

Pilotierungsuntersuchung 2003 zu Orientierungsarbeiten in Jahrgangsstufe 2

**Bericht über die Ergebnisse
der Pilotierung der Testaufgaben
für die Jahrgangsstufe 2
in den Ländern Brandenburg und Berlin**

Prof. Dr. Hartmut Ditton
Dr. Jan Krüsken
Jutta Reich, M.A.

Lehrstuhl für Allgemeine Pädagogik
Erziehungs- und Sozialisationsforschung
Ludwig-Maximilians-Universität München



INHALTSVERZEICHNIS

1. Einleitung	2
2. Stichprobe und Durchführung	2
3. Aufgabenanalysen	4
3.1 Leseverständnis	5
3.2 Leiselesen (Lesegeschwindigkeit).....	7
3.3 Rechtschreibung	8
3.4 Mathematik.....	9
3.5 Zusammenhänge zwischen den Testteilen	11
4. Hintergrundmerkmale	12
4.1 Schülerfragebogen.....	12
4.2 Elternfragebögen	14
4.3 Zusammenhänge zwischen Hintergrundmerkmalen und Testleistungen.....	17

1. Einleitung

Zum Ende des Schuljahres 2002/03 beteiligten sich die Länder Berlin und Brandenburg an einer Pilotierungsuntersuchung des Landes Bayern zur Entwicklung von Orientierungsarbeiten in der 2. Jahrgangsstufe der Grundschule. Die Pilotierung der entwickelten Aufgaben wurde in Berlin und Brandenburg an zwei eigenen Schulstichproben von insgesamt 49 Schulklassen aus beiden Ländern vorgenommen. Das Ziel aller drei Länder war es, neue Aufgabensätze auf ihre Eignung für den landesweiten Einsatz bei Orientierungsarbeiten im Schuljahr 2003/04 zu überprüfen. Im vorliegenden Bericht werden die zentralen Ergebnisse der Erhebung in Berlin und Brandenburg dargestellt.

Die getesteten Aufgabensätze für die Jahrgangsstufe 2 wurden vom Institut für Schulpädagogik und Bildungsforschung in München (ISB) für den Einsatz in Bayern 2004 unter wissenschaftlicher Begleitung des Instituts für Pädagogik der Universität München entwickelt. Sie decken die Bereiche Lesen, Rechtschreibung und Mathematik ab. Neben den Leistungsdaten wurden auch Kontextvariablen über Schüler- und Elternfragebögen erhoben.

Die Stichprobenziehung, Vorbereitung und Durchführung der Erhebung erfolgte durch die zuständigen Referate der Kultusministerien beider Länder. Dies waren in Brandenburg das Referat für schulische Qualitätsentwicklung, Schulforschung und pädagogische Schulentwicklung des Ministeriums für Bildung, Jugend und Sport unter Leitung von Herrn Hans-Jürgen Kuhn und in Berlin das Referat Schulreform und Qualitätssicherung unter Leitung von Herrn Tom Stryck. Schließlich erfolgte die Dateneingabe und -auswertung der Universität München durch eine Arbeitsgruppe am Lehrstuhl für allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung unter Leitung von Prof. Hartmut Ditton.

2. Stichprobe und Durchführung

In Brandenburg wurden 25 Schulen und in Berlin 23 Schulen für die Teilnahme an der Untersuchung ausgewählt; jede Schule nahm mit je einer Klasse der 2. Jahrgangsstufe an der Erhebung teil; eine Berliner Grundschule war mit zwei Klassen vertreten. Die Auswahl erfolgte durch die Kultusministerien der beiden Länder. In beiden Stichproben war die regionale Verteilung der Schulen ein wichtiges Auswahlkriterium; in Berlin wurde darüber hinaus auf eine breite Streuung des Schulmerkmals „Anteil von nichtdeutschen Muttersprachlern“ geachtet; in Brandenburg spielten „Schulgröße“ und „Urbanität“ in der Stichprobe eine größere Rolle. Die Schulen nahmen freiwillig teil und wurden im Vorfeld eingehend über Ziel und Durchführung der Untersuchung informiert.

Die Erhebung wurde in beiden Ländern von geschulten Testleitern durchgeführt, die an zwei Tagen mit den Testmaterialien die ausgewählten Klassen aufsuchten und die Testsitzungen in Beisein der Klassenlehrkräfte durchführten. In der nachfolgenden Übersicht sind die Erhebungsbestandteile und Bearbeitungszeiten an beiden Testtagen in Berlin und Brandenburg dargestellt.

1. Testtag , Aufgabenabfolgen und Zeitangaben in Minuten

<i>Pilotierungsaufgaben:</i>	Gesamt	Instruktion	Bearbeitung
	25	5	20
	10	5-6	4
	30	5	25
	25	5	20
	120	20	69

2. Testtag, Aufgabenabfolgen und Zeitangaben in Minuten

Mathematikaufgaben	45	5	40
Pause	15		
Lesetext 2	25	5	20
Gesamtzeit 2. Testtag	85	10	60

Abbildung 1: Übersicht zum Erhebungsablauf an beiden Testtagen in Brandenburg und Berlin

Die einzelnen Erhebungsbestandteile wurden dabei in der gleichen Art und Weise eingesetzt, wie dies auch in einer möglichen Vollerhebung geschehen sollte: Nach dem Austeilen der Testhefte und einer kurzen Instruktion begann die Bearbeitungszeit für die Klasse. Die Instruktion bestand beim Leseverständnis nur aus einer kurzen Erläuterung zu den Multiple-Choice-Aufgaben („Nur eine Antwort ist richtig“). Für den Testteil „Leiselesen“ (Lesegeschwindigkeit) erfolgte eine Erläuterung der Aufgabe und des Ankreuzmodus. Antworten auf inhaltliche Fragen waren nicht zugelassen. Um die Arbeitsruhe nicht zu stören, konnten Schülerinnen und Schüler, die bereits frühzeitig mit der Testbearbeitung abschlossen, ihre Testhefte vor Ende der Bearbeitungszeit abgeben. Die Ausnahme bildete der Lesegeschwindigkeitstest, der innerhalb eines konstanten Zeitraumes von vier Minuten von allen Schülern zu bearbeiten war.

Nach der Erhebung wurden die bearbeiteten Testmaterialien in den Kultusministerien gesammelt und von dort an das Institut für Pädagogik in München zur Codierung und Auswertung weitergeleitet. Die Elternfragebögen wurden an den Schulen gesammelt und von diesen ebenfalls zur Auswertung nach München weitergeleitet.

Die Stichprobe wird in Tabelle 1 dargestellt. Insgesamt befanden sich in den ausgewählten 49 Klassen 1104 Schüler; davon waren 49 % Mädchen und 51 % Jungen. Die Testteilnahme an beiden Tagen fällt etwas unterschiedlich aus, wobei die Differenzen allerdings gering bleiben (1. Testtag 96 %, 2. Testtag 95 %). Sowohl am 1. als auch am 2. Testtag fehlten nur knapp 3%

der Schüler¹. Der Altersdurchschnitt der Schüler zum Erhebungszeitpunkt lag in der Gesamtstichprobe bei 9,09 Jahren.

Tab. 1: Stichprobenbeschreibung der Pilotierungsuntersuchung nach Bundesländern

	Brandenburg	Berlin
Klassen	25	24
Klassengröße	$M = 21.48, SD = 4.39$ $Min. = 13; Max. = 32$	$M = 23.62, SD = 3.18$ $Min. = 18; Max = 29$
Gesamtstichprobenzahl	537	567
davon Jungen	266 (50 %)	298 (53 %)
davon Mädchen	271 (50 %)	269 (47 %)
Nicht-Teilnahme am 1. Testtag	16 (3 %)	22 (4 %)
Nicht-Teilnahme am 2. Testtag	23 (4 %)	24 (4 %)
Abwesend an beiden Testtagen	13 (2.6 %)	15 (2.8 %)
Muttersprache nicht Deutsch	24 (4.5 %)	124 (22 %)
Anerkannte Legastheniker	11 (2 %)	8 (1.5%)
Klassenwiederholung	26 (5 %)	22 (4 %)
Einschulungszeitpunkt		
früh	41 (8 %)	32 (6 %)
zurückgestellt	31 (6 %)	24 (4 %)
Alter der Schüler in Jahren im Durchschnitt zum Erhebungszeitpunkt	$M = 9.06 (SD = 0.77)$	$M = 9.12 (SD = 0.78)$

Angaben der Klassenliste zufolge wurden 88 % der teilnehmenden Schüler regulär eingeschult (Angabe „zurückgestellt“ 5 %, Angabe „früh eingeschult“ 7 %). Etwa 4 % der Stichprobe hatten eine Klasse wiederholt. Der Anteil Schüler mit dem Eintrag „anerkannte LRS / Legasthenie“ in der Klassenliste lag bei etwa 2 %. Alle vorstehenden Werte sind in beiden Länderstichproben etwa gleich hoch.

Die Angabe „Deutsch ist nicht Muttersprache“ lag in der Gesamtstichprobe bei 13 %. Aus der Tabelle wird ersichtlich, wie stark die Verteilung der nicht-muttersprachlichen Schüler in den beiden Länderstichproben differiert. Während in der Berliner Stichprobe 22 % der teilnehmenden Schüler keine deutschen Muttersprachler sind, liegt dieser Anteil in der Brandenburger Stichprobe bei nur knapp 5 %. Darüber hinaus waren die ausgewählten Klassen der Berliner Stichprobe im Durchschnitt etwas größer als die teilnehmenden Klassen in Brandenburg.

3. Aufgabenanalysen

Die Dateneingabe einschließlich Kodierung der offenen Angaben erfolgte am Lehrstuhl für allgemeine Pädagogik, Erziehungs- und Sozialisationsforschung der Ludwig-Maximilians-Universität München. Die meisten Aufgabenformate erlaubten eine direkte Eingabe der Antworten (Geschlossene Frageformate, Lesegeschwindigkeitsaufgaben sowie die Mehrzahl der Mathematik- und Rechtschreibaufgaben).

¹ Die Stichprobengröße je Untersuchungsinstrument schwankt etwas, da an den Testtagen einige Schüler während der Testungen ausfielen. Die Stichprobengröße je Instrument wird bei der Darstellung der Auswertung der jeweiligen Erhebungsinstrumente dargestellt.

Für die Kodierung der offenen Fragen wurden vom Institut für Schulpädagogik und Bildungsforschung (ISB) Kodieranweisungen erarbeitet. Die Schülerantworten zu diesen Formaten wurden entsprechend der Kodieranweisungen kodiert und eingegeben. Anhand einer Substichprobe von 100 Schülerinnen und Schülern je Land wurde die Übereinstimmung zwischen je zwei unabhängigen Kodierern bei den offenen Fragen (zwei Fragen beim Leseverständnis, eine Frage im Mathematiktest) überprüft. Die Übereinstimmung lag im Schnitt bei etwa 95 Prozent und ist damit zufriedenstellend.

Im ersten Analyseschritt wurde für jede Aufgabe ihre Schwierigkeit (Itemmittelwert) bestimmt. Im Falle mehrkategorialer Items (bspw. Mit drei Abstufungen 0, 1, 2) wurden zusätzlich die Häufigkeiten für jede Kategorie ermittelt. Auf diese Weise konnten sofort zu schwere Items aussortiert werden. Zu schwer bedeutet, dass weniger als 15 Prozent der Schüler die Aufgabe erfolgreich gelöst hatten. Diese Items sind wenig zur angemessenen Differenzierung innerhalb der vorliegenden Stichprobe geeignet, Gleiches gilt für extrem leichte Items.

Der restliche Pool von Aufgaben ging in die Analysen zur Skalierung auf der Basis probabilistischer Testmodelle (Item Response Modelle) ein. Sie wurden darauf überprüft, ob sie „modellkonforme“ Eigenschaften im Sinne des gewählten Auswertungsmodell (Raschmodell) haben. Aufgaben, die aufgrund mathematisch-statistischer Kriterien schlechte Konformitätswerte aufwiesen, wurden aus dem Itempool ausgeschlossen.

3.1 Leseverständnis

3.1.1 Inhaltliche Anforderungen und Itemschwierigkeiten

Zur Erfassung des Leseverständnisses wurden zwei verschiedene Texte (ein Sachtext und eine Geschichte) ausgewählt, die in zwei verschiedenen Testsitzungen zum Einsatz kamen. Die Texte waren recht kurz (Umfang etwa 250 Wörter). Zu jedem Text wurden 12 Fragen von den Arbeitskreisen des ISB entwickelt. Alle Fragen waren einfach codiert (gelöst = „1“). Von 1.022 Schülern wurden beide Tests bearbeitet.

Die Fragen zum Leseverständnis beziehen sich auf zwei Prozesse: (1) Einerseits wird die Entnahme von im Text mehr oder minder **explizit genannter Information** verlangt, andererseits geht es um die **Nutzung von textlicher Information zu Schlussfolgerungen** über den Text bzw. darum, **implizite Inhalte aus dem Text erschließen**. (2) Der zweite Prozess baut auf den erstgenannten auf; die Komplexität der verlangten Schlüsse reicht vom **Erschließen nicht wörtlich genannter Inhalte** bis zu **einfachen reflektierenden Betrachtungen des Gelesenen**.

Die den Fragen zugrundeliegenden Prozesse sollten an das in IGLU formulierte Kompetenzmodell anschlussfähig sein – mit der Einschränkung, dass mit den vorliegenden Fragen stärker der Aspekt des „textbasierten Verstehens“ gegenüber dem „wissensbasierten Verstehen“ erfasst wird. Dies ergibt sich einerseits aufgrund der Kürze der verwendeten Texte, die entsprechend weniger Raum für Reflexionen und textübergreifende Analysen erlauben, andererseits aber auch aus der ausgewählten Jahrgangsstufe. Bei den „beginnenden Lesern“ der Klasse 2 interessiert auch eine genauere Betrachtung der basalen Kompetenzen zur Informationsentnahme und der Verwendung der Informationen im Bereich des textbasierten Verstehens.

Beide Texte erweisen sich als geeignet, das Leseverständnis adäquat zu erfassen; allerdings sind die Texte mit insgesamt nur 250 Wörtern je Text relativ kurz und die Fragen erscheinen alles in allem zu leicht. So wurden bei den Texten jeweils vier Fragen ausgeschlossen.

In Tabelle 2 werden die verbleibenden Fragen für beide Texte zusammengestellt. Über beide pilotierten Texte hinweg ist erkennbar, dass die Aufgaben (Items) zum Bereich „Implizite Inhalte erschließen“ erwartungsgemäß von weniger Schülern gelöst werden als Fragen zum Bereich „Explizite Informationsentnahme“, so dass sich auch beim Einsatz der einzelnen Texte vermutlich Schülergruppen auf diesen beiden unterschiedlichen Kompetenzniveaus identifizieren lassen.

Tab. 2: Mittlere Schwierigkeiten der geeigneten Fragen nach erfasstem Prozess und Text

		Explizite Informationen finden	Implizite Information erschließen / Schlussfolgerungen	Summe
Text I	Anzahl Aufgaben	4	4	8
	mittlere Lösungshäufigkeit	71 %	47 %	59 %
Text II	Anzahl Aufgaben	3	5	8
	mittlere Lösungshäufigkeit	60 %	53 %	56 %
Beide Texte	Anzahl Aufgaben	7	9	16
	mittlere Lösungshäufigkeit	66 %	50 %	57 %

3.1.2 Empfehlungen zur Erfassung des Leseverständnisses

Für einen Einsatz in der Vollerhebung der Orientierungsarbeiten wird im Bereich Leseverständnis empfohlen, einen der beiden Texte auszuwählen und um einige Fragen zu ergänzen, so dass in etwa zwischen 10 und 12 Fragen je Text resultieren. Diese Menge scheint in der vorgesehenen Arbeitszeit von 20 bis 25 Minuten gut zu bewältigen. Da die geeigneten Fragen insgesamt recht leicht sind, empfiehlt es sich weiterhin, die neuen Fragen zum Bereich „implizite Textinhalte erschließen / Schlussfolgern über den Text“ etwas anspruchsvoller bzw. schwieriger zu gestalten. Insgesamt betrachtet erweist sich die Mehrzahl der getesteten Aufgaben als geeignet für einen flächendeckenden Einsatz bei Orientierungsarbeiten in der Jahrgangsstufe 2 in beiden Ländern. Für diejenigen Aufgaben, deren Format sich als problematisch erwies, konnten aufgrund der Untersuchung Hinweise zur Verbesserung gewonnen werden.

3.1.3 Ausblick: Leistungsgruppen der Schülerinnen und Schüler

Im Anschluss an die Prüfung der Aufgaben auf ihre Eignung wurden für die Schüler Leistungswerte über beide eingesetzten Texte hinweg bestimmt. Zugrundegelegt wurden die erzielte Punktsomme über die ausgewählten Fragen der beiden Texte. Der gebildete Testwert für das Leseverständnis wurde in den weiteren Analysen verwendet (s.u. Abschnitt 3.5 ff.). Weitere Aufgabenanalysen zeigten, dass sich drei Leistungsgruppen im Bereich Leseverständnis identifizieren lassen. Die Gruppen können anhand der jeweils bewältigten Anforderungen näher beschrieben werden:

(1) **Schwache Leser:** Diese Schüler hatten Probleme, den Texten Informationen zu entnehmen: Wörtliche Informationen, die nicht an herausragenden Textstellen positioniert waren (bspw. am Textanfang oder Beginn eines Absatzes) konnten sie nicht zuverlässig auffinden. Wurde die Information nicht wörtlich, sondern sinngemäß abgefragt, oder war diese über mehrere Textstellen verteilt enthalten, so scheiterten diese Schüler. Dementsprechend konnten sie auch nicht die Fragen beantworten, in denen es um die Nutzung der Informationen des Textes ging, d.h. sie scheiterten an den eigentlichen „Verständnisfragen“ zum Text.

(2) **Durchschnittliche Leser:** Diesen Schülern gelang die Entnahme von Informationen aus den Texten gut; sie konnten auch die meisten der nicht explizit dargestellten Inhalte aus kurzen Texten erschließen und einfachere weiterführende Fragen zum Text (einfache Schlussfolgerungen) beantworten.

(3) **Starke Leser:** Diesen Schülern gelang es, die meisten Aufgaben zu den Texten zu lösen, d.h. nicht nur die Informationsentnahme, sondern auch die Beantwortung von weiterführenden Fragen gelang ihnen gut. Damit konnten diese Schüler die textliche Information relativ vollständig nutzen und auch einige, gemessen an der Jahrgangsstufe, schwierigere Schlussfolgerungen ziehen.

3.2 Leiselesen (Lese- bzw. Decodiergeschwindigkeit)

Die Erfassung der Lese- bzw. Decodiergeschwindigkeit stellt neben der Untersuchung des Leseverständnisses eine wichtige Information zur Beurteilung der allgemeinen Leseleistung dar. In der Pilotierungsuntersuchung wurde die Lese- bzw. Decodiergeschwindigkeit der Schüler mit einem sogenannten „Leiselesetest“ erhoben. Dabei wird untersucht, wie viele Wörter innerhalb eines festgesetzten Zeitraums (in der Pilotierungsuntersuchung 4 Minuten) gelesen und den passenden Bildern aus einer Reihe von vorgegebenen Bildern zugeordnet werden können. Die Bilderreihen bestehen dabei aus „lautverwandten“ Wörtern, so dass die richtige Zuordnung nur bei korrekter Decodierung möglich ist. Das leise Lesen entspricht der natürlichen Lesesituation und ermöglicht somit – im Gegensatz zum lauten Vortragen – Aussagen über alltags- und schulalltagsrelevante Lesefertigkeiten. Der Vergleich der Geschwindigkeitswerte mit den Verständnisleistungen beim Lesen in den Klassen zeigt, ob Verständnisprobleme beim Lesen mit Problemen in der Worterkennung einhergehen oder nicht. Der Testwert besteht aus der „Summe der richtig gelesenen Wörter“, also ohne Wortauslassungen und Wortfehler in einer Zeile und kann als Indikator für die Lese- bzw. Decodiergeschwindigkeit dienen.

Durchführung und Auswertung. Für die Durchführung ist es erforderlich, dass sich die Testleiter vor der Testung mit dem Vorgehen vertraut machen, da eine Instruktionsphase notwendig ist, in der die Schüler die Bearbeitung des Tests üben müssen. Damit die Testwerte interpretierbar bleiben, muss sichergestellt werden, dass alle Schüler die Aufgabenstellung und die Art des Ankreuzens verstanden haben und dass die Zeitvorgaben eingehalten werden (Stoppuhr erforderlich). In der Pilotierungsuntersuchung wurde den Schülern das Verfahren als „Lesespiel“ vorgestellt. Die Durchführung schien den Schülern in fast allen Klassen mehr Spaß gemacht zu haben als die übrigen Testteile.

Die Auswertung ist mit Schablonen möglich; allerdings werden für den Test etwa 8 Testseiten und 8 verschiedenen Schablonen benötigt. Die Mittelwerte der Lese- bzw. Decodiergeschwindigkeitstests

sind in Tabelle 3 dargestellt. Der Mittelwert des Tests lag in beiden Ländern bei knapp 47 korrekt erkannten Worten. Als Orientierungsmaßstab für die Identifikation von extrem „schnellen“ und extrem „langsamen“ Lesern kann das Intervall von +/- einer Standardabweichung (SD) um den Mittelwert (M) dienen, in dem gut zwei Drittel der Testleistungen zu finden sind.

Tab. 3: Mittelwerte zum Lesegeschwindigkeitstest in beiden Ländern ($N = 1063$)

	Gesamtstichprobe	Brandenburg	Berlin
Auslassungsfehler (nur ein Wort in der Zeile erkannt)	$M = 0.53$ $SD = 0.95$	$M = 0.52$ $SD = 0.94$	$M = 0.53$ $SD = 0.95$
Verwechslungsfehler (falsches Wort/Bild erkannt)	$M = 5.10$ $SD = 4.63$	$M = 5.22$ $SD = 4.80$	$M = 4.98$ $SD = 4.46$
Anzahl richtig erkannter Worte (Lesegeschwindigkeitswert)	$M = 46.72$ $SD = 19.42$	$M = 46.36$ $SD = 19.01$	$M = 47.05$ $SD = 19.81$

3.3 Rechtschreibung

3.3.1 Inhaltliche Anforderungen, Bepunktung, Schwierigkeiten der Items

Die Rechtschreibleistung wurde mit einer vom Arbeitskreis Deutsch des ISB entwickelten Aufgabensammlung untersucht. Diese umfasste sieben unterschiedliche Aufgabenformate, in denen jeweils mehrere, auf die Jahrgangsstufe abgestimmte Worte unterschiedlichen Schwierigkeitsgrades entweder komplett geschrieben werden mussten, oder Formate, bei denen an kritischen Stellen Ergänzungen einzufügen waren. Insgesamt mussten in der getesteten Version gut 50 Wörter bearbeitet werden, davon war jedoch nur die eine Hälfte komplett von den Schülern selbst zu schreiben; in der anderen Hälfte des Materials waren Ergänzungen oder Ähnliches gefragt. Für die Bearbeitung standen 25 Minuten zur Verfügung; der Zeitraum erwies sich als ausreichend.

Die Aufgabenanforderungen orientierten sich zum einen an den bayrischen Lehrplaninhalten; zum anderen enthielten die Aufgaben zur Hälfte Wörter, die Bestandteil des bayrischen Grundwortschatzes (GWS) waren, und zur anderen Hälfte vom Arbeitskreis als kindgemäß eingestufte Wörter, die nicht im GWS enthalten sind. Hinter dieser Auswahl stand die Absicht, die Schwierigkeit des Wortmaterials besser kontrollieren zu können. Aufgabenanalysen zeigen, dass dies auch für die vorliegende Stichprobe in Brandenburg und Berlin gelungen ist, d.h. die durchschnittliche Schwierigkeit der eingesetzten Rechtschreibaufgaben auch für diese beiden Länderstichproben angemessen war.

Inhaltlich decken die Aufgabenzusammenstellungen verschiedene Anforderungen ab:

1. Lautgetreue Wörter schreiben
2. Regelgeleitetes Schreiben
3. Schreiben von unregelmäßigen Wörtern

Für einen möglichen Einsatz in einer Vollerhebung ergaben sich durch die Pilotierungsuntersuchung auch im Bereich Rechtschreiben einige wichtige Hinweise auf Verbesserungsmöglichkeiten hinsichtlich der getesteten Aufgabenformate und der Auswahl des geeigneten Wortmaterials.

3.3.2 Ausblick: Die Leistungsgruppen der Schülerinnen und Schüler

Im Hinblick auf die Anforderungen im Test ließen sich drei Leistungsgruppen in der Stichprobe identifizieren. Die Gruppen lassen sich anhand der Fehlerhäufigkeiten bei den Aufgaben beschreiben, in denen es galt, mehrere ganze Wörter zu schreiben.

(1) Schüler mit **unterdurchschnittlichen** Testleistungen hatten größere Schwierigkeiten in allen Aufgabenbereichen: Bei den lautgetreuen Wörtern lag ihre Fehlerquote im Schnitt bei 30% falsch geschriebenen Wörtern (mindestens 1 Fehler je Wort oder mehr), bei den übrigen regelmäßigen Wörtern und den Merkwörtern zu gleichermaßen bei etwa 50%.

(2) Schüler mit **durchschnittlichen** Testleistungen hatten bei den vorliegenden Aufgaben eine Fehlerquote von 13% bei den lautgetreuen Wörtern, von 15% bei übrigen regelmäßigen Wörtern und von 21% bei den Merkwörtern zu verzeichnen.

(3) Schüler mit **überdurchschnittlicher** Testleistung schrieben in allen Bereichen die ausgewählten Worte relativ sicher: Ihre Fehlerquote lag bei 5 % bei den lautgetreuen Wörtern, und war mit 9% bzw. 8% bei übrigen regelmäßigen und den Merkwörtern ungefähr gleich niedrig.

3.4 Mathematik

3.4.1 Inhaltliche Anforderungen, Schwierigkeiten und Eignung der Items

Im Bereich Mathematik wurden 15 Aufgabentypen getestet; die meisten davon enthielten eine oder mehrere unabhängige Teilaufgaben, so dass insgesamt 25 Einzelaufgaben in der Pilotierung eingesetzt wurden. Für die Bearbeitung standen 45 Minuten Bearbeitungszeit zur Verfügung.

Inhaltlich werden vier Bereiche abgedeckt:

1. Grundrechenarten
2. Zahlenbegriff und -verständnis bis 100
3. Geometrie
4. sachbezogene Mathematik

Zu den jeweiligen Bereichen wurden sowohl einfachere als auch komplexere Aufgaben eingesetzt. Einfachere Aufgaben lassen sich mit dem Schlagwort „**Reproduktion**“ bzw. „**Anwendung von Routinefertigkeiten**“ kennzeichnen: Die Anforderungen umfassen relativ einfachen und / oder im Unterricht recht übliche Anwendungen des gelernten Stoffes. Die Lösung dieser Aufgaben gibt darüber Auskunft, ob zumindest ein Teilverständnis der betreffenden mathematischen Konzepte vorhanden ist. Die komplexeren Aufgaben verlangten Verknüpfungen von Inhalten und Operationen oder deren flexiblere Anwendung. Sie weichen von häufiger vorkommenden, routinemäßigen Anwendungen ab und zeichnen sich entweder durch einen Transferaspekt oder einen problemlösenden Aspekt aus, d.h. der Lösungsweg ist weniger offensichtlich. Die Bewältigung dieser Aufgaben lässt den Schluss auf ein sichereres Beherrschen bzw. ein vollständigeres Verständnis der jeweiligen Inhalte zu.

Für einen Einsatz in einer Vollerhebung verbleiben 19 geeignete Aufgaben; davon sind 11 dem einfacheren, reproduktiven Anforderungsniveau zuzuordnen und 8 dem komplexeren

bzw. problemlösenden Anforderungsniveau. Die Verteilung der Aufgaben auf die Inhaltsgebiete zeigt Tabelle 4. Es ist nicht vollständig gelungen, in allen Gebieten eine ausreichende Anzahl von Aufgaben auf beiden Anforderungsniveaus zu erhalten. Da die Fachbereiche unterschiedliche Schwierigkeit haben, finden sich in der vorliegenden Auswahl in der Geometrie eher komplexere Aufgaben und im Bereich Grundrechenarten fast ausschließlich reproduktive Aufgaben. Abgesehen davon scheint die vorliegende Aufgabenauswahl ohne Änderungen für einen Einsatz in der Vollerhebung geeignet. Falls sich die Testzeit in der Vollerhebung auf 30 Min. reduziert, besteht außerdem die Möglichkeit, von den 19 verbleibenden Aufgaben eine Auswahl zu treffen.

Tab. 4: Durchschnittliche Lösungshäufigkeit der ausgewählten Mathematikaufgaben nach Gebiet und Aufgabentyp

	Anzahl reproduktive Aufgaben	Anzahl komplexe Aufgaben / Problemlösen	Summe der Aufgaben	*
Grundrechenarten	6	1	7	66 %
Zahlenverständnis	2	2	4	41 %
Geometrie	-	3	3	46 %
Sachbezogene Mathematik	3	2	5	49 %
Summe	11	8	19	
*Durchschnittliche Lösungshäufigkeit der Gebiete	66 %	33 %		50 %

3.4.2 Ausblick: Drei Leistungsgruppen der Schülerinnen und Schüler

Auch im Bereich Mathematik wurden drei Gruppen mit unterschiedlichen Leistungen gebildet und bestimmt, wie sich deren Testleistung anhand der bewältigten Anforderungen charakterisieren lässt:

(1) Schüler mit **unterdurchschnittlicher Testleistung** hatten bei den reproduktiven Aufgaben zu den Grundrechenarten einige Probleme. Im Bereich Zahlenbegriff bis 100 scheiterten sie in der Mehrzahl der Aufgaben. Auch bei den Aufgaben zur Geometrie und sachbezogenen Mathematik gelang es der Mehrzahl dieser Schüler nicht, die einfacheren (reproduktiven) Aufgaben zu lösen.

(2) Schüler mit **durchschnittlicher Testleistung** beherrschten die Aufgaben zu den Grundrechenarten auf beiden Niveaus. Im Bereich Zahlenverständnis bis 100 lösten sie die reproduktiven Aufgaben gut, hatten aber noch etwas Schwierigkeiten mit den komplexeren Aufgaben aus diesem Bereich. Bei den Aufgaben zur Geometrie und der sachbezogenen Mathematik scheiterten sie bei der Mehrzahl der Transferaufgaben.

(3) Schüler mit **überdurchschnittlicher Testleistung** lösten auch die komplexeren Aufgaben zu allen Inhaltsgebieten des Tests recht gut, d.h. im Unterschied zu den durchschnittlichen Schülern auch in den Bereichen des Zahlenraums bis 100, der Geometrie und der sachbezogenen Mathematik.

3.5 Zusammenhänge zwischen den Testteilen

Tabelle 5 gibt darüber Auskunft, ob und wie die einzelnen Testteile miteinander zusammenhängen, wobei zusätzlich noch das Lehrer- und das Elternurteil zur Lesefähigkeit Berücksichtigung finden. Die in der Tabelle aufgeführten Korrelationskoeffizienten sind Zusammenhangsmaße, die Werte zwischen -1 und $+1$ annehmen können. Das *Vorzeichen* gibt dabei die Richtung des Zusammenhangs an (ein inverser Zusammenhang zwischen zwei Größen drückt sich in negativem Vorzeichen aus²); der Betrag informiert über die Enge des Zusammenhangs (sehr enge Zusammenhänge zwischen zwei Größen nähern sich dem Wert 1.00, besteht hingegen gar kein Zusammenhang zwischen zwei Variablen, so nähert sich der Betrag dem Wert 0).

Durchweg zeigen sich starke Zusammenhänge der Testteile untereinander. Ebenso korrelieren sie sehr stark mit dem per Klassenliste erfassten Lehrerurteil, der (in Jahrgangsstufe 2 oft noch hypothetischen) Fachnote in Deutsch und Mathematik bzw. der Einschätzung der Lesefähigkeit durch die Lehrkräfte. Etwas geringer, aber immer noch recht hoch fällt die Korrelation zwischen dem entsprechenden Urteil der Eltern und den Testleistungen aus.

Kernaussage der Tabelle 5 ist, dass es – über alle Schülerinnen und Schüler hinweg betrachtet – keine speziellen Fähigkeiten und keine Teilbegabungen gibt. Tendenziell gilt: Bessere Leistungen in der einen Domäne gehen einher mit besseren Leistungen in den anderen Bereichen – und dies wird i.d.R. von den Lehrkräften und den Eltern auch so wahrgenommen.

Tab. 5: Korrelationen der Testteile untereinander, zu den Fähigkeitsurteilen von Lehrern und Eltern sowie zum Geschlecht

	Lesegeschwindigkeit	Rechtschreiben	Mathematik	Deutschnote	Mathematiknote	*Lehrerurteil Lesefähigkeit	*Elternurteil Lesefähigkeit	Geschlecht
Leseverständnis	.59	.63	.62	-.56	-.51	-.60	-.45	.08
Lesegeschwindigkeit		.63	.54	-.51	-.44	-.55	-.46	.14
Rechtschreiben			.65	-.64	-.55	-.61	-.49	.16
Mathematik				-.53	-.60	-.49	-.33	-.06**
Deutschnote					.70	.83	.53	-.14
Mathematiknote						.63	.35	.09
Lehrerurteil* Lesefähigkeit							.61	-.12
Elternurteil* Lesefähigkeit								-.12

** Alle Korrelationen bis auf die markierte sind zuverlässig, $p < .01$. N je nach Kombination zwischen 1.053 und 1.021 bei den Tests und 1.003 und 853. Geschlecht (1) = Junge, (2) = Mädchen.

² Z. B. ist die Korrelationen der Testleistungen mit den Zensuren negativ, da gute Testergebnisse (=hohe Punktzahlen) mit gute Leistungsbewertungen (= niedrige Notenwerte) einhergehen.

Daneben zeigen sich sehr geringfügige, aber in der Gesamtstichprobe zuverlässige Zusammenhänge der Testleistungen im Bereich der sprachbezogenen Testleistungen mit dem Geschlecht der Schüler. Jungen schneiden in der Gesamtstichprobe etwas schlechter bei den Tests ab als Mädchen. Ein Hinweis auf eine Ursache dieses bekannten Effekts beim Leseverständnis und der Lesegeschwindigkeit ergibt sich aus den Angaben des Schülerfragebogens (s.u.). Jungen berichteten hier ein geringeres Interesse und eine geringere Häufigkeit des Lesens. Der Zusammenhang zwischen Geschlecht und Leseverständnis ist nicht mehr zuverlässig, wenn man diese beiden Variablen kontrolliert, d.h. bei etwa gleichem Interesse und gleicher Lesegeschwindigkeit nähern sich die Testleistungen der Jungen und der Mädchen an.

4. Hintergrundmerkmale

4.1 Schülerfragebogen

Als weiteres Erhebungsinstrument wurde ein dreiseitiger Schülerfragebogen eingesetzt, der von den Schülern während des ersten Testtages in Beisein des Testleiters auszufüllen war. Mit dem Fragebogen wurden Angaben zur Familien- bzw. Haushaltsgröße, der Umgangssprache im Haushalt, zum Lese- und Freizeitverhalten sowie zur schulbezogenen Motivation erhoben. Als schwierig erwies sich für die Zweitklässler die Frage nach der Familienstruktur bzw. der jeweiligen Haushaltsgröße oder auch dem Geburtsland der Eltern. Für diese Items sind vergleichsweise hohe Missing-Quoten zu verzeichnen, was vermutlich einerseits auf das Frageformat zurückzuführen ist, andererseits aber auch auf die wohl für Zweitklässler ungewohnten Fragestellungen.

Wenig Probleme bereiteten den Schülern offensichtlich die meisten im Ratingskalenformat gestellten Fragen mit gestuften Antwortvorgaben (Beispiel: „Wie gerne liest Du?“, s.u.), was sich in weniger fehlenden Werten zeigte. Die zugehörigen Fragen wurden dabei mit vierstufigen Antwortmöglichkeiten erfasst, wobei 4 die höchste bzw. positivste Ausprägung bedeutet und 1 die geringste bzw. negativste³.

Die Skalen „Schulmotivation“, „schulisches Selbstkonzept“ und „Anstrengungsbereitschaft“ bestehen aus den Mittelwerten von 4 – 5 inhaltlich zugehörigen Fragen. Die Skala „Schullust“ wurde dabei aus den Items zu Punkt 14 „Einige Fragen zur Schule“ gebildet (Beispielitem: „Wie gerne gehst Du zur Schule?“). Die Fragen zu Punkt 17 „Wie ist das bei Dir in der Schule?“ gaben Auskunft über die grundlegende „Anstrengungsbereitschaft“ des befragten Schülers (Beispielitem: „Im Unterricht gebe ich mir viel Mühe alles zu verstehen“). Mit der Skala „schulische Selbstwirksamkeitserwartung“ wurde das Zutrauen in die eigenen schulleistungsbezogenen Fähigkeiten erfasst (Beispielitem „Ich verstehe das Meiste, was wir in der Schule lernen müssen“).

Tabelle 6 zeigt die mittlere Ausprägung dieser motivationalen Hintergrundmerkmale in beiden Substichproben; dabei ist kein Länderunterschied in den Variablen festzustellen. Auffällig ist die außerordentlich hohe Ausprägung aller drei motivationsbezogenen Skalenwerte im Jahrgang.

³ Die Items wurden auf einer Skala von 1 – 4 bewertet; wobei: 1 = gar nicht gerne, 2 = weniger gerne, 3 = ziemlich gerne; 4 = sehr gerne.

Wie nicht zuletzt differenzierte Analysen im Kontext von PISA bestätigen, nimmt das Leseinteresse sowie die Lesepraxis bedeutenden Einfluss auf die tatsächliche Leseleistung. Tabelle 6 zeigt die Mittelwerte des Leseinteresses und der Häufigkeit des Lesens. Das Leseinteresse ist ähnlich wie die motivationalen Hintergrundmerkmale unabhängig von der Substichprobe sehr hoch ausgeprägt. Insgesamt zeigen ca. 65% der Schülerinnen und Schüler ein starkes bis mittleres Leseinteresse, nur 17% lesen weniger gerne und 6% gar nicht gerne. Nahezu 70% der Schüler geben an, ein- bis zweimal die Woche oder gar täglich zu lesen; lediglich knapp ein Fünftel der Schüler liest nur ein- bis zweimal im Monat oder weniger. Die folgende Tabelle zeigt das mittlere Leseinteresse.

Tab. 6: Motivationsbezogene Variablen in den Stichproben

	Gesamtstichprobe	Brandenburg	Berlin
Skalenwerte Schulmotivation			
Schullust ($N = 994$)	$M = 3.35$ $SD = 0.58$	$M = 3.37$ $SD = 0.56$	$M = 3.34$ $SD = 0.62$
Anstrengungsbereitschaft ($N = 970$)	$M = 3.55$ $SD = 0.53$	$M = 3.57$ $SD = 0.53$	$M = 3.53$ $SD = .539$
schulisches Selbstkonzept ($N = 965$)	$M = 3.44$ $SD = 0.60$	$M = 3.43$ $SD = 0.62$	$M = 3.45$ $SD = 0.59$
Fragen zur Lesemotivation			
Leseinteresse ($N = 987$)	$M = 3.12$ $SD = 0.95$	$M = 3.11$ $SD = 0.98$	$M = 3.13$ $SD = 0.91$
Lesehäufigkeit ($N = 968$)	$M = 3.15$ $SD = 1.06$	$M = 3.10$ $SD = 1.08$	$M = 3.20$ $SD = 1.03$

Neben diesen motivationalen Faktoren wurde weiterhin die durchschnittliche Zeit erfragt, die die Schüler für ihre Hausaufgaben aufwenden. Dabei zeigt sich ein zuverlässiger, aber schwacher Zusammenhang zwischen den Testleistungen und der durchschnittlichen Hausaufgabenzeit: schwächere Schüler brauchten länger für die Bearbeitung der Hausaufgaben (zwischen $r = -.14$ und $r = -.16$). Die Angabe wurde in fünf aufsteigenden Kategorien in Zeitabschnitten von je 15 Minuten erhoben. Die häufigste Nennung liegt bei „etwa 15 Minuten“. Tabelle 7 zeigt die Nennung zur Frage in beiden Ländern.

Tab. 7: Angaben zur Hausaufgabenzeit nach Ländern ($N = 959$)

	Gesamt	Brandenburg	Berlin
(1) Weniger als 15 Min.	315 (33 %)	172 (34 %)	143 (32 %)
(2) 15 Min.	360 (37 %)	192 (38 %)	168 (37 %)
(3) 30 Min.	209 (21 %)	107 (21 %)	102 (23 %)
(4) 45 Min.	51 (5 %)	23 (4 %)	28 (6 %)
(5) 60 Min.	24 (3 %)	14 (3 %)	10 (2 %)

4.2 Elternfragebögen

Mit den Elternfragebögen wurden Informationen über den sozialen Hintergrund sowie persönliche Merkmale der Schüler erfasst. Erhoben wurden einerseits soziale und ökonomische Verhältnisse wie bspw. Berufstätigkeit, Haushaltseinkommen, Bildungsabschlüsse und Migrationsstatus der Eltern. Darüber hinaus wurden Auskünfte zur elterlichen Unterstützung des Schülers in schulischen Angelegenheiten erhoben. Aufgrund zahlreicher vorangegangener Studien ist zu vermuten, dass diese Hintergrundvariablen als Determinanten der jeweiligen Schulleistung fungieren. Zu den zentralen Einflussfaktoren zählen Merkmale wie Muttersprache, Migrationsstatus, Qualifikationsniveau sowie Erwerbstätigkeit und Erwerbsstatus der Eltern.

Innerhalb der Bildungsforschung gilt der schulische Bildungsabschluss in der Familie seit langem als „durchschlagender Faktor“, der auf den gesamten Bildungsweg des Kindes Einfluss nimmt. Dabei erweist sich das schulische Qualifikationsniveau innerhalb der Familie als weitaus bedeutsamer als die später erworbenen beruflichen Abschlüsse. Für die Berechnungen innerhalb der Pilotierungsstudie wurde aus den Angaben zum schulischen Qualifikationsniveau der weiblichen sowie der männlichen Bezugsperson der „höchste Bildungsabschluss in der Familie“ herangezogen. Abbildung 2 zeigt die Verteilung der befragten Familien auf das jeweilige Niveau der Bildungsabschlüsse getrennt für Brandenburg und Berlin. Unterschieden wurde zwischen den Qualifikationsniveaus „kein schulischer Abschluss“, „Volksschul-, Hauptschulabschluss oder Abschluss der 8. Klasse POS“, „Realschule bzw. mittlere Reife“ und „Hochschul- bzw. Fachhochschulreife“. Dabei zeigt sich – mit der Ausnahme der etwas größeren Bedeutung mittlerer Bildungsabschlüsse in Brandenburg – eine ähnliche Verteilung der Qualifikationsniveaus in den beiden Bundesländern.

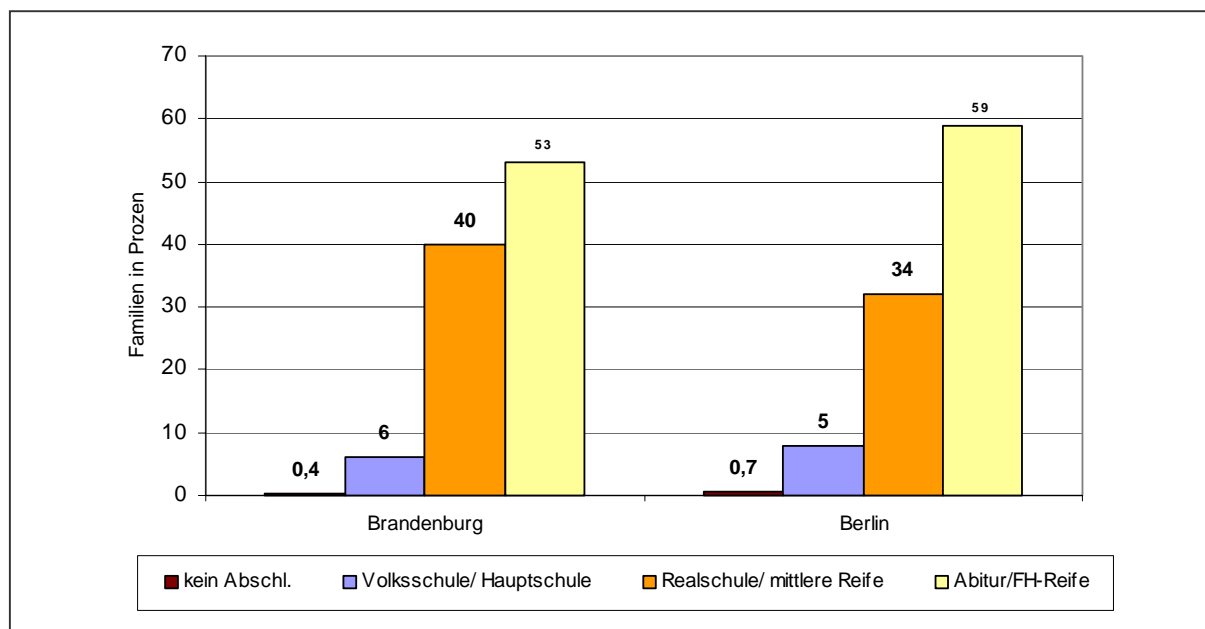


Abbildung 2: Prozentuale Verteilung der höchsten Bildungsabschlüsse in der Familie in Brandenburg und Berlin ($N = 899$)

Neben dem höchsten Schulabschluss innerhalb der Elternschaft ist es weiterhin der Migrationsstatus, der sich – in engem Zusammenhang mit dem beruflichen Status – stark auf die

Schulleistung auswirkt. Die Variable „Migrationsstatus“ unterscheidet drei Ausprägungen der familiären Migrationsgeschichte: Familien, in denen beide Elternteile aus Deutschland stammen (Familie ohne Migrationshintergrund), Familien, in denen ein Elternteil im Ausland geboren wurde und schließlich Familien, in denen beide Elternteile im Ausland geboren wurden. Unter die Ausprägung „Familien ohne Migrationshintergrund“ werden auch Familien gefasst, bei denen Personen aus der Generation der Großeltern in Deutschland einwanderten und somit zur so genannten „dritten Generation“ zählen. In der Gesamtstichprobe stammen etwa 17% der befragten Schüler aus Familien mit Migrationshintergrund, davon gut 7% aus Familien, in denen beide Elternteile im Ausland geboren sind. Abbildung 3 zeigt die Verteilung der Familien hinsichtlich des Migrationsstatus getrennt für die untersuchten Länderstichproben.

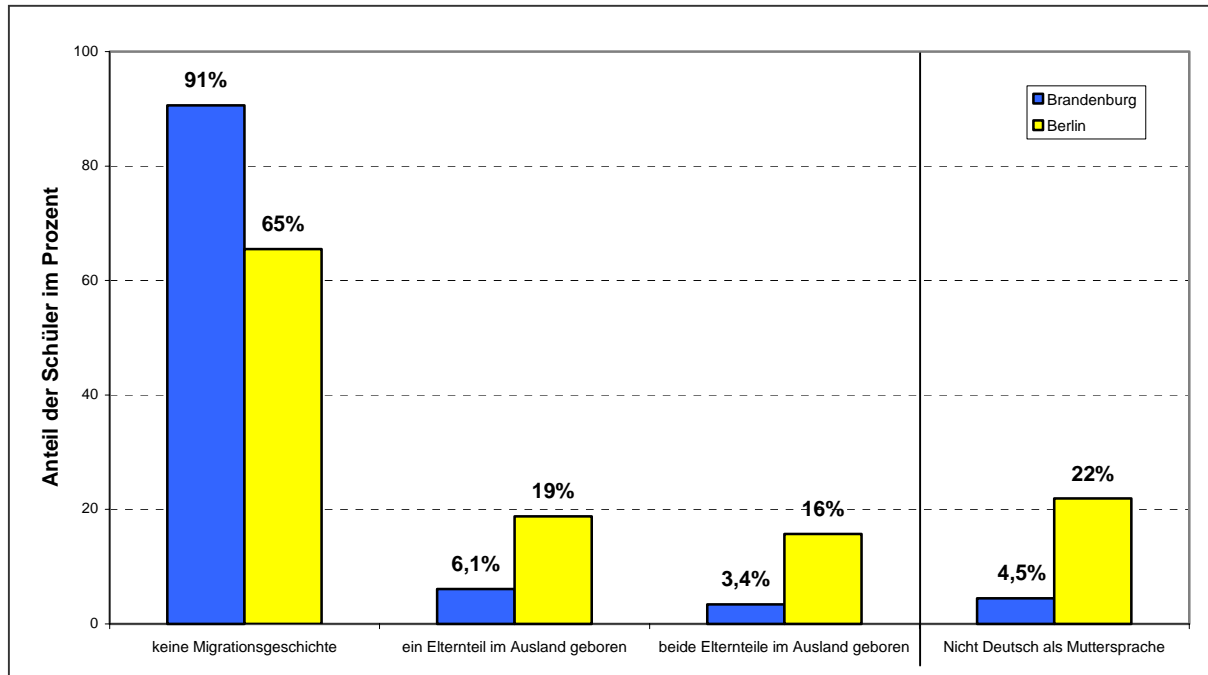


Abbildung 3: Anzahl der Schüler mit und ohne Migrationshintergrund ($N = 847$) sowie Anzahl der Muttersprachler in Berlin und Brandenburg (Angabe der Klassenliste: $N = 1102$)

Dabei zeichnet sich deutlich der höhere Anteil von Familien mit Migrationsgeschichte für das Bundesland Berlin ab. Starken Einfluss auf die Schulleistung nimmt weiterhin die jeweilige Muttersprache des Kindes, da das Beherrschen der „Majoritätssprache“ Deutsch auf einem hohen Niveau zu den Grundvoraussetzungen schulischer Leistungsfähigkeit zählt. Die Erfassung der Muttersprachlichkeit ergänzt dabei den Indikator des Migrationsstatus, da auch in der dritten Migrationsgeneration Kinder mit einer anderen Muttersprache oder zumindest bilingual aufwachsen. Die jeweilige Muttersprache des befragten Schülers wurde über entsprechende Angaben im Elternfragebogen („welche Sprache wird in Ihrer Familie überwiegend gesprochen?“), der Klassenliste („Muttersprache Deutsch?“) sowie des Schülerfragebogens („welche Sprache hast Du sprechen gelernt, als Du noch ganz klein warst?“) ermittelt. In der Gesamtstichprobe sprechen 86 % der teilnehmenden Schüler in der zweiten Jahrgangsstufe Deutsch als Muttersprache. Dieser Anteil ist der Stichprobenziehung entsprechend in Berlin deutlich geringer als in der Brandenburger Stichprobe (s.o. Abschnitt *Stichprobe und Durchführung*).

In den begleitenden Fragebögen wurde sowohl an die Eltern als auch an die Schüler verschiedene Fragen gerichtet, anhand derer Indikatoren für das kulturelle, soziale und ökonomische

Kapital gewonnen werden konnten. Zur Abschätzung des sozialen Kapitals des Elternhauses, das Befunden der PISA- sowie der IGLU-Studie zufolge in engem Zusammenhang mit der Bildungsbeteiligung einerseits und der Schulleistung andererseits steht, wurde in den Elternfragebögen der jeweilige Berufsstatus der Eltern erfasst. Die Selbstzuordnung zu insgesamt elf Berufsstatusgruppen wurde anschließend mit dem Klassifikationsindex EGP (nach Erikson, Goldthorpe und Portocarero) zu sechs Berufs- oder auch EGP-Klassen verdichtet. Das *EGP-Schema* (oder auch: *Dienstklassenindex*) als Sozialschicht-Indikator lässt eine anschauliche Klassifizierung unterschiedlichster Berufe zu und hat darüber hinaus den Vorteil, nicht direkt mit dem formalen Bildungsabschluss verbunden zu sein. So können auch Befragte mit niedrigerer Schulbildung bei entsprechendem beruflichen Engagement oder entsprechendem Tätigkeitsprofil zu den höheren Klassen gezählt werden. Die Einteilung der befragten Eltern (herangezogen wird jeweils die EGP-Klasse der männlichen Bezugsperson) in EGP-Klassen trägt damit sowohl dem (Berufs-) Prestige als auch den ökonomischen Verhältnisse der Familien Rechnung.

In Anlehnung an die PISA- und IGLU-Studie werden die Befragten sechs EGP-Klassen zugeordnet. In der Gesamtstichprobe wächst nahezu ein Drittel der Kinder in Familien aus der oberen Dienstklasse auf; erstaunlicherweise spielt der Anteil der Eltern aus der unteren Dienstklasse mit 3% eine deutlich untergeordnete Rolle. Auch zu den an- und ungelerten Arbeitern sind insgesamt lediglich 3% der Familien zu zählen. Abbildung 4 zeigt die Verteilung der befragten Eltern auf die sechs EGP-Klassen im Ländervergleich. Die zugrunde liegende Sozialstruktur beider Substichproben ist weitgehend ähnlich.

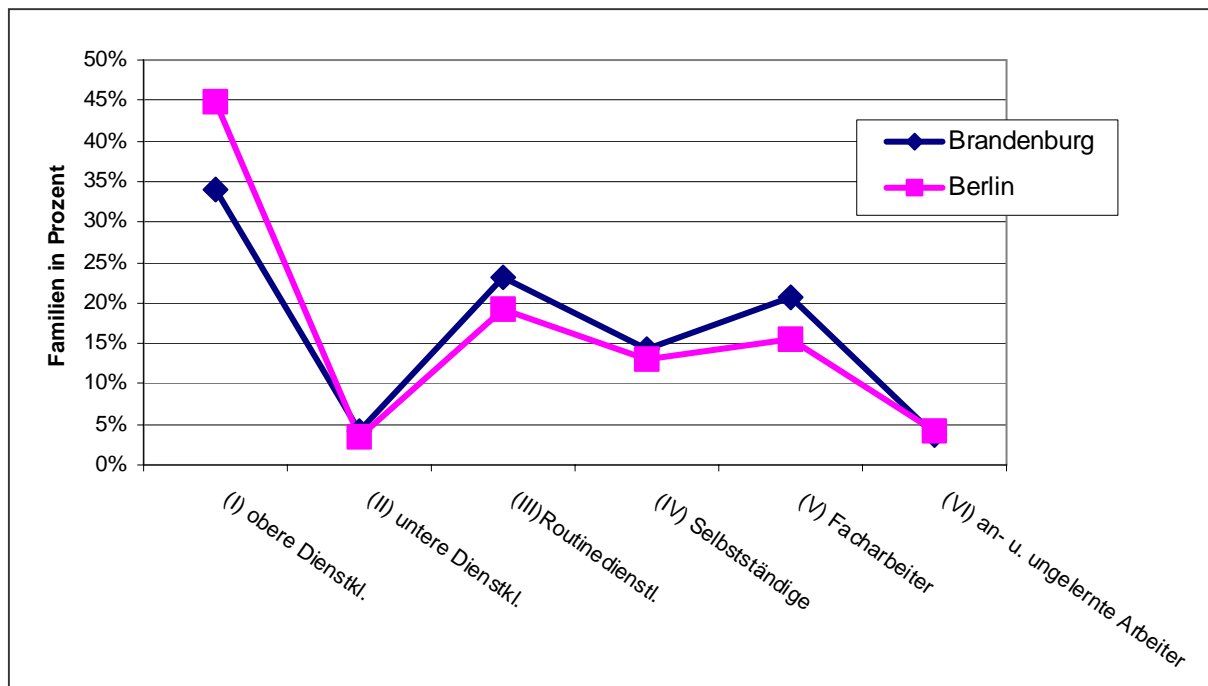


Abbildung 4: Sozialschichtzugehörigkeit der befragten Familien nach dem EGP-Klassifikationsindex in Brandenburg und Berlin ($N = 864$).

Zum kulturellen Kapital, das – wie zahlreiche Studien in der Tradition der Schulleistungsvergleichsforschung bestätigen – ähnlich wie der ökonomische Hintergrund in engem Zusammenhang mit der schulischen Leistungsfähigkeit steht, zählt neben den schulischen und beruflichen Qualifikationen im Elternhaus (institutionelles kulturelles Kapital) auch das so genannte „verobjektivierte“ kulturelle Kapital. Letzteres, erfassbar bspw. über die Anzahl der Lexika

im Haushalt, das Verfügen über einen eigenen PC oder auch der Art der Kunstgegenstände im Haus, wurde im Schüler- wie auch im Elternfragebogen über das Item „Anzahl der Bücher im Haushalt“ operationalisiert. Die häufigste Nennung in der Gesamtstichprobe lag bei „über 101-200 Bücher“ (42%), zuverlässige Länderunterschiede zeigen sich nicht.

Tab. 8: Anzahl der Bücher im Haushalt in Berlin und Brandenburg ($N = 908$)

	Brandenburg	Berlin
(1) 0-10 Bücher	14 (3 %)	30 (7 %)
(2) 11-25 Bücher	35 (7 %)	158 (6 %)
(3) 26-100 Bücher	129 (27 %)	119 (26 %)
(4) 101-200 Bücher	99 (21 %)	89 (20 %)
(5) über 200 Bücher	198 (42 %)	190 (42 %)

4.3 Zusammenhänge zwischen Hintergrundmerkmalen und Testleistungen

Die meisten der beschriebenen Hintergrundmerkmale stehen in Zusammenhang mit den jeweiligen Schulleistungen der Schülerinnen und Schüler. Für die vorliegende Untersuchung liefert Tabelle 9 eine Übersicht über Korrelationen der dargestellten Hintergrundmerkmale mit den eingesetzten Testsummenwerten in der Gesamtstichprobe.⁴

Die Hintergrundmerkmale können herangezogen werden, um die Testleistungen besser verstehen und erklären zu können. Soll von einer Schülerin oder einem Schüler ohne zusätzliche Informationen die Testleistung vorhergesagt werden, so wird man als Schätzung den Gruppenmittelwert verwenden (müssen). Sind hingegen in einer Stichprobe die Beziehungen zwischen weiteren Merkmalen und der Testleistung bekannt, lässt sich die Testleistung durch die Berücksichtigung dieser zusätzlichen Information besser erklären bzw. vorhersagen. Bei diesem Vorgehen werden die Hintergrundvariablen als sogenannte *Prädiktoren* für die Vorhersage der Testleistung verwendet. Dabei interessiert auch das Ausmaß, in dem ein oder mehrere Prädiktoren in der Lage sind, Leistungsunterschiede zu erklären. Dazu wird bestimmt, in welchem Ausmaß die vorgefundenen Leistungsunterschiede auf die Prädiktoren zurückzuführen sind. Dies wird im sogenannten Anteil *aufgeklärter Varianz* ausgedrückt, der in den letzten beiden Zeilen in der Tabelle 9 jeweils für die verschiedenen Testteile dargestellt wird. Zwischen 17% und 22% der vorgefundenen Leistungsunterschiede lassen sich demnach allein auf die untersuchten Hintergrundmerkmale zurückführen.

In der Untersuchung wurde auch bestimmt, wie hoch der Varianzanteil ist, den nur die stärksten Prädiktoren allein aufklären können (letzte Zeile in der Tabelle 9). Dies sind die Variablen „höchster schulischer Abschluss der Familie“, „Migrationshintergrund“ und „Muttersprache“.

⁴ Bei der Interpretation der dargestellten Zusammenhänge ist zu berücksichtigen, dass hierbei noch nicht die Korrelationen der Hintergrundmerkmale untereinander abgebildet sind. Beispielsweise gehen höhere Schulabschlüsse in den Herkunftsfamilien mit höherem Einkommen, höherer EGP-Klasse und höherem Buchbestand usw. einher. Aussagen über die *relative* Varianzaufklärung der Faktoren lassen sich anhand der dargestellten Korrelationen allein noch nicht treffen. Regressionsanalytisch wurde zusätzlich der durch die dargestellten Hintergrundmerkmale gemeinsam aufgeklärte Anteil an der Leistungsvarianz bestimmt. Dieser Wert ist in der vorletzten Zeile der Tabelle 9 für die einzelnen Leistungstests dargestellt.

Tab. 9: Korrelationen zwischen Schulleistungen und zentralen Hintergrundvariablen

	Lesever- ständnis	Lesege- schwindig- keit	Rechtschrei- ben	Mathematik
Muttersprache Deutsch? (Klassenliste)	.25	.26	.19	.22
Migrationsstatus (beide Fragebögen)	-.24	-.20	-.18	-.25
höchster schulischer Bildungsabschluss in der Familie (EFB)	.34	.26	.29	.29
Einkommen, monatlich (EFB)	.34	.28	.31	.33
EGP-Klasse der Bezugsperson (EFB)	-.20	-.18	-.14	-.16
Bücher im Haushalt (beide Fragebögen)	.38	.28	.30	.32
Lesen: Wie gerne? (SFB)	.20	.23	.22	.08
Lesen: Wie oft? (SFB)	.23	.23	.22	.20
Konzept eigener Begabung (SFB)	.19	.18	.24	.19
Anstrengungsbereitschaft (SFB)	.13	.11	.18	.13
Lust auf Schule (SFB)	.01	.01	.06	-.03*
Gemeinsam aufgeklärter Varianzanteil <i>aller</i> dargestellten Hintergrundmerkmale	22 %	19 %	18 %	17 %
Nur durch die Variablen „höchster Schulabschluss“, „Migration“ / „Muttersprache“ aufgeklärte Varianz	14 %	11 %	9 %	11 %

* In der vorliegenden Stichprobe sind Korrelationen in der absoluten Höhe größer als .08 zuverlässig, $p > .01$; EFB = Elternfragebogen; SFB = Schülerfragebogen.

Die nachfolgenden Abbildungen 5 bis 7 zeigen die Zusammenhänge zwischen den Testleistungen im Bereich Lesen und Mathematik exemplarisch für diese beiden bedeutsamsten Prädiktoren. Zur besseren Veranschaulichung wurden die erzielten Testleistungen im Lesen und der Mathematik dafür auf eine Skala mit Mittelwert 250 und Standardabweichung von 50 transformiert. Dargestellt werden die mittleren Testleistungen in Schülergruppen mit unterschiedlicher Ausprägung der Hintergrundmerkmale.

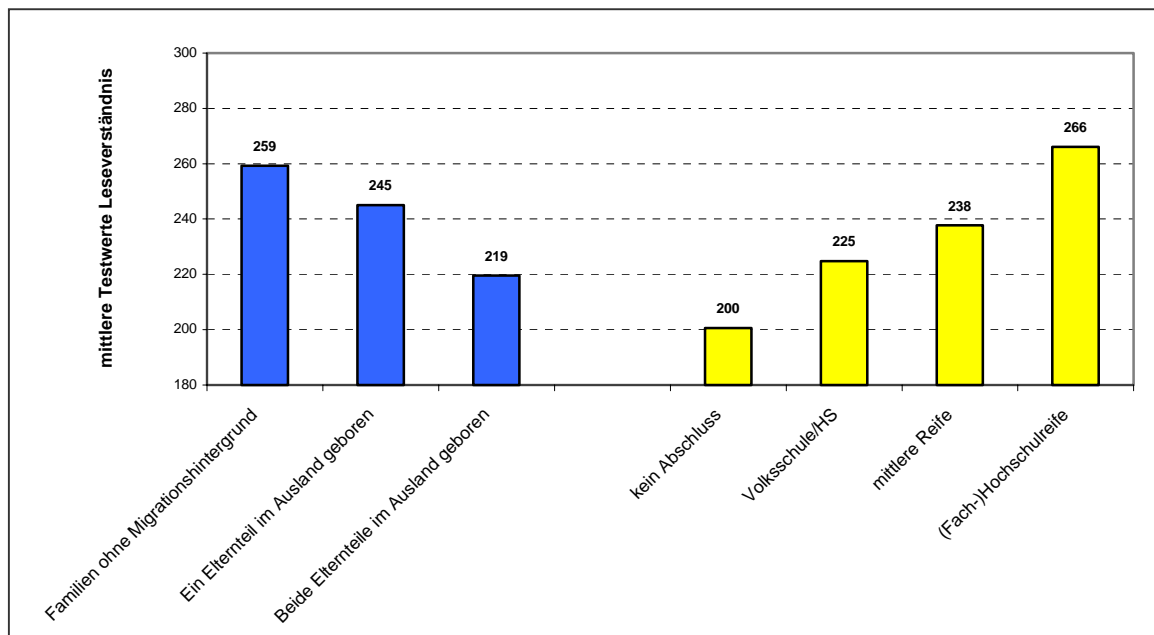


Abbildung 5: Leseverständnisleistung in Abhängigkeit von Migrationsstatus und schulischem Abschluss der Eltern

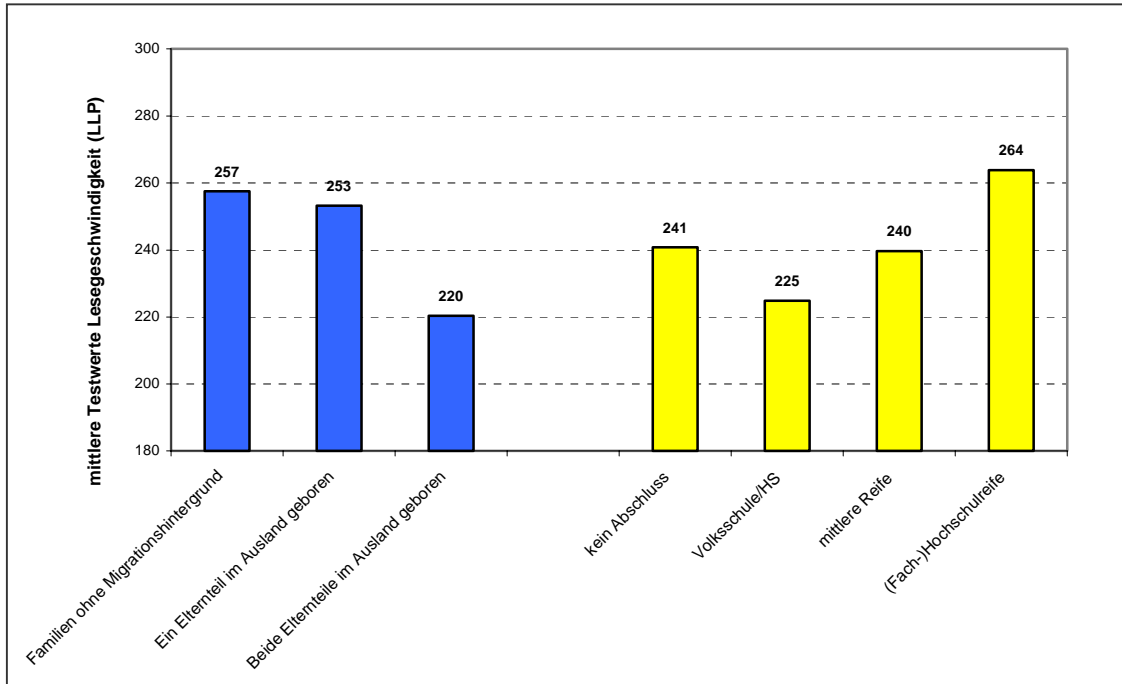


Abbildung 6: Lesegeschwindigkeit in Abhängigkeit von Migrationsstatus und schulischem Abschluss der Eltern

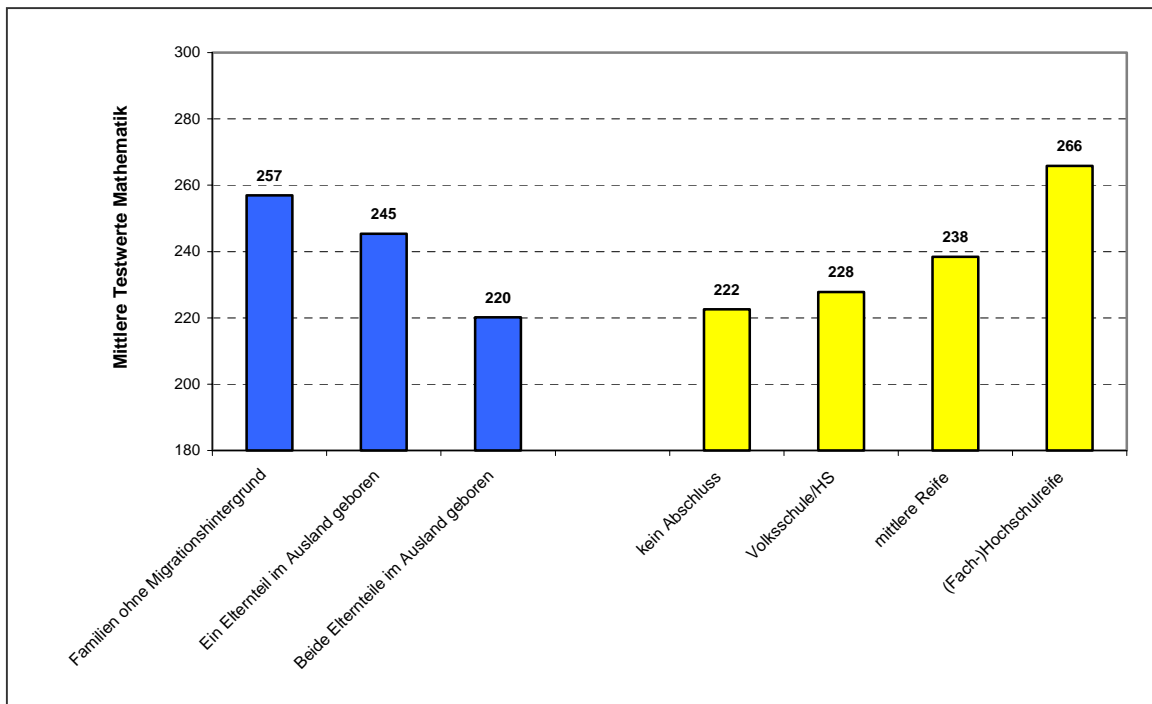


Abbildung 7: Mathematikleistung in Abhängigkeit von Migrationsstatus und schulischem Abschluss der Eltern