

Rainer H. Lehmann, Rainer Peek,  
Rüdiger Gänsfuß und Vera Husfeldt

# **LAU 9**

## **Aspekte der Lernausgangslage und der Lernentwicklung - Klassenstufe 9 -**

Ergebnisse einer längsschnittlichen  
Untersuchung in Hamburg



## **Vorwort**

Die Hamburger Schulbehörde setzt seit 1996 im Rahmen des Projekts *„Aspekte der Lernausgangslage und der Lernentwicklung - LAU“* eine Form des School-monitoring um, die auch weit über die Stadtgrenzen hinaus große Beachtung gefunden hat.

Im Mittelpunkt steht dabei eine Längsschnittstudie, mit der das Amt für Schule erreichte Lernstände, Lernentwicklungen und schulbezogene Einstellungen vom Beginn der Sekundarstufe I bis in die Vorstufe erheben lässt. Mit der wissenschaftlichen Leitung für die Erhebungen wurde Herr Prof. Dr. Dr. Rainer H. Lehmann, Humboldt Universität zu Berlin, beauftragt. Die LAU-Ergebnisse geben Auskunft über das Leistungsvermögen des Hamburger Schulwesens in bestimmten Bereichen und dienen der Behörde und den Schulen als externer Referenzrahmen für die geleistete Arbeit.

Damit werden Teile des neuen Steuerungsmodells umgesetzt, das einen Wechsel von einer Input- zu einer Output-Orientierung vollzieht, wie dies der Hamburger Staatsrat Herr Dr. h.c. Hermann Lange in seinem Artikel zur Qualitätssicherung in Schulen in „Die Deutsche Schule“, 91 Jahrgang 1999, Heft 2 ausführt. Diese neue Sichtweise betont die Eigenverantwortung der Schulen als Organisationseinheiten für die Qualität ihrer Arbeit im Rahmen von Vorgaben und fordert zugleich eine Rechenschaftslegung über die geleistete Arbeit. Das Wechselspiel zwischen der Eigenständigkeit der Schulen und der staatlichen Verantwortung wird für das Hamburger Schulwesen in dem Bericht *„Eigenständigkeit der Schulen in staatlicher Verantwortung“* der Behörde für Bildung und Sport beschrieben.

Die Ergebnisse der Erhebungen LAU 5 und LAU 7 haben zu Veränderungen im Hamburger Schulsystem geführt, z. B. zur Einführung von Springerklassen, zum Projekt PriMa, das der Qualitätsverbesserung des Mathematikunterrichts in der Grundschule dient, oder zur Einführung von Vergleichsarbeiten.

Nun liegt der Bericht über die Klassenstufe 9 vor. Er gibt mancherlei Anlass zum Nachdenken über weitere Veränderungen im Hamburger Schulwesen und leistet damit einen weiteren Beitrag zur Qualitätssicherung.

Dr. Ingrid Steinbach  
Projektleitung

## **Danksagung**

Dank gilt allen, die seit nunmehr gut sechs Jahren die Arbeitsgruppe an der Humboldt-Universität zu Berlin bei der wissenschaftlichen Auswertung des Hamburger Projekts „Aspekte der Lernausgangslage und der Lernentwicklung“ unterstützen, und denen, die an der erfolgreichen Durchführung der LAU-9-Erhebung mitgewirkt haben.

Im Amt für Schule hat Frau Dr. Steinbach als Projektleiterin in allen Phasen der Untersuchung wichtige Impulse für die konzeptionelle Entwicklung gegeben und war dabei stets eine kritische und konstruktive Partnerin für die Berliner Forschungsgruppe. Diesen Impulsen wie auch den Präzisierungen und Anregungen von Herrn Vieluf als Leiter des Referats für Bildungsberichterstattung, der die Studie aufmerksam begleitet hat, verdankt der jetzt vorgelegte Bericht wesentliche Ergebnisse. Frau Brose und Frau Nielsen haben als Projektmitarbeiterinnen in Hamburg wichtige Beiträge geleistet.

Die Hamburger 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, Lehrkräfte 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.

Frau Gebauer hat in Berlin für die Deutsch-Testkomponente „Textproduktion“ das Beurteilungsschema optimiert sowie den Auswertungsprozess koordiniert, kontrolliert und dokumentiert.

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

Rainer H. Lehmann  
wissenschaftlicher Leiter

## **Inhaltsverzeichnis**

<b>1</b>	<b>Ziele der Untersuchung</b>	<b>7</b>
<b>2</b>	<b>Anlage und Durchführung der Untersuchung</b>	<b>8</b>
2.1	Organisation der Erhebungen	8
2.2	Instrumentierung	9
2.3	Datenstruktur im Längsschnitt	17
<b>3</b>	<b>Fachleistungen und Problemlösekompetenz am Ende der Klassenstufe 8</b>	<b>21</b>
3.1	Methodische Vorüberlegungen	21
3.2	Analyse der fachbezogenen Leistungstests	25
3.2.1	Fachleistung Mathematik	27
3.2.2	Fachleistung Deutsch	34
3.2.2.1	Leseverständnis	34
3.2.2.2	Sprache	41
3.2.2.3	Rechtschreiben	45
3.2.2.4	Textproduktion	50
3.2.2.5	Zusammenfassung der Fachleistung Deutsch	54
3.2.3	Fachleistung Englisch	56
3.2.4	Fachleistung Latein	61
3.2.5	Fachleistung Französisch	64
3.2.6	Zusammenfassung der allgemeinen Fachleistung am Ende der Klassenstufe 8	67
3.3	Problemlösekompetenz	74
3.4	Zusammenfassung	77

<b>4</b>	<b>Entwicklung der Fachleistungen in den Klassenstufen 7 und 8</b>	<b>80</b>
4.1	Methodische Vorüberlegungen	80
4.2	Entwicklung der Fachleistung Mathematik	83
4.3	Entwicklung der Fachleistung Deutsch	96
4.4	Entwicklung der Fachleistung im Bereich des Fremdsprachenunterrichts	110
4.5	Zusammenfassung	119
<b>5</b>	<b>Individueller und sozialer Kontext der Schülerleistungen</b>	<b>124</b>
5.1	Kognitive Voraussetzungen	124
5.2	Alter und Geschlecht	128
5.3	Außerschulischer Hintergrund	139
5.4	Migrantenstatus	146
5.5	Lern- und schulbezogene Einstellungen der Schülerinnen und Schüler	158
<b>6</b>	<b>Schlussbetrachtung</b>	<b>163</b>

## **Literaturverzeichnis**

## **Glossar**

# 1 Ziele der Untersuchung

Die Untersuchung *Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9* ist Bestandteil einer Längsschnittstudie, die flächendeckend die Hamburger Fünftklässlerinnen und Fünftklässler des Schuljahres 1996/97 sowie die später in diesen Jahrgang eingetretenen Schülerinnen und Schüler in staatlichen Schulen mit allgemein bildenden Lehrplänen erfasst. Sie dokumentiert und analysiert deren jeweils erreichte Lernstände, Lernentwicklungen und schulbezogene Einstellungen vom Ende der Grundschulzeit bis in die Klassenstufe 9.

Die erste Erhebung der Längsschnittstudie im September 1996 hatte sich auf den Übergang von der Grundschule in die Sekundarstufe I und auf ausgewählte Aspekte der Lernausgangslage konzentriert, auf die der Unterricht an den weiterführenden Schulen aufbaut (vgl. LEHMANN & PEEK 1997). Die zweite Erhebung im September 1998 war auf die Ausprägung und Entwicklung von Fachleistungen, Problemlösekompetenz und schulbezogenen Einstellungen desselben Schülerjahrgangs in den Beobachtungsstufen der Haupt- und Realschulen bzw. der Gymnasien sowie in den Klassenstufen 5 und 6 der Gesamtschulen am Ende der Klassenstufe 6 – und damit zugleich auf die zu Beginn des Unterrichts in der Klassenstufe 7 erreichten Lernstände – gerichtet (vgl. LEHMANN, GÄNSFUß & PEEK 1999). Mit der LAU-9-Erhebung schließt sich nunmehr die Beschreibung und Analyse der weiteren Lernentwicklung der Hamburger Schülerinnen und Schüler bis zum Übergang in die Klassenstufe 9 an (vgl. LEHMANN 1999).

Für den hier vorgelegten Bericht steht die Aufklärung von Fragestellungen zu ***Fachleistungen und Problemlösekompetenz, Ausprägungen und Entwicklungen in den Klassenstufen 7 und 8*** im Vordergrund:

- Welche Lernstände in ausgewählten *fachbezogenen Lernbereichen* (Mathematik; Deutsch: Sprache, Leseverständnis, Rechtschreiben, Textproduktion; erste Fremdsprache Englisch bzw. Latein bzw. Französisch) und in dem *fachübergreifenden Bereich* Problemlösen haben die Schülerinnen und Schüler am Ende der Klassenstufe 8 erreicht?
- Wie haben sich die Lernstände der Schülerinnen und Schüler in den Klassen (bzw. in den Kursen der integrierten Gesamtschulen sowie den Lerngruppen der Integrierten Haupt- und Realschulen), Schulen und Schulformen bzw. Kursniveaus seit Eintritt in die Klassenstufe 7 entwickelt bzw. verändert?
- Welche Zusammenhänge bestehen zwischen der Lernentwicklung und außerschulischen Merkmalen, insbesondere der Bildungsnähe des Eltern-

hauses? Inwiefern zeigen sich Besonderheiten in der Lernentwicklung von Mädchen und Jungen und von Schülerinnen und Schülern aus Migrantenfamilien?

Diesen Zielsetzungen entsprechend ist der vorliegende Bericht aufgebaut: Nach Ausführungen zur Anlage und zur Durchführung der empirischen Untersuchung (Organisation der Erhebungen, Instrumentarium, Datenstruktur im Längsschnitt) in Kapitel 2 werden in den Kapiteln 3 und 4 Ergebnisse zu den am Ende der Klassenstufe 8 erreichten Lernständen sowie zur bisherigen Lernentwicklung der Schülerinnen und Schüler vorgestellt. In Kapitel 5 werden die Lernstände bzw. -entwicklungen unter Berücksichtigung des individuellen und sozialen Kontextes der Schülerinnen und Schüler thematisiert. Der Bericht endet mit einer Zusammenfassung der Ergebnisse in Kapitel 6.

## **2 Anlage und Durchführung der Untersuchung**

In diesem Kapitel werden Informationen zur Organisation der Erhebung gegeben (2.1), es werden die grundlegenden Konstruktionsprinzipien und der Aufbau der Erhebungsinstrumente vorgestellt (2.2), und es werden die erzielten Teilnahme- und Bearbeitungsquoten im Längsschnitt mitgeteilt (2.3).

### **2.1 Organisation der Erhebungen**

Bei der Organisation der Erhebung wurde auf erprobte und bewährte Verfahren aus der LAU-5- bzw. der LAU-7-Untersuchung zurückgegriffen.

Die allgemeine Steuerung der Untersuchung lag in der Verantwortung einer Begleitkommission (BEKO), die koordinierende und beratende Funktion gegenüber den Entscheidungsträgern im Amt für Schule und gegenüber der wissenschaftlichen Forschungsgruppe hat. Die Information der Schulen über die Ziele und Verfahren der Studie und die Organisation der Datenerhebung (Versorgung der Schulen mit Test- und Befragungsmaterialien; Beratung der Testleiterinnen und Testleiter in den Schulen) erfolgte über ein Planungsbüro im Amt für Schule.

Für die schulinterne Organisation war jeweils eine von der Schulleitung benannte Ansprechperson aus dem Kollegium verantwortlich. Die Tests und Befragungen wurden von (Klassen-) Lehrkräften, darunter auch Referendarinnen bzw. Referendaren, durchgeführt.

Die Datenerfassung, -aufbereitung und -auswertung sowie die Abfassung des Berichts über die Untersuchung lag in der Verantwortung der wissenschaftlichen Forschungsgruppe an der Humboldt-Universität zu Berlin.



## 2.2 Instrumentierung

Im Rahmen der Untersuchung wurden standardisierte Fachleistungstests, ein spezielles Verfahren zur Ermittlung von Problemlösekompetenzen, eine Aufgabe zur Textproduktion sowie ein Schüler- und ein Elternfragebogen eingesetzt. Diese Instrumente wurden ergänzt durch einen nonverbalen Test zum schlussfolgernden Denken. Darüber hinaus wurden Angaben aus den amtlichen Schülerakten entnommen.<sup>1</sup>

### Schulleistungstests

Der Bestimmung des erreichten Leistungsstandes dienten zwei standardisierte Schulleistungstests: der *Hamburger Schulleistungstest für achte und neunte Klassen – SL-HAM 8/9* (BEHÖRDE FÜR SCHULE, JUGEND UND BERUFSBILDUNG, AMT FÜR SCHULE, HAMBURG 2000) und die *Hamburger Schreibprobe für fünfte bis neunte Klassen – HSP 5-9* (MAY 1994).

Der *SL-HAM 8/9* umfasst Untertests zu den fachbezogenen Lernbereichen *Mathematik*, *Deutsch (Sprache, Leseverständnis und Rechtschreibwissen)* und erste Fremdsprache *Englisch* bzw. *Latein* bzw. *Französisch* sowie zu dem fachübergreifenden Bereich *Problemlösen*. Die Aufgaben der fachbezogenen Untertests wurden von Hamburger Lehrerteams entwickelt, mit der wissenschaftlichen Forschungsgruppe abgestimmt und nach einer breit angelegten Voruntersuchung für die Haupterhebung unter inhaltlichen, teststatistischen und auswertungsökonomischen Gesichtspunkten optimiert.

(1) Um die curricularen Anforderungen der unterschiedlichen Schulformen bzw. Kursniveaus mit dem *Mathematiktest* möglichst breit abzudecken, erhielt jede Schülerin bzw. jeder Schüler über einen gemeinsamen Kernbestand von 20 Aufgaben hinaus (Aufgaben der Schwierigkeitsstufe 3 in der folgenden Tabelle 1) weitere 20 bis 24 Aufgaben, die nach Schulform bzw. Kursniveau differenziert waren. In allen Testversionen wurden jeweils die drei Stoffgebiete

---

<sup>1</sup> Aufgrund datenschutzrechtlicher Vereinbarungen wurden Daten aus der LAU-5-Erhebung im Jahre 1999 gelöscht. Für Analysen im Rahmen der LAU-9-Untersuchung liegen allein den Schülerbögen entnommene Organisationsdaten über die Schullaufbahnen der Schülerinnen und Schüler sowie die individuellen Schülerwerte über Lernzuwächse in Mathematik, Sprache und Leseverständnis von LAU 5 nach LAU 7 vor (vgl. dazu LEHMANN, GÄNSFUß & PEEK 1999). Die Lernzuwachswerte gehen auf Daten zurück, die 1996 bzw. 1998 ermittelt wurden: Neben den Ergebnissen im *Hamburger Schulleistungstest für 4. und 5. Klassen – HST 4/5* (MIETZEL & WILLENBERG 2000) sind das die Ergebnisse der Schülerinnen und Schüler im *Hamburger Schulleistungstest für 6. und 7. Klassen – SL-HAM 6/7* (BEHÖRDE FÜR SCHULE, JUGEND UND BERUFSBILDUNG, AMT FÜR SCHULE, HAMBURG 1998).

Algebra, Arithmetik und Geometrie – in den Gymnasien und in den Mathematikkursen I der Gesamtschulen<sup>2</sup> zusätzlich Stochastik – berücksichtigt. Hinsichtlich der Anforderungen handelt es sich um typische Mathematikaufgaben der entsprechenden Schul- bzw. Kursform. Tabelle 1 verdeutlicht das Design. Durch vielfache Überschneidungen zwischen den fünf eingesetzten Testformen ist es möglich, allen Schülerinnen und Schülern unabhängig von der besuchten Schulform bzw. dem Kursniveau einen Lernstandswert auf einer einheitlichen Skala („Metrik“) zuzuweisen.

Tabelle 1 Design des Mathematiktests im *SL-HAM 8/9*

<i>Schwierigkeitsstufe 1</i> 10 Aufgaben	<i>Schwierigkeitsstufe 2</i> 10 Aufgaben	<i>Schwierigkeitsstufe 3</i> 20 Aufgaben	<i>Schwierigkeitsstufe 4</i> 10 Aufgaben	<i>Schwierigkeitsstufe 5</i> 14 Aufgaben
10 Aufgaben    10 Aufgaben    20 Aufgaben <i>Testform für Hauptschulklassen</i>				
5 Aufgaben    10 Aufgaben    20 Aufgaben    5 Aufgaben <i>Testform für Kurse II an Gesamtschulen</i>				
	10 Aufgaben    20 Aufgaben    10 Aufgaben <i>Testform für Realschulklassen</i>			
	5 Aufgaben    20 Aufgaben    10 Aufgaben    8 Aufgaben <i>Testform für Kurse I an Gesamtschulen</i>			
		20 Aufgaben    10 Aufgaben    14 Aufgaben <i>Testform für Gymnasialklassen</i>		

Zehn der insgesamt 64 Aufgaben wurden bereits 1998 in der LAU-7-Untersuchung eingesetzt; sie ermöglichen eine Verkopplung der Ergebnisse mit den Befunden aus den Untersuchungen des Längsschnitts. Insgesamt enthält der Mathematiktest 26 Algebra-, 18 Arithmetik-, 15 Geometrie- und 5 Stochastikaufgaben. Die Aufgaben wurden in Form von Mehrfachwahlantworten (Multiple-Choice-Items mit vier Antwortalternativen) gestellt. Die Bearbeitungszeit betrug einschließlich einer organisatorischen Einführung in den Test in jeder Schulform 45 Minuten.

(2) Der Untertest *Sprache* ist neben dem Untertest *Leseverständnis* Bestandteil des auf zwei Unterrichtsstunden angelegten Deutschtests. Der Deutschttest enthält neben einem Prosa- und einem Gebrauchstext mit insge-

<sup>2</sup> In integrierten Gesamtschulen werden Fachleistungskurse auf der höheren Anspruchsebene als Kurs I, Fachleistungskurse auf der niedrigeren Anspruchsebene als Kurs II bezeichnet.

samt zehn Leseverständnisaufgaben aus der LAU-7-Untersuchung und zehn – ebenfalls aus der LAU-7-Untersuchung stammenden – Sprachverständnisaufgaben drei Prosatexte, zwei Sachtexte und vier Gebrauchstexte, auf die sich weitere 76 Aufgaben zum Sprachverständnis und 62 Aufgaben zum Leseverständnis beziehen. Die Verteilung der neuen Texte bzw. der dazugehörigen Aufgaben erfolgte aufgrund von Voruntersuchungen nach demselben Prinzip wie beim Mathematiktest, wobei den Schülerinnen und Schülern neben einem gemeinsamen Kernbestand nach Schulform bzw. Kursniveau differenzierte Texte bzw. Aufgaben vorgelegt wurden. Bei einem ähnlichen Umfang von jeweils sieben oder acht Texten mit 42 (Hauptschulversion) bis 49 (Version für Deutschkurse I an Gesamtschulen) Sprachaufgaben bzw. 42 (Version für Realschulen und Deutschkurse I an Gesamtschulen) bis 48 (Version für Hauptschulen) Leseverständnisaufgaben unterscheiden sich die Varianten des Deutschtests in ihrem Anforderungsniveau. Die Aufgaben zum Untertest Sprache beziehen sich zum Teil auf formale Kenntnisse, also Grammatikwissen im Sinne eines adäquaten Umgangs mit Fachtermini. Daneben werden Aufgaben gestellt, die auf den reflexiven Umgang mit Sprache zielen. Als Ankeritems für die Ermittlung der Lernentwicklung im Bereich Sprache dienten die bereits erwähnten zehn Aufgaben, bei denen in fünf Fällen Bedeutungsunterschiede erkannt werden müssen und in den fünf übrigen Fällen die korrekte Verwendung von Artikeln, Pronomen, Nomen, Verben, Präpositionen und Konjunktionen sowie die regelgeleitete Konstruktion von Sätzen beherrscht werden müssen.

(3) Die Aufgaben zum Untertest *Leseverständnis* beziehen sich auf dieselben Texte wie die Aufgaben zum Sprachverständnis. Das nach Schulform bzw. Kursniveau differenzierende Design des Tests enthält insgesamt 72 – überwiegend im Multiple-Choice-Format gestellte – Aufgaben, von denen sich 28 auf Prosa-, 18 auf Sach- und 26 auf Gebrauchstexte beziehen. Den Aufgaben zum Leseverständnis liegt ein Testkonzept zugrunde, das bereits in der LAU-5- und in der LAU-7-Untersuchung leitend war: Es werden – auch innerhalb der schulform- bzw. kursniveauspezifischen Testformen – bewusst unterschiedlich schwierige Items berücksichtigt, um den gesamten Anforderungsbereich von einfacher Informationsentnahme bis hin zum eigenständigen schlussfolgernden bzw. interpretatorischen Umgang mit Texten zu erfassen. Die beiden Ankertexte mit ihren zehn Ankeraufgaben als Berechnungsgrundlage für Längsschnittanalysen stammen aus dem 1998 eingesetzten *Hamburger Schulleistungstest für sechste und siebte Klassen – SL-HAM 6/7*.

(4) Neben Kompetenzen in den Bereichen Sprachverständnis und Leseverständnis prüft der *SL-HAM 8/9* eine dritte wesentliche Dimension des Um-

gangs mit der deutschen Sprache: Kenntnisse der Rechtschreibung. Das **Rechtschreibwissen** wird hier in der anspruchsvollen Form der Fehlersuche in einem vorgegebenen Text erhoben. Der Bezugstext zum Freizeitverhalten von Jugendlichen mit insgesamt 285 Wörtern enthält 29 Wörter mit typischen Rechtschreibfehlern. Die Fehlschreibungen sollten unterstrichen werden, wobei in der Auswertung fälschlich unterstrichene Richtigschreibungen berücksichtigt werden. Der Untertest zum Rechtschreibwissen enthält keine schulform- bzw. kursniveaudifferenzierenden Sonderformen; derselbe Test wurde in allen Schulformen eingesetzt. Da Rechtschreibwissen kein Bestandteil der LAU-7-Untersuchung war, gibt es hier keine Ankeritems.<sup>3</sup>

(5) Der eine Unterrichtsstunde umfassende *Englishtest* erhebt Verständnisleistungen sowie Vokabel-, Rechtschreib- und Grammatikkenntnisse der Schülerinnen und Schüler, die Englisch als erste Fremdsprache gewählt haben. Darüber hinaus prüft er das Hörverständnis.<sup>4</sup>

Sämtlichen Schülerinnen und Schülern wurde im Rahmen des *Hörverständnistests* über Kassette eine ca. drei Minuten lange Sequenz vorgespielt, zu der sie anschließend vier in deutscher und drei in englischer Sprache vorgegebene Verständnisfragen (Multiple-Choice-Format) beantworten sollten. Der Text „The Interview“, in dem zwei Jugendliche in einem Radiointerview über Erfahrungen am Arbeitsplatz Auskunft geben, und die darauf bezogenen Aufgaben sind schulformübergreifend bewusst einfach gewählt, um allen Schülerinnen und Schülern gleichermaßen einen motivierenden Einstieg in den nachfolgenden C-Test zu geben.

Die Feststellung von Verständnisleistungen aus dem Textzusammenhang – genauer: der globalen fremdsprachlichen Kompetenz im Unterschied zu speziellen Kompetenzen – erfolgte im Rahmen eines *C-Tests*. Die Testkonstruktion beruht auf einer Variante des Cloze-Prinzips. Im Unterschied zum Cloze-Test werden – abgesehen vom ersten und letzten Satz – nicht ganze Wörter, sondern nur die Buchstaben der zweiten Hälfte jedes vierten Wortes im Text gelöscht, die dann von den Schülerinnen und Schülern zu vervollständigen ist. Der in der Literatur formulierten Anforderung (vgl. RAATZ & KLEIN-BRALEY

---

<sup>3</sup> Zum Zeitpunkt der LAU-7-Erhebung war gerade die Rechtschreibreform eingeführt worden. Auf den Einsatz eines Tests zum Rechtschreibwissen wurde deshalb verzichtet.

<sup>4</sup> Eine Stichprobe von 13 Gymnasial- und Gesamtschulklassen hat zusätzlich den Test „English in Use“ bearbeitet. Hier sollten die Schülerinnen und Schüler in einer Schulstunde Fragen und Antworten zu Alltagsthemen formulieren und einen Brief an eine Austauschschülerin bzw. einen Austauschschüler schreiben. Ergebnisse zu den in englischer Sprache formulierten Sätzen und zur Textaufgabe werden zu einem späteren Zeitpunkt veröffentlicht.

1983; GROTHJAHN 1995), dass ein C-Test – um die Reliabilität zu sichern – mindestens 100 Tilgungen aufweisen, aus verschiedenen kurzen Texten bestehen und exakt auswertbar sein soll, kommt der eingesetzte Englisch-C-Test so nach: Er enthält sechs zusammenhängende Texte mit jeweils 20 bis 25 Wortergänzungen, von denen zwei von allen Schülerinnen und Schülern bearbeitet, die übrigen vier nach Schulform bzw. Kursniveau differenziert eingesetzt wurden. Die insgesamt 142 Wortergänzungen sind eindeutig. Einer der beiden Texte aus dem Kernbereich war bereits Bestandteil des LAU-7-Instrumentariums. Die hierauf bezogenen 25 Aufgaben bilden die Ankeritems, die den Schlüssel für Aussagen zur Lernentwicklung bilden.

(6) Für die Schülerinnen und Schüler aus Gymnasien mit Latein als erster Fremdsprache ist ein ebenfalls eine Unterrichtsstunde umfassender *Lateintest* konstruiert worden. Der Test mit Multiple-Choice-Aufgaben sowie mit Wortergänzungs- und Markieraufgaben zu zwei lateinischen Texten erhebt Kenntnisse und Fähigkeiten in den Bereichen Lexik, Grammatik, Textverständnis und Stilistik. In einem zweiten Teil des Tests sind Aufgaben zum römischen Leben gestellt.

(7) Der ebenfalls einstündige *Französischtest* für die Schülerinnen und Schüler mit Französisch als erster Fremdsprache ist analog zum *Englischtest* konzipiert. Hier bearbeiteten die Schülerinnen und Schüler einen Hörtest mit insgesamt 14 offenen Aufgaben zu zwei Texten, die über Kassette präsentiert wurden, und einen C-Test mit insgesamt vier Texten und 118 Wortergänzungen.

Die Konzeption der fachbezogenen *SL-HAM 8/9-Leistungstests* war darauf ausgerichtet, durch die Autorenschaft von Hamburger Lehrkräften eine deutliche Nähe zur tatsächlichen Unterrichtspraxis herzustellen, also das sog. implementierte Curriculum möglichst gut abzubilden. Gleichzeitig sollte durch die Einbeziehung von Experten der Behörde für Schule, Jugend und Berufsbildung in die Begutachtung der Aufgaben die Lehrplanvalidität weitgehend gewährleistet sein.

(8) Im Zusammenhang der Untersuchung wird Problemlösen als übergreifende Schlüsselqualifikation verstanden, die es den Schülerinnen und Schülern ermöglicht, komplexe Anforderungen zu bewältigen. Der fachübergreifende **Problemlösetest** wurde vom Institut didaktik & diagnostik, Gesellschaft für angewandte Bildungsforschung mbH, entwickelt. Er ist auf eine Unterrichtsstunde Testzeit angelegt und umfasst insgesamt 20 gleich gewichtete Arbeitsschritte. Der Test besteht aus drei Teilen (sog. Projekten), die jeweils in fünf Arbeitsschritte mit einer oder mehreren zugehörigen Fragen untergliedert sind. Jeder Teil simuliert einen Projektauftrag aus dem schulischen Alltag („Schülerdisco“, „Einrichten eines Aufenthaltsraums“, „Energie- und Wassersparen in der Schule“). Über eine Serie von Informationen und Fragestellungen werden die Schülerinnen und Schüler durch den Ablauf eines jeden Projekts hindurchgeleitet: vom ersten Verstehen der Ausgangssituation und der Zielsetzung über das Beschaffen relevanter Informationen und die Koordination von Arbeitsschritten bis hin zur Analyse auftretender Fehler. Konzeptionell wichtig ist, dass die Schülerinnen und Schüler die drei Projekte jeweils als zusammenhängende Einheiten wahrnehmen. Deshalb beginnt jedes Projekt mit einer konkreten Situationsbeschreibung und einer Übersicht über die einzelnen Arbeitsschritte.<sup>5</sup>

Während die bisher vorgestellten Untertests des *SL-HAM 8/9* von sämtlichen Neuntklässlerinnen und Neuntklässlern obligatorisch bearbeitet wurden (wobei freilich jede Schülerin bzw. jeder Schüler nur einen der drei Fremdsprachentests zu bearbeiten hatte), wurden der zusätzliche Rechtschreibtest *Hamburger Schreibprobe für fünfte bis neunte Klassen (HSP 5-9)* und eine Aufgabe zur *Textproduktion* – ebenfalls verpflichtend – Auswahlstichproben gestellt.

---

<sup>5</sup> Im Anschluss an den Problemlösetest wurde ein Gruppentest eingesetzt, in dem jeweils drei Schülerinnen und Schüler bei gegenseitiger Berücksichtigung unterschiedlicher Interessen gemeinsam ein vorgegebenes Problem lösen sollten. Ergebnisse dieser kooperativen Planungsaufgabe werden zu einem späteren Zeitpunkt veröffentlicht.

Die *Hamburger Schreibprobe für fünfte bis neunte Klassen – HSP 5-9* (MAY 1994) zielt auf die Ermittlung des aktiven Rechtschreibkönnens der Schülerinnen und Schüler. Die *HSP 5-9* besteht aus 14 Einzelwörtern und fünf Sätzen. Die Bedeutung der Wörter und Sätze wird durch Illustrationen veranschaulicht. Die Schülerinnen und Schüler konnten in ihrem individuellen Tempo schreiben, nachdem ihnen die Wörter bzw. die Kurztex te vorgelesen worden waren. Als Zeitrahmen standen 20 Minuten zur Verfügung. Die Auswertung erfolgte aus auswertungsökonomischen Gründen nicht – wie vom Testkonzept her vorgesehen – auf der Ebene von Graphemen, sondern auf der Ebene von Wörtern. Insgesamt konnten maximal 57 Rohpunkte erreicht werden.

Die Aufgabe *Textproduktion*, für die 25 Minuten Bearbeitungszeit vorgesehen waren, besteht darin, einen Reklamationsbrief zu verfassen, der auf die unzulängliche Lieferung eines Computers bzw. auf die dieser Lieferung zugrunde liegende Zeitungsanzeige Bezug nimmt.

### **Test zum schlussfolgernden Denken**

Die Schulleistungstests wurden durch den nonverbalen *Culture Fair Intelligence Test – CFT 20* in seiner Kurzform ergänzt. Dieser Test zielt auf die Messung von Aspekten des schlussfolgernden Denkens, also der allgemeinen kognitiven Lernvoraussetzungen. Es handelt sich bei diesem Verfahren um einen bundesweit geeichten Test, was den Vergleich mit Altersnorm- bzw. Klassenstandardwerten ermöglicht (vgl. WEIß 1997). 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, Übereinstimmungen und 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 erfordert ca. 35 Minuten. Anders als die Teilnahme an den Schulleistungstests war die Bearbeitung des CFT 20 nicht obligatorisch, sondern ausdrücklich an das schriftliche Einverständnis eines Erziehungsberechtigten gebunden.

## Fragebögen

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. Er umfasst insgesamt 27 Aussagen zu den Themenbereichen *Schulzufriedenheit*, *leistungsbezogenes Selbstkonzept* und *Anstrengungsbereitschaft*, die jeweils auf einer vierstufigen Skala eingeschätzt werden sollen. Darüber hinaus erfasst der Schülerfragebogen zentrale Aspekte der außerschulischen Lern- und Lebenswelt der Schülerinnen und Schüler, vor allem Informationen zu ihrer Wohnsituation und zu Lernangeboten zu Hause, zu ihrer Muttersprache, zu ihrer außerschulischen Alltagsgestaltung, zur Nutzung lernförderlicher Angebote (z. B. Nachhilfekurse, Förderkurse, Zusatzangebote für Leistungsstarke, Büchereien / Bücherhallen) und zu ihren Hausaufgabenzeiten. Die Bearbeitung des Schülerfragebogens bedurfte des schriftlichen Einverständnisses einer oder eines Erziehungsberechtigten.

Der eingesetzte *Elternfragebogen* bezieht sich vor allem auf die Bildungsnähe des Elternhauses und den sozioökonomischen Hintergrund der Familie (Erwerbsstatus, Schul- und Ausbildungsabschlüsse der Eltern). Weitere Schwerpunkte des Elternfragebogens bilden Aussagen zu den Leistungserwartungen der Eltern an ihre Kinder und Erwartungen an ihre Schullaufbahn. Durch diese Informationen und durch die Angaben der Schülerinnen und Schüler in ihrem Fragebogen konnte in die nachfolgenden Analysen der erreichten Lernstände der Schülerinnen und Schüler auch ihr jeweiliges soziales Umfeld einbezogen werden.

## Schülerbogen

Als Quelle amtlich verfügbarer Informationen wurde wie 1996 und 1998 der in den Schulen vorhandene Schülerbogen genutzt. Neben wenigen Daten zur persönlichen Situation der Schülerinnen und Schüler (Alter, Geschlecht, Nationalität) ging es hier in erster Linie um Aspekte des Lernerfolgs und der Schullaufbahn in den letzten beiden Jahren vor der Datenerhebung in LAU 9: Wiederholen bzw. Überspringen einer Klasse, die Noten in den letzten Zeugnissen und – für Schülerinnen und Schüler aus integrierten Schulformen – die Kurszugehörigkeit in den Fächern Mathematik, Deutsch und Englisch. Hinzu kommen Organisationsdaten zur Rekombination der ehemaligen achten Klassen.

### 2.3 Datenstruktur im Längsschnitt



Diese dritte Datenerhebung fand zwischen dem 25. und dem 29. September 2000 statt, also unmittelbar nach dem Übergang in die Klassenstufe 9. Es wurde angestrebt, zu Beginn der Klassenstufe 9 Informationen über möglichst alle Schülerinnen und Schüler zu erhalten, die bereits an der Untersuchung in den Klassenstufen 5 und 7 teilgenommen hatten, und zusätzlich auch diejenigen in die Studie einzubeziehen, die aus verschiedenen Gründen (z. B. Zuzug, Klassenwiederholung) 1996 bzw. 1998 noch nicht teilgenommen hatten. Während die staatlichen Haupt- bzw. Realschulen, Gymnasien und Gesamtschulen obligatorisch in die Längsschnittstudie eingebunden sind, wurde den staatlichen Sonderschulen und den Schulen in privater Trägerschaft wie in LAU 5 und in LAU 7 das Angebot gemacht, sich an der LAU-9-Untersuchung zu beteiligen. Von diesem Angebot haben neun Privatschulen und eine Sonderschule Gebrauch gemacht.<sup>6</sup> Die folgende Tabelle 2 gibt schulformspezifisch an, wie viele Schulen, Klassen und Schülerinnen bzw. Schüler in die Untersuchung einbezogen wurden, für wie viele Schülerinnen und Schüler aus Klassen bzw. Schulen also mindestens Angaben aus den Schülerbögen vorliegen.

Tabelle 2 Datenstruktur der Erhebungen: Anzahl der Schulen, Klassen sowie Schülerinnen und Schüler, nach Schulform

	Schulen	Klassen	Schülerinnen und Schüler
Hauptschule	65	73	1.478
Realschule		80	2.070
Gymnasium	70	196	4.660
integrierte Haupt- und Realschule (IHR-Schule)	15	31	709
Gesamtschule	integriert	37	164
	kooperativ <sup>7</sup>	1	6
<i>insgesamt</i>	<i>188</i>	<i>550</i>	<i>12.620</i>

Die nachfolgende Tabelle 3 zeigt schulformbezogen die Teilnahmequoten für die Bestandteile der Untersuchung, die sich unmittelbar an die Schülerinnen

<sup>6</sup> Die Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Integrationsklassen, für die zwar Daten vorliegen, aber jeweils individuelle Bedingungen gelten, bleiben im vorliegenden Bericht unberücksichtigt. Ebenso werden die Ergebnisse der Sonderschulklasse aufgrund fehlender Vergleichbarkeit nicht ausgewiesen.

<sup>7</sup> Eine der einbezogenen Gesamtschulen arbeitet nicht integriert, sondern ist eine kooperative Gesamtschule. Aus dieser Schule sind eine Klasse der Schulform Hauptschule, drei Klassen der Schulform Realschule und zwei Klassen der Schulform Gymnasium zuzuordnen. Sie werden im Folgenden unter diesen Schulformen berücksichtigt, sind also nicht bei den Gesamtschulen aufgeführt.

und Schüler bzw. an ihre Eltern richteten. Die Bearbeitung der Schulleistungstests war verbindlich. Die Beantwortung des Schülerfragebogens und die Bearbeitung des CFT 20 waren davon abhängig, ob die Erziehungsberechtigten vorab schriftlich ihr Einverständnis gegeben hatten. Die Bearbeitung des Elternfragebogens war freiwillig.

Tabelle 3 Bearbeitungsquote der Instrumente (Schülerinnen und Schüler, Eltern), nach Schulform (in Prozent)

	<i>SL HAM 8/9</i>				
	<b>Mathematik</b>	<b>Leseverständnis</b>	<b>Sprache</b>	<b>Rechtschreibwissen</b>	<b>Englisch *</b>
Hauptschule	86,3	94,9	94,9	91,3	82,1
Realschule	92,8	97,7	97,7	95,1	90,9
Gymnasium	94,2	98,2	98,2	96,8	95,6
IHR-Schule	89,7	95,8	95,8	93,2	87,7
Gesamtschule	82,7	90,5	90,5	87,1	80,2
<i>insgesamt</i>	<i>89,4</i>	<i>95,3</i>	<i>95,3</i>	<i>92,8</i>	<i>88,1</i>

\* Die Quoten beziehen sich ausschließlich auf die 12.313 Schülerinnen und Schüler, die ab Klassenstufe 5 Englisch als erste Fremdsprache gewählt haben.

Tabelle 3 Fortsetzung

	<i>SL HAM 8/9</i>			<b>HSP 5-9 ***</b>	<b>Textproduktion ***</b>
	<b>Latein **</b>	<b>Französisch **</b>	<b>Problemlösungen</b>		
Hauptschule	---	---	88,8	86,7	86,7
Realschule	---	---	93,2	89,3	91,0
Gymnasium	94,9	93,0	95,5	93,7	94,1
IHR-Schule	---	---	91,0	86,7	87,5
Gesamtschule	---	100	85,0	82,5	81,8
<i>insgesamt</i>	<i>94,9</i>	<i>94,9</i>	<i>91,0</i>	<i>89,3</i>	<i>89,5</i>

\*\* Die Quoten beziehen sich ausschließlich auf die 256 Schülerinnen und Schüler, die ab Klassenstufe 5 Latein bzw. auf die 51 Schülerinnen und Schüler, die ab Klassenstufe 5 Französisch als erste Fremdsprache gewählt haben.

\*\*\* Für die *HSP 5-9* und die Aufgabe *Textproduktion* umfasst die angestrebte Stichprobe 1.650 Schülerinnen und Schüler aus 71 neunten Klassen, die schulformbezogen ausgewählt wurden.

Tabelle 3 Fortsetzung

	<b>CFT 20</b>	<b>Schülerfragebogen</b>	<b>Elternfragebogen</b>
Hauptschule	73,7	78,4	69,1
Realschule	85,4	88,0	80,3
Gymnasium	91,1	91,9	87,7
IHR-Schule	78,3	83,4	76,6
Gesamtschule	64,9	66,7	63,8
<i>insgesamt</i>	<i>79,7</i>	<i>81,8</i>	<i>76,7</i>

Angesichts dieser Rücklaufquoten kann insgesamt von einer soliden Datengrundlage gesprochen werden, auf die sich die weiteren Analysen stützen. Wie ebenfalls erkennbar ist, war (abgesehen von dem Sonderfall Französisch) die Mitwirkungsbereitschaft an den Gymnasien am höchsten und an den Gesamtschulen am niedrigsten.<sup>8</sup> Der Umstand, dass die Untersuchung als Vollerhebung konzipiert worden ist, ermöglicht es, Merkmalsausprägungen am Ende der zuletzt besuchten Schulklassen zu rekonstruieren. Dieses Verfahren wurde bereits bei den Erhebungen 1996 für den Übergang von der Primarstufe in die Sekundarstufe I und 1998 für den Übergang von den Beobachtungsstufen bzw. von der Klassenstufe 6 an Gesamtschulen in die Klassenstufe 7 eingesetzt. Die sich teilweise ergebenden Veränderungen in der Zusammensetzung von Klassen bzw. Kursen bringen nur geringfügige Auswertungsprobleme im Längsschnitt mit sich, weil in den meisten Fällen die Organisationsdaten aus den Schülerbögen den Abgleich zwischen Klassen vom Beginn der Klassenstufe 5 bis zum Beginn der Klassenstufe 9 erlauben. Da jedoch ein nicht unbeachtlicher Teil der Schülerinnen und Schüler zwischen den Klassenstufen 5 und 8 entweder eine Klasse wiederholt oder – seltener – übersprungen hat oder aus Hamburg weggezogen ist, während dafür andere in der Population neu aufgetaucht sind (z. B. Kinder aus zugezogenen Familien, Wiederholer aus höheren bzw. Springer aus unteren Klassen), und vor dem Hintergrund, dass sich

<sup>8</sup> Die nach Schulform unterschiedlichen Rücklaufquoten, die zudem noch zwischen den einzelnen Untertests variieren, haben zu einer geringfügigen Verschiebung der Gewichte geführt. So beziehen sich – statt eines Gymnasialanteils von 36,8 Prozent laut amtlicher Schulformzuordnung – 39,9 Prozent der allgemeinen Fachleistungsdaten auf das Gymnasium. Umgekehrt stammen nur 26,5 Prozent dieser Fachleistungsdaten von Gesamtschulen, obwohl der Gesamtschulanteil in Hamburg 29,3 Prozent umfasst. Um unterschiedliche Gewichtungsfunktionen zu vermeiden, wurde für diesen Bericht gänzlich auf eine Gewichtung verzichtet, zumal die genannten Abweichungen keinen Einfluss auf Mittelwertvergleiche zwischen den Schulformen haben und sich nur geringfügig auf andere Analysen auswirken.

Sonderschulen, Privatschulen und siebenstufige Gymnasien 1998 oder 2000 einmalig oder erstmals an der Untersuchung beteiligt haben, war im Abgleich der Daten im Längsschnitt mit einigen Schwankungen zu rechnen (vgl. dazu vor allem die Abschnitte 3.1 und 4.1). In den Untersuchungen zu Beginn der Klassenstufen 5, 7 und 9 wurden zusammen 16.362 verschiedene Schülerinnen und Schüler erfasst. Darüber, wie viele zu welchen Erhebungszeitpunkten einbezogen werden konnten, gibt die folgende Tabelle 4 Auskunft.

Tabelle 4 Einbezogene Schülerinnen und Schüler zu den drei Erhebungszeitpunkten 2000 (LAU 9), 1998 (LAU 7) und 1996 (LAU 5)

LAU 9	LAU 7	LAU 5	Anzahl	Prozent
ja	ja	ja	9.864	60,3 %
ja	ja	nein	1.210	7,4 %
ja	nein	nein	1.574	9,6 %
nein	ja	ja	1.756	10,7 %
nein	ja	nein	479	2,9 %
nein	nein	ja	1.479	9,0 %
<i>insgesamt</i>			<i>16.362</i>	<i>100 %</i>

Tabelle 5 schließlich gibt schulformspezifisch an, wie viele Schülerinnen und Schüler aus der LAU-9-Untersuchung in LAU 7 bzw. in LAU 5 in derselben oder in einer anderen Schulform unterrichtet wurden bzw. in einer der beiden Vorgängeruntersuchungen nicht einbezogen waren.

Tabelle 5 Schülerinnen und Schüler im Längsschnitt, nach Schulform

Schülerinnen und Schüler in LAU 9		Schulform in LAU 7			Schulform in LAU 5		
		dieselbe	eine andere	nicht in LAU 7	dieselbe	eine andere	nicht in LAU 5 und 7
Hauptschule	1.478	918	307	253	881	109	488
Realschule	2.070	1.342	291	437	874	450	746
Gymnasium	4.660	3.994	64	602	3.561	176	923
IHR-Schule	709	555	68	86	447	101	161
Gesamtschule	3.703	3.359	153	191	2.945	308	450
<i>insgesamt</i>	<i>12.620</i>	<i>10.168</i>	<i>883</i>	<i>1.569</i>	<i>8.708</i>	<i>1.144</i>	<i>2.768</i>
<i>insgesamt (in %)</i>	<i>100 %</i>	<i>80,6 %</i>	<i>7,0 %</i>	<i>12,4 %</i>	<i>68,8 %</i>	<i>9,1 %</i>	<i>22,0 %</i>

### **3 Fachleistungen und Problemlösekompetenz am Ende der Klassenstufe 8**

#### **3.1 Methodische Vorüberlegungen**

Die Erhebungen zu den Fachleistungen und der Problemlösekompetenz fanden am Anfang der Klassenstufe 9 statt, zu einem Zeitpunkt also, zu dem seit Ende der Klassenstufe 8 keine nennenswerten Lernfortschritte stattgefunden hatten. Der Umstand, dass es sich bei LAU 9 – wie bereits bei LAU 5 und LAU 7 – um eine Vollerhebung handelt, gestattet es in den meisten Fällen, die daran beteiligten Schülerinnen und Schüler derjenigen Schulform zuzuordnen, die sie ausweislich der Schülerakten in der Klassenstufe 8 besucht hatten. Wo im Folgenden nicht ausdrücklich anderes vermerkt ist, beziehen sich die Analysen auf diese Zuordnung. Dabei wird zwischen folgenden Kategorien unterschieden:

- Hauptschulklassen der traditionellen Haupt- und Realschulen,
- Realschulklassen der traditionellen Haupt- und Realschulen,
- Gymnasialklassen,
- H-Kurse der Integrierten Haupt- und Realschulen,
- R-Kurse der Integrierten Haupt- und Realschulen,
- Kurse II der Gesamtschulen,
- Kurse I der Gesamtschulen.

Danach werden z. B. Schülerinnen und Schüler, die in der Klassenstufe 8 das Gymnasium besucht haben, zu Beginn der Klassenstufe 9 aber an einer Realschule angetroffen und getestet wurden, dem Gymnasium zugeordnet. Allgemein gilt: *Alle Aussagen, die in diesem Bericht über Unterschiede zwischen den Schulformen bzw. Kursniveaus getroffen werden, beziehen sich auf die Schulform- bzw. Kursniveauezugehörigkeit in der Klassenstufe 8.* Der Anlage der Studie entsprechend können dabei nur diejenigen Schülerinnen und Schüler berücksichtigt werden, die im Herbst 2000 eine Klasse 9 einer Hamburger Schule besucht haben. Unschärfen, die sich aus einer von der genannten Einteilung abweichenden Schulform- oder Kursniveauezugehörigkeit zu Beginn der Klassenstufe 7 ergeben, werden in Kauf genommen.

Generell haben Schulformwechsel, aber auch Klassenwiederholung, Klassenüberspringen und regionale Mobilität der Familien, in geringerem Umfang auch Datenausfälle durch Nichtteilnahme einer Schule, einer Klasse oder eines Kurses Konsequenzen für die Interpretation der nachfolgend vorgelegten Befunde. So können die im Folgenden berichteten Werte nicht ohne Weiteres als Indikatoren für die Lernausgangslage der Klassenstufe 9 einer Schulform interpretiert werden, weil sowohl Schulformwechsel als auch Klassenwiederholung

am Ende der Klassenstufe 8 zu einer systematischen Anhebung der Lernstände in den jeweils neu konstituierten Lerngruppen führt. Eine genaue quantitative Erfassung der Auswirkungen der Klassenwiederholung kann hier nicht vorgelegt werden, weil die dazu erforderliche Unterscheidung zwischen den verschiedenen möglichen Gründen für das völlige Ausscheiden eines Schülers oder einer Schülerin aus der Untersuchungsgruppe bei der gegebenen Datenlage nur teilweise möglich ist. Eindeutig ist, dass die Abgänge aus der Untersuchungsgruppe, bezogen auf die letzte bekannte Schulform, regelmäßig und wie zu erwarten eine negativ ausgelesene Gruppe darstellen, die Zugänge hingegen wenigstens innerhalb der ersten zwei Jahre im Bezugsrahmen der neuen Lerngruppe eher durchschnittliche Leistungen erbringen. Dagegen ist für diejenigen, bei denen ein Schulformwechsel ohne gleichzeitige oder anschließende Klassenwiederholung erfolgt ist, eine Abschätzung der Effekte möglich. Exemplarisch soll dies für die drei traditionellen Schulformen<sup>9</sup> gezeigt werden, und zwar unter Vorgriff auf einen Index für die allgemeine Fachleistung, der im Abschnitt 3.1.6 diskutiert wird (vgl. Tabelle 6).

Tabelle 6 Allgemeine Fachleistung der Teilnehmerinnen und Teilnehmer an allen drei Erhebungen (1996, 1998 und 2000) nach Schulform und Schulformwechsel, Anfang Klassenstufe 9 (traditionelle Schulformen); Effektstärke der Differenz zwischen Schülerinnen und Schülern ohne Schulformwechsel und der Gesamtgruppe

Ursprüngliche Schulformzugehörigkeit	Schulformwechsel					Effektstärke <i>d</i>	Anteil Schulformwechsler
	nein		ja		<i>insgesamt</i>		
	$\bar{x}$	<i>s</i>	$\bar{x}$	<i>s</i>			
Hauptschule	38,9	4,7	44,0	8,0	39,1	-0,04	4,0 %
Realschule	47,7	5,1	45,6	7,0	47,1	0,11	29,2 %
Gymnasium	58,1	5,3	49,1	5,8	56,7	0,22	15,7 %

An den in der Tabelle aufgezeigten Werten wird sichtbar, dass die Effekte des Schulformwechsels in den Klassenstufen 5 bis Anfang 9 in den drei traditionellen Schulformen recht unterschiedlich sind. Die Schülerinnen und Schüler von Hauptschulen können entsprechend der Anlage der vorliegenden Untersu-

<sup>9</sup> Entsprechende Analysen für die integrierten Schulformen müssten erheblich komplexer angelegt werden, weshalb in diesem Bericht darauf verzichtet wird. Ohnehin ist Schulformwechsel hier selten: 4,7 Prozent der zu Anfang der Klassenstufe 7 in Integrierten Haupt- und Realschulklassen unterrichteten Schülerinnen und Schüler (ohne Klassenwiederholer und Wegzüge) wurden zu Anfang der Klassenstufe 9 in einer anderen Schulform angetroffen, und 94,5 Prozent der zu Anfang der Klassenstufe 5 an Gesamtschulen Unterrichteten (wiederum ohne Klassenwiederholer und Wegzüge) befanden sich auch am Anfang der Klassenstufe 9 noch in dieser Schulform, was einer Schulformwechslerquote von 5,5 Prozent entspricht.

chung – die Förderschulen sind nicht erfasst – im Hamburger System nur „auf- oder umsteigen“. 4,0 Prozent waren davon betroffen, von denen 60 Prozent den Weg in die Realschule gefunden haben und die übrigen entweder auf eine Integrierte Haupt- und Realschule oder auf eine Gesamtschule übergewechselt sind. Dementsprechend handelt es sich bei diesen Schulformwechslern überwiegend um eine positiv ausgelesene Gruppe, die im Mittel im Jahr 2000 eine Fachleistung gezeigt hat, die um mehr als eine Standardabweichung (5,1 Skalenpunkte) über dem durchschnittlichen Hauptschulniveau lag. Auffällig ist zugleich die große Streuung der Leistungen in dieser (insgesamt sehr kleinen) Gruppe. Die durch den Schulformwechsel bedingte Verminderung (und Homogenisierung) der Fachleistung macht sich wegen der kleinen Zahl der davon Betroffenen in einer nur geringen Absenkung des Durchschnitts (von 39,1 auf 38,9 Skalenpunkte) bemerkbar.

Die Schulformwechsler aus den Realschulen stellen hingegen eine im Durchschnitt eher leistungsschwache Gruppe innerhalb dieser Schulform dar. Das liegt vor allem daran, dass unter den Schulformwechslern die Rückstufungen in die Hauptschule mit 46,6 Prozent deutlich zahlreicher sind als die Übergänge ans Gymnasium mit 38,7 Prozent. Die restlichen 14,7 Prozent sind an Integrierte Haupt- und Realschulen und Gesamtschulen übergegangen. Angesichts der divergenten Gründe für eine Abwanderung aus der Realschule ist die starke Leistungsstreuung in der Gruppe der Schulformwechsler erwartungskonform. Zugleich wird in dieser Schulform eine doppelte Tendenz zur Leistungshomogenisierung in Form des Schulformwechsels sichtbar. Wegen der Dominanz der Übergänge in Hauptschulklassen ergibt sich zugleich eine Anhebung des mittleren Leistungsniveaus (von 47,1 auf 47,7 Skalenpunkte).

Im Gegensatz dazu gibt es für die Gymnasiastinnen und Gymnasiasten keine Aufwärtsmobilität im Bildungssystem durch Schulformwechsel. Die insgesamt 15,7 Prozent Betroffenen sind in den seltensten Fällen in Hauptschulklassen gelangt (6,9 Prozent). 55,2 Prozent der Schulformwechsler wurden im Jahr 2000 in einer Realschulklasse angetroffen, die restlichen an einer Integrierten Haupt- und Realschule oder einer Gesamtschule. Hier hat also eine Homogenisierung durch das Ausscheiden tendenziell besonders leistungsschwacher Schülerinnen und Schüler stattgefunden, was sich auch in der deutlich höheren Fachleistung der am Gymnasium Verbliebenen gegenüber den zwischen 1996 und 2000 an eine andere Schulform Gewechselten zeigt (58,1 gegenüber 49,1 Skalenpunkten).

Nun stehen aber hinter den aus Tabelle 6 abzulesenden Differenzen nicht nur Ausleseeffekte, sondern auch kumulierte Einflüsse unterschiedlicher Bil-

dungsbiografien, die zusätzlich durch den Unterricht in einer neuen Lerngruppe, in einigen Fällen seit mehr als drei Jahren, bedingt sein können. Hinzu kommen, wie bereits erwähnt, vor allem die Auswirkungen von Klassenwiederholungen, die jedoch nur über die dadurch entstandenen Zugänge erfasst sind. Für die mit negativen Ausleseeffekten verbundenen Abgänge liegen im Gegensatz dazu keine Leistungsdaten vor. Um dennoch zumindest eine partielle Bilanz der Verschiebungen für den hier maßgeblichen Untersuchungszeitraum, also die Klassenstufen 7 und 8, vorzulegen, soll an dieser Stelle – wiederum nur für die drei traditionellen Schulformen – auf die Differenzen zwischen Lernstandswerten, die auf die Schulformzugehörigkeit zum Ende der Klassenstufe 8 bezogen sind, und Angaben zur Lernausgangslage zu Beginn der Klassenstufe 9 eingegangen werden. Dabei sind die Einflüsse von Klassenwiederholung, Schulformwechsel, regionaler Mobilität etc. miteinander verwoben. Hier interessiert allerdings der Gesamteffekt, der sich mit der Neuzusammensetzung der Lerngruppen einstellt (vgl. Tabelle 7). Es sei dabei nochmals betont, dass für Klassenwiederholer der Klassenstufen 7 und 8 keine Leistungsdaten vorliegen, so dass die Lernstandswerte für das Ende der Klassenstufe 8 überschätzt sind, und zwar umso mehr, je höher die Wiederholerquote ist. Diese wird daher hier mit aufgelistet.<sup>10</sup>

Wie man in Tabelle 7 sieht, sind die Verschiebungen, die sich im Zuge des Übergangs von der Klassenstufe 8 nach Klassenstufe 9 ergeben, insgesamt – zumindest auf den ersten Blick betrachtet – recht gering. Allerdings sind dabei die Abgänge zu berücksichtigen, die überwiegend durch Nichtversetzung oder (vor allem wohl an der Hauptschule) sogar durch Verlassen der Schule ohne Schulabschluss bedingt sein werden, auch wenn in den Quoten der Abgänge z.B. auch der leistungsneutrale Wohnortwechsel enthalten ist. In Wirklichkeit sind die erreichten Lernstände zum Ende der Klassenstufe 8 also niedriger als die in diesem Kapitel berichteten. Dies gilt es im Folgenden zu beachten.

---

<sup>10</sup> Die Angaben in Abschnitt 5.2 weichen von den hier verwendeten zum Teil ab, weil hier für die drei Erhebungen die Schulformzugehörigkeiten aus den Testprotokollen, dort aber die Wiederholerquoten aus den Schülerakten rekonstruiert worden sind. Letzteres hat den Nachteil, dass insbesondere an den Hauptschulen die vorzeitigen Abgänge nicht erfasst sind, dagegen den Vorteil, dass bei den jeweiligen Zugängen zwischen Klassenwiederholungen und Wohnortwechsel unterschieden werden kann.



Tabelle 7 Allgemeine Fachleistung der Teilnehmerinnen und Teilnehmer an der Erhebung 2000 nach Schulform: Werte für die rekombinierten Lerngruppen der Klassenstufe 8 und der zu Anfang der Klassenstufe 9 neu konstituierten Lerngruppen im Vergleich (traditionelle Schulformen); schulformspezifische Quoten der Abgänge durch Verlassen der Schule ohne Abschluss, Klassenwiederholung, Springen, Wohnortwechsel, Testausfall

	Schulformwechsel				Effektstärke $d$	Anteil der Abgänge
	Ende Klassenstufe 8		Beginn Klassenstufe 9			
	$\bar{x}$	$s$	$\bar{x}$	$s$		
Hauptschule	39,6	5,2	39,8	5,3	0,04	32,8 %
Realschule	48,0	5,4	48,6	5,3	0,10	25,8 %
Gymnasium	57,4	5,5	57,5	5,5	0,02	13,1 %

Man kann diesen Ansatz auf der Basis der vorhandenen Daten noch weiter verfolgen und versuchen, die methodisch bedingte Überschätzung der Lernstände zum Ende der Klassenstufe 8 (und damit der Lernentwicklung in den Klassenstufen 7 und 8) wenigstens näherungsweise quantitativ zu bestimmen. Dies wird zu Beginn des Kapitels 4 geschehen.

### 3.2 Analyse der fachbezogenen Leistungstests

Es war das Ziel, für sämtliche getesteten Schülerinnen und Schüler auf der Grundlage jeweils *aller* zur Verfügung stehenden schulform- bzw. kursniveauspezifischen Aufgabenlösungen die erreichten Lernstände insgesamt zu bestimmen. Im Rahmen der probabilistischen Testtheorie (Item Response Theory; vgl. FISCHER & MOLENAAR 1995; HAMBLETON, SWAMINATHAN & ROGERS 1991) wurden die Aufgaben der fachbezogenen Untertests *Mathematik*, *Deutsch: Sprache*, *Deutsch: Leseverständnis* und *Englisch: C-Test* im *SL-HAM 8/9* unter Nutzung des einparametrischen Rasch-Modells skaliert. Im Unterschied zu klassischen Methoden der Testauswertung wird bei diesem Verfahren nicht vorausgesetzt, dass alle gestellten Aufgaben gleich schwierig sind. Vielmehr wird das Leistungsniveau (die ‚Fähigkeit‘) einer Schülerin oder eines Schülers unter Berücksichtigung der Schwierigkeit der gelösten Aufgaben ermittelt.<sup>11</sup> Das Verfahren erlaubt es,

<sup>11</sup> Eine gut nachvollziehbare Erklärung der Raschskalierung findet sich bei BAUMERT, BOS & LEHMANN (2000, S. 60 ff.).

- die Einschlägigkeit von Testaufgaben für die theoretisch bestimmte Fähigkeitsdimension empirisch zu prüfen und ungeeignete Items auszusondern,
- den Schwierigkeitsgrad 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, wobei die Festsetzung der Aufgabenschwierigkeiten 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 für die einzelne Schülerin bzw. für den einzelnen Schüler bestimmbar Schwierigkeit gelöst werden, können für alle Aufgaben und Fähigkeitsgruppen die Lösungswahrscheinlichkeiten berechnet werden.

Nachdem die Aufgabenschwierigkeiten und die Schülerfähigkeiten innerhalb des Rasch-Modells auf dieselbe Skala projiziert sind, 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 Trefferquote (Wahrscheinlichkeit von 0,5) erbracht wird.<sup>12</sup> Übersteigt die auf solcher Grundlage geschätzte Fähigkeit einer Schülerin oder eines Schülers die Schwierigkeit einer Aufgabe, so wird sie bzw. er diese umso eher lösen, je größer die Differenz zwischen Fähigkeits- und Schwierigkeitsparameter ist. Unterschreitet umgekehrt die geschätzte Fähigkeit des Schülers oder der Schülerin die Aufgabenschwierigkeit, so wird er oder sie umso eher scheitern, je größer die Differenz ist.

Um Aussagen über die Lernentwicklung treffen zu können, wurden in den *SL-HAM-8/9-Untertests Mathematik, Deutsch: Sprache und Leseverständnis* und *Englisch: C-Test* die Daten aus LAU 9 gemeinsam mit denen aus LAU 7

---

<sup>12</sup> Die Festlegung dieser Lösungswahrscheinlichkeit für die Zuordnung von Fähigkeits- und Schwierigkeitsparametern beruht auf einer pragmatisch begründeten Entscheidung. Mit der Wahl eines strengeren Maßstabs (vgl. z. B. die in der TIMS-Studie festgelegte Lösungswahrscheinlichkeit von 0,65) verringert sich schülerseitig das Fähigkeitsniveau nur formal. Die „Beherrschung“ eines bestimmten Schwierigkeitsniveaus ist dann an niedrigere Fehlerquoten gebunden. Die Unterschiede innerhalb der Personengruppen und zwischen ihnen ändern sich dabei nicht.

skaliert. Dies war möglich, weil es in den Untertests für die Klassenstufen 7 und 9 jeweils einen gemeinsamen Bestand von Testaufgaben (sog. ‚Ankeritems‘) gab, über den die vorhandene Metrik gleichsam ‚verlängert‘ werden konnte, und zwar so, dass nun die Ergebnisse für alle drei Erhebungszeitpunkte 1996 (LAU 5), 1998 (LAU 7) und 2000 (LAU 9) auf einer einzigen Skala verortet sind. Ausgenommen bleibt hier die Fachleistung *Englisch*, die zu Beginn der Klassenstufe 5 noch nicht gemessen werden konnte, weshalb sich die Verkoppelung hier auf die Klassenstufen 7 und 9 beschränken musste.

Der Mittelwert und die Standardabweichung für die Leistungen der Siebtklässlerinnen und Siebtklässler in den Bereichen *Mathematik*, *Sprache* und *Leseverständnis* lagen bei der Skalenbildung 2000 bereits vor (vgl. LEHMANN, GÄNSFUß & PEEK 1999); deshalb wurde die nunmehr erweiterte Metrik so geeicht, dass diese Werte (also Mittelwert und Streuung der Schülerleistungen in der siebten Klassenstufe) gleich geblieben sind, während sie für die neunte Klassenstufe diesem Maßstab entsprechend berechnet werden konnten. Für den Bereich *Englisch: C-Test* musste ein entsprechender Bezugsrahmen neu festgelegt werden. Um größenordnungsmäßig eine Entsprechung zu den Ergebnissen aus der LAU 7 in den übrigen Untertests herzustellen, wurde für die siebte Klassenstufe im Untertest *Englisch* die durchschnittliche Schülerleistung auf 100 und die Standardabweichung auf 25 festgelegt.<sup>13</sup>

### 3.2.1 Fachleistung Mathematik

Die Teststruktur auf der einen Seite und die Fähigkeitsstruktur der Schülerinnen und Schüler auf der anderen Seite lassen sich – nach dem in Abschnitt 3.1 aufgezeigten Auswertungsverfahren – anschaulich darstellen. Unter der Voraussetzung der Eindimensionalität (vgl. BAUMERT, BOS & LEHMANN 2000, S. 94 ff.) war es sinnvoll (und möglich), die Verteilung der Aufgabenschwierigkeiten derjenigen der gemessenen Testleistungen grafisch gegenüberzustellen und dabei gleichzeitig die mit dem Test erfassten Themenbereiche kenntlich zu machen. Abbildung 1 zeigt dies für den Mathematiktest, ergänzt um Beispielaufgaben aus dem Bereich Algebra.

---

<sup>13</sup> Die Wahl der Transformationsregel unterliegt einzig Gesichtspunkten der Zweckmäßigkeit. 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 die Standardabweichung gewählt, und zwar so, dass die einzelnen Fähigkeitsparameter im Rahmen der gegebenen Messgenauigkeit als ganze positive Zahlen mit „rundem Mittelwert“ dargestellt werden können.

Auf der rechten Seite der Skala sind die 64 Testaufgaben nach ihrem statistisch bestimmten Schwierigkeitsgrad geordnet eingetragen. So liegen z. B. sieben Aufgaben zu *Geometrie*, *Arithmetik* und *Algebra* mit geringerer Schwierigkeit im Skalenbereich 100 bis 110, während sich eine Aufgabe zur *Stochastik* mit relativ hoher Schwierigkeit im Skalenbereich 155 bis 160 befindet. Bei den 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 Alternativen, Einfluss nicht im engeren Sinne mathematischer Fähigkeiten wie z. B. des Leseverständnisses u. a.) führt jedoch in der Regel dazu, dass sich eine Verteilung der Aufgabenschwierigkeiten mit relativ vielen Aufgaben im mittleren Bereich einstellt. Die Aufgaben der verschiedenen Themenbereiche sind weitgehend über die gesamte Skala verteilt; allein die Aufgaben aus dem Bereich *Arithmetik* sind deutlicher auf die untere Hälfte, die wenigen (vor allem für die leistungsstärkeren Schülerinnen und Schüler bestimmten) Aufgaben der *Stochastik* auf die obere Hälfte der Skala konzentriert.

Die Verteilung der mathematischen Testleistungen der Neuntklässlerinnen und Neuntklässler ist auf der linken Seite der Skala abgebildet. Wie die Abbildung 1 zeigt, sind die Ergebnisse annähernd normalverteilt: Nur wenige Schülerinnen und Schüler haben extrem schwache bzw. (nahezu) perfekte Leistungen gezeigt, während viele Leistungen im Bereich um den Skalenmittelwert 125 (Standardabweichung  $s = 19,6$ ) liegen.

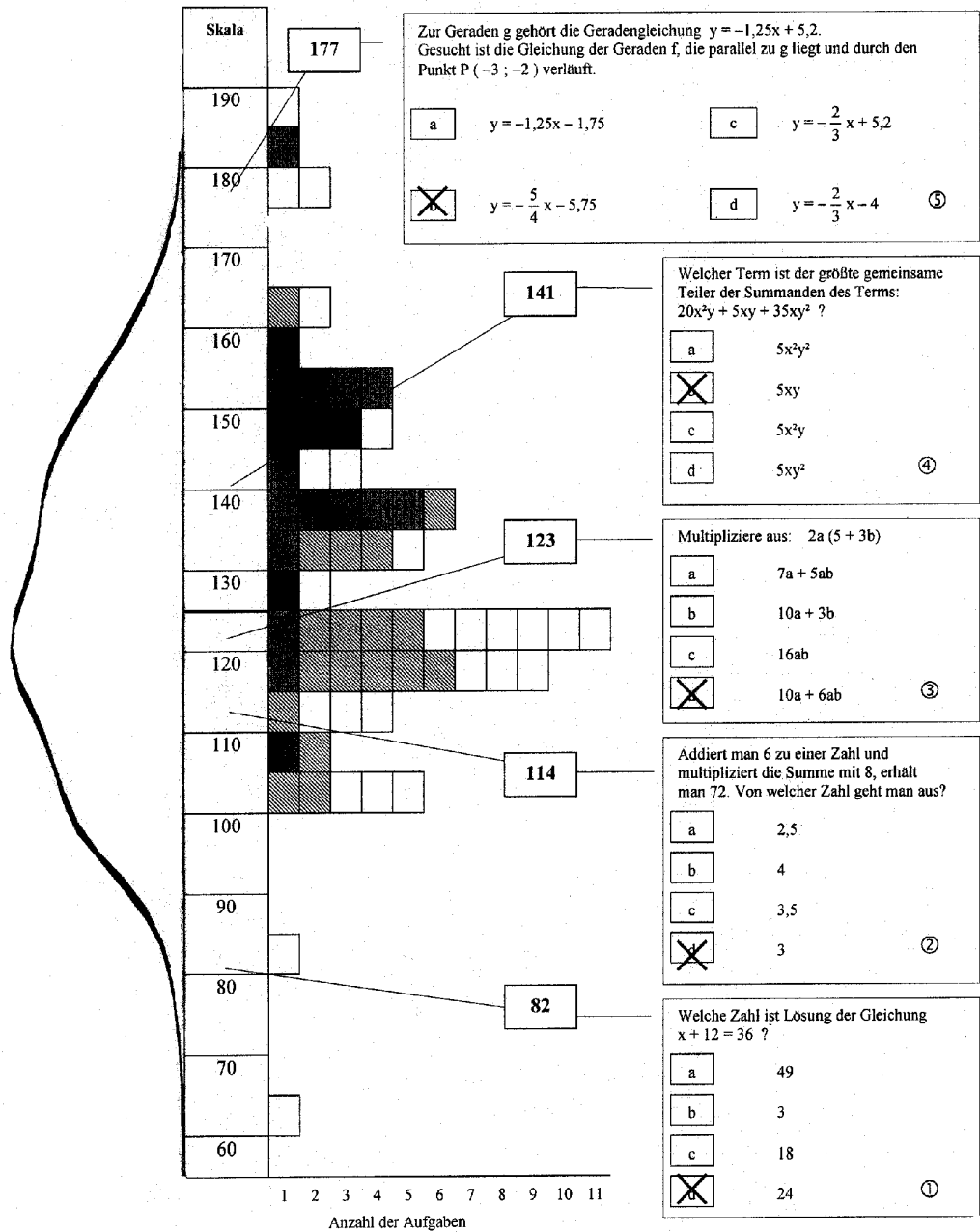
In *Algebra* sollten die Schülerinnen und Schüler sichere Kenntnisse beim Lösen von einfachen Gleichungen und bei Grundrechenoperationen mit Dezimalbrüchen nachweisen, Rechengesetze im Umgang mit Klammern und mit Variablen beherrschen sowie Teilbarkeitsregeln anwenden und Geradengleichungen lösen können. Um nun einen Eindruck von den Aufgabenstellungen des Tests und ihren unterschiedlichen Anforderungsniveaus zu geben, sind in die Abbildung 1 Beispielaufgaben aus dem Themenbereich *Algebra* eingetragen. Die Darstellung ist so gewählt, dass die Aufgaben bestimmten Fähigkeitsniveaus zugeordnet werden, die erreicht sein müssen, um sie mit einiger Sicherheit – d. h. nach dem oben Gesagten: mit einer Wahrscheinlichkeit von 0,5 – lösen zu können.

Abbildung 1 Verteilung der Schülerleistungen im Fach Mathematik im Vergleich mit den Schwierigkeiten der Testaufgaben im *SL-HAM 8/9-Mathematiktest* (nach Themenbereichen)

Verteilung der Schülerleistungen

Testaufgaben

Aufgaben aus dem Bereich *Algebra*



Gesamtstichprobe ( $N = 11.304$ )

Themenbereiche

■ Stochastik ■ Geometrie ■ Arithmetik □ Algebra

Die Datenauswertung hat ergeben, dass z. B. das Leistungsniveau mit dem Skalenwert 82 nahezu von allen Schülerinnen und Schülern (99,5 Prozent) mit einiger Sicherheit erreicht oder überschritten wurde. Die Beispielaufgabe 1 mit der Anforderung, eine einfache Gleichung mit einer Unbekannten zu lösen, ist eine Aufgabe dieser Art und Schwierigkeit.<sup>14</sup> Das Fähigkeitsniveau 114 mit der Beispielaufgabe 2, bei der von den Schülerinnen und Schülern selbst eine Gleichung aufgestellt und anschließend gelöst werden musste, haben ca. 68 Prozent der Schülerinnen und Schüler erreicht bzw. übertroffen. Die Lösung einer Aufgabe, die den Umgang mit Variablen verlangt und die Beherrschung von Grundrechenarten bei Berücksichtigung der Klammerregeln erfordert, gelang mit einiger Sicherheit ab einem Skalenwert von 123; die Beispielaufgabe 3 spiegelt das Leistungsniveau wider, das ca. 52 Prozent der Hamburger Neuntklässlerinnen und Neuntklässler erreicht oder überschritten haben. Die Beispielaufgabe 4 (Skalenwert 141), bei der Teilbarkeitsregeln auf Variablen angewandt werden müssen, ist eine typische Aufgabe für ein mathematisches Leistungsniveau, das ca. 23 Prozent der Schülerinnen und Schüler erzielt oder überschritten haben. Nur ein geringer Anteil von 0,3 Prozent der Hamburger Neuntklässlerinnen und Neuntklässler hat ein Leistungsniveau erreicht, das dem Skalenwert von 177 und damit dem Schwierigkeitsgrad der fünften Beispielaufgabe entspricht.

### **Fachleistung Mathematik, differenziert nach Schulformen bzw. Kursniveaus**

Ab dem Beginn der Klassenstufe 7 teilt sich die Hamburger Schülerschaft in fünf Schulformen, in denen nach Regellehrplänen unterrichtet wird: Hauptschulen, Realschulen, Gymnasien, Gesamtschulen sowie – als Schulversuch – Integrierte Haupt- und Realschulen (IHR-Schule). Innerhalb der Integrierten Haupt- und Realschulen und innerhalb der Gesamtschulen ist zwischen fachspezifischen Kursniveaus zu unterscheiden: an den Integrierten Haupt- und Realschulen zwischen den H- und R-Kursen, an den (integrierten) Gesamtschulen zwischen dem Kursniveau II und dem anspruchsvolleren Kursniveau I. Ähnlich wie in den beiden vorangegangenen LAU-Berichten (vgl. LEHMANN & PEEK 1997; LEHMANN, GÄNSFUß & PEEK 1999) sollen zunächst die spezifischen Fachleistungen berichtet werden, ehe ein Index für die allgemeine Fachleistung gebildet wird. Dabei sind – wo immer möglich – nicht die Testrohwerte Grundlage der Analyse, sondern die skalierten Lernstandswerte, um den

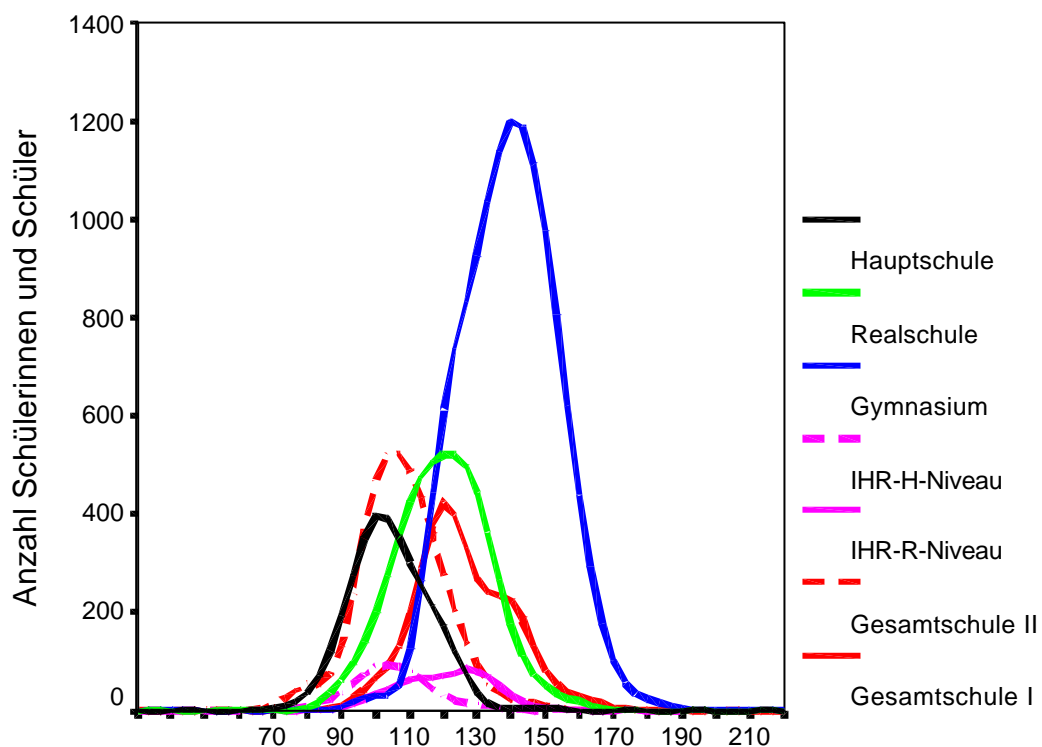
---

<sup>14</sup> Die hier genannten Prozentangaben beziehen sich nicht auf die einzelnen Aufgaben, sondern auf das erreichte Schwierigkeitsniveau, für das die Beispielaufgaben stellvertretend stehen.

Bezug zur Darstellung von Lernfortschritten in den Klassenstufen 7 und 8 zu sichern.

Abbildung 2 zeigt die Verteilung der mathematischen Fachleistung am Ende der Klassenstufe 8, und zwar getrennt nach Schulformen bzw. Kursniveaus. Hier ist zunächst auffällig, dass sich die Streuung der Leistungen im Hamburger Gesamtsystem gegenüber dem Stand zu Beginn der Klassenstufe 7 offenbar weiter verringert hat, obwohl in den Klassenstufen 7 und 8 mit fünf verschiedenen Schulformen (bzw. mit insgesamt sieben Differenzierungsgraden) ein deutlich höheres Maß an äußerer Differenzierung gegeben ist als in den beiden Klassenstufen davor.

Abbildung 2 Fachleistung Mathematik: Verteilung der Lernstände zum Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Fachleistung Mathematik Ende Klassenstufe 8

Wie bereits zu Beginn der Klassenstufe 7 sind in der LAU-9-Untersuchung die Überschneidungen zwischen den Schulformen bzw. Kursniveaus im Fach Mathematik sehr beträchtlich: Bezeichnet man – ähnlich wie in der früheren Untersuchung – den Leistungsbereich oberhalb des Schnittpunkts zwischen den Verteilungen der Realschulen und der Gymnasien (d. h. Lernstandswerte  $> 118$ ) als den *typisch gymnasialen Leistungsbereich*, so waren am Ende der achten Klassenstufe 14,5 Prozent der Schülerinnen und Schüler aus Hauptschulen und 56,6 Prozent aus Realschulen in diesem Segment zu finden. Die

Werte für die Integrierten Haupt- und Realschulen sind ähnlich: 13,6 Prozent in den H-Kursen und 58,3 Prozent in den R-Kursen haben Ergebnisse erzielt, die mit der größten Wahrscheinlichkeit an Gymnasien anzutreffen sind. An den Gesamtschulen waren es 20,7 Prozent (Kurse II) bzw. 68,2 Prozent (Kurse I). 9,7 Prozent der Schülerinnen und Schüler von Gymnasien haben niedrigere Ergebnisse erreicht, als es der von ihnen besuchten Schulform typischerweise entspricht.

Bei der Interpretation dieser Zahlen ist indessen zu beachten, dass der hohe Gymnasialanteil (40,6 Prozent der insgesamt getesteten Schülerinnen und Schüler) und die geringe Realschulquote die Abgrenzung der Segmente auf einen vergleichsweise niedrigen Wert festlegen. Würde man von einer Gleichgewichtung aller Schulformen und Kursniveaus ausgehen, läge die kritische Schwelle wesentlich höher, und demnach wäre auch ein geringerer Anteil der nichtgymnasialen Schülerinnen und Schüler im typisch gymnasialen Leistungsbereich vertreten; freilich wäre so auch der Anteil der Gymnasiastinnen und Gymnasiasten mit Leistungen unterhalb des typisch gymnasialen Bereichs wesentlich höher. Auch ist zu bedenken, dass sich außerhalb des Gymnasiums unter den Getesteten die Anteile derer, die jeweils den anspruchsvolleren Mathematikunterricht erhalten haben, deutlich unterscheiden: An den Haupt- und Realschulen befanden sich am Ende der achten Klassenstufe 62,3 Prozent der getesteten Schülerinnen und Schüler in Realschulklassen, an den Integrierten Haupt- und Realschulen 52,4 Prozent in R-Kursen für das Fach Mathematik und an den Gesamtschulen 45,6 Prozent in Mathematikkursen I. So betrachtet, hatten sich die Integrierten Haupt- und Realschulen auf der Fachebene selektiver verhalten als die traditionellen. Die Ähnlichkeit der jeweiligen Anteile im typisch gymnasialen Bereich deutet darauf hin, dass die Standards in beiden Fällen dennoch ähnlich sind; tatsächlich sind sie an den Integrierten Haupt- und Realschulen etwas höher. Die Gesamtschulen hatten bei der Einstufung der Schülerinnen und Schüler in die Kurse I die höchsten Standards angelegt. Dadurch sind auf *beiden* Kursniveaus die Anteile im typisch gymnasialen Bereich relativ hoch.<sup>15</sup>

Ein weiterer aufschlussreicher Kennwert ist mit dem gymnasialen Leistungsdurchschnitt in *Mathematik* gegeben, der einem Lernstandswert von 139 entspricht. Dieser Wert wurde von 0,9 Prozent der Schülerinnen und Schüler

---

<sup>15</sup> Die Schnittpunkte der Verteilungen für das untere bzw. obere Differenzierungsniveau liegen bei den Lernstandswerten 107 (traditionelle Haupt- und Realschulen), 112 (Integrierte Haupt- und Realschulen) und 117 (Gesamtschulen). Die Korrelationen zwischen Niveaueingetragtheit und Fachleistung betragen  $r = 0,77$ ,  $r = 0,81$  und  $r = 0,79$ , unterscheiden sich also nur geringfügig.



an Hauptschulen, 10,8 Prozent an Realschulen, 0,3 Prozent in den H-Kursen der IHR-Schulen, 8,7 Prozent in den R-Kursen der IHR-Schulen, 1,7 Prozent in den Kursen II der Gesamtschulen und 18,5 Prozent in den Kursen I der Gesamtschulen erreicht oder überschritten. In der nachstehenden Tabelle 8 werden die Unterschiede in den erreichten Mathematikleistungen am Ende der Klassenstufe 8 in numerischer Form nach Schulformen bzw. nach Kursniveaus differenziert ausgewiesen.

Tabelle 8 Fachleistung Mathematik am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	104,5	12,3	-1,04	1.158
Realschule	120,4	14,2	-0,22	1.911
Gymnasium	139,2	14,5	0,73	4.459
IHR-Schule – H	104,5	12,4	-1,04	271
IHR-Schule – R	120,6	13,8	-0,21	298
Gesamtschule II	107,8	13,3	-0,87	1.573
Gesamtschule I	125,7	15,4	0,05	1.316
<i>insgesamt</i>	<i>124,8</i>	<i>19,6</i>	<i>---</i>	<i>10.986</i>

Wie diesen Zahlen zu entnehmen ist, bestanden am Ende der Klassenstufe 8 beträchtliche Unterschiede zwischen den Schulformen bzw. Kursniveaus: Die Gymnasiastinnen und Gymnasiasten haben Durchschnittsleistungen erzielt, die trotz des mit 40,6 Prozent hohen Anteils an der hier berücksichtigten Schülerschaft um 0,73 Standardabweichungen über dem allgemeinen Mittelwert lagen. Der Vergleich der Schulformen bzw. Kursniveaus zeigt außerdem, dass zwischen den durchschnittlich erreichten Mathematikleistungen der Schülerinnen und Schüler aus Haupt- und Realschulen in der Regel- bzw. der integrierten Form praktisch keine Unterschiede auf den jeweiligen Differenzierungsniveaus feststellbar waren. Die Durchschnittsleistungen auf den beiden Kursniveaus der Gesamtschulen waren jeweils höher als die in den beiden Differenzierungsformen an den Haupt- und Realschulen<sup>16</sup>, wobei die an den Gesamtschulen praktizierte höhere Selektivität zu berücksichtigen ist.

<sup>16</sup> Wie bereits berichtet und den in der Tabelle 2 in der Spalte *N* aufgeführten Werten zur Anzahl der Schülerinnen und Schüler zu entnehmen ist, ist die Beteiligungsquote an den unterschiedlichen Bildungsgängen unterschiedlich hoch.

Wie aus den Standardabweichungen (LAU 9:  $s = 19,6$ ; LAU 7:  $s = 23,8$ ) nochmals ersichtlich wird, hat sich die Streuung der Mathematikleistungen im Hamburger Schulsystem insgesamt vom Ende der Klassenstufe 6 zum Ende der Klassenstufe 8 verringert, und zwar – wie die Vergleiche für einzelne Schulformen zeigen (vgl. hierzu LEHMANN, GÄNSFUß & PEEK 1999, S. 114) – jeweils in analogen Verhältnisgrößen. Um ein Artefakt, das von Deckeneffekten der einzelnen Testversionen herrühren würde, kann es sich dabei – ausweislich der Verteilungsformen, die in Abbildung 2 sichtbar waren – nicht handeln. Für das Ende der Klassenstufe 8 ist in Hamburg im Fach Mathematik demnach eine auffällige Homogenität der Leistungen innerhalb der Schulformen bzw. Kursniveaus festzustellen, die auch die Streuung innerhalb des Gesamtsystems vergleichsweise gering erscheinen lässt und als Konsequenz ein hohes Maß an Durchlässigkeit zwischen den Schulformen aufrechterhält bzw. ermöglicht.

### **3.2.2 Fachleistung Deutsch**

Der *SL-HAM 8/9* enthält drei Untertests, die für den Deutschunterricht relevante Leistungen erfassen: einen Leseverständnistest, einen Sprachverständnistest und einen Rechtschreibtest zur Ermittlung des Rechtschreibwissens. Darüber hinaus wurden einer Stichprobe die *Hamburger Schreibprobe (HSP 5-9)* als Test zur Messung des Rechtschreibkönnens und eine Aufgabe zur *Textproduktion* gestellt. Die hierauf bezogenen Befunde für die Schülerinnen und Schüler am Ende der Klassenstufe 8 werden in den folgenden Abschnitten vorgestellt.

#### **3.2.2.1 Leseverständnis**

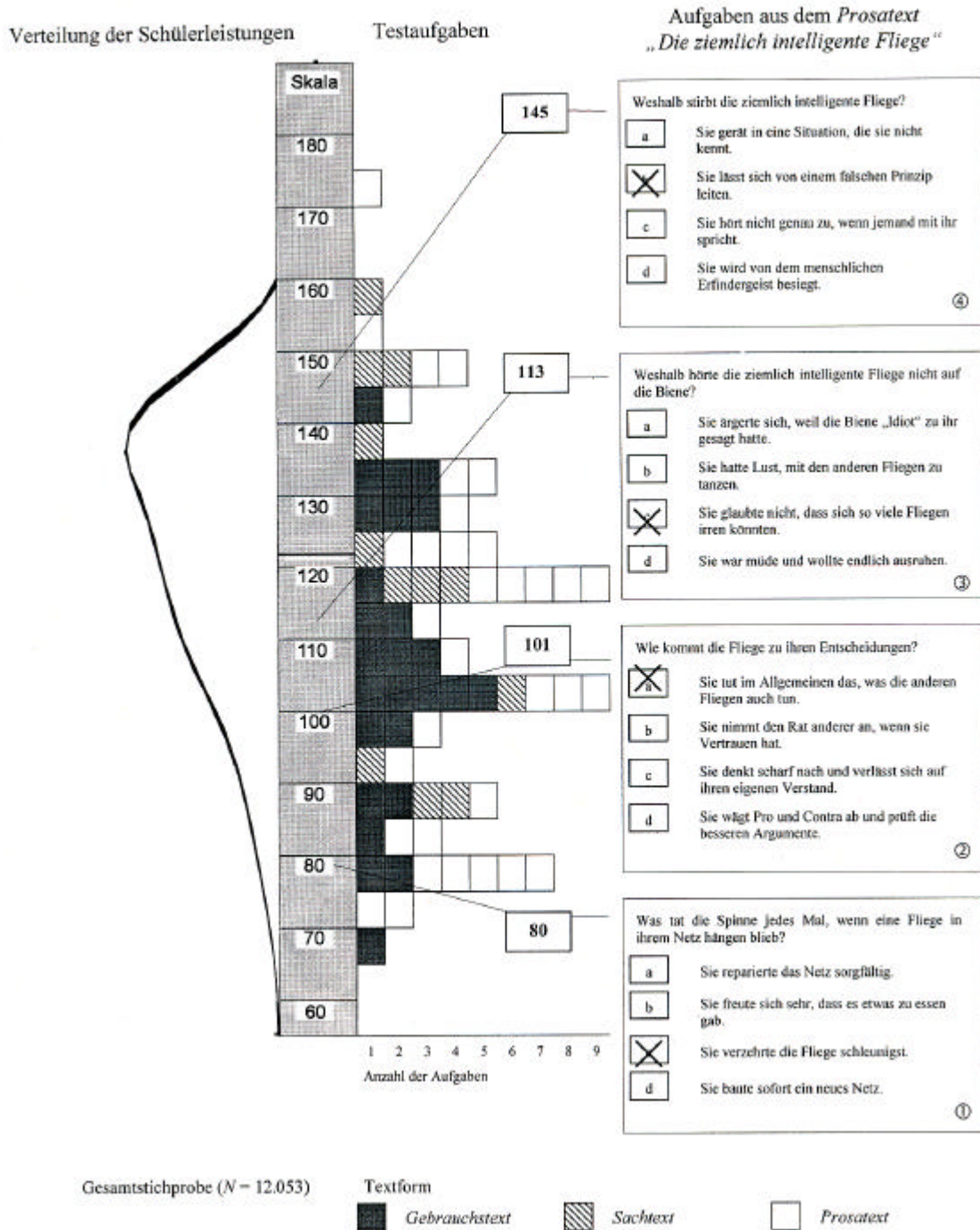
Die Eigenschaften der Aufgaben des *SL-HAM 8/9-Leseverständnistests* und die bei den Schülerinnen und Schülern gegebenen Kompetenzen im Leseverständnis sind – wie im Abschnitt 3.1 für den Untertest Mathematik beschrieben – über die Rasch-Skalierung zugleich miteinander verschränkt und aufeinander projizierbar. Die Gesamtskala für das Leseverständnis umfasst aufgabenseitig unterschiedliche Textformen (Prosa-, Sach- und Gebrauchstexte) und Schwierigkeitsgrade, schülerseitig unterschiedliche Kompetenzen und Leistungsniveaus. Dabei ist es wiederum möglich und sinnvoll, die Verteilung der Aufgabenschwierigkeiten derjenigen des gemessenen Leseverständnisses grafisch gegenüberzustellen. Abbildung 3 zeigt das Ergebnis für die Gesamtgruppe der getesteten Schülerinnen und Schüler, wobei die einzelnen Aufgaben nach den unterschiedlichen Textformen ausgewiesen sind, auf die sie sich beziehen.

Hinsichtlich der Aufgabenschwierigkeiten hat sich bei einem Mittelwert von 110 eine breite Verteilung von 65 bis 175 Skalenpunkten ergeben. Die Aufgaben, die sich auf Gebrauchstexte beziehen, sind im Vergleich zu den Prosa- und Sachtextaufgaben deutlicher auf die mittlere und untere Hälfte der Skala konzentriert. Dies liegt offensichtlich in der Textform selbst begründet: Gebrauchstexte werden verfasst, um elementare Informationen möglichst unmissverständlich mitzuteilen. Es war zu erwarten, dass die Schülerinnen und Schüler mit solchen Texten verhältnismäßig geringe Schwierigkeiten haben würden. Dagegen besteht kein eindeutiges Schwierigkeitsgefälle zwischen literarischen Texten und Sachtexten. In beiden Fällen können die entsprechenden Aufgaben sehr spezielle Verständnisleistungen erfordern, was sich in der breiteren Verteilung der Aufgabenschwierigkeiten über den gesamten Skalenbereich zeigt.

Das Leseverständnis ist – wie auf der linken Seite der Grafik ersichtlich – keineswegs „normal“ verteilt: Einige Schülerinnen und Schüler zeigten extrem schwache Verständnisleistungen, während verhältnismäßig viele Schülerinnen und Schüler gute und sehr gute Werte erzielen konnten. Der Mittelwert für das Leseverständnis liegt mit dem Raschwert 122 (Standardabweichung: 19,6) deutlich höher als die mittlere Aufgabenschwierigkeit (110). Besonders an den Gymnasien ist nicht ganz auszuschließen, dass ein Teil der Schülerinnen und Schüler unterfordert war.

Dass bei der Testkonzeption bewusst unterschiedlich schwierige Items berücksichtigt wurden, um möglichst den gesamten Anforderungsbereich (von einfachen Dekodierungsaufgaben bis hin zum eigenständigen schlussfolgernden Umgang mit Texten) zu erfassen, bestätigt sich empirisch, wenn man die jeweils geforderten Leistungen genauer untersucht. Die Tendenz von einfachen Aufgaben des (Wieder-) Erkennens zu komplexen mentalen Operationen ist unverkennbar. Immer erscheinen solche Aufgaben am leichtesten, die mit dem Auffinden eines Wortes, einer kürzeren Passage oder auch eines Symbols lediglich einen direkten Abgleich zwischen Text und Frage verlangen. Die Schwierigkeit steigt in dem Maße, wie komplexere Randbedingungen präsent gehalten werden müssen und wie der Gehalt umfangreicherer Passagen zunächst verdichtet und gegebenenfalls weiterverarbeitet werden muss.

Abbildung 3 Verteilung der Schülerleistungen im Leseverständnis im Vergleich mit den Schwierigkeiten der Testaufgaben im *SL-HAM 8/9-Leseverständnistest* (nach Textformen)



In der Abbildung 3 sind – beispielhaft – aus dem auf der folgenden Seite abgedruckten Prosatext *„Die ziemlich intelligente Fliege“* vier Aufgaben wiedergegeben, die das Prinzip der Hierarchie zunehmend schwierigerer Anforderungen verdeutlichen.

Die Frage: *„Was tat die Spinne jedes Mal, wenn eine Fliege in ihrem Netz hängen blieb?“* (Beispielaufgabe 1) erfordert auf einer niedrigen Anforderungsstufe (Rasch-Skalenwert 80), eine im selben Wortlaut gegebene Information aufzufinden, die direkt im ersten Absatz des Textes steht. 97 Prozent der getesteten Schülerinnen und Schüler haben diese Fähigkeitsstufe der einfachen Informationsentnahme erreicht bzw. übertroffen.

Die anspruchsvollere Beispielaufgabe 2 (*„Wie kommt die Fliege zu ihren Entscheidungen?“*) mit dem Rasch-Skalenwert 101 verlangt die Fähigkeit, aus dem Text einen Sachverhalt zu rekonstruieren, wobei die Formulierungen in Frage und Text sprachlich nicht identisch sind und eine Verallgemeinerung aus vorgegebenen Einzelsituationen geleistet werden muss. Oberhalb dieses Skalenwertes lagen 84 Prozent der getesteten Schülerinnen und Schüler.

Die dritte Beispielfrage: *„Weshalb hörte die ziemlich intelligente Fliege nicht auf die Biene?“* (Skalenwert 113) steht für ein Fähigkeitsniveau, das immer noch von vielen Schülerinnen und Schülern erreicht bzw. übertroffen und nur von 30 Prozent nicht erreicht worden ist. Hier geht es darum, im Text gegebene Informationen und Handlungs- bzw. Hintergrundmotive zu rekonstruieren und sodann miteinander zu kombinieren.

Ein Beispiel für eine vergleichsweise anspruchsvolle Aufgabe mit dem Rasch-Skalenwert 145 – dieser Wert wurde nur von 10 Prozent der getesteten Schülerinnen und Schüler erreicht bzw. überschritten bzw. von 90 Prozent nicht erreicht – ist die Frage: *„Weshalb stirbt die ziemlich intelligente Fliege?“* (Beispielaufgabe 4). Die richtige Antwort (*„Sie lässt sich von einem falschen Prinzip leiten.“*) erfordert auf einem hohen Niveau – über die vorgelegten Leseverständnisstufen hinaus – die Fähigkeit, Schlussfolgerungen zu treffen und Verknüpfungen vorzunehmen.

## **Die ziemlich intelligente Fliege**

Eine große Spinne hatte in einem alten Haus ein schönes Netz gewoben, um Fliegen zu fangen. Jedes Mal, wenn eine Fliege sich auf dem Netz niederließ und darin hängen blieb, verzehrte die Spinne sie schleunigst, damit andere Fliegen, die vorbeikamen, denken sollten, das Netz sei ein sicherer und gemütlicher Platz.

Eines Tages schwirrte eine ziemlich intelligente Fliege so lange um das Netz herum, ohne es zu berühren, dass die Spinne schließlich hervorkroch und sagte: „Komm, ruh’ dich bei mir ein bisschen aus.“ Aber die Fliege ließ sich nicht übertölpeln: „Ich setz’ mich nur an Stellen, wo ich andere Fliegen sehe“, antwortete sie, „und ich sehe bei dir keine anderen Fliegen.“

Damit flog sie weiter, bis sie an eine Stelle kam, wo sehr viele Fliegen saßen. Sie wollte sich gerade zu ihnen gesellen, als eine Biene ihr zurief: „Halt, du Idiot, hier ist Fliegenleim. Alle diese Fliegen sitzen rettungslos fest.“ – „Red’ keinen Unsinn“, sagte die Fliege. „Sie tanzen doch.“ Damit ließ sie sich nieder und blieb auf dem Fliegenleim kleben wie all die anderen Fliegen.

Moral: Der Augenschein kann ebenso trügerisch sein wie die Sicherheit, in der man sich wiegt.

## **Leseverständnis, differenziert nach Schulformen bzw. Kursniveaus**

Im Folgenden soll – wieder differenziert für die Schülerinnen und Schüler aus den verschiedenen Schulformen bzw. Kursniveaus Hauptschule, Realschule, Gymnasium, Integrierte Haupt- und Realschule (H-Kurse bzw. R-Kurse) und Gesamtschule (Deutsch-Kurse II und I) – die Verteilung der erreichten Leistungen im Leseverständnistest am Ende der Klassenstufe 8 dargestellt und diskutiert werden. In der nachstehenden Tabelle 9 sind die statistischen Kennwerte für das am Ende der Klassenstufe 8 erreichte Leseverständnis nach Schulform bzw. nach Kursniveau aufgeführt.

Die in der Tabelle berichteten Werte zeigen deutliche Unterschiede zwischen den Schulformen bzw. Kursniveaus am Ende der Klassenstufe 8 auf, wobei sich die Verhältnisse ähnlich gestalten wie bei den Mathematikleistungen: Wiederum liegen die durchschnittlichen Leistungen der Gymnasiastinnen und Gymnasiasten (Mittelwert: 134) bei der insgesamt geringsten Leistungsstreuung ( $s = 13,3$ ) deutlich vor den Leistungen der Schülerinnen und Schüler aus den Kursen I der Gesamtschulen. Es folgen – mit etwas schwächer ausgeprägten Durchschnittsleistungen – die Schülerinnen und Schüler aus den Realschulen bzw. den R-Kursen der Integrierten Haupt- und Realschulen, deren Mittelwerte und Streuungsmaße nahezu übereinstimmen.

Tabelle 9 Leseverständnis am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

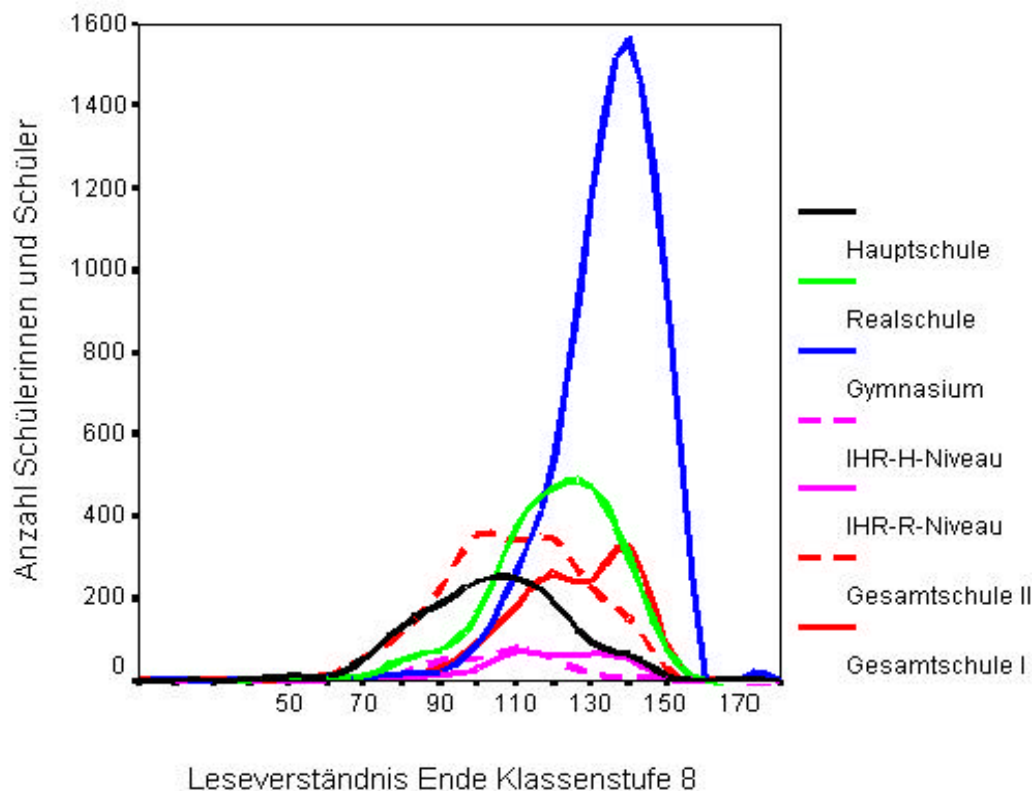
	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	104,6	19,4	-0,91	1.271
Realschule	120,7	16,5	-0,09	2.017
Gymnasium	133,9	13,3	0,58	4.647
IHR Schule – H	106,3	16,8	-0,83	300
IHR Schule – R	119,7	16,0	-0,14	314
Gesamtschule II	109,6	18,9	-0,66	1.877
Gesamtschule I	124,9	17,0	0,12	1.257
<i>insgesamt</i>	<i>122,5</i>	<i>19,6</i>	<i>---</i>	<i>11.683</i>

Die Schülerinnen und Schüler aus Hauptschulklassen bzw. den H-Kursen der Integrierten Haupt- und Realschulen zeigen mit 105 bzw. 106 durchschnittlich erreichten Skalenpunkten die schwächsten Leistungen. Die Schülerinnen und Schüler aus den Deutschkursen II der Gesamtschulen liegen im Vergleich etwas höher.<sup>17</sup> In der folgenden Abbildung 4 sind für die insgesamt 11.683 berücksichtigten Schülerinnen und Schüler die schulform- bzw. kursniveau-bezogenen Verteilungen eingezeichnet.

Was bereits in der Abbildung 3 für die Gesamtgruppe der Schülerinnen und Schüler zu Beginn der Klassenstufe 9 durch die insgesamt schiefe Verteilung deutlich wurde, bestätigt sich hier für die Schulformen bzw. Kursniveaus am Ende der Klassenstufe 8: Bei einer mittleren Testschwierigkeit von 110 Skalenpunkten zeigt eine deutliche Mehrheit der Schülerinnen und Schüler aus Realschulklassen, den R-Kursen der Integrierten Haupt- und Realschulen, den Deutschkursen I an Gesamtschulen und den Gymnasialklassen höhere Testleistungen.

<sup>17</sup> Dabei ist wieder zu beachten, dass an den Gesamtschulen 40,1 Prozent der Schülerinnen und Schüler im Fach Deutsch dem höheren Kursniveau zugeordnet sind, während es an den Integrierten Haupt- und Realschulen 51,1 Prozent sind und an den Haupt- und Realschulen 61,3 Prozent dem Realschulzweig angehören.

Abbildung 4 Leseverständnis: Verteilung der Lernstände am Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Eine asymmetrische Verteilung der Testergebnisse zeigt sich vor allem im Gymnasialbereich, obwohl „perfekte“ Ergebnisse selten sind und daher kein sog. „Deckeneffekt“ vorliegt (vgl. auch unten, Abbildung 5). Wie für das Fach Mathematik sind die Überschneidungen zwischen den Schulformen bzw. Kursniveaus im Leseverständnis sehr beträchtlich: Bezeichnet man wieder den Bereich oberhalb des Schnittpunkts zwischen den Verteilungen der Realschulen und der Gymnasien (Lernstandswerte  $> 116$ ) als den *typisch gymnasialen Leistungsbereich*, so waren am Ende der achten Klassenstufe 29,0 Prozent der Schülerinnen und Schüler aus Hauptschulklassen bzw. 27,7 Prozent aus H-Kursen an Integrierten Haupt- und Realschulen in diesem Segment zu finden. Die Werte für die Schülerinnen und Schüler aus Realschulklassen (65,8 Prozent) bzw. aus R-Kursen an Integrierten Haupt- und Realschulen (59,6 Prozent) liegen deutlich höher. In den Gesamtschulen haben 40,0 Prozent (Kursniveau II) bzw. 73,6 Prozent (Kursniveau I) einen Stand im Leseverständnis erreicht, der mit der höchsten Wahrscheinlichkeit im Gymnasium anzutreffen ist. Demgegenüber liegen 9,1 Prozent der Schülerinnen und Schüler an den Gymnasien mit ihren Testergebnissen im Leseverständnis unterhalb der Grenze, mit der in der von ihnen besuchten Schulform typischerweise zu rechnen ist.



Schließlich soll auch hier – als weiterer aufschlussreicher Kennwert – der gymnasiale Leistungsdurchschnitt betrachtet werden, der einem Skalenwert von 134 entspricht. Dieser Wert wurde von 2,1 Prozent der Schülerinnen und Schüler an Hauptschulen, 16,6 Prozent an Realschulen, 1,8 Prozent in den H-Kursen der IHR, 18,6 Prozent in den R-Kursen der IHR, 6,8 Prozent in den Kursen II der Gesamtschulen und 22,6 Prozent in den Kursen I der Gesamtschulen erreicht oder überschritten. Die asymmetrische Verteilung, die an den Gymnasien für den Untertest *Leseverständnis* festzustellen ist, zeigt sich auch an der hohen Quote von 68,1 Prozent Schülerinnen und Schülern dieser Schulform, die oberhalb des Skalenwertes 134 liegen.

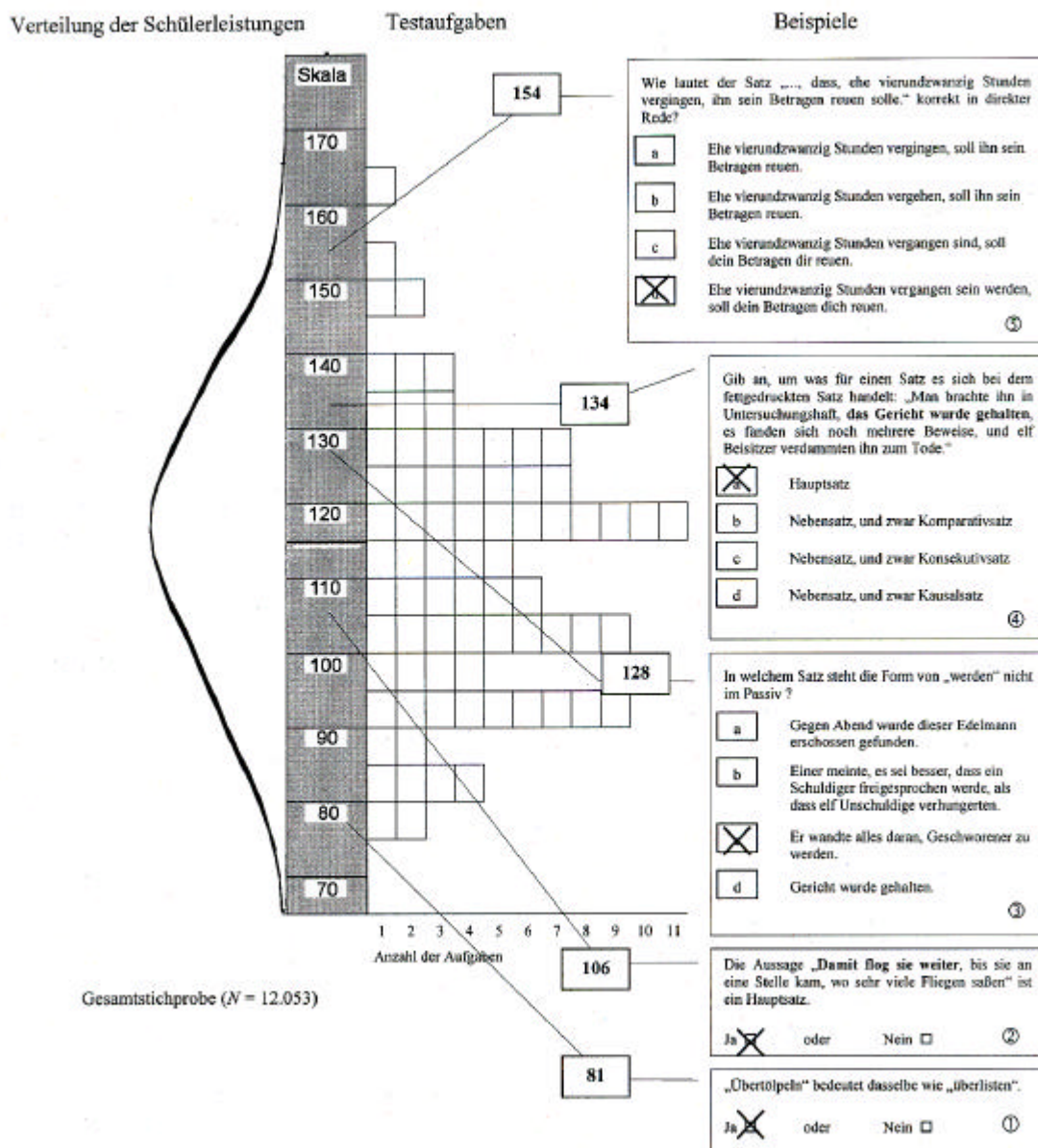
### 3.2.2.2 Sprache

Bei insgesamt neun der elf Texte, die Grundlage des Leseverständnistests waren, wurden 76 Aufgaben zum Sprachverständnis gestellt. Hinzu kommen zehn Ankeritems aus der LAU-7-Untersuchung, die nicht unmittelbar textgebunden sind. Wie für die Untertests *Mathematik* und *Deutsch: Leseverständnis* zeichnet sich auch für den Untertest *Deutsch: Sprache* ein breites Spektrum an Aufgabenschwierigkeiten ab, zu denen die Schülerleistungen unmittelbar in Beziehung gesetzt werden können. Abbildung 5 zeigt die Verteilungen auf der Rasch-Skala. Die durchschnittliche Aufgabenschwierigkeit des Sprachtests liegt mit dem Skalenwert 112 nahe beim Mittelwert für den Leseverständnistests, wobei die Aufgabenschwierigkeiten über den Bereich 75 bis 165 wieder relativ breit streuen.<sup>18</sup> Die Verteilung der Schülerleistungen ist bei einem Skalenmittelwert 115 (Standardabweichung: 19,0) annähernd „normal“ und gegenüber der Verteilung im Leseverständnistest insgesamt nur wenig in den oberen Leistungsbereich verschoben. Es zeigt sich damit, dass der Sprachtest für die Schülerinnen und Schüler insgesamt etwas schwerer zu lösen war als der Leseverständnistest. In der Abbildung sind beispielhaft Aufgaben mit den für sie geltenden Rasch-Skalenwerten aufgeführt, die einen Einblick in die Anforderungsstruktur des Tests geben und darüber hinaus die zunehmende Schwierigkeit der Aufgaben deutlich machen.

---

<sup>18</sup> Die Sprachaufgaben sind im Unterschied zu den Leseverständnisaufgaben in der Abbildung nicht textformspezifisch ausgewiesen.

Abbildung 5 Verteilung der Schülerleistungen im *Sprachverständnis* im Vergleich mit den Schwierigkeiten der Testaufgaben im *SL-HAM 8/9-Sprachverständnis*test



Ein Fähigkeitsniveau, dem Aufgaben wie das mit dem Rasch-Skalenwert 81 verbundene Beispiel 1 zur Wortbedeutung von „*übertölpeln*“ entsprechen, haben 96,5 Prozent der Schülerinnen und Schüler erreicht oder übertroffen. Hier gilt es lediglich, eine Aussage zu bestätigen oder zu verneinen. Beispielaufgabe 2, bei der – ebenfalls nur mit „ja“ oder „nein“ – die Frage beantwortet werden muss, ob es sich bei dem Satz „*Damit flog sie weiter, ...*“ um einen Hauptsatz handelt, steht mit dem Skalenwert 106 für ein sprachliches Kompetenzniveau, das von 69,3 Prozent der Schülerinnen und Schüler erreicht bzw. überschritten wurde. Bei der dritten Beispielaufgabe mit dem Rasch-

Skalenwert 128 – dieses Fähigkeitsniveau haben 26,6 Prozent der Schülerinnen und Schüler erreicht bzw. übertroffen – muss aus vier vorgegebenen Sätzen derjenige bestimmt werden, der sprachlich nicht im Passiv, sondern im Aktiv verfasst ist. Beispielaufgabe 4 liegt mit dem Skalenwert 134 im oberen Anforderungsbereich, der von 15,2 Prozent der Schülerinnen und Schüler erreicht bzw. übertroffen worden ist. Hier gilt es, einen Teilsatz als Hauptsatz zu bestimmen, wobei alternativ drei unterschiedliche Nebensatztypen angeboten sind. Schließlich liegt beim Skalenwert 154, den nur 1,3 Prozent der Schülerinnen und Schüler erreicht bzw. übertroffen haben, die Beispielaufgabe 5, bei der aus vier Alternativvorgaben diejenige ausgewählt werden muss, die anzeigt, wie die vorgegebene indirekte Aussage sprachlich in direkter Rede lautet.

### **Sprachverständnis, differenziert nach Schulformen bzw. Kursniveaus**

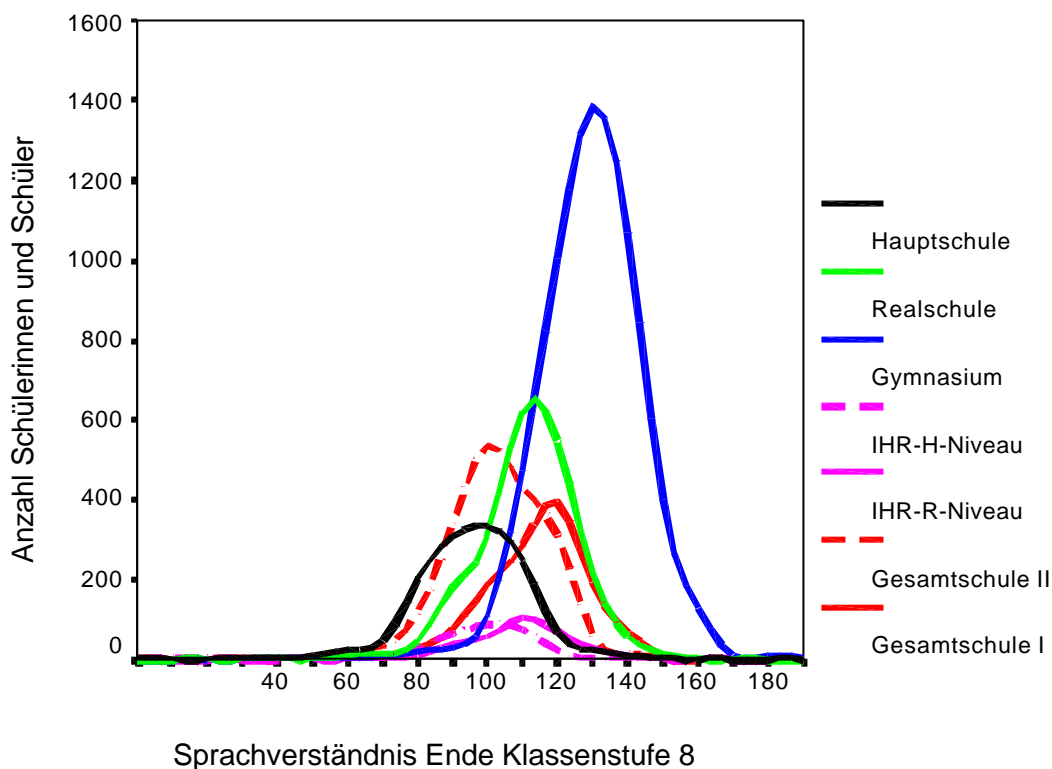
Insgesamt zeigt sich im Bereich *Deutsch: Sprache* – dies ist in der folgenden Tabelle 10 an den standardisierten Effektstärken ablesbar – bei einer dem Leseverständnis analogen Rangfolge der Testleistungen in den Schulformen bzw. Kursniveaus ein größerer Abstand der Gymnasiastinnen bzw. Gymnasiasten von den anderen Gruppen.

Tabelle 10 Sprachverständnis am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	95,9	15,0	-1,02	1.271
Realschule	111,1	13,7	-0,22	2.017
Gymnasium	129,2	13,8	0,74	4.647
IHR Schule – H	97,9	15,9	-0,92	300
IHR Schule – R	109,1	12,7	-0,32	314
Gesamtschule II	102,4	15,0	-0,68	1.877
Gesamtschule I	114,0	14,6	-0,06	1.257
<i>insgesamt</i>	<i>115,2</i>	<i>18,9</i>	<i>---</i>	<i>11.683</i>

Darüber hinaus ist es erwähnenswert, dass hier die Leistungen innerhalb der Schulformen bzw. Kursniveaus ähnlich breit wie beim Leseverständnistest gestreut sind. In der folgenden Abbildung 6 sind für die insgesamt 11.683 berücksichtigten Schülerinnen und Schüler die Verteilungen der schulform- bzw. kursniveaubezogenen Ergebnisse dargestellt.

Abbildung 6 Sprachverständnis: Verteilung der Lernstände am Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Oberhalb des Schnittpunkts zwischen den Verteilungen der Realschulen und der Gymnasien (Lernstandswerte  $> 112$ ) befinden sich 14,4 Prozent der Schülerinnen und Schüler aus Hauptschulklassen, 49,0 Prozent aus Realschulklassen, 14,0 Prozent aus H-Kursen und 41,1 Prozent aus R-Kursen an Integrierten Haupt- und Realschulen sowie 25,7 Prozent (Kurse II) bzw. 60,1 Prozent (Kurse I) aus Gesamtschulen. An den Gymnasien haben 91,4 Prozent der Schülerinnen und Schüler diesen „typisch gymnasialen“ Kennwert erreicht oder übertroffen.<sup>19</sup>

### 3.2.2.3 Rechtschreiben

<sup>19</sup> Den gymnasialen Leistungsdurchschnitt (Skalenwert 129) haben 49,3 Prozent der Gymnasiastinnen und Gymnasiasten übertroffen. Die entsprechenden Werte für die Schülerschaft der Hauptschulklassen (1,4 Prozent), der Realschulklassen (7,7 Prozent), der H-Kurse (1,3 Prozent) und der R-Kurse (4,8 Prozent) an Integrierten Haupt- und Realschulen sowie der Gesamtschulen (Kurse II: 2,6 Prozent; Kurse I: 12,1 Prozent) sind deutlich niedriger als die entsprechenden Werte für die Fachleistung Mathematik und das Leseverständnis.

Die Rechtschreibkenntnisse der Schülerinnen und Schüler wurden in LAU 9 in der Gesamtgruppe als *Rechtschreibwissen* und zusätzlich in einer Stichprobe als *Rechtschreibkönnen* gemessen.

Das *Rechtschreibwissen* wurde im *SL-HAM 8/9* über einen schulform- bzw. kursniveauübergreifenden Test für alle Schülerinnen und Schüler in der anspruchsvollen Form der Suche von Fehlschreibungen geprüft. Der vorgegebene Text mit insgesamt 285 Wörtern ist ein Sachtext zum Thema „*Jugend und Freizeit an der Jahrtausendwende*“, in dem das Freizeitverhalten von Jugendlichen unter soziologischen Aspekten thematisiert wird. Insgesamt enthalten 29 der 285 Wörter einen Rechtschreibfehler; die Fehlschreibungen sollten unterstrichen werden.

Um einen Einblick in die Anforderungsstruktur des Rechtschreibtests zu gewinnen, ist im Folgenden für eine Auswahl von sechs der falsch vorgegebenen Wörter aufgeführt, wie viel Prozent der Neuntklässlerinnen und Neuntklässler sie als Fehlschreibungen erkannt haben.<sup>20</sup> Von den insgesamt 11.378 Schülerinnen und Schülern, die den Rechtschreibtest bearbeitet haben und einer der sieben Schulform- bzw. Kursniveaugruppierungen der Klassenstufe 8 zugeordnet werden konnten, haben ca. 75 Prozent die folgenden beiden fett und kursiv gesetzten Wörter als Fehlschreibungen erkannt:

- .. und vor allem ohne **Leistungsdruk** (77,2 %),
- .. stärker an der Arbeit und der Familie **orientirt**, .. (73,6 %).

Ca. 50 Prozent der Schülerinnen und Schüler war geläufig, dass das Partizip I bei „*anstrengend*“ mit dem Suffix *-d* endet; ebenfalls ca. die Hälfte der Neuntklässlerinnen und Neuntklässler hat die fälschliche Zusammenschreibung zweier getrennt zu schreibender Wörter erkannt:

- *Freizeit ist für die heutige Jugend ganz schön **anstrengent***. (52,7 %);
- **Dazugehört** dann die entsprechend modische Sportkleidung, ... (52,7 %).

---

<sup>20</sup> Da der Rechtschreibtest für alle Schulformen bzw. Kursniveaus einheitlich war, konnte auf eine Rasch-Skalierung verzichtet werden; sie hätte bei der großen Zahl der Items auch zu Problemen geführt. Stattdessen wurden bei der Auswertung fälschlich unterstrichene Richtigschreibungen so berücksichtigt, dass jeder erkannte Rechtschreibfehler mit dem Faktor „Anzahl richtiger Markierungen“ / „Anzahl aller Markierungen“ gewichtet wurde.

Schließlich gibt es Fehlschreibungen, die nur von ungefähr einem Viertel der getesteten Schülerinnen und Schüler erkannt wurden. In dem folgenden Fall:

- Eine starke *Enttäuschung* am Staat ... (28,2 %)

gilt es zu wissen, dass bei der Zusammensetzung von Wörtern die Morpheme an der Fugengrenze beibehalten werden, und bei einem der schwierigsten Wörter muss das Stammprinzip beherrscht werden:

- „...“, *ergenz* Horst Opaschowski. (25,7 %).

74,3 Prozent der Schülerinnen und Schüler haben nicht erkannt, dass in der Ableitung der Stammvokal *-a-* erhalten bleibt und in *-ä-* umgelautet wird.

Bei anteiliger Berücksichtigung fälschlich unterstrichener Richtigschreibungen ergibt sich – bei insgesamt 29 vorgegebenen Fehlern – für die Schülerinnen und Schüler insgesamt ein Durchschnitt von 10,4 richtig erkannten Fehlschreibungen. Die Gymnasiastinnen und Gymnasiasten als leistungsstärkste Gruppe haben die Hälfte der vorgegebenen Fehlschreibungen identifiziert, die Hauptschülerinnen und Hauptschüler als leistungsschwächste Gruppe etwa jedes fünfte falsch geschriebene Wort.

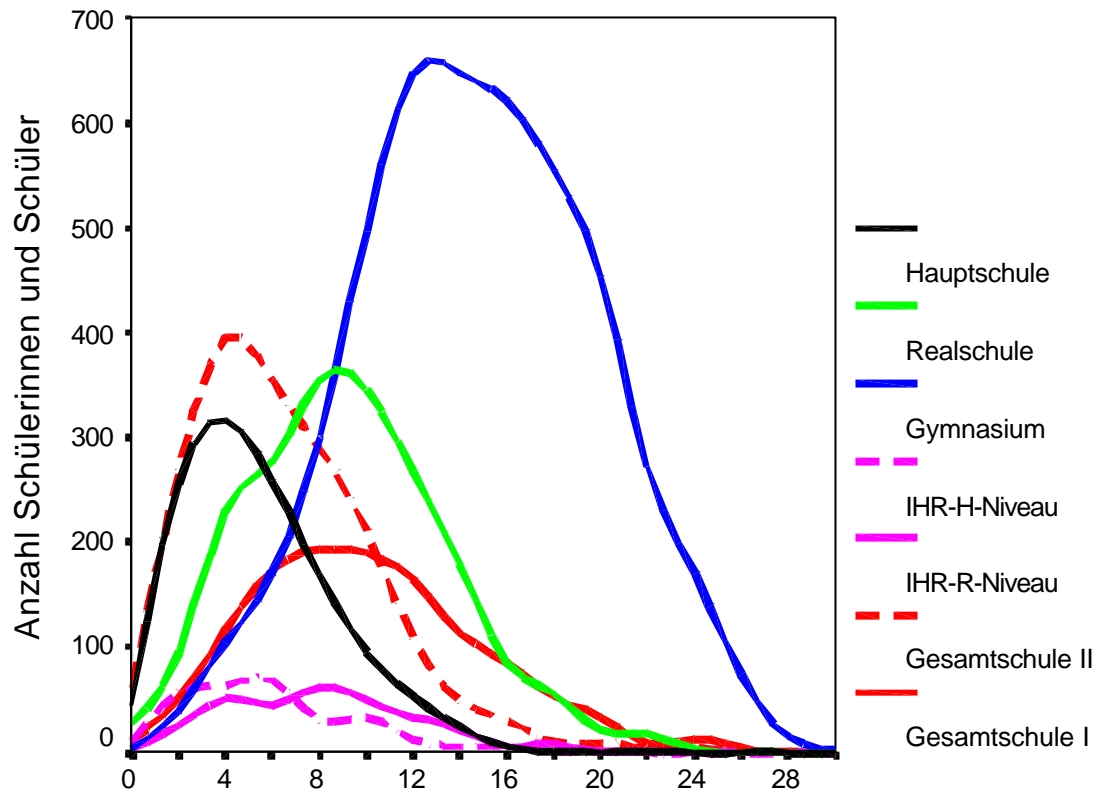
Tabelle 11 Rechtschreibwissen am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	5,4	3,2	-0,86	1.227
Realschule	9,1	4,3	-0,22	1.961
Gymnasium	14,5	5,1	0,71	4.574
IHR Schule – H	5,9	3,7	-0,78	290
IHR Schule – R	8,0	4,0	-0,41	309
Gesamtschule – Kurs II	6,5	3,9	-0,67	1.803
Gesamtschule – Kurs I	10,0	4,8	-0,07	1.214
<i>insgesamt</i>	<i>10,4</i>	<i>5,8</i>	<i>---</i>	<i>11.378</i>

Es zeigen sich deutlich linkssteile Verteilungen, die die Aufgaben insgesamt als eher schwierig für die nichtgymnasiale Schülerschaft kennzeichnen (vgl. dazu Abbildung 7). Den gymnasialen Leistungsdurchschnitt (14,5 Punkte) haben 0,8 Prozent der Schülerinnen und Schüler aus Hauptschulklassen, 9,7 Prozent aus Realschulklassen, 48,0 Prozent aus Gymnasialklassen, 3,1 Prozent aus H-Kursen und 5,5 Prozent aus R-Kursen an Integrierten Haupt- und Realschulen

sowie 3,7 Prozent (Kurse II) bzw. 16,9 Prozent (Kurse I) aus Gesamtschulen erreicht oder übertroffen.

Abbildung 7 Rechtschreibwissen: Verteilung der Lernstände am Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Passives Rechtschreibwissen Ende Klassenstufe 8

Der Schnittpunkt zwischen den Verteilungskurven für die Realschülerinnen und Realschüler auf der einen Seite, für die Gymnasiastinnen und Gymnasiasten auf der anderen Seite – hier wiederum gleichbedeutend mit der unteren Grenze des „typisch gymnasialen Bereichs“ – liegt bei 8,4 Punkten. Darüber befanden sich 17,7 Prozent der Schülerinnen und Schüler aus Hauptschulklassen, 53,2 Prozent aus Realschulklassen, 88,0 Prozent aus Gymnasialklassen, 23,8 Prozent aus H-Kursen und 43,8 Prozent aus R-Kursen an Integrierten Haupt- und Realschulen sowie 26,9 Prozent (Kurse II) bzw. 57,5 Prozent (Kurse I) aus Gesamtschulen. Die Relationen der Schulformen bzw. Kursniveaus zueinander liegen also in dem für die bereits vorgestellten Fachleistungen üblichen Rahmen.

Für den Bereich des *Rechtschreibkönnens* wurde die *Hamburger Schreibprobe für fünfte bis neunte Klassen – HSP 5-9* eingesetzt. Hier liegt keine Vollerhebung vor, sondern eine Stichprobe von 1.473 getesteten Schü-

lerinnen und Schülern.<sup>21</sup> Im Folgenden ist aus der Gesamtzahl der 57 diktierten Wörter die Liste derjenigen 15 Wörter aufgeführt, die von weniger als 75 Prozent der Schülerinnen und Schüler richtig geschrieben worden sind und damit als eher schwierig gelten können.

*Dieses Wort wurde von . . . Prozent der Schülerinnen und Schüler richtig geschrieben:*

• Fußballmannschaft	38,1 %	• dass	67,8 %
• Reißverschluss	41,4 %	• Nilpferd	67,9 %
• Sekretärin	52,8 %	• Tierärztin	70,3 %
• Fahrradschloss	53,2 %	• Fernsehprogramm	71,2 %
• Gießkanne	62,2 %	• Päckchen	71,6 %
• Dauernd	64,2 %	• kaputt	73,4 %
• Bohrmaschine	66,4 %	• stöhnt	74,4 %
• Schiedsrichter	67,2 %		

Die folgende Tabelle 12 zeigt, dass – über alle getesteten Schülerinnen und Schüler gerechnet – im Durchschnitt ca. neun der 57 diktierten Wörter fehlerhaft geschrieben wurden. Bei den Gymnasiastinnen und Gymnasiasten als der leistungsstärksten Gruppe waren es ca. sechs, bei den Hauptschülerinnen und Hauptschülern, der leistungsschwächsten Gruppe, ca. 19 Wörter.

Tabelle 12 Rechtschreibkönnen am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

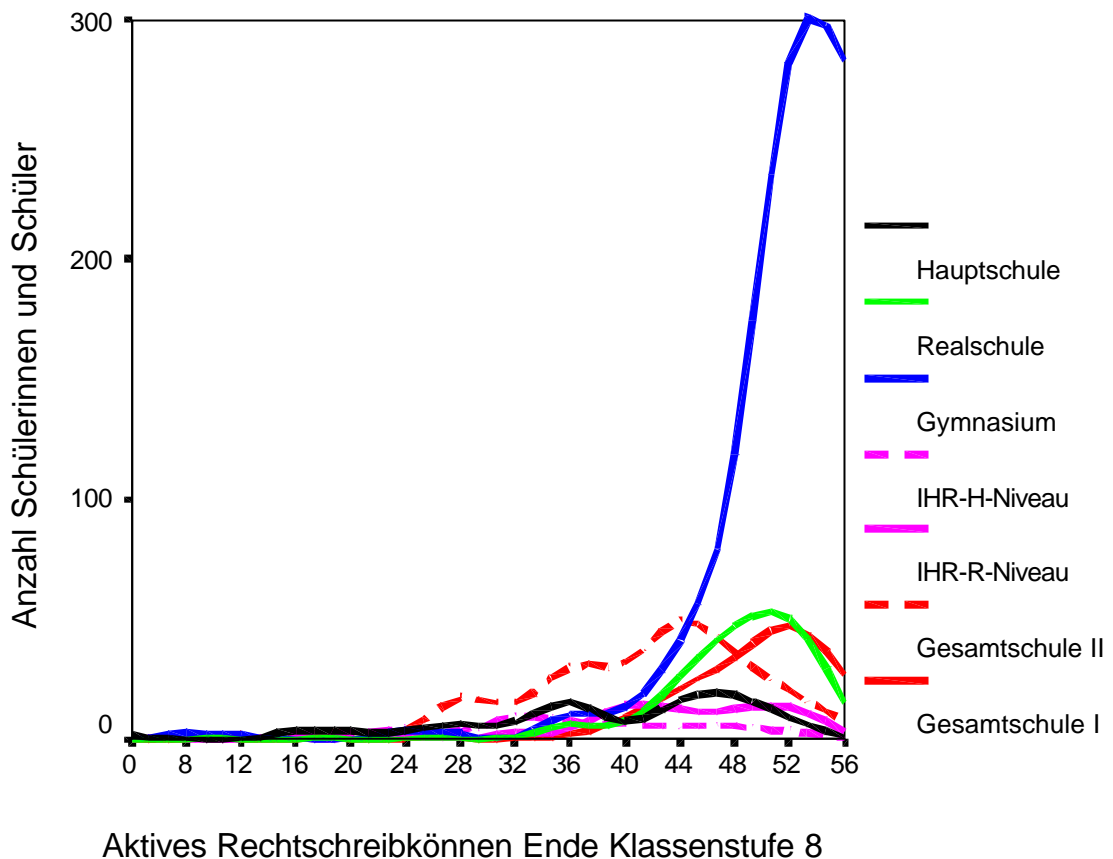
	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	38,5	10,5	-1,14	89
Realschule	48,0	5,0	0,05	152
Gymnasium	51,2	5,4	0,45	756
IHR Schule – H	35,7	8,8	-1,49	45
IHR Schule – R	44,3	6,6	-0,41	61
Gesamtschule II	40,8	8,2	-0,85	214
Gesamtschule I	49,1	4,8	0,19	140
<i>insgesamt</i>	<i>47,6</i>	<i>8,0</i>	<i>---</i>	<i>1.457</i>

<sup>21</sup> Wie Kontrollrechnungen an Hand der sonstigen Leistungen im Fach Deutsch ergeben haben, kann diese Stichprobe nur für die Gymnasien und die H-Kurse der Integrierten Haupt- und Realschulen als repräsentativ gelten. Vor allem für die Realschulen ( $d = 0,29$ ), aber auch für die R-Kurse der Integrierten Haupt- und Realschulen ( $d = 0,16$ ) und die Gesamtschulen (Kurse II:  $d = 0,08$ ; Kurse I:  $d = 0,07$ ) werden die tatsächlichen Leistungen durch die Stichprobenergebnisse systematisch überschätzt.



Zu beachten ist hier, dass die Fallzahlen in einzelnen Gruppen – insbesondere im Haupt- und Realschulbereich – teilweise recht niedrig sind. Auch war in einigen Fällen die Schulform bzw. das Kursniveau am Ende der Klassenstufe 8 nicht eindeutig zu ermitteln. Die schulform- bzw. kursniveaubezogenen Ergebnisse, deren Verteilungen in der folgenden Abbildung 8 noch einmal aufgeführt sind, sind deshalb mit relativ großen Stichprobenfehlern behaftet.<sup>22</sup>

Abbildung 8 Rechtschreibwissen: Verteilung der Lernstände zum Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Unter den gekennzeichneten Bedingungen ist es nicht möglich, wie für die vorigen Fachleistungen einen so genannten „typisch gymnasialen Bereich“ zu be-

<sup>22</sup> Für eine grobe Abschätzung des Standardstichprobenfehlers für die Mittelwerte (SSF) kann die Formel  $SSF = \text{Standardabweichung} / \sqrt{N/3}$  benutzt werden. Hiernach wäre beispielsweise SSF (Hauptschulklassen) ca. 1,9 und SSF (IHR-Hauptschulkurse) ca. 2,3. Angesichts dieser Schätzungenauigkeit kann es nicht als gesichert gelten, dass sich, um bei den Beispielen zu bleiben, die Rechtschreibleistungen in Hauptschulklassen und in IHR-Hauptschulkursen in den beiden Gesamtpopulationen unterscheiden, da sich die Fehlerbereiche deutlich überschneiden. Diese Faustformel ist selbstverständlich bei Bedarf durch präzisere statistische Techniken zu ersetzen. Im Übrigen ist noch der systematische Stichprobenfehler (vgl. die vorige Anmerkung) zu berücksichtigen.

stimmen. Alternativ dazu soll hier vor dem Hintergrund, dass die Verteilung der Hauptschülerinnen und Hauptschüler zweigipflig ist und bei dem Skalenwert 40 eine Teilung zwischen leistungsstärkeren und leistungsschwächeren Jugendlichen anzeigt, der Bereich unterhalb bzw. oberhalb von 40 Richtigschreibungen beleuchtet werden. Mehr als 40 Richtigschreibungen – d. h. weniger als 17 Fehler in den 57 diktierten Wörtern – haben 55,1 Prozent der Schülerinnen und Schüler aus den Hauptschulklassen erzielt, 92,1 Prozent aus den Realschulklassen, 96,7 Prozent aus den Gymnasialklassen, 35,6 Prozent aus den H-Kursen und 77,1 Prozent aus den R-Kursen an Integrierten Haupt- und Realschulen sowie 62,2 Prozent (Kurse II) bzw. 97,9 Prozent (Kurse I) aus den Gesamtschulen. Auch diese Prozentwertschätzungen sind stichprobenfehlerbehaftet.

Nimmt man den gymnasialen Durchschnitt von 51,2 Richtigschreibungen als Grenzwert, so haben diesen 4,5 Prozent der Schülerinnen und Schüler aus Hauptschulklassen übertroffen, 27,6 Prozent aus Realschulklassen, 58,1 Prozent aus Gymnasialklassen („Deckeneffekt“), 2,2 Prozent aus H-Kursen und 16,4 Prozent aus R-Kursen an Integrierten Haupt- und Realschulen sowie 7,5 Prozent (Kurse II) bzw. 38,8 Prozent (Kurse I) aus Gesamtschulen.

### 3.2.2.4 Textproduktion

Einer nur teilweise repräsentativen<sup>23</sup> Stichprobe von Schulklassen – denselben, die auch den Rechtschreibtest *HSP 5-9* bearbeitet haben – wurde die Aufgabe gestellt, einen Beschwerdebrief an eine fiktive Computerfirma zu schreiben, die ein von ihrem Angebot abweichendes Produkt mit einer überhöhten Rechnung geliefert hatte. Die Auswertung der Schülertexte erfolgte jeweils durch mindestens zwei Kodiererinnen bzw. Beurteilerinnen, in 20 Prozent der Fälle zusätzlich durch eine dritte. Bewertet wurden:

- die formale Vollständigkeit und Korrektheit des Schreibens ([ggf. fiktive] Absenderanschrift, Datum, Empfängeranschrift, Anrede, Gruß, Zeichnung u. a., jeweils „korrekt vorhanden“ [Code 1] vs. „fehlerhaft“ bzw. „fehlend“ [Code 0]), unberücksichtigt blieben Interpunktion und Grammatik.

---

<sup>23</sup> Es handelt sich hier um dieselbe Stichprobe wie zur Erhebung des Rechtschreibkönnens. Damit gelten hier dieselben Einschränkungen bzgl. der Repräsentativität der Ergebnisse.

- die inhaltliche Vollständigkeit und Korrektheit des Schreibens (Liste der Mängel, konkrete Forderung nach Nachbesserung oder Rücknahme sowie Wertberichtigung nach analogem Kodierschema),
- die Angemessenheit des Aufbaus, des Ausdrucks und des Stils (jeweils auf einer Skala von 1 bis 5).

Die paarweisen Übereinstimmungen zwischen den ausgezählten Codes waren ausgesprochen hoch (Cronbachs Alpha > 0,94), diejenige zwischen den qualitativen Beurteilungen deutlich geringer (Cronbachs Alpha > 0,68). Es ließ sich jedoch im Einvernehmen mit Experten der Behörde für Schule, Jugend und Berufsbildung eine psychometrisch akzeptable Gesamtskala bilden, die die drei genannten Kodierungs- bzw. Beurteilungsaspekte gleich gewichtet und dabei jeweils auf allen verfügbaren Auswerterdatensätzen beruht. Der Wertebereich dieser Skala reicht von 1 bis 100.

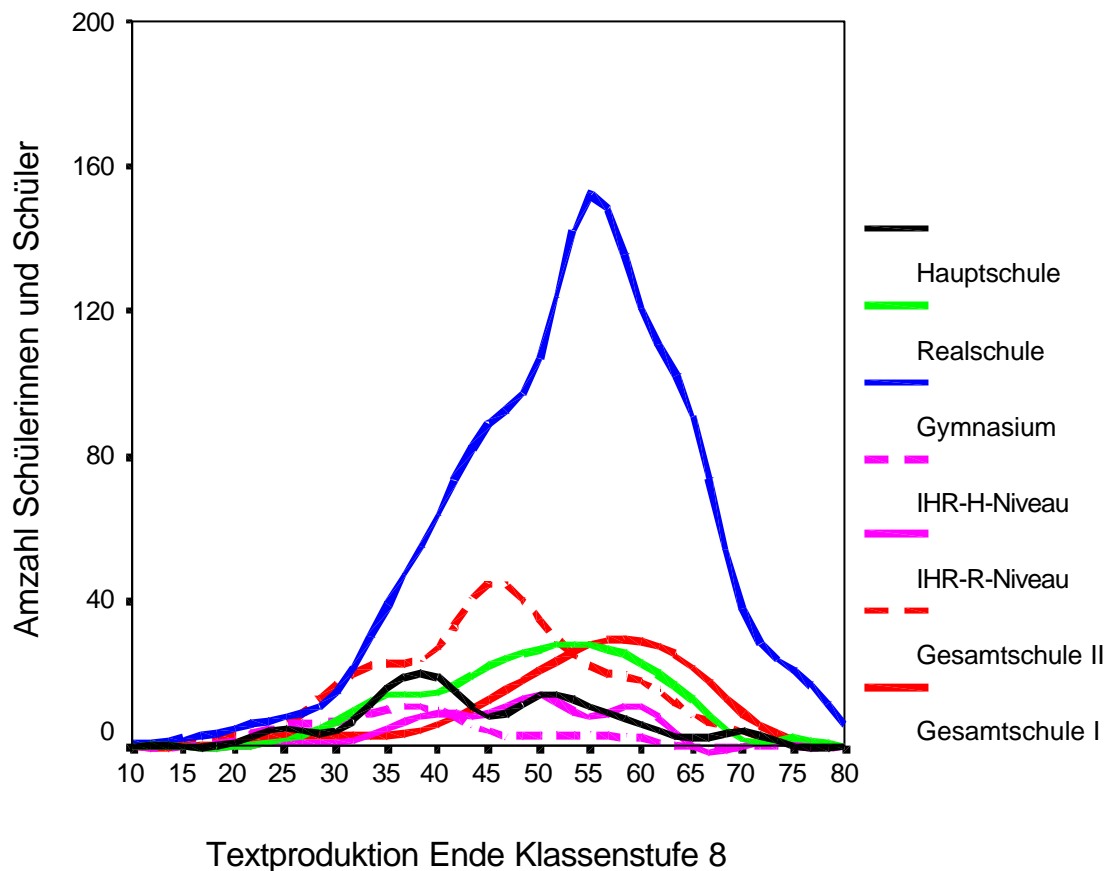
Die Verteilung der so beurteilten Textqualität ist, getrennt nach Schulform bzw. Kursniveau, in Abbildung 9 dargestellt. Bereits ein erster Blick auf die Abbildung zeigt, dass es, ähnlich wie beim Rechtschreibkönnen, nicht möglich ist, einen ‚typisch gymnasialen Leistungsbereich‘ zu definieren, weil die Verteilung für die Gymnasien alle übrigen Verteilungen einhüllt. Gleichzeitig ist die Leistungsverteilung in den Kursen I der Gesamtschulen unter diesem Aspekt offenbar noch günstiger als die an den Gymnasien (vgl. die jeweiligen Gipfelpunkte).<sup>24</sup>

Danach erscheint es wiederum als sinnvoll, zwischen denjenigen zu unterscheiden, die ausweislich der zweigipfligen Verteilung für die Hauptschulen erhebliche Schwierigkeiten mit der Textproduktion haben (Skalenwerte < 45), und denen, für die dies offenbar nicht zutrifft. Gemessen an diesem Kriterium sind zu der ersten Gruppe 44,2 Prozent der Hauptschülerinnen und Hauptschüler, 20,4 Prozent der Realschülerinnen und Realschüler sowie 13,8 Prozent der Gymnasiastinnen und Gymnasiasten zu rechnen. Ferner gehören 72,5 Prozent in den H-Kursen und 20,6 Prozent in den R-Kursen der Integrierten Haupt- und Realschulen sowie 35,0 Prozent in den Kursen II und 10,1 Prozent in den Kursen I der Gesamtschulen dazu. Zu beachten ist allerdings wieder neben den systematischen Verzerrungen in der Stichprobe die äußerst niedrige Probandenzahl an den Integrierten Haupt- und Realschulen ( $N = 106$ ), die hier mit besonders großen Stichprobenfehlern verbunden ist.

---

<sup>24</sup> Auch unter Berücksichtigung der systematischen Verzerrungen in der Stichprobe ist der Mittelwert für das Deutsch-Kursniveau I der Gesamtschulen etwas höher als der für die Gymnasien.

Abbildung 9 Textproduktion: Verteilung der Lernstände zum Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Die wichtigsten statistischen Kennwerte der schulform- bzw. kursniveau-spezifischen Verteilungen sind in Tabelle 13 erfasst.

Tabelle 13 Textproduktion am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	44,7	11,4	-0,50	90
Realschule	49,9	10,5	-0,06	155
Gymnasium	53,1	11,5	0,22	759
IHR Schule – H	37,5	9,7	-1,11	47
IHR Schule – R	48,5	9,3	-0,18	59
Gesamtschule II	45,5	11,0	-0,43	211
Gesamtschule I	54,6	10,6	0,34	140
<i>insgesamt</i>	<i>50,6</i>	<i>11,8</i>	<i>---</i>	<i>1.461</i>

Abweichend von allen bisher betrachteten Aspekten der fachbezogenen Schulleistungen, zeigt sich hier eine relative Überlegenheit der Schülerinnen und Schüler aus den Deutsch-Kursen I der Gesamtschulen gegenüber den Gymnasiastinnen und Gymnasiasten. Dies mag darin begründet sein, dass das Verfassen formal korrekter Schreiben an den Gymnasien wenig üblich ist. Es gibt für diese Vermutung immerhin ein indirektes Indiz. Betrachtet man nämlich nur die Angemessenheit des Aufbaus, des Ausdrucks und des Stils (vgl. Tabelle 14), so ist die gewohnte Reihenfolge wieder

63+, hergestellt. Ähnlich, wenn auch schwächer ausgeprägt, sieht es hinsichtlich der inhaltlichen Qualität der Texte aus (ohne Tabelle). Demnach darf angenommen werden, dass in manchen Schulformen am Ende der Klassenstufe 8 bereits eine gewisse Vertrautheit mit den Formalia eines Geschäftsbriefes die Gesamtqualität der Texte positiv beeinflusst hat.

Dieser Effekt hat seinen Niederschlag auch in den Werten für die Anteile derjenigen gefunden, die den gymnasialen Durchschnitt (53,1 Skalenpunkte) erreicht oder überschritten haben. Dies gilt für 25,6 Prozent an den Hauptschulen, 41,3 Prozent an den Realschulen, 53,8 Prozent an den Gymnasien, 6,4 Prozent in den H-Kursen und 30,5 Prozent in den R-Kursen der Integrierten Haupt- und Realschulen sowie für 23,7 Prozent in den Kursen II und 60,0 Prozent in den Kursen I der Gesamtschulen.

Tabelle 14 Textproduktion am Ende der Klassenstufe 8 (Aufbau, Ausdruck und Stil): Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	<b>Mittelwert</b>	<b>Standard- abweichung</b>	<b>Effektstärke <i>d</i></b>	<b><i>N</i></b>
Hauptschule	10,2	3,3	-0,95	90
Realschule	13,4	3,1	-0,12	155
Gymnasium	15,3	3,3	0,39	759
IHR Schule – H	9,5	2,6	-1,16	47
IHR Schule – R	13,6	3,3	-0,07	59
Gesamtschule II	11,4	3,1	-0,66	211
Gesamtschule I	14,0	3,0	0,05	140
<i>insgesamt</i>	<i>13,8</i>	<i>3,7</i>	<i>---</i>	<i>1.461</i>

### 3.2.2.5 Zusammenfassung der Fachleistung Deutsch

Angesichts der Vielzahl gemessener Leistungsaspekte im Fach Deutsch ist die Frage sinnvoll, ob sich hier ein besonderer Fachindex bilden lässt, der die verschiedenen Teilleistungen zusammenfasst. Erste Voraussetzung dafür ist es, dass die verschiedenen erfassten Aspekte ausreichend miteinander korrelieren. Wie die Tabelle 15 zeigt, ist dies ganz überwiegend der Fall.

Tabelle 15 Fachleistung Deutsch: Interkorrelationen

	Lesever- ständnis	Sprache	Rechtschreib- wissen	Rechtschreib- können	Text- produktion
Leseverständnis	---	0,77	0,57	0,54	0,43
Sprache	---	---	0,65	0,60	0,41
Recht- schreibwissen	---	---	---	0,63	0,37
Recht- schreibkönnen	---	---	---	---	0,46
Textproduktion	---	---	---	---	---

Die niedrigeren Korrelationen vor allem der Textproduktion mit den übrigen Variablen deuten darauf hin, dass den Schülerinnen und Schülern mit der entsprechenden Aufgabe zusätzliche, eigenständige Kompetenzen abverlangt wurden. Der Textproduktion kommt also eine gewisse Sonderstellung zu, die wohl nicht nur auf den in den Bewertungen enthaltenen Messfehler zurückgeht, sondern zumindest auch darauf, dass für das Schreiben eines formal, inhaltlich und stilistisch überzeugenden Briefes zusätzliche Kompetenzen gefordert sind. In jedem Falle erscheint es nicht ratsam, die Textproduktion und die Ergebnisse aus der *HSP 5-9* in dem gesuchten Index mit zu berücksichtigen, da jeweils nur für knapp 1.500 Schülerinnen und Schüler einschlägige Werte vorliegen. Der zu bildende *Index für die Fachleistung Deutsch* erfasst somit nur einen Teil der für dieses Fach maßgeblichen Leistungen.

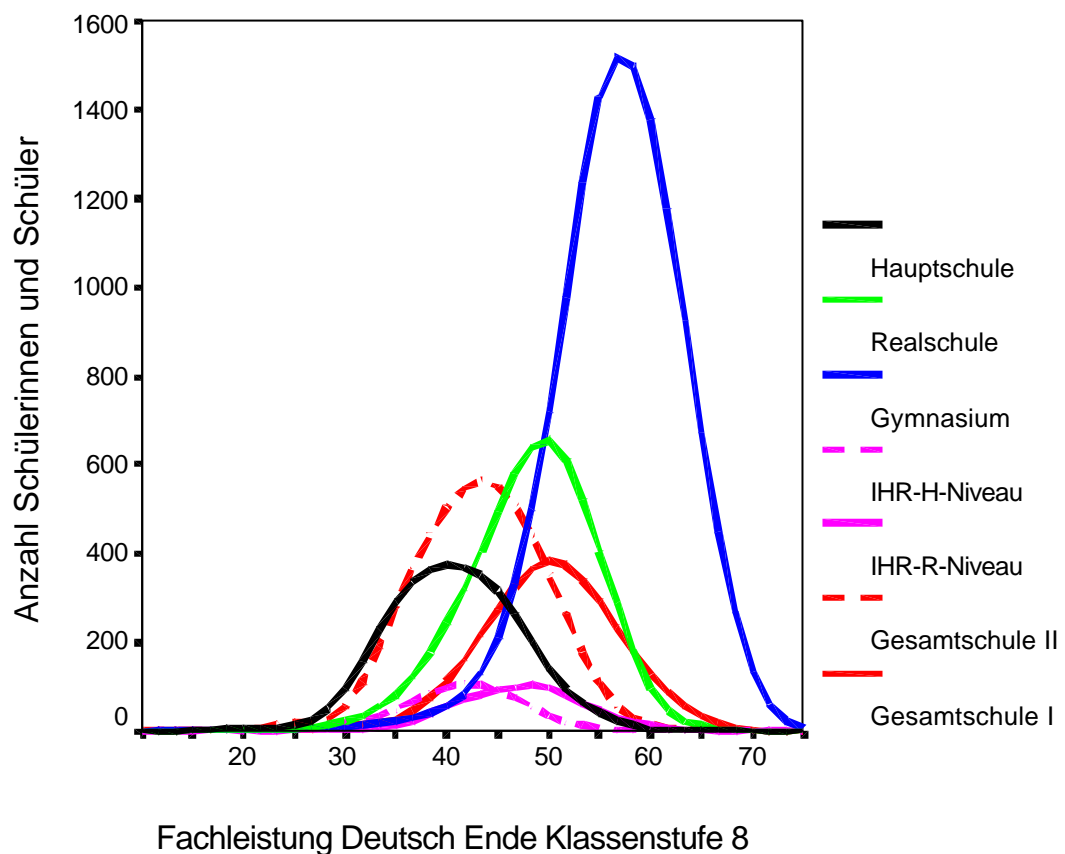
Da die einbezogenen drei Skalen *Leseverständnis*, *Sprache* und *Rechtschreibwissen* unterschiedliche Wertebereiche haben, sind sie zunächst vergleichbar zu machen.<sup>25</sup> Weiterhin sind ein neuer Werteabstand und Wertebereich zu definieren. Gewählt wurde hier ein Mittelwert von 50 und eine

<sup>25</sup> Diese Standardisierung wurde über eine sog. „z-Transformation“ erzielt. Die Berücksichtigung fehlender Werte erfolgte so, dass jeweils der Durchschnitt aus den maximal drei vorhandenen Werten gebildet wurde.

Standardabweichung von 10 Punkten.<sup>26</sup> Abbildung 10 zeigt das Ergebnis des Schulformen- bzw. Kursniveauvergleichs in grafischer Form.

Wegen der Abhängigkeit der Befunde von den zugehörigen Untertests ist auf der Grundlage dieser Indexbildung nicht mit neuen Einsichten zu rechnen. Die entsprechenden numerischen Befunde (vgl. Tabelle 16) sind hingegen insofern von Bedeutung, als es die Leistung im Fach Deutsch insgesamt und nicht diejenige in einem bestimmten Teilleistungsbereich ist, die ggf. den Ausschlag für eine Kurszuweisung gibt.

Abbildung 10 Fachleistung Deutsch (Index aus Leseverständnis, Sprache, Rechtschreibwissen): Verteilung der Lernstände am Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



<sup>26</sup> Dass für die Standardabweichung der Sollwert von 10 Punkten nicht exakt getroffen wurde, hängt mit dem Auftreten fehlender Werte zusammen.

Tabelle 16 Fachleistung Deutsch (Index aus Leseverständnis, Sprache, Rechtschreibwissen) am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	40,8	6,4	-1,04	1.281
Realschule	48,2	6,3	-0,20	2.017
Gymnasium	56,8	6,4	0,77	4.651
IHR Schule – H	41,6	6,3	-0,94	306
IHR Schule – R	47,1	5,8	-0,32	315
Gesamtschule II	43,3	6,5	-0,76	1.895
Gesamtschule I	49,9	7,0	-0,01	1.266
<i>insgesamt</i>	<i>50,0</i>	<i>8,9</i>		<i>11.731</i>

Gemessen an der unteren Grenze des *typisch gymnasialen Bereichs* ( $\geq 49$  Skalenpunkte) haben an den Hauptschulen 10,0 Prozent, an den Realschulen 48,0 Prozent, an den Gymnasien 90,6 Prozent, an den Integrierten Haupt- und Realschulen 10,5 Prozent (H-Kurse) bzw. 40,0 Prozent (R-Kurse) und an den Gesamtschulen 18,6 Prozent (Kurse II) bzw. 56,2 Prozent (Kurse I) diesen Wert erreicht oder überschritten. Verglichen mit dem anspruchsvolleren Kriterium des gymnasialen Durchschnitts ( $\geq 56,8$  Skalenpunkte) gilt dies für 0,5 Prozent der Schülerinnen und Schüler an den Hauptschulen, 6,6 Prozent an den Realschulen, 52,1 Prozent an den Gymnasien, 1,6 Prozent (H-Kurse) bzw. 5,4 Prozent (R-Kurse) an den Integrierten Haupt- und Realschulen und 1,6 Prozent (Kurse II) bzw. 15,1 Prozent (Kurse I) an den Gesamtschulen.

### 3.2.3 Fachleistung Englisch

Der eine Unterrichtsstunde umfassende *SL-HAM 8/9-Englishtest* erhebt das Hörverständnis und – im Format des C-Tests – Wort-, Rechtschreib-, Syntax- und Grammatikkenntnisse der Schülerinnen und Schüler im Textzusammenhang.

Den Schülerinnen und Schülern wurde im Rahmen des *Hörverständnistests* über Kassette ein in englischer Sprache geführtes ca. drei Minuten langes Gespräch zwischen einem Radiomoderator und zwei Jugendlichen vorgespielt, zu dem sie anschließend vier in deutscher und drei in englischer Sprache vorgegebene Verständnisfragen (Multiple-Choice-Format) beantworten sollten. In



der folgenden Tabelle 17 ist für die Schülerinnen und Schüler insgesamt und schulform- bzw. kursniveauspezifisch aufgeführt, wie viele der sieben Fragen durchschnittlich richtig beantwortet wurden und wie breit die Leistungen um den Mittelwert streuen. Über die Abstände des jeweiligen schulform- bzw. kursniveauspezifischen Mittelwerts zum Mittelwert insgesamt gibt die standardisierte Effektstärke  $d$  Auskunft.

Insgesamt wurden durchschnittlich knapp vier der sieben Fragen zum Text richtig beantwortet. Schulform- bzw. kursniveauspezifisch zeigen sich hier – das lässt sich gut über den Vergleich der Effektstärken belegen – tendenziell ähnliche Verhältnisse wie in den vorher berichteten Fachleistungen. Wegen der beschränkten Anzahl an Testaufgaben, vor allem aber wegen technischer Durchführungsschwierigkeiten, ergab sich insgesamt eine sehr niedrige Reliabilität des Kurztests (Cronbachs Alpha = 0,44).

Tabelle 17 Englisch: Hörverständnis am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke $d$	$N$
Hauptschule	2,9	1,2	-0,63	1.119
Realschule	3,4	1,4	-0,31	1.864
Gymnasium	4,8	1,5	0,56	4.233
IHR Schule – H	2,9	1,4	-0,63	262
IHR Schule – R	3,3	1,4	-0,38	300
Gesamtschule II	3,0	1,3	-0,56	1.542
Gesamtschule I	3,7	1,5	-0,13	1.248
<i>insgesamt</i>	<i>3,9</i>	<i>1,6</i>	<i>---</i>	<i>10.568</i>

Auch die Validität der Messung ist nicht zufriedenstellend, wie ein Vergleich der Bestimmtheitsmaße  $Eta^2$  für die Schulform bzw. das Kursniveau zeigt: Dieses Maß liefert für den Untertest *Englisch Hörverständnis* mit  $Eta^2 = 0,25$  einen wesentlich geringeren Wert als für den Untertest *Englisch C-Test* ( $Eta^2 = 0,46$ ). Die zu erwartenden Leistungsdifferenzen zwischen den verschiedenen Bildungsgängen wurden also durch den Hörverständnistest nur unzulänglich erfasst. Deshalb werden hier keine weiteren Analysen vorgenommen.

Die Feststellung der globalen Verständnisleistungen sowie der Wort- und Grammatikkenntnisse erfolgte mit insgesamt sechs Texten, in denen – abgesehen vom ersten und letzten Satz – jedes vierte Wort nur zur Hälfte ausgeschrieben war und von den Schülerinnen und Schülern vervollständigt werden musste. Insgesamt enthält der *Englisch: C-Test* 142 Wortergänzungen, von denen 133 bei der Auswertung den teststatistischen Anforderungen der Rasch-Skalierung genügten.<sup>27</sup> In der Abbildung 11 sind auf der rechten Seite der Rasch-Skala die 133 Aufgaben – geordnet nach ihrem Schwierigkeitsgrad – und auf der linken Seite die darauf bezogene Verteilung der Schülerleistungen eingetragen.

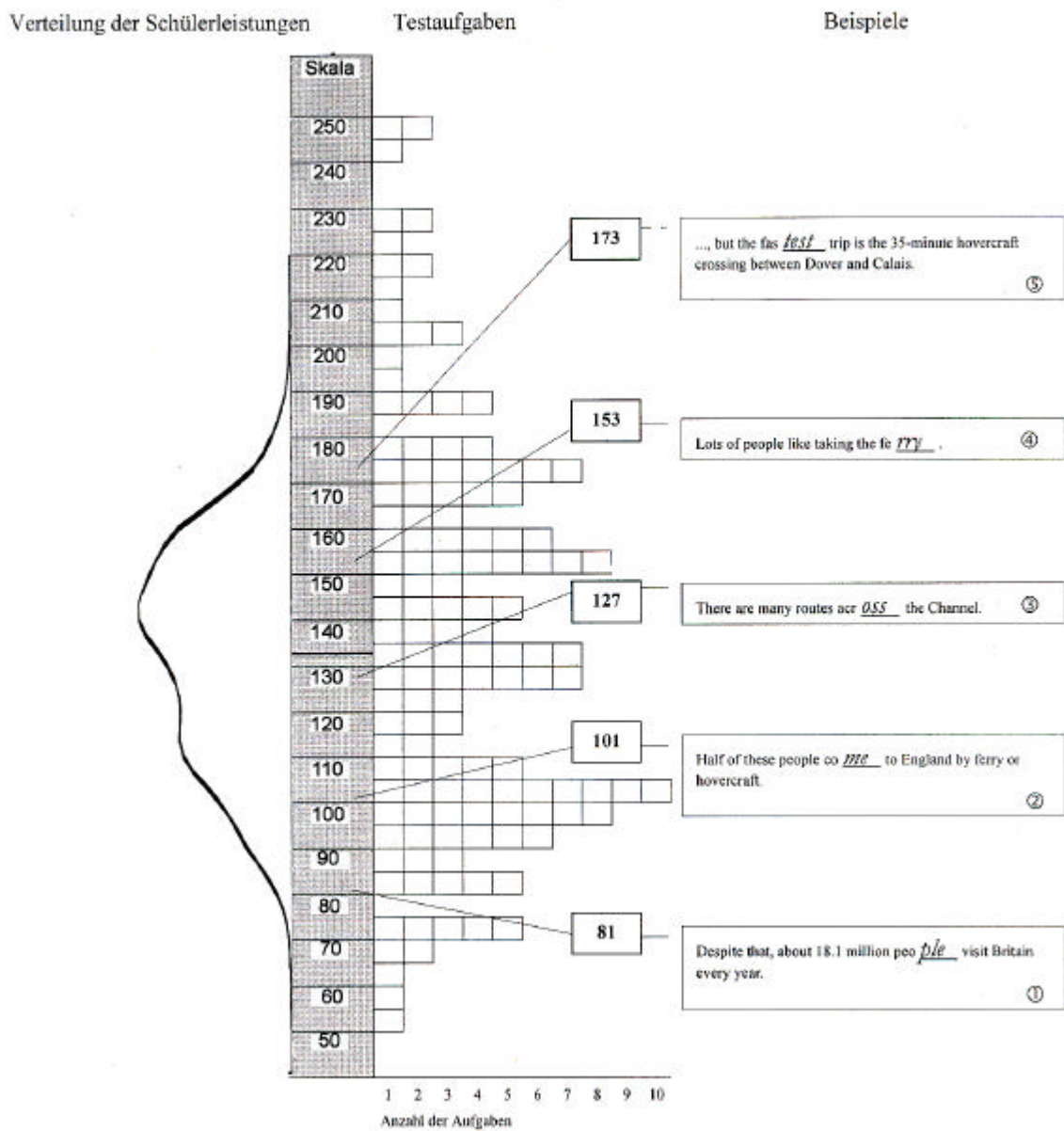
Die Testaufgaben verteilen sich bei einer mittleren Schwierigkeit um den Wert 120 breit über den Bereich von 50 bis 250 Skalenpunkten; der Mittelwert für die Schülerleistungen liegt mit 133 Skalenpunkten ( $s = 24,7$ ) demgegenüber etwas höher. Wie man der Abbildung entnehmen kann, sind – bei angenäherter Normalverteilung in den Schülerleistungen – die Schwierigkeiten der Testaufgaben und die gegebenen Fähigkeiten bzw. Leistungen der Schülerinnen und Schüler insgesamt gut aufeinander abgestimmt.

Um einen Eindruck davon zu bekommen, welche Leistungsniveaus mit bestimmten erreichten Skalenwerten assoziiert sind, sind wieder fünf Beispielaufgaben – aus einem Text mit dem Titel „*Crossing the Channel*“ – in die Abbildung eingetragen. Beispiel 1 – die korrekte Ergänzung des Wortes „*people*“ im Satz- bzw. Textzusammenhang – steht für ein Leistungsniveau (Skalenwert 81), das mit einer Wahrscheinlichkeit von 0,5 von 98,7 Prozent der getesteten Schülerinnen und Schüler erreicht bzw. übertroffen wurde. Die weiteren Beispiele aus demselben Text – auch hier sind die geforderten Wortergänzungen kursiv geschrieben – stehen für zunehmend anspruchsvolle Leistungsniveaus, die von 88,8 Prozent (Skalenwert 101), von 60,3 Prozent (Skalenwert 127), von 22,2 Prozent (Skalenwert 153) und von 3,2 Prozent der Schülerinnen und Schüler (Skalenwert 173) erreicht bzw. übertroffen wurden.

---

<sup>27</sup> Die C-Tests wurden zunächst durch die Testleitungen an den Schulen ausgewertet. Um systematische Fehlkodierungen auszuschließen, wurden aus jeder Klasse bzw. aus jedem Kurs zwei Tests durch Mitarbeiterinnen der Forschungsgruppe in Berlin nachkodiert, bei Auftreten von Irrtümern noch mindestens zwei weitere. Zeigten sich dann systematische Fehlertendenzen, wurde der ganze Klassensatz nachkodiert. Außerdem wurden alle fehlenden Kodierungen an Hand der Testunterlagen vervollständigt.

Abbildung 11 Verteilung der Schülerleistungen in *Englisch* im Vergleich mit den Schwierigkeiten der Testaufgaben im *SL-HAM 8/9-Englisch: C-Test*



Gesamtstichprobe (N = 10.863)

### Fachleistung Englisch: C-Test, differenziert nach Schulformen bzw. Kursniveaus

Die folgende Tabelle 18 zeigt die schulform- bzw. kursniveauspezifischen Ergebnisse für den *Englisch: C-Test*.

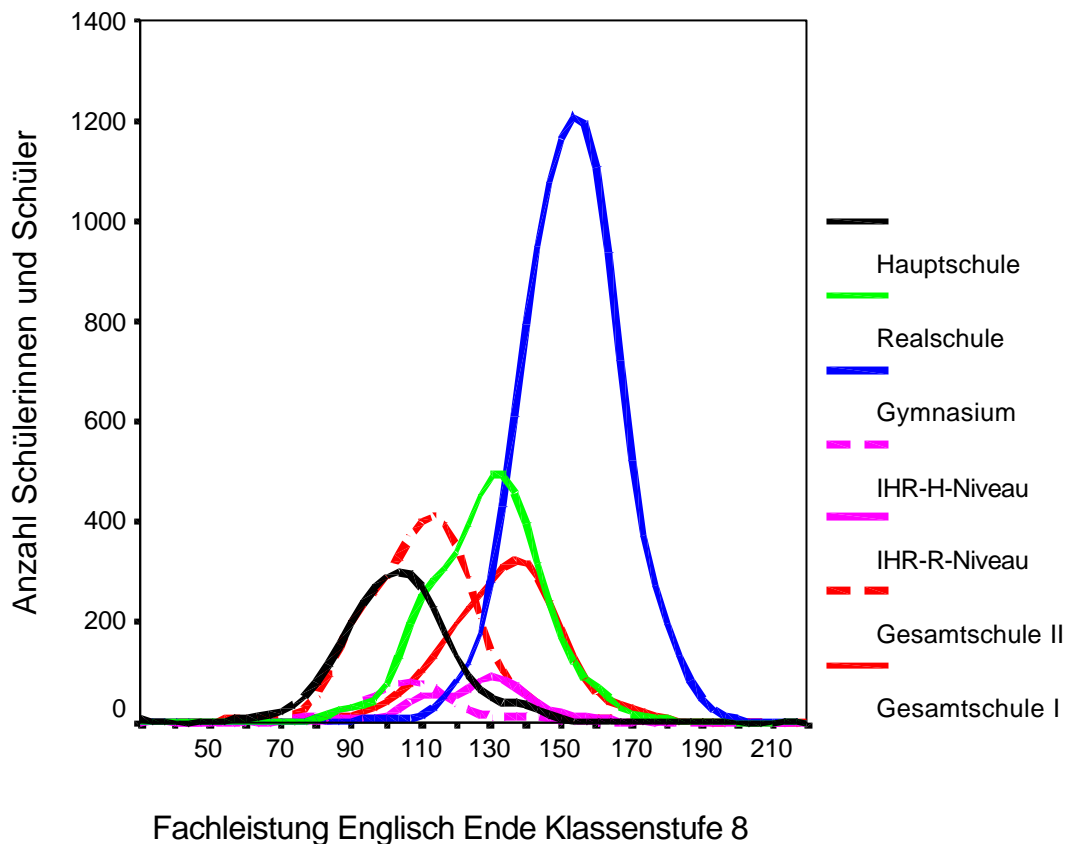
Tabelle 18 Fachleistung Englisch C-Test am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	103,5	17,1	-1,20	1.108
Realschule	128,5	16,8	-0,18	1.863
Gymnasium	153,0	14,8	0,82	4.234
IHR Schule – H	107,4	15,6	-1,04	261
IHR Schule – R	125,9	14,9	-0,28	300
Gesamtschule II	109,4	16,2	-0,96	1.542
Gesamtschule I	133,3	17,2	0,02	1.248
<i>insgesamt</i>	<i>132,9</i>	<i>24,6</i>	<i>---</i>	<i>10.556</i>

Die Schülerinnen und Schüler, die in den Gesamtschulen einen Englischkurs I besuchen, weisen mit durchschnittlich 133 Rasch-Skalenpunkten fast denselben Mittelwert auf wie die Gesamtgruppe. Die Gymnasiastinnen und Gymnasiasten zeigen bei der vergleichsweise geringsten Leistungsstreuung mit durchschnittlich 153 Skalenpunkten die mit Abstand höchsten Fachleistungen, die Schülerinnen und Schüler aus Englischkursen II an Gesamtschulen ( $\bar{x} = 109$  Punkte), aus Hauptschulkursen der Integrierten Haupt- und Realschulen ( $\bar{x} = 107$  Punkte) und aus Hauptschulklassen ( $\bar{x} = 104$  Punkte) die niedrigsten. Die Abbildung 12 veranschaulicht diesen Befund grafisch.

Zur Orientierung darüber, wie viel Prozent der Schülerinnen und Schüler den Kennwert 133 und damit die Scheidelinie von Realschul- und Gymnasialleistungen überschreiten, sollen hier wieder die entsprechenden Quoten für den „typisch gymnasialen Bereich“ berichtet werden: Es sind dies 4,9 Prozent aus Hauptschulklassen bzw. 6,5 Prozent aus IHR-Hauptschulkursen, 39,6 Prozent aus Realschulklassen bzw. 33,0 Prozent aus IHR-Realschulkursen, 6,0 Prozent (Kurse II) bzw. 52,6 Prozent (Kurse I) aus Gesamtschulen und schließlich mit 92,5 Prozent die überwiegende Mehrheit aus Gymnasialklassen. Für keine der bisher berichteten Fachleistungen zeigten sich für das hier gewählte Kriterium solch deutliche Leistungsvorteile der Gymnasiastinnen bzw. Gymnasiasten.

Abbildung 12 Englisch C-Test: Verteilung der Lernstände zum Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Das bestätigt sich auch in Folgendem: 50,3 Prozent der Schülerinnen und Schüler aus Gymnasialklassen haben den Skalenmittelwert für die Schulform Gymnasium ( $\bar{x} = 153$  Punkte) übertroffen; aus den übrigen Schulformen sind es zwischen 9,1 Prozent (Kurse I an Gesamtschulen) und 0 Prozent (IHR-Hauptschulkurse).

### 3.2.4 Fachleistung Latein

Entsprechend der allgemeinen Anlage der LAU wurden diejenigen Schülerinnen und Schüler an Gymnasien, die den Fremdsprachenunterricht nicht mit Englisch, sondern mit Latein begonnen haben, nur in Latein, nicht aber mit den Englisch-Untertests des *SL-HAM 8/9* auf ihre fremdsprachliche Leistung hin überprüft.

Vor der Analyse der Leistungsdaten wurde der Frage nachgegangen, ob es sich bei diesen Klassen um kognitiv ausgelesene Lerngruppen innerhalb der gymnasialen Schülerschaft handelt. Statistische Vergleiche zeigen, dass dies in kaum nennenswertem Maße der Fall ist: Das Bestimmtheitsmaß  $\eta^2$  zeigt (bei keineswegs allgemeiner Überlegenheit gegenüber den Englischklassen an den

Gymnasien) an, dass dieser Ausleseeffekt im Vorzeichen wechselt und außerordentlich gering ist. Für das schlussfolgernde Denken (CFT) gilt  $Eta^2 = 0,001$  (positiv), für den Index Fachleistung Deutsch ebenfalls  $Eta^2 = 0,001$  (positiv), für die Fachleistung Mathematik  $Eta^2 = 0,000$  (positiv), hingegen für die beurteilte Textproduktion  $Eta^2 = 0,030$  (negativ).

Die geringe Anzahl getesteter Schülerinnen und Schüler aus Lateinklassen ( $N = 243$ ) lässt es nicht zu, die Testergebnisse aus der LAU 7 mit den Ergebnissen der LAU 9 so zu verkoppeln, dass eine gemeinsame Metrik für Vergleiche über die Klassenstufen hinweg zur Verfügung steht. Zudem stieß bereits die klassische Skalierung des Lateintests in seiner Teilkomponente *Sprache* auf Schwierigkeiten, und zwar überwiegend dadurch, dass auch bei vielen Schülerinnen und Schülern mit relativ besseren Lateinkenntnissen offenbar das Hintergrundwissen nicht ausreichte, um das dargebotene Textmaterial angemessen zu erschließen. Ein Beispiel hierfür ist eine Aufgabe, in der alle Wörter zu markieren waren, die sich auf Christoph Kolumbus beziehen, und bei der sich gerade auch die sonst Leistungsstärkeren für „*Rex ipse, qui ... contulerat*“ entschieden, d. h. Kolumbus offenbar für einen König hielten, und somit das auf Kolumbus verweisende Dativobjekt (*viventi ... mortuo*) mit einem davon abhängigen Genitiv-Attribut (*eius*) nicht zu erkennen vermochten. Es ist daher zu bezweifeln, dass sich der Sinn des Ausgangstextes vielen eröffnet hat. Es gibt aber auch eine Reihe von rein grammatikalischen Items, bei denen viele der Leistungsstärkeren offenbar an den Mehrdeutigkeiten der lateinischen Sprache gescheitert sind (z. B. *iudices* zugleich als Nom. Pl. von *iudex* und als 2. Pers. Sing. Konj. Präsens Akt. von *iudicare*). Deshalb konnten insgesamt nur 17 Items in die Auswertung einbezogen werden, was eine vergleichsweise niedrige interne Konsistenz des lateinischen Sprachtests zur Folge hatte (Cronbachs Alpha = 0,64).

Die Testkomponente „*Sachwissen über das römische Leben*“ hingegen war mit klassischen Methoden gut zu skalieren: hier wurde trotz der nur 14 Items eine Zuverlässigkeit von Cronbachs Alpha = 0,73 erreicht. Allenfalls auffällig war, dass „Augustus“ auch viele leistungsstärkere Schülerinnen und Schüler für einen römischen Vornamen hielten und dass vielen von ihnen der Zusammenhang zwischen „Konzession“ und „*cedere*“ nicht deutlich war.

So kann hier zunächst nur das Verhältnis der beiden Teiltests *Latein: Sprache* und *Sachwissen über das römische Leben* im Kontext einiger weiterer kognitiver Variablen betrachtet werden, um sodann rein deskriptiv die erzielten Ergebnisse im Lateintest zu beleuchten. Wie die nachstehende Korrelationstabelle 19 zeigt, sind die Zusammenhänge zwischen den beiden Kom-

ponenten des Lateintests untereinander und mit anderen kognitiven Variablen nicht sehr eng.

Tabelle 19 Fachleistung Latein: Interkorrelation und Korrelationen mit anderen kognitiven Variablen

	Latein Sprache	Römisches Leben	Deutsch: Sprache	Mathe- matik	schlussfolgerndes Denken
<b>Latein Sprache</b>	---	0,40	0,41	0,29	0,17
<b>Römisches Leben</b>	---	---	0,45	0,33	0,16

Die Korrelationen sind also nur mäßig ausgeprägt und legen beispielsweise keine ausgeprägten Transferleistungen des lateinischen Sprachunterrichts für das Sprachverständnis im Deutschen oder das schlussfolgernde Denken nahe. Doch darf nicht übersehen werden, dass – wie oben dargelegt – eine Reihe der anspruchsvolleren Testaufgaben von der Analyse hat ausgeschlossen werden müssen. Für diejenigen Items, die berücksichtigt werden konnten, gilt, dass es sich um Items mit stark variierender Schwierigkeit (mit Lösungshäufigkeiten zwischen 14 und 74 Prozent) handelt.

Durchschnittlich wurden 5,8 Punkte erreicht (d. h. 34,1 Prozent richtige Aufgabenlösungen) bei einer Standardabweichung von 2,9 Punkten für *Latein: Sprache* und 6,9 Punkte (d. h. 49,4 Prozent richtige Lösungen) bei einer Standardabweichung von 3,1 Punkten für *Sachwissen über das römische Leben*. Aufschlussreich sind Vergleiche zwischen den Schulen und Schulklassen im Hinblick auf die erreichten Lernstände in der Testkomponente *Latein: Sprache*: Während an einem Gymnasium mit vier Lateinklassen in drei Klassen hoch überdurchschnittliche Mittelwerte erreicht wurden, variierte das Leistungsniveau an zwei weiteren Gymnasien mit mehr als einer Lateinklasse sehr beträchtlich. Bei jeweils nur einer Klasse von überdurchschnittlichem Leistungsniveau (6,0 bzw. 7,0 Punkte im Klassendurchschnitt) waren in den übrigen Klassen dieser beiden Schulen mit Durchschnittswerten zwischen 4,4 und 4,8 Punkten überraschend niedrige Leistungsstände zu verzeichnen. Da sich hierfür in den schul- und klassenspezifischen Verteilungen kognitiver Lernvoraussetzungen (*CFT 20*-Werte) keine Entsprechung findet, ist zu vermuten, dass der Lateinunterricht in Hamburg sehr unterschiedlich effizient ist.

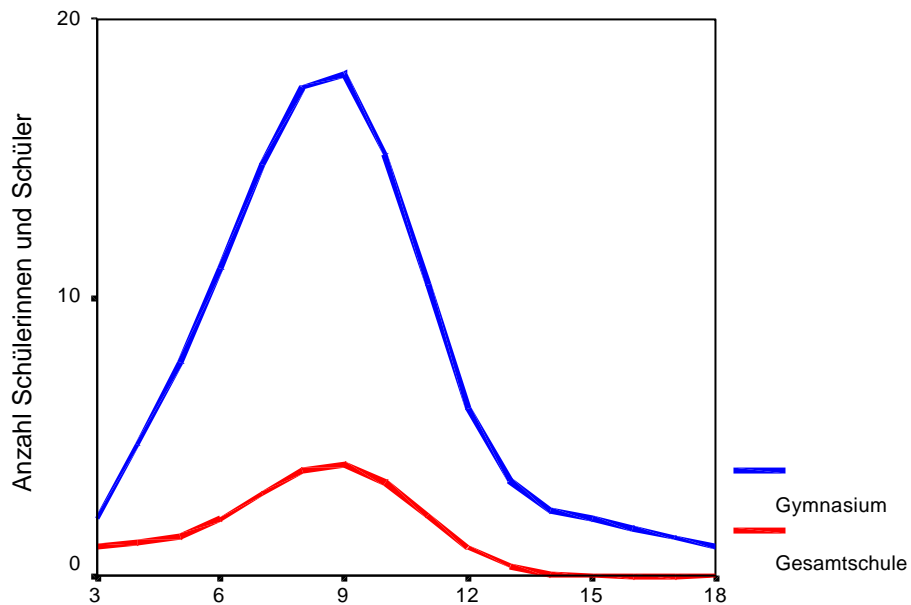
### 3.2.5 Fachleistung Französisch

Eine kleine Zahl von Schülerinnen und Schülern in Hamburg hat zu Beginn des Schuljahres 1996/97 von der Möglichkeit Gebrauch gemacht, den Fremdsprachenunterricht statt mit Englisch oder Latein mit Französisch zu beginnen. Es handelte sich dabei um 40 Schülerinnen und Schüler aus der gymnasialen Beobachtungsstufe und acht an einer Gesamtschule. Wie eine Überprüfung der Lernstände im Fach Deutsch zeigt, handelt es sich dabei, anders als im Fach Latein, zumindest im sprachlichen Bereich um eine im positiven Sinne stark vorausgelesene Gruppe. Der Fachleistungsindex Deutsch ist bei den Schülerinnen und Schülern aus den Gymnasialklassen um 0,62 Standardabweichungen und bei denen aus der Gesamtschule sogar um 0,99 Standardabweichungen höher als der jeweils entsprechende schulformspezifische Gesamtmittelwert. Die durchschnittlichen Werte für das schlussfolgernde Denken (*CFT 20*) und den Mathematiktest, die die Schülerinnen und Schüler der Französischgruppe an der Gesamtschule erreichten, entsprechen dem allgemeinen Durchschnitt der Gymnasien, im Fach Deutsch blieb der Mittelwert jedoch hinter dem gymnasialen Durchschnitt zurück.

Analog zu dem Vorgehen im Englischtest wurden die fremdsprachlichen Leistungen auch in den Französischklassen zunächst mit einem *Hörverständnistest* und dann mit einem *C-Test* überprüft. Beide Tests erwiesen sich als skalierbar. Der Hörverständnistest weist mit 25 Items insgesamt eine interne Konsistenz von Cronbachs Alpha = 0,67 auf. Von den 25 Items wurden im Durchschnitt 8,7 Items, d. h. 34,7 Prozent, richtig gelöst; er war also für die untersuchte Gruppe recht schwer. Die Standardabweichung betrug 3,2 Punkte. Schlösse man 16 leichtere, weniger gut diskriminierende Aufgaben aus, so ließe sich die interne Konsistenz zwar auf Cronbachs Alpha = 0,80 erhöhen, doch der Anteil der dann richtig gelösten Aufgaben würde sich bei solchem Vorgehen auf 12,4 Prozent vermindern. Hier soll deshalb der Hörverständnistest in seiner ganzen Länge Berücksichtigung finden. Der *C-Test* hatte, wie auch im Fach Englisch, gute psychometrische Eigenschaften: Mit 118 auszufüllenden Lücken wurde eine Zuverlässigkeit von Cronbachs Alpha = 0,95 erzielt. Dabei wurden von den Schülerinnen und Schülern im Mittel 61,6 korrekte Lösungen gefunden bzw. 52,2 Prozent des theoretischen Maximums erreicht. Die Standardabweichung lag hier bei 18,5 Punkten. Diese Testkomponente war demnach gut an den vorfindlichen Leistungsstand angepasst. In Abbildung 13 sind zunächst die Leistungsverteilungen für den Hörtest, getrennt für das Gymnasium und die Gesamtschule, dargestellt, wobei die geringen Fallzahlen zu beachten sind.



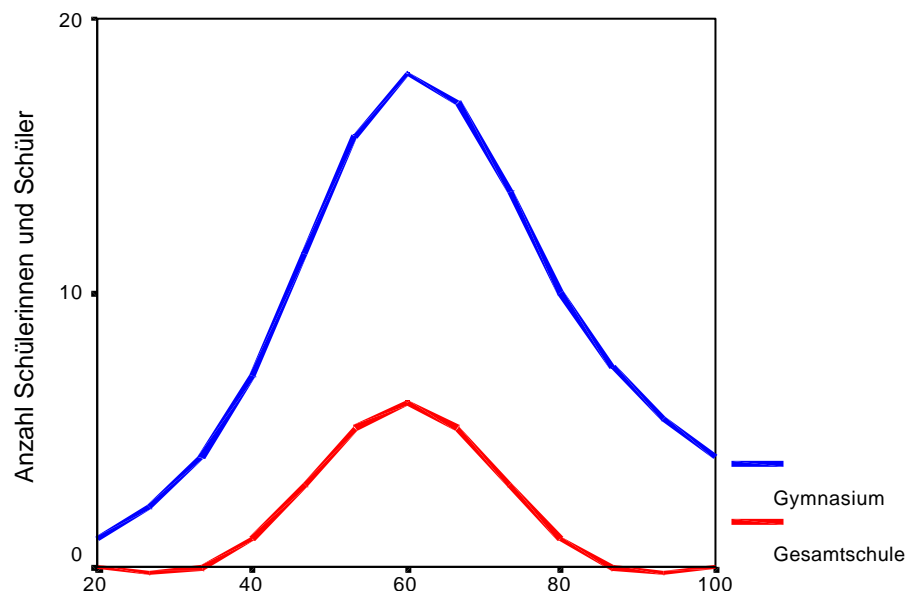
Abbildung 13 Französisch: Hörverständnis am Ende der Klassenstufe 8, nach Schulform



Fachleistung Französisch (Hörverständnis) Ende Klassenstufe 8

Ihnen seien die Verteilungen der Ergebnisse im modifizierten *C-Test* gegenübergestellt (Abbildung 14).

Abbildung 14 Französisch: Sprachlicher Lernstand (modifizierter C-Test) am Ende der Klassenstufe 8, nach Schulform



Fachleistung Französisch (C-Test) Ende Klassenstufe 8

Zusätzlich seien die wichtigsten in diesen beiden Grafiken enthaltenen Informationen auch in numerischer Form dargeboten (vgl. Tabelle 20).

Tabelle 20 Französisch: Hörverständnis und Sprache (C-Test) am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
<b>Hörverständnis</b>				
Gymnasium	8,9	3,2	0,07	40
Gesamtschule	7,5	2,8	-0,37	8
<i>insgesamt</i>	8,7	3,2	---	48
<b>C-Test</b>				
Gymnasium	62,8	19,6	0,06	40
Gesamtschule	55,9	10,7	-0,30	8
<i>insgesamt</i>	61,6	18,5	---	48

Angesichts der geringen Fallzahlen, insbesondere angesichts der kleinen Gesamtschulgruppe, sollten diese Befunde nicht überbewertet werden, zumal wirklich aussagekräftige Vergleichsgrößen, „*benchmarks*“, fehlen. Es ist jedoch auffällig, dass die beobachteten niedrigeren Leistungen in der Gesamtschulgruppe im Französischtest mit  $d = -0,37$  bzw.  $d = -0,30$  wesentlich geringer ausgeprägt sind als der für diese kleine Gruppe analog berechnete Rückstand im Fachindex Deutsch ( $d = -1,31$ ). Der Umstand, dass es sich an den Gesamtschulen bei den Französischschülerinnen und -schülern nach Maßgabe der kognitiven Voraussetzungen um eine hoch positiv ausgelesene Gruppe handelt – während dies an den Gymnasien kaum nachweisbar ist<sup>29</sup> – hat sich demnach nur im Französischtest, nicht aber zugleich im Deutschtest auswirken können. Abschließend seien deshalb wie für den Lateintest einige Korrelationen berichtet, die für die Interpretation der Befunde hilfreich sein können (Tabelle 21). Diese Werte sind – soweit sie den in der Tabelle 19 für den Lateintest berichteten Werten strukturell vergleichbar sind – insoweit interessant, als sie für die Fachleistung Französisch mindestens so enge, teilweise vielleicht als Indiz für Transferleistungen zu deutende Zusammenhänge mit dem Sprachverständnis nahe legen wie für das Fach Latein.

<sup>29</sup> Die CFT-20-Werte der Französischgruppen an der Gesamtschule und an den beiden Gymnasien sind nahezu identisch.

Tabelle 21 Fachleistung Französisch: Interkorrelation und Korrelationen mit anderen kognitiven Variablen

	Französisch Hörverst.	Französisch Sprache	Deutsch Leseverst.	Deutsch Sprache	schlussfolgerndes Denken
<b>Französisch Hörverständnis</b>	---	0,72	0,21	0,39	0,25
<b>Französisch C-Test</b>	---	---	0,28	0,53	0,32

Wegen der geringen Fallzahlen sind die Fehlermargen auch für diese Korrelationskoeffizienten groß, und die zu Eingang dieses Abschnittes geschilderte Selbstselektion der Lerngruppen beim Eintritt in eine Französischklasse kann ebenfalls diese Befunde beeinflusst haben.

### 3.2.6 Zusammenfassung der allgemeinen Fachleistung am Ende der Klassenstufe 8

Es erscheint sinnvoll, zum Abschluss der Diskussion der fachbezogenen Lernstände am Ende der Klassenstufe 8 die Frage nach den erreichten Lernständen noch einmal in allgemeinerer Form zu stellen, und zwar vor allem deshalb, weil die horizontale Durchlässigkeit des Hamburger Schulsystems bzw. der am Ende der Sekundarstufe I vermittelten Übergangsberechtigungen nicht an die einzelnen Schulfächer, sondern an den allgemeinen Leistungsstand eines Schülers bzw. einer Schülerin geknüpft ist. In diesem Zusammenhang sind zwei Vorfragen zu klären. Die eine betrifft den Umgang mit der Zugehörigkeit von Schülerinnen und Schülern zu unterschiedlichen Kursniveaus an den integrierten Schulformen, die andere die Regeln für die Bildung eines allgemeinen Fachleistungsindex. In der nachstehenden Tabelle 22 ist für die beiden integrierten Schulformen dargestellt, wie sich bei den Schülerinnen und Schülern die Kursniveauzuordnung über die Fächergrenzen hinweg gestaltet.

Tabelle 22 Zuordnung der Schülerinnen und Schüler an integrierten Schulformen zum jeweils höheren Kursniveau in den Fächern Deutsch, Englisch und Mathematik (in Prozent)

<i>höheres Kursniveau in . . .</i>	<b>IHR-Schule</b>	<b>Integrierte Gesamtschule</b>
keinem Fach	43,1	40,7
einem Fach	6,0	18,3
zwei Fächern	4,1	14,7
drei Fächern	46,8	26,3

Diesen Zahlen ist zu entnehmen, dass die Kursniveauzuordnung an den integrierten Haupt- und Realschulen über die Fächergrenzen hinweg wesentlich stabiler ist als an den Gesamtschulen und insoweit dem Zuordnungsmuster der Haupt- und Realschulen nicht unähnlich ist, als nur etwa 10 Prozent der Schülerinnen und Schüler in verschiedenen Fächern auf unterschiedlichem Niveau unterrichtet werden. Dagegen sind die Anteile der Schülerinnen und Schüler, die an den integrierten Schulformen durchgehend dem niedrigeren Niveau zugeordnet sind, der Hauptschulquote an den nicht integrierten Haupt- und Realschulen (39,3 Prozent) vergleichbar. Wenn oben mehrfach davon die Rede war, dass die Gesamtschulen bei der Grenzziehung zