Und sie lassen sich bereits beim Schuleintritt zu einem beträchtlichen Teil durch individuelle Lernvoraussetzungen, wie Erstsprache und soziale Herkunft der Schülerinnen und Schüler, erklären. Soziale Ungleichheiten entstehen, längst bevor die Kinder ins Bildungssystem eintreten. Die Auseinandersetzung mit sozialen Ungleichheiten im Bildungssystem konzentrierte sich lange Zeit fast ausschliesslich auf Ungleichheiten der Bildungsbeteiligung zu einem bestimmten Zeitpunkt der Schullaufbahn. Beispielsweise wurde die Diskussion vor allem am Anteil an Kindern aus sozioökonomisch benachteiligten Familien im Gymnasium geführt, der weit kleiner ist als der Anteil an Kindern aus sozioökonomisch privilegierten Familien (Maaz, Baumert & Trautwein, 2009, S. 13). Soziale Ungleichheiten können allerdings bereits vor dem Eintritt der Kinder ins Bildungssystem aufgedeckt werden (Bayer & Moser, 2009).

Zur Erklärung von herkunftsbedingten Unterschieden ist Boudons Theorie (1974) über den Zusammenhang zwischen sozialer Herkunft und Bildungsungleichheit zentral. Soziale Ungleichheiten der Bildungsbeteiligung werden als das Ergebnis von Entscheidungen gesehen, die in einem institutionellen Rahmen der Schule getroffen werden müssen. Bildungsentscheidungen ergeben sich aufgrund der schulischen Leistungen, der Selektionsverfahren im Schulsystem und der familiären Bewertung von Bildung. Für die Erklärung der Bildungsentscheidungen führt Boudon die Unterscheidung zwischen primären und sekundären Herkunftseffekten ein. Diese Unterscheidung ist für die empirische Überprüfung von Herkunftseffekten grundlegend (Maaz, Baumert & Trautwein, 2009, S. 14ff.).

Als *primäre Herkunftseffekte* werden jene Einflüsse der sozialen Herkunft bezeichnet, die sich direkt auf die Kompetenzentwicklung der Schülerinnen und Schüler auswirken und in den Schulleistungen sichtbar werden. Die auf die soziale Herkunft zurückführbaren Leistungsunterschiede sind eine Folge der unterschiedlichen Ausstattung von Familien mit ökonomischem, sozialem und kulturellem Kapital. Die nach Herkunft unterschiedlichen Ressourcen und Anregungsmilieus schlagen sich bereits zu Beginn der Schullaufbahn nieder. Primäre soziale Ungleichheiten sind Unterschiede zwischen den Schülerinnen und Schülern verschiedener sozialer Herkunft in den bis zu einer Schnittstelle erworbenen und für den Übergang

vorausgesetzten Leistungen. Sie werden als durch Leistungen gedeckte soziale Ungleichheiten bezeichnet (Baumert & Schümer, 2001, S. 358).

Als *sekundäre Herkunftseffekte* werden dagegen jene sozialen Ungleichheiten bezeichnet, die – unabhängig von der Kompetenzentwicklung und den erreichten schulischen Leistungen – auf unterschiedliche Bildungsaspirationen und Entscheidungsverhalten nach sozialer Herkunft zurückzuführen sind. Sekundäre soziale Ungleichheiten sind Unterschiede zwischen den Schülerinnen und Schülern verschiedener sozialer Herkunft, die – bei gleichen schulischen Leistungen – aus einem je nach sozialer Herkunft der Familie spezifischen Übergangszustand kommen. «Der sekundäre Herkunftseffekt ist demnach auch ein Ausdruck der verinnerlichten Sozialschichtzugehörigkeit» (Maaz et al., 2009, S. 15). Beispielsweise besteht für ein Kind von Eltern ohne Bildungsabschluss eine grössere Distanz, ein Gymnasium zu besuchen, als für ein Kind aus einer Akademikerfamilie. Soziale Ungleichheiten entstehen allerdings auch innerhalb des Bildungssystems. Beispielsweise können sozial selektive Erwartungs- und Wertschätzungsstrukturen der Schule oder Übertrittsempfehlungen von Lehrpersonen dafür verantwortlich sein (Tillmann, 2002).

Ungleichheiten lassen sich auch auf andere individuelle Merkmale, wie die Erstsprache oder das Geschlecht, zurückführen. Die Frage, inwieweit die Schule beziehungsweise das Bildungssystem für Bildungsungleichheiten verantwortlich ist, lässt sich allerdings nicht so einfach beantworten (Ditton, 2007, S. 60f.). Die Beschreibung des Leistungszuwachses vom Schuleintritt bis zum Ende der Primarschule gibt zumindest einen Hinweis darauf, ob und wie sich die sozialen Ungleichheiten bis zur ersten Schnittstelle, dem Übergang in die Sekundarstufe I, verändern.

Mit der Darstellung des Leistungszuwachses vom Schuleintritt bis zum Ende der Primarstufe lässt sich aufzeigen, wie sich die Leistungen in Abhängigkeit von individuellen Merkmalen, wie Erstsprache, Geschlecht oder soziale Herkunft, entwickeln. Bei den folgenden Darstellungen wird unterschieden zwischen dem (faktischen) Leistungszuwachs und dem nach Lernvoraussetzungen statistisch kontrollierten Leistungszuwachs. Die Darstellung ohne statistische Kontrolle der Lernvoraussetzungen gibt ein Bild davon, wie der Leistungszuwachs für verschiedene Gruppen – beispielsweise Knaben oder Schülerinnen und Schüler mit Deutsch als Zweitsprache – verläuft.

Die Darstellung des Leistungszuwachses vom Schuleintritt bis zum Ende der Primarstufe mit statistischer Kontrolle der Lernvoraussetzungen zeigt, welche Bedeutung ein spezifisches Merkmal, wie Geschlecht, Erstsprache oder soziale Herkunft, für den Leistungszuwachs hat. Diese Differenzierung in der Darstellung erlaubt es, die Ursachen des Leistungszuwachses präziser auf individuelle Merkmale zurückzuführen. Bei der Darstellung mit statistischer Kontrolle der Lernvoraussetzungen werden jeweils die Effekte aller Merkmale ausser jenem von Interesse konstant gehalten. Zu den Lernvoraussetzungen werden das Alter, die Erstsprache, das Geschlecht, die kognitive Grundfähigkeit und die soziale Herkunft gezählt.

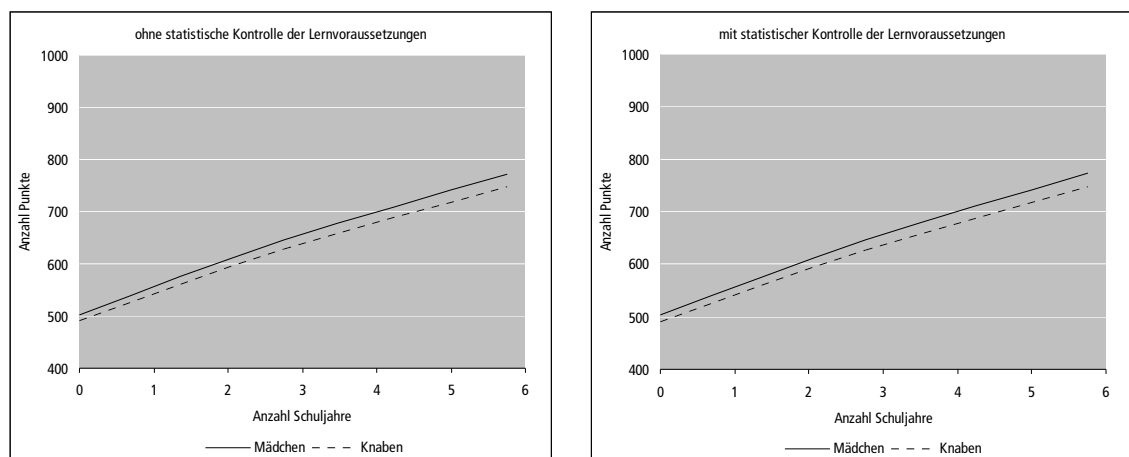
2.2 Leistungszuwachs nach Geschlecht

Die Abbildungen 2.1 bis 2.3 zeigen den Leistungszuwachs der Mädchen und Knaben im Lesen, im Wortschatz und in der Mathematik vom Schuleintritt bis zum Ende der 6. Klasse ohne statistische und mit statistischer Kontrolle der Lernvoraussetzungen.

Leistungszuwachs im Lesen

Im linken Teil der Abbildung 2.1 sind die Ergebnisse im Lesen ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Zu Beginn der 1. Klasse schneiden die Mädchen im Lesen statistisch signifikant besser ab als die Knaben. Die Mädchen erzielen 503 Punkte und die Knaben 490 Punkte. Bis zum Ende der 6. Klasse steigt dieser Vorsprung von 13 Punkten um nahezu das Zweifache auf 25 Punkte an. Die Mädchen weisen bis zum Ende der 6. Klasse einen leicht grösseren Leistungszuwachs auf als die Knaben. Der Unterschied beträgt jedoch lediglich 2 Punkte pro Schuljahr und ist statistisch nicht signifikant. Am Ende der 6. Klasse erzielen die Mädchen 773 Punkte und die Knaben 748 Punkte. Der Unterschied von 25 Punkten ist statistisch signifikant.

Abbildung 2.1: Leistungszuwachs im Lesen nach Geschlecht



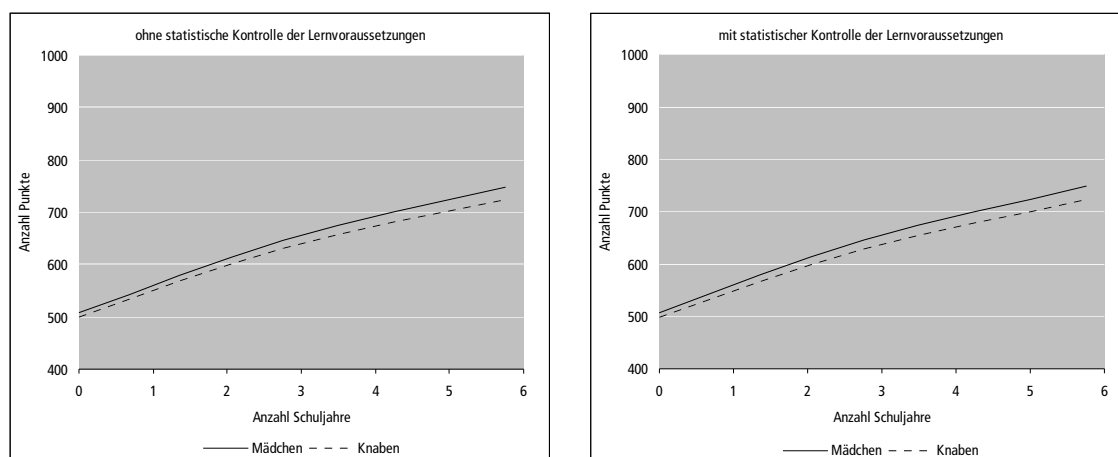
Im rechten Teil der Abbildung 2.1 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen dargestellt. Die Kurvenverläufe ändern sich kaum im Vergleich zur unkontrollierten Darstellung. Die Mädchen weisen auch nach statistischer Kontrolle der Lernvoraussetzungen die Tendenz nach einen leicht grösseren Leistungszuwachs auf (rund 2 Punkte pro Schuljahr). Demzufolge beträgt der Leistungszuwachs bis zum Ende der 6. Klasse 258 Punkte bei den Knaben und 270 Punkte bei den Mädchen. Am Ende der 6. Klasse erreichen Mädchen gegenüber Knaben statistisch signifikant bessere Leseleistungen. Ihr Vorsprung beträgt rund 25 Punkte.

Leistungszuwachs im Wortschatz

Im linken Teil der Abbildung 2.2 sind die Ergebnisse im Wortschatz ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Im Wortschatz erreichen die Mädchen zu Beginn der

1. Klasse 507 Punkte und die Knaben 500 Punkte. Der Unterschied von 7 Punkten zugunsten der Mädchen ist statistisch nicht signifikant. Demgegenüber ist der Leistungszuwachs der Mädchen bis zum Ende der 6. Klasse mit rund 3 Punkten pro Schuljahr statistisch signifikant grösser als jener der Knaben. Während der Leistungszuwachs der Mädchen bis zum Ende der 6. Klasse 241 Punkte beträgt, verbessern sich die Knaben im selben Zeitraum lediglich um 223 Punkte. Am Ende der 6. Klasse erzielen Mädchen eine um rund 25 Punkte bessere Wortschatzleistungen als Knaben. Der Unterschied ist statistisch signifikant.

Abbildung 2.2: Leistungszuwachs im Wortschatz nach Geschlecht

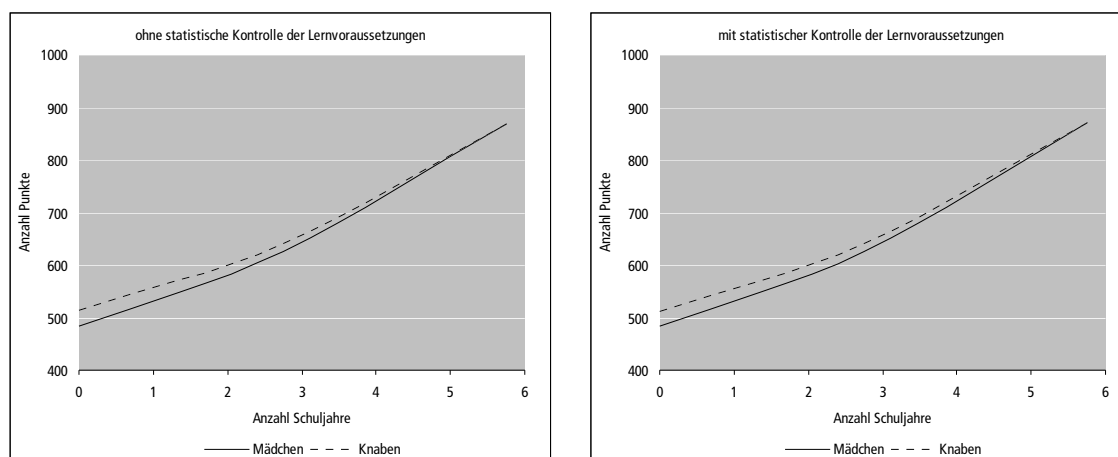


Wie aus der Darstellung im rechten Teil der Abbildung 2.2 hervorgeht, zeigt sich bei statistischer Kontrolle der Lernvoraussetzungen ein weitgehend ähnliches Bild wie bei der unkontrollierten Darstellung. Zu Beginn der 1. Klasse finden sich lediglich geringe Unterschiede in den Wortschatzleistungen zwischen den Geschlechtern. Der Leistungszuwachs der Mädchen bleibt im Vergleich zu den Knaben auch nach Berücksichtigung der Lernvoraussetzungen statistisch signifikant grösser (rund 3 Punkte pro Schuljahr). Bei den Knaben beträgt der Leistungszuwachs bis zum Ende der Primarstufe 224 Punkte, bei den Mädchen 241 Punkte. Am Ende der 6. Klasse erzielen die Mädchen 749 Punkte und die Knaben 723 Punkte. Der Unterschied von 26 Punkten ist statistisch signifikant.

Leistungszuwachs in der Mathematik

Der linke Teil der Abbildung 2.3 zeigt die Ergebnisse in der Mathematik für die Mädchen und die Knaben ohne statistische Kontrolle der Lernvoraussetzungen. Im Gegensatz zu den sprachlichen Kompetenzbereichen schneiden die Knaben in der Mathematik beim Schuleintritt statistisch signifikant besser ab als die Mädchen. Die Knaben erreichen 514 Punkte und die Mädchen 483 Punkte. Weil der Leistungszuwachs der Mädchen um rund 5 Punkte pro Schuljahr statistisch signifikant grösser ist als jener der Knaben, verschwindet der anfängliche Vorsprung von 31 Punkten bis zum Ende der 6. Klasse vollständig.

Abbildung 2.3: Leistungszuwachs in der Mathematik nach Geschlecht



Im rechten Teil der Abbildung 2.3 sind die Ergebnisse in der Mathematik mit statistischer Kontrolle der Lernvoraussetzungen abgebildet. Dabei zeigt sich nahezu dasselbe Bild wie bei der unkontrollierten Darstellung. Zu Beginn der 1. Klasse erzielen die Knaben 512 Punkte und die Mädchen 485 Punkte. Bis zum Ende der 6. Klasse gleichen die Mädchen diesen Rückstand aus. Der Leistungszuwachs der Mädchen ist statistisch signifikant grösser, rund 5 Punkte pro Schuljahr. Seit Beginn der 1. Klasse beträgt der Leistungszuwachs bei den Mädchen 385 Punkte und bei den Knaben 358 Punkte. Am Ende der 6. Klasse erreichen sowohl Mädchen als auch Knaben rund 870 Punkte.

2.3 Leistungszuwachs nach Erstsprache

Die Abbildungen 2.4 bis 2.6 zeigen den Leistungszuwachs der Schülerinnen und Schüler mit Deutsch als Erstsprache (DaE) und mit Deutsch als Zweitsprache (DaZ) vom Schuleintritt bis zum Ende der 6. Klasse ohne statistische und mit statistischer Kontrolle der Lernvoraussetzungen. Die Erstsprache entspricht der Sprache, die die Eltern am häufigsten mit ihrem Kind sprechen. Der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache beträgt nach dieser Definition zum ersten Testzeitpunkt im Jahr 2003 rund 27 Prozent. Von diesen Kindern sprechen rund 20 Prozent eine südslawische Sprache (Bosnisch, Kroatisch, Mazedonisch, Montenegroinisch, Serbisch und Slowenisch) und je 15 Prozent Italienisch oder Albanisch. Andere Sprachen, wie Türkisch, Spanisch, Portugiesisch oder Tamilisch, werden von weniger als 10 Prozent der Kinder mit «Deutsch als Zweitsprache» gesprochen. Die Angaben zur Erstsprache stammen aus dem Elternfragebogen.

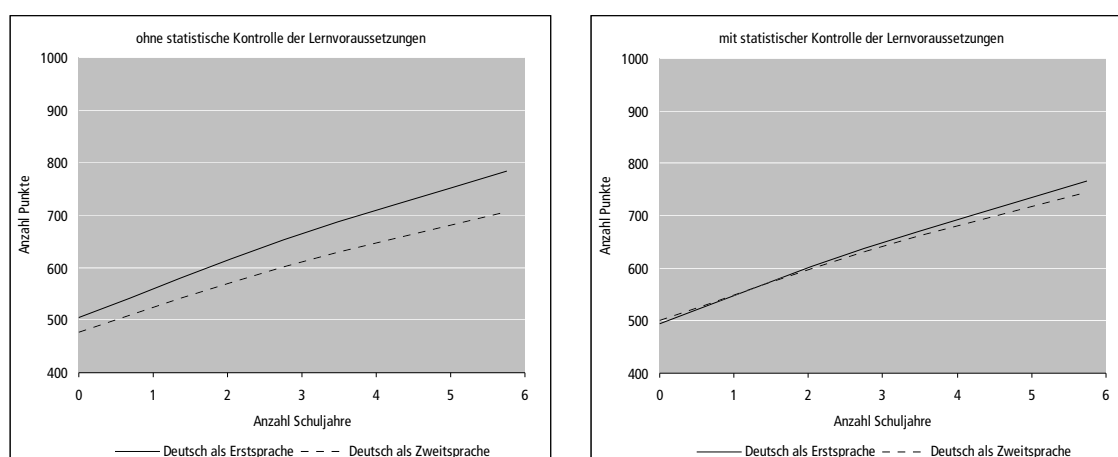
Leistungszuwachs im Lesen

Im linken Teil der Abbildung 2.4 sind die Ergebnisse im Lesen ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Zu Beginn der 1. Klasse erzielen die Schülerinnen und Schüler mit Deutsch als Erstsprache 505 Punkte und jene mit Deutsch als Zweitsprache 478 Punkte. Zudem ist der Leistungszuwachs der Schülerinnen und Schüler mit Deutsch als Zweitsprache bis zum Ende der 6. Klasse mit rund 9 Punkten pro Schuljahr statistisch signifikant kleiner als jener der Schülerinnen und Schüler mit Deutsch als Erstsprache. Der anfängliche Rück-

stand von 27 Punkten steigt deshalb bis zum Ende der 6. Klasse auf 77 Punkte an. Der Unterschied ist statistisch signifikant. Am Ende der 6. Klasse erreichen die Schülerinnen und Schüler mit Deutsch als Erstsprache im Lesen 784 Punkte, jene mit Deutsch als Zweitsprache 707 Punkte.

Im rechten Teil der Abbildung 2.4 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen dargestellt. Im Gegensatz zur unkontrollierten Darstellung zeigt sich ein leicht verändertes Bild, insbesondere was die Unterschiede beim Eintritt in die 1. Klasse betrifft. Nach statistischer Kontrolle der Lernvoraussetzungen unterscheiden sich die Leseleistungen zwischen Schülerinnen und Schülern mit und ohne Deutsch als Zweitsprache nicht. Der Leistungszuwachs der Schülerinnen und Schüler mit Deutsch als Erstsprache bleibt allerdings nach Berücksichtigung der Lernvoraussetzungen statistisch signifikant grösser als jener der Schülerinnen und Schüler mit Deutsch als Zweitsprache. Die Differenz im Leistungszuwachs beträgt rund 5 Punkte pro Schuljahr.

Abbildung 2.4: Leistungszuwachs im Lesen nach Erstsprache



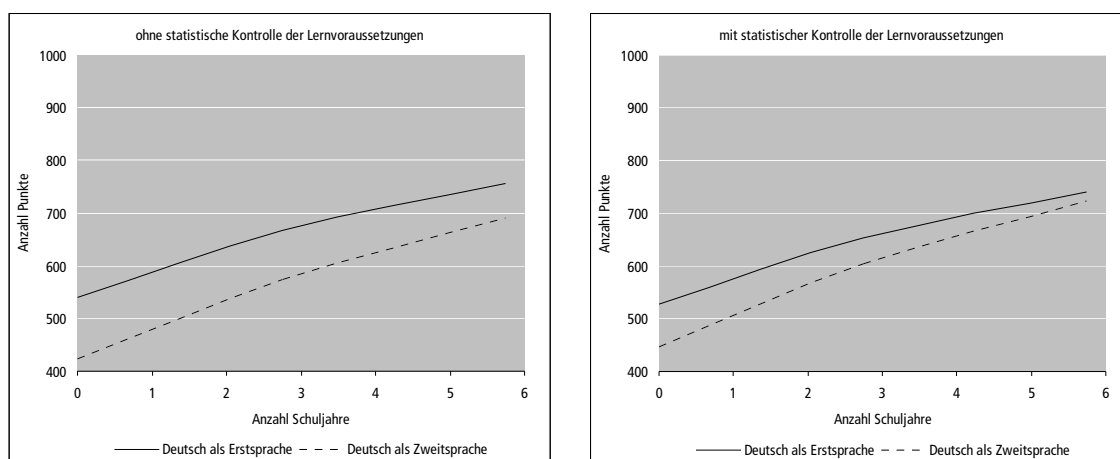
Der Vergleich der beiden Darstellungen in Abbildung 2.4 zeigt, dass der Effekt der Erstsprache bei der statistischen Kontrolle der Lernvoraussetzungen deutlich geringer ausfällt als ohne Berücksichtigung der Lernvoraussetzungen. Während sich die Schülerinnen und Schüler mit Deutsch als Erstsprache bis zur 6. Klasse um 273 Punkte verbessern, beträgt der Leistungszuwachs der Schülerinnen und Schüler mit Deutsch als Zweitsprache 245 Punkte.

Am Ende der 6. Klasse erreichen die Schülerinnen und Schüler mit Deutsch als Zweitsprache rund 22 Punkte tiefere Leseleistungen als jene mit Deutsch als Erstsprache. Der deutlich geringere Effekt der Erstsprache nach Berücksichtigung der Lernvoraussetzungen lässt sich dadurch erklären, dass für die Schülerinnen und Schüler mit Deutsch als Zweitsprache primär nicht der bilinguale Spracherwerb ein Problem darstellt, sondern vor allem die soziale Herkunft. Schülerinnen und Schüler mit Deutsch als Zweitsprache stammen überdurchschnittlich oft aus sozial benachteiligten Verhältnissen.

Leistungszuwachs im Wortschatz

Im linken Teil der Abbildung 2.5 sind die Ergebnisse im Wortschatz ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Im Wortschatz sind die Leistungsunterschiede zwischen den Schülerinnen und Schülern mit Deutsch als Erstsprache und jenen mit Deutsch als Zweitsprache mit 117 Punkten zu Beginn der 1. Klasse wesentlich grösser als im Lesen. Die Kinder mit Deutsch als Erstsprache erreichen 539 Punkte, jene mit Deutsch als Zweitsprache 422 Punkte. Am Ende der 6. Klasse sind die Unterschiede deutlich geringer. Die Kinder mit Deutsch als Zweitsprache weisen einen statistisch signifikant grösseren Leistungszuwachs auf und holen deshalb einen Teil ihres Rückstandes auf. Mit 65 Punkten sind die Unterschiede am Ende der 6. Klasse aber immer noch sehr gross und statistisch signifikant.

Abbildung 2.5: Leistungszuwachs im Wortschatz nach Erstsprache

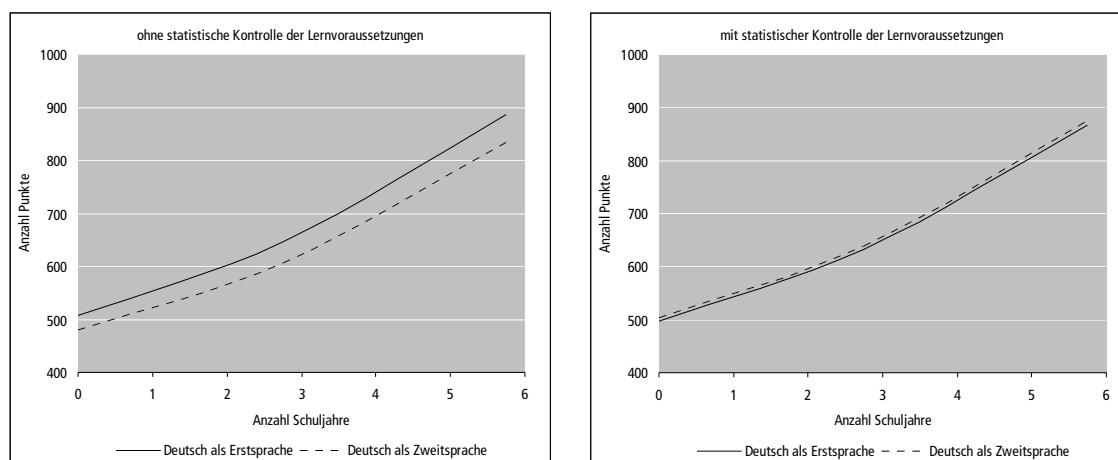


Im rechten Teil der Abbildung 2.5 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen abgebildet. Dabei zeigt sich ein ähnliches Bild wie bei der unkontrollierten Darstellung. Der Unterschied zwischen den beiden Gruppen fällt aber wesentlich geringer aus. Zu Beginn der 1. Klasse weisen die Schülerinnen und Schüler mit Deutsch als Erstsprache einen Vorsprung von 81 Punkten auf. Der Vorsprung reduziert sich bis zum Ende der 6. Klasse deutlich und beträgt nur noch 17 Punkte, ist jedoch statistisch signifikant. Während sich die Kinder mit Deutsch als Erstsprache um 213 Punkte verbessern, sind es bei den Kindern mit Deutsch als Zweitsprache 277 Punkte.

Leistungszuwachs in der Mathematik

Im linken Teil der Abbildung 2.6 sind die Ergebnisse in der Mathematik ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Zu Beginn der 1. Klasse erreichen die Schülerinnen und Schüler mit Deutsch als Erstsprache 507 Punkte und jene mit Deutsch als Zweitsprache 479 Punkte. Weil die Schülerinnen und Schüler mit Deutsch als Zweitsprache einen statistisch signifikant geringeren Leistungszuwachs aufweisen (rund 4 Punkte pro Schuljahr), steigt der anfängliche Rückstand von 28 Punkten bis zum Ende der 6. Klasse auf 51 Punkte.

Abbildung 2.6: Leistungszuwachs in der Mathematik nach Erstsprache



Wie im rechten Teil der Abbildung 2.6 ersichtlich ist, zeigen sich nach statistischer Kontrolle der Lernvoraussetzungen keine Unterschiede mehr zwischen den beiden Gruppen. Zu Beginn der 1. Klasse erreichen die Schülerinnen und Schüler mit Deutsch als Erstsprache 503 Punkte und jene mit Deutsch als Zweitsprache 497 Punkte. Der Unterschied von 6 Punkten ist statistisch nicht signifikant. Auch im Leistungszuwachs zeigen sich keine statistisch signifikanten Unterschiede zwischen den beiden Gruppen. Unabhängig von der Erstsprache verbessern sich die Schülerinnen und Schüler bis zum Ende der 6. Klasse um rund 371 Punkte. Am Ende der 6. Klasse zeigen sich zwischen den beiden Gruppen entsprechend keine statistisch signifikanten Unterschiede in den Mathematikleistungen.

2.4 Leistungszuwachs nach sozialer Herkunft

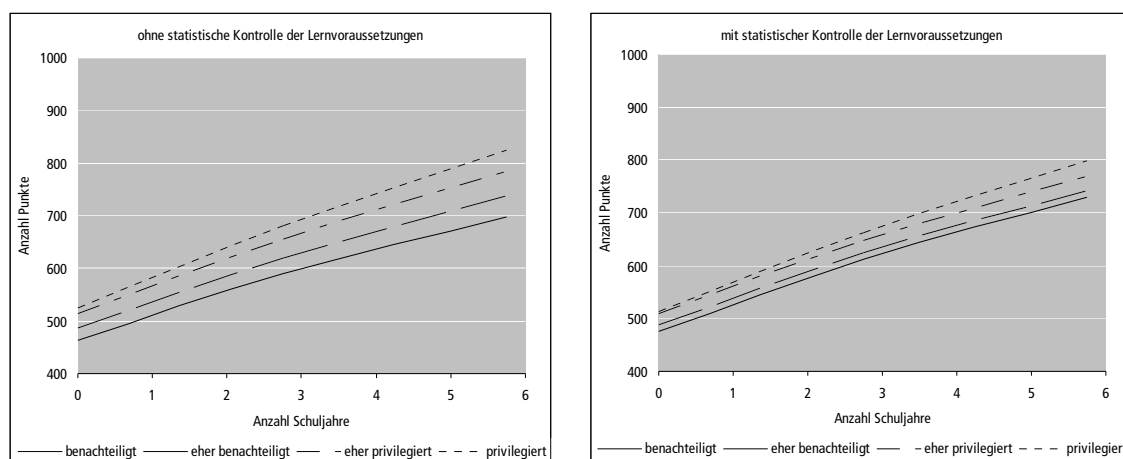
Die Abbildungen 2.7 bis 2.9 zeigen den Leistungszuwachs der Schülerinnen und Schüler vom Schuleintritt bis zum Ende der 6. Klasse nach sozialer Herkunft. Die soziale Herkunft der Schülerinnen und Schüler wurde anhand von Indikatoren zu einer Reihe von ökonomischen und kulturellen Ressourcen des Elternhauses erfasst. Für die Darstellung der Ergebnisse in den Abbildungen 2.7 bis 2.9 wurden aufgrund des Index zur sozialen Herkunft vier gleich große Gruppen gebildet: Schülerinnen und Schüler mit benachteiligter, eher benachteiligter, eher privilegierter und privilegierter sozialer Herkunft.

Leistungszuwachs im Lesen

In der linken Hälfte der Abbildung 2.7 sind die Ergebnisse im Lesen ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Bereits zu Beginn der 1. Klasse unterscheiden sich die Leseleistungen nach der sozialen Herkunft: Schülerinnen und Schüler mit benachteiligter sozialer Herkunft erreichen 463 Punkte, jene mit eher benachteiligter sozialer Herkunft 486 Punkte, jene mit eher privilegierter sozialer Herkunft 514 Punkte und jene mit privilegierter sozialer Herkunft 524 Punkte. Beim Schuleintritt beträgt der Leistungsunterschied zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft 61 Punkte. Wie aus der Abbildung 2.7 hervorgeht, steigt dieser Unterschied bis zum Ende der 6. Klasse um mehr als das Zweifache auf 127 Punkte an. Je privilegierter die soziale

Herkunft ist, desto grösser ist auch der Leistungszuwachs im Lesen. Im Vergleich zu den Schülerinnen und Schülern mit benachteiligter sozialer Herkunft ist der Leistungszuwachs der Schülerinnen und Schüler mit eher privilegierter sozialer Herkunft rund 6 Punkte pro Schuljahr, jener der Schülerinnen und Schüler mit privilegierter sozialer Herkunft rund 11 Punkte pro Schuljahr grösser. Zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit eher benachteiligter sozialer Herkunft hingegen zeigen sich keine statistisch signifikanten Unterschiede im Leistungszuwachs.

Abbildung 2.7: Leistungszuwachs im Lesen nach sozialer Herkunft



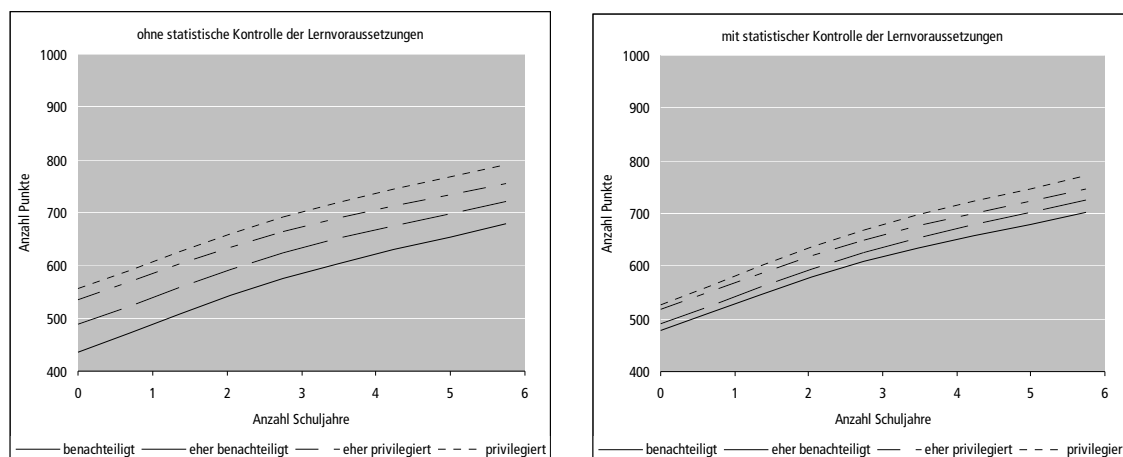
Im rechten Teil der Abbildung 2.7 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen dargestellt. Dabei zeigt sich ein ähnliches Verlaufsmuster wie bei der unkontrollierten Darstellung. Die Unterschiede zwischen den Gruppen sind jedoch deutlich geringer. Zu Beginn der 1. Klasse beträgt der Unterschied zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft 36 Punkte. Der Leistungszuwachs unterscheidet sich lediglich zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft statistisch signifikant (rund 6 Punkte pro Schuljahr) sowie zwischen den Schülerinnen und Schülern mit eher privilegierter und jenen mit privilegierter sozialer Herkunft (rund 4 Punkte pro Schuljahr). Je höher die soziale Herkunft ist, desto höher ist auch der Leistungszuwachs im Lesen. Am Ende der 6. Klasse erreichen die Schülerinnen und Schüler mit benachteiligter sozialer Herkunft 730 Punkte, jene mit eher benachteiligter sozialer Herkunft 743 Punkte, jene mit eher privilegierter sozialer Herkunft 770 Punkte und jene mit privilegierter sozialer Herkunft 799 Punkte. Die Unterschiede zwischen den Gruppen unterschiedlicher sozialer Herkunft sind statistisch signifikant. In den ersten sechs Schuljahren steigt somit der Leistungsrückstand zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft um nahezu das Zweifache auf 69 Punkte an.

Leistungszuwachs im Wortschatz

Im linken Teil der Abbildung 2.8 sind die Ergebnisse im Wortschatz ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Die Unterschiede zwischen Schülerinnen und Schü-

lern unterschiedlicher sozialer Herkunft sind zu Beginn der 1. Klasse im Wortschatz grösser als im Lesen. Der Leistungsunterschied zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft beträgt 120 Punkte. Weil sich der Leistungszuwachs zwischen den beiden Gruppen nicht statistisch signifikant unterscheidet, reduziert sich dieser Unterschied bis zum Ende der 6. Klasse nur geringfügig auf 113 Punkte.

Abbildung 2.8: Leistungszuwachs im Wortschatz nach sozialer Herkunft



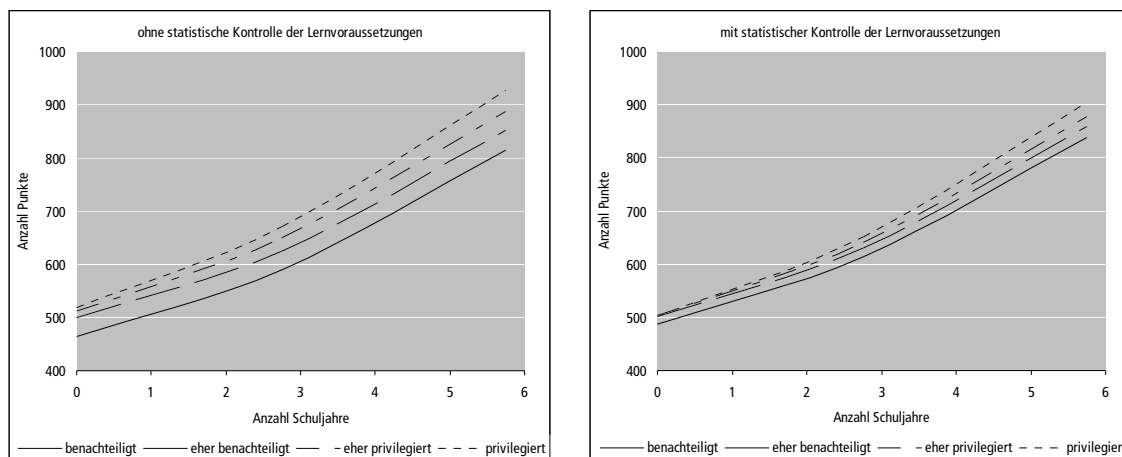
Im rechten Teil der Abbildung 2.8 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen dargestellt. Die Verlaufsmuster des Lernzuwachses ändern sich kaum im Vergleich zur unkontrollierten Darstellung. Wie beim Lesen sind die Unterschiede zwischen den Schülerinnen und Schülern unterschiedlicher sozialer Herkunft deutlich geringer ausgeprägt. Zu Beginn der 1. Klasse beträgt der Unterschied zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft 49 Punkte. Weil der Leistungszuwachs der Schülerinnen und Schüler mit privilegierter sozialer Herkunft statistisch signifikant grösser ist (rund 4 Punkte pro Schuljahr), steigt dieser Unterschied bis zum Ende der 6. Klasse auf 71 Punkte an. Auch im Wortschatz unterscheiden sich die Leistungen am Ende der 6. Klasse statistisch signifikant zwischen den Gruppen unterschiedlicher sozialer Herkunft.

Leistungszuwachs in der Mathematik

Im linken Teil der Abbildung 2.9 sind die Ergebnisse in der Mathematik ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Zu Beginn der 1. Klasse sind die Unterschiede nach sozialer Herkunft ähnlich stark ausgeprägt wie im Lesen. Die Schülerinnen und Schüler mit benachteiligter sozialer Herkunft erreichen im Vergleich zu jenen mit privilegierter sozialer Herkunft 53 Punkte weniger. Bis zum Ende der 6. Klasse steigt dieser Unterschied auf 113 Punkte an. Je privilegierter die soziale Herkunft ist, desto grösser ist auch der Leistungszuwachs. Im Vergleich zu den Schülerinnen und Schülern mit benachteiligter sozialer Herkunft ist der Leistungszuwachs der Schülerinnen und Schüler mit eher privilegierter sozialer Herkunft rund 5 Punkte pro Schuljahr, jener der Schülerinnen und Schüler mit privilegierter sozialer Herkunft rund 10 Punkte pro Schuljahr grösser. Zwischen den Schülerinnen und Schülern

mit benachteiligter und solchen mit eher benachteiligter sozialer Herkunft finden sich keine Unterschiede im Leistungszuwachs.

Abbildung 2.9: Leistungszuwachs in der Mathematik nach sozialer Herkunft



Im rechten Teil der Abbildung 2.9 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen dargestellt. Wiederum sind die Verlaufsmuster ähnlich wie bei der unkontrollierten Darstellung. Die Unterschiede zwischen den Gruppen sind jedoch deutlich geringer, insbesondere zu Beginn der 1. Klasse. Nach statistischer Kontrolle der Lernvoraussetzungen beträgt der Unterschied zwischen den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft lediglich 14 Punkte. Die Unterschiede zwischen den Gruppen vergrößern sich jedoch bis zum Ende der 6. Klasse. Die Schülerinnen und Schüler mit eher privilegierter und jene mit privilegierter sozialer Herkunft weisen im Vergleich zu den Schülerinnen und Schülern mit benachteiligter sozialer Herkunft einen statistisch signifikant grösseren Leistungszuwachs auf (4 beziehungsweise 9 Punkte pro Schuljahr). Bis zum Ende der 6. Klasse verbessern sich die Schülerinnen und Schüler mit benachteiligter sozialer Herkunft mit 351 Punkten um 53 Punkte weniger als jene mit privilegierter sozialer Herkunft (404 Punkte). Die Leistungsunterschiede zwischen den Gruppen unterschiedlicher sozialer Herkunft am Ende der 6. Klasse sind statistisch signifikant.

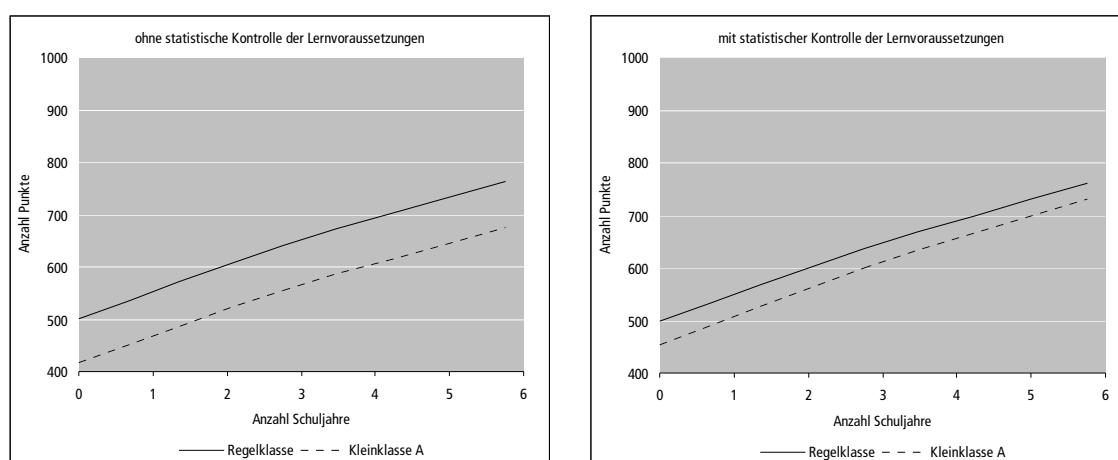
2.5 Leistungszuwachs nach Einschulungsform

Die Abbildungen 2.10 bis 2.12 zeigen den Leistungszuwachs der Schülerinnen und Schüler der Regelklassen und der Kleinklassen A vom Schuleintritt bis zum Ende der 6. Klasse beziehungsweise der 5. Klasse. In der Kleinklasse A wird der Unterrichtsstoff der 1. Klasse auf zwei Jahre verteilt, weshalb die Schülerinnen und Schüler nach sechs Jahren in der Regel die 5. Klasse abschliessen. Diese Schülerinnen und Schüler hatten zum Erhebungszeitpunkt die Unterrichtsinhalte der 6. Klasse des Lehrplans noch nicht behandelt.

Leistungszuwachs im Lesen

Im linken Teil der Abbildung 2.10 sind die Ergebnisse im Lesen ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Zu Beginn der 1. Klasse schneiden die Schülerinnen und Schüler der Regelklassen mit 500 Punkten statistisch signifikant besser ab als jene der Kleinklassen A mit 418 Punkten. Nach sechs Schuljahren erzielen die Schülerinnen und Schüler der Regelklassen 764 Punkte und die Schülerinnen und Schüler, die beim Schuleintritt die Kleinklasse A besuchten, 676 Punkte. Der Unterschied von 88 Punkten ist statistisch signifikant. Der Leistungszuwachs hingegen ist für beide Gruppen gleich gross.

Abbildung 2.10: Leistungszuwachs im Lesen nach Einschulungsform

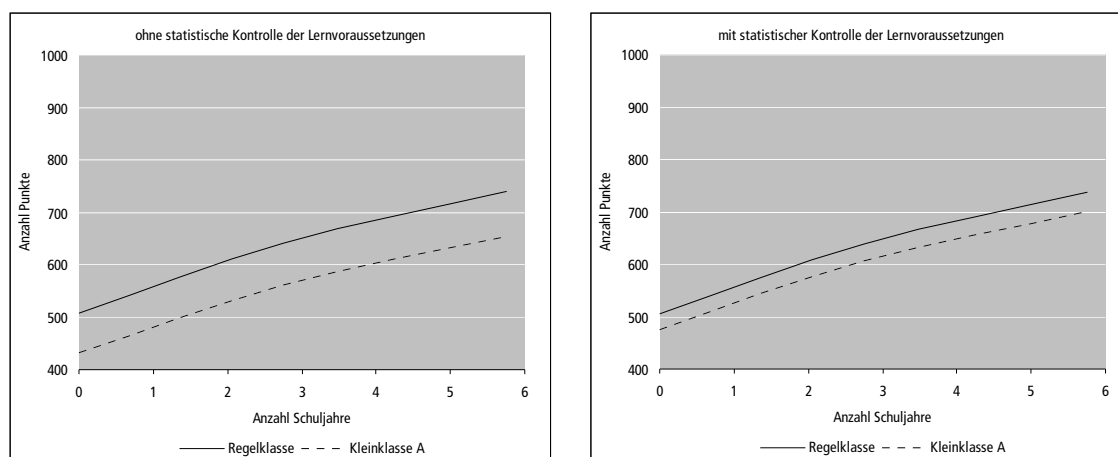


Im rechten Teil der Abbildung 2.10 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen dargestellt. Die Kurvenverläufe ändern sich kaum im Vergleich zur unkontrollierten Darstellung. Die Unterschiede zu Beginn der 1. Klasse sind bei gleichen Lernvoraussetzungen erwartungsgemäss geringer. Die Schülerinnen und Schüler der Regelklassen erreichen 498 Punkte und jene der Kleinklassen A 435 Punkte. Nach sechs Schuljahren erreichen die Schülerinnen und Schüler der Regelklassen 762 Punkte gegenüber 732 Punkten der Schülerinnen und Schüler, die beim Schuleintritt die Kleinklasse A besuchten. Der Leistungszuwachs unterscheidet sich in den sechs Jahren zwischen den beiden Gruppen statistisch nicht signifikant.

Leistungszuwachs im Wortschatz

Im linken Teil der Abbildung 2.11 sind die Ergebnisse im Wortschatz ohne statistische Kontrolle der Lernvoraussetzungen dargestellt. Zu Beginn der 1. Klasse sind die Leistungsunterschiede zwischen den Schülerinnen und Schülern der Regelklassen und jenen der Kleinklassen A im Wortschatz nur geringfügig kleiner als im Lesen. Die Kinder der Regelklassen erreichen 507 Punkte, jene der Kleinklassen A 432 Punkte. Nach sechs Schuljahren erreichen die Kinder der Regelklassen 740 Punkte gegenüber 655 Punkten der Schülerinnen und Schüler, die beim Schuleintritt die Kleinklasse A besuchten. Der Leistungszuwachs unterscheidet sich statistisch nicht signifikant zwischen den beiden Gruppen.

Abbildung 2.11: Leistungszuwachs im Wortschatz nach Einschulungsform



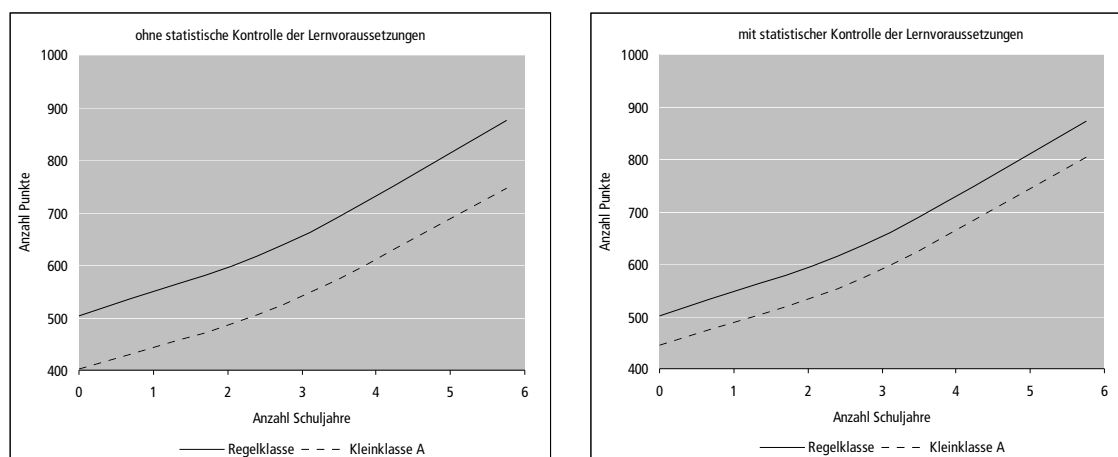
Im rechten Teil der Abbildung 2.11 sind die Ergebnisse mit statistischer Kontrolle der Lernvoraussetzungen abgebildet. Dabei zeigt sich wiederum ein ähnliches Verlaufsmuster wie bei der unkontrollierten Darstellung. Die Unterschiede zwischen den beiden Gruppen fallen aber geringer aus. Zu Beginn der 1. Klasse erreichen die Schülerinnen und Schüler der Regelklassen 505 Punkte und jene der Kleinklassen A 475 Punkte. Bis zum Ende der 6. Klasse verbessern sich die Schülerinnen und Schüler der Regelklassen um 233 Punkte, jene der Kleinklassen A bis zum Ende der 5. Klasse um 226 Punkte. Wiederum unterscheidet sich der Leistungszuwachs in den sechs Jahren zwischen den beiden Gruppen statistisch nicht signifikant.

Leistungszuwachs in der Mathematik

Der linke Teil der Abbildung 2.12 zeigt die Ergebnisse in der Mathematik ohne statistische Kontrolle der Lernvoraussetzungen. Zu Beginn der 1. Klasse erreichen die Schülerinnen und Schüler der Regelklassen 503 Punkte und jene der Kleinklassen A 403 Punkte. Der beträchtliche Unterschied von 100 Punkten steigt im Verlaufe der Primarschule auf 129 Punkte an, weil die Schülerinnen und Schüler der Regelklassen einen vergleichsweise grösseren Leistungszuwachs aufweisen (rund 5 Punkte pro Schuljahr). Nach sechs Schuljahren erreichen die Schülerinnen und Schüler der Regelklassen 876 Punkte und jene, die beim Schuleintritt die Kleinklasse A besuchten, 747 Punkte.

Im rechten Teil der Abbildung 2.12 sind die Ergebnisse in der Mathematik mit statistischer Kontrolle der Lernvoraussetzungen abgebildet. Zu Beginn der 1. Klasse erzielen die Schülerinnen und Schüler der Regelklassen 501 Punkte gegenüber 445 Punkten der Schülerinnen und Schüler der Kleinklassen A. Nach der statistischen Kontrolle der Lernvoraussetzungen zeigen sich im Hinblick auf den Leistungszuwachs keine statistisch signifikanten Unterschiede zwischen den beiden Gruppen. Nach sechs Schuljahren erreichen die Schülerinnen und Schüler der Regelklassen 873 Punkte und jene, die beim Schuleintritt die Kleinklasse A besuchten, 803 Punkte. Der Unterschied von 70 Punkten ist statistisch signifikant.

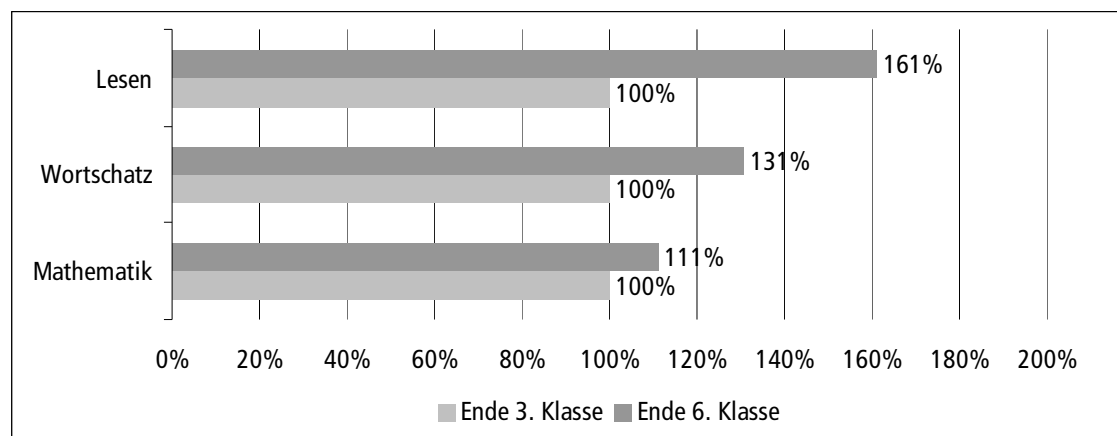
Abbildung 2.12: Leistungszuwachs in der Mathematik nach Einschulungsform



2.6 Leistungszuwachs auf der Mittelstufe

Der Leistungszuwachs unterscheidet sich nach verschiedenen individuellen Merkmalen. Weil die Testinstrumente für die 3. und 6. Klasse miteinander verbunden wurden, kann berechnet werden, wie sich die Leistungen insgesamt verändern. Abbildung 2.13 zeigt die Zunahme der Leistungsunterschiede (Streuung) zwischen den beiden Testzeitpunkten in Prozent.

Abbildung 2.13: Zunahme der Streuung der Leistungen vom Ende der 3. Klasse bis zum Ende der 6. Klasse



Die Streuung der Leistungen nimmt zwischen der 3. und 6. Klasse zu. In allen drei Bereichen sind die Leistungsunterschiede zwischen den Schülerinnen und Schülern auf der Mittelstufe grösser geworden. Am stärksten ist die Zunahme im Lesen (plus 61 Prozent). Eher gering ist die Zunahme in der Mathematik (plus 11 Prozent).

Die Zunahme der Leistungsunterschiede lässt sich auch mit anderen Massen veranschaulichen. Die Differenz zwischen den 5 Prozent schwächsten und den 5 Prozent stärksten Schüle-

rinnen und Schülern beträgt im Lesen am Ende der 3. Klasse 332 Punkte, am Ende der 6. Klasse 407 Punkte. Im Wortschatz beträgt diese Differenz am Ende der 3. Klasse 326 Punkte, am Ende der 6. Klasse 365 Punkte und in der Mathematik am Ende der 3. Klasse 335 Punkte, am Ende der 6. Klasse 349 Punkte.

2.7 Leistungen am Ende der 6. Klasse nach Vorwissen

Neben den individuellen Merkmalen Geschlecht, Erstsprache, soziale Herkunft und kognitive Grundfähigkeiten ist für den Leistungszuwachs auf der Primarstufe auch das Vorwissen von Bedeutung. Wer bereits beim Schuleintritt über gute Kompetenzen im Lesen und im Wortschatz sowie in der Mathematik verfügt, erreicht am Ende der 6. Klasse bessere Leistungen. Dieser Zusammenhang unterscheidet sich je nach Kompetenzbereich. Damit die Bedeutung des Vorwissens interpretiert werden kann, wurde untersucht, wie wichtig die individuellen Merkmale für die Leistungen am Ende der 6. Klasse im Vergleich zum Vorwissen sind.

Abbildung 2.14: Erklärung der Leistungsunterschiede am Ende der 6. Klasse: Vorwissen beim Schuleintritt

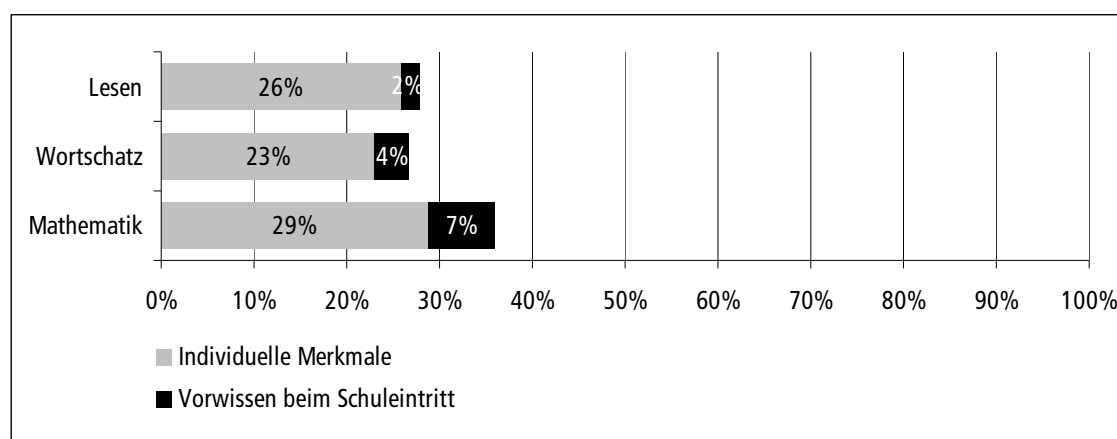
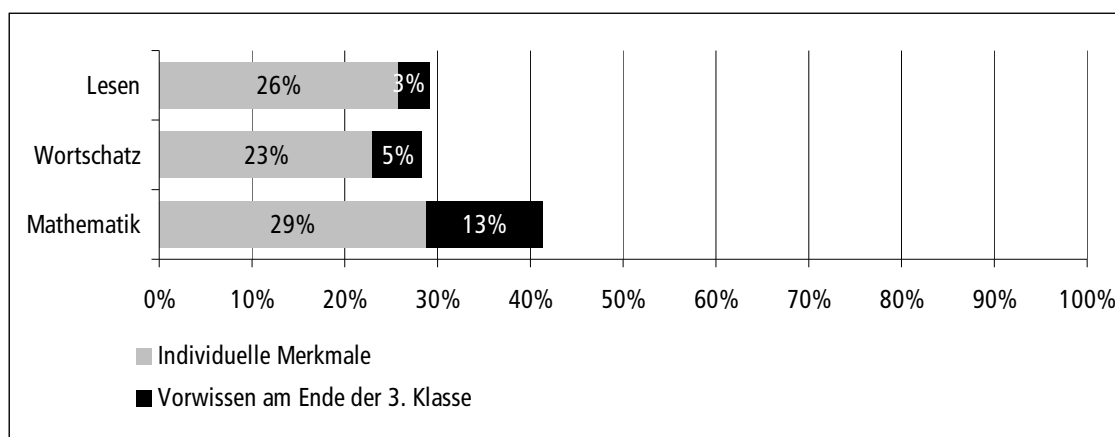


Abbildung 2.14 zeigt, dass die individuellen Merkmale Geschlecht, Erstsprache, soziale Herkunft und kognitive Grundfähigkeiten beim Schuleintritt im Lesen 26 Prozent, im Wortschatz 23 Prozent und in der Mathematik 29 Prozent der Leistungsunterschiede zwischen den Schülerinnen und Schülern am Ende der 6. Klasse erklären. Im Vergleich dazu ist die zusätzliche Bedeutung des Vorwissens beim Schuleintritt eher gering. Bei gleichen Lernvoraussetzungen erklärt das Vorwissen beim Schuleintritt rund 2 Prozent der Leistungen im Lesen am Ende der 6. Klasse. Das Vorwissen im Wortschatz erklärt 4 Prozent der Wortschatzleistungen am Ende der 6. Klasse und das Vorwissen in der Mathematik 7 Prozent der Mathematikleistungen am Ende der 6. Klasse. Das Vorwissen beim Schuleintritt hat für die Mathematikleistungen die grösste, für das Lesen die geringste Bedeutung.

Abbildung 2.15: Erklärung der Leistungsunterschiede am Ende der 6. Klasse: Vorwissen am Ende der 3. Klasse



Zusätzlich wurde überprüft, welche Bedeutung das Vorwissen am Ende der 3. Klasse für die Leistungen am Ende der 6. Klasse hat. Abbildung 2.15 zeigt wiederum, dass die individuellen Merkmale Geschlecht, Erstsprache, soziale Herkunft und kognitive Grundfähigkeiten beim Schuleintritt im Lesen 26 Prozent, im Wortschatz 23 Prozent und in der Mathematik 29 Prozent der Leistungsunterschiede zwischen den Schülerinnen und Schülern am Ende der 6. Klasse erklären.

Die Bedeutung des Vorwissens am Ende der 3. Klasse ist für die Leistungen am Ende der 6. Klasse etwas grösser als die Bedeutung des Vorwissens beim Schuleintritt. Bei gleichen Lernvoraussetzungen erklärt das Vorwissen am Ende der 3. Klasse im Lesen rund 3 Prozent der Leistungen im Lesen am Ende der 6. Klasse. Das Vorwissen im Wortschatz erklärt 5 Prozent der Wortschatzleistungen am Ende der 6. Klasse und das Vorwissen in der Mathematik 13 Prozent der Mathematikleistungen am Ende der 6. Klasse.

2.8 Fazit

Die Beschreibung des Leistungszuwachses vom Schuleintritt bis zum Ende der Primarstufe verdeutlicht einmal mehr, wie stark der Lernerfolg durch die soziale Herkunft geprägt wird. Je höher die soziale Herkunft ist, desto grösser ist der Leistungszuwachs. Aufgrund dieses Zusammenhangs vergrössern sich auch die Leistungsunterschiede zwischen den Schülerinnen und Schülern unterschiedlicher sozialer Herkunft im Lesen und in der Mathematik; Leistungsunterschiede, die bereits beim Schuleintritt bestehen. Diese primären Herkunftseffekte verstärken sich im Verlaufe der Primarschule. Die Schere öffnet sich vor allem zwischen den beiden Extremgruppen: den Schülerinnen und Schülern mit benachteiligter und jenen mit privilegierter sozialer Herkunft.

Die Erstsprache der Schülerinnen und Schüler ist demgegenüber für den Leistungszuwachs im Lesen und in der Mathematik kaum von Bedeutung. Der Wortschatz der Schülerinnen und Schüler mit Deutsch als Zweitsprache ist zwar beim Schuleintritt deutlich geringer als jener der Schülerinnen und Schüler mit Deutsch als Erstsprache. Weil die Schülerinnen und Schüler

mit Deutsch als Zweitsprache im Verlaufe der Primarschule bei gleichen Lernvoraussetzungen jedoch mehr hinzulernen, verringert sich dieser Rückstand deutlich bis zum Ende der 6. Klasse. Die Darstellung der Ergebnisse nach statistischer Kontrolle der Lernvoraussetzungen zeigt, dass das Aufwachsen mit zwei Sprachen an sich nicht für den geringeren Leistungszuwachs der Schülerinnen und Schüler mit Deutsch als Zweitsprache verantwortlich gemacht werden kann, sondern die soziale Herkunft. Kinder mit Deutsch als Zweitsprache stammen überproportional aus sozial benachteiligten Verhältnissen, weshalb ihr Leistungszuwachs weniger steil verläuft als jener der Kinder mit Deutsch als Erstsprache.

Eher geringe Unterschiede zeigen sich zwischen den Knaben und Mädchen. Zu Beginn der 1. Klasse schneiden die Mädchen im Lesen und im Wortschatz besser ab als die Knaben. In der Mathematik sind hingegen die Knaben im Vorsprung. Die Unterschiede im Leistungszuwachs sind gering. Der Tendenz nach lernen die Mädchen in der Primarschule etwas mehr als die Knaben. Im Lesen wird der anfängliche Rückstand der Knaben bis zum Ende der Primarschule nur unbedeutend grösser. Die Fortschritte der Knaben und Mädchen sind nahezu gleich gross. Im Wortschatz steigt hingegen der Vorsprung der Mädchen bis zum Ende der 6. Klasse weiter an. Ihr Leistungszuwachs ist im Wortschatz statistisch signifikant grösser als jener der Knaben. Auch in der Mathematik lernen die Mädchen mehr als die Knaben. Dadurch vermögen sie bis zum Ende der 6. Klasse ihren anfänglichen Rückstand auszugleichen.

Wer beim Schuleintritt bereits über ein grosses Vorwissen verfügt, profitiert davon auch während der Primarschulzeit. Das Vorwissen hängt statistisch signifikant mit den Leistungen am Ende der Primarschule zusammen. Noch besser werden die Leistungen am Ende der Primarschule durch das Wissen am Ende der 3. Klasse erklärt. Im Vergleich zu den individuellen Merkmalen, wie soziale Herkunft oder Erstsprache, ist das Vorwissen aber deutlich weniger wichtig für den Leistungszuwachs in der Primarschule.

Wer eine Kleinklasse A besucht, benötigt in der Regel für die Primarschule sieben statt sechs Jahre. Die Eingangsleistungen der Schülerinnen und Schüler der Kleinklasse A sind zwar statistisch signifikant tiefer, der Leistungszuwachs ist hingegen nicht geringer als jener der Schülerinnen und Schüler, die in die 1. Regelklasse eintreten.

Dass die herkunftsbedingten Bildungsungleichheiten während der Primarschule eher zu-, mit Sicherheit jedoch nicht abnehmen, ist ein wichtiges Ergebnis der Längsschnittstudie. Dass die Leistungsunterschiede zwischen den Schülerinnen und Schülern bis zum Ende der Primarschule generell stark zunehmen, ist das andere wichtige Ergebnis. Die generelle Zunahme der Leistungsunterschiede ist bis zu einem gewissen Grad systembedingt, weil nach sechs Jahren über 14 Prozent der Kinder – zumindest derjenigen unserer Längsschnittstichprobe – nicht in der 6. Regelklasse unterrichtet werden. Die Heterogenität eines Schülerjahrgangs ist demzufolge grösser, als aufgrund der Leistungsstreuung innerhalb einer Klassenstufe vermutet wird. Je älter die Schülerinnen und Schüler werden, desto grösser sind die Leistungsunterschiede in den sprachlichen und mathematischen Kompetenzen. Bis zum Ende der 6. Klasse der Primarschule öffnet sich die Schere zwischen leistungsschwachen und leistungsstarken Schülerinnen und Schülern.

3 Fachleistungen am Ende der 6. Klasse

Urs Moser & Domenico Angelone

3.1 Ergebnisse nach Anforderungsniveaus

Um die Fachleistungen der Schülerinnen und Schüler am Ende der 6. Klasse mit Bezug zum Lehrplan beurteilen zu können, wurden für die beiden Fachbereiche Deutsch und Mathematik Anforderungsniveaus gebildet. Der Lehrplan für die Volksschule des Kantons Zürich diente deshalb bereits bei der Entwicklung der Testaufgaben als Referenzrahmen.

Jede Testaufgabe musste einem im Voraus bestimmten Anforderungsniveau zugeteilt werden. Dazu konnte der Lehrplan nur beschränkt genutzt werden, weil nicht jedes Ziel beziehungsweise jeder Inhalt differenziert nach der Schwierigkeit dargestellt ist. Ziele und Inhalte lassen sich auch nicht präzise einer Klassenstufe zuordnen. Unterschieden wird im Lehrplan hingegen zwischen Verbindlichkeit, Intensität und Gründlichkeit, in denen die Ziele und Inhalte bearbeitet werden. Beispielsweise wird das Sachrechnen mit Grössen in der 4. Klasse aufgegriffen. Die Schülerinnen und Schüler erkennen die Problemstellung, experimentieren und machen erste Erfahrungen mit diesem Unterrichtsinhalt. In der 5. Klasse arbeiten sich die Schülerinnen und Schüler gründlich ins Thema ein und in der 6. Klasse werden die Kenntnisse und Fertigkeiten durch Übungsphasen gefestigt. Die Testaufgaben wurden deshalb von Lehrpersonen entwickelt und mithilfe des Lehrplans vier Anforderungsniveaus zugeordnet:

Niveau I: Das Lösen der Aufgaben mit minimalen Anforderungen wird von allen Schülerinnen und Schülern am Ende der 5. Klasse erwartet. Das Lösen dieser Aufgaben genügt noch nicht, damit die Ziele des Lehrplans am Ende der 6. Klasse als «erreicht» bezeichnet werden.

Niveau II: Das Lösen der Aufgaben mit grundlegenden Anforderungen wird von allen Schülerinnen und Schülern am Ende der 6. Klasse erwartet. Das Lösen der Aufgaben ist ein Hinweis dafür, dass die Ziele im geprüften Bereich mehrheitlich erreicht werden.

Niveau III: Das Lösen der Aufgaben mit erweiterten Anforderungen ist ein Hinweis dafür, dass die Ziele des Lehrplans im geprüften Bereich vollständig erreicht werden.

Niveau IV: Das Lösen der Aufgaben mit höheren Anforderungen ist ein Hinweis dafür, dass die Ziele des Lehrplans im geprüften Bereich übertroffen werden.

Die Bildung von Anforderungsniveaus wurde anhand der Aufgabenschwierigkeit nach der Testdurchführung überprüft. Aufgaben, die im Voraus einem zu tiefen oder einem zu hohen Niveau zugeordnet worden waren, wurden entsprechend der Schwierigkeit eingeteilt.

Aufgrund der Testergebnisse lassen sich die Schülerinnen und Schüler einem der vier Anforderungsniveaus zuordnen. Auf einem Anforderungsniveau verfügen die Schülerinnen und Schüler nicht nur über die dem Niveau zugeordneten Fähigkeiten, sondern auch über die in den darunterliegenden Niveaus geforderten Fähigkeiten. Alle Schülerinnen und Schüler, die beispielsweise den Anforderungen von Niveau III genügen, werden demzufolge auch die Anforderungen von Niveau II erfüllen. Von allen Schülerinnen und Schülern, die ein bestimmtes Niveau erreichen, wird erwartet, dass sie mindestens die Hälfte der Aufgaben dieses Ni-

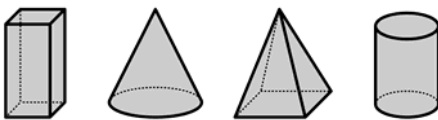
veaus richtig lösen. Aufgaben eines tieferen Niveaus werden mit einer grösseren Wahrscheinlichkeit richtig gelöst, Aufgaben eines höheren Niveaus werden mit einer kleineren Wahrscheinlichkeit richtig gelöst.

3.2 Anforderungsniveaus in Mathematik

Niveau I: Minimale Anforderungen

Schülerinnen und Schüler, die Niveau I erreichen, können Additionen und Multiplikationen mit ganzen Zahlen im Kopf oder halbschriftlich lösen. Sie kennen die schriftlichen Rechenverfahren für die Grundoperationen und lösen schriftliche Additionen mit mehreren Summanden. Einfache Sachrechnungen mit Gewichten und Geldwerten werden richtig gelöst. Die Schreibweise für Brüche beziehungsweise der Bruch als Teil einer gewählten Einheit (Grössenkonzept) ist bekannt. Die Anzahl Ecken von geometrischen Figuren wird korrekt angegeben; der Zylinder wird richtig benannt. Regeln bei einfachen arithmetischen Zahlenfolgen, wie das fortlaufende Addieren einer Zahl, die jeweils um 10 zunimmt, werden erkannt.

Aufgabenbeispiele: 650–749 Punkte

Arithmetik	$440 + 588 =$ $800 \cdot 40 =$ $134706 + 248000 + 7289 =$ $46201 + 308769 + 5008 =$
Brüche	Schreibe das passende Beziehungszeichen $<$, $>$, $=$ ins Kästchen. $\frac{1}{12} \square \frac{1}{15}$
Geometrie	Benenne die folgenden Körper und gib die Anzahl der Flächen, Kanten und Ecken an. <div style="text-align: center;">  </div>
Grössen	Eistee kann man aus Wasser und Pulver zubereiten. Man braucht 50 g Pulver pro Liter Wasser. Wie viel Eistee wird man mit 150 g Pulver zubereiten können?
Zahlenfolgen	Suche die Regel und führe die Reihe weiter. 30 40 60 90 130 1110 1130 1150 1170 1190

Niveau II: Grundlegende Anforderungen

Schülerinnen und Schüler, die Niveau II erreichen, können additive und multiplikative Grundoperationen mit ganzen Zahlen im Kopf oder halbschriftlich lösen. Sie kennen die schriftlichen Rechenverfahren für die Grundoperationen und lösen Subtraktionen mit mehreren Subtrahenden und Multiplikationen, bei denen ein Faktor höchstens zweistellig ist. Einfache Sachrechnungen mit Gewichten, Geldwerten, Längen und Zeitmassen werden richtig gelöst. Einfache proportionale Verhältnisse werden erkannt und zur Lösung von Sachaufgaben genutzt. Der Bruch als Operator (Operatorenkonzept) ist bekannt. Brüche können in Dezimalzahlen und abbrechende Dezimalzahlen in Brüche umgewandelt werden, sofern es sich um einfache Brüche wie $\frac{1}{4}$ oder Dezimalzahlen im Bereich von Zehnteln handelt. Winkelhalbierende und Mittelsenkrechte können konstruiert werden. Begriffe zur Beschreibung des Dreiecks (spitzwinklig, rechtwinklig, gleichseitig, stumpfwinklig) sind bekannt und die Pyramide wird richtig benannt. Regeln bei einfachen arithmetischen Zahlenfolgen, wie das fortlaufende Subtrahieren einer Zahl, die jeweils um 1 abnimmt, werden erkannt.

Aufgabenbeispiele: 750–849 Punkte

Arithmetik	$1034 - 543 =$ $1056 : 8 =$ $42 \cdot 6509 =$ $800780 - 432609 - 47257 =$
Brüche	$\frac{3}{5} + \frac{4}{5} =$ Schreibe die Brüche als Dezimalzahlen. $\frac{1}{2}$ $\frac{1}{4}$ Verwandle in die nächstkleinere Sorte. $\frac{1}{4}$ Fr. = $\frac{1}{8}$ km = Schreibe die Dezimalzahlen als Brüche. Kürze so weit wie möglich. $0.1 =$ $0.7 =$
Grössen	Für vierzehn Tage Ferien im Hotel muss Herr Meier 1820 Fr. bezahlen. Wie teuer kommt Frau Müllers Aufenthalt von 20 Tagen in demselben Hotel zu stehen? Auf dem Markt erhält man einen Dreierpack Blumentöpfe für 12.50 Fr. Herr Blume benötigt 12 Blumentöpfe. Wie viel bezahlt er?

Grössen



Wie viele Stunden und Minuten sind seit 3 Uhr vergangen?

Zahlenfolgen Suche die Regel und führe die Reihe weiter.

198 188 179 171 164

Niveau III: Erweiterte Anforderungen

Schülerinnen und Schüler, die Niveau III erreichen, lösen schriftlich Divisionen, bei denen der Divisor zweistellig ist, sowie additive Grundoperationen mit ganzen Zahlen, bei denen die Klammerregel «Operationen in Klammern haben Vorrang» angewendet werden muss. Operationen mit einfachen Brüchen und Dezimalzahlen können gelöst werden. Die dezimale Stellenwertschreibweise mit den Begriffen Zehntel, Hundertstel und Tausendstel ist bekannt. Komplexere Sachrechnungen mit Grössen, Brüchen und Dezimalzahlen können gelöst werden, ebenso Sachrechnungen, deren Lösung das Erkennen der direkten Proportionalität erfordert. Komplexere Brüche im Bereich von Hundertsteln und Tausendsteln können in Dezimalzahlen umgewandelt werden. Abbrechende Dezimalzahlen im Bereich von Hundertsteln und Tausendsteln können in Brüche umgewandelt werden. Winkel können mithilfe des Zirkels übertragen werden; der Kegel wird richtig benannt. Regeln bei komplexeren arithmetischen Zahlenfolgen werden erkannt.

Aufgabenbeispiele: 850–949 Punkte

Arithmetik $113624 : 56 =$
 $450820 - (13277 + 148278) =$

Brüche $4.5 + 2\frac{3}{4} =$
 $6.3 : 100 =$
Schreibe die Brüche als Dezimalzahlen.

$$\frac{11}{50} = \frac{17}{1000} = \frac{4}{25} = \frac{1}{20} = \frac{14}{25} =$$

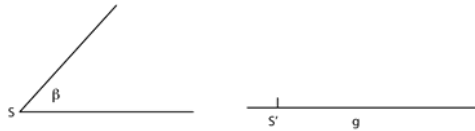
Schreibe die Dezimalzahlen als Brüche. Kürze so weit wie möglich.

$$0.031 = \quad 0.4 =$$

Verwandle in die nächstkleinere Sorte.

$$3\frac{3}{4} \text{Fr.} =$$

Geometrie Übertrage mithilfe des Zirkels den Winkel β auf die Gerade g .



Grössen 750 g Kaffee kosten 16.65 Fr. Wie viel kostet 1 kg derselben Sorte?
Vier Knaben haben beim Fussballspielen eine Fensterscheibe eingeschlagen. An den Schaden von 540 Fr. bezahlt der «Schütze» einen Drittel. Den Rest teilen die drei Mitspieler unter sich auf. Wie viel bezahlt somit jeder der drei?

Zahlenfolgen Suche die Regel und führe die Reihe weiter.
12 4 24 8 48

Niveau IV: Höhere Anforderungen

Schülerinnen und Schüler, die Niveau IV erreichen, lösen komplexe schriftliche Grundoperationen auch dann, wenn Klammerregeln zu befolgen sind und der Term Dezimalzahlen, Brüche oder Masse enthält. Einfache Brüche werden von den Schülerinnen und Schülern gleichnamig gemacht und danach addiert; Brüche werden mit natürlichen Zahlen multipliziert. Brüche in Dezimalzahlen und abbrechende Dezimalzahlen in Brüche umzuwandeln gelingt bei erhöhtem Schwierigkeitsgrad. Die Schülerinnen und Schüler lösen anspruchsvolle Sachaufgaben, denen die direkte oder indirekte Proportionalität zugrunde liegt. In der Geometrie sind die Schülerinnen und Schüler fähig, Dreiecke mit dem Zirkel zu konstruieren. Winkel an geometrischen Formen werden erkannt und können mit dem Geodreieck gemessen werden. Komplexe arithmetische Zahlenfolgen, bei denen zwei Aspekte berücksichtigt werden müssen (die Summe der beiden vorhergehenden Zahlen beziehungsweise Fibonacci-Zahlen), werden erkannt.

Aufgabenbeispiele: ab 950 Punkte

Arithmetik $765321 - 8915 - 777 - 63 - 248867 =$
 $(272376 : 78) \cdot 39 =$

$$27\text{h}32\text{min} + 46\text{min} - 17\frac{5}{12}\text{h}$$

$$95.6\text{m}^2 + 134\frac{2}{5}\text{m}^2 - 90\text{m}^28\text{dm}^2$$

$$\left(11 \cdot 3\frac{2}{5}\text{kg}\right) + \square = 132.5\text{kg} - 27\frac{5}{8}\text{kg} - 19\text{kg}75\text{g}$$

Brüche

$$8\frac{3}{7} - \frac{5}{7} =$$

Schreibe die Dezimalzahlen als Brüche. Kürze so weit wie möglich.

$$0.375 = \quad 0.08 =$$

Verwandle in die nächstkleinere Sorte.

$$5\frac{7}{8} \text{ kg} =$$

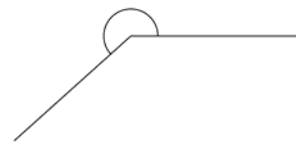
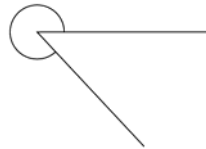
$$2.05 = \frac{7}{100} + \frac{1}{2} =$$

$$6 \cdot 1\frac{1}{3} - (10 - 8.3) =$$

Geometrie

Konstruiere das Dreieck ABC mit den Seitenlängen $\overline{AB} = 3.6 \text{ cm}$, $\overline{BC} = 1.8 \text{ cm}$, $\overline{AC} = 3 \text{ cm}$.

