Rainer H. Lehmann, Rainer Peek,
Rüdiger Gänsfuß, Sabine Lutkat,
Stephan Mücke und Ingola Barth

QuaSUM

Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik

Ergebnisse einer repräsentativen
Untersuchung im Land Brandenburg

Einleitung

Nationale und Internationale Schulleistungsuntersuchungen der letzten Jahre haben unter Eltern, Lehrkräften, Bildungspolitikern und Erziehungswissenschaftlern eine höchst emotionale und polarisierte Diskussion ausgelöst. Auf der einen Seite werden derartige Untersuchungen als längst überfällige empirische Wende der Schulforschung begrüßt, auf der anderen Seite wird der Vorwurf erhoben, die mit solchen Studien verbundenen testtheoretischen und empirisch-methodischen Probleme würden zu einem völlig verzerrten Bild der Wirklichkeit führen.

Aus der Sicht des Ministeriums für Bildung, Jugend und Sport sind weder Dramatisierung noch allzu große Euphorie die angemessene Antwort auf die zunehmende Bedeutung externer Qualitätsuntersuchungen. Empirische Forschung hat den Vorteil, dass sie an die Stelle des bloßen Vermutens, der unterschiedlichen Meinungen und Einschätzungen relativ gesicherte Befunde setzen kann. Angesichts der vielfältigen Ziele, die durch Unterricht und Erziehung in der Schule bewirkt werden sollen, ist die erfolgreiche Aneignung bestimmter Kenntnisse durch Schülerinnen und Schüler natürlich eine grundlegende „Leistung“, allerdings erschöpft sich Schulleistung nicht in der Konzentration auf Wissensbestände, sie umfasst weitere Kompetenzbereiche.

Bei der im Land Brandenburg im Jahr 1998 geplanten und 1999 durchgeführten Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik handelt es sich um eine Untersuchung, die den Stand der Aneignung von mathematischen Kenntnissen am Ende der Jahrgangsstufen 5 und 9 bilanziert. Angesichts der kritischen Diskussion um die Validität der TIMSS-Aufgaben für den deutschen Mathematikunterricht und vor dem Hintergrund der Auseinandersetzungen zwischen Schulbehörde und Lehrerschaft bei der Hamburger Lernausgangslagenuntersuchung (LAU 5) hat das Ministerium für Bildung, Jugend und Sport sich bemüht, bei der Konzeption und Durchführung der Studie diese kritischen Hinweise aufzugreifen. Das Brandenburger Vorgehen ist daher durch einige Besonderheiten gekennzeichnet:

- **Spezifische Testentwicklung für Brandenburg**

Die Aussagekraft und Vertrauenswürdigkeit der Ergebnisse aus Schulleistungstests steht und fällt mit den gewählten Testaufgaben, ihrer Lehrplan- und Unterrichtsvalidität, der angemessenen Transparenz der getesteten Fähigkeiten, dem Antwortformat der Aufgaben und der für die untersuchten Jahrgangsstufen geeigneten Formulierung der Aufgaben. Zur Entwicklung eines rahmenplanvaliden Mathematiktests wurden Brandenburger Lehrkräfte als Aufgabenentwickler herangezogen, die über 600 erarbeiteten Aufgaben wurden unter Anwendung teststatistischer Verfahren „eingedampft“ und im Sommer und Herbst 1998 in zwei Pilotstudien erprobt. Der QuaSUM-Test ist damit ein Spiegel des in Brandenburg tatsächlich erteilten Mathematikunterrichts.

- **Akzeptanz durch Transparenz und Mitwirkung**

In begleitenden Gesprächen und Informationsveranstaltungen mit der GEW, dem Landesschulbeirat, dem Hauptpersonalrat und den ausgewählten Schulen ist es im Vorfeld der Hauptuntersuchung gelungen, eine überwiegend positive Grundstimmung bei den Beteiligten zum Projekt herzustellen. Die Ziele der Untersuchung und die für den Einsatz erstellten Fragebögen zum Schul- und Unterrichtskontext waren rechtzeitig bekannt. Statt der in einigen alten Bundesländern vor allem nach der TIMSS-Debatte feststellbaren Polarisierung und Ablehnung quantitativer Schulforschung gab es in Brandenburg eher eine kritische Mitwirkung und ein gespanntes Interesse.

Ein Grund dafür war unter Umständen auch die Auswahl, Ausbildung und der Einsatz von 30 Lehrkräften aus allen Landkreisen zu „Multiplikatoren für Schulqualitätsuntersuchungen“. Die erfolgreiche Vermittlung der Zielsetzung und Anlage des Projektes sowie die Datenerhebung war nur mit Hilfe dieser Lehrkräfte möglich.

- **Klassen- bzw. kursbezogene Rückmeldung**

Neben der Funktion des Systemmonitorings war ein weiteres Ziel der Untersuchung, Schulen durch exakte Ergebnisrückmeldungen auf der Ebene der Klasse oder des Kurses in die Lage zu versetzen, eine datengestützte Diskussion über die Ergebnisse des eigenen Mathematik-Unterrichts zu führen. Zur Vorbereitung auf diese neue Form der Ergebnisbetrachtung fanden in den Monaten Oktober und November 1999 in allen Landkreisen Beratungen mit den betroffenen Mathematiklehrkräften und Schulleitungen aller Stichprobenschulen statt. Die an der Erhebung der QuaSUM-Daten beteiligten Schulen erhielten dann im Januar 2000 ausführliche Ergebnisrückmeldungen für jede teilnehmende Klasse bzw. jeden Gesamtschulkurs.

Die in dieser Studie verwandte Form der Rückmeldung von ausgewählten Ergebnissen an Klassen und Kursgruppen im fairen Vergleich schafft die Voraussetzung für eine intensive schulinterne Diskussion. Dabei wird erwartet, dass zwischen den betroffenen Lehrkräften eine offene und kollegiale Diskussion über die jeweiligen Ergebnisse und die dafür maßgeblichen innerschulischen Bedingungen geführt wird. Die Rahmenbedingungen des QuaSUM-Projekts sind bewusst so gestaltet worden, dass Entwicklungspotentiale der Schulen bei der Aufklärung der Ursachen und Hintergründe von Lernergebnissen genutzt werden müssen, wenn es zu Verbesserungen kommen soll. Die Ergebnisse der Klassen und Kurse sind vor allem für die schulinternen Diskussionen zur Steigerung der Effizienz vorgesehen. Die Klassendaten „gehören“ der Schule, Eltern und Schülerinnen und Schüler sind jedoch zu informieren.

- **Keine Rückmeldung der konkreten Schul- und Klassenergebnisse an die Schulaufsicht**

Externe Schulqualitätsuntersuchungen mit Ergebnissen auf Klassen- bzw. Kursbasis lösen auch in Brandenburg bei den beteiligten Lehrkräften große Ängste aus. Diese beziehen sich sowohl auf die Information der Schulleitung, auf die Schulaufsicht und auf allgemein öffentliche Diskussionen über Einzelschulergebnisse. Nach längeren Diskussionen und in großer Übereinstimmung auch mit den örtlich zuständigen

Schulrätinnen und Schulräten wurde für Brandenburg festgelegt, dass die Schulaufsicht keine klassen- bzw. kursbezogenen Ergebnisse erfährt (Ausnahmen wurden lediglich für eine kleine Zahl von Modellschulen festgelegt, die zusätzlich in die Studie aufgenommen wurden).

Durch die Festlegung, dass – von wenigen Ausnahmen abgesehen – weder das Ministerium für Bildung, Jugend und Sport noch die Schulaufsicht in den staatlichen Schulämtern die Einzelergebnisse der Schulen erfährt, besteht die Chance zu einer offenen Diskussion in den Schulen. Schulen können selbstverständlich ihre internen Ergebnisse und Schlussfolgerungen mit ihren zuständigen Schulrätinnen und Schulräten besprechen, es gibt jedoch keine Verpflichtung. Angesichts der offenen Diskussion über die Probleme der Balance zwischen Kontrolle und Beratung im schulaufsichtlichen Handeln ist diese Entscheidung vor allem von den Schulen sehr begrüßt worden. Selbstverständlich wird sich die Schulaufsicht mit den zentralen Landesergebnissen und den Hinweisen auf die Bedingungsfaktoren gelingenden Fachunterrichts intensiv beschäftigen.

- **Systematische Evaluation der Wirkungen der Rückmeldungen**

Bei allen bisher in Deutschland durchgeführten externen Untersuchungen mit Schulrückmeldungen bleibt offen, wie und an welchen Sachfragen orientiert sich die Schulen intern mit den Ergebnissen auseinander gesetzt haben. Ebenso gibt es bisher keine empirischen Befunde, welchen Beitrag die Rückmeldung extern erhobener Daten auf Schulentwicklungsprozesse hat. Brandenburg hat daher in Kooperation mit der Humboldt-Universität zu Berlin im Sommer 2000 eine zweite QuaSUM-Studie „Schulmonitoring und Schulentwicklung“ durchgeführt. Mit Ergebnissen kann im Frühjahr 2001 gerechnet werden. Das Ziel dieses zweiten QuaSUM-Projektes besteht darin, die Auseinandersetzung der QuaSUM-Schulen mit den Ergebnisrückmeldungen unter formalen, thematischen und handlungsorientierten Aspekten zu evaluieren und damit Rezeptionsstrategien sowie Folgewirkungen derartiger Ergebnisrückmeldungen für interne Schulevaluation und Unterrichtsverbesserungen festzustellen. Zusätzlich wurde im November 2000 eine qualitative Fallstudie an sechs Gymnasien begonnen. Mit diesem alternativen methodischen Verfahren sollen im Rahmen von Gruppengesprächen mit den beteiligten Lehrkräften einer Schule diejenigen Fragen genauer ausgeleuchtet werden, die aus den Fragebogendaten der Evaluationsstudie nicht eindeutig interpretiert werden können. Die Zusammenschau dieser unterschiedlichen Vorgehensweisen und der jeweils unterschiedlichen Ergebnisse soll in erster Linie den Schulen, aber auch der Bildungsverwaltung helfen, einen vertieftes und differenziertes Verständnis zu den Voraussetzungen und Bedingungen erfolgreichen Unterrichts zu erlangen. Das QuaSUM-Projekt kann damit einen eigenständigen Beitrag zu einem zu entwickelnden Gesamtkonzept zur Qualitätssicherung und -entwicklung leisten.

Hans-Jürgen Kuhn
Projektleiter

Danksagung

Das Ministerium für Bildung, Jugend und Sport des Landes Brandenburg hat im Januar 1999 eine wissenschaftliche Untersuchung zur Qualität des Mathematikunterrichts an Brandenburger Schulen in Auftrag gegeben, deren Ergebnisse mit diesem Bericht einer breiteren wissenschaftlichen und schulischen Öffentlichkeit bekannt gemacht werden.

Dank gilt allen, die die QuaSUM-Studie ermöglicht haben und daran mitgewirkt haben. Herr Kuhn als Projektleiter im Ministerium für Bildung, Jugend und Sport war in allen Phasen des Projekts ein kritischer und konstruktiver Partner, der durch umsichtige Organisation die Voraussetzungen dafür geschaffen hat, dass ein aussagekräftiger Datensatz gewonnen und ausgewertet werden konnte.

In allen mathematikdidaktischen Fragen haben Herr Dr. Bieber, Frau Ludwig, Herr Rädsch und Herr Dr. Steinberg vom Pädagogischen Landesinstitut des Landes Brandenburg (PLIB) kompetente Unterstützung geleistet.

Die Schülerinnen und Schüler haben fast ohne Ausnahme mit Eifer und Ernsthaftigkeit die Mühe auf sich genommen, in bis zu sechs Unterrichtsstunden die Tests und einen Fragebogen zu bearbeiten. Schulleitungen, Mathematiklehrkräfte und Eltern haben in großer Mehrheit durch aktive Mitarbeit, insbesondere durch ihre Antworten auf die zahlreich gestellten Fragen, die Studie unterstützt.

Mit hohem Engagement haben Schulleitungen und Lehrkräfte die eigentlichen Erhebungen in den Schulen durchgeführt. Sie wurden dabei von den besonders geschulten Multiplikatoren für Schuluntersuchungen inhaltlich und organisatorisch beraten.

Frau Dr. Husfeldt hat die Skalierung der Mathematik-Leistungsdaten durchgeführt. Zahlreiche studentische Hilfskräfte beteiligten sich mit großer Sorgfalt an der Bewältigung der umfangreichen technischen Arbeiten.

Ohne diese vielfältige Zusammenarbeit hätte die Studie nicht durchgeführt und erfolgreich abgeschlossen werden können.

Rainer H. Lehmann
wissenschaftlicher Leiter

Inhaltsverzeichnis

1 Ziele der Untersuchung

Das Ministerium für Bildung, Jugend und Sport des Landes Brandenburg hat im Februar 1999 die Durchführung und Auswertung der *Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM)* in Auftrag gegeben. Zentrale Anliegen dieser wissenschaftlichen Untersuchung sind die Erhebung von Lernständen am Ende der Klassenstufen 5 und 9 im Fach Mathematik sowie die Analyse von inner- und außerschulischen Bedingungen, die unterschiedliche Lernstände auf den Ebenen Schulformen bzw. Kursniveaus, Schulen innerhalb von Schulformen und Klassen bzw. Kursen innerhalb von Schulen erklären können.

Maßgeblich für die Untersuchung sind folgende Fragestellungen:

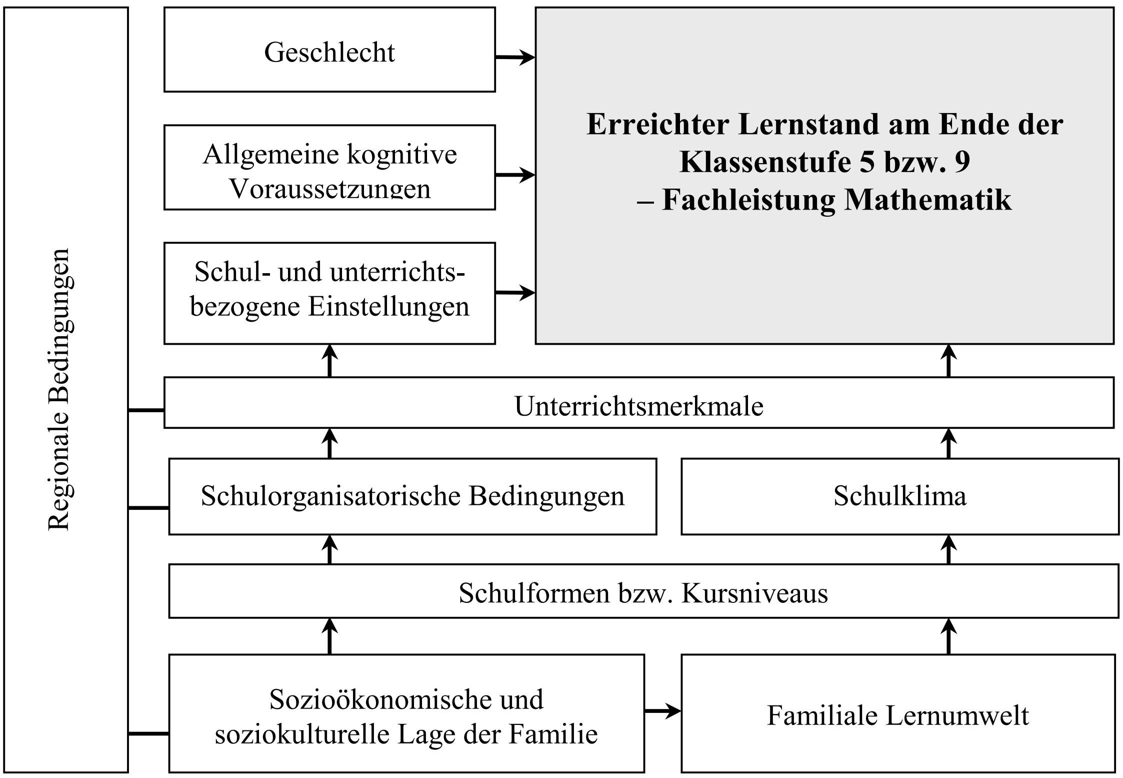
- Welche Lernstände sind im Land Brandenburg am Ende der Klassenstufen 5 und 9 im Fach Mathematik erreicht worden, differenziert nach Regionen, Schulformen (ggf. Kursniveaus), Schulen und Klassen bzw. Kursen?
- Welcher Zusammenhang besteht zwischen außerschulischen Faktoren wie allgemeinen kognitiven Voraussetzungen, familialen Merkmalen (z. B. Sozialstatus, Bildungsnähe) oder dem Geschlecht und den erhobenen Lernständen?
- Welcher Zusammenhang besteht zwischen schul- und unterrichtsbezogenen Einstellungen der Schülerinnen und Schüler und den gemessenen Mathematikleistungen?
- In welchem Zusammenhang stehen die erzielten Mathematikleistungen zu schulorganisatorischen und schulstrukturellen Merkmalen und insbesondere zu Merkmalen der Schul- und Unterrichtsqualität?
- In welchem Zusammenhang stehen schulische Beurteilungen (Noten, Leistungseinschätzungen der Lehrkräfte) sowie die bisherige Schullaufbahn der Schülerinnen und Schüler und ihre erzielten Mathematikleistungen?

Angesichts dessen, dass über die Qualität der Unterrichtsarbeit in Mathematik für Brandenburg bislang keine empirisch begründeten Aussagen gemacht werden können, hat die wissenschaftliche Untersuchung eine doppelte Zielstellung: Zum einen sollen abgesicherte Erkenntnisse über die tatsächliche Umsetzung der Rahmenpläne erzielt und dabei insbesondere die Leistungsheterogenität innerhalb von Klassen (Kursen), Schulen und Schulformen aufgedeckt werden. Darüber hinaus gilt es, Grunddaten zu den Bedingungen hoher Mathematikleistungen bereitzustellen und dabei den relativen Stellenwert schulischer bzw. unterrichtlicher Merkmale zu bestimmen.

Mit den Ergebnissen dieser Untersuchung soll dem Ministerium eine wesentlich verbesserte Grundlage für die Beratung und Unterstützung der Schulen und für die Weiterentwicklung der Rahmenpläne für Mathematik zur Verfügung gestellt werden. Darüber hinaus sollen die im Bericht mitgeteilten Ergebnisse für die pädagogische Arbeit in den Schulen genutzt werden: Schulen verfügen mit den Untersuchungsergebnissen über verlässliche Informationen zu den durchschnittlich erreichten Leistungsständen im Fach Mathematik und zu deren Streuung als zusätzliche Grundlage für die Planung und Durchführung des Unterrichts bzw. für die Reflexion ihrer bisherigen pädagogischen Arbeit.

Abbildung 1 zeigt in grafischer Form die Anlage der Studie. Zur Erklärung der im Zentrum der Untersuchung stehenden Fachleistungen in Mathematik werden schulische bzw. unterrichtliche, individuelle und familiale Merkmale berücksichtigt.

Abbildung 1 Rahmenmodell der Untersuchung



Den Zielsetzungen entsprechend ist der vorliegende Bericht aufgebaut: Im Abschnitt 2 werden die methodische Anlage und die Durchführung der Untersuchung vorgestellt. Abschnitt 3 enthält die Beschreibung der in Brandenburg ermittelten Mathematikleistungen – zunächst unter testanalytischen Gesichtspunkten, darauf aufbauend als Ergebnisse für Regionen, Schulformen (Kursniveaus), Schulen und Klassen (bzw. Kurse). Im Abschnitt 4 wird dann die Bedeutung außerunterrichtlicher bzw. außerschulischer Merkmale für die erzielten Mathematikleistungen beleuchtet. Abschnitt 5 geht analog den schulischen und unterrichtlichen Einflüssen nach und thematisiert hier zusätzlich die Zusammenhänge zwischen der Schullaufbahn sowie der diagnostischen Praxis von Lehrkräften und den Mathematikleistungen der Schülerinnen und Schüler. Schließlich wird im Abschnitt 6 – ausgehend vom in Abbildung 1 dargestellten Rahmenmodell der Untersuchung, das auch das familiale, schulische und unterrichtliche Umfeld der Schülerinnen und Schüler berücksichtigt – der relativen Bedeutung einzelner Bedingungsfaktoren für die erreichten Leistungen im Zusammenhang nachgegangen.

2 Anlage und Durchführung der Untersuchung

In diesem zweiten Abschnitt des Berichts werden die grundlegenden Konstruktionsprinzipien und der Aufbau der Erhebungsinstrumente vorgestellt (2.1), es werden Angaben zur Gewinnung und zur Struktur der Stichprobe gemacht (2.2), und es werden die Verfahren der Datenerhebung sowie die erzielten Teilnahme- bzw. Bearbeitungsquoten mitgeteilt (2.3).

2.1 Instrumentierung

Zur Erhebung der in Abschnitt 1 genannten Leitfragestellungen wurde eine Kombination verschiedener Tests und Fragebögen eingesetzt. Darüber hinaus wurden Angaben aus den amtlichen Schülerakten entnommen.

Testverfahren

Zur Bestimmung der Mathematikleistungen dienten insgesamt drei standardisierte Tests: Der *QuaSUM-Mathematiktest für die Klassenstufe 5* und der *QuaSUM-Mathematiktest für die Klassenstufe 9* sind speziell für diese Untersuchung entwickelt worden. Beim *Mathe40-Test* handelt sich sich um ausgewählte Testaufgaben aus der *Third International Mathematics and Science Study* (*TIMSS*/Population II; vgl. BEATON ET AL. 1996; BAUMERT ET AL. 1997), die zusätzlich zum QuaSUM-Test in der Klassenstufe 9 eingesetzt wurden.

Als erster Schritt zur Konstruktion der ***QuaSUM-Mathematiktests*** wurde am Pädagogischen Landesinstitut des Landes Brandenburg (PLIB) eine Datenbank angelegt, in der potenziell geeignete Testaufgaben systematisch gesammelt wurden. In die Datenbank wurden ausschließlich neu entwickelte Aufgaben von Brandenburger Mathematiklehrkräften aufgenommen, wobei sich die Aufgaben eng an den geltenden Rahmenplänen und der tatsächlichen Unterrichtspraxis in den Klassenstufen 5/6 bzw. 9/10 orientieren sollten. Nach einer vorläufigen Überprüfung der curricularen Validität der 647 Aufgaben für die Klassenstufe 5/6 und der 613 Aufgaben für die Klassenstufe 9/10 durch Expertinnen bzw. Experten im PLIB wurden sämtliche Aufgaben – ebenfalls von Brandenburger Mathematiklehrkräften – hinsichtlich ihrer Eignung und ihres Schwierigkeitsgrades für einzelne Schulformen eingeschätzt. Zur Validierung der Lehrereinschätzungen wurde in ca. 20 Brandenburger Klassen bzw. Kursen ein eingeschränkter Aufgabensatz von jeweils 40 Aufgaben pro Klassenstufe eingesetzt.

Auf der Grundlage der Erkenntnisse, die über die Lehrereinschätzungen gewonnen worden waren, wurde sodann ein Satz von 80 Aufgaben für die

Klassenstufe 5 und von 240 Aufgaben für die Klassenstufe 9 ausgewählt und in 16 fünften Klassen bzw. in 30 neunten Klassen pilotiert. In der Klassenstufe 9 wurden die Schulformen und Kursniveaus angemessen berücksichtigt. Dabei ergaben sich hinreichende Gründe für die Erwartung, dass bei Preisgabe teststatistisch weniger befriedigender Aufgaben für jede Klassenstufe ein von der Länge her zumutbarer, methodisch abgesicherter und aussagekräftiger Test würde entwickelt werden können.

Für die Hauptuntersuchung in der Klassenstufe 5 wurden schließlich 40 Aufgaben zusammengestellt, deren Schwierigkeitsindizes laut Lehrereinschätzung zwischen $p = 0{,}40$ und $p = 0{,}90$ lagen und die im Pilottest insgesamt eine durchschnittliche Testschwierigkeit von 50 Prozent und eine Reliabilität von $\alpha > 0{,}90$ aufwiesen. In der Klassenstufe 9 wurde unter der Vorgabe, schulformdifferenzierende Tests einzusetzen, prinzipiell das gleiche Verfahren gewählt. Ein Kernbereich von 40 Aufgaben, den alle Schülerinnen und Schüler bearbeiten sollten, wies schulformspezifisch Reliabilitäten von $\alpha = 0{,}86$ (Mathematik-Grundkurse an Gesamtschulen), $\alpha = 0{,}88$ (Mathematik-Erweiterungskurse an Gesamtschulen und Realschulklassen) und $\alpha = 0{,}80$ (Gymnasialklassen) auf.

Um die curricularen Anforderungen der unterschiedlichen Schulformen bzw. Kursniveaus in dem QuaSUM-Mathematiktest für die Klassenstufe 9 möglichst breit abzudecken und die Neuntklässlerinnen und Neuntklässler nicht ungebührlich zu belasten, erhielt jede Schülerin bzw. jeder Schüler über den gemeinsamen Kernbestand von 40 Aufgaben hinaus weitere 40 Aufgaben, die nach Schulform bzw. Kursniveau differenziert waren. So wurde eine Testform für die Mathematik-Grundkurse an Gesamtschulen entwickelt, eine zweite – gemeinsame – für die Erweiterungskurse an Gesamtschulen und die Realschulen und eine dritte für die Gymnasien. In allen Testversionen wurden jeweils die vier Stoffgebiete berücksichtigt, die hinsichtlich ihrer Anforderungen typische Mathematikaufgaben der entsprechenden Schul- bzw. Kursform darstellen. Tabelle 1 verdeutlicht das Design.

Dass den Schülerinnen und Schülern der Erweiterungskurse an Gesamtschulen dieselben Aufgaben wie den Realschulklassen, nicht aber Aufgaben der höchsten Schwierigkeitsstufe vorgelegt wurden, ergab sich aus den Befunden der Vorerprobungen, wonach zwischen diesen beiden Gruppen praktisch keine Leistungsunterschiede bestanden. So konnten negative Auswirkungen vermieden werden, die sich aus einer Überforderung durch zu schwierige Aufgaben ergeben hätten. Gleichzeitig war durch das Skalierungsverfahren sichergestellt, dass die Mathematikleistungen in den Erweiterungskursen nicht unterschätzt werden würden (vgl. hierzu Abschnitt 3.1).

Tabelle 1 Design des QuaSUM-Mathematiktests, Klassenstufe 9

Schwierigkeitsstufe 1	*Schwierigkeitsstufe 2*	*Schwierigkeitsstufe 3*	*Schwierigkeitsstufe 4*	*Schwierigkeitsstufe 5*
20 Aufgaben	20 Aufgaben	40 Aufgaben	20 Aufgaben	20 Aufgaben
Testform für G-Kurse an Gesamtschulen				
	Testform für Realschulklassen und für E-Kurse an Gesamtschulen			
		Testform für Gymnasialklassen		

Insgesamt weist der QuaSUM-Mathematiktest für die neunte Klassenstufe somit 120 Aufgaben auf. Die durchschnittlichen Testschwierigkeiten der unterschiedlichen Formen mit jeweils 80 Aufgaben (Kernbereich und schulformspezifische Erweiterungen) lagen nach den Pilotergebnissen für alle Schulformen bzw. Kursniveaus jeweils zwischen 50 und 60 Prozent. Der folgenden Tabelle 2 ist die Anzahl der in den Tests eingesetzten Aufgaben nach Stoffgebieten zu entnehmen.

Tabelle 2 Stoffgebiete und Aufgabenzahl der QuaSUM-Mathematiktests, Klassenstufen 5 und 9

Klassenstufe 5		**Klassenstufe 9**	
Stoffgebiet	Anzahl der Aufgaben	Stoffgebiet	Anzahl der Aufgaben
Zahlenbereiche/Rechnen	10	Funktionen	30
Größen	10	Gleichungen/Ungleichungen	30
Verhältnisgleichungen/ Proportionalität	10	Zahlen/Variablen	30
Geometrie	10	Geometrie	30
insgesamt	*40*	*insgesamt*	*120*

Die Aufgaben haben unterschiedliche Formate: Neben Aufgaben mit Mehrfachwahlantworten (Multiple-Choice-Items) wurden offene Aufgabenstellungen verwendet[1]. Die Bearbeitungszeit des *QuaSUM-Mathematiktest für die*

[1] In jeder Multiple-Choice-Aufgabe wurden fünf Anwortmöglichkeiten angeboten. Die

Klassenstufe 5 betrug 40 Minuten (10 Minuten pro Themenbereich), die des QuaSUM-Mathematiktest für die neunte Klassenstufe 80 Minuten (20 Minuten pro Themenbereich).

Zusammenfassend zeichnen sich beide QuaSUM-Mathematiktests dadurch aus, dass sie aufgrund der Autorenschaft von Brandenburger Lehrkräften eine deutliche Nähe zur tatsächlichen Unterrichtspraxis aufweisen, also das sog. implementierte Curriculum vergleichsweise gut abbilden dürften. Gleichzeitig ist durch den Einbezug von Fachdidaktikern in die Begutachtung der Aufgaben die Lehrplanvalidität weitgehend gewährleistet. Durch den Einsatz klassischer und probabilistischer Testanalysen konnte als notwendige Voraussetzung ein hohes Maß an interner Konsistenz der in der Hauptuntersuchung verwendeten Tests nachgewiesen werden.

Der ***Mathe40-Test*** enthält insgesamt 40 aus 158 Aufgaben, die 1996 im Rahmen der *Third International Mathematics and Science Study (TIMSS)* bei Siebt- und Achtklässlern im Bereich Mathematik eingesetzt wurden. Die Aufgaben entstammen sämtlich dem zur Veröffentlichung freigegebenen Aufgabensatz (vgl. BAUMERT ET AL. 1998) und wurden so ausgewählt, dass sie erstens möglichst breit die Themenbereiche von TIMSS abdecken und dass zweitens keine allzu engen Überschneidungen mit den QuaSUM-Aufgaben auftreten. Die folgende Tabelle 3 zeigt in der zweiten Spalte, wie sich die 40 Aufgaben des Mathe40-Tests auf die in TIMSS ausgewiesenen Stoffbereiche verteilen. In der dritten Spalte ist zusätzlich die Gesamtanzahl von Aufgaben des entsprechenden Stoffgebiets im TIMSS-Test angegeben. Die Verteilung der Aufgaben des Mathe40-Tests auf die verschiedenen Bereiche spiegelt also weitgehend die Proportionen im TIMSS-Test wider.

Bei den ausgewählten Aufgaben handelt es sich ausschließlich um Multiple-Choice-Items; die Bearbeitungszeit des Mathe40-Tests betrug 40 Minuten.

insgesamt fünf (Jahrgangsstufe 5) bzw. zehn (Jahrgangsstufe 9) unterschiedlichen offenen Aufgaben wurden „richtig"/„falsch" kodiert.

Tabelle 3 Verteilung der Mathe40-Aufgaben auf die TIMSS-Stoffgebiete

Stoffgebiet	**Aufgaben im Mathe40-Test**	**Aufgaben im TIMSS-Test**
Zahlen und Zahlenverständnis	14	52
Messen und Maßeinheiten	5	21
Algebra	6	29
Geometrie	5	23
Proportionaliät	2	12
Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung	8	21
insgesamt	*40*	*158*

Die Mathematik-Leistungstests wurden in beiden Klassenstufen durch den nonverbalen ***Culture Fair Intelligence Test – CFT 20*** in seiner Kurzform ergänzt. Dieser Test zielt auf die Messung der sog. „flüssigen Grundintelligenz" und erfasst Aspekte des schlussfolgernden Denkens, also allgemeine kognitive Lernvoraussetzungen. Es handelt sich bei diesem Verfahren um einen bundesweit geeichten Test, was den Vergleich mit Altersnorm- bzw. Klassenstandardwerten ermöglicht (vgl. WEIß 1997; siehe auch Abschnitt 4.1). Die Aufgaben des CFT 20 sind in vier Subtests (Reihen fortsetzen; Klassifikationen; Matrizen; topologische Schlussfolgerungen) untergliedert. Die einzelnen Subtests bestehen aus sprachfreien, in zeichnerischer Form dargestellten und nach Schwierigkeiten geordneten Einzelaufgaben im Multiple-Choice-Format. Seine anschaulich-figurale Gestaltung macht diesen standardisierten Test vom Grad der Beherrschung der deutschen Sprache sowie von fachgebundenem Wissen und Können weitgehend unabhängig und gestattet es, Diskrepanzen zwischen dem kognitiven Potenzial einer Schülerin bzw. eines Schülers und ihren oder seinen fachgebundenen Fähigkeiten zu erkennen. Die Durchführung des CFT 20 in seiner Kurzform mit insgesamt 46 Aufgaben erforderte ca. 35 Minuten.

Befragungsinstrumente

Mit Hilfe eines ***Schülerfragebogens*** wurden schul- und unterrichtsbezogene Einstellungen der Schülerinnen und Schüler untersucht. Der Schülerfragebogen setzt sich aus Items bzw. Skalen zusammen, die bereits in anderen Studien zur Messung schul- und unterrichtsbezogener Einstellungen eingesetzt wurden und bei denen teilweise ein Zusammenhang zur Mathematikleistung angenommen werden konnte (vgl. LEHMANN & PEEK 1997; LEHMANN, GÄNSFUß & PEEK 1999; BAUMERT ET AL. 1997; ROEDER & BAUMERT 1997; MEL-

ZER & STENKE 1996). Das Instrument richtet sich sowohl an die Fünft- als auch an die Neuntklässler und enthält insgesamt 85 Aussagen zu den Themenbereichen *Schulzufriedenheit*, *leistungsbezogenes Selbstkonzept*, *Sach- und Fachinteresse an Mathematik*, *Anstrengung als Beitrag zum Schulerfolg* und *Wahrnehmung des eigenen Mathematikunterrichts*, die jeweils auf einer vierstufigen Skala eingeschätzt werden sollten.

Der eingesetzte ***Elternfragebogen*** bezieht sich auf die außerschulischen Lebens- und Lernwelten der Schülerinnen und Schüler. Er umfasst insgesamt 26 geschlossene Fragen vor allem zur Bildungsnähe des Elternhauses, zum sozioökonomischen Hintergrund der Familie und zur außerschulischen Alltagsgestaltung der Schülerinnen und Schüler. Einen weiteren Schwerpunkt des Elternfragebogens bilden Aussagen zu den Leistungserwartungen der Eltern an ihre Kinder und Erwartungen an ihre Schullaufbahn. Durch diese Informationen kann der erreichte Lernstand der Schülerinnen und Schüler im Fach Mathematik auf ihr jeweiliges soziales Umfeld rückbezogen werden.

Neben dem Schülerfragebogen und dem Elternfragebogen wurden Fragebögen für die Mathematiklehrkräfte und die Schulleitungen eingesetzt. Der ***Fragebogen für Mathematiklehrkräfte*** beinhaltet drei Fragenblöcke: Der erste Block bezieht sich auf die Unterrichtssituation in der getesteten Klasse bzw. in dem Testkurs sowie auf die Zielsetzungen und Methoden im Mathematikunterricht der Klasse bzw. des Kurses. Der zweite Frageblock erhebt allgemeine Einschätzungen von Aspekten des Mathematikunterrichts. Der abschließende Frageblock erfasst berufsbiografische Angaben. Der Aufbau und die Inhalte des Fragebogens orientieren sich an Vorlagen, die bereits in vorherigen Untersuchungen zur Schulleistung erfolgreich eingesetzt wurden (vgl. z. B. LEHMANN ET AL. 1995; BAUMERT ET AL. 1997). Darüber hinaus wurden die Mathematiklehrkräfte im Rahmen eines ***Ratingbogens*** gebeten, die einzelnen Aufgaben aus den Leistungstests hinsichtlich ihrer Schwierigkeit für die eigene Klasse bzw. den eigenen Kurs einzuschätzen, d. h. Angaben zur vermuteten Lösungshäufigkeit der Aufgaben in ihrer Klasse bzw. in ihrem Kurs vorzunehmen.

Der ***Fragebogen für Schulleiterinnen und Schulleiter*** erhebt in einem ersten Bereich allgemeine Angaben zum schulischen Umfeld wie z. B. eine Einschätzung zur sozialen Lage im Einzugsbereich der Schule. Der zweite Bereich erfasst strukturelle Rahmendaten der Schule wie z. B. die materielle und personelle Ausstattung. Der dritte Bereich bezieht sich auf Einschätzungen von Aspekten des Schulalltags, z. B. der Leistungsbereitschaft der Schülerinnen und Schüler und des Arbeitsklimas unter den Lehrkräften. Im vierten

Bereich wird das berufliche Selbstkonzept der Schulleitungen sowie ihre Einschätzung von Unterrichts- und Erziehungszielen erfragt. Dem Fragebogen liegen ebenfalls erprobte Vorlagen aus anderen Schulleistungsuntersuchungen zugrunde (vgl. LEHMANN ET AL. 1995; BAUMERT ET AL. 1997).

Schülerakte

Als Quelle amtlich verfügbarer Informationen wurde die in den Schulen vorhandene Schülerakte genutzt. Neben einigen wenigen Angaben zur persönlichen Situation der Schülerinnen und Schüler (Alter, Geschlecht, Wohnbezirk) geht es in erster Linie um Aspekte des bisherigen Lernerfolgs und der Schullaufbahn: Wiederholen bzw. Überspringen einer Klasse, die Noten in den letzten Zeugnissen und – für die Gesamtschülerinnen und Gesamtschüler – die Zugehörigkeit zu Grund- der Erweiterungskursen in den Fächern Mathematik, Deutsch und Englisch.

2.2 Stichprobe

Der *Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM)* liegen Daten zu den unter 2.1 aufgezeigten Instrumenten für eine repräsentative Stichprobe von Brandenburger Fünft- und Neuntklässlern aus staatlichen Regelschulen zugrunde. Zusätzlich zu einer Zufallsauswahl aus den regulären Schulen wurden alle Fünft- und Neuntklässler der Schulen aus Schul- und Modellversuchen und der Schulen mit besonderer Prägung mit in die Untersuchung aufgenommen.

Zur Bestimmung der Fünftklässler-Stichprobe wurde folgendes Verfahren gewählt: Aus der Grundgesamtheit von 723 eigenständigen Grundschulen bzw. Gesamtschulen mit angegliedertem Grundschulbereich[2] im Schuljahr 1998/99 wurden zunächst die Förderschulen, die Schulen in privater Trägerschaft und diejenigen Schulen ausgeschlossen, die als Schulen mit besonderer Prägung oder als Schulen aus Modellversuchen über die Zufallsstichprobe hinaus in die Untersuchung einbezogen werden sollten[3]. Die so definierte Zielpopulation wurde in einem zweiten Schritt in eine Teilpopulation „Grundschulen" (543 Schulen mit 1.338 Klassen und 30.860 Schülerinnen

[2] Im Land Brandenburg werden in Grundschulen Schülerinnen und Schüler der Jahrgangsstufen 1 bis 6 unterrichtet. Gesamtschulen mit angegliedertem Grundschulbereich umfassen – unter einer gemeinsamen Schulleitung – Grundschulklassen bis zur Jahrgangsstufe 6 und daran anschließend einen Sekundarbereich (Jahrgangsstufen 7 bis 10), der als Gesamtschule geführt wird.

[3] Hierzu gehören die Schulen des Landes-Modellversuchs „Kleine Grundschule" sowie die Schulversuche „Jenaplanschule" und „Montessori-Pädagogik in der Primarstufe".

und Schüler) und eine Teilpopulation „Gesamtschulen mit angegliedertem Grundschulbereich” (85 Schulen mit 184 Klassen und 4.141 Schülerinnen und Schülern) unterteilt. Unter der Maßgabe, mindestens 2.000 Fünftklässlerinnen und Fünftklässler in die Untersuchung einzubeziehen, wurden anhand einer nach Landkreisen und Schulformen geordneten Aufstellung aller Brandenburger allgemeinbildenden staatlichen Regelschulen nach dem „Random-Start-Equal-Steps“-Verfahren 34 Grundschulen gezogen. Die zusätzliche Ziehung der Gesamtschulen mit angegliedertem Grundschulbereich stellte sich so dar: In der vorliegenden Teilstichprobe „Gesamtschulen, Klassenstufe 9” (siehe unten) mit fünften Klassen wurden die Schulen mit angegliedertem Grundschulbereich automatisch auch in die Stichprobe der Klassenstufe 5 aufgenommen (17 Schulen mit 37 Klassen und 790 Schülerinnen und Schülern). Aus den verbleibenden 68 Gesamtschulen mit angegliedertem Grundschulbereich wurden wiederum unter Anwendung des „Random-Start-Equal-Steps“-Verfahrens weitere 5 Gesamtschulen (11 Klassen mit 240 Schülerinnen und Schülern) gezogen.

Die Ziehung der Stichprobe für die Klassenstufe 9 folgte prinzipiell dem gleichen Muster, nun aber unter Berücksichtigung der Aufgliederung der Schülerschaft nach Schulformen: Die Zufallsstichprobe von Schulen[4] mit Schülerinnen und Schülern in neunten Klassen wurde wiederum anhand einer nach Landkreisen und Schulformen geordneten Aufstellung aller Brandenburger allgemeinbildenden staatlichen Regelschulen mit neunten Klassen gezogen. Diese Auflistung – exklusive Förderschulen, Schulen in privater Trägerschaft, Schulen des zweiten Bildungsweges und Schulen, die obligatorisch untersucht werden sollten (Schulen des Modellversuchs SINUS sowie die Schulen mit besonderer Prägung) – ergibt bei insgesamt 420 Schulen mit 1.352 Klassen und 34.782 Schülerinnen und Schülern im Schuljahr 1998/99 eine nach Schulformen geschichtete Grundgesamtheit („definierte Zielgruppe”) von 95 Gymnasien (356 Klassen mit 9.684 Schülerinnen und Schülern), 74 Realschulen (213 Klassen mit 5.587 Schülerinnen und Schülern) und 251 Gesamtschulen (783 Klassen mit 19.511 Schülerinnen und Schülern). Auf der Ebene der Gesamtschulen wurden abgesicherte Ergebnisse sowohl für Schülerinnen und Schüler aus Mathematik-Grundkursen als auch für solche aus Mathematik-Erweiterungskursen angestrebt.

---

4 Die Begründung für eine Stichprobenziehung auf Schul- und nicht auf Klassen- oder Schülerebene liegt in der zusätzlichen zentralen Zielsetzung des Projekts, den beteiligten Schulen anhand vergleichender klassen- bzw. kursbezogener Ergebnisrückmeldungen Informationen für die schulinterne Evaluation an die Hand zu geben (vgl. dazu MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT DES LANDES BRANDENBURG 1999a).

Entsprechend der vorab bestimmten Stichprobengröße von mindestens 2.000 Schülerinnen und Schülern für jede der vier Untersuchungsgruppen[5] wurden schließlich 25 Gymnasien, 25 Realschulen und 50 Gesamtschulen mit sämtlichen Schülerinnen und Schülern in neunten Klassen gezogen. Hierzu wurde schulformspezifisch das „Random-Start-Equal-Steps"-Verfahren angewandt. Die folgende Tabelle 4 gibt schulformspezifisch an, wie viele Schulen, Klassen und Schülerinnen bzw. Schüler Aufnahme in die Stichprobe fanden[6].

Tabelle 4 Datenstruktur der Erhebungen: Anzahl der Schulen, Klassen sowie Schülerinnen und Schüler, nach Schulform

	Schulen	**Klassen**	**Schüler**
Klassenstufe 5			
Grundschulen	44	99	2.249
Gesamtschulen mit angegliedertem Grundschulbereich	22	49	1.017
insgesamt	*66*	*148*	*3.266*
Klassenstufe 9			
Gesamtschulen	59	188	4.544
Realschulen	27	78	2.030
Gymnasien	28	105	2.780
insgesamt	*114*	*371*	*9.354*

Die nachstehende Tabelle 5 zeigt für die einzelnen Klassenstufen bzw. Schulformen innerhalb der Stichprobe Durchschnittsangaben zu der Anzahl fünfter bzw. neunter Klassen pro Schule, von Schülerinnen bzw. Schüler pro Klasse und von Schülerinnen bzw. Schülern pro Schule. Die Gegenüberstellung mit entsprechenden Durchschnittswerten für die Grundgesamtheit der Brandenburger Schulen ergibt eine deutliche Übereinstimmung der Kennwerte; die

5 Diese Stichprobengröße gewährleistet einen Stichprobenfehler (95-Prozent-Konfidenzintervall) in der Größe von maximal einem Zehntel Standardabweichung.

6 Nach der Ziehung der Stichprobe wurden die ausgewählten Schulen angeschrieben, um die Teilnahme zu bestätigen. Es fielen wenige Schulen aus, wenn die komplette neunte bzw. fünfte Klassenstufe während des zweiwöchigen Test- und Nachtestzeitraumes im Betriebspraktikum oder auf Klassenfahrt war. Für diese Schulen (ein Gymnasium, eine Real- und eine Gesamtschule) wurden die in der Auflistung jeweils folgenden Schulen nachgezogen. Eine ausfallende Grundschule wurde nicht ersetzt, da für sie als Modellversuch „Kleine Grundschule" keine Schule gleichen Typs als Ersatz zur Verfügung stand.

Grundgesamtheit wird durch die geschichteten Zufallsstichproben in allen Fällen sehr gut abgebildet.

Tabelle 5 Durchschnittswerte (Anzahl der Klassen, Schul- und Klassengröße) in der Grundgesamtheit und in der Stichprobe, nach Schulform

	Grundgesamtheit			**Stichprobe**		
	Klassen/ Schule	**Schüler/ Klasse**	**Schüler/ Schule**	**Klassen/ Schule**	**Schüler/ Klasse**	**Schüler/ Schule**
Klassenstufe 5						
Grundschulen	2,5	23,1	57,4	2,5	22,8	57,7
Gesamtschulen mit angegliedertem Grundschulbereich	2,2	22,5	48,7	2,2	21,5	46,8
Klassenstufe 9						
Gesamtschulen	3,1	24,9	78,0	3,0	24,6	73,9
Realschulen	2,9	26,3	76,3	2,9	26,6	77,6
Gymnasien	3,7	27,2	101,9	3,8	27,5	104,6

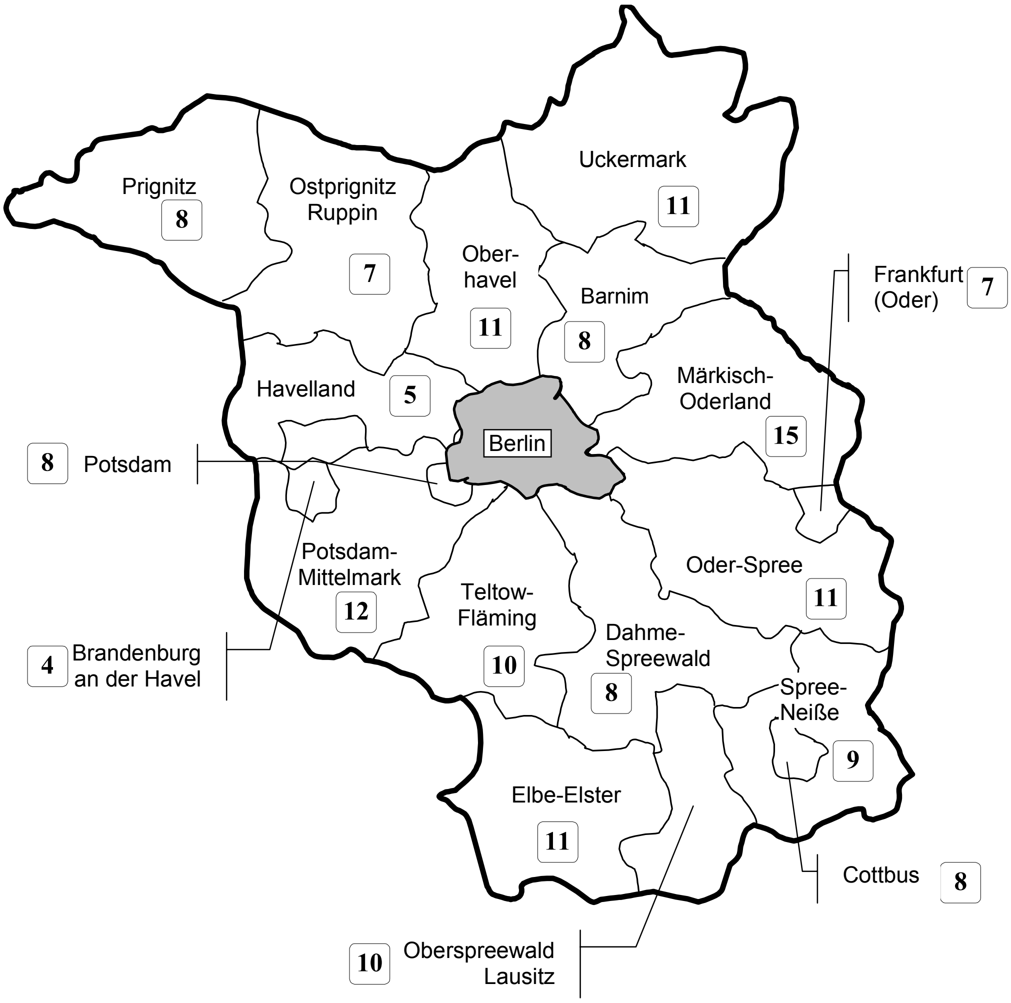
Eine möglichst gleichmäßige Berücksichtigung der Landkreise bei der Auswahl der Schulen wurde durch das Ziehungsverfahren ebenfalls gut gewährleistet (vgl. dazu Abbildung 2).

Insofern die Stichprobe die Grundgesamtheit unter den Aspekten „Verteilung der Schulen nach Schulformen“ und „vollständige Einbeziehung von Schulen aus Schul- und Modellversuchen sowie Schulen mit besonderer Prägung“ nicht genau abbildet, wurde der Datensatz entsprechend der tatsächlichen Verteilungen im Land Brandenburg gewichtet[7].

Die in diesem Bericht formulierten Aussagen über erreichte Lernstände in Mathematik erfolgen somit auf der Grundlage einer repräsentativen Zufallsstichprobe, deren Struktur in angebbaren Grenzen präzise Aussagen über die Brandenburger Schülerschaft in fünften und neunten Klassen – insgesamt und schulformspezifisch – erlaubt.

7 Aufgrund eines Übermittlungsfehlers wurde anstatt einer zusätzlichen Schule im Schulversuch eine nach normalen Lehrplänen unterrichtende Grundschule in das Oversampling aufgenommen. Da diese Schule nicht über die Zufallsstichprobe in die Stichprobe gelangt ist, wurde sie hinsichtlich der Gesamtstichprobe wie die Schulen aus Modell- bzw. Schulversuchen gewichtet.

Abbildung 2 Verteilung der Stichprobenschulen (N = 163) nach Landkreisen sowie kreisfreien Städten (in Absolutwerten)



2.3 Datenerhebung und Datenaufbereitung

Die Information der beteiligten Schulen über die Ziele und Verfahren der Untersuchung und die Organisation der Datenerhebung (Versorgung der Schulen mit Test- und Befragungsmaterialien; Beratung der Testleiterinnen und Testleiter in den Schulen) erfolgte durch sog. Multiplikatoren für Schulqualitätsuntersuchungen – Brandenburger Lehrkräfte, die für diese Aufgabe besonders geschult wurden und als Ansprechpartner die beteiligten Schulen vor und während der Erhebung unterstützten.

Jede Schule benannte einen sog. Schulkoordinator, der für die sachgemäße Durchführung der Erhebung innerhalb der Schule verantwortlich war. Die Tests und die Befragungen der Schülerinnen und Schüler wurden von Lehrkräften der beteiligten Schulen nach standardisierten Anleitungen durchgeführt, wobei die Lehrkräfte nicht selbst in der Klasse unterrichten sollten. Die Datenerhebungen erfolgten innerhalb eines festen Zeitfensters zwischen dem 14. und 25. Juni 1999 nach einem vorgegebenen Testplan. Die Tests und die Schülerbefragung wurden im Klassenverband während der Unterrichtszeit durchgeführt. Der Elternfragebogen wurde von den Schülerinnen und Schülern mit nach Hause genommen und dort von ihnen gemeinsam mit den Eltern ausgefüllt. Die Bearbeitung der Fragebögen bzw. Ratings durch die Lehrkräfte und die Schulleitungen erfolgte außerhalb der normalen Unterrichtszeit.

Die nachfolgende Tabelle zeigt die Teilnahmequoten für die Bestandteile der Untersuchung, die sich unmittelbar an die Schülerinnen und Schüler bzw. an ihre Eltern richteten. Die Bearbeitung der Mathematiktests war verbindlich. Die Beantwortung des Schülerfragebogens und die Bearbeitung des CFT 20 war davon abhängig, ob die Erziehungsberechtigten vorab schriftlich ihr Einverständnis gegeben hatten. Die Bearbeitung des Elternfragebogens war freiwillig.

Tabelle 6 Bearbeitungsquote der Instrumente (Schülerinnen und Schüler, Eltern), nach Schulform (in Prozent)

	QuaSUM-Mathematik-Test	**Mathe40-Test**	**Schüler-fragebogen**	**CFT 20**	**Eltern-fragebogen**
Klassenstufe 5					
Grundschulen	96,4	---	87,2	86,5	90,8
Gesamtschulen mit angegliedertem Grundschulbereich	97,1	---	90,0	86,9	91,1
insgesamt	*96,6*	---	*88,1*	*86,7*	*90,9*
Klassenstufe 9					
Gesamtschulen, Grundkurse	91,7	92,2	77,1	74,0	74,7
Gesamtschulen, Erweiterungskurse	97,5	98,0	89,1	87,7	87,2
Realschulen	97,5	97,8	92,7	92,0	94,8
Gymnasien	97,2	98,3	95,3	94,6	89,5
insgesamt	*95,9*	*96,6*	*88,7*	*87,1*	*86,8*

Bei der Bearbeitung der Fragebögen für Mathematiklehrkräfte bzw. Schulleiter gab es nur geringfügige Ausfälle. Die Rückmeldung der Einschätzbögen bzw. Ratings zur Schwierigkeit der Mathematikaufgaben für die eigene Klasse war deutlich eingeschränkter.

Tabelle 7 Bearbeitungsquote der Instrumente (Mathematiklehrkräfte, Schulleitungen), nach Schulform (in Prozent)

	Fragebogen Lehrkräfte	**Fragebogen Schulleitungen**	**Ratings Lehrkräfte**
Klassenstufe 5			
Grundschulen	100,0	96,0	66,7
Gesamtschulen mit angegliedertem Grundschulbereich	100,0	95,9	63,3
insgesamt	*100,0*	*95,9*	*65,5*
Klassenstufe 9			
Gesamtschulen, Grundkurse	94,2	100,0	51,1
Gesamtschulen, Erweiterungskurse	92,9	100,0	54,5
Realschulen	100,0	100,0	57,7
Gymnasien	98,1	100,0	68,6
insgesamt	*95,9*	*100,0*	*57,4*

Die insgesamt hohen Teilnahmequoten gerade bei den freigestellten Testteilen spricht für eine hohe Akzeptanz der Studie unter den Beteiligten. Die vergleichsweise geringe Teilnahmequote an den Ratings liegt vermutlich daran, dass diese zeitlich verschoben drei Monate nach der Testdurchführung und den Befragungen durchgeführt wurden.

Die Konstruktion bzw. Zusammenstellung der Erhebungsinstrumente, die Datenerfassung, -aufbereitung und -auswertung, der Zwischenbericht (vgl. LEHMANN ET AL. 1999) und der vorliegende Endbericht über die Untersuchung waren Aufgabe der wissenschaftlichen Forschungsgruppe an der Humboldt-Universität zu Berlin[8].

[8] Mit Rückmeldungen von klassen- bzw. kursbezogenen Ergebnissen an die Schulen wurde – unter inhaltlicher Verantwortung der Forschungsgruppe an der Humboldt-Universität zu Berlin – ein auf EDV-Auftragsarbeiten spezialisiertes Unternehmen (Firma research & value, Berlin) beauftragt. Diese Rückmeldungen wurden im Dezember 1999 zusammen mit dem Zwischenbericht an die beteiligten Schulen versandt.

3 Brandenburger Schülerinnen und Schüler und ihre Fachleistungen in Mathematik

Das Anliegen dieses dritten Abschnitts ist es, die mathematischen Fachleistungen bei der repräsentativen Stichprobe Brandenburger Fünft- und Neuntklässler zu beschreiben, wie sie mittels des QuaSUM-Mathematiktests und des Mathe40-Tests gemessen wurden. Dabei konzentrieren sich die Ausführungen in einem ersten Teil auf methodische Grundlagen der Testanalyse und auf die Berichterstattung über Verteilungen von Fähigkeitsniveaus innerhalb der Stichprobe (3.1). Daran anschließend werden die erreichten Lernstände bzw. Fähigkeitsniveaus in regionalen, schulform-, schul- und klassen- bzw. kursspezifischen Zusammenhängen beleuchtet (3.2). Hier werden im Rahmen der Berichterstattung über die Ergebnisse im Mathe40-Test zusätzlich Vergleichswerte aus der TIMS-Studie vorgestellt und diskutiert.

3.1 Analyse der Leistungstests

Der QuaSUM-Mathematiktest für die Klassenstufe 5 umfasst insgesamt 40 Aufgaben. Die jeweils 10 Aufgaben zu den Themenbereichen *Zahlenbereiche/Rechnen*, *Größen*, *Verhältnisgleichungen/Proportionalität* und *Geometrie* beziehen sich auf den Lehrstoff der Klassenstufen 5 und 6. Sämtliche beteiligten Fünftklässlerinnen und Fünftklässler haben dieselben Aufgaben im Rahmen einer Unterrichtsstunde bearbeitet.

Dem QuaSUM-Mathematiktest für die Klassenstufe 9 liegt das Konzept zugrunde, schulform- bzw. kursniveauspezifische Anforderungsprofile des Mathematikunterrichts angemessen zu berücksichtigen: Die insgesamt 120 Aufgaben entstammen dem Lehrstoff der Klassenstufen 9 und 10 in den Themenbereichen *Funktionen*, *Gleichungen/Ungleichungen*, *Zahlen/Variablen* und *Geometrie*. Die drei unterschiedlichen Testversionen für die Neuntklässlerinnen und Neuntklässler aus Mathematik-Grundkursen an Gesamtschulen, aus Mathematik-Erweiterungskursen an Gesamtschulen sowie aus Realschulen und aus Gymnasien sind durch die 40 gemeinsamen Aufgaben, die hier als sog. Ankeritems fungieren, miteinander verbunden.

Es war das Ziel, für alle Schülerinnen und Schüler der einzelnen Klassenstufen auf der Grundlage jeweils *aller* zur Verfügung stehenden Aufgabenlösungen das Leistungsniveau im Fach Mathematik insgesamt zu bestimmen. Im Rahmen der probabilistischen Testtheorie (Item Response Theory; vgl. FISCHER & MOLENAAR 1995; HAMBLETON, SWAMINATHAN & ROGERS 1991) wurden die Testaufgaben unter Nutzung des einparametrigen Rasch-Modells skaliert.

Diese Technik erlaubt es,

- die Einschlägigkeit von Testaufgaben für die theoretisch bestimmte Fähigkeitsdimension empirisch zu prüfen und ungeeignete Items auszusondern[9],
- die Schwierigkeit von Testaufgaben und die Fähigkeit von Personen auf demselben Maßstab abzubilden und
- die Fähigkeit einer Person zuverlässig zu schätzen, auch wenn nur eine Teilmenge der Aufgaben bearbeitet wurde.

Bei dem Skalierungsverfahren wird die Fähigkeit einer Schülerin bzw. eines Schülers über Wahrscheinlichkeiten definiert. Dabei steigt die geschätzte Fähigkeit mit der Gesamtzahl der jeweils richtigen Antworten, während umgekehrt die Aufgabenschwierigkeit von der Anzahl richtiger Lösungen in der Stichprobe abhängt. Wenn sich nun die Antwortmuster der Schülerinnen und Schüler in guter Näherung so darstellen lassen, dass die Aufgaben stets nur bis zu einer individuell bestimmten Schwierigkeit gelöst werden, können für alle Aufgaben und Fähigkeitsgruppen die gesuchten Lösungswahrscheinlichkeiten berechnet werden. Dabei wird vorausgesetzt, dass die Kombination von Fähigkeiten, die erforderlich ist, den Test insgesamt zu bearbeiten, auch für die Schwierigkeit der einzelnen Aufgaben maßgeblich ist. Insofern die Aufgabenschwierigkeiten und die Schülerfähigkeiten innerhalb des Rasch-Modells auf dieselbe numerische Skala projiziert werden können, lassen sich die Aufgaben eines bestimmten Schwierigkeitsgrades als diejenige Leistung interpretieren, die von einer Schülerin bzw. einem Schüler des entsprechenden Fähigkeitsniveaus mit einer 50-prozentigen Wahrscheinlichkeit[10] noch gemeistert wird. Übersteigt die auf solcher Grundlage geschätzte Fähigkeit einer Schülerin oder eines Schülers die Schwierigkeit einer Aufgabe, so wird sie bzw. er diese wahrscheinlich lösen, und zwar mit um so höherer Wahrscheinlichkeit, je größer die Differenz ist.

---

[9] Sämtliche eingesetzten QuaSUM-Aufgaben entsprechen den teststatistischen Anforderungen, sodass der gesamte Aufgabensatz für die weiteren Auswertungen genutzt werden konnte.

[10] Die Festlegung einer 50-prozentigen Lösungswahrscheinlichkeit beruht auf einer a priori getroffenen Entscheidung, die voraussetzungsfrei ist. Mit der Wahl eines strengeren Maßstabs (vgl. z. B. die in der TIMS-Studie bestimmte Vorgabe einer 65-prozentigen Lösungswahrscheinlichkeit, bei der eine Aufgabe einem bestimmten Fähigkeitsniveau zugeordnet wird) verringert sich schülerseitig rein formal das entsprechend definierte Fähigkeitsniveau, ohne dass sich die Relationen innerhalb der Personengruppen und zwischen den Personengruppen ändern.

Die Ergebnisse der nach dem Rasch-Modell ausgewerteten QuaSUM-Mathematiktests wurden für die weitere Darstellung linear so transformiert, dass die durchschnittliche Schülerleistung auf 150 und die Standardabweichung auf 25 gesetzt sind[11]. Für die Berechnung der Ergebnisse im Mathe40-Test wurden – analog zu den Auswertungen der TIMS-Studie (vgl. BAUMERT ET AL. 1997), deren Ergebnisse bei der Interpretation vergleichend berücksichtigt werden sollen – die Zuordnungsregeln für eine 65-prozentige Lösungswahrscheinlichkeit und die internationalen Schwierigkeitsparameter (internationaler Mittelwert 500, Standardabweichung 100) übernommen.

3.1.1 QuaSUM-Mathematiktest – Klassenstufe 5

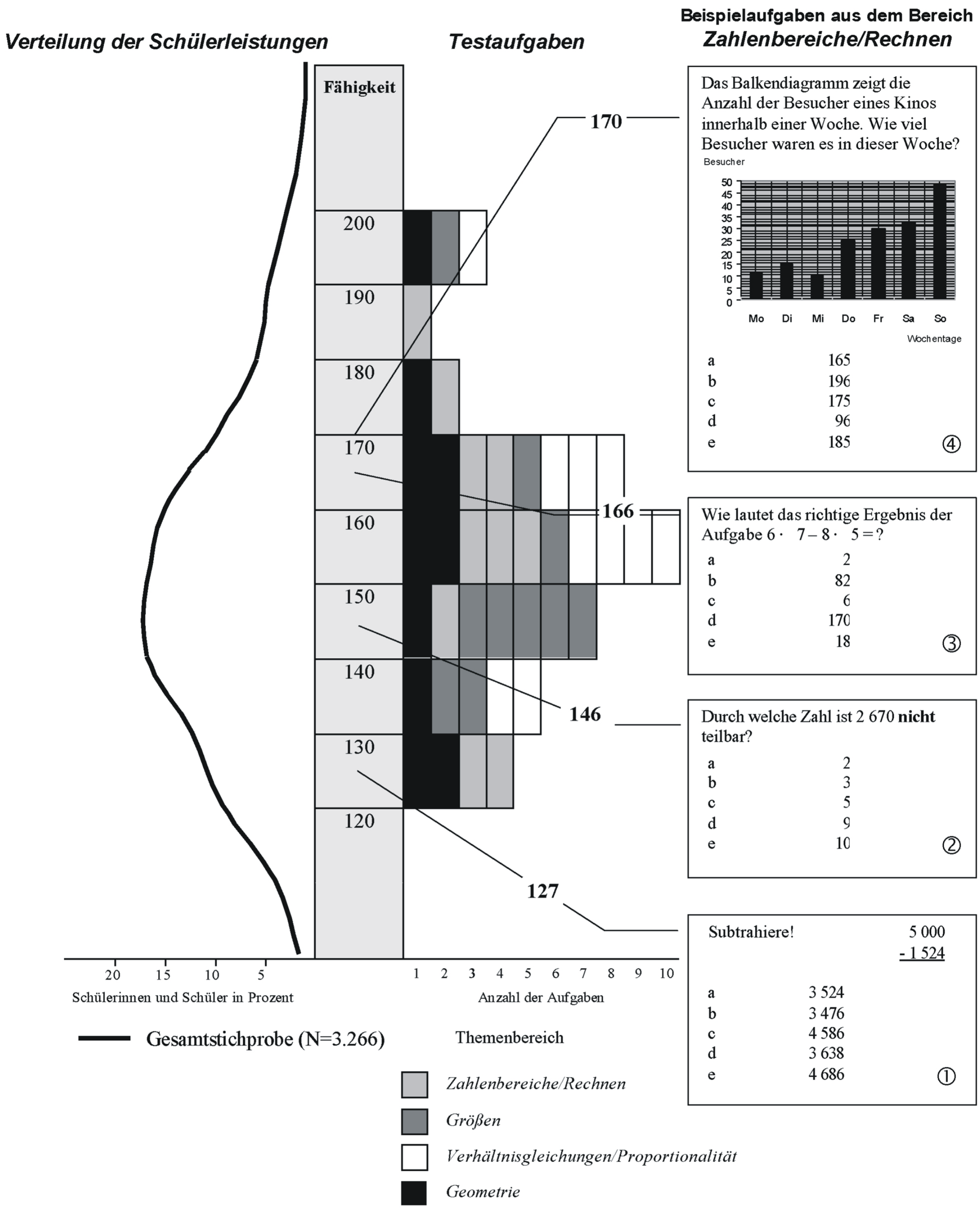
In der genannten Weise sind die Eigenschaften der QuaSUM-Mathematiktests für die Klassenstufen 5 bzw. 9 und die mathematischen Fähigkeiten der getesteten Schülerinnen und Schüler miteinander verschränkt und aufeinander bezogen.

Die Gesamtskala für die Klassenstufe 5 umgreift auf der Testseite unterschiedliche Stoffgebiete bzw. Themenbereiche, auf der Fähigkeitsseite unterschiedliche Leistungsdispositionen. Es ist deshalb sinnvoll (und möglich), die Verteilung der Aufgabenschwierigkeiten derjenigen der gemessenen mathematischen Fähigkeiten grafisch gegenüberzustellen und dabei die verschiedenen Themenbereiche kenntlich zu machen.

Abbildung 3 zeigt die Ergebnisse im QuaSUM-Mathematiktest für die Klassenstufe 5.

[11] Die Wahl der Transformationsregel unterliegt im Rahmen der Rasch-Skalierung einzig Gesichtspunkten der Zweckmäßigkeit. Die Standardabweichung ist der übliche Kennwert für die Streuung eines Merkmals. Je größer die Standardabweichung einer Verteilung ist, desto weiter liegen die Einzelwerte auseinander; je geringer sie ist, desto enger scharen sie sich um den Mittelwert. Die bei der Rasch-Skalierung zunächst nicht festliegenden „Maßeinheiten“ werden üblicherweise unter Bezug auf dieses Maß gewählt, und zwar so, dass die einzelnen Fähigkeitsparameter im Rahmen der gegebenen Messgenauigkeit als ganze positive Zahlen mit „rundem Mittelwert“ dargestellt werden können, ohne dass Verwechslungen mit anderen Skalen auftreten können.

Abbildung 3 Verteilung der Schülerleistungen im Fach Mathematik (Klassenstufe 5) im Vergleich mit den Schwierigkeiten der Testaufgaben im QuaSUM-Mathematiktest (nach Themenbereichen)



Auf der rechten Seite der Fähigkeitsskala sind die 40 Testaufgaben nach ihrer Schwierigkeit geordnet eingetragen. So liegen z. B. fünf Aufgaben zu *Geometrie*, *Größen* und *Verhältnisgleichungen/Proportionalität* mit niedrigerer Schwierigkeit im Skalenbereich 130 bis 140, während sich eine Aufgabe zu *Zahlenbereiche/Rechnen* mit relativ hoher Schwierigkeit im Skalenbereich 180 bis 190 befindet. Auf der Seite der Aufgabenschwierigkeiten ist eine Gleichverteilung ideal, wenn über den gesamten Leistungsbereich hinweg mit gleicher Präzision gemessen werden soll. Die Vielzahl der Faktoren, die letztlich die Aufgabenschwierigkeit mitbestimmen (Komplexität der Aufgaben, Plausibilität der angegebenen Alternativantworten, Einbeziehung nicht im engeren Sinne mathematischer Fähigkeiten wie z. B. des Leseverständnisses u. a.) führt jedoch dazu, dass sich eine Verteilung der Aufgabenschwierigkeiten mit relativ vielen Aufgaben im mittleren Bereich einstellt. Immerhin lässt sich für den QuaSUM-Mathematiktest der Klassenstufe 5 der Erfolg bei der Erfassung eines breiten Spektrums messbarer Mathematikleistungen erkennen. Dabei sind die Aufgaben der verschiedenen Themenbereiche weitgehend über die gesamte Skala verteilt; allein die Aufgaben aus dem Bereich *Größen* sind deutlicher auf die untere Hälfte der Skala konzentriert.

Die mathematischen Fähigkeiten der Schülerinnen und Schüler aus fünften Klassen sind demgegenüber auf der linken Seite der Skala abgebildet. Wie Abbildung 3 zeigt, sind die Ergebnisse angenähert normalverteilt: Nur wenige Schülerinnen und Schüler zeigen extrem schwache oder (nahezu) perfekte Leistungen, während viele im mittleren Bereich liegen.

Um einen Eindruck von den Aufgabenstellungen des Tests und seiner unterschiedlichen Schwierigkeit zu geben, sind in die Abbildung 3 Beispielaufgaben aus dem Themenbereich *Zahlenbereiche/Rechnen* eingetragen[12]. Die Darstellung ist so gewählt, dass die Aufgaben bestimmten Fähigkeitsniveaus zugeordnet werden, die erreicht werden müssen, um sie mit einiger Sicherheit – d. h. hier: mit 50-prozentiger Wahrscheinlichkeit – lösen zu können.

Im Themenbereich *Zahlenbereiche/Rechnen* sollten die Schülerinnen und Schüler sichere Kenntnisse bei Grundrechenoperationen im Bereich der natürlichen Zahlen nachweisen, Rechengesetze und vorteilhaftes Rechnen beherrschen sowie Rundungs- und Teilbarkeitsregeln anwenden, ihr Wissen über die römischen Zahlzeichen zeigen, mit Diagrammen umgehen und Sachaufgaben lösen können. Abbildung 3 zeigt, dass z. B. das Fähigkeitsniveau

[12] Für die drei anderen Themenbereiche sind Beispielaufgaben in der Abbildung 4 aufgeführt.

mit dem Skalenwert 127 von 85 Prozent der Schülerinnen und Schüler mit einiger Sicherheit erreicht oder überschritten wird. Die Beispielaufgabe 1 mit der Anforderung, die Grundrechenart Subtraktion zu beherrschen, ist eine Aufgabe dieser Art und Schwierigkeit. Das Fähigkeitsniveau 146 mit der Beispielaufgabe 2, deren Lösung die Teilbarkeitsregeln voraussetzt, erreichen bzw. übertreffen ca. 55 Prozent der Schülerinnen und Schüler. Die Lösung einer Aufgabe, die die Beherrschung von Grundrechenarten bei Berücksichtigung der Klammerregeln erfordert, gelingt mit einiger Sicherheit erst bei einem Skalenwert von 166; die Beispielaufgabe 3 spiegelt das Fähigkeitsniveau, welches ca. 23 Prozent der Brandenburger Fünftklässler erreichen oder überschreiten. Die ebenfalls relativ schwierige Beispielaufgabe 4 (Skalenwert 170), bei der anwendungsbezogen Informationen aus einem Diagramm verarbeitet werden müssen, ist eine typische Aufgabe für ein mathematisches Fähigkeitsniveau, das 20 Prozent der Schülerinnen und Schüler erzielt haben oder überschreiten[13].

Die folgende Abbildung 4 zeigt für die drei Themenbereiche *Größen*, *Verhältnisgleichungen/Proportionalität* und *Geometrie* jeweils vier Beispielaufgaben für die Fähigkeitsniveaus, die mit einer 50-prozentigen Wahrscheinlichkeit bei dem angegebenen Skalenwert erreicht oder überschritten werden. Für die insgesamt 12 Beispielaufgaben ist jeweils zusätzlich die tatsächliche Lösungsquote mit angegeben.

[13] Die genannten Beispielaufgaben markieren bestimmte Schwierigkeitsniveaus, die mit angebbarer Wahrscheinlichkeit erreicht oder übertroffen werden; die Abbildung 3 gibt aber noch keine Information über die tatsächliche Lösungsquote der Aufgaben. So wurde Beispielaufgabe 1 von 71 Prozent, Beispielaufgabe 2 von 56 Prozent, Beispielaufgabe 3 von 36 Prozent und Beispielaufgabe 4 von 33 Prozent der Schülerinnen und Schüler richtig gelöst. Diese Lösungsquoten, die sich jeweils nur auf eine Aufgabe beziehen, unterscheiden sich etwas von den im Text genannten Prozentsätzen, die jeweils das gesamte Lösungsmuster, also auch „Flüchtigkeitsfehler“ und „Zufallstreffer“, berücksichtigen.

Abbildung 4 Beispielaufgaben des QuaSUM-Mathematiktests (Klassenstufe 5) aus den Themenbereichen *Größen*, *Verhältnisgleichungen/Proportionalität* und *Geometrie*

Im Themenbereich ***Größen*** wurden die Kenntnisse der Schülerinnen und Schüler im Umwandeln und Rechnen mit den Einheiten der Masse, der Länge und der Zeit überprüft. In Sachaufgaben mussten diese Kenntnisse angewendet werden.

Rasch-Skalenwert: 136

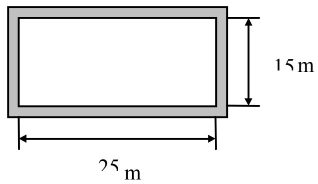
Rechne $3\frac{1}{2}$ Stunden in Minuten um!

a 160 min
b 180 min
c 210 min
d 350 min
e 230 min

Lösungshäufigkeit dieser Aufgabe: 62 %

Rasch-Skalenwert: 149

Angela schwimmt viermal die kurze Bahn, Horst dreimal die lange. Anne schwimmt einmal am Rand herum.



Wie viel Meter schwimmt jeder?

Lösungshäufigkeit dieser Aufgabe: 51 %

Rasch-Skalenwert: 164

Auf dem Weg zur Schule geht Ulf zunächst 640 m zu Fuß. Dann fährt er 3,2 km mit dem Bus. Zu Fuß legt er nochmals 260 m zurück.
Wie viel **km** beträgt sein Schulweg?

Lösungshäufigkeit dieser Aufgabe: 36 %

Rasch-Skalenwert: 190

20 t Kartoffeln sollen in 5 kg–Portionen abgepackt werden. Wie viele Beutel werden benötigt?

Lösungshäufigkeit dieser Aufgabe: 18 %

Im Themenbereich ***Verhältnisgleichungen/Proportionalität*** wurden Aufgaben eingesetzt, die das Erkennen von Zuordnungen in Tabellen und das inhaltliche Lösen von Gleichungen und Ungleichungen überprüfen. Die Kenntnisse der Schülerinnen und Schüler aus diesem Themenbereich wurden darüber hinaus in einer Sachaufgabe getestet.

Rasch-Skalenwert: 139

Aus 1 kg Frischpilzen lassen sich etwa 200 g Trockenpilze herstellen. Wie viel Frischpilze werden für die Herstellung von 500 g Trockenpilzen benötigt?

a 10 kg
b 5 kg
c 100 g
d 500 g
e 2,5 kg

Lösungshäufigkeit dieser Aufgabe: 60 %

Rasch-Skalenwert: 151

Bestimme die Lösung der folgenden Gleichung!
(4 – x) _ 3 = 12 $(x \in N)$

Lösungshäufigkeit dieser Aufgabe: 50 %

Rasch-Skalenwert: 166

Welche Zuordnung ist **falsch** ?

	a	b	c	d	e
r	0	1	2	3	4
$3 \cdot r + 1$	1	5	7	10	13

Lösungshäufigkeit dieser Aufgabe: 34 %

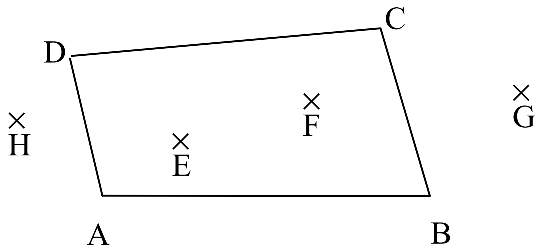
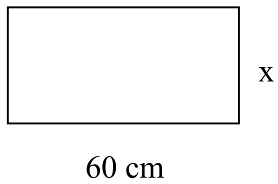
Rasch-Skalenwert: 193

Gegeben ist die Ungleichung: $11 + 2 \cdot a < 21$ $(a \in N)$
Wie lautet die richtige Lösung?

a $L_1 = \{1; 2; 3; 4; 5\}$
b $L_2 = \{0; 1; 2\}$
c $L_3 = \{0; 1; 2; 3; 4; 5\}$
d $L_4 = \{1; 2; 3\}$
e $L_5 = \{0; 1; 2; 3; 4\}$

Lösungshäufigkeit dieser Aufgabe: 16 %

Abbildung 4 (Fortsetzung)

Im Themenbereich ***Geometrie*** sollten die Schülerinnen und Schüler symmetrische Figuren erkennen, die Lagebeziehungen von Geraden, Punkten bzw. Flächen bestimmen, eine Klassifizierung von Dreiecken vornehmen, Bestimmungsstücke von Flächen und Körpern berechnen können und deren Anwendung beherrschen.	
Rasch-Skalenwert: 123 Welche Aussage ist wahr?  a Die Punkte H, E, F, G liegen auf einer Geraden. b Die Punkte A, D, G, B bilden das gezeichnete Viereck. c Die Punkte E, F liegen innerhalb des gezeichneten Vierecks. d Die Punkte H, G liegen innerhalb des gezeichneten Vierecks. e Die Punkte E, F liegen außerhalb des Vierecks. Lösungshäufigkeit dieser Aufgabe: 73 %	**Rasch-Skalenwert: 152** Wie groß ist die Länge x des Rechteckes, wenn der Flächeninhalt 2 400 cm² beträgt?  a 4 cm b 2 280 cm c 40 cm d 1 140 cm e 400 cm Lösungshäufigkeit dieser Aufgabe: 47 %
Rasch-Skalenwert: 167 Katrins Hund bekommt einen neuen rechteckigen Hundezwinger. Er ist 6 m lang und 4 m breit. Wie viel Meter Maschendrahtzaun werden benötigt? a 20 m b 24 m c 16 m d 10 m e 14 m Lösungshäufigkeit dieser Aufgabe: 34 %	**Rasch-Skalenwert: 197** Wie groß ist das Volumen eines Würfels mit der Kantenlänge von 4 cm? a 16 cm³ b 64 cm³ c 12 cm³ d 32 cm³ e 60 cm³ Lösungshäufigkeit dieser Aufgabe: 13 %

3.1.2 QuaSUM-Mathematiktest – Klassenstufe 9

Der QuaSUM-Mathematiktest für die Klassenstufe 9 mit insgesamt 120 Aufgaben wurde in schulform- bzw. kursniveaudifferenzierenden Fassungen eingesetzt. Sämtliche Schülerinnen und Schüler haben einen gemeinsamen Satz von 40 Aufgaben bearbeitet, die auf der Basis von Voruntersuchungen als Aufgaben mittlerer Schwierigkeit klassifiziert wurden (Stufe 3). Den Schülerinnen und Schülern aus Mathematik-Grundkursen an Gesamtschulen wurden weitere 40 Aufgaben der Anforderungsstufen 1 und 2, den Schülerinnen und Schülern aus Realschulklassen und aus Mathematik-Erweiterungskursen an Gesamtschulen weitere 40 Aufgaben der Anforderungsstufen 2 und 4 und den Schülerinnen und Schülern aus Gymnasialklassen weitere 40 Aufgaben der Anforderungsstufen 4 und 5 vorgelegt (vgl. dazu Tabelle 1). Die drei Testformen enthalten jeweils die gleiche Anzahl an Aufgaben aus den Themenbereichen *Funktionen*, *Gleichungen/Ungleichungen*, *Zahlen/Variablen* und *Geometrie*.

Auch wenn Schülerinnen und Schüler nur eine Teilmenge der insgesamt eingesetzten Aufgaben bearbeitet haben, kann auf der Grundlage der Rasch-Skalierung eine zuverlässige Schätzung ihrer mathematischen Fähigkeiten getroffen werden, die den Gesamtaufgabenbestand zu Grunde legt. Analog zur Abbildung 3 ist in der folgenden Abbildung 5 die Verteilung der Aufgabenschwierigkeiten der Verteilung der Schülerfähigkeiten für die Gesamtstichprobe gegenübergestellt.

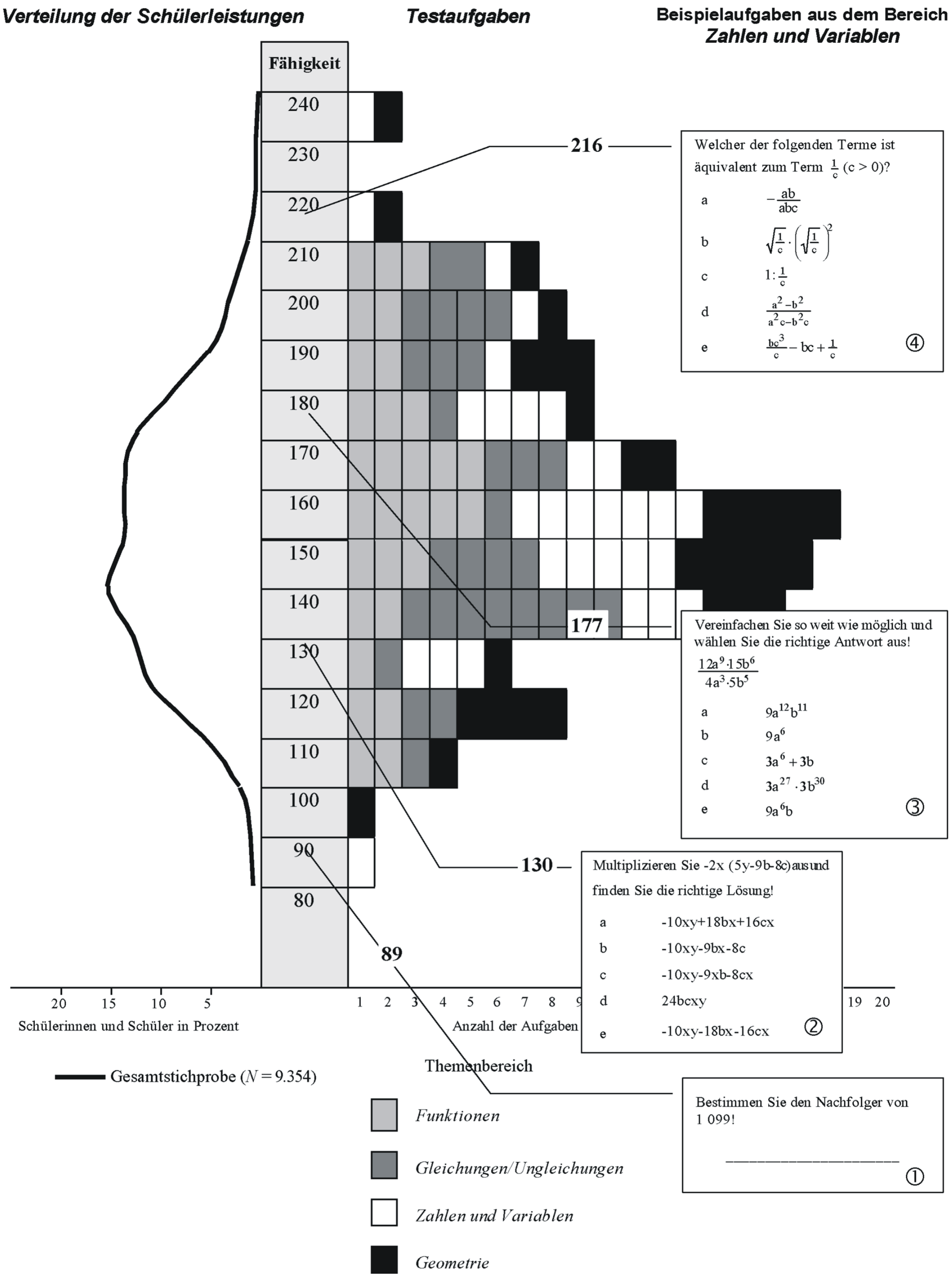
Wie beim Test für die Fünftklässlerinnen und Fünftklässler streuen die Aufgaben der vier Themenbereiche über die gesamte Skala, wobei sich auch hier eine Konzentration der Aufgabenschwierigkeiten im mittleren Skalenbereich zeigt. Auf der linken Seite der Skala ist für die Gesamtstichprobe die Leistungsverteilung eingetragen. Bei dem gesetzten Mittelwert 150 und der gesetzten Standardabweichung 25 zeigt sich eine Verteilung über den gesamten Skalenbereich von 80 bis 240; im unteren und oberen Leistungsbereich flacht die Verteilung deutlich ab.

In die Abbildung 5 sind vier Beispielaufgaben – hier aus dem Themenbereich *Zahlen/Variablen*[14] – aufgenommen, die die Fähigkeitsniveaus von Schülergruppen kennzeichnen. Die Beispielaufgabe 1 markiert ein Fähigkeitsniveau, das 99 Prozent aller Neuntklässlerinnen und Neuntklässler mit 50-prozentiger Wahrscheinlichkeit erreichen oder überschreiten. Die weiteren Beispielaufgaben stehen für Fähigkeitsniveaus, die von 77 Prozent (Beispiel 2), 15 Prozent (Beispiel 3) bzw. 1 Prozent (Beispiel 4) der gesamten Schülerschaft erreicht oder überschritten werden.

---

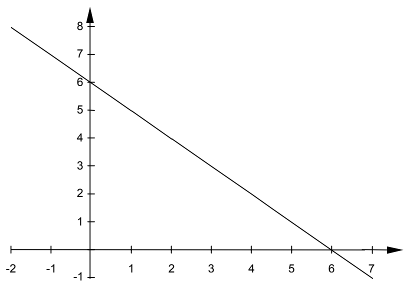
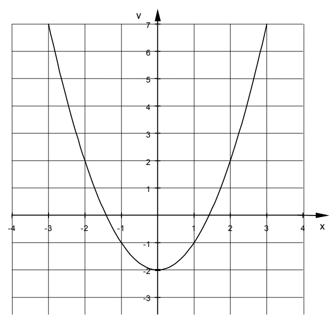
[14] Die Testaufgaben zum Themenbereich *Zahlen/Variablen* enthalten folgende Anforderungen: mit rationalen Zahlen rechnen und diese vergleichen, Terme erkennen, vereinfachen, berechnen, aufstellen und umformen, Aussagen zur Wurzeldefinition treffen und die Potenzgesetze kennen.

Abbildung 5 Verteilung der Schülerleistungen im Fach Mathematik (Klassenstufe 9) im Vergleich mit den Schwierigkeiten der Testaufgaben im QuaSUM-Mathematiktest (nach Themenbereichen)



In Abbildung 6 sind Beispielaufgaben aus den drei Themenbereichen *Funktionen*, *Gleichungen/Ungleichungen* und *Geometrie* genannt, die gleichfalls für die in der Abbildung 5 aufgeführten Schwierigkeitsniveaus stehen[15].

Abbildung 6 Beispielaufgaben des QuaSUM-Mathematiktests (Klassenstufe 9) aus den Themenbereichen *Funktionen*, *Gleichungen/ Ungleichungen* und *Geometrie*

<table>
<tr><td colspan="2">Im Themenbereich Funktionen sollten die Schülerinnen und Schüler Klassen von Funktionen bestimmen und typische Eigenschaften dieser kennen. Es waren Funktionsgleichungen zu ermitteln, Aussagen zu den Nullstellen, zum Anstieg und zum absoluten Glied zu treffen sowie geordnete Paare zu finden, eine Diskriminantenuntersuchung vorzunehmen, Scheitelpunkts- und Normalform zu bilden, Nullstellen zu finden, Koordinaten markanter Punkte zu bestimmen, Monotoniebetrachtungen durchzuführen und die Lage von Funktionsbildern anzugeben.</td></tr>
<tr><td>
Rasch-Skalenwert: 109
Finden Sie heraus, welche Funktionsgleichung zum abgebildeten Graphen gehört!

a f (x) = 2x – 4
b f (x) = -3x + 2
c f (x) = 3x
d f (x) = -3
e f (x) = 6 – x
</td><td>
Rasch-Skalenwert: 138
Im Koordinatensystem ist das Bild einer quadratischen Funktion dargestellt. Wählen Sie die zugehörige Funktionsgleichung aus!

a $f(x) = x^2$
b $f(x) = x^2 + 2$
c $f(x) = (x-2)^2$
d $f(x) = -x^2 - 2$
e $f(x) = x^2 - 2$
</td></tr>
<tr><td>
Rasch-Skalenwert: 177
Welche der vorgegebenen Lösungen ist die Nullstelle der quadratischen Funktion $f(x) = x^2 + 4x + 4$?
a $x_0 = 2$
b $x_0 = -2$
c $x_0 = -1$
d $x_0 = 0$
e $x_0 = 1$
</td><td>
Rasch-Skalenwert: 208
Für welchen Wert p hat die Funktion
$f(x) = x^2 + px + 6{,}25$ ($x \in$ R) genau eine Nullstelle?
a p = -2,5
b p = 5
c p =-6,25
d p = 6,25
e p = -12,5
</td></tr>
</table>

15 In der Abbildung 6 sind für die Beispielaufgaben keine Lösungsquoten angegeben, da bei dem gewählten Testdesign in der Jahrgangsstufe 9 (vgl. Tabelle 1) nur ein ausgewählter Kernbereich von 40 Aufgaben von sämtlichen Schülerinnen und Schülern bearbeitet wurde.

Abbildung 6 (Fortsetzung)

Im Themenbereich ***Gleichungen/Ungleichungen*** sollten die Schülerinnen und Schüler Zahlen und Größen vergleichen, Zuordnungen zu Zahlenbereichen vornehmen, den Maßstab anwenden, Lösungen und Lösungsmengen bei Gleichungen und Ungleichungen sowie Gleichungssysteme ermitteln, Funktionsgleichungen aufstellen, Aussagen zu Sachverhalten finden, die Prozentrechnung anwenden und Formeln umstellen.

Rasch-Skalenwert: 119

Wie heißt die Lösung der Gleichung $2x - 3 = 11$?

a	4
b	-7
c	7
d	-4
e	10

Rasch-Skalenwert: 139

Welche der folgenden Zahlen gehört **nicht** zur Lösungsmenge der Ungleichung $-2x < 10$?

a	$x = 0$
b	$x = -5$
c	$x = 4$
d	$x = -1$
e	$x = 5$

Rasch-Skalenwert: 181

Welche Werte x_1 und x_2 sind Lösungen der Gleichung $x^2 + 8x + 12 = 0$?

a	$x_1 = -2$ und $x_2 = -6$
b	$x_1 = 2$ und $x_2 = -6$
c	$x_1 = 1$ und $x_2 = -7$
d	$x_1 = -1$ und $x_2 = -7$
e	$x_1 = 4$ und $x_2 = -4$

Rasch-Skalenwert: 204

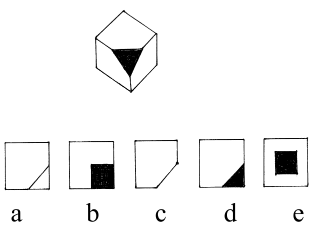
Finden Sie ohne Berechnung der Lösungsmenge heraus, welches der Gleichungssysteme genau eine Lösung hat!

a	$y = 2x - 1$ $y = x + 1$
b	$2y = x - 2$ $4y - 2x = -4$
c	$2y - 2 = x$ $y = 0{,}5x + 1$
d	$2x = 3y$ $6x = 9y$
e	$y = x + 1$ $-3y = -3x - 3$

Im Themenbereich ***Geometrie*** sollten die Schülerinnen und Schüler mit Größen arbeiten, Winkelgrößen und Beziehungen kennen, Kenntnisse zur Kongruenz nachweisen, geometrische Flächen und Körper mit ihren Eigenschaften kennen, Kenntnisse zum Strahlensatz besitzen, Eigenschaften der Zweitafelprojektion sowie geometrische Flächen und Körper kennen, geometrische Sätze beherrschen, Berechnungen an Flächen und Körpern mit ihren Anwendungen vornehmen und mit geometrischen Darstellungen und Beschreibungen arbeiten.

Rasch-Skalenwert: 109

Die Abbildung stellt einen Holzwürfel dar, bei dem man eine Ecke abgeschnitten hat. Die Schnittfläche ist dunkel. Welche der Zeichnungen in der Reihe darunter zeigt, wie der Würfel von oben betrachtet aussieht?



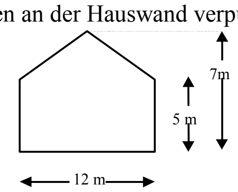
Rasch-Skalenwert: 139

In einem Viereck sind die Innenwinkel mit $\alpha = 80^\circ$, $\beta = 35^\circ$ und $\gamma = 125^\circ$ gegeben. Berechnen Sie die Größe des fehlenden Innenwinkels δ !

a	205°
b	215°
c	240°
d	120°
e	60°

Rasch-Skalenwert: 175

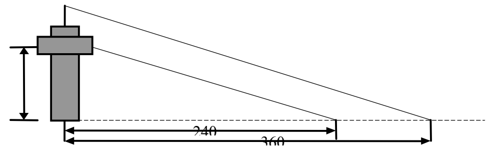
Wie viel m² müssen an der Hauswand verputzt werden?



a	72 m²
b	60 m²
c	42 m²
d	84 m²
e	53 m²

Rasch-Skalenwert: 200

Ein Fernsehturm ist 270 Meter hoch. In welcher Höhe befindet sich die Aussichtsplattform?



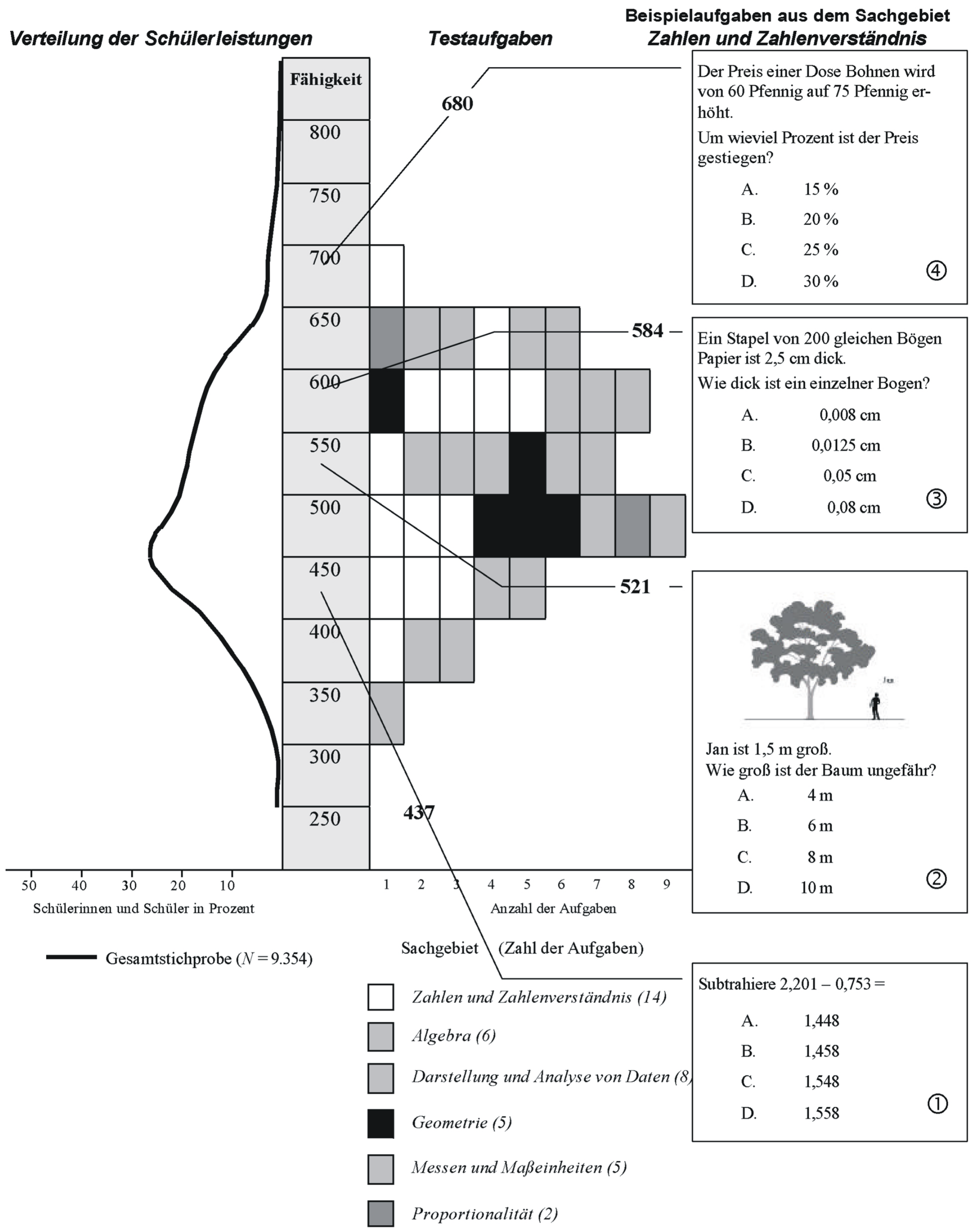
a	150 m
b	180 m
c	200 m
d	220 m
e	300 m

3.1.3 Mathe40-Test – Klassenstufe 9

Der Mathe40-Test besitzt ein anderes Anforderungsprofil als der QuaSUM-Test, was u. a. in der unterschiedlichen Art und Weise der Entwicklung der Tests begründet liegt: Der Mathe40-Test wurde aus Items zusammengestellt, die dem Aufgabenpool von TIMSS II in den Jahrgangsstufen 7 und 8 entstammen. Es handelt sich hierbei um Aufgaben, die auf einer internationalen curricularen Matrix beruhen, welche von einer etwas anderen Einteilung in Sachgebiete ausgeht und vor allem auch solche Ziele betont, die die Kontextualisierung mathematischer Operationen einschließen. Dagegen basiert der QuaSUM-Test auf eigens für diese Untersuchung entworfenen Aufgabensammlungen von Brandenburger Lehrerinnen und Lehrern (vgl. Abschnitt 2.1) mit einer stärkeren Gewichtung innermathematischer Operationen.

Die Unterschiedlichkeit der beiden Tests für die Schülerinnen und Schüler der QuaSUM-Stichprobe in der Klassenstufe 9 zeigt sich demnach nicht prinzipiell in den grundlegenden Anforderungsarten der Aufgaben: Sowohl der QuaSUM-Test als auch der Mathe40-Test enthalten Anforderungen auf den Ebenen *Kenntnisse*, *Beherrschung von Routineverfahren*, *Beherrschung von komplexen Verfahren* und *Lösung anwendungsbezogener und mathematischer Probleme* (zur gewählten Taxonomie vgl. BAUMERT ET AL. 1997). Der Unterschied liegt vielmehr in der Verteilung der Anforderungsarten. Quantitativ ist der QuaSUM-Test deutlich mehr durch Aufgaben gekennzeichnet, die Kenntnisse und Routineverfahren überprüfen; hingegen enthält er weniger Aufgaben zum Anwenden komplexer Verfahren sowie zu anwendungsbezogenen Problemen. Der Mathe40-Test weicht davon ab. Hier überwiegen Aufgaben aus dem eher anwendungsbezogenen Bereich (zur Qualifizierung der Mathe40-Aufgaben im Vergleich mit den QuaSUM-Testaufgaben vgl. BIEBER 1999). Bevor in Abbildung 7 die Verteilung der Aufgabenschwierigkeiten im Mathe40-Test der Verteilung der darauf bezogenen Schülerfähigkeiten gegenübergestellt wird, bleibt festzustellen, dass – nimmt man die Merkmale des QuaSUM-Tests als Maßstab – der Mathe40-Test das implementierte Curriculum von Brandenburger Lehrkräften nur bedingt nachzeichnet. Hinsichtlich des intendierten Curriculums ist zu beachten, dass acht der 40 Aufgaben im Mathe40-Test dem Bereich „Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung“ zuzurechnen sind (vgl. Tabelle 3); dieses Stoffgebiet ist im Brandenburger Rahmenplan Mathematik von 1991 nicht enthalten und wird damit in der Regel auch nicht Unterrichtsgegenstand für die getesteten Schülerinnen und Schüler gewesen sein.

Abbildung 7 Verteilung der Schülerleistungen im Fach Mathematik (Klassenstufe 9) im Vergleich mit den Schwierigkeiten der Testaufgaben im Mathe40-Test (nach Sachgebieten)



Auf der linken Seite der Skala ist wieder – wie für den QuaSUM-Test in Abbildung 5 – für die Gesamtstichprobe die Leistungsverteilung eingetragen. Analog zu dem in TIMSS gewählten Maßstab ist hier für die Gesamtstichprobe der Mittelwert auf 500 bei einer Standardabweichung von 100 gesetzt. Für die Gesamtstichprobe ergibt sich eine Verteilung über den gesamten Skalenbereich von 300 bis 750, wobei der Leistungsbereich knapp unterhalb des Mittelwertes stärker besetzt ist als andere Bereiche. In die Abbildung sind auf der rechten Seite der Skala die 40 Aufgaben nach ihrer Schwierigkeit geordnet eingetragen. Insgesamt verteilen sich die Aufgaben aller sechs Sachgebiete – auch die des Gebiets *Darstellung und Analyse von Daten, Wahrscheinlichkeitsrechnung* – relativ gleichmäßig über die Gesamtskala, wobei allein die Geometrieaufgaben deutlicher auf den mittleren Schwierigkeitsbereich konzentriert bleiben.

Zur Illustration der gemessenen Fähigkeitsniveaus sind am rechten Rand der Abbildung vier Beispielaufgaben aus dem Sachgebiet *Zahlen und Zahlenverständnis* nach Schwierigkeit geordnet wiedergegeben[16]. Es zeigt sich, dass die Grundrechenarten in der neunten Jahrgangsstufe auf dem durchschnittlichen Fähigkeitsniveau der Schülerinnen und Schüler bei einem Skalenwert von 437 mit „hinreichender Sicherheit“[17] (d. h. mit 65-prozentiger Lösungswahrscheinlichkeit) beherrscht werden (Beispiel 1), während anwendungsbezogene Multiplikation erst oberhalb eines Wertes von 521 Skalenpunkten erwartet werden kann (Beispiel 2). Eingekleidete Aufgaben, die zur Lösung eine komplexe Folge von Rechenoperationen voraussetzen, entsprechen einer Schwierigkeit, die dem durchschnittlichen Fähigkeitsniveau von 584 Skalenpunkten gleichkommen (Beispiel 3). Über einige Sicherheit in der Prozentrechnung schließlich verfügt nur der besonders leistungsstarke Teil der Schülerinnen und Schüler mit mehr als 680 Skalenpunkten (Beispiel 4)[18]. Die für den Bereich *Zahlen und Zahlenverständnis* aufgezeigte Struktur, wonach die Aufgaben der Mathe40-Skala nicht nur bezogen auf ihren Lösungsgrad, sondern auch unter inhaltlich-anforderungsbezogenen Gesichts-

---

[16] Die übrigen fünf Sachgebiete bzw. die Verteilung ihrer Aufgaben auf der Skala werden wegen der teilweise sehr geringen Aufgabenanzahl nicht besonders ausgeführt.

[17] Um Vergleichbarkeit zu den TIMSS-Ergebnissen herstellen zu können wurde bei der Auswertung des Mathe40-Tests die in TIMSS verwendete Vorgabe einer 65-prozentigen Lösungswahrscheinlichkeit gewählt, ab der eine Aufgabe einem bestimmten Fähigkeitsniveau zugeordnet wird.

[18] Die Beispielaufgaben bzw. ihre Schwierigkeitsprofile entsprechen weitgehend denen, die für die bundesweite TIMS-Studie berichtet bzw. berechnet wurden (vgl. dazu ausführlich BAUMERT ET AL. 1997, S. 68f.).

punkten eine zunehmend höhere Schwierigkeit repräsentieren, gilt für sämtliche Aufgaben im Mathe40-Test.

Obwohl die Korrelation zwischen den Ergebnissen im QuaSUM-Test und dem Mathe40-Test über alle Schulformen bzw. Kursniveaus mit r = 0,79 recht hoch ist, lassen sich die besonderen Anforderungen des letzteren mit faktorenanalytischen Methoden gut nachweisen. Derjenigen Testdimension (der sog. „ersten Hauptkomponente“), die am ehesten als „allgemeine mathematische Fähigkeit“ interpretiert werden kann, ist die überwiegende Zahl der Aufgaben aus dem Mathe40-Test zuzuordnen (je nach der in den Vergleich einbezogenen QuaSUM-Testversion 70 Prozent (Gesamtschule, Grundkurse), 93 Prozent (Gesamtschule, Erweiterungskurse und Realschulen), 95 Prozent (Gymnasien)), während ein vergleichsweise hoher Anteil der QuaSUM-Testaufgaben davon unabhängige, spezifischere Anforderungen stellt (Grundkursversion 59 Prozent, Erweiterungskurs- und Realschulversion 66 Prozent, Gymnasialversion 61 Prozent).

Zwar ließen sich die Aufgaben aus beiden Tests, rein messtechnisch betrachtet, auch zu einer einzigen Skala vereinigen; angesichts dieser Befunde empfiehlt es sich jedoch, im Folgenden von zwei verschiedenen Maßen für die mathematischen Fähigkeiten der Neuntklässlerinnen und Neuntklässler auszugehen. Sie unterscheiden sich nicht hinsichtlich ihres Zusammenhangs mit der Fähigkeit zum schlussfolgernden Denken, etwa in dem Sinne, dass der Mathe40-Test stärker darauf bezogen wäre als der QuaSUM-Test (r = 0,60 für ersteren gegenüber r = 0,59 für letzteren). Wohl aber hängt der QuaSUM-Test mit Eta^2 = 0,56 deutlich enger mit den unterschiedlichen Schulformzugehörigkeiten zusammen als der Mathe40-Test (Eta^2 = 0,38), und Ähnliches gilt für seine Eignung, Leistungsunterschiede zwischen einzelnen Schulen einer Schulform sichtbar zu machen. In dem Maße, wie solche Fragestellungen im Vordergrund des Interesses stehen, stellt er also eine besonders wichtige Erkenntnisquelle dar.

3.2 Mathematische Fachleistungen, differenziert nach Regionen, Schulformen, Schulen und Klassen bzw. Kursen

Im Folgenden werden zunächst die Leistungen der Fünftklässlerinnen und Fünftklässler (Abschnitt 3.2.1), anschließend die der Neuntklässlerinnen und Neuntklässler (Abschnitte 3.2.2 und 3.2.2) unter dem Gesichtspunkt regionaler Besonderheiten in Brandenburg und unter schul(-form)strukturellen Aspekten beleuchtet.

3.2.1 Mathematikleistungen in der Klassenstufe 5

Die mit dem QuaSUM-Test in der Klassenstufe 5 festgestellten Schulleistungen im Fach Mathematik entsprechen, wie im vorigen Abschnitt dargestellt wurde, in guter Näherung den Erwartungen der Brandenburger Lehrkräfte, von denen die Testaufgaben entwickelt worden waren. Die folgende Tabelle 8 gibt noch einmal für ausgewählte Skalenwerte zwischen 100 und 220 an, wie viel Prozent der Fünftklässlerinnen und Fünftklässler das entsprechende Fähigkeitsniveau erreicht haben oder überschreiten.

Tabelle 8 Verteilung der Fähigkeitsniveaus der Schülerinnen und Schüler im QuaSUM-Mathematiktest (Klassenstufe 5) für ausgewählte Skalenwerte (in Prozent)

	Skalenwert						
	bis 100	bis 120	bis 140	bis 160	bis 180	bis 200	über 200
Grundschulen/ Gesamtschulen mit angegliedertem Grundschulbereich	99,5	93,9	68,9	35,2	12,1	4,4	3,1

Zusammen mit den Beispielaufgaben in den Abbildungen 3 und 4 ergibt sich hiermit ein aufschlussreiches Bild über die breite Leistungsstreuung in den mathematischen Fähigkeiten bei den Fünftklässlerinnen und Fünftklässlern.

Eine Bewertung dieser Ergebnisse kann im vorliegenden Bericht nicht erfolgen; denn unmittelbare Vergleichswerte, etwa für andere Bundesländer, stehen nicht zur Verfügung. Unter dem Gesichtspunkt der bildungspolitischen Gestaltung – der „Systemsteuerung“ – käme für eine Bewertung am ehesten ein Vergleich der erreichten Mathematikleistungen mit den im Lehrplan festgelegten Anforderungen in Frage, wobei allerdings die außerschulischen Bedingungszusammenhänge und die sich daraus ergebenden Streuungen zwischen Schülerinnen bzw. Schülern, Klassen und Schulen immer mit zu berücksichtigen sind. Wenn man unterstellt, dass die hier gemessenen Kompetenzen, deren inhaltliche Bedeutung sich allgemein aus dem Bezug der Testaufgaben, speziell der Beispielitems, zu ihren Schwierigkeiten ergibt, prinzipiell von allen oder jedenfalls möglichst vielen Schülerinnen und Schülern erreicht werden können und sollten, ergeben sich aus dem Vergleich von Teilstichproben Ansatzpunkte und Prioritäten für konkrete Interventionen.

In diesem Abschnitt werden zunächst, diesseits normativer Fragestellungen, Differenzierungen beschrieben. Dabei wird zunächst untersucht, ob sich

Leistungsunterschiede feststellen lassen, die mit der Sozialtopografie des Landes Brandenburg oder der Organisation des Grundschulwesens zusammenhängen.

Regionale Differenzierung

Was die regional unterschiedlichen ökonomischen und damit auch sozialen Bedingungen anbetrifft, unter denen in den Grundschulen bzw. Primarstufen gearbeitet wird, so werden in den Befunden der Untersuchung kaum systematische Effekte sichtbar. Es war erwartet worden, dass in Gebieten mit relativ günstiger wirtschaftlicher Entwicklung höhere Fachleistungen angetroffen werden, und zwar teils wegen des Zuzugs von Familien mit vergleichsweise hohen Bildungsabschlüssen in der Elterngeneration, teils wegen der motivationalen Anreize, die mit einer positiven Wahrnehmung künftiger Berufsaussichten verbunden sein könnten.

Allerdings ergaben sich bei der Unterscheidung zwischen Schulen, die im „engeren Verflechtungsraum“ liegen, und solchen des „ländlichen Entwicklungsraumes“ keine statistisch signifikanten Differenzen. Es gibt in dieser Allgemeinheit wenig Anlass für die Vermutung, dass durch den Zuzug ökonomisch erfolgreicher Familien in Gemeinden mit neu erschlossenem Bauland in der Umgebung Berlins Schulen mit generell besonders hohem Leistungsniveau entstanden sind. Dies schließt nicht aus, dass für einzelne Schulen die Einzugsgebiete besonders günstig zugeschnitten sind, so dass sich hier ein vergleichsweise hohes Leistungsbild ergibt.

Demgegenüber sind zwischen den Landkreisen Unterschiede festzustellen. Aus mehreren Gründen verbietet es sich jedoch, daraus auf Unterschiede in den Fachleistungen auf Landkreisebene zu schließen. Erstens ist die Anzahl der in die Untersuchung einbezogenen Schulen je Landkreis zu gering, um im Einzelfall seriöse Vergleiche anzustellen. Zweitens ist der einzelnen Schule ein je spezifisches Einzugsgebiet zuzuordnen, das sich im Leistungsbild wesentlich stärker widerspiegeln dürfte als die Zugehörigkeit zu einem bestimmten, in sich heterogenen Landkreis. Drittens und vor allem zeigt der Vergleich mit analogen „Rangordnungen der Landkreise“ für die unabhängig gezogenen Stichproben der Klassenstufe 9, dass es nicht einmal näherungsweise eine Tendenz gibt, der zufolge einzelne Landkreise im Durchschnitt erfolgreichere oder weniger erfolgreiche Schülerinnen bzw. Schüler aufzuweisen hätten.

Nachweisen lässt sich indessen, dass in den Grundschulen aus besonders kleinen Ortschaften (Einwohnerzahl der Gemeinde unter 1.000) mit einem skalierten Testergebnis von 144,3 im Durchschnitt deutlich niedrigere Leistungen erzielt wurden als in den größeren Orten (im Durchschnitt 151,5; Effektstärke d = 0,26, also mehr als eine Viertel Standardabweichung)[19]. Dies wiederum steht in Zusammenhang mit den auf der Ebene der Landkreise erhobenen Sozialindikatoren: Die Schulen aus kleinen Ortschaften befinden sich schwerpunktmäßig in Landkreisen mit überdurchschnittlicher Arbeitslosigkeit und überdurchschnittlich vielen Sozialhilfeempfängern. Die Erwartung, dass sich die mit diesen beiden Indikatoren beschriebenen Kontexte (und gerade nicht die administrativen Zuordnungen als solche) in den gemessenen Fachleistungen widerspiegeln, trifft also offenbar zu. Sie konnte auf Individualebene mit der Methode der Regressionsanalyse bestätigt werden, auch wenn diese Effekte – bezogen auf die Fachleistungen der einzelnen Schülerinnen und Schüler – nicht sehr stark ausgeprägt sind (R^2 = 0,01) und im statistischen Sinne nicht als signifikant gelten können[20].

---

[19] Die Effektstärke d ist ein standardisiertes Maß für Merkmalsunterschiede zwischen zwei Gruppen. Sie wird berechnet, indem die Differenz der Mittelwerte der Gruppen durch die gemeinsame Standardabweichung dividiert wird. Je größer die dimensionslose, also auch über verschiedene Variablen hin vergleichbare Größe d ist, desto stärker sind die Unterschiede zwischen zwei Vergleichsgruppen hinsichtlich des untersuchten Merkmals ausgeprägt. Auch gilt: Je größer d ist, desto geringer ist der Überschneidungsbereich zwischen den beiden verglichenen Merkmalsverteilungen. Dies sei anhand einer beliebigen gruppenintern normalverteilten Variable für einige ausgewählte Werte von d veranschaulicht. Für d = 0,10 betrifft der Überschneidungsbereich 96,0 Prozent jeder der beiden Verteilungen, für d = 0,25 88,1 Prozent, für d = 0,50 80,2 Prozent, für d = 0,75 70,8 Prozent und für d = 1,00 61,7 Prozent. Mit wachsender Effektstärke d sind also die beiden verglichenen Verteilungen bezüglich des gemessenen Merkmals immer deutlicher unterschieden.

[20] Bei Aussagen über die statistische Signifikanz im Leistungsbereich wird in diesem Bericht in Rechnung gestellt, dass die Daten nicht aus einer einfachen Zufallsstichprobe stammen, sondern schulweise erhoben worden sind. Es wird deshalb stets eine Korrektur für die damit verbundenen „Design-Effekte“ (*Deff*) vorgenommen. Dieser Korrekturfaktor, der sich als das Verhältnis tatsächlich getesteter Schülerinnen und Schüler zum Äquivalent einer einfachen Zufallsstichprobe interpretieren lässt, beträgt für die Grundschulen *Deff* = 4,18. Der Wert für die Grundkurse der Gesamtschulen liegt bei *Deff* = 3,37, für die Erweiterungskurse der Gesamtschulen bei *Deff* = 4,04, für die Realschulen bei *Deff* = 6,17 und für die Gymnasien bei *Deff* = 4,82. Je größer der Wert *Deff*, desto homogener sind die Schulklassen im Verhältnis zur Gesamtvarianz.

Differenzierung nach Schulform

Hinsichtlich der Schulform sind in der Klassenstufe 5 keine nennenswerten Differenzen gefunden worden. Die fünften Klassen aus Grundschulen haben praktisch die gleichen Fachleistungen erzielt wie die Klassen aus Gesamtschulen mit eigener Primarstufe. Die durchschnittlichen Leistungen betragen sowohl für die Schülerinnen und Schüler aus fünften Klassen an Gesamtschulen als auch für diejenigen aus eigenständigen Grundschulen jeweils 150 Skalenpunkte. Lediglich die Streuung ist in den fünften Klassen der Gesamtschulen mit einer Standardabweichung von 25 Skalenpunkten etwas niedriger als in den fünften Klassen der Grundschulen (27 Skalenpunkte): Besonders gute, aber auch besonders schwache Leistungen sind in den Gesamtschulen etwas seltener vertreten[21]. Die Unterschiede sind allerdings statistisch nicht signifikant und müssen deshalb, bezogen auf die vorliegende Stichprobe, als zufallsbedingt gelten. Im Folgenden wird daher nicht mehr zwischen den beiden Organisationsformen unterschieden; der Begriff „Grundschule" schließt somit die fünften Klassen von Gesamtschulen mit ein.

Differenzierung nach Schulen

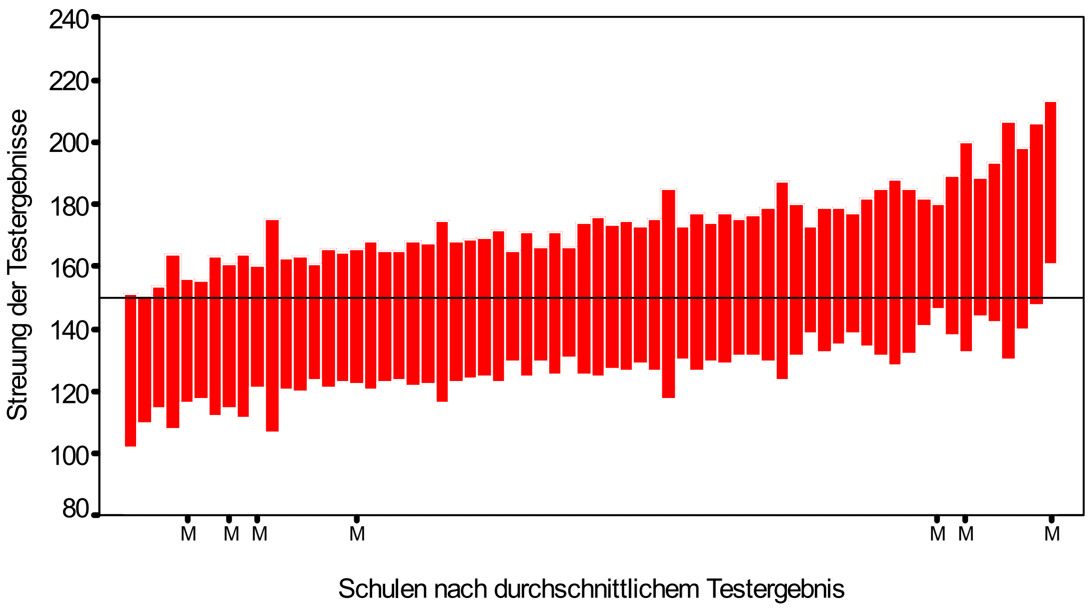
Nach allen bisherigen Erfahrungen war zu erwarten, dass sich das Leistungsbild von Schule zu Schule erheblich unterscheidet. Diese Erscheinung ist aus anderen Untersuchungen bekannt (vgl. z. B. LEHMANN & PEEK 1997, S. 56, die diesen Befund für Hamburg nachgewiesen haben), und die Zwischen-Schul-Varianz ist in Brandenburg offenbar ebenfalls hoch: Es gibt Schulen, in denen die besten Mathematikleistungen kaum an die schwächsten Lernergebnisse in denjenigen heranreichen, wo auf besonders hohem Niveau gearbeitet werden konnte (vgl. Abbildung 8).

In der Abbildung sind die untersuchten Schulen mit fünften Klassen nach der erreichten Durchschnittsleistung aufsteigend geordnet, wobei die Länge jeden „Balkens" den Bereich des Mittelwerts plus/minus eine Standardabweichung markiert, also eine Streuweite, innerhalb derer etwa zwei Drittel der Schülerinnen und Schüler zu finden sind. Vergleichsweise leistungshomogene Schulen sind somit durch kürzere Balken gekennzeichnet, solche mit stärker streuenden durch längere. Die horizontale Linie bezeichnet den Mittelwert über alle Schulen hinweg. Demzufolge haben an der leistungsstärksten Schule fast alle Schülerinnen und Schüler diesen Gesamtmittelwert

[21] Beide Gruppen haben eine durchschnittliche Lösungshäufigkeit von 18 Aufgaben bei einer Standardabweichung von 8 (fünfte Klassen in Grundschulen) bzw. 9 Aufgaben (fünfte Klassen in Gesamtschulen).

übertroffen, während an der leistungsschwächsten Schule nur sehr wenige in diesem Sinne überdurchschnittliche Resultate erzielt haben.

Abbildung 8 Verteilung der Ergebnisse im QuaSUM-Mathematiktest (Klassenstufe 5), nach Schulen (Durchschnittswerte plus/minus eine Standardabweichung)[22]



Bemerkenswert ist, dass sich die Gründe für solche Differenzen zwischen den Einzelschulen in Brandenburg stark von jenen unterscheiden, die für die Mathematikleistung in der Hamburger Studie festzustellen waren (vgl. LEHMANN & PEEK 1997). In Hamburg konnten 70 Prozent der *Zwischen-Schul-Varianz* durch Sozialindikatoren (schulbezogene Durchschnitte der Bildungsabschlüsse der Eltern und der Buchbestände in den Elternhäusern) aufgeklärt werden. Im Land Brandenburg erreicht der entsprechende Wert kaum 15 Prozent. Ein höherer Erklärungsgrad (41 Prozent; zum Vergleich: In Hamburg sind es 83 Prozent) ergibt sich erst, wenn man zusätzlich den Schulmittelwert für die Fähigkeit zum schlussfolgernden Denken (Ergebnis im CFT 20) als Prädiktor heranzieht. Auf individueller Ebene sind die Zusammenhänge zwischen den genannten Sozialindikatoren und der Mathematikleistung in Brandenburg mit R^2 = 0,13 sogar noch etwas stärker ausgeprägt als in Hamburg (R^2 = 0,10). Es ist deshalb davon auszugehen, dass die hohen aggregierten Effekte in Hamburg durch soziale Segregation zustande kommen, die sich in relativ homogenen Schuleinzugsgebieten widerspiegelt, während die

[22] Die sieben in der Fußleiste der Abbildung mit „M“ gekennzeichneten Schulen sind Schulen aus dem Modellversuch „Kleine Grundschule“.

Grundschulen bzw. die Gesamtschulen mit Grundschulbereich in Brandenburg jeweils eine sozial stärker durchmischte Schülerschaft zu versorgen haben. Wenn sich nun gleichwohl die Zwischen-Schul-Varianz zwischen Hamburg und Brandenburg nur wenig unterscheidet (Hamburg: Eta^2 = 0,15; Brandenburg: Eta^2 = 0,12), so müssen hier in stärkerem Maße noch andere Faktoren als die sozialräumliche Schichtung eine Rolle spielen. Auch der Umstand, dass knapp 23 Prozent der untersuchten Schulen einzügig sind, wobei hier die Zwischen-Schul-Varianz mit der (höheren) Zwischen-Klassen-Varianz zusammenfällt, erklärt allenfalls einen kleinen Teil der für Brandenburg unerwartet starken Unterschiede im mathematischen Leistungsbild zwischen den Schulen. Hinter diesen Leistungsunterschieden zwischen den Schulen verbergen sich also in erheblichem Maße einstweilen ungeklärte Bedingungsfaktoren, darunter vielleicht auch Differenzen im Anspruchsniveau und in der Qualität des Mathematikunterrichts, die in Hamburg wiederum durch den sozialräumlichen Kontext verdeckt werden.

Exkurs: Ergebnisse der Schulen aus Schul- und Modellversuchen

Von besonderem Belang sind in diesem Zusammenhang die Ergebnisse, auf die sich die Schulversuche „Jenaplan-Schule“ und „Montessori-Schule“ sowie die Schulen des Modellversuchs „Kleine Grundschule“ beziehen.

Zwar zeigen die Schülerinnen und Schüler der „Jenaplan-Schule“ deutlich höhere Mathematikleistungen als ihre Altersgenossen in den Vergleichsgruppen (d = 0,62), aber wegen der kleinen Zahl Betroffener (N = 18) ist dieser Befund zufallskritisch nicht zu sichern, und er wird halbiert, wenn man die soziale und kognitive Eingangsselektivität dieser Schule in Rechnung stellt (d = 0,29). Hinsichtlich der Einstellungen gegenüber Schule und Unterricht gibt es keine erwähnenswerten Differenzen. Das Gleiche gilt durchgehend für die „Montessori-Schule“, die ebenfalls vorwiegend von Kindern aus eher bildungsnahen Elternhäusern besucht wird. Nur wären hier aus eben diesem Grunde eher überdurchschnittliche Testleistungen zu erwarten gewesen, was nicht der Fall ist.

Anders verhält es sich mit den sieben Schulen aus dem Modellversuch „Kleine Grundschule“ (vgl. die in der Abbildung 8 durch den Buchstaben „M“ gekennzeichneten Schulen). Obwohl sich solche Schulen sowohl am oberen als auch am unteren Rand der Rangreihe befinden, die sich aus den Testergebnissen Mathematik ergibt, überwiegt das positive Moment: Insgesamt wurden hier tendenziell überdurchschnittliche Mathematikleistungen festgestellt, was sich weniger im Oberflächenbefund zeigt als vielmehr bei

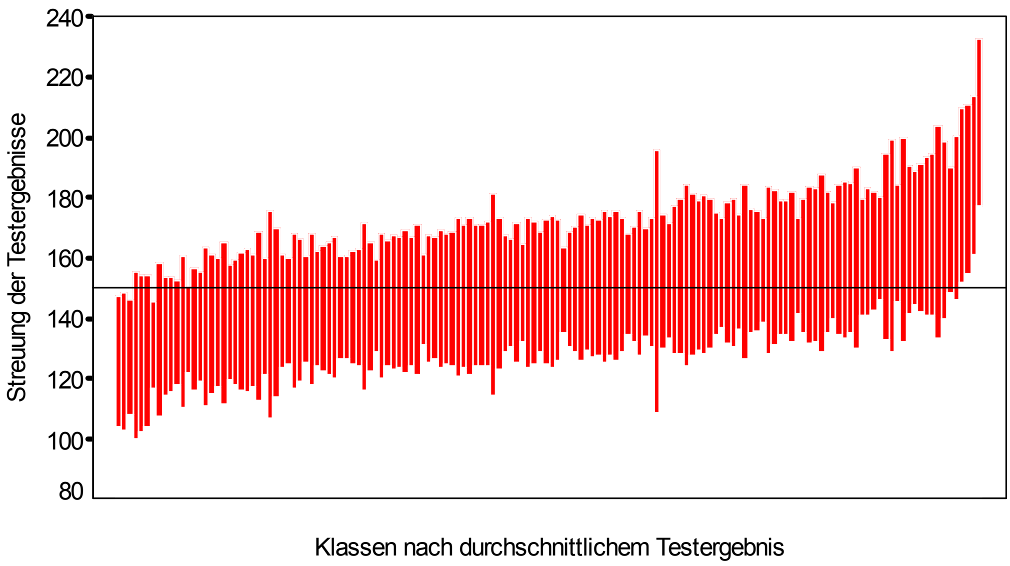
Berücksichtigung der in diesem Falle eher ungünstigen außerschulischen Lernvoraussetzungen. Besonders bemerkenswert ist es, dass der an diesen Schulen erteilte Mathematikunterricht von den Schülerinnen und Schülern in ungewöhnlichem Maße als zielgerichtet, transparent und strukturiert wahrgenommen wird (d = 0,41) und dass hier auch ein überdurchschnittliches Maß an Schulzufriedenheit geäußert wird (d = 0,52; vgl. dazu auch Abschnitt 4.4). Alles deutet darauf hin, dass dieser Reformansatz in seinem Modellversuchsstadium erfolgreich verläuft, auch wenn damit noch keine Aussagen über die Erfolgsaussichten im Falle einer flächendeckenden Umsetzung gemacht werden können.

Differenzierung nach Schulklassen

Die Leistungsunterschiede zwischen den untersuchten Schulklassen sind zwangsläufig höher als die Differenzen zwischen den Schulen: Die Parallelklassen in einer Schule werden sich so gut wie niemals auf exakt gleichem Leistungsstand befinden, und so wird die Gesamtvarianz der individuellen Testergebnisse in höherem Maße durch die Klassenzugehörigkeit als durch die Schulzugehörigkeit bestimmt.

Grafisch zeigt dies die Abbildung 9, die analog zur Abbildung 8 aufgebaut ist. In dieser Abbildung sind die Klassen ohne Rücksicht auf die Zugehörigkeit zu einer bestimmten Schule nach ihrer durchschnittlichen Mathematikleistung angeordnet.

Abbildung 9 Verteilung der Ergebnisse im QuaSUM-Mathematiktest (Klassenstufe 5), nach Klassen (Durchschnittswerte plus/minus eine Standardabweichung)

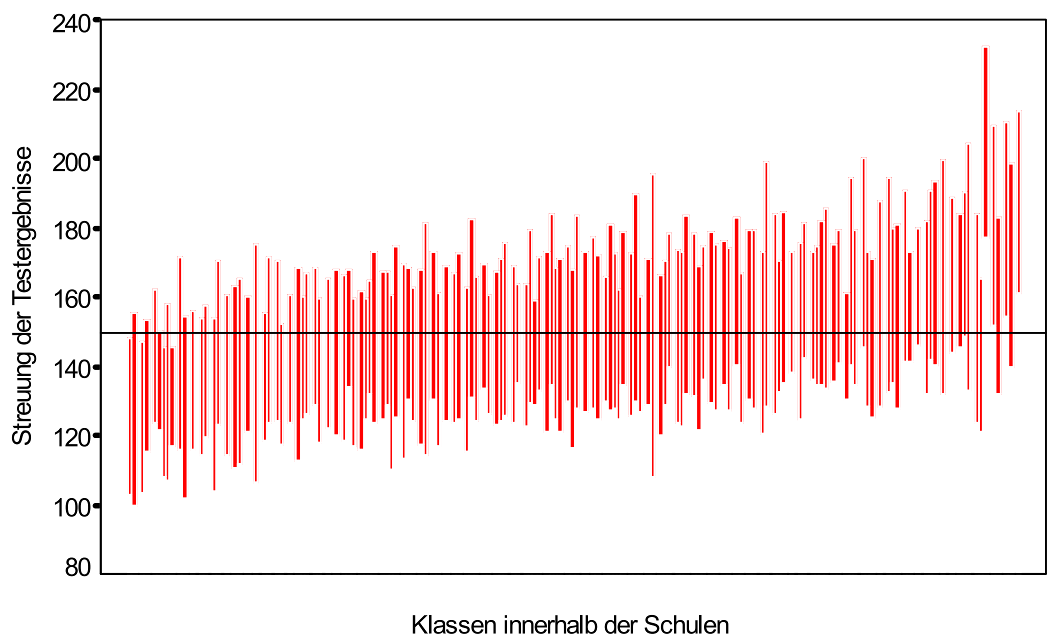


Wären die Klassenmittelwerte miteinander verbunden, so ergäbe sich eine am linken Rand nach unten und am rechten Rand nach oben abgebogene Kurve; denn die Klassenmittelwerte sind in guter Näherung mit einer Standardabweichung von 11,7 Skalenpunkten normalverteilt um den Gesamtmittelwert von 150, der durch die waagerechte Linie markiert wird. Das Bestimmtheitsmaß Eta^2 für die Zwischen-Klassen-Varianz beträgt 0,20 und ist damit um 8 Prozentpunkte höher als die Zwischen-Schul-Varianz. Hierin liegt – jenseits des Einflusses außerschulischer Faktoren – ein erneuter Hinweis auf die grundlegende Bedeutung des Unterrichts für das jeweils erzielte Leistungsniveau in einer Klasse.

An zwei schon optisch auffälligen Beispielen sei dies demonstriert. Etwas rechts von der Mitte der Grafik ragt eine Klasse durch eine ungewöhnlich breite Streuung der Leistungen heraus. Es handelt sich dabei um eine Lerngruppe mit 17 Schülerinnen und Schülern, von denen sich acht im unteren Leistungsviertel und acht im oberen Leistungsviertel der Brandenburger Gesamtverteilung befinden. Von den sozialen Eingangsvoraussetzungen her handelt es sich dabei insgesamt um eine eher benachteiligte Gruppe. Es ist der Lehrkraft aber offenbar gelungen, bei fast der Hälfte der Schülerinnen und Schüler sehr gute bis hervorragende Ergebnisse zu ermöglichen, freilich um den Preis, dass die andere Hälfte der Klasse auf dem eigentlich zu erwartenden niedrigen Leistungsniveau verblieben ist. Das leistungsbezogene Selbstkonzept, die Schulzufriedenheit und Wahrnehmung des Mathematikunterrichts als einer zielgerichteten und wohlstrukturierten Aktivität sind in dieser Klasse insgesamt überdurchschnittlich günstig ausgeprägt. In teilweisem Gegensatz dazu ist am rechten Rand der Grafik eine Klasse zu erkennen, in der 19 von 20 Schülerinnen und Schülern Testergebnisse erzielt haben, die für das obere Leistungsviertel in Brandenburg charakteristisch sind, davon 18 sogar Ergebnisse oberhalb des 85. Perzentils. Die Verteilung der Leistungen innerhalb der Klasse ist nicht weiter auffällig, suggeriert also keine Unregelmäßigkeiten in der Testdurchführung. Auch hier handelt es sich keineswegs um eine sozial hoch ausgelesene (wenn auch nicht um eine benachteiligte, sondern eher durchschnittliche) Gruppe. Und wiederum zeigen die schul- und unterrichtsbezogenen Einstellungen – leistungsbezogenes Selbstkonzept, Schulzufriedenheit und Wahrnehmung des Unterrichts – ein positives, hier sogar außergewöhnlich günstiges Bild. Vor allem diese zweite exemplarisch beleuchtete Klasse zeigt also, dass es gelingen kann, hervorragende Erfolge im Mathematikunterricht zu erzielen, ohne die Schülerinnen und Schüler übermäßig zu belasten, vielleicht sogar durch eben diese Erfolge die Lernfreude und das Selbstvertrauen zu steigern.

Dass solche Erfolge ihre Wurzeln am ehesten auf der Ebene des Fachunterrichts im Klassenverband haben, legt die nachstehende Abbildung 10 nahe. Hier sind die Schulen – jeweils getrennt um eine Balkenbreite – wieder nach aufsteigender Durchschnittsleistung angeordnet, nun aber aufgelöst in die Ergebnisse aus den zugehörigen Parallelklassen. So wird beispielsweise sichtbar, dass sich die Ausnahmeklasse mit den besten Ergebnissen überhaupt an einer Schule mit weiteren zwei Parallelklassen befindet, in denen unter- bzw. durchschnittliche Mathematikleistungen gezeigt wurden, und zwar obwohl die Eingangsvoraussetzungen dort nicht wesentlich ungünstiger waren, wie zusätzliche Analysen zeigen. Die Klasse, in der nur die eine Hälfte der Schülerinnen und Schüler ungewöhnlich gute Ergebnisse erreicht hat, besitzt eine Parallelklasse, in der die Eingangsvoraussetzungen insgesamt eher günstiger waren und für die die Testergebnisse weitgehend den Erwartungen entsprechen. Im übrigen wird in der Grafik sichtbar, dass sich an vielen Schulen die Parallelklassen eher unähnlich sind, und eben diese Erscheinung spiegelt sich in der erwähnten Steigerung der Zwischen-Klassen-Varianz gegenüber der Zwischen-Schul-Varianz wider.

Abbildung 10 Verteilung der Ergebnisse im QuaSUM-Mathematiktest (Klassenstufe 5), nach Klassen innerhalb von Schulen (Durchschnittswerte plus/minus eine Standardabweichung)



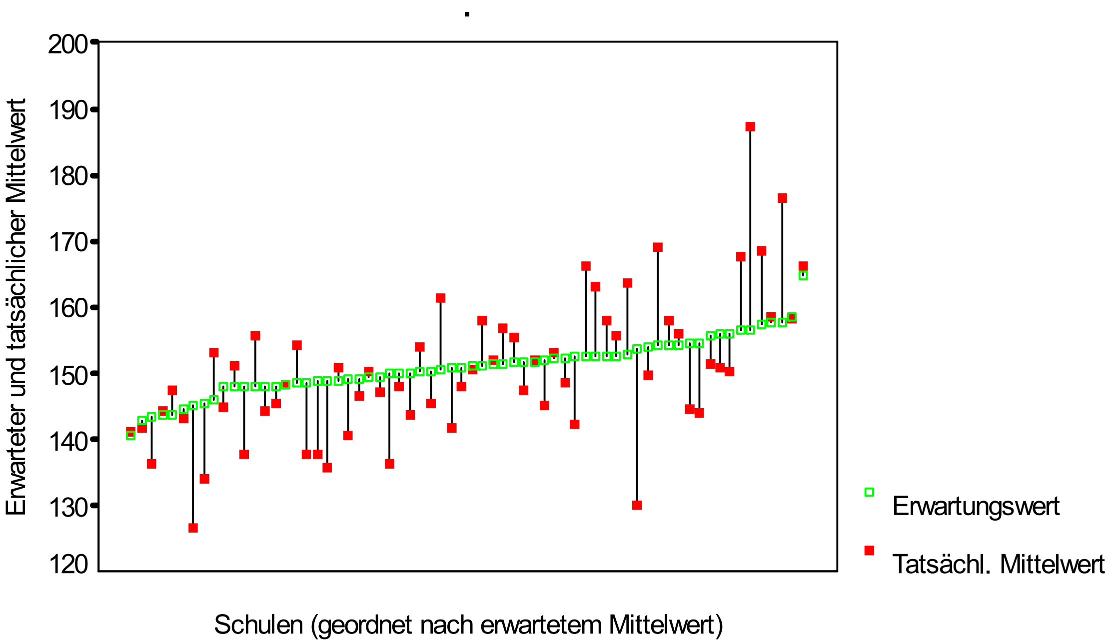
Gewissermaßen als Sonderfall im Rahmen dieses Ansatzes kann man das sog. „*Benchmarking*“ verstehen: Schulen und Klassen, die – gemessen an ihren Eingangsvoraussetzungen – erwartungswidrig günstige Testergebnisse erzielt haben, können als Maßstäbe dafür dienen, was unter realistischen Bedingungen erreichbar ist.

Insgesamt liegt nahe, dass die Unterschiede in den Testleistungen über schulische und unterrichtliche Merkmale hinaus wesentlich durch Kontextfaktoren mitbedingt sind, die außerhalb von Schule und Unterricht liegen. Um dem nachzugehen, wurden auf Schülerebene entsprechende Berechnungen durchgeführt. Als beste Prädiktoren auf individueller Ebene erwiesen sich in nachstehender Reihenfolge: die Fähigkeit zum schlussfolgernden Denken (Gesamtwert im CFT 20), das leistungsbezogene Selbstkonzept, der höchste erreichte Schulabschluss der Eltern bzw. des erziehenden Elternteils und der Buchbestand im Elternhaus. Alle diese Merkmale besitzen einen spezifischen, hochsignifikanten Einfluss auf die Testleistung. Gemeinsam erklären sie 42 Prozent der Schülervarianz im QuaSUM-Mathematiktest.

Somit lassen sich die erzielten Mathematikleistungen in einzelnen Schulen und in einzelnen Klassen zu denjenigen in Beziehung setzen, die nach Brandenburger Maßstäben bei der gegebenen Zusammensetzung der Schule bzw. Klasse hinsichtlich der Schülermerkmale *Fähigkeit zum schlussfolgernden Denken*, *leistungsbezogenes Selbstkonzept und Bildungsnähe des Elternhauses (Schulabschluss, Bücherbestand)* zu erwarten gewesen wäre. Damit werden, soweit es bei der gegebenen Datenlage möglich ist, außerschulische Kontexteinflüsse angemessen berücksichtigt. Abweichungen zwischen zu erwartender und tatsächlich beobachteter schul- bzw. klassenspezifischer Testleistung geben zumindest Hinweise auf besondere Erfolge (oder auch relative Misserfolge) in den untersuchten Aspekten grundschulpädagogischer Arbeit.

Ordnet man die Grundschulen bzw. die Gesamtschulen mit angegliedertem Grundschulbereich nach der erwarteten durchschnittlichen Testleistung ihrer Fünftklässlerinnen und Fünftklässler in Mathematik, so ergibt sich das in Abbildung 11 gezeichnete Bild.

Abbildung 11 Testergebnisse im QuaSUM-Mathematiktest (Klassenstufe 5) nach Grundschulen bzw. Gesamtschulen mit angegliedertem Grundschulbereich: Vergleich der erwarteten und der tatsächlichen Schulmittelwerte

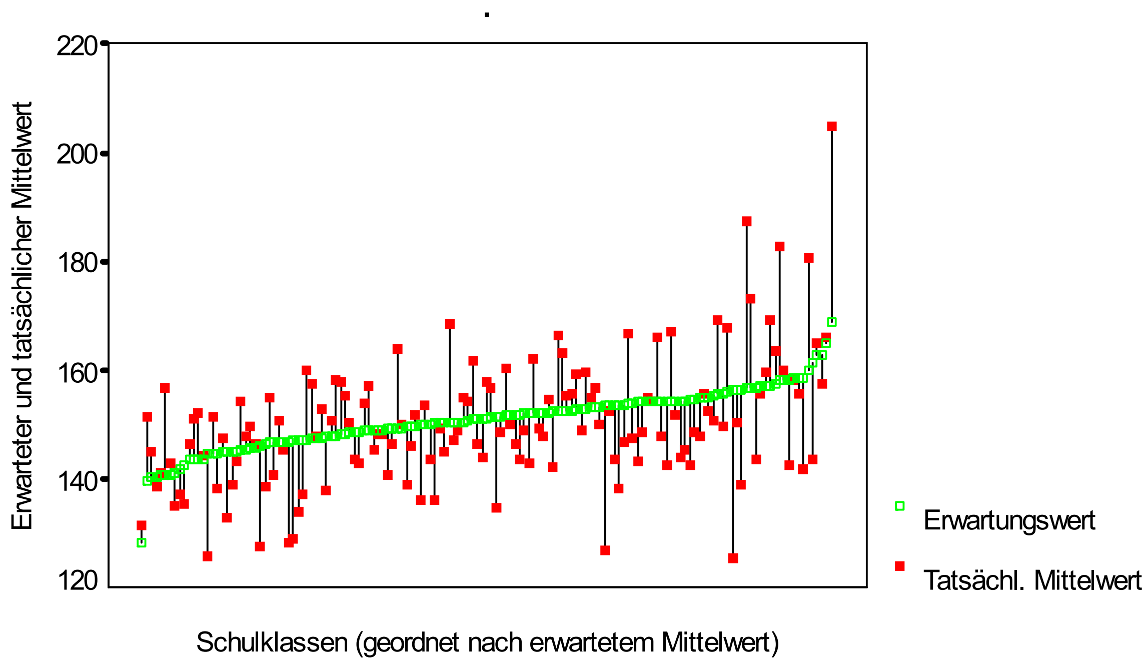


Es wird sichtbar, dass in nicht wenigen Schulen Durchschnittsleistungen angetroffen werden, die deutlich über den eigentlich zu erwartenden Stand hinausragen. In einigen anderen Schulen wurden Durchschnittsleistungen festgestellt, die erheblich unter dem Erwartungswert liegen[23].

Auch auf der Ebene von Klassen lässt sich auf der Grundlage der berücksichtigten Merkmale des außerschulischen Kontextes der Schülerinnen und Schüler für das Gesamtergebnis ein Durchschnittswert berechnen und dem ermittelten tatsächlichen Wert gegenüberstellen (vgl. Abbildung 12).

[23] Aus mess- und stichprobentheoretischen Gründen können nur Abweichungen zwischen beobachteten und erwartetem Leistungsstand, die mehr als 10 Skalenpunkte betragen, als aussagekräftig gelten.

Abbildung 12 Testergebnisse im QuaSUM-Mathematiktest (Klassenstufe 5) nach Klassen: Vergleich der erwarteten und der tatsächlichen Klassenmittelwerte



Der Frage, welche schulischen und unterrichtlichen Merkmale die hier aufgezeigten Diskrepanzen bedingen könnten, soll im Kapitel 5 dieses Berichts nachgegangen werden.

3.2.2 Mathematikleistungen in der Klassenstufe 9

Für Leistungsvergleiche der Schülerinnen und Schüler in der Klassenstufe 9 nach Regionen des Landes, nach Schulformen, Schulen und Schulklassen werden vor allem die Ergebnisse im QuaSUM-Mathematiktest genutzt.

Regionale Differenzierung

Wie in den Grundschulen, so sind auch in der Klassenstufe 9 keine starken Tendenzen zu einer regionalen Differenzierung in Brandenburg erkennbar. Wenn man davon ausgeht, dass zumindest auf der Ebene der Landkreise trotz verschiedener Wirtschaftsstruktur und unterschiedlich dichter Besiedlung insgesamt ein etwa gleichwertiges Bildungsangebot bereitgestellt werden soll, so dürfen – über alle Schulformen hinweg gerechnet – die durchschnittlichen Mathematikleistungen zwischen den Landkreisen nicht allzu stark voneinander abweichen. Da für die vorliegende Untersuchung keine Vollerhebung möglich war, kann die Überprüfung der Erreichung dieses Ziels mit den vorhandenen Daten nur eingeschränkt erfolgen. Obwohl aus den 18

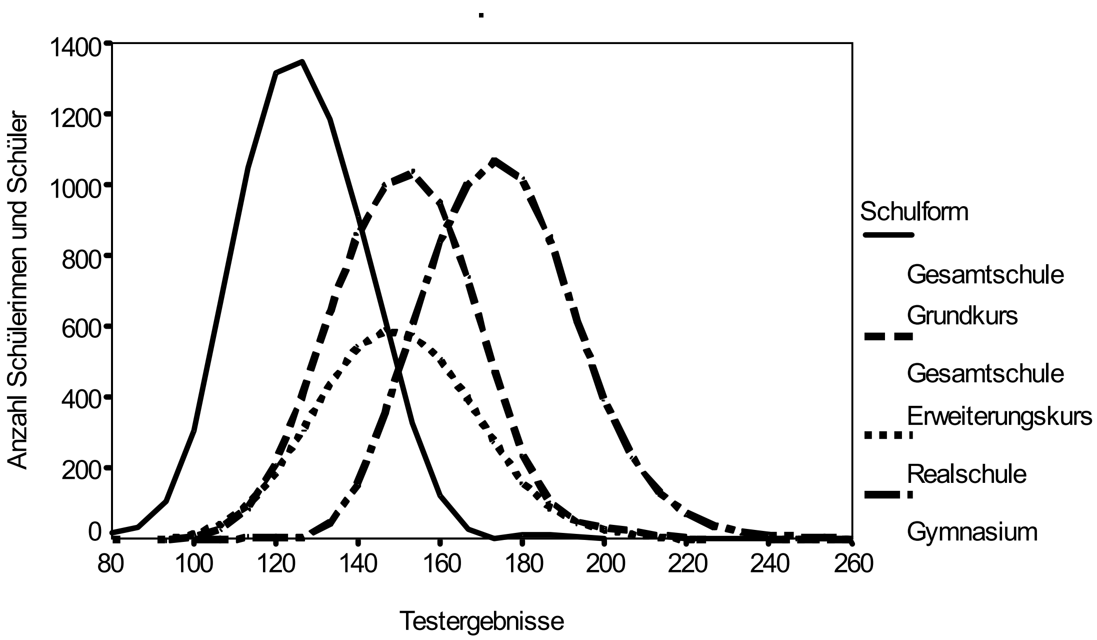
Landkreisen jeweils knapp 500 Schülerinnen und Schüler aus sechs bis sieben Schulen getestet wurden, sind doch die Fehlergrenzen so erheblich, dass Einzelvergleiche nicht sinnvoll sind. Auch lässt sich kaum ein inhaltlich interpretierbares Muster erkennen: Insgesamt sind mit der Lage der Schulen in einem bestimmten Landkreis nur 1,6 Prozent der Schülervarianz assoziiert ($Eta^2 = 0,02$).

Zu berücksichtigen ist allerdings auch, dass innerhalb der Landkreise durchaus unterschiedliche Übergangsquoten in die verschiedenen Schulformen bzw. Kursniveaus gegeben sein können, was anhand der Stichprobe kaum sinnvoll überprüft werden kann. In der Untersuchung variiert z. B. der Gymnasialanteil, teilweise sicherlich zufällig, zwischen 15 Prozent (Teltow-Fläming) und 41 Prozent (Stadt Brandenburg). In dieser Stichprobe aber gilt, dass die gemessene Durchschnittsleistung in einem Landkreis zwar erwartungsgemäß von dem dort erfassten Gymnasiastenanteil abhängt ($r = 0,25$), noch stärker allerdings von der Realschulquote in der untersuchten Schülerschaft des Kreises ($r = 0,50$). Ein Beispiel dafür liefert auch eine lokale Vollerhebung aus Bad Freienwalde, wo die 40,0 Prozent Gymnasiastinnen bzw. Gymnasiasten Ergebnisse über dem Landesdurchschnitt des gymnasialen Anteils von 27,9 Prozent erzielt haben ($d = 0,22$) und wo die 28,2 Prozent Schülerinnen und Schüler aus Realschulen sogar fast eine halbe Standardabweichung ($d = 0,49$) über dem schulformspezifischen Landesdurchschnitt liegen, ohne dass das örtliche Leistungsniveau der Schülerinnen und Schüler aus Gesamtschulen – etwa durch den sog. „Absahn-Effekt“ – nennenswert unter dem Landesdurchschnitt liegt ($d = -0,13$). Die naheliegende Folgerung, die mit 15,8 Prozent der Schülerschaft in Brandenburg am schwächsten vertretene Schulform in der Fläche zu stärken, um ein vergleichsweise leistungsfähiges und für bestimmte Schülergruppen attraktives Bildungsangebot anzubieten, lässt sich mit diesen Daten allein allerdings nicht hinreichend begründen. Im Übrigen gilt, dass den Gesamtschulen besondere Aufgaben bei der schulischen Versorgung auch kleinerer Ortschaften zukommen, wobei für die Mathematik-Grundkurse an dörflichen Schulen in Orten mit weniger als 1.000 Einwohnern keine Leistungsrückstände gegenüber jenen in größeren Ortschaften und Städten festgestellt werden konnten. Allenfalls für die Mathematik-Erweiterungskurse war eine Tendenz zugunsten der Schulen an größeren Orten zu beobachten. Diese Überlegungen führen zwingend zu der Frage, welche Schulform- und Kursniveauunterschiede im Fach Mathematik in Brandenburg in der Klassenstufe 9 festzustellen sind.

Differenzierung nach Schulform bzw. Kursniveau

Die folgende Abbildung 13 zeigt, dass die Unterschiede zwischen den Schulformen bzw. Kursniveaus – den Intentionen eines gegliederten bzw. gestuften Bildungswesens entsprechend – teilweise beträchtlich sind: Während die Leistungsverteilungen der Realschülerinnen und -schüler und die der Schülerinnen und Schüler aus Erweiterungskursen an Gesamtschulen bei einer mittleren Leistung von 150 (Standardabweichung: 18) bzw. 151 Skalenpunkten (Standardabweichung: 16) mit Ausnahme eines Proportionalfaktors einander stark ähneln, sind die Werte der Schülerinnen und Schüler aus Grundkursen an Gesamtschulen – ihre durchschnittliche Leistung liegt auf der Rasch-Skala bei 126 Punkten (Standardabweichung: 15) – demgegenüber zum unteren Bereich hin verschoben. Die dargestellten Gymnasialleistungen konzentrieren sich bei einem Mittelwert von 175 (Standardabweichung: 18) auf den oberen Bereich der Skala.

Abbildung 13 Verteilung der Schülerleistungen im QuaSUM-Mathematiktest (Klassenstufe 9), nach Schulform bzw. Kursniveau



In der folgenden Tabelle 9 wird schulform- bzw. kursniveauspezifisch für ausgewählte Skalenwerte zwischen 100 und 220 angegeben, wie viele Schülerinnen und Schüler die entsprechenden Fähigkeitsniveaus mit 50-prozentiger Wahrscheinlichkeit erreicht oder übertroffen haben.

Tabelle 9 Verteilung der Fähigkeitsniveaus der Schülerinnen und Schüler im QuaSUM-Mathematiktest (Klassenstufe 9) für ausgewählte Skalenwerte, nach Schulform bzw. Kursniveau (in Prozent)

	Skalenwert						
	bis 100	bis 120	bis 140	bis 160	bis 180	bis 200	über 200
Gesamtschulen, Grundkurse	98,0	70,1	18,8	1,9	0,2	0,0	0,0
Gesamtschulen, Erweiterungskurse	100,0	98,9	76,6	31,1	4,8	0,6	0,4
Realschulen	99,9	96,8	70,8	29,7	6,3	0,9	0,8
Gymnasien	100,0	100,0	99,0	82,8	40,3	9,5	8,2
insgesamt	99,4	89,9	62,9	35,3	12,5	2,9	0,5

Die Prozentwerte zeigen noch einmal deutlich die hohen Übereinstimmungen zwischen dem Leistungsprofil der Schülerinnen und Schüler aus Realschulen und aus Erweiterungskursen an Gesamtschulen sowie die Abstände zu den leistungsschwächeren Schülerinnen und Schülern aus Grundkursen an Gesamtschulen und zu den leistungsstärkeren Gymnasiastinnen und Gymnasiasten. Eine Bewertung der Abstände zwischen den Schulformen ist angesichts fehlender Vergleichsdatensätze mit Schwierigkeiten behaftet. Aus den Befunden kann insgesamt nicht der Schluss gezogen werden, dass in Brandenburg eine geringere Durchlässigkeit zwischen den Schulformen gegeben sei als in anderen Bundesländern[24].

Abbildung 13 ist weiterhin zu entnehmen, dass es zwischen den Leistungsverteilungen der einzelnen Schulformen und Kursniveaus deutliche Überschneidungen gibt. Um den Zusammenhang zwischen mathematischer Fähigkeit und Schul- bzw. Kursformzugehörigkeit angemessen charakterisieren

---

[24] Solche Vergleiche werden die Ergebnisse der PISA-Untersuchung bereitstellen. Für die Klassenstufe 8 haben bislang die IEA-Lesestudie (vgl. LEHMANN ET AL. 1995) und die TIMS-Studie (vgl. BEATON ET AL. 1996; BAUMERT ET AL. 1997) nahezu übereinstimmend einen Abstand des Leistungsniveaus der Hauptschulen von 0,71 bzw. 0,75 Standardabweichungen gegenüber dem allgemeinen Durchschnitt ermittelt, verglichen mit 0,96 Standardabweichungen für die Grundkurse an Gesamtschulen in der vorliegenden Untersuchung. Umgekehrt wurde in der IEA-Lesestudie ein Leistungsvorsprung der Gymnasiasten von 0,82 Standardabweichungen gegenüber dem Gesamtmittelwert gefunden, verglichen mit 0,72 Standardabweichungen in der TIMS-Studie und 1,01 Standardabweichungen in der vorliegenden Untersuchung. Hierbei ist indessen zu berücksichtigen, dass in den genannten bundesweit repräsentativen Studien wegen der bekannten erheblichen Unterschiede zwischen den Bundesländern hinsichtlich der Schulorganisation und der Bildungsbeteiligung in der Sekundarstufe I die Leistungsdifferenzen zwischen den Schulformen vermindert erscheinen.

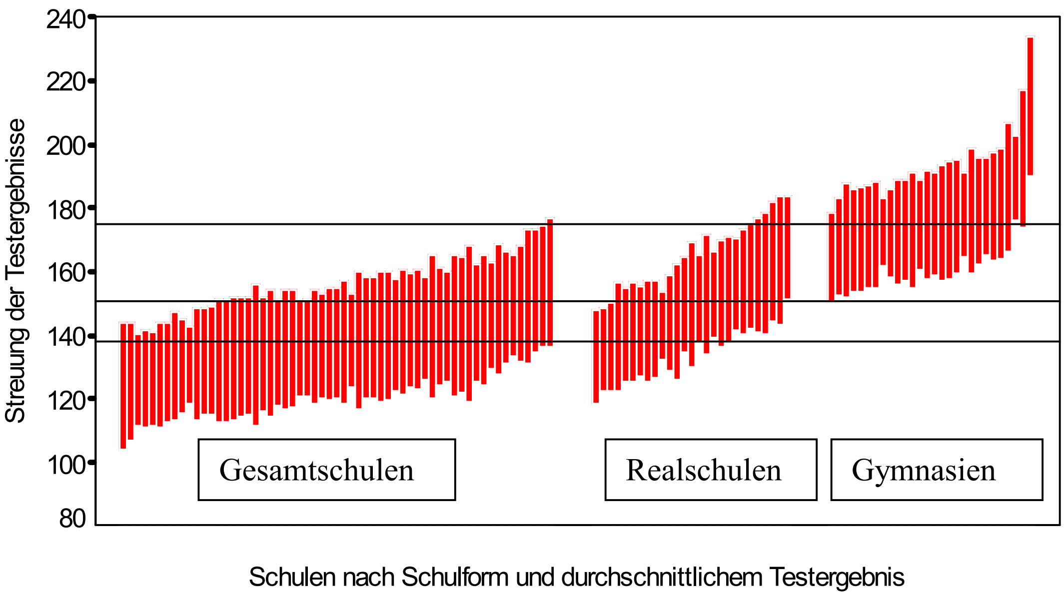
zu können, sollen hier ein – für Brandenburg – „*typisch gymnasialer Leistungsbereich*“ und ein „*typischer G-Kurs-Leistungsbereich*“ definiert werden. „*Typisch gymnasial*“ soll derjenige Leistungsbereich heißen, der durch die Schnittstelle zwischen der Leistungsverteilung in Gymnasien und jener in den Erweiterungskursen der Gesamtschulen bestimmt ist. Analog soll der „*typische G-Kurs-Leistungsbereich*“ durch die Schnittstelle der Leistungsverteilung von Mathematik-Grund- und Erweiterungskursen an Gesamtschulen festgelegt sein. In den Mathematik-Grundkursen an Gesamtschulen haben 82,7 Prozent der Schülerinnen und Schüler Testleistungen im typischen G-Kurs-Leistungsbereich erbracht, und 1,2 Prozent haben die Schwelle zum typisch gymnasialen Bereich überschritten. Die übrigen 16,1 Prozent hätte man nach den für Brandenburg charakteristischen Standards eher in Erweiterungskursen (oder an Realschulen) erwartet. Von den Schülerinnen und Schülern der Erweiterungskurse haben 25,5 Prozent das typisch gymnasiale Niveau erreicht; allerdings haben 25,0 Prozent auch die Schwelle zwischen Grund- und Erweiterungskursen unterschritten. Mithin sind 49,5 in jenem Mittelbereich zu finden, der definitionsgemäß für Erweiterungskurse (und Realschulen) charakteristisch ist. Ähnlich stellt sich die Situation an den Realschulen dar: 24,9 Prozent der Schülerinnen und Schüler dieser Schulform haben die Schwelle zum typisch gymnasialen Leistungsbereich überschritten, während 30,9 Prozent Testleistungen erbracht haben, die man in den Grundkursen der Gesamtschulen erwarten würde. So verbleiben 44,9 Prozent im mittleren Bereich. An den Gymnasien lagen 78,2 Prozent der Testleistungen auf dem schulformspezifischen Niveau, 17,9 Prozent im Mittelbereich und 3,9 Prozent sogar im typischen G-Kurs-Leistungsbereich. Auch für diese Überschneidungszonen zwischen den Schulformen bzw. Kursniveaus gilt die oben getroffene Einschränkung, dass Vergleichswerte aus bundesweit repräsentativen Untersuchungen oder aus anderen Bundesländern nicht zur Beurteilung der Frage herangezogen werden können, ob die Brandenburger Befunde hinsichtlich der Balance zwischen der Förderung unterschiedlich leistungsstarker Schülerinnen und Schüler und der Durchlässigkeit des Bildungssystems Anlass für Veränderungen auf bildungspolitischer oder pädagogisch-didaktischer Ebene bieten. Unter dem Gesichtspunkt einer gerechten Verteilung der Bildungschancen innerhalb des Landes ist diese Frage zumindest im vorliegenden Zusammenhang offen zu halten.

Differenzierung nach Schulen

Innerhalb der Schulformen gibt es in Brandenburg – wie auch anderenorts – erhebliche Unterschiede in der Mathematikleistung zwischen den Einzelschulen. Wie im Falle der Grundschulen lassen sich diese Differenzen am besten grafisch veranschaulichen (vgl. Abbildung 14).

In der Grafik sind am linken Rand die Gesamtschulen (ohne Rücksicht auf den jeweiligen Anteil der Erweiterungskurse) zusammengefasst, in der Mitte die Realschulen und am rechten Rand die Gymnasien. Jeder „Balken“ entspricht wieder einer Schule. Die Balken sind innerhalb der jeweiligen Schulform aufsteigend nach der Durchschnittsleistung angeordnet; die Länge eines Balkens entspricht der Streuung der Mathematikleistungen an dieser Schule (Mittelwert plus/minus eine Standardabweichung). Die Schulformmittelwerte sind durch waagerechte Geraden gekennzeichnet.

Abbildung 14 Verteilung der Ergebnisse im QuaSUM-Mathematiktest (Klassenstufe 9), nach Schulform und Schulen (Durchschnittswerte plus/minus eine Standardabweichung)



Auf diese Weise zeigen sich noch einmal deutlich die im vorigen Abschnitt beschriebenen Schulformunterschiede, nun aber vor allem auch die Differenzen im Einzelnen, die so groß sind, dass die schwächsten Mathematikleistungen an einer Schule den besten an einer anderen derselben Schulform entsprechen können. Dabei stellt sich selbstverständlich die durch den optischen Eindruck allein nicht zu beantwortende Frage, ob diese Heterogenität für alle Schulformen in gleichem Maße charakteristisch ist. Prüfen lässt sich dies, indem man für jede Schulform einzeln das Bestimmtheitsmaß Eta^2 berechnet, das angibt, welcher Anteil der Schülervarianz durch die Schulzugehörigkeit bestimmt ist: Je größer dieses Eta^2, desto stärker unterscheiden sich die Schulen einer Schulform hinsichtlich der Mathematikleistung voneinander. Dabei zeigt sich nun, dass die Realschule in Brandenburg mit Eta^2 = 0,24 die bei weitem heterogenste Schulform darstellt, gefolgt von der Gesamtschule und dem Gymnasium (jeweils Eta^2 = 0,13). Unterscheidet man innerhalb der Gesamtschulen das Grundkurs- vom Erweiterungskursniveau, so ist dort die Heterogenität mit Eta^2 = 0,15 bzw. Eta^2 = 0,20 erkennbar stärker ausgeprägt als an den Gymnasien, vor allem im letztgenannten Fall. Die scheinbare Ähnlichkeit der Gesamtschulen als organisatorischer Einheiten bei relativer Verschiedenartigkeit der Kursniveaus resultiert, wie sich zeigen lässt, aus zwei einander überlagernden Effekten: erstens aus dem je spezifischen Anteil der Schülerinnen und Schüler in den Erweiterungskursen, der in vielen Fällen wohl begründet sein wird, und zweitens aus unterschiedlichen schulspezifischen Standards bei der Zuweisung der Schüler zu den Kursniveaus. Somit bleibt es einerseits bei der Sonderstellung der Realschulen, deren durchschnittliches Leistungsniveau zwischen 133 Skalenpunkten („*typischer G-Kurs-Leistungsbereich*“) und 168 Skalenpunkten („*typisch gymnasialer Leistungsbereich*“) schwanken kann. Andererseits erscheint das Gymnasium als diejenige Schulform, in der am ehesten nach einheitlichen Standards gearbeitet wird.

Exkurs: Ergebnisse der Schulen mit besonderer Prägung sowie der Schulen aus Modellversuchen

Bei der Betrachtung von Schulen besonderer Prägung, von denen eine Sorbisch als Unterrichtssprache verwendet, vier einen sportlichen und zwei einen mathematischen Schwerpunkt haben, ist es ebenso wie bei den Schulen, an denen ein Modellversuch des SINUS-Programms läuft, erforderlich, zwischen den jeweiligen Schulformen zu unterscheiden. In Tabelle 10 werden die Ergebnisse im QuaSUM-Mathematiktest und zusätzlich die im Mathe40-Test nach Schulform und jeweiliger Sonderform zusammengestellt.

Tabelle 10 Ergebnisse im QuaSUM-Mathematiktest und im Mathe40-Test (in Klammern), nach Schulform und jeweiliger Sonderform

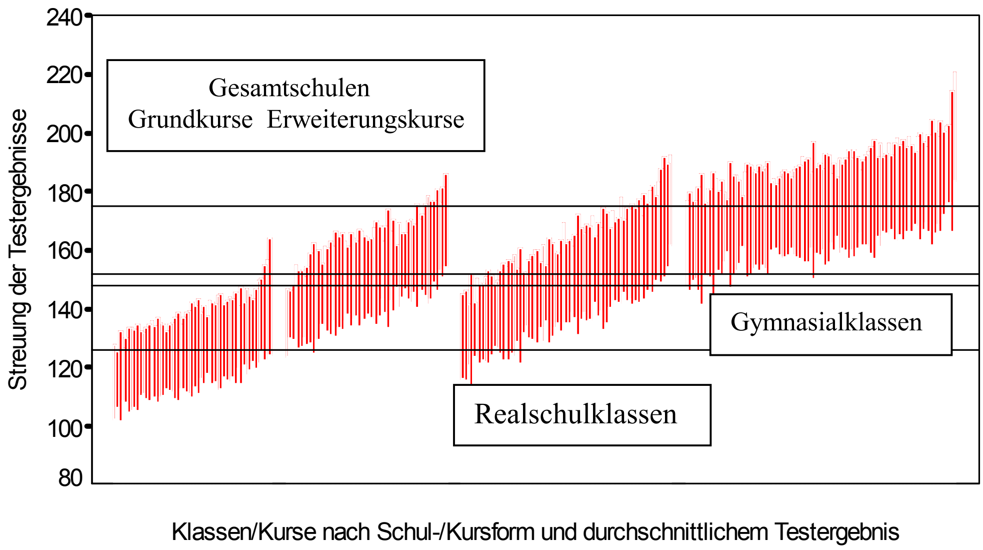
	Gesamtschulen		Realschulen	Gymnasien
	Grundkurse	**Erweiterungskurse**		
SINUS-Modellversuch	129,4 (442,9)	150,9 (507,2)	143,7 (505,7)	
nichtmathematischer Schwerpunkt	129,8 (450,2)	157,8 (531,3)	---	173,0 (592,7)
mathematischer Schwerpunkt	---	---	---	212,1 (714,1)
mathematischer Schwerpunkt, gleichzeitig SINUS-Modellversuch	---	---	---	195,5 (647,2)
Landesdurchschnitt	*126,0 (434,6)*	*151,1 (501,2)*	*149,5 (495,9)*	*175,2 (576,5)*

Im Hinblick auf die Modellversuche des SINUS-Programms – diese haben erst im Schuljahr 1998/99 in der Jahrgangsstufe 7 begonnen – konnten in der QuaSUM-Untersuchung noch keine Effekte erwartet werden. Tatsächlich lassen sich hier noch keine besonderen Erfolge verzeichnen und statistisch eindeutig absichern. Immerhin aber kann im Grundkursbereich eine vergleichsweise erfolgreiche Tendenz festgestellt werden ($d = 0{,}23$), und zwar hier vor allem bei dem stärker auf mathematische Routineoperationen abgestimmten QuaSUM-Mathematiktest. Interessanterweise lassen aber die Mathe40-Testergebnisse für alle Schulen mit einem SINUS-Modellversuch auf mathematische Fähigkeiten über dem jeweiligen Landesdurchschnitt schließen. Für die Schulen besonderer Prägung mit sportlicher bzw. auf die Pflege des Sorbischen ausgerichteter Orientierung lassen sich keine Besonderheiten bezogen auf die Mathematikleistungen der Schülerschaft feststellen. An den beiden Gymnasien mit einem mathematischen Schwerpunkt ist ein Leistungsniveau zu beobachten, das weit über dem gymnasialen Landesdurchschnitt liegt. Dies wurde in Abbildung 14 bereits optisch erkennbar, in der diese Schulen am rechten Rand der Gymnasialverteilung liegen. An einer dieser beiden Schulen wurden im QuaSUM-Mathematiktest durchschnittliche Ergebnisse erzielt, die den gymnasialen Landesmittelwert um $d = 1{,}15$ Standardabweichungen übertreffen, an der anderen beträgt der entsprechende Abstand sogar $d = 2{,}10$ Standardabweichungen. Dies entspricht einem Lernvorsprung von mehreren Schuljahren. Nun hätte man vielleicht erwartet, dass die Vorteile in dem Bereich, der weniger an das Brandenburger Curriculum gebunden ist, also im Mathe40-Test, noch größer wären. Dies ist indessen nicht der Fall: Die entsprechenden Effektstärken betragen hier $d = 0{,}88$ bzw. $d = 1{,}71$ Standardabweichungen.

Differenzierung nach Schulklassen bzw. Kursen

Wie sich die Mathematikleistungen der Schulklassen innerhalb der Schulformen verteilen, zeigt Abbildung 15.

Abbildung 15 Verteilung der Ergebnisse im QuaSUM-Mathematiktest (Klassenstufe 9), nach Schulform/Kursniveau und Klassen/Kursen (Durchschnittswerte plus/minus eine Standardabweichung)



Ein geeigneter Indikator für die Bedeutung der Zugehörigkeit zu einer bestimmten Schulklasse bzw. zu einem bestimmten Kurs ist – wie auf der Schulebene – das Bestimmtheitsmaß Eta^2, das sich auf den jeweiligen Anteil an der Schülervarianz bezieht. Notwendigerweise nimmt dieses Maß im Falle der Schulklassen bzw. Kurse einen höheren Wert als auf der Schulebene an. Die Differenz kann man als obere Grenze dafür interpretieren, was auf klassen- bzw. kurs- und unterrichtsspezifische Besonderheiten zurückgeführt werden kann. In der nachfolgenden Tabelle 11 sind deshalb für die Schulformen bzw. Kursniveaus die entsprechenden Werte zusammengestellt.

Es zeigt sich an den Differenzen zwischen schul- und klassenspezifischer Varianz – mit der sogleich noch zu diskutierenden Ausnahme der Erweiterungskurse an Gesamtschulen –, dass ein Anteil von ca. 8 Prozent der Schülervarianz auf Klassenebene anzusiedeln ist. Dieser Wert ist zwar geringer als das Eta^2 auf der Ebene der Einzelschulen, doch gilt es zu berücksichtigen, dass in der Zwischen-Schul-Varianz ein erheblicher Anteil verborgen ist, der auf das örtliche Einzugsgebiet und die lokale „Wettbewerbssituation“ der Schule verweist.

Tabelle 11 Anteile an der Schülervarianz (QuaSUM-Mathematiktest, Klassenstufe 9), die durch die Schul- und Klassen- bzw. Kurszugehörigkeit festgelegt sind (Bestimmtheitsmaß Eta^2), nach Schulform bzw. Kursniveau

	Schulebene	**Klassenebene**	**Differenz**
Gesamtschulen, Grundkurse	0,15	0,23	0,08
Gesamtschulen, Erweiterungskurse	0,20	0,32	0,12
Realschulen	0,24	0,33	0,09
Gymnasien	0,13	0,20	0,07

Deshalb bestätigen diese Daten die Vermutung, dass im Fach Mathematik – und mutmaßlich auch in anderen Fächern – solchen Bemühungen um eine Verbesserung des Bildungsangebots, die vor allem auf „Gestaltung durch Organisationsentwicklung“ (ROLFF 1995) setzen, dabei aber die Ebene des Unterrichts selbst nicht erreichen, der Erfolg letztlich versagt bleiben wird. Dies lässt sich illustrieren am Beispiel der beiden Gymnasien besonderer Prägung mit mathematischem Schwerpunkt, von denen bereits die Rede war. An einer dieser Schulen existieren vier Klassen, die im QuaSUM-Mathematiktest mit durchschnittlich 201 Skalenpunkten Spitzenergebnisse erzielt haben, neben einer Parallelklasse, deren mittlere Punktzahl im Bereich des gymnasialen Landesdurchschnitts liegt und die also nichts von den Erfolgen in den Nachbarklassen erkennen lässt (d = 1,35). Im Mathe40-Test, wo eine der erstgenannten Klassen das beste Ergebnis überhaupt erreicht hat, klaffen die Resultate ähnlich weit auseinander (d = 1,12). Nicht ganz so krass stellen sich die Befunde für die andere gymnasiale Spezialschule dar, aber auch hier bleibt eine der Klassen im QuaSUM-Mathematiktest um etwa eine Standardabweichung hinter den Parallelklassen zurück (d = 1,13), nicht hingegen im Mathe40-Test (d = 0,02). Von den Eingangsvoraussetzungen her, die über die kognitiven Grundfähigkeiten (Ergebnisse im CFT 20) geprüft wurden, sind diese Diskrepanzen nicht zu erklären. Dann aber liegt die Annahme nahe, dass durch die Unterrichtsführung in den Klassen einer Schule unter ähnlichen Ausgangsbedingungen sehr unterschiedliche Leistungsniveaus und -profile erzielt werden können.

Eine besondere Bemerkung verdient noch der auffällig hohe kursspezifische Varianzanteil in den Erweiterungskursen der Gesamtschulen. Es ist bereits darauf hingewiesen worden, dass sich die Gesamtschulen – nahezu unabhängig von der Größe der Schule – deutlich darin unterscheiden, welcher Anteil der Schülerschaft den Erweiterungskursen zugewiesen wird. Die Anforderungen, die eine Schülerin oder ein Schüler erfüllen muss, um einen solchen Kurs besuchen zu können, sind an den kleineren Gesamtschulen eher höher als an den größeren (r = -0,23). Außerdem sind – erwartungsgemäß – die Anforderungen bei einem höheren Anteil der Schülerschaft in Erweiterungskursen tendenziell niedriger (r = -0,27). So gilt zwar ceteris paribus, dass die Mathematikleistungen an Gesamtschulen mit dem Anteil der Schülerschaft in Erweiterungskursen und mit den geforderten Standards bei der Kurszuweisung steigen. Angesichts der negativen Korrelation zwischen diesen beiden Kriterien entsteht aber die bereits auf Schulebene beschriebene und nunmehr auf Klassenebene bestätigte Heterogenität.

Unter dem Gesichtspunkt schulorganisatorischer Merkmale ergibt sich zusätzlich die Frage, ob und inwieweit sich die Mathematikleistungen der Schülerinnen und Schüler aus Gesamtschulen mit bzw. ohne gymnasiale Oberstufe unterscheiden. Vergleicht man die Leistungen unter dem Aspekt des Vorhandenseins (13 Schulen) bzw. Nichtvorhandenseins (46 Schulen) einer gymnasialen Oberstufe, zeigen sich keine Unterschiede: Sowohl die Grundkursler als auch die Erweiterungskursler weisen im QuaSUM-Mathematiktest nahezu gleiche Ergebnisse auf (vgl. Tabelle 12). Der insgesamt etwas höhere Mittelwert für die untersuchten Schülerinnen und Schüler an Gesamtschulen mit gymnasialer Oberstufe hängt auch damit zusammen, dass hier der relative Anteil von Erweiterungskurslern höher liegt[25].

Tabelle 12 Testleistungen der Schülerinnen und Schüler aus Gesamtschulen im QuaSUM-Mathematiktest (Klassenstufe 9), nach Vorhandensein bzw. Nichtvorhandensein einer gymnasialen Oberstufe an der Schule (mit standardisierten Effektstärken *d*)

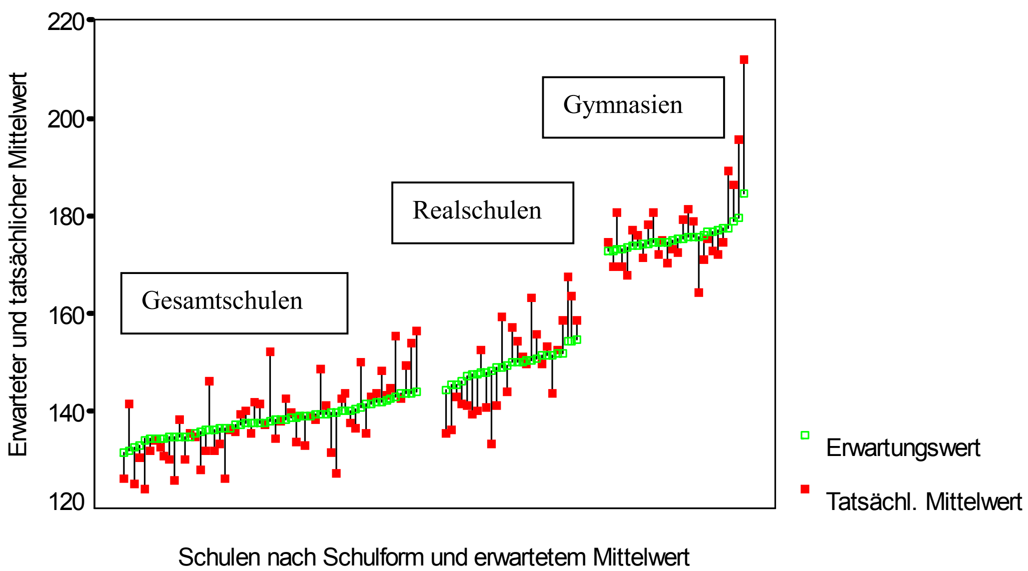
Gesamtschulen	**mit gymnasialer Oberstufe**		**ohne gymnasiale Oberstufe**		**Effektstärke *d***
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung	
Mathematik-Grundkurse	127	15	126	15	0,04
Mathematik-Erweiterungskurse	152	18	151	16	0,05
insgesamt	142	21	136	19	0,32

[25] Das gleiche Ergebnis zeigt sich im Mathe40-Test. Auch hier liegen die erreichten Durchschnittwerte der Grundkursler (436 bzw. 434) und der Erweiterungskursler (503 bzw. 500) eng beieinander; insgesamt haben die Schülerinnen und Schüler aus Gesamtschulen mit gymnasialer Oberstufe einen Vorteil von 478 gegenüber 461 Testpunkten.

Das Verhältnis von Mathematik-Erweiterungskurslern (61,3 Prozent) gegenüber Mathematik-Grundkurslern (38,7 Prozent) an Gesamtschulen mit gymnasialer Oberstufe dreht sich an den Gesamtschulen ohne gymnasiale Oberstufe um.

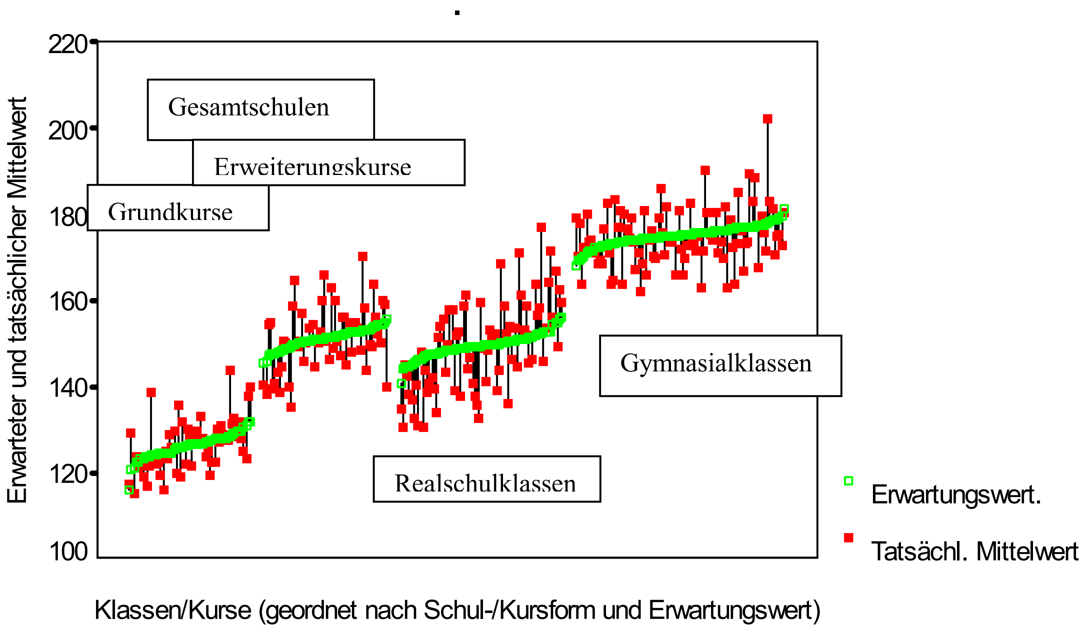
Wie für die Klassenstufe 5 lassen sich auch für die Klassenstufe 9 die erzielten Mathematikleistungen in einzelnen Schulen und in einzelnen Klassen bzw. Kursen zu denjenigen in Beziehung setzen, die nach Brandenburger Maßstäben zu erwarten gewesen wäre. Durch die außerschulischen bzw. außerunterrichtlichen Schülermerkmale Fähigkeit zum schlussfolgernden Denken (Gesamtwert im CFT 20), leistungsbezogenes Selbstkonzept, höchster erreichter Schulabschluss der Eltern bzw. des erziehenden Elternteils und Buchbestand im Elternhaus können je nach Schulform/Kursniveau bis zu 34 Prozent der Unterschiede zwischen den Mathematikleistungen der Neuntklässlerinnen und Neuntklässler erklärt werden. In der folgenden Abbildung 16 sind – nach dem Ordnungsmuster der Schulformen – die tatsächlichen Schulergebnisse den bei der gegebenen Zusammensetzung der Schülerschaft erwarteten Ergebnissen gegenübergestellt, wobei die Erwartungswerte für die Schulen bzw. Klassen jeweils getrennt nach Schulform bzw. Kursniveau berechnet sind. Während auf der einen Seite für viele Schulen der erwartete und der tatsächliche Leistungswert eng beieinander liegen, weisen mehrere Schulen deutliche Abweichungen auf. Die tatsächlichen Leistungen liegen also deutlich höher bzw. niedriger, als es für die Schülerschaft dieser Schule erwartbar war.

Abbildung 16 Testergebnisse im QuaSUM-Mathematiktest (Klassenstufe 9) nach Schulform und Schulen: Vergleich der erwarteten und der tatsächlichen Schulmittelwerte



Auf der Ebene von Klassen bzw. Kursen einer Schulform zeigt sich ein sehr ähnliches Bild: In allen vier untersuchten Schul- bzw. Kursformen bestehen neben weitgehenden Übereinstimmungen zwischen erwartetem und tatsächlichem Testergebnis auch deutliche Abweichungen (vgl. Abbildung 17).

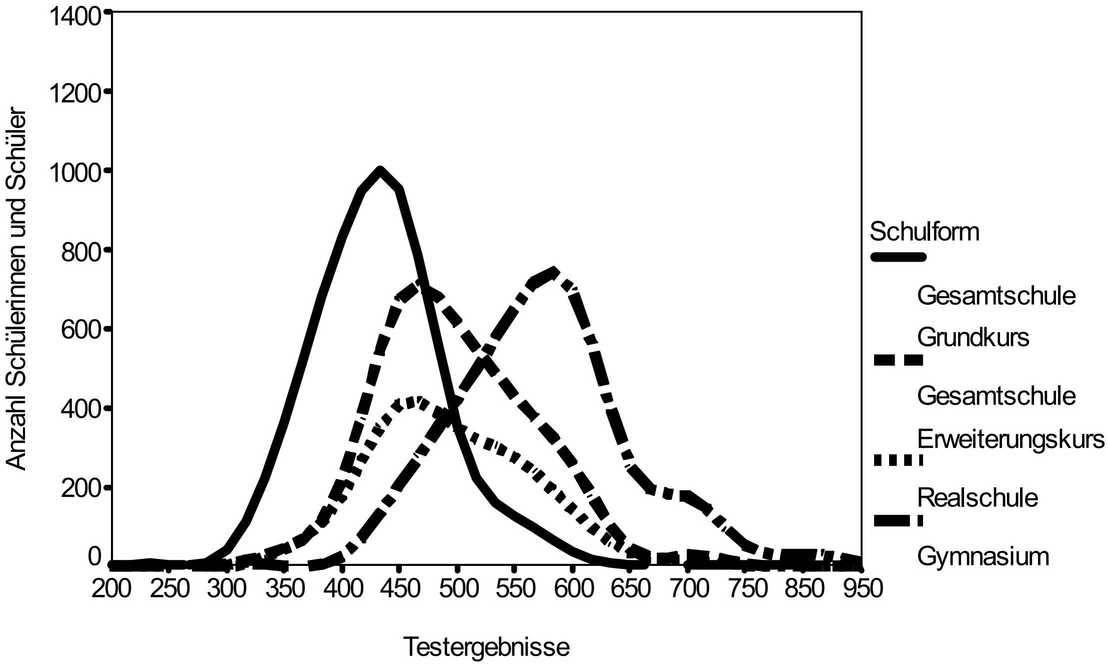
Abbildung 17 Testergebnisse im QuaSUM-Mathematiktest (Klassenstufe 9) nach Schulform/Kursform und Klassen/Kursen: Vergleich der erwarteten und der tatsächlichen Klassen- bzw. Kursmittelwerte



3.2.3 Ergebnisse der Neuntklässler im Mathe40-Test

Die Ergebnisse im QuaSUM-Mathematiktest und im Mathe40-Test, der auf veröffentlichten Aufgaben der TIMS-Studie beruht, korrelieren relativ hoch miteinander (r = 0,79). Deshalb zeigen sich im Mathe40-Test, bezogen auf die Schulformen bzw. Kursniveaus, die bereits bekannten Differenzen, wenn auch wegen der geringeren Curriculumnähe des letzteren ein wenig abgeschwächt: Die Leistungsverteilungen der Realschülerinnen und -schüler (Mittelwert: 496; Standardabweichung: 71) und die der Schülerinnen und Schüler aus Erweiterungskursen an Gesamtschulen (Mittelwert: 501; Standardabweichung: 67) liegen wieder eng beieinander, wenn man den Proportionalitätsfaktor berücksichtigt, der sich aus den unterschiedlichen Schülerzahlen ergibt. Die Leistungen der Schülerinnen und Schüler aus Grundkursen an Gesamtschulen sind auch im Mathe40-Test nach unten verschoben (Mittelwert: 435; Standardabweichung: 57), die der Gymnasiastinnen und Gymnasiasten nach oben (Mittelwert: 577; Standardabweichung: 80). Die Leistungsüberschneidungen zwischen den einzelnen Schul- bzw. Kursformen sind hier – wie im QuaSUM-Test – bemerkenswert.

Abbildung 18 Verteilung der Schülerleistungen im Mathe40-Test (Klassenstufe 9), nach Schulform bzw. Kursniveau



In tabellarischer Form kann gezeigt werden, wie viel Prozent der Schülerinnen und Schüler insgesamt und nach Schul- bzw. Kursformen bestimmte Fähigkeitssniveaus erreicht bzw. überschritten haben (vgl. Tabelle 13).

Tabelle 13 Verteilung der Fähigkeitsniveaus der Schülerinnen und Schüler im Mathe40-Test (Klassenstufe 9) für ausgewählte Skalenwerte, nach Schulform bzw. Kursniveau (in Prozent)

	Skalenwert								
	bis 350	bis 400	bis 450	bis 500	bis 550	bis 600	bis 650	bis 700	über 700
Gesamtschulen, Grundkurse	96,9	78,3	46,3	14,5	4,2	0,9	0,1	0,1	0,1
Gesamtschulen, Erweiterungskurse	99,6	96,4	84,4	52,9	27,5	11,3	3,3	1,4	0,7
Realschulen	99,3	94,5	80,5	50,4	27,5	9,7	3,6	1,2	0,6
Gymnasien	100,0	99,8	98,0	87,8	68,4	40,6	20,8	10,7	6,6
insgesamt	98,8	91,6	76,1	50,7	32,0	16,1	7,3	3,6	2,1

Innerhalb des brandenburgischen Kontexts sind diese Zahlen von hohem Interesse; denn sie ermöglichen es, die Leistungen der Schülerinnen und Schüler nach Maßgabe der national und international definierten Fähigkeitsniveaus zu interpretieren, so wie dies in erster Annäherung bereits oben für die Ergebnisse aus dem QuaSUM-Mathematiktest geschehen ist. So verfügen etwa drei Viertel der Schülerinnen und Schüler aus Grundkursen an Brandenburger Gesamtschulen mit hinreichender Sicherheit – d. h. mit einer Lösungswahrscheinlichkeit von 0,65 oder darüber – über die „elementaren Rechenkenntnisse" im Sinne von TIMSS (bis 405 Skalenpunkte; vgl. BAUMERT ET AL. 1997, S. 79). Etwas mehr als ein Viertel der Schülerinnen und Schüler aus Erweiterungskursen der Gesamtschulen bzw. aus Realschulen beherrscht mit derselben Zuverlässigkeit die in TIMSS so bezeichneten „Routineverfahren" (bis 550 Skalenpunkte) und ca. ein Drittel der Schülerinnen und Schüler aus Gymnasien besitzt ein „Verständnis von mathematischen Konzepten und Verfahren" (550 Skalenpunkte oder mehr).

Ein über die Grenzen des Landes Brandenburg hinausgehender Vergleich mit den verfügbaren TIMSS-Ergebnissen, die für das Ende der Klassenstufe 8 bundesweit repräsentativ vorliegen, ist aus einer Reihe von Gründen – zumindest auf dem gegenwärtigen Kenntnisstand – *unzulässig*. Naheliegende methodische Einwände gegen einen solchen Versuch ergeben sich vor allem aus den bekannten Unterschieden zwischen den Bundesländern hinsichtlich der Schulorganisation und der Bildungsbeteiligung, ferner aus dem unterschiedlichen Testzeitpunkt in den beiden Studien. Wenn auch die Brandenburger Ergebnisse (Klassenstufe 9) den bundesweiten Befunden für das Ende der Klassenstufe 8 in guter Näherung entsprechen, so kann daraus vorläufig nicht der Schluss gezogen werden, die brandenburgischen Schülerinnen und Schüler seien hinter ihren Altersgenossen in anderen Bundesländern in der Entwicklung mathematischer Fähigkeiten zurück. Denkbar ist, dass der TIMSS-Test, der international für die Zielgruppe der 13-jährigen entwickelt worden ist, für die Klassenstufen 7 und 8 valider ist als für die neunte Klassenstufe. Gewisse – im internationalen Vergleich erkennbare – Schwächen deutscher Schülerinnen und Schüler beim kumulativen Lernen sowie Vergesseneffekte können zu einer Unterschätzung der Leistungsstände am Ende der Klassenstufe 8 führen, wenn man es unternähme, die mit dem Mathe40-Test festgestellten Leistungsstände aus der Klassenstufe 9 auf der Basis der in TIMSS beobachteten durchschnittlichen jährlichen Zuwachsrate für die Entwicklung mathematischer Fähigkeit einfach rückzuprojizieren.

So muss trotz der inhaltlich ergiebigen Verzahnung der QuaSUM-Untersuchung mit TIMSS die Frage nach der relativen Position Brandenburger Schülerinnen und Schüler im nationalen Vergleich vorerst unbeantwortet bleiben. Mit den von der PISA-Studie zu erwartenden Ergebnissen wird auch für das Fach Mathematik eine geeignetere Vergleichsgrundlage zur Verfügung stehen, und es bleibt abzuwarten, welche Schlussfolgerungen dann gezogen werden können, sowohl hinsichtlich des Vergleichs zwischen den Bundesländern als auch im Hinblick auf das Verhältnis zwischen der ausgesprochen curriculumnahen Leistungsmessung bei QuaSUM und der stärker transferorientierten Untersuchung mathematischer Fähigkeiten bei PISA.

4 Individueller und sozialer Kontext der Mathematikleistungen

Im Kapitel 4 wird die Stichprobe hinsichtlich individueller und sozialer Merkmale – konkret hinsichtlich kognitiver Voraussetzungen der Schülerinnen und Schüler, ihres Alters, ihres Geschlechts, ihrer familialen Situation und schließlich hinsichtlich ihrer lern- und schulbezogenen Einstellungen – beschrieben, und es werden Zusammenhänge zwischen diesen Merkmalen und den mathematischen Fähigkeiten der Schülerinnen und Schüler analysiert.

4.1 Kognitive Voraussetzungen

Die allgemeinen kognitiven Lernvoraussetzungen wurden in beiden Klassenstufen durch den *Culture Fair Intelligence Test – CFT 20* in seiner Kurzform mit Aufgaben zum schlussfolgernden Denken erhoben (vgl. Abschnitt 2.1). Die Reliabilität der Testbatterie mit 46 Items erreichte für beide Klassenstufen einen Wert von r_{tt} = 0,81[26], wobei die Kennwerte für die Substichproben *Gymnasien* und *Gesamtschulen, Erweiterungskurse* – vor allem wegen der eingeschränkten Varianz und in Übereinstimmung mit den Eichuntersuchungen – mit jeweils r_{tt} = 0,73 niedriger lagen.

Tabelle 14 enthält die durchschnittliche Anzahl gelöster CFT 20-Aufgaben in der QuaSUM-Untersuchung im Vergleich zu den Klassenstandardwerten der Eichstichprobe.

Tabelle 14 Vergleich der Testergebnisse im CFT 20 in der QuaSUM-Stichprobe und in der bundesweiten Eichstichprobe (Mittelwerte und Standardabweichungen)

	QuaSUM-Stichprobe		**Eichstichprobe**	
	Mittelwert	Standardabweichung	Mittelwert	Standardabweichung
Klassenstufe 5	28,3	6,4	28,6	5,8
Klassenstufe 9				
Gesamtschulen, Grundkurse	28,6	6,3	---	---
Gesamtschulen, Erweiterungskurse	33,1	5,0	---	---
Realschulen	32,4	5,4	---	---
Gymnasien	35,6	4,7	---	---
insgesamt	*32,5*	*6,0*	*31,0*	*6,5*

[26] Für eine bundesweite Eichstichprobe liegen Normentabellen und Schulstandardwerte für das 3. bis 10. Schuljahr vor. Die bei WEIß (1997) für die Klassenstufe 5 berichtete Reliabilität liegt bei r_{tt} = 0,90, diejenige für die Klassenstufe 9 vermutlich in vergleichbarer Höhe.

Der Vergleich der Durchschnittswerte zeigt, dass sich die Brandenburger Schülerinnen und Schüler wenig von der bundesweiten Eichstichprobe unterscheiden, tendenziell aber in der 9. Klassenstufe etwas günstigere Resultate erzielt haben ($d = 0{,}23$). Im Übrigen fällt auf, dass – abgesehen von der allgemeinen altersabhängigen Zunahme – die Streuung der kognitiven Grundfähigkeiten in der Sekundarstufe, verglichen mit der Grundschule, im brandenburgischen *Gesamtsystem* nicht zunimmt. *Innerhalb der Schulformen* sind mit Ausnahme der Grundkurse an den Gesamtschulen im Zuge der verschiedenen Übergänge an weiterführende Schulen – wie es zu erwarten war – deutlich homogenere Lerngruppen entstanden.

In der Klassenstufe 5 lösen die Mädchen im CFT 20 durchschnittlich eine Aufgabe mehr als die Jungen[27]. In der Klassenstufe 9 zeigen sich hinsichtlich der erreichten Ergebnisse im CFT 20 *insgesamt* keine Geschlechterunterschiede; das schließt aber Unterschiede innerhalb der Schulformen bzw. Kursniveaus nicht aus. Die Überrepräsentanz der Mädchen an den Gymnasien – sie machen dort 59,8 Prozent der Schülerschaft aus – geht einher mit einem etwas höheren CFT-20-Wert der Jungen an dieser Schulform ($d = -0{,}08$); dieser Effekt wird übertroffen von den Unterschieden in den anderen Teilstichproben (Grundkurse: $d = -0{,}14$; Erweiterungskurse: $d = -0{,}16$; Realschulen: $d = -0{,}15$). Offenbar ist also die ungleiche Verteilung der beiden Geschlechter auf die Schulformen und Kursniveaus auch von anderen Merkmalen als den nichtverbalen kognitiven Fähigkeiten bestimmt. Die beobachteten schulformspezifischen Differenzen kommen dann so zustande, dass für die Jungen im Hinblick auf die gemessene Intelligenz die „Schwelle" für den Übergang in eine anspruchsvollere Schulform oder ein anspruchsvolleres Kursniveau höher liegt. Will man nicht eine Benachteiligung der Jungen unterstellen, so müssen umgekehrt die Mädchen Fähigkeiten und Verhaltensweisen aufweisen, die von den Lehrkräften besonders geschätzt werden und die vielleicht auch in manchen Fällen kompensierend eingesetzt werden können.

Wie erwartet hängen die allgemeinen kognitiven Voraussetzungen deutlich mit den erreichten Mathematikleistungen der Schülerinnen und Schüler zusammen. Die folgende Tabelle 15 gibt für die getesteten Schülerinnen und Schüler die schulformspezifischen Korrelationen zwischen dem CFT-20-Wert und dem erreichten Wert im QuaSUM- bzw. Mathe40-Test an.

---

[27] Möglichen Gründen hierfür wurde systematisch nachgegangen. Es konnten keine Unterschiede zwischen den Jungen und den Mädchen in der Arbeitsgeschwindigkeit (Bearbeitungsgrad der Testaufgaben) festgestellt werden. Somit kommen als Gründe differenzielle Reifung und/oder Unterschiede in der Arbeitseinstellung in Betracht.

Tabelle 15 Korrelationen zwischen kognitiven Voraussetzungen (Ergebnis im CFT 20) und Mathematikleistung, nach Schulform bzw. Kursniveau

	QuaSUM-Mathematiktest Klassenstufe 5	QuaSUM-Mathematiktest Klassenstufe 9	Mathe40-Test
Klassenstufe 5	0,57	--	--
Klassenstufe 9			
Gesamtschulen, Grundkurse	--	0,45	0,45
Gesamtschulen, Erweiterungskurse	--	0,42	0,45
Realschulen	--	0,50	0,51
Gymnasien	--	0,42	0,47
insgesamt	--	*0,53*	*0,60*

Der Mathe40-Test ist nur am Gymnasium und damit auch in der Gesamtstichprobe etwas höher mit den kognitiven Grundfähigkeiten korreliert als der QuaSUM-Mathematiktest. Dies ist ein Indiz dafür, dass die Ergebnisse in *beiden* Tests nicht nur von einer allgemeinen und sprachfreien Fähigkeit zum schlussfolgernden Denken, sondern wohl auch von unterrichtsnahen Lernstrategien beeinflusst sind. Umgekehrt spielt „Begabung“ hinsichtlich der erfolgreichen Bearbeitung des QuaSUM- und des Mathe40-Tests eine ähnlich gewichtige Rolle.

4.2 Alter und Geschlecht

Zum Testzeitpunkt waren die Schülerinnen und Schüler der Klassenstufe 5 durchschnittlich 11 Jahre und 7 Monate (Standardabweichung: 5 Monate) und in der Klassenstufe 9 im Durchschnitt 15 Jahre und 9 Monate (Standardabweichung: 6 Monate) alt. Im Vergleich der Schul- bzw. Kursformen in der Klassenstufe 9 treten im Durchschnitt charakteristische Altersunterschiede auf: Wegen der in den Grundkursen an Gesamtschulen häufiger festzustellenden Klassenwiederholungen (und vielleicht auch verspäteter Einschulungen) sind die Schülerinnen und Schüler mit 15 Jahren und 11 Monaten hier durchschnittlich etwa 4 Monate älter als die in den anderen Schulformen bzw. Kursniveaus. Die Altersspannen reichen über die Klassenstufen und die Schulformen bzw. Kursniveaus hinweg jeweils über vier bis fünf Jahre, wobei die Streuung in den Grundkursen der Gesamtschulen am höchsten ist (vgl. Tabelle 16).

Tabelle 16 Statistische Kennwerte der Stichproben für das Merkmal Alter

	Mittelwert	**Standard-abweichung**	**Minimum**	**Maximum**
Klassenstufe 5	11; 7*	0; 5	9; 9	14; 1
Klassenstufe 9				
Gesamtschulen, Grundkurse	15; 11	0; 8	14; 3	18; 5
Gesamtschulen, Erweiterungskurse	15; 8	0; 5	13; 10	18; 0
Realschulen	15; 8	0; 6	14; 2	18; 4
Gymnasien	15; 7	0; 4	13; 0	18; 1
insgesamt	*15; 9*	*0; 6*	*13; 0*	*18; 5*

* (Jahre; Monate)

In der Regel handelt es sich in beiden Klassenstufen bei den deutlich älteren Schülerinnen und Schülern um solche, die – offensichtlich bedingt durch schwache Lernleistungen – in ihrer bisherigen Schullaufbahn einmal oder mehrfach ein Schuljahr wiederholt haben. In den Klassenstufen 3, 4 und 5 waren dies 2,6 Prozent[28], in den Klassenstufe 7, 8 und 9 mit 6,3 Prozent der Schülerinnen und Schüler deutlich mehr (13,6 Prozent aus den Mathematik-Grundkursen der Gesamtschulen, 2,4 Prozent aus den Mathematik-Erweiterungskursen, 7,4 Prozent aus den Realschulen und 1,3 Prozent aus den Gymnasien). Zwar sind auch die Schülerinnen und Schüler, die nicht die deutsche Staatsangehörigkeit besitzen, im Durchschnitt je nach Schulform um bis zu ein Jahr älter als ihre deutschen Klassenkameraden, doch mit einem Umfang von 0,4 Prozent (Klassenstufe 5) bzw. 0,6 Prozent (Klassenstufe 9) ist diese Gruppe so klein, dass die Auswirkungen auf die Altersstruktur insgesamt vernachlässigt werden können.

Vor diesem Hintergrund war zu erwarten, dass die älteren Schülerinnen und Schüler geringere Mathematikleistungen zeigen. In der Tat belegen die korrelativen Zusammenhänge zwischen Alter und Mathematikleistung(en) für die Klassenstufe 5 (QuaSUM-Mathematiktest: $r = -0,15$) und für die Klassenstufe 9 (QuaSUM-Mathematiktest: $r = -0,26$; Mathe40-Test: $r = -0,23$), dass die jüngeren Schülerinnen und Schüler durchschnittlich höhere Leistungen erbringen.

Das Merkmal *Geschlecht* ist in den Stichproben für beide Klassenstufen erwartungsgemäß fast gleich verteilt: In der Klassenstufe 5 sind 50,7 Prozent und in der Klassenstufe 9 sind 50,1 Jungen; dementsprechend sind 49,3 bzw. 49,9 Prozent Mädchen. Im Vergleich der Schul- und Kursformen in der Klas-

[28] Im Land Brandenburg ist für Grundschülerinnen und -schüler kein sog. Sitzenbleiben vorgesehen. Die Schülerinnen und Schüler können bei „schwachen“ Leistungen auf Antrag der Eltern zurückgestellt werden.

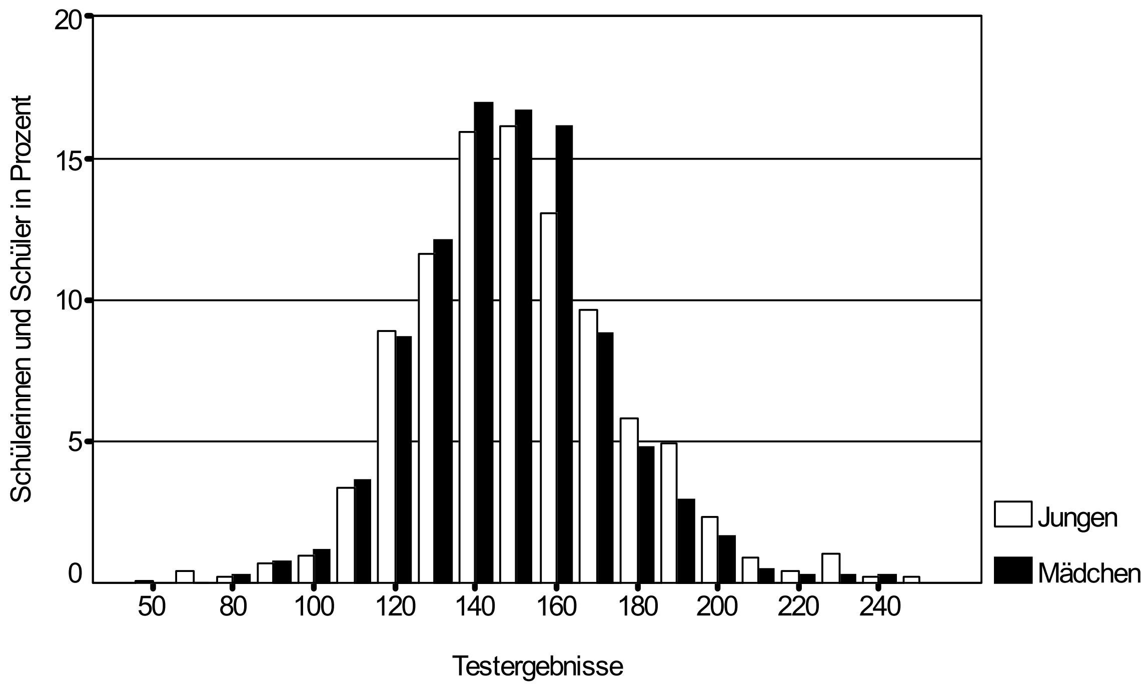
senstufe 9 sind Mädchen in den leistungsschwächeren Grundkursen an Gesamtschulen unter- und in den erheblich leistungsstärkeren Klassen an Gymnasien überrepräsentiert (vgl. Tabelle 17).

Tabelle 17 Schülerinnen und Schüler der Klassenstufe 9, nach Schul- bzw. Kursform und Geschlecht (in Prozent)

	Jungen	**Mädchen**
Gesamtschulen, Grundkurse	57,2	42,8
Gesamtschulen, Erweiterungskurse	51,6	48,4
Realschulen	50,8	49,2
Gymnasien	40,2	59,8
insgesamt	*50,7*	*49,3*

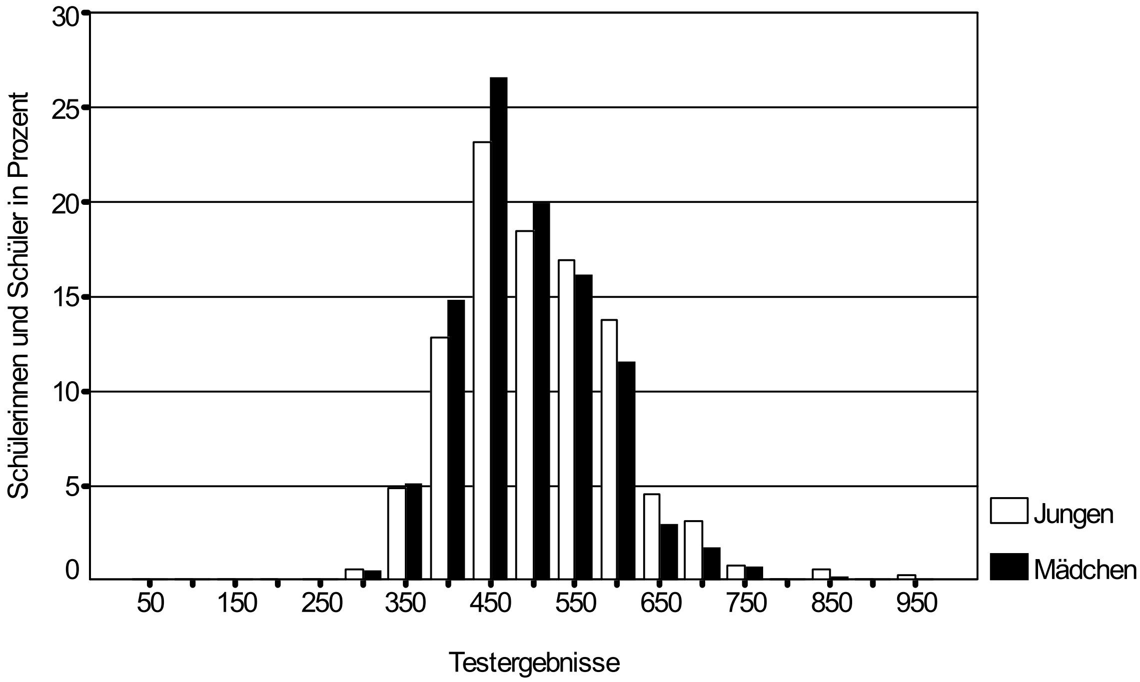
Im Hinblick auf die Ergebnisse im QuaSUM-Mathematiktest erreichen in der Klassenstufe 5 die Jungen (Rasch-Skalenmittelwert: 151) im Durchschnitt eine etwas höhere Testleistung als die Mädchen (Rasch-Skalenmittelwert: 149; d = 0,10). Ein ähnlicher Leistungsunterschied (d = 0,18) war bereits in der Hamburger Lernausgangslagenuntersuchung für Schülerinnen und Schüler am Ende der Klassenstufe 4 festgestellt worden (vgl. LEHMANN & PEEK 1997, S. 111). Die folgende Abbildung 19 zeigt die Verteilungen der Rasch-Skalenwerte nach Geschlecht, wobei auch die etwas größere Leistungsstreuung bei den Jungen auffällt: Die Standardabweichungen betragen 26,5 Skalenpunkte für die Jungen und 23,2 Skalenpunkte für die Mädchen.

Abbildung 19 Verteilung der Schülerleistungen im QuaSUM-Mathematiktest (Klassenstufe 5), nach Geschlecht



In der Klassenstufe 9 ergeben sich für den QuaSUM-Mathematiktest über die Schulformen hinweg keine Leistungsunterschiede zwischen Jungen und Mädchen. Dagegen weisen die Jungen im Mathe40-Test (Rasch-Skalenmittelwert: 508) etwas bessere Ergebnisse auf als die Mädchen (Rasch-Skalenmittelwert: 495); der Unterschied beläuft sich auf d = 0,15. Bei beiden Tests sind die Streuungen für die Jungen wieder etwas stärker ausgeprägt. Abbildung 20 zeigt die Verteilungen der Leistungen im Mathe40-Test nach Geschlecht.

Abbildung 20 Verteilung der Schülerleistungen im Mathe40-Test (Klassenstufe 9), nach Geschlecht



Beim Vergleich *innerhalb der Schulformen bzw. Kursniveaus* ergeben sich in der Klassenstufe 9 zum Teil sehr deutliche Leistungsunterschiede zugunsten der Jungen. In der folgenden Tabelle 18 sind die Mittelwerte und Standardabweichungen (in Klammern) sowohl für den QuaSUM-Test als auch für den Mathe40-Test aufgeführt. In der letzten Spalte der Tabelle sind die standardisierten Effektstärken angegeben, die durchgängig auf einen deutlicheren Geschlechterunterschied beim Mathe40-Test hinweisen. Die deutlichsten Unterschiede zeigen sich am Gymnasium, die geringsten in den Mathematik-Grundkursen an Gesamtschulen. Die in der Tabelle 18 aufgezeigten Befunde sind in der gleichen Richtung bereits in TIMSS für die Klassenstufen 7 und 8 (vgl. BAUMERT ET AL. 1997) und in der Hamburger Lernausgangslagenuntersuchung für die Klassenstufen 5, 6 und 7 (vgl. LEHMANN & PEEK 1997; LEHMANN, GÄNSFUß & PEEK 1999) aufgezeigt worden.

Tabelle 18 Vergleich der Leistungen im QuaSUM-Mathematiktest und im Mathe40-Test von Jungen und Mädchen in der Klassenstufe 9, nach Schulform bzw. Kursniveau, Mittelwerte und Standardabweichungen (in Klammern)

	QuaSUM-Mathematiktest			**Mathe40-Test**		
	Jungen	**Mädchen**	**Effektstärke *d***	**Jungen**	**Mädchen**	**Effektstärke *d***
Gesamtschulen, Grundkurse	127 (15)	125 (14)	0,12	440 (58)	428 (53)	0,22
Gesamtschulen, Erweiterungskurse	153 (17)	149 (15)	0,26	514 (69)	487 (62)	0,40
Realschulen	152 (19)	147 (17)	0,28	511 (76)	481 (63)	0,42
Gymnasien	179 (19)	172 (16)	0,39	601 (83)	560 (74)	0,52
insgesamt	*150 (26)*	*150 (24)*	*0,00*	*508 (92)*	*495 (82)*	*0,15*

Insgesamt sind also für die Klassenstufe 5 nur sehr geringe und für die Klassenstufe 9 keine Unterschiede in den Mathematikleistungen (QuaSUM-Test) zwischen Jungen und Mädchen festzustellen. Wenn nun in der Klassenstufe 9 *innerhalb* der Schul- bzw. Kursformen trotzdem bedeutende Geschlechterdifferenzen zugunsten der Jungen auftreten, so wird dies mit dem bevorzugten Übergang der Mädchen in die anspruchsvolleren Bildungsgänge zusammenhängen, und zwar bemerkenswerterweise auch an den Gesamtschulen, wo die Differenzierung nach der spezifischen Fachleistung erfolgen soll. Dieser Effekt beruht vor allem auf den geschlechtsspezifisch differenten Leistungsschwellen beim Übergang in eine anspruchsvollere Schulform oder in eine anspruchsvollere Kursform. Außerhalb der Gesamtschule wird er vermutlich verstärkt durch die bekanntermaßen überlegenen Leistungen der Mädchen im sprachlichen Bereich bedingt sein. Auf diese Weise fließen den Realschulen und Gymnasien (und bemerkenswerter Weise eben auch den Mathematik-Erweiterungskursen der Gesamtschulen) verhältnismäßig viele Mädchen mit niedrigerem mathematischem Lernstand zu, während vielen Jungen mit vergleichbaren Fähigkeiten solcher Zugang verwehrt bleibt. Damit aber sinkt die durchschnittliche Leistung der Mädchen in allen Schulformen, während sie für die Jungen steigt. Andererseits erhalten die Mädchen im Rahmen eines anspruchsvolleren Unterrichtsangebots in den Erweiterungskursen, Realschulen und Gymnasien *insgesamt* die Möglichkeit zum Ausgleich des ursprünglichen – geringen – Leistungsdefizits (vgl. dazu auch LEHMANN & PEEK 1997).

Dass nicht nur der Unterricht, sondern auch z. B. motivationale Faktoren zu diesem Befund beitragen, belegen zusätzlich die verschiedenen Effektstärken zwischen dem QuaSUM-Test und dem Mathe40-Test: Angesichts der größeren Nähe zum Curriculum sind die Aufgaben aus dem QuaSUM-Test für die Mädchen relativ einfacher zu bearbeiten, während bei den vergleichsweise ungewohnten Aufgaben des Mathe40-Tests die kompensierenden Effekte des bevorzugten Übergangs in anspruchsvollere Schul- bzw. Kursformen weniger zur Geltung kommen.

4.3 Familiale Merkmale

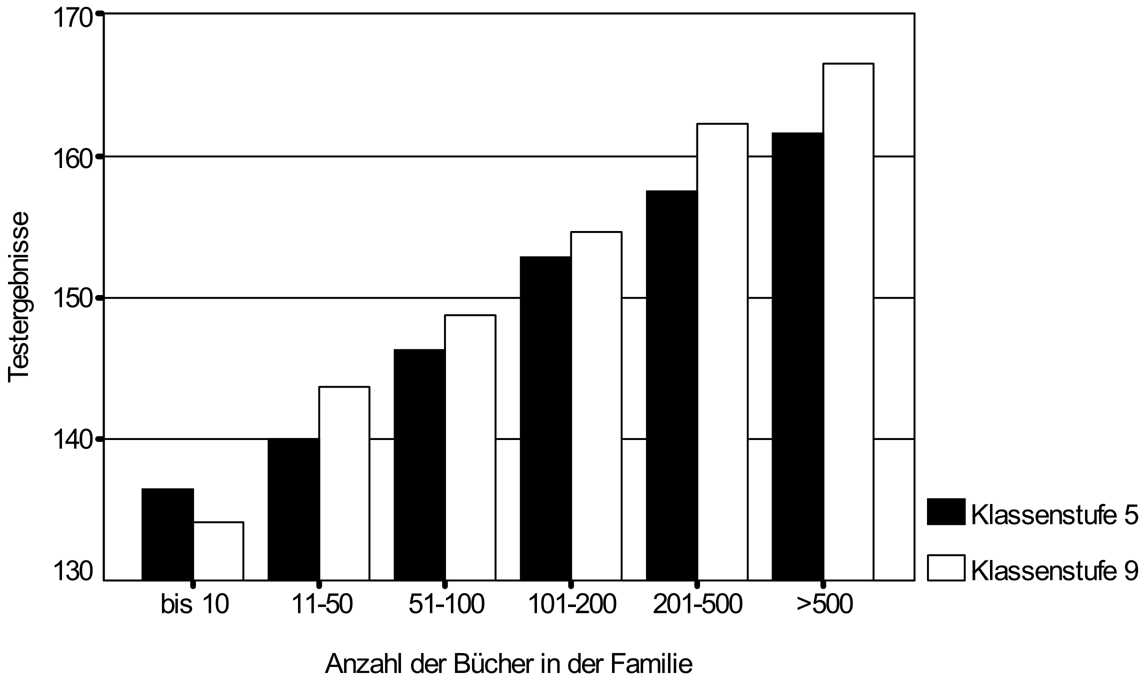
In diesem Abschnitt geht es darum, Zusammenhänge zwischen Merkmalen der Familien, in denen die Schülerinnen und Schüler aufwachsen, und den Schülerleistungen in Mathematik aufzuzeigen. Grundlage für die Analysen sind Angaben aus dem Elternfragebogen (vgl. Abschnitt 2.1) zur außerschulischen Lern- und Lebenssituation der Schülerinnen und Schüler. Der Elternfragebogen wurde zu Hause von den Schülerinnen und Schülern zusammen mit ihren Eltern ausgefüllt.

Der familiale *Buchbestand im Elternhaus* und der *höchste erreichte Schulabschluss der Eltern*[29] sind geeignete Indikatoren für das sozio-kulturelle und sozio-ökonomische Milieu der Familien (vgl. FEND 1982). Die Familien der Fünft- und der Neuntklässler machen sehr ähnliche Angaben über ihren Bücherbesitz; der typische Haushalt verfügt in beiden Fällen über mehr als 100 Bücher, und Familien mit höchstens 10 Büchern sind mit ca. 3 Prozent selten. Die folgende Abbildung 21 zeigt für beide Klassenstufen, dass die erreichten Werte im QuaSUM-Mathematiktest mit zunehmendem Buchbestand in den Familien steigen, wobei der Zusammenhang für die Klassenstufe 5 ($Eta^2 = 0{,}09$) niedriger ist als für die Klassenstufe 9 ($Eta^2 = 0{,}12$).

Schülerinnen und Schüler, in deren Elternhäusern sehr wenige Bücher vorhanden sind, erreichen im Durchschnitt einen Rasch-Skalenwert unter 137; solche, deren Eltern über 500 Bücher besitzen, erreichen durchschnittlich einen Rasch-Skalenwert von 162 bzw. 167. Dieses Ergebnis, wonach die Bildungsnähe des Elternhauses – indiziert über den Buchbesitz – einen deutlich positiven Einfluss auf die Mathematikleistungen hat, zeigt sich bei den Neuntklässlern auch im Mathe40-Test (ohne Abbildung). Die Korrelationen zwischen der Anzahl der Bücher zu Hause und dem Ergebnis im QuaSUM-Test betragen bei den Fünftklässlern $r = 0{,}30$ und bei den Neuntklässlern $r = 0{,}34$ bzw. bei den Neuntklässlern im Mathe40-Test $r = 0{,}31$.

[29] Sowohl in der Klassenstufe 5 ($r = 0{,}46$) als auch in der Klassenstufe 9 ($r = 0{,}43$) stehen der höchste erreichte Schulabschluss der Eltern und die Bücheranzahl zu Hause in einem engen Zusammenhang.

Abbildung 21 Ergebnisse im QuaSUM-Mathematiktest, nach Buchbestand in der Familie und Klassenstufe



Ebenso wie bei der Anzahl der Bücher in der Familie zeigen sich deutliche Zusammenhänge zwischen der Schulbildung der Eltern und den Mathematikleistungen der Schülerinnen und Schüler. Die folgende Tabelle 19 veranschaulicht sowohl für die Fünft- als auch für die Neuntklässler einen kontinuierlichen Anstieg der erreichten Testleistungen bei höherem Schulabschluss der Eltern.

Tabelle 19 Ergebnisse im QuaSUM-Mathematiktest, nach höchstem Schulabschluss der Eltern und Klassenstufe (Ergebnisse für die Klassenstufe 9 in Klammern)

Schulabschluss	**der Mutter**		**des Vaters**	
	Mittelwert der Kinder	*N*	Mittelwert der Kinder	*N*
ohne Abschluss	136 (134)	18 (64)	132 (136)	25 (70)
Polytechnische Oberschule (8. Klasse)	137 (138)	190 (703)	140 (142)	311 (995)
Haupt-/ Volksschulabschluss	140 (143)	69 (242)	143 (145)	75 (239)
Realschulabschluss/mittlere Reife/Polytechnische Oberschule (10. Klasse)	149 (151)	1.975 (5.056)	151 (152)	1.601 (4.009)
Fachhochschulreife	157 (155)	190 (466)	156 (157)	115 (284)
allgemeine bzw. fachgebundene Hochschulreife/Abitur	164 (167)	429 (1.158)	162 (166)	464 (1.289)

Die Korrelationen zwischen dem zusammengefassten Wert für den höchsten erreichten Schulabschluss der Eltern und der Mathematikleistung sind ähnlich hoch wie bei der Anzahl der Bücher: Für den QuaSUM-Test liegen sie bei den Fünftklässlern bei $r = 0{,}31$ und bei den Neuntklässlern bei $r = 0{,}35$; für den Mathe40-Test ergibt sich bei den Neuntklässlern ein Wert von $r = 0{,}31$[30].

Im Elternfragebogen wurden weitere Aspekte der außerschulischen Umgebung der Schülerinnen und Schüler erhoben und in den Zusammenhang mit der Leistung im QuaSUM-Mathematiktest gestellt. Für die *Schülerinnen und Schüler der Klassenstufe 5* zeigen sich zwischen der materiellen Ausstattung der Familien und der Leistung im Mathematiktest insgesamt kaum Bezüge[31], die Korrelationen liegen bei maximal $r = 0{,}11$. Enger sind die Zusammenhänge zwischen der Mathematikleistung und den Aktivitäten außerhalb der Schule, und zwar weniger die unterschiedlichen Freizeitaktivitäten[32], als vielmehr den auf Schule und Unterricht bezogenen häuslichen Tätigkeiten. Der deutlichste Zusammenhang mit der Mathematikleistung zeigt sich für das Merkmal Nachhilfeunterricht. Insgesamt geben in der fünften Klassenstufe 16,4 Prozent der Schülerinnen und Schüler an, wöchentlich Nachhilfeunterricht zu erhalten, wobei für die meisten zwischen 1 und 3 Stunden in der Woche angesetzt sind. Speziell in Mathematik bekommen 10,8 Prozent der Schülerinnen und Schüler Nachhilfe, die meisten davon bis maximal 3 Stunden pro Woche. Die Korrelationen zwischen der Testleistung und der Häufigkeit von Nachhilfe insgesamt ($r = -0{,}27$) bzw. von Nachhilfe im Fach Mathematik ($r = -0{,}26$) zeigen – wie erwartet – deutlich, dass die Schülerinnen und Schüler mit eher schlechten Leistungen Nachhilfe erhalten. Nachhilfeunterricht ist also in der hier untersuchten Klassenstufe 5 in erster Linie ein Instrument, krisenhaften Leistungsentwicklungen gegenzusteuern, nicht so sehr ein Mittel, auch bei mittleren und günstigen Lernständen das individuelle Leistungsniveau durch zusätzliche Lernangebote anzuheben und so vielleicht höhere Bildungsaspirationen zu verwirklichen.

Die Zeit, die die Schülerinnen und Schüler der fünften Klassen mit Hausaufgaben und Lernen insgesamt verbringen, beträgt für die meisten Schüle-

---

[30] Sehr ähnliche Werte ergeben sich, wenn man anstelle der Schulabschlüsse die Form der Berufsausbildung betrachtet, was wegen der engen Koppelung von Schulabschluss und Berufsausbildung nicht überrascht.

[31] Es wurde erhoben, ob die Schülerinnen und Schüler selbst folgende Dinge haben: ein eigenes Zimmer; einen eigenen Schreibtisch; ein eigenes Nachschlagewerk, Lexikon; eigene Bücher (ohne Schulbücher); einen eigenen Computer; einen eigenen Kassettenrekorder oder CD-Player; einen eigenen Fernseher; einen eigenen Taschenrechner.

[32] Bei den Freizeitaktivitäten steht die Häufigkeit des Lesens zum eigenen Vergnügen am deutlichsten mit der Mathematikleistung in Zusammenhang ($r = 0{,}23$).

rinnen und Schüler bis zu 6 Stunden wöchentlich; 20 Prozent benötigen dafür mehr Zeit. Im Durchschnitt sind es etwas mehr als 4 Stunden pro Woche, also knapp eine Stunde an den Schultagen. Der Anteil der Mathematikhausaufgaben umfasst für 88 Prozent der Schülerinnen und Schüler wöchentlich bis zu 3 Stunden. Nur 12 Prozent brauchen wöchentlich mehr Zeit für die Hausaufgaben in Mathematik; im Mittel sind es etwas weniger als 2 Stunden in der Woche. Die Korrelation zwischen der QuaSUM-Testleistung und dem Zeitaufwand für die Mathematikhausaufgaben beträgt $r = -0{,}10$, die zwischen der Testleistung und dem Zeitaufwand für Hausaufgaben insgesamt $r = -0{,}03$. Ein erhöhter Zeitaufwand bei den Hausaufgaben in Mathematik ist damit zunächst und vor allem eine Begleiterscheinung niedriger Mathematikleistungen in der Schule. Eine leistungssteigernde Wirkung der mit der Vergabe von Hausaufgaben beabsichtigten Übung lässt sich in der Klassenstufe 5 auf individueller Ebene nicht nachweisen. Etwas anders sieht es aus, wenn man diese Zusammenhänge auf Klassenebene untersucht. Hier gilt, dass die Schülerinnen und Schüler aus Klassen mit höherem Leistungsniveau auch von einem höheren Aufwand für die Hausaufgaben berichten ($r = -0{,}13$). Das Verhältnis zwischen dem schwach negativen Individualeffekt und dem tendenziell positiven Klasseneffekt bedarf noch einer eingehenderen Untersuchung, die im Rahmen des vorliegenden Berichts nicht geleistet werden kann.

Die Fünftklässler wurden gefragt, wie oft ihre Eltern die Hausaufgaben nachsehen. Die Korrelation von $r = -0{,}29$ zwischen Testleistung und Häufigkeit des Nachsehens passt in das Bild, wonach ein erhöhtes Elternengagement vor allem bei schwachen Lernleistungen der Kinder gegeben ist. 23 Prozent der Schülerinnen und Schüler geben an, dass ihre Eltern jeden Tag die Hausaufgaben nachsehen, 30 Prozent fast jeden Tag, 22 Prozent zwei- bis dreimal die Woche. Nur 19 Prozent versichern, dass die Eltern die Hausaufgaben höchstens einmal die Woche nachsehen, 6 Prozent, dass sie es nie tun. Für fast die Hälfte (48 Prozent) der Schülerinnen und Schüler aus fünften Klassen gilt, dass ihre Eltern jeden Tag mit ihnen über die Schule reden, und 36 Prozent sagen, dass sie es fast jeden Tag tun. Nur 12 Prozent der Eltern reden mit ihren Kindern zwei- bis dreimal die Woche über die Schule, 4 Prozent seltener oder gar nicht.

Auf die Frage, wie wichtig es ist, in den „Kernfächern“ gut zu sein, antworten jeweils 69 Prozent der Fünftklässlerinnen und Fünftklässler für die Fächer Deutsch und Mathematik mit „wichtig“ und weitere 29 Prozent mit „eher wichtig“, für das Fach Englisch sind es 65 Prozent bzw. 30 Prozent. Nur zwischen 2 und 5 Prozent der Schülerinnen und Schüler dieser Klassenstufe halten es für „unwichtig“ bzw. „eher unwichtig“, in diesen Fächern gut zu sein. Von den Eltern wird die Bedeutung der Leistung in diesen Fächern noch höher eingeschätzt: 84 Prozent bzw. 85 Prozent halten es für wichtig, dass ihr Kind in Mathematik bzw. Deutsch gut ist, und für 73 Prozent ist wichtig, dass es in Englisch gut ist. Vier Fünftel aller befragten Eltern der Fünftklässlerinnen und Fünftklässler geben an, mit den allgemeinen Schulleistungen ihrer Tochter/ihres Sohnes „sehr zufrieden“ oder „eher zufrieden“ zu sein. 18 Prozent sind „eher unzufrieden“, aber nur 3 Prozent sind „sehr unzufrieden“. Die Zufriedenheit der Eltern mit den Leistungen ihrer Tochter/ihres Sohnes im Fach Mathematik verteilt sich ähnlich. Laut Angaben der Eltern gehen fast alle „meistens“ oder „immer“ zu den Elternversammlungen, nur 7 Prozent geben an, „selten“ oder „nie“ zu den Elternversammlungen zu gehen.

Wie in der Klassenstufe 5 steht in der *Klassenstufe 9* die materielle Ausstattung der Schülerinnen und Schüler in fast keinem Zusammenhang mit der Testleistung im QuaSUM-Mathematiktest. Allein die Tatsache, dass zu Hause ein Computer zur Verfügung steht, ist positiv mit der Fachleistung korreliert ($r = 0{,}22$ für den QuaSUM- und für den Mathe40-Test). Dass sich dieser Zusammenhang gegenüber den fünften Klassen ($r = 0{,}13$) verstärkt hat und im Übrigen für den Rechner im Elternhaus weitaus enger ist als für den *eigenen* Computer, ist ein Indiz für die mit steigendem Alter und steigenden Lernanforderungen zunehmende Bedeutung dieser Geräte in ihren seriösen Anwendungen für die häusliche Lernumwelt der Schülerinnen und Schüler[33]. Die über die Schulformen und Bildungsniveaus der Eltern ungleiche Verbreitung von Computern beleuchtet daneben die Gefahr, dass in diesem Bereich neue Bildungsdisparitäten entstehen. Die Präferenz für bestimmte Freizeitaktivitäten, soweit sie sonst im Schülerfragebogen erfasst wurden, hat nahezu keine Bedeutung für die Mathematikleistung[34].

In der Klassenstufe 9 geben 15 Prozent der Schülerinnen und Schüler aus den Mathematik-Grundkursen an Gesamtschulen an, Zeit für Nachhilfe aufzubringen; in den Erweiterungskursen sind es 10 Prozent, in den Realschulen

---

[33] Auffällig sind in diesem Zusammenhang auch die – mit aufsteigender Schulform zunehmenden – Korrelationswerte zwischen Testleistung und Computerspielen bzw. Internetsurfen. Für Grundkurse beträgt die Korrelation zwischen der Leistung im QuaSUM-Test und der Häufigkeit des Computerspielens bzw. Internetsurfens $r = 0{,}10$, für die Erweiterungskurse $r = 0{,}15$, für die Realschulen $r = 0{,}18$ und für die Gymnasien $r = 0{,}22$.

[34] Die Korrelationen zwischen Testleistung und der Häufigkeit von Freizeitaktivitäten wie „einem bezahlten Job nachgehen“, „mit Freunden etwas unternehmen“, „Arbeiten, die zu Hause anfallen“, „Sport treiben“, Lesen zum eigenen Vergnügen“, „mit der Familie etwas unternehmen“, „Freizeitgruppen, Vereine“, „Lesen einer Tageszeitung“ sowie die Häufigkeit der „Buchausleihe aus öffentlichen Büchereien oder der Schulbücherei“ überschreiten nicht die Grenzen $r = \pm\, 0{,}10$.

15 Prozent und in den Gymnasien 11 Prozent. Nachhilfe in Mathematik bekommen 13 Prozent der Schülerinnen und Schüler aus Gesamtschul-Grundkursen, 9 Prozent aus Gesamtschul-Erweiterungskursen, 13 Prozent aus Realschulen und 8 Prozent aus Gymnasien. Die Korrelationen zwischen der Testleistung und der Häufigkeit von Nachhilfe insgesamt bzw. der Nachhilfe in Mathematik liegen für die unterschiedlichen Schul- bzw. Kursformen zwischen r = -0,03 und r = -0,18 bzw. r = 0,05 und r = -0,17. Diese Werte sind niedriger als die in der Klassenstufe 5 beobachteten. Es ist denkbar, dass die Bereitschaft der Eltern, ihre Kinder gegebenenfalls durch Nachhilfeunterricht zu unterstützen, vor dem wichtigen Übergang in die weiterführende Schulform größer ist als danach, wo es in der subjektiven Wahrnehmung vor allem um Zensuren und Klassenwiederholung gehen mag. Möglich ist daneben aber auch ein Zusammenhang mit der auch sonst zu beobachtenden Abnahme des direkten elterlichen Engagements mit zunehmendem Alter der Jugendlichen.

Was die wöchentlich aufgebrachte Zeit für Hausaufgaben anbelangt, so ist in den Mathematik-Grundkursen an Gesamtschulen der Anteil der Schülerinnen und Schüler, die keine Hausaufgaben machen (müssen), am höchsten, während die Schülerinnen und Schüler in den Gymnasien im Vergleich mit den anderen Schulformen die meiste Zeit für Mathematikhausaufgaben aufbringen. Die folgende Tabelle gibt Auskunft über die Werte für die wöchentlich aufgebrachte Zeit für Hausaufgaben insgesamt und für Mathematik.

Tabelle 20 Wöchentliche Zeit für Hausaufgaben insgesamt und für Mathematik in der Klassenstufe 9, nach Schulform bzw. Kursniveau (in Prozent)

	Hausaufgaben in Stunden					
	0	**< 1**	**1 bis 3**	**4 bis 6**	**7 bis 9**	**> 9**
insgesamt						
Gesamtschule, Grundkurse	8,0	28,9	40,4	15,5	5,1	2,1
Gesamtschule, Erweiterungskurse	2,2	19,8	40,5	25,9	7,9	3,7
Realschulen	1,9	19,2	40,6	25,3	9,3	3,8
Gymnasien	0,6	6,2	30,2	34,1	18,1	10,8
in Mathematik						
Gesamtschule, Grundkurse	13,3	49,3	30,7	5,5	0,9	0,0
Gesamtschule, Erweiterungskurse	5,2	49,9	38,0	5,1	1,3	0,6
Realschulen	4,9	48,4	39,0	6,2	1,4	0,1
Gymnasien	2,2	42,4	46,3	7,7	1,2	0,2

Zu berücksichtigen ist in diesem Zusammenhang allerdings, dass der von den Schülerinnen und Schülern angegebene Aufwand für die Hausaufgaben in der Klassenstufe 9 insgesamt geringer ist als in der Klassenstufe 5. Insgesamt beträgt der wöchentliche Zeitaufwand für Hausaufgaben in der Klassenstufe 5 durchschnittlich 4,3 Stunden, in der Klassenstufe 9 hingegen 3,6 Stunden, wobei der Wert für die Grundschulen nur an den Gymnasien mit 5,0 Stunden übertroffen, überall sonst aber unterschritten wird. Ein ähnliches Bild zeigt sich, wenn man lediglich die Hausaufgaben im Fach Mathematik betrachtet: einem Aufwand von 1,9 Stunden wöchentlich in den fünften Klassen stehen 1,4 Stunden in der Klassenstufe 9 gegenüber, wobei hier sogar an den Gymnasien durchschnittlich nur 1,6 Stunden in der Woche für die Hausaufgaben im Fach Mathematik aufzuwenden sind oder jedenfalls nach Schülerauskunft aufgewendet werden. Warum die Anforderungen an häusliche Arbeiten bei den älteren Schülerinnen und Schülern niedriger sind als bei den jüngeren, lässt sich aus den vorhandenen Daten nicht ermitteln. Denkbar ist, dass die Lehrkräfte in der Klassenstufe 9 Rücksicht nehmen auf die umfangreicheren Stundenpläne und die durchschnittlich längeren Schulwege. In jedem Falle lässt sich auch für die Klassenstufe 9 – wie für die Grundschulklassen – auf der Ebene der Individualdaten zunächst kein positiver Einfluss des Aufwands für die Hausaufgaben auf die Mathematikleistung nachweisen. Es soll aber nicht verschwiegen werden, dass wie an den Grundschulen, so auch an den Gesamt- und Realschulen in Kursen bzw. Schulklassen mit besseren Testergebnissen offenbar auch ein höherer Zeitaufwand für die Hausaufgaben verlangt wird, eine Tendenz, die deswegen bemerkenswert ist, weil sie sich trotz des auf individueller Ebene *negativen* Zusammenhangs zwischen Zeitaufwand und Leistungsstand – schwächere Schülerinnen und Schüler benötigen für die Bearbeitung derselben Aufgabenmenge mehr Zeit als leistungsstarke – durchsetzt.

Dass hier verhältnismäßig komplexe Zusammenhänge vorliegen, die noch intensiver Untersuchung bedürfen, zeigt sich noch an einem weiteren Befund. In allen Schulformen bzw. Kursniveaus ist der Aufwand der Mädchen für die Hausaufgaben im Fach Mathematik nach eigenen Angaben höher als der der Jungen. Bemerkenswert ist nun, dass die geschlechtsbezogenen Leistungsdifferenzen im Fach Mathematik überall unter denjenigen am geringsten sind, die ein mittleres Maß Zeit (etwa eine bis drei Stunden pro Woche) für mathematische Hausaufgaben aufwenden. In den Randbereichen (gar keine Hausaufgaben oder weniger als eine Stunde pro Woche bzw. mehr als drei Stunden pro Woche) wachsen die Abstände zu den Extremen hin. Es könnte also sein, dass bei maßvollem Zeiteinsatz die regelmäßige Erledigung von

Hausaufgaben durchaus eine leistungshomogenisierende Funktion übernehmen kann; auch hier sind vertiefende Analysen erforderlich.

Erwartungsgemäß sinkt von der Klassenstufe 5 nach Klassenstufe 9 die Häufigkeit, mit der die Eltern Hausaufgaben nachsehen. In den Grundkursen geben 42 Prozent der Schülerinnen und Schüler an, dass ihre Eltern nie ihre Hausaufgaben nachsehen, in den Erweiterungskursen sind es 54 Prozent, in den Realschulen 50 Prozent und in den Gymnasien 68 Prozent. Entsprechend sinken auch die Korrelationswerte zwischen der Häufigkeit des Hausaufgabennachschauens und der Mathematikleistung; sie liegen in der Klassenstufe 9 für die unterschiedlichen Schul- bzw. Kursformen zwischen $r = -0{,}11$ und $r = -0{,}17$.

Wie die elterliche Durchsicht der Hausaufgaben, so nimmt gegenüber den fünften Klassen auch die Häufigkeit ab, mit der die Eltern mit ihrer Tochter/ihrem Sohn über die Schule reden. Dabei unterscheiden sich die Schulformen bzw. Kursniveaus kaum voneinander: Insgesamt sprechen nach Angaben der Jugendlichen 26 Prozent der Eltern jeden Tag mit ihren Kindern über die Schule, 35 Prozent fast jeden Tag, 21 Prozent zwei- bis dreimal die Woche, 15 Prozent höchstens einmal die Woche und 2 Prozent nie.

Die drei Kernfächer Mathematik, Deutsch und Englisch haben einen hohen Stellenwert bei den Schülerinnen und Schülern der Klassenstufe 9. Auf die Frage, für wie wichtig sie es halten, in diesen Fächern jeweils gute Leistungen zu erbringen, antworten in den verschiedenen Schulformen bzw. Kursniveaus jeweils über 90 Prozent der Schülerinnen und Schüler, dass es „wichtig“ bzw. „eher wichtig“ ist, in Mathematik, Deutsch bzw. Englisch gut zu sein. Der niedrigste Wert liegt in den Mathematik-Grundkursen der Gesamtschulen im Fach Englisch bei immer noch 86 Prozent. Nur jeweils maximal 2,6 Prozent der Eltern halten es für „unwichtig“ oder „eher unwichtig“, dass ihre Kinder in den drei Fächern gut sind. Die einzige Ausnahme bildet wiederum das Fach Englisch bei den Eltern der Schülerinnen und Schüler aus den Grundkursen; hier steigt der Wert auf 6,1 Prozent.

Die Zufriedenheit der Eltern mit den Schulleistungen ihrer Kinder nimmt von der Klassenstufe 5 zur Klassenstufe 9 stark ab. Erstaunlicherweise unterscheiden sich aber die Schul- bzw. Kursformen darin nur wenig, sieht man von den ungünstigeren Verhältnissen in den Mathematik-Grundkursen der Gesamtschulen ab. Nur hier deutet sich eine Tendenz zur Resignation unter den Eltern an. Die folgende Tabelle spiegelt die Zufriedenheit der Eltern der Neuntklässlerinnen und Neuntklässler mit deren Schulleistungen insgesamt bzw. mit den Leistungen im Fach Mathematik wider. In der letzten Spalte der Tabelle sind – schul- bzw. kursniveauspezifisch – die korrelativen Zusammenhänge zwischen der elterlichen Zufriedenheit mit der Schulleistung ihrer Kinder und deren Ergebnis im QuaSUM-Mathematiktest angegeben. An den Realschulen und Gymnasien ist demnach die Zufriedenheit der Eltern mit den Lernerfolgen ihrer Kinder deutlich stärker durch die tatsächliche *Fachleistung* in Mathematik bestimmt als an den Gesamtschulen.

Tabelle 21 Zufriedenheit der Eltern mit den Schulleistungen insgesamt und mit den Mathematikleistungen ihrer Kinder in der Klassenstufe 9, nach Schulform bzw. Kursniveau (in Prozent); Korrelationen mit den Ergebnissen im QuaSUM-Mathematiktest

	sehr unzufrieden	**eher unzufrieden**	**eher zufrieden**	**sehr zufrieden**	***r***
mit den Schulleistungen insgesamt					
Gesamtschule, Grundkurse	11	43	41	6	0,17
Gesamtschule, Erweiterungskurse	2	24	55	20	0,18
Realschulen	4	30	53	13	0,32
Gymnasien	2	21	57	20	0,34
mit den Schulleistungen in Mathematik					
Gesamtschule, Grundkurse	12	44	37	7	0,23
Gesamtschule, Erweiterungskurse	3	27	51	19	0,25
Realschulen	7	35	43	15	0,40
Gymnasien	5	30	47	18	0,44

Die Frage danach, wie häufig Elternversammlungen besucht werden, beantwortet auch in der Klassenstufe 9 der überwiegende Teil der Eltern mit „meistens“ oder „immer“. Nur durchschnittlich 12,6 Prozent aller Eltern geben hier „selten“ oder „nie“ an. Eltern von Schülerinnen und Schülern aus Mathematik-Grundkursen an Gesamtschulen versichern zu 80 Prozent, „meistens“ oder „immer“ zu den Elternversammlungen zu gehen, in den Mathematik-Erweiterungskursen und den Realschulen zu je 88 Prozent und in den Gymnasien zu 93 Prozent. Es besteht kein erkennbarer Zusammenhang zwischen der Häufigkeit, mit der Eltern zu den Elternveranstaltungen gehen, und den Mathematikleistungen ihrer Kinder.

4.4 Lern- und schulbezogene Einstellungen der Schülerinnen und Schüler

Im Schülerfragebogen sollten die Schülerinnen und Schüler Angaben dazu machen, wie sie sich in der Schule fühlen (*Schulzufriedenheit*), wie sie ihre Lernfähigkeit einschätzen (*leistungsbezogenes Selbstkonzept*), wie hoch ihr mathematisches Interesse ist (*Sach- und Fachinteresse an Mathematik*), wie wichtig Anstrengung ihrer Meinung nach für den Schulerfolg ist (*Anstrengung als Beitrag zum Schulerfolg*) und wie sie ihren Mathematikunterricht unter der Perspektive von Zielgerichtetheit, Transparenz, Strukturiertheit und lernförderlicher Unterstützung wahrnehmen (*Wahrnehmung des eigenen Mathematikunterrichts*). Auf jeder der insgesamt fünf Skalen tendieren die befragten Schülerinnen und Schüler zu einer eher positiven Antwort: Bei einem Wertebereich, der bei den für die Skalen zusammengefassten Einzelaussagen von 1 (für positiv formulierte Items: „trifft überhaupt nicht zu") bis 4 (für positiv formulierte Items: „trifft völlig zu") reicht, liegen die Mittelwerte für beide Klassenstufen und für alle Schulformen zwischen 2,5 (dem theoretischen Skalenmittelwert) und 3,5. Insoweit äußern sowohl die Fünft- als auch die Neuntklässler im Ganzen überwiegend positive lern- und schulbezogene Einstellungen, wobei sich im einzelnen zwischen den Klassenstufen und zwischen den Schulformen bzw. Kursniveaus sowie zwischen den Geschlechtern Unterschiede zeigen.

Schulzufriedenheit

Die erste Dimension der erfragten Einstellungen[35] drückt sich in der Bewertung von insgesamt 14 Aussagen aus wie z. B.: „Ich fühle mich in meiner Schule wohl", „In meiner Klasse fühle ich mich richtig wohl" und „Ich gehe gerne zur Schule". Für beide Klassenstufen und für die unterschiedlichen Schul- bzw. Kursformen in der Klassenstufe 9 liegen die Mittelwerte zwischen 2,60 und 3,00[36]. Die Streuungen um den jeweiligen Mittelwert sind in den einzelnen Vergleichsgruppen sehr ähnlich (vgl. Tabelle 22). Im Vergleich der Schul- bzw. Kursformen in der Klassenstufe 9 fühlen sich die Gymnasi-

---

[35] Die interne Konsistenz der Skala „Schulzufriedenheit" liegt für die untersuchten Klassenstufen bei $\alpha = 0{,}77$ (Fünftklässler) bzw. $\alpha = 0{,}78$ (Neuntklässler). Für die vier Schul- bzw. Kursformen in der Klassenstufe 9 ergeben sich Reliabilitäten zwischen $\alpha = 0{,}76$ (Gesamtschulen, Erweiterungskurse) und $\alpha = 0{,}80$ (Gymnasien).

[36] Die Ergebnisse zur allgemeinen Schulzufriedenheit der Brandenburger Schülerinnen und Schüler entsprechen weitgehend denen von Hamburger Schülerinnen und Schülern aus fünften bzw. aus siebten Klassen (vgl. LEHMANN & PEEK 1997; LEHMANN, GÄNSFUß & PEEK 1999), die im Rahmen von Schülerbefragungen einen fast identischen Fragensatz beantwortet haben.

astinnen und Gymnasiasten in der Schule am wohlsten, die befragten Schülerinnen und Schüler aus den Mathematik-Grundkursen an Gesamtschulen haben hier die niedrigsten Werte. In allen hier unterschiedenen Gruppen äußern die Mädchen ein höheres Maß an Schulzufriedenheit als die Jungen, wobei der Unterschied in den fünften Klassen d = 0,24 Standardabweicungen beträgt, während auf der Klassenstufe 9 differenzielle Effekte zu beobachten sind: In den Grundkursen der Gesamtschulen beläuft sich der Abstand auf d = 0,34 Standardabweichungen, in den Erweiterungskursen auf d = 0,30, an den Realschulen auf d = 0,13 und an den Gymnasien auf d = 0,20. Warum die Jungen an den Gesamtschulen hinsichtlich der Schulzufriedenheit anscheinend eine ungünstigere Entwicklung als die anderen Teilgruppen durchlaufen (soweit man aus dem Kohortenvergleich Alterstrends ableiten darf), ist auf der Grundlage der vorliegenden Daten nicht zu entscheiden.

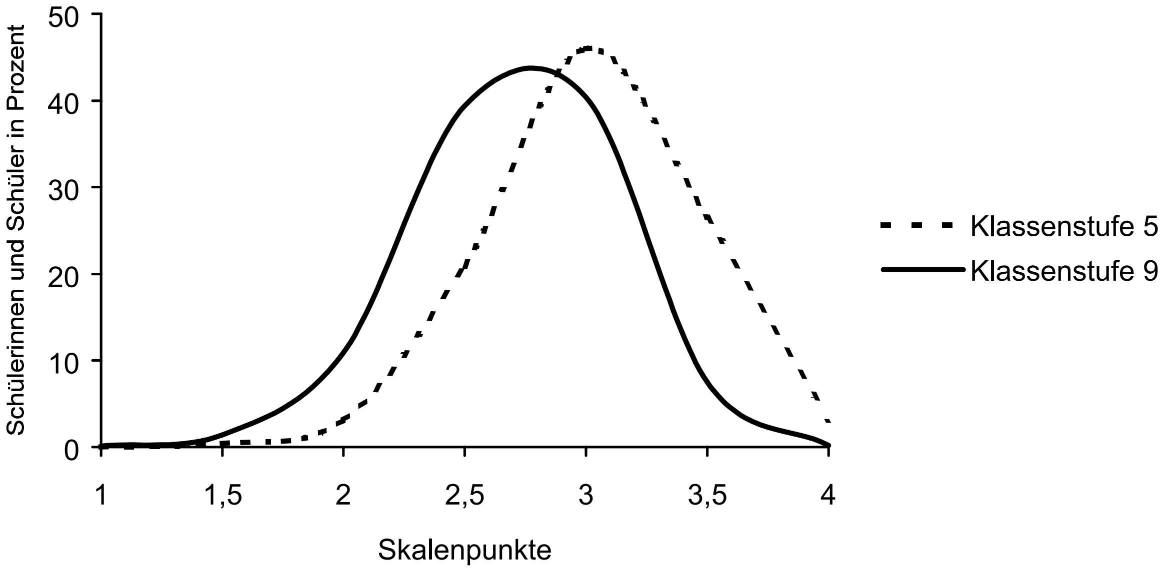
Tabelle 22 Statistische Kennwerte der Skala „Schulzufriedenheit“

	Mittelwert*	Standardabweichung
Klassenstufe 5	3,03	0,41
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,61	0,41
Gesamtschulen, Erweiterungskurse	2,71	0,37
Realschulen	2,71	0,39
Gymnasien	2,79	0,39
insgesamt	*2,71*	*0,40*

* Ausprägungen hier und in den Tabellen 23 bis 26: 1 = sehr gering bis 4 = sehr hoch

Abbildung 22 zeigt die Verteilungen für die Schülerinnen und Schüler aus den fünften und neunten Klassen, wobei die einzelnen Schulformen bzw. Kursniveaus in der Klassenstufe 9 wegen ihrer weitgehenden Übereinstimmung nicht gesondert aufgenommen sind. Der erkennbare allgemeine Rückgang der Schulzufriedenheit von Klassenstufe 5 nach 9 dürfte mit der bekannten Umorientierung der Jugendlichen auf außerschulische Lebensbereiche in der Adoleszenz zusammenhängen.

Abbildung 22 Verteilung der Werte auf der Skala „Schulzufriedenheit“, nach Klassenstufe



Leistungsbezogenes Selbstkonzept

Die zweite Dimension der erfragten Einstellungen der Schülerinnen und Schüler bezieht sich auf insgesamt 11 Aussagen wie z. B.: „Auch bei Aufgaben, von denen ich glaube, dass ich sie kann, habe ich Angst zu versagen“ (negativ) und „Häufig denke ich: Ich bin nicht so klug wie die Anderen“ (negativ) [37]. Die Ausprägungen auf den beiden Ebenen Klassenstufe und Schulform bzw. Kursniveau sind in Tabelle 23 aufgeführt. Aus den dort aufgelisteten Werten ist eine Sonderentwicklung bei denjenigen Schülerinnen und Schülern abzulesen, die in Gesamtschulen einen Grundkurs im Fach Mathematik besuchen: Wiederum mit der Einschränkung, dass aus dem Kohortenvergleich nur bedingt altersbezogene Tendenzen abgeleitet werden können, ist nur hier ein (geringfügiger) Rückgang des leistungsbezogenen Selbstkonzepts festzustellen. Überall sonst findet ein Zuwachs des Selbstvertrauens statt.

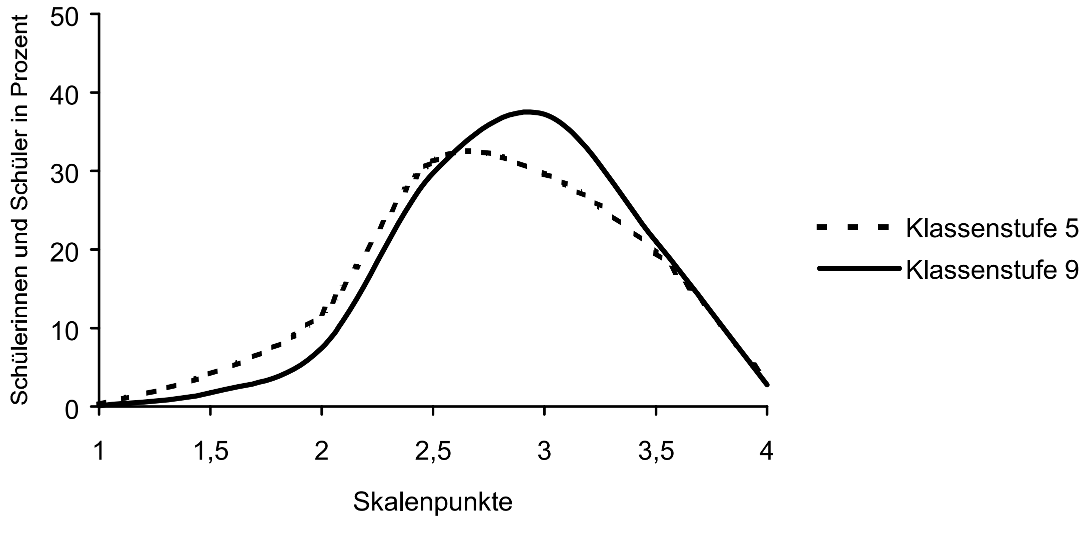
[37] Die interne Konsistenz der Skala liegt für die untersuchten Klassenstufen bei $\alpha = 0{,}83$ (Fünftklässler) bzw. $\alpha = 0{,}81$ (Neuntklässler). Für die Schul- bzw. Kursformen in der Klassenstufe 9 ergeben sich Reliabilitäten zwischen $\alpha = 0{,}78$ (Gesamtschulen, Erweiterungskurse) und $\alpha = 0{,}82$ (Gymnasien).

Tabelle 23 Statistische Kennwerte der Skala „leistungsbezogenes Selbstkonzept“

	Mittelwert	Standardabweichung
Klassenstufe 5	2,80	0,56
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,75	0,49
Gesamtschulen, Erweiterungskurse	2,95	0,44
Realschulen	2,89	0,48
Gymnasien	2,95	0,49
insgesamt	*2,88*	*0,48*

Abbildung 23 veranschaulicht noch einmal in graphischer Form den allgemeinen (durch die Studie von HORSTKEMPER (1987) auch im Längsschnitt nachgewiesenen) Zuwachs des Selbstvertrauens der Jugendlichen. Zu beachten ist die ebenfalls von HORSTKEMPER (1987) berichtete und mit zunehmendem Alter wachsende Differenz im Selbstvertrauen der Jungen und Mädchen. Bereits in der Klassenstufe 5 sind die Jungen merklich selbstbewusster (d = 0,21 Standardabweichungen), und der Abstand vergrößert sich – an allen Schulformen bzw. Kursniveaus etwa gleichmäßig – auf d = 0,40.

Abbildung 23 Verteilung der Werte auf der Skala „leistungsbezogenes Selbstkonzept“, nach Klassenstufe



Immerhin auffällig ist es, dass zwar die in Form von Zensuren rückgemeldete Fachleistung Mathematik das leistungsbezogene Selbstkonzept wesentlich stärker bestimmt als die Leistungen in anderen Schulfächern, dass aber die Mädchen die erreichten Lernerfolge im Fach Mathematik weniger in Selbstvertrauen umzusetzen vermögen als die Jungen (ohne Tabelle). Der Umstand, dass, wie oben dargestellt, die Mädchen innerhalb der Schulformen bzw. Kursniveaus durchschnittlich geringere Mathematikleistungen zeigen als die Jungen, dürfte diesen Effekt verstärken.

Sach- und Fachinteresse an Mathematik

Die dritte Dimension bezieht sich auf insgesamt 14 Aussagen wie z. B.: „Mathematik ist eines meiner Lieblingsfächer", „Ich freue mich immer auf die Stunden im Fach Mathematik" und „Es macht mir einfach Spaß, an einem mathematischen Problem zu knobeln"[38]. Ergebnisse für die untersuchten Klassenstufen bzw. Schul- und Kursformen sind in der Tabelle 24 und der Abbildung 24 aufgeführt.

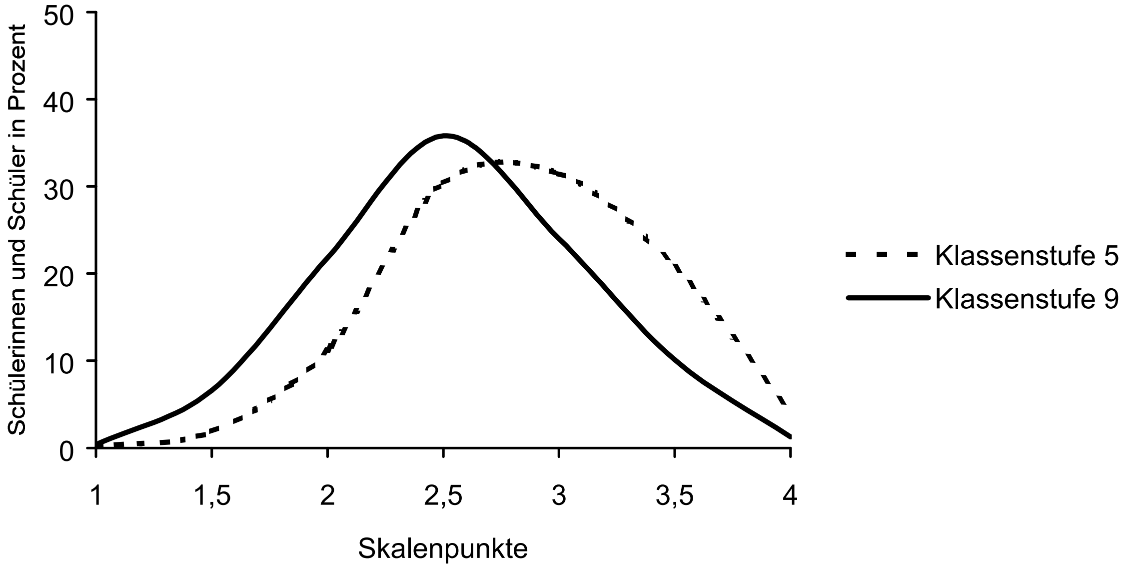
Tabelle 24 Statistische Kennwerte der Skala „Sach- und Fachinteresse an Mathematik"

	Mittelwert	Standardabweichung
Klassenstufe 5	2,85	0,54
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,43	0,50
Gesamtschulen, Erweiterungskurse	2,65	0,50
Realschulen	2,56	0,57
Gymnasien	2,60	0,59
insgesamt	*2,56*	*0,55*

Der von diesen Befunden nahegelegte Rückgang an mathematischem Sach- und Fachinteresse zwischen den Klassenstufen 5 und 9 hat Parallelen z. B. in dem durch frühere Studien belegten Rückgang des Fachinteresses im naturwissenschaftlichen Bereich. Die Verringerung des fachlichen Interesses ist vermutlich nicht auf das Fach Mathematik beschränkt, denn sie geht einher mit der Abnahme der allgemeinen Schulzufriedenheit und ist im Übrigen auch für andere Schulfächer wie Physik (vgl. HOFFMANN & LEHRKE 1986) und Geschichte (vgl. VON BORRIES 1995) belegt.

[38] Die interne Konsistenz der Skala liegt für die untersuchten Klassenstufen bei $\alpha = 0{,}87$ (Fünftklässler) bzw. $\alpha = 0{,}89$ (Neuntklässler). Für die Schul- bzw. Kursformen in der Klassenstufe 9 ergeben sich Reliabilitäten zwischen $\alpha = 0{,}86$ (Gesamtschulen, Grundkurse) und $\alpha = 0{,}91$ (Gymnasien).

Abbildung 24 Verteilung der Werte auf der Skala „Sach- und Fachinteresse an Mathematik“, nach Klassenstufe



Besondere Beachtung verdienen in diesem Zusammenhang wieder die sehr unterschiedlichen Einstellungen bei Jungen und Mädchen. Bereits in der Klassenstufe 5 sind die Jungen erheblich interessierter ($d = 0{,}48$), und dieser Abstand vermindert sich bis zur Klassenstufe 9 kaum ($d = 0{,}38$). Bei den älteren Mädchen befinden sich die Durchschnittswerte an allen Schulformen bzw. Kursniveaus in der Nähe des theoretischen Skalenmittelwerts, meist sogar knapp darunter, während das mathematische Sach- und Fachinteresse der jüngeren tendenziell noch so positiv ausgeprägt ist wie bei den meisten älteren Jungen. Nur in den Mathematik-Grundkursen der Gesamtschulen sinkt das Fachinteresse der Jungen bis zur Klassenstufe 9 bis zum theoretischen Skalenmittelwert, also gewissermaßen auf ein neutrales, distanziertes Niveau. So überrascht es nicht, wenn in diesen Kursen auf dem allgemein niedrigen Niveau des Fachinteresses auch die Geschlechterdifferenz mit $d = 0{,}27$ vergleichsweise gering ist. Dass aber an den Gymnasien mit $d = 0{,}52$ der größte Interessensabstand zwischen Jungen und Mädchen überhaupt auftritt, und zwar nicht auf besonders hoher Ebene, deutet auf eine frühe Festigung traditioneller weiblicher Fächerpräferenzen hin.

Anstrengung als Beitrag zum Schulerfolg

Die vierte Dimension der erfragten Einstellungen der Schülerinnen und Schüler bezieht sich auf insgesamt 16 Aussagen wie z. B.: „Um in der Schule Erfolg zu haben, ist es wichtig, dass man sich auf Arbeiten gut vorbereitet“, „...

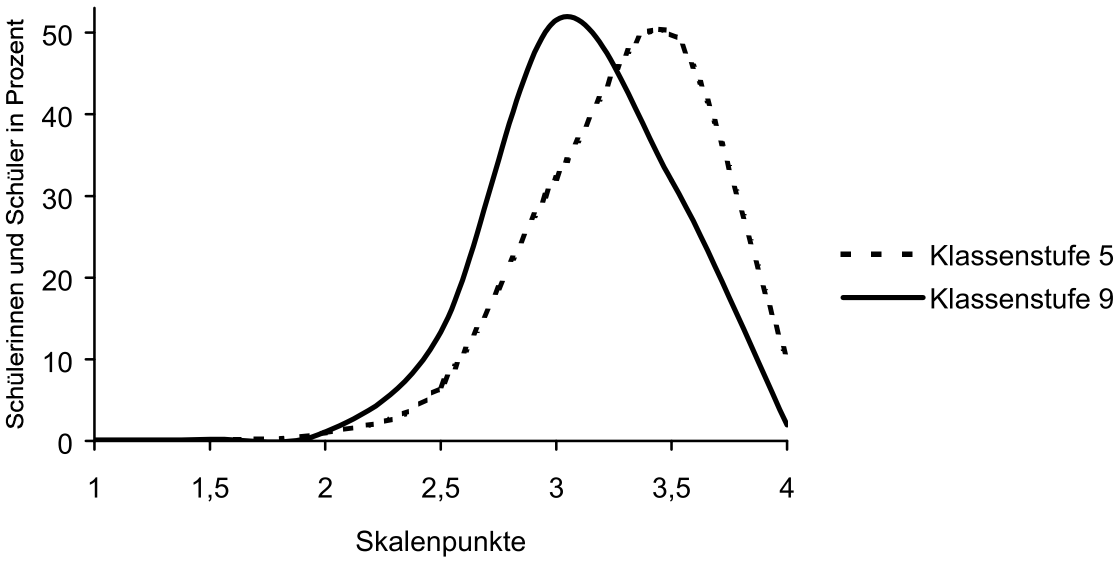
dass man immer sein Bestes gibt“ und „... dass man regelmäßig Hausaufgaben macht“[39].

Tabelle 25 Statistische Kennwerte der Skala „Anstrengung als Beitrag zum Schulerfolg“

	Mittelwert	Standardabweichung
Klassenstufe 5	3,27	0,38
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,99	0,39
Gesamtschulen, Erweiterungskurse	3,09	0,32
Realschulen	3,08	0,34
Gymnasien	3,11	0,32
insgesamt	*3,06*	*0,35*

Dass schulischer Erfolg in starkem Maße (auch) von der persönlichen Anstrengung abhängt, gehört nach diesen Ergebnissen zu den Grundüberzeugungen der Schülerinnen und Schüler (vgl. Abbildung 25).

Abbildung 25 Verteilung der Werte auf der Skala „Anstrengung als Beitrag zum Schulerfolg“, nach Klassenstufe



Auffällig sind der Rückgang der Betonung von Anstrengung von der fünften zur neunten Klassenstufe und die relativ hohe Übereinstimmung der Werte in den unterschiedlichen Schul- und Kursformen der Klassenstufe 9. Einzig in

[39] Die interne Konsistenz der Skala liegt für die Klassenstufen bei $\alpha = 0{,}81$ (Fünftklässler) bzw. $\alpha = 0{,}77$ (Neuntklässler). Für die Schul- bzw. Kursformen in der Klassenstufe 9 ergeben sich Reliabilitäten zwischen $\alpha = 0{,}72$ (Gymnasien) und $\alpha = 0{,}81$ (Gesamtschulen, Grundkurse).

den Mathematik-Grundkursen der Gesamtschulen sind merklich niedrigere Werte anzutreffen (*d* = -0,22). Möglicherweise werden in diesem Befund vermehrt auftretende resignative Tendenzen gerade in dieser Gruppe erkennbar, und zwar unter den Jungen stärker als unter den Mädchen. In allen Schulformen bzw. Kursniveaus sind es nämlich die letzteren, die die Bedeutung der Anstrengung für den schulischen Erfolg höher einschätzen (Klassenstufe 5: *d* = 0,21; Klassenstufe 9: *d* = 0,23). Dem entspricht es, dass die Mädchen mehr Zeit für die Hausaufgaben aufwenden als die Jungen, wobei sich der Abstand im Laufe der Schulzeit offenbar vergrößert (Hausaufgaben insgesamt, Klassenstufe 5: *d* = 0,10; Klassenstufe 9: *d* = 0,23).

Wahrnehmung des eigenen Mathematikunterrichts

Die fünfte Dimension der erfragten Einstellungen – die Wahrnehmung des eigenen Mathematikunterrichts unter den Gesichtspunkten Zielgerichtetheit, Transparenz, Strukturiertheit und lernförderliche Unterstützung – bezieht sich auf insgesamt 12 Aussagen wie z. B.: „Der Lehrer achtet sehr darauf, dass wir aufpassen“, „Unser Lehrer sagt uns zu Beginn der Stunde, was wir durchnehmen werden“ und „Unser Lehrer lobt auch die schwachen Schüler, wenn er merkt, dass sie sich verbessern“[40].

Tabelle 26 Statistische Kennwerte der Skala „Wahrnehmung des eigenen Mathematikunterrichts“

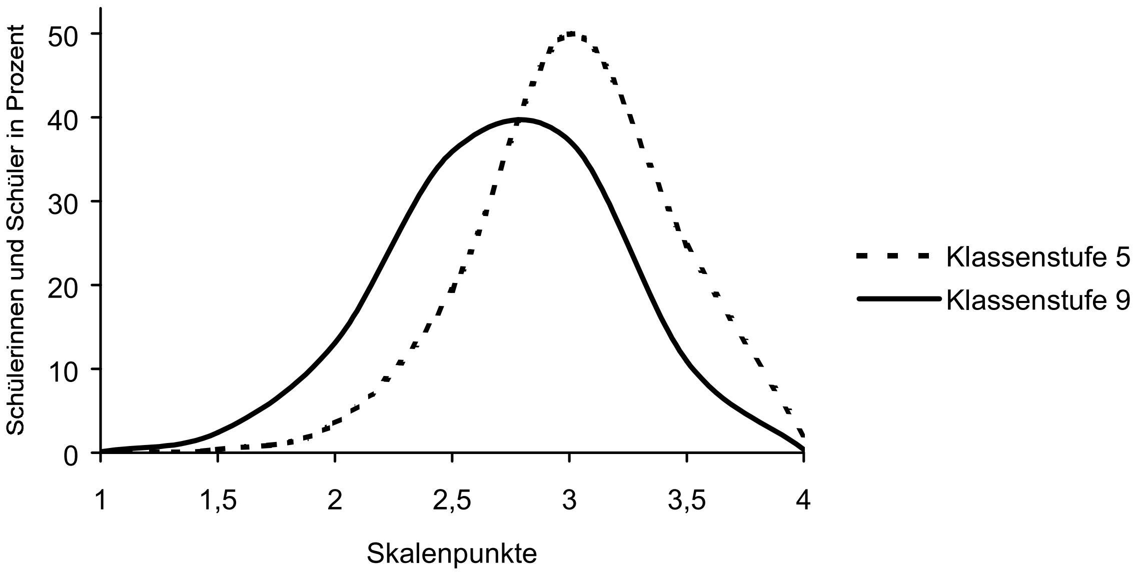
	Mittelwert	Standardabweichung
Klassenstufe 5	2,97	0,39
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,60	0,48
Gesamtschulen, Erweiterungskurse	2,73	0,45
Realschulen	2,68	0,46
Gymnasien	2,68	0,43
insgesamt	*2,67*	*0,45*

Die Schülerinnen und Schüler in der Klassenstufe 9 schätzen den Unterricht allgemein kritischer ein als die Fünftklässlerinnen und Fünftklässler (vgl. hierzu auch Abbildung 26 sowie die Untersuchungsergebnisse von STURZBECHER 1997). Nennenswerte Unterschiede zwischen Jungen und Mädchen in der Wahrnehmung des Mathematikunterrichts sind nicht zu beobachten. Sie wären auch nicht plausibel, jedenfalls dann nicht, wenn diesen Schülereinschätzungen primär Merkmale des Unterrichts und nicht nur subjektive Fakto-

40 Für die untersuchten Klassenstufen liegt die interne Konsistenz der Skala bei $\alpha = 0,72$ (Fünftklässler) bzw. $\alpha = 0,79$ (Neuntklässler).

ren zu Grunde liegen. Dass die Einschätzungen in diesem Sinne valide sind, lässt sich varianzanalytisch belegen (Klassenstufe 5: $Eta^2 = 0{,}26$ für die Lerngruppe; Klassenstufe 9: $Eta^2 = 0{,}36$).

Abbildung 26 Verteilung der Werte auf der Skala „Wahrnehmung des eigenen Mathematikunterrichts“, nach Klassenstufe



Zwischen den Schulformen gibt es hinsichtlich der verschiedenen hier verwendeten Skalen ein relativ stabiles Muster. Im Allgemeinen sind die schul- und unterrichtsbezogenen Wahrnehmungen, Bewertungen und Einstellungen an den Gymnasien und in den Erweiterungskursen der Gesamtschulen am günstigsten oder doch besonders positiv ausgeprägt, während sie sich umgekehrt bei den Schülerinnen und Schülern der Grundkurse im Vergleich – aber auch nur im Vergleich – am wenigsten positiv darstellen. Letzteres mag auch mit den schulinternen Auslesemechanismen der Gesamtschulen zusammenhängen: Bei der Aufteilung der Schülerschaft von Gesamtschulen in Grund- und Erweiterungskurse spielen außer dem dominanten Aspekt der Fachleistung nachweislich auch affektive Schülermerkmale eine eigenständige Rolle, was gleichbedeutend ist mit einer relativ ungünstigen Motivationslage in den Grundkursen einerseits und einer vergleichsweise positiven Situation in den Erweiterungskursen andererseits. Trotzdem ist es bemerkenswert, dass in den Erweiterungskursen der Unterricht als besonders zielgerichtet, transparent, wohlstrukturiert und lernförderlich wahrgenommen wird.

Die erfragten Aspekte der lern- und schulbezogenen Einstellungen der Schülerinnen und Schüler sind nicht voneinander unabhängig; in einigen Fällen geht die positive Einschätzung der Aussagen in einer Dimension mit der positiven Beurteilung in einer anderen einher (vgl. Tabelle 27). Dabei zeigt sich eine über beide Klassenstufen und alle Schulformen sowie über die Geschlechter hin stabile Struktur. Zum einen hängt das Interesse am Fach Mathematik relativ eng mit dem leistungsbezogenen Selbstkonzept zusammen. Diejenigen, die sich besonders für Mathematik interessieren, sind gleichzeitig selbstbewusster und umgekehrt. Anscheinend ist es besonders schwierig, Schülerinnen und Schüler mit geringem Selbstvertrauen und hoher Leistungsangst für das Fach Mathematik zu interessieren. Zum anderen bilden die drei Einstellungsdimensionen „Anstrengung als Erfolgsfaktor", „Schulzufriedenheit" und „Wahrnehmung des eigenen Mathematikunterrichts" eine Gruppe verwandter Einstellungen. Diejenigen äußern allgemein ein höheres Maß an Schulzufriedenheit, die stärker die Bedeutung eigener Anstrengung betonen und die auch den eigenen Mathematikunterricht als transparenter und strukturierter erleben.

Tabelle 27 Lern- und schulbezogene Einstellungen bzw. Selbsteinschätzungen: Interkorrelationen

	leistungs-bezogenes Selbstkonzept	Sach-/ Fachinteresse an Mathematik	Anstrengung als Beitrag zum Schulerfolg	Wahrnehmung des eigenen Mathematik-unterrichts
Klassenstufe 5				
Schulzufriedenheit	0,20	0,36	0,47	0,49
leistungsbezogenes Selbstkonzept	---	0,50	0,17	0,08
Sach-/ Fachinteresse an Mathematik	---	---	0,41	0,28
Anstrengung als Beitrag zum Schulerfolg	---	---	---	0,34
Klassenstufe 9				
Schulzufriedenheit	0,16	0,31	0,40	0,37
leistungsbezogenes Selbstkonzept	---	0,43	0,05	0,10
Sach-/ Fachinteresse an Mathematik	---	---	0,25	0,31
Anstrengung als Beitrag zum Schulerfolg	---	---	---	0,28

Selbstverständlich ist es von hohem Interesse, wie sich die gemessenen Leistungen im Fach Mathematik zu den erfassten Einstellungsmerkmalen verhalten. In der Tabelle 28 werden die lern- und schulbezogenen Einstellungen zu den Ergebnissen im QuaSUM-Mathematiktest und im Mathe40-Test in Beziehung gesetzt.

Tabelle 28 Lern- und schulbezogene Einstellungen bzw. Selbsteinschätzungen: Korrelationen mit der Mathematikleistung im QuaSUM-Mathematiktest und im Mathe40-Test

	Schul-zufriedenheit	leistungs-bezogenes Selbst-konzept	Sach-/ Fachinteresse an Mathematik	Anstrengung als Beitrag zum Schul-erfolg	Wahrneh-mung des eigenen Mathematik-unterrichts
Klassenstufe 5	0,11	0,41	0,42	0,20	-0,01
Klassenstufe 9 (QuaSUM-Mathematiktest)					
Gesamtschulen, Grundkurse	0,03	0,14	0,24	0,10	0,06
Gesamtschulen, Erweiterungskurse	0,05	0,23	0,38	-0,01	0,12
Realschulen	0,12	0,32	0,47	0,06	0,20
Gymnasien	0,08	0,33	0,50	-0,03	0,10
insgesamt	*0,17*	*0,28*	*0,36*	*0,11*	*0,12*
Klassenstufe 9 (Mathe40-Test)					
Gesamtschulen, Grundkurse	0,00	0,15	0,21	0,06	-0,01
Gesamtschulen, Erweiterungskurse	-0,03	0,25	0,28	-0,04	0,01
Realschulen	0,04	0,28	0,37	-0,01	0,09
Gymnasien	0,03	0,31	0,43	-0,09	0,05
insgesamt	*0,11*	*0,28*	*0,33*	*0,06*	*0,06*

In der Klassenstufe 5 sind die Zusammenhänge zwischen mathematischem Sach- und Fachinteresse sowie einem positiven leistungsbezogenen Selbstkonzept und der gezeigten Mathematikleistung besonders deutlich ausgeprägt. Derselbe Zusammenhang zeichnet sich – wenn auch nicht so stark wie bei den Grundschülerinnen und -schülern – auch bei den Neuntklässlern ab. Die Korrelationen sind bei den Schülerinnen und Schülern aus Gesamtschulen geringer ausgeprägt als bei denen aus Gymnasien und Realschulen; insbesondere die Schülerinnen und Schüler aus Mathematik-Grundkursen zeigen niedrigere Zusammenhänge zwischen Interesse und Leistung sowie zwischen der Einschätzung des eigenen Leistungsvermögens und gezeigter Mathematikleistung. Dabei ist zu berücksichtigen, dass sich hier das Einstellungsprofil und die Fachleistungen gleichermaßen ungünstig darstellen, letztere also auf niedrigem – und zugleich vergleichsweise homogenem – Niveau. Das erklärt zumindest teilweise, weshalb hier die statistischen Zusammenhänge abgeschwächt erscheinen.

Einen besonderen Hinweis verdient der Umstand, dass die Bedeutung, die die Schülerinnen und Schüler der eigenen Anstrengung beimessen, in der Klassenstufe 9, gemessen an der Klassenstufe 5, nur noch wenig lernerfolgsrelevant zu sein scheint. Hier ist daran zu erinnern, dass bei den Jugendlichen die Anstrengungen selbst im Vergleich mit den Fünftklässlerinnen und Fünftklässlern nachgelassen haben, wie die Angaben zum Zeitaufwand für die Hausaufgaben belegen. Offenkundig gelingt es in den neunten Klassen aller Schulformen bzw. Kursniveaus nicht (mehr) hinreichend, insbesondere den weniger leistungsfähigen Schülerinnen und Schülern – aber eben nicht nur ihnen – die Erfahrung zu vermitteln, dass eigene Anstrengungen Erfolge zeitigen können. Dies legt bei den Jugendlichen ein Attributionsmuster nahe, demzufolge letztlich unbeeinflussbare externe Faktoren und „Begabungsstrukturen“ maßgeblich für den Schulerfolg sind. Dass sich eine solche „implizite Handlungstheorie“ negativ auf die Lernentwicklung auswirkt, darf als gesicherte Erkenntnis der Pädagogischen Psychologie gelten.

5 Schulischer und unterrichtlicher Kontext der Mathematikleistungen

5.1 Schulische Rahmenbedingungen des Mathematikunterrichts

In diesem Abschnitt sollen ausgewählte Ergebnisse aus der Befragung der Schulleiterinnen und Schulleiter ($N = 180$) vorgestellt werden, die parallel zu dem Test- und Befragungsprogramm bei den Schülerinnen und Schülern durchgeführt wurde. Dabei soll die Aufmerksamkeit weniger auf der ausführlichen Beschreibung der Schulen des Landes Brandenburg liegen, als vielmehr auf der Prüfung möglicher Zusammenhänge mit der gemessenen Schülerleistung im Fach Mathematik.

5.1.1 Äußere schulische Rahmenbedingungen

Der Blick richtet sich zunächst auf ausgewählte äußere Rahmenbedingungen: die Größe und Ausstattung der Schulen sowie die – nach der Einschätzung der befragten Schulleiterinnen und Schulleiter – gegebene soziale Lage der Familien im Einzugsgebiet der Schule.

Es wird vielfach vermutet, dass die *Größe der Schulen* mit der Vielfalt der Lernangebote und auf diese Weise indirekt positiv mit dem Lernerfolg in Zusammenhang steht. Im Schuljahr 1998/99 besuchten 60,7 Prozent der Brandenburger Schülerinnen und Schüler der Klassenstufe 5 Schulen kleinerer und mittlerer Größe (101 bis 400 Schüler), für die nur zum Teil Mehrzügigkeit angenommen werden kann. Schließt man die angegliederten Grundschulen an Gesamtschulen aus dieser Berechnung aus, so ist der Wert mit 64,9 Prozent sogar noch etwas höher[41]. Tabelle 29 gibt für die Fünftklässlerinnen und Fünftklässler und für die Neuntklässlerinnen und Neuntklässler – hier getrennt für die Schulformen Gesamtschule, Realschule und Gymnasium – die Verteilung der darin unterrichteten Schülerinnen und Schüler an.

[41] Im Jahr 1998 gab es im Land Brandenburg 610 Grundschulen. Durch den starken Rückgang der Schülerzahlen – der eigentliche Tiefpunkt wird für das Schuljahr 2000/01 erwartet – waren viele „normale" Grundschulen im ländlichen Raum von einer Schließung betroffen. Noch im Oktober 1998 konnten davon 65 Grundschulen im Rahmen des Modellversuchs „Kleine Grundschule" erhalten bleiben (vgl. MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT DES LANDES BRANDENBURG 1999b).

Tabelle 29 Größe der Schulen, nach Klassenstufe und Schulform (Schüleranteile in Prozent)

	Klassenstufe 5	**Klassenstufe 9**		
Anzahl der Schülerinnen und Schüler	Grundschule	Gesamtschule	Realschule	Gymnasium
bis 100	1,1	0,0	0,0	0,0
101 bis 200	17,0	2,6	7,4	0,0
201 bis 400	43,7	49,8	68,3	4,7
401 bis 600	24,3	27,7	24,3	9,9
601 bis 800	13,4	15,9	0,0	59,7
801 bis 1000	0,6	4,0	0,0	19,5
über 1000	0,0	0,0	0,0	6,2
durchschnittliche Anzahl der Schülerinnen und Schüler pro Schule	*376,0*	*442,4*	*346,0*	*753,6*

Bezogen auf die Schulgröße lässt sich mit einer Ausnahme kein Effekt auf die erreichte Testleistung im QuaSUM-Mathematiktest beobachten: Nur an den Gesamtschulen ist – Grund- und Erweiterungskurse zusammengenommen – an großen Schulen das Leistungsniveau höher als an kleinen ($r = 0,14$)[42].

Im Folgenden geht es um die *räumlich-apparative Ausstattung* der Brandenburger Schulen und um deren Zusammenhänge mit den erreichten Lernständen in Mathematik. In der folgenden Tabelle 30 sind ausgewählte Ausstattungsmerkmale aufgeführt, prozentuiert über die Anteile der davon betroffenen Schülerinnen und Schüler. In diesen Befunden zeigt sich ein Ausstattungsmuster, das sich primär an schulstufenspezifischen Unterrichtsformen orientiert: Für die Brandenburger Fünftklässlerinnen und Fünftklässler sind Klassenräume mit Funktionsecken typisch. Lernwerkstätten finden sich immerhin an etwa jeder vierten untersuchten Schule, und zwar tendenziell eher an den kleineren; nur jede fünfte Schülerin bzw. jeder fünfte Schüler hat potenziellen Zugang dazu. Fachräume für Mathematik sind dagegen erwartungsgemäß für die Sekundarstufe I, insbesondere für das Gymnasium, charakteristisch. Die Ausstattung der Schulen mit Computern steckt hingegen

---

42 Mit der Größe der Schule hängt unmittelbar die Zahl der dort beschäftigten Lehrkräfte zusammen. Die Schulformen unterscheiden sich stark hinsichtlich des Anteils der Teilzeitkräfte: An den reinen Grundschulen scheint er besonders hoch zu sein, gefolgt von den Gesamtschulen. Da von den Schulleitungen keine Umrechnungen der Teilzeitstellen in Vollzeitäquivalente erbeten wurde, scheidet eine saubere Berechnung von Schüler-Lehrer-Relationen aus, die für eine Berücksichtigung der je Schüler eingesetzten personellen Ressourcen wichtig gewesen wäre. Angaben liegen hingegen zu den Klassen- bzw. Kursgrößen vor, die im Abschnitt 5.2.1 näher untersucht werden.

noch in den Anfängen: nur etwa jede dritte Grundschülerin bzw. jeder dritte Grundschüler und etwa jede zweite Schülerin bzw. jeder zweite Schüler der Sekundarstufe I (48 Prozent) hat Zugang dazu.

Tabelle 30 Räumlich-apparative Ausstattung an Brandenburger Schulen, nach Klassenstufe und Schulform (in Prozent der davon betroffenen Schülerinnen und Schüler)

	Fachräume für Mathematik	**Klassenräume mit Funktionsecken**	**Lernwerkstatt**	**Computer**
Klassenstufe 5	14,2	74,2	19,8	33,0
Klassenstufe 9				
Gesamtschulen	69,0	15,3	7,6	53,4
Realschulen	55,8	5,5	0,0	56,6
Gymnasien	80,7	10,2	0,0	33,5

Hinsichtlich eines Zusammenhangs der räumlichen Ausstattung mit der erreichten Testleistung im QuaSUM-Mathematiktest lassen sich für die Schülerinnen und Schüler in der Klassenstufe 5 keine Effekte nachweisen. Offenbar kommt es weniger darauf an, dass eine bestimmte Einrichtung vorhanden ist, als vielmehr darauf, wie sie genutzt wird. Einzig die Ausstattung der Grundschulen mit Computern geht mit höheren Mathematikleistungen einher, wie sich selbst dann zeigt, wenn man die soziale und kognitive Eingangsselektivität der Schulen kontrolliert. In der Klassenstufe 9 zeigt sich ein ähnliches Bild. An Gesamt- und Realschulen mit Fachräumen wurden geringfügig höhere Leistungen angetroffen; für die rund 20 Prozent der Gymnasien, die über keinen solchen Raum verfügen, kann hingegen kein Leistungsnachteil festgestellt werden. Hinsichtlich der Ausstattung der Schulen mit Computern sind die Befunde im Sekundarbereich uneinheitlich. Nur unter den Gesamtschulen können nach Berücksichtigung der Zusammensetzung der Schülerschaft Leistungsvorteile für diejenigen Schulen aufgezeigt werden, die über Computer verfügen. Diese Vorteile beschränken sich aber auf Erweiterungskurse (d = 0,15). Es gibt demnach keineswegs einen Automatismus, über den sich bestimmte Investitionen unmittelbar in höhere Fachleistungen umsetzen.

Die Schulleiterinnen und Schulleiter wurden befragt, wie sie insgesamt die *soziale Lage der Familien im Einzugsgebiet der Schule* einschätzen. Für diese Einschätzung wurde eine 5-stufige Skala von „weit unter dem Brandenburger Durchschnitt“ bis „weit über dem Brandenburger Durchschnitt“ eingesetzt. Die Extremkategorien wurden von den Schulleitungen nicht genutzt. In der Klassenstufe 9 lässt sich zwar erwartungsgemäß beobachten, dass die Einzugsgebiete – vielleicht besser: die soziale Herkunft der Schülerschaft – der Gymnasien als günstiger eingeschätzt wurden als die der anderen Schulformen. Der Zusammenhang zwischen Herkunftsmilieu der Schüler und Fachleistung Mathematik ist aber primär auf Individualebene vermittelt, so dass das globale Urteil der Schulleitung im Hinblick auf die Testleistung nur wenig zu erklären vermag. So konnte lediglich für die Gesamtschulen auf dieser hoch aggregierten Ebene ein nennenswerter Zusammenhang (r = 0,09) festgestellt werden.

5.1.2 Innere schulische Rahmenbedingungen

In diesem Abschnitt werden die Ergebnisse zu den inneren schulischen Rahmenbedingungen aus der Sicht der befragten Schulleiterinnen und Schulleiter vorgestellt. Hierbei stehen Fragestellungen im Vordergrund, nach denen insbesondere ein störungsarmes, lernzentriertes Schulklima ein relevantes Merkmal für eine erfolgreiche Schule ist (vgl. dazu BAUMERT 1989).

In dem an sie gerichteten Fragebogen sollten die Schulleiterinnen und Schulleiter zu einer Reihe von Aussagen Stellung nehmen, die sich auf Aspekte des Schullebens und der eigenen Tätigkeit beziehen. Es ist bemerkenswert, dass sich dabei die subjektiven Wahrnehmungen zwischen den Schulleitungen der Grundschulen und der Schulen im Sekundarbereich merklich unterscheiden. Während sich an den Schulen der Sekundarstufe I drei Hauptaspekte unterscheiden lassen – erstens das allgemeine Schulklima und die Kooperation der Lehrkräfte, zweitens bestimmte Beeinträchtigungen wie Raumnot, Ausstattungsmangel und Schülerdevianz, drittens die Betonung des eigenen pädagogischen und administrativen Engagements –, ordnen sich die einzelnen Aspekte an den Grundschulen anders. Zwar besitzt auch hier die Betonung des eigenen pädagogischen Engagements das größte Gewicht, doch das Schulklima wird stärker im Zusammenhang mit der Ausstattung und dem Schülerverhalten beurteilt, und die Kooperation der Lehrkräfte untereinander wird gerade nicht im Zusammenhang mit dem allgemeinen Schulklima beurteilt. Aus den Antworten der Schulleiterinnen und Schulleiter konnten insgesamt neun konsistente Skalen gebildet werden, die jeweils einen bestimmten Einzelaspekt der inneren schulischen Rahmenbedingungen beschreiben[43]. In jedem Falle wurden auch die Zusammenhänge mit der gemessenen Mathe-

[43] Neben „Schulklima“ und „Beeinträchtigung des Unterrichts durch abweichendes Verhalten“, auf die im Folgenden näher eingegangen wird, konnten aus den Einschätzungen der Schulleitungen Skalen zu den Aspekten „ Zusammenarbeit der Lehrkräfte“, „Ausstattungsmangel“, „Raummangel“, „Verwaltungstätigkeiten“, „Unterstützung der Lehrkräfte“, „Bedeutung von Kritikfähigkeit und Kreativität bei Schülern“ und „Beratung von Lehrern, Eltern und Schülern“ gebildet werden.

matikleistung und den schul- und unterrichtsrelevanten Einstellungen der Schülerschaft geprüft. Insgesamt sind die korrelativen Zusammenhänge relativ schwach ausgeprägt. Im Folgenden werden deshalb nur für zwei dieser (Summen-)Skalen Ergebnisse vorgestellt und diskutiert[44].

Die Skala „*Schulklima aus der Sicht der befragten Schulleiterinnen und Schulleiter*" gründet sich auf die Bewertung von insgesamt 7 Aussagen wie z. B.: „Die Beziehung zwischen Schülern und Lehrkräften ist gut", „Es herrscht ein guter Gemeinschaftsgeist an dieser Schule" und – in umgekehrter Richtung – „In dieser Schule sind Disziplinschwierigkeiten (Schwänzen, zu spät Kommen, Stören des Unterrichts) an der Tagesordnung". Die befragten Schulleiterinnen und Schulleiter tendieren bei der Beurteilung des Schulklimas zu einer eher positiven Einschätzung: Für beide Klassenstufen und für die unterschiedlichen Schulformen liegen die Mittelwerte zwischen 2,76 und 3,13 (vgl. Tabelle 31).

Tabelle 31 Statistische Kennwerte der Skala „Schulklima aus der Sicht der Schulleitung"

	Mittelwert*	Standardabweichung
Klassenstufe 5	3,13	0,31
Klassenstufe 9		
Gesamtschulen	2,76	0,36
Realschulen	2,89	0,24
Gymnasien	3,03	0,23
insgesamt	*2,86*	*0,33*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Im Bereich der Sekundarstufe I ist nur an den Gymnasien ein Schulklima zu verzeichnen, das dem vergleichsweise hohen Wert für die Grundschulen wenigstens näherungsweise entspricht. In den übrigen Schulformen wird dieser Aspekt von den Schulleitungen deutlich kritischer eingeschätzt. Die (fast) allgemeine „Verschlechterung" des Schulklimas im Anschluss an die Grundschulzeit entspricht zunächst durchaus der Einschätzung der Situation durch die Schülerinnen und Schüler (vgl. die Ergebnisse zur Skala „Schulzufriedenheit" in Abschnitt 4.4), und sie könnte auch mit dem häufiger registrierten abweichenden Verhalten von Schülerinnen und Schüler an den Schulen der

[44] Die Zusammenfassung der Aussagen zu einer Skala erfolgte selbstverständlich unter der Voraussetzung, dass positiv und negativ formulierte Aussagen bei der Verrechnung gleichsinnig gepolt werden. Die interne Konsistenz der Skala liegt für die untersuchten Klassenstufen jeweils höher als $\alpha = 0{,}72$. Für die außerdem in diesem Abschnitt diskutierten Skalen gelten die gleichen Voraussetzungen.

Sekundarstufe I zusammenhängen[45]. Zu beachten bleibt aber, dass innerhalb der Klassenstufen und Schulformen die Korrelationen zwischen den Einschätzungen der Schulleitungen und der von den Schülerinnen und Schülern geäußerten Schulzufriedenheit vernachlässigbar gering sind. Sehr Ähnliches gilt übrigens für das Leistungskriterium: Korrelative Zusammenhänge zwischen der Wahrnehmung des Schulklimas aus der Sicht der Schulleitung und der Testleistung der Schülerinnen und Schüler erweisen sich weitgehend als schulformbedingt. Die einzige Ausnahme stellen mit $r = 0{,}13$ die Realschulen dar, die allerdings untereinander wesentlich heterogener sind als die übrigen Schulformen.

Eine zweite Skala der erfragten inneren schulischen Rahmenbedingungen bezieht sich auf Aussagen zur Frage „Wie stark beeinträchtigen die folgenden Umstände Ihrer Ansicht nach den Unterricht an Ihrer Schule?“ (u. a.: „Schüler, die den Unterricht stören“, „uninteressierte Schüler“ und „Gewalt an der Schule“. Es geht also um die Einschätzung der *Beeinträchtigung des Unterrichts durch abweichendes Schülerverhalten*. In Tabelle 32 sind hierzu statistische Kennwerte für die beiden untersuchten Klassenstufen aufgeführt.

Tabelle 32 Statistische Kennwerte der Skala „Beeinträchtigung des Unterrichts durch abweichendes Schülerverhalten“

	Mittelwert*	Standardabweichung
Klassenstufe 5	2,10	0,46
Klassenstufe 9		
Gesamtschulen	2,62	0,84
Realschulen	2,37	0,91
Gymnasien	2,58	1,04
insgesamt	*2,56*	*0,92*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Für die Klassenstufe 5 liegt der Mittelwert in einem Bereich, wo Aussagen über Beeinträchtigungen des Unterrichts durch abweichendes Schülerverhalten eher verneint werden. Für die unterschiedlichen Schulformen in der Klassenstufe 9 dagegen wird außer an den Realschulen der theoretische Skalenmittelwert überschritten. Außerdem fällt hierbei auf, dass die Standardabweichungen wesentlich höher liegen als in der Klassenstufe 5 oder auch für die

[45] In den Klassenstufen 5 und 9 zeigen sich für die Einschätzungen der befragten Schulleiterinnen und Schulleiter deutlich negative Zusammenhänge zwischen der Skala „Schulklima“ und der Skala „Beeinträchtigung des Unterrichts durch abweichendes Schülerverhalten“ ($r = -0{,}43$ bzw. $r = -0{,}45$): Je höher die Beeinträchtigung des Unterrichts durch abweichendes Schülerverhalten eingeschätzt wird, desto ungünstiger ist die Einschätzung des Schulklimas.

ebenfalls 4-stufige Skala „Schulklima aus der Sicht der Schulleitung“ in der Klassenstufe 9 (vgl. Tabelle 31). Diese vergleichsweise hohe Streuung bedeutet, dass von Schule zu Schule manchmal eine starke und manchmal auch fast keine Beeinträchtigung konstatiert wird. Es zeigen sich jedoch – ebenso wenig wie für die Skala „Schulklima“ und die übrigen sieben Schulleiterskalen – keine konsistenten Zusammenhänge mit den Leistungsvariablen und den Schülereinstellungen.

Im Hinblick auf einen Teil der vorliegenden Literatur zur Schulentwicklung und zum Schulmanagement mögen diese Befunde als enttäuschend empfunden werden, zumal auch das Konstrukt „Zusammenarbeit der Lehrkräfte aus der Sicht der Schulleitung“ nur an den Gymnasien mit $r = 0{,}11$ mit dem Fachleistungsniveau erkennbar – allerdings schwach – zusammenhängt (bei nur drei Items und entsprechend geringer Messgenauigkeit: $\alpha = 0{,}62$).

Die berichteten Ergebnisse bedeuten nicht, dass die inneren Rahmenbedingungen einer Schule für die Entwicklung eines erfolgreichen Unterrichts gleichgültig wären. Klar ist nur – wie auch bei der Analyse der Effekte von äußeren Rahmenbedingungen –, dass die eigentlichen Mechanismen für die Verbesserung des Unterrichts sehr viel genauer rekonstruiert werden müssen, als es im Rahmen einer notgedrungen globalen Schulleiterbefragung möglich war.

5.2 Merkmale des Mathematikunterrichts

Im Kontext der QuaSUM-Untersuchung fand auch eine Befragung der Mathematiklehrkräfte ($N = 582$) statt. In diesem Abschnitt soll vor allem untersucht werden, ob die so gewonnenen Äußerungen von Lehrerinnen und Lehrern zu den äußeren und inneren Bedingungen des Mathematikunterrichts einen Beitrag zur Erklärung der erreichten Lernstände im Fach Mathematik an Brandenburger Schulen leisten.

5.2.1 Äußere Merkmale des Mathematikunterrichts

Zu den wichtigsten äußeren Merkmalen schulischen Unterrichts gehört die Größe der Lerngruppe – der Klasse oder des Kurses –, weil gemeinhin angenommen wird, dass in kleineren Gruppen mehr und besser gelernt wird. Auch eine adäquate materielle Ausstattung, von der die Schülerinnen und Schüler unmittelbar betroffen sind, könnte sich positiv auf das jeweilige Leistungsniveau auswirken. Und schließlich gehört aus der Sicht der Lehrkräfte auch das Schulklima und die Unterstützung durch die Schulleitung zu denjenigen äußeren Rahmenbedingungen, von denen angenommen wird, dass sie einen positiven Einfluss auf den Unterrichtserfolg haben. Hier soll deshalb untersucht werden, ob bzw. in welchem Maße diese Vermutungen zutreffen.

Hinsichtlich der *Klassen- bzw. Kursgrößen* bestehen charakteristische Unterschiede zwischen den verschiedenen Schulformen, wie die nachstehende Tabelle 33 belegt.

Tabelle 33 Klassen- bzw. Kursgröße (Anzahl der Schülerinnen und Schüler), nach Schulform

	Mittelwert	Standardabweichung
Klassenstufe 5	22,6	3,2
Klassenstufe 9		
Gesamtschulen, Grundkurse	16,0	3,5
Gesamtschulen, Erweiterungskurse	18,3	3,7
Realschulen	26,0	3,1
Gymnasien	26,8	2,5
insgesamt	*20,4*	*5,8*

In der Klassenstufe 5 gibt es durchschnittlich etwas weniger als 23 Schülerinnen und Schüler pro Klasse; die kleinsten Klassen bestehen aus 13, die größten aus 30 Schülerinnen und Schülern. Ein positiver Effekt kleiner Klassen auf die erreichte Testleistung der Schülerinnen und Schüler im QuaSUM-Mathematiktest lässt sich für die Schülerinnen und Schüler in der Klassenstufe 5 nicht nachweisen (r = -0,02).

In der Klassenstufe 9 gibt es durchschnittlich etwas mehr als 20 Schülerinnen und Schüler pro Klasse bzw. Kurs. Auffällig ist die personell aufwändigere Organisationsform der Gesamtschulen: In den Grundkursen (16,0 Schüler) und in den Erweiterungskursen (18,3 Schüler) der Gesamtschulen sind die Lerngruppen wesentlich kleiner als an den Realschulen (26,0 Schüler) und an den Gymnasien (26,8 Schüler). Diese Werte stimmen gut mit der amtlichen Schulstatistik des Landes Brandenburg überein (vgl. MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT DES LANDES BRANDENBURG 1999b). Insgesamt ist auch in der Klassenstufe 9 ein günstiger Einfluss kleiner Lerngruppen auf die Fachleistung Mathematik nicht nachweisbar. In den Erweiterungskursen an Gesamtschulen und in den Realschulklassen geht der in größeren Klassen bzw. Kursen erteilte Mathematikunterricht sogar mit etwas höheren durchschnittlichen Ergebnissen der Schülerinnen und Schüler im QuaSUM-Mathematiktest einher (r = 0,15 bzw. r = 0,16). Nur in den Gymnasialklassen gibt es einen, allerdings schwachen, umgekehrten Trend (r = -0,07), während für die Schülerinnen und Schüler in den Grundkursen an den Gesamtschulen keinerlei Effekt erkennbar ist.

Interessant ist in diesem Zusammenhang die Stellungnahme der Lehrkräfte. Zu erwarten ist nach diesen Befunden, dass die Lehrerinnen und Lehrer von Realschulen und Gymnasien häufiger über zu große Lerngruppen klagen als ihre Kolleginnen und Kollegen von Gesamtschulen. Dies ist in der Tat der Fall. Ausgesprochen erwartungswidrig ist es hingegen, dass bei den Lehrkräften der Gesamtschulen gerade jene in hohen Schülerzahlen eine Beeinträchtigung des Unterrichts sehen, die in eher leistungsstarken Klassen unterrichten (r = 0,12). Aus alledem ist der Schluss zu ziehen, dass das Verhältnis von Lerngruppengröße und Lernerfolg jedenfalls erheblich komplexer ist, als es die Formel „je kleiner, desto besser“ nahe legt.

Zu den möglichen Auswirkungen der *materiellen Ausstattung* gibt es Einzelaussagen, die sich jedoch nicht zu einem kohärenten Bild zusammenfügen. Nur in den Realschulen gehen niedrigere Durchschnittsleistungen tendenziell auch mit Beschwerden der Lehrkräfte über eine unzureichende materielle Ausstattung einher (r = -0,11); in allen anderen Gruppen gibt es keinen Zusammenhang zwischen solchen Klagen und der erreichten Fachleistung. Ebenfalls nur an den Realschulen äußern die Lehrkräfte von eher leistungsschwachen Klassen, dass sie in ihrer pädagogischen Arbeit durch zu viel „alltäglichen Kleinkram“ beeinträchtigt würden (r = -0,16). Auch in dieser Hinsicht kann also nicht von einfachen Zusammenhängen zwischen den objektiven Rahmenbedingungen, deren subjektiver Wahrnehmung durch die Lehrkräfte und dem Erfolg des Unterrichts ausgegangen werden.

Das *Schulklima aus der Sicht der Mathematiklehrkräfte* konnte über eine Summenskala erfasst werden, die auf der Bewertung von Aussagen beruht, die auch die Grundlage für die Skala „Schulklima aus der Sicht der Schulleitungen“ war. Die Befunde enthält die nachstehende Tabelle 34. Das Schulklima wird von der Mehrzahl der Lehrkräfte positiv eingeschätzt[46]. Für beide Klassenstufen und für die unterschiedlichen Schul- bzw. Kursformen in der Klassenstufe 9 liegen die Mittelwerte zwischen 2,77 und 3,13, wobei sowohl die Schulstufen- als auch die Schulformeffekte auffällig sind.

[46] In den Klassenstufen 5 und 9 lassen sich enge Zusammenhänge zwischen dem Lern- und Arbeitsklima in der Klasse bzw. im Kurs und der Beurteilung des Schulklimas aus der Sicht der Lehrkräfte beobachten. In der Klassenstufe 5 (r = 0,50) ist dieser Zusammenhang deutlicher ausgeprägt als in der Klassenstufe 9 (r = 0,41). Bemerkenswert sind in der Differenzierung nach Schulform bzw. Kursniveau auch die Ähnlichkeiten mit den Beurteilungen des Schulklimas aus der Sicht der Schulleitungen (vgl. Tabelle 31).

Tabelle 34 Statistische Kennwerte der Skala „Schulklima aus der Sicht der Mathematiklehrkräfte“

	Mittelwert*	Standardabweichung
Klassenstufe 5	3,17	0,25
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,77	0,33
Gesamtschulen, Erweiterungskurse	2,85	0,32
Realschulen	3,02	0,32
Gymnasien	3,13	0,23
insgesamt	*2,93*	*0,34*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Im Vergleich der Schul- bzw. Kursformen in der Klassenstufe 9 bewerten die befragten Mathematiklehrkräfte an den Gymnasien das Schulklima am günstigsten. Die befragten Mathematiklehrkräfte, die Grundkurse an den Gesamtschulen unterrichten, geben hier im Durchschnitt den niedrigsten Wert an. Hinsichtlich des Zusammenhangs zwischen dem Schulklima aus der Sicht der befragten Mathematiklehrkräfte und der erreichten Leistung der Schülerinnen und Schüler im QuaSUM-Mathematiktest lässt sich in der Klassenstufe 9 ($r = 0{,}36$) über alle Schulformen bzw. Kursniveaus hinweg ein ausgeprägterer positiver Zusammenhang beobachten als in der Klassenstufe 5 ($r = 0{,}10$). Bei differenzierterer Betrachtung erweist sich der Zusammenhang im Bereich der Sekundarstufe I allerdings als ebenso schwach wie an den Primarschulen. An den Gesamtschulen beträgt er $r = 0{,}14$, an den Gymnasien $r = 0{,}05$ und nur an den Realschulen $r = 0{,}23$. Bemerkenswert ist, dass sich die Einschätzungen des Schulklimas aus der Sicht der Schulleitung und aus der Sicht der Lehrkräfte auf der Ebene der Einzelschule nur an den Gesamtschulen mit $r = 0{,}33$ einigermaßen decken. An den Realschulen ($r = -0{,}05$) und an den Gymnasien ($r = 0{,}14$) haben sie offenbar wenig miteinander gemein.

Ein weiterer externer Faktor ist die *Unterstützung durch die Schulleitung*. Entsprechende Einschätzungen wurden durch die Mathematiklehrkräfte mit der Bewertung von insgesamt 4 Aussagen ausgedrückt wie z. B.: „Die Schulleitung fördert Kontakte zu den Fachkollegen“ und „Die Schulleitung initiiert Lehrerfort- und -weiterbildungen“. Im Hinblick auf das Engagement der Schulleitungen tendieren die befragten Mathematiklehrkräfte in der Klassenstufe 5 zu einer positiveren Antwort als ihre Kolleginnen und Kollegen der Klassenstufe 9 (vgl. Tabelle 35).

Tabelle 35 Statistische Kennwerte der Skala „Unterstützung durch die Schulleitung“

	Mittelwert*	Standardabweichung
Klassenstufe 5	2,85	0,55
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,58	0,53
Gesamtschulen, Erweiterungskurse	2,48	0,63
Realschulen	2,62	0,47
Gymnasien	2,31	0,56
insgesamt	*2,48*	*0,57*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Die Schulleitungen an den Gymnasien erhalten hier von ihren Mathematiklehrkräften im Durchschnitt den niedrigsten Wert. Indessen fehlt hier auch jeder nennenswerte Zusammenhang zwischen den Selbsteinschätzungen der Schulleitungen und den Fremdeinschätzungen durch die befragten Lehrkräfte (Gymnasien: r = 0,05; Realschulen: r = 0,17; Gesamtschulen: r = 0,19). Ein nennenswerter Zusammenhang mit der Fachleistung in Mathematik ist mit r = 0,25 wiederum nur an den Realschulen zu beobachten.

5.2.2 Innere Merkmale des Mathematikunterrichts

Die über den Lehrerfragebogen erhobenen Daten haben es ermöglicht, vier weitere Skalen zu bilden, die als Repräsentanten innerer Merkmale des Mathematikunterrichts aus Lehrerperspektive gelten können. Eingeschätzt werden sollte jeweils das Lern- und Arbeitsklima in der Klasse, die Bedeutung der Eigenständigkeit und Kreativität der Lernenden, der Stellenwert eng kontrollierender Unterrichtsführung und die Bedeutung des Lehrbuchs für den Unterricht.

Die Skala zum *Lern- und Arbeitsklima in der Klasse bzw. im Kurs* bezieht sich auf insgesamt 9 Aussagen (z. B.: „In dieser Klasse bzw. in diesem Kurs sind Disziplinschwierigkeiten an der Tagesordnung“, „Die Schüler dieser Klasse bzw. dieses Kurses haben eine positive Einstellung zu schulischer Leistung“ und „Die Beziehung zwischen Schülern und Lehrkräften ist in dieser Klasse bzw. in diesem Kurs gut“). Das Lern- und Arbeitsklima in den Klassen und Kursen wird von der Mehrzahl der Lehrkräfte insgesamt eher positiv eingeschätzt, doch zeigt sich eine differenzielle Entwicklung zwischen den Schulformen bzw. Kursniveaus, sofern man den Kohortenvergleich als Quasi-Längsschnitt interpretiert (vgl. Tabelle 36).

Tabelle 36 Statistische Kennwerte der Skala „Lern- und Arbeitsklima in der Klasse bzw. im Kurs“

	Mittelwert*	Standardabweichung
Klassenstufe 5	3,14	0,38
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,53	0,43
Gesamtschulen, Erweiterungskurse	3,13	0,33
Realschulen	2,73	0,50
Gymnasien	3,08	0,41
insgesamt	*2,84*	*0,49*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

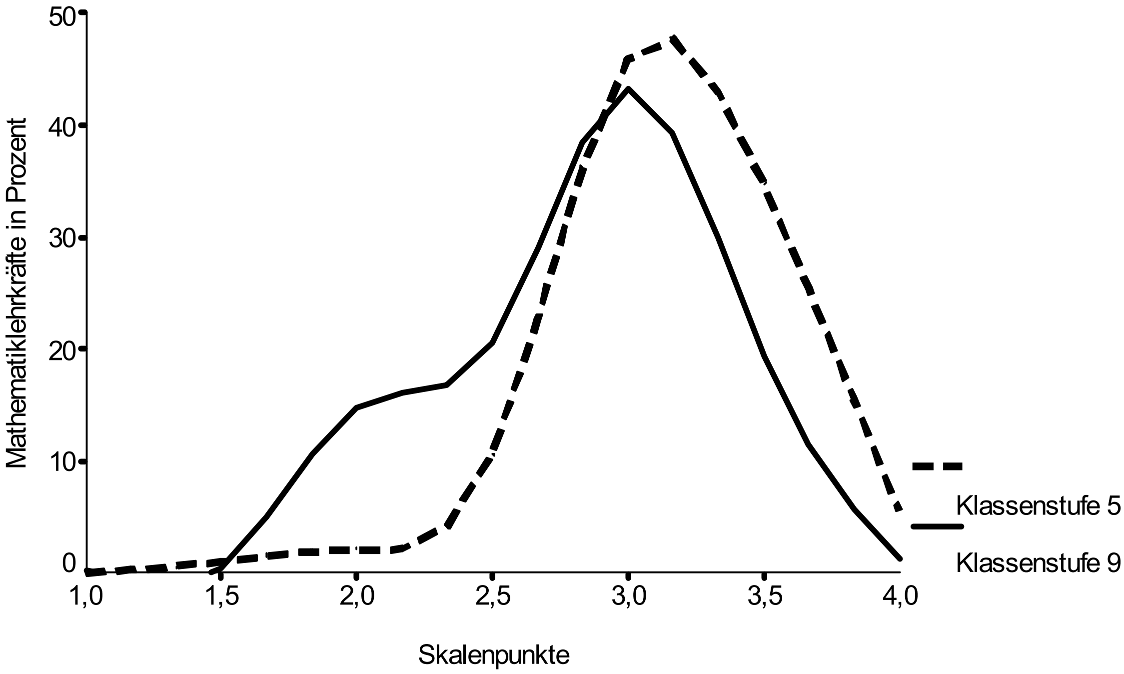
Auffällig ist wieder der schon mehrfach beobachtete Unterschied zwischen den Grundschulen und den Schulen mit Klassen im Sekundarbereich, wobei hier aber nicht nur in den Gymnasien, sondern auch in den Erweiterungskursen der Gesamtschulen das günstige Ausgangsniveau der Grundschulen aufrechterhalten ist.

Dies entspricht auch den bemerkenswert positiven Angaben der Schülerinnen und Schüler aus Erweiterungskursen und aus Gymnasien im Hinblick auf die eigenen schul- und unterrichtsbezogenen Einstellungen (vgl. oben, Abschnitt 4.4) Gleichzeitig liegt ein Vergleich mit den Auskünften der Schulleitungen zur „Beeinträchtigung des Unterrichts durch abweichendes Schülerverhalten“ nahe (vgl. oben, Tabelle 32). Berücksichtigt man die unterschiedliche Polung der beiden Skalen, so sind die Lehrerauskünfte über das Lern- und Arbeitsklima bzw. seine Beeinträchtigungen durch abweichendes Schülerverhalten durchgehend positiver als die der Schulleitungen, und zwar am stärksten an den Gymnasien und am wenigsten an den Realschulen.

Interessant ist hier ein grafischer Vergleich der Verteilungen der Lehrerantworten zwischen den beiden Klassenstufen (vgl. Abbildung 27). Aus der Abbildung wird nicht nur die bereits numerisch belegte allgemeine Verschlechterung des Lern- und Arbeitsklimas sichtbar, die sich z. B. auch in den Schülerangaben zur Schulzufriedenheit spiegelt, sondern in den Antworten der Lehrkräfte, die in den neunten Klassen unterrichten, erscheint eine Häufung ausgesprochen negativer Einschätzungen unterhalb des theoretischen Skalenmittels.

Zusatzanalysen haben ergeben, dass diese Erscheinung nur insoweit mit den Schulformen bzw. Kursniveaus zusammenhängt, als sie in den Erweiterungskursen der Gesamtschulen und an den Gymnasien merklich seltener ist; mit dem Alter, Geschlecht oder der Ausbildung der Lehrkräfte zeigen sich keine systematischen Zusammenhänge. Zu vermuten ist, dass die – verglichen mit den Grundschulen – offenkundig schwierigeren Arbeitsbedingungen an den weiterführenden Schulen einen immerhin auffälligen Teil der Lehrerschaft hoch belasten.

Abbildung 27 Verteilung der Werte auf der Skala „Arbeits- und Lernklima in der Klasse bzw. im Kurs“, nach Klassenstufe



Von Interesse ist die Frage nach dem Zusammenhang zwischen der Einschätzung des Arbeits- und Lernklimas durch die Lehrkräfte und der erreichten Testleistung der Schülerinnen und Schüler im QuaSUM-Mathematiktest. In der Klassenstufe 5 ist dieser Zusammenhang schwach ausgeprägt ($r = 0{,}10$), ebenso wie in drei der vier Teilgruppen der Klassenstufe 9. Nur an den Realschulen ist eine Beziehung, der zufolge in Klassen mit günstigerem Lern- und Arbeitsklima mehr gelernt wird, auch statistisch abzusichern ($r = 0{,}17$).

Die Skala „Bedeutung der Kreativität und Selbständigkeit von Schülerinnen und Schülern“ erfasst ein übergeordnetes Unterrichts- bzw. Erziehungsziel der Mathematiklehrkräfte. Sie bezieht sich auf insgesamt 7 Zielformulierungen wie z. B.: „kreatives Denken der Schüler fördern“, „Eigen- und Selbständigkeit fördern“ und „kritisches Denken bei den Schülern fördern“. Aus Tabelle 37 lässt sich ablesen, dass die Mehrheit der befragten Mathematiklehrkräfte sowohl in der Klassenstufe 5 als auch in der Klassenstufe 9 diesem übergeordneten Erziehungsziel große Bedeutung beimisst: Für beide Klassenstufen und für die unterschiedlichen Schul- bzw. Kursformen liegen die Mittelwerte im hoch zustimmenden Bereich zwischen 3,49 und 3,74.

Tabelle 37 Statistische Kennwerte der Skala „Betonung der Kreativität und Selbständigkeit von Schülerinnen und Schülern“

	Mittelwert*	Standardabweichung
Klassenstufe 5	3,74	0,33
Klassenstufe 9		
Gesamtschulen, Grundkurse	3,49	0,37
Gesamtschulen, Erweiterungskurse	3,50	0,39
Realschulen	3,53	0,45
Gymnasien	3,65	0,36
insgesamt	*3,54*	*0,40*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Die befragten Mathematiklehrkräfte in der Klassenstufe 5 weisen der Förderung der Kreativität und Selbständigkeit eine noch höhere Bedeutung zu als die in neunten Klassen unterrichtenden Kolleginnen und Kollegen. Betrachtet man zusätzlich die Schulformunterschiede innerhalb der Klassenstufe 9, so ergibt sich erneut der Befund, dass die Gymnasien an die optimistischen Einschätzungen der Grundschulen anzuknüpfen vermögen, während die anderen Schulformen stärkere Korrekturen gegenüber den Idealen vornehmen. Wie eng aber gerade diese Einstellungsdimension mit Idealvorstellungen verknüpft ist, zeigt die im Grunde einhellige Zustimmung zu den entsprechenden Items. Angesichts dessen konnten hohe Korrelationen mit der Testleistung von vornherein nicht erwartet werden; einzig für die Realschulen lässt sich mit einer Korrelation von $r = 0{,}13$ zeigen, dass Lehrkräfte, die den Aspekten der Selbständigkeit und Kreativität ihrer Schülerinnen und Schüler mehr Gewicht geben, in etwas leistungsfähigeren Klassen unterrichten. Dabei muss offen bleiben, ob sie damit auf günstigere Eingangsvoraussetzungen reagieren oder ob ein entsprechender Unterricht günstigere Ergebnisse zeitigt.

Eine ganze Reihe von Items im Lehrerfragebogen bezogen sich auf das methodische Konzept, das auf eine eng *kontrollierende Unterrichtsführung* setzt. Daraus konnte eine Skala gebildet werden, die eine solche Vorgehensweise der Stofferarbeitung und Stofffestigung beschreibt. Die entsprechenden Skalenwerte beruhen auf insgesamt 17 Aussagen wie z. B.: „Ich lasse die Schüler die Abfolge der Lösungsschritte genau im Heft notieren“, „Ich lasse Regeln und Merksätze wiederholen, damit die Schüler Sicherheit gewinnen“ und „Ich lege Wert darauf, dass es im Unterricht absolut ruhig ist“. Hohe Werte auf diesem Konstrukt charakterisieren demnach eine Form der Instruktion, bei der die Lehrkraft das Vorgehen eng und kleinschrittig lenkt und auf sich selbst zentriert. Die Schülerrolle erscheint in einer so gesteuerten Unterrichtssituation als rezeptiv und reaktiv[47]. Die Verbreitung einer solchen Einstellung der befragten Mathematiklehrkräfte ist aus der folgenden Tabelle 38 abzulesen.

Tabelle 38 Statistische Kennwerte der Skala „Kontrollierende Unterrichtsführung"

	Mittelwert*	Standardabweichung
Klassenstufe 5	3,09	0,27
Klassenstufe 9		
Gesamtschulen, Grundkurse	3,16	0,27
Gesamtschulen, Erweiterungskurse	3,04	0,26
Realschulen	3,14	0,23
Gymnasien	3,01	0,29
insgesamt	*3,09*	*0,27*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Trotz des oben dargestellten (vgl. Tabelle 37) Bekenntnisses zur Förderung der Eigenständigkeit der Lernenden praktiziert – nach eigener Auskunft – offenbar der überwiegende Teil der Brandenburger Mathematiklehrkräfte eine eher kontrollierende Unterrichtsführung. Auffällig sind das Fehlen eines Unterschieds zwischen den beiden Klassenstufen und die durchweg geringen Streuungen. Zwischen den Schulformen bzw. Kursniveaus sind die Differenzen mit bis zu einer halben Standardabweichung zwar merklich, und sie liegen mit einem etwas freizügigeren Unterrichtsstil in den Erweiterungskursen der Gesamtschulen und in den Gymnasien auch in der zu erwartenden Richtung. In der Substanz aber wird im Mathematikunterricht in Brandenburg „meistens" so vorgegangen.

Hinsichtlich eines Zusammenhanges zwischen der erreichten Testleistung der Schülerinnen und Schüler im QuaSUM-Mathematiktest und der Häufigkeit der Anwendung dieser Unterrichtsform zeigt sich in der Klassenstufe 5 ein geringfügig positiver Zusammenhang (r = 0,11): Je häufiger die Lehrkräfte diese Unterrichtsform anwenden, desto besser ist die erreichte Testleistung der Schülerinnen und Schüler im QuaSUM-Mathematiktest.

In der einschlägigen Literatur, vor allem in der Tradition der Münchner SCHOLASTIK-Studie (vgl. WEINERT & HELMKE 1995), wird häufig ange-

[47] „Diese Form ist für das Erarbeiten mathematischer Erkenntnisse und Zusammenhänge weniger geeignet, weil durch die dominierende Stellung des Lehrers und durch seine Strenge, auf kleinste Schritte beschränkte Unterrichtsführung den Schülern wenig Raum gelassen wird zur Entfaltung, zum Nachdenken und zum Entwickeln eigener Lösungsideen" (NEUNZIG 1995, S. 275).

nommen, dass eine eng geführte „direkte Instruktion“ vor allem für lernschwache Schülerinnen und Schüler günstiger sei als offenere Unterrichtsformen. Wenn man die Ergebnisse im CFT 20 zur Überprüfung dieser Hypothese heranzieht und die Schülerinnen und Schüler danach in vier Gruppen (sog. „Quartile“) einteilt, so bestätigen die vorliegenden Daten diesen Befund nicht: In etwa gleicher Weise lassen sich in allen vier Gruppen leichte Vorteile der kontrollierenden Unterrichtsform zeigen (ohne Tabelle).

In der Klassenstufe 9 lässt sich ein analoger Vorteil für die direkte Instruktion nur für die Schülerinnen und Schüler an den Realschulen feststellen (r = 0,11), und zwar wiederum in allen „Fähigkeitsgruppen“ nach Maßgabe des CFT 20. Während an den Gesamtschulen die Tendenzen zumeist schwach und ausgesprochen uneinheitlich sind, geht an den Gymnasien eine eng kontrollierende Unterrichtsführung gerade bei den Schülerinnen und Schülern, die über eher ungünstige kognitive Eingangsvoraussetzungen verfügen, mit niedrigeren Lernerfolgen einher.

Die *Bedeutung des Lehrbuchs für den Unterricht* konnte mit einer Kurzskala erfasst werden, die sich auf drei Aussagen stützt („Ich verwende das Lehrbuch als methodischen Leitfaden“, „Ich verwende das Lehrbuch zur Strukturierung des Stoffes“ und „Ich gehe genau nach Lehrbuch vor“). Wie die nachstehende Tabelle 39 zeigt, besitzt das Lehrbuch in Brandenburg einen wichtigen Stellenwert.

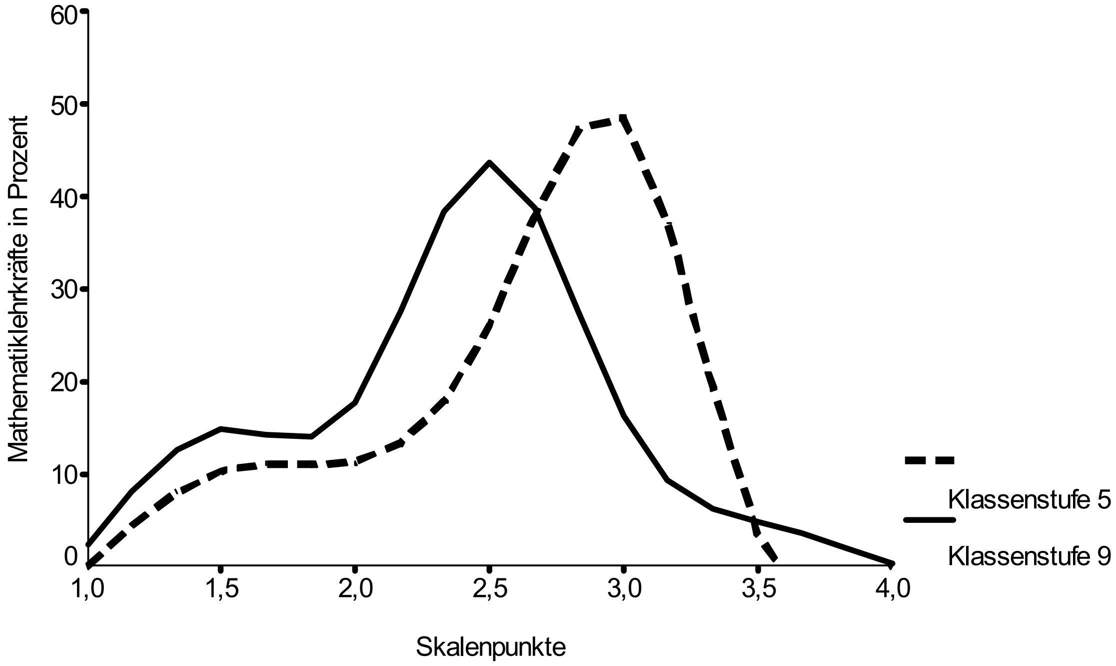
Tabelle 39 Statistische Kennwerte der Skala „Bedeutung des Lehrbuchs für den Unterricht“

	Mittelwert*	Standardabweichung
Klassenstufe 5	2,64	0,51
Klassenstufe 9		
Gesamtschulen, Grundkurse	2,21	0,56
Gesamtschulen, Erweiterungskurse	2,40	0,53
Realschulen	2,43	0,49
Gymnasien	2,44	0,53
insgesamt	*2,35*	*0,54*

* Ausprägungen: 1 = sehr gering bis 4 = sehr hoch

Nach diesen Befragungsergebnissen nimmt die Bedeutung des Lehrbuchs von Klassenstufe 5 nach Klassenstufe 9 ab, auch wenn sich alle Werte nahe am theoretischen Skalenmittelwert befinden. Gemessen an der Vorliebe der Lehrkräfte für eher kleinschrittiges, systematisches Vorgehen hätte man vielleicht eine vergleichsweise enge Orientierung an den Lehrbüchern erwartet. Tatsächlich aber werden offenbar häufig auch andere Quellen für die methodische Planung des Unterrichts genutzt. Zudem gibt es auf beiden Klassenstufen eine beträchtliche Gruppe von Lehrerkräften, die mit dem Lehrbuch ausgesprochen „selten“ arbeiten, wie die nachstehende Abbildung zeigt.

Abbildung 28 Verteilung der Werte auf der Skala „Bedeutung des Lehrbuchs für den Unterricht“, nach Klassenstufe



Interessant ist nun die Frage, was über diese Gruppe von Lehrkräften aus den Daten zu erfahren ist. Bei den Grundschullehrkräften, die (fast) ohne Lehrbuch auskommen, sind eher schwächere Ergebnisse im QuaSUM-Mathematiktest zu verzeichnen (d = -0,12). In allen Schulformen bzw. Kursniveaus der Sekundarstufe I geht dagegen der (weitgehende) Verzicht auf ein Lehrbuch mit eher *besseren* Ergebnissen einher; die Effektstärken liegen hier zwischen d = 0,08 am Gymnasium und d = 0,23 in den Erweiterungskursen der Gesamtschulen. Es bleibt also letztlich unklar, welche Auswirkungen die Arbeit mit „alternativen“ Grundlagen für die Unterrichtsplanung hat. Die Zusammenhänge mit den unterrichtsbezogenen Einstellungen und methodischen Präferenzen der Lehrkräfte sind indessen uneinheitlich zwischen den verschiedenen befragten Gruppen, so dass vorerst ebenso offen bleiben muss, welche Motivlagen hinter der Entscheidung für oder gegen eine Lösung vom Lehrbuch als dem traditionellen didaktischen Leitmedium stehen.

Schließlich seien kurz die Angaben der Lehrkräfte zu den *Hausaufgaben* zusammengefasst und zu den übrigen Daten in Beziehung gesetzt. Wie für die Schülerinnen und Schüler, so sind auch für die Lehrkräfte Mathematikhausaufgaben fester Bestandteil des Unterrichts: In beiden Klassenstufen geben die befragten Mathematiklehrkräfte regelmäßig Hausaufgaben auf.

Tabelle 40 Zeitaufwand für die Erledigung der Hausaufgaben im Fach Mathematik (in Minuten pro Woche)

	Lehrerangaben		**Schülerangaben**	
	Mittelwert	SD*	Mittelwert	SD
Klassenstufe 5	57,9	30,9	111,1	97,4
Klassenstufe 9				
Gesamtschulen, Grundkurse	40,9	23,7	74,3	87,1
Gesamtschulen, Erweiterungskurse	46,1	24,5	85,7	92,7
Realschulen	54,1	24,8	87,6	88,4
Gymnasien	54,1	33,2	98,3	88,9
insgesamt	*48,1*	*27,7*	*86,8*	*89,8*
insgesamt		*2,84*		*0,49*

* SD = Standardabweichung

An dieser Zusammenstellung ist zweierlei sehr auffällig: Erstens schätzen die Lehrkräfte – wie die Schülerinnen und Schüler – die Belastung durch Mathematik-Hausaufgaben in der Klassenstufe 9 gegenüber der Klassenstufe 5 geringer ein, und zwar um etwa 20 Prozent. Zweitens unterschätzen sie in allen Teilgruppen den tatsächlichen Aufwand für die Hausaufgaben erheblich. An den Grundschulen arbeiten die Schülerinnen und Schüler (jedenfalls nach eigenen Angaben) nahezu doppelt so lange an den Aufgaben, wie ihre Lehrkräfte annehmen, an den Realschulen – wo die Unterschiede in der Einschätzung am geringsten sind – immer noch um mehr als 60 Prozent länger. Dass dabei die Schülerauskünfte stark streuen, war zu erwarten, weil der Aufwand für die Hausaufgaben innerhalb einer Schulklasse je nach Leistungsniveau und Arbeitsgeschwindigkeit sehr unterschiedlich sein kann. Beunruhigend erscheint, dass die Schülerangaben praktisch nichts mit den Lehrerauskünften gemein haben. Auf dieser Grundlage sind konsistente Zusammenhänge zwischen dem vermuteten Aufwand für die Hausaufgaben und der Fachleistung nicht zu erwarten.

Weitere Daten aus der Befragung der Mathematiklehrkräfte wie z. B. Angaben über die Grundlagen der Unterrichtsvorbereitung und die im Unterricht praktizierten Sozialformen stehen in keiner systematischer Beziehung zu den erhobenen Mathematikleistungen der Schülerinnen und Schüler.

5.3 Diagnosepraxis der Lehrkräfte – Zusammenhänge zwischen Noten bzw. Leistungseinschätzungen und Testleistungen

5.3.1 Testleistung und Noten

Die aus den Schülerakten entnommenen Angaben enthalten u. a. auch die Noten der letzten drei Zeugnisse. Im Folgenden sollen die Zusammenhänge zwischen den Mathematiknoten im Halbjahreszeugnis 1998/99 und den Ergebnissen im QuaSUM-Mathematiktest bzw. im Mathe40-Test untersucht werden.

Klassenstufe 5

Die durchschnittlichen Zensuren im Halbjahreszeugnis 1998/99 betragen für das Fach Mathematik 2,5, für das Fach Deutsch 2,4 und für die erste Fremdsprache 2,3. In der Klassenstufe 5 wird also im Fach Mathematik ein wenig „strenger" zensiert als im Fach Deutsch und in der ersten Fremdsprache[48]. Berechnet man die Korrelation zwischen den Mathematiknoten im Schulhalbjahreszeugnis 1998/99 und den Leistungen im QuaSUM-Mathematiktest, so ergibt sich über alle Schülerinnen und Schüler gerechnet ein Gesamtwert von $r = -0{,}62$[49]. Dabei ist zu beachten, dass es innerhalb der einzelnen Klassen große Unterschiede gibt. Die *innerhalb der Klassen* berechneten Korrelationen reichen von 0,27 bis -0,92. Demnach gibt es Klassen, in denen recht leistungsfähige Schüler schlechte Noten oder weniger leistungsfähige Schüler gute Noten erhalten haben. Die folgende Abbildung 29 stellt den Zusammenhang zwischen den erreichten Leistungen im QuaSUM-Mathematiktest (ausgedrückt als Raschwert) und den gegebenen Mathematiknoten dar.

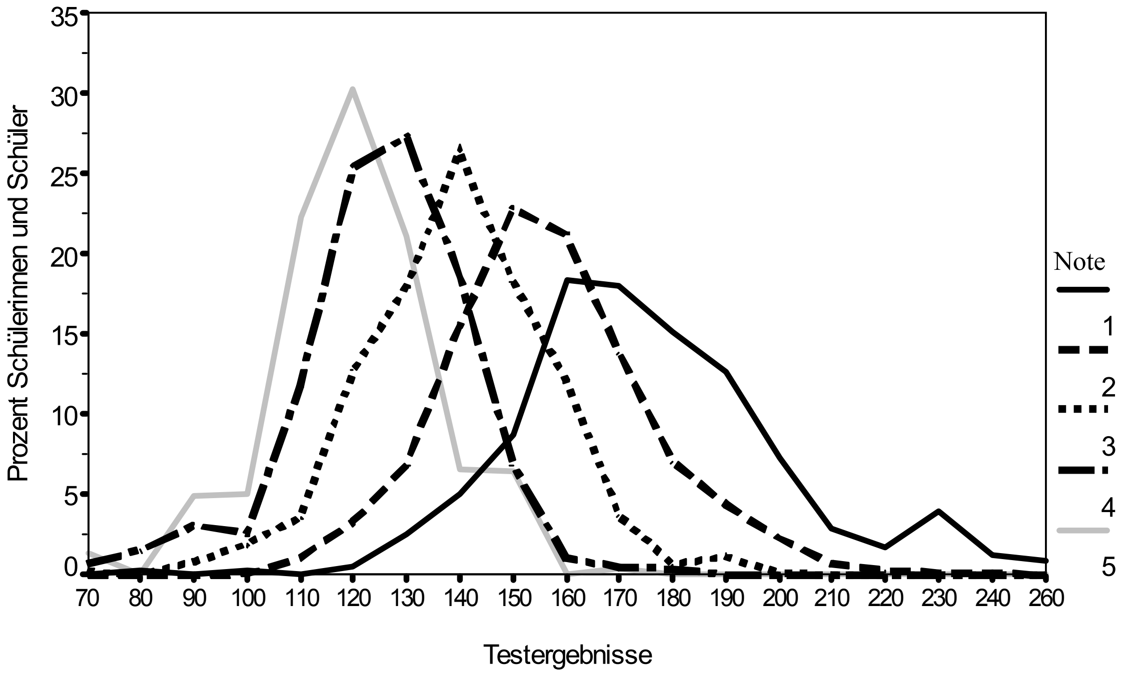
Wie ersichtlich, gibt es breite Überschneidungszonen zwischen den einzelnen Noten. So kann ein Schüler oder eine Schülerin bei einem (dem Durchschnitt entsprechenden) Testergebnis von 150 Punkten je nach örtlichem Standard alle Noten von 1 bis 5 erhalten haben. Es gilt lediglich die in der genannten Korrelation von $r = -0{,}62$ ausgedrückte allgemeine Tendenz, dass in der Regel höhere Testleistungen auch mit besseren Noten einhergehen.

---

48 Die Mittelwerte betragen am Ende des Schuljahres 1997/98, also am Ende der vierten Klasse, für Mathematik 2,3 und für Deutsch 2,5, die Mittelwerte am Ende des Schuljahres 1996/97, also am Ende der dritten Klasse, soweit vorhanden, für Mathematik 2,2 und für Deutsch 2,2.

49 Das negative Vorzeichen der Korrelation bedeutet, dass hohe Testleistungen mit ziffernmäßig niedrigen, also guten Noten einher gehen.

Abbildung 29 Notenverteilung nach durchschnittlichem Testergebnis (Raschwert) in der Klassenstufe 5 [50]



Man kann dies besonders gut erkennen, wenn man die Prozentzahlen in Wahrscheinlichkeiten umrechnet. So ist bei der mittleren Testleistung von 150 Punkten die Wahrscheinlichkeit am größten, die Note 2 im Zeugnis zu haben. Wichtig sind deshalb die Schnittpunkte der Kurven, aus deren Lage gegenüber der verwendeten Raschskala typische Notenbereiche ermittelt werden können; ein durchschnittlich erreichter Raschwert von 135 bis 147 entspricht z. B. der Note 3. Die Leistungsbereiche für die einzelnen Noten sind der folgenden Tabelle 41 zu entnehmen.

Tabelle 41 Typische Leistungsbereiche nach Noten und Geschlecht, Klassenstufe 5 (Halbjahreszeugnis 1998/99)

	Leistungsbereich (Raschwerte im QuaSUM-Mathematiktest)		
Note	Jungen	Mädchen	*insgesamt*
1	über 172	über 160	über 164
2	150 – 172	146 – 160	148 – 164
3	138 – 149	130 – 145	136 – 147
4	123 – 137	127 – 129	123 – 135
5	unter 123	unter 127	unter 123

[50] Da die Note 6 in der Jahrgangsstufe 5 so gut wie nicht vorkommt, wurde sie in der Abbildung der Übersichtlichkeit wegen weggelassen.

Im Zusammenhang mit den großen Überschneidungsbereichen stellt sich die Frage, welche Faktoren außer der Fachleistung die Mathematiknote bestimmen. Regressionsanalytisch ergeben sich für die Klassenstufe 5 die folgenden Einflussfaktoren (in der angegebenen Reihenfolge, ausgedrückt durch die sog. „Beta-Gewichte"), wobei diese 9 Merkmale zusammen 53 Prozent der Schülervarianz in den Mathematiknoten in der Klassenstufe 5 erklären:

- die Testleistung im QuaSUM-Mathematiktest (β = -0,46),
- das allgemeine Leistungsniveau der Klasse (β = 0,16),
- ein hohes Sach- und Fachinteresse der Schülerinnen und Schüler an Mathematik (β = -0,16),
- die Fähigkeit zum schlussfolgernden Denken (β = -0,14),
- das Geschlecht (weiblich) (β = -0,14),
- der Bildungsstand der Eltern (β = -0,12),
- das positive leistungsbezogene Selbstkonzept der Schülerinnen und Schüler (β = -0,12),
- die Wahrnehmung des Mathematikunterrichts als eines straff gelenkten, wohlstrukturierten Geschehens (β = 0,06),
- die positive Einschätzung der eigenen Anstrengung als Beitrag zum Schulerfolg (β = -0,05).

Negative Koeffizienten bedeuten positive Einflüsse auf die Mathematiknoten, da hohe Notenwerte schwachen Leistungen entsprechen. Demnach ist eine hohe Leistung im QuaSUM-Mathematiktest der mit Abstand wichtigste Prädiktor für gute Noten. Dagegen bedeutet das positive Vorzeichen für das Leistungsniveau der Klasse, dass es – unter Berücksichtigung aller übrigen Einflüsse – in leistungsstarken Klassen schwerer ist, eine gute Mathematiknote zu erzielen als in leistungsschwachen Klassen. Dieser Befund deckt sich mit Ergebnissen aus der Hamburger Lernausgangslagen-Untersuchung (vgl. LEHMANN & PEEK 1997). Darüber hinaus werden ceteris paribus bei einem hohen Sach- und Fachinteresse an Mathematik bessere Noten erzielt. Das Gleiche gilt für das leistungsbezogene Selbstkonzept, die positive Einschätzung der eigenen Anstrengung als Beitrag zum Schulerfolg und die Fähigkeit zum schlussfolgernden Denken. Gleichzeitig wird unter sonst gleichen Voraussetzungen in einem transparenten, straff geführten und wohlstrukturierten Mathematikunterricht in der Regel auch strenger zensiert. Ein weiteres Ergebnis ist es, dass Kinder aus bildungsnahen Elternhäusern unabhängig von ihrer Testleistung und den übrigen berücksichtigten Faktoren bessere Zensuren erhalten; auch dieser Befund hatte sich bereits in der Hamburger Untersuchung gezeigt.

Das negative Vorzeichen bei der Variable „Geschlecht“ zeigt an, dass Mädchen unter sonst gleichen Voraussetzungen bessere Noten bekommen als Jungen. Geschlechtsunterschiede werden auch durch einen anderen Vergleich deutlich: Obwohl die Jungen im Halbjahreszeugnis 1998/99 tendenziell schlechtere Mathematiknoten als die Mädchen haben, ist die erreichte Leistung der Jungen im QuaSUM-Mathematiktest höher als die der Mädchen. Die Auflistung der typischen Notenbereiche nach Geschlecht (vgl. dazu Tabelle 41) veranschaulicht diesen Befund. Die Jungen müssen vor allem im oberen Bereich höhere Testleistungen als die Mädchen erbringen, um die gleiche sehr gute oder gute Note zu erhalten. Selbst im „befriedigenden“ Bereich sind noch deutliche Vorteile für die Mädchen zu erkennen. Andererseits ist auffällig, wie schmal für die Mädchen das Segment für die Note „ausreichend“ ist: Die Schwelle zwischen den Noten 5 und 4 liegt für die Mädchen eindeutig höher[51].

Klassenstufe 9

Wenn man die Durchschnittsnoten im Fach Mathematik in der Klassenstufe 9 zwischen den Schulformen bzw. Kursniveaus miteinander vergleicht, so zeigen sich nur geringe Unterschiede zwischen den Erweiterungskursen an Gesamtschulen (Mittelwert 3,0), den Realschulen (Mittelwert 3,1) und den Gymnasien (Mittelwert 2,8). Somit hat, in Anbetracht der wesentlich höheren Fachleistungen in dieser Schulform, das Gymnasium offenbar seine eigenen, von den beiden anderen Bezugsgruppen verschiedenen Bewertungsstandards. Bei den Grundkursen der Gesamtschulen (Mittelwert 4,0) ist zunächst zu vermuten, dass – schulformbedingt – hier eine gewisse Verschränkung mit den Bewertungsstandards in den Erweiterungskursen gegeben ist.

Die folgende Tabelle 42 gibt für die untersuchte Stichprobe der Neuntklässlerinnen und Neuntklässler aus Gesamtschulen, Realschulen und Gymnasien für die Schuljahreszeugnisse in der 7. (1996/97) und 8. (1997/98) sowie für das Halbjahreszeugnis in Klassenstufe 9 (1998/99) über die durchschnittlich erzielten Mathematiknoten hinaus die Noten in Deutsch und in der ersten Fremdsprache an.

[51] Diese Befunde sagen noch nichts darüber aus, warum es zu einer derart unterschiedlichen Behandlung der Geschlechter kommt, zumal wichtige Einstellungsunterschiede zwischen den Geschlechtern in dem hier verwendeten Regressionsmodell bereits verrechnet worden sind.

Tabelle 42 Verteilung der Noten in den Fächern Deutsch, Mathematik und erste Fremdsprache in den Klassenstufen 7, 8 und 9, nach Schulform bzw. Kursniveau

	Gesamtschulen G-Kurse	**Gesamtschulen E-Kurse**	**Realschulen**	**Gymnasien**
Halbjahreszeugnis 1998/99				
Deutsch	3,7	2,9	2,9	2,6
Mathematik	4,0	3,0	3,1	2,8
erste Fremdsprache	3,8	3,1	3,1	2,8
Schuljahreszeugnis 1997/98				
Deutsch	3,6	2,8	2,9	2,4
Mathematik	3,9	2,8	3,1	2,7
erste Fremdsprache	3,7	2,9	3,0	2,7
Schuljahreszeugnis 1996/97				
Deutsch	3,5	2,6	2,8	2,4
Mathematik	3,7	2,6	2,8	2,6
erste Fremdsprache	3,6	2,8	2,9	2,6

Auffällig sind zwei Aspekte. Zum Einen fällt auf, dass sich in allen vier Teilgruppen die Noten von der Klassenstufe 7 zur Klassenstufe 9 kontinuierlich verschlechtern., und zwar jeweils um durchschnittlich eine Zehntel Note. Ob dies als eine Verschlechterung der Schulleistungen oder eine Verschärfung der Notenvergabe zu werten ist, muss offen bleiben. Zum Zweiten finden sich, ebenfalls für fast alle Teilgruppen und Klassenstufen, in Deutsch die relativ besten und in Mathematik die relativ schlechtesten Noten, während die Noten für die erste Fremdsprache dazwischen liegen.

Im Folgenden werden die Beziehungen zwischen Testleistung und Benotungspraxis näher untersucht. Ähnlich wie in der Klassenstufe 5 erreicht die Korrelation zwischen den Mathematiknoten und den Leistungen im QuaSUM-Mathematiktest über alle Schülerinnen und Schüler (also auch alle Schulformen) hinweg einen Gesamtwert von r = -0,57. Die Korrelation zwischen den Mathematiknoten und den Leistungen im Mathe40-Test liegt mit r = -0,46 niedriger, worin sich erneut die engere Curriculumnähe des QuaSUM-Tests ausdrückt.

Die einzelnen Schulformen unterscheiden sich hinsichtlich des Verhältnisses von Testleistung und Benotung deutlich voneinander (vgl. Tabelle 43).

Tabelle 43 Korrelationen zwischen Testleistungen und Mathematiknoten in Klassenstufe 9, nach Schulform bzw. Kursniveau

	QuaSUM-Mathematiktest	***Korrelationen reichen*** *von*	*bis*	**Mathe40-Test**
Gesamtschulen, Grundkurse	-0,30	*0,58*	*-0,80*	-0,15
Gesamtschulen, Erweiterungskurse	-0,40	*0,10*	*-0,88*	-0,25
Realschulen	-0,57	*0,14*	*-0,90*	-0,42
Gymnasien	-0,55	*-0,02*	*-0,88*	-0,43

Aus diesen Zahlen kann man zunächst schließen, dass für alle Gruppen gleichermaßen der QuaSUM-Test – verglichen mit dem Mathe40-Test – als unterrichtsnäher zu gelten hat. Weiterhin ist auffällig, dass an den Gesamtschulen – und hier besonders in den Grundkursen – die Beziehungen zwischen Testleistung und Note deutlich gelockert sind.

Untersucht man auf der Grundlage des QuaSUM-Tests auch für die Klassenstufe 9 die Standards, die für die einzelnen Notenstufen gelten, so zeigen sich für alle vier Schulformen ähnlich große Überlappungsbereiche bei den Noten wie in der Klassenstufe 5 (ohne Abbildung). Vor allem aber zeigt ein Vergleich der typischen Leistungsbereiche nach Notenstufen – wiederum ermittelt über die Schnittpunkte der Kurven für die Notenwahrscheinlichkeiten – erwartungsgemäß, dass man in den unterschiedlichen Schulformen unterschiedliche Leistungsniveaus erreichen muss, um eine bestimmte Note zu erhalten. Die folgende Tabelle 44 zeigt diese Leistungsbereiche getrennt nach Schulformen bzw. Kursniveaus.

Tabelle 44 Typische Leistungsbereiche nach Noten in der Klassenstufe 9, nach Schulform bzw. Kursniveau

	Leistungsbereich (Raschwerte im QuaSUM-Mathematiktest)			
Note	Gesamtschulen Grundkurse	Gesamtschulen Erweiterungskurse	Realschulen	Gymnasien
1	--	über 166	über 172	über 192
2	über 144	157 – 166	160 – 172	179 – 192
3	127 – 144	152 – 156	146 – 159	169 – 178
4	119 – 126	142 – 151	137 – 145	165 – 168
5	71 – 118	unter 142	132 – 136	unter 165
6	unter 71	--	unter 132	--

Erwartungsgemäß wird deutlich, dass die Note 3 im Gymnasium leistungsmäßig nicht das Gleiche bedeutet wie die Note 3 in den anderen Schulformen bzw. Kursniveaus. Mit einem Punktwert von 135 bekommt man in den Grundkursen der Gesamtschulen durchschnittlich eine „3", in den Erweiterungskursen und an der Realschule dagegen wahrscheinlich eine „5", während im Gymnasium mit einem solchen Niveau im Grunde nicht mehr zu rechnen ist.

Wegen der denkbaren Übergänge zwischen den Kursniveaus in den Gesamtschulen und eventuell auch den Schulformen sind die Äquivalenzrelationen von Bedeutung. Zwischen Grund- und Erweiterungskursen unterscheiden sich die Noten bei gleicher Mathematikleistung an den Brandenburger Gesamtschulen de facto um etwa 2 Stufen; der Abstand zwischen der unteren Grenze für eine „ausreichende" Zensur und der oberen Grenze für die Note „gut" ist etwa gleich (Grundkurse: 26 Punkte, Erweiterungskurse: 24 Punkte; zum Vergleich Gymnasien: 27 Punkte). Die Benotungsskala der Realschulen ist demgegenüber mit 36 Punkten zwischen den genannten Grenzen merklich breiter angelegt, so dass es hier keine gleichmäßige Äquivalenzrelation gegenüber den Zensuren der anderen Schulformen geben kann.

Testleistungen, die eine mindestens „ausreichende" Zensur im Erweiterungskurs nahe legen, haben 12,6 Prozent der Schülerinnen und Schüler aus Grundkursen erbracht. In den Erweiterungskursen hingegen befinden sich 28,3 Prozent, die im QuaSUM-Mathematiktest kein für dieses Kursniveau „ausreichendes" Leistungsniveau nachgewiesen haben. Auf mindestens „ausreichendem" gymnasialem Niveau befinden sich in den Erweiterungskursen 19,2 Prozent der Schülerinnen und Schüler, an den Realschulen 20,1 Prozent. Unter den Gymnasiasten wiederum haben 26,9 Prozent die Schwelle unterschritten, unterhalb derer nach Brandenburger Standards die Note „mangelhaft" vergeben werden müsste.

Die Suche nach Bedingungsfaktoren für die Mathematiknote in der Klassenstufe 9 erbringt nur teilweise ein ähnliches Bild wie in der Klassenstufe 5 (vgl. Tabelle 45). Die gleichen 9 Merkmale wie in Klassenstufe 5 erklären in der Klassenstufe 9 je nach Schulform zwischen 17,4 und 44,2 Prozent der Unterschiede in den Mathematiknoten. Die verschiedenen Selektionsformen zwischen den Schulformen und auch innerhalb der Gesamtschulen relativieren also nicht nur, was durchaus intendiert ist, die Bedeutung der Noten über die Schulformen hinweg, sondern sie gehen auch einher mit einer erhöhten Unsicherheit *innerhalb* der Schulformen. Bis zu einem gewissen Grade ist dies verständlich, weil anders als in den Grundschulen in der Sekundarstufe I jeweils nur noch ein Teil des gesamten Leistungsspektrums den Bezugsrahmen für die Benotung liefert.

Tabelle 45 Einfluss von Faktoren auf die Mathematiknote in Klassenstufe 9, nach Schulform bzw. Kursniveau (Beta-Gewichte)

	Gesamtschulen		**Realschulen**	**Gymnasien**
	G-Kurse	**E-Kurse**		
Testleistung im QuaSUM-Mathematiktest	-0,20	-0,36	-0,46	-0,47
allgemeines Leistungsniveau der Klasse/des Kurses	0,04	0,10	0,12	0,16
Sach- und Fachinteresse an Mathematik	-0,28	-0,27	-0,35	-0,27
Fähigkeit zum schlussfolgernden Denken	-0,02	0,01	-0,05	-0,01
Geschlecht (weiblich)	-0,20	-0,19	-0,18	-0,17
Bildungsstand der Eltern	0,01	-0,03	-0,01	-0,05
leistungsbezogenes Selbstkonzept	-0,00	0,01	-0,04	-0,10
Wahrnehmung des Mathematikunterrichts als strukturiertes Geschehen	-0,05	0,07	0,05	0,03
positive Einschätzung der eigenen Anstrengung als Beitrag zum Schulerfolg	0,01	0,04	0,00	0,03
erklärte Varianz (in Prozent)	*17,4*	*23,3*	*44,2*	*40,9*

Angesichts dessen ist es auffällig, dass an den Realschulen und Gymnasien der Zusammenhang zwischen Fachleistung und Benotung gegenüber der Klassenstufe 5 gewahrt bleibt, während er an den Gesamtschulen – in den Grundkursen ausgeprägter als in den Erweiterungskursen – stark gelockert erscheint (vgl. auch oben, Tabelle 43). Einen Anhaltspunkt zur Erklärung dieser Unterschiede könnte man darin sehen, dass die Gewichte für die Fachleistung offenbar in Zusammenhang mit der Bedeutung des allgemeinen Klassenniveaus für die Notengebung stehen; dies könnte ein Indiz für die oben geäußerte Vermutung sein, dass vor allem in den Grundkursen das implizite Vergleichsprinzip im Bezugsrahmen der Lerngruppe, das eine fachleistungsbestimmte Notengebung zweifellos erleichtert, tendenziell vermieden wird.

Einfluss gewonnen hat dagegen generell das Sach- und Fachinteresse an Mathematik, was vielleicht – gleichsam im Umkehrschluss – auf die zunehmenden Motivationsprobleme in der Adoleszenz verweist, von denen bereits mehrfach die Rede war und denen die Lehrkräfte möglicherweise mit dem Instrument der Benotung zu begegnen suchen. Zugenommen haben ebenfalls in allen Schulformen die Vorteile der Mädchen bei der Benotung, wobei erneut darauf hingewiesen sei, dass dabei deren positivere Einstellungen zu Schule und Unterricht bereits berücksichtigt sind. Abgenommen haben dagegen der eigenständige Beitrag der Fähigkeit zum schlussfolgernden Denken und der eines positiven Selbstkonzepts (mit Ausnahme an den Gymnasien), vor allem auch die Bedeutung des Bildungsstandes der Eltern. Eine nennenswerte soziale Benachteiligung ist in der Benotungspraxis der Klassenstufe 9 in keiner der untersuchten Schulformen nachweisbar.

5.3.2 Testleistung und Leistungseinschätzung durch die Lehrkräfte

Wie sich gezeigt hat, sind die Mathematiknoten, die die Lehrkräfte vergeben, vor allem abhängig von der Fachleistung der Schülerinnen und Schüler, so wie sie in den standardisierten Mathematiktests sichtbar wird. Der Einfluss der Fachleistung auf die Zensuren variiert jedoch sehr stark, und nicht alle zusätzlichen „fachfremden“ Motive bei der Benotung sind notwendig illegitim. Prinzipiell ist es zwar möglich, den Zusammenhang zwischen Fachleistung und Note nicht nur innerhalb von Schulformen und Kursniveaus zu untersuchen, sondern auch auf Klassenebene zu betrachten. Wegen der (durch die Befunde zumindest nahegelegten) intervenierenden Einflüsse fachfremder Bewertungsnormen ist es jedoch nicht unproblematisch, den Grad, in dem das in der Note ausgedrückte Urteil einer Lehrkraft mit den Testergebnissen ihrer Klasse übereinstimmt, als Indikator für die pädagogisch-diagnostische Kompetenz zu interpretieren. Im QuaSUM-Projekt ist deshalb folgender Weg beschritten worden: Um personenbezogene Urteilskomponenten möglichst auszuschließen, wurde versucht, die diagnostische Urteilsfähigkeit der Lehrkräfte unter Bezug auf die Testaufgaben und die jeweilige Lerngruppe als Ganze zu erfassen.

Im Anschluss an die Tests und Befragungen in den Schulen wurden die Mathematiklehrkräfte gebeten, für ihre Klasse bzw. ihren Kurs einzuschätzen, wie viel Prozent der Schülerinnen und Schüler die einzelnen Aufgaben vermutlich richtig gelöst haben. Geht man davon aus, dass die Übereinstimmung zwischen tatsächlicher Testleistung der Schülerinnen und Schüler und der Leistungseinschätzung der Lehrkräfte etwas über die Fähigkeit der Lehrkräfte aussagt, ihre Schülerinnen und Schüler zu beurteilen, so kann man die Klassenbinnenkorrelation zwischen Testleistung und Leistungseinschätzung als Kennzahl für die diagnostische Kompetenz der Lehrkräfte betrachten. Denn in dieses Maß geht vor allem die *relative* Schwierigkeit der Testaufgaben für die Lerngruppe ein, die beurteilen zu können für die Planung und Durchführung von Unterricht unerlässlich erscheint. Man könnte deshalb erwarten, dass die Klassen von Lehrkräften, die über ein sicheres diagnostisches Urteilsvermögen verfügen, auch leistungsmäßig ein eher günstiges Bild bieten.

Für die Klassenstufe 5 beträgt die mittlere Klassenbinnenkorrelation zwischen Testleistung und Leistungseinschätzung $r = 0{,}36$, für die Klassenstufe 9 $r = 0{,}43$. Diese Werte sind nicht sehr hoch, insbesondere dann nicht, wenn man sich vergegenwärtigt, dass bei der Berechnung der Koeffizienten etwaige Neigungen zu einer allgemeinen Über- oder Unterschätzung der Klasse ohne Einfluss bleiben. Zu beachten ist allerdings, dass die Korrelationen für die einzelnen Klassen bzw. Kurse eine große Bandbreite aufweisen. In der Jahrgangsstufe 5 reichen die Werte von -0,09 bis 0,73, in der Jahrgangsstufe 9 von 0,01 bis 0,77. Die nachstehende Tabelle 46 enthält die ermittelten Klassenbinnenkorrelationen zwischen Testleistung und Lehrereinschätzung in der Klassenstufe 9, getrennt nach Schulformen.

Tabelle 46 Relative Übereinstimmung zwischen Leistungseinschätzung und Testleistung in Klassenstufe 9 (Klassenbinnenkorrelation), nach Schulform bzw. Kursniveau

	niedrigste Korrelation	**höchste Korrelation**	**durchschnittliche Korrelation**
Gesamtschulen, Grundkurse	0,04	0,60	0,33
Gesamtschulen, Erweiterungskurse	0,01	0,61	0,41
Realschulen	0,05	0,68	0,44
Gymnasien	0,21	0,77	0,55
insgesamt	*0,01*	*0,77*	*0,43*

Die Unterschiede zwischen den Schulformen bzw. Kursniveaus in den Korrelationen werden noch deutlicher, wenn man jeweils die besten 10 bzw. 20 Prozent und die schlechtesten 10 bzw. 20 Prozent der Übereinstimmungen zwischen Einschätzungen und Testleistungen betrachtet (vgl. Tabelle 47). Während z. B. nur 20 Prozent der Lehrkräfte in den Grundkursen eine Korrelation zwischen Leistungseinschätzung und Testleistung von $r = 0{,}44$ und darüber aufweisen, gibt es in den Gymnasien nur 10 Prozent der Lehrkräfte, die eine Korrelation von 0,41 und darunter haben. Demnach haben die Lehrkräfte an den Grundschulen, gefolgt von denjenigen, die in Grundkursen tätig sind, die größten Schwierigkeiten, das tatsächliche (relative) Anspruchsniveau der Testaufgaben für die eigene Klasse bzw. den eigenen Kurs einzuschätzen. Die geringsten Differenzen sind an den Gymnasien zu verzeichnen.

Tabelle 47 Relative Übereinstimmung (Klassenbinnenkorrelation) zwischen Testleistung und Leistungseinschätzung: 10., 20., 80. und 90. Perzentil, nach Klassenstufe und Schulform bzw. Kursniveau

		Klassenstufe 9			
	Klassen-stufe 5	**Gesamtschule G-Kurs**	**Gesamtschule E-Kurs**	**Realschule**	**Gymnasium**
10 Prozent haben eine Korrelation unter	0,10	0,17	0,25	0,23	0,41
20 Prozent haben eine Korrelation unter	0,20	0,22	0,29	0,33	0,48
20 Prozent haben eine Korrelation über	0,53	0,44	0,51	0,57	0,64
10 Prozent haben eine Korrelation über	0,59	0,48	0,56	0,62	0,70

Korreliert man wiederum diesen Indikator für die diagnostische Präzision, d. h. die Klassenbinnenkorrelation zwischen Testleistung und Leistungseinschätzung, mit der Testleistung selbst, so ergibt sich für die Klassenstufe 5 kein Zusammenhang (r = 0,03). In der Klassenstufe 9 zeigen sich zwischen den Schulformen sehr unterschiedliche Werte: In den Grundkursen beträgt die Korrelation r = 0,28, in den Erweiterungskursen r = 0,01, in den Realschulen r = -0,01 und in den Gymnasien r = 0,24. Es gibt demnach Hinweise darauf, dass zumindest in bestimmten Untergruppen die Fähigkeit zur Beurteilung von Aufgabenschwierigkeiten in Bezug auf die eigene Schülerschaft positiv mit den Leistungen der Schülerinnen und Schüler zusammenhängt. Warum die Ausgangshypothese über denkbare förderliche Auswirkungen einer hohen diagnostischen Kompetenz nur teilweise bestätigt werden kann, muss vorerst offen bleiben.

Unberücksichtigt blieben bei dieser Betrachtung, wie gesagt, individuelle Tendenzen zur Über- oder Unterschätzung der eigenen Schülerschaft, die ebenfalls als Indikatoren für die pädagogisch-diagnostische Kompetenz herangezogen werden können. Hierbei geht es um die *absolute* Übereinstimmung zwischen den Einschätzungen der Lehrkräfte und den tatsächlichen Testergebnissen in der eigenen Klasse bzw. im eigenen Kurs (vgl. Tabelle 48).

Tabelle 48 Absolute Übereinstimmung zwischen den Leistungseinschätzungen durch die Lehrkräfte und den Testleistungen in den Lerngruppen, nach Klassenstufe und Schulform (in Prozent)

	durchschnittl. Überschätzung	**höchste Überschätzung**	**höchste Unterschätzung**
Klassenstufe 5	14,6	41,1	11,4
Klassenstufe 9			
Gesamtschulen, Grundkurse	12,4	69,7	21,6
Gesamtschulen, Erweiterungskurse	20,3	52,2	19,6
Realschulen	18,0	39,5	9,0
Gymnasien	20,3	53,3	11,5
insgesamt	*16,7*	*69,7*	*21,6*

Nach den bisherigen Befunden sind diese Ergebnisse zunächst überraschend. Hinsichtlich der Beurteilung des globalen Leistungsniveaus der eigenen Lerngruppe, die sich hinter diesen Zahlen verbirgt, unterscheiden sich die verschiedenen Gruppen von Lehrkräften keineswegs in derselben Weise wie hinsichtlich ihrer differenzierten Einschätzung der relativen Aufgabenschwierigkeiten. Insgesamt zeigt sich bei den Lehrkräften sowohl für die Klassenstufe 5 als auch für alle Schul- bzw. Kursformen der Klassenstufe 9 eine allgemeine, freilich maßvolle, Tendenz zur Überschätzung der Testleistung der eigenen Schülerinnen und Schüler. Diese liegt bei 16,7 Prozent. Die Unterschiede zwischen den Klassenstufen und Schulformen bzw. Kursniveaus folgen jedenfalls nicht den bisher gefundenen Mustern. Es ist aber vorstellbar, dass die in Bezug auf die relativen Einschätzungen aufgedeckten größeren Schwierigkeiten in den Grundschulen und in den Mathematik-Grundkursen an Gesamtschulen weniger mit (mangelnder) diagnostischer Professionalität und statt dessen mehr mit unterschiedlichen Ausdifferenzierungen des Mathematikunterrichts zu tun haben, die es vor allem den Gymnasiallehrkräften erleichtern, die relativen Aufgabenschwierigkeiten zu erkennen.

Bei den Leistungseinschätzungen hatten die Lehrkräfte die Möglichkeit, die einzelnen Aufgaben zu kommentieren. Die meisten dieser Kommentare bezogen sich auf die Aufgabenstellung und den Behandlungsgrad der Aufgaben. Hieraus lässt sich ein weiteres Kriterium für die Validität der eingesetzten QuaSUM-Mathematiktests ableiten. Gleichzeitig würde eine größere Anzahl in Zweifel zu ziehender Aufgaben möglicherweise die Ergebnisse der nachträglichen Leistungseinschätzung verfälschen. Eine Inhaltsanalyse der Kommentare zu den einzelnen Aufgaben ergab, dass im Test für die Klassenstufe 5 nur zwei von den insgesamt 40 Aufgaben enthalten waren, bei denen nach

Auskunft von mehr als einem Drittel der befragten Lehrkräfte der entsprechende Inhalt nicht behandelt worden war[52]. In der Klassenstufe 9 gab es nur in der Gymnasialversion des QuaSUM-Mathematiktests eine Aufgabe, bei der mehr als ein Drittel der Lehrkräfte, die die Leistungseinschätzung ihrer Schülerinnen und Schüler vorgenommen haben, angaben, dass der Inhalt dieser Aufgabe nicht Gegenstand des Unterrichts gewesen war. In den Versionen für die Gesamtschul-Grundkurse, die Gesamtschul-Erweiterungskurse und die Realschulen waren nach Angabe der Lehrkräfte keine solchen Aufgaben gestellt worden.

5.4 Schullaufbahn – Zusammenhänge zwischen Testleistung und Sitzenbleiben, Springen, Schulform- und Kurswechsel

Für die Auswertung der Schullaufbahn der Schülerinnen und Schüler standen aus den Schülerakten Angaben zu den drei Schuljahren 1998/99, 1997/98 und 1996/97 zur Verfügung. Für das Schuljahr 1995/96 lagen solche Daten nur für die getestete Klassenstufe 5 und in Ausnahmefällen für die Klassenstufe 9 vor. Das bedeutet, dass bei normalem Schulverlauf (ohne Überspringen oder Wiederholen) für die Klassenstufe 5 Angaben zu den Klassenstufen 2 bis 5 vorhanden sind und für die Klassenstufe 9 Angaben zu den Klassenstufen 7 bis 9, zur Klassenstufe 6 nur in Ausnahmefällen.

Klassenstufe 5

In der Klassenstufe 5 gibt es nur einen extrem geringen Anteil Schülerinnen und Schüler, die während der letzten vier Schuljahre eine Klasse *übersprungen* haben (0,1 Prozent). Eine eigene Analyse ist daher für diese Gruppe nicht sinnvoll. Der Anteil der Schülerinnen und Schüler, die eine Klasse zwischen der 3. und 5. Klasse *wiederholt* haben, beträgt 2,6 Prozent. Dabei haben 0,8 Prozent die 3. Klasse wiederholt, 0,6 Prozent die 4. Klasse und 1,1 Prozent die 5. Klasse. Unterscheidet man die Wiederholer von den Schülerinnen und Schülern, die keine Klassenstufe wiederholt haben, so ergeben sich zwischen den beiden Gruppen hinsichtlich der Ergebnisse im QuaSUM-Mathematiktest die in Tabelle 49 aufgezeigten Unterschiede.

[52] Folgende Kommentare der Lehrkräfte wurden als „Inhalt der Aufgabe nicht behandelt" bewertet: „Behandlung ist zu lange her", „(noch) nicht behandelt", „kaum behandelt", „Stoff der nächsten Klassenstufe", „nicht in der Form behandelt", „Stoff eines anderen Fachs".

Tabelle 49 Ergebnisse im QuaSUM-Mathematiktest, nach Klassenwiederholung in den Klassenstufen 2 bis 5 (durchschnittliche Rasch-Skalenwerte)

	Mittelwert	Standardabweichung
Wiederholer	131,4	25,0
Nichtwiederholer	150,4	17,7

Diejenigen, die in den letzten vier Jahren mindestens einmal eine Klasse wiederholt haben, stellen eine besonders lernschwache Gruppe dar, die entweder durch die Klassenwiederholung nicht den Anschluss an ein mittleres Leistungsniveau hat finden können oder – falls die Wiederholung bereits mehrere Jahre zurück liegt – inzwischen wieder auf einen Lernstand zurückgefallen ist, der dem typischen Leistungsbereich der Note „ausreichend“ zuzuordnen ist. Der Unterschied in der Testleistung gegenüber den Nichtwiederholern ist hoch, die Effektstärke beträgt $d = 0{,}76$.

Klassenstufe 9

Das Phänomen des *Überspringens* einer Klasse tritt in Brandenburg in den Klassenstufen 7 bis 9 nicht mehr auf. Dagegen nimmt die *Wiederholung von Klassenstufen* etwas zu: Die siebte Klassenstufe haben 0,8 Prozent der Schülerinnen und Schüler wiederholt, die achte Klassenstufe 1,5 Prozent und die neunte Klassenstufe 4,1 Prozent. Insgesamt haben 6,3 Prozent der Schülerinnen und Schüler mindestens eine der Klassenstufen 7 bis 9 wiederholt. Getrennt nach den unterschiedlichen Schulformen haben 13,3 Prozent der Schülerinnen und Schüler aus Mathematik- Grundkursen, 2,4 Prozent Schülerinnen und Schüler aus Mathematik-Erweiterungskursen, 7,3 Prozent Schülerinnen und Schüler aus Realschulen und 1,3 Prozent Schülerinnen und Schüler aus Gymnasien mindestens eine der Klassen 7 bis 9 wiederholt. Setzt man das Wiederholen einer Klasse in Beziehung zu den Mathematikleistungen, ergibt sich für die Klassenstufe 9 das in Tabelle 50 dargestellte Bild.

Die Werte in der Tabelle verweisen deutlich auf die allgemeine Tendenz, dass – ähnlich wie an den Grundschulen – die Schülerinnen und Schüler, die eine Klassenstufe wiederholen, auch danach eher zu den Leistungsschwächeren gehören. Eine Ausnahme bilden die Gesamtschul-Erweiterungskursler, was insofern überrascht, als Klassenwiederholungen erstens im Gesamtschulsystem bis zum Ende der Klassenstufe 8 nur auf Elternwunsch erfolgen können und zweitens unter solchen Bedingungen Klassenwiederholer nicht unbedingt auf dem anspruchsvolleren Niveau der Erweiterungskurse vermutet werden.

Tabelle 50 Ergebnisse im QuaSUM-Mathematiktest und im Mathe40-Test, nach Wiederholung (W) und Nichtwiederholung (NW) in den Klassenstufen 7 bis 9 (durchschnittliche Rasch-Skalenwerte und Effektstärken *d*)

	QuaSUM-Mathematiktest			**Mathe40-Test**		
	W	**NW**	***d***	**W**	**NW**	***d***
Gesamtschulen, Grundkurse	122,9	126,3	0,23	429	435	0,12
Gesamtschulen, Erweiterungskurse	154,6	150,8	-0,23	521	501	-0,30
Realschulen	138,2	150,3	0,66	460	499	0,56
Gymnasien	165,3	175,2	0,57	560	577	0,22
insgesamt	132,2	151,0	0,75	454	504	0,58

Denkbar ist allerdings, dass wegen der Bedeutung der Kurseinstufung für das Abschlusszeugnis an Gesamtschulen die Klassenwiederholung häufiger auch bei „besseren“ Schülerinnen und Schülern in der Erwartung erfolgt, über ein – anders mutmaßlich nicht erreichbares – günstiges Leistungsprofil den anspruchsvolleren Schulabschluss zu sichern. Angemerkt sei, dass die erkennbaren Effekte einer Klassenwiederholung wieder deutlich niedriger ausfallen, wenn man statt des curriculumnahen QuaSUM-Mathematiktests den Mathe40-Test als Vergleichsgrundlage wählt.

In der Klassenstufe 9 sind neben der Frage nach der Wiederholung einer Klasse vor allem Aspekte des Schulformwechsels bzw. des Kursniveauwechsels von Interesse. Insgesamt erfolgt ein *Schulformwechsel* in Brandenburg zwischen der 7. und der 9. Klasse sehr selten: 97,7 Prozent der Schülerinnen bleiben zwischen der 7. und 9. Klasse in der Schulform, in die sie nach Beendigung der Grundschule eingetreten sind. Auch hier ist wegen der geringen Gruppenumfänge kaum eine spezielle Analyse angezeigt. Was die *Kurszugehörigkeit* der Schülerinnen und Schüler in den Gesamtschulen in den Fächern Deutsch[53], Mathematik und erste Fremdsprache für die Klassenstufen 7 bis 9 anbelangt, so zeigen sich ähnlich stabile Tendenzen wie bei der Schulformzugehörigkeit: 70,2 Prozent der Schülerinnen und Schüler behalten während dieser drei Schuljahre ihre jeweilige Kurszuordnung in allen drei Fächern bei, und bei 22,4 Prozent der Schülerinnen und Schüler gibt es lediglich in einem Fach einen Kurswechsel, der entweder endgültig ist oder wieder rückgängig gemacht wird. Bei 2,3 Prozent der Schülerinnen und Schüler ist eine ständige

[53] Das Kurssystem für das Fach Deutsch beginnt erst in der Klassenstufe 8, während Mathematik und die erste Fremdsprache bereits ab Klassenstufe 7 in Kursen unterrichtet werden.

Verbesserung, bei weiteren 2,3 Prozent eine ständige Verschlechterung zu verzeichnen. Nur bei den restlichen 2,8 Prozent der Schülerinnen und Schüler ist das Bild der Kurszugehörigkeit inkonsistent, da diese ständig wechselt.Im vorliegenden Zusammenhang ist vor allem die Kursniveauzugehörigkeit für das Fach Mathematik bedeutsam (vgl. Tabelle 51).

Tabelle 51 Kurszugehörigkeit der Schülerinnen und Schüler im Fach Mathematik an Gesamtschulen über die Klassenstufen 7 bis 9

im Schuljahr 1996/97	**1997/98**	**1998/99**	**Prozentsatz der Schülerinnen und Schüler**
besuchtes Kursniveau in Mathematik			
Grundkurs	Grundkurs	Grundkurs	43,3
Grundkurs	Grundkurs	Erweiterungskurs	3,2
Grundkurs	Erweiterungskurs	Grundkurs	0,7
Grundkurs	Erweiterungskurs	Erweiterungskurs	3,7
Erweiterungskurs	Grundkurs	Grundkurs	5,1
Erweiterungskurs	Grundkurs	Erweiterungskurs	0,9
Erweiterungskurs	Erweiterungskurs	Grundkurs	3,3
Erweiterungskurs	Erweiterungskurs	Erweiterungskurs	39,9

Wie man der Auflistung entnehmen kann, ist – bezogen auf das Fach Mathematik – die Kurseinstufung noch stabiler, als dies für die leistungsdifferenzierten Fächer Deutsch, erste Fremdsprache und Mathematik im Verbund gilt: 83,2 Prozent der Schülerinnen und Schüler wurden im Fach Mathematik stets auf demselben Kursniveau unterrichtet. Ob die großen Überschneidungsbereiche in den landesweiten Leistungsverteilungen der beiden Kursniveaus auch auf der Ebene der Einzelschulen – und, wenn ja, welcher – zu verzeichnen sind, bedürfte eigener umfangreicher Untersuchungen, die den Rahmen des vorliegenden Berichts sprengen würden. Auffällig ist jedenfalls, dass in den drei Jahren nach der Grundschule nur 6,9 Prozent der Schülerschaft den Anschluss vom ursprünglichen Grundkursniveau an einen Erweiterungskurs bewältigt. Dem stehen 8,4 Prozent „Absteiger" gegenüber.

Informativ sind in diesem Zusammenhang auch die durchschnittlichen Testergebnisse, differenziert nach der Kursniveauzugehörigkeit über die Zeit. Betrachtet man die Zahlen in Tabelle 52, so bestätigen die Testergebnisse insgesamt die Plausibilität der Kurszuordnung: Sieht man von den klaren Verhältnissen in den Extremgruppen ab, so haben Schülerinnen und Schüler, die bereits seit zwei Jahren den Mathematik-Erweiterungskurs besucht haben, die günstigsten Werte, gefolgt von den „Aufsteigern" des letzten Jahres, die wegen des erst jüngst erfolgten Wechsels offenbar teilweise noch den vollen

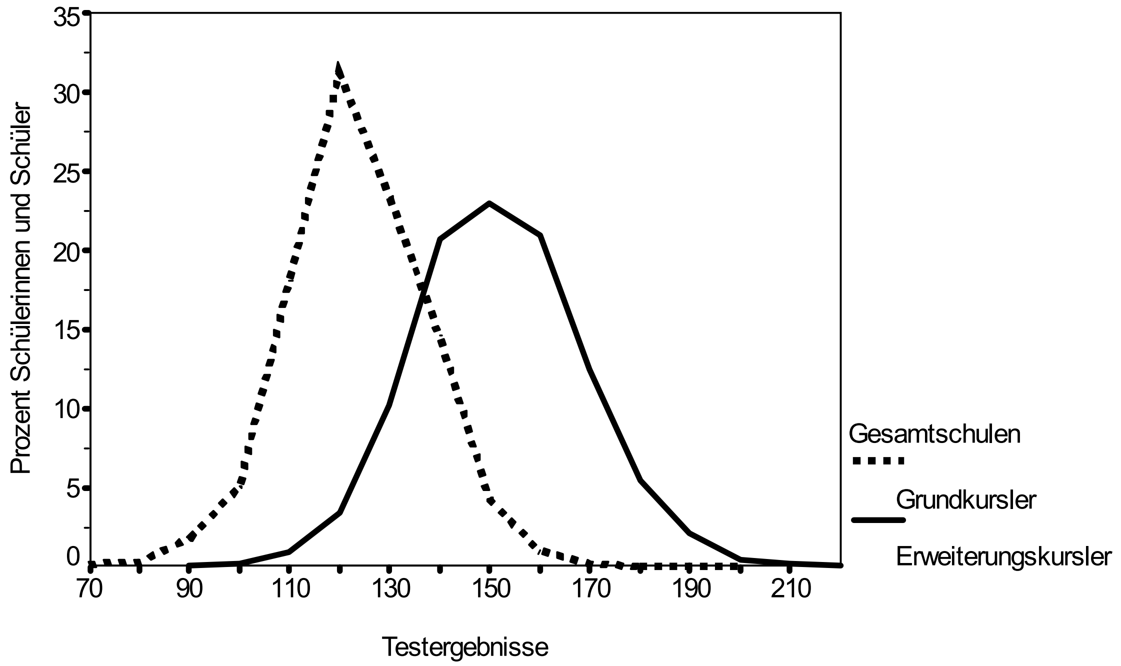
Anschluss an das höhere Kursniveau gewinnen müssen. „Absteiger“ dagegen befinden sich zumeist eindeutig in dem Leistungsbereich, der in Brandenburg keine „ausreichende“ Note im Erweiterungskurs mehr rechtfertigt.

Tabelle 52 Durchschnittliches Testergebnis im QuaSUM-Mathematiktest und im Mathe40-Test nach Kurszugehörigkeit an Gesamtschulen über die Klassenstufen 7 bis 9

im Schuljahr 1996/97	1997/98	1998/99	QuaSUM-Mathematiktest	Mathe40-Test
besuchtes Kursniveau in Mathematik				
Grundkurs	Grundkurs	Grundkurs	123,6	427
Grundkurs	Grundkurs	Erweiterungskurs	138,9	462
Grundkurs	Erweiterungskurs	Grundkurs	131,2	451
Grundkurs	Erweiterungskurs	Erweiterungskurs	147,1	484
Erweiterungskurs	Grundkurs	Grundkurs	134,3	464
Erweiterungskurs	Grundkurs	Erweiterungskurs	140,0	485
Erweiterungskurs	Erweiterungskurs	Grundkurs	133,7	466
Erweiterungskurs	Erweiterungskurs	Erweiterungskurs	151,7	503

Vergleicht man die beiden Extremgruppen, also die Schülerinnen und Schüler, die drei Schuljahre lang im Mathematik-Grundkurs waren, mit denen, die drei Schuljahre lang im Mathematik-Erweiterungskurs waren, bezüglich ihrer Leistung im QuaSUM-Mathematiktest, so ergeben sich doch trotz der allgemein plausiblen Kurseinstufungspraxis Rückfragen (vgl. Abbildung 30).

Abbildung 30 Ergebnisse im QuaSUM-Mathematiktest an Gesamtschulen, nach Kursniveauzugehörigkeit in Mathematik (unter Ausschluss der Schülerinnen und Schüler mit Kurswechsel)



Erkennbar wird in dieser Abbildung ein nicht unerheblicher Überschneidungsbereich zwischen den beiden Gruppen mit stabiler Kurszugehörigkeit: 15,2 Prozent derjenigen, die immer einen Mathematik-Grundkurs besucht haben, fallen leistungsmäßig in das Spektrum, das nach Brandenburger Standards den Mathematik-Erweiterungskursen zugeordnet werden kann, und sogar 19,6 Prozent derjenigen, die immer die günstigere Einstufung erhalten haben, erbringen – ausweislich des QuaSUM-Mathematiktests – nur Leistungen auf Grundkursniveau. Somit stellen sich auch innerhalb des Bezugsrahmens der Gesamtschule grundsätzlich Fragen nach der Chancengerechtigkeit bei den Selektionsentscheidungen.

Als sinnvoll erscheint in dem Zusammenhang ein Vergleich der Mathematik-Grund- und Erweiterungskurse des Schuljahres 1998/99 unter der Frage, welche Kriterien für die Zuweisung zu einem Erweiterungskurs Mathematik maßgeblich sind. Eine entsprechende Regressionsanalyse, die 42 Prozent der Varianz aufklärt, liefert die nachstehenden Einflussfaktoren in der Reihenfolge ihres (Beta-) Gewichts:

- eine hohe Testleistung im QuaSUM-Mathematiktest (β = 0,60),
- das Wiederholen einer Klasse zwischen der 7. und 9. Klassenstufe (β = -0,11),
- das leistungsbezogene Selbstkonzept (β = 0,06),
- die Wahrnehmung des Mathematikunterrichts als eines straff gelenkten, wohlstrukturierten Geschehens (β = 0,06),
- das Geschlecht (weiblich) (β = 0,05),
- der Bildungsabschluss der Eltern (β = -0,04),
- die positive Einschätzung der eigenen Anstrengung als Beitrag zum Schulerfolg (β = 0,03).

Die mit dem QuaSUM-Mathematiktest gemessene Fachleistung besitzt somit die überragende Bedeutung bei der Niveaueinstufung der Schülerinnen und Schüler an den Gesamtschulen. Daneben erhöht es die Wahrscheinlichkeit für eine solchen Einstufung, wenn keine Klasse zwischen der 7. und 9. Klassenstufe wiederholt wurde, wenn ein günstiges leistungsbezogenes Selbstkonzept gegeben ist, wenn es sich um eine Schülerin bzw. einen Schüler handelt, die oder der den Mathematikunterricht als straff geführt, transparent und wohlstrukturiert empfindet, wenn es sich um ein Mädchen handelt, wenn die Eltern ceteris paribus einen niedrigeren Bildungsabschluss haben, und wenn die eigene Anstrengung positiv als Beitrag zum Schulerfolg bewertet wird. Man kann diese Ergebnisse so deuten, dass auch sie noch einmal die Plausibilität der Praxis bei der Kurseinstufung an den Gesamtschulen bestätigt. Es ist jedoch denkbar, dass auch noch ganz andere, schulorganisatorische Faktoren an der Kurseinstufung beteiligt sind. Falls der Erweiterungskurs bereits viele Schülerinnen und Schüler enthält, wird z. B. die Neigung abnehmen, neue aus den Grundkursen aufsteigen zu lassen; ist der Erweiterungskurs sehr klein, wird diese Neigung zunehmen.

6 Bedingungen von Mathematikleistungen: ein Strukturmodell

In diesem Abschnitt sollen die wichtigsten Befunde aus den vorangegangenen Analysen miteinander verbunden werden, und zwar nicht als eine einfache Zusammenfassung, sondern in Form eines Strukturmodells, in dem die verschiedenen, insgesamt recht komplexen Zusammenhänge simultan verrechnet und interpretiert werden können. Dabei wird angeknüpft an das in Abschnitt 1 skizzierte Rahmenmodell, das für die Anlage der QuaSUM-Untersuchung von Anfang an maßgeblich war.

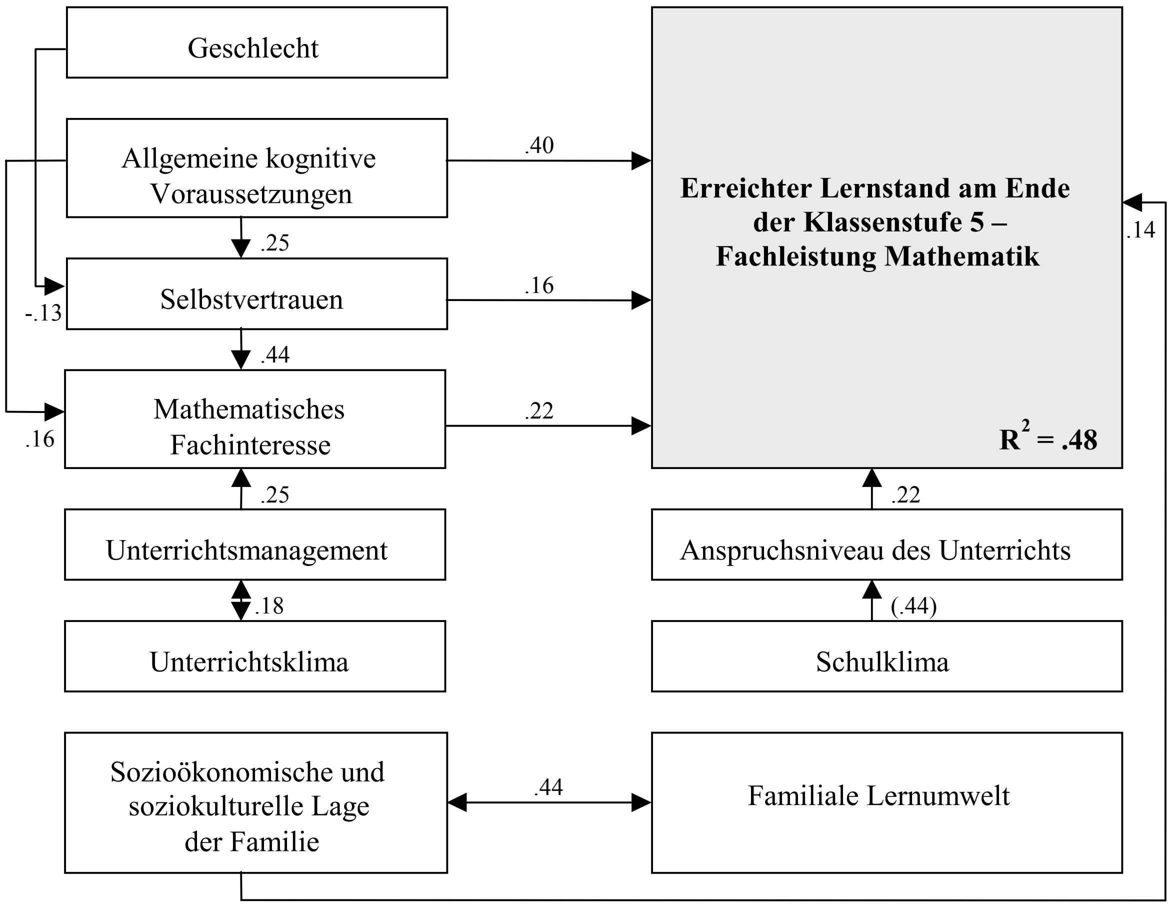
Die verwendete statistische Technik ist die der *Pfadanalyse*. Sie umfasst eine ganze Sequenz von Regressionsanalysen der Art, wie sie bereits mehrfach vorgestellt worden sind. Indem diese Regressionsanalysen nun nacheinander abgearbeitet werden, gelingt es, korrelative Zusammenhänge so aufzulösen, dass zwischen eigenständigen *direkten Effekten* – grafisch als einfacher Pfeil markiert – und den durch intervenierende Variablen vermittelten *indirekten Effekten* unterschieden werden kann, die sich aus der Kombination von zwei oder mehr direkten Effekten zusammensetzen. Die Stärke eines indirekten Effekts entspricht dem Produkt der daran beteiligten direkten Effekte. Die Summe der direkten und indirekten Effekte zwischen zwei Variablen ergibt die (einfache) Korrelation. Nicht weiter analysierte korrelative Zusammenhänge werden durch Doppelpfeile symbolisiert. Dabei ist es wichtig zu betonen, dass die Strukturbeziehungen selbst nicht Ergebnis der Rechnungen sind, sondern diesen als ein System von theoretisch begründeten und überwiegend auch in anderen Untersuchungen empirisch gut bestätigten Hypothesen vorausgehen.

Insgesamt wurde für jede der beiden Klassenstufen eine große Zahl von möglichen Varianten iterativ gerechnet. Kleine Effekte ($\beta < 0{,}10$) wurden dabei weggelassen, um zu einer sparsamen, übersichtlichen Modellstruktur zu gelangen. In einzelnen Fällen wurde das Rahmenmodell leicht modifiziert, weil sich die erwarteten Zusammenhänge so nicht belegen ließen, dafür aber vergleichbar bedeutsame Hintergründe für die Ausprägung eines bestimmten Konzepts im Modell ermittelt werden konnten.

6.1 Bedingungen von Mathematikleistungen in der Klassenstufe 5

Zentrales Anliegen der QuaSUM-Studie und damit auch der hier entwickelten Analyse ist es, die Unterschiede zwischen den Schülerinnen und Schülern hinsichtlich ihrer Mathematikleistung zu erklären. Dabei entfällt für die Klassenstufe 5 das Element der Unterscheidung von Schulformen bzw. Kursniveaus. Dennoch können über die in der Abbildung 31 erfassten Zusammenhänge rund 48 Prozent der Zwischen-Schüler-Varianz aufgeklärt werden.

Abbildung 31 Bedingungen der Mathematikleistungen der Schülerinnen und Schüler aus fünften Klassen



Die „Pfadkoeffezienten", d. h. die bei den einzelnen Pfeilen stehenden Beta-Gewichte, zeigen an, dass in der Klassenstufe 5 die Fachleistung Mathematik – gemessen über den QuaSUM-Mathematiktest – von den erhobenen Merkmalen mit Abstand am stärksten mit allgemeinen kognitiven Voraussetzungen zusammenhängt: Schülerinnen und Schüler mit höheren Fähigkeiten im schlussfolgernden Denken – gemessen über den CFT 20 – haben auch die besseren Ergebnisse im QuaSUM-Mathematiktest (β = 0,40). Dies entspricht den Erwartungen. Übrigens könnte an dieser Stelle – dies zeigen vergleichbare Studien – mit ähnlichem Einfluss auch die Leistung in einem anderen Schulfach, besser aber noch die Fachleistung Mathematik zu einem früheren Zeitpunkt stehen. Hinzu kommen die indirekten Effekte, die sich aus dem Zusammenhang zwischen der Fähigkeit zum schlussfolgernden Denken und positiven schul- und unterrichtsbezogenen Einstellungen der Schülerinnen und Schüler ergeben. Unabhängig vom allgemeinen Lernpotenzial wirken sich ein ausgeprägtes Sach- und Fachinteresse für die Mathematik ($\beta = 0{,}22$) und ein starkes Selbstvertrauen ($\beta = 0{,}16$) günstig auf die Fachleistung aus. Beides ist bei den Jungen höher entwickelt ($r = -0{,}12$ bzw. $r = -0{,}11$), womit die Leistungsrückstände der Mädchen zumindest teilweise erklärt sind (vgl. dazu den fehlenden direkten Zusammenhang zwischen Geschlecht und Fachleistung Mathematik). Der Einfluss der sozialen Herkunft der Schülerinnen und Schüler auf die Fachleistung Mathematik ist nach Berücksichtigung des allgemeinen Fähigkeits- und Einstellungsprofils weniger stark ausgeprägt ($\beta = 0{,}14$), als man vielleicht erwartet. Schülerinnen und Schüler aus bildungsnahen Elternhäusern haben nur geringfügig bessere Testergebnisse erreicht, wenn man die anderen – gewichtigeren – Determinanten der Fachleistung berücksichtigt.

Von besonderem Interesse ist die Frage, ob sich aus dem Datenmaterial Aussagen darüber gewinnen lassen, mit welchen Lehrereinstellungen oder -verhaltensweisen ein hohes Fachinteresse für die Mathematik einher geht; denn von den bislang vier der insgesamt fünf herausgearbeiteten direkten Bedingungen für eine hohe mathematische Fachleistung ist diese wohl am ehesten durch eine angemessene Gestaltung des Unterrichts zu beeinflussen. In der Tat zeigen diejenigen Schülerinnen und Schüler, die den Unterricht als gut strukturiert und transparent gestaltet empfinden, ein höheres mathematisches Interesse ($\beta = 0{,}25$). Hier geht es also um das Unterrichtsmanagement aus der Sicht der Schülerinnen und Schüler, was auch damit zu begründen ist, dass aus den Lehrerangaben kein nennenswerter Anhaltspunkt für die Erklärung von Unterschieden in der Ausprägung des Sach- und Fachinteresses der Schülerinnen und Schüler zu gewinnen war. Darauf, dass es sich hierbei nicht nur um Schülerwahrnehmungen handelt, die mit dem tatsächlichen Verhalten der Lehrkraft wenig oder nichts zu tun haben, deutet jedoch der Umstand hin, dass die Lehrkräfte solcher Klassen selber das Arbeitsklima in der Klasse als überdurchschnittlich positiv empfinden ($r = 0{,}18$).

Zudem zeigt sich ein ähnlicher Zusammenhang, wenn man das Anspruchsniveau des Unterrichts näher untersucht. Zu diesem Zweck ist für jede Klasse der Benotungsstandard in der Weise bestimmt worden, dass die geringste Testleistung eines Schülers oder einer Schülerin mit „befriedigender" Zensur im Halbjahreszeugnis als charakteristischer Schwellenwert für ein mittleres Leistungsniveau in dieser Lerngruppe ermittelt wurde. Es gilt nun, dass die individuelle Fachleistung um so höher ist, je höhere Ansprüche – gemessen an diesem Kriterium – die jeweilige Lehrkraft stellt ($\beta = 0{,}22$). Darin drückt sich einerseits der oben dargestellte Sachverhalt aus, dass es in leistungsstarken

Klassen vergleichsweise schwieriger ist, eine „gute“ oder auch nur „befriedigende“ Note zu erhalten. Gleichzeitig aber liegt es durchaus nahe anzunehmen, dass sich in einer solchen Lernumgebung die Anstrengungsbereitschaft und damit die Fachleistung der Schülerinnen und Schüler erhöht.

Für diese Deutung spricht, dass an den Grundschulen die Benotungsstandards der Lehrkräfte offenbar nicht unabhängig von den Rahmenbedingungen auf der Schulebene sind. Je günstiger das Schulklima aus der Sicht der Lehrkräfte – ebenso wie aus der Sicht der Schulleiterinnen bzw. Schulleiter – ausgeprägt ist, je intensiver sich die Zusammenarbeit der Lehrkräfte gestaltet, je weniger direktiv die Schulleitung verfährt und je effektiver die Lehrkraft die Lernprozesse durch einen sorgfältig geplanten Unterricht steuert, je mehr aber dabei auch die Kreativität und Eigenständigkeit der Lernenden gewahrt bleiben, desto höher sind offenbar die Standards (R = 0,44). Diese Erkenntnis verdankt sich einer Analyse, mit der den Rahmenbedingungen für einen an der Benotungspraxis erkennbaren anspruchsvollen Unterricht im Einzelnen nachgegangen wurde, und zwar auch unter Berücksichtigung der unterschiedlichen kognitiven und sozialen Voraussetzungen auf Seiten der Schülerinnen und Schüler. In der Reihenfolge der je eigenständigen Bedeutung dieser Rahmenbedingungen, nach Maßgabe der Beta-Gewichte also, sind die folgenden Aspekte zu nennen:

- eine positive Wahrnehmung des Schulklimas durch die Lehrkraft (β = 0,23),
- eine gute Zusammenarbeit unter den Lehrkräften (β = 0,21),
- eine eng kontrollierende Unterrichtsführung (β = 0,19),
- die Betonung der Kreativität und Eigenständigkeit von Schülerinnen und Schülern durch die Lehrkraft (β = 0,17),
- eine vergleichsweise geringe Betonung administrativer Pflichten der Schulleitung (β = -0,13),
- eine vergleichsweise geringe Betonung von Beratungsaufgaben der Schulleitung (β = -0,12),
- eine positive Wahrnehmung des Schulklimas durch die Schulleitung (β = 0,12).

Damit wird eine Reihe von Hypothesen der Schulentwicklungsforschung – jedenfalls für die Grundschulen in Brandenburg bzw. die in den fünften Klassen tätigen Lehrkräfte – bestätigt, insoweit die genannten Rahmenbedingungen – wenn auch nur mittelbar – mit dem an einer Schule erreichten Fachleistungsniveau zusammenhängen.

Bemerkenswert ist der Mechanismus der Vermittlung, nämlich die Bedeutung, die offenbar einem auch über die Beurteilungspraxis vermittelten hohen Leistungsanspruch zukommt. Gemeinsam mit dem Befund, dass ein relativ straffes Unterrichtsmanagement ceteris paribus für ein hohes Leistungsniveau verantwortlich zu sein scheint – hier vermittelt über den positiven Einfluss auf das Sach- und Fachinteresse –, spricht vieles dafür, die Lernanforderungen insbesondere dort zu erhöhen, wo im Grunde eine höhere Mitarbeits- und Leistungsbereitschaft seitens der Schülerinnen und Schüler vorausgesetzt werden kann.

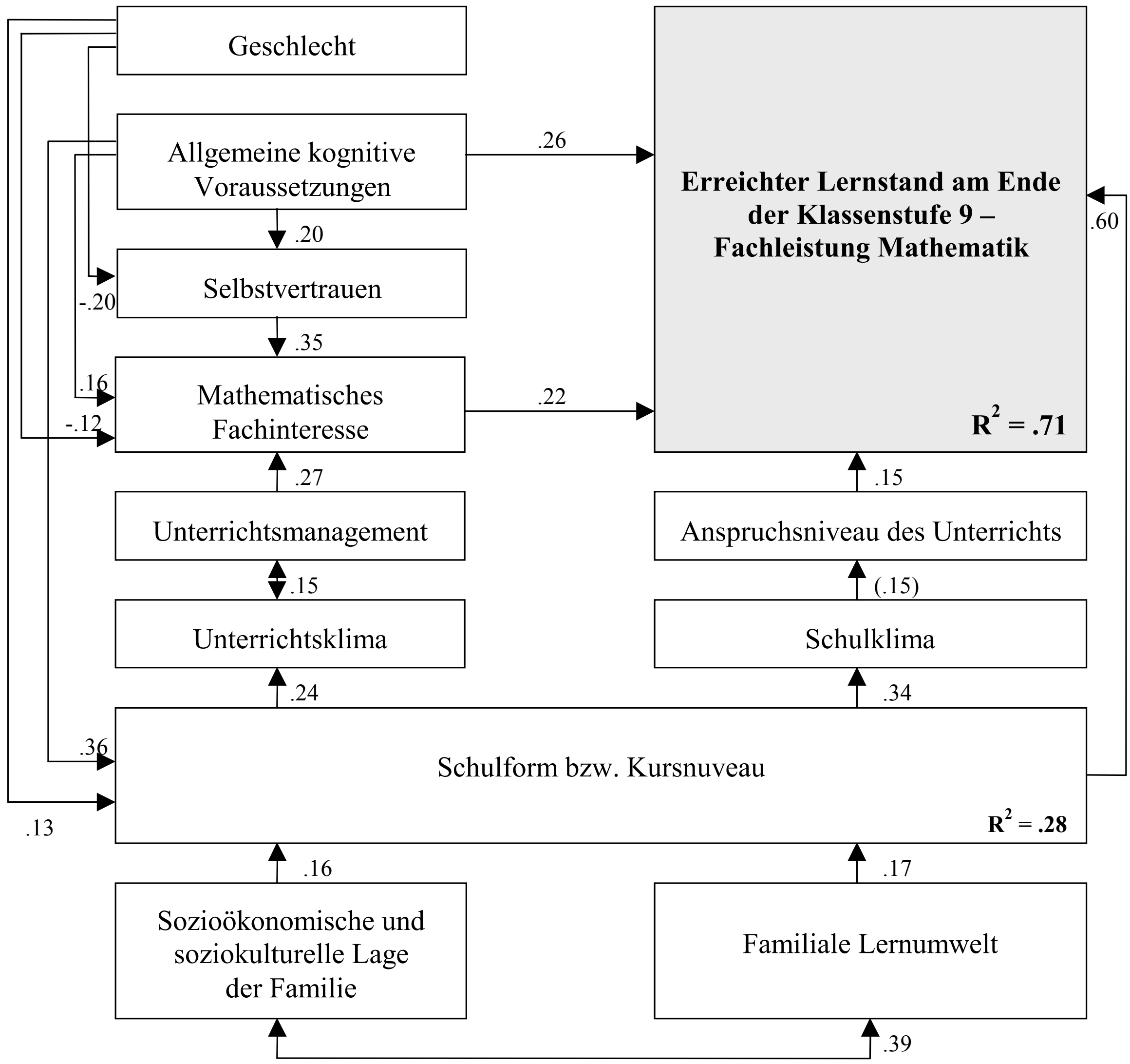
6.2 Bedingungen von Mathematikleistungen in der Klassenstufe 9

Die Befunde der Analysen für die Klassenstufe 5 werden in wesentlichen Teilen durch analoge Untersuchungen der Neuntklässlerinnen und Neuntklässler bestätigt. Insofern in der Klassenstufe 9 der Unterscheidung verschiedener Schulformen bzw. Kursniveaus zentrale Bedeutung zukommt, wird ein neues Strukturelement in das Rahmenmodell eingefügt, das – wie sich zeigen wird – das Bedingungsgefüge insgesamt entscheidend beeinflusst.

Die in diesem Modell berücksichtigten Variablen sind in der Regel ebenso definiert wie im Falle der Klassenstufe 5. Die Variable „Schulform- bzw. Kursniveau" wurde so behandelt, dass keine Unterscheidung zwischen den Mathematik-Erweiterungskursen der Gesamtschulen und den Realschulen getroffen wurde. Die Rechtfertigung dafür liegt zum einen in den kaum unterscheidbaren Leistungsprofilen der beiden Schülergruppen begründet, zum anderen darin, dass Realschulen in Brandenburg verhältnismäßig selten sind, eine Unterscheidung zwischen den beiden Gruppen also stark von den lokalen Gegebenheiten beeinflusst wäre. Da die Leistungsabstände zwischen den drei so gebildeten Gruppen (Grundkurse vs. Erweiterungskurse/Realschulen vs. Gymnasien) im Rahmen der Fehlergrenzen gleich sind, spricht unter der Leitfrage nach den Hintergründen der Fachleistung Mathematik wenig dagegen, diese Variable als quasi-intervallskaliert in die Rechnungen einzuführen.

Fragt man nach denjenigen Faktoren, die je für sich mit der im QuaSUM-Mathematiktest gemessenen Fachleistung in Zusammenhang stehen, so wird zunächst und erwartungsgemäß die überragende Bedeutung der institutionellen Aufteilung der Schülerschaft in drei Leistungsniveaus erkennbar: Insgesamt können 71 Prozent der Varianz zwischen den Schülerinnen und Schülern aufgeklärt werden, davon der weitaus größte Anteil über die Variable „Schulform bzw. Kursniveau" in dem eben beschriebenen Sinne ($\beta = 0{,}60$).

Abbildung 32 Bedingungen der Mathematikleistungen der Schülerinnen und Schüler aus neunten Klassen



Nach jener Aufteilung der Schülerschaft in unterschiedliche Schulformen und Kursniveaus, die mit den allgemeinen kognitiven Lernvoraussetzungen eng verbunden ist (β = 0,36), schwindet notwendig der statistische Einfluss der Testleistung im CFT 20: Unter Berücksichtigung der Schulform (also *innerhalb* der drei Gruppen, gleichsam im Durchschnitt) beträgt der mit der Fähigkeit zum schlussfolgernden Denken assoziierte Effekt nur mehr β = 0,26. Auch innerhalb der drei hier unterschiedenen Gruppen gilt also – wenn auch aus dem genannten Grunde abgeschwächt –, dass günstige allgemeine Lernvoraussetzungen mit hohen Testleistungen einhergehen. Erhalten bleibt hingegen das Gewicht des Sach- und Fachinteresses an der Mathematik (β = 0,22): Die interessierteren Schülerinnen und Schüler sind in der Regel auch die leistungsstärkeren. Nachweisbar ist ebenfalls auch in der Klassenstufe 9 ein positiver Einfluss des jeweiligen Anspruchsniveaus der Lehrkraft, wiederum über den Minimalstandard für eine „befriedigende" Note erfasst, und zwar relativ zu den Standards der jeweiligen Schulform (β = 0,15). In allen Schulformen bzw. Kursniveaus sind in den Klassen, in denen es – unter Berücksichtigung der übrigen Rahmenbedingungen – schwerer ist, eine „befriedigende" oder „gute" Note zu erhalten, auch höhere Testleistungen anzutreffen.

Besonderes Augenmerk richtet sich auf das leistungsbezogene Selbstkonzept (in der Abbildung als „Selbstvertrauen" bezeichnet), weil es nicht nur von den allgemeinen kognitiven Voraussetzungen abhängt (β = 0,20), sondern sogar noch stärker als bei den Jüngeren vom Geschlecht (β = -0,20). Wie aus der Literatur bekannt, verfügen leistungsstarke Schülerinnen und Schüler über ein ausgeprägteres Selbstbewusstsein als leistungsschwache, und im Allgemeinen haben Jungen ein günstigeres Verhältnis zur eigenen Leistung als Mädchen. Auch das Interesse an der Mathematik ist bei Schülerinnen und Schülern mit besseren Ergebnissen im CFT 20 in der gleichen Höhe wie in der Klassenstufe 5 stärker entwickelt (β = 0,16). Nachweisbar sind die Jungen an Mathematik interessierter als die Mädchen (β = -0,12).

Wie schon in der Klassenstufe 5 ist ein hohes Fachinteresse eher in solchen Klassen anzutreffen, die – vor allem im Hinblick auf die Transparenz, die Strukturiertheit und das Zeitmanagement – von der Schülerschaft als kompetent geführt empfunden werden, verglichen mit solchen, in denen das weniger der Fall ist (β = 0,27), und wie an den Grundschulen gilt, dass ein vergleichsweise kompetentes Unterrichtsmanagement mit einem nach Auffassung der Lehrkräfte eher positiven Arbeitsklima einher geht (r = 0,15).

Ein Effekt, der sich in der Klassenstufe 9 eindeutig anders darstellt als in der Klassenstufe 5, ist der Zusammenhang zwischen hohen Leistungsansprüchen bzw. Benotungsstandards und den Angaben der Schulleitungen zum Schulklima. Die multiple Korrelation für alle Variablen in der oben angeführten ‚Schulklima-Gruppe' beträgt lediglich R = 0,15, und keiner der Einzeleffekte erfüllt das Einschlusskriterium ($\beta \geq 0{,}10$). *Warum* es an den weiterführenden Schulen nicht gelingt, durch Lehrerkooperation und ein angenehmes Arbeitsklima, durch Kollegialität und durch wenig direktives Verhalten der Schulleitung zu anspruchsvollen Zielsetzungen für den Mathematikunterricht zu gelangen, kann auf der Basis der vorliegenden Daten nicht beantwortet werden. Denkbar ist, dass die ausgeprägte Arbeitsteilung zwischen den Gruppen von Fachlehrkräften die hier in Frage stehenden Zusammenhänge erheblich komplexer gestaltet, als es in dem verwendeten statistischen Modell berücksichtigt werden kann. Auch kommt als Erklärung in Betracht, dass die

Schulen mit einer Sekundarstufe I zumeist größer sind als die Grundschulen, so dass sich hier die Zusammenarbeit und ein für alle günstiges Schulklima schwieriger herstellen lassen.

Auf die zentrale Bedeutung der institutionellen Trennung verschiedener Leistungsniveaus in der Sekundarstufe wurde bereits hingewiesen. Dass an den Gymnasien höhere Leistungsstände zu verzeichnen sind als an den Realschulen und in den Erweiterungskurse, und an diesen wiederum höhere als in den Grundkursen ($\beta = 0{,}60$), ist weitgehend trivial. Auch kann es nicht wirklich überraschen, dass die Schulform- bzw. Kursniveauzugehörigkeit nicht nur von den kognitiven Lernvoraussetzungen ($\beta = 0{,}36$) abhängt, sondern auch und darüber hinaus von der sozioökonomischen und soziokulturellen Lage der Familie, gemessen über die Bildungsabschlüsse der Eltern ($\beta = 0{,}16$), sowie der familialen Lernumwelt, bestimmt über die Zahl der im Haushalt vorhandener Bücher ($\beta = 0{,}17$). Kinder von Eltern mit hohen Bildungsabschlüssen und solche aus Haushalten, in denen zahlreiche Bücher vorhanden sind, besuchen unter sonst gleichen Voraussetzungen mit höherer Wahrscheinlichkeit ein Gymnasium und sind mit der geringsten Wahrscheinlichkeit im Grundkurs einer Gesamtschule anzutreffen. Die Schulform- bzw. Kursniveauzugehörigkeit beeinflusst ihrerseits das Arbeitsklima in der Lerngruppe – letzteres wird von den Lehrkräften der Gymnasialklassen am günstigsten eingeschätzt ($\beta = 0{,}24$) – und in dem gleichen Sinne nach Überzeugung der Schulleitungen das Schulklima ($\beta = 0{,}34$).

Abgesehen davon stellen die Daten der Klassenstufe 9 eine unabhängige Bestätigung für die aus der Analyse der Grundschuldaten gewonnenen Erkenntnis dar, dass ein durch gutes Unterrichtsmanagement gesicherter verantwortlicher Umgang mit der Lernzeit der Schülerinnen und Schüler und ein angemessenes Anspruchsniveau gute Voraussetzungen für eine Verbesserung der Mathematikleistungen an den Schulen in Brandenburg sind.

Literaturverzeichnis

BAUMERT, J. (1989): Was bewirken Schulleiter? Empirische Befunde und theoretische Pespektiven. (Beiträge aus dem Forschungsbereich Schule und Unterricht des Max-Planck-Instituts für Bildungsforschung). Berlin.

BAUMERT, J., LEHMANN, R., LEHRKE, M., SCHMITZ, B., CLAUSEN, M., HOSENFELD, I., KÖLLER, O. & NEUBRAND, J. (1997): TIMSS - Mathematisch-naturwissen-schaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. Opladen.

BAUMERT, J., LEHMANN, R., LEHRKE, M., CLAUSEN, M., HOSENFELD, I., NEUBRAND, J., PATJENS, S., JUNGCLAUS, H. & GÜNTHER, W. (Hrsg.) (1998): Testaufgaben Mathematik TIMSS 7./8. Klasse (Population 2). (= Materialien aus der Bildungsforschung des Max-Planck-Instituts für Bildungsforschung, Band 60). Berlin.

BEATON, A. E., MARTIN, M. O., MULLIS, I. V. S., GONZALESZ, E. J., SMITH, T. A. & KELLY, D. A. (1996): Mathematics Achievement in the Middle School Years. IEA's Third International Mathematics and Science Study. Chestnut Hill, MA.

BIEBER, G. (1999): Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik. Mathe40- und QuaSUM-Testaufgaben im Vergleich (unveröffentlichtes Manuskript). Ludwigsfelde.

BORRIES, B. VON (1995): Das Geschichtsbewußtsein Jugendlicher. Eine repräsentative Untersuchung über Vergangenheitsdeutungen, Gegenwartswahrnehmungen und Zukunftserwartungen von Schülerinnen und Schülern in Ost- und Westdeutschland. Weinheim, München.

FEND, H. (1982): Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs. Weinheim, Basel.

FISCHER, G. H. & MOLENAAR, I. H. (1995): Rasch Models – Foundations, Recent Developments, and Applications. New York.

HAMBLETON, R. K., SWAMINATHAN, H. & ROGERS, H. J. (Hrsg.) (1991): Fundamentals of Item Response Theory. Newbury Park, CA.

HOFFMANN, L. & LEHRKE, M. (1986): Eine Untersuchung über Schülerinteressen an Physik und Technik. In: Zeitschrift für Pädagogik, 32. Jg., Nr. 2/86, S. 189-204.

HORSTKEMPER, M. (1987): Schule, Geschlecht und Selbstvertrauen. Eine Längsschnittstudie über Mädchensozialisation in der Schule. Weinheim, München.

LEHMANN, R. H., PEEK, R., PIEPER, I. & STRITZKY, R. VON (1995): Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen. Weinheim, Basel.

LEHMANN, R. H. & PEEK, R. (1997): Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Bericht über die Untersuchung im September 1996 (unveröffentlichter Forschungsbericht). Hamburg.

LEHMANN, R. H., BARTH, I., GÄNSFUß, R., LUTKAT, S., MÜCKE, S. & PEEK, R. (1999): Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik – QuaSUM. Zwischenbericht über die Untersuchung an Brandenburger Schulen im Juni 1999 (unveröffentlichter Forschungsbericht). Potsdam.

LEHMANN, R. H., GÄNSFUß, R. & PEEK, R. (1999): Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern an Hamburger Schulen – Klassenstufe 7. Bericht über die Untersuchung im September 1998 (unveröffentlichter Forschungsbericht). Hamburg.

MELZER, W. & STENKE, D. (1996): Schulentwicklung und Schulforschung in den ostdeutschen Bundesländern. In: H.-G. ROLFF, K.-O. BAUER, K. KLEMM & H. PFEIFFER (Hrsg.): Jahrbuch der Schulentwicklung. Daten, Beispiele und Perspektiven. Band 9. Weinheim, München. S. 307-337.

MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT DES LANDES BRANDENBURG (1999a): Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik (QuaSUM). Projektskizze. Potsdam.

MINISTERIUM FÜR BILDUNG, JUGEND UND SPORT DES LANDES BRANDENBURG (1999b): Schuldatenerhebung, Datengrundlage 1998/99. Stichtag: 15.09.98. Potsdam.

NEUNZIG, WALTER (1995): Methodisch-mediales Handeln im Lernbereich Mathematik. In: G. OTTO & W. SCHULZ (Hrsg.): Enzyklopädie Erziehungswissenschaft. Band 4: Methoden und Medien der Erziehung und des Unterrichts. Stuttgart, Dresden. S. 275.

ROEDER, P. M. & BAUMERT, J. (1997): BIJU. Bildungsverläufe und psychosoziale Entwicklung im Jugendalter. Skalenhandbuch (unveröffentlichter Bericht). Berlin.

ROLFF, H.-G. (1995): Wandel durch Selbstorganisation. Theoretische Grundlagen und praktische Hinweise für eine bessere Schule. Weinheim, München.

STURZBECHER, D. (1997): Jugend und Gewalt in Ostdeutschland. Lebenserfahrungen in Schule, Freizeit und Familie. Göttingen.

WEINERT, F. E. & HELMKE, A. (Hrsg.) (1995): Entwicklung im Grundschulalter. Weinheim.

WEIß, R. (1997): Grundintelligenztest Skala 2 (CFT 20) mit Wortschatztest (WS) und Zahlenfolgentest (ZF). 4. überarbeitete Auflage. Göttingen.

Glossar

Aggregieren, Aggregation, aggregierte Ebene, aggregierte Effekte

Das Aggregieren bezeichnet einen datentechnischen Vorgang, bei dem mehrere Fälle einer Gruppe zu einem neuen Fall zusammengefasst („aggregiert“) werden. Beispielsweise lassen sich in der vorliegenden Untersuchung die 3.266 Fälle (Schülerinnen und Schüler) der Jahrgangstufe 5 klassenweise zu 148 neuen Fällen (Klassen) oder zu 66 neuen Fällen (Schulen) aggregieren.

Alpha

Alpha (genauer: „Cronbachs Alpha“) ist eine Kennzahl für die interne Konsistenz einer ⇒ Skala. Werden inhaltlich zusammenpassende Aufgaben oder Fragen zu einer Skala zusammengefasst, z. B. 14 Fragen aus dem Schülerfragebogen zur „Schulzufriedenheit“, so gibt Alpha Auskunft darüber, wie gut sich die Antworten zu einer Gesamttendenz zusammenfügen. Alpha kann maximal den Wert 1 annehmen. Als Faustregel gilt, dass bei Skalen ab $\alpha = 0{,}75$ eine befriedigende und ab $\alpha = 0{,}85$ eine gute Skalenqualität vorliegt.

arithmetisches Mittel, arithmetischer Mittelwert, Durchschnittswert

⇒ Mittelwert.

Beta(-Gewicht)

⇒ Regressionsanalyse.

d*, Effektstärke *d*, standardisierte Effektstärke *d

ist ein standardisiertes Maß für Merkmalsunterschiede zwischen zwei Gruppen. Sie wird berechnet, indem die Differenz der ⇒ Mittelwerte der Gruppen durch die gemeinsame ⇒ Standardabweichung dividiert wird.

Design-Effekte, *Deff*

Bei der Ziehung der QauSUM-Stichprobe wurden nicht einzelne Schülerinnen und Schüler zufällig gezogen, sondern ganze Schulen. Damit verbunden ist ein Klumpeneffekt, auch Design-Effekt genannt: Schülerinnen und Schüler innerhalb einer Schule bzw. Klasse sind sich ähnlicher als Schülerinnen und Schüler verschiedener Schulen bzw. Klassen. Um diese Design-Effekte angemessen zu berücksichtigen, berechnet man einen Korrekturfaktor *Deff*. Liegt überhaupt kein Klunpeneffekt vor, nimmt *Deff* seinen Minimalwert 1 an. Erfahrungsgemäß überschreitet *Deff* selten 5; in der QuaSUM-Stichprobe ist dies nur bei den Realschulen der Fall. Formal bezeichnet *Deff* das Verhältnis der tatsächlich untersuchten Fälle zum „effektiven Stichprobenumfang“, also dem Äquivalent einer einfachen Zufallsstichprobe.

erklärte Varianz

⇒ Regressionsanalyse, ⇒ Varianzanalyse

Eta²*, Bestimmtheitsmaß *Eta²

ist die Maßzahl für die erklärte Varianz bei der ⇒ Varianzanalyse. *Eta²* bezeichnet den Varianzanteil, den man erhält, wenn man die Einzelwerte durch den jeweiligen Gruppenmittelwert ersetzt, daraus die Varianz berechnet und durch die ursprüngliche Varianz teilt; es variiert zwischen 0 und 1. Z. B. gibt das mit der Schulform verbundene *Eta²* für ein Leistungsmerkmal Auskunft darüber, welchen Anteil Schulformunterschiede (Mittelwertdifferenzen) an den Leistungsunterschieden zwischen den Schülerinnen und Schülern überhaupt haben.

Gewichtung

Wenn eine Stichprobe die tatsächliche Verteilung in der Grundgesamtheit (z. B. alle Neuntklässler in Land Brandenburg) nicht exakt abbildet (also verzerrt ist, was z. B. durch ⇒ Oversampling entstehen kann), gibt es die Möglichkeit, die tatsächliche Verteilung in der Grundgesamtheit durch Gewichtung wiederherzustellen. Hierbei werden unterrepräsentierte Fälle mit Werten über 1 gewichtet, überrepräsentierte Fälle mit Werten unter 1. Die Gesamtfallzahl in der Stichprobe bleibt dabei erhalten, es werden nur die Proportionen entsprechend den Verhältnissen in der Zielgruppe im Sinne einer Korrektur verschoben.

interne Konsistenz

⇒ Alpha.

intervallskalierte Variable

Wenn Gleichheit der Abstände zwischen zwei benachbarten Werten einer Variable angenommen werden darf, spricht man von einer intervallskalierten Variable.

Item

⇒ Variable

Koeffizient

Ein Koeffizient ist ein statistischer, ein mathematischer Kennwert. ⇒ Pearsons r ist z.B. ein Korrelationskoeffizient, d. h. ein statistisches Zusammenhangsmaß.

Korrelation, korrelieren

Zusammenhang zwischen zwei ⇒ Variablen (Merkmalen). Für ⇒ intervallskalierte Daten ist das Korrelationsmaß der Pearsonsche Produkt-Moment-Korrelationskoeffizient r (kurz „Pearsons r“ oder nur „r“). Er variiert zwischen minus eins und plus eins. Ein hohes negatives r besagt: Je höher das eine Merkmal ausgeprägt ist, desto niedriger das andere Merkmal, und je niedriger das eine Merkmal, desto höher das andere Merkmal. Ein hohes positives r besagt sinngemäß entsprechend: Je höhere Werte das eine Merkmal annimmt, desto höhere auch das andere (bzw. je niedriger, desto niedriger). Ein r nahe Null sagt aus, daß zwischen den beiden Merkmalen kein Zusammenhang besteht. r^2 gibt direkt die so ⇒ erklärte Varianz an.

Kriterium(s-Variable)

abhängige Variable, ⇒ Regressionsanalyse.

Lösungswahrscheinlichkeit

Die Lösungswahrscheinlichkeit einer Aufgabe gibt an, wie groß die Wahrscheinlichkeit ist, dass irgend ein Schüler bzw. irgendeine Schülerin diese Aufgabe löst. Die Lösungswahrscheinlichkeit wird mit dem Wert p (vom englischen probability) angegeben und liegt zwischen 0 und 1. Eine Lösungswahrscheinlichkeit von $p = 0{,}47$ beispielsweise besagt, dass „auf lange Sicht“ 47 Prozent der Schülerinnen und Schüler einer definierten Gruppe diese Aufgabe lösen.

Mittelwert

Kurzbezeichnung für den arithmetischen Mittelwert. Der Mittelwert darf berechnet werden, wenn mindestens ⇒ intervallskalierte Daten vorliegen. Er ist die Summe der Einzelwerte aller Fälle dividiert durch die Fallzahl.

multiple Regressionsanalyse

⇒ Regressionsanalyse.

multiples *R*, multiples *R*²

⇒ Regressionsanalyse.

Objektivität

Objektivität ist ein Gütekriterium für sozialwissenschaftliche Messungen. Objektivität in Schulleistungsuntersuchungen ist gegeben, wenn für alle Schülerinnen und Schüler gleiche Aufgabenstellungen, Bearbeitungszeiten, Erläuterungen der Aufgaben, Arbeitsmaterialien u. ä. gelten; die Auswertung und die Interpretation erfolgen nach klaren Kriterien, die unabhängig von der Person des Auswerters sind.

Oversampling

Die QuaSUM-Zufallsstichprobe wurde ergänzt durch die vollständige Einbeziehung von Schulen aus Schul- und Modellversuchen sowie von Schulen mit besonderere Prägung. Dies nennt man Oversampling. Um bei allgemeinen Analysen für das ganze Land Brandenburg diese Überrepräsentanz der Schulen aus dem Oversampling wieder auf ihren Anteil an den tatsächlichen Verteilungen im Land Brandenburg zu reduzieren, wurde der Datensatz ⇒ gewichtet.

***p*, *p*-Wert**

Abkürzung für ⇒ Lösungswahrscheinlichkeit.

Prädiktor(variable)

unabhängige Variable, ⇒ Regressionsanalyse.

Perzentil

Allgemeiner Begriff für eine prozentuale Unterteilung einer Merkmalsverteilung; z. B. bezeichnet das 20. Perzentil denjenigen Wert, der von 20 Prozent der Fälle erreicht oder unterschritten worden ist. Entsprechend wurde dieser Wert von 80 Prozent der Fälle überschritten.

Quartil

Quartile entstehen bei der Aufteilung einer Stichprobe hinsichtlich eines sortierten Merkmals (z. B. QuaSUM-Mathematiktest-Rohpunktwert nach ansteigender Punktzahl sortiert) in vier gleich starke Gruppen. Die Bestimmung der Quartile richtet sich dabei nach der Verteilung der Fälle auf das Merkmal (z. B. haben beim QuaSUM-Mathematiktest in der Jahrgangsstufe 9 25 Prozent der Schülerinnen und Schüler 45 und mehr Punkte insgesamt erreicht, oberstes Quartil, gleichbedeutend mit dem 75. ⇒ Perzentil).

r Abkürzung für Pearsons *r*, ⇒ Korrelation.

Random-Start-Equal-Verfahren

Verfahren zur Ziehung einer Zufallsstichprobe. Hierbei wird aus einer Liste, die alle Elemente der Grundgesamtheit enthält (in QuaSUM: Schulliste) mittels Zufallszahl das erste Element für die Stichprobe aus der Liste bestimmt („Random-Start“, z. B. das fünfte Element der Liste), und dann jedes x-te weitere gezogen („Equal-Steps“, z. B. jedes zwanzigste Element).

Regressionsanalyse

Die (multiple) Regressionsanalyse ist ein Analyseverfahren, das den Zusammenhang zwischen einer ⇒ intervallskalierten abhängigen (zu erklärenden) Variable (dem sogenannten Kriterium) und mehreren, ebenfalls intervallskalierten unabhängigen (erklärenden) Variablen (den sogenannten Prädiktoren) aufdeckt. Bei der Berechnung der Regressionsgleichung werden die ⇒ Korrelationen der Prädiktoren untereinander berücksichtigt. Die Maßzahl für den Zusammenhang zwischen allen Prädiktoren einerseits und dem Kriterium andererseits ist das „multiple R“. Es kann wie ⇒ Pearsons *r* interpretiert werden, kann allerdings keine negativen Werte annehmen. Das quadrierte multiple R entspricht der erklärten Varianz. Der eigenständige Beitrag jedes einzelnen Prädiktors (bei Konstanthaltung der anderen Prädiktoren) zur Aufklärung der Unterschiede im Kriterium wird mit den ⇒ Beta-Gewichten angegeben.

Reliabilität

Reliabilität ist ein Gütekriterium für sozialwissenschaftliche Messungen. Sie entspricht der Zuverlässigkeit einer Messung. Reliabel ist ein Test oder eine Skala, wenn nur geringe Messfehler auftreten. Es gibt verschiedene Methoden, die Reliabilität zu überprüfen, z. B. ⇒ Cronbachs Alpha.

Skala

1. Kurzbezeichnung für die Ausprägungen einer Einschätzskala (Ratingskala). Bei der Vorgabe der Antwortmöglichkeiten von 1 = „nein, trifft gar nicht zu“ bis 4 = „ja, trifft voll und ganz zu“ im Schülerfragebogen spricht man z. B. von einer vierstufigen Skala.
2. Inhaltlich zusammenpassende Einzelitems können, z. B. durch Aufsummieren oder Mittelwertbildung, zu einer Skala zusammengefasst werden. Ein Beispiel ist die Skala „Schulzufriedenheit“, die auf dem Wege der individuellen Mittelwertbildung über 14 Fragen aus dem Fragebogen für Schülerinnen und Schüler ermittelt wurde.

Standardabweichung, s

Die Standardabweichung ist ein sogenanntes Streuungsmaß, das für intervallskalierte Daten Auskunft darüber gibt, wie homogen oder heterogen eine Merkmalsverteilung ist. Je kleiner die Standardabweichung ist, desto enger gruppieren sich die Werte der einzelnen Fälle um den Mittelwert, je größer sie ist, desto weiter streuen sie um den Mittelwert.
Liegt eine Normalverteilung vor, so lässt sich über die Verteilung folgendes sagen: Im Bereich Mittelwert plus/minus eine Standardabweichung liegen etwa 68 Prozent der Fälle; im Bereich Mittelwert plus/minus zwei Standardabweichungen liegen etwa 95 Prozent der Fälle.

Streuung

⇒ Standardabweichung.

Validität

Validität ist ein Gütekritrium für sozialwissenschaftliche Messungen. Validität gibt die Gültigkeit eines Messinstruments, z. B. eines Tests, an. Ein Test ist valide, wenn er das misst, was er zu messen vorgibt. Validität lässt sich an Außenkriterien überprüfen; z. B. lassen sich die Testergebnisse des QuaSUM-Mathematiktests mit den Ergebnissen im Mathe40-Test validieren.

Variable

Ein erhobenes Merkmal, das mehrere Ausprägungen haben kann, wird als Variable bezeichnet, z. B. Alter oder erreichter Rasch-Skalenwert im Mathe40-Test. Die Ausprägungen nennt man auch Merkmalsausprägungen.

Varianz

Die Varianz ist das Quadrat der ⇒ Standardabweichung. Mathematisch ist die Varianz der Durchschnitt aus den quadrierten Abweichungen aller Einzelwerte vom Gesamtmittelwert.

Varianzanalyse

Dieses Analyseverfahren ist mit der ⇒ Regressionsanalyse verwandt. Auch hier geht es um die Aufklärung individueller Unterschiede in einer abhängigen Variable durch eine oder mehrere unabhängige Variablen. Der Hauptunterschied zur Regressionsanalyse besteht darin, dass die unabhängigen Variablen nicht ⇒ intervallskaliert zu sein braucht. Die Maßzahl für die erklärte Varianz heißt ⇒ Eta^2.