Miss die folgenden Winkel mit dem Geodreieck und schreibe ins graue Feld, wie gross sie sind.



Grössen

Guido und Isabel fahren einander auf dem Velo entgegen. Guido startet in Zürich und Isabel gleichzeitig im 36 km entfernten Rapperswil. Guido fährt mit einer durchschnittlichen Geschwindigkeit von 21 km/h, Isabel mit einer solchen von 24 km/h. Wie weit sind sie nach 10 Minuten voneinander entfernt?

In einer Fabrik brauchen 18 Roboter zur Herstellung einer bestimmten Anzahl Autos 37 Tage. Wie lange dauert es, wenn nach 12 Tagen 8 Roboter ausfallen?

Rolf wirft von der Quaibrücke in Zürich ein Holzstück in die Limmat und verfolgt es bis zur Walchebrücke. Für diese Fahrt benötigt das «Schifflein» 16 Minuten. Auf dem Stadtplan, der im Massstab 1:20000 gezeichnet ist, sind die beiden Brücken 64 mm voneinander entfernt. Gib die Fliessgeschwindigkeit der Limmat in km/h an.

Zahlenfolgen

Suche die Regel und führe die Reihe weiter.

1 2 6 24

0 3 3 6 9 15

1137 1197 399 459 153 213

3.3 Anforderungsniveaus in Deutsch

Niveau I: Minimale Anforderungen

Schülerinnen und Schüler, die Niveau I erreichen, können aus kurzen, einfachen Texten Informationen ermitteln, die im Text ausdrücklich erwähnt sind. Das heisst, dass sie im Text wörtliche Informationen finden. Werden die Schülerinnen und Schüler aufgefordert, Wortarten im Text zu bestimmen, dann gelingt dies bei Nomen. Kommas zwischen Teilsätzen werden dann richtig gesetzt, wenn die Aufgabe spezifisch danach ausgerichtet ist. Der Wortschatz beschränkt sich auf Begriffe aus dem schulischen Kontext. Die Inhalte und Ziele des Lehrplans werden von den Schülerinnen und Schülern dann erreicht, wenn die Texte sehr einfach sind, wenn sich der Wortschatz auf den Alltag beschränkt und wenn die Aufgabe sehr spezifisch gestellt und das Erkennen eines Wortes, einer Textstelle oder einer einfachen Regel verlangt wird. Diese klar definierten rezeptiven Fähigkeiten können bei einer produktiven Aufgabenstellung zu den gleichen Themen noch nicht genutzt werden.

Aufgabenbeispiele: 555–654 Punkte

Sprachbetrachtung	Setze bei den folgenden Sprichwörtern und Redensarten das Komma zwischen die beiden Teilsätze. Hat es keines, so markiere das Feld rechts.
	Wer zuletzt lacht lacht am besten. <input type="checkbox"/>
	Wenn zwei sich streiten freut sich der dritte. <input type="checkbox"/>

Wortschatz	Streiche in jeder Zeile das Wort durch, das nicht zu den andern passt.
	läuten, klingeln, tönen, erschallen, dröhnen, spotten
	krumm, bogenförmig, gebogen, gerade, rund

Niveau II: Grundlegende Anforderungen

Schülerinnen und Schüler, die Niveau II erreichen, können aus kurzen, einfachen Texten Informationen dazu nutzen, Fragen zum Inhalt des Textes zu beantworten, die über die Identifizierung von wörtlichen Informationen hinausgehen. Das heisst, dass sie Informationen aus dem Text nutzen, um auf eine Fragestellung die passende Antwort aus verschiedenen Vorgaben zu finden. Der Wortschatz beschränkt sich auf den schulnahen und alltäglichen Kontext, ist aber in diesen Bereichen so differenziert, dass Synonyme erkannt werden. Schülerinnen und Schüler, die Niveau II erreichen, beherrschen die Zeichensetzung bei der direkten Rede. Das heisst, dass sie nach einer Ankündigung einen Doppelpunkt und danach die direkte Rede zwischen Anfangs- und Schlusszeichen setzen. Zudem werden Punkt, Ausrufe- und Fragezeichen bei den entsprechenden Satzarten gesetzt. Es gelingt den Schülerinnen und Schülern, ihre Kenntnisse über Rechtschreibregeln für die Überarbeitung eines Textes zu nutzen. Diese Fähigkeit beschränkt sich allerdings auf Verdoppelungen.

Aufgabenbeispiele: 655–754 Punkte

Sprachbetrachtung	Setze im nachfolgenden Witz die Satzzeichen. Die Lehrerin fragt Weswegen bewundern wir die alten Römer Weil sie fließend Latein sprechen antwortet ein Kind. Bestimme die Wortart. RAUBEN SIE DICH AUS UND DU <u>MERKST</u> ES GAR NICHT. DER IHM UNTER <u>ÄNGSTLICHEN</u> GEBÄRDEN DIE UHR AUSLIEFERT. Bestimme die Zeitform der unterstrichenen Verben, indem du das richtige Kästchen ankreuzt. Allerdings <u>kann</u> ich kein Billett lösen. Ich <u>lief</u> bis in die hinterste Ecke des Busses. Es hat eine richtige Verfolgungsjagd <u>gegeben</u> .
-------------------	--

Texte überarbeiten	häte → hätte / Riter → Ritter / kent → kennt
--------------------	--

Wortschatz	Streiche in jeder Zeile das Wort durch, das nicht zu den andern passt. Hinweise, Tipp, Ratschlag, Auskunft, Erklärung, Anfrage Gang, Korridor, Zimmer, Diele, Flur, Durchgang
------------	---

Niveau III: Erweiterte Anforderungen

Schülerinnen und Schüler, die Niveau III erreichen, können aus literarischen Texten Informationen ermitteln, die im Text umschrieben sind. Aus Erzählungen und Sachtexten können einfache Schlussfolgerungen gezogen werden, indem Informationen aus verschiedenen Textteilen miteinander verbunden und somit Zusammenhänge zwischen verschiedenen Textstellen hergestellt werden. Der Text wird zur Beantwortung von Fragen genutzt. Der (passive) Wortschatz ist differenziert und beschränkt sich nicht mehr auf den Alltag beziehungsweise auf den schulischen Kontext. Wortarten und Zeitformen werden bei offener Fragestellung erkannt. Das bedeutet beispielsweise, dass in einem Text die Verben in einer bestimmten Zeitform erkannt werden. Schülerinnen und Schüler, die Niveau III erreichen, korrigieren die Rechtschreibfehler in einem Text (ck, tz, v/f) und wenden ihr grammatikalisches Wissen zur Überarbeitung eines Textes an.

Aufgabenbeispiele: 755–854 Punkte

Sprachbetrachtung Die Verben im Text stehen in verschiedenen Zeitformen. Suche im Text vier Verben, die in den folgenden Zeitformen stehen. Führe jeweils auch die Grundform an.

Texte überarbeiten Lüke → Lücke / setzte → setzte / einferstanden → einverstanden
bekam ich angst → bekam ich Angst
im stau stehen müssen → im Stau stehen müssen

Wortschatz Suche jeweils das Wort, das am ehesten die gleiche Bedeutung wie das fettgedruckte Wort hat.

dreist frech, tapfer, fest, schwach, zaghaft

übertölpeln überlisten, überfordern, überreden, überfallen

Wo werden die untenstehenden Dinge versorgt? Folgende Wörter stehen zur Auswahl:

Armbrust, Depot, Etui, Film, Flugsteig, Garage, Gürtel, Haltestelle, Hangar, Hülle, Keller, Köcher, Optiker, Portmonee, Post, Scheide

Fotoapparat _____

Brille _____

Niveau IV: Höhere Anforderungen

Schülerinnen und Schüler, die Niveau IV erreichen, verstehen komplexe Sachtexte und beantworten Fragen zu den Texten, die das Verbinden mehrerer Textteile und schlussfolgerndes Denken aufgrund der Nutzung verschiedener Informationen verlangen. Der Wortschatz ist differenziert, sodass auch dem Alter entsprechende literarische Texte verstanden werden. Die grammatikalischen Kenntnisse wie über Vergleichsformen bei Adjektiven oder Vergangenheitsformen sind vorhanden und werden bei der Überarbeitung von Texten – wie auch die Regeln der Rechtschreibung – angewendet. Kommas werden vor Konjunktionen gesetzt, die zwei Teilsätze verbinden. Die Satzzeichen bei der direkten Rede werden bei den vier Satzarten (Aussagesatz, Fragesatz, Ausrufesatz, Befehlssatz) richtig gesetzt.

Aufgabenbeispiele: 855–954 Punkte

Sprachbetrachtung

SAUBERER DACHS

Ein junger Dachs hat in England die Schleudertour in einer Waschmaschine überlebt. Das Tier war in Petersfield durch die Katzenklappe in eine Küche gelangt und hatte es sich in der Waschmaschinentrommel gemütlich gemacht. Die Hausfrau, die davon nichts bemerkt hatte, startete darauf die Maschine. Als sie schliesslich die saubere Wäsche herausholte, kroch ihr ein nasses, erschöpftes Bündel entgegen. Ein Tierarzt stellte erstaunt fest: Der sechs Monate alte Dachs hat die 60-Grad-Wäsche und den Schleudergang von 800 Umdrehungen pro Minute unverletzt überlebt.

Suche ein Adjektiv, bei dem die drei Vergleichsformen sinnvoll gebildet werden können. Bilde die drei Vergleichsformen.

Suche im Text ein Verb, das im Plusquamperfekt (Vorvergangenheit) steht.

Texte überarbeiten

Ich war traurig den der letzte Tag → traurig, denn der ...

Wir müssen gehen rief Carlo nun → «Wir müssen gehen!» rief Carlo nun.

Zum Glück war Carlo dabei den er → ... dabei, denn er ...

Wortschatz

Streiche in jeder Zeile das Wort durch, das nicht zu den andern passt.

stöhnen, jauchzen, wimmern, ächzen, winseln, seufzen

heiter, lustig, freundlich, aufgeräumt, übermütig, freudig

Wo werden die untenstehenden Dinge versorgt?

Folgende Wörter stehen zur Auswahl:

Armbrust, Depot, Etui, Film, Flugsteig, Garage, Gürtel, Haltestelle, Hangar, Hülle, Keller, Köcher, Optiker, Portmonee, Post, Scheide

Schwert _____

Pfeil _____

Ergänze die Verben mit einem Anfangsbaustein (einer Vorsilbe), sodass der ganze Satz Sinn macht.

Wer wird nun für diese Kosten _____ kommen?

Die Gäste setzen sich zu Tisch, man kann _____ richten.

3.4 Verteilung der Schülerinnen und Schüler nach den Anforderungsniveaus

Abbildung 3.1 zeigt die Verteilung der Schülerinnen und Schüler nach den Anforderungsniveaus. 17 Prozent beziehungsweise 18 Prozent der Schülerinnen und Schüler erreichen in Deutsch beziehungsweise in Mathematik lediglich Niveau I und erfüllen somit im geprüften Bereich des Lehrplans die Ziele noch nicht vollständig. Die Lücken können für Mathematik aufgrund der detaillierten Beschreibung der Inhalte im Lehrplan eindeutiger eruiert werden als für Deutsch. Beispielsweise werden schriftliche Rechenverfahren für die Grundoperationen erst zum Teil beherrscht; insbesondere die Division bereitet noch Mühe. In Deutsch werden beispielsweise Lehrplanvorgaben zu den grammatikalischen Kenntnissen nicht erreicht und einfache Kommaregeln nicht wie erwartet angewendet.

Rund 35 Prozent beziehungsweise 36 Prozent der Schülerinnen und Schüler erreichen in Deutsch beziehungsweise in Mathematik Niveau II und somit die Ziele des Lehrplans auf der Stufe der Grundanforderungen. Weitere 32 Prozent beziehungsweise 31 Prozent erreichen Niveau III und erfüllen die Ziele des Lehrplans auf der Stufe der erweiterten Anforderungen. 16 Prozent beziehungsweise 15 Prozent erreichen Niveau IV und übertreffen die Ziele des Lehrplans.

Abbildung 3.1: Verteilung der Schülerinnen und Schüler nach den Anforderungsniveaus

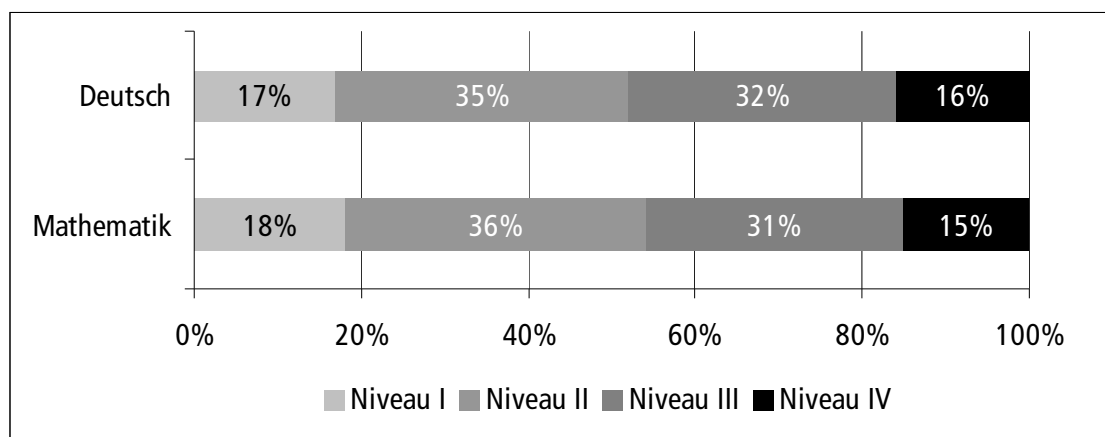


Tabelle 3.1 zeigt, wie gross der Anteil an Schülerinnen und Schülern ist, die in beiden Fächern das gleiche Niveau erreichen, beziehungsweise wie gross die Anteile der Schülerinnen und Schüler sind, die in einem der Fächer höhere beziehungsweise tiefere Anforderungen erfüllen.

Rund 8 Prozent der Schülerinnen und Schüler erreichen in Deutsch und in Mathematik Niveau I, rund 16 Prozent Niveau II, rund 15 Prozent Niveau III und rund 7 Prozent Niveau IV. Insgesamt erreichen rund 46 Prozent der Schülerinnen und Schüler in Deutsch und Mathematik das gleiche Niveau.

Rund 8 Prozent der Schülerinnen und Schüler erreichen beispielsweise in Mathematik Niveau II und in Deutsch Niveau I oder umgekehrt. Insgesamt liegen bei rund 45 Prozent der Schüle-

rinnen und Schüler die Fähigkeiten in den beiden Fachbereichen um ein Niveau auseinander. Bei rund 7 Prozent liegen die Fähigkeiten um zwei Niveaus auseinander.

Tabelle 3.1: Verteilung nach Niveaus in Deutsch und Mathematik

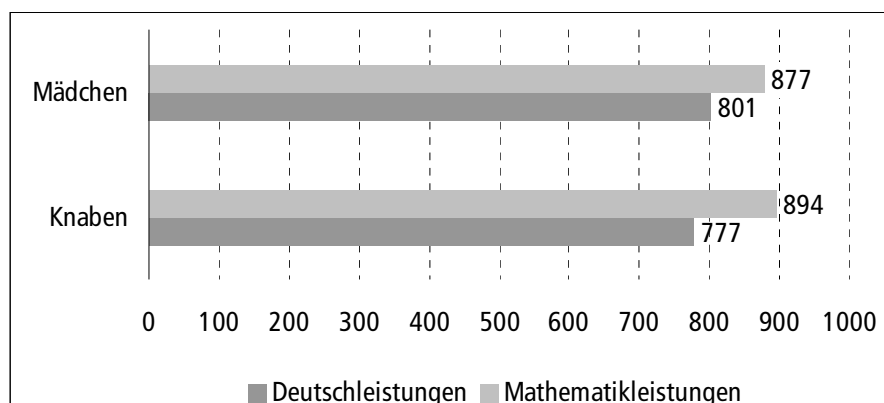
	Mathematik				
Deutsch	Niveau I	Niveau II	Niveau III	Niveau IV	
Niveau I	8.0%	7.7%	1.6%	0.0%	17.3%
Niveau II	8.0%	16.2%	8.3%	1.9%	34.4%
Niveau III	1.3%	10.2%	15.2%	5.6%	32.3%
Niveau IV	0.0%	2.6%	6.4%	7.0%	16.0%
	17.3%	36.7%	31.5%	14.5%	100.0%

3.5 Leistungen nach individuellen Merkmalen

Leistungen nach Geschlecht

In Abbildung 3.2 sind die Leistungen in Deutsch und Mathematik am Ende der 6. Klasse nach dem Geschlecht dargestellt. In Deutsch erreichen die Mädchen 801 Punkte und die Knaben 777 Punkte. Die Differenz von 24 Punkten ist als klein zu beurteilen. In Mathematik erreichen die Knaben bessere Leistungen als die Mädchen. Der Rückstand der Mädchen beträgt 17 Punkte. Auch in Mathematik ist der Unterschied zwischen den Geschlechtern als gering zu beurteilen.

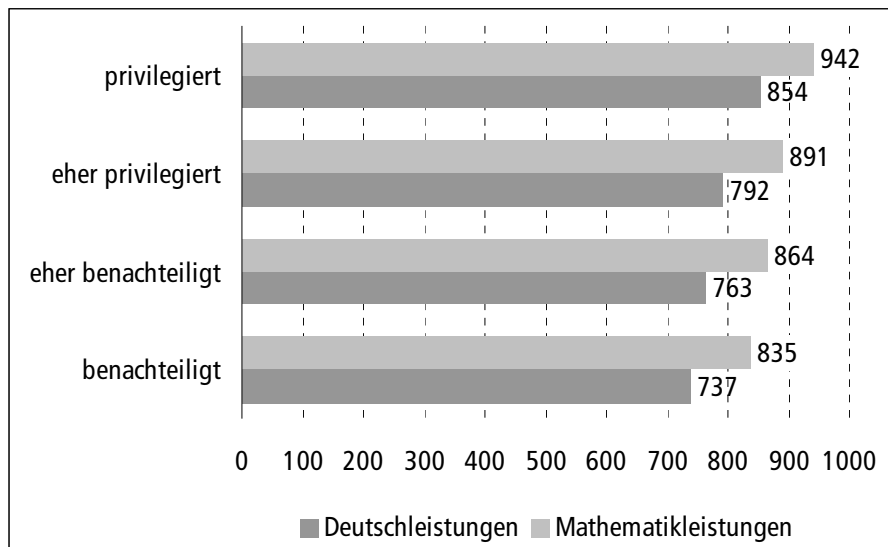
Abbildung 3.2: Leistungen am Ende der 6. Klasse nach Geschlecht



Leistungen nach sozialer Herkunft

Abbildung 3.3 zeigt die Leistungen in Deutsch und in der Mathematik am Ende der 6. Klasse nach der sozialen Herkunft. Die Schülerinnen und Schüler wurden aufgrund ihrer sozialen Herkunft in vier gleich grosse Gruppen eingeteilt.

Abbildung 3.3: Leistungen am Ende der 6. Klasse nach sozialer Herkunft

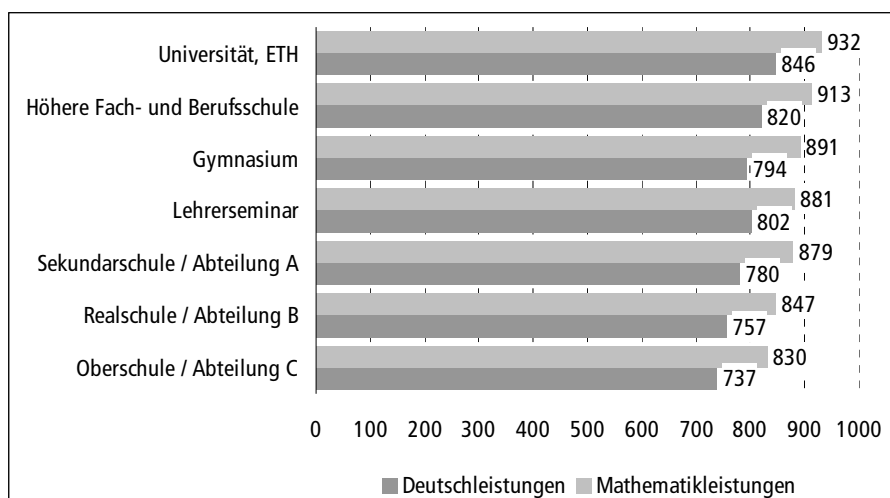


In Deutsch beträgt der Mittelwert der Schülerinnen und Schüler mit privilegierter sozialer Herkunft 854 Punkte, der Mittelwert der Schülerinnen und Schüler mit benachteiligter sozialer Herkunft 737 Punkte. Der Unterschied zwischen den beiden Gruppen ist mit 117 Punkten als sehr gross zu beurteilen ($d = 1.2$). In der Mathematik erreichen die Schülerinnen und Schüler mit privilegierter sozialer Herkunft 942 Punkte, die Schülerinnen und Schüler mit benachteiligter sozialer Herkunft 835 Punkte. Auch in der Mathematik sind die Leistungsunterschiede zwischen den beiden nach der sozialen Herkunft gebildeten Gruppen als sehr gross zu beurteilen (107 Punkte, $d = 1.1$).

Leistungen nach dem Bildungsabschluss der Eltern

Ein wichtiger Indikator für die soziale Herkunft der Schülerinnen und Schüler ist der Bildungsabschluss der Eltern. Abbildung 3.4 zeigt die Leistungen in Deutsch und Mathematik am Ende der 6. Klasse nach dem höchsten Bildungsabschluss der Eltern. Wie erwartet, zeigt sich ein positiver Zusammenhang zwischen den Leistungen der Schülerinnen und Schüler und dem formalen Bildungsabschluss der Eltern. Je höher der Bildungsabschluss der Eltern ist, desto höher sind die Leistungen der Schülerinnen und Schüler. In Deutsch erreichen beispielsweise die Schülerinnen und Schüler von Eltern mit einem universitären Abschluss 846 Punkte, jene von Eltern mit dem niedrigsten Abschluss der Sekundarschule (Abteilung C) 737 Punkte. In der Mathematik zeigt sich ein nahezu gleiches Bild.

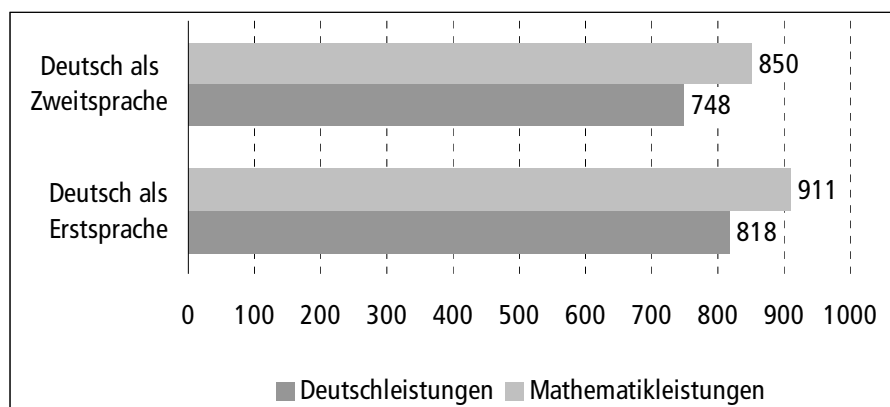
Abbildung 3.4: Leistungen am Ende der 6. Klasse nach Ausbildungsgrad der Eltern



Leistungen nach der Erstsprache

Abbildung 3.5 zeigt die Leistungen in Deutsch und Mathematik am Ende der 6. Klasse nach der Erstsprache der Schülerinnen und Schüler. In Deutsch erzielen Schülerinnen und Schüler mit Deutsch als Erstsprache 818 Punkte, jene mit Deutsch als Zweitsprache 748 Punkte. Der Unterschied von 70 Punkten ist als gross zu beurteilen ($d = 0.9$). In der Mathematik ist der Unterschied zwischen den beiden Gruppen mit 61 Punkten geringer und kann als mittelgross beurteilt werden ($d = 0.6$).

Abbildung 3.5: Leistungen am Ende der 6. Klasse nach Erstsprache

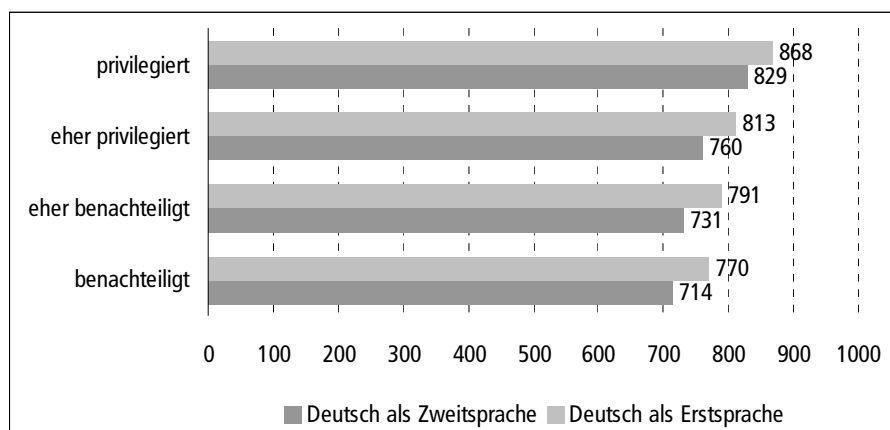


Leistungen nach sozialer Herkunft und Erstsprache

Um zu überprüfen, ob die Bedeutung der Erstsprache für die schulischen Leistungen unabhängig von der sozialen Herkunft ist, wurden die Differenzen zwischen den Schülerinnen und

Schülern mit Deutsch als Erstsprache beziehungsweise mit Deutsch als Zweitsprache innerhalb der vier nach sozialer Herkunft gebildeten Gruppen dargestellt. Abbildung 3.6 zeigt die Leistungen in Deutsch am Ende der 6. Klasse nach sozialer Herkunft und Erstsprache.

Abbildung 3.6: Leistungen in Deutsch am Ende der 6. Klasse nach sozialer Herkunft und Erstsprache



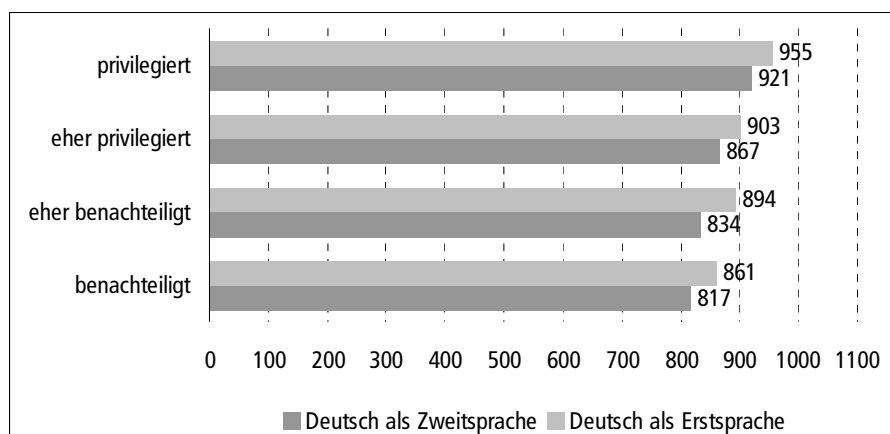
In der Gruppe mit privilegierter sozialer Herkunft beträgt der Rückstand der Schülerinnen und Schüler mit Deutsch als Zweitsprache 39 Punkte, in der Gruppe mit eher privilegierter sozialer Herkunft 53 Punkte, in der Gruppe mit eher benachteiligter sozialer Herkunft 60 Punkte und in der Gruppe mit benachteiligter sozialer Herkunft 56 Punkte. Der Rückstand der Schülerinnen und Schüler mit Deutsch als Zweitsprache ist in der Gruppe mit privilegierter sozialer Herkunft statistisch signifikant geringer als in der Gruppe mit eher privilegierter sozialer Herkunft. Zwischen den Schülerinnen und Schülern mit eher privilegierter, eher benachteiligter und benachteiligter sozialer Herkunft unterscheiden sich die Effekte der Erstsprache hingegen nicht statistisch signifikant.

Abbildung 3.7 zeigt die Leistungen in Mathematik am Ende der 6. Klasse nach sozialer Herkunft und Erstsprache. In der Gruppe mit privilegierter sozialer Herkunft beträgt der Rückstand der Schülerinnen und Schüler mit Deutsch als Zweitsprache 34 Punkte, in der Gruppe mit eher privilegierter sozialer Herkunft 36 Punkte, in der Gruppe mit eher benachteiligter sozialer Herkunft 60 Punkte und in der Gruppe mit benachteiligter sozialer Herkunft 44 Punkte. Die Erstsprache wirkt sich erstaunlicherweise in der Gruppe der Schülerinnen und Schüler mit eher benachteiligter sozialer Herkunft am stärksten negativ auf die Mathematikleistungen aus. Zwischen den Schülerinnen und Schülern mit benachteiligter, eher privilegierter und privilegierter sozialer Herkunft zeigen sich keine statistisch signifikanten Unterschiede.

Insgesamt zeigen sich geringe Wechselwirkungen zwischen der Erstsprache und der sozialen Herkunft auf die Leistungen der Schülerinnen und Schüler. In Deutsch unterscheidet sich der Leistungsrückstand der Schülerinnen und Schüler mit Deutsch als Zweitsprache lediglich zwischen den Schülerinnen und Schülern mit privilegierter sozialer Herkunft und jenen mit eher privilegierter sozialer Herkunft. In der Mathematik zeigen sich keine plausiblen Unter-

schiede im Leistungsrückstand der Schülerinnen und Schüler mit Deutsch als Zweitsprache nach sozialer Herkunft.

Abbildung 3.7: Leistungen in Mathematik am Ende der 6. Klasse nach sozialer Herkunft und Erstsprache



3.6 Vergleich von Ergebnissen in den Jahren 1998 und 2009

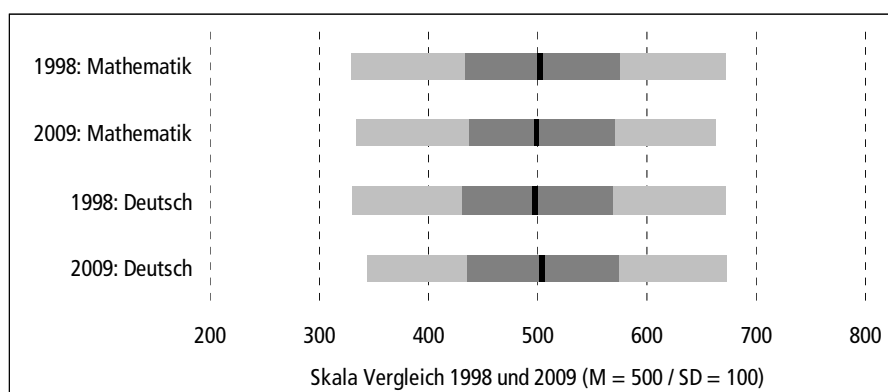
Fachliche Leistungen

Für den Vergleich der Leistungen zwischen den Jahren 1998 und 2009 wurde eine separate Skala entwickelt, die auf den Mittelwert $M = 500$ und die Standardabweichung $SD = 100$ normiert wurde. Für diesen Vergleich wurden nur jene Aufgaben genutzt, deren Schwierigkeit sich innerhalb der letzten elf Jahre nicht verändert hatte.

Abbildung 3.8 zeigt die Ergebnisse des Leistungsvergleichs zwischen 1998 und 2009. Die Gesamtlänge des Balkens umfasst 90 Prozent der Schülerleistungen. 50 Prozent der Schülerleistungen liegen innerhalb der dunkelgrauen Balken. Der kleine schwarze Balken stellt jenen Bereich dar, in dem der Mittelwert mit einer statistischen Sicherheit von 95 Prozent liegt. Die Mittelwerte liegen sehr nahe beieinander und unterscheiden sich nicht statistisch signifikant.

Das Ergebnis ändert sich auch dann nicht, wenn Kontextfaktoren, wie der Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache oder die soziale Herkunft, statistisch kontrolliert werden. Das bedeutet, dass sich das Leistungsniveau in den letzten elf Jahren nicht verändert hat.

Abbildung 3.8: Leistungen in den Lernstandserhebungen 1998 und 2009



Selbsteinschätzungen von Schul- und Unterrichtsmerkmalen

Neben den fachlichen Leistungen können auch Einschätzungen der Schülerinnen und Schüler von Schul- und Unterrichtsmerkmalen zwischen den Jahren 1998 und 2009 verglichen werden. In beiden Erhebungen wurden den Schülerinnen und Schülern Fragen zur Zufriedenheit mit der Schule, zum Schulklima, zum sozialen Verhalten an der Schule – als Indikator für aggressives und gewalttätiges Verhalten – sowie zum Arbeitsklima im Deutsch- und Mathematikunterricht gestellt.

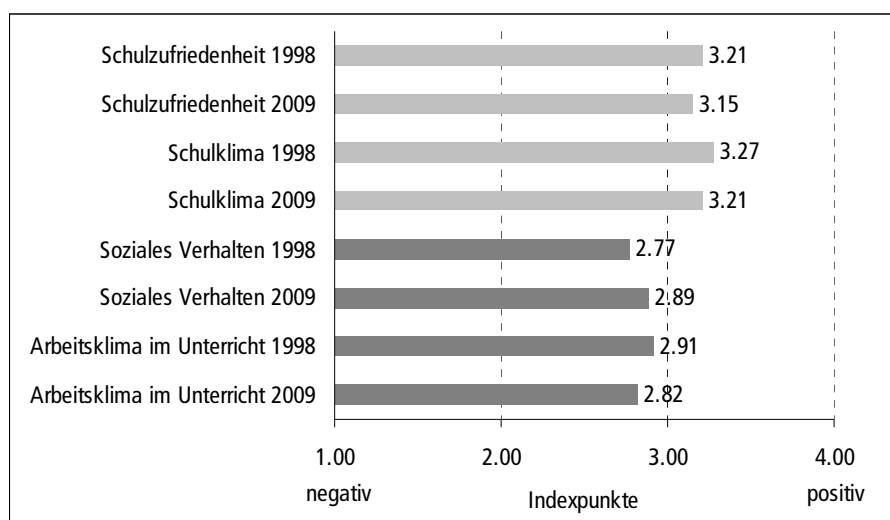
Die Schulzufriedenheit wurde mit sechs Aussagen erfasst wie «Ich gehe gerne in unsere Schule» oder «Mir gefällt es in unserer Schule». Zur Erfassung des Schulklimas wurden den Schülerinnen und Schülern sechs Aussagen vorgelegt, wie «Wir geben uns alle Mühe, dass wir eine gute Schule sind» oder «An unserer Schule fühlt man sich einfach wohl». Beispiele zu den sechs Aussagen, mit denen das soziale Verhalten an der Schule erhoben wurde, sind «In unserer Schule kommt es oft vor, dass mehrere Schüler einander verprügeln» und «In unserer Schule sind einige Schüler brutal zueinander». Das Arbeitsklima im Deutsch- und Mathematikunterricht wurde mit jeweils vier Aussagen pro Unterrichtsfach erfasst, wie «Im Deutschunterricht geht es sehr diszipliniert zu und her» oder «Im Mathematikunterricht sind die Schülerinnen und Schüler aufmerksam».

Zur Einschätzung der Aussagen standen folgende vier Antwortkategorien zur Verfügung: «stimmt überhaupt nicht», «stimmt eher nicht», «stimmt eher» und «stimmt genau». Aus den Antworten zu den erfassten Merkmalen wurden Indizes gebildet, indem die Mittelwerte aus den Antworten zu einem Merkmal gebildet wurden. Hohe Indexwerte entsprechen dadurch einer positiven Einschätzung, tiefe Indexwerte einer negativen Einschätzung der Schul- und Unterrichtsmerkmale.

Abbildung 3.9 zeigt die mittlere Einschätzung der Schul- und Unterrichtsmerkmale in den Jahren 1998 und 2009. Die Bewertungen der Schulzufriedenheit und des Schulklimas unterscheiden sich statistisch nicht signifikant zwischen den beiden Erhebungen. Das soziale Verhalten der Schülerinnen und Schüler an der Schule als Indikator für aggressives und gewalttätiges Verhalten hingegen wird im Jahr 2009 positiver beurteilt als im Jahr 1998. Der Unterschied von rund 0.1 Indexpunkten zwischen den beiden Erhebungen ist zwar statistisch sig-

nifikant, jedoch als gering zu beurteilen ($d = 0.17$). Auch die Bewertung des Arbeitsklimas im Deutsch- und Mathematikunterricht unterscheidet sich statistisch signifikant zwischen den beiden Erhebungen. Im Gegensatz zum sozialen Verhalten an der Schule fällt die Bewertung jedoch im Jahr 2009 negativer aus als im Jahr 1998. Auch dieser Unterschied ist als gering zu beurteilen ($d = 0.19$).

Abbildung 3.9: Schülerbewertung von Schul- und Unterrichtsmerkmalen in den Lernstandserhebungen 1998 und 2009



Anmerkung: Statistisch signifikante unterschiedliche Bewertungen zwischen den Jahren 1998 und 2009 sind dunkel gekennzeichnet.

3.7 Fazit

Die Leistungen der Schülerinnen und Schüler am Ende der 6. Klasse sind im Jahr 2009 gleich zu beurteilen wie im Jahr 1998. Dies zeigt der Vergleich der Mittelwerte in Deutsch und in der Mathematik, die aufgrund von zwei repräsentativen Stichproben berechnet wurden. Etwas schwieriger ist der Vergleich der Leistungen aufgrund der Verteilung der Schülerinnen und Schüler auf die Anforderungsniveaus, weil die Bildung von solchen Niveaus mit einer gewissen Willkür behaftet ist. Wo die Grenzen zwischen den Niveaus gezogen werden, lässt sich zwar inhaltlich begründen. Allerdings würden vermutlich verschiedene Lehrplanexperten und Lehrpersonen zu unterschiedlichen Einteilungen gelangen, zumal der Lehrplan bei der Einteilung nur bedingt genutzt werden kann.

Die Darstellung nach Anforderungsniveaus hat jedoch den entscheidenden Vorteil, dass die Leistungen inhaltlich umschrieben und mit Aufgabenbeispielen illustriert werden können. Von daher kann ohne Bedenken festgestellt werden, dass knapp ein Fünftel der Schülerinnen und Schüler die Anforderungen des Lehrplans in den geprüften Bereichen nur teilweise erfüllt. Zu diesem Ergebnis kam bereits die Lernstandserhebung im Jahr 1998. Insgesamt ist der Anteil an Schülerinnen und Schülern, für die bereits vergleichsweise einfache Aufgaben grosse Schwierigkeiten darstellen, in beiden Erhebungen gleich gross.

Dass diese Schülerinnen und Schüler in den geprüften Bereichen die Lehrplanziele zum Teil nicht erreichen, lässt sich ohne Einschränkungen feststellen. Selbst einfache Subtraktionen und Divisionen, wie sie bereits in der 5. Klasse geübt werden, stellen diese Schülerinnen und Schüler zum Teil vor Schwierigkeiten. Ohne zusätzliche Massnahmen zur Füllung der Lücken ist die Wahrscheinlichkeit hoch, dass diese Schülerinnen und Schüler auch am Ende der obligatorischen Schulbildung die Ziele des Lehrplans nicht erreichen und der Übertritt in die Berufsbildung nicht ohne Schwierigkeiten verlaufen wird.

Was ebenfalls deutlich wird, ist die Zunahme der Leistungsunterschiede zwischen den Schülerinnen und Schülern auf der Mittelstufe. Der Anteil an Schülerinnen und Schülern, die die Ziele des Lehrplans nicht erfüllen, ist am Ende der 3. Klasse deutlich geringer (rund 10 Prozent) als am Ende der 6. Klasse (vgl. Keller & Moser, 2008, S. 41ff.). Die Zunahme der Leistungsdifferenzen zwischen den leistungsstarken und den leistungsschwachen Schülerinnen und Schülern erfolgt erst auf der Mittelstufe. Diese Zunahme wird als Schereneffekt bezeichnet und kann – wie die Ergebnisse im vorliegenden Bericht zeigen – zum Teil durch die soziale Herkunft der Schülerinnen und Schüler erklärt werden.

Dass gewisse Aspekte von Schule und Unterricht im Jahr 2009 anders eingeschätzt werden als noch vor zehn Jahren, sollte nicht überbewertet werden. Selbsteinschätzungen müssen nicht zwingend mit den objektiven Gegebenheiten übereinstimmen. Zudem ändert sich auch die Interpretation der vorgelegten Aussagen über die Zeit hinweg. Der Vergleich von Ergebnissen aus zwei Erhebungen, zwischen denen mehr als zehn Jahre liegen, zeigt, dass empirisch nachweisbare Qualitätssteigerungen in der Schule nur sehr langsam erfolgen.

4 Die Bedeutung der Klassenzusammensetzung

Domenico Angelone & Urs Moser

4.1 Segregation nach bildungsrelevanten Merkmalen

Die Ergebnisse in Kapitel 2 zur Bedeutung individueller Merkmale für den Leistungszuwachs machen eines deutlich: Die soziale Herkunft erklärt die Leistungsunterschiede zwischen den Schülerinnen und Schülern besser als jedes andere Merkmal. Kinder, die in einer anregungsreichen Umgebung aufwachsen, sind auf die Schule besser vorbereitet und erfahren auch während der Schulzeit mehr Unterstützung von zu Hause aus. Dies bedeutet zugleich bessere Schulleistungen und eine erfolgreichere Schullaufbahn. Jugendliche aus sozioökonomisch benachteiligten Verhältnissen haben hingegen mehr Schwierigkeiten, die Lehrplanziele zu erreichen, besuchen auf der Sekundarstufe I überdurchschnittlich häufig die Schultypen mit geringen Anforderungen und bekunden beim Übertritt in die Berufsbildung häufiger Mühe als Jugendliche aus sozioökonomisch privilegierten Verhältnissen (vgl. Stalder, Meyer & Hupka-Brunner, 2008).

Der Zusammenhang zwischen sozialer Herkunft und Schulerfolg wird seit längerer Zeit als eines der grundlegendsten Probleme der Schule erwähnt. Er wird für die Schule auch deshalb zum Problem, weil sich die Kinder verschiedener sozialer Herkunft nicht gleichmässig auf die Schulen verteilen. Vor allem in Städten und in Agglomerationen ist die Segregation der Bevölkerung nach bildungsrelevanten Merkmalen, wie Bildungsnähe des Elternhauses, ökonomische, soziale und kulturelle Ressourcen der Eltern oder Migrationshintergrund, so weit fortgeschritten, dass sich die Lehr-Lern-Bedingungen in den Klassen bereits zu Beginn der Schulzeit stark unterscheiden. Je privilegierter die soziale Zusammensetzung einer Klasse ist, desto höher sind die durchschnittlichen Kompetenzen der Klasse beim Schuleintritt (Moser, Berweger & Stamm, 2005).

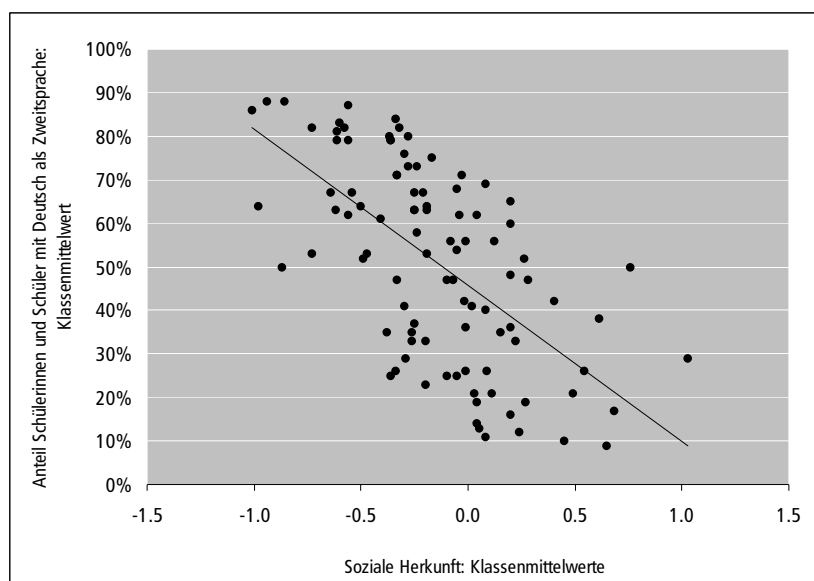
Die Aufteilung der Bevölkerung nach bildungsrelevanten Merkmalen führt nicht nur zu einer Gettoisierung in gewissen städtischen Quartieren, sondern vor allem auch zu ungünstigen Lernbedingungen in Schulklassen. Die Folge davon ist eine doppelte Benachteiligung von Schülerinnen und Schülern aus sozial benachteiligten Verhältnissen. Zum einen fehlt ihnen die Unterstützung zu Hause. Zum andern zeigt die Bildungsforschung in den letzten zehn Jahren, dass Kontextmerkmale wie die Zusammensetzung der Klasse für das Lernen und die Schulleistungen von Bedeutung sein können (Baumert, Stanat & Watermann, 2006).

Ob in einer Schulklasse vorwiegend Schülerinnen und Schüler mit Deutsch als Erstsprache aus privilegierten Verhältnissen oder vorwiegend Schülerinnen und Schüler mit Deutsch als Zweitsprache aus benachteiligten Verhältnissen unterrichtet werden, ist für das Vorgehen im Unterricht bedeutsam. Zudem besteht die Gefahr, dass aufgrund der Klassenzusammensetzung unterschiedliche Lern- und Entwicklungsmöglichkeiten entstehen. Schülerinnen und Schüler können vom Leistungsniveau der Klasse profitieren oder aber benachteiligt werden, was zu einer Zunahme der Leistungsdifferenzen zwischen den Klassen beziehungsweise zwischen den Schülerinnen und Schülern führen kann (Baumert, Stanat & Watermann, 2006).

4.2 Zusammensetzung der Klassen

Bei der Klassenzusammensetzung sind zwei Merkmale für den Leistungszuwachs relevant: Die soziale Zusammensetzung der Klasse und der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache (Moser, Keller & Zimmermann, 2008). Abbildung 4.1 zeigt den Zusammenhang zwischen der sozialen Zusammensetzung der Klassen und dem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache. Jeder Punkt entspricht einer Klasse, deren Position sich durch den durchschnittlichen Index der sozialen Herkunft und den Prozentanteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache ergibt. Es gibt Klassen mit einem Mittelwert des Index zur sozialen Herkunft von -1 und Klassen mit einem Mittelwert des Index zur sozialen Herkunft von $+1$. Und es gibt Klassen mit einem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache von mehr als 80 Prozent und Klassen mit einem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache von weniger als 10 Prozent. Es gilt zu berücksichtigen, dass in Abbildung 4.1 die Klassen der Stichprobe abgebildet sind, was nicht der Verteilung der Klassen in der Population entspricht. Für die Klassenstichprobe wurden überproportional viele Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache gewählt.

Abbildung 4.1: Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache nach sozialer Klassenzusammensetzung



Anmerkung: Die Ergebnisse stammen aus der Klassenstichprobe.

Weil Schülerinnen und Schüler mit Deutsch als Zweitsprache überproportional häufig aus sozial benachteiligten Verhältnissen stammen, besteht ein negativer Zusammenhang zwischen der sozialen Herkunft und der Erstsprache. Je höher der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache in einer Klasse ist, desto tiefer ist die soziale Zusammensetzung der Klasse. Der Zusammenhang zwischen sozialer Herkunft und Erstsprache ist in anderer Form auch in Tabelle 4.1 dargestellt. Von den Schülerinnen und Schülern mit

Deutsch als Erstsprache stammen rund 32 Prozent aus Familien mit privilegierten sozialen Verhältnissen und 17 Prozent aus Familien mit benachteiligten sozialen Verhältnissen. Von den Schülerinnen und Schülern mit Deutsch als Zweitsprache stammen rund 16 Prozent aus Familien mit privilegierten sozialen Verhältnissen und 31 Prozent aus Familien mit benachteiligten sozialen Verhältnissen.

Tabelle 4.1: Soziale Herkunft der Schülerinnen und Schüler nach Erstsprache

	soziale Herkunft				Total
	benachteiligt	eher benachteiligt	eher privilegiert	privilegiert	
Deutsch als Erstsprache	17.4%	21.0%	29.6%	32.1%	100.0%
Deutsch als Zweitsprache	31.4%	28.1%	24.9%	15.6%	100.0%

4.3 Klassenzusammensetzung und Schulleistungen

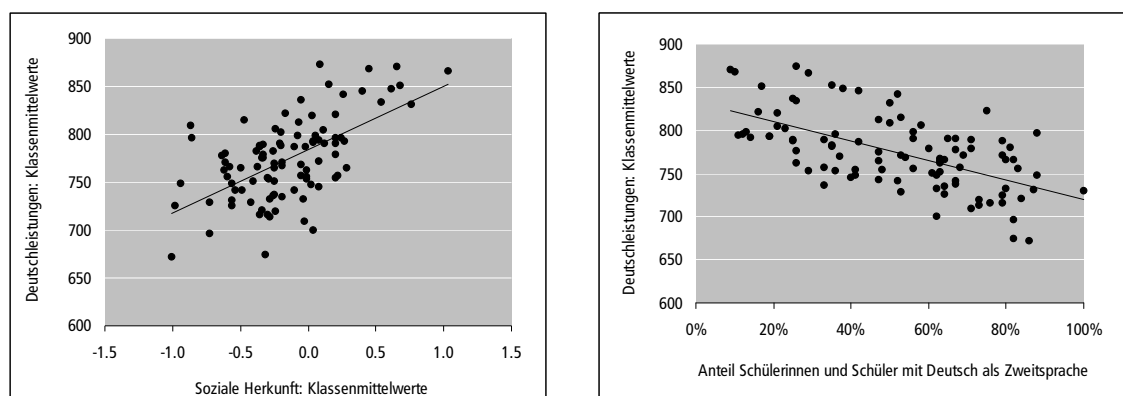
Deutschleistungen nach sozialer Zusammensetzung und nach Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache

Aufgrund des engen Zusammenhangs zwischen der sozialen Herkunft und den Leistungen einerseits und der Segregation der Schülerpopulation nach der sozialen Herkunft andererseits unterscheiden sich die durchschnittlichen Leistungen der Klassen. Je privilegierter die soziale Zusammensetzung einer Klasse ist (durchschnittlicher Index zur sozialen Herkunft), desto besser sind die durchschnittlichen Leistungen der Klasse. Und je höher der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache ist, desto geringer sind die durchschnittlichen Leistungen der Klasse.

Abbildung 4.2 zeigt die durchschnittlichen Deutschleistungen der Klassen in Abhängigkeit der Zusammensetzung der Klasse. Wiederum sind in der Abbildung die Klassen der Stichprobe dargestellt, für die überproportional viele Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache ausgewählt wurden. Die linke Grafik zeigt, dass die durchschnittlichen Deutschleistungen der Klassen mit dem Anstieg des Index der sozialen Zusammensetzung besser werden. Die rechte Grafik zeigt, dass die durchschnittlichen Deutschleistungen der Klassen mit dem Anstieg des Anteils an Schülerinnen und Schülern mit Deutsch als Zweitsprache sinken.

Die Zusammenhänge zwischen der Zusammensetzung der Klasse und den fachlichen Leistungen sind gross: Der Korrelationskoeffizient beträgt für den Zusammenhang zwischen der sozialen Zusammensetzung und den Deutschleistungen $r = 0.60$, für den Zusammenhang zwischen dem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache und den Deutschleistungen $r = -0.60$. Die Koeffizienten für die analogen Zusammenhänge in der Mathematik betragen $r = 0.54$ und $r = -0.52$.

Abbildung 4.2: Zusammensetzung der Klasse und Deutschleistungen



In Deutsch erreicht die beste Klasse einen Mittelwert von rund 874 Punkten, die schwächste einen Mittelwert von 673 Punkten. Zwischen den beiden Extremmittelwerten liegen 201 Punkte. Sieht man von den 10 Prozent besten und schwächsten Klassen ab und betrachtet nur die 80 Prozent mittleren Klassen, dann reduziert sich die Spannbreite von 201 Punkten auf 117 Punkte. Die Spannbreite der 50 Prozent mittleren Klassen beträgt noch 50 Punkte.

In der Mathematik erreicht die beste Klasse einen Mittelwert von rund 979 Punkten, die schwächste einen Mittelwert von 795 Punkten. Zwischen den beiden Extremmittelwerten liegen 184 Punkte. Sieht man wiederum von den 10 Prozent besten und schwächsten Klassen ab und betrachtet nur die 80 Prozent mittleren Klassen, dann reduziert sich die Spannbreite von 184 Punkten auf 95 Punkte. Die Spannbreite der 50 Prozent mittleren Klassen beträgt nur noch 39 Punkte. Die Varianz zwischen den Klassen ist in der Mathematik folglich etwas geringer als in Deutsch.

Wirkungen der Klassenzusammensetzung auf die Schulleistungen

Die bisherigen Darstellungen zeigen, dass sich (1) die soziale Zusammensetzung und der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache zwischen den Klassen unterscheiden und (2) die Leistungen der Klassen mit der sozialen Zusammensetzung und dem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache zusammenhängen.

Die bisherigen Darstellungen sagen allerdings nichts darüber aus, ob sich die durchschnittlichen Leistungen der Klassen einzig aufgrund des engen Zusammenhangs zwischen der sozialen Herkunft und den Leistungen beziehungsweise zwischen der Erstsprache und den Leistungen erklären lassen oder ob sich darüber hinaus die Klassenzusammensetzung noch zusätzlich auf die Leistungsentwicklung der Kinder auswirkt. Es könnte beispielsweise sein, dass in Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache das Lernen im Unterricht eingeschränkt ist oder dass der Lehr-Lern-Prozess aufgrund der sozialen Zusammensetzung einer Klasse zusätzlich erschwert oder erleichtert wird.

Mit einer Mehrebenenanalyse wurde deshalb untersucht, durch welche Merkmale sich die Leistungsunterschiede zwischen den Klassen erklären lassen. Grundsätzlich lassen sich die Leistungsunterschiede durch Merkmale der Schülerinnen und Schüler (individuelle Merkmale) und durch Merkmale der Klassen (Klassenmerkmale) erklären.

Vier individuelle Merkmale erklären die Leistungsunterschiede zwischen den Schülerinnen und Schülern am Ende der 6. Klasse besonders gut: Alter, Erstsprache, Geschlecht und soziale Herkunft. (1) Ältere Schülerinnen und Schüler erzielen leicht schlechtere Leistungen am Ende der 6. Klassen. Der negative Effekt des Alters auf die Leistungen kommt deshalb zustande, weil die Klassenstichprobe auch Schülerinnen und Schüler enthält, die repetiert oder eine Kleinklasse A besuchten, in der der Unterrichtsstoff der 1. Klasse auf zwei Schuljahre verteilt wird. Diese Schülerinnen und Schüler sind die leistungsschwächeren und zugleich die älteren Schülerinnen und Schüler. (2) Schülerinnen und Schüler mit Deutsch als Zweitsprache erzielen schlechtere Leistungen als solche mit Deutsch als Erstsprache. (3) Knaben erzielen in der Mathematik, Mädchen in Deutsch bessere Leistungen. (4) Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto besser sind ihre Schulleistungen.

Auch die beiden Klassenmerkmale – die soziale Zusammensetzung und der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache – erklären Leistungsunterschiede zwischen den Schülerinnen und Schülern. Allerdings sind die Zusammenhänge zwischen der sozialen Zusammensetzung und den Leistungen nicht linear im Sinne «je höher, desto besser».

Die soziale Zusammensetzung der Klasse wirkt sich erst dann positiv auf die Leistungen der Schülerinnen und Schüler aus, wenn die durchschnittliche soziale Herkunft besonders hoch ist. Dies trifft im Kanton Zürich für rund 14 Prozent der Klassen zu. Schülerinnen und Schüler profitieren dann von der sozialen Zusammensetzung einer Klasse, wenn die Klasse vorwiegend von Kindern mit privilegierter sozialer Herkunft besucht wird. Für die Schülerinnen und Schüler bedeutet dies, dass ihre Deutschleistungen rund 45 Punkte und ihre Mathematikleistungen rund 24 Punkte besser ausfallen. Diese besseren Leistungen sind einzig auf die soziale Zusammensetzung der Klasse zurückzuführen.

Der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache wirkt sich im Gegensatz zur sozialen Zusammensetzung der Klasse linear auf die Leistungen aus: Je höher der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache ist, desto tiefer sind die Leistungen. Mit einem Anstieg des Anteils an Schülerinnen und Schülern mit Deutsch als Zweitsprache in einer Klasse um 10 Prozent verschlechtern sich die Deutschleistungen der Schülerinnen und Schüler in Deutsch um 5 Punkte und in der Mathematik um 3.5 Punkte.

Beispiel

Wenn beispielsweise eine Schülerin durchschnittlichen Alters mit Deutsch als Erstsprache und mit durchschnittlicher sozialer Herkunft eine Klasse mit einem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache von 80 Prozent besucht, dann fallen ihre Deutschleistungen um rund 30 Punkte und ihre Mathematikleistungen um rund 20 Punkte tiefer aus, als wenn diese Schülerin eine Klasse mit einem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache von 20 Prozent besucht.

4.4 Die Situation von QUIMS-Schulen

Den Herausforderungen, die sich für die Schule aufgrund der sozialen Zusammensetzung der Schülerschaft ergeben können, wird im Kanton Zürich unter anderem mit dem Programm QUIMS begegnet. Schulen mit einem hohen Anteil an Kindern und Jugendlichen mit Deutsch als Zweitsprache und aus sozial benachteiligten Verhältnissen haben den Auftrag, zusätzliche Massnahmen durchzuführen, um die Qualität in multikulturellen Schulen (QUIMS) sicherzu-

stellen. Ziel ist es, das Leistungsniveau an diesen Schulen zu steigern und so die Chancen der Schülerinnen und Schüler für eine erfolgreiche Schullaufbahn und die Integration zu erhöhen. Dazu sind in der Handreichung «Umsetzung Volksschulgesetz – Qualität in multikulturellen Schulen (QUIMS)» Massnahmen zu drei Handlungsfeldern konkretisiert worden: (1) Sprachförderung, (2) Förderung des Schulerfolgs und (3) Förderung der Integration.

Zwischenziel der Umsetzung bis 2010 ist es, dass Schulen, die in einer zweiten Phase (2007–2009) in QUIMS aufgenommen wurden, ausgewählte Massnahmen zu den drei Handlungsfeldern eingeführt haben. Dabei kann es sich sowohl um befristete Entwicklungsprojekte als auch um feste Angebote handeln. Für die Umsetzung des QUIMS-Auftrags setzen die Schulen einen QUIMS-Beauftragten ein, dem dafür zwischen zwei und fünf Wochenstunden zur Verfügung stehen. Den QUIMS-Schulen steht zudem ein finanzieller Beitrag zur Verfügung, dessen Grösse unter anderem vom Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache bestimmt wird und der im Durchschnitt CHF 40'000 beträgt (www.quims.ch).

In einer ersten freiwilligen Phase (1999–2006) beteiligten sich 21 Schulen an QUIMS. In der zweiten Phase (2007–2009) wurden alle Schulen in das Programm QUIMS aufgenommen, die die Aufnahmekriterien erfüllten. Entsprechend dieser Einführungszeitpunkte von QUIMS wurde die Stichprobe für die Lernstandserhebung gebildet. Verglichen werden Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, die bereits vor 2006 QUIMS angehörten, und solche, die erst nach 2008 von QUIMS profitieren konnten.

Tabelle 4.2 zeigt, dass der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache in Klassen, die von QUIMS bereits vor 2006 unterstützt wurden, mit knapp 70 Prozent am höchsten ist. Demgegenüber ist dieser Anteil bei den Klassen ohne besonders hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache mit rund 35 Prozent etwa halb so gross. Die soziale Zusammensetzung (durchschnittlicher Index der sozialen Herkunft) ist in Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, die von QUIMS nach 2008 unterstützt werden, am tiefsten. Sie beträgt -0.39 Indexpunkte gegenüber -0.29 Indexpunkten der Klassen mit QUIMS vor 2006. Der Unterschied von 0.10 Indexpunkten ist statistisch nicht signifikant.

Tabelle 4.2: Zusammensetzung der Klassen nach den drei Vergleichsgruppen

	Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache		
	nicht hoch ohne QUIMS	hoch QUIMS vor 2006	hoch QUIMS nach 2008
Deutsch als Erstsprache	65%	30%	36%
Deutsch als Zweitsprache	35%	70%	64%
Index der sozialen Herkunft	0.04	-0.29	-0.39
Total	100%	100%	100%

In Abbildung 4.3 sind die Deutschleistungen, in Abbildung 4.4 die Mathematikleistungen der drei Vergleichsgruppen dargestellt. Die Spannweite der Leistungen ist in Form eines Balkens dargestellt. Die Gesamtlänge des Balkens umfasst 90 Prozent der Schülerleistungen. 50 Prozent der Schülerleistungen liegen innerhalb des dunkelgrauen Balkens. Der kleine schwarze Balken stellt jenen Bereich dar, in dem der Mittelwert mit einer statistischen Sicherheit von 95 Prozent liegt. Je kleiner der schwarze Balken ist, desto zuverlässiger ist die Schätzung des Mittelwerts. In den beiden Gruppen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache ist der schwarze Balken jeweils grösser, weil die Mittelwerte aufgrund einer geringeren Anzahl Klassen geschätzt wurden.

In Deutsch erzielen die Klassen mit einem geringen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache einen Mittelwert von 793 Punkten. Die Mittelwerte der Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, die von QUIMS (vor oder nach 2008) unterstützt werden, liegen je bei rund 756 Punkten.

Abbildung 4.3: Deutschleistungen nach den drei Vergleichsgruppen

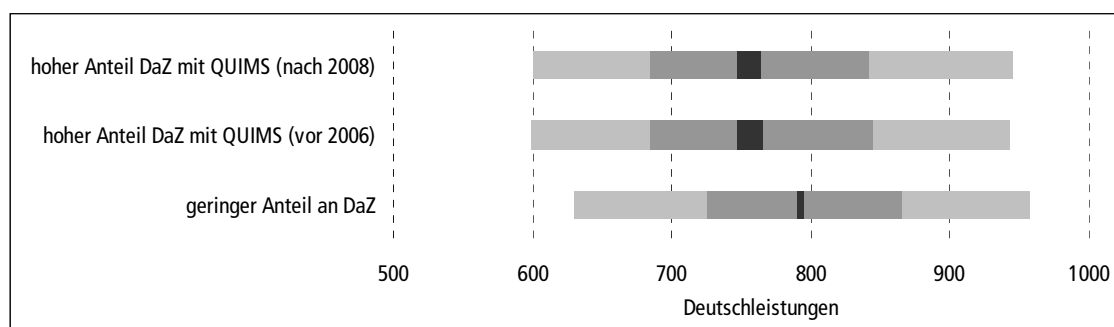
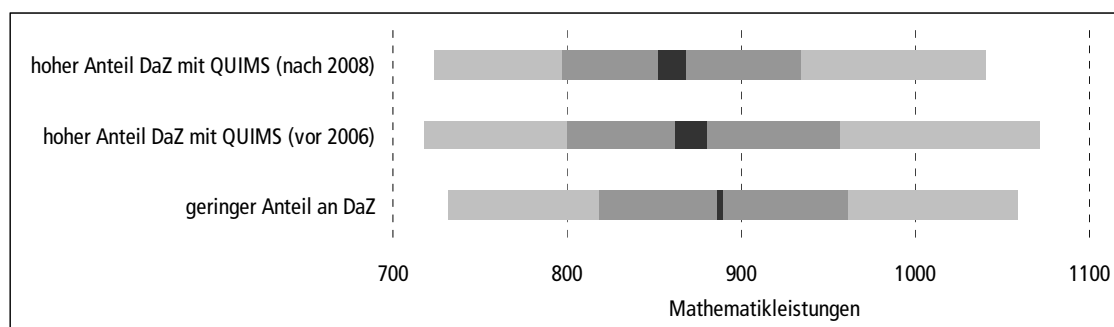


Abbildung 4.4: Mathematikleistungen nach den drei Vergleichsgruppen



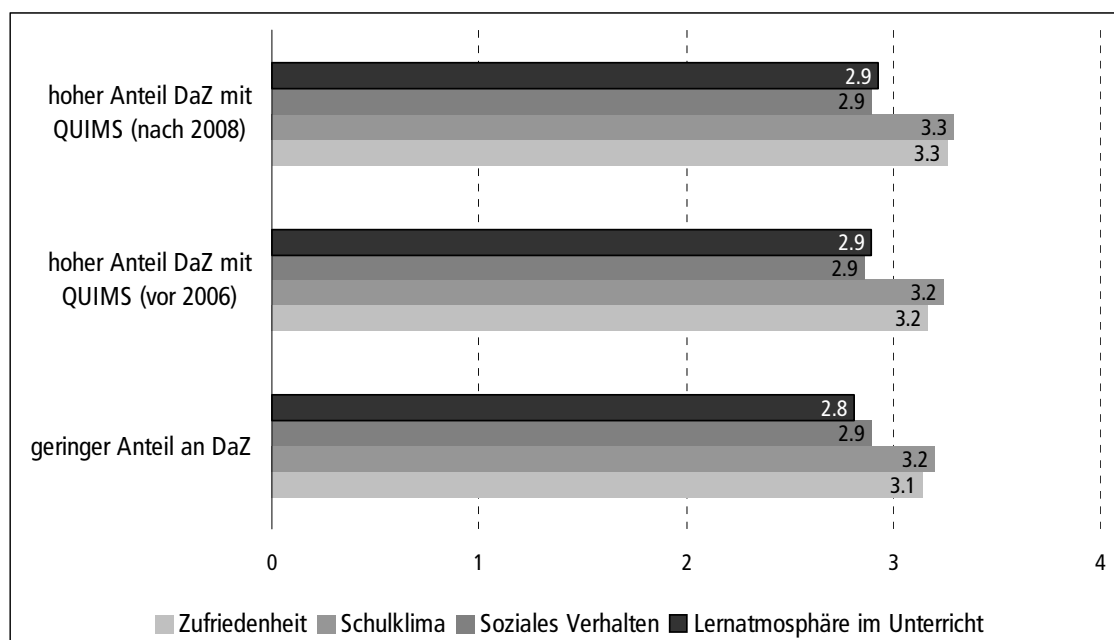
In der Mathematik erzielen die Klassen mit einem geringen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache einen Mittelwert von 888 Punkten. Die Mittelwerte der Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache liegen mit Unterstützung von QUIMS vor 2006 bei 872 Punkten und mit Unterstützung von QUIMS nach 2008 bei 860 Punkten. Die geringe Differenz von 12 Punkten ist statistisch nicht signifikant.

Die Leistungen in Deutsch und Mathematik unterscheiden sich am Ende der 6. Klasse somit nicht statistisch signifikant zwischen Schulen mit QUIMS-Unterstützung vor 2006 und solchen mit QUIMS-Unterstützung nach 2008. Bei diesem Ergebnis gilt es aber zu berücksichtigen, dass aufgrund der Klassenstichprobe nur Analysen im Querschnitt möglich sind. Im Rahmen der Längsschnittstudie konnten zwischen den Erhebungen am Ende der 3. und am Ende der 6. Klasse keine Klassen verfolgt werden, weshalb der Leistungszuwachs für Klassen mit QUIMS-Unterstützung vor 2006 und Klassen mit QUIMS-Unterstützung nach 2008 nicht ausgewiesen werden kann.

Die Ergebnisse am Ende der 6. Klasse zeigen aber auch, dass der Leistungsrückstand der Schülerinnen und Schüler in Schulen mit QUIMS-Unterstützung gegenüber solchen ohne QUIMS-Unterstützung nicht besonders gross ausfällt. In Deutsch beträgt der Rückstand rund 40 Punkte, in Mathematik rund 30 Punkte. Die Unterschiede sind als mittelstark bis eher klein zu beurteilen ($d = \max. 0.36$) und können vor dem Hintergrund des deutlich höheren Anteils an Schülerinnen und Schülern mit Deutsch als Zweitsprache in den Schulen mit QUIMS-Unterstützung als positives Ergebnis gewertet werden.

Zwischen den drei Vergleichsgruppen können neben den Leistungen auch die subjektiven Einschätzungen der Schülerinnen und Schüler zu Merkmalen der Schul- und Unterrichtskultur verglichen werden. Abbildung 4.5 zeigt für die drei Vergleichsgruppen die Mittelwerte zur Schulzufriedenheit, zum Schulklima, zum sozialen Verhalten und zum Arbeitsklima im Deutsch- und Mathematikunterricht.

Abbildung 4.5: Schul- und Unterrichtsmerkmale nach den drei Vergleichsgruppen



Die Schulzufriedenheit wird in Schulen mit QUIMS-Unterstützung nach 2008 geringfügig höher eingeschätzt als in Schulen mit QUIMS-Unterstützung vor 2006 oder in Schulen ohne

QUIMS-Unterstützung. Zwischen letzteren beiden Gruppen sind diesbezüglich keine Unterschiede feststellbar. Das Schulklima wird in Schulen mit QUIMS-Unterstützung nach 2008 geringfügig positiver eingeschätzt als in Schulen ohne QUIMS-Unterstützung. Darüber hinaus finden sich keine Unterschiede in der Bewertung des Schulklimas. Das soziale Verhalten zwischen den Schülerinnen und Schülern an der Schule wird in allen drei Vergleichsgruppen gleich bewertet. Schliesslich wird das Arbeitsklima im Deutsch- und Mathematikunterricht in Schulen mit QUIMS-Unterstützung geringfügig positiver bewertet als in Schulen ohne QUIMS-Unterstützung. Zwischen Schulen mit QUIMS vor 2006 und nach 2008 hingegen unterscheiden sich die Bewertungen des Schulklimas nicht statistisch signifikant.

Insgesamt sind die Unterschiede zwischen den drei Vergleichsgruppen auch mit Blick auf die Einschätzung von Merkmalen zur Schul- und Unterrichtskultur nur sehr gering ausgeprägt.

4.5 Die Bedeutung der Förderung von Deutsch als Zweitsprache

Im Gegensatz zu den Angaben über Förderprogramme an QUIMS-Schulen liegen für alle Schülerinnen und Schüler der Längsschnittstichprobe differenzierte Angaben über die Nutzung des Förderangebots Deutsch als Zweitsprache vor (DaZ-Unterricht). Der DaZ-Unterricht bildet einen Schwerpunkt der Sprachförderung an QUIMS-Schulen. Er wird an QUIMS-Schulen besonders häufig angeboten. Die Angaben zur Nutzung des Förderangebots werden neu von der Bildungsstatistik erfasst und konnten zur Überprüfung der Wirkung des DaZ-Unterrichts genutzt werden.

Damit der Lernerfolg auf den DaZ-Unterricht zurückgeführt werden konnte, wurden sogenannte statistische Zwillinge gebildet. Im Idealfall unterscheiden sich die statistischen Zwillinge einzig darin, dass sie DaZ-Unterricht besuchen oder nicht. Erstsprache, soziale Herkunft, Geschlecht, kognitive Grundbildung, Alter und Schulleistungen vor dem Besuch des DaZ-Unterrichts sind hingegen identisch. Zwillinge, die den DaZ-Unterricht besuchen, werden zur Experimentalgruppe zusammengefasst. Zwillinge, die keinen DaZ-Unterricht besuchen, werden zur Kontrollgruppe zusammengefasst. Danach werden die Unterschiede im Leistungszuwachs beziehungsweise in den Deutschleistungen der Zwillinge *mit* DaZ-Unterricht (Experimentalgruppe) und der Zwillinge *ohne* DaZ-Unterricht (Kontrollgruppe) verglichen.

Mit dieser Methode können statistisch signifikante positive Effekte des DaZ-Unterrichts auf die Deutschleistungen der Schülerinnen und Schüler nachgewiesen werden. Allerdings tritt dieser Effekt erst bei einer bestimmten Anzahl Stunden auf. Schülerinnen und Schüler, die den DaZ-Unterricht in der 3. Klasse während 76 Lektionen besuchen, was durchschnittlich zwei Wochenlektionen entspricht, schneiden am Ende der 3. Klasse im Deutshtest um rund 30 Punkte besser ab als vergleichbare Kinder, die den DaZ-Unterricht nicht besuchen. Dieser Effekt kann als mittelstark beurteilt werden. Die Deutschleistungen von Schülerinnen und Schülern, die den DaZ-Unterricht in der 3. Klasse nur während 38 Lektionen besuchen, was durchschnittlich einer Wochenlektion entspricht, erreichen keine statistisch signifikant besseren Deutschleistungen.

Die positive Wirkung des DaZ-Unterrichts konnte nur für jene Schülerinnen und Schüler nachgewiesen werden, die dieses Förderangebot in der 3. Klasse nutzten. Zwischen dem Besuch des DaZ-Unterrichts in der 4., 5. oder 6. Klasse und dem Leistungszuwachs bis zum Ende der 6. Klasse konnten keine statistisch signifikanten Effekte nachgewiesen werden. Der Grund

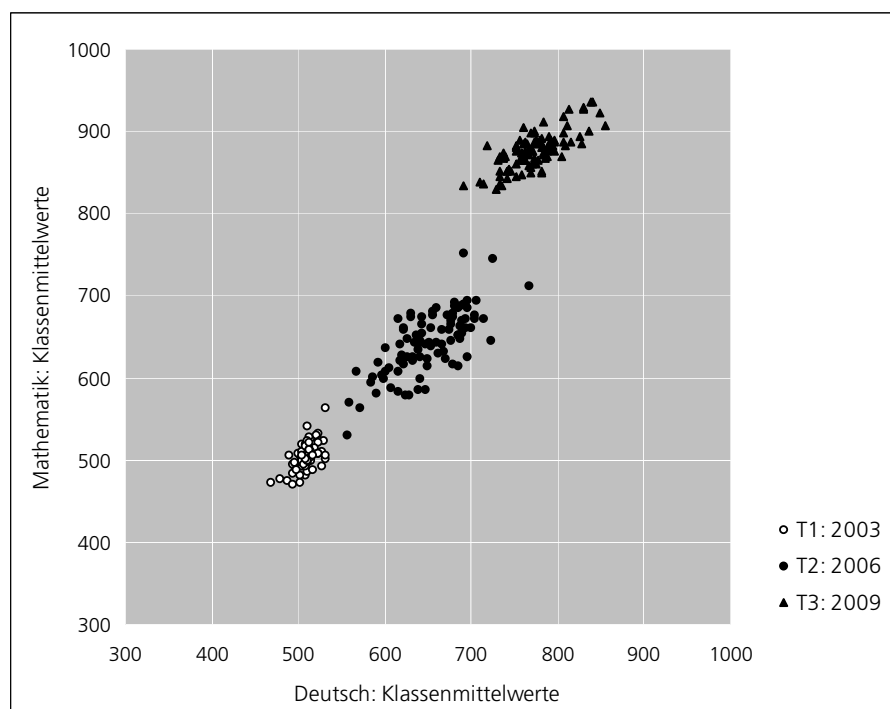
dafür liegt vermutlich nicht bei der fehlenden Wirkung des DaZ-Unterrichts, sondern bei der geringen Anzahl Schülerinnen und Schüler, die in der Mittelstufe den DaZ-Unterricht besuchen. Aus methodischer Perspektive lässt sich der Vergleich erst bei einer bestimmten Anzahl Schülerinnen und Schüler mit DaZ-Unterricht zuverlässig durchführen.

Insgesamt sind die Ergebnisse zur Wirkung des DaZ-Unterrichts ermutigend. Für eine bessere Beurteilung müsste allerdings die Wirkung des DaZ-Unterrichts im Kindergarten und auf der Unterstufe überprüft werden. Diese Angaben stehen jedoch nicht zur Verfügung.

4.6 Leistungsunterschiede zwischen den Klassen

Die Klassenzusammensetzung gehört zu jenen Merkmalen, die die Leistungsunterschiede zwischen den Schülerinnen und Schülern erklären. Zusätzlich werden die Leistungsunterschiede auch durch andere Klassenmerkmale, wie der Unterrichtsqualität, oder durch Erwartungshaltungen von Lehrpersonen bestimmt.

Abbildung 4.6: Klassenmittelwerte zu den drei Lernstandserhebungen 2003, 2006, 2009



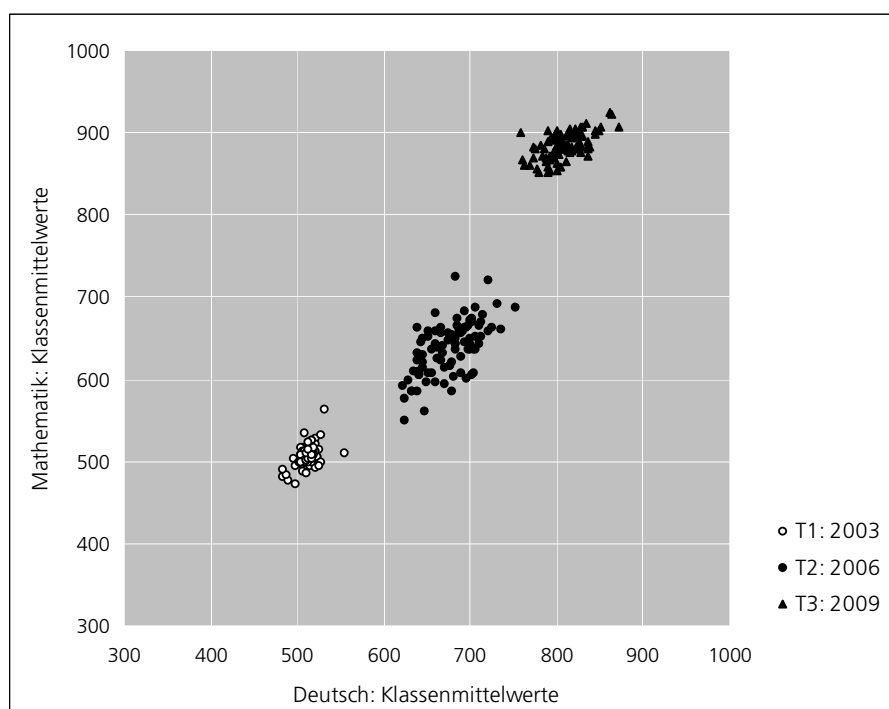
Anmerkung: Die Mittelwerte wurden mit HLM berechnet (shrunken means).

Abbildung 4.6 zeigt die Klassenmittelwerte zu den drei Lernstandserhebungen in Deutsch und Mathematik. Die Position der Punkte ergibt sich aus den Klassenmittelwerten in Deutsch und Mathematik. Die weissen Punkte zeigen die Klassenmittelwerte zu Beginn der 1. Klasse (Testzeitpunkt T1). Dabei handelt es sich bei den Deutschleistungen um erste Lesefertigkeiten, in der Mathematik um mathematisch-numerische Vorläuferfertigkeiten. Die schwarzen Punkte zeigen die Klassenmittelwerte am Ende der 3. Klasse (T2) in Deutsch und Mathematik. Dargestellt sind die Mittelwerte jener Klassen, die bereits zu Beginn der 1. Klasse getestet

wurden. Die schwarzen Dreiecke zeigen die Klassenmittelwerte am Ende der 6. Klasse (T3) in Deutsch und Mathematik. Dargestellt sind die Klassen der Klassenstichprobe. Es handelt sich folglich um andere Klassen als zu den beiden ersten Testzeitpunkten.

Die Klassenmittelwerte liegen beim ersten Testzeitpunkt relativ eng beieinander. In Deutsch beträgt der Unterschied zwischen den 5 Prozent besten und den 5 Prozent schlechtesten Klassen 53 Punkte, in der Mathematik 36 Punkte. Nach drei Jahren ist die Streuung der Klassenmittelwerte deutlich grösser. Beim zweiten Testzeitpunkt beträgt der Unterschied zwischen den 5 Prozent besten und den 5 Prozent schlechtesten Klassen in Deutsch 128 Punkte, in der Mathematik 114 Punkte. Am Ende der 6. Klasse ist die Streuung der Klassenmittelwerte etwas kleiner als am Ende der 3. Klasse. Am Ende der 6. Klasse beträgt der Unterschied zwischen den 5 Prozent besten und den 5 Prozent schlechtesten Klassen in Deutsch 108 Punkte, in der Mathematik 90 Punkte.

Abbildung 4.7: Klassenmittelwerte zu den drei Lernstandserhebungen 2003, 2006, 2009 nach statistischer Kontrolle der Lernvoraussetzungen



Berechnet man die Klassenmittelwerte nach der statistischen Kontrolle der Lernvoraussetzungen (Geschlecht, soziale Herkunft, Erstsprache, Alter und kognitive Grundfähigkeit), dann reduzieren sich die Unterschiede zwischen den einzelnen Klassen. Die statistisch kontrollierten Klassenmittelwerte sind in Abbildung 4.7 dargestellt. Beim ersten Testzeitpunkt beträgt der Unterschied zwischen den 5 Prozent besten und den 5 Prozent schlechtesten Klassen in Deutsch 41 Punkte, in der Mathematik 32 Punkte. Beim zweiten Testzeitpunkt beträgt dieser Unterschied in Deutsch 91 Punkte, in der Mathematik 103 Punkte, beim dritten Testzeitpunkt in Deutsch 78 Punkte, in der Mathematik 50 Punkte.

4.7 Fazit

Die Analysen zeigen, dass Kontextmerkmale, wie die soziale Zusammensetzung oder der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache einer Klasse, für die Leistungen der Schülerinnen und Schüler von Bedeutung sind, unabhängig von ihrer eigenen sozialen Herkunft oder ihrer Erstsprache. Bei gleichen individuellen Merkmalen erzielten Schülerinnen und Schüler in Klassen mit einer vorwiegend privilegierten sozialen Zusammensetzung bessere Deutsch- und Mathematikleistungen als solche in Klassen mit einer weniger privilegierten sozialen Zusammensetzung, und je höher der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache ist, desto tiefer sind auch die Deutsch- und Mathematikleistungen. Beide Klassenmerkmale hängen relativ stark miteinander zusammen. Je höher die soziale Zusammensetzung einer Klasse ist, desto tiefer ist auch der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache. Deshalb ist es schwierig zu eruieren, welches der beiden Klassenmerkmale genau für die geringeren Leistungen verantwortlich gemacht werden kann. Die Analysen sprechen jedoch dafür, dass die soziale Zusammensetzung einer Klasse der entscheidende Faktor ist und weniger der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache. Kontextmerkmale, wie die soziale Zusammensetzung einer Klasse, wirken sich typischerweise nicht unmittelbar auf die schulischen Lern- und Entwicklungsmöglichkeiten aus. Vermittelt über Wertorientierungen innerhalb der Eltern- und Schülerschaft (z.B. Leistungserwartungen und -normen), Unterstützungsmöglichkeiten, Erwartungen der Lehrkräfte und Aspekte der Unterrichtsgestaltung können sie jedoch die Lern- und Entwicklungsmöglichkeiten beeinflussen (vgl. Baumert, Stanat & Watermann, 2006).

Mit der Lernstandserhebung ist keine Evaluation zur Wirkung des QUIMS-Programms auf die Leistungen der Schülerinnen und Schüler möglich. So lassen sich beispielsweise die Klassen mit QUIMS-Unterstützung nicht über mehrere Testzeitpunkte miteinander vergleichen, weil zwischen der 3. und der 6. Klasse keine ganzen Klassen verfolgt werden konnten. Die repräsentative Stichprobe der 6. Klassen erlaubt es jedoch, die Leistungen der Klassen, die bereits vor 2006 von QUIMS unterstützt wurden, mit den Leistungen der Klassen zu vergleichen, die erst nach 2008 von QUIMS unterstützt wurden. Die Ergebnisse zeigen, dass die Deutsch- und Mathematikleistungen in den QUIMS-Klassen im Vergleich zu den Klassen mit einem geringen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache schlechter ausfallen. Die Unterschiede sind jedoch als mittelstark bis eher klein zu beurteilen, insbesondere vor dem Hintergrund des deutlich höheren Anteils an Schülerinnen und Schülern mit Deutsch als Zweitsprache in den Klassen mit QUIMS-Unterstützung. Zwischen den Klassen, die vor 2006 von QUIMS unterstützt wurden, und den Klassen mit einem vergleichbar hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, die erst nach 2008 von QUIMS unterstützt wurden, finden sich keine Unterschiede in den Leistungen. Dabei gilt es zu berücksichtigen, dass die Zusammenhänge zwischen den QUIMS-Massnahmen und den Leistungen wahrscheinlich komplexer sind, als anhand der vorliegenden Datenbasis überprüft werden kann. Zudem verfolgt QUIMS neben der Förderung der schulischen Leistungen auch weitere Ziele, wie die Förderung des Schulerfolgs und der Integration.

Die Datenbasis der Längsschnittstichprobe konnte mit individuellen Angaben zur Nutzung des Förderangebots Deutsch als Zweitsprache (DaZ-Unterricht) seit Beginn der 3. Klasse verknüpft werden. Dadurch sind kausale Analysen zur Wirkung des DaZ-Unterrichts auf die Entwicklung der Deutschleistungen der Schülerinnen und Schüler mit Deutsch als Zweitsprache möglich. Die Analysen zeigen eine positive Wirkung des DaZ-Unterrichts auf die Leistungs-

entwicklung in Deutsch vom Schuleintritt bis zum Ende der 3. Klasse. Dies gilt jedoch nur für Schülerinnen und Schüler, die den DaZ-Unterricht in der 3. Klasse während durchschnittlich zwei Wochenlektionen besucht haben. Für Schülerinnen und Schüler, die den DaZ-Unterricht weniger intensiv in Anspruch genommen haben, konnte keine positive Wirkung auf die Leistungsentwicklung nachgewiesen werden. Einschränkend gilt es dabei zu erwähnen, dass für die Analysen keine Angaben zur Nutzung des DaZ-Unterrichts vor der 3. Klasse zur Verfügung standen. Es konnte somit nicht berücksichtigt werden, ob eine Schülerin oder ein Schüler bereits im Verlaufe der ersten zwei Schuljahre eine entsprechende Förderung erhalten hatte. Dies kann zu einer Unterschätzung der Wirkung des DaZ-Unterrichts führen, weil Fortschritte von Schülerinnen und Schülern aufgrund der Förderung bis zum Ende der 2. Klasse in den vorliegenden Analysen nicht erkennbar sind.

Die Leistungsunterschiede zwischen den Klassen sind am Ende der 3. Klasse etwas grösser als am Ende der 6. Klasse. Es scheint, dass die erwarteten Leistungen am Ende der 6. Klasse klarer sind als am Ende der 3. Klasse und dass der Übertritt der Schülerinnen und Schüler in die Schultypen der Sekundarstufe I zu einer Vereinheitlichung der Leistungen führt. Die Klassenmittelwerte streuen zwar auch am Ende der 6. Klasse noch beträchtlich, aber doch weniger als am Ende der 3. Klasse. Dieses Ergebnis darf allerdings nicht darüber hinwegtäuschen, dass die Leistungsunterschiede zwischen den guten und den schwachen Schülerinnen und Schülern auf der Mittelstufe insgesamt zunehmen.

5 Leistungsbeurteilung

Domenico Angelone & Urs Moser

5.1 Leistungsbeurteilung durch Noten

Während in den 1990er-Jahren in verschiedenen Kantonen Alternativen zur Beurteilung mit Noten diskutiert und zum Teil auch eingeführt wurden, werden die Noten mittlerweile wieder in allen Kantonen bereits auf der Primarstufe zur Beurteilung der Schülerinnen und Schüler genutzt. Allerdings setzt die Beurteilung anhand von Noten je nach Kanton zu verschiedenen Zeitpunkten ein. Im Kanton Aargau beispielsweise ab der 1. Klasse, im Kanton Zürich ab der 2. Klasse, im Kanton Bern ab der 3. Klasse, im Kanton Solothurn ab der 4. Klasse, im Kanton Waadt ab der 5. Klasse und im Kanton Neuenburg ab der 6. Klasse (EDK, 2010).

Grundlage der Leistungsbeurteilung ist der Vergleich. Prüfungsergebnisse werden mit den gesetzten Lernzielen (sachliche Bezugsnorm), mit früheren Lernergebnissen (individuelle Bezugsnorm) und mit den Ergebnissen anderer Schülerinnen und Schüler (soziale Bezugsnorm) verglichen. Alle drei Bezugsnormen sind für die Beurteilung und Förderung der Schülerinnen und Schüler notwendig. Sofern die Beurteilung mit der Vergabe von Berechtigungen beziehungsweise mit der Einteilung der Schülerinnen und Schüler in Schultypen verbunden ist, kommt der sozialen Bezugsnorm allerdings eine besondere Bedeutung zu. Fairerweise müssten in allen Klassen gleiche Leistungen mit gleichen Noten beurteilt werden. Die Bildungsforschung zeigt immer wieder, dass dieser Anspruch nicht eingehalten werden kann.

Die Kritik an den Noten beruht vor allem auf Erkenntnissen über Fehlerquellen und Verzerrungen bei der Leistungsbeurteilung in der Schule. Zum einen kommt es bei der Leistungsbeurteilung zu sekundären sozialen Ungleichheiten. Damit ist gemeint, dass die Noten nicht vollständig auf die Schulleistungen, sondern zum Teil auch auf herkunftsspezifische Bewertungsgrundlagen von Lehrpersonen zurückzuführen sind. Bei gleichen Leistungen werden beispielsweise Schülerinnen und Schüler aus sozial benachteiligten Verhältnissen (z.B. aus bildungsfernen Familien) schlechter beurteilt als Schülerinnen und Schüler aus sozial privilegierten Verhältnissen (vgl. Kronig, 2007; Maaz & Nagy, 2009; Rhyn & Moser, 2002).

Zum andern orientieren sich Lehrpersonen bei der Leistungsbeurteilung in der Regel an einem klasseninternen Bewertungssystem. Als soziale Bezugsnorm dienen die Leistungen der Schülerinnen und Schüler innerhalb der Klasse. Dies ist insofern problematisch, als sich Klassen in ihrem Leistungsniveau zum Teil beträchtlich unterscheiden, weshalb auch vom Referenzgruppenfehler gesprochen wird (Kronig, 2007). Im Extremfall können Referenzgruppenfehler zur Folge haben, dass Schülerinnen und Schüler mit einer bestimmten Leistung in leistungsstarken Klassen schlechte Noten und in leistungsschwachen Klassen gute Noten erhalten.

Der Nachweis von Beurteilungsfehlern lässt sich aufgrund der vorliegenden Daten nicht abschliessend erbringen, weil die eingesetzten Leistungstests nicht vollumfänglich die Bewertungsgrundlagen der Lehrpersonen abbilden. Während Leistungstests die Fähigkeiten zu einem bestimmten Zeitpunkt erfassen, erstreckt sich die Notenbewertung über ein ganzes Semester. Zudem fließen in die Note auch Faktoren wie mündliche Beiträge im Unterricht oder – speziell in Deutsch – produktive Fähigkeiten wie Schreiben und Lesen ein. Diese Fähigkeiten wurden mit den eingesetzten Leistungstests nicht erfasst. Die Beschreibung der

Leistungsbeurteilung durch Noten kann allerdings Hinweise darauf geben, welche Mechanismen bei der Notengebung in der Schule wirken.

5.2 Noten, individuelle Merkmale und Bezugsnorm

Vergleich von Testleistungen und Zeugnisnoten

Der Vergleich der Zeugnisnoten mit den Ergebnissen in den Leistungstests zeigt, dass die mit den Testaufgaben erfassten Fähigkeiten stark mit jenen Fähigkeiten zusammenhängen, die von den Lehrpersonen mit Noten beurteilt werden. Tabelle 5.1 enthält die Korrelationskoeffizienten zwischen den Testleistungen und den Zeugnisnoten. Für den Fachbereich Deutsch beträgt der Korrelationskoeffizient $r = 0.69$, für den Fachbereich Mathematik beträgt der Korrelationskoeffizient $r = 0.71$. Der Zusammenhang zwischen den Zeugnisnoten für Deutsch und Mathematik beträgt $r = 0.73$. Die Testleistungen untereinander hängen etwas weniger stark zusammen.

Tabelle 5.1: Korrelation zwischen den Testleistungen und den Zeugnisnoten

	Testleistung in Deutsch	Testleistung in Mathematik	Zeugnisnoten in Deutsch
Zeugnisnoten in Deutsch	0.69		
Zeugnisnoten in Mathematik		0.71	0.73
Testleistung in Deutsch		0.61	0.69

Die Korrelationen zwischen den Testleistungen und den Noten sind stark, aber nicht perfekt. Bei der Interpretation muss berücksichtigt werden, dass es sich um zwei unterschiedliche Leistungsbeurteilungen handelt. Die Note bezieht sich auf die Leistungen während eines Semesters. Zudem werden mit der Note mehr Fähigkeiten beurteilt, als mit den Leistungstests erfasst wurden. Beispielsweise beschränken sich die Leistungstests in beiden Fachbereichen auf reproduktive Kompetenzen. In Deutsch wurde beispielsweise weder das Hörverstehen noch das Schreiben geprüft.

Noten nach individuellen Merkmalen

Tabelle 5.2 zeigt die Zeugnisnoten in den beiden Fachbereichen nach individuellen Merkmalen der Schülerinnen und Schüler. Die Darstellung der Noten nach Erstsprache, Geschlecht und sozialer Herkunft entspricht den Leistungen dieser Gruppen. Die Mädchen erzielen in Deutsch bessere Leistungen als die Knaben und werden im Durchschnitt auch etwas besser beurteilt. In der Mathematik fallen die Leistungs- und Notenunterschiede zugunsten der Knaben aus. Schülerinnen und Schüler mit Deutsch als Erstsprache erzielen in beiden Fachbereichen bessere Testleistungen, was sich auch in den Noten niederschlägt. Je höher die soziale Herkunft, desto besser sind die Testleistungen und die Noten.

Tabelle 5.2: Zeugnisnoten nach individuellen Merkmalen

	Deutsch	Mathematik
Notendurchschnitt	4.60	4.60
Geschlecht		
– Knaben	4.54	4.62
– Mädchen	4.67	4.53
Erstsprache		
– Deutsch als Erstsprache	4.79	4.77
– Deutsch als Zweitsprache	4.32	4.33
Soziale Herkunft		
– benachteiligt	4.22	4.15
– eher benachteiligt	4.47	4.41
– eher privilegiert	4.66	4.66
– privilegiert	4.98	5.00

Die Darstellung der Noten nach individuellen Merkmalen ist ein Indikator für primäre soziale Ungleichheiten; also Ungleichheiten, die sich aufgrund der Leistungen der Schülerinnen und Schüler erklären lassen. Darüber hinaus lässt sich überprüfen, ob beispielsweise Knaben bei gleichen Leistungen schlechter oder besser beurteilt werden als Mädchen. Wenn gleiche Leistungen unterschiedlich bewertet werden, dann ist dies ein Hinweis darauf, dass sekundäre soziale Ungleichheiten vorhanden sind und sich die Beurteilung nicht ausschliesslich durch die Leistungen erklären lässt.

Tabelle 5.3: Bedeutung individueller Merkmale für Zeugnisnoten

	Deutschnoten	Mathematiknoten
Testleistungen: 100 Punkte	0.40	0.46
Knaben	nicht signifikant	nicht signifikant
Alter in Monaten	-0.01	-0.01
Deutsch als Zweitsprache	-0.15	-0.10
Soziale Herkunft: privilegiert versus benachteiligt	0.26	0.26

Tabelle 5.3 zeigt die Bedeutung verschiedener individueller Merkmale für die Beurteilung anhand der Zeugnisnoten. Ein Anstieg der Testleistungen um 100 Punkte geht in Deutsch mit einem Anstieg von 0.40 Notenpunkten einher, in der Mathematik mit einem Anstieg von

0.46 Notenpunkten. Bei gleichen Leistungen erhalten Knaben und Mädchen die gleichen Noten. Alter, Erstsprache und soziale Herkunft sind hingegen für die Beurteilung anhand der Noten von Bedeutung.

Ältere Schülerinnen und Schüler werden bei gleichen Leistungen etwas schlechter beurteilt. Eine Altersdifferenz von 12 Monaten führt bei gleichen Leistungen zu einer Note, die rund 0.12 Punkte tiefer liegt, sowohl in Deutsch als auch in der Mathematik. Die Zeugnisnoten der Schülerinnen und Schüler mit Deutsch als Zweitsprache fallen bei gleichen Testleistungen in Deutsch um 0.15 Notenpunkte und in der Mathematik um 0.10 Notenpunkte tiefer aus. Und auch die soziale Herkunft spielt für die Beurteilung eine Rolle. Je höher die soziale Herkunft ist, desto höher sind die Noten. Schülerinnen und Schüler aus sozial privilegierten Verhältnissen werden bei gleichen Testleistungen um 0.26 Notenpunkte höher beurteilt als solche aus benachteiligten Verhältnissen.

Beispiel

Ein Schüler mit Deutsch als Zweitsprache aus sozial benachteiligten Verhältnissen, der 13 Jahre alt ist, wird bei gleicher Testleistung in der Mathematik eine halbe Note, in Deutsch sogar mehr als eine halbe Note tiefer beurteilt als ein Schüler mit Deutsch als Erstsprache, der aus sozial privilegierten Verhältnissen stammt und 12 Jahre alt ist.

Dass Alter, Erstsprache und soziale Herkunft für die Notenbeurteilung eine Rolle spielen, könnte ein Hinweis auf sekundäre Ungleichheiten sein. Allerdings gilt es zu berücksichtigen, dass mit den Testleistungen verschiedene Fähigkeiten nicht abgedeckt werden. Zudem stellen Testleistungen ein «einmaliges» Ergebnis dar, die Zeugnisnoten hingegen beruhen auf einer Beurteilung, die sich auf ein halbes Schuljahr bezieht. Es könnte auch sein, dass die Effekte von Alter, Erstsprache und sozialer Herkunft bei der Berücksichtigung mündlicher Kompetenzen nicht mehr vorhanden wären. Der Alterseffekt kann auch dadurch erklärt werden, dass ältere Schülerinnen und Schüler meist eine Klasse repetiert haben und deshalb eher zu den schwächeren Schülerinnen und Schülern innerhalb einer Klasse gehören.

Noten nach durchschnittlichen Fähigkeiten der Klasse

Für die Beurteilung der Leistungen aufgrund von Noten gibt es kein standardisiertes Vorgehen. Zwar sind aufgrund der Ziele im Lehrplan und des Schwierigkeitsgrads der Aufgaben in den Lehrmitteln relativ klare Anhaltspunkte darüber vorhanden, was einer guten oder weniger guten Leistung entspricht. Eine direkte Zuordnung von Fähigkeiten und Noten fehlt hingegen. Lehrpersonen orientieren sich deshalb bei der Beurteilung aufgrund der Noten am Leistungsniveau der Klasse.

In den Abbildungen 5.1 und 5.2 sind die Leistungsverteilungen in Deutsch und Mathematik nach den Zeugnisnoten dargestellt. Die Gesamtlänge des Balkens umfasst 90 Prozent der Schülerleistungen. 50 Prozent der Schülerleistungen liegen innerhalb der dunkelgrauen Balken. Der kleine schwarze Balken stellt jenen Bereich dar, in dem der Mittelwert mit einer statistischen Sicherheit von 95 Prozent liegt.

Abbildung 5.1: Leistungsverteilung nach Deutschnoten

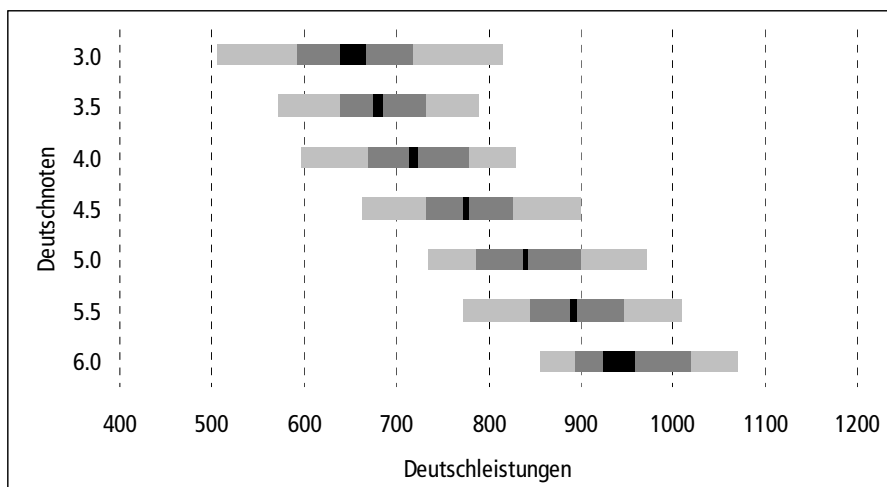
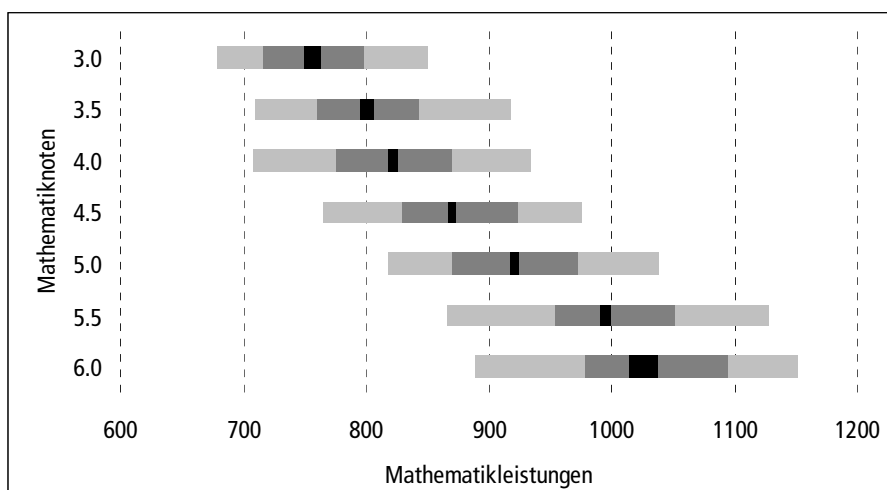


Abbildung 5.2: Leistungsverteilung nach Mathematiknoten

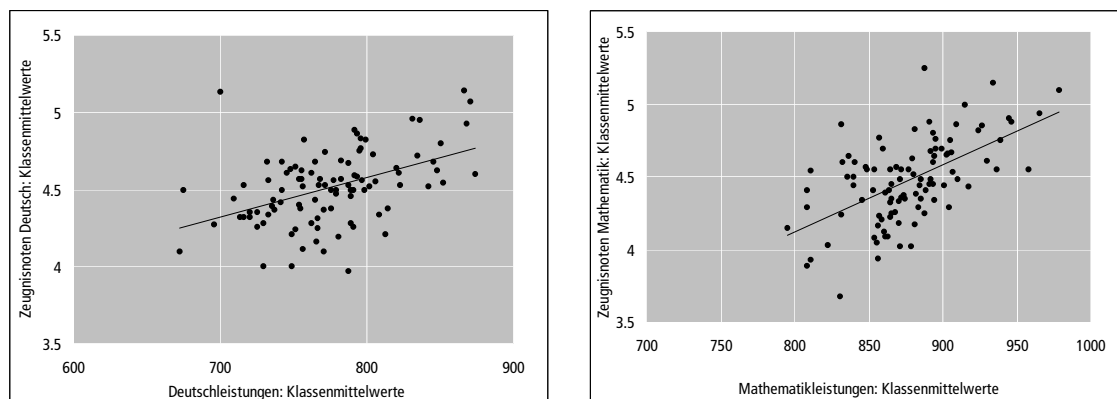


Mit steigenden Noten nehmen auch die durchschnittlichen Leistungen zu. Die Balken zeigen aber auch, dass beträchtliche Leistungsüberschneidungen zwischen den einzelnen Notenkategorien vorzufinden sind. In Deutsch können beispielsweise Schülerinnen und Schüler, die 800 Punkte erreichen, Noten zwischen 4 und 5.5 erhalten. In der Mathematik liegen die Noten von Schülerinnen und Schülern, die 900 Punkte erzielen, zwischen 3.5 und 6.

Abbildung 5.3 zeigt den Zusammenhang zwischen Testleistung und Noten, zusammengefasst pro Klasse. Jeder Punkt repräsentiert eine Klasse. Die Position einer Klasse ergibt sich aus der durchschnittlichen Klassenleistung (Horizontale) und der durchschnittlichen Klassennote (Vertikale). Die durchschnittlichen Klassennoten streuen in Deutsch zwischen den Noten 4 und 5.1, in der Mathematik zwischen den Noten 3.7 und 5.2. Die mittleren 90 Prozent der durchschnittlichen Klassennoten liegen zwischen 4 und 5, die mittleren 50 Prozent der durch-

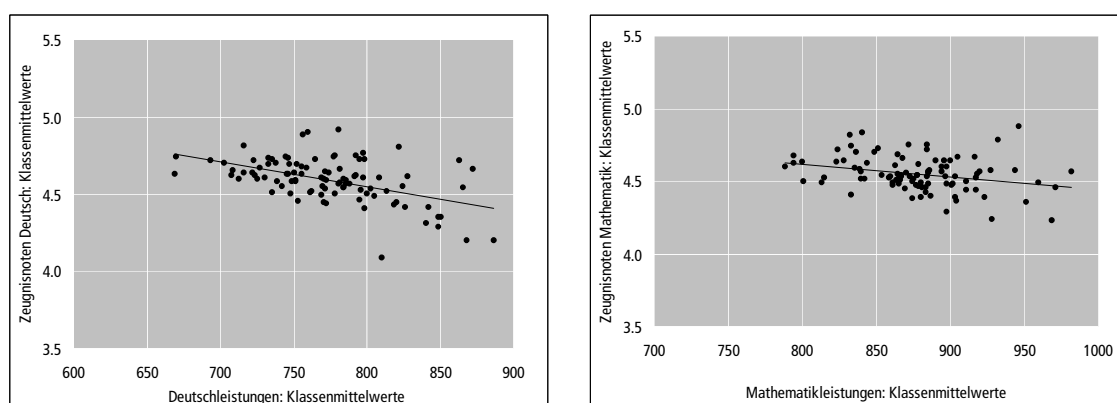
schnittlichen Klassennoten liegen zwischen 4.3 und 4.7. Die Verteilungen der Zeugnisnoten zeigen, dass die durchschnittlichen Klassenmittelwerte zum Teil beträchtlich von den Gesamtmittelwerten aller Klassen abweichen, die bei 4.5 liegen.

Abbildung 5.3: Testleistungen und Noten: Klassenmittelwerte



Die Korrelation zwischen den durchschnittlichen Klassenleistungen und den durchschnittlichen Klassennoten beträgt für Deutsch $r = .45$, für Mathematik $r = .56$. Je höher die durchschnittliche Klassenleistung ist, desto höher ist auch die durchschnittliche Klassennote. Dieser Zusammenhang ist zwar relativ stark, aber relativ weit vom idealen Zusammenhang «je höher, desto höher» entfernt. Beispielsweise liegen die durchschnittlichen Noten der Klassen mit einer durchschnittlichen Mathematikleistung von 900 Punkten mehrheitlich zwischen 4 und 5.

Abbildung 5.4: Testleistungen nach statistischer Kontrolle der individuellen Merkmale und Noten: Klassenmittelwerte



Wenn die Bezugsgruppenorientierung der Lehrpersonen bei der Notengebung zutrifft, dann müssten die durchschnittlichen Noten in Klassen mit vergleichsweise guten Leistungen tiefer ausfallen als in Klassen mit vergleichsweise schlechten Leistungen. Um diese Hypothese überprüfen zu können, wurde untersucht, welche Bedeutung das Leistungsniveau einer Klas-

se für den Beurteilungsmassstab von Lehrpersonen hat. Die Ergebnisse dieser Analysen sind in Abbildung 5.4 grafisch umgesetzt. Abbildung 5.4 zeigt wieder den Zusammenhang zwischen Testleistung und Noten, zusammengefasst pro Klasse. Jeder Punkt repräsentiert eine Klasse. Die Position der Klasse ergibt sich aus der durchschnittlichen Klassenleistung (Horizontale) und der durchschnittlichen Klassennote (Vertikale). Bei dieser Darstellung wurde allerdings die Bedeutung der individuellen Merkmale statistisch kontrolliert. Je höher die durchschnittliche Klassenleistung ist, desto tiefer ist auch die durchschnittliche Klassennote. In Klassen mit besseren Leistungen wird also ein strengerer Beurteilungsmassstab angewendet.

Beispiele

Ein Schüler erhält für seine Leistungen in einer Klasse mit einem Mittelwert von 750 Punkten eine um 0.27 Punkte bessere Deutschnote als in einer Klasse mit einem Mittelwert von 850 Punkten. Eine Schülerin erhält für ihre Leistungen in einer Klasse mit einem Mittelwert von 800 Punkten eine um 0.21 Punkte bessere Mathematiknote als in einer Klasse mit einem Mittelwert von 900 Punkten.

Berücksichtigen wir nun noch die individuellen Merkmale, die für die Note eine Bedeutung haben, dann erhält beispielsweise ein Schüler mit Deutsch als Zweitsprache aus sozial benachteiligten Verhältnissen, der 13 Jahre alt ist, in einer sehr guten Klasse im Vergleich zu einem Schüler mit Deutsch als Erstsprache aus sozial privilegierten Verhältnissen, der 12 Jahre alt ist und eine sehr schwache Klasse besucht (200 Punkte Unterschied), für die gleichen Leistungen rund eine Note weniger.

5.3 Fazit

Die vorliegenden Ergebnisse weisen darauf hin, dass die Beurteilung der Schülerinnen und Schüler zu den schwierigen Aufgaben der Lehrpersonen gehört. Zwar ist der Zusammenhang zwischen der Note und den Testleistungen gross und Unterschiede in den Noten sind vorwiegend auf fachliche Leistungen zurückzuführen. Es gibt aber auch Anzeichen dafür, dass sich Lehrpersonen bei der Beurteilung systematisch an gewissen individuellen Merkmalen ihrer Schülerinnen und Schüler orientieren. Zudem hängt die Leistungsbeurteilung von Lehrpersonen davon ab, ob sie eine leistungsstarke oder eine leistungsschwache Klasse unterrichten. Lehrpersonen von leistungsstarken Klassen beurteilen strenger als solche von leistungsschwachen Klassen.

Die Orientierung der Leistungsbeurteilung an individuellen Merkmalen führt für Schülerinnen und Schüler aus sozial benachteiligten Verhältnissen zu einem Nachteil. Schülerinnen und Schüler aus sozial benachteiligten Verhältnissen werden bei gleichen Testleistungen etwas tiefer beurteilt als solche aus sozial privilegierten Verhältnissen. Diese Benachteiligung gleicht sich dann aus, wenn diese Schülerinnen und Schüler eine eher leistungsschwache Klasse besuchen, in der der Beurteilungsmassstab offensichtlich milder ist als in leistungsstarken Klassen.

Die Orientierung der Leistungsbeurteilung am Klassenverband führt für Schülerinnen und Schüler in leistungsstarken Klassen vor allem deshalb zu einer Benachteiligung, weil die Noten für den Übertritt in die Schultypen der Sekundarstufe I genutzt werden. Für den Übertritt ins Langgymnasium zählen beispielsweise zur Hälfte die Erfahrungsnoten. Wer eine leistungsstarke Klasse besucht, lernt zwar aufgrund der Klassenzusammensetzung eher etwas mehr, wird jedoch strenger beurteilt.

Weil Schülerinnen und Schüler aus sozial privilegierten Verhältnissen überproportional häufig in leistungsstarken Klassen unterrichtet werden, kann es zu einer indirekten herkunftsbedingten Benachteiligung kommen: Schülerinnen und Schüler aus sozial privilegierten Verhältnissen können ebenfalls benachteiligt werden; dann nämlich, wenn ihre Leistungen aufgrund eines zu strengen Beurteilungsmassstabes im klassenübergreifenden Vergleich zu tief beurteilt werden. Benachteiligt sind allerdings alle Schülerinnen und Schüler, die für gleiche Leistungen in leistungsstarken Klassen tiefer benotet werden als in leistungsschwachen.

Insgesamt stimmt die Leistungsbeurteilung der Lehrpersonen erstaunlich gut mit den Testleistungen überein. Die Ergebnisse zeigen, dass das Urteil der Lehrpersonen am Ende der Primarstufe weit zuverlässiger ist, als aufgrund der öffentlichen Diskussion des Themas angenommen werden musste. Auch die nachgewiesenen Verzerrungen sollten mit Vorsicht interpretiert werden. Vielleicht würde das Ergebnis anders ausfallen, wenn mit der Lernstandserhebung auch mündliche beziehungsweise produktive Kompetenzen erfasst worden wären.

Zudem ist es verständlich, dass sich Lehrpersonen bei der Beurteilung am Klassenverband orientieren, weil bis anhin keine geeigneten Instrumente für eine klassenübergreifende Beurteilung vorliegen. Mit den geplanten Leistungsmessungen zur Überprüfung der Bildungsstandards wird die Diskussion über Anforderungen und Erwartungen an die Kompetenzen der Schülerinnen und Schüler zunehmen – eine Diskussion, die für die Leistungsbeurteilung hilfreich sein könnte, weil Ziele in Form von Kompetenzerwartungen und konkreten Aufgabenbeispielen umschrieben werden.

6 Übertritt von der Primarschule in die Sekundarstufe I

Domenico Angelone & Urs Moser

6.1 Kritische Schnittstelle

Am Ende der 6. Klasse besteht für die Schülerinnen und Schüler im Kanton Zürich die Möglichkeit, nach bestandener Aufnahmeprüfung in ein Langgymnasium zu wechseln mit dem Ziel, die Maturität zu erreichen. Die übrigen Schülerinnen und Schüler werden aufgrund einer Gesamtbeurteilung einer Abteilung der Sekundarstufe I zugeteilt. Je nach Gemeinde stehen dabei zwei oder drei Abteilungen zur Verfügung (A und B oder A, B und C). Die Zuteilung zu den Abteilungen wird von den betroffenen Lehrpersonen, der Schulleitung und den Eltern festgelegt. Neben den kognitiven Fähigkeiten werden auch das Arbeits-, Lern- und Sozialverhalten sowie die persönliche Entwicklung beurteilt.

Der Übertritt in die Abteilungen der Sekundarstufe I wird durch einen Bildungsentscheid geregelt. Die Grundlage für die Entscheidung bilden vor allem die kognitiven Fähigkeiten, jedoch auch fächerübergreifende Kompetenzen. Die Bedeutung der schulischen Leistungen ist zentral, vermutlich auch deshalb, weil ein einheitliches Notensystem zur Verfügung steht. Die Beurteilung der Leistungen kann von den Lehrpersonen allerdings bereits aufgrund von sozial selektiven Erwartungs- und Wertschätzungsstrukturen der Schule verzerrt sein. Zudem fließt auch die familiäre Bewertung von Bildung in den Übertrittsentscheid ein. Inwiefern Bildungsentscheidungen tatsächlich durch fachliche Leistungen gedeckt sind, lässt sich folglich nicht so einfach belegen. Divergenzen zwischen Noten, fachlichen Leistungen und der Zuteilung zu den Abteilungen der Sekundarstufe I können deshalb nicht als Diskriminierung interpretiert werden. Was hingegen ohne Einschränkung nachgewiesen werden kann, ist, ob sich soziale Ungleichheiten bei den Schnittstellen verfestigen oder gar vergrössern (Maaz, Baumert & Trautwein, 2009, S. 13).

Der Übergang von der Primarstufe in die Sekundarstufe I bedeutet zudem für den grossen Teil der Schülerinnen und Schüler bereits eine zentrale Weichenstellung für die Berufswahl. Wer sich für die Berufsbildung entscheidet, wird sich spätestens im zweiten Jahr der Sekundarstufe I mit der Berufswahl auseinandersetzen. Die Weichen für die Berufswahl werden allerdings bereits beim Übertritt von der Primarstufe in die Sekundarstufe I gestellt. Für Jugendliche, die auf der Sekundarstufe I die Abteilung B oder C besuchen, ist die Berufswahl bereits erheblich eingeschränkt. Mit der Einteilung der Schülerinnen und Schüler in Schultypen leistet die Schule einen (vor-)entscheidenden Beitrag für die Berufslaufbahn der Jugendlichen. Der Übertritt von der Primarschule in die Sekundarstufe I gilt deshalb als kritische Schnittstelle, auch im Hinblick auf die Verfestigung von sozialen Ungleichheiten (Baumert, Stanat & Watermann, 2006).

6.2 Leistungen nach zugewiesenem Schultyp

Die dritte Lernstandserhebung fand am Ende der 6. Klasse statt. Zu diesem Zeitpunkt standen die Übertrittsentscheide in die Schultypen der Sekundarstufe I bereits fest. Die Schülerinnen und Schüler wurden deshalb im Schulfragebogen gebeten, die Übertrittsentscheide anzugeben (Langgymnasium beziehungsweise Abteilung A, B oder C).

Tabelle 6.1 zeigt die Verteilung der Schülerinnen und Schüler auf die Abteilungen der Sekundarstufe I. In der zweiten Spalte sind die prozentualen Anteile der Schülerinnen und Schüler nach Abteilung des Schuljahres 2007/2008 angegeben. Die Zahlen stammen von der Bildungsstatistik der Bildungsdirektion des Kantons Zürich. In der dritten Spalte sind die prozentualen Anteile der Schülerinnen und Schüler nach Abteilung der Stichprobe der Lernstandserhebung im Schuljahr 2008/2009 enthalten. Obwohl zwei aufeinanderfolgende Schuljahrgänge miteinander verglichen werden und obwohl es sich zum einen um Daten der Bildungsstatistik, zum andern um hochgerechnete Daten aus der Lernstandserhebung handelt, ist die Verteilung der Schülerinnen und Schüler auf die Abteilungen sehr ähnlich. Dies spricht für die gute Repräsentativität der Stichprobe der Lernstandserhebung. Rund 16 Prozent der befragten Schülerinnen und Schüler haben den Übertritt ins Langgymnasium geschafft. Rund 48 Prozent der Schülerinnen und Schüler wurden in die Abteilung A, rund 31 Prozent in die Abteilung B und rund 4 Prozent in die Abteilung C eingeteilt. Rund 1 Prozent der Schülerinnen und Schüler repetiert die 6. Klasse der Primarstufe.

Tabelle 6.1: Verteilung der Schülerinnen und Schüler auf die Schultypen der Sekundarstufe I: Bildungsstatistik 2008 und Stichprobe der Lernstandserhebung 2009

	Bildungsstatistik Schuljahr 2007/08	Stichprobe Lernstandserhebung 2009 ¹
Langgymnasium	16.0%	16.1%
Abteilung A	45.5%	47.5%
Abteilung B	32.7%	31.1%
Abteilung C	5.7%	4.2%
6. Klasse		1.0%
Total	100%	100%

Anmerkung: ¹Auf die Population hochgerechnete Angaben

Für den Übertrittsentscheid in die Schultypen der Sekundarstufe I sind vor allem die Leistungen in den Fächern Deutsch und Mathematik entscheidend. Die im Rahmen der Lernstandserhebung eingesetzten Leistungstests erfassen zwar nur einen Teil der schulischen Leistungen, die als Entscheidungsgrundlage für den Übertritt dienen, namentlich im Fach Deutsch. Die Lehrplanbereiche «Hören und Sprechen» sowie «Schreiben» beispielsweise wurden mit dem Deutschttest nicht erfasst. Trotzdem wird erwartet, dass die Testleistungen mit der vorgesehenen Einteilung in die Schultypen der Sekundarstufe I eng zusammenhängen.

Die Mittelwerte der Deutschleistungen unterscheiden sich zwischen den Schultypen unterschiedlicher Anforderungen entsprechend den Erwartungen statistisch signifikant. Schülerinnen und Schüler, die die Prüfung in das Langgymnasium bestanden, erreichten einen Mittelwert von 907 Punkten. Jene, die der Abteilung A zugeteilt wurden, erreichten mit einem Mittelwert von 811 Punkten 96 Punkte weniger. Schülerinnen und Schüler, die in die Abtei-

lung B eingeteilt wurden, erreichten einen Mittelwert von 721 Punkten und jene, die in die Abteilung C eingeteilt wurden, 653 Punkte.

Für die Mittelwerte der Mathematikleistungen gilt das Gleiche: Sie unterscheiden sich zwischen den Schultypen mit unterschiedlichen Anforderungen statistisch signifikant und deutlich. Der Mittelwert der Schülerinnen und Schüler, die die Prüfung ins Langgymnasium bestanden, beträgt 1004 Punkte. Schülerinnen und Schüler, die in die Abteilung A der Sekundarschule eingeteilt wurden, erreichten einen Mittelwert von 906 Punkten. Schülerinnen und Schüler, die in die Abteilung B eingeteilt wurden, erreichten einen Mittelwert von 813 Punkten und jene, die in die Abteilung C eingeteilt wurden, 751 Punkte.

Trotz dieser deutlichen Unterschiede in den durchschnittlichen Leistungen finden sich Überschneidungsbereiche in den Leistungen der Schülerinnen und Schüler nach den zugewiesenen Schultypen. In den Abbildungen 6.1 und 6.2 sind die Verteilungen der Testleistungen nach dem zugewiesenen Schultyp dargestellt. Die Gesamtlänge des Balkens umfasst 90 Prozent der Schülerleistungen. 50 Prozent der Schülerleistungen liegen innerhalb der dunkelgrauen Balken. Der kleine schwarze Balken stellt jenen Bereich dar, in dem der Mittelwert mit einer statistischen Sicherheit von 95 Prozent liegt.

Abbildung 6.1: Deutschleistungen nach zugewiesenem Schultyp

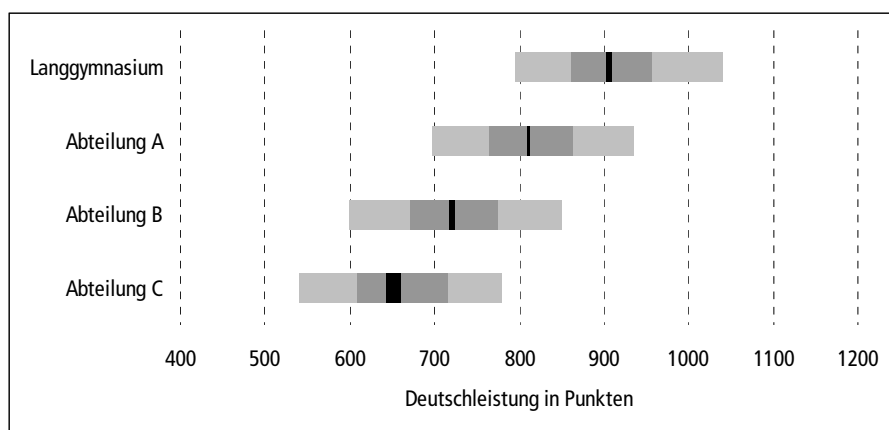
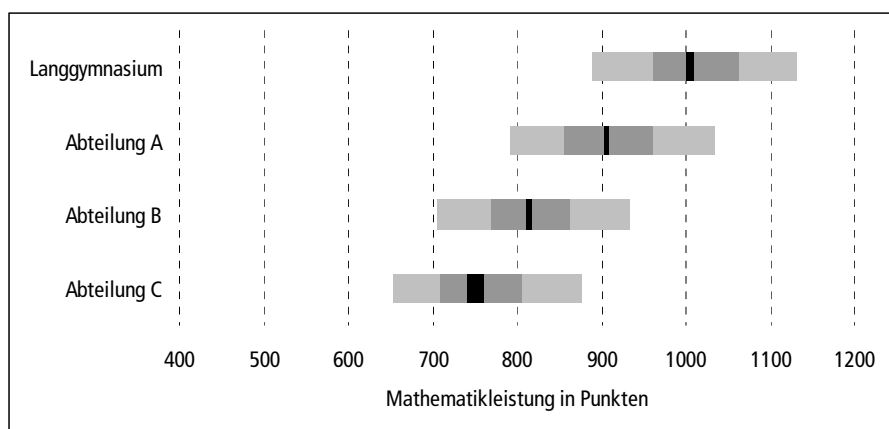


Abbildung 6.2: Mathematikleistungen nach zugewiesenem Schultyp



Um die Bedeutung der Überschneidungsbereiche besser beurteilen zu können, sind in der Tabelle 6.2 die Prozentanteile an Schülerinnen und Schülern ausgewiesen, die bessere Leistungen aufweisen als die schwächsten 10 Prozent (Prozentrang 10), als die schwächsten 25 Prozent (Prozentrang 25) und als die schwächere Hälfte (Prozentrang 50) der Schülerinnen und Schüler eines anspruchsvolleren Schultyps.

Tabelle 6.2: Prozentanteile an Schülerinnen und Schülern der Abteilungen C, B und A, die ein bestimmtes Leistungsniveau im anspruchsvolleren Schultyp erreichen

	Abteilung B			Abteilung A			Langgymnasium		
	Prozentrang	Prozentrang	Prozentrang	Prozentrang	Prozentrang	Prozentrang	Prozentrang	Prozentrang	Prozentrang
Abteilung C	10	25	50	10	25	50	10	25	50
Deutsch	69%	40%	14%						
Mathematik	60%	34%	15%						
Abteilung B	10	25	50	10	25	50	10	25	50
Deutsch				48%	28%	12%			
Mathematik				46%	24%	10%			
Abteilung A	10	25	50	10	25	50	10	25	50
Deutsch							44%	24%	9%
Mathematik							46%	22%	10%

In Deutsch erreichen rund 69 Prozent der Schülerinnen und Schüler der Abteilung C bessere Leistungen als die schwächsten 10 Prozent der Abteilung B. Rund 40 Prozent erreichen bessere Leistungen als die schwächsten 25 Prozent und rund 15 Prozent erreichen bessere Leistungen als die schwächere Hälfte der Schülerinnen und Schüler der Abteilung B.

In der Abteilung B erreichen rund 48 Prozent der Schülerinnen und Schüler bessere Deutschleistungen als die schwächsten 10 Prozent der Abteilung A. Rund 28 Prozent erreichen bessere Leistungen als die schwächsten 25 Prozent und rund 12 Prozent erreichen bessere Leistungen als die schwächere Hälfte der Schülerinnen und Schüler der Abteilung A.

In der Abteilung A erreichen rund 44 Prozent der Schülerinnen und Schüler bessere Deutschleistungen als die schwächsten 10 Prozent des Langgymnasiums. Rund 24 Prozent erreichen bessere Leistungen als die schwächsten 25 Prozent und rund 9 Prozent erreichen bessere Leistungen als die schwächere Hälfte der Schülerinnen und Schüler des Langgymnasiums.

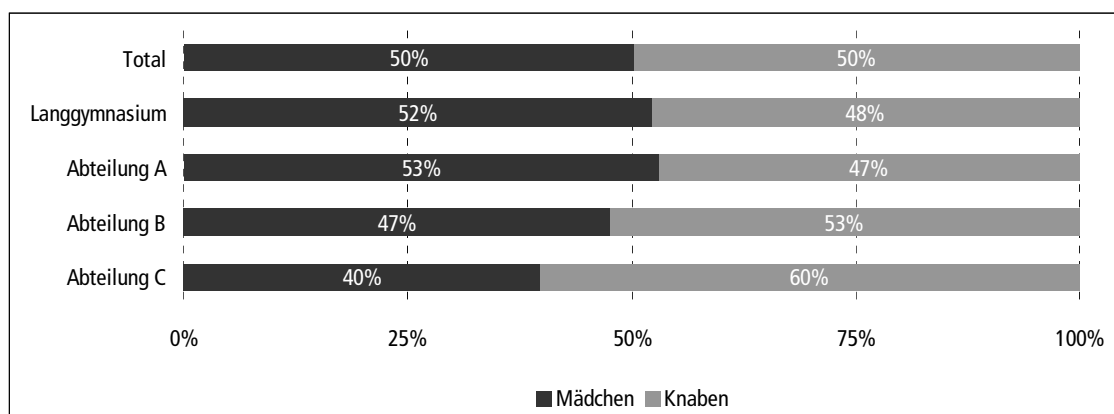
Die Überschneidungen der Leistungen verschiedener Schultypen sind in der Mathematik ähnlich gross wie in Deutsch, weshalb auch die Anteile an Schülerinnen und Schülern, die ein bestimmtes Leistungsniveau im anspruchsvolleren Schultyp erreichen, in der Mathematik und in Deutsch ähnlich gross sind.

6.3 Individuelle Merkmale nach zugewiesenem Schultyp

Geschlecht

Abbildung 6.3 zeigt, wie sich die Knaben und Mädchen auf die Schultypen der Sekundarstufe I verteilen. Im Langgymnasium und in der Abteilung A der Sekundarschule sind die Mädchen mit rund 52 Prozent und rund 53 Prozent knapp in der Mehrheit. In der Abteilung B sind hingegen mit 53 Prozent und in der Abteilung C mit 60 Prozent die Knaben in der Mehrheit.

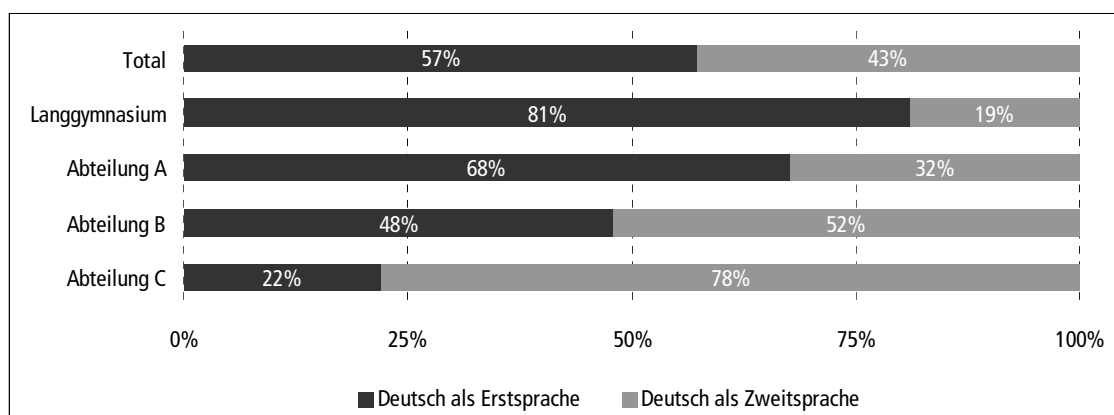
Abbildung 6.3: Anteil an Schülerinnen und Schülern nach Geschlecht und Schultyp



Erstsprache

Abbildung 6.4 zeigt, wie sich die Schülerinnen und Schüler nach der Erstsprache auf die Schultypen der Sekundarstufe I verteilen.

Abbildung 6.4: Anteil an Schülerinnen und Schülern nach Erstsprache und Schultyp



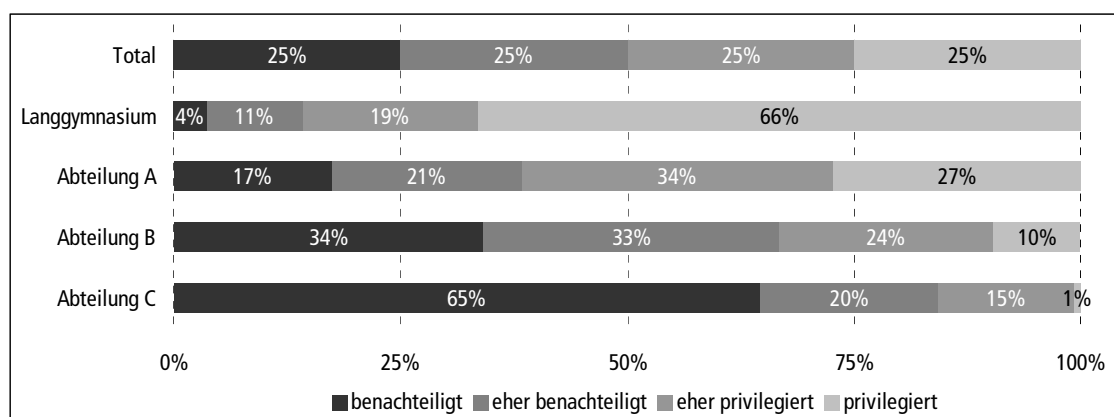
In der Abteilung C beträgt der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache 78 Prozent, im Langgymnasium 19 Prozent. Nahezu umgekehrt ist die Verteilung der Schülerinnen und Schüler mit Deutsch als Erstsprache. Für 81 Prozent der Schülerinnen und

Schüler, die die Prüfung ins Langgymnasium bestanden, entspricht Deutsch der Erstsprache. In der Abteilung C beträgt der Anteil an Schülerinnen und Schülern mit Deutsch als Erstsprache 22 Prozent. Je anspruchsvoller der Schultyp ist, desto geringer ist der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache.

Soziale Herkunft

Abbildung 6.5 zeigt, wie sich die Schülerinnen und Schüler nach der sozialen Herkunft auf die zugeteilten Schultypen der Sekundarstufe I verteilen. Die Schülerinnen und Schüler wurden aufgrund des Index zur sozialen Herkunft in vier gleich grosse Gruppen eingeteilt (privilegiert, eher privilegiert, eher benachteiligt, benachteiligt). 66 Prozent der Schülerinnen und Schüler, die die Aufnahmeprüfung ins Langgymnasium bestanden, stammen aus sozial privilegierten Verhältnissen. Demgegenüber stammen 65 Prozent der Schülerinnen und Schüler, die der Abteilung C zugewiesen wurden, aus sozial benachteiligten Verhältnissen.

Abbildung 6.5: Anteil an Schülerinnen und Schülern nach sozialer Herkunft und Schultyp



6.4 Übertritt ins Langgymnasium

Zusatzunterricht zur Prüfungsvorbereitung

Zusatzunterricht zur Vorbereitung der Aufnahmeprüfung ins Langgymnasium wird vor allem von privaten Bildungsinstitutionen, zum Teil aber auch von öffentlichen Primarschulen angeboten. Während private Institutionen Kurse zur Prüfungsvorbereitung in Kleingruppen ausserhalb der Schule anbieten, findet die Prüfungsvorbereitung an der öffentlichen Schule einerseits mit der ganzen Klasse, andererseits auch in kleineren Gruppen statt.

Die vorliegenden Ergebnisse beziehen sich auf die Prüfungsvorbereitung ausserhalb der Schule. Den Schülerinnen und Schülern wurde die Frage gestellt, ob sie ausserhalb der Schule einen Unterricht zur Prüfungsvorbereitung besuchten. In der 6. Klasse besuchten 21 Prozent der Schülerinnen und Schüler einen Zusatzunterricht zur Vorbereitung auf die Prüfung ins Langgymnasium. Tabelle 6.3 zeigt die Verteilung der Schülerinnen und Schüler, die eine zusätzliche Prüfungsvorbereitung besuchten, nach den Zeugnisnoten in Deutsch und Mathematik sowie nach Geschlecht, Erstsprache und sozialer Herkunft.

Tabelle 6.3: Anteil Schülerinnen und Schüler mit Unterricht zur Prüfungsvorbereitung nach Zeugnisnoten und individuellen Lernvoraussetzungen

Zeugnisnote Deutsch	mit Prüfungsvorbereitung	ohne Prüfungsvorbereitung
6	60.0%	40.0%
5.5	50.1%	49.9%
5	35.2%	64.8%
4.5	9.3%	90.7%
Zeugnisnote Mathematik		
6	58.0%	42.0%
5.5	48.0%	52.0%
5	33.1%	66.9%
4.5	10.9%	89.1%
Geschlecht		
Mädchen	21.9%	78.1%
Knaben	20.1%	79.9%
Erstsprache		
Deutsch als Erstsprache	23.3%	76.7%
Deutsch als Zweitsprache	18.6%	81.4%
soziale Herkunft		
privilegiert	38.6%	61.4%
eher privilegiert	21.1%	78.9%
eher benachteiligt	12.9%	87.1%
benachteiligt	8.0%	92.0%

Je höher die Zeugnisnoten in Deutsch oder Mathematik sind, desto grösser ist der Anteil an Schülerinnen und Schülern, die ausserhalb der Schule einen Unterricht zur Prüfungsvorbereitung besuchten. Von den Schülerinnen und Schülern mit einer Deutschnote 5.5 besuchte beispielsweise rund die Hälfte einen Vorbereitungsunterricht, von jenen mit einer Deutschnote 5 sind es rund 35 Prozent, während nur noch 9 Prozent mit der Deutschnote 4.5 einen Vorbereitungsunterricht besuchten.

Knaben und Mädchen besuchten gleich häufig einen Vorbereitungsunterricht. Schülerinnen und Schüler mit Deutsch als Erstsprache besuchten etwas häufiger einen Vorbereitungsunterricht als solche mit Deutsch als Zweitsprache (23 Prozent zu 19 Prozent). Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto grösser ist der Anteil an Schülerinnen und Schülern, die Vorbereitungskurse fürs Langgymnasium besuchten. Von den Schülerinnen und Schülern mit privilegierter sozialer Herkunft besuchten 38.6 Prozent einen Vorbereitungsunterricht. Im Vergleich dazu sind es bei den Schülerinnen und Schülern mit benachteiligter sozialer Herkunft lediglich 8 Prozent.

Tabelle 6.4: Anteil Schülerinnen und Schüler mit Unterricht zur Prüfungsvorbereitung nach Zeugnisnoten und sozialer Herkunft

	soziale Herkunft			
	benachteiligt	eher benachteiligt	eher privilegiert	privilegiert
Zeugnisnote Deutsch				
5.5	44.3%	63.8%	44.8%	49.5%
5	23.4%	20.7%	34.0%	47.8%
Zeugnisnote Mathematik				
5.5	48.0%	44.9%	47.6%	49.1%
5	24.6%	25.5%	31.6%	42.0%

Tabelle 6.4 zeigt die Anteile an Schülerinnen und Schülern mit Vorbereitungsunterricht nach Zeugnisnoten und nach sozialer Herkunft. Bei gleichen Zeugnisnoten besuchten Schülerinnen und Schüler mit einer höheren sozialen Herkunft vergleichsweise öfters einen Unterricht zur Prüfungsvorbereitung. Bei einer Deutschnote 5 besuchten beispielsweise rund 21 Prozent der Schülerinnen und Schüler mit eher benachteiligter, rund 34 Prozent mit eher privilegierter und rund 48 Prozent mit privilegierter sozialer Herkunft einen Vorbereitungsunterricht. Für diesen Zusammenhang sind zwei Erklärungen denkbar. Zum einen steigen mit höherer sozialer Herkunft auch die Bildungsaspirationen der Schülerinnen und Schüler (und jene der Eltern). Zum anderen stehen den Schülerinnen und Schülern aus privilegierten sozialen Verhältnissen auch eher die finanziellen Ressourcen für den Besuch eines Unterrichts zur Prüfungsvorbereitung zur Verfügung.

Übertritt ins Langgymnasium

Für den Übertritt ins Langgymnasium zählen für die Schülerinnen und Schüler der öffentlichen Schule je zur Hälfte das Ergebnis in der Aufnahmeprüfung und die Erfahrungsnoten. Massgebend sind jeweils die Leistungen in Deutsch und Mathematik. Diese werden nicht nur von der Schule, sondern vor allem auch von Herkunftsmerkmalen bestimmt (primäre soziale Ungleichheiten). Beispielsweise hängen die Leistungen positiv mit der sozialen Herkunft der

Schülerinnen und Schüler zusammen. Als Folge davon werden Schülerinnen und Schüler mit einer tieferen sozialen Herkunft beim Übertritt in die Sekundarstufe I überproportional häufig in die Schultypen mit geringeren Ansprüchen eingeteilt. Die soziale Herkunft der Schülerinnen und Schüler ist dabei nicht direkt, sondern vermittelt über die Leistungen für den Übertritt von Bedeutung. Darüber hinaus können auch sekundäre soziale Ungleichheiten für den Prüfungserfolg von Bedeutung sein. Beispielsweise entscheiden sich Schülerinnen und Schüler aus sozial benachteiligten Verhältnissen bei gleichen Leistungen weniger häufig fürs Langgymnasium. Bei gleichen Testleistungen und bei gleichen Noten sind es drei Merkmale, die für den Prüfungserfolg von Bedeutung sind:

Soziale Herkunft: Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto grösser ist die Wahrscheinlichkeit, dass sie das Langgymnasium besuchen werden.

Alter: Je älter die Schülerinnen und Schüler sind, desto geringer ist die Wahrscheinlichkeit, dass sie das Langgymnasium besuchen werden.

Zusätzlicher Unterricht zur Vorbereitung der Aufnahmeprüfung: Schülerinnen und Schüler, die sich zusätzlich auf die Aufnahmeprüfung vorbereiten, erhöhen ihre Erfolgswahrscheinlichkeit bei der Aufnahmeprüfung.

In Abbildung 6.6 sind die Wahrscheinlichkeiten für den Übertritt in das Langgymnasium dargestellt. Die Wahrscheinlichkeiten sind in Abhängigkeit von den Zeugnisnoten in Deutsch und in Mathematik dargestellt. Die vier Kurven zeigen beispielhaft für Schülerinnen und Schüler mittleren Alters mit Deutsch als Erstsprache sowie ohne Besuch eines Unterrichts zur Prüfungsvorbereitung die Wahrscheinlichkeit für den Übertritt in das Langgymnasium nach der sozialen Herkunft (Schülerinnen und Schüler aus benachteiligten, eher benachteiligten, eher privilegierten und privilegierten Verhältnissen). Bei gleichen Schulnoten steigt die Wahrscheinlichkeit für den Übertritt in das Langgymnasium mit dem Anstieg der sozialen Herkunft.

Wie Abbildung 6.6 zeigt, ist die Bedeutung der sozialen Herkunft für die Übertrittschancen vorwiegend bei einer Deutsch- und Mathematiknote von 5.5 stark ausgeprägt. Bei Schülerinnen und Schülern mit einer 5 oder einer 6 zeigen sich nahezu keine Unterschiede zwischen Schülerinnen und Schülern unterschiedlicher sozialer Herkunft. Die zusätzliche Erklärungskraft der sozialen Herkunft ist in beiden Fällen gering. Offenbar entscheiden sich Schülerinnen und Schüler mit einer Durchschnittsnote 5 zu einem grossen Teil für die Sekundarschule. Bei einer Durchschnittsnote 6 hingegen entscheidet sich die grosse Mehrheit für das Langgymnasium. Kommt hinzu, dass der Prüfungserfolg für den Übertritt ins Langgymnasium zur Hälfte durch die Vornoten bestimmt wird.

Beispiel

Für ein Mädchen mittleren Alters mit Deutsch als Erstsprache, mit sehr guten Deutsch- und Mathematiknoten (5.5), ohne Besuch eines Unterrichts zur Prüfungsvorbereitung und mit privilegierter sozialer Herkunft beträgt die Wahrscheinlichkeit für den Übertritt in das Langgymnasium rund 71 Prozent, mit eher privilegierter sozialer Herkunft 53 Prozent, mit eher benachteiligter sozialer Herkunft 41 Prozent und mit benachteiligter sozialer Herkunft 29 Prozent.

Abbildung 6.6: Wahrscheinlichkeit für den Übertritt ins Langgymnasium nach sozialer Herkunft

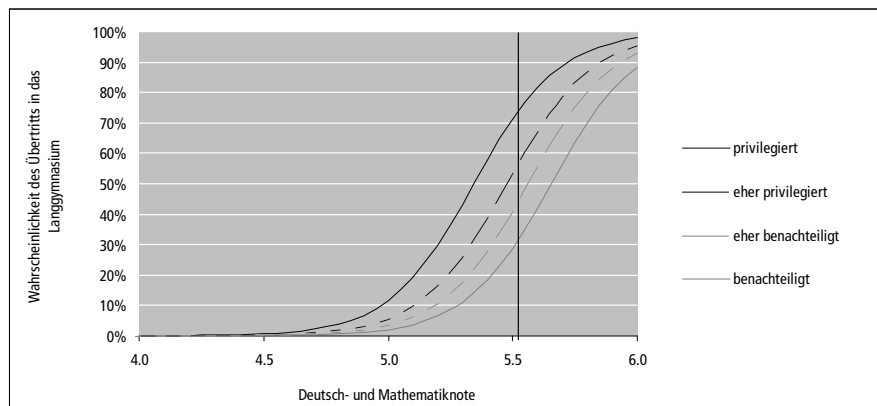
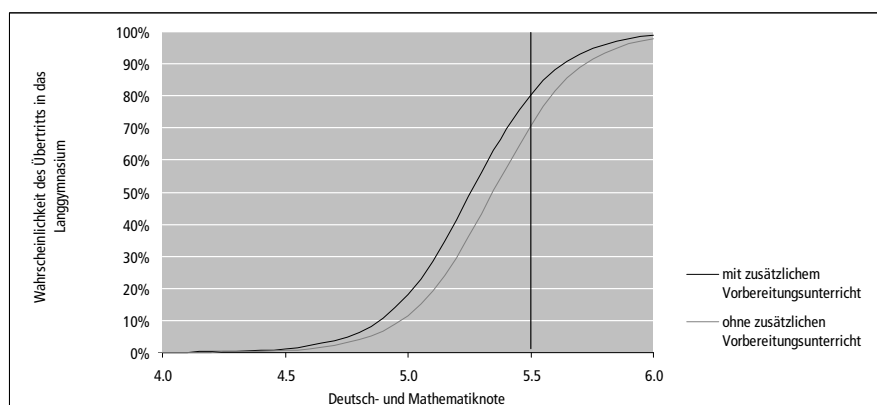


Abbildung 6.7 zeigt die Wahrscheinlichkeiten für den Übertritt in das Langgymnasium in Abhängigkeit der Deutsch- und Mathematiknoten für Schülerinnen und Schüler mit zusätzlichem und ohne zusätzlichen Unterricht zur Prüfungsvorbereitung. Wer ausserhalb der Schule einen Zusatzunterricht zur Prüfungsvorbereitung besucht, erhöht die Übertrittschancen.

Beispiel

Die Wahrscheinlichkeit für den Übertritt beträgt für ein Mädchen mittleren Alters mit Deutsch als Erstsprache, mit sehr guten Deutsch- und Mathematiknoten (5.5) und mit privilegierter sozialer Herkunft ohne Prüfungsvorbereitung rund 71 Prozent und mit Prüfungsvorbereitung rund 80 Prozent.

Abbildung 6.7: Wahrscheinlichkeit für den Übertritt ins Langgymnasium nach Vorbereitungsunterricht



6.5 Übertritt in die Abteilung A der Sekundarschule

Analog zu den Analysen über einen erfolgreichen Übertritt in das Langgymnasium lässt sich auch die Einteilung der Schülerinnen und Schüler in die Abteilung A gegenüber der Einteil-

lung in die Abteilungen B oder C der Sekundarschule überprüfen. Die Schülerinnen und Schüler werden am Ende der Primarschule aufgrund einer Gesamtbeurteilung einer der Abteilungen der Sekundarschule zugeteilt. Die fachlichen Leistungen spielen vor allem für die fachweise Zuteilung zu einer Anforderungsstufe eine Rolle, sind allerdings auch Bestandteil einer Gesamtbeurteilung.

Bei gleichen Testleistungen und bei gleichen Noten sind es vier Merkmale, die für die Zuteilung in die Abteilung A von Bedeutung sind:

Soziale Herkunft: Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto grösser ist die Wahrscheinlichkeit, dass sie der Abteilung A zugeteilt werden.

Alter: Je älter die Schülerinnen und Schüler sind, desto geringer ist die Wahrscheinlichkeit, dass sie der Abteilung A zugeteilt werden.

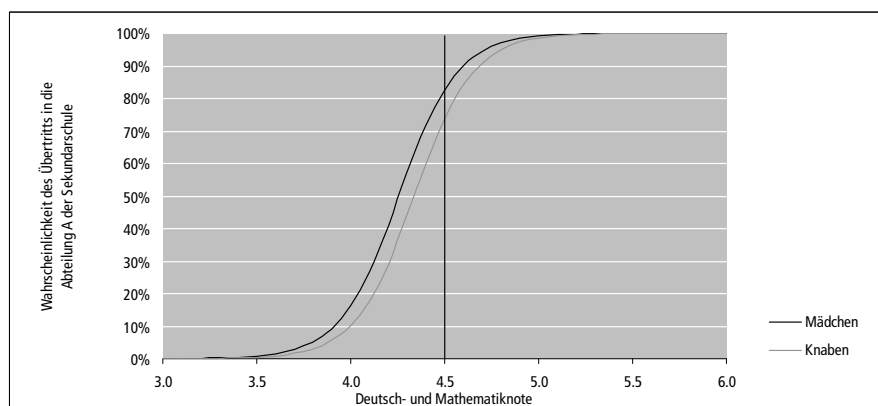
Erstsprache: Schülerinnen und Schüler mit Deutsch als Zweitsprache haben eine grössere Wahrscheinlichkeit, dass sie der Abteilung A zugeteilt werden.

Geschlecht: Mädchen haben eine grössere Wahrscheinlichkeit, dass sie der Abteilung A zugeteilt werden.

Die Analysen zeigen zudem, dass die Deutschnote für den Übertritt in die Abteilung A von grösserer Bedeutung ist als die Mathematiknote.

Abbildung 6.8 zeigt die Wahrscheinlichkeiten für den Übertritt in die Abteilung A in Abhängigkeit der Zeugnisnoten in Deutsch und in Mathematik für Mädchen und Knaben.

Abbildung 6.8: Wahrscheinlichkeit für den Übertritt in die Abteilung A der Sekundarschule nach Geschlecht



Beispiel

Für ein Mädchen mittleren Alters mit Deutsch als Erstsprache, mit durchschnittlichen Deutsch- und Mathematiknoten (4.5) und mit privilegierter sozialer Herkunft beträgt die Wahrscheinlichkeit für die Zuteilung in die Abteilung A rund 83 Prozent, für einen Knaben mit gleichen Merkmalen hingegen nur 74 Prozent.

6.6 Fazit

Die vorliegenden Ergebnisse zeigen, dass es beim Übertritt von der Primarstufe in die Abteilungen der Sekundarstufe I zu beträchtlichen sekundären sozialen Ungleichheiten kommt. Mit sekundären sozialen Ungleichheiten sind Einflüsse der sozialen Herkunft auf den Übertritt gemeint, die unabhängig von den erreichten schulischen Leistungen auftreten. Bei gleichen Schulleistungen treten Schülerinnen und Schüler aus privilegierten sozialen Verhältnissen (z.B. aus bildungsnahen Familien) öfters ins Langgymnasium über als Schülerinnen und Schüler aus weniger privilegierten sozialen Verhältnissen (z.B. aus bildungsfernen Familien). Auch der Übertritt in die verschiedenen anspruchsvollen Abteilungen der Sekundarschule hängt mit der sozialen Herkunft der Schülerinnen und Schüler zusammen, wenn auch in einem geringeren Ausmass als der Übertritt ins Langgymnasium. Bei gleichen Schulleistungen treten Schülerinnen und Schüler aus privilegierten sozialen Verhältnissen öfter in die Abteilung A der Sekundarschule ein als Schülerinnen und Schüler aus weniger privilegierten sozialen Verhältnissen.

Nach Berücksichtigung der sozialen Herkunft und der Schulleistungen wird der Übertritt ins Langgymnasium nicht zusätzlich durch die Erstsprache der Schülerinnen und Schüler beeinflusst. Schülerinnen und Schüler mit Deutsch als Zweitsprache weisen die gleiche Chance auf, ins Langgymnasium überzutreten, wie Schülerinnen und Schüler mit Deutsch als Erstsprache. Wird nur der Übertritt in die Abteilungen der Sekundarschule betrachtet, weisen Schülerinnen und Schüler mit Deutsch als Zweitsprache sogar eine höhere Chance auf, in die Abteilung A überzutreten, als Schülerinnen und Schüler mit Deutsch als Erstsprache. Eine Erklärung für dieses Ergebnis wäre, dass bei Selektionsentscheidungen die Fremdsprachigkeit und das Mehrwissen in der heimatlichen Sprache und Kultur berücksichtigt werden, wie das beispielsweise auch von der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren zur Schulung der fremdsprachigen Kinder empfohlen wird.

Das Geschlecht ist für den Übertritt ins Langgymnasium nicht von Bedeutung. Mädchen weisen jedoch gegenüber Knaben eine grössere Chance auf, in die Abteilung A statt in die Abteilungen B oder C der Sekundarschule überzutreten.

Von Bedeutung für den Übertritt ins Langgymnasium ist hingegen, wie sich die Schülerinnen und Schüler für die Aufnahmeprüfung vorbereiten. Durch den Besuch eines zusätzlichen Unterrichts zur Prüfungsvorbereitung lassen sich die Chancen für den Übertritt ins Langgymnasium verbessern. Eltern, die sich eine zusätzliche Prüfungsvorbereitung ausserhalb der Schule leisten können, verhelfen ihren Kindern zu Vorteilen.

Die Ergebnisse machen deutlich, wie kritisch die Schnittstelle zwischen Primar- und Sekundarstufe für die Entstehung und Verstärkung sozialer Ungleichheiten ist. Mit den vorliegenden Daten sind keine Aussagen zu den dahinterliegenden Wirkungsmechanismen möglich. Die Forschung zeigt aber, dass sekundäre soziale Ungleichheiten unter anderem ein Ergebnis individueller Bildungsentscheidungen sind und auf herkunftsspezifische Kosten- und Nutzenbewertungen von Bildung zurückgeführt werden können (vgl. etwa Ditton, 2007; Maaz, Baumert und Trautwein, 2009). Studien weisen zudem darauf hin, dass Übertritte auch durch herkunftsspezifische Schullaufbahneempfehlungen von Lehrpersonen beeinflusst werden können (vgl. etwa Maaz und Nagy, 2009).

7 Motivation und lernbezogene Emotionen nach sechs Jahren Primarstufe in Mathematik und Deutsch

Alex Buff

7.1 Einleitung

Theoretischer Hintergrund

Angebots-Nutzungs-Modelle (Fend, 1998; Helmke, 2003; Pauli & Reusser, 2006) machen deutlich, dass der «Ertrag» von Schule, wie z.B. akademische Leistungen, nicht allein von der Qualität des schulischen Angebots abhängig ist, sondern in hohem Masse auch davon, wie Schülerinnen und Schüler dieses Angebot nutzen können und wollen. Für die «*Nutzungsqualität*» sind aufseiten der Schülerinnen und Schüler neben Intelligenz und Vorwissen insbesondere «*motivational-affektive*» Merkmale (Helmke & Weinert, 1997a) zentral, so namentlich fachspezifische lern-leistungsbezogene *Kompetenz- und Valenzüberzeugungen*. Kompetenzbezogene Überzeugungen manifestieren sich, wie etwa Pintrich und Schunk (2002) schreiben, in den subjektiven Antworten auf die Frage «*Kann ich es?*», die valenzbezogenen in denjenigen auf die Frage «*Will ich es und warum?*». Unter kompetenzbezogene Überzeugungen fallen Merkmale wie Fähigkeitsselbstkonzept, Selbstwirksamkeitserwartungen usw. (zur Differenzierung verschiedener kompetenzbezogener Konstrukte vgl. Bong & Skolvik, 2003; Buff, Reusser & Pauli, 2010a; Pietsch, Walker und Chapman, 2003; Skinner, 1996). Valenzüberzeugungen umfassen Konstrukte wie etwa die subjektive Wahrnehmung der Wichtigkeit bzw. des Nutzens eines Gegenstands (bspw. das Fach Mathematik, Lernen von Mathematik), dessen positiver oder negativer Anreiz oder auch das Interesse am Gegenstand (zur Differenzierung verschiedener Facetten der Valenzüberzeugungen vgl. Eccles, 2005). Kompetenz- und valenzbezogene Überzeugungen sind als wichtige Bedingung und Prozessvoraussetzung wie auch als Ergebnis von schulischem Lern- und Leistungshandeln zu sehen. Zudem sind sie als zentrale Ingredienzien einer lern-leistungsbezogenen *motivationalen Orientierung* zu betrachten, die selbstinitiiertem, eigenständigem Lernen (nicht nur während der Schulzeit) zugrunde liegt. Der Erhalt bzw. die Förderung positiver kompetenz- und valenzbezogener Überzeugungen wird damit, insbesondere mit Blick auf das immer wichtiger werdende lebenslange Lernen, zu einem zentralen, eigenständigen Erziehungs- und Bildungsziel der Schule (Bandura, 2003; Krapp, 2003; Spinath & Spinath, 2005).

Pekrun (2000, 2006) macht in der «Control-Value Theory of Achievement Emotions» (CVTAE) aufseiten der motivational-affektiven Merkmale auf einen weiteren Variablenkomplex aufmerksam, welcher in der pädagogisch-psychologischen Forschung im schulischen Lern-Leistungskontext (mit Ausnahme der Prüfungsangstforschung) bisher eher stiefmütterlich behandelt wurde (vgl. Boekaerts, 2001; Krapp, 2005; Maehr, 2004; Pekrun, 1998, 2000; Pintrich, 2003; Schutz & Pekrun, 2007): *lern-leistungsbezogene Emotionen*. Hierbei handelt es sich um Emotionen, welche direkt mit Lern-Leistungshandeln oder dessen Ergebnissen verbunden sind. In einer Reihe explorativer Studien (Pekrun, 1998; Pekrun & Hofmann, 1999; Pekrun, Goetz, Titz & Perry, 2002), zeigt sich u.a., dass Emotionen im schulischen Lern-Leistungskontext äusserst vielfältig sind und sich keineswegs auf Angst be-

schränken (Pekrun, 1998; Pekrun & Hofmann, 1999). Unterschieden wird in der CVTAE zwischen verschiedenen distinkten Emotionen – positiv-aktivierend (z.B. Freude), negativ-aktivierend (z.B. Angst), positiv-deaktivierend (z.B. Erleichterung), negativ-deaktivierend (z.B. Langeweile) –, welche auf ein Ereignis/Ergebnis oder auf eine Aktivität fokussieren (vgl. etwa die Taxonomien bei Pekrun, 1998, 2000; Pekrun, Frenzel, Goetz & Perry, 2007). Lernleistungsbezogenen Emotionen wird zugeschrieben, dass sie Lernprozesse nachhaltig beeinflussen (Pekrun & Schutz, 2007; Schutz, Hong, Cross & Osbon, 2006). Von Interesse sind nachfolgend die auf die Aktivität des Lernens bezogenen aktivierenden Emotionen *Freude* und *Angst*.

Der vorliegende Beitrag fokussiert diese nutzungsbezogenen individuellen motivational-affektiven Lernvoraussetzungen und fragt nach deren Qualität gegen Ende der Primarstufe sowie deren Entwicklung in der zweiten Hälfte der Primarschulzeit. Im Zentrum des Interesses stehen dabei die Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler bezogen auf die beiden Fächer Mathematik und Deutsch.

Bemerkungen zur Ergebnispräsentation

In den Analysen berücksichtigt wurden Daten der Lernstandserhebungen 2006/2007 (Ende der 3. Klasse) sowie 2009, da die Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler lediglich zu diesen beiden Messzeitpunkten erhoben wurden.

Um einen ersten Eindruck davon zu gewinnen, wie es um die Motivation und die lernbezogenen Emotionen der Schülerinnen und Schüler nach sechs Jahren Primarstufe steht und wie sich die Motivation seit den letzten Erhebungen in den Jahren 2006 resp. 2007 bis 2009 entwickelt hat, werden die interessierenden Merkmale zunächst jeweils in einem deskriptiven Sinne dargestellt. Für die deskriptive Ergebnisdarstellung wurde aus Gründen der Anschaulichkeit, wie schon anlässlich der Lernstandserhebung am Ende der 3. Klasse 2006/2007, das Antwortspektrum (1–4) in fünf gleich grosse Abschnitte unterteilt (tief, eher tief, ambivalent, eher hoch, hoch). Dargestellt werden die prozentualen Anteile der Schülerinnen und Schüler in den jeweiligen Abschnitten.

Hinsichtlich der Kompetenz- und Valenzüberzeugung folgen dann detailliertere Präsentationen für die einzelnen Entwicklungsverläufe. Da die beiden lernbezogenen Emotionen Freude und Angst lediglich in der Lernstandserhebung im Jahr 2009 erfasst wurden, können keine Entwicklungen aufgezeigt werden, bei den diesbezüglich dargestellten Ergebnissen handelt es sich um «Momentaufnahmen» zu diesem Messzeitpunkt. Alle Analysen zu Entwicklungsverläufen bzw. im Falle der lernbezogenen Emotionen zu Gruppenunterschieden basieren auf den kontinuierlichen Skalenwerten 1 bis 4.

Die Ergebnispräsentation für die einzelnen Merkmale erfolgt getrennt nach zwei Gruppen von Schülerinnen und Schülern: diejenigen, bei denen die letzte Erhebung drei Jahre zurückliegt (Stichprobe 06/09), und diejenigen, bei denen die letzte Erhebung zwei Jahre zurückliegt (Stichprobe 07/09). Im Falle Letzterer handelt es sich um Kinder, die das Ende der 3. Klasse erst nach vier Schuljahren erreicht hatten (Sonderklasse A, Repetition usw.). In der Stichprobe 06/09 wird jeweils zusätzlich geprüft, ob sich bei Schülerinnen und Schüler, welche primär mit Blick auf die Sekundarstufe I unterschiedliche schulische Laufbahnen ein-

schlagen (Langgymnasium, Abteilung A oder B/C der Sekundarschule, Primarstufe), differente Entwicklungen bezogen auf die interessierenden Merkmale ergeben. In diese Analysen werden also lediglich diejenigen Schülerinnen und Schüler einbezogen, welche 2006 am Ende 3. Klasse standen, da bei standardmässigem weiterem Verlauf nur sie nach der Erhebung 2009 in die Sekundarstufe I hätten übertreten können.

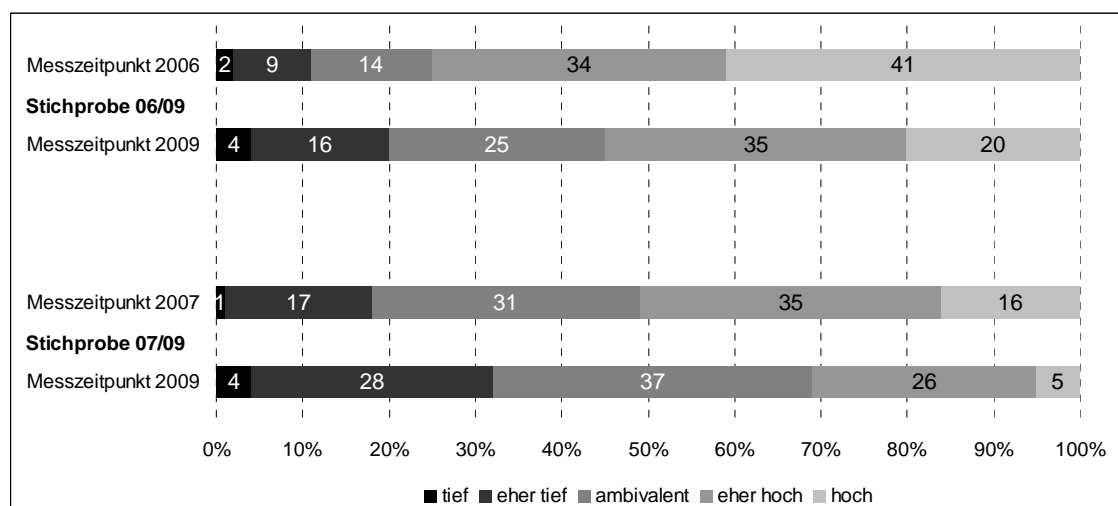
Nähere methodische Angaben (Stichprobe, Instrumente usw.) sowie eine detailliertere Darstellung der einzelnen Analyseergebnisse finden sich in Kapitel 7, Teil 2.

7.2 Kompetenzüberzeugung in Mathematik und Deutsch

Deskription

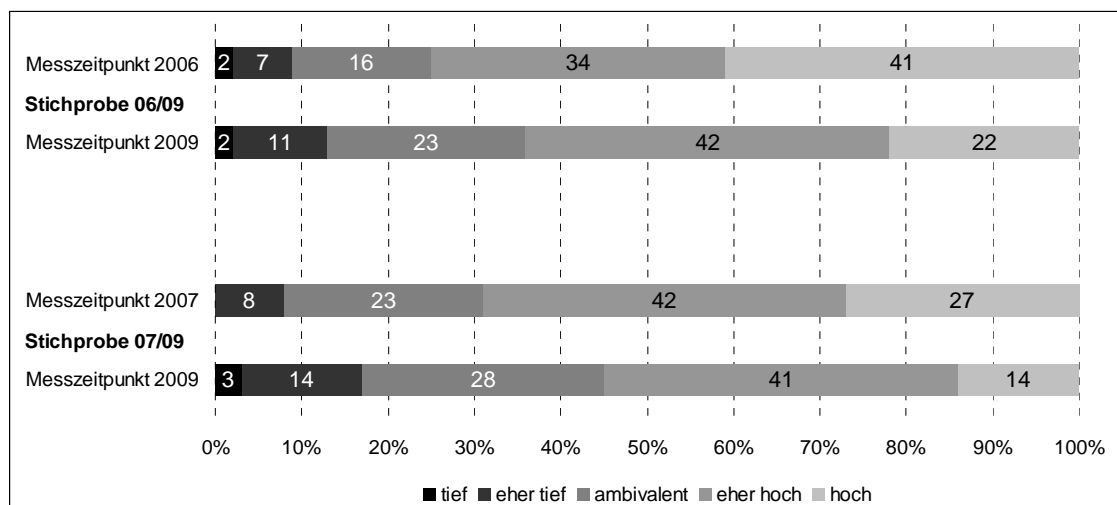
Abbildung 7.1 zeigt die Ergebnisse für *Mathematik*. Für beide Stichproben (06/09 bzw. 07/09) zeigt sich ein ähnliches Bild. Nimmt man die beiden positiven Abschnitte (hoch, eher hoch) zusammen, so sinkt der prozentuale Anteil vom früheren zum späteren Messzeitpunkt in beiden Gruppen. Während bspw. in der Stichprobe 07/09 im Jahr 2007 noch 51% der Schülerinnen und Schüler über eine (eher) hohe Kompetenzüberzeugung verfügen, sind es 2009 noch 31%. Genau umgekehrt verhält es sich am anderen Ende des Spektrums (tief, eher tief). Hier ist vom früheren zum späteren Messzeitpunkt jeweils eine Zunahme der prozentualen Besetzung der beiden Abschnitte zu verzeichnen; wiederum in der Stichprobe 07/09 bspw. von 18% auf 32%.

Abbildung 7.1: Kompetenzüberzeugung in Mathematik



Für *Deutsch* (Abbildung 7.2) zeigt sich in beiden Stichproben ein zur Mathematik paralleles Ergebnismuster: eine Abnahme der prozentualen Besetzung am positiven Ende des Spektrums (hoch, eher hoch) und eine Zunahme am anderen Ende (tief, eher tief). Wiederum in der Stichprobe 07/09 bspw. sinkt der Anteil derjenigen Schülerinnen und Schüler mit hoher bzw. eher hoher Kompetenzüberzeugung von 69% (2007) auf 55% (2009), während derjenige von Schülerinnen und Schüler mit tiefer bzw. eher tiefer Kompetenzüberzeugung von 8% auf 17% ansteigt.

Abbildung 7.2: Kompetenzüberzeugung in Deutsch



Insgesamt entsteht der Eindruck, dass sich die Kompetenzüberzeugung in Mathematik wie auch in Deutsch im Verlaufe der Zeit verschlechtert, d.h., dass sich die Kinder 2009 insgesamt weniger zutrauen als noch zwei bzw. drei Jahre früher.

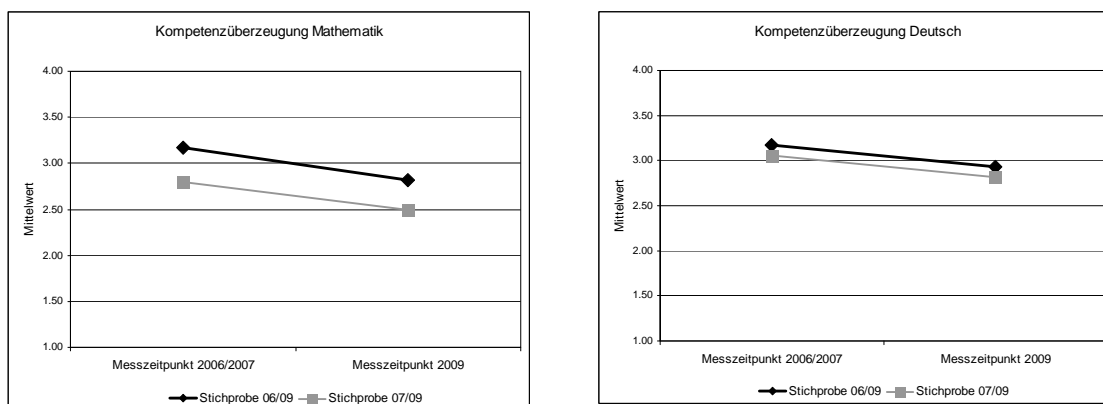
Entwicklungsverläufe

Wie Analysen zeigen, trägt der aufgrund der deskriptiven Darstellungen gewonnene Eindruck eines allgemeinen Negativtrends der Kompetenzüberzeugung nicht. Alle nachfolgend erwähnten Trends über die Zeit hinweg bzw. Differenzen zwischen Gruppen sind – sofern nicht speziell vermerkt – statistisch bedeutsam. Abbildung 7.3 zeigt die Ergebnisse für die beiden Stichproben 06/09 bzw. 07/09 in Mathematik und Deutsch.

In *Mathematik* wird deutlich, dass die Schülerinnen und Schüler der Stichprobe 07/09 zu ihrem ersten Messzeitpunkt 2007 ($M = 2.80$)¹ über eine geringere Kompetenzüberzeugung verfügen als diejenigen der Stichprobe 06/09 im Jahr 2006 ($M = 3.17$). An dieser Differenz ändert sich im Verlaufe der Zeit nichts, da sich bis zum zweiten Messzeitpunkt die Kompetenzüberzeugung in beiden Gruppen gleichermaßen verringert. 2009 liegt die Kompetenzüberzeugung in der Stichprobe 06/09 bei $M = 2.82$, in der Stichprobe 07/09 bei $M = 2.50$.

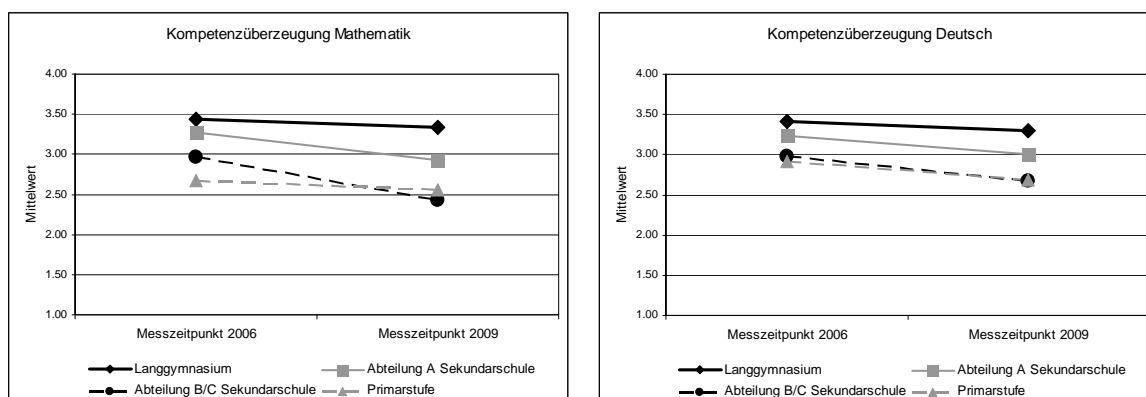
Ein ähnliches Bild zeigt sich in *Deutsch*. Hier unterscheiden sich die beiden Stichproben in der Höhe ihrer Kompetenzüberzeugung zum jeweils ersten Messzeitpunkt nicht: Stichprobe 06/09 bei $M = 3.17$ vs. Stichprobe 07/09 bei $M = 3.06$. Im weiteren Verlauf verschlechtert sich das Vertrauen in die eigene Kompetenz jedoch auch hier in beiden Gruppen gleichermaßen. 2009 ist die Kompetenzüberzeugung in der Stichprobe 06/09 bei $M = 2.93$, in der Stichprobe 07/09 bei $M = 2.81$.

Abbildung 7.3: Entwicklung der Kompetenzüberzeugung in Mathematik und Deutsch getrennt nach Stichproben



Die Ergebnisse hinsichtlich der Frage allfällig divergenter Entwicklungen von Schülerinnen und Schüler mit unterschiedlichen schulischen Laufbahnen nach der Lernstandserhebung 2009 (Langgymnasium, Abteilung A oder B/C der Sekundarschule, Primarstufe) sind in Abbildung 7.4 dargestellt. In diese Analysen einbezogen wurde, wie bereits erwähnt, lediglich die Stichprobe 06/09.

Abbildung 7.4: Entwicklung der Kompetenzüberzeugung in Mathematik und Deutsch getrennt nach schulischen Laufbahnen



In *Mathematik* verfügen alle vier Gruppen 2006 über unterschiedliche Kompetenzüberzeugungen: Am höchsten ist diese bei denjenigen Schülerinnen und Schülern, die nach der Lernstandserhebung 2009 ins Langgymnasium wechseln, am niedrigsten bei denjenigen, die nach 2009 in der Primarstufe verbleiben (Langgymnasium $M = 3.44$, Abteilung A $M = 3.27$, Abteilung B/C $M = 2.97$, Primarstufe $M = 2.67$). Im weiteren Verlauf verschlechtert sich das Vertrauen in die eigenen Fähigkeiten zwar in allen Gruppen, jedoch in unterschiedlichem Ausmasse. Im Falle der Schülerinnen und Schüler, die ins Langgymnasium übertreten bzw. in der Primarstufe verbleiben, zeigt sich der schwächste Negativtrend. 2009 liegt der Mittelwert in der Gruppe Langgymnasium bei $M = 3.33$, im Falle der Gruppe Primarstufe bei $M = 2.56$. Verglichen mit dem Langgymnasium fällt der Trend bei denjenigen Schülerinnen und

Schüler deutlich stärker aus, welche in die Abteilung A der Sekundarschule wechseln, 2009 weist diese Gruppe noch einen Mittelwert von $M = 2.93$ auf. Der stärkste Rückgang der Kompetenzüberzeugung findet sich in der Gruppe, welche in die Abteilung B/C der Sekundarschule wechselt. Hier liegt die Kompetenzüberzeugung 2009 bei $M = 2.43$. Diese Entwicklung führt dazu, dass diese Gruppe 2009 über eine negativere Kompetenzüberzeugung verfügt als die Gruppe Primarstufe. Alle anderen Unterschiede zwischen den Gruppen sind 2009 ebenfalls statistisch bedeutsam.

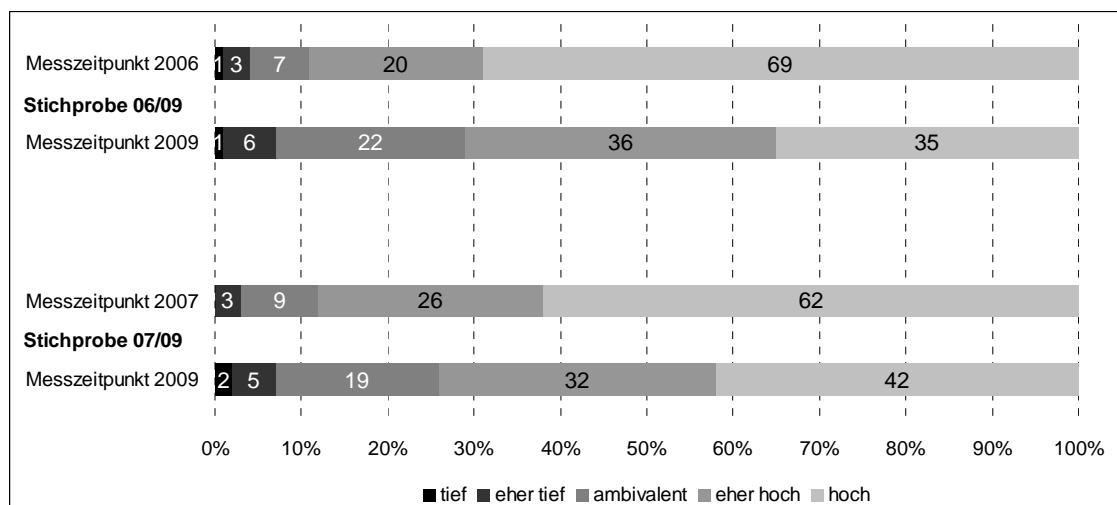
In *Deutsch* zeigt sich ein über weite Strecken analoges Bild. Mit Ausnahme derjenigen Schülerinnen und Schüler in den Gruppen Abteilung B/C der Sekundarschule sowie Primarstufe unterscheiden sich alle anderen in der Höhe ihrer Kompetenzüberzeugung 2006 voneinander (Langgymnasium $M = 3.41$, Abteilung A $M = 3.24$, Abteilung B/C $M = 2.98$, Primarstufe $M = 2.91$). Auch hier verschlechtert sich die Kompetenzüberzeugung insgesamt, jedoch wiederum in unterschiedlichem Ausmasse. Die Gruppe Langgymnasium weist 2009 noch eine Kompetenzüberzeugung von $M = 3.30$ auf, die Gruppe Primarstufe eine von $M = 2.67$. Der Negativtrend ist in der Gruppe Primarstufe etwas grösser als in der Gruppe Langgymnasium, jedoch nicht statistisch bedeutsam davon verschieden. Im Falle Abteilung A der Sekundarschule jedoch erweist sich der Abwärtstrend stärker als im Falle des Langgymnasiums. Hier findet sich 2009 ein Mittelwert von $M = 3.00$. Gleiches gilt für die Gruppe Abteilung B/C der Sekundarschule. Hier ist der Abwärtstrend am deutlichsten. Die mittlere Kompetenzüberzeugung beträgt 2009 noch $M = 2.67$. Mit Ausnahme der Gruppen Primarstufe sowie Abteilung B/C der Sekundarschule unterscheiden sich alle Mittelwerte 2009 voneinander.

7.3 Valenzüberzeugung in Mathematik und Deutsch

Deskription

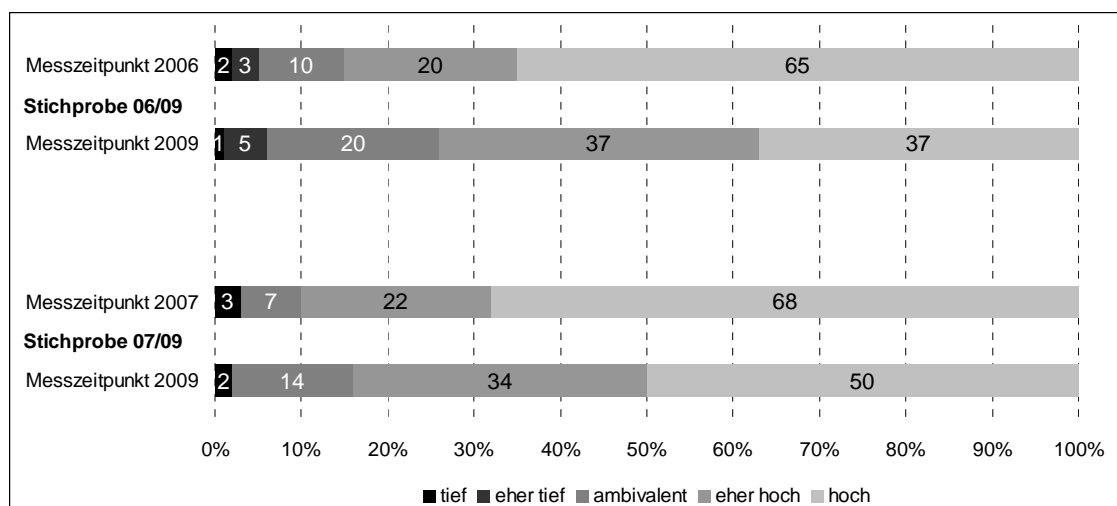
Abbildung 7.5 zeigt die Ergebnisse für *Mathematik*. In den beiden Stichproben 06/09 bzw. 07/09 ergibt sich ein ähnliches Bild. Die prozentuale Besetzung der beiden positiven Abschnitte (hoch, eher hoch) zusammen nimmt vom früheren zum späteren Messzeitpunkt in beiden Gruppen ab. In der Stichprobe 06/09 bspw. schreiben 2006 noch 89% der Schülerinnen und Schüler Mathematik (eher) hohe Bedeutung zu, erachten das Fach als (eher) nützlich und es hat für sie einen (eher) positiven Anreiz. 2009 sind dies noch 71%. Am anderen Ende des Spektrums erscheint die Veränderung nicht so ausgeprägt. Sind es bspw. in der Stichprobe 06/09 im Jahr 2006 4%, steigt dieser Anteil bis 2009 auf 7%. Auffällig ist in beiden Stichproben die Zunahme des Anteils der Schülerinnen und Schülern, die der Mathematik gegenüber ambivalent eingestellt sind; in der Stichprobe 06/09 bspw. verdreifacht sich diese Gruppe von 2006 bis 2009.

Abbildung 7.5: Valenzüberzeugung in Mathematik



In *Deutsch* (Abbildung 7.6) zeigt sich in beiden Stichproben ein ziemlich analoges Ergebnis-
muster zu dem in Mathematik: eine Abnahme der prozentualen Besetzung am positiven
Ende des Spektrums der Valenzüberzeugung (hoch, eher hoch) und eine deutliche Zunahme
des Anteils der dem Fach Deutsch gegenüber ambivalent eingestellten Schülerinnen und
Schüler. In der Stichprobe 06/09 bspw. sinkt der prozentuale Anteil der Schülerinnen und
Schüler mit (eher) hoher Valenzüberzeugung von 85% (2006) auf 74% (2009), während der-
jenige der Ambivalenten von 10% (2006) auf 20% (2009) ansteigt. Verglichen mit den Wer-
ten von Mathematik verändert sich jedoch im Fach Deutsch am negativen Ende des Spekt-
rums kaum etwas.

Abbildung 7.6: Valenzüberzeugung in Deutsch

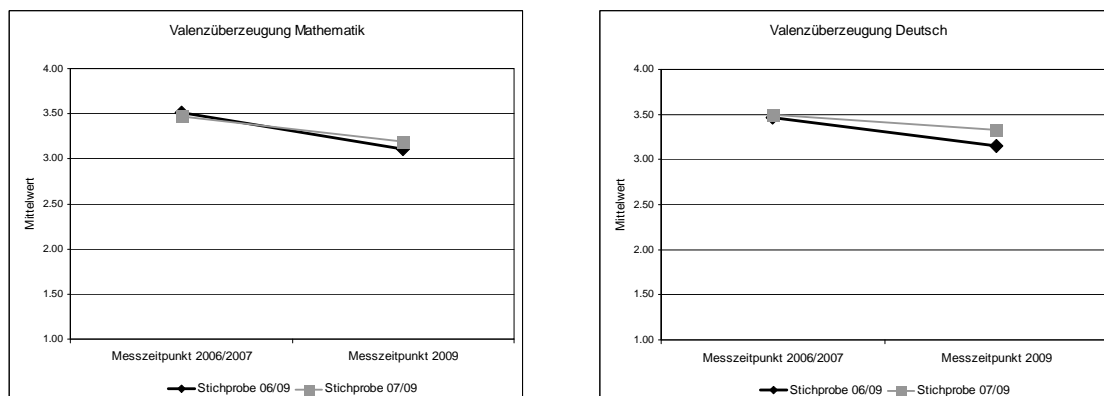


Alles in allem zeichnet sich auch hinsichtlich der Valenzüberzeugung in Mathematik und
Deutsch eine Abwärtstendenz ab, die sich auch, wie die nachfolgend dargestellten Analysen
zeigen, statistisch absichern lässt.

Entwicklungsverläufe

Alle erwähnten Trends über die Zeit hinweg bzw. Differenzen zwischen Gruppen sind – sofern nicht speziell vermerkt – auch hier statistisch bedeutsam. Abbildung 7.7 zeigt die Ergebnisse für die beiden Stichproben 06/09 bzw. 07/09 in Mathematik und Deutsch.

Abbildung 7.7: Entwicklung der Valenzüberzeugung in Mathematik und Deutsch getrennt nach Stichproben

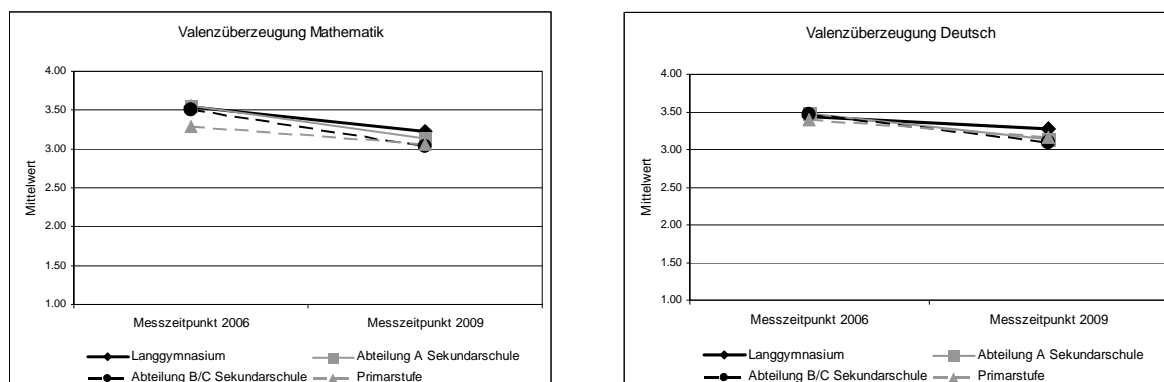


Für *Mathematik* wird deutlich, dass sich die beiden Stichproben zu ihrem jeweils ersten Messzeitpunkt 2006 resp. 2007 hinsichtlich ihrer Valenzüberzeugung nicht voneinander unterscheiden. In beiden Fällen ist diese recht hoch (Stichprobe 06/09 bei $M = 3.51$, Stichprobe 07/09 bei $M = 3.47$). Die weitere Entwicklung verläuft in beiden Gruppen mehr oder weniger parallel, wobei der Abwärtstrend in der Stichprobe 07/09 etwas schwächer ausfällt als in der Stichprobe 06/07. Statistisch bedeutsam ist dies jedoch nicht. 2009 liegt die Valenzüberzeugung in beiden Gruppen deutlich tiefer (Stichprobe 06/09 bei $M = 3.11$, Stichprobe 07/09 bei $M = 3.19$), sie unterscheiden sich jedoch nach wie vor nicht voneinander.

Auch in *Deutsch* unterscheiden sich die beiden Stichproben hinsichtlich ihrer Valenzüberzeugung zum jeweils ersten Messzeitpunkt 2006 resp. 2007 nicht voneinander (Stichprobe 06/09 bei $M = 3.46$ vs. Stichprobe 07/09 bei $M = 3.50$). Im weiteren Verlauf verringert sich jedoch die in beiden Gruppen anfänglich hohe Valenzüberzeugung. Dieser Abwärtstrend fällt in der Stichprobe 06/09 jedoch stärker aus. 2009 liegt die Valenzüberzeugung in der Stichprobe 06/09 bei $M = 3.15$, in der Stichprobe 07/09 bei $M = 3.33$. In der Stichprobe 07/09 wird Deutsch ein höherer Stellenwert zugeschrieben als in der Stichprobe 06/09.

Die Ergebnisse hinsichtlich divergenter Entwicklungen von Schülerinnen und Schüler mit unterschiedlichen schulischen Laufbahnen nach der Lernstandserhebung 2009 (Langgymnasium, Abteilung A oder B/C der Sekundarschule, Primarstufe) sind in Abbildung 7.8 dargestellt. In die Analysen einbezogen wurde auch hier lediglich die Stichprobe 06/07.

Abbildung 7.8: Entwicklung der Valenzüberzeugung in Mathematik und Deutsch getrennt nach schulischen Laufbahnen



2006 unterscheiden sich die drei Gruppen Langgymnasium ($M = 3.54$), Abteilung A ($M = 3.55$) resp. Abteilung B/C der Sekundarschule ($M = 3.50$) hinsichtlich ihrer Valenzüberzeugung in *Mathematik* nicht voneinander. In der Gruppe Primarstufe findet sich zu diesem Zeitpunkt jedoch verglichen mit allen anderen eine niedrigere Valenzüberzeugung ($M = 3.29$). Bis 2009 sinkt die Valenzüberzeugung in allen Gruppen, wobei dieser Trend in der Gruppe Langgymnasium und Primarstufe am wenigsten ausgeprägt ist. Deutlicher ist dies bei den Schülerinnen und Schüler der beiden anderen Gruppen. 2009 liegt der Mittelwert in der Gruppe Langgymnasium bei $M = 3.23$, im Falle der Gruppe Abteilung A bei $M = 3.13$, in der Gruppe Abteilung B/C bei $M = 3.04$ sowie in der Gruppe Primarstufe bei $M = 3.06$. Zu diesem Zeitpunkt ist die Valenzüberzeugung bei der Gruppe Langgymnasium höher als diejenige aller anderen Gruppen. Die Gruppe Primarstufe unterscheidet sich nun nicht mehr von den Gruppen Abteilung A resp. B/C. Hinsichtlich letzterer beiden findet sich jedoch ein Unterschied zugunsten der Abteilung A.

In *Deutsch* ist 2006 die Valenzüberzeugung in allen Gruppen mehr oder weniger identisch (Langgymnasium $M = 3.44$, Abteilung A $M = 3.48$, Abteilung B/C $M = 3.46$, Primarstufe $M = 3.40$). Auch die Valenzüberzeugung in Deutsch sinkt bis 2009, allerdings wiederum in den einzelnen Gruppen nicht gleich stark. Am schwächsten ist der Trend in den Gruppen Langgymnasium sowie Primarstufe. Diese beiden Gruppen weisen 2009 hinsichtlich ihrer Valenzüberzeugung noch Werte von $M = 3.27$ (Langgymnasium) sowie $M = 3.16$ (Primarstufe) auf. Deutlicher erweist sich der Rückgang in der Gruppe Abteilung A und noch etwas stärker in der Gruppe Abteilung B/C der Sekundarschule. Die Mittelwerte betragen hier 2009 $M = 3.14$ (Abteilung A) resp. $M = 3.09$ (Abteilung B/C). Die Valenzüberzeugung Deutsch ist 2009 in der Gruppe Langgymnasium höher als in den Gruppen Abteilung A sowie B/C der Sekundarschule, unterscheidet sich jedoch nicht von derjenigen der Primarstufe. Zwischen den anderen drei Gruppen lassen sich keine Unterschiede feststellen.

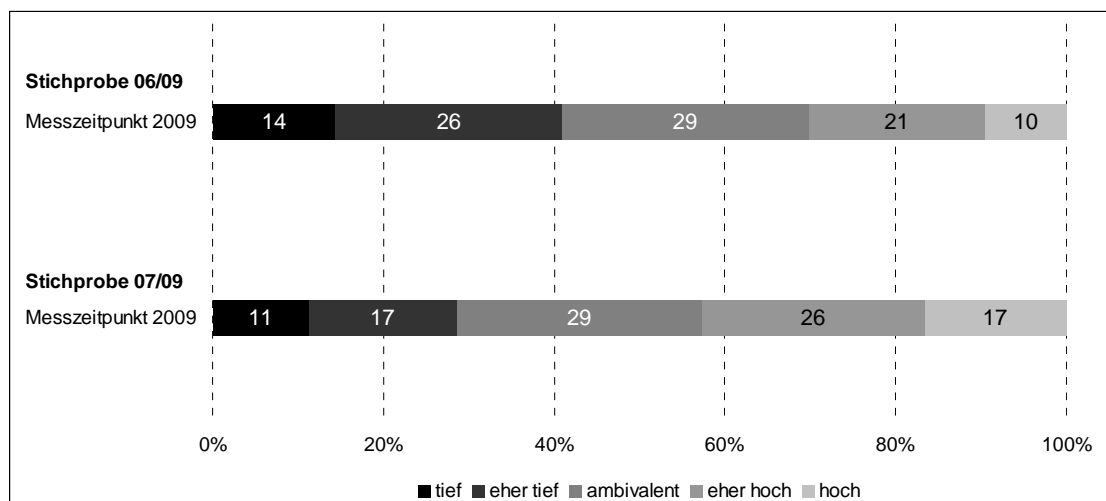
7.4 Lernbezogene Emotionen in Mathematik und Deutsch

Lernbezogene Emotionen wurden, wie bereits erwähnt, lediglich in der Erhebung 2009 erfasst, womit es sich hier um Momentaufnahmen nach sechs Jahren Primarschule handelt. Die nachfolgend erwähnten Differenzen zwischen Gruppen sind auch hier – sofern nicht speziell vermerkt – statistisch bedeutsam.

Freude beim Mathematik- und Deutschlernen

Abbildung 7.9 zeigt das Ergebnis für *Mathematik* in den beiden Stichproben 06/09 resp. 07/09. Es deutet sich an, dass die Schülerinnen und Schüler der Stichprobe 07/09, verglichen mit der Stichprobe 06/09, mehr Freude beim Mathematiklernen empfinden. Einerseits ist die prozentuale Besetzung der beiden positiven Abschnitte (hoch, eher hoch) grösser (43% vs. 31%), andererseits diejenige der beiden Abschnitte am anderen Ende des Spektrums (tief, eher tief) kleiner (28% vs. 40%). In einer entsprechenden Analyse bestätigt sich dies: Stichprobe 07/09 ($M = 2.73$) zeigt mehr Freude beim Mathematiklernen als Stichprobe 06/09 ($M = 2.50$).

Abbildung 7.9: Freude beim Mathematiklernen

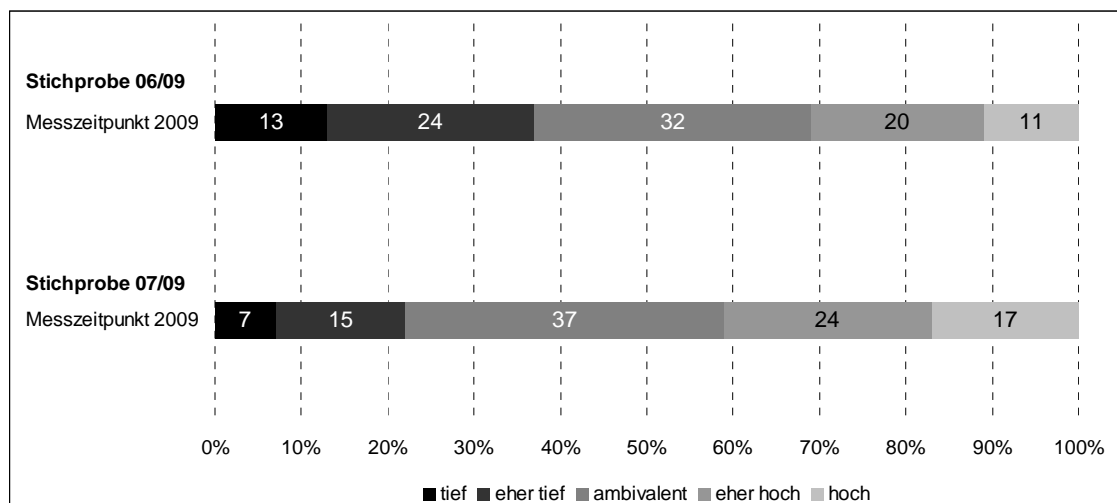


Hinsichtlich der Stichprobe 06/09 zeigt sich bezogen auf die unterschiedlichen schulischen Laufbahnen, dass in der Gruppe Langgymnasium die Freude am Mathematiklernen am stärksten ausgeprägt ist ($M = 2.61$). In den Gruppen Primarstufe sowie Abteilung A der Sekundarschule liegen die Mittelwerte etwas ($M = 2.56$ resp. $M = 2.52$), in der Gruppe Abteilung B/C jedoch deutlich niedriger ($M = 2.39$). Statistisch bedeutsam sind lediglich die Unterschiede zwischen der Gruppe Abteilung B/C der Sekundarschule und allen übrigen Gruppen.

Ein nahezu identisches Bild ergibt sich für *Deutsch* (Abbildung 7.10). Auch hier ist in der Stichprobe 07/09 das positive Ende des Spektrums stärker besetzt als in der Stichprobe 06/09 (41% vs. 31%), das negative jedoch weniger stark (22% vs. 37%) und auch hier liegt der Mittelwert der Stichprobe 06/09 ($M = 2.55$) niedriger als derjenige der Stichprobe 07/09 ($M = 2.80$).

Bezogen auf die unterschiedlichen schulischen Laufbahnen ergeben sich folgende Mittelwerte für die verschiedenen Gruppen: Langgymnasium $M = 2.64$, Abteilung A der Sekundarschule $M = 2.53$, Abteilung B/C der Sekundarschule $M = 2.52$ sowie Primarstufe $M = 2.67$. Statistisch bedeutsame Unterschiede zwischen den Gruppen existieren hier lediglich zwischen den Schülerinnen und Schülern, die ins Langgymnasium wechseln, sowie denjenigen, die in die Abteilungen A bzw. B/C der Sekundarschule übertreten.

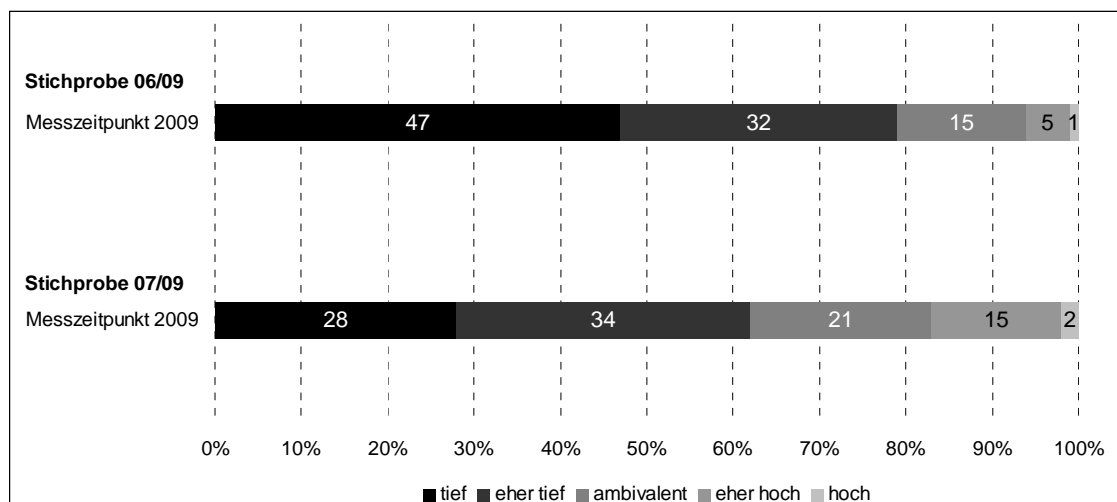
Abbildung 7.10: Freude beim Deutschlernen



Angst beim Mathematik- und Deutschlernen

Die Abbildung 7.11 zeigt das Ergebnis für *Mathematik* in den beiden Stichproben 06/09 bzw. 07/09. Insgesamt ist die Angst beim Mathematiklernen in beiden Stichproben gering ausgeprägt. In die Bereiche «wenig Angst» und «eher wenig Angst» fallen 79% der Schülerinnen und Schüler der Stichprobe 06/09 und 62% der Stichprobe 07/09.

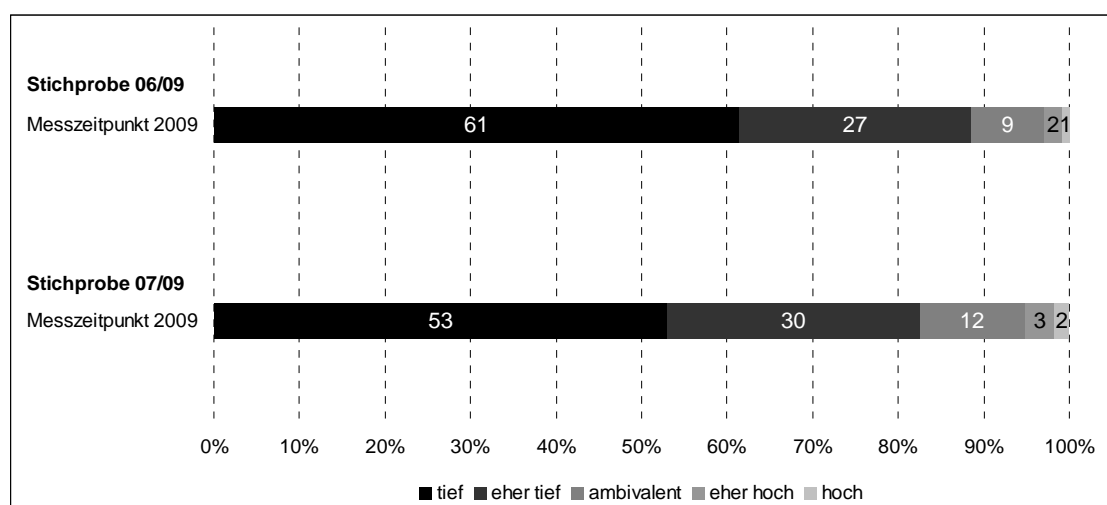
Abbildung 7.11: Angst beim Mathematiklernen



Betrachtet man die andere Seite des Spektrums, so geben in der Stichprobe 06/09 6% der Schülerinnen und Schüler an, dass die Angst beim Mathematiklernen (eher) hoch sei, in der Stichprobe 07/08 hingegen sind es immerhin 17%. In einem Vergleich der Mittelwerte erweist sich die Angst beim Mathematiklernen in der Stichprobe 07/09 ($M = 2.15$) als stärker ausgeprägt als in der Stichprobe 06/09 ($M = 1.82$).

Bezogen auf Schülerinnen und Schüler mit unterschiedlichen schulischen Laufbahnen weisen die vier Gruppen folgende Mittelwerte auf: Langgymnasium $M = 1.38$, Abteilung A der Sekundarschule $M = 1.72$, Abteilung B/C der Sekundarschule $M = 2.17$ sowie Primarstufe $M = 2.07$. Allein letztere beiden unterscheiden sich nicht bedeutsam voneinander.

Abbildung 7.12: Angst beim Deutschlernen



Im Falle von *Deutsch* ist die Angst in beiden Stichproben noch geringer ausgeprägt als in Mathematik (Abbildung 7.12). In der Stichprobe 06/09 berichteten 88%, in der Stichprobe 07/09 83% der Schülerinnen und Schülern, (eher) wenig Angst zu haben. Umgekehrt sind es nur sehr wenige Schülerinnen und Schüler, die beim Deutschlernen (eher) hohe Angstwerte aufweisen (Stichprobe 06/07 bei 3%; Stichprobe 07/09 bei 5%). In der Stichprobe 07/09 ($M = 1.73$) liegen die Angstwerte höher als in der Stichprobe 06/07 ($M = 1.61$).

Bezüglich der Gruppen von Schülerinnen und Schüler mit unterschiedlichen schulischen Laufbahnen zeigt sich ein ähnliches Bild wie in Mathematik. Die tiefsten Angstwerte finden sich in der Gruppe Langgymnasium ($M = 1.30$). Etwas höher liegen die Werte in der Gruppe Abteilung A der Sekundarschule ($M = 1.54$). Die höchsten Werte finden sich in der Gruppe Abteilung B/C der Sekundarschule ($M = 1.85$) sowie in der Gruppe Primarstufe ($M = 1.84$). Mit Ausnahme letzterer beiden Gruppen erweisen sich alle anderen Mittelwertsdifferenzen als statistisch bedeutsam.

7.5 Fazit

Die Kompetenzüberzeugung in Mathematik und Deutsch verringert sich in den letzten Jahren der Primarschulzeit in statistisch bedeutsamer Weise, d.h., das leistungsmässige Selbstvertrauen der Schülerinnen und Schüler sinkt. Ein ähnliches Bild zeigt sich auch hinsichtlich

der Valenzüberzeugung. Der Stellenwert, den Mathematik und Deutsch für die Schülerinnen und Schüler hat, schwindet mit zunehmender Schuldauer und es spielt dabei nahezu keine Rolle, ob die Schülerinnen und Schüler 2006 oder 2007 am Ende der 3. Klasse standen.

Diese allgemeinen Abwärtstrends sollten jedoch nicht überbewertet werden. Solche Entwicklungen sind keineswegs ein neues Phänomen, man kennt sie auch aus anderen Ländern mit teilweise nicht so stark selektierenden Schulsystemen (vgl. Fredericks & Eccles, 2002; Helmke, 1997; Jacobs, Laza, Osgood, Eccles & Wigfield, 2002; Wigfield & Eccles, 2002; Wigfield et al., 1997). Die Ursachen für den Rückgang von Kompetenz- und Valenzüberzeugungen sind vielfältig. Es wäre verfehlt, die Gründe hierfür allein bei der Schule zu suchen. Mögliche Ursachen wurden bereits anlässlich der letzten Lernstandserhebung genannt (detaillierter vgl. Buff, 2008a, S. 91f.). Im Falle der Kompetenzüberzeugung ist dies u.a. die kognitive Entwicklung. Die Fähigkeit zu sozialen Vergleichen steigt, womit zumindest einige Schülerinnen und Schüler ihre anfänglich zu optimistischen Einschätzungen nach unten korrigieren (müssen). Man muss sich in diesem Zusammenhang vor Augen halten, dass Ende der 3. Klasse 06/07 die Kompetenz- wie auch die Valenzüberzeugung insgesamt (sehr) hoch waren. Auch Kontexteinflüsse können eine Rolle spielen, wie etwa eine verstärkte Orientierung der Lehrpersonen an einer sozialen Bezugsnorm bei der Leistungsbeurteilung u.a. im Zusammenhang mit den bevorstehenden Übertritten in die Sekundarstufe I im deutschsprachigen Raum. Das Absinken von Valenzüberzeugungen wird häufig als eine Folge sinkender Kompetenzüberzeugungen betrachtet, jedoch auch als Folge der zunehmenden Ausbildung einer eigenen Identität mit spezifischen Vorlieben und Abneigungen.

In die Analysen hinsichtlich der Frage allfällig differenter Entwicklungstrends in Gruppen mit unterschiedlichen schulischen Laufbahnen nach der Lernstandserhebung 2009 wurden lediglich diejenigen Schülerinnen und Schüler einbezogen, welche 2006 am Ende der 3. Klasse standen, da bei standardmässigem weiterem Verlauf nur diese nach der Erhebung 2009 in die Sekundarstufe I hätten übertreten können. Hinsichtlich der Kompetenzüberzeugung zeigt sich in beiden Fächern ein mehr oder weniger analoges Bild. Die Kompetenzüberzeugung sinkt zwar in allen Gruppen in den drei Jahren nach Ende der 3. Klasse in statistisch bedeutsamer Weise. Der Rückgang ist aber bei denjenigen Schülerinnen und Schülern, welche nach der Lernstandserhebung 2009 in die Abteilungen A resp. B/C der Sekundarschule übertraten, deutlich ausgeprägter als in den anderen beiden Gruppen (Langgymnasium, Primarstufe). Am stärksten fällt der Trend bei denjenigen aus, welche in die Abteilung B/C der Sekundarschule wechselten. Ein nahezu identisches Entwicklungsmuster zeigt sich auch bezüglich der Valenzüberzeugung in Mathematik und Deutsch. Der Unterschied zwischen Kompetenz- und Valenzüberzeugung liegt darin, dass sich die vier Gruppen hinsichtlich ihrer Kompetenzüberzeugung zu beiden Messzeitpunkten deutlich pointierter voneinander unterscheiden. Bezüglich der Valenzüberzeugung sind kaum grosse Unterschiede zwischen den Gruppen feststellbar. Hinsichtlich der Kompetenzüberzeugung zeigt sich mehr oder weniger das Bild, das man erwarten würde: Am höchsten ist diese bei denjenigen Schülerinnen und Schülern, welche ins Langgymnasium übertreten, am niedrigsten bei denjenigen, die in die Abteilung B/C der Sekundarschule wechseln oder in der Primarschule verbleiben.

Es liegt nahe, die sich zeigenden Entwicklungsmuster mit dem Übertritt in die Sekundarstufe I bzw. mit den während der Zuweisungsphase und/oder mit dem Zuweisungsentscheid ge-

machten Erfahrungen in Zusammenhang zu bringen. Auffällig ist bspw., dass die Abwärtstrends in der Gruppe derjenigen Schülerinnen und Schülern, die ins Langgymnasium – also in den prestigeträchtigsten Schultyp der Sekundarstufe I – wechseln, tendenziell am schwächsten, in der Gruppe derjenigen, die in die Abteilung B/C der Sekundarschule übertreten, jedoch am stärksten ausfällt. Ähnliche Befundmuster ergaben sich im deutschsprachigen Raum hinsichtlich kompetenzbezogener Überzeugungen auch bei Buff (1991a, 1991b) und Wagner (2001). Die Antwort auf die Frage, ob und wie stark das Faktum *Selektionsprozess in verschiedene Schultypen* bzw. die damit gemachten Erfahrungen zu den differierenden Entwicklungen beitragen, muss offenbleiben. Eine akkurate Klärung dieser Frage ist anhand der Lernstandserhebungsdaten nicht möglich, da keine Vergleichsgruppe ohne Selektionserfahrung vorhanden ist. Aus den vorliegenden Ergebnissen zur Entwicklung von Kompetenz- und Valenzüberzeugung lassen sich jedenfalls keine Argumente für oder gegen die Selektion ableiten.

Hinsichtlich der beiden lernbezogenen Emotionen zeigt sich alles in allem im Falle der Angst ein erfreuliches Bild. Nur ein kleiner Teil der Schülerinnen und Schüler berichtet, dass die Angst beim Mathematik- bzw. Deutschlernen eher gross oder gross sei. Das heisst jedoch nicht, dass umgekehrt die Freude am Lernen in den beiden Fächern entsprechend hoch wäre. Diese hält sich eher in Grenzen.

Das Ergebnis hinsichtlich der Lernfreude sieht auf den ersten Blick nicht sonderlich positiv aus. Es gilt hier jedoch zu bedenken, dass Lernen mit zunehmender Schuldauer und insbesondere in Fächern wie Mathematik und Deutsch für immer mehr Schülerinnen und Schüler unweigerlich (auch) harte Arbeit bedeutet, die nicht immer (nur) Spass macht. Buff et al. (2010a) sind der Ansicht, dass es wohl eine Illusion wäre, viele oder gar die Mehrheit der Schülerinnen und Schüler dazu bringen zu wollen, dass sie rein aus Freude an der Sache lernen. Wichtiger erscheine es längerfristig viel eher, dass den Schülerinnen und Schüler die Einsicht in die Wichtigkeit und Nützlichkeit eines Lerngegenstandes nicht verloren gehe (vgl. hierzu auch Koestner & Losier, 2002; Lepper & Henderlong, 2000). Der hohe Stellenwert, den die Mehrheit der Schülerinnen und Schüler Mathematik wie auch Deutsch nach sechs Jahren Primarschule immer noch einräumt, kann als Indiz dafür gesehen werden, dass es in der Primarschulzeit gelungen zu sein scheint, mitunter die Einsicht der Schülerinnen und Schüler in die Wichtigkeit und Nützlichkeit der beiden Fächer auf einem hohen Niveau zu halten.

7.6 Anmerkungen

- ¹ Allfällige kleinere Abweichungen zwischen den hier und nachfolgend im Text ausgewiesenen Mittelwerten und denjenigen, die sich aus den Berechnungen aufgrund der Angaben in den entsprechenden Tabellen (vgl. Kapitel 7, Teil 2) ergeben, sind in der Rundung der Werteangaben in den Tabellen auf zwei Kommastellen begründet.

8 Motivation und Leistung in Mathematik und Deutsch

Alex Buff

8.1 Einleitung

Theoretischer Hintergrund

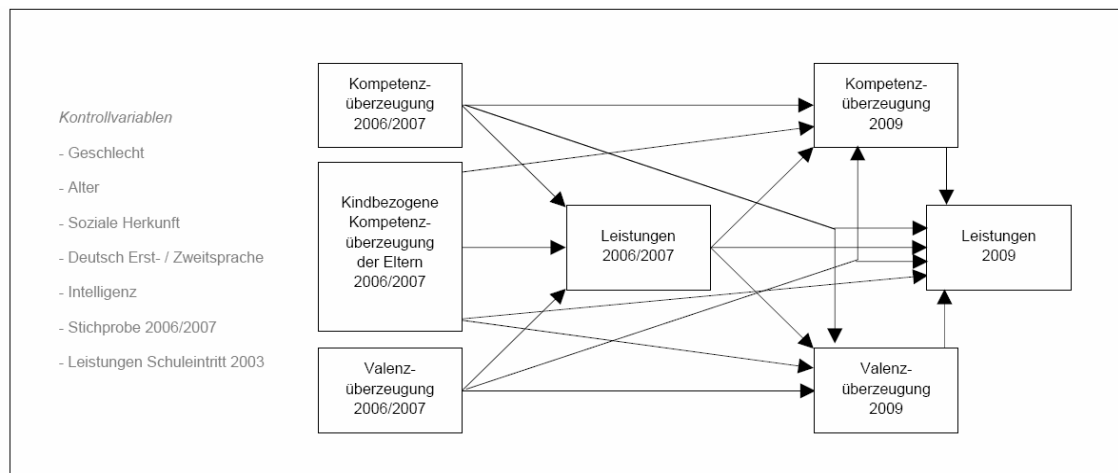
Wie in Kapitel 7, Teil 1 einleitend dargestellt, werden fachspezifische kompetenz- und valenzbezogene Überzeugungen als zentrale Ingredienzien einer lern-leistungsbezogenen *motivationalen Orientierung* betrachtet. Sie stellen aufseiten der Schülerinnen und Schüler individuelle motivational-affektive Lernvoraussetzungen dar, welche massgeblich darüber entscheiden, ob und wie schulische Bildungsangebote genutzt werden (Fend, 1998; Helmke, 2003; Pauli & Reusser, 2006). Helmke (2003) bspw. erachtet kompetenz- und valenzbezogene Überzeugungen als entscheidende individuelle Eingangsvoraussetzungen dafür, «ob, wie lange, wie erfolgreich jemand lernt und was er leistet» (S. 43). Kompetenz- und valenzbezogene Überzeugungen beeinflussen schulische Lernaktivitäten in konkreten Situationen und über deren Qualität (Anstrengung, Zeitinvestment, Einsatz spezifischer Lernstrategien) mittelbar auch die Leistungen (Pekrun & Schiefele, 1996; Schiefele & Pekrun, 1996; Trautwein, Lüdtke, Schnyder & Niggli, 2006). Mit anderen Worten: Neben kognitiven Lernvoraussetzungen im engeren Sinne (Intelligenz, Vorwissen usw.) spielt für die Leistungserbringung einerseits eine zentrale Rolle, was sich die Schülerinnen und Schüler leistungsmässig selbst zutrauen (Kompetenzüberzeugungen), andererseits jedoch auch, welche Bedeutung und welchen Nutzen sie einem Lerngegenstand (z.B. einem Fach) zuschreiben und ob für sie mit diesem Lerngegenstand ein eher positiver oder negativer Anreiz verbunden ist (Valenzüberzeugungen).

Insbesondere internationale Schulleistungsstudien, wie etwa TIMSS (Moser, Ramseier, Keller & Huber, 1997), PISA (Baumert et al., 2001) oder IGLU (Bos et al., 2003), machen deutlich, dass Leistungen von Kindern und Jugendlichen in Mathematik, Lesen oder Naturwissenschaften in hohem Masse auch von Disparitäten der sozialen bzw. familiären Herkunft beeinflusst werden: sozioökonomischer Status der Familie, Migrationshintergrund usw. Will man jedoch leistungsbezogenes Verhalten und Schulerfolg von Kindern und Jugendlichen besser verstehen, wird man nicht umhin kommen, familiären Faktoren mit ihren fördernden oder hemmenden Wirkungen stärker als bisher Beachtung zu schenken (Bleeker & Jacobs, 2004; Goni-da & Urdan, 2007; Hirsjärvi & Perälä-Littunen, 2001). Hierbei sollte es sich insbesondere auch um Faktoren jenseits struktureller Merkmale familiärer Lebensverhältnisse (sozioökonomischer Status, Migrationshintergrund usw.) handeln, denn es ist davon auszugehen, dass Grössen wie etwa kindbezogene elterliche Überzeugungen oder auch schulbezogenes elterliches Verhalten letztlich den Schulerfolg im weitesten Sinne mit beeinflussen (Davis-Kean, 2005; Jodl, Michael, Malanchuk, Eccles & Sameroff, 2001; Lorenz & Wild, 2007). Eccles (2007) bspw. ist der Ansicht, «parents' beliefs and psychological and social resources can override the effects of even the most stressful demographic characteristics on children's school achievement and motivation» (S. 671), denn das familiäre Umfeld lässt sich mit Blick auf Angebots-Nutzungs-Modelle (Fend, 1998; Helmke, 2003; Pauli & Reusser, 2006) insgesamt als ein Entwicklungsmilieu auffassen, das unterschiedliche Ressourcen für die Ausbil-

derung derjenigen individuellen Merkmale aufseiten des Kindes bereitstellt, die es ihm erst ermöglichen, schulische Bildungsangebote auch wirklich zu nutzen, so zum Beispiel selbstbezogene Kompetenz- und Valenzüberzeugungen.

Vor diesem Hintergrund fokussiert der vorliegende Beitrag die Frage der Determinanten von Schulleistung in Mathematik und Deutsch unter einer längsschnittlichen Perspektive. Primär interessieren Effekte und wechselseitige Beziehungen zwischen den Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler sowie deren Leistungen in der zweiten Hälfte der Primarschulzeit. Als möglicher Einflussfaktor jenseits struktureller Merkmale familiärer Lebensverhältnisse interessieren inhaltlich auch die kindbezogenen elterlichen Kompetenzüberzeugungen. Ihnen wird von verschiedenen Seiten eine besondere Bedeutung hinsichtlich der Motivation und der Leistungen der Schülerinnen und Schüler zugeschrieben (Aunola, Nurmi, Lerkkanen & Rasku-Puttonen, 2003; Dai, 2002; Eccles, 2007; Fredricks & Eccles, 2002; Frome & Eccles, 1998; Helmke, Schrader & Lehneis-Klepper, 1991; Phillips, 1987).

Abbildung 8.1: Primär interessierende Variablenzusammenhänge



In Abbildung 8.1 sind die primär interessierenden Variablenzusammenhänge grafisch dargestellt. Deutlich wird, dass im Variablennetzwerk von wechselseitigen Beziehungen über die Zeit hinweg ausgegangen wird. Das heisst im Falle der Kompetenzüberzeugung 2006/2007 bspw., dass von dieser ein Effekt auf die Leistungen 2006/2007 erwartet wird, die Leistungen jedoch wiederum die Kompetenzüberzeugung 2009 beeinflussen sollten. Abgebildet sind zudem lediglich die erwarteten *direkten* Effekte. Dies allerdings ist keineswegs der einzige Weg, über den eine Grösse eine andere beeinflussen kann. Zu rechnen ist immer auch mit *indirekten*, medierten bzw. vermittelten Effekten. Im vorliegenden Zusammenhang stehen diejenigen der beiden Motivationskomponenten Kompetenz- und Valenzüberzeugung 2006/2007 auf die Leistungen 2009 im Fokus. Hier wäre theoretisch zu erwarten, dass bei Berücksichtigung der Leistungen 2006/2007 sowie der beiden Motivationskomponenten 2009 die Kompetenz- und Valenzüberzeugung 2006/2007 kaum mehr in direkter Weise Wirkung auf die Leistungen 2009 zeitigen, sondern indirekt, vermittelt über die Leistungen 2006/2007 sowie die Pendanten der beiden Komponenten 2009. Diese Annahme soll ebenfalls geprüft werden.

Bemerkungen zur Ergebnispräsentation

Die Ergebnisse werden in Form von Pfadmodellen präsentiert. Abgebildet und kommentiert sind jeweils lediglich die Effekte bezüglich der primär interessierenden Variablenzusammenhänge zwischen Kompetenz- und Valenzüberzeugung der Schülerinnen und Schüler, deren Leistungen sowie der elterlichen kindbezogenen Kompetenzüberzeugung. Hinsichtlich dieser Variablen konzentrieren sich die Analysen auf die Lernstandserhebung Ende der 3. Klasse 2006/2007 sowie auf diejenige von 2009, da die Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler lediglich zu diesen beiden Messzeitpunkten erhoben wurden.

In den verschiedenen Abbildungen in grau bzw. aufgehellt dargestellt sind jeweils diejenigen Grössen, welche in den entsprechenden Analysen nicht primär interessieren, oder Effekte, die in vorangehenden Ergebnisdarstellungen bereits erwähnt wurden. In allen Analysen als Kontrollvariablen berücksichtigt wurden auch das Geschlecht, das Alter, die soziale Herkunft, Deutsch als Erst- oder Zweitsprache, die Intelligenz, die leistungsmässigen Voraussetzungen der Schülerinnen und Schüler bei Schuleintritt im jeweiligen Fach sowie der Umstand, ob sie 2006 oder 2007 an der Lernstandserhebung am Ende der 3. Klasse teilnahmen. Im Falle Letzterer handelt es sich um Schülerinnen und Schüler, die das Ende der 3. Klasse erst nach vier Schuljahren erreicht hatten (Sonderklasse A, Repetition usw.). Bei den nachfolgend hinsichtlich einzelner interessierender Variablen berichteten Effekte handelt es sich also um eigenständige Erklärungsbeiträge, welche diese leisten, unabhängig von Effekten anderer, in den Modellen mitberücksichtigter Grössen.

Sofern nichts anderes erwähnt, sind alle berichteten Zusammenhänge statistisch bedeutsam. In den Abbildungen sind hinsichtlich der interessierenden Zusammenhänge auch ausschliesslich die statistisch bedeutsamen Effekte dargestellt. Angegeben werden dazu jeweils die standardisierten Regressionskoeffizienten, aufgrund derer auch die Bedeutsamkeit der einzelnen Effekte untereinander verglichen werden kann. Der Wertebereich der standardisierten Regressionskoeffizienten liegt zwischen -1 und 1 . Je stärker der Wert von 0 verschieden ist, desto bedeutsamer ist der Effekt der entsprechenden Grösse hinsichtlich des Merkmals, welches es zu erklären gilt. Inhaltlich indizieren Regressionskoeffizienten mit einem positiven Vorzeichen einen Zusammenhang im Sinne von «je höher, desto höher bzw. je tiefer, desto tiefer», solche mit einem negativen einen im Sinne von «je höher, desto tiefer bzw. je tiefer, desto höher».

Nähere methodische Angaben (Stichprobe, Instrumente usw.) sowie eine detailliertere Darstellung der einzelnen Analyseergebnisse finden sich in Kapitel 8, Teil 2.

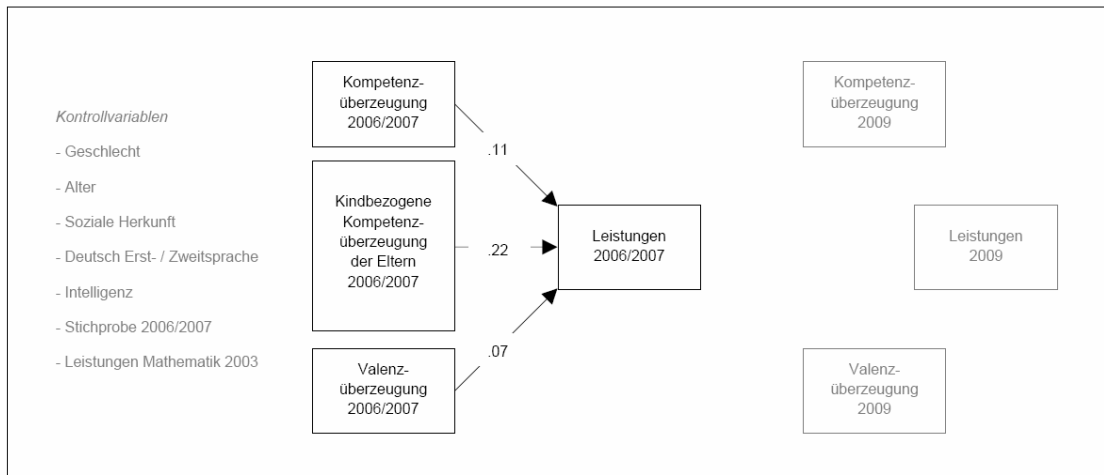
8.2 Motivation und Leistung in Mathematik

Leistungen in Mathematik 2006/2007

Abbildung 8.2 zeigt das Ergebnis bezogen auf die Mathematikleistungen Ende der 3. Klasse 2006/2007. Hinsichtlich der beiden Motivationskomponenten Kompetenz- und Valenzüberzeugung findet sich jeweils theoriekonform ein positiver Effekt bezüglich der Leistungen: Je höher die Kompetenz- resp. die Valenzüberzeugung, desto besser waren die Leistungen. Gleiches gilt hinsichtlich der elterlichen kindbezogenen Kompetenzüberzeugung: Je positiver die

Eltern die Kompetenz ihrer Kinder 2006/2007 einschätzten, desto höher waren deren Leistungen.

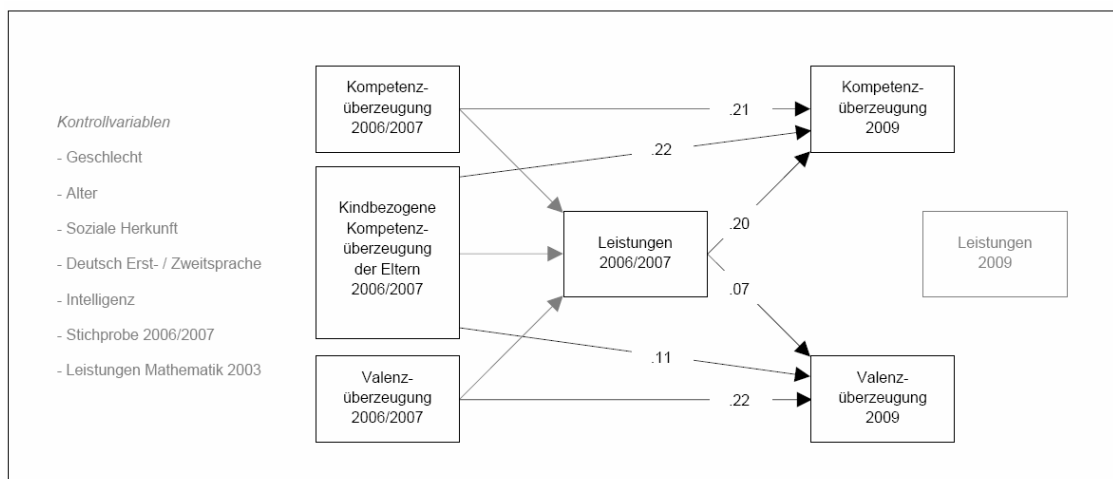
Abbildung 8.2: Vorhersage der Leistungen in Mathematik 2006/2007



Kompetenz- und Valenzüberzeugung in Mathematik 2009

In Abbildung 8.3 wird deutlich, dass sich hinsichtlich der Kompetenzüberzeugung 2009 von den hier primär interessierenden Grössen – Kompetenz- und Valenzüberzeugung, Mathematikleistungen sowie die kindbezogene elterliche Kompetenzüberzeugung 2006/2007 – mit Ausnahme der Valenzüberzeugung alle übrigen als relevante Prädiktoren erweisen: Je höher die Werte in diesen Variablen 2006/2007, desto höher war die Kompetenzüberzeugung zwei resp. drei Jahre später.

Abbildung 8.3: Vorhersage der Kompetenz- und Valenzüberzeugung in Mathematik 2009



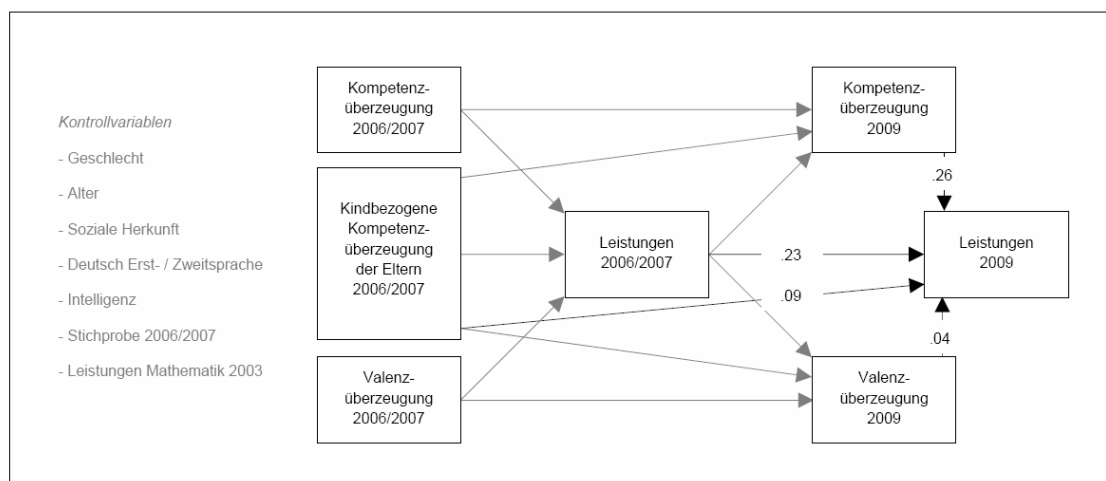
Hinsichtlich der Valenzüberzeugung 2009 zeigt sich, dass diese primär von der zwei bzw. drei Jahre früher gemessenen Valenzüberzeugung beeinflusst wird. Deutlich schwächer sind die Effekte der elterlichen kindbezogenen Kompetenzüberzeugung und insbesondere der Ma-

thematische Leistungen. Dies bedeutet insgesamt, dass der Stellenwert, den Schülerinnen und Schülern der Mathematik 2009 zuschreiben, um so höher ist, je positiver deren Valenzüberzeugung und je höher deren Leistungen 2006/2007 waren und je positiver die Eltern die Kompetenzen ihrer Kinder Ende der 3. Klasse einschätzten. Die Kompetenzüberzeugung 2006/2007 erweist sich hinsichtlich der Valenzüberzeugung 2009 als irrelevanter Prädiktor.

Leistungen in Mathematik 2009

In einer ersten Analyse (ohne Abbildung) wurde geprüft, inwieweit die am Ende der 3. Klasse erhobene Kompetenz- und Valenzüberzeugung der Schülerinnen und Schüler, deren damaligen Leistungen sowie die kindbezogene Kompetenzüberzeugung der Eltern zur Erklärung der Mathematikleistungen 2009 beitragen. Von den zentral interessierenden Größen sind es die Mathematikleistungen 2006/2007, welche sich als bedeutsamster Prädiktor der Leistungen 2009 erweisen: Je höher die Leistungen am Ende der 3. Klasse, desto besser waren die Leistungen 2009. Von den beiden Motivationskomponenten ist es allein die Kompetenzüberzeugung Ende der 3. Klasse, welche ebenfalls einen positiven Effekt auf die Mathematikleistungen 2009 ausübt: Je höher die Schülerinnen und Schüler ihre Kompetenzen 2006/2007 einschätzen, desto besser waren die Leistungen 2009. Ein positiver Effekt findet sich auch hinsichtlich der elterlichen kindbezogenen Kompetenzüberzeugung 2006/2007. Eine höhere elterliche kindbezogene Kompetenzüberzeugung wirkt sich also nicht allein günstig auf die Kompetenz- wie auch Valenzüberzeugung der Schülerinnen und Schüler zwei bzw. drei Jahre später aus (vgl. oben), sondern auch auf deren Leistungen.

Abbildung 8.4: Vorhersage der Leistungen in Mathematik 2009



In einer zweiten Analyse wurden als zusätzliche Prädiktoren die Kompetenz- und die Valenzüberzeugung der Schülerinnen und Schüler 2009 berücksichtigt (vgl. Abbildung 8.4). Verglichen mit der ersten Analyse führt dies dazu, dass sich hinsichtlich der am Ende der 3. Klasse erhobenen Kompetenzüberzeugung nun kein statistisch bedeutsamer Effekt bezüglich der Mathematikleistungen 2009 mehr finden lässt. Die mit den Leistungen 2009 zeitgleich erhobene Kompetenz- wie auch Valenzüberzeugung erweisen sich nun jedoch beide – neben den bereits erwähnten Mathematikleistungen 2006/2007 sowie der elterlichen kindbezogenen

Kompetenzüberzeugung 2006/2007 – als relevante eigenständige Prädiktoren im Sinne von «je höher, desto höher».

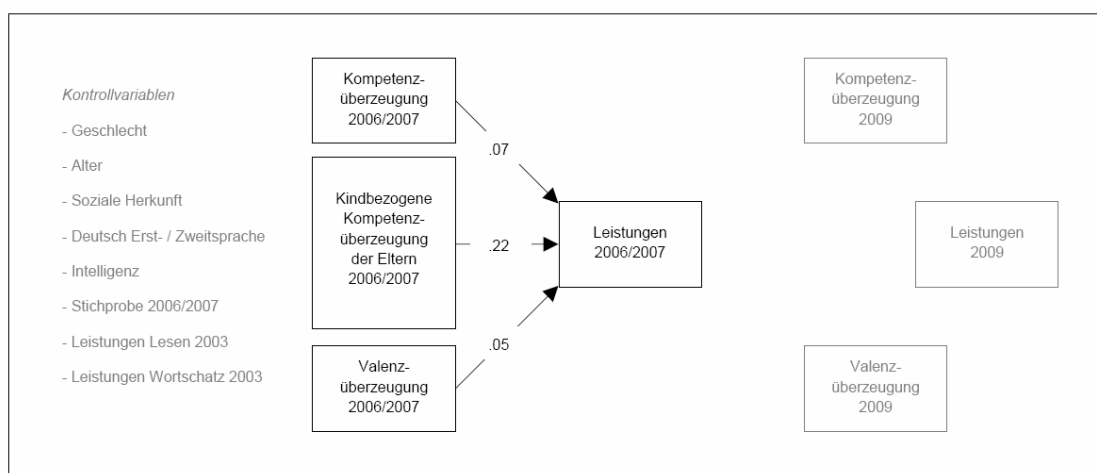
Bei allen bisher dargestellten Ergebnissen handelt es sich um *direkte* Effekte. Theoretisch wurde jedoch erwartet, dass die Effekte der Motivation 2006/2007 auf die Leistungen 2009 eigentlich nicht direkter, sondern indirekter Natur sind, vermittelt über die Leistungen 2006/2007 und die Motivation 2009. Diese indirekten Effekte der beiden Motivationskomponenten 2006/2007 auf die Leistungen 2009 wurden daher geprüft. Im Falle der Kompetenzüberzeugung findet sich in der Tat ein statistisch bedeutsamer indirekter positiver Effekt der Kompetenzüberzeugung 2006/2007 auf die Leistungen 2009. Dies bedeutet, dass eine positive Kompetenzüberzeugung Ende der 3. Klasse 2006/2007 die Leistungen 2009 – also zwei bzw. drei Jahre später – auf indirekte Weise günstig beeinflussen, vermittelt über die Leistungen Ende der 3. Klasse sowie die Kompetenzüberzeugungen 2009. Hinsichtlich der zweiten Motivationskomponente Valenzüberzeugung zeigt sich jedoch kein entsprechender Effekt.

8.3 Motivation und Leistung in Deutsch

Leistungen in Deutsch 2006/2007

Wie schon im Falle von Mathematik erweisen sich auch in Deutsch die Kompetenz- und die Valenzüberzeugung 2006/2007 als jeweils bedeutsame Prädiktoren der Leistungen 2006/2007: Eine höhere Kompetenz- bzw. Valenzüberzeugung geht einher mit besseren Leistungen. Ebenso manifestiert sich ein positiver Effekt der elterlichen kindbezogenen Kompetenzüberzeugung. Damit zeigt sich auch in Deutsch: Je mehr sich die Schülerinnen und Schüler leistungsmässig zutrauen, je grösser der Stellenwert ist, den sie dem Fach zuschreiben, und insbesondere je höher ihre Kompetenzen von den Eltern eingeschätzt werden, desto besser sind ihre Leistungen.

Abbildung 8.5: Vorhersage der Leistungen in Deutsch 2006/2007

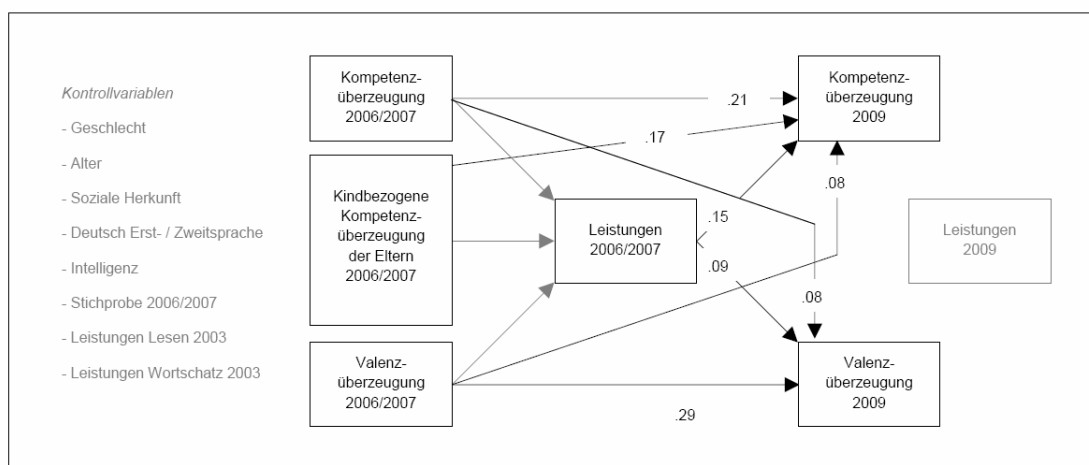


Kompetenz- und Valenzüberzeugung in Deutsch 2009

Auch im Falle von Deutsch erweisen sich die Kompetenzüberzeugung, die Leistungen sowie die kindbezogene elterliche Kompetenzeinschätzung 2006/2007 als relevante Prädiktoren hinsichtlich der Kompetenzüberzeugung der Schülerinnen und Schüler 2009. Eine höhere Kompetenzüberzeugung und bessere Leistungen sowie eine positivere elterliche kindbezogene Kompetenzeinschätzung 2006/2007 wirken sich günstig auf die Kompetenzüberzeugung der Schülerinnen und Schüler 2009 aus. Im Unterschied zu Mathematik ist es hier jedoch auch die Valenzüberzeugung, welche einen statistisch bedeutsamen positiven Effekt zeitigt. Je mehr Relevanz, Nutzen und Anreiz, also je grösser der Stellenwert, der dem Fach Deutsch 2006/2007 zugeschrieben wurde, desto höher ist die Kompetenzüberzeugung zwei bzw. drei Jahre später.

Bezogen auf die Valenzüberzeugung 2009 finden sich – wie in Mathematik – positive Effekte der Valenzüberzeugung wie auch der Leistungen 2006/2007 im Sinne von «je besser, desto höher». Im Gegensatz zu Mathematik erweist sich in Deutsch auch die Kompetenzüberzeugung 2006/2007 als relevanter Prädiktor der Valenzüberzeugung 2009: Schülerinnen und Schüler, die sich 2006/2007 in Deutsch mehr zutrauen, schreiben 2009 Deutsch mehr Bedeutung und Nützlichkeit zu und das Fach hat für sie einen positiveren Anreiz. Hinsichtlich der bisher prädiktiv immer bedeutsamen elterlichen kindbezogenen Kompetenzüberzeugung 2006/2007 zeigt sich im Falle der Valenzüberzeugung 2009 in Deutsch jedoch kein statistisch bedeutsamer Effekt.

Abbildung 8.6: Vorhersage der Kompetenz- und Valenzüberzeugung in Deutsch 2009



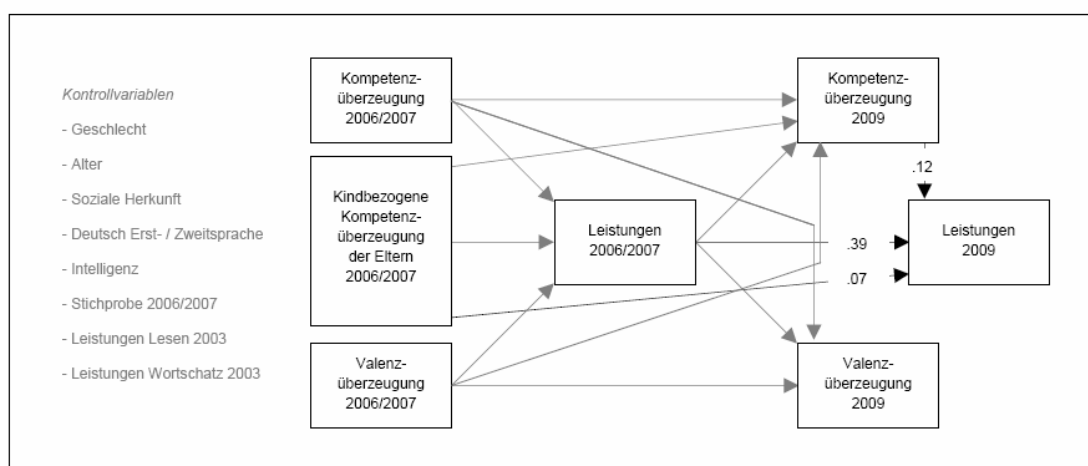
Leistungen in Deutsch 2009

Wiederum wurde in einer ersten Analyse (ohne Abbildung) geprüft, inwieweit die am Ende der 3. Klasse erhobene Kompetenz- und Valenzüberzeugung der Schülerinnen und Schüler, deren Leistungen sowie die kindbezogene Kompetenzüberzeugung der Eltern zur Erklärung der Deutschleistungen 2009 beitragen. Hinsichtlich der Leistungen in Deutsch 2009 zeigt sich verglichen mit Mathematik ein etwas anderes Bild. Sowohl die Kompetenz- als auch die Valenzüberzeugung 2006/2007 erweisen sich hier als irrelevante Prädiktoren. Wiederum sind es

jedoch die Leistungen der Schülerinnen und Schüler in Deutsch 2006/2007, aufgrund derer sich die Leistungen 2009 vorhersagen lassen. Konsistent mit dem Ergebnis in Mathematik ist jedoch der positive Effekt der elterlichen kindbezogenen Kompetenzüberzeugung. Auch in Deutsch zeigt sich, dass diejenigen Kinder höhere Leistungen 2009 erbringen, die von ihren Eltern 2006/2007 bezogen auf ihrer Fähigkeiten in Deutsch besser eingeschätzt wurden.

Werden wiederum in einer zweiten Analyse die Kompetenz- und die Valenzüberzeugung 2009 zusätzlich als Prädiktoren berücksichtigt, so findet sich, wie in Abbildung 8.7 dargestellt, lediglich bei der Kompetenzüberzeugung ein signifikanter positiver Effekt hinsichtlich der Leistungen 2009. An den bereits in der ersten Analyse beschriebenen beiden anderen Effekten ändert sich nichts. Alles in allem zeigt sich: Je besser die elterliche kindbezogene Kompetenzüberzeugung sowie insbesondere die Deutschleistungen der Schülerinnen und Schüler 2006/2007 und deren Kompetenzüberzeugung 2009 sind, desto höher sind die Leistungen 2009.

Abbildung 8.7: Vorhersage der Leistungen in Deutsch 2009



Aufgrund der Resultate stellt sich die Frage, ob die Kompetenz- und Valenzüberzeugung 2006/2007 für die Leistungen 2009 als Einflussgrößen irrelevant sind, was angesichts der fehlenden direkten Effekte (vgl. oben) vermutet werden könnte. Wie im Abschnitt zum Fach Mathematik erwähnt, kann diese Frage nicht allein aufgrund der direkten Effekte beantwortet werden, denn es ist auch mit indirekten, vermittelten Einflüssen zu rechnen. Geprüft wurde auch hier der aufgrund theoretischer Annahmen primär interessierende indirekte Effekt der Motivation 2006/2007 auf die Leistungen 2009 via die Leistungen 2006/2007 und die Motivation 2009. Die Prüfung dieses indirekten Effekts beschränkt sich allerdings auf die Kompetenzüberzeugung, da hinsichtlich der Valenzüberzeugung 2009 kein direkter Einfluss auf die Leistungen 2009 festgestellt werden konnte. Wie schon in Mathematik findet sich auch in Deutsch ein statistisch bedeutsamer positiver indirekter Effekt der Kompetenzüberzeugung 2006/2007 – vermittelt über die Leistungen 2006/2007 und die Kompetenzüberzeugung 2009 – auf die Leistungen 2009: Je mehr sich die Schülerinnen und Schüler 2006/2007 zutrauten, desto besser waren ihre Leistungen zu diesem Zeitpunkt. Dies wiederum wirkt sich

positiv auf ihr leistungsmässiges Selbstvertrauen 2009 aus, was seinerseits die Leistungen 2009 günstig beeinflusst.

8.4 Fazit

Einmal mehr erweisen sich insbesondere die Kompetenzüberzeugungen der Schülerinnen und Schüler sowie in eingeschränktem Masse auch die Valenzüberzeugungen als individuelle Merkmale mit eigenständiger Erklärungskraft hinsichtlich der Leistungen in den Fächern Mathematik und Deutsch, dies selbst bei Kontrolle einer ganzen Reihe leistungsrelevanter familiärer und individueller Merkmale. Je mehr sich die Schülerinnen und Schüler selbst zutrauen, desto besser sind ihre Leistungen. Dieser Effekt zeigt sich in beiden Fächern kurzfristig, d.h. innerhalb der jeweiligen Erhebungen 2006/2007 und 2009, im Falle von Mathematik jedoch auch längerfristig über zwei bzw. drei Jahre hinweg. Es kommt also nicht nur darauf an, was man effektiv kann, sondern auch darauf, dass man daran glaubt, etwas zu können. Hinsichtlich der Valenzüberzeugungen sind die Ergebnisse weniger eindeutig. Wenn sich bei gleichzeitiger Berücksichtigung der Kompetenzüberzeugungen überhaupt ein positiver Effekt finden lässt, dann nur innerhalb der einzelnen Messzeitpunkte. In Mathematik ist dies jeweils der Fall, in Deutsch lediglich Ende der 3. Klasse 2006/2007. Längerfristige Effekte der Valenzüberzeugungen auf spätere Leistungen sind nicht feststellbar.

Der Befund, dass bei gleichzeitiger Berücksichtigung von Kompetenz- und Valenzüberzeugungen (bzw. ihnen verwandten Grössen) vielfach allein Erstere sich hinsichtlich der Leistungen als relevant erweisen, ist nicht neu (vgl. etwa Buff, Reusser & Pauli, 2010a; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2004; Wigfield & Eccles, 2000). Wird jedoch, wie bereits anlässlich der letzten Lernstandserhebung ausgeführt (vgl. Buff 2008a, S. 112f.), leistungsbezogenes Verhalten in einem etwas umfassenderen Sinne verstanden (d.h. nicht allein im Sinne von Testleistungen und Noten), so präsentiert sich ein etwas anderes Bild. Geht es bspw. um Dinge wie etwa das Engagement in einem Fach, um Wahlverhalten hinsichtlich Leistungskursen, Fächern, Studien- und Berufsrichtungen, dann erweisen sich valenzbezogene Überzeugungen (Wichtigkeit, Nützlichkeit, Anreiz usw.) selbst bei simultaner Berücksichtigung kompetenzbezogener Überzeugungen durchaus als relevant (vgl. etwa Buff et al., 2010a; Eccles, Barber, Updegraff & O'Brien, 1998). Es deutet sich hier an, dass valenzbezogenen Überzeugungen vor allem dann Bedeutung zukommt, wenn auf freiwilliger Basis gehandelt werden kann, wenn jemand also grundsätzlich die Wahl hat, sich für oder gegen bspw. einen Leistungskurs oder ein Engagement in einem Fach zu entscheiden. Welchen Wert, welchen Nutzen und/oder welchen Anreiz ein Lerngegenstand oder ein Fach für die Schülerinnen und Schüler hat, ist also für das Lernen und die Schulkarriere insgesamt keineswegs belanglos. Nur wer etwas als nützlich und wichtig erachtet und/oder diesem Gegenstand auch etwas Positives abzugewinnen vermag, wird sich freiwillig engagiert damit auseinandersetzen. So gesehen dürften leistungsmässige Effekte von Valenzüberzeugungen (gemessen anhand von Testleistungen oder Noten) lediglich bedingt ein relevanter Indikator dafür sein, um deren Bedeutung für schulische Lernprozesse insgesamt beurteilen zu können. Alles in allem zeigt sich hinsichtlich der Leistungen in Mathematik und Deutsch, dass insbesondere Kompetenz-, in abgeschwächtem Masse jedoch auch Valenzüberzeugungen wichtige individuelle Ressourcen darstellen, die es (auch) in der Schule zu erhalten und zu fördern gilt.

Wie die Ergebnisse weiter deutlich machen, handelt es sich bei den Zusammenhängen zwischen Kompetenz- und Valenzüberzeugungen einerseits sowie Leistungen in den beiden Fächern andererseits nicht um einseitig gerichtete Effekte in dem Sinne, dass Erstere die Leistungen beeinflussen, in umgekehrter Richtung jedoch keine Effekte vorhanden wären. Über die Zeit hinweg existieren vielmehr wechselseitige Beziehungen. Leistungen sind bspw. nicht nur von den Kompetenzüberzeugungen abhängig, sondern beeinflussen ihrerseits nachfolgende Kompetenzüberzeugungen und diese wiederum die Leistungen. Dies ist in beiden Fächern der Fall. Gezeigt werden kann auch, dass in Mathematik und Deutsch eine höhere Kompetenzüberzeugung der Schülerinnen und Schüler 2006/2007 zwar bei Einbezug der Kompetenzüberzeugung 2009 keinen direkten günstigen Effekt mehr auf die Leistungen 2009 hat, diese jedoch *indirekt* – vermittelt über die Leistungen 2006/2007 sowie die Kompetenzüberzeugung 2009 – nach wie vor positiv beeinflusst. Dies ist auch ein Indiz dafür, wie wichtig das Zutrauen in die eigene Leistungsfähigkeit längerfristig sein kann. Hinsichtlich der Valenzüberzeugungen finden sich keine entsprechenden indirekten Effekte.

Das wohl interessanteste Ergebnis zeigt sich bezüglich der elterlichen kindbezogenen Kompetenzüberzeugung in Mathematik resp. Deutsch am Ende der 3. Klasse 2006/2007. Deutlich wird, dass diese auch zwei bzw. drei Jahre später noch Einfluss auf die Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler nehmen und sich ebenso auf deren Leistungen in den beiden Fächern auswirken. Das Muster ist in beiden Fächern und bezogen auf alle drei Grössen Kompetenz- und Valenzüberzeugung sowie Leistungen mit einer Ausnahme – Valenzüberzeugung 2009 in Deutsch – immer dasselbe. Je mehr die Eltern ihren Kindern am Ende der 3. Klasse leistungsmässig zutrauen, desto mehr trauen sich die Schülerinnen und Schüler zwei bzw. drei Jahre später selbst zu, desto höher ist der Stellenwert, den sie Mathematik beimessen, und desto besser sind die Leistungen in den beiden Fächern. Diese Ergebnisse machen deutlich, welche Bedeutung familiären Charakteristiken jenseits struktureller Merkmale, wie etwa sozioökonomischer Status, Migrationshintergrund usw., insbesondere auch hinsichtlich des Lernerfolgs zukommt. Das Vertrauen der Eltern in die Leistungsfähigkeit ihrer Kinder scheint hierbei eine ganz zentrale Rolle zu spielen. Dai (2002) vermerkt denn auch entsprechend, «confidence in children's ability can be a great source of social-emotional support» (S. 638). Solchen familiären Merkmalen wird künftig mehr Beachtung zu schenken sein, wenn Schulerfolg oder auch -misserfolg besser verstanden werden sollen.

Bei den obigen Aussagen, welche Effekte der Kompetenz- und Valenzüberzeugung 2009 auf die Leistungen 2009 einschliessen, muss einschränkend darauf hingewiesen werden, dass es sich hier um Querschnittsdaten handelt. Alle drei Grössen wurden am gleichen Halbttag erhoben. Dies beeinträchtigt die Aussagekraft der Ergebnisse, denn es ist unter diesen Umständen letztlich nicht zu sagen, was «Ursache» und was «Wirkung» ist. Die Einflüsse könnten auch umgekehrt verlaufen. Die Modellierung der dargestellten Wirkrichtungen ist theoretisch begründet. Hinsichtlich der Zusammenhänge innerhalb der Lernstandserhebung 2006/2007 ist dies anders. Hier wurden Kompetenz- und Valenzüberzeugungen zeitlich vor den Leistungen erhoben.

9 Unterricht, Motivation und Leistung in Mathematik und Deutsch

Alex Buff

9.1 Einleitung

Theoretischer Hintergrund

«Guter Unterricht» wird als Vehikel betrachtet, mit welchem der Lernertrag der Schülerinnen und Schüler gesteigert werden kann. Wie dieser Unterricht aussehen sollte bzw. was bei dessen Gestaltung zu berücksichtigen wäre, ist seit längerem Gegenstand der Unterrichtsforschung (Helmke, 2003; Helmke, Hosenfeld, Schrader & Wagner, 2002; Helmke & Weinert, 1997b). Guter Unterricht stellt jedoch – wie «Angebots-Nutzungs-Modelle» (Fend, 1998; Helmke, 2003; Pauli & Reusser, 2006) deutlich machen – lediglich ein Angebot von Seiten der Schule dar, das von der anderen Seite, d.h. den Schülerinnen und Schülern, entsprechend genutzt werden muss, damit sich die intendierten leistungsmässigen Effekte auch wirklich einstellen. Mit anderen Worten: Das schulische Angebot kann zwar vieles zum Lernerfolg der Schülerinnen und Schüler beitragen, «bewirken» bzw. «garantieren» kann es ihn letztlich nur bedingt oder in gewissen Fällen gar nicht.

Ob und wie das Angebot genutzt wird, ist, wie an anderer Stelle ausgeführt (vgl. hierzu Kapitel 7.1, Teil 1, und Kapitel 8.1, Teil 1), von vielerlei Faktoren abhängig, auf welche die Schule teilweise kaum Einfluss hat bzw. die sich dem Einfluss der Schule gar gänzlich entziehen. Dennoch, die Qualität des Unterrichts ist ein wichtiger Faktor, über den die Schule die Leistungen der Schülerinnen und Schüler beeinflusst. Die Unterrichtsgestaltung zeitigt jedoch keineswegs allein leistungsmässige Effekte, sondern wirkt sich auch auf «motivational-affektive» Merkmale der Schülerinnen und Schüler aus, welche ihrerseits zentral für die Qualität der Nutzung schulischer Angebote sind; so im vorliegenden Zusammenhang auf die Kompetenzüberzeugung der Schülerinnen und Schüler in einem Fach sowie den intrinsischen Anreiz des Faches. Kurz gesagt wirkt sich die Unterrichtsgestaltung darauf aus, was sich die Schülerinnen und Schüler im Fach leistungsmässig selbst zutrauen (Kompetenzüberzeugung), sowie auf die Freude, die sie bei der Beschäftigung mit dem Fach empfinden (intrinsischer Anreiz). Helmke (2003) bspw. macht in seinem «Angebots-Nutzungs-Modell des Unterrichts» deutlich, dass Unterricht, nicht nur, aber insbesondere, über solcherlei Grössen leistungsmässige Wirkung entfaltet. Kompetenzüberzeugung und intrinsischer Anreiz stellen damit zentrale vermittelnde individuelle Merkmale dar, über die sich der Unterricht indirekt auf die Leistungen auswirkt. Welche Unterrichtsgestaltung beeinflusst nun aber das Vertrauen in die eigenen Fähigkeiten und die Freude an einem Fach günstig? Auch hierzu existieren verschiedenste theoretische Ansätze. Hinsichtlich des intrinsischen Anreizes eines Unterrichtsgegenstandes bzw. eines Faches wäre etwa die Selbstbestimmungstheorie zu nennen (Deci & Ryan, 1993; Ryan & Deci, 2002), bezüglich des Selbstvertrauens z.B. der Mastery-Klima-Ansatz (Satow, 1999, 2002; Satow & Schwarzer, 2003).

Die Auswahl der in der Lernstandserhebung 2009 zu erfassenden Unterrichtsmerkmale erfolgte hinsichtlich ihrer postulierten positiven Effekte unter Berücksichtigung aller drei Perspektiven: bezogen auf die Leistungen, die Kompetenzüberzeugung und den intrinsischen

Anreiz. Erhoben wurden aus der Sicht der Schülerinnen und Schüler: *Regelklarheit, Lernunterstützung, Diagnosekompetenz der Lehrperson, Strukturierungshilfen, Erklärkompetenz der Lehrperson, Motivierung thematisch, Motivierung instrumentell* sowie *Beziehung zur Lehrperson*. Hinsichtlich der Leistungen werden alle acht Qualitätsmerkmale des Unterrichts als relevant erachtet (Helmke, 2003; Helmke et al., 2002; Helmke & Weinert, 1997b). Bezogen auf den intrinsischen Anreiz werden der Beziehung zur Lehrperson, der Motivierung, der Erklärkompetenz sowie den Strukturierungshilfen und der individuellen Lernunterstützung besondere Bedeutung zugeschrieben (vgl. etwa Buff, Reusser & Pauli, 2010a, 2010b; Prenzel & Drechsel, 1996). Für die Kompetenzüberzeugung schliesslich werden die Beziehung zur Lehrperson, ein motivierender und individuell unterstützender Unterricht als besonders wichtig angesehen (vgl. etwa Buff et al., 2010a; Satow, 1999; Satow & Schwarzer, 2003).

Ziel der nachfolgenden Analysen ist es, die postulierten positiven motivationalen – bezogen auf die Kompetenzüberzeugung und den intrinsischen Anreiz – und leistungsmässigen Effekte der Unterrichtsqualitätsmerkmale in den Fächern Mathematik und Deutsch zu prüfen.

Bemerkungen zur Ergebnispräsentation

Ein zentrales und für die weiteren Analysen «folgeschweres» Ergebnis in den Skalenanalysen (vgl. Abschnitt 9.2, Teil 2) war, dass sich die den erhobenen Unterrichtsmerkmalen unterstellte theoretische Struktur mit acht Unterrichtsmerkmalen in beiden Fächern zwar statistisch absichern liess, die korrelativen Beziehungen der acht Merkmale untereinander jedoch so hoch waren, dass es äusserst zweifelhaft erschien, inwieweit – wie beabsichtigt – verschiedene Unterrichtsmerkmale erhoben wurden. Viel eher scheint es so zu sein, dass sich die Schülerinnen und Schüler in ihrem gesamten Antwortverhalten von einem pauschalen, mehr oder weniger positiven «Eindruck» leiten liessen, was die Lehrperson tut oder unterlässt. Unter diesen Voraussetzungen ist es wenig sinnvoll, ja eher irreführend, Analysen auf der Basis der Einzelmerkmale durchzuführen. Für die definitiven Auswertungen wurden daher die acht Unterrichtsmerkmale pro Fach zu einer «Globalskala» *Unterrichtswahrnehmung* zusammengefasst.

Die nachfolgend dargestellten Ergebnisse basieren auf den gewichteten Daten der repräsentativen Klassenstichprobe 2009.

Um einen ersten Eindruck davon zu gewinnen, wie die Schülerinnen und Schüler den Unterricht am Ende ihrer Primarschulzeit (rückblickend) wahrnehmen bzw. beurteilen, wird die Unterrichtswahrnehmung zuerst jeweils in einem deskriptiven Sinne dargestellt. Für die deskriptive Ergebnisdarstellung wurde aus Gründen der Anschaulichkeit wie schon anlässlich der Lernstandserhebung am Ende der 3. Klasse das Antwortspektrum (1–4) in fünf gleich grosse Abschnitte unterteilt (negativ, eher negativ, ambivalent, eher positiv, positiv). Dargestellt werden die prozentualen Anteile der Schülerinnen und Schüler in den jeweiligen Abschnitten.

Es folgt dann die Präsentation der primär interessierenden Variablenzusammenhänge zwischen *Unterrichtswahrnehmung*, *Kompetenzüberzeugung*, *intrinsischem Anreiz* sowie *Leistungen* in Form von Pfadmodellen. Alle diesbezüglichen Analysen basieren auf den kontinuierlichen Skalenwerten 1 bis 4. Dargestellt und kommentiert werden lediglich die Zusammenhänge zwischen diesen vier Grössen. In allen Analysen als Kontrollvariablen – in den Abbildungen in grau bzw. aufgehellt dargestellt – berücksichtigt, inhaltlich jedoch nicht weiter von Interesse, wurden das Geschlecht, das Alter, die soziale Herkunft, Deutsch als Erst- oder Zweitsprache sowie die auf Klassenebene aggregierte soziale Herkunft sowie Deutsch als Erst- bzw. Zweitsprache.¹ Bei Letzterem handelt es sich um den Anteil von Schülerinnen und Schüler in einer Klasse mit Deutsch als Zweitsprache. Bei den hinsichtlich Unterrichtswahrnehmung, Kompetenzüberzeugung sowie intrinsischem Anreiz berichteten Effekten handelt es sich entsprechend jeweils um eigenständige Effekte dieser Variablen, unabhängig von Effekten anderer, in den Modellen berücksichtigter Grössen.

Sofern nichts anderes erwähnt, sind alle berichteten Zusammenhänge statistisch bedeutsam. In den Abbildungen sind – mit einer Ausnahme (vgl. unten) – hinsichtlich der interessierenden Zusammenhänge ausschliesslich die statistisch bedeutsamen Effekte dargestellt. Angegeben werden dazu jeweils die standardisierten Regressionskoeffizienten, anhand derer auch die Bedeutsamkeit der einzelnen Effekte untereinander verglichen werden kann. Der Wertebereich der standardisierten Regressionskoeffizienten liegt zwischen -1 und 1 . Je stärker der Wert von 0 verschieden ist, desto bedeutsamer ist der Effekt der entsprechenden Grösse hinsichtlich des Merkmals, welches es zu erklären gilt. Inhaltlich indizieren Regressionskoeffizienten mit einem positiven Vorzeichen einen Zusammenhang im Sinne von «je höher, desto höher bzw. je tiefer, desto tiefer», solche mit einem negativen einen im Sinne von «je höher, desto tiefer bzw. je tiefer, desto höher».

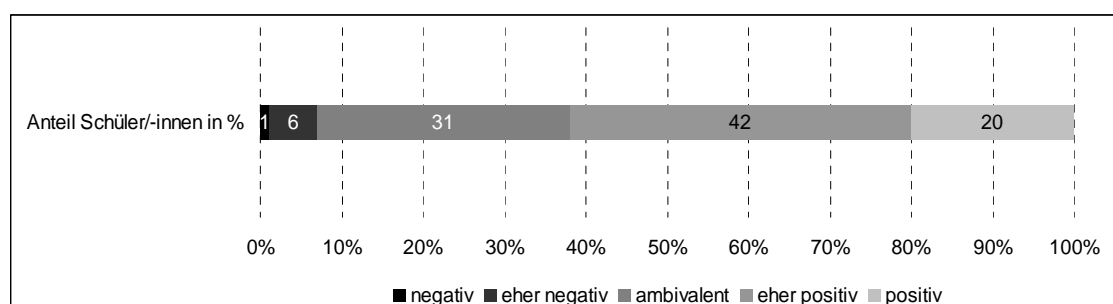
Nähere methodische Angaben (Stichprobe, Instrumente usw.) sowie eine detailliertere Darstellung der einzelnen Analyseergebnisse finden sich in Kapitel 9, Teil 2.

9.2 Unterricht, Motivation und Leistung in Mathematik

Unterrichtswahrnehmung in Mathematik

In Abbildung 9.1 wird deutlich, dass rund 60% der Schülerinnen und Schüler den Mathematikunterricht positiv oder eher positiv beurteilen.

Abbildung 9.1: Wahrnehmung des Unterrichts in Mathematik



Rund ein Drittel äussert sich ambivalent und nur sehr wenige – rund 7% – (eher) negativ. Es zeigt sich also ein erfreuliches Bild am Ende der Primarschulzeit: Eine Mehrheit der Schülerinnen und Schüler erachtet den Unterricht (rückblickend) als motivierend und unterstützend; man ist der Ansicht, dass die Lehrperson gut erkläre, Strukturierungshilfen gebe und merke, wenn jemand etwas nicht verstehe. Verhaltensregeln seien klar und die Beziehung zur Lehrperson sei gut.²

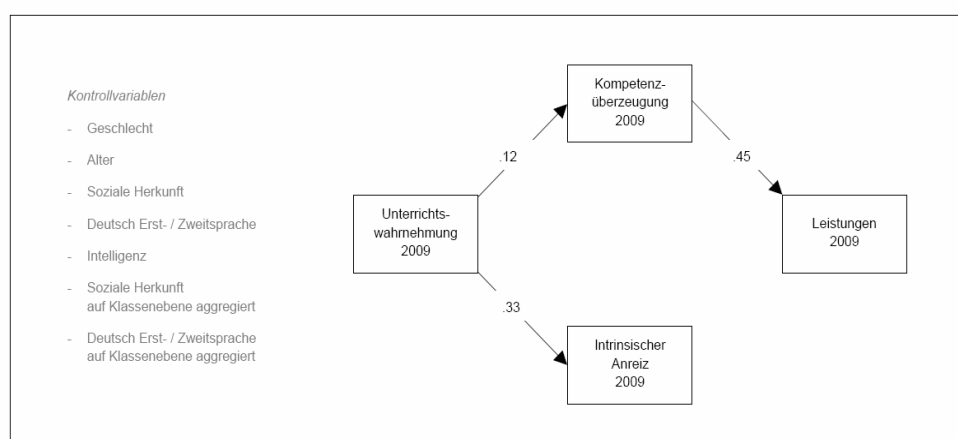
Unterrichtswahrnehmung, intrinsischer Anreiz, Kompetenzüberzeugung und Leistungen in Mathematik

Wie in Abbildung 9.2 dargestellt, erweist sich die Unterrichtswahrnehmung theoriekonform als bedeutsame Einflussgrösse hinsichtlich des intrinsischen Anreizes von Mathematik. Deutlich wird: Je positiver die Unterrichtswahrnehmung, desto grösser ist die Freude am Fach. Auch hinsichtlich der Kompetenzüberzeugung zeigt sich: Je günstiger die Unterrichtswahrnehmung, desto mehr trauen sich die Schülerinnen und Schüler zu.

Hinsichtlich der Leistungen zeitigt die Unterrichtswahrnehmung keine direkten Effekte. Gleiches gilt für den intrinsischen Anreiz. Ersichtlich wird, dass es die Kompetenzüberzeugung ist, die einen positiven erklärenden Effekt hinsichtlich der Leistungen hat; je mehr sich die Schülerinnen und Schüler zutrauen, desto besser sind ihre Leistungen.

Angesichts des fehlenden direkten Effekts der Unterrichtswahrnehmung hinsichtlich der Leistungen stellt sich die Frage, ob sich allenfalls ein indirekter positiver Effekt via die beiden Motivationskomponenten finden lässt. Als mögliche vermittelnde Variable kommt hier lediglich die Kompetenzüberzeugung in Frage, denn nur diese leistet einen Beitrag zur Erklärung der Leistungen. In einer entsprechenden Analyse erweist sich der indirekte Effekt als statistisch bedeutsam. Mit einer positiveren Unterrichtswahrnehmung gehen indirekt – vermittelt über eine günstigere Kompetenzüberzeugung – höhere Leistungen einher.

Abbildung 9.2: Unterricht, Motivation und Leistung in Mathematik

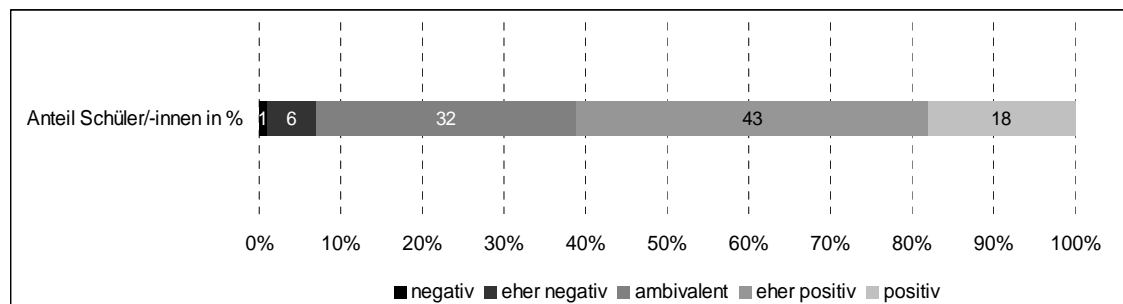


9.3 Unterricht, Motivation und Leistung in Deutsch

Unterrichtswahrnehmung in Deutsch

Auch in Deutsch (vgl. Abbildung 9.3) zeigt sich hinsichtlich der Unterrichtswahrnehmung mehr oder weniger das gleiche erfreuliche Bild wie in Mathematik. Rund 60% der Schülerinnen und Schüler beurteilen den Deutschunterricht (rückblickend) positiv oder eher positiv. Rund ein Drittel äussert sich ambivalent und wiederum nur sehr wenige – rund 7% – eher negativ oder negativ.³

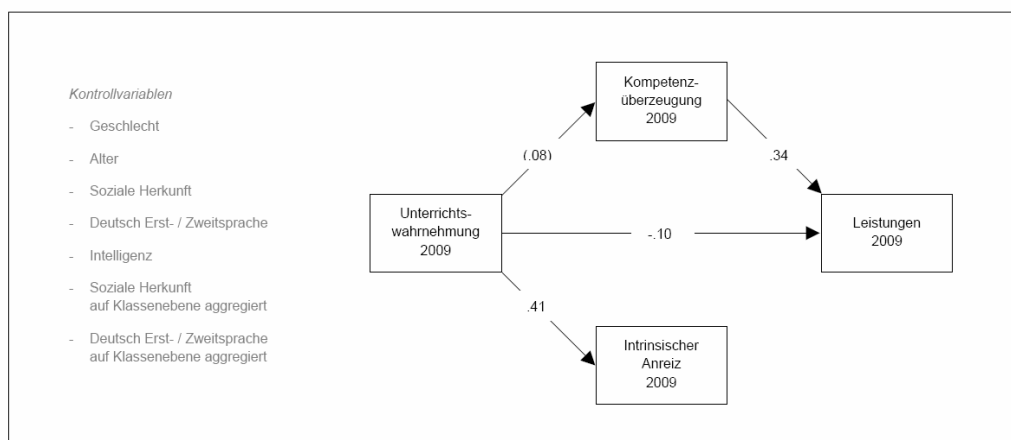
Abbildung 9.3: Wahrnehmung des Unterrichts in Deutsch



Unterrichtswahrnehmung, intrinsischer Anreiz, Kompetenzüberzeugung und Leistungen in Deutsch

Abbildung 9.4 zeigt die Ergebnisse hinsichtlich der interessierenden Variablenzusammenhänge in Deutsch. Wie schon in Mathematik erweist sich auch in Deutsch die Unterrichtswahrnehmung als bedeutsam hinsichtlich des intrinsischen Anreizes des Faches. Eine positivere Unterrichtswahrnehmung geht auch hier einher mit mehr Freude an Deutsch. Im Gegensatz zu Mathematik findet sich hinsichtlich der Kompetenzüberzeugung in Deutsch jedoch lediglich ein statistisch tendenziell bedeutsamer positiver Effekt der Unterrichtswahrnehmung.

Abbildung 9.4: Unterricht, Motivation und Leistung in Deutsch



Hinsichtlich der Leistungen zeigt sich einmal mehr: Je höher die Kompetenzüberzeugung, desto besser sind die Leistungen. Und auch hier bleibt – wie schon im Falle von Mathematik – der intrinsische Anreiz des Faches ohne statistisch bedeutsame leistungsmässige Effekte. Im Unterschied zu Mathematik manifestiert sich bezüglich der Leistungen in Deutsch nun auch ein direkter Effekt der Unterrichtswahrnehmung; eine günstigere Unterrichtswahrnehmung geht mit schlechteren Leistungen einher. Der negative Effekt der Unterrichtswahrnehmung überrascht, entspricht er doch nicht dem, was theoretisch zu erwarten wäre: ein positiver bzw. bei Berücksichtigung der beiden Motivationskomponenten eher gar kein Effekt. Auf dieses etwas irritierende Ergebnis wird in Abschnitt 9.4 noch näher eingegangen.

Hinsichtlich eines allfälligen indirekten Einflusses der Unterrichtswahrnehmung auf die Leistungen käme als vermittelnde Motivationskomponente wiederum nur die Kompetenzüberzeugung in Frage. Angesichts des lediglich tendenziellen positiven Zusammenhangs zwischen Unterrichtswahrnehmung und Kompetenzüberzeugung in Deutsch zeigt sich denn auch lediglich ein tendenziell bedeutsamer indirekter Effekt, der jedoch in die gleiche Richtung weist wie in Mathematik. Eine positivere Unterrichtswahrnehmung in Deutsch zeitigt tendenziell – vermittelt über eine höhere Kompetenzüberzeugung – positive leistungsmässige Effekte.

9.4 Fazit

Wie sich zeigt, wird der Unterricht in globalem Sinne von einer Mehrheit der Schülerinnen und Schüler in beiden Fächern am Ende der sechsten Klasse eher positiv oder positiv beurteilt. Der Mathematik- wie auch der Deutschunterricht werden als motivierend und unterstützend erachtet, man ist der Ansicht, dass die Lehrperson gut erkläre, Strukturierungshilfen gebe, merke, wenn jemand etwas nicht verstehe, dass die Verhaltensregeln klar seien und eine gute Beziehung zur Lehrperson vorhanden sei. Ein ähnlich positives Ergebnis zeigt sich Ende der sechsten Klasse in beiden Fächern auch bezogen auf die Kompetenzüberzeugung und den intrinsischen Anreiz. Bei jeweils mehr als der Hälfte der Schülerinnen und Schüler sind das leistungsmässige Selbstvertrauen und die Freude am Fach noch immer eher hoch oder hoch. Nicht nur, aber insbesondere auch den Lehrpersonen scheint es trotz des Übertrittsprozederes in die Sekundarstufe I, das für etliche Schülerinnen und Schüler nicht unbedingt das gewünschte Ergebnis zeitigt und mit Enttäuschungen verbunden sein kann, zu gelingen, dass eine Mehrheit der Schülerinnen und Schüler nach wie vor an ihr Leistungsvermögen in den beiden Fächern glaubt und die Freude daran nicht verloren hat.

Hinsichtlich der Kompetenzüberzeugung und des intrinsischen Anreizes bestätigen sich die postulierten positiven Effekte der Unterrichtswahrnehmung in Mathematik und Deutsch grösstenteils. Je positiver die Unterrichtswahrnehmung, desto grösser ist die Freude am Fach und desto höher die Kompetenzüberzeugung. Letzteres ist in Deutsch lediglich tendenziell der Fall. Aufgrund der Ergebnisse könnte die Unterrichtsgestaltung bzw. deren Wahrnehmung durch die Schülerinnen und Schüler auch dazu beitragen, dass diesen in beiden Fächern das Selbstvertrauen wie auch die Freude nicht abhanden kommen.

Bezogen auf die Leistungen erweist sich in Mathematik und Deutsch übereinstimmend die Kompetenzüberzeugung als wichtige Einflussgrösse: Je mehr sich die Schülerinnen und Schüler selbst zutrauen, desto besser sind ihre Leistungen. Die Freude am Fach ist bei gleichzeitiger Berücksichtigung der Kompetenzüberzeugung in den Analysen als eigenständiger Ein-

flussfaktor irrelevant. Dies ist nicht ganz überraschend. Wenn Leistungen (Testleistungen oder Noten) erklärt werden sollen, sind bei simultaner Berücksichtigung beider Motivationskomponenten meist lediglich die Kompetenzüberzeugungen oder andere kompetenzbezogene Überzeugungen als erklärende Faktoren bedeutsam. Dies heisst jedoch nicht, dass die Freude an einem Fach hinsichtlich leistungsmässigem Verhalten – verstanden in einem umfassenderen Sinne – gänzlich belanglos wären. Geht es etwa bspw. um das Engagement in einem Fach oder um Wahlverhalten hinsichtlich Leistungskursen, Fächern, Studien- und Berufsrichtungen, so spielen die Freude oder das Interesse sehr wohl eine wichtige Rolle (vgl. etwa Buff et al. 2010a; Eccles, 2005; Fredricks & Eccles, 2002; Eccles, Barber, Updegraff & O'Brien, 1998).

Die Bedeutung der Kompetenzüberzeugungen im Zusammenhang mit schulischen Leistungen wird dadurch unterstrichen, dass es auch diese sind, über welche sich die Unterrichtswahrnehmung auf indirekte Weise positiv in den Leistungen niederschlägt. In Mathematik ist dieser indirekte Effekt statistisch bedeutsam, in Deutsch ist dies lediglich tendenziell der Fall. Die Richtung des Einflusses ist jedoch in beiden Fällen gleich: Mit einer positiveren Unterrichtswahrnehmung gehen indirekt – vermittelt über eine günstigere Kompetenzüberzeugung – höhere Leistungen einher.

Direkte Effekte der Unterrichtswahrnehmung hinsichtlich der Leistungen sind in Mathematik bei gleichzeitiger Berücksichtigung von Kompetenzüberzeugung und intrinsischem Anreiz keine vorhanden. Bezogen auf die Leistungen in Deutsch zeigt sich indes unerwartet ein statistisch bedeutsamer negativer Effekt: Je positiver die Unterrichtswahrnehmung, desto schlechter waren die Leistungen. Ein ähnlich «irritierendes» Ergebnis (ebenfalls im Fach Deutsch) fand sich bereits in der Lernstandserhebung am Ende der 3. Klasse. Es zeigte sich dort: Je mehr die Lehrpersonen angaben, leistungsdifferenzierende Massnahmen in ihrem Unterricht zu realisieren, desto schlechter waren die Leistungen der Schülerinnen und Schüler (vgl. Buff, 2008b, S. 223f.). Ein möglicher und durchaus plausibler Grund für das Ergebnis wurde damals in der Anlage der Studie gesehen. Es handelte sich – wie auch im vorliegenden Falle – um Querschnittsdaten; d.h., alle in den Zusammenhängen interessierenden Grössen wurden zum gleichen Zeitpunkt erfasst. Was unter solchen Bedingungen letztlich «Ursache» und was «Wirkung» ist, lässt sich bei Zusammenhängen zwischen Variablen, wie sie auch im vorliegenden Falle interessieren (Leistungen, Unterrichtswahrnehmung, Kompetenzüberzeugung usw.), nicht klären. Die in den einzelnen Analysen unterstellten «Wirkrichtungen» sind lediglich theoretisch begründbar, mit Querschnittsdaten jedoch keinesfalls «belegbar». Wie schon anlässlich der letzten Lernstandserhebung dargelegt (vgl. Buff, 2008b, S. 224), ist davon auszugehen, dass zwischen Unterricht und Leistungen über die Zeit hinweg Wechselwirkungen existieren. Unterricht ist also nicht nur als Ursache von Leistung, sondern auch als Folge davon zu sehen. Denkbar ist, dass sich bei schlechteren Leistungen die Lehrpersonen intensiver bemühen, den Unterricht entsprechend der erhobenen Qualitätsmerkmale zu gestalten, und die Schülerinnen und Schüler dies auch wahrnahmen und berichteten. Bis zum Erhebungszeitpunkt hatten sich jedoch (noch) keine günstigen leistungsmässigen Effekte eingestellt. Die Datenlage lässt eine akkurate Prüfung dieser Vermutung allerdings nicht zu.

9.5 Anmerkungen

- ¹ Die auf Klassenebene aggregierte Unterrichtswahrnehmung erwies sich als Prädiktor in allen Analysen als irrelevant und wurde aus Gründen der Sparsamkeit in die definitiven Analysemodellen nicht einbezogen.
- ² Hinsichtlich der beiden Motivationskomponenten Kompetenzüberzeugung und intrinsischer Anreiz zeigt sich in Mathematik (wiederum das Antwortspektrum 1–4 in fünf gleich grosse Abschnitte unterteilt: hoch, eher hoch, ambivalent, eher tief, tief), dass rund 52% der Schülerinnen und Schüler über eine (eher) hohe Kompetenzüberzeugung verfügen, 25% liegen im ambivalenten Bereich und 23% berichten eine (eher) niedrige Kompetenzüberzeugung. Im Falle des intrinsischen Anreizes betragen die Werte: 54% eher hoch und hoch, 20% ambivalent und 26% eher tief und tief.
- ³ Hinsichtlich der beiden Motivationskomponenten Kompetenzüberzeugung und intrinsischer Anreiz (wiederum das Antwortspektrum 1–4 in fünf gleich grosse Abschnitte unterteilt) betragen die Werte im Fach Deutsch im Falle der Kompetenzüberzeugung 64% in den Bereichen eher hoch und hoch, 21% im ambivalenten Bereich und 15% in den Bereichen eher tief und tief. Hinsichtlich des intrinsischen Anreizes des Faches ergeben sich folgende Werte: 53% eher hoch und hoch, 23% ambivalent und 24% eher tief und tief.

Teil 2

Analysen zu den Ergebnissen

1 Die Zürcher Lernstandserhebung im Überblick

Domenico Angelone & Urs Moser

1.1 Grundgesamtheit und Stichprobe

Längsschnittstichprobe 1. bis 6. Klasse

Die Grundgesamtheit für die Stichprobe der Zürcher Längsschnittstudie bilden alle Schülerinnen und Schüler, die im Sommer 2003 im Kanton Zürich in die Regelklasse (ohne Mehrklassenabteilungen) eingetreten sind. Die Stichprobe wurde so gebildet, dass repräsentative Aussagen (1) für die Regelklassen des Kantons Zürich, (2) für die Klassen des Projekts «Teilautonome Volksschule» (TaV), (3) für die Klassen des Programms «Qualität in multikulturellen Schulen» (QUIMS) und (4) für die Kleinklassen A beziehungsweise für die Einschulungsklassen möglich sind.

Dementsprechend wurde die Grundgesamtheit in vier Subpopulationen (Strata) eingeteilt. Innerhalb jeder Subpopulation wurde eine bestimmte Anzahl Klassen zufällig ausgewählt. Von den ausgewählten Klassen wurden sämtliche Schülerinnen und Schüler einbezogen. Die unterschiedliche Auswahlwahrscheinlichkeit der Schülerinnen und Schüler der vier Subpopulationen wurde bei der Berechnung der Ergebnisse für den Kanton Zürich berücksichtigt (Gewichtung der Daten).

Tabelle 1.1 gibt einen Überblick über die Grundgesamtheit und die Längsschnittstichprobe der Zürcher Lernstandserhebung nach den vier Strata. Im Sommer 2003 sind im Kanton Zürich 11118 Schülerinnen und Schüler in insgesamt 650 Klassen in die Primarschule eingetreten (ohne Mehrklassenabteilungen). Die Längsschnittstichprobe umfasst 2046 Schülerinnen und Schüler aus insgesamt 120 Klassen.

Tabelle 1.1: Grundgesamtheit und Längsschnittstichprobe der Zürcher Lernstandserhebung

	Grundgesamtheit Schuljahr 2003/04 ¹⁾		Längsschnittstichprobe der Lernstandserhebung	
	Anzahl Schülerinnen und Schüler	Anzahl Klassen	Anzahl Schülerinnen und Schüler	Anzahl Klassen
Regelklassen	6226	316	815	40
QUIMS	562	30	517	30
TaV	3164	159	609	30
Kleinklasse A	1166	145	105	20
Total	11118	650	2046	120

Anmerkungen: ¹⁾Schuljahr 2003/04. Berücksichtigt wurden nur Schülerinnen und Schüler, die nicht in Mehrklassenabteilungen unterrichtet wurden.

Bei einer Längsschnittuntersuchung muss aus verschiedenen Gründen mit einem Stichprobenausfall gerechnet werden. Ein Teil der Schülerinnen und Schüler zieht während der Schulzeit aus dem Kanton Zürich weg und wechselt die Schule. Ein Teil wechselt in eine Sonderklasse oder Sonderschule und einige Schülerinnen und Schüler können deshalb nicht einbezogen werden, weil ihre Lehrpersonen oder Eltern eine Teilnahme verweigern.

Tabelle 1.2 zeigt die Anzahl Schülerinnen und Schüler der Stichproben zu den drei Testzeitpunkten. Im Jahr 2003 zu Beginn der 1. Klasse (Testzeitpunkt T1) wurden 2046 Schülerinnen und Schüler einbezogen. Davon besuchten 1941 (94.9 Prozent) Kinder eine Regelklasse und 105 Kinder (5.1%) eine Kleinklasse A.

Tabelle 1.2: Stichprobenausfall zwischen 2003, 2006 und 2009

	T1			T2		T3	T1 bis T3
	Stichprobe 2003	weggezogen	ausgeschlossen	Stichprobe 2006	weggezogen	Stichprobe 2009	Stichprobenausfall
Total	2046	81	5	1960	157	1803	11.9%
Regelklasse	1941			1868		1754	
Kleinklasse	105			92		40	
Sekundarstufe I						9	

Nach drei Jahren, am Ende der 3. Klasse (Testzeitpunkt T2), hatte sich die ursprüngliche Stichprobe um 86 Fälle auf insgesamt 1960 Schülerinnen und Schüler reduziert. 81 Kinder besuchten im Jahr 2006 die Schule nicht mehr im Kanton Zürich und 5 Kinder wurden aus der Stichprobe ausgeschlossen, weil sie nach vier Schuljahren und zweimaliger Repetition die 2. Klasse besuchten. Von den 1960 Schülerinnen und Schülern der Stichprobe zum Testzeitpunkt T2 besuchten 1868 Kinder (95.3%) eine Regelklasse und 92 Kinder (4.7%) eine Kleinklasse.

Wesentlich grösser ist der Stichprobenausfall nach weiteren drei Jahren ausgefallen. Am Ende der 6. Klasse im Jahr 2009 (Testzeitpunkt T3) hatte sich die Stichprobe um weitere 157 Kinder, die zu diesem Zeitpunkt die Schule nicht mehr im Kanton Zürich besuchten, auf insgesamt 1803 Schülerinnen und Schüler reduziert. Davon wurden 1754 Kinder (97.3%) in einer Regelklasse und 40 Kinder (2.3%) in einer Kleinklasse unterrichtet. 9 Kinder besuchten im Jahr 2009 die 1. Klasse der Sekundarstufe I. Insgesamt hat sich die Ausgangsstichprobe somit um 243 Schülerinnen und Schüler reduziert, was einem Stichprobenausfall von rund 12 Prozent entspricht.

Von den 1803 ausgewählten Schülerinnen und Schülern für den Testzeitpunkt T3 am Ende der 6. Klasse haben 124 Kinder (Krankheit, Verweigerung usw.) nicht an der Erhebung teilgenommen. Diese entspricht einer Rücklaufquote zum Testzeitpunkt T3 von 93.1 Prozent. Für Analysen, die sich auf den gesamten Untersuchungszeitraum der Längsschnittstichprobe be-

ziehen, das heisst vom Schuleintritt bis zum Ende der 6. Klasse, können somit 82.1 Prozent (N = 1679) aller Kinder der Ausgangsstichprobe (N = 2046) berücksichtigt werden.

Schullaufbahn bis zum Ende der 6. Klasse

In Tabelle 1.3 sind die unterschiedlichen Schullaufbahnen aller Schülerinnen und Schüler zusammengefasst, die nach sechs Schuljahren und drei Testzeitpunkten in der Stichprobe der Zürcher Lernstandserhebung verblieben sind. Die Angaben geben somit keine Auskunft über die Schullaufbahnen aller Kinder, die im Schuljahr 2003/04 in die 1. Klasse des Kantons Zürich eingetreten sind. Von den Schülerinnen und Schülern, die am Ende der 6. Klasse nicht mehr an der Längsschnittstudie teilnahmen, waren überdurchschnittlich viele ein Schuljahr zurück (also in der 5. Klasse). Der Stichprobenausfall ist nicht zufällig, sondern konzentriert sich auf Schülerinnen und Schüler, deren Schullaufbahn verzögert verläuft.

Die erste Spalte zeigt, ob der Schuleintritt im Jahr 2003 in eine Regelklasse oder in eine Kleinklasse A erfolgte. Die Schattierungen in den Spalten 2 bis 8 zeigen, welche Klassen die Schülerinnen und Schüler zum Testzeitpunkt T3 im Mai 2009 besuchten. Die Spalte 9 informiert darüber, ob die Schülerinnen und Schüler im Sommer 2009 eine Regelklasse oder eine Kleinklasse besuchten. Die Spalten 10 und 11 zeigen die Anzahl Schülerinnen und Schüler der Stichprobe beziehungsweise die prozentualen Stichprobenanteile nach verschiedenen Schullaufbahnen.

Tabelle 1.3: Schullaufbahn nach 6 Schuljahren

Zeitpunkt T1 2003	1. Kl.	2. Kl.	3. Kl.	4. Kl.	5. Kl.	6. Kl.	Sek.	Zeitpunkt T3 2009	Anzahl	Anteil
Regelklasse								Regelklasse	9	0.5%
Regelklasse								Regelklasse	1545	85.7%
Regelklasse								Regelklasse	138	7.7%
Regelklasse								Regelklasse	1	0.1%
Regelklasse								Kleinklasse	19	1.1%
Regelklasse								Kleinklasse	12	0.7%
Regelklasse								Kleinklasse	1	0.1%
Kleinklasse A								Regelklasse	3	0.2%
Kleinklasse A								Regelklasse	65	3.6%
Kleinklasse A								Regelklasse	2	0.1%
Kleinklasse A								Kleinklasse	8	0.4%
Total									1803	100.0%

Anmerkungen: Berücksichtigt wurden nur Schülerinnen und Schüler, die nach sechs Schuljahren in der Längsschnittstichprobe verblieben sind. Die Angaben zur besuchten Klasse stammen aus der offiziellen Bildungsstatistik.

9 Schülerinnen und Schüler (0.5%) der Stichprobe sind bei Schuleintritt im Jahr 2003 in die 1. Klasse der Primarschule eingetreten und besuchten nach sechs Schuljahren bereits die 1. Klasse der Sekundarstufe I (Akzeleration). Die grosse Mehrheit der Stichprobe, 1545 Schülerinnen und Schüler (85.7%), sind zu Beginn ihrer Schullaufbahn in eine Regelklasse eingetreten und besuchten nach sechs Jahren die 6. Regelklasse. 138 Schülerinnen und Schüler (7.7 %) sind im Jahr 2003 in eine Regelklasse eingetreten und besuchten sechs Jahre später nach einmaliger Repetition die 5. Regelklasse. 1 Schülerin beziehungsweise 1 Schüler (0.1%) wurde in eine Regelklasse eingeschult und besuchte im Jahr 2009 nach zweimaliger Repetition die 4. Klasse (Regelklasse).

32 Schülerinnen und Schüler wurden zu Beginn ihrer Schullaufbahn in eine Regelklasse eingeschult und besuchten sechs Schuljahre später eine Kleinklasse. Davon wurden 19 Kinder in der 6. Klasse, 12 Kinder in der 5. Klasse (eine Repetition) und 1 Kind in der 4. Kleinklasse (zwei Repetitionen) unterrichtet.

78 Schülerinnen und Schüler wurden in eine Kleinklasse A eingeschult. Davon sind 70 Kinder bis im Jahr 2009 in eine Regelklasse übergetreten. 3 Kinder haben kein Jahr verloren und besuchten die 6. Klasse, 65 Kinder besuchten die 5. Klasse und 2 Kinder die 4. Klasse. 8 Kinder wurden in eine Kleinklasse A eingeschult und wurden nach sechs Schuljahren in einer Kleinklasse unterrichtet.

Klassenstichprobe 6. Klassen

Die Längsschnittstichprobe ermöglicht keine repräsentativen Aussagen zu den Schülerinnen und Schüler beziehungsweise Klassen der 6. Klassenstufe im Kanton Zürich. Zum einen ist die Längsschnittstichprobe mit einem Stichprobenausfall behaftet und zum anderen verteilen sich die Schülerinnen und Schüler der Längsschnittstichprobe nach sechs Jahren nicht mehr auf 120 Klassen, wie beim Schuleintritt, sondern auf 334 Klassen. Deshalb wurde die Zürcher Lernstandserhebung um eine zusätzliche, repräsentative Stichprobe der 6. Klassen im Schuljahr 2008/09 des Kantons Zürich ergänzt.

Die Grundgesamtheit der repräsentativen Klassenstichprobe bilden alle 6. Klassen des Kantons Zürich im Jahr 2009, ohne Mehrklassenabteilungen und ohne Kleinklassen. Die Stichprobe wurde nach einem wissenschaftlichen Auswahlverfahren so gebildet, dass repräsentative Aussagen zu drei verschiedenen Schulungsprogrammen der Regelklassen möglich sind: (1) 6. Klassen mit einem geringen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, (2) 6. Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, die von QUIMS (Qualität in multikulturellen Schulen) bereits vor 2006 unterstützt werden und (3) 6. Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, die von QUIMS nach 2008 unterstützt werden.

Dementsprechend wurde die Grundgesamtheit in drei Strata (Subpopulationen) eingeteilt. Innerhalb des ersten Stratums, den 6. Klassen des Kantons Zürich mit einem geringen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache, wurden 50 Klassen zufällig ausgewählt. Das zweite Stratum umfasst alle 6. Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache aus Schulen, die seit dem Schuljahr 2003/04 von QUIMS unterstützt werden ($n = 26$). Das dritte Stratum umfasst alle 6. Klassen mit einem hohen Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache aus Schulen, die erst seit Sommer 2008 von QUIMS unterstützt werden. Von den ausgewählten Klassen wur-

den jeweils sämtliche Schülerinnen und Schüler einbezogen. Die unterschiedliche Auswahlwahrscheinlichkeit der Klassen in den drei Subpopulationen wurde bei der Berechnung der Ergebnisse für den Kanton Zürich berücksichtigt (Gewichtung der Daten).

Tabelle 1.4 zeigt die Anzahl ausgewählter und teilnehmender Klassen beziehungsweise Schülerinnen und Schüler der Klassenstichprobe. Die Stichprobe umfasst 100 Klassen mit insgesamt 2093 Schülerinnen und Schülern. Zwei Regelklassen verweigerten die Teilnahme an der Lernstandserhebung. Von den ausgewählten Schülerinnen und Schülern haben 112 Kinder (Krankheit, Verweigerung usw.) nicht an der Erhebung teilgenommen. Für die Analysen verbleiben somit 98 Klassen mit insgesamt 1981 Schülerinnen und Schülern. Dies entspricht einer Rücklaufquote von 94.6 Prozent.

Tabelle 1.4: Repräsentative Stichprobe der 6. Klassen: Schuljahr 2008/09

Klassen		Schülerinnen und Schüler		
ausgewählt	teilgenommen	ausgewählt	teilgenommen	Rücklaufquote
100	98	2093	1981	94.6%

Tabelle 1.5 zeigt den Vergleich der Anzahl Schülerinnen und Schüler der Grundgesamtheit mit derjenigen der gewichteten Klassenstichprobe. Bei der gewichteten Stichprobe werden für die Hochrechnung der Ergebnisse auf die Grundgesamtheit die Auswahlwahrscheinlichkeiten der Klassen sowie der Stichprobenausfall berücksichtigt.

Tabelle 1.5: Grundgesamtheit und Stichprobe der 6. Klassen: Schuljahr 2009

Grundgesamtheit		Stichprobe		
Klassen	Schülerinnen und Schüler	ausgewählt	teilgenommen	gewichtet
503	10758	2093	1981	10387

Anmerkungen: Die Angaben zur Grundgesamtheit stammen von der BISTA aus dem Schuljahr 2008/09.

Im Schuljahr 2008/09 besuchten im Kanton Zürich 10758 Schülerinnen und Schüler eine 6. Regelklasse (ohne Mehrklassenabteilungen und Sonderklassen). Die Anzahl Schülerinnen und Schüler der gewichteten Stichprobe beträgt 10387 Schülerinnen und Schüler. Damit wird die Grundgesamtheit durch die Stichprobe relativ genau abgebildet. Abweichungen entstehen dadurch, dass die Liste der Bildungsstatistik, auf deren Grundlage die Stichprobe gezogen wurde, aufgrund von Mutationen während des laufenden Schuljahrs nicht dem aktuellsten Stand entspricht.

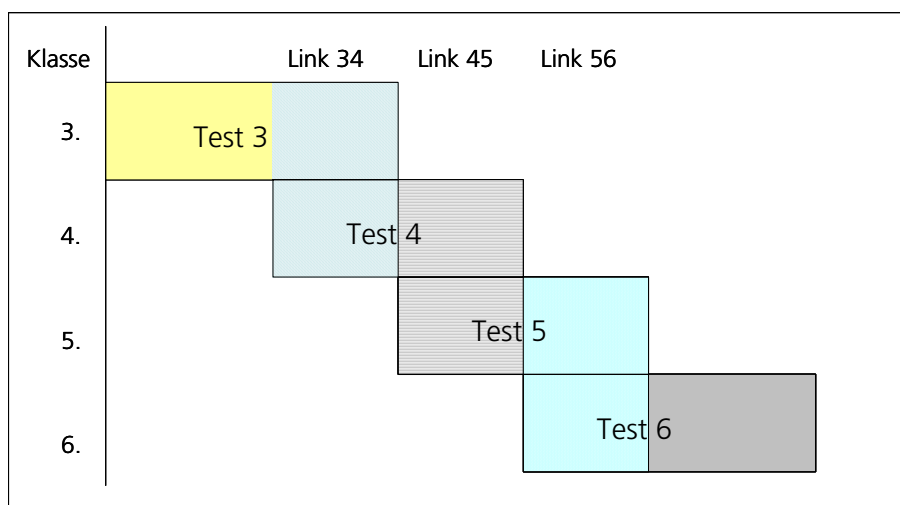
1.2 Leistungsskalen

Leistungszuwachs von der 1. bis zur 6. Klasse

Während die Lernstandserhebungen zu Beginn der 1. Klasse und am Ende der 3. Klasse aufgrund des grossen Leistungszuwachs der Kinder in den ersten drei Schuljahren mit vollständig verschiedenen Testinstrumenten durchgeführt wurden, konnten die Lernstandserhebungen am Ende der 3. und der 6. Klasse durch eine Skalierungstechnik miteinander verbunden werden. Dadurch kann der Leistungszuwachs von der 3. bis zur 6. Klasse auf einer einheitlichen Kompetenzskala entsprechend den Zielen des Lehrplans des Kantons Zürich ausgewiesen werden.

Die Verbindung der Leistungstests von der 3. bis zur 6. Klasse verlangt ein spezielles methodisches Vorgehen auf der Basis der Item Response Theorie, das in Abbildung 1.1 skizziert ist. In einer besonderen Studie ein Jahr vor der Lernstandserhebung am Ende der 6. Klasse wurden dazu sich überlappende Leistungstests bei Schülerinnen und Schülern der 3., 4., 5. und 6. Klasse eingesetzt. Die verschiedenen Leistungstests wurden durch gleiche Aufgaben – sogenannte Link-Aufgaben – miteinander verbunden. Dieses Verfahren ermöglicht es, den Leistungszuwachs zwischen der 3. und 6. Klasse inhaltlich zu beschreiben und auf einer einheitlichen Kompetenzskala abzubilden.

Abbildung 1.1: Verbindung von Tests verschiedener Klassenstufen



Durch die Verbindung der Tests anhand der Link-Aufgaben kann der Leistungszuwachs quantifiziert werden. Die Ergebnisse der Lernstandserhebung zu Beginn der 1. Klasse wurden auf den Mittelwert $M = 500$ und die Standardabweichung $SD = 100$ normiert. Weil die Leistungstests zu Beginn der 1. Klasse und am Ende der 3. Klasse nicht mit Link-Aufgaben miteinander verknüpft wurden, lassen sich lediglich die relativen Positionen der Schülerinnen und Schüler zu den beiden Testzeitpunkten miteinander vergleichen. Damit trotzdem eine Interpretation im Sinne eines Leistungszuwachses während den ersten drei Schuljahren möglich ist, wurde der Mittelwert der Leistungstests am Ende der 3. Klasse auf 650 Punkte festgelegt. Damit wird ein durchschnittlicher Leistungszuwachs in den ersten drei Schuljahren von 150 Punkten

angenommen. Ein Anstieg von rund 50 Punkten pro Schuljahr entspricht den bisherigen Erfahrungswerten bei Schulleistungsstudien, bei denen ein ähnliches Skalierungsverfahren angewendet wurde. Die Annahme des Leistungszuwachses von durchschnittlich 150 Punkten in den ersten drei Schuljahren der Primarschule stellt die bestmögliche Annäherung für den tatsächlichen Leistungszuwachs der Schülerinnen und Schüler dar.

Nach weiteren drei Jahren, am Ende der 6. Klasse, konnten die Mittelwerte für Deutsch und Mathematik aufgrund der einheitlichen Skala berechnet werden. Die Mittelwerte sind je nach Fachbereich unterschiedlich hoch. Sie hängen aber nicht nur vom eigentlichen Leistungszuwachs der Schülerinnen und Schüler ab, sondern auch von anderen Faktoren wie der Diskrimination der Tests.

Leistungstest am Ende der 6. Klasse

Die Tabellen 1.6 und 1.7 geben einen Überblick über die Leistungstests in Deutsch und in Mathematik, die zur Erfassung der Leistungen am Ende der 6. Klasse eingesetzt wurden. In der ersten Spalte sind die verschiedenen Teilbereiche der Tests angegeben, in der zweiten Spalte die Anzahl Aufgaben, in der dritten Spalte die mittlere Schwierigkeit der Aufgaben und in der vierten Spalte die mittlere Trennschärfe (punktbiseriale Korrelation zwischen dem Erfolg bei einer Aufgabe und dem Testergebnis).

Tabelle 1.6: Aufgaben für den Deutschtest in der 6. Klasse

Lehrplanbereich	Anzahl Testaufgaben	mittlere Schwierigkeit	mittlere Trennschärfe
Leseverständnis	89	65.3%	.34
Wortschatz	93	68.2%	.30
Sprachbetrachtung	99	67.5%	.35
Texte überarbeiten	45	38.0%	.46
Total	326	59.8%	.36

Für den Deutschtest am Ende der 6. Klasse wurden insgesamt 326 Testaufgaben eingesetzt. Der Lehrplanbereich «Leseverständnis» umfasst 89 Aufgaben, der Bereich «Wortschatz» 93 Aufgaben, der Bereich «Sprachbetrachtung» 99 Aufgaben und der Bereich «Texte überarbeiten» 45 Aufgaben. Von den Bereichen «Leseverständnis», «Wortschatz» und «Sprachbetrachtung» wurden im Durchschnitt rund zwei Drittel der Aufgaben richtig gelöst, vom Bereich «Texte überarbeiten» nur 38 Prozent. Die durchschnittliche Trennschärfe der Testaufgaben ist im Bereich «Texte überarbeiten» mit $r = .46$ am höchsten.

Tabelle 1.7: Aufgaben für den Mathematiktest in der 6. Klasse

Lehrplanbereich	Anzahl Testaufgaben	mittlere Schwierigkeit	mittlere Trennschärfe
Arithmetik	41	72.1%	.38
Größen/Sachrechnen/Proportionen	37	43.9%	.51
Brüche	38	64.7%	.55
Problemlösen	27	53.0%	.43
Geometrie	36	61.1%	.35
Total	179	59.0%	.44

Für den Mathematiktest am Ende der 6. Klasse wurden insgesamt 179 Aufgaben eingesetzt. Der Lehrplanbereich «Arithmetik» umfasst 41 Aufgaben, der Bereich «Größen/Sachrechnen/Proportionen» 37 Aufgaben, der Bereich «Brüche» 38 Aufgaben, der Bereich «Problemlösen» 27 Aufgaben und der Bereich «Geometrie» 36 Aufgaben. Am besten wurden die Testaufgaben zum Bereich «Arithmetik» gelöst. Im Durchschnitt konnten die Schülerinnen und Schüler 72 Prozent der Aufgaben richtig lösen. Die mittlere Trennschärfe ist mit $r = .55$ im Bereich «Brüche» am höchsten und mit $r = .38$ im Bereich «Arithmetik» am geringsten. Die Trennung zwischen leistungsstarken und leistungsschwachen Schülerinnen und Schüler wird in der Mathematik am besten aufgrund der Aufgaben mit Brüchen erreicht.

1.3 Soziale Herkunft, Erstsprache und kognitive Grundfähigkeiten

Die soziale Herkunft und die Erstsprache der Schülerinnen und Schüler wurde je nach verwendeter Datenbasis – Längsschnittstichprobe 1. bis 6. Klassen oder repräsentative Stichprobe der 6. Klassen – unterschiedlich erfasst und operationalisiert. Die Indikatoren der Längsschnittstichprobe stammen aus dem Elternfragebogen zum Testzeitpunkt T2, jene der repräsentativen Stichprobe der 6. Klassen aus dem Schülerfragebogen zum Testzeitpunkt T3.

Längsschnittstichprobe 1. bis 6. Klasse

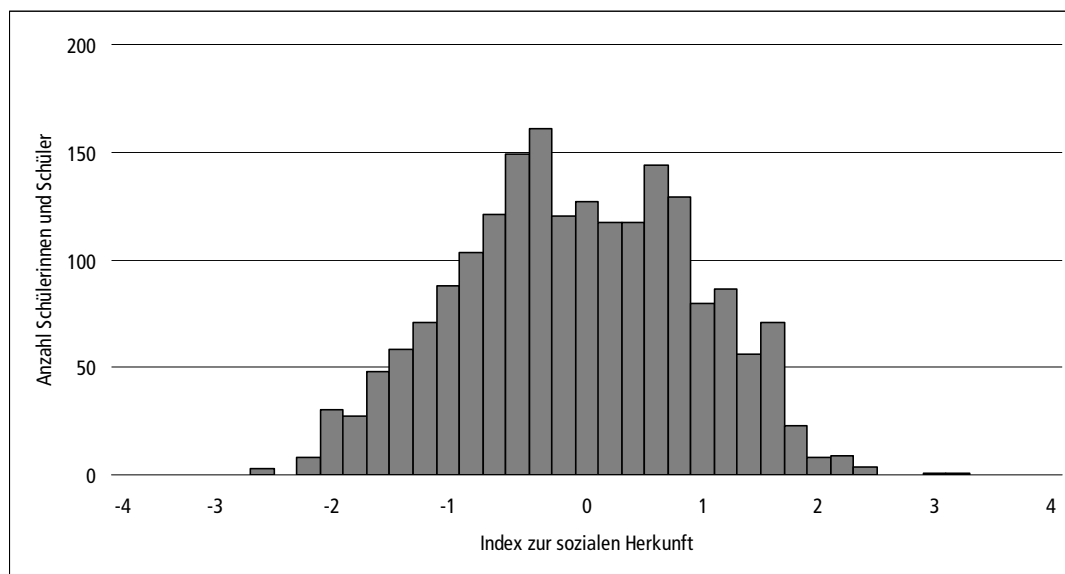
Soziale Herkunft der Schülerinnen und Schüler

Zur Bestimmung der sozialen Herkunft der Schülerinnen und Schüler wurde ein Index gebildet, der Indikatoren zu den ökonomischen und bildungsrelevanten Ressourcen des Elternhauses berücksichtigt. Als Indikator für die ökonomischen Ressourcen des Elternhauses wurde die Anzahl Zimmer der Wohnung oder des Hauses pro darin lebende Person berücksichtigt. Die bildungsrelevanten Ressourcen wurden anhand der Anzahl Bücher im Haushalt und dem höchsten Bildungsabschluss des Elternhauses erfasst. Aus diesen Angaben wurde ein additiver Index zur sozialen Herkunft der Schülerinnen und Schüler gebildet. Die Angaben zur Bildung des Indexes stammen alle aus dem Elternfragebogen zum Testzeitpunkt T2.

Abbildung 1.2 zeigt die Verteilung des Indexes zur sozialen Herkunft. Auf der x-Achse ist die soziale Herkunft der Schülerinnen und Schüler in Form von Indexpunkten beziehungsweise

Standardabweichungen abgetragen, auf der y-Achse die entsprechenden Häufigkeiten. Der Index zur sozialen Herkunft weist einen Mittelwert von 0 und eine Standardabweichung von 1. Die Streubreite reicht von -2.77 bis 3.51 Indexpunkten. Der Index ist annähernd normalverteilt. Deshalb liegt im Bereich zwischen -1 und 1 Standardabweichungen die soziale Herkunft von rund zwei Dritteln aller Schülerinnen und Schüler.

Abbildung 1.2: Verteilung des Indexes zur sozialen Herkunft der Schülerinnen und Schüler



Für Analysen, bei denen die soziale Herkunft der Schülerinnen und Schüler nicht als stetige sondern als kategoriale Variable verwendet wurde, wurden die Schülerinnen und Schüler aufgrund ihrer sozialen Herkunft in vier gleich grosse Gruppen eingeteilt, sogenannte Quartilsgruppen: Schülerinnen und Schüler mit (1) *benachteiligter*, (2) *eher benachteiligter*, (3) *eher privilegiert* und (4) *privilegiert sozialer Herkunft*.

Die Gruppe der Schülerinnen und Schüler mit *benachteiligter sozialer Herkunft* umfasst die Kinder mit den 25 Prozent tiefsten Werten beim Index zur sozialen Herkunft, die Gruppe der Kinder mit *eher benachteiligter sozialer Herkunft* jene mit den Prozentwerten 26 bis 50, die Gruppe der Kinder mit *eher privilegiert sozialer Herkunft* jene mit den Prozentwerten 51 bis 75 und die Gruppe der Kinder mit privilegiert sozialer Herkunft jene mit den 25 Prozent höchsten Werten.

Erstsprache der Schülerinnen und Schüler

Die Variable zur Erstsprache der Schülerinnen und Schüler unterscheidet zwischen Kindern mit Deutsch als Erstsprache (DaE) und solchen mit Deutsch als Zweitsprache (DaZ). Die Erstsprache wurde als Sprache erfasst, welche die Eltern am häufigsten mit ihrem Kind sprechen. Sie kann somit dahingehend interpretiert werden, dass bei den Kindern mit «Deutsch als Erstsprache» die Erstsprache auch der Unterrichtssprache entspricht. Die Angaben stammen aus dem Elternfragebogen zum Testzeitpunkt T2.

Kognitive Grundfähigkeiten

Die kognitive Grundfähigkeit der Schülerinnen und Schüler wurden beim Testzeitpunkt T1 zu Beginn der 1. Klasse anhand des «Culture Fair Intelligence Test» (CFT 1) erhoben (Weiss & Osterland, 1997). Mit diesem Test lässt sich die kognitive Grundfähigkeit frei von sprachlichen und kulturellen Einflüssen erfassen.

Stichprobe 6. Klassen

Soziale Herkunft der Schülerinnen und Schüler

Zur Bestimmung der sozialen Herkunft der Schülerinnen und Schüler wurde mit Hilfe einer Hauptkomponentenanalyse Index gebildet, der aus der Anzahl Bücher im Haushalt und dem höchsten Bildungsabschluss der Eltern besteht. Der Index weist einen Mittelwert von 0 und eine Standardabweichung von 1 auf. Die Streubreite reicht von -2.45 bis 1.96 Indexpunkten. Die Angaben zur Bildung des Indexes stammen aus dem Schülerfragebogen zum Testzeitpunkt T3.

In den Analysen, bei denen die soziale Herkunft der Schülerinnen und Schüler nicht als stetige sondern als kategoriale Variable verwendet wurde, wurden die Schülerinnen und Schüler aufgrund ihres Indexes zur sozialen Herkunft in vier gleich grosse Gruppen eingeteilt, sogenannte Quartilsgruppen: Schülerinnen und Schüler mit (1) *benachteiligter*, (2) *eher benachteiligter*, (3) *eher privilegierter* und (4) *privilegierter sozialer Herkunft*.

Erstsprache der Schülerinnen und Schüler

Die Variable Erstsprache unterscheidet zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache (DaE) und solchen mit Deutsch als Zweitsprache (DaZ). Die Erstsprache der Schülerinnen und Schüler entspricht der Sprache, welche die Kinder zu Hause als erste gelernt haben und kann dahingehend interpretiert werden, dass bei den Kindern mit «Deutsch als Erstsprache» die Erstsprache auch der Unterrichtssprache entspricht. Die Angaben zur Erstsprache stammen aus dem Schülerfragebogen zum Testzeitpunkt T3.

2 Analysen zum Leistungszuwachs vom Schuleintritt bis zum Ende der Primarstufe

Domenico Angelone

2.1 Einleitung

In diesem Kapitel werden die statistischen Analysen zum Leistungszuwachs der Schülerinnen und Schüler in Deutsch, im Wortschatz und in der Mathematik von Beginn der 1. Klasse bis zum Ende der 6. Klasse vorgestellt. Im Zentrum der Analysen stehen zwei Fragestellungen: (1) Wie entwickeln sich die beim Schuleintritt feststellbaren Leistungsdisparitäten zwischen Schülerinnen und Schülern mit unterschiedlichen Lernvoraussetzungen in den ersten sechs Schuljahren? Verstärken sich primäre Herkunftseffekte im Bildungserfolg oder vermag die Schule die anfänglichen Disparitäten auszugleichen? (2) Welche Bedeutung kommt dem Vorwissen beim Schuleintritt für die Leistungen am Ende der Primarschule zu?

In Kapitel 2.2 werden die Datenbasis sowie die statistischen Auswertungsmethoden zur Untersuchung der interessierenden Fragestellungen vorgestellt. In Kapitel 2.3 und 2.4 werden die Ergebnisse der Analysen zum Leistungszuwachs der Schülerinnen und Schüler während den ersten sechs Schuljahren und die Ergebnisse zur Bedeutung des Vorwissens für die Leistungen im späteren Schulverlauf präsentiert.

2.2 Methode

Datenbasis

Grundlage für die statistischen Analysen bilden die Daten der Längsschnittstichprobe der Zürcher Lernstandserhebung. Mit der Längsschnittstudie wurden die Leistungen im Lesen, im Wortschatz und in der Mathematik von rund 2000 Schülerinnen und Schülern, die im Sommer 2003 in die Primarschule des Kantons Zürich eingetreten waren, im Abstand von drei Jahren getestet und befragt. Getestet wurden die Leistungen der Schülerinnen und Schüler zu Beginn der 1. Klasse, am Ende der 3. Klasse und kurz vor dem Übertritt in die Sekundarstufe I am Ende der 6. Klasse.

Statistische Auswertungsmethode

Die Analyse der Entwicklung von individuellen Merkmalen wie den schulischen Leistungen kann mit Hilfe von hierarchischen Mehrebenenmodellen erfolgen (vgl. Raudenbush & Bryk, 2002). Dabei wird auf Ebene 1 die Entwicklung der untersuchten Schülerinnen und Schüler anhand von individuellen Wachstumspfaden repräsentiert. Auf der Ebene 2 stellen diese Wachstumspfade die zu erklärende Variable dar, die von individuellen Eigenschaften der Schülerinnen und Schüler abhängt. Die hierarchische Modellierung besteht formal gesprochen darin, dass die einzelnen Beobachtungseinheiten (Ebene 1), im vorliegenden Fall die Leistungen zu den verschiedenen Testzeitpunkten, pro Schüler beziehungsweise pro Schülerin (Ebene 2) gruppiert sind. Weil im Rahmen der Zürcher Lernstandserhebung mehr als zwei Beobachtungszeitpunkte erfasst wurden, lässt sich ein nicht-lineares Zwei-Ebenenmodell schätzen, ein sogenanntes quadratisches Wachstumsmodell. Im Unterschied zu einem linea-

ren Wachstumsmodell wird dabei angenommen, dass die Wachstumsrate der interessierenden Variablen – hier die schulischen Leistungen – über die Zeit nicht konstant, sondern zunehmend oder abnehmend ist. Das heisst, dass sich ein zusätzliches Schuljahr entweder stärker oder schwächer auf die Leistungen auswirkt. Dementsprechend wurden für die in Kapitel 2.3 geschätzten quadratischen Wachstumsmodelle mit drei Beobachtungszeitpunkten (zu Beginn der 1. Klasse, am Ende der 3. Klasse und am Ende der 6. Klasse) auf der Ebene 1 als Prädiktoren für die Leistungsentwicklung (1) die Anzahl Schuljahre seit Beginn der 1. Klasse sowie (2) ein quadratischer Term der Anzahl Schuljahre berücksichtigt. Der erste Term repräsentiert den linearen Leistungszuwachs für ein zusätzliches Schuljahr und der zweite Term beschreibt, ob die Zuwachsraten über die Zeit abnehmend (falls negativ) beziehungsweise zunehmend (falls positiv) sind. Als Prädiktoren auf der Ebene 2 zur Vorhersage des individuellen Leistungszuwachses wurden das Geschlecht, das Alter, die kognitiven Grundfähigkeiten, die Erstsprache und die soziale Herkunft der Schülerinnen und Schüler einbezogen.

Die Analysen zur Bedeutung des Vorwissens für die Leistungen am Ende der 3. und am Ende der 6. Klasse in Kapitel 2.4 wurden mit Hilfe von klassischen linearen Regressionsanalysen durchgeführt. Als Prädiktoren für die Leistungen am Ende der 3. und am Ende der 6. Klasse wurden zunächst nur die Lernvoraussetzungen der Schülerinnen und Schüler (Geschlecht, Alter, kognitive Grundfähigkeiten, Erstsprache und soziale Herkunft) einbezogen. In einem zweiten Schritt wurde auch das Vorwissen, gemessen an den Leistungen beim Schuleintritt oder am Ende der 3. Klasse, in die Analyse einbezogen, damit die zusätzliche Erklärungskraft des Vorwissens quantifiziert werden konnte.

2.3 Ergebnisse zum Leistungszuwachs von der 1. bis 6. Klasse

Tabelle 2.1 zeigt die Ergebnisse der statistischen Schätzmodelle zum Einfluss der Lernvoraussetzungen der Schülerinnen und Schüler auf den Leistungszuwachs vom Beginn der 1. Klasse bis zum Ende der 6. Klasse im Lesen, im Wortschatz und in der Mathematik.

In den mit «b-Koeff.» beschrifteten Spalten sind die unstandardisierten Regressionskoeffizienten ausgewiesen. Die b-Koeffizienten zeigen, wie gross die Effekte der entsprechenden Variablen auf die Leistungen zu Beginn der 1. Klasse (vgl. «Fixe Effekte für die Konstante») sowie auf den Leistungszuwachs bis zum Ende der 6. Klasse (vgl. «Fixe Effekte für den linearen Zeiteffekt») sind, unter der Annahme, dass alle anderen Variablen konstant gehalten werden. Im Teil «Fixe Effekte für die Konstante» sind die Effekte der einbezogenen Variablen auf die Leistungen zu Beginn der 1. Klasse gezeigt. Die erste aufgeführte Variable «Konstante» entspricht dem Leistungsmittelwert der Referenzgruppe beim Schuleintritt. Die Referenzgruppe bilden die Mädchen mittleren Alters, mit Deutsch als Erstsprache, durchschnittlichen kognitiven Grundfähigkeiten und mit benachteiligter sozialer Herkunft (1. Quartil). Im Lesen beispielsweise beträgt der Mittelwert der Referenzgruppe zu Beginn der 1. Klasse rund 482 Punkte.

Die folgenden Variablen zu den Lernvoraussetzungen beschreiben die Effekte der individuellen Lernvoraussetzungen auf die Leistungen zu Beginn der 1. Klasse: Die Variable «Knaben» entspricht dem Geschlechtereffekt. In Tabelle 2.1 ist für die drei Fachbereiche jeweils aufgeführt, wie gross die Abweichung der Leistungen der Knaben im Vergleich zu den Mädchen ist. Die Variable «Alter in Monaten» gibt die Leistungsveränderung an, wenn das Alter um einen Monat ansteigt. Die Variable «Kognitive Grundfähigkeiten» zeigt den Effekt der kogni-

tiven Grundfähigkeiten der Schülerinnen und Schüler auf die Leistungen. Die Variable «Deutsch als Zweitsprache» zeigt, wie gross der Leistungsunterschied der Schülerinnen und Schüler mit Deutsch als Zweitsprache im Vergleich zu jenen mit Deutsch als Erstsprache ist. Die drei Dummy-Variablen «benachteiligt», «eher privilegiert» und «privilegiert» entsprechen dem Effekt der sozialen Herkunft der Schülerinnen und Schüler auf die Leistungen. Sie geben jeweils die Leistungsdifferenz der entsprechenden Gruppe im Vergleich zur Gruppe der Schülerinnen und Schüler mit benachteiligter sozialer Herkunft an.

Im Teil «Fixe Effekte für den linearen Zeiteffekt (Schuljahre)» sind die Effekte der einbezogenen Variablen auf den Leistungszuwachs von Beginn der 1. Klasse bis zum Ende der 6. Klasse gezeigt. Die erste aufgeführte Variable «Konstante» entspricht dem durchschnittlichen Leistungszuwachs pro Schuljahr der Referenzgruppe (linearer Effekt). Im Lesen beispielsweise beträgt der durchschnittliche lineare Leistungszuwachs der Referenzgruppe pro Schuljahr rund 56 Punkte. Die folgenden Variablen zu den individuellen Lernvoraussetzungen zeigen die entsprechenden Effekte auf den jährlichen Leistungszuwachs bis zum Ende der 6. Klasse. Die Variable «Knaben» beispielsweise zeigt, um wie viele Punkte sich der (lineare) Leistungszuwachs der Knaben im Vergleich zu den Mädchen unterscheidet.

Schliesslich entspricht die Konstante im Teil «Fixe Effekte für den quadratischen Zeiteffekt» dem quadratischen Effekt der Anzahl Schuljahre auf den Leistungszuwachs. Bei einem positiven Koeffizienten ist die Zuwachsrate über die Zeit zunehmend, bei einem negativen Koeffizienten abnehmend. Im Lesen und im Wortschatz ist die Leistungszuwachsrate mit einem Koeffizienten von -1.6 Punkten und -2.9 Punkten abnehmend, in der Mathematik hingegen mit einem Koeffizienten von 5.2 Punkten zunehmend. Im Lesen und im Wortschatz wirkt sich ein zusätzliches Schuljahr über den Schulverlauf immer weniger stark (positiv) auf die Leistungen aus. In der Mathematik hingegen wirkt sich ein zusätzliches Schuljahr immer stärker (positiv) auf die Leistungen aus. Um den Leistungszuwachs vorauszusagen, muss der lineare und der quadratische Effekt gemeinsam berücksichtigt werden. Im Lesen beispielsweise beträgt der Leistungszuwachs für die Referenzgruppe nach einem Schuljahr 56.4 Punkte ($1 \cdot 56.4 + 1^2 \cdot -1.6$) und nach zwei Schuljahren 106.4 Punkte ($2 \cdot 56.4 + 2^2 \cdot -1.6$).

Geschlecht

Zu Beginn der 1. Klasse erreichen die Knaben im Lesen und im Wortschatz statistisch signifikant tiefere Leistungen als die Mädchen. Im Lesen sind die Leistungen der Knaben rund 14 Punkte und im Wortschatz rund 9 Punkte tiefer als jene der Mädchen. Demgegenüber sind beim Schuleintritt die Mathematikleistungen der Knaben um rund 27 Punkte höher als jene der Mädchen.

Der Leistungszuwachs bis zum Ende der 6. Klasse unterscheidet sich lediglich im Wortschatz und in der Mathematik statistisch signifikant zwischen den Geschlechtern. Im Wortschatz ist der Leistungszuwachs der Knaben jährlich 3 Punkte und in der Mathematik jährlich 5 Punkte geringer als jener der Mädchen.

Tabelle 2.1: Ergebnisse der Mehrebenenanalysen zum Einfluss der Lernvoraussetzungen auf den Leistungszuwachs von Beginn der 1. Klasse bis zum Ende der 6. Klasse

	Leistungen im Lesen	Leistungen im Wortschatz	Leistungen in der Mathematik
	b-Koeff.	b-Koeff.	b-Koeff.
Fixe Effekte für die Konstante			
Konstante	481.6***	508.3***	472.2***
Knaben	-13.9**	-8.8*	26.6***
Alter (in Monaten)	1.5**	1.1*	2.3***
Kognitive Grundfähigkeiten	2.0***	1.7***	3.2***
Deutsch als Zweitsprache	5.3	-82.8***	6.4
Soziale Herkunft: eher benachteiligt	12.3*	11.9+	14.8*
Soziale Herkunft: eher privilegiert	32.9***	39.1***	16.1**
Soziale Herkunft: privilegiert	36.3***	48.8***	13.5*
Fixe Effekte für den linearen Zeiteffekt (Schuljahre)			
Konstante	56.4***	53.3***	34.0***
Knaben	-2.1+	-3.0**	-4.8***
Alter (in Monaten)	-0.3*	-0.1	-0.4***
Kognitive Grundfähigkeiten	0.1*	0.1*	0.0
Deutsch als Zweitsprache	-5.7***	11.1***	-0.5
Soziale Herkunft: eher benachteiligt	0.1	2.1	0.8
Soziale Herkunft: eher privilegiert	1.2	0.9	4.0**
Soziale Herkunft: privilegiert	5.7**	3.8*	9.1***
Fixe Effekte für den quadratischen Zeiteffekt (Schuljahre ²)			
Konstante	-1.6***	-2.9***	5.2***
Variable Effekte (Varianzen)			
Konstante (Leistungen 1. Klasse)	983***	667**	2919***
Linearer Zeiteffekt (Schuljahre)	76***	64***	52***
Ebene 1	6910***	5415***	4279***
Modellparameter			
N (Ebene 2)	1872	1872	1872
R ² (Ebene 2)	55.3%	83.5%	46.5%

Anmerkungen: Quadratisches Wachstumsmodell (Ebene 1: Zeitpunkte, Ebene II: Schülerinnen Schüler), Random-Intercept-Model, Schätzverfahren: Restricted Maximum Likelihood. Abhängige Variable ist der Leistungszuwachs im Lesen, im Wortschatz und in der Mathematik (Zeitpunkt T1= 0, Zeitpunkt T2 = 2.75 Monate und Zeitpunkt T3 = 5.75 Monate). Die Variablen «kognitive Grundfähigkeiten» und «Alter in Monaten» wurden am «Grand Mean» zentriert. Signifikanzniveau: †p < .10; *p < .05; **p < .01; ***p < .001 (zweiseitige Tests).

Alter

Das Alter der Schülerinnen und Schüler steht beim Schuleintritt in einer statistisch signifikant positiven Beziehungen zu den Leistungen in den drei Fächern. Je älter die Schülerinnen und Schüler sind, desto höher sind auch ihre Leistungen. In der Mathematik führt ein Anstieg des

Alters um einen Monat zu einem Anstieg der Eingangsleistungen im Lesen und in der Mathematik um rund 2 Punkte, im Wortschatz um rund 1 Punkt.

Das Alter wirkt sich im Lesen und in der Mathematik auch statistisch signifikant auf den Leistungszuwachs auf. Je älter die Schülerinnen und Schüler sind, desto geringer ist auch ihr Leistungszuwachs. Bei einem Anstieg des Alters um einen Monat verringert sich der Leistungszuwachs im Lesen jährlich um 0.3 Punkte und in der Mathematik um jährlich 0.4 Punkte.

Kognitive Grundfähigkeiten

Auch die kognitiven Grundfähigkeiten der Schülerinnen und Schüler hängen mit den Leistungen zu Beginn der 1. Klasse zusammen. Je höher die kognitiven Grundfähigkeiten sind, desto höher sind auch die Leistungen. Bei einem Anstieg der kognitiven Grundfähigkeiten um 15 Punkte, was einer Standardabweichung entspricht, steigen die Leistungen im Lesen um 30 Punkte, im Wortschatz um 26 Punkte und in der Mathematik um 48 Punkte an.

Im Lesen und in der Mathematik stehen die kognitiven Grundfähigkeiten in einem statistisch signifikant positiven Zusammenhang mit dem Leistungszuwachs. Die Effekte sind jedoch eher gering. Bei einem Anstieg der kognitiven Grundfähigkeiten um rund 15 Punkte (eine Standardabweichung) steigt der jährliche Leistungszuwachs um 1.5 Punkte an.

Deutsch als Zweitsprache

Die Erstsprache der Schülerinnen und Schüler hängt zu Beginn der 1. Klasse lediglich mit den Leistungen im Wortschatz zusammen. Die Schülerinnen und Schüler mit Deutsch als Zweitsprache erreichen mehr als 80 Punkte weniger als jene mit Deutsch als Erstsprache. Dieser Unterschied in den Wortschatzleistungen muss als sehr gross interpretiert werden.

Im Hinblick auf den Leistungszuwachs ist die Erstsprache im Lesen und im Wortschatz von Bedeutung. Die Effekte gehen jedoch in zwei unterschiedliche Richtungen. Im Lesen ist der Leistungszuwachs der Schülerinnen und Schüler mit Deutsch als Zweitsprache im Vergleich zu jenem der Schülerinnen und Schüler mit Deutsch als Erstsprache geringer (6 Punkte pro Schuljahr). Dies führt dazu, dass die Kinder mit Deutsch als Zweitsprache am Ende der 6. Klasse im Lesen einen Rückstand aufweisen. Im Wortschatz hingegen ist der Leistungszuwachs der Kinder mit Deutsch als Zweitsprache grösser (rund 11 Punkte pro Schuljahr), so dass sich die beim Schuleintritt beobachteten Leistungsunterschiede bis zum Ende der Primarschule auf weniger als 20 Punkte reduzieren. Für den Erwerb von mathematischen Fähigkeiten hat die Erstsprache bei gleichen Lernvoraussetzungen keine eigenständige Bedeutung.

Soziale Herkunft

Die soziale Herkunft der Schülerinnen und Schüler spielt erwartungsgemäss eine relativ starke Rolle für die Leistungen beim Schuleintritt, namentlich im Lesen und im Wortschatz. Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto höher sind auch die Leistungen zu Beginn der 1. Klasse. Im Lesen erreichen die Kinder mit eher benachteiligter sozialer Herkunft rund 12 Punkte, jene mit eher privilegierter sozialer Herkunft rund 33 Punkte und jene mit privilegierter sozialer Herkunft rund 36 Punkte mehr als jene mit benachteiligter sozialer Herkunft. Im Wortschatz erzielen die Kinder mit eher privilegierter sozialer Herkunft rund 39 Punkte und jene mit privilegierter sozialer Herkunft rund 49 Punkte mehr als jene mit benachteiligter sozialer Herkunft. Zwischen den Schülerinnen und Schülern mit benachteiligter sozialer Herkunft und jenen mit eher benachteiligter sozialer Herkunft finden sich im

Wortschatz keine Unterschiede. In der Mathematik schliesslich zeigt sich zu Beginn der 1. Klasse ein gleichmässigeres Bild. Im Vergleich zu den Kindern mit benachteiligter sozialer Herkunft erzielen die Kinder mit eher benachteiligter, mit eher privilegierter und mit privilegierter sozialer Herkunft ein um rund 14 bis 16 Punkte besseres Ergebnis.

Die soziale Herkunft der Schülerinnen und Schüler ist auch für den Leistungszuwachs von Beginn der 1. Klasse bis zum Ende der Primarschule von Bedeutung, wenn auch in einem geringeren Ausmass als für die Leistungen beim Schuleintritt. Im Lesen und im Wortschatz zeigen sich beim Leistungszuwachs keine Unterschiede zwischen Schülerinnen und Schülern mit eher privilegierter sowie mit eher benachteiligter sozialer Herkunft und jenen mit benachteiligter sozialer Herkunft. Demgegenüber ist im Lesen der jährliche Leistungszuwachs der Kinder mit privilegierter sozialer Herkunft um rund 6 Punkte und im Wortschatz um rund 4 Punkte grösser als jener der Kinder mit benachteiligter sozialer Herkunft. In der Mathematik weisen die Kinder mit eher privilegierter (rund 9 Punkten pro Schuljahr) und jene mit privilegierter (rund 4 Punkte pro Schuljahr) sozialer Herkunft einen statistisch signifikant grösseren Leistungszuwachs auf als die Kinder mit benachteiligter sozialer Herkunft.

2.4 Ergebnisse zu den Leistungen am Ende der 6. Klasse nach Vorwissen

Tabelle 2.2 zeigt die Ergebnisse der Regressionsanalysen zur Erklärung der Leistungen am Ende der 6. Klasse. In der ersten Spalte sind die einzelnen Variablen aufgeführt, die in die Analysen einbezogen wurden. Aufgeführt sind jeweils die unstandardisierten Regressionskoeffizienten (b-Koeff). Diese geben an, wie die Einflüsse zu interpretieren sind, wenn der Einfluss aller anderen Variablen konstant gehalten wird. Pro Fachbereich wurden drei getrennte Analysen durchgeführt. Die Modelle I berücksichtigen jeweils nur die Lernvoraussetzungen der Schülerinnen und Schüler (Geschlecht, Alter, kognitive Grundfähigkeiten, Erstsprache und soziale Herkunft) zur Erklärung der Leistungen am Ende der 6. Klasse. Die Modelle II beziehen zusätzlich das Vorwissen der Schülerinnen und Schüler ein, gemessen an den Leistungen zu Beginn der 1. Klasse. Die Modelle III schliesslich berücksichtigen zusätzlich zu den Lernvoraussetzungen die Leistungen am Ende der 3. Klasse.

Bei der Interpretation der Ergebnisse stehen im Folgenden nicht die einzelnen Effekte der Prädiktoren im Vordergrund, sondern die Verbesserung der Modellerklärungskraft, wenn zusätzlich zu den Lernvoraussetzungen der Schülerinnen und Schüler auch ihr Vorwissen, das heisst ihre Leistungen zu Beginn der 1. beziehungsweise am Ende der 3. Klasse, berücksichtigt werden.

Leistungen im Lesen

Im Lesen erklären die Lernvoraussetzungen der Schülerinnen und Schüler 25.8 Prozent der Leistungsunterschiede am Ende der 6. Klasse (vgl. Modell I). Die Berücksichtigung der Leseleistungen zu Beginn der 1. Klasse verbessert die Erklärungskraft des Modells um 2 Prozent auf 27.8 Prozent (vgl. Modell II). Je höher die Leseleistungen beim Schuleintritt sind, desto höher sind die Leistungen am Ende der 6. Klasse. Bei einem Anstieg der Eingangsleistungen um 100 Punkte verbessern sich die Leistungen am Ende der 6. Klasse um rund 22 Punkte.

Ein ähnliches Bild zeigt sich, wenn zur Vorhersage der Leistungen am Ende der 6. Klasse nicht die Eingangsleistungen, sondern die Leistungen am Ende der 3. Klasse verwendet wer-

den. Nach Berücksichtigung der Lernvoraussetzungen erklären diese im Lesen weitere 3.3 Prozent der Leistungsunterschiede (vgl. Modell III).

Leistungen im Wortschatz

Im Wortschatz erklären die Lernvoraussetzungen der Schülerinnen und Schüler 23 Prozent der Leistungsunterschiede am Ende der 6. Klasse (Modell I). Die zusätzliche Berücksichtigung der Leistungen zu Beginn der 1. Klasse verbessert die Erklärungskraft des Modells um 3.6 Prozent auf 26.6 Prozent (vgl. Modell II). Auch im Wortschatz hängen die Eingangsleistungen mit den Leistungen am Ende der 6. Klasse zusammen. Bei einem Anstieg der Leistungen zu Beginn der 1. Klasse um 100 Punkte verbessern sich die Wortschatzleistungen am Ende der 6. Klasse um rund 28 Punkte.

Werden zur Vorhersage der Leistungen am Ende der 6. Klasse nicht die Eingangsleistungen, sondern die Leistungen am Ende der 3. Klasse verwendet, dann beträgt der zusätzliche Erklärungsgehalt 5.4 Prozent (vgl. Modell III).

Leistungen in der Mathematik

In der Mathematik erklären die Lernvoraussetzungen der Schülerinnen und Schüler 28.9 Prozent der Leistungsunterschiede am Ende der 6. Klasse (vgl. Modell I). Die zusätzliche Berücksichtigung der Leistungen beim Schuleintritt erhöht die Erklärungskraft des Modells um rund 7 Prozent (Modell II). Erwartungsgemäss hängen auch in der Mathematik die Eingangsleistungen positiv mit den Leistungen am Ende der 6. Klasse zusammen. Bei einem Anstieg der Leistungen zu Beginn der 1. Klasse um 100 Punkte steigen die Mathematikleistungen am Ende der 6. Klasse um rund 34 Punkte an.

Nach Berücksichtigung der Lernvoraussetzungen erklären die Leistungen am Ende der 3. Klasse weitere 12.5 Prozent der Leistungsunterschiede am Ende der 6. Klasse (vgl. Modell III).

Tabelle 2.2: Ergebnisse der Regressionsanalysen zum Einfluss der Leistungen beim Schuleintritt auf die Leistungen am Ende der 6. Klasse

	Leistungen im Lesen			Leistungen im Wortschatz			Leistungen in der Mathematik		
	Modell I	Modell II	Modell III	Modell I	Modell II	Modell III	Modell I	Modell II	Modell III
	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.
Konstante	777.60***	778.37***	773.75***	761.36***	754.14***	754.87***	869.83***	876.89***	873.74***
Knaben	-17.04**	-15.48**	-8.77+	-20.11***	-20.24***	-11.96*	3.65	-7.43+	-1.13
Alter (in Monaten)	-0.03	-0.42	0.00	-0.05	-0.43	-0.09	-0.19	-1.09*	-0.34
Deutsch als Zweitsprache	-18.92**	-23.06**	-15.60*	-35.01***	-11.67+	-24.65***	3.77	-0.34	3.61
Kognitive Grundfähigkeiten	2.87***	2.32***	2.39***	1.75***	1.33***	1.07***	2.77***	1.77***	1.41***
Soziale Herkunft	26.34***	22.69***	21.07***	26.29***	19.35***	18.19***	26.22***	23.33***	19.70***
Leistungen 1. Klasse (in 100 Punkten)		21.73***			28.01***			33.74***	
Leistungen Ende 3. Klasse (in 100 Punkten)			26.56***			32.96***			44.15***
R2 (Gesamt)	25.8%	27.8%	29.1%	23.0%	26.6%	28.4%	28.9%	35.9%	41.4%
N (listwise delete)	1673	1617	1652	1673	1617	1652	1675	1619	1654

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Regressionsmodelle. Abhängige Variable sind die Leistungen im Lesen, im Wortschatz und in der Mathematik am Ende der 6. Klasse. Für die Leistungen wurde jeweils der 1. Plausible Value verwendet. Signifikanzniveaus: +p < .10; *p < .05; **p < .01; ***p < .001

3 Methodisches Vorgehen und Analysen zu den Fachleistungen am Ende der 6. Klasse

Urs Moser & Domenico Angelone

3.1 Bildung von Anforderungsniveaus

Testtheoretische Grundlagen

Um die Fachleistungen der Schülerinnen und Schüler am Ende der 6. Klasse mit Bezug zum Lehrplan beurteilen zu können, wurden für die beiden Fachbereiche Deutsch und Mathematik Anforderungsniveaus gebildet. Der Lehrplan für die Volksschule des Kantons Zürich diente deshalb bereits bei der Entwicklung der Testaufgaben als Referenzrahmen.

Jede Testaufgabe musste einem im Voraus bestimmten Anforderungsniveau zugeteilt werden (Niveau I: minimale Anforderungen, Niveau II: grundlegende Anforderungen, Niveau III: erweiterte Anforderungen, Niveau IV: höhere Anforderungen). Inwieweit die Einteilung der Aufgaben nach den vier formal definierten Niveaus zutrifft, zeigte sich in den Schwierigkeiten der Aufgaben, die aus der Lösungshäufigkeit berechnet wurde. Grundlage der Beurteilung der Leistungen aufgrund von lehrplanorientierten Anforderungsniveaus ist die probabilistische Testtheorie, nach der ein (probabilistischer) Bezug zwischen den Schwierigkeiten der Aufgaben und den Fähigkeiten der Schülerinnen und Schüler hergestellt wird. Die Wahrscheinlichkeit, dass ein Schüler eine Aufgabe richtig löst, lässt sich als Funktion der *Schwierigkeit* der Aufgabe und der *Fähigkeit* eines Schülers bestimmen. Je fähiger beispielsweise eine Schülerin ist, desto wahrscheinlicher ist es, dass sie eine Aufgabe mit einer bestimmten Schwierigkeit richtig löst. Und je schwieriger eine Aufgabe ist, desto unwahrscheinlicher ist es, dass eine Schülerin mit einer bestimmten Fähigkeit die Aufgabe richtig löst.

Die Beziehung zwischen Lösungswahrscheinlichkeit, Fähigkeit einer Person und Schwierigkeit einer Aufgabe ist für die lehrplanorientierte Interpretation der Ergebnisse von entscheidender Bedeutung. Sofern die Aufgaben dem probabilistischen Testmodell entsprechen, können die Ergebnisse in den Leistungstests mit Bezug zu den Aufgaben interpretiert werden. Fähigkeit eines Schülers (Testergebnis) und Schwierigkeit einer Aufgabe (Testaufgabe) lassen sich auf der gleichen Skala abbilden. Konkret heisst das, dass sich aufgrund des individuellen Ergebnisses für jede Aufgabe bestimmen lässt, wie wahrscheinlich es ist, dass ein Schüler oder eine Schülerin die Aufgabe richtig löst (Rost, 2004, S. 89ff.).

Bildung von Anforderungsniveaus

In einem ersten Schritt wurde überprüft, ob die Aufgaben dem Testmodell entsprechen und davon ausgegangen werden kann, dass sämtliche Aufgaben als Indikatoren für die Fähigkeiten in Deutsch und Mathematik genutzt werden können. In einem zweiten Schritt wurden Aufgaben, deren Schwierigkeit nahe beieinander liegen, analysiert und aufgrund ihrer Anforderungen kategorisiert. Durch dieses Vorgehen konnten die vier Anforderungsniveaus inhaltlich unterschieden werden.

Die gleichzeitige Verortung der Aufgabenschwierigkeit einerseits und des Ergebnisses in den Leistungstests auf der gleichen Skala andererseits ermöglicht nicht nur die Beurteilung des Lernstandes nach Anforderungsniveau, sondern auch eine Beschreibung der Fähigkeiten.

Jedem der aufeinander aufbauenden Anforderungsniveaus entsprechen Aufgaben mit ansteigendem Schwierigkeitsgrad. Die Aufgaben eines Anforderungsniveaus wurden auf Ähnlichkeiten hin überprüft. Zudem wurde analysiert, wie sich die Aufgaben eines Anforderungsniveaus von den Aufgaben des nächsthöheren Niveaus unterscheiden.

Die Übereinstimmung der zum Voraus durchgeführten Einteilung der Aufgaben nach Anforderungsniveau mit der Schwierigkeit der Aufgaben war für die Mathematik grösser als für Deutsch. Massgebend für die Bildung der Anforderungsniveaus waren die Schwierigkeiten der Aufgaben, wie sie aus der Lösungshäufigkeit pro Aufgabe ermittelt wurden.

3.2 Leistungstests für den Vergleich zwischen 1998 und 2009

Im Sommer 1998 wurden – gleich wie im Sommer 2009 – die Schulleistungen am Ende der 6. Klasse bei einer repräsentativen Stichprobe des Kantons Zürich erhoben. Um die Ergebnisse der beiden Lernstandserhebungen vergleichen zu können, wurde in der Lernstandserhebung 2009 ein Teil der Testaufgaben aus dem Jahr 1998 wieder eingesetzt. Für den Vergleich der Deutschleistungen wurden 52 Aufgaben aus dem Jahr 1998 verwendet, für den Vergleich der Mathematikleistungen 50 Aufgaben.

Auf den ersten Blick scheint einem validen und zuverlässigen Schulleistungsvergleich nichts im Wege zu stehen, schliesslich werden identische Testaufgaben bei zwei für den Kanton Zürich repräsentativen Stichproben zu zwei Zeitpunkten eingesetzt. Auch der internationale Schulleistungsvergleich PISA versucht, dank der Verwendung gleicher Testaufgaben in den verschiedenen Erhebungen zu Trendaussagen über die Leistungsentwicklung zu gelangen. Trotz gleicher Testaufgaben ist ein zuverlässiger Vergleich aber alles andere als einfach. In einer Zeitspanne von elf Jahren kann sich in der Schule einiges ändern. Zwar wurden im Jahr 1998 der gleiche Lehrplan und das gleiche obligatorische Mathematiklehrmittel eingesetzt wie im Jahr 2009. Ob dieses intendierte Curriculum jedoch zu beiden Zeitpunkten gleich umgesetzt wurde, ist ungewiss. Wie die Mathematik im Unterricht verstanden und vermittelt wird, hängt kaum nur von offiziellen Lehrplänen und Lehrmitteln ab.

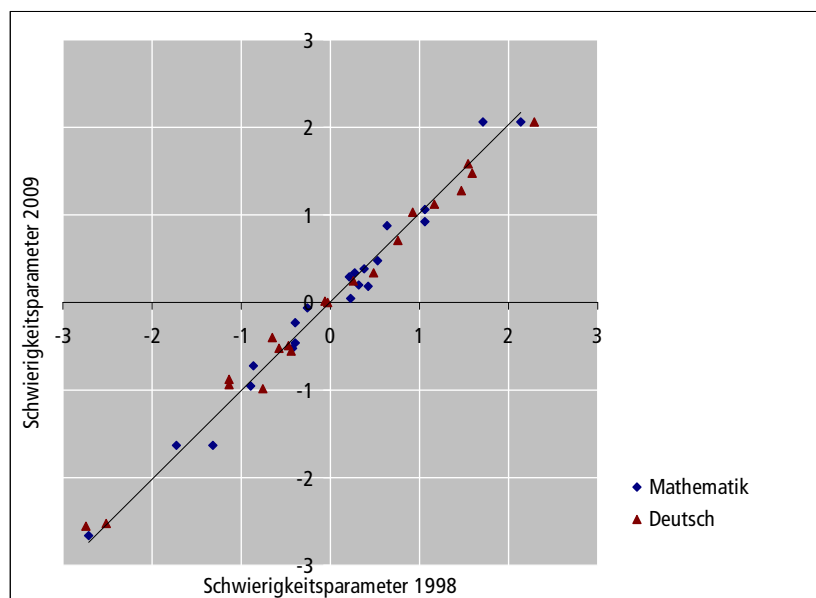
Der Vergleich der Ergebnisse der beiden Lernstandserhebungen ist allerdings nur dann sinnvoll, wenn die Aufgaben zu beiden Zeitpunkten etwa gleich bedeutsam für den Unterricht sind und wenn sich deshalb die (relative) Schwierigkeit der Aufgaben während den Erhebungszeitpunkten nicht ändert (vgl. PISA-Konsortium Deutschland, 2004). Selbstverständlich mögen beispielsweise Aufgaben im Jahr 2009 faktisch schwieriger sein als im Jahr 1998, weil sie von den Schülerinnen und Schülern im Jahr 2009 schlechter gelöst werden. Ein solches Ergebnis kann aber nur dann sinnvoll interpretiert werden, wenn Effekte von unkontrollierbaren Einflüssen wie neue Lehrmittel ausgeschlossen werden können und die (relative) Schwierigkeit der Aufgaben über die Zeit hinweg gleich geblieben ist. Unabhängig von den Fähigkeiten der Schülerinnen und Schülern, die die Aufgabe lösen, müssen einfache Aufgaben sowohl im Jahr 1998 als auch im Jahr 2009 einfach sein, schwierige sowohl im Jahr 1998 als auch im Jahr 2009 schwierig sein.

Neben curricular bedingten Änderungen im Unterricht kann der Vergleich aber auch deshalb beeinträchtigt sein, weil die Leistungstests der beiden Lernstandserhebungen nicht exakt gleich sein konnten. Während im Jahr 1998 sämtliche Schülerinnen und Schüler alle Testaufgaben bearbeiteten, wurde im Jahr 2009 ein Multi-Matrix-Testdesign eingesetzt, weil die

Testzeit auf zwei Lektionen beschränkt war. Für die Lernstandserhebung 2009 wurden in Deutsch und Mathematik je sechs verschiedene Testhefte eingesetzt. Trotz kurzer Testzeit können so die Schulleistungen valide erfasst und beschrieben werden. Das unterschiedliche Vorgehen hatte jedoch zur Folge, dass die gleichen Testaufgaben jeweils nicht an der gleichen Stelle im Test eingesetzt werden konnten. Wird eine Aufgabe am Ende eines Tests eingesetzt, so wird sie von weniger Schülerinnen und Schülern gelöst, als wenn sie am Anfang eines Tests eingesetzt wird. Die unterschiedliche Stellung einer Aufgabe in einem Leistungstests könnte sich – gleich wie curriculare Änderungen – auf die Schwierigkeit einer Aufgabe auswirken.

Die Überprüfung der Aufgabenschwierigkeit zu den beiden Zeitpunkten führte dazu, dass von den 102 Aufgaben 41 Aufgaben uneingeschränkt für den Vergleich genutzt werden konnten. Abbildung 3.2 zeigt die Schwierigkeitsparameter der 41 Aufgaben, die in beiden Lernstandserhebungen nahezu gleich schwierig waren. Die Quadrate in den Abbildungen repräsentieren die Mathematikaufgaben, die Dreiecke die Deutschaufgaben. Die Position der Aufgaben (Quadrate und Dreiecke) ergibt sich durch die Schwierigkeitsparameter der Jahre 1998 (x-Achse) und 2009 (y-Achse). Die Spannweite der Schwierigkeitsparameter reicht von -3 bis $+3$. Die Berechnung der Schwierigkeitsparameter wurde so festgelegt, dass die durchschnittliche Schwierigkeit in beiden Jahren bei «0» lag.

Abbildung 3.1: Schwierigkeitsparameter der Aufgaben für den Leistungsvergleich



Bei der Auswahl der Aufgaben für den Vergleich der Leistungen in den Jahren 1998 und 2009 wurde zudem kontrolliert, ob prozentual gleich viele Schülerinnen und Schüler die Aufgabe bearbeiteten, unabhängig davon, ob sie die Aufgabe richtig oder falsch lösten. Für die Schätzung der Fähigkeiten wurde ein Hintergrundmodell mit Informationen zu den Schülerinnen und Schülern wie soziale Herkunft oder Schulform auf der Sekundarstufe genutzt. Damit kann die Messgenauigkeit erhöht werden (vgl. beispielsweise PISA-Konsortium Deutschland, 2004, S. 380ff.)

3.3 Ergebnisse zum Leistungsvergleich zwischen 1998 und 2009

Die Datengrundlage für die Analysen zum Vergleich der Deutsch- und Mathematikleistungen zwischen 1998 und 2009 bilden die beiden repräsentativen Stichproben der 6. Klassen. Für den Leistungsvergleich wurden nur diejenigen Testaufgaben verwendet, die in beiden Leistungstests dieselbe Schwierigkeit aufweisen (vgl. Kapitel 3.2). Die Analysen wurden mit Hilfe von linearen Zwei-Ebenenmodellen durchgeführt (Schülerinnen und Schüler sind in Klassen gruppiert), wobei lediglich die Variation der Konstanten auf der Ebene 2 zwischen den Klassen zugelassen wurde (random-intercept-model). Um mögliche Unterschiede in der Zusammensetzung der Schülerschaft zwischen den beiden Stichproben zu berücksichtigen, wurden die Analysen unter Kontrolle der Lernvoraussetzungen der Schülerinnen und Schüler durchgeführt.

Tabelle 3.1 zeigt die Ergebnisse zum Vergleich der Deutschleistungen der Jahre 1998 und 2009. In der ersten Spalte sind die Variablen aufgeführt, die in die Analysen einbezogen wurden. In den mit «b-Koeff.» bezeichneten Spalten sind die unstandardisierten Regressionskoeffizienten ausgewiesen. Diese zeigen, wie die Effekte der entsprechenden Variablen auf die Deutschleistungen zu interpretieren sind, wenn alle anderen Variablen konstant gehalten werden. Die Variable «Knaben» entspricht der Abweichung der Leistungen der Knaben von den Leistungen der Mädchen. Die Variable «Deutsch als Zweitsprache» zeigt, wie gross die Abweichung der Schülerinnen und Schüler mit Deutsch als Zweitsprache im Vergleich zu den Schülerinnen und Schülern mit Deutsch als Erstsprache ist. Die Variable «Soziale Herkunft» gibt die Leistungsveränderung an, wenn die soziale Herkunft der Schülerinnen und Schüler um eine Standardabweichung ansteigt. Schliesslich zeigt die Variable «Jahr 2009», wie gross die Leistungsabweichung im Jahr 2009 gegenüber dem Jahr 1998 ist.

In Modell I wurde lediglich die Dummy-Variable «Jahr 2009» in die Analyse einbezogen. Die Konstante entspricht somit den durchschnittlichen Leistungen im Jahr 1998, die 503 Punkte betragen. Der Effekt der Variablen «Jahr 2009» ist statistisch nicht signifikant und zeigt, dass sich die Deutschleistungen zwischen den Jahren 1998 und 2009 nicht unterscheiden. In Modell II wurden zusätzlich die Lernvoraussetzungen der Schülerinnen und Schüler in die Analyse einbezogen. Die b-Koeffizienten der Lernvoraussetzungen entsprechen den durchschnittlichen Effekten der Lernvoraussetzungen für die Jahre 1998 und 2009. Die Ergebnisse zeigen, dass auch bei gleichen Lernvoraussetzungen keine Leistungsunterschiede zwischen 1998 und 2009 feststellbar sind. Die Deutschleistungen unterscheiden sich auch nicht zwischen Mädchen und Knaben. Die Erstsprache der Schülerinnen und Schüler hingegen ist von Bedeutung für die Leistungen. Schülerinnen und Schüler mit Deutsch als Erstsprache erreichen statistisch signifikant höhere Deutschleistungen als Schülerinnen und Schüler mit Deutsch als Zweitsprache. Ihr Vorsprung beträgt rund 30 Punkte. Schliesslich hängen die Leistungen auch mit der sozialen Herkunft der Schülerinnen und Schüler zusammen. Bei einem Anstieg der sozialen Herkunft um eine Standardabweichung steigen die Deutschleistungen um rund 21 Punkte an.

Weiter wurde überprüft, ob sich die Effekte der Lernvoraussetzungen zwischen den Jahren 1998 und 2009 unterscheiden. Beim Geschlechtereffekt und beim Effekt der Erstsprache zeigen sich keine statistisch signifikanten Unterschiede zwischen den beiden Erhebungen. Der Effekt der sozialen Herkunft hingegen ist im Jahr 2009 weniger stark ausgeprägt als im Jahr 1998. Dieses Ergebnis ist in Modell III ausgewiesen, in den zusätzlich zu den bisherigen Lern-

voraussetzungen auch ein Interaktionsterm «Soziale Herkunft·Jahr 2009» berücksichtigt wurde. Der statistisch signifikante Interaktionsterm zwischen der sozialen Herkunft und dem Erhebungsjahr zeigt, dass der Effekt der sozialen Herkunft im Jahr 2009 rund 15 Punkte geringer ausfällt als im Jahr 1998. Im Jahr 2009 steigen die Deutschleistungen bei einem Anstieg der sozialen Herkunft um eine Standardabweichung um rund 28 Punkte an, im Jahr 2009 um rund 13 Punkte.

Tabelle 3.1: Ergebnisse der Mehrebenenanalysen zum Leistungsvergleich in Deutsch

	Modell I	Modell II	Modell III
Fixe Effekte	b-Koeff.	b-Koeff.	b-Koeff.
Konstante	503.0***	507.8***	507.1***
Knaben		5.6	5.9
Deutsch als Zweitsprache		-29.7***	-30.0***
Soziale Herkunft (Index)		20.8***	28.3***
Soziale Herkunft * Jahr 2009			-14.7*
Jahr 2009	-5.4	-3.5	-3.6
Variable Effekte			
Varianz Klassenebene	664.7***	403.4***	379.2***
Varianz Schülerebene	9367.1***	8733.5***	8695.3***
Erklärte Varianz			
R ² (Schülerebene)	0.2%	39.4%	43.1%
R ² (Klassenebene)	0.0%	6.8%	7.2%
R ² (Gesamt)	0.0%	8.9%	9.6%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler sind in Klassen gruppiert), Random-Intercept. Abhängige Variable sind die Deutschleistungen am Ende der sechsten Klasse. Die Variable «Soziale Herkunft (Index)» wurde am «Grand Mean» zentriert. Gewichtete Daten. Signifikanzniveaus: *p < .10; **p < .05; ***p < .01; ****p < .001.

Tabelle 3.2 zeigt die Ergebnisse zum Leistungsvergleich in Mathematik zwischen den Jahren 1998 und 2009. In Modell I wurde wiederum nur die Dummy-Variable «Jahr 2009» in die Analyse einbezogen. Die Konstante entspricht somit den durchschnittlichen Leistungen in Mathematik im Jahr 1998. Diese betragen rund 500 Punkte. Auch in Mathematik sind keine Leistungsunterschiede zwischen den Jahren 1998 und 2009 feststellbar. Der Effekt der Variable «Jahr 2009» ist statistisch nicht signifikant. In Modell II wurden zusätzlich die Lernvoraussetzungen der Schülerinnen und Schüler in die Analyse einbezogen. Auch bei gleichen Lernvoraussetzungen sind keine Leistungsunterschiede zwischen den beiden Erhebungen feststellbar.

Die Knaben erzielten in Mathematik gegenüber den Mädchen bessere Leistungen. Ihr Leistungsvorsprung von rund 13 Punkten ist jedoch als gering zu beurteilen. Schülerinnen und Schüler mit Deutsch als Erstsprache erreichen ein um rund 23 Punkte besseres Ergebnis als

Schülerinnen und Schüler mit Deutsch als Zweitsprache. Schliesslich nehmen die Mathematikleistungen bei einem Anstieg der sozialen Herkunft um eine Standardabweichung um rund 22 Punkte zu. Weitere Analysen haben gezeigt, dass sich die Effekte der Lernvoraussetzungen auf die Mathematikleistungen nicht statistisch signifikant zwischen den beiden Erhebungen unterscheiden.

Tabelle 3.2: Ergebnisse der Mehrebenenanalysen zum Leistungsvergleich in Mathematik

	Modell I	Modell II
Fixe Effekte	b-Koeff.	b-Koeff.
Konstante	499.6***	498.7***
Knaben		12.6*
Deutsch als Zweitsprache		-22.5***
Soziale Herkunft (Index)		21.7***
Jahr 2009	1.6	5.1
Variable Effekte		
Varianz Klassenebene	635.6***	463.6***
Varianz Schülerebene	9377.9***	8828.9***
Erklärte Varianz		
R ² (Schülerebene)	0.8%	26.4%
R ² (Klassenebene)	0.0%	5.9%
R ² (Gesamt)	0.1%	7.1%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler sind in Klassen gruppiert), Random-Intercept. Abhängige Variable sind die Mathematikleistungen am Ende der sechsten Klasse. Die Variable «Soziale Herkunft (Index)» wurde am «Grand Mean» zentriert. Gewichtete Daten. Signifikanzniveaus: *p < .10; **p < .05; ***p < .01; ****p < .001.

3.4 Ergebnisse zum Vergleich der Selbsteinschätzungen von Schul- und Unterrichtsmerkmalen zwischen 1998 und 2009

Neben den Leistungen können auch Einschätzungen der Schülerinnen und Schüler von Schul- und Unterrichtsmerkmalen zwischen den Jahren 1998 und 2009 verglichen werden. Dazu gehören Einschätzungen zur Schulzufriedenheit, zum Schulklima, zum sozialen Verhalten an der Schule als Indikator für aggressives und gewalttätiges Verhalten sowie zum Arbeitsklima im Deutsch- und Mathematikunterricht.

Tabelle 3.3 zeigt die Ergebnisse der Mehrebenenanalysen zum Vergleich der Selbsteinschätzungen von Schul- und Unterrichtsmerkmalen zwischen 1998 und 2009. In der Spalte, die mit «b-Koeff.» bezeichnet ist, sind die unstandardisierten Regressionskoeffizienten ausgewiesen. Um Effekte unterschiedlicher Schülerzusammensetzungen zwischen den beiden Stichproben zu berücksichtigen, wurden die Lernvoraussetzungen der Schülerinnen und Schüler in die Analysen einbezogen. Die Konstante entspricht somit der durchschnittlichen Einschätzung von Mädchen mit Deutsch als Erstsprache und mittlerer sozialer Herkunft im Jahr 1998. Die

Variable «Jahr 2009» zeigt, wie gross die Abweichung der Einschätzung im Jahr 2009 gegenüber jener im Jahr 1998 ist.

Tabelle 3.3: Ergebnisse der Mehrebenenanalysen zum Vergleich der Selbsteinschätzungen zu Schul- und Unterrichtsmerkmalen

	Zufriedenheit mit Schule	Schulklima	Soziales Verhalten	Arbeitsklima im Unterricht
Fixe Effekte	b-Koeff	b-Koeff.	b-Koeff.	b-Koeff.
Konstante	3.21***	3.29***	2.83***	2.87***
Knaben	-0.11***	-0.07**	-0.12***	0.03
Deutsch als Zweitsprache	0.12***	0.04 ⁺	0.00	0.11***
Soziale Herkunft (Index)	0.02	0.00	0.00	-0.04**
Jahr 2009	-0.05	-0.06	0.11*	-0.11*
Variable Effekte				
Varianz Klassenebene	0.07***	0.05***	0.28***	0.07***
Varianz Schülerebene	0.35***	0.20***	0.57***	0.26***
Erklärte Varianz				
R ² (Schülerebene)	0.0%	2.6%	5.5%	5.6%
R ² (Klassenebene)	2.3%	2.3%	0.5%	2.2%
R ² (Gesamt)	1.2%	2.4%	1.5%	2.9%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler sind in Klassen gruppiert), Random-Intercept. Abhängige Variable ist die Bewertung der Schul- und Unterrichtskultur am Ende der sechsten Klasse. Die Variable «Soziale Herkunft (Index)» wurde am «Grand Mean» zentriert. Gewichtete Daten. Signifikanzniveaus: ⁺p < .10; *p < .05; **p < .01; ***p < .001.

Die Ergebnisse zeigen, dass sich die Bewertung der Schulzufriedenheit zwischen den beiden Erhebungen nicht unterscheidet. Knaben sind mit der Schule weniger zufrieden als Mädchen und Schülerinnen und Schüler mit Deutsch als Zweitsprache sind mit der Schule zufriedener als Schülerinnen und Schüler mit Deutsch als Erstsprache. Die Unterschiede zwischen den Schülergruppen sind als gering zu beurteilen ($d = 17$, $d = 0.18$). Die soziale Herkunft der Schülerinnen und Schüler hingegen steht in keiner systematischen Beziehung mit der Schulzufriedenheit.

Auch die Bewertung des Schulklimas unterscheidet sich nicht zwischen 1998 und 2009. Wiederum fällt die Einschätzung der Knaben gegenüber jener der Mädchen geringfügig negativer aus ($d = 0.07$). Die Erstsprache und die soziale Herkunft der Schülerinnen und Schüler hingegen sind für die Einschätzung des Schulklimas nicht von Bedeutung.

Die Selbsteinschätzungen zum sozialen Verhalten an der Schule, als Indikator für aggressives und gewalttätiges Verhalten, fällt im Jahr 2009 positiver aus als im Jahr 1998, wobei der Unterschied mit 0.11 Punkten als gering zu beurteilen ist ($d = 0.17$). Weiter zeigen die Ergebnisse, dass Knaben gegenüber Mädchen das soziale Verhalten geringfügig weniger aggressiv einschätzen ($d = 0.19$).

Werden die Schülerinnen und Schüler nach dem Arbeitsklima im Deutsch- und Mathematikunterricht gefragt, so fällt die Bewertung im Jahr 2009 statistisch signifikant negativer aus als im Jahr 1998. Der Unterschied beträgt jedoch lediglich 0.11 Punkte und ist als gering zu beurteilen ($d = 0.19$). Die Bewertung des Arbeitsklimas im Unterricht wird auch durch die Erstsprache und die soziale Herkunft der Schülerinnen und Schüler bestimmt. Schülerinnen und Schüler mit Deutsch als Erstsprache schätzen das Arbeitsklima geringfügig positiver ein als Schülerinnen und Schüler mit Deutsch als Erstsprache ($d = 0.19$). Die soziale Herkunft der Schülerinnen und Schüler steht in einer negativen Beziehung mit der Bewertung des Arbeitsklimas. Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto negativer wird das Arbeitsklima im Unterricht eingeschätzt. Bei einem Anstieg der sozialen Herkunft um eine Standardabweichung sinkt die Bewertung um 0.04 Punkte. Der Effekt der sozialen Herkunft ist zwar statistisch signifikant, jedoch unbedeutsam.

Weitere Analysen haben gezeigt, dass sich die Effekte der Lernvoraussetzungen auf die Bewertung der Schulzufriedenheit, des Schulklimas, des sozialen Verhaltens und des Arbeitsklimas im Mathematik- und Deutschunterricht zwischen den beiden Erhebungen nicht unterscheiden.

4 Analysen zur Bedeutung der Klassenzusammensetzung für die Leistungen

Domenico Angelone

4.1 Einleitung

In diesem Kapitel werden die statistischen Analysemodelle zur Bedeutung der Klassenzusammensetzung für die Leistungen am Ende der 6. Klasse einerseits und zur Wirkung des DaZ-Unterrichts andererseits präsentiert. Mit Blick auf die Bedeutung der Klassenzusammensetzung stehen zwei Fragestellungen im Vordergrund: (1) Gibt es einen Zusammenhang zwischen der sozialen Zusammensetzung einer Klasse und den Testleistungen der Schülerinnen und Schüler? (2) Gibt es einen Zusammenhang zwischen dem Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache einer Klasse und den Testleistungen? Die Analysen zum DaZ-Unterricht hingegen überprüfen die Frage, inwiefern sich der DaZ-Unterricht positiv auf die Deutschleistungen der Schülerinnen und Schüler auswirkt.

4.2 Methode

Datenbasis und statistische Auswertungsmethode zur Analyse von Klassenmerkmalen

Grundlage für die statistischen Analysen zur Bedeutung von Klassenmerkmalen bilden die Daten der repräsentativen Stichprobe der 6. Klassen der Zürcher Lernstandserhebung. Die Stichprobendaten erlauben neben repräsentativen Aussagen auf der Ebene der Schülerinnen und Schüler auch repräsentative Aussagen zu den 6. Klassen des Kantons Zürich.

Zur Analyse der Bedeutung von Klassenmerkmalen für die Leistungen der Schülerinnen und Schüler am Ende der 6. Klasse werden Variablen unterschiedlicher Messebenen herangezogen. Auf der Schülerebene (Ebene 1) interessiert der Einfluss von individuellen Merkmalen wie dem Geschlecht, dem Alter, der Erstsprache und der soziale Herkunft der Schülerinnen und Schüler. Zudem wird überprüft, ob unabhängig davon auch Klassenmerkmale (Ebene 2) einen Einfluss auf die Testleistungen ausüben. Die korrekte und gleichzeitige Analyse von Einflussfaktoren auf unterschiedlichen Messebenen erfolgt mit Hilfe der Mehrebenenanalyse (Raudenbush & Bryk, 2002). Bei den folgenden Mehrebenenmodellen werden die Schätzkoeffizienten der erklärenden Variablen alle fix gehalten, lediglich die Variation der Konstanten auf der Ebene 2 zwischen den Klassen wird zugelassen (random-intercept-model).

Datenbasis und statistische Auswertungsmethode zur Wirkungsanalyse des DaZ-Unterrichts

Die Datengrundlage für die Wirkungsanalysen zum DaZ-Unterricht bildet die Längsschnittstichprobe der Zürcher Lernstandserhebung zu Beginn der 1. Klasse bis zum Ende der 6. Klasse der Primarschule. In den Daten sind neben Angaben zu den schulischen Leistungen zu Beginn der 1. Klasse, am Ende der 3. Klasse und am Ende der 6. Klasse auch Angaben zum Besuch des DaZ-Unterrichts während der 3., 4., 5. und 6. Klasse enthalten.

Ziel der Analysen zum DaZ-Unterricht ist es, die (positive) kausale Wirkung des DaZ-Unterrichts auf die Testleistungen in Deutsch zu überprüfen. Bei der methodischen Umsetzung dieser Fragestellung stellt sich jedoch ein grundlegendes Problem (vgl. Caliendo & Kopeinig

2008; Rubin & Thomas, 1983). Ein Kind kann nicht gleichzeitig in beiden Zuständen, das heisst mit und ohne DaZ-Unterricht, beobachtet werden. In sogenannten kontrollierten Experimenten, bei denen die Probanden vor einer bestimmten Massnahme *zufällig* in eine Experimental- und in eine Kontrollgruppe zugeteilt werden, stellt sich dieses Problem nicht. Die Experimentalgruppe würde – am Beispiel der vorliegenden Fragestellung – aus Schülerinnen und Schülern mit Deutsch als Zweitsprache bestehen, die einen DaZ-Unterricht besuchen, die Kontrollgruppe hingegen aus Schülerinnen und Schülern mit Deutsch als Zweitsprache, die keinen DaZ-Unterricht besuchen. Weil die Zuteilung in eine der beiden Gruppen rein zufällig geschieht, kann (mit hoher Wahrscheinlichkeit) davon ausgegangen werden, dass sich die beiden Gruppen in ihrer Zusammensetzung der Schülerschaft sehr ähnlich sind. Im optimalen Fall unterscheiden sich die Kinder der beiden Gruppen einzig dadurch, dass jene der Kontrollgruppe keinen DaZ-Unterricht besuchen werden. Nach erfolgter Massnahme lassen sich die Experimental- und die Kontrollgruppe mit Blick auf die interessierenden Variablen, beispielsweise die Entwicklung der Deutschleistungen, miteinander vergleichen. Weil die Zuteilung zur Experimental- und zur Kontrollgruppe rein zufällig geschieht, können beobachtbare Unterschiede mit hoher Wahrscheinlichkeit als eine kausale Folge der erfolgten Massnahme – in diesem Beispiel als eine kausale Folge des besuchten DaZ-Unterrichts – interpretiert werden.

Im Rahmen von nicht-randomisierten Untersuchungsdesigns hingegen geschieht die Selektion in eine Experimental- und in eine Kontrollgruppe (mit hoher Wahrscheinlichkeit) nicht zufällig. Die Zusammensetzung der Schülerschaft kann sich in den beiden Gruppen deshalb systematisch voneinander unterscheiden. So ist beispielsweise bekannt, dass die Zuteilung zu sonderpädagogischen Massnahmen mit den schulischen Vorleistungen oder der sozialen Herkunft der Schülerinnen und Schüler zusammenhängt (vgl. z.B. Cloerkes 2003; Wagner & Powell 2003; Kronig, Eckhart & Haeberlin 2007). Wenn nun solche selektionsrelevante Merkmale auch einen Einfluss auf die interessierende Ergebnisvariable – beispielsweise die schulischen Leistungen – ausüben, dann führt ein einfacher Vergleich (z.B. Mittelwertvergleich) zwischen den beiden Gruppen zu verzerrten Ergebnissen. Aufgrund der Selbstselektion – beispielsweise werden Schülerinnen und Schüler mit geringen Vorschulleistungen eher einer sonderpädagogischen Massnahme zugewiesen – werden sich die beiden Gruppen auch ohne Teilnahme an einer Massnahme unterscheiden. Der direkte Vergleich der Experimental- und der «naiven» Kontrollgruppe kann deshalb irreführend sein.

Weil der vorliegenden Untersuchung kein experimentelles Untersuchungsdesign zugrunde liegt, werden die Wirkungsanalysen zum DaZ-Unterricht mit Hilfe eines sogenannten *Matching*-Ansatzes vorgenommen. *Matching* stellt eine Methode dar, bei der aus einer Gruppe von potenziellen Vergleichsprobanden diejenigen ausgewählt werden, die den Probanden der Experimentalgruppen möglichst ähnlich sind, sogenannte «statistische Zwillinge». Damit wird eine Kontrollgruppe gebildet, die in ihrer Zusammensetzung möglichst gut mit derjenigen der Experimentalgruppe übereinstimmt. Auf die vorliegende Untersuchung bezogen heisst das, dass jedem Kind mit Deutsch als Zweitsprache mit DaZ-Unterricht (Experimentalgruppe) ein möglichst ähnliches Kind (z.B. hinsichtlich schulischen Vorleistungen, kognitiven Fähigkeiten, sozialer Herkunft und Geschlecht) ohne DaZ-Unterricht zugeteilt wird (Kontrollgruppe). Die Suche nach statistischen Zwillingen beziehungsweise das *Matching* wird jedoch ab einer bestimmten Anzahl selektionsrelevanter Merkmale – die zwischen den beiden Gruppen konstant gehalten werden müssen – schnell undurchführbar. Mit $k = 4$ dichotomen selektionsrelevanten Variablen beispielsweise entstehen bereits 16 (2^4) *Matching*-Kategorien.

Rubin und Thomas (1983) konnten jedoch zeigen, dass das *Matching* mit Hilfe des sogenannten *Propensity scores* möglich ist. Der *Propensity score* bezeichnet die Wahrscheinlichkeit, dass ein Proband an einem Massnahmenprogramm teilnimmt. Für die vorliegende Untersuchung entspricht der *Propensity score* der Wahrscheinlichkeit, dass ein Kind mit Deutsch als Zweitsprache dem DaZ-Unterricht zugeteilt wurde. Für das eigentliche *Matching* wird dann nur noch der *Propensity score* verwendet. Das heisst, den Schülerinnen und Schülern mit DaZ-Unterricht werden Kinder ohne DaZ-Unterricht zugeordnet, deren *Propensity scores* möglichst gleich sind. Der Hauptvorteil dieses Ansatzes (gegenüber der kategorialen Bildung der Kontrollgruppe) besteht darin, dass ohne grösseren Aufwand gleichzeitig mehrere selektionsrelevante Variablen berücksichtigt werden können.

Zur Schätzung des *Propensity scores* kann die binär logistische Regression verwendet werden (vgl. z.B. Tabachnik & Fidell, 2007). Durch Einbezug möglichst vieler selektionsrelevanter Variablen –beispielsweise der schulischen Vorleistungen – soll die Wahrscheinlichkeit der Zuteilung zum DaZ-Unterricht möglichst gut geschätzt werden. Dabei stehen nicht die einzelnen Effekte der einbezogenen Variablen im Vordergrund, sondern der Erklärungsgehalt des logistischen Regressionsmodells. Im Rahmen der logistischen Regression wird die Erklärungskraft mittels des sogenannten Pseudo-R² – in Anlehnung an das bekannte R² der linearen Regression – ausgewiesen. Von Interesse ist somit ein möglichst hohes Pseudo-R². Je höher der Erklärungsgehalt des Schätzmodells ist, desto besser passt das Modell zu den Daten. Ein höherer Erklärungsgehalt geht deshalb mit einer besseren Schätzung der *Propensity scores* und deshalb mit einem besseren *Matching* einher. Bei den in den folgenden logistischen Regressionsmodellen ausgewiesenen Pseudo-R² handelt es sich um das sogenannte *McFaddens* Pseudo-R², das Werte zwischen 0 und 1 einnehmen kann. Im Allgemeinen kann ein Pseudo-R² von 0.2 bis 0.4 bereits als sehr gute Erklärungskraft interpretiert werden (Tabachnik & Fidell, 2007). Gemäss Rubin und Thomas (1996) sollten möglichst alle verfügbaren Variablen in die Schätzung der *Propensity scores* einbezogen werden, sofern nicht eindeutig nachgewiesen werden kann, dass sie keinen Einfluss auf die Zuteilung in die Experimentalgruppe ausüben. Die Autoren empfehlen zudem, auch jene Variablen in der Schätzung beizubehalten, die keinen statistisch signifikanten Einfluss auf die Selektion aufweisen.

Nach der Schätzung der *Propensity scores* kann das eigentliche *Matching* vorgenommen werden, das heisst die Bildung der Kontrollgruppe. Für das *Matching* mit Hilfe der *Propensity scores* werden in der Literatur verschiedene Verfahren vorgeschlagen (für eine Übersicht vgl. Caliendo & Kopeinig, 2008). Für die vorliegende Untersuchung wurde das *Nearest neighbour matching with replacement*-Verfahren verwendet. Dabei wird jedem Kind aus der Experimentalgruppe jenes Kind aus der Kontrollgruppe als *Matching*-Partner zugeordnet, dessen *Propensity score* am nächsten liegt. Die *Replacement*-Option erlaubt, dass ein Kontrollkind gleichzeitig für mehrere Kinder aus der Experimentalgruppe als *Matching*-Partner verwendet werden kann. Dadurch erhöht sich einerseits die *Matching*-Qualität, andererseits führt dies zu einer Reduktion der Anzahl Beobachtungen und deshalb zu einer kleineren Varianz innerhalb der Kontrollgruppe. Falls ein Kind mehrmals als *Matching*-Partner ausgewählt wird, muss es ein entsprechendes Gewicht in der Kontrollgruppe erhalten.¹

¹ Wird ein Kontrollkind beispielsweise zweimal verwendet, erhält es ein Gewicht von 2.

Das *Matching* wird im Folgenden nur im Bereich des sogenannten gemeinsamen Stützbereichs (*common support*) vorgenommen. Damit wird sichergestellt, dass nur Kinder der Kontrollgruppe berücksichtigt werden, die tatsächlich auch mit Kindern der Experimentalgruppe vergleichbar sind. Die Definition des gemeinsamen Stützbereichs wird mit Hilfe der Minimum-Maximum-Regel definiert. Diese besagt, dass jene Kinder für das *Matching* ausgeschlossen werden, deren *Propensity score* kleiner als das Minimum und grösser als das Maximum der *Propensity scores* der Kontroll- beziehungsweise der Experimentalgruppe ist.²

Nach dem *Matching* gilt es noch die *Matching*-Qualität zu überprüfen, das heisst die Vergleichbarkeit der Experimental- und der Kontrollgruppe. Dafür können anhand von einfachen t-Tests die Mittelwerte der selektionsrelevanten Variablen zwischen den beiden Gruppen verglichen werden. Während sich vor dem *Matching* die beiden Gruppen unterscheiden können, dürfen sich nach dem *Matching* keine statistisch signifikanten Unterschiede mehr zeigen. Eine weitere Qualitätsprüfung ist mit Hilfe des Pseudo-R² und der Gesamtsignifikanz möglich. Nach dem *Matching* kann das logistische Regressionsmodell zur Schätzung der *Propensity scores* nochmals geschätzt und die Pseudo-R² vor und nach dem *Matching* können miteinander verglichen werden. Während das Pseudo-R² vor dem *Matching* möglichst hoch sein sollte, muss es nach dem *Matching* möglichst gering ausfallen – die beiden Gruppen sollten sich hinsichtlich der berücksichtigten Variablen nicht mehr unterscheiden. Weiter kann mit Hilfe eines *Likelihood-ratio*-Tests die globale Nullhypothese getestet werden, gemäss derer alle in das Regressionsmodell einbezogenen Schätzparameter gleich 0 sind. Während dieser Test vor dem *Matching* statistisch signifikant ausfallen sollte, sollte er nach dem *Matching* nicht mehr statistisch signifikant sein, das heisst, die Nullhypothese darf nicht mehr verworfen werden.

Nach erfolgreichem *Matching* ist anhand der Experimental- und der neu gebildeten Kontrollgruppe die eigentliche Schätzung eines kausalen Effekts einer Massnahme möglich. Weil die beiden Gruppen nun miteinander vergleichbar sind, kann dies mittels einfachen Mittelwertvergleichen in den interessierenden Variablen erfolgen. Zwar lässt sich nicht für jedes Kind der individuelle Effekt des DaZ-Unterrichts schätzen. Möglich ist aber die Schätzung des durchschnittlichen Effekts des DaZ-Unterrichts auf die Experimentalgruppe (*Average Effect of Treatment on the Treated [ATT]*).

Tabelle 4.1 fasst die einzelnen Schritte des angewandten *Matching*-Verfahrens zur Schätzung des Effekts des DaZ-Unterrichts nochmals zusammen.

² Wenn beispielsweise der *Propensity score* in der Experimentalgruppe zwischen .07 und .94 liegt und in der Kontrollgruppe zwischen .04 und .89, werden für das *Matching* nur Beobachtungen berücksichtigt, deren *Propensity score* zwischen .07 und .89 liegt (= *Common support region*).

Tabelle 4.1 Vorgehen bei der *Matching*-Analyse

Schritt 1	Schätzung der <i>Propensity scores</i> mit Hilfe der binär logistischen Regression
Schritt 2	Bildung von statistischen Zwillingen (Kontrollgruppe) mit dem Verfahren des <i>Nearest neighbour matching</i> mit Hilfe der geschätzten <i>Propensity scores</i> und Überprüfung der <i>Matching</i> -Qualität
Schritt 3	Schätzung des kausalen Effekts (<i>ATT</i>) des DaZ-Unterrichts auf die interessierenden Variablen mittels Mittelwertsvergleichen zwischen der Experimental- und der neu gebildeten Kontrollgruppe.

4.3 Ergebnisse zur Bedeutung der Klassenzusammensetzung für die Leistungen am Ende der 6. Klasse

Ergebnisse zum Einfluss der Klassenzusammensetzung auf die Leistungen in Deutsch

Tabelle 4.2 zeigt die Ergebnisse der Mehrebenenanalysen zur Bedeutung der Klassenzusammensetzung für die Testleistungen in Deutsch am Ende der 6. Klasse. In der ersten Spalte sind die Variablen aufgeführt, deren Effekte auf die Leistungen überprüft wurden. Die folgenden Spalten zeigen die unstandardisierten Regressionskoeffizienten.

Die Ergebnisse des Nullmodells sind in der zweiten Spalte aufgeführt. Beim Nullmodell werden keine erklärenden Variablen ins Modell eingeführt. Es enthält lediglich eine Regressionskonstante, die zwischen den Kontexteinheiten (Klassen) variiert. Die Konstante entspricht den durchschnittlichen Klassenleistungen am Ende der 6. Klasse (rund 787 Punkte). Anhand des Nullmodells lässt sich die Gesamtvarianz in den Leistungen in zwei Bestandteile zerlegen: die Varianz zwischen den Schülerinnen und Schülern und die Varianz zwischen den Klassen. Daraus lässt sich der Varianzanteil in den Leistungen bestimmen, der auf Unterschiede zwischen den Klassen zurückzuführen ist. Die Varianz der klassenspezifischen Regressionskonstante, die der Varianz der Klassenmittelwerte entspricht, beträgt 1249 Punkte. Die Varianz innerhalb der Klassen beträgt 8677 Punkte. Entsprechend können rund 13 Prozent der Gesamtvarianz in den Deutschleistungen auf Unterschiede zwischen Klassen und rund 87 Prozent auf Unterschiede zwischen Schülerinnen und Schülern innerhalb der Klassen zurückgeführt werden. Dies bedeutet, dass maximal 13 Prozent der Leistungsunterschiede anhand von Klassenmerkmalen erklärt werden können. Die restlichen Unterschiede müssen anhand von Merkmalen der Schülerinnen und Schüler erklärt werden.

Tabelle 4.2: Ergebnisse der Mehrebenenanalysen zum Einfluss der Klassenzusammensetzung auf die Leistungen in Deutsch am Ende der 6. Klasse

	Nullmodell	Modell I	Modell II	Modell III	Modell IV
Fixe Effekte für die Konstante	786.2***	807.9***	801.9***	800.3***	797.5***
Schülerinnen- und Schülermerkmale					
Knaben		-22.4***	-22.6***	-22.4***	-22.6***
Alter in Monaten		-3.0***	-2.9***	-3.0***	-2.9***
Deutsch als Zweitsprache (DaZ)		-38.8***	-38.8***	34.9***	-35.7***
Soziale Herkunft (Index)		35.6***	33.9***	35.2***	33.7***
Klassenmerkmale					
Privilegierte soziale Zusammensetzung (86.5 bis 100 Perzentil)			44.6***		40.4***
Anteil Deutsch als Zweitsprache (DaZ) (in 10 Prozent)				-4.8*	-2.9†
Variable Effekte					
Klassenebene	1249***	461***	237***	387***	214**
Schülerinnen- und Schülerebene	8677	6215	6208	6211	6206
Erklärte Varianz					
R ² (Individualmerkmale)		28.4%	28.5%	28.4%	28.5%
R ² (Klassenmerkmale)		63.1%	81.0%	69.0%	82.9%
R ² (Gesamt)		32.7%	35.1%	33.5%	35.3%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler in Klassen, Random-Intercept). Abhängige Variable sind die Leistungen in Deutsch (5 «Plausible Values»). Alle metrischen Variablen wurden am «Grand Mean» zentriert. Zwei Klassen aus Analyse ausgeschlossen (Ausreisser). Gewichtete Daten. Signifikanzniveaus: †p < .10; *p < .05; **p < .01; ***p < .001.

In Modell I wurden die Lernvoraussetzungen der Schülerinnen und Schüler in die Analyse einbezogen, das heisst, deren Wirkung auf die Leistungen überprüft. Als individuelle Lernvoraussetzungen wurden das Geschlecht, das Alter, die Erstsprache sowie die soziale Herkunft der Schülerinnen und Schüler berücksichtigt. Am Ende der 6. Klasse sind die Deutschleistungen der Knaben rund 22 Punkte tiefer als jene der Mädchen. Je älter die Kinder sind, desto tiefer sind auch ihre Leistungen. Ein Altersunterschied von einem Monat entspricht einem Rückstand von rund 3 Punkten. Die älteren Kinder in der Stichprobe sind eher die lernschwächeren Schülerinnen und Schüler. Zum einen weil sie später eingeschult wurden, zum anderen weil sie Klassen wiederholen mussten. Auch die Erstsprache ist für die Leistungen in Deutsch von Bedeutung. Schülerinnen und Schüler mit Deutsch als Zweitsprache erzielen ein um rund 39 Punkte schlechteres Ergebnis als jene Schülerinnen und Schüler mit Deutsch als Erstsprache. Schliesslich hängen die Deutschleistungen auch mit der sozialen Herkunft der Schülerinnen und Schüler zusammen. Je höher die soziale Herkunft, desto besser sind auch

die Deutschleistungen. Bei einem Anstieg des Indexes zur sozialen Herkunft um eine Standardabweichung (eine Einheit) steigen die Deutschleistungen um rund 36 Punkte an.

Der Einbezug der individuellen Lernvoraussetzungen erklärt rund 33 Prozent der Gesamtvarianz in den Deutschleistungen. Von der Varianz zwischen den Schülerinnen und Schülern innerhalb der Klassen konnten rund 28 Prozent erklärt werden. Weil sich die Klassen in ihrer Zusammensetzung aufgrund individueller Lernvoraussetzungen stark voneinander unterscheiden, konnte in Modell I auch ein beträchtlicher Teil der Varianz zwischen den Klassen erklärt werden (rund 63 Prozent). Nach Berücksichtigung der Lernvoraussetzungen der Schülerinnen und Schüler wird die Varianz zwischen den Klassen noch auf 461 Punkte geschätzt. Die Deutschleistungen der Klassen unterscheiden sich auch bei gleichen Lernvoraussetzungen ihrer Schülerschaft weiterhin voneinander.

Zusätzlich wurde überprüft, ob Klassenmerkmale wie die soziale Zusammensetzung einer Klasse oder der Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache einen Beitrag zur Erklärung der verbleibenden Leistungsunterschiede zwischen den Klassen leisten können. Weil die beiden Klassenmerkmale relativ stark miteinander zusammenhängen, wurde deren Einfluss auf die Leistungen in einem ersten Schritt jeweils getrennt überprüft.

Abbildung 4.1: Durchschnittliche Klassenleistungen in Deutsch, bereinigt um individuelle Lernvoraussetzungen, nach der sozialen Zusammensetzung einer Klasse

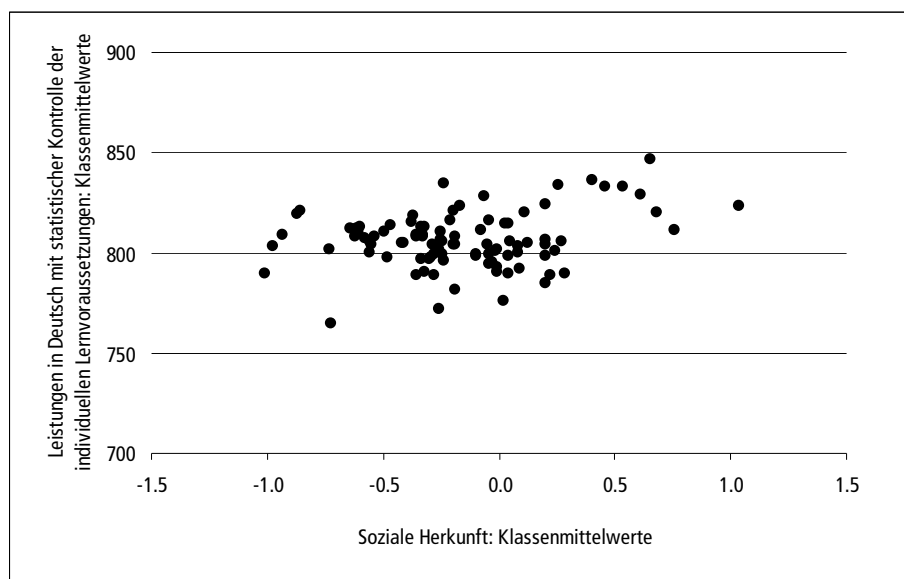


Abbildung 4.1 zeigt, wie die Deutschleistungen nach statistischer Kontrolle der Lernvoraussetzungen mit der sozialen Zusammensetzung der Klassen zusammenhängen. Auf der y-Achse sind die anhand des Modells I vorausgesagten Deutschleistungen einer Klasse abgebildet. Die x-Achse repräsentiert die soziale Zusammensetzung der Klassen, gemessen an der durchschnittlichen sozialen Herkunft ihrer Schülerschaft. Abbildung 4.1 deutet darauf hin, dass sich die soziale Zusammensetzung einer Klasse erst ab einem kritischen Schwellenwert, nämlich bei einem Wert mehr als 0.40 Indexpunkten auf die Leistungen in Deutsch auswirkt. Ein Wert höher als 0.40 Indexpunkte entspricht in der Population der 6. Klassen des Kantons Zürich den 13.5 Prozent Klassen mit privilegiertester sozialer Zusammensetzung. Anders aus-

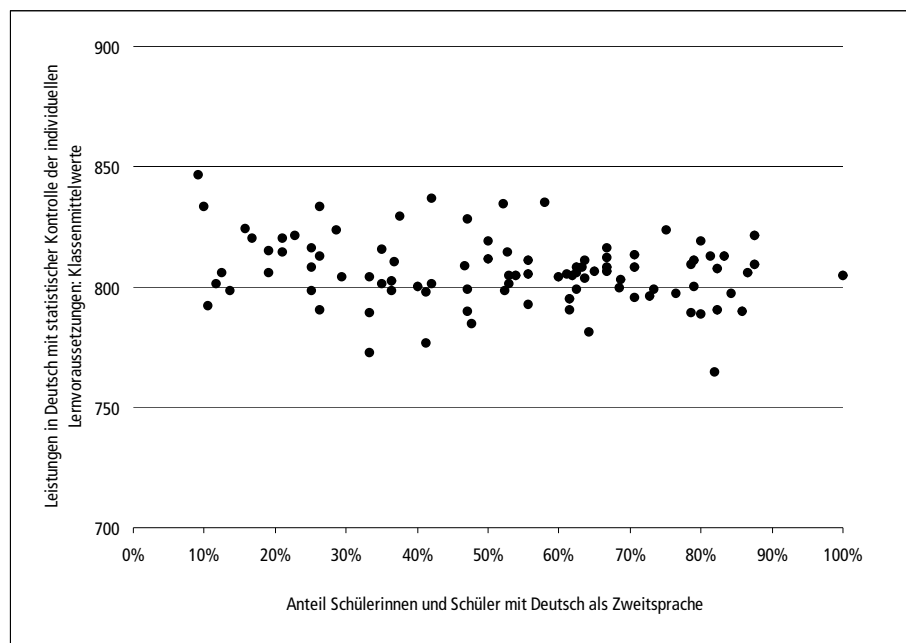
gedrückt: 13.5 Prozent aller Klassen weisen einen Indexwert zur sozialen Zusammensetzung auf, der grösser als 0.40 ist und entsprechend weisen 86.5 Prozent aller Klassen einen Indexwert zur sozialen Zusammensetzung auf, der kleiner als 0.40 ist.

In Modell II wurde deshalb der Effekt der sozialen Zusammensetzung einer Klasse anhand einer Dummy-Variablen überprüft, bei der Klassen mit einer sozialen Zusammensetzung grösser als 0.40 Indexpunkte mit Klassen mit einer sozialen Zusammensetzung kleiner als 0.40 Indexpunkte verglichen wurden. Dabei zeigt sich ein deutlicher und statistisch signifikanter Effekt der sozialen Zusammensetzung einer Klasse auf die Leistungen. Die Deutschleistungen der Schülerinnen und Schüler in Klassen mit einer sozialen Zusammensetzung höher als 0.40 Indexpunkte – was in der Population der 6. Klassen denjenigen Klassen mit dem 13.5 Prozent höchsten Werten bei der sozialen Zusammensetzung entspricht – sind rund 45 Punkte höher als jene der Kinder in Klassen mit einer tieferen sozialen Zusammensetzung. Dabei muss jedoch berücksichtigt werden, dass der beobachtete Effekt aufgrund von nur neun Klassen in der Stichprobe zustande gekommen ist. Vertiefte Analysen anhand einer Stichprobe mit mehreren Klassen mit einer solch hohen sozialen Zusammensetzung wären von Vorteil. Der Einbezug der sozialen Zusammensetzung der Klassen vermag rund 81 Prozent der Varianz in den Deutschleistungen zwischen den Klassen zu erklären.

Der Effekt der sozialen Zusammensetzung einer Klasse auf die Deutschleistungen zeigt sich unabhängig von individuellen Lernvoraussetzungen wie der sozialen Herkunft oder der Erstsprache der Schülerinnen und Schüler und kommt deshalb zu allfälligen weiteren Effekten hinzu. Wenn beispielsweise eine Schülerin mittleren Alters mit Deutsch als Erstsprache und mit durchschnittlicher sozialer Herkunft eine Klasse mit einer sozialen Zusammensetzung von weniger als 0.40 Indexpunkten besucht, fallen ihre Leistungen um rund 45 Punkte tiefer aus, als wenn dieselbe Schülerin eine Klasse mit einer sozialen Zusammensetzung von grösser als 0.40 Indexpunkte besucht.

Abbildung 4.2 zeigt die durchschnittlichen Deutschleistungen einer Klasse nach statistischer Kontrolle der Lernvoraussetzungen nach dem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache. Auf der y-Achse sind die anhand des Modells I vorausgesagten durchschnittlichen Deutschleistungen einer Klasse abgebildet, bereinigt um die Lernvoraussetzungen der Schülerinnen und Schüler. Die x-Achse repräsentiert den Anteil Schülerinnen und Schüler mit DaZ einer Klasse. Die Abbildung 4.2 deutet darauf hin, dass mit steigendem Anteil Schülerinnen und Schüler mit DaZ auch die Deutschleistungen einer Klasse sinken. Im Gegensatz zur sozialen Herkunft einer Klasse (vgl. Abbildung 4.1) kann von einem linearen Zusammenhang zwischen dem Anteil Schülerinnen und Schülern mit Deutsch als Zweitsprache und den Deutschleistungen einer Klasse ausgegangen werden.

Abbildung 4.2: Durchschnittliche Klassenleistungen in Deutsch, bereinigt um individuelle Lernvoraussetzungen, nach dem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache



Modell III überprüft den Effekt des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache einer Klasse auf die Leistungen. Dabei bestätigen sich die bisherigen Ergebnisse. Je höher der Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache ist, desto tiefer sind die Leistungen. Mit einem Anstieg des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache einer Klasse um 10 Prozent verschlechtern sich die durchschnittlichen Leistungen um rund 5 Punkte. Wenn beispielsweise eine Schülerin mittleren Alters mit Deutsch als Erstsprache und mit durchschnittlicher sozialer Herkunft eine Klasse mit einem Anteil Schülerinnen und Schüler mit DaZ von 80 Prozent besucht, fallen ihre Deutschleistungen um rund 29 Punkte tiefer aus ($-29 = 6 \cdot -4.8$), als wenn dieselbe Schülerin eine Klasse mit einem Anteil Schülerinnen und Schüler mit DaZ von 20 Prozent besucht. Mit rund 69 Prozent erklärter Varianz zwischen den Klassen ist der Erklärungsgehalt des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache einer Klasse im Vergleich zur sozialen Zusammensetzung (81 Prozent) deutlich geringer.

In Modell IV schliesslich wurden beide Klassenmerkmale gemeinsam in die Analyse einbezogen. Dabei reduziert sich der Effekt der sozialen Zusammensetzung einer Klasse nur geringfügig und ist statistisch signifikant. Der Effekt des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache hingegen ist deutlich stärker zurückgegangen (um rund 40%) und ist statistisch nicht mehr signifikant ($p = .07$). Für die Stichprobe zeigt sich jedoch weiterhin ein negativer Effekt. Mit einem Anstieg des Anteils Schülerinnen und Schüler mit DaZ einer Klasse um 10 Prozent verschlechtern sich die durchschnittlichen Leistungen um rund 3 Punkte. Dieses Ergebnis lässt sich jedoch aufgrund der fehlenden statistischen Signifikanz nicht ohne Bedenken auf die Grundgesamtheit verallgemeinern.

Ergebnisse zum Einfluss der Klassenzusammensetzung auf die Leistungen in der Mathematik

Tabelle 4.3 zeigt die Ergebnisse der Mehrebenenanalysen zur Bedeutung der Klassenzusammensetzung für die Testleistungen in Deutsch am Ende der 6. Klasse. Die Varianzkomponenten des Nullmodells zeigen, dass sich die Leistungen zwischen den Klassen auch in der Mathematik unterscheiden. Die Varianz zwischen den Klassen beträgt 884 Punkte, jene zwischen den Schülerinnen und Schülern innerhalb der Klassen 8650 Punkte. In der Mathematik können deshalb maximal rund 9 Prozent der Gesamtvarianz in den Leistungen aufgrund von Klassenmerkmalen erklärt werden. Das sind rund 4 Prozent weniger als bei den Leistungen in Deutsch.

Tabelle 4.3: Ergebnisse der Mehrebenenanalysen zum Einfluss der Klassenzusammensetzung auf die Leistungen in Mathematik am Ende der 6. Klasse

	Nullmodell	Modell I	Modell II	Modell III	Modell IV
Fixe Effekte für die Konstante	884.0***	881.9***	878.1***	876.0***	875***
Schülerinnen- und Schülermerkmale					
Knaben		19.7***	19.7***	19.8***	19.8***
Alter in Monaten		-3.4***	-3.4***	-3.4***	-3.4***
Deutsch als Zweitsprache (DaZ)		-29.5***	-29.1***	-26.1***	-26.6***
Soziale Herkunft (Index)		31.5***	30.3***	31.1***	30.1***
Klassenmerkmale					
Privilegierte soziale Zusammensetzung (86.5 bis 100 Perzentil)			24.1*		20.6 ⁺
Anteil Deutsch als Zweitsprache (DaZ) (in 10 Prozent)				-3.5 ⁺	-2.5
Variable Effekte					
Klassenebene	884***	344***	280***	306	262
Schülerinnen- und Schülerebene	8650	6866	6863	6866	6864
Erklärte Varianz					
R ² (Individualmerkmale)		20.6%	20.7%	20.6%	20.6%
R ² (Klassenmerkmale)		61.0%	68.4%	65.4%	70.4%
R ² (Gesamt)		24.4%	25.1%	24.8%	25.2%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler in Klassen, Random-Intercept). Abhängige Variable sind die Leistungen in der Mathematik (5 «Plausible Values»). Alle metrischen Variablen wurden am «Grand Mean» zentriert. Eine Klasse aus Analyse ausgeschlossen (Ausreisser). Gewichtete Daten. Signifikanzniveaus: *p < .10; **p < .05; ***p < .01; ****p < .001.

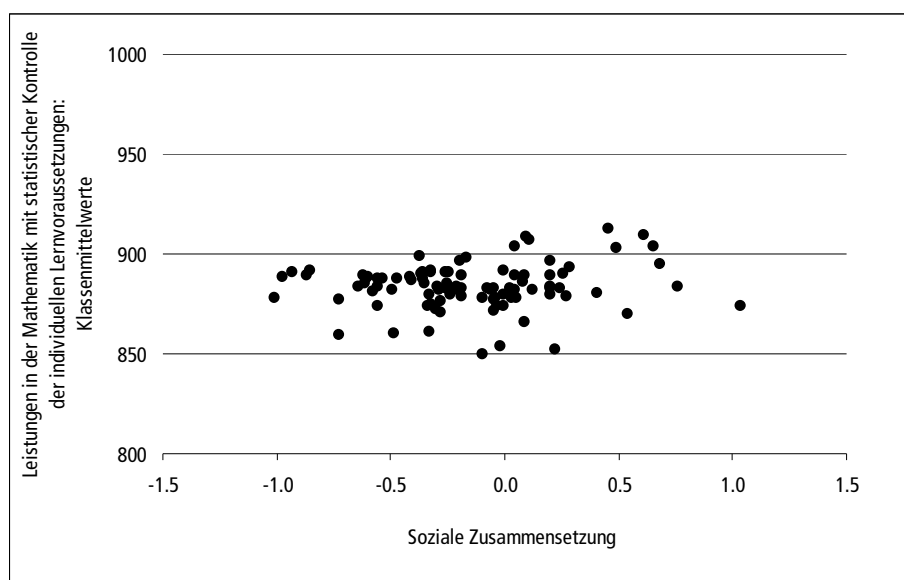
In Modell I wurden die Lernvoraussetzungen der Schülerinnen und Schüler in die Analyse einbezogen. Die Mathematikleistungen der Knaben fallen rund 20 Punkte höher aus als jene der Mädchen. Auch in der Mathematik zeigt sich ein statistisch signifikant negativer Effekt des Alters auf die Leistungen der Schülerinnen und Schüler. Ein Altersunterschied von einem

Monat entspricht einem Rückstand von rund 3 Punkten. Die Erstsprache wirkt sich in der Mathematik weniger stark aus als im Deutsch. Der Effekt bleibt jedoch beträchtlich. Die Leistungen der Schülerinnen und Schüler mit Deutsch als Zweitsprache sind rund 30 Punkte tiefer als jene mit Deutsch als Erstsprache. Schliesslich unterscheiden sich die Mathematikleistungen auch nach der sozialen Herkunft der Schülerinnen und Schüler. Je höher die soziale Herkunft, desto höher sind auch die Leistungen. Bei einem Anstieg des Indexes zur sozialen Herkunft um 1 Standardabweichung steigen die Leistungen in Mathematik um rund 32 Punkte an.

Insgesamt vermögen die Lernvoraussetzungen der Schülerinnen und Schüler rund 24 Prozent der Gesamtvarianz in den Leistungen zu erklären. Zwischen den Schülerinnen und Schüler innerhalb der Klassen beträgt die erklärte Varianz rund 21 Prozent. Auch nach Berücksichtigung der Lernvoraussetzungen der Schülerinnen und Schülern können noch Unterschiede in den Mathematikleistungen zwischen den Klassen festgestellt werden.

Abbildung 4.3 zeigt die durchschnittlichen Mathematikleistungen einer Klasse nach statistischer Kontrolle der Lernvoraussetzungen nach der sozialen Zusammensetzung der Klassen. Auf der y-Achse sind wiederum die anhand des Modells I geschätzten durchschnittlichen Mathematikleistungen einer Klasse abgebildet. Die x-Achse repräsentiert die soziale Zusammensetzung einer Klasse. Dabei zeigt sich ein ähnliches Zusammenhangsmuster wie im Deutsch. Auch in der Mathematik scheint sich die soziale Zusammensetzung einer Klasse erst ab einem kritischen Schwellenwert auf die Leistungen auszuwirken.

Abbildung 4.3: Durchschnittliche Klassenleistungen in Mathematik, bereinigt um individuelle Lernvoraussetzungen, nach der sozialen Zusammensetzung einer Klasse

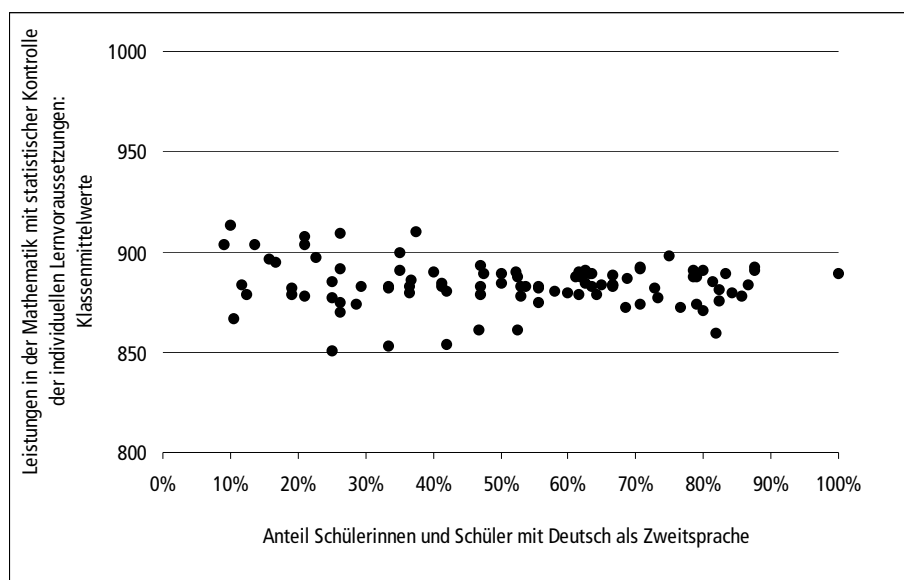


In Modell II wurde deshalb der Effekt der sozialen Zusammensetzung einer Klasse auf die Mathematikleistungen wie im Deutsch anhand einer Dummy-Variablen überprüft, bei der die Klassen mit einer sehr privilegierten sozialen Zusammensetzung von grösser als 0.40 Indexpunkte mit den restlichen Klassen kontrastiert werden. Die Ergebnisse zeigen, dass die Mathematikleistungen der Schülerinnen und Schüler in Klassen mit einer sozialen Zusammensetzung

zung von höher als 0.40 Indexpunkte rund 24 Punkte höher ausfallen als jene der Schülerinnen und Schüler in Klassen mit einer sozialen Zusammensetzung von weniger als 0.40 Indexpunkte. Wenn beispielsweise eine Schülerin mittleren Alters mit Deutsch als Erstsprache und durchschnittlicher sozialer Herkunft eine Klasse mit einer sozialen Zusammensetzung von weniger als 0.40 Indexpunkte besucht – was in der Population den Klassen mit der 13.5 Prozent höchsten sozialen Zusammensetzung entspricht –, fallen ihre Mathematikleistungen um rund 24 Punkte tiefer aus, als wenn dieselbe Schülerin eine Klasse mit einer tieferen sozialen Zusammensetzung besucht. Wiederum muss dabei in Betracht gezogen werden, dass der postulierte Effekt nur auf der Basis von neun Klassen in der Stichprobe geschätzt wurde. Die Berücksichtigung der sozialen Zusammensetzung einer Klasse vermag rund 68 Prozent der Varianz zwischen den Klassen zu erklären.

Abbildung 4.4 zeigt, wie die Mathematikleistungen nach Berücksichtigung der Lernvoraussetzungen mit dem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache zusammenhängen. Auf der y-Achse sind die anhand des Modells I geschätzten durchschnittlichen Klassenleistungen in Mathematik dargestellt, bereinigt um die Lernvoraussetzungen der Schülerinnen und Schüler. Die x-Achse repräsentiert den Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache einer Klasse. Daraus geht hervor, dass ein linear negativer Zusammenhang zwischen dem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache und den Mathematikleistungen einer Klasse plausibel ist. Je höher der Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache, desto geringer sind auch die Mathematikleistungen.

Abbildung 4.4: Durchschnittliche Klassenleistungen in Mathematik, bereinigt um individuelle Lernvoraussetzungen, nach dem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache



In Modell III wurde der Effekt des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache auf die Leistungen in Mathematik überprüft. Die Ergebnisse deuten relativ stark darauf hin, dass mit steigendem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache einer Klasse auch die Mathematikleistungen sinken. Bei einem Anstieg des Anteils Schülerin-

nen und Schüler mit Deutsch als Zweitsprache um 10 Prozent sinken die Leistungen um rund 3.5 Punkte. Wenn beispielsweise eine Schülerin mittleren Alters mit Deutsch als Erstsprache und mit durchschnittlicher sozialer Herkunft eine Klasse mit einem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache von 80 Prozent besucht, fallen ihre Mathematikleistungen um rund 29 Punkte tiefer aus ($-21 = 6^* - 3.5$), als wenn dieselbe Schülerin eine Klasse mit einem Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache von 20 Prozent besucht. Der Effekt verfehlt jedoch mit einer Irrtumswahrscheinlichkeit von $p = .055$ knapp die statistische Signifikanz. Die Berücksichtigung des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache vermag rund 65 Prozent der Leistungsunterschiede zwischen den Klassen zu erklären. In der Mathematik ist die Erklärungskraft des Anteils Schülerinnen und Schüler mit Deutsch als Zweitsprache im Vergleich zu jener der sozialen Zusammensetzung einer Klasse nur geringfügig kleiner.

Werden beiden Klassenmerkmale gleichzeitig in die Analyse einbezogen, ist keiner der beiden Effekte mehr statistisch signifikant (Modell IV). Für die Stichprobe zeigt sich jedoch weiterhin ein deutlicher Effekt der sozialen Zusammensetzung einer Klasse ($p = .09$) auf die Mathematikleistungen. Die Ergebnisse der Schülerinnen und Schüler in Klassen mit einer sehr privilegierten sozialen Zusammensetzung (> 0.40 Indexpunkte) sind rund 21 Punkte besser als jene der Kinder in Klassen mit einer weniger privilegierten sozialen Zusammensetzung. Das Ergebnis zum Anteil Schülerinnen und Schüler mit Deutsch als Zweitsprache einer Klasse verfehlt die statistische Signifikanz deutlich stärker ($p = .18$).

4.4 Ergebnisse zum DaZ-Unterricht

Das folgende Kapitel enthält die Ergebnisse der Analysen zur Wirkung des «Deutsch als Zweitsprache (DaZ)»-Unterrichts auf die Testleistungen in Deutsch. Tabelle 4.4 zeigt für die Längsschnittstichprobe der Zürcher Lernstandserhebung, wie viele Kinder mit Deutsch als Zweitsprache den DaZ-Unterricht besucht haben. Im Schuljahr 2005/06 (3. Klasse) besuchten 279 Kinder, im Schuljahr 2006/07 (4. Klasse) 112 Kinder, im Schuljahr 2007/08 (5. Klasse) 59 Kinder und im Schuljahr 2008/09 (6. Klasse) noch 33 Kinder den DaZ-Unterricht.

Tabelle 4.4 Anzahl Kinder mit Deutsch als Zweitsprache mit DaZ-Unterricht (Längsschnittstichprobe)

	Anzahl Schülerinnen und Schüler mit DaZ-Unterricht	Prozent aller Schülerinnen und Schüler mit Deutsch als Zweitsprache, die den DaZ-Unterricht besuchen
Schuljahr 2005/06: 3. Klasse	279	48.9%
Schuljahr 2006/07: 4. Klasse	112	18.8%
Schuljahr 2007/08: 5. Klasse	59	9.9%
Schuljahr 2008/09: 6. Klasse	33	5.5%

Anmerkungen: Grundgesamtheit sind alle Schülerinnen und Schüler mit Deutsch als Zweitsprache (N=596), von denen auch gültige Angaben zur Erstsprache vorhanden sind.

Die im Kapitel 4.2 beschriebenen Untersuchungsschritte werden in einer ersten Analyse zur Wirkung des DaZ-Unterrichts in der 3. Klasse auf die Leistungen am Ende der 3. Klasse nochmals vorgestellt und kommentiert: (1) Schätzung der *Propensity scores*, (2) *Matching* und Überprüfung der *Matching*-Qualität und (3) Schätzung des Massnahmeneffekts. Bei den weiteren Analysen werden im Interesse der Lesbarkeit nur noch die Ergebnisse zu den Effekten des DaZ-Unterrichts präsentiert.

Schätzung der Propensity Scores

Zur Schätzung des Effekts des DaZ-Unterrichts in der 3. Klasse auf die Leistungen am Ende der 3. Klasse müssen zunächst für die Kinder mit DaZ-Unterricht (Experimentalgruppe) vergleichbare Kinder mit Deutsch als Zweitsprache ohne DaZ-Unterricht gesucht werden. Dafür werden mit Hilfe der logistischen Regression für alle Schülerinnen und Schüler mit Deutsch als Zweitsprache die *Propensity scores* für die Zuteilung zum DaZ-Unterricht in der 3. Klasse geschätzt, das heisst die Wahrscheinlichkeiten, dass ein Kind in der 3. Klasse dem DaZ-Unterricht zugeteilt wurde. Anhand der geschätzten *Propensity scores* findet dann das eigentliche *Matching* statt, das heisst die Suche nach statistischen Zwillingen.

Tabelle 4.5 zeigt die Ergebnisse der logistischen Regression zum Besuch des DaZ-Unterrichts in der 3. Klasse. In der Tabelle ist die geschätzte relative Chance beziehungsweise das *odds ratio* ausgewiesen, falls der Einfluss aller anderen Faktoren statistisch kontrolliert wird. Ein *odds ratio* grösser 1 bedeutet, dass ein hoher Wert auf der entsprechenden unabhängigen Variablen die Chance erhöht, den DaZ-Unterricht zu besuchen. Demgegenüber bedeutet ein Wert kleiner 1, dass ein hoher Wert auf der entsprechenden Variablen die Chance reduziert, den DaZ-Unterricht zu besuchen. Ein *odds ratio* von 1 bedeutet, dass die Variable keinen Einfluss auf die Chance hat, den DaZ-Unterricht zu besuchen.

Als unabhängige Variablen wurden das Geschlecht, die kognitive Grundfähigkeiten, die soziale Herkunft und die Testleistungen im Wortschatz, im Lesen und in der Mathematik beim Schuleintritt sowie als Merkmale der besuchten Klassen, die soziale Zusammensetzung einer Klasse, gemessen an der durchschnittlichen sozialen Herkunft und der Anteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache in die logistische Regression einbezogen (vgl. Kapitel Teil 2, Kapitel 1.3.1). Für die Leistungen in Deutsch beim Schuleintritt konnten lediglich zwei Fachbereiche des Deutschunterrichts berücksichtigt werden: (1) Wortschatz und (2) Lesen. Am Ende der 3. Klasse setzt sich der Deutschtest hingegen aus vier Teilbereichen zusammen: 1) Texte lesen und verstehen, (2) Sprachbetrachtung, (3) Texte überarbeiten und (4) Texte für sich und andere schreiben.

Für die Wirkungsanalysen zum DaZ-Unterricht wurden nur Schülerinnen und Schüler mit Deutsch als Zweitsprache einbezogen, die keine Klasse repetiert haben und in Regelklassen unterrichtet wurden. An dieser Stelle sei nochmals erwähnt, dass es bei der Schätzung der *Propensity scores* nicht primär um die einzelnen Effekte der einbezogenen Variablen geht, sondern darum, die Selektion der Schülerinnen und Schüler mit Deutsch als Zweitsprache zum DaZ-Unterricht so gut wie möglich vorherzusagen. Je höher der Erklärungsgehalt des Regressionsmodells ist, desto besser kann auch das *Matching* anhand der daraus geschätzten *Propensity scores* vorgenommen werden.

Tabelle 4.5 Ergebnisse der logistischen Regression zur Vorhersage des Besuchs des DaZ-Unterrichts in der 3. Klasse

Variable	odds ratio
Knaben	1.288
Kognitive Grundfähigkeit	0.999
Soziale Herkunft	0.757*
Leistung in Mathematik zu Beginn 1. Klasse	1.027
Leistung in Lesen zu Beginn 1. Klasse	0.893
Leistung in Wortschatz zu Beginn 1. Klasse	0.420***
Klassenmittelwert soziale Herkunft	2.251*
Klassenanteil Deutsch als Zweitsprache (%)	1.147
Konstante	48.294***
Pseudo-R ² (Mc-Faddens)	0.070
Prob > chi2	0.000
N	514

Anmerkungen: Binär logistische Regression. Abhängige Variable ist der Besuch des DaZ-Unterrichts in der 3. Klasse (Nein = 0, Ja = 1). Grundgesamtheit sind alle Schülerinnen und Schüler mit Deutsch als Zweitsprache der Längsschnittstichprobe, die nicht repetiert haben und in Regelklassen unterrichtet wurden.
Signifikanzniveaus: *p < .10; **p < .05; ***p < .01; ****p < .001 (zweiseitige Tests).

Die Ergebnisse zeigen, dass von den Schülermerkmalen einzig die soziale Herkunft und die Leistungen im Wortschatz beim Schuleintritt einen eigenständigen und statistisch signifikanten Einfluss auf den Besuch des DaZ-Unterrichts in der 3. Klasse ausüben. Je höher die soziale Herkunft oder je höher die Wortschatzleistungen, desto geringer ist auch die Chance, dass ein Kind dem DaZ-Unterricht zugeteilt wird. Von den Klassenmerkmalen ist unabhängig davon auch die soziale Zusammensetzung einer Klasse von Bedeutung. Je privilegiierter die soziale Zusammensetzung einer Klasse ist, desto höher ist auch die Chance, den DaZ-Unterricht zu besuchen. Das relativ geringe Pseudo-R² von rund 7 Prozent deutet darauf hin, dass mit Hilfe der einbezogenen Variablen lediglich ein kleiner Teil der selektionsrelevanten Kriterien für die Zuteilung zum DaZ-Unterricht berücksichtigt werden konnte. Der Besuch des DaZ-Unterrichts wird somit mehrheitlich von weiteren Merkmalen beeinflusst, die im Rahmen der Lernstandserhebung nicht erhoben wurden.

Matching und Überprüfung der Matching-Qualität

Mit Hilfe der geschätzten *Propensity scores* kann nun für die Gruppe der Schülerinnen und Schüler mit DaZ-Unterricht (Experimentalgruppe) eine geeignete Kontrollgruppe gebildet werden. Als *Matching*-Methode kommt dabei das *Nearest-neighbour*-Verfahren zur Anwendung. Jedem Kind mit Deutsch als Zweitsprache aus der Experimentalgruppe wird jenes Kind mit Deutsch als Zweitsprache aus der Kontrollgruppe als *Matching*-Partner zugeordnet, dessen *Propensity score* seinem am nächsten liegt.

Die Qualität des *Matching* lässt sich überprüfen, indem die Mittelwerte bei den zur Schätzung der *Propensity scores* verwendeten Variablen zwischen der Experimental- und der Kon-

trollgruppe miteinander verglichen werden. Während zwischen den beiden Gruppen vor dem *Matching* aufgrund der Selektion Unterschiede bestehen können, dürfen sich nach erfolgreichem *Matching* keine statistisch signifikanten Unterschiede mehr zeigen.

Tabelle 4.6 zeigt die Ergebnisse der Mittelwertsvergleiche für die verwendeten Variablen zur Schätzung der *Propensity scores* für die Experimental- und für die Kontrollgruppe, vor und nach dem *Matching*. Vor dem *Matching* unterscheiden sich die beiden Gruppen mit Ausnahme der Geschlechterverteilung und dem Klassenanteil an Schülerinnen und Schülern mit Deutsch als Zweitsprache statistisch signifikant voneinander. Die Kinder mit DaZ-Unterricht weisen bei den kognitiven Grundfähigkeiten, der sozialen Herkunft und den Testleistungen signifikant geringere Mittelwerte aus als die Kinder ohne DaZ-Unterricht. Im Wortschatz beispielsweise beträgt der Leistungsunterschied zwischen den beiden Gruppen rund 35 Punkte zugunsten der Schülerinnen und Schüler ohne DaZ-Unterricht. Nach dem *Matching* zeigen sich keine statistisch signifikanten Unterschiede mehr zwischen den beiden Gruppen. Für das erfolgreiche *Matching* sprechen auch das geringe Pseudo-R² von weniger als 2 Prozent nach dem *Matching* sowie der statistisch nicht mehr signifikante *Likelihood-ratio*-Test zur globalen Nullhypothese. Im Hinblick auf die berücksichtigten Merkmale sind die beiden Gruppen nun miteinander vergleichbar.

Tabelle 4.6 *Matching*-Qualität: Mittelwerte der in die logistische Regression einbezogenen Variablen vor und nach dem *Matching* (DaZ-Unterricht in 3. Klasse: Ja / Nein)

Variable	Sample	Experimental-Gruppe	Kontroll-Gruppe	p
Knaben (%)	Unmatched	0.50	0.46	0.343
	Matched	0.50	0.54	0.372
Kognitive Grundfähigkeiten	Unmatched	93.72	96.47	0.023
	Matched	93.72	91.72	0.100
Soziale Herkunft	Unmatched	-0.84	-0.61	0.002
	Matched	-0.84	-0.83	0.905
Leistung im Lesen: 1. Klasse	Unmatched	475.94	496.10	0.009
	Matched	475.94	484.09	0.307
Leistung im Wortschatz: 1. Klasse	Unmatched	400.21	435.86	0.000
	Matched	400.21	393.19	0.215
Leistung in der Mathematik: 1. Klasse	Unmatched	481.43	497.87	0.044
	Matched	481.43	485.50	0.624
Klassenanteil Deutsch als Zweitsprache (%)	Unmatched	0.46	0.47	0.525
	Matched	0.46	0.48	0.183
Klassenmittelwert soziale Herkunft	Unmatched	-0.26	-0.28	0.774
	Matched	-0.26	-0.34	0.060

Anmerkungen: Grundgesamtheit sind alle Schülerinnen und Schüler mit Deutsch als Zweitsprache der Längsschnittstichprobe, die nicht repetiert haben und in Regelklassen unterrichtet wurden.

Effekte des DaZ-Unterrichts auf die Leistungen in Deutsch

Nach erfolgreichem *Matching* kann nun der kausale Effekt des DaZ-Unterrichts in der 3. Klasse auf die Deutschleistungen am Ende der 3. Klasse überprüft werden. Zwar lässt sich nicht für jedes Kind der individuelle Effekt des DaZ-Unterrichts schätzen. Möglich ist aber die Schätzung eines durchschnittlichen Effekts des DaZ-Unterrichts auf die Experimentalgruppe.³ Dies geschieht mit Hilfe eines einfachen Mittelwertvergleichs zwischen der Experimental- und der neu gebildeten Kontrollgruppe.

Der Effekt des DaZ-Unterrichts in der 3. Klasse auf die Leistungen in Deutsch am Ende der 3. Klasse ist in Tabelle 4.7 ausgewiesen. Neben den erreichten Leistungsmittelwerten der Experimentalgruppe und der neu gebildeten Kontrollgruppe ist auch der kausale Effekt des DaZ-Unterrichts mit der entsprechenden Irrtumswahrscheinlichkeit (p) angegeben. Der Effekt des DaZ-Unterrichts berechnet sich aus der Differenz zwischen den Leistungsmittelwerten der Experimental- und der Kontrollgruppe.

Die Ergebnisse zeigen, dass sich kein Effekt des DaZ-Unterrichts in der 3. Klasse auf die Deutschleistungen am Ende der 3. Klasse nachweisen lässt (vgl. 4. Zeile in Tabelle 4.7). Zwar beträgt der Mittelwert der Experimentalgruppe mit rund 447 Punkten rund 11 Punkte mehr als in der Kontrollgruppe – was für einen positiven Effekt des DaZ-Unterrichts sprechen würde – der Effekt ist jedoch statistisch nicht signifikant.

Tabelle 4.7 Geschätzte Effekte des DaZ-Unterrichts auf die Leistungen in Deutsch

DaZ-Unterricht	Mittelwerte in den Deutschleistungen				Anzahl Schülerinnen und Schüler		
	Experimentalgruppe	Kontrollgruppe	DaZ-Effekt ¹	p	Experimentalgruppe	Kontrollgruppe	Total
Effekt auf Leistungen am Ende der 3. Klasse							
DaZ-Unterricht 3. Klasse	447	436	11	.695	250	264	514
DaZ-Unterricht 3. Klasse: 38 Lektionen	440	448	-8	.677	63	264	327
DaZ-Unterricht 3. Klasse: 76 Lektionen	447	436	25	.062	155	264	419
Effekt auf Leistungen am Ende der 6. Klasse							
DaZ-Unterricht 6. Klasse	706	691	15	.605	21	418	439

Anmerkungen: ¹Entspricht dem *Average Effect of Treatment on the Treated (ATT)*. Standardfehler durch Bootstrapping mit 200 Replikationen ermittelt (unter der Annahme einer Normalverteilung). Grundgesamtheit sind Schülerinnen und Schüler mit Deutsch als Zweitsprache der Längsschnittstichprobe, die nicht repetiert haben und in Regelklassen unterrichtet wurden. Signifikanzniveaus: zweiseitige Tests.

Im Folgenden wird die Analyse präzisiert, indem sich die Experimentalgruppe jeweils auf die Gruppe der Kinder mit DaZ-Unterricht in der 3. Klasse beschränkt, die 38 oder 76 Lektionen Unterricht erhalten haben. Dies entspricht den beiden grössten Gruppen. Von den Kindern mit Deutsch als Zweitsprache, die in der 3. Klasse den DaZ-Unterricht besucht haben, haben

³ *Average Effect of Treatment on the Treated (ATT)*.

rund 12 Prozent (N = 71) 38 und rund 61 Prozent (N = 171) 76 Lektionen DaZ-Unterricht erhalten.⁴ Die Ergebnisse zeigen, dass sich auch kein Effekt des DaZ-Unterrichts in der 3. Klasse mit 38 Lektionen auf die Deutschleistungen am Ende der 3. Klasse nachweisen lässt (vgl. Tabelle 4.7).⁵ Für die Experimentalgruppe wird der Mittelwert im Deutsch auf 440 Punkte geschätzt, in der Kontrollgruppe auf 448 Punkte. Der Unterschied von 8 Punkten ist statistisch nicht signifikant. Einzig der Besuch des DaZ-Unterrichts in der 3. Klasse während 76 Lektionen führt zu statistisch signifikanten Unterschieden in den Deutschleistungen am Ende der 3. Klasse (vgl. Tabelle 4.7). Die Schülerinnen und Schüler mit 76 Lektionen DaZ-Unterricht, was ungefähr 2 Wochenlektionen entspricht, erzielen mit rund 447 Punkten ein um rund 25 Punkte besseres Ergebnis als die vergleichbaren Kinder ohne DaZ-Unterricht (436 Punkte).

Weiter wurde überprüft, ob der Besuch des DaZ-Unterrichts in der 6. Klasse einen Effekt auf die Deutschleistungen am Ende der 6. Klasse ausübt. Dabei gilt es zu beachten, dass für diese Analysen lediglich noch 21 Kinder der Längsschnittstichprobe für die Experimentalgruppe berücksichtigt werden konnten.⁶ Die Ergebnisse zeigen, dass sich kein Effekt des DaZ-Unterrichts in der 6. Klasse auf die Deutschleistungen am Ende der Primarschule nachweisen lässt (vgl. Tabelle 4.7). Der Mittelwert der Experimentalgruppe beträgt 706 Punkte, jener der Kontrollgruppe 691 Punkte. Der Unterschied von 15 Punkten zugunsten der Kinder mit DaZ-Unterricht ist jedoch statistisch nicht signifikant. Auch weitere Analysen zum Besuch des DaZ-Unterrichts, der weiter als nur 1 Schuljahr vor dem Leistungstest am Ende der 6. Klasse zurückliegt, führen zu keinen neueren Erkenntnissen.

⁴ 8 beziehungsweise 16 Schülerinnen und Schüler konnten für die Analysen nicht berücksichtigt werden, weil sie entweder eine Klasse repetiert haben oder in Sonderklassen unterrichtet wurden.

⁵ Auf die Schätzungen der *Propensity scores* und die Überprüfung der *Matching*-Qualität wird an dieser Stelle im Interesse der Lesbarkeit nicht mehr eingegangen.

⁶ 12 Schülerinnen und Schüler konnten für die Analysen nicht berücksichtigt werden, weil sie entweder eine Klasse repetiert haben oder in Sonderklassen unterrichtet wurden.

5 Analysen zur Leistungsbeurteilung

Domenico Angelone

5.1 Einleitung

In diesem Kapitel werden die statistischen Schätzmodelle zur Analyse der Zeugnisnoten in Deutsch und Mathematik am Ende der 6. Klasse vorgestellt. Im Zentrum der Untersuchung stehen zwei Fragen: (1) Hängen die Zeugnisnoten, unabhängig von den schulischen Leistungen, auch mit sozialen Merkmalen wie dem Geschlecht, der Erstsprache oder der sozialen Herkunft der Schülerinnen und Schüler zusammen? (2) Wie gut sind Zeugnisnoten über verschiedene Klassen hinweg vergleichbar oder orientiert sich die Notengebung an einer internen Klassennorm wie dem durchschnittlichen Leistungsniveau einer Klasse?

5.2 Methode

Datenbasis

Grundlage für die statistischen Analysen der Zeugnisnoten bildet die Datenbasis der repräsentativen Stichprobe der 6. Klassen der Zürcher Lernstandserhebung. Zum Zeitpunkt der Erhebung standen die Zeugnisnoten bereits fest und konnten deshalb mit dem Schülerfragebogen erfasst werden.

Statistische Auswertungsmethode

Zur Analyse der Zeugnisnoten wurden Merkmale verwendet, die unterschiedliche Messebenen zugeordnet werden können. Auf der Schülerebene (Ebene 1) interessierte der Einfluss von individuellen Merkmalen wie Geschlecht, Alter, Erstsprache, soziale Herkunft und Testleistungen der Schülerinnen und Schüler. Auf der Klassenebene (Ebene 2) wurde überprüft, ob unabhängig davon auch das durchschnittliche Leistungsniveau einer Klasse einen Einfluss auf die Noten ausübt. Die Mehrebenenanalyse erlaubte es, den Einfluss von Merkmalen unterschiedlicher Ebenen gleichzeitig zu schätzen (Raudenbush & Bryk, 2002). Weil die abhängige Variable – die Zeugnisnoten am Ende der 6. Klasse – metrisch ist, wurden lineare Zweiebenenmodelle geschätzt. Dabei wurden die Schätzkoeffizienten der erklärenden Variablen alle fix gehalten, lediglich die Variation der Konstanten zwischen den Klassen wurde zugelassen (random-intercept-model).

5.3 Ergebnisse zur Erklärung der Zeugnisnoten in Deutsch

Tabelle 5.1 zeigt die Ergebnisse der Mehrebenenanalysen zur Erklärung der Zeugnisnoten am Ende der 6. Klasse in Deutsch. In der ersten Spalte sind die Variablen aufgeführt, die in die Analysen einbezogen wurden. In den mit «b-Koeff.» bezeichneten Spalten sind die unstandardisierten Regressionskoeffizienten ausgewiesen. Diese zeigen, wie die Effekte der entsprechenden Variablen auf die Zeugnisnoten zu interpretieren sind, wenn alle anderen Variablen konstant gehalten werden. Die Variable «Knaben» entspricht dem Geschlechtereffekt. Zudem ist aufgeführt, wie gross die Abweichung der Noten der Knaben von den Noten der Mädchen ist. Die Variable «Alter in Monaten» gibt die Notenveränderung an, wenn das Alter um einen Monat ansteigt. Die Variable «Deutsch als Zweitsprache» zeigt, wie gross der No-

tenunterschied zwischen den Schülerinnen und Schülern mit Deutsch als Zweitsprache und jenen mit Deutsch als Erstsprache ist. Die Variable «Soziale Herkunft» gibt die Notenveränderung an, wenn die soziale Herkunft der Schülerinnen und Schüler um eine Standardabweichung ansteigt. Die Variable «Leistung in Deutsch» zeigt die Notenveränderung, wenn die Testleistungen in Deutsch um 100 Punkte ansteigen.

Tabelle 5.1: Ergebnisse der Mehrebenenanalysen zur Erklärung der Zeugnisnoten in Deutsch

	Modell I	Modell II	Modell III	Modell IV
Fixe Effekte	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.
Konstante	4.60***	4.52***	4.59***	4.64***
Schülermerkmale				
Knaben			-0.04	-0.04
Alter in Monaten			-0.01***	-0.01***
Deutsch als Zweitsprache			-0.13***	-0.15***
Soziale Herkunft (Index)			0.09***	0.10***
Leistung in Deutsch (in 100 Punkten)		0.46***	0.39***	0.40***
Klassenmerkmale				
Durchschnittliche Leistung in Deutsch (in 100 Punkten)				-0.27***
Variable Effekte				
Klassenebene	0.03***	0.03***	0.03***	0.02***
Schülerebene	0.39***	0.19***	0.18***	0.18***
Erklärte Varianz				
R ² (Schülerebene)		50.3%	54.6%	54.7%
R ² (Klassenebene)				37.4%
R ² (Gesamt)		45.6%	50.5%	53.6%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler sind in Klassen gruppiert), Random-Intercept. Abhängige Variable ist die Zeugnisnote in Deutsch am Ende der 6. Klasse. Alle metrischen Variablen wurden am «Grand Mean» zentriert. Für die Testleistungen der Schülerinnen und Schüler wurden sogenannte Warm Estimates verwendet. Gewichtete Daten. N (listwise delete) = 1676.
Signifikanzniveaus: *p < .10; **p < .05; ***p < .01; ****p < .001.

In Modell I sind die Ergebnisse des Nullmodells ausgewiesen. Beim Nullmodell werden keine erklärenden Variablen ins Modell einbezogen. Es enthält lediglich eine Regressionskonstante, die zwischen den Kontexteinheiten (Klassen) variiert. Die Konstante entspricht deshalb dem Mittelwert der durchschnittlichen Klassennoten in Deutsch (rund 4.60). Das Nullmodell zerlegt die Gesamtvarianz in den Zeugnisnoten in zwei Bestandteile, die Varianz zwischen Schülerinnen und Schülern und die Varianz zwischen Klassen. Daraus lässt sich der Varianzanteil in den Zeugnisnoten bestimmen, der auf Unterschiede zwischen Klassen zurückzuführen ist. Die Varianz der klassenspezifischen Regressionskonstante, die der Varianz der Klassenmittelwerte entspricht, wird auf rund 0.03 Notenpunkte geschätzt. Die Varianz innerhalb der Klassen beträgt rund 0.39 Notenpunkte. Entsprechend können rund 6 Prozent der Gesamtvarianz in den Deutschnoten auf Unterschiede zwischen Klassen und rund 94 Prozent auf

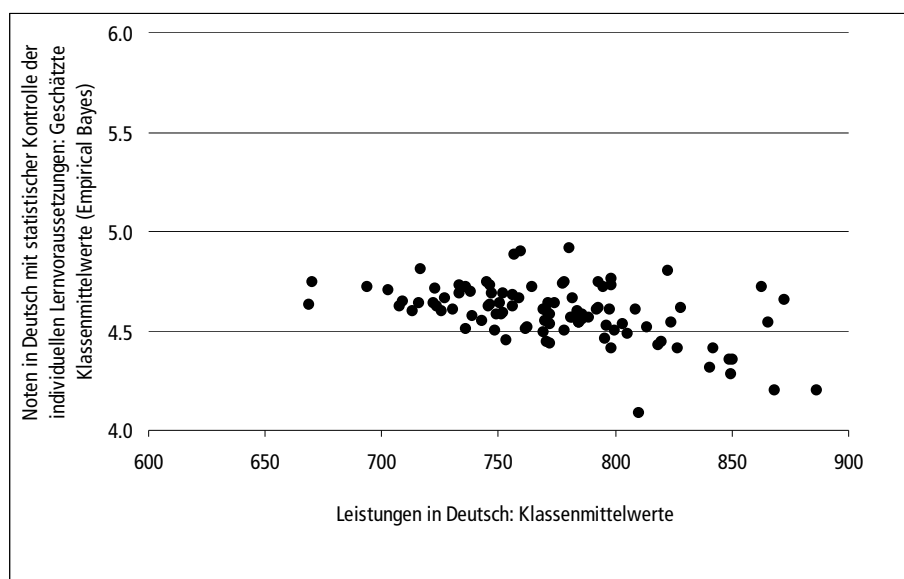
Unterschiede zwischen Schülerinnen und Schülern innerhalb der Klassen zurückgeführt werden.

In Modell II wurden die Testleistungen in Deutsch einbezogen. Je höher die Testleistungen sind, desto höher sind die Zeugnisnoten. Eine Verbesserung der Testleistungen in Deutsch um 100 Punkte, was in etwa einer Standardabweichung entspricht, erhöht die Deutschnoten um rund 0.46 Notenpunkte. Die Varianz der Noten zwischen den Schülerinnen und Schülern innerhalb der Klassen reduziert sich auf 0.19 Notenpunkte. Der Einbezug der Testleistungen erklärt somit rund 50 Prozent der Varianz zwischen den Schülerinnen und Schülern innerhalb der Klassen. Die Varianz zwischen den Klassen wird auf 0.03 Notenpunkte geschätzt.

Modell III zeigt die Ergebnisse, nachdem zusätzlich zu den Testleistungen auch die Lernvoraussetzungen der Schülerinnen und Schüler einbezogen wurden. Bei gleichen Leistungen sind das Alter, die Erstsprache und die soziale Herkunft der Schülerinnen und Schüler von Bedeutung. Zwischen Mädchen und Knaben hingegen zeigen sich keine Unterschiede in den Zeugnisnoten. Die Zeugnisnoten der Schülerinnen und Schüler mit Deutsch als Zweitsprache fallen rund 0.13 Notenpunkte tiefer aus als jene der Schülerinnen und Schüler mit Deutsch als Erstsprache. Je höher die soziale Herkunft ist, desto höher sind auch die Zeugnisnoten. Bei einem Anstieg der sozialen Herkunft um eine Standardabweichung steigen die Deutschnoten um rund 0.10 Notenpunkte an. Schliesslich stehen die Zeugnisnoten in einem negativen Zusammenhang mit dem Alter der Schülerinnen und Schüler. Ein Anstieg des Alters um ein Jahr führt zu einer Verschlechterung der Noten um rund 0.13 Notenpunkte ($-0.13 = 12 \cdot -0.011$). Die Varianz der Schülerinnen und Schüler innerhalb der Klassen reduziert sich auf rund 0.18 Notenpunkte. Die zusätzliche Berücksichtigung der Lernvoraussetzungen erhöht somit die erklärte Varianz zwischen den Schülerinnen und Schülern innerhalb der Klassen auf rund 55 Prozent. Die Varianz zwischen den Klassen wird auf 0.03 Notenpunkte geschätzt und ist statistisch signifikant. Auch bei gleichen Lernvoraussetzungen der Schülerinnen und Schüler zeigen sich noch Unterschiede in den Zeugnisnoten zwischen den Klassen.

In einem nächsten Schritt wurde überprüft, ob das Leistungsniveau einer Klasse, gemessen an der durchschnittlichen Deutschleistung einer Klasse, einen Beitrag zur Erklärung der verbleibenden Unterschiede zwischen den Klassen leisten kann. Abbildung 5.1 zeigt, wie die durchschnittlichen Deutschnoten nach statistischer Kontrolle der Lernvoraussetzungen der Schülerinnen und Schüler mit dem Leistungsniveau einer Klasse zusammenhängen. Auf der y-Achse sind die anhand des Modells III vorausgesagten durchschnittlichen Deutschnoten einer Klasse abgebildet. Die x-Achse repräsentiert das Leistungsniveau der Klassen, gemessen an den durchschnittlichen Deutschleistungen einer Klasse. Nach Kontrolle der Lernvoraussetzungen zeigt sich ein negativer Zusammenhang zwischen den durchschnittlichen Zeugnisnoten und dem Leistungsniveau einer Klasse. Je höher das Leistungsniveau einer Klasse ist, desto tiefer sind auch die durchschnittlichen Deutschnoten.

Abbildung 5.1: Noten in Deutsch, bereinigt um individuelle Lernvoraussetzungen, nach der durchschnittlichen Leistung von Klassen (lineare Mehrebenenanalyse)



Dieses Ergebnis bestätigt sich in Modell IV, das als weitere Variable das Leistungsniveau einer Klasse in die Analyse einbezieht. Bei einem Anstieg der durchschnittlichen Klassenleistung um 100 Punkte sinkt die Deutschnote um rund 0.27 Notenpunkte. Wenn beispielsweise eine Schülerin mittleren Alters mit Deutsch als Erstsprache und durchschnittlicher sozialer Herkunft eine Klasse mit dem 25 Prozent höchstem Leistungsniveau (798 Punkte) besucht, sind ihre Deutschnoten um 0.14 Notenpunkte tiefer, als wenn dieselbe Schülerin eine Klasse mit dem 25 Prozent tiefstem Leistungsniveau (746 Punkte) besucht. Insgesamt vermag die Berücksichtigung des Leistungsniveaus einer Klasse rund 37 Prozent der Varianz zwischen den Klassen zu erklären.

5.4 Ergebnisse zur Erklärung der Zeugnisnoten in Mathematik

Tabelle 5.2 zeigt die Ergebnisse der Mehrebenenanalysen zur Erklärung der Zeugnisnoten am Ende der 6. Klasse in Mathematik. In der ersten Spalte sind wiederum die Variablen aufgeführt, die in die Analysen einbezogen wurden. Die nachfolgenden Spalten zeigen die unstandardisierten Regressionskoeffizienten.

In Modell I sind die Ergebnisse des Nullmodells ausgewiesen. Die Varianz zwischen den Klassen wird auf rund 0.03 Notenpunkte geschätzt, jene zwischen den Schülerinnen und Schülern innerhalb der Klassen auf rund 0.52 Notenpunkte. Somit lassen sich rund 5 Prozent der Gesamtvarianz in den Mathematiknoten durch Klassenmerkmale und rund 95 Prozent durch Schülermerkmale erklären.

Tabelle 5.2: Ergebnisse der Mehrebenenanalysen zur Erklärung der Zeugnisnoten in Mathematik

	Modell I	Modell II	Modell III	Modell IV
	b-Koeff.	b-Koeff.	b-Koeff.	b-Koeff.
Fixe Effekte				
Konstante	4.57***	4.53***	4.56***	4.58***
Schülermerkmale				
Knaben			0.01	0.01
Alter in Monaten			-0.01**	-0.01*
Deutsch als Zweitsprache			-0.08*	-0.10*
Soziale Herkunft (Index)			0.12***	0.12***
Leistung in Mathematik (in 100 Punkten)		0.52***	0.45***	0.46***
Klassenmerkmale				
Durchschnittliche Leistung in Mathematik (in 100 Punkten)				-0.21***
Variable Effekte				
Klassenebene	0.03***	0.03***	0.02**	0.02***
Schülerebene	0.52***	0.25***	0.24***	0.24***
Erklärte Varianz				
R ² (Schülerebene)		50.6%	54.1%	54.1%
R ² (Klassenebene)				46.6%
R ² (Gesamt)		47.6%	52.5%	53.7%

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Schülerinnen und Schüler in Klassen, Random-Intercept). Abhängige Variable sind die Zeugnisnoten in Mathematik. Alle metrischen Variablen wurden am «Grand Mean» zentriert. Für die Testleistungen der Schülerinnen und Schüler wurden sogenannte Warm Estimates verwendet. Gewichtete Daten. N (listwise delete) = 1682
Signifikanzniveaus: *p < .10; **p < .05; ***p < .01; ****p < .001.

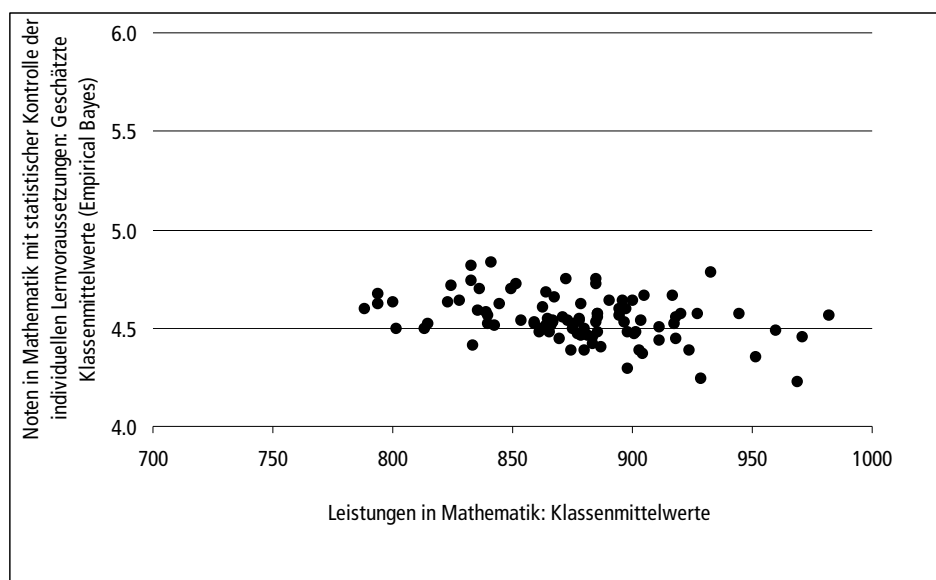
In Modell II wurden lediglich die Testleistungen der Schülerinnen und Schüler in Mathematik in die Analyse einbezogen. Erwartungsgemäss hängen die Testleistungen positiv mit den Zeugnisnoten in Mathematik zusammen. Je höher die Testleistungen sind, desto höher sind die Zeugnisnoten. Der Zusammenhang ist nahezu gleich stark ausgeprägt wie zwischen den Testleistungen und den Zeugnisnoten in Deutsch. Bei einem Anstieg der Testleistungen um 100 Punkte, was etwa einer Standardabweichung entspricht, verbessern sich die Mathematiknoten um rund 0.52 Notenpunkte. Die Varianz der Noten zwischen den Schülerinnen und Schülern innerhalb der Klassen ist auf 0.25 Notenpunkte gesunken. Die Testleistungen vermögen somit rund 51 Prozent der Varianz zwischen den Schülerinnen und Schülern innerhalb der Klassen zu erklären. Die Varianz zwischen den Klassen wird auf 0.03 Notenpunkte geschätzt.

In Modell III wurden auch die Lernvoraussetzungen der Schülerinnen und Schüler in die Analyse einbezogen. Dabei zeigt sich ein ähnliches Bild wie bei der Erklärung der Deutschnoten. Bei gleichen Testleistungen hängen die Zeugnisnoten in Mathematik mit dem Alter, der Erstsprache und der sozialen Herkunft der Schülerinnen und Schüler zusammen. Das Geschlecht hingegen spielt keine Rolle. Die Zeugnisnoten der Schülerinnen und Schüler mit Deutsch als

Zweitsprache sind rund 0.08 Notenpunkte tiefer als jene der Schülerinnen und Schüler mit Deutsch als Erstsprache. Je höher die soziale Herkunft der Schülerinnen und Schüler ist, desto höher sind auch die Noten. Bei einem Anstieg der sozialen Herkunft um eine Standardabweichung steigen die Mathematiknoten um rund 0.12 Notenpunkte. Schliesslich hängen die Noten negativ mit dem Alter der Schülerinnen und Schüler zusammen. Ein Anstieg des Alters um 1 Jahr führt zu einer Verschlechterung der Noten um rund 0.10 Notenpunkte ($-0.10 = 12 \cdot -0.008$). Nach Berücksichtigung der Lernvoraussetzungen hat sich die geschätzte Varianz zwischen den Schülerinnen und Schülern innerhalb der Klassen auf 0.24 Notenpunkte reduziert. Dadurch erhöht sich die erklärte Varianz zwischen den Schülerinnen und Schülern auf rund 54 Prozent. Die Varianz zwischen den Klassen wird auf 0.02 Notenpunkte geschätzt. Auch in Mathematik zeigen sich bei gleichen Lernvoraussetzungen der Schülerinnen und Schüler weiterhin noch Unterschiede in den Zeugnisnoten zwischen den Klassen.

Wiederum wurde überprüft, ob das Leistungsniveau einer Klasse einen Beitrag zur Erklärung der verbleibenden Klassenunterschiede zu leisten vermag. Abbildung 5.2 zeigt, wie die durchschnittlichen Mathematiknoten, nach statistischer Kontrolle der Lernvoraussetzungen der Schülerinnen und Schüler, mit dem Leistungsniveau einer Klasse zusammenhängen. Auf der y-Achse sind die anhand des Modells III geschätzten durchschnittlichen Mathematiknoten einer Klasse abgebildet. Die x-Achse repräsentiert das Leistungsniveau der Klassen, gemessen an den durchschnittlichen Mathematikleistungen einer Klasse. Dabei zeigt sich dasselbe Bild wie bei den Deutschnoten. Je höher das Leistungsniveau einer Klasse, desto tiefer sind auch die durchschnittlichen Mathematiknoten.

Abbildung 5.2: Noten in Mathematik, bereinigt um individuelle Lernvoraussetzungen nach der durchschnittlichen Leistung von Klassen (lineare Mehrebenenanalyse)



Modell IV überprüft auch in Mathematik den Einfluss des durchschnittlichen Leistungsniveaus einer Klasse auf die Zeugnisnoten. Die Ergebnisse zeigen, dass bei gleichen schulischen Leistungen und Lernvoraussetzungen der Schülerinnen und Schüler die Zeugnisnoten mit stei-

gendem Leistungsniveau einer Klasse sinken. Bei einem Anstieg der durchschnittlichen Mathematikleistungen einer Klasse um 100 Punkte sinken die Mathematiknoten um rund 0.21 Notenpunkte. Wenn beispielsweise eine Schülerin mittleren Alters mit Deutsch als Erstsprache und durchschnittlicher sozialer Herkunft eine Klasse mit dem 25 Prozent höchstem Leistungsniveau (900 Punkte) besucht, sind ihre Deutschnoten um 0.09 Notenpunkte tiefer, als wenn dieselbe Schülerin eine Klasse mit dem 25 Prozent tiefstem Leistungsniveau (858 Punkte) besucht. In Mathematik erklärt das Leistungsniveau einer Klasse rund 47 Prozent der Varianz in den Zeugnisnoten zwischen den Klassen, rund 10 Prozent mehr als in Deutsch.

6 Analysen zum Übertritt von der Primarschule in die Sekundarstufe I

Domenico Angelone

6.1 Einleitung

In diesem Kapitel werden die statistischen Analysen zur Bedeutung von sozialen Merkmalen wie dem Geschlecht oder der sozialen Herkunft der Schülerinnen und Schüler beim Übertritt in die Sekundarstufe I vorgestellt. Die Analysen beziehen sich (1) auf den Übertritt in das Langgymnasium und (2) auf den Übertritt in die unterschiedlich anspruchsvollen Abteilungen der Sekundarschule.

6.2 Methode

Datenbasis

Die Datenbasis für die vorliegenden Analysen bildet die repräsentative Stichprobe der 6. Klassen der Zürcher Lernstandserhebung. Die Daten enthalten neben den fachlichen und überfachlichen Kompetenzen der Schülerinnen und Schüler am Ende der 6. Klasse der Primarschule auch Angaben zu den Übertrittsentscheiden in die verschiedenen Schultypen der Sekundarstufe I. Die Angaben zu den Übertrittsentscheiden wurden mit dem Schülerfragebogen erfasst.

Statistische Auswertungsmethode

Die Analyse der Bedeutung von sekundären Herkunftseffekten für den Übertritt in die Sekundarstufe I erfolgt mit Hilfe der logistischen Mehrebenenanalyse (Raudenbush & Bryk, 2002). Die logistische Mehrebenenanalyse erlaubt es, die eigenständigen Effekte von mehreren Einflussfaktoren auf eine binäre abhängige Variable – beispielsweise Übertritt in das Langgymnasium gegenüber Übertritt in die Sekundarschule – zu untersuchen und gleichzeitig die hierarchische Datenstruktur zu berücksichtigen (Schülerinnen und Schüler sind in Schulklassen gruppiert). Die logistische Regressionsanalyse bezieht sich jedoch nicht direkt auf die Wahrscheinlichkeit für die Zugehörigkeit zu einer bestimmten Gruppe, in diesem Beispiel für den Übertritt in das Langgymnasium, sondern auf die sogenannte Chance (odds) für den Übertritt in das Langgymnasium. Die Chance entspricht dabei dem Verhältnis der Wahrscheinlichkeit in das Langgymnasium überzutreten zur Gegenwahrscheinlichkeit. Dies sei an einem Beispiel erläutert. Im Kanton Zürich sind im Jahr 2006 rund 22 Prozent der Schülerinnen und Schüler mit Deutsch als Erstsprache in ein Langgymnasium übergetreten (eigene Berechnungen aufgrund der Stichprobe). Ihre Chance für den Besuch des Langgymnasiums beträgt somit 22:78 beziehungsweise 0.282. Demgegenüber sind lediglich rund 8 Prozent der Schülerinnen und Schüler mit Deutsch als Zweitsprache ins Langgymnasium übergetreten. Ihre Chance den Besuch des Langgymnasiums beträgt somit 8:92 beziehungsweise 0.087. Sie ist um den Faktor 0.309 ($0.087 : 0.282$) kleiner als jene der Schülerinnen und Schüler mit Deutsch als Erstsprache. Dieser Faktor wird als Chancenverhältnis beziehungsweise relative Chance (odds ratio) bezeichnet und beschreibt die Bedeutung des Zusammenhangs zwischen den beiden untersuchten Variablen. Eine relative Chance grösser als 1 geht mit einer höheren Chance für den Übertritt in das Langgymnasium der betreffenden Gruppe einher; eine relative Chance

kleiner als 1 mit einer geringeren Chance. Eine relative Chance von 1 bedeutet, dass das betreffende Merkmal oder die betreffende Gruppe in keinem Zusammenhang mit der Chance für den Übertritt steht. Die Interpretation der Ergebnisse der logistischen Regression anhand der relativen Chancen für das Auftreten eines bestimmten Ereignisses (z.B. Übertritt in das Langgymnasium) ist wenig anschaulich. Mit Hilfe der geschätzten Modellparameter lassen sich jedoch auch die (vertrauten) Wahrscheinlichkeiten für das Auftreten eines bestimmten Ereignisses berechnen.

6.3 Ergebnisse zum Übertritt in das Langgymnasium

Tabelle 6.1 zeigt die Ergebnisse der Mehrebenenanalysen zum Einfluss von sozialen Merkmalen der Schülerinnen und Schüler auf die Chance für den Übertritt in das Langgymnasium.

Tabelle 6.1: Ergebnisse der logistischen Mehrebenenanalysen zum Einfluss von sozialen Merkmalen auf den Übertritt in das Langgymnasium

	Modell I	Modell II	Modell III
	Odds ratio	Odds ratio	Odds ratio
Fixe Effekte für die Konstante	0.003***	0.008***	0.001***
Knaben	1.105	0.856	1.026
Alter (in Monaten)	0.951*	0.940**	0.947*
Deutsch als Zweitsprache	0.955	0.715	1.050
Soziale Herkunft	2.022***	2.020***	1.854***
Zeugnisnote in Deutsch	19.035***		11.131***
Zeugnisnote in Mathematik	18.485***		5.835***
Deutschleistung (in 100 Punkten)		4.755***	2.911***
Mathematikleistung (in 100 Punkten)		4.897***	2.834***
Unterricht in Deutsch und/oder Mathematik zur Prüfungsvorbereitung für das Langgymnasium	1.678*	3.058***	1.794***
Variable Effekte (Varianzen)			
Ebene 2	0.52*	0.34 ⁺	0.01
Modellparameter			
N (listwise delete)	1813	1830	1784

Anmerkungen: Binär logistische Mehrebenenanalyse (Schülerinnen und Schüler sind in Klassen gruppiert), Random-Intercept-Model, Schätzverfahren: EM Laplace Iterations (100). Abhängige Variable ist der Übertritt in das Langgymnasium (0 = Übertritt in Sekundarschule Typ A, B, C, 1 = Übertritt in das Langgymnasium). Alle metrischen Variablen wurden am «Grand Mean» zentriert. Signifikanzniveaus: ⁺p < .10; *p < .05; **p < .01; ***p < .001 (zweiseitige Tests). Gewichtete Daten.

In der ersten Spalte sind die Variablen aufgeführt, die in die Analysen einbezogen wurden. Als soziale Merkmale wurden das Geschlecht, das Alter, die Erstsprache und die soziale Herkunft der Schülerinnen und Schüler berücksichtigt, als schulische Vorleistungen die Deutsch- und Mathematiknoten am Ende der 6. Klasse und/oder die Testleistungen in Deutsch- und in Mathematik (WLE). Schliesslich wurde auch berücksichtigt, ob die Schülerinnen und Schüler einen Unterricht in Deutsch- und/oder Mathematik zur Vorbereitung für die Prüfung für das Langgymnasium besucht haben. In den folgenden Spalten sind die zugehörigen relativen Chancen beziehungsweise odds ratios ausgewiesen, falls der Einfluss aller anderen Faktoren konstant gehalten wird.

Modell I bezieht die sozialen Merkmale der Schülerinnen und Schüler sowie ihre Deutsch- und Mathematiknoten als Indikatoren für die schulischen Leistungen in die Analyse ein. Dadurch lässt sich die Frage beantworten, inwiefern schulische Leistungen und individuelle Lernvoraussetzungen für den Übertritt in das Langgymnasium von Bedeutung sind. Die Ergebnisse in Modell I zeigen, dass neben dem Alter auch die soziale Herkunft der Schülerinnen und Schüler von Bedeutung sind. Bei einem Anstieg der sozialen Herkunft um eine Standardabweichung steigt die relative Chance für den Übertritt in das Langgymnasium um das Zweifache an. Je höher das Alter der Schülerinnen und Schüler ist, desto geringer ist auch die Chance für den Übertritt. Der Effekt des Alters ist jedoch relativ schwach ausgeprägt. Bei einem Anstieg des Alters um einen Monat verringert sich die Chance um den Faktor 0.951. Zwischen Mädchen und Knaben sowie zwischen Schülerinnen und Schülern mit Deutsch als Erstsprache und solchen mit Deutsch als Zweitsprache zeigen sich keine statistisch signifikanten Unterschiede in der Chance für den Übertritt in das Langgymnasium. Auch der Besuch eines Deutsch- und/oder Mathematikunterrichts zur Prüfungsvorbereitung erhöht die Chance eines Übertritts. Der Besuch eines Unterrichts zur Prüfungsvorbereitung verbessert die Chance für den Übertritt um den Faktor 1.7. Erwartungsgemäss erhöhen auch die Deutsch- oder Mathematiknoten die Chance für den Übertritt in das Langgymnasium. Bei einem Anstieg der Deutsch- oder Mathematiknoten um 1 Notenpunkt verbessern sich die Chancen um rund den Faktor 19. Der zusätzliche Effekt der Noten auf die Übertrittschancen ist vergleichsweise gross. Dies erstaunt nicht, wird doch der Übertritt in erster Linie durch die schulischen Vorleistungen bestimmt. Bereits bei einem Notendurchschnitt von 5 ist die Wahrscheinlichkeit relativ gering, in das Langgymnasium überzutreten.

Modell II bezieht als Indikator für die schulischen Leistungen nicht die Schulnoten, sondern die Testleistungen in Deutsch und in Mathematik ein. Diese ermöglichen gegenüber den Schulnoten eine klassen- und lehrpersonunabhängige Beurteilung der Fähigkeiten. Allerdings erfassen die Testleistungen nicht die gesamten schulischen Leistungen, sondern decken lediglich einen Ausschnitt davon ab. Beispielweise wurden im Deutschttest keine produktiven Fähigkeiten getestet. In Modell II bestätigen sich die bisher beobachteten Ergebnisse mit geringen Abweichungen in den Zahlen. Auch unter Berücksichtigung der Testleistungen steigt mit einer höheren sozialen Herkunft die Chance für den Übertritt in das Langgymnasium markant an. Der Effekt der sozialen Herkunft fällt mit einer relativen Chance von 2 gleich stark aus wie unter Berücksichtigung der Schulnoten.

Modell III schliesslich bezieht sowohl die Schulnoten als auch die Testleistungen in die Analyse ein. Erwartungsgemäss reduzieren sich die Effekte der Schulnoten und der Testleistungen, weil ein Teil der Erklärungskraft auf beide Merkmale gemeinsam zurückführbar ist. Die Verringerung der Effekte ist vorwiegend bei den Mathematiknoten stark ausgeprägt. Dies deutet

darauf hin, dass die getesteten Deutschleistungen im Vergleich zu den Mathematikleistungen weniger gut die Fähigkeiten abdecken, die anhand der Schulnoten beurteilt wurden. Im Deutshtest wurden beispielsweise – im Gegensatz zu den Schulnoten – keine produktiven Fähigkeiten bewertet. Die Mathematikleistungen hingegen wurden sehr viel breiter getestet. Der Einfluss der sozialen Herkunft der Schülerinnen und Schüler auf die Übertrittschancen verringert sich nur unwesentlich.

Interpretation der Effekte in Wahrscheinlichkeiten

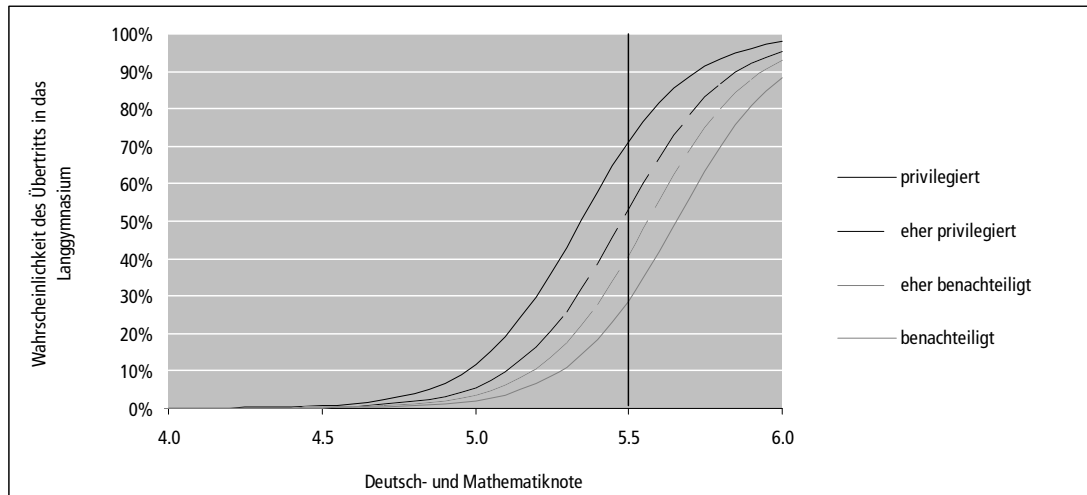
Die Interpretation der relativen Chancen ist nicht sehr anschaulich. Anhand der Koeffizienten des Analysemodells lassen sich auch die Übertrittswahrscheinlichkeiten für bestimmte Ausprägungen in den unabhängigen Variablen berechnen. In Abbildung 6.1 sind die Wahrscheinlichkeiten (Modell I) für den Übertritt in das Langgymnasium nach den Zeugnisnoten in Deutsch und in Mathematik für Schülerinnen und Schüler unterschiedlicher sozialer Herkunft grafisch dargestellt. Auf der x-Achse sind die Deutsch- und Mathematiknoten angegeben. Die y-Achse repräsentiert die Wahrscheinlichkeiten für den Übertritt in das Langgymnasium.

Die vier Kurven zeigen beispielhaft für Schülerinnen mittleren Alters mit Deutsch als Erstsprache, ohne Besuch eines Unterrichts zur Prüfungsvorbereitung und unterschiedlicher sozialer Herkunft den Verlauf der Wahrscheinlichkeiten in Abhängigkeit der erzielten Zeugnisnoten in Deutsch und in Mathematik. Im Hinblick auf die soziale Herkunft wird der Vergleich zwischen Schülerinnen mit benachteiligter, mit eher benachteiligter, mit eher privilegierter und mit privilegierter sozialer Herkunft dargestellt.¹ Daraus geht hervor, dass bei gleichen Schulnoten mit steigender sozialer Herkunft auch die Wahrscheinlichkeit für den Übertritt in das Langgymnasium ansteigt. Beispielsweise beträgt bei einem Mädchen mittleren Alters, mit Deutsch als Erstsprache, mit sehr guten Deutsch- und Mathematiknoten (5.5), ohne Besuch eines Unterrichts zur Prüfungsvorbereitung und mit privilegierter sozialer Herkunft die Wahrscheinlichkeit für den Übertritt in das Langgymnasium rund 71 Prozent; mit eher privilegierter sozialer Herkunft 53 Prozent; mit eher benachteiligter sozialer Herkunft 41 Prozent; mit benachteiligter sozialer Herkunft noch 29 Prozent.

Wie Abbildung 6.1 zeigt, ist die Bedeutung der sozialen Herkunft für die Übertrittschancen vorwiegend bei einer Deutsch- und Mathematiknote von 5.5 stark ausgeprägt. Bei Schülerinnen und Schülern mit einer 5 oder einer 6 zeigen sich nahezu keine Unterschiede zwischen Schülerinnen und Schülern unterschiedlicher sozialer Herkunft. Der Grund dafür ist, dass bei einer Note 5 die Wahrscheinlichkeiten für den Übertritt bereits sehr gering (unter 10 Prozent) beziehungsweise bei einer Note 6 bereits sehr hoch (über 90 Prozent) sind. Die zusätzliche Erklärungskraft der sozialen Herkunft ist in beiden Fällen nur sehr gering.

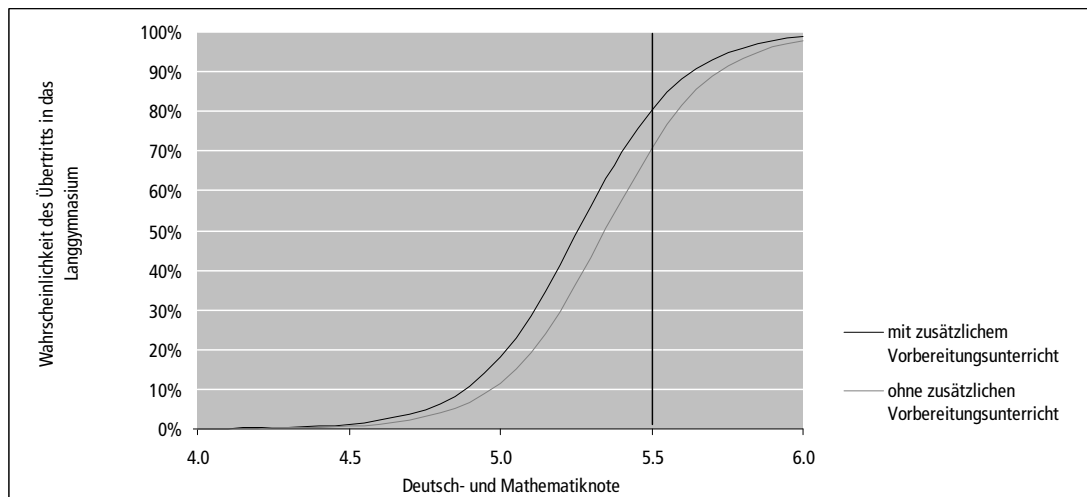
¹ Die Einteilung der Schülerinnen und Schüler nach sozialer Herkunft in vier Gruppen entspricht der Einteilung nach Quartilsgruppen. Zur Vorhersage der Übertrittswahrscheinlichkeiten für Schülerinnen und Schüler unterschiedlicher Gruppen sozialer Herkunft wurde für den Index zur sozialen Herkunft jeweils der Mittelwert der entsprechenden Gruppe verwendet.

Abbildung 6.1: Wahrscheinlichkeiten des Übertritts in das Langgymnasium nach Deutsch- und Mathematiknoten für Schülerinnen unterschiedlicher sozialer Herkunft (logistische Mehrebenenanalyse)



Anmerkungen: Die Wahrscheinlichkeiten beziehen sich jeweils auf ein Mädchen mittleren Alters, mit Deutsch als Erstsprache und ohne Besuch eines Unterrichts zur Prüfungsvorbereitung.

Abbildung 6.2: Wahrscheinlichkeiten des Übertritts in das Langgymnasium nach Deutsch- und Mathematiknoten für Schülerinnen mit und ohne Besuch eines Unterrichts zur Prüfungsvorbereitung (logistische Mehrebenenanalyse)



Anmerkungen: Die Wahrscheinlichkeiten beziehen sich jeweils auf ein Mädchen mittleren Alters mit Deutsch als Erstsprache und mit privilegierter sozialer Herkunft.

Abbildung 6.2 zeigt die Wahrscheinlichkeiten für den Übertritt in Abhängigkeit der Zeugnisnoten in Deutsch und in Mathematik für Schülerinnen und Schüler, die einen Unterricht zur Prüfungsvorbereitung besucht haben, im Vergleich zu solchen, die keinen Unterricht zur Prüfungsvorbereitung besucht haben. Beispielsweise beträgt die Wahrscheinlichkeit für den Übertritt bei einem Mädchen mittleren Alters, mit Deutsch als Erstsprache, mit sehr guten

Deutsch- und Mathematiknoten (5.5) und mit privilegierter sozialer Herkunft, das keinen Unterricht zur Prüfungsvorbereitung besucht hat rund 71 Prozent; bei Besuch eines Unterrichts zur Prüfungsvorbereitung rund 80 Prozent.

6.4 Ergebnisse zum Übertritt in die Abteilung A der Sekundarschule

In diesem Kapitel wird die vorangehende Analyse auf die Übertrittschancen in die verschiedenen Schultypen der Sekundarschule repliziert. Dabei wird der Übertritt in die Abteilung A der Sekundarschule dem Übertritt in die weniger anspruchsvolleren Abteilung B und C gegenübergestellt. Die Schülerinnen und Schüler, die in das Langgymnasium übergetreten sind, werden aus der Analyse ausgeschlossen. Im Jahr 2009 sind am Ende der 6. Klasse 47.5 Prozent der Schülerinnen und Schüler in die Abteilung A, 31.1 Prozent in die Abteilung B und 4.2 Prozent in die Abteilung C übergetreten (1 Prozent mussten die 6. Klasse wiederholen und 16.1 Prozent sind in das Langgymnasium übergetreten).

Tabelle 6.2: Ergebnisse der logistischen Mehrebenenanalysen zum Einfluss von sozialen Merkmalen auf den Übertritt in die Abteilung A der Sekundarschule

	Modell I	Modell II	Modell III
	Odds ratio	Odds ratio	Odds ratio
Fixe Effekte für die Konstante	1.277	1.639**	1.203
Knaben	0.588*	0.671**	0.596*
Alter (in Monaten)	0.951**	0.944***	0.959*
Deutsch als Zweitsprache	1.262	0.781	1.471 [†]
Soziale Herkunft	1.356*	1.645***	1.289*
Zeugnisnote in Deutsch	37.034***		27.325***
Zeugnisnote in Mathematik	16.455***		11.340***
Deutschleistung (in 100 Punkten)		3.514***	1.578**
Mathematikleistung (in 100 Punkten)		3.816***	1.966***
Variable Effekte			
Klassenebene	0.23*	0.32***	0.15
Modellparameter			
N (listwise delete)	1414	1440	1392

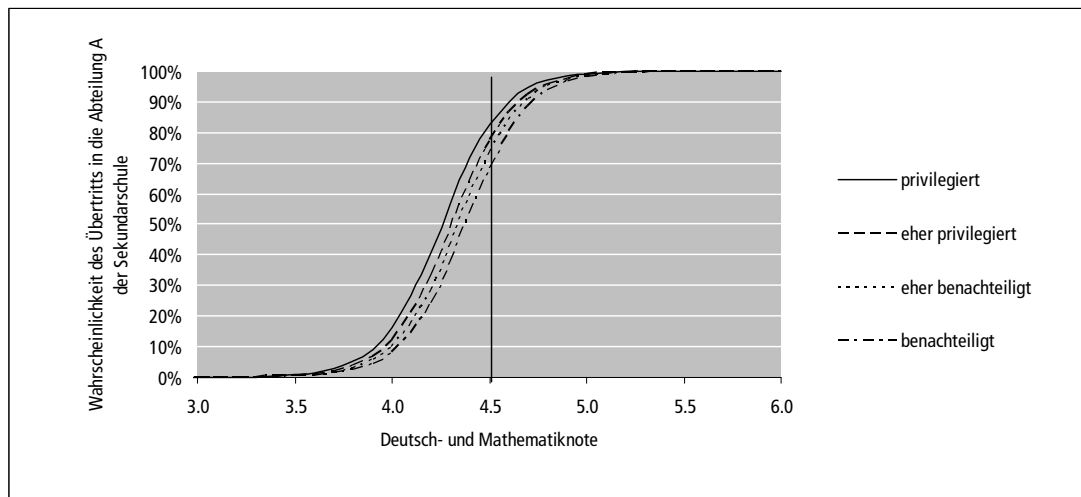
Anmerkungen: Binär logistische Mehrebenenanalyse (Schülerinnen und Schüler sind in Klassen gruppiert), Random-Intercept-Model, Schätzverfahren: EM Laplace Iterations (100). Abhängige Variable ist der Übertritt in die Abteilung A der Sekundarschule (Sekundarschule Abteilung B/C = 0, Sekundarschule Abteilung A = 1). Alle metrischen Variablen wurden am «Grand Mean» zentriert. Gewichtete Daten. Signifikanzniveaus: [†]p < .10; *p < .05; **p < .01; ***p < .001 (zweiseitige Tests).

Tabelle 6.2 zeigt die Ergebnisse der Mehrebenenanalysen zum Einfluss von sozialen Merkmalen der Schülerinnen und Schüler auf die Chance für den Übertritt in die Abteilung A der Sekundarschule. In der ersten Spalte sind die Variablen aufgeführt, die in die Analysen einbezogen wurden. Als soziale Merkmale wurden wiederum das Geschlecht, das Alter, die Erstsprache und die soziale Herkunft der Schülerinnen und Schüler berücksichtigt, als schulische Vorleistungen die Zeugnisnoten in Deutsch und in Mathematik am Ende der 6. Klasse und/oder die Testleistungen in Deutsch- und in Mathematik (WLE). In den folgenden Spalten sind die zugehörigen relativen Chancen beziehungsweise odds ratios ausgewiesen, falls der Einfluss aller anderen Faktoren konstant gehalten wird.

In Modell I wurden als Indikatoren für die schulischen Leistungen lediglich die Schulnoten, in Modell II lediglich die Testleistungen und in Modell III beide Indikatoren gleichzeitig berücksichtigt. Die Ergebnisse zeigen, dass die Mädchen im Vergleich zu Knaben bei sonst gleichen Lernvoraussetzungen eine deutlich höhere Chance für den Besuch der Abteilung A der Sekundarschule aufweisen. Die Knaben haben im Vergleich zu den Mädchen eine rund 0.6-mal kleinere Chance für den Übertritt in die Abteilung A. Je höher das Alter der Schülerinnen und Schüler ist, desto geringer ist auch die Chance für den Übertritt in die Abteilung A. Der Alterseffekt ist jedoch relativ schwach ausgeprägt. Bei einem Anstieg des Alters um einen Monat verringert sich die Chance um rund den Faktor 0.9. Im Hinblick auf die Erstsprache der Schülerinnen und Schüler zeigt sich lediglich nach Kontrolle der Schulnoten und der Testleistungen ein statistisch signifikanter Effekt ($p = .070$) auf den Übertritt. Schülerinnen und Schüler mit Deutsch als Zweitsprache weisen eine rund 1.5-mal höhere Chance auf für den Besuch der Abteilung A als solche mit Deutsch als Erstsprache. Je höher die soziale Herkunft, desto höher ist auch die Chance für den Übertritt in die Abteilung A. Am stärksten ist dieser Effekt mit einer relativen Chance von rund 1.6 in Modell II ausgeprägt, das als Indikator für die schulischen Leistungen lediglich die Testleistungen einbezieht. Wird lediglich (Modell I) oder zusätzlich (Modell III) um die Schulnoten kontrolliert, verringert sich der Effekt der sozialen Herkunft auf eine relative Chance von rund 1.3. Erwartungsgemäss hängt die Chance für den Übertritt in die Abteilung A auch mit den Schulnoten und den Testleistungen in den beiden Fächern zusammen. Je höher die Schulnoten oder die Testleistungen, desto höher ist auch die Chance für den Übertritt in die Abteilung A. Die Ergebnisse zeigen, dass die Deutschnoten im Vergleich zu den Mathematiknoten einen deutlich stärkeren Effekt auf die Chance für den Übertritt in die Abteilung A aufweisen. Während ein Anstieg der Deutschnote um 1 Einheit (= 1 Note) die Chance um den Faktor 37 erhöht, bewirkt ein vergleichbarer Anstieg der Mathematiknote eine Erhöhung der Chancen um den Faktor 16.5.

In Abbildung 6.3 ist der Effekt der sozialen Herkunft auf den Übertritt in die Abteilung A der Sekundarschule grafisch dargestellt. Abgebildet sind die Wahrscheinlichkeiten (Modell I) für den Übertritt nach Deutsch- und Mathematiknoten für Schülerinnen und Schüler unterschiedlicher sozialer Herkunft. Beispielsweise beträgt die Wahrscheinlichkeit für den Übertritt in die Abteilung A bei einem Mädchen mittleren Alters mit Deutsch als Erstsprache und einer Note von 4.5 in beiden Fächern und mit privilegierter sozialer Herkunft rund 83 Prozent; mit eher privilegierter sozialer Herkunft 77 Prozent; mit eher benachteiligter sozialer Herkunft 73 Prozent; mit benachteiligter sozialer Herkunft noch 69 Prozent.

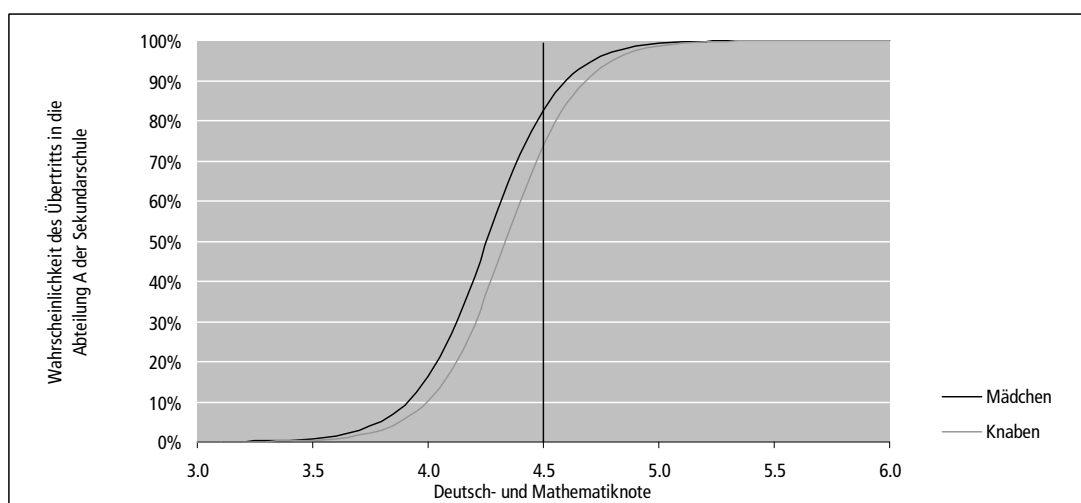
Abbildung 6.3: Wahrscheinlichkeiten des Übertritts in die Abteilung A der Sekundarschule nach Deutsch- und Mathematiknoten für Schülerinnen unterschiedlicher sozialer Herkunft (logistische Mehrebenenanalyse)



Anmerkungen: Die Wahrscheinlichkeiten beziehen sich jeweils auf ein Mädchen mittleren Alters mit Deutsch als Erstsprache.

Abbildung 6.4 zeigt die Wahrscheinlichkeiten für den Übertritt in die Abteilung A in Abhängigkeit der Deutsch- und Mathematiknoten für Mädchen und Knaben. Beispielsweise beträgt die Wahrscheinlichkeit für den Übertritt in den Typ A bei einem Mädchen mittleren Alters mit Deutsch als Erstsprache und einer Note von 4.5 in den Fächern Deutsch und Mathematik und mit privilegierter sozialer Herkunft rund 83 Prozent, bei einem Knaben rund 74 Prozent.

Abbildung 6.4: Wahrscheinlichkeiten des Übertritts in die Abteilung A der Sekundarschule nach Deutsch- und Mathematiknoten für Mädchen und Knaben (logistische Mehrebenenanalyse)



Anmerkungen: Die Wahrscheinlichkeiten beziehen sich auf eine Schülerin oder ein Schüler mittleren Alters mit Deutsch als Erstsprache und mit privilegierter sozialer Herkunft.

7 Analysen zu Motivation und lernbezogene Emotionen nach sechs Jahren Primarstufe in Mathematik und Deutsch

Alex Buff

7.1 Einleitung

Angebots-Nutzungs-Modelle (Fend, 1998; Helmke, 2003; Pauli & Reusser, 2006) machen deutlich, dass der «Ertrag» von Schule, wie z.B. akademische Leistungen, nicht allein von der Qualität des schulischen Angebots abhängig ist, sondern in hohem Masse auch davon, wie Schülerinnen und Schüler dieses Angebot nutzen können und nutzen wollen. Wichtige individuelle Eingangsvoraussetzungen für die «Nutzungsqualität» sind auf Seiten der Schülerinnen und Schüler neben Intelligenz und Vorwissen insbesondere auch «motivational-affektive» Merkmale (Helmke & Weinert, 1997), und hier speziell das Vertrauen, das die Schülerinnen und Schüler in ihre eigene Leistungsfähigkeit in einem Fach haben (Kompetenzüberzeugung) sowie der Stellenwert, den sie dem Fach zuschreiben (Valenzüberzeugung).

Der Beitrag des Kapitels fokussiert die beiden motivational-affektiven Merkmale Kompetenz- und Valenzüberzeugungen in den Fächern Mathematik und Deutsch und analysiert deren Entwicklung in der zweiten Hälfte der Primarschulzeit. Von Interesse sind hierbei einerseits allgemeine Trends, andererseits speziell auch allfällige unterschiedliche Entwicklungsverläufe in Gruppen von Schülerinnen und Schülern mit unterschiedlichen schulischen Laufbahnen vom Ende der 3. Klasse bis zum Beginn der Sekundarstufe I. Unterschieden wird hierbei zwischen Schülerinnen und Schülern, die nach der Lernstandserhebung 2009 ins Langgymnasium oder in die Abteilungen A resp. B/C der Sekundarschule übertreten, sowie solchen, die in der Primarstufe verbleiben. Erhoben wurden 2009 auch die beiden lernbezogenen Emotionen Freude und Angst. Bei den diesbezüglich dargestellten Ergebnissen handelt es sich um Momentaufnahmen nach sechs Jahren Primarschulzeit.

Im nachfolgenden Abschnitt 7.2 finden sich Angaben zur Datenbasis, zum eingesetzten Instrumentarium wie auch zu den Analyseverfahren. Abschnitt 7.3 beinhaltet nähere Angaben zu den hinsichtlich der interessierenden Zusammenhängen durchgeführten Analysen sowie eine Darstellung der Ergebnisse.

7.2 Methode

Datenbasis und Datenerhebungen

Detaillierte Angaben zur Stichprobe finden sich in Kapitel 1, Teil 2. In die nachfolgenden Analysen einbezogen wurden diejenigen 1'679 Schülerinnen und Schüler aus 406 Klassen, welche sich 2009 zum dritten Mal an den Lernstandserhebungen beteiligten. Berücksichtigung in den Analysen finden die Angaben der Schülerinnen und Schüler aus den Erhebungen 2006 resp. 2007 sowie 2009. Der Grund für die alleinige Berücksichtigung der Daten aus der zweiten und dritten Erhebung liegt darin, dass die hier zentral interessierenden Variablen Kompetenz- und Valenzüberzeugung erstmals 2006 ($N = 1'564$) bzw. 2007 ($N = 115$) am Ende des dritten Schuljahres erfasst wurden. Bei den 2007 befragten Kindern handelt es sich um Schü-

lerinnen und Schüler, die das Ende der 3. Klasse erst nach vier Schuljahren erreicht hatten (Sonderklasse A, Repetition usw.).

Instrumente

Kompetenz- und valenzbezogene Überzeugungen: Die Erhebung der entsprechenden Konstrukte erfolgte 2006/2007 sowie 2009 identisch. In Mathematik und Deutsch wurden die gleichen Items verwendet, die sich lediglich in der Nennung der Fächer unterschieden, und in beiden Skalen war die Hälfte der Items positiv, die andere negativ formuliert. Verwendet wurde ein vierstufiges Antwortformat: stimmt genau (1) – stimmt eher (2) – stimmt eher nicht (3) – stimmt überhaupt nicht (4). Hinsichtlich der valenzbezogenen Überzeugungen wurde einerseits die Facette intrinsischer Anreiz sowie andererseits die Facette Relevanz/Nutzen mit vier resp. acht Items erhoben. Sowohl in Mathematik als auch in Deutsch ist die Korrelation der Skalen intrinsischer Anreiz sowie Relevanz/Nutzen 2009 wie auch 2006/2007 erheblich: Mathematik $r = .63$ resp. $r = .69$; Deutsch $r = .68$ resp. $r = .71$. In beiden Fächern wurden daher die beiden Skalen zu einer Skala Valenzüberzeugung zusammengefasst. In Tabelle 7.1 sind die Anzahl Items, Beispielitems, die Reliabilitäten sowie Mittelwerte und Standardabweichungen zu beiden Messzeitpunkten hinsichtlich der einzelnen Skalen dargestellt.

Tabelle 7.1: Skalen zu den kompetenz- und valenzbezogenen Überzeugungen

Skalen	n Items	Beispielitems	Reliabilität α		M (SD)	
			06/07	09	06/07	09
Kompetenzüberzeugung M	6	Mathematik ist einfach.	.87	.88	3.14 (.68)	2.80 (.64)
Kompetenzüberzeugung D	6	Ich habe Probleme in Deutsch.	.85	.87	3.17 (.63)	2.92 (.61)
Relevanz/Nutzen M	8	Mathematik ist mir unwichtig.	.83	.84	3.64 (.47)	3.35 (.51)
Relevanz/Nutzen D	8	Deutsch kann ich brauchen.	.84	.87	3.59 (.53)	3.36 (.50)
Intrinsischer Anreiz M	4	Mathematik macht mir Spass.	.88	.90	3.37 (.77)	2.89 (.78)
Intrinsischer Anreiz D	4	Deutsch ist langweilig.	.89	.88	3.34 (.77)	2.96 (.73)
Valenzüberzeugung M	Intrinsischer Anreiz und Relevanz/Nutzen		.76	.74	3.51 (.57)	3.12 (.58)
Valenzüberzeugung D	Intrinsischer Anreiz und Relevanz /Nutzen		.80	.78	3.46 (.60)	3.16 (.56)

Anmerkungen: M = Mittelwert, SD = Standardabweichung. M = Mathematik, D = Deutsch.

Lernbezogene Emotionen: In beiden Fächern wurden zur Erhebung der beiden lernbezogenen Emotionen Freude und Angst identische Items basierend auf den entsprechenden Skalen von Pekrun, Goetz und Perry (2005) verwendet, die sich wiederum nur in der Nennung der Fächer unterschieden. Das Antwortformat war auch hier vierstufig: stimmt genau (1) – stimmt eher (2) – stimmt eher nicht (3) – stimmt überhaupt nicht (4). *Freude* wurde mittels jeweils fünf Items erfasst (z.B.: „Ich lerne gerne Neues in Mathematik.“ oder „Wenn ich Deutsch lerne, bin ich richtig happy.“). Die Reliabilitäten, Mittelwerte und Standardabweichungen betragen für Mathematik resp. Deutsch 2009: $\alpha = .87$, $M = 2.51$, $SD = .72$ resp. $\alpha = .89$, $M = 2.57$, $SD = .74$. Auch im Falle von *Angst* basieren die Skalen für Mathematik und Deutsch auf fünf

Items (z.B.: „Wenn ich Mathematik lerne, werde ich nervös.“ oder „Beim Deutschlernen muss ich immer wieder mit meiner Angst kämpfen.“). Die Reliabilitäten, Mittelwerte und Standardabweichungen betragen für Mathematik resp. Deutsch 2009: $\alpha = .84$, $M = 1.84$, $SD = .64$ resp. $\alpha = .84$, $M = 1.62$, $SD = .57$.

Die Skalenbildung erfolgte in allen Fällen durch Summation der (falls notwendig vorher umgepolten) Einzelitems, wobei die Skalenwerte jeweils auf das ursprüngliche Antwortformat transformiert wurden. Höhere Werte bedeuten nachfolgend bessere Kompetenzüberzeugungen, mehr Angst usw..

Schulische Laufbahnen Schuljahr 2009/2010: Die effektive Zuteilung der Schülerinnen und Schüler im auf die Erhebung 2009 folgenden Schuljahr 2009/2010 basiert auf Angaben der Bildungsstatistik aus dem Februar 2010.

Fehlende Werte und Datenauswertung

Fehlende Werte wurden mittels des Programmpakets NORM 2.03 imputiert (vgl. Allison, 2001; Lüdtke, Robitzsch, Trautwein & Köller, 2007). Bei den Analysen zu den Entwicklungsverläufen handelt es sich um Mehrebenenanalysen unter Verwendung des Programms HLM 6.03 (2004). Als Maß für den globalen Modellfit wird für die einzelnen Analysemodelle jeweils das Maddala-Maximum-Likelihood- R^2 (Maddala, 1986) angegeben. Zusätzliche Paarvergleiche sowie Analysen bezüglich der lernbezogenen Emotionen wurden mittels Mplus 5.2 (2007) durchgeführt, deskriptive Analysen mit SPSS 15.0 (2006). Bei allen Signifikanzangaben handelt es sich um zweiseitige Tests.

7.3 Ergebnisse

Die Ergebnispräsentation hinsichtlich der Entwicklungen von Kompetenz- und Valenzüberzeugungen beschränkt sich im Falle der vier Gruppen mit unterschiedlichen Schullaufbahnen nach der Lernstandserhebung 2009 (Langgymnasium, Abteilung A oder B/C der Sekundarschule, Primarstufe) auf die Darstellung der Ergebnisse der Analysen in denen die Gruppe Langgymnasium die Referenzgruppe (0) darstellte. Zusatzanalysen, in denen andere Gruppen als Basiskategorie dienten, oder zusätzlich durchgeführte Paarvergleiche werden aus Platzgründen nicht dargestellt. Auch hinsichtlich der Präsentation der Ergebnisse bezüglich der lernbezogenen Emotionen beschränkt sich diese auf die Befunde mit Langgymnasium als Referenzgruppe. Ebenso werden auch hier keine zusätzlich durchgeführten Paarvergleiche berichtet. Inhaltlich ausführlich kommentiert sind die Ergebnisse in Kapitel 7, Teil 1.

Entwicklung der Kompetenzüberzeugung in Mathematik und Deutsch

Tabelle 7.2 zeigt die Ergebnisse hinsichtlich der Entwicklung der Kompetenzüberzeugung in Mathematik und Deutsch für die beiden Stichproben 06/09 sowie 07/09. In Tabelle 7.3 ist gleiches für die vier Gruppen von Schülerinnen und Schülern mit unterschiedlichen Schullaufbahnen nach der Lernstandserhebung 2009 dargestellt. In den Spalten «b-Koeff.» handelt es sich jeweils um unstandardisierte Regressionskoeffizienten.

Tabelle 7.2: Ergebnisse der Mehrebenenanalysen zur Entwicklung der Kompetenzüberzeugung in Mathematik und Deutsch getrennt nach Stichproben

	Mathematik	Deutsch
Fixe Effekte für die Konstante	b-Koeff.	b-Koeff.
Konstante	3.17***	3.17***
Stichprobe 07/09	-.37***	-.11 ⁺
Fixe Effekte für die linearen Zeiteffekte		
Konstante	-.34***	-.24***
Stichprobe 07/09	.04	-.01
Variable Effekte (Varianzen)		
Schülerinnen und Schüler	.18***	.15***
Zeit	.24***	.23***
Erklärte Varianz		
R ² (Gesamt)	17.0 %	11.6 %

Anmerkungen: b-Koeff. = unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Random-Intercept Modelle).
Referenzgruppe: Stichprobe 06/09 = 0. N = 1679.
Signifikanzniveaus: ⁺p < .10; *p < .05; **p < .01; ***p < .001.

Die Werte in der Rubrik «Fixe Effekte für die Konstante» in Tabelle 7.2, Zeile «Konstante» zeigen, wie hoch die Kompetenzüberzeugungen Ende der 3. Klasse in der Referenzgruppe (Stichprobe 06/09) waren. In der Zeile «Stichprobe 07/09» kommt zum Ausdruck, ob und wie stark sich die Stichprobe 07/09 hinsichtlich ihrer Kompetenzüberzeugungen Ende der 3. Klasse von der Referenzgruppe unterscheidet. Im Falle von Mathematik bspw., lag der Mittelwert der Kompetenzüberzeugung in der Stichprobe 06/07 Ende der 3. Klasse bei 3.17, derjenige in der Stichprobe 07/09 jedoch tiefer bei 2.80 (3.17 - 0.37).

In den Werten der Rubrik «Fixe Effekte für die linearen Zeiteffekte» werden die Veränderungen der Kompetenzüberzeugungen bis zur Lernstandserhebung 2009 zum Ausdruck gebracht. In der Zeile «Konstante» zeigt sich, ob und wie stark sich die Kompetenzüberzeugungen im Falle der Referenzgruppe (Stichprobe 06/09) verbessern oder verschlechtern. In der Zeile «Stichprobe 07/09» wird deutlich, ob und inwieweit sich die Stichprobe 07/09 hinsichtlich der Veränderungen in den Kompetenzüberzeugungen von der Referenzgruppe unterscheidet. Wiederum im Falle von Mathematik findet sich bei der Referenzgruppe ein Absinken der Kompetenzüberzeugung um 0.34 Einheiten. Die Kompetenzüberzeugung liegt 2009 in dieser Gruppe noch bei 2.83 (3.17 - 0.34). Auch in der Stichprobe 07/09 sinkt die Kompetenzüberzeugung jedoch etwas weniger stark als in der Referenzgruppe um 0.30 Einheiten (-0.34 + 0.04). In der Stichprobe 07/09 liegt der Mittelwert der Kompetenzüberzeugung in Mathematik bei 2.53 (2.83 - 0.30).

Tabelle 7.3: Ergebnisse der Mehrebenenanalysen zur Entwicklung der Kompetenzüberzeugung in Mathematik und Deutsch getrennt nach schulischen Laufbahnen

	Mathematik	Deutsch
Fixe Effekte für die Konstante	b-Koeff.	b-Koeff.
Konstante	3.44***	3.41***
Abteilung A Sekundarschule	-.18***	-.17***
Abteilung B/C Sekundarschule	-.48***	-.43***
Primarstufe	-.78***	-.50***
Fixe Effekte für die linearen Zeiteffekte		
Konstante	-.11**	-.11**
Abteilung A Sekundarschule	-.22***	-.13**
Abteilung B/C Sekundarschule	-.43***	-.21***
Primarstufe	.00	-.12
Variable Effekte (Varianzen)		
Schülerinnen und Schüler	.13***	.12***
Zeit	.23***	.23***
Erklärte Varianz		
R^2 (Gesamt)	27.1 %	18.3 %

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Random-Intercept Modelle). Referenzgruppe: Langgymnasium = 0. $N = 1564$. Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

Die Werte in der Rubrik «Fixe Effekte für die Konstante» in Tabelle 7.3, Zeile «Konstante» zeigen, wie hoch die Kompetenzüberzeugungen Ende der 3. Klasse in der Referenzgruppe (Langgymnasium) waren. In der Zeilen «Abteilung A Sekundarschule», «Abteilung B/C Sekundarschule» und «Primarstufe» kommt jeweils zum Ausdruck, ob und wie stark sich die drei Gruppen hinsichtlich ihrer Kompetenzüberzeugungen Ende der 3. Klasse von der Referenzgruppe unterscheiden. Im Falle von Mathematik bspw. lag der Mittelwert der Kompetenzüberzeugung bei Schülerinnen und Schülern, welche später das Langgymnasium besuchen, am Ende der 3. Klasse bei 3.44, derjenige in der Gruppe Abteilung B/C der Sekundarschule zum Beispiel mit $M = 2.96$ ($3.44 - 0.48$) jedoch tiefer.

In den Werten der Rubrik «Fixe Effekte für die linearen Zeiteffekte» werden die Veränderungen der Kompetenzüberzeugungen bis zur Lernstandserhebung 2009 zum Ausdruck gebracht. In der Zeile «Konstante» zeigt sich ob und wie stark sich die Kompetenzüberzeugungen im Falle der Referenzgruppe (Langgymnasium) verbessern oder verschlechtern. In der Zeilen «Abteilung A Sekundarschule», «Abteilung B/C Sekundarschule» und «Primarstufe» wird deutlich, ob und inwieweit sich diese Gruppen hinsichtlich der Veränderungen in den Kompetenzüberzeugungen von der Referenzgruppe unterscheiden. Wiederum in Mathematik findet sich im Falle der Referenzgruppe ein Absinken der Kompetenzüberzeugung um 0.11

Einheiten. Die Kompetenzüberzeugung liegt 2009 in dieser Gruppe noch bei 3.33 (3.44 – 0.11). Mit Ausnahme der Gruppe Primarstufe sinkt in den beiden anderen Gruppen Sekundarschule die Kompetenzüberzeugung jedoch noch stärker als in der Referenzgruppe. Bei denjenigen Schülerinnen und Schülern, die in die Abteilung B/C der Sekundarschule wechseln, bspw. um 0.54 Einheiten (-0.11 + -0.43). Der Mittelwert in Mathematik liegt hier 2009 bei 2.42 (2.96 - 0.54).

Entwicklung der Valenzüberzeugung in Mathematik und Deutsch

Tabelle 7.4 zeigt die Ergebnisse hinsichtlich der Entwicklung der Valenzüberzeugung in Mathematik und Deutsch für die beiden Stichproben 06/09 sowie 07/09. In Tabelle 7.5 ist gleiches für die vier Gruppen von Schülerinnen und Schülern mit unterschiedlichen Schullaufbahnen nach der Lernstandserhebung 2009 dargestellt. In den Spalten «b-Koeff.» handelt es sich jeweils um unstandardisierte Regressionskoeffizienten. Die Interpretation der Tabellen ist analog derjenigen zu den Entwicklungen der Kompetenzüberzeugungen (vgl. oben Tabellen 7.2 und 7.3).

Tabelle 7.4: Ergebnisse der Mehrebenenanalysen zur Entwicklung der Valenzüberzeugung in Mathematik und Deutsch getrennt nach Stichproben

	Mathematik	Deutsch
Fixe Effekte für die Konstante	b-Koeff.	b-Koeff.
Konstante	3.51***	3.46***
Stichprobe 07/09	-.04	.04
Fixe Effekte für die linearen Zeiteffekte		
Konstante	-.40***	-.31***
Stichprobe 07/09	.12	.14*
Variable Effekte (Varianzen)		
Schülerinnen und Schüler	.09***	.13***
Zeit	.24***	.21***
Erklärte Varianz		
R ² (Gesamt)	14.0 %	13.4 %

Anmerkungen: b-Koeff. = unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Random-Intercept Modelle). Referenzgruppe: Stichprobe 06/09 = 0. N=1674. Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$.

Tabelle 7.5: Ergebnisse der Mehrebenenanalysen zur Entwicklung der Valenzüberzeugung in Mathematik und Deutsch getrennt nach schulischen Laufbahnen

	Mathematik	Deutsch
Fixe Effekte für die Konstante	b-Koeff.	b-Koeff.
Konstante	3.54***	3.44***
Abteilung A Sekundarschule	.01	.04
Abteilung B/C Sekundarschule	-.04	.02
Primarstufe	-.25***	-.04
Fixe Effekte für die linearen Zeiteffekte		
Konstante	-.30***	-.17***
Abteilung A Sekundarschule	-.12*	-.16**
Abteilung B/C Sekundarschule	-.15**	-.19***
Primarstufe	.08	-.06
Variable Effekte (Varianz)		
Schülerinnen und Schüler	.09***	.13***
Zeit	.24***	.21***
Erklärte Varianz		
R ² (Gesamt)	15.1 %	14.1 %

Anmerkungen: Unstandardisierte Regressionskoeffizienten linearer Zwei-Ebenenmodelle (Random-Intercept Modelle). Referenzgruppe: Langgymnasium = 0. $N = 1564$. Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

Lernbezogene Emotionen – Freude beim Mathematik- und Deutschlernen

Tabelle 7.6 zeigt die Ergebnisse hinsichtlich der Freude beim Mathematik- und Deutschlernen für die beiden Stichproben 06/09 sowie 07/09. In Tabelle 7.7 ist gleiches für die vier Gruppen von Schülerinnen und Schülern mit unterschiedlichen Schullaufbahnen nach der Lernstandserhebung 2009 dargestellt. In den Spalten «b-Koeff.» handelt es sich jeweils um unstandardisierte Regressionskoeffizienten.

Die Werte in der Zeile «Konstante» in Tabelle 7.6 zeigen, wie gross die Freude beim Mathematik- resp. Deutschlernen 2009 in der Referenzgruppe (Stichprobe 06/09) war. In der Zeile «Stichprobe 07/09» kommt zum Ausdruck, ob und wie stark sich die Stichprobe 07/09 hinsichtlich der Freude beim Mathematik- resp. Deutschlernen 2009 von der Referenzgruppe unterscheidet. Im Falle von Mathematik bspw. findet sich im Falle der Referenzgruppe ein Mittelwert von 2.50, in der Stichprobe 07/09 liegt dieser Wert höher, bei 2.73 (2.50 + 0.23).

Tabelle 7.6: Ergebnisse der Regressionsanalysen zu Freude beim Mathematik- und Deutschlernen getrennt nach Stichproben

	Mathematik	Deutsch
	b-Koeff.	b-Koeff.
Konstante	2.50***	2.55***
Stichprobe 07/09	.23**	.25***
Erklärte Varianz		
R^2	1.0 %	1.0 %

Anmerkungen: b-Koeff. = unstandardisierte Regressionskoeffizienten linearer Einebenen-Modelle.
Referenzgruppe: Stichprobe 06/09 = 0. $N = 1679$.
Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

Die Werte in der Zeile «Konstante» in Tabelle 7.7 zeigen, wie hoch die Freude beim Mathematik- resp. Deutschlernen 2009 in der Referenzgruppe (Langgymnasium) war. In der Zeilen «Abteilung A Sekundarschule», «Abteilung B/C Sekundarschule» und «Primarstufe» kommt jeweils zum Ausdruck, ob und wie stark sich die drei Gruppen hinsichtlich ihrer Freude beim Mathematik- resp. Deutschlernen 2009 von der Referenzgruppe unterschieden. Im Falle von Mathematik bspw. lag der Mittelwert der Kompetenzüberzeugung bei den Schülerinnen und Schülern, welche später das Langgymnasium besuchen, bei 2.61, derjenige in der Gruppe Abteilung B/C der Sekundarschule zum Beispiel jedoch tiefer bei 2.39 (2.61 - .22).

Tabelle 7.7: Ergebnisse der Regressionsanalysen zu Freude beim Mathematik- und Deutschlernen getrennt schulischen Laufbahnen

	Mathematik	Deutsch
	b-Koeff.	b-Koeff.
Konstante	2.61***	2.64***
Abteilung A Sekundarschule	-.09	-.11*
Abteilung B/C Sekundarschule	-.22***	-.12*
Primarstufe	-.05	.03
Erklärte Varianz		
R^2	1.0 %	1.0 %

Anmerkungen: b-Koeff. = unstandardisierte Regressionskoeffizienten linearer Einebenen-Modelle.
Referenzgruppe: Stichprobe 06/09 = 0. $N = 1564$.
Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

Lernbezogene Emotionen – Angst beim Mathematik- und Deutschlernen

Tabelle 7.8 zeigt die Ergebnisse hinsichtlich der Angst beim Mathematik- und Deutschlernen für die beiden Stichproben 06/09 sowie 07/09. In Tabelle 7.9 ist gleiches für die vier Gruppen von Schülerinnen und Schüler mit unterschiedlichen Schullaufbahnen nach der Lernstandserhebung 2009 dargestellt. In den Spalten «b-Koeff.» handelt es sich jeweils um unstandardisierte Regressionskoeffizienten. Die Interpretation der Tabellen ist analog zu derjenigen der Tabellen 7.6 und 7.7.

Tabelle 7.8: Ergebnisse der Regressionsanalysen zu Angst beim Mathematik- und Deutschlernen getrennt nach Stichproben

	Mathematik	Deutsch
	b-Koeff.	b-Koeff.
Konstante	1.82***	1.61***
Stichprobe 07/09	.33**	.12*
Erklärte Varianz		
R^2	2.0 %	1.0 %

Anmerkungen: b-Koeff. = unstandardisierte Regressionskoeffizienten linearer Einebenen-Modelle.
Referenzgruppe: Stichprobe 06/09 = 0. $N = 1679$.
Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

Tabelle 7.9: Ergebnisse der Regressionsanalysen zu Angst beim Mathematik- und Deutschlernen getrennt nach schulischen Laufbahnen

	Mathematik	Deutsch
	b-Koeff.	b-Koeff.
Konstante	1.38***	1.30***
Abteilung A Sekundarschule	.34***	.24***
Abteilung B/C Sekundarschule	.79***	.55***
Primarstufe	.69***	.53***
Erklärte Varianz		
R^2	19.8 %	12.2 %

Anmerkungen: b-Koeff. = unstandardisierte Regressionskoeffizienten linearer Einebenen-Modelle.
Referenzgruppe: Stichprobe 06/09 = 0. $N = 1564$.
Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

8 Analysen zu Motivation und Leistung in Mathematik und Deutsch

Alex Buff

8.1 Einleitung

Was sich Schülerinnen und Schüler leistungsmässig selbst zutrauen (Kompetenzüberzeugung) und welchen Stellenwert sie einem Lerngegenstand bzw. einem Fach beimessen, wird für deren Lernerfolg als zentral erachtet. Kompetenz- und Valenzüberzeugungen sind wichtige Ingredienzien einer lern-leistungsbezogenen motivationalen Orientierung, welche entscheidend ist für die Qualität schulischer Lernaktivitäten und damit auch für die Leistungen. Der Beitrag fokussiert das Zusammenspiel von Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler einerseits und ihren Leistungen in den Fächern Mathematik und Deutsch andererseits. In den Analysen berücksichtigt werden eine Reihe leistungsrelevanter individueller Merkmale (z.B. Intelligenz, Eingangsleistungen zu Beginn der Schulzeit) sowie strukturelle Merkmale des familiären Hintergrundes, welche sich in relativ konsistenter Weise als mitentscheidend für den Lernerfolg erwiesen haben, z.B. die soziale Herkunft. Vermutet wird jedoch, dass neben strukturellen auch noch andere Charakteristiken der Familie leistungsrelevant sein könnten, so etwa elterliche Überzeugungen oder auch lernbezogene Unterstützung. Konkret interessieren im vorliegenden Falle die Vorstellungen, welche sich Eltern von der Leistungsfähigkeit ihrer Kinder machen (elterliche kindbezogene Kompetenzüberzeugungen). Untersucht wird, welche Bedeutung diesen hinsichtlich der Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler sowie deren Leistungen zukommt.

Im nachfolgenden Abschnitt 8.2 finden sich Angaben zur Datenbasis, zum eingesetzten Instrumentarium wie auch zu den Analyseverfahren. Abschnitt 8.3 beinhaltet nähere Angaben zu den hinsichtlich der interessierenden Zusammenhänge durchgeführten Analysen sowie eine Darstellung der Ergebnisse.

8.2 Methode

Datenbasis und Datenerhebungen

Detaillierte Angaben zur Stichprobe finden sich in Kapitel 1, Teil 2. In die nachfolgenden Analysen einbezogen wurden diejenigen 1'679 Schülerinnen und Schüler aus 406 Klassen, welche sich 2009 zum dritten Mal an den Lernstandserhebungen beteiligten. Bei $n = 1'564$ Schülerinnen und Schülern fand die zweite Erhebung im Jahr 2006 statt (Stichprobe 06), bei $n = 115$ Schülerinnen und Schülern 2007 (Stichprobe 07). Im Falle letzterer handelt es sich um Schülerinnen und Schüler, die das Ende der 3. Klasse erst nach vier Schuljahren erreicht hatten (Sonderklasse A, Repetition usw.). Die erste Lernstandserhebung erfolgte bei allen im Jahr 2003, kurz nach Schuleintritt.

Instrumente

Leistungstests und Intelligenz: Die in Mathematik und Deutsch 2009 eingesetzten Leistungstests sind in Kapitel 1, Teil 2 dargestellt (bezüglich der Erhebung 2003 vgl. Moser, Berweger & Stamm, 2005a, 2005b, 2005c, bezüglich der Erhebungen 2006/2007 vgl. Keller & Moser,

2008a, 2008b). Bei den Leistungsmassen Mathematik und Deutsch handelt es sich um Globalscores, errechnet aus allen geprüften Teilaspekten. Die Intelligenz wurde zu Beginn der 1. Klasse mit dem CFT 1 (Weiss & Osterland, 1997) erfasst.

Soziale Herkunft sowie Deutsch als Erst- resp. Zweitsprache: Zur Bestimmung der sozialen Herkunft wurden Elternangaben aus den Erhebungen 2006/2007 verwendet (vgl. Keller & Moser, 2008a). Gleiches gilt für die Angaben bezüglich Deutsch als Erst- resp. Zweitsprache.

Elterliche kindbezogene Kompetenzüberzeugungen: Die Erhebung erfolgte 2006/2007 für Mathematik und Deutsch getrennt mit jeweils drei Items, die sich lediglich in der Nennung der Fächer unterschieden (z.B.: „Unser / das Kind ist gut in Mathematik.“ oder „Unser / das Kind hat keine Probleme in Deutsch.“). Verwendet wurde ein sechsstufiges Antwortformat: stimmt ganz genau (6) – (5) – (4) – (3) – (2) – (1) stimmt überhaupt nicht. Die Reliabilitäten, Mittelwerte und Standardabweichungen betragen in Mathematik resp. Deutsch: $\alpha = .91$, $M = 4.57$, $SD = 1.17$ resp. $\alpha = .91$, $M = 4.51$, $SD = 1.17$.

Kompetenz- und valenzbezogene Überzeugungen: Die Erhebung der entsprechenden Konstrukte erfolgte 2006/2007 sowie 2009 identisch. In Mathematik und Deutsch wurden die gleichen Items verwendet, die sich lediglich in der Nennung der Fächer unterschieden, und in beiden Skalen war die Hälfte der Items positiv, die andere negativ formuliert. Verwendet wurde ein vierstufiges Antwortformat: stimmt genau (1) – stimmt eher (2) – stimmt eher nicht (3) – stimmt überhaupt nicht (4). Hinsichtlich der valenzbezogenen Überzeugungen wurden einerseits die Facette intrinsischer Anreiz sowie andererseits die Facette Relevanz/Nutzen mit vier resp. acht Items erhoben. Sowohl in Mathematik als auch in Deutsch ist die Korrelation der Skalen intrinsischer Anreiz sowie Relevanz/Nutzen 2009 wie auch 2006/2007 erheblich: Mathematik $r = .63$ resp. $r = .69$; Deutsch $r = .68$ resp. $r = .71$. In beiden Fächern wurden daher die beiden Skalen zu einer Skala Valenzüberzeugung zusammengefasst. In Tabelle 8.1 sind die Anzahl Items, Beispielitems, die Reliabilitäten sowie Mittelwerte und Standardabweichungen zu beiden Messzeitpunkten hinsichtlich der einzelnen Skalen dargestellt.

Die Skalenbildung erfolgte bei den elterlichen kindbezogenen Kompetenzüberzeugungen sowie bei den Kompetenz- und Valenzüberzeugungen der Schülerinnen und Schüler durch Summation der (falls notwendig vorher umgepolten) Einzelitems, wobei die Skalenwerte jeweils auf das ursprüngliche Antwortformat transformiert wurden. Höhere Werte bedeuten nachfolgend bessere Kompetenzüberzeugungen, günstigere elterliche kindbezogene Kompetenzüberzeugungen usw.

Tabelle 8.1: Skalen zu den kompetenz- und valenzbezogenen Überzeugungen

Skalen	n Items	Beispielitems	Reliabilität α		M (SD)	
			06/07	09	06/07	09
Kompetenzüberzeugung M	6	Mathematik ist einfach.	.87	.88	3.14 (.68)	2.80 (.64)
Kompetenzüberzeugung D	6	Ich habe Probleme in Deutsch.	.85	.87	3.17 (.63)	2.92 (.61)
Relevanz/Nutzen M	8	Mathematik ist mir unwichtig.	.83	.84	3.64 (.47)	3.35 (.51)
Relevanz/Nutzen D	8	Deutsch kann ich brauchen.	.84	.87	3.59 (.53)	3.36 (.50)
Intrinsischer Anreiz M	4	Mathematik macht mir Spass.	.88	.90	3.37 (.77)	2.89 (.78)
Intrinsischer Anreiz D	4	Deutsch ist langweilig.	.89	.88	3.34 (.77)	2.96 (.73)
Valenzüberzeugung M	Intrinsischer Anreiz und Relevanz/Nutzen		.76	.74	3.51 (.57)	3.12 (.58)
Valenzüberzeugung D	Intrinsischer Anreiz und Relevanz /Nutzen		.80	.78	3.46 (.60)	3.16 (.56)

Anmerkungen: M = Mittelwert, SD = Standardabweichung. M = Mathematik, D = Deutsch.

Fehlende Werte und Datenauswertung

Fehlende Werte wurden mit dem Programm NORM 2.03 (2000) imputiert (vgl. Allison, 2001; Lüdtke, Robitzsch, Trautwein & Köller, 2007); fünf imputierte Datensätze. Bei den Analysen handelt es sich um Regressionsanalysen unter Verwendung des Programms Mplus 5.2 (2007). Indirekte Effekte wurden mittels «product-of-coefficient tests» auf statistische Signifikanz geprüft. Zur Anwendung kam eine Erweiterung des Sobel-Tests für Fälle mit zwei Mediatorvariablen (vgl. Taylor, MacKinnon & Tein, 2008). Bei allen Signifikanzangaben handelt es sich um zweiseitige Tests.

8.3 Ergebnisse

Die hinsichtlich der interessierenden Zusammenhänge durchgeführten Analysen zu den direkten und indirekten Effekten werden nachfolgend erläutert und deren Ergebnisse dargestellt. Inhaltlich ausführlich kommentiert sind die Befunde in Kapitel 8, Teil 1.

Direkte Effekte

Tabelle 8.2 zeigt die Ergebnisse der Analysen hinsichtlich Mathematik, Tabelle 8.3 hinsichtlich Deutsch. Von Interesse waren in beiden Fällen fünf Analysemodelle:

Modell 1: Prädiktion der Leistungen 2006/2007 aufgrund der Kompetenz- und Valenzüberzeugungen 2006/2007 sowie der elterlichen kindbezogenen Kompetenzüberzeugungen 2006/2007 in Mathematik/Deutsch.

Modell 2: Prädiktion der Kompetenzüberzeugungen 2009 aufgrund der Kompetenz- und Valenzüberzeugungen 2006/2007, der elterlichen kindbezogenen Kompetenzüberzeugungen 2006/2007 sowie der Leistungen 2006/2007 in Mathematik/Deutsch.

Modell 3: Prädiktion der Valenzüberzeugungen 2009 aufgrund der Kompetenz- und Valenzüberzeugungen 2006/2007, der elterlichen kindbezogenen Kompetenzüberzeugungen 2006/2007 sowie der Leistungen 2006/2007 in Mathematik/Deutsch.

Modell 4: Prädiktion der Leistungen 2009 aufgrund der Kompetenz- und Valenzüberzeugungen 2006/2007, der elterlichen kindbezogenen Kompetenzüberzeugungen 2006/2007 sowie der Leistungen 2006/2007 in Mathematik/Deutsch.

Modell 5: Prädiktion der Leistungen 2009 aufgrund der Kompetenz- und Valenzüberzeugungen 2006/2007, der elterlichen kindbezogenen Kompetenzüberzeugungen 2006/2007, der Leistungen 2006/2007 sowie der Kompetenz- und Valenzüberzeugungen 2009 in Mathematik/Deutsch.

Tabelle 8.2: Ergebnisse der Regressionsanalysen zu Kompetenz- und Valenzüberzeugung sowie Leistungen in Mathematik

	Leistungen Mathematik 2006/2007	Kompetenz- überzeugung 2009	Valenz- überzeugung 2009	Leistungen Mathematik 2009	Leistungen Mathematik 2009
<i>Prädiktoren</i>	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5
Kompetenzüberzeugung 06/07	.11***	.21***	-.00	.08***	.03
Valenzüberzeugung 06/07	.07**	-.01	.22***	.01	.01
Mathematikleistungen 06/07	•	.20***	.07*	.29***	.23***
Kompetenzüberzeugung 09	•	•	•	•	.26***
Valenzüberzeugung 09	•	•	•	•	.04*
Elterliche kindbezogene Kompetenz- überzeugung 06/07	.22***	.22***	.11***	.15***	.09**
Geschlecht	-.04 ⁺	.08***	.09***	-.05**	-.08***
Alter	-.01	-.06**	-.01	-.04*	-.03
Soziale Herkunft	.12***	.05 ⁺	-.05	.18***	.17***
Deutsch Erst- / Zweitsprache	-.10***	-.01	.07*	-.04	-.04 ⁺
Intelligenz	.27***	.07**	-.01	.12***	.10***
Stichprobe 06/07	.08***	-.04 ⁺	.03	-.18***	-.18***
Leistungen Mathematik 03	.21***	.00	-.05	.11***	.11***
Erklärte Varianz					
R^2	43.4%	32.9%	11.7%	51.7%	57.2%

Anmerkungen: Standardisierte Regressionskoeffizienten linearer Ebenen-Modelle. Geschlecht: Mädchen = 0, Knaben = 1. Deutsch als Erst- resp. Zweitsprache: Erstsprache = 0, Zweitsprache = 1. Stichprobe 06/07: 2006 = 0, 2007 = 1. $N = 1679$. Signifikanzniveaus: * $p < .10$; ** $p < .05$; *** $p < .01$; **** $p < .001$.

Tabelle 8.3: Ergebnisse der Regressionsanalysen zu Kompetenz- und Valenzüberzeugung sowie Leistungen in Deutsch

	Leistungen Deutsch 2006/2007	Kompetenz- überzeugung 2009	Valenz- überzeugung 2009	Leistungen Deutsch 2009	Leistungen Deutsch 2009
<i>Interessierende Prädiktoren</i>	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5
Kompetenzüberzeugung 06/07	.07**	.21***	.08**	.02	-.01
Valenzüberzeugung 06/07	.05*	.08**	.29***	.01	-.00
Deutschleistungen 06/07	•	.15***	.09**	.41***	.39***
Kompetenzüberzeugung 09	•	•	•	•	.12***
Valenzüberzeugung 09	•	•	•	•	.02
Elterliche kindbezogene Kompetenzüberzeugung 06/07	.22***	.17***	.03	.09***	.07**
Geschlecht	-.14***	-.03	-.05*	-.04*	-.04*
Alter	-.03	-.10***	-.05*	-.04*	-.02
Soziale Herkunft	.13***	.03	.02	.16***	.15***
Deutsch Erst- / Zweitsprache	-.08**	-.05 ⁺	.03	.03	.04
Intelligenz	.19***	-.07*	-.12***	.14***	.15***
Stichprobe 06/07	.02	.00	.06*	-.07***	-.07***
Leistungen Lesen 03	.21***	.08**	.01	.09***	.08***
Leistungen Wortschatz 03	.12***	.08*	-.05 ⁺	.13***	.12***
Erklärte Varianz					
R^2	44.7%	28.5%	18.1%	56.3%	57.5%

Anmerkungen: Standardisierte Regressionskoeffizienten linearer Einebenen-Modelle. Geschlecht: Mädchen = 0, Knaben = 1. Deutsch als Erst- resp. Zweitsprache: Erstsprache = 0, Zweitsprache = 1. Stichprobe 06/07: 2006 = 0, 2007 = 1. $N = 1679$. Signifikanzniveaus: ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Bei den weiteren, in den Analysen zu Zusammenhängen der zentral interessierenden Grössen – Kompetenz- und Valenzüberzeugungen, Leistungen und elterliche kindbezogene Kompetenzüberzeugungen – zusätzlich berücksichtigten Prädiktoren (Geschlecht, Alter usw.), handelt es sich um Kontrollvariablen, welche inhaltlich nicht näher interessieren. Hinsichtlich der zentral interessierenden Grössen bedeutet die Mitberücksichtigung der Kontrollvariablen in den Analysen, dass Effekte ersterer um die Einflüsse letzterer «bereinigt» sind; anders ausgedrückt: Zeigt sich bspw. ein Effekt der Kompetenzüberzeugung auf die Leistungen, dann handelt es sich um einen eigenständigen Erklärungsbeitrag der Kompetenzüberzeugung hinsichtlich der Leistungen unabhängig von den Erklärungsbeiträgen aller übrigen im Modell mitberücksichtigten Variablen.

Bei den in den beiden Tabellen 8.2 und 8.3 dargestellten Koeffizienten handelte es sich um standardisierte Regressionskoeffizienten. Der Wertebereich standardisierter Regressionskoeffizienten liegt zwischen -1 und 1. Je stärker der Wert von 0 verschieden ist, desto bedeutsa-

mer ist der Effekt der entsprechenden Variablen hinsichtlich des Merkmals, welches es zu erklären gilt. Inhaltlich indizieren Regressionskoeffizienten mit einem positiven Vorzeichen einen Zusammenhang im Sinne von «je höher, desto höher bzw. je tiefer, desto tiefer», solche mit einem negativen einen Zusammenhang im Sinne «je höher, desto tiefer bzw. je tiefer, desto höher». In Modell 1 (Tabelle 8.2) zeigt sich bspw., dass die Kompetenzüberzeugung der Kinder 2006/2007 hinsichtlich der Vorhersage der Leistungen in Mathematik 2006/2007 etwa gleich bedeutsam ist, wie Deutsch als Erst- resp. Zweitsprache. Allerdings sind die Vorzeichen anders. Hinsichtlich der Kompetenzüberzeugung 2006/2007 gilt: Je höher die Kompetenzüberzeugung, desto besser sind die Leistungen; «je höher, desto höher». Hinsichtlich Deutsch als Erst- resp. Zweitsprache gilt: Kinder mit Deutsch als Erstsprache (0) erzielen die besseren Leistungen als Kinder mit Deutsch als Zweitsprache (1); «je höher, desto tiefer». Als bedeutsamste Prädiktoren der Leistungen in Modell 1 erweisen sich die Leistungen bei Schuleintritt 2003, die elterliche kindbezogene Kompetenzüberzeugung 2006/2007 sowie insbesondere die bei Schuleintritt 2003 ebenfalls erhobene Intelligenz. Je besser die Leistungen bei Schuleintritt, je positiver die elterliche kindbezogene Kompetenzüberzeugung und je höher die Intelligenz, desto besser die Leistungen in Mathematik 2006/2007.

Indirekte Effekte

Von den im Falle von Mathematik und Deutsch interessierenden beiden indirekten Effekten – Kompetenzüberzeugung 2006/2007 → Leistungen 2006/2007 → Kompetenzüberzeugung 2009 → Leistungen 2009 resp. Valenzüberzeugung 2006/2007 → Leistungen 2006/2007 → Valenzüberzeugung 2009 → Leistungen 2009 – erwies sich in beiden Fächern jeweils lediglich der indirekte Effekt der Kompetenzüberzeugung 2006/2007 als statistisch bedeutsam (Sobel Teststatistik Mathematik: $z = 3.28$, $p < .01$; Sobel Teststatistik Deutsch: $z = 2.38$, $p < .05$). In beiden Fächern beeinflussen die Kompetenzüberzeugung Ende der 3. Klasse 2006/2007 die Leistungen 2009 – also zwei bzw. drei Jahre später – auf indirekte Weise positiv, vermittelt über die Leistungen Ende der 3. Klasse sowie die Kompetenzüberzeugung 2009.

9 Analysen zu Unterricht, Motivation und Leistung in Mathematik und Deutsch

Alex Buff

9.1 Einleitung

Der Beitrag fokussiert Zusammenhänge zwischen Qualitätsmerkmalen des Unterrichts in Mathematik und Deutsch einerseits sowie leistungsmässigem Selbstvertrauen (Kompetenzüberzeugung), Freude am Fach (intrinsischer Anreiz) und Leistungen der Schülerinnen und Schüler in den beiden Fächern andererseits. Erhoben wurden in beiden Fächern die Unterrichtsmerkmale *Regelklarheit, Lernunterstützung, Diagnosekompetenz der Lehrperson, Strukturierungshilfen, Erklärkompetenz der Lehrperson, Motivierung thematisch, Motivierung instrumentell* sowie *Beziehung zur Lehrperson* aus Sicht der Schülerinnen und Schüler. Für die Qualitätsmerkmale «guten Unterrichts» werden nicht nur positive leistungsmässige Effekte postuliert, sondern auch positive Effekte hinsichtlich der beiden «motivational-affektiven» Grössen Kompetenzüberzeugung und intrinsischer Anreiz. Diese beiden Merkmale stellen wichtige Ingredienzen einer lern-leistungsbezogenen motivationalen Orientierung dar und werden im «Angebots-Nutzungs-Modell des Unterrichts» von Helmke (2003) als individuelle Merkmale erachtet, über welche Unterricht leistungsmässige Wirkung entfaltet. Geprüft wurde daher u. a., ob sich direkte positive Effekte des Unterrichts auf die beiden «motivational-affektiven» Merkmale sowie indirekte positive Effekte hinsichtlich der Leistungen finden lassen.

Im nachfolgenden Abschnitt 9.2 finden sich Angaben zur Datenbasis, zum eingesetzten Instrumentarium wie auch zu den Analyseverfahren. Abschnitt 9.3 beinhaltet nähere Angaben zu den hinsichtlich der interessierenden Zusammenhänge durchgeführten Analysen sowie eine Darstellung der Ergebnisse.

9.2 Methode

Datenbasis und Datenerhebungen

Detaillierte Angaben zur Stichprobe finden sich in Kapitel 1, Teil 2. In die nachfolgenden Analysen einbezogen wurden diejenigen 1'981 Schülerinnen und Schüler aus 98 Klassen, die sich 2009 in der repräsentativen Klassenstichprobe an der Lernstandserhebung beteiligten. Rund die Hälfte jeweils einer Klasse beantwortete den Fragebogen für das Fach Mathematik, die andere Hälfte denjenigen für das Fach Deutsch.

Instrumente

Leistungstests: Die in Mathematik und Deutsch 2009 eingesetzten Leistungstests sind in Kapitel 1, Teil 2 dargestellt. Bei den Leistungsmassen Mathematik und Deutsch handelt es sich um Globalscores, errechnet aus den geprüften Teilaspekten.

Soziale Herkunft sowie Deutsch als Erst- resp. Zweitsprache: Zur Bestimmung der sozialen Herkunft wurden die Schülerangaben 2009 verwendet; gleiches gilt für die Angaben bezüglich Deutsch als Erst- resp. Zweitsprache (vgl. Kapitel 1, Teil 2).

Kompetenzüberzeugung und intrinsischer Anreiz: Die Erhebung der entsprechenden Konstrukte erfolgte mit den gleichen Instrumenten wie im Falle der Längsschnittstichprobe (vgl. Kapitel 7.2 resp. 8.2, Teil 2). In Mathematik und Deutsch wurden identische Items verwendet, die sich lediglich in der Nennung der Fächer unterschieden, und in beiden Skalen war die eine Hälfte der Items positiv, die andere negativ formuliert. Verwendet wurde ein vierstufiges Antwortformat: stimmt genau (1) – stimmt eher (2) – stimmt eher nicht (3) – stimmt überhaupt nicht (4). Die *Kompetenzüberzeugungen* wurden mit jeweils sechs Items erfasst (z.B.: „Mathematik ist einfach.“). Die Reliabilitäten, Mittelwerte und Standardabweichungen betragen für Mathematik resp. Deutsch: $\alpha = .90$, $M = 2.80$, $SD = .71$ resp. $\alpha = .87$, $M = 2.92$, $SD = .61$. Der *intrinsische Anreiz* der beiden Fächer wurde mittels jeweils vier Items erhoben (z.B.: „Mathematik macht mir Spass.“). Die Reliabilitäten, Mittelwerte und Standardabweichungen betragen für Mathematik resp. Deutsch: $\alpha = .93$, $M = 2.88$, $SD = .89$ resp. $\alpha = .91$, $M = 2.85$, $SD = .82$. Die Skalenbildung erfolgte in beiden Fällen durch Summation der (falls notwendig vorher umgepolten) Einzelitems, wobei die Skalenwerte jeweils auf das ursprüngliche Antwortformat transformiert wurden. Höhere Werte bedeuten nachfolgend bessere Kompetenzüberzeugungen und höheren intrinsischen Anreiz.

Unterrichtsmerkmale: Erhoben wurden acht Unterrichtsmerkmale aus der Sicht der Schülerinnen und Schüler: *Regelklarheit, Lernunterstützung, Diagnosekompetenz der Lehrperson, Strukturierungshilfen, Erklärkompetenz der Lehrperson, Motivierung (thematisch), Motivierung (instrumentell)* sowie *Beziehung zur Lehrperson*. Wiederum unterscheiden sich die Items zu Mathematik und Deutsch lediglich in der Nennung der Fächer. Die Itemauswahl für die einzelnen Skalen basiert auf Kunter et al., (2002), Rakoczy, Buff und Lipowski (2005) sowie dem Skalenhandbuch DESI (T1+T2) – Schülerfragebögen (2005). Verwendet wurde auch hier ein vierstufiges Antwortformat: stimmt genau (1) – stimmt eher (2) – stimmt eher nicht (3) – stimmt überhaupt nicht (4). Die Skalenbildung erfolgte wiederum durch Summation der (falls notwendig vorher umgepolten) Einzelitems, wobei die Skalenwerte jeweils auf das ursprüngliche Antwortformat transformiert wurden. Höhere Werte bedeuten mehr Lernunterstützung, bessere Beziehung zur Lehrperson usw. In Tabelle 9.1 sind die Anzahl Items, Beispielitems, die Reliabilitäten sowie Mittelwerte und Standardabweichungen hinsichtlich der einzelnen Skalen dargestellt.

In einer konfirmatorischen Faktorenanalyse mittels des Programms Mplus 5.2 (2007) auf der Basis der Originaldaten (ohne Fälle mit fehlenden Werten) bestätigt sich in beiden Fächern die theoretisch unterstellte achtfaktorielle Struktur (Modelle ohne korrelierte Fehlerterme). Sowohl für Mathematik als auch für Deutsch resultieren akzeptable Fitindizes (Mathematik: $\chi^2 = 826.83$, $df = 406$, $p = .001$; $CFI = .95$; $TLI = .95$; $RMSEA = .04$; Deutsch: $\chi^2 = 882.91$, $df = 406$, $p = .001$; $CFI = .96$; $TLI = .95$; $RMSEA = .04$).

Tabelle 9.1: Skalen zu den acht Unterrichtsmerkmalen 2009

Skalen	n Items	Beispielitems	Reliabilität α	M (SD)
Regelklarheit M	3	Im Mathematikunterricht ist klar, was man machen darf und was nicht.	.67	3.23 (.61)
Regelklarheit D	3		.69	3.26 (.60)
Lernunterstützung M	6	Im Deutschunterricht gibt mir die Lehrerin/der Lehrer Hinweise, wie ich mich verbessern kann.	.85	3.19 (.59)
Lernunterstützung D	6		.87	3.18 (.60)
Diagnosekompetenz M	3	Unsere Deutschlehrerin/unsere Deutschlehrer merkt sofort, wenn jemand etwas nicht richtig versteht.	.79	2.95 (.67)
Diagnosekompetenz D	3		.83	2.91 (.69)
Strukturierungshilfen M	5	Unsere Mathematiklehrerin/unsere Mathematiklehrer gibt Hinweise, worauf es in der Stunde besonders ankommt.	.77	3.06 (.59)
Strukturierungshilfen D	5		.76	3.06 (.57)
Erklärkompetenz M	4	Unsere Mathematiklehrerin/unsere Mathematiklehrer erklärt Dinge klar und verständlich.	.78	3.11 (.63)
Erklärkompetenz D	4		.77	3.08 (.62)
Motivierung thematisch M	3	Unsere Deutschlehrerin/unsere Deutschlehrer gestaltet den Unterricht spannend.	.81	2.75 (.74)
Motivierung thematisch D	3		.82	2.72 (.74)
Motivierung instrumentell M	3	Unsere Mathematiklehrerin/unsere Mathematiklehrer sagt immer wieder, wie nützlich gute Mathematikkenntnisse im Alltag sind.	.81	2.81 (.75)
Motivierung instrumentell D	3		.85	2.88 (.76)
Beziehung zur Lehrperson M	4	Ich komme gut mit meiner Deutschlehrerin/meinem Deutschlehrer aus.	.85	3.04 (.68)
Beziehung zur Lehrperson D	4		.86	3.03 (.70)

Anmerkungen: M = Mittelwert, SD = Standardabweichung. M = Mathematik, D = Deutsch.

Eine genauere Betrachtung der Interkorrelationen der latenten Faktoren zeigt in beiden Fächern allerdings, dass diese hoch, teilweise sogar extrem hoch sind. In Mathematik beträgt die durchschnittliche Korrelation $r = .62$ (min. $r = .36$, max. $r = .79$), der Median liegt bei $r = .65$. In Deutsch liegen die Werte noch höher: durchschnittliche Korrelation $r = .68$ (min. $r = .45$, max. $r = .81$), Median $r = .71$. Es macht den Anschein, als ob die Schülerinnen und Schüler in diesem Alter oder zumindest in unserer Stichprobe nicht gross zwischen einzelnen Merkmalen differenzieren, sondern sich in ihrem Antwortverhalten eher von einem globalen «Eindruck» leiten lassen. Die hohen Interkorrelationen lassen den Sinn getrennter Analysen für die einzelnen Unterrichtsmerkmale (mehr als) fraglich erscheinen. Angesichts dieser hohen Interkorrelationen und damit der fehlenden diskriminanten Validität im Sinne einer relativ klaren inhaltlichen Unterscheidbarkeit der einzelnen Unterrichtsmerkmale wurden die acht Einzelskalen für die Analysen pro Fach zu einer «Globalskala» *Unterrichtswahrnehmung* zusammengefasst (Reliabilität, Mittelwert und Standardabweichung für Mathematik resp. Deutsch: $\alpha = .88$, $M = 3.02$, $SD = .49$ resp. $\alpha = .90$, $M = 3.01$, $SD = .51$). Höhere Werte indizieren eine positivere Wahrnehmung des Unterrichts insgesamt.

Fehlende Werte und Datenauswertung

Fehlende Werte wurden mit dem Programm NORM 2.03 (2000) imputiert (vgl. Allison, 2001; Lüdtke, Robitzsch, Trautwein & Köller, 2007); fünf imputierte Datensätze. Deskriptive Analysen wurden mit SPSS 15.0 (2006) durchgeführt. In den Regressionsanalysen kam angesichts der hierarchischen Datenstruktur – Einbezug ganzer Klassen in die Studie – das Programm Mplus 5.2 (2007) zur Anwendung. Dessen «Cluster»-Option berücksichtigt (mögliche) Design-Effekte bei der Berechnung der Standardfehler (vgl. Muthén & Muthén, 2007). Indirekte Effekte wurden mittels des Sobel-Tests auf statistische Signifikanz geprüft. Bei allen Signifikanzangaben handelt es sich um zweiseitige Tests.

9.3 Ergebnisse

Die hinsichtlich der interessierenden Zusammenhänge durchgeführten Analysen zu den direkten und indirekten Effekten werden nachfolgend erläutert und deren Ergebnisse dargestellt. Inhaltliche ausführlich kommentiert sind die Befunde in Kapitel 9, Teil 1.

Direkte Effekte

Tabelle 9.2 zeigt die Ergebnisse der Analysen hinsichtlich Mathematik und Deutsch. Von Interesse sind in beiden Fällen drei Analysemodelle:

Modell 1: Prädiktion des intrinsischen Anreizes aufgrund der Unterrichtswahrnehmung in Mathematik/Deutsch.

Modell 2: Prädiktion der Kompetenzüberzeugung aufgrund der Unterrichtswahrnehmung in Mathematik/Deutsch.

Modell 3: Prädiktion der Leistungen in Mathematik und Deutsch. Den theoretischen Überlegungen von Helmke (2003) folgend (vgl. oben), werden hier die beiden Motivationskomponenten intrinsischer Anreiz sowie Kompetenzüberzeugung als zusätzliche Prädiktoren zusammen mit der Unterrichtswahrnehmung in den Analysen simultan berücksichtigt, da Kompetenzüberzeugung und intrinsischer Anreiz a) als unmittelbare Determinanten von Leistungen angesehen werden und b) zentrale vermittelnde Größen hinsichtlich leistungsmässiger Effekte der Unterrichtsgestaltung darstellen.

In allen Analysen zu den Zusammenhängen der zentral interessierenden Größen – Unterrichtswahrnehmung, intrinsischer Anreiz, Kompetenzüberzeugung, und Leistungen – als Kontrollvariablen ebenfalls berücksichtigt wurden die Individualmerkmale soziale Herkunft, Alter, Geschlecht und Deutsch als Erst- bzw. Zweitsprache, als Klassenmerkmale die auf Klassenebene aggregierte soziale Herkunft sowie Deutsch als Erst- bzw. Zweitsprache. Bei Letzterem handelt es sich um den Anteil von Schülerinnen und Schülern in einer Klasse mit Deutsch als Zweitsprache. Die auf Klassenebene aggregierte Unterrichtswahrnehmung erwies sich als Prädiktor in allen Analysen als irrelevant und wurde aus Gründen der Sparsamkeit in den definitiven Analysen nicht einbezogen. Inhaltlich interessieren allfällige Effekte der Kontrollvariablen nicht näher. Hinsichtlich der zentral interessierenden Größen bedeutet die Berücksichtigung der Kontrollvariablen in den Analysen, dass Effekte ersterer um die Einflüsse letzterer «bereinigt» sind; anders ausgedrückt: Zeigt sich bspw. ein Effekt der Kompetenzüberzeugung auf die Leistungen, dann handelt es sich um einen eigenständigen Erklärungsbeitrag

der Kompetenzüberzeugung hinsichtlich der Leistungen unabhängig von den Erklärungsbeiträgen aller übrigen im Modell berücksichtigten Variablen.

Sämtliche in Tabelle 9.2 dargestellten Analysen basieren auf gewichteten Daten.

Tabelle 9.2: Ergebnisse der Regressionsanalysen zu Unterrichtswahrnehmung, intrinsischem Anreiz, Kompetenzüberzeugung sowie Leistungen in Mathematik und Deutsch

	Intrinsischer Anreiz 2009		Kompetenzüberzeugung 2009		Leistungen 2009	
	Mathematik	Deutsch	Mathematik	Deutsch	Mathematik	Deutsch
<i>Prädiktoren</i>	Modell 1		Modell 2		Modell 3	
Unterrichtswahrnehmung M / D	.33***	.41***	.12**	.08 ⁺	-.03	-.10**
Intrinsischer Anreiz M / D	•	•	•	•	-.03	-.02
Kompetenzüberzeugung M / D	•	•	•	•	.45***	.34***
Geschlecht	.13**	-.18***	.23***	-.15***	.01	-.08**
Soziale Herkunft	.10*	.14***	.25***	.27***	.21***	.26***
Deutsch Erst- / Zweitsprache	.08*	-.01	-.02	-.12**	-.12**	-.09*
Alter in Monaten	-.09*	-.02	-.16***	-.16***	-.14***	-.13***
Soziale Herkunft (aggregiert)	-.01	-.02	-.00	-.04	.02	.04
Anteil Deutsch als Zweitsprache (aggregiert)	.05	.04	.03	-.03	-.06	-.12*
Erklärte Varianz						
R^2	15.0%	22.7 %	15.9 %	16.9 %	39.9 %	43.4 %

Anmerkungen: Standardisierte Regressionskoeffizienten unter Berücksichtigung des allfälligen Clustereffekts (Schülerinnen und Schüler in Klassen). M = Mathematik, D = Deutsch. Geschlecht: Mädchen = 0, Knaben = 1. Deutsch als Erst- resp. Zweitsprache: Erstsprache = 0, Zweitsprache = 1. Mathematik: $N=965$ (Eine Klasse ausgeschlossen: Ausreisser). Deutsch: $N=966$ (zwei Klasse ausgeschlossen: Ausreisser). Gewichtete Daten. Signifikanzniveaus: ⁺ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Bei den in Tabellen 9.2 dargestellten Koeffizienten handelte es sich um standardisierte Regressionskoeffizienten. Der Wertebereich standardisierter Regressionskoeffizienten liegt zwischen -1 und 1. Je stärker der Wert von 0 verschieden ist, desto bedeutsamer ist der Effekt der entsprechenden Variablen hinsichtlich des Merkmals, welches es zu erklären gilt. Inhaltlich indizieren Regressionskoeffizienten mit einem positiven Vorzeichen einen Zusammenhang im Sinne von «je höher, desto höher bzw. je tiefer, desto tiefer», solche mit einem negativen einen Zusammenhang im Sinne «je höher, desto tiefer bzw. je tiefer, desto höher». In Modell 2 (Tabelle 9.2) zeigt sich bspw. im Falle von Mathematik, dass die Unterrichtswahrnehmung hinsichtlich der Vorhersage der Kompetenzüberzeugung etwas weniger bedeutsam ist als das Alter der Schülerinnen und Schüler in Monaten. Zudem sind die Vorzeichen der beiden Koeffizienten verschieden. Hinsichtlich der Unterrichtswahrnehmung gilt, je positiver die Unterrichtswahrnehmung, desto höher die Kompetenzüberzeugung; «je höher, desto

höher». Hinsichtlich des Alters in Monaten gilt: Je älter die Schülerinnen und Schüler, desto geringer ist die Kompetenzüberzeugung; «je höher, desto tiefer». Als bedeutsamste Prädiktoren der Kompetenzüberzeugung erweisen sich das Geschlecht und die soziale Herkunft. Mädchen (0) sind von ihren Kompetenzen in Mathematik weniger überzeugt als Knaben (1). Je höher die soziale Herkunft, desto positiver ist die Kompetenzüberzeugung der Schülerinnen und Schüler.

Indirekte Effekte

Den theoretischen Annahmen entsprechend, interessierten im Falle der Unterrichtswahrnehmung auch allfällige indirekte Effekte hinsichtlich der Leistungen in Mathematik und Deutsch, vermittelt über die beiden Motivationskomponenten. Konkret handelt es sich um die Effekte Unterrichtswahrnehmung → Kompetenzüberzeugung → Leistungen sowie Unterrichtswahrnehmung → intrinsischer Anreiz → Leistungen.

In beiden Fächern liess sich jeweils der indirekte Effekt der Unterrichtswahrnehmung vermittelt über den intrinsischen Anreiz des Faches statistisch nicht absichern. In Mathematik fand sich indes ein statistisch bedeutsamer positiver indirekter Effekt, vermittelt über die Kompetenzüberzeugung auf die Leistungen (Sobel Teststatistik Mathematik: $z = 2.78$, $p < .01$). In Deutsch zeigte sich angesichts des lediglich tendenziellen positiven Zusammenhangs zwischen Unterrichtswahrnehmung und Kompetenzüberzeugung (vgl. Tabelle 9.2, Modell 2) nur ein tendenziell statistisch bedeutsamer indirekter Effekt (Sobel Teststatistik: $z = 1.85$, $p < .10$), der jedoch in die gleiche Richtung weist wie in Mathematik. Eine positivere Unterrichtswahrnehmung zeitigt (tendenziell) – vermittelt über höhere Kompetenzüberzeugung – günstige leistungsmässige Effekte.

Literatur

- Allison, P. D. (2001). *Missing data*. Thousand Oaks, CA: Sage.
- Aunola, K., Nurmi, J.-E., Lerkkanen, M.-K. & Rasku-Puttonen, H. (2003). The role of achievement-related behaviours and parental beliefs in children's mathematical performance. *Educational Psychology, 23*, 403-421.
- Bandura, A. (2003). *Self-efficacy: The exercise of control* (7th ed.). New York: Freeman and Company.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiss, M. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K. J. Tillmann & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–407). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit: Vertiefende Analysen im Rahmen von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften | GWV Fachverlage GmbH Wiesbaden.
- Bayer N. & Moser, U. (2009). Wirkungen unterschiedlicher Modelle der Schuleingangsstufe auf den Lern- und Entwicklungsstand: Erste Ergebnisse einer Längsschnittstudie. *Zeitschrift für Grundschulforschung. Bildung im Elementar- und Primarbereich, 2* (1).
- Boudon, R. (1974). *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. New York: Wiley.
- Bleeker, M. M. & Jacobs, J. E. (2004). Achievement in Math and Science: Do mothers' beliefs matter 12 years later? *Journal of Educational Psychology, 96*, 97-109.
- Boekaerts, M. (2001). Context sensitivity: Activated motivational beliefs, current concerns and emotional arousal. In S. Volet & S. Järvelä (Eds.), *Motivation in learning contexts: Theoretical advances and methodological implications* (p. 17–31). Amsterdam: Pergamon.
- Bong, M. & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Reviews, 15*, 1-40.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.

- Buff, A. (1991a). *Persönlichkeitsentwicklung im Umfeld des Übertritts in die Sekundarstufe I*. Unveröffentlichte Dissertation, Universität Zürich.
- Buff, A. (1991b). Schulische Selektion und Selbstkonzeptentwicklung. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung* (S. 100-114). Stuttgart: Enke.
- Buff, A. (2008a). Kann ich es? Will ich es und warum? In Moser, U. & Hollenweger, J. (Hrsg.), *Drei Jahre danach* (S. 89-113). Oberentfelden: Sauerländer.
- Buff, A. (2008b). Schulisches «Angebot» und individueller «Ertrag». In Moser, U. & Hollenweger, J. (Hrsg.), *Drei Jahre danach* (S. 205-232). Oberentfelden: Sauerländer.
- Buff, A., Reusser, K. & Pauli, C. (2010a). Selbstvertrauen ist wichtig, aber nicht ausreichend - Die Bedeutung von Unterricht, Selbstvertrauen, Qualität der Lernmotivation für Engagement und Leistung im Fach Mathematik. In K. Reusser, C. Pauli & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität - Ergebnisse einer internationalen und schweizerischen Video-Studie zum Mathematikunterricht* (S. 279-308). Münster: Waxmann.
- Buff, A., Reusser, K. & Pauli, C. (2010b). Die Qualität der Lernmotivation in Mathematik auf der Basis freier Äusserungen: Welches Bild präsentiert sich bei Deutschschweizer Schülerinnen und Schülern im 8. und 9. Schuljahr. In K. Reusser, C. Pauli & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität - Ergebnisse einer internationalen und schweizerischen Video-Studie zum Mathematikunterricht* (S. 253-278). Münster: Waxmann.
- Caliendo, M. & Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22 (1), 31-72.
- Cloerkes, G. (Hrsg.). (2003). *Wie man behindert wird. Texte zur Konstruktion einer sozialen Rolle und zur Lebenssituation betroffener Menschen*. Heidelberg: Universitätsverlag Winter.
- Dai, Y. D. (2002). Incorporating parental perceptions: A replication and extension study of the internal-external frame of reference model of self-concept development. *Journal of Adolescent Research*, 17, 617-645.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19, 294-304.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39, 223•238.
- Ditton, H. (Hrsg.). (2007). *Kompetenzaufbau und Laufbahnen im Schulsystem. Eine Längsschnittuntersuchung an Grundschulen*. Münster: Waxmann.

- Eccles, J. S. (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105-121). New York: Guilford Press.
- Eccles, J. S. (2007). Families, schools, and developing achievement-related motivation and engagement. In J. E. Grusec & P. D. Hastings (Eds.), *Handbook of socialization: theory and Research* (pp. 665-691). New York, NY and London: The Guilford Press.
- Eccles, J. S., Barber, B. L. Updegraff, K., & O'Brien, K. M. (1998). An expectancy-value model of achievement choices: The role of ability self-concepts, perceived task utility and Interest in predicting activity choice and course enrollment. In L. Hoffmann, A. Krapp, K. A. Renninger & J. Baumert (Eds.), *Interest and learning* (pp. 267-279). Kiel: Leibnitz-Institut für die Pädagogik der Naturwissenschaften (IPN).
- EDK (2010). <http://www.edk.ch/dyn/15780.php>
- Fend, H. (1998). *Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung*. Weinheim: Juventa.
- Fredricks, J. A. & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, 4, 519-533.
- Frome, P. M. & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, 74, 435-452.
- Gonida, E. N. & Urdan, T. (2007). Parental influences on student motivation, affect and academic behaviour: Introduction to the special issue. *European Journal of Psychology of Education*, 22, 3-6.
- Helmke, A. (1997). Entwicklung lern- und leistungsbezogener Motive und Einstellungen: Ergebnisse aus dem SCHOLASTIK-Projekt. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 59-76). Weinheim: PVU.
- Helmke, A. (2003). *Unterrichtsqualität erfassen, bewerten, verbessern*. Seelze: Kallmeyer.
- Helmke, A., Hosenfeld, I., Schrader, F.-W. & Wagner, W. (2002). Unterricht aus der Sicht der Beteiligten. In A. Helmke & R. S. Jäger (Hrsg.), *Das Projekt MARKUS. Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext* (S. 325-411). Landau: Empirische Pädagogik.
- Helmke, A., Schrader, F.-W. & Lehneis-Klepper, G. (1991). Zur Rolle des Elternverhaltens für die Schulleistungsentwicklung ihrer Kinder. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 13, 1-22.
- Helmke, A. & Weinert, F. E. (1997a). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, D, Serie Pädagogische Psychologie, Bd. 3; S. 71-176). Göttingen: Hogrefe.

- Helmke, A. & Weinert, F. E. (1997b). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. W. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241-251). Weinheim: Beltz.
- Hirsjärvi, S. & Perälä-Littunen, S. (2001). Parental beliefs and their role in child-rearing. *European Journal of Psychology of Education, 16*, 87-116.
- HLM 6.03 (2004). [Computer Programm]. Lincolnwood, IL: Scientific Software International Inc.
- Koestner, R. & Losier, G. F. (2002). Distinguishing three ways of being highly motivated: A closer look at introjection, identification, and intrinsic motivation. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 101-121). Rochester, NY: The University of Rochester Press.
- Jacobs, J. E., Lanza; S., Osgood, D. W., Eccles, J. S. & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*, 509-527.
- Jodl, K. M., Michael, A., Malanchuk, O., Eccles, J. S. & Sameroff, A. (2001). Parents' role in shaping early adolescents' occupational aspirations. *Child Development, 72*, 1247-1265.
- Keller, F. & Moser, U. (2008a). Die Untersuchung im Überblick. In U. Moser & J. Hollenweger (Hrsg.), *Drei Jahre danach*. (S. 13-40). Oberentfelden: Sauerländer.
- Keller, F. & Moser, U. (2008b). Fachleistungen am Ende der 3. Klasse. In U. Moser & J. Hollenweger (Hrsg.), *Drei Jahre danach*. (S. 41-87). Oberentfelden: Sauerländer.
- Koestner, R. & Losier, G. F. (2002). Distinguishing three ways of being highly motivated: A closer look at introjection, identification, and intrinsic motivation. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 101-121). Rochester, NY: The University of Rochester Press.
- Krapp, A. (2003). Die Bedeutung der Lernmotivation für die Optimierung des schulischen Bildungssystems. *Politische Studien, 3*, 91-105.
- Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instructions, 15*, 381-395.
- Kronig, W. (2007). *Die systematische Zufälligkeit des Bildungserfolgs. Theoretische Erklärungen und empirische Untersuchungen zur Lernentwicklung und zur Leistungsbewertung in unterschiedlichen Schulklassen*. Bern, Stuttgart, Wien: Haupt Verlag.
- Kronig, W., Eckhart, M. & Haeblerlin, U. (2007). *Immigrantenkinder und schulische Selektion: pädagogische Visionen, theoretische Erklärungen und empirische Untersuchungen zur Wirkung integrierender und separierender Schulformen in den Grundschuljahren*. Bern, Stuttgart, Wien: Haupt Verlag AG.

- Kunter, M., Schümer, G., Artlet, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiss, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Lepper, M. R. & Henderlong, J. (2000). Turning «play» into «work» and «work» into «play»: 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 257-307). San Diego: Academic Press.
- Lorenz, F. & Wild, E. (2007). Parental involvement in schooling. Results concerning its structure and impact on students' motivation. In M. Prenzel (Ed.), *Studies on the educational quality of schools. The final report on the DFG Priority Programme* (pp. 299-316). Münster u.a.: Waxmann.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung: Probleme und Lösungen. *Psychologische Rundschau*, 58, 103-117.
- Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.
- Maehr, M. L. (2004). Paul Pintrich: A once and continuing influence. *Educational Psychologist*, 40, 129–133.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2004). *Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering*. Paper presented at the 3rd International Biennial SELF Research Conference, in Berlin, Germany – July 4-7, 2004.
- Maaz, K., Baumert, J. & Trautwein, U. (2009). Genese sozialer Ungleichheiten im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen*. (Zeitschrift für Erziehungswissenschaften, Sonderheft 12, S. 11–46). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maaz, K. & Nagy, G. (2009). Primäre und sekundäre Herkunftseffekte. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaften, Sonderheft 12, S. 153–182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Moser, U., Berweger, S. & Stamm, M. (2005a). Lesekompetenzen bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz, Mathematik und soziale Kompetenzen beim Schuleintritt* (S. 37-58). Oberentfelden: Sauerländer.
- Moser, U., Berweger, S. & Stamm, M. (2005b). Wortschatz bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz, Mathematik und soziale Kompetenzen beim Schuleintritt* (S. 59-76). Oberentfelden: Sauerländer.

- Moser, U., Berweger, S. & Stamm, M. (2005c). Mathematische Kompetenz bei Schuleintritt. In U. Moser, M. Stamm & J. Hollenweger (Hrsg.), *Für die Schule bereit? Lesen, Wortschatz, Mathematik und soziale Kompetenzen beim Schuleintritt* (S. 77-98). Oberentfelden: Sauerländer.
- Moser, U. Keller F. & Zimmermann P. (2008). Soziale Ungleichheiten und Fachleistungen. In U. Moser & J. Hollenweger (Hrsg.), *Drei Jahre danach. Lesen, Wortschatz, Mathematik und soziale Kompetenzen am Ende der dritten Klasse* (S. 115–151). Oberentfelden: Sauerländer Verlage AG.
- Moser, U., Ramseier, E., Keller C. & Huber, M. (1997). *Schule auf dem Prüfstand. Eine Evaluation der Sekundarstufe I auf der Grundlage der «Third International Mathematics and Science Study»*. Chur: Rüegger.
- Mplus 5.2 (2007). [Computer Programm]. Los Angeles, CA: Muthén & Muthén.
- Muthén, L. K. & Muthén, B. O. (2007). *Mplus•User's guide*. Los Angeles, CA: Muthén & Muthén.
- NORM 2.03 (2000). [Computer Programm]. Verfügbar unter <http://www.multiple-imputation.com/> [6.10.2009].
- Pauli, C. & Reusser, K. (2006). Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und –entwicklung. *Zeitschrift für Pädagogik*, 52, 774-798.
- Pekrun, R. (1998). Schüleremotionen und ihre Förderung: Ein blinder Fleck der Unterrichtsforschung. *Psychologie in Erziehung und Unterricht*, 44, 230-248.
- Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (ed.), *Motivational psychology of human development* (pp. 143-163). Amsterdam: Elsevier.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315-341.
- Pekrun, R., Frenzel, A.C., Goetz, T., & Perry, R.P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P.A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 13-36). Amsterdam: Elsevier.
- Pekrun, R., Goetz, T. & Perry, R. P. (2005). *Academic Emotions Questionnaire (AEQ) - User's Manual*. Unpubl. manuscript, Universities of Munich (Germany) and Manitoba (Canada), Departments of Psychology.
- Pekrun, R., Goetz, T., Titz, W., & Perry, R.P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37, 91-105.

- Pekrun, R., & Hofmann (1999). Lern- und Leistungsemotionen: Erste Befunde eines Forschungsprogramms. Emotions in learning and achievement: First results of a program of research. In M. Jerusalem & R. Pekrun (Hrsg.), *Emotion, Motivation und Leistung* (S. 248-267). Göttingen: Hogrefe.
- Pekrun, R. & Schiefele, U. (1996). Emotions- und motivationspsychologische Bedingungen der Lernleistung. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion* (Enzyklopädie der Psychologie, D, Serie Pädagogische Psychologie, Bd. 2; S. 153-180). Göttingen: Hogrefe.
- Pekrun, R., & Schutz, P. A. (2007). Where do we go from here? Implications and future directions for inquiry on emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 313-331). Amsterdam: Elsevier.
- Phillips, D. A. (1987). Socialization of perceived academic competence among highly competent children. *Child Development*, *58*, 1308-1320.
- Pietsch, J., Walker, R. & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, *3*, 589-603.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching. *Journal of Educational Psychology*, *95*, 667-686.
- Pintrich, P.R. & Schunk, D.H. (2002). *Motivation in education: Theory, research and application* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- PISA-Konsortium Deutschland (Hrsg.). (2004). *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann.
- Prenzel, M. & Drechsel, B. (1996). Ein Jahr kaufmännische Erstausbildung: Veränderungen in Lernmotivation und Interesse. *Unterrichtswissenschaft*, *24*, 217-234.
- Rakoczy, K., Buff, A. & Lipowsky, F. (2005). Befragungsinstrumente. In E. Klieme, C. Pauli, & K. Reusser (Hrsg.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie «Unterrichtsqualität, Lernverhalten und mathematisches Verständnis»*. Materialien zur Bildungsforschung, Band 13. Frankfurt am Main: GPF.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, London, New Delhi: Sage Publications.
- Rhyn, H. & Moser, U. (2002). Schülerinnen und Schüler beurteilen. In H. Rhyn (Hrsg.), *Beurteilung macht Schule* (S. 25–34). Bern: Haupt Verlag AG.
- Rost, J. (2004). *Lehrbuch Testtheorie, Testkonstruktion*. Bern u.a.: Hans Huber.
- Rubin, D. B. & Thomas, N. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrics*, *70*(1), 41-55.

- Ryan, R. M. & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3-33). Rochester, NY: The University of Rochester Press.
- Satow, L. (1999). *Klassenklima und Selbstwirksamkeitsentwicklung: Eine Längsschnittstudie in der Sekundarstufe I*. Dissertation, Freie Universität Berlin.
- Satow, L. (2002). Unterrichtsklima und Selbstwirksamkeitsdynamik. In M. Jerusalem, & D. Hopf (Hrsg.), *Zeitschrift für Pädagogik. Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen. 44. Beiheft* (S. 174-191).
- Satow, L. & Schwarzer, R. (2003). Entwicklung schulischer und sozialer Selbstwirksamkeitserwartung. *Psychologie in Erziehung und Unterricht, 50*, 168-181.
- Schiefele, U. & Pekrun, R. (1996). Psychologische Modelle des fremdgesteuerten und selbstgesteuerten Lernens. In F. E. Weinert (Hrsg.), *Psychologie des Lernens und der Instruktion* (Enzyklopädie der Psychologie, D, Serie Pädagogische Psychologie, Bd. 2; S. 249-278). Göttingen: Hogrefe.
- Schutz, P. A., Hong, J. Y., Cross, D. I., & Osbon, J. N. (2006). Reflections on investigating emotion in educational activity settings. *Educational Psychology Review, 18*, 343-360.
- Schutz, P. A. & Pekrun, R. (2007). Introduction to emotion in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 3-10). Amsterdam: Elsevier.
- Skalenhandbuch DESI (T1+T2) – Schülerfragebögen* (2005, internes Arbeitspapier Stand 09.06.05). Universität Dortmund: Institut für Schulentwicklung.
- Skinner, E. (1996). A guide to constructs of control. *Journal of Personality and Social Psychology, 3*, 549-570.
- Spinath, B. & Spinath, F. (2005). Longitudinal analysis of the link between learning motivation and competence beliefs among elementary school children. *Learning and Instruction, 15*, 87-102.
- SPSS 15 (2006). [Computer Programm]. Chicago, IL: SPSS Inc.
- Stalder, B. E., Meyer T. & Hupka-Brunner S. (2008). Leistungsschwach - Bildungsarm? PISA-Kompetenzen als Prädiktoren für nachobligatorische Bildungschancen. *Die Deutsche Schule, 100*(4), 436-448.
- Tabachnik, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*. Boston: Pearson/Allyn and Bacon.
- Taylor, A. B., MacKinnon, D. P., & Tein, J.-Y. (2008). Tests of the three-path mediated effect. *Organizational Research Methods, 11*, 241-269.
- Tillmann, K. J. (2002). Die doppelte Benachteiligung. *Erziehung & Wissenschaft, 9*, 6-10.

- Trautwein, U., Lüdtke, O., Schnyder, I. & Niggli, A. (2006). Predicting homework effort: Support for a domain-specific, multilevel homework model. *Journal of Educational Psychology, 98*, 438-456.
- Wagner, J. W. (2001). *Leistungsvergleiche in der Schule. Bezugsgruppeneffekte und Fähigkeitsselbstkonzept*. Landau: Verlag Empirische Pädagogik.
- Wagner, S. & Powell, J. (2003). Ethnisch-kulturelle Ungleichheit im deutschen Bildungssystem. Zur Überrepräsentanz von Migrantenjugendlichen an Sonderschulen. In G. Cloerkes (Hrsg.), *Wie man behindert wird. Texte zur Konstruktion einer sozialen Rolle und zur Lebenssituation betroffener Menschen* (S. 183–208). Heidelberg: Universitätsverlag Winter.
- Weiss, R. & Osterland, J. (1997). *Grundintelligenztest Skala 1 – CFT*. Göttingen: Hogrefe.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81.
- Wigfield, A. & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91-120). San Diego, CA: Academic Press.
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arberton, A., Freedman-Doan, C. & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology, 89*, 451-469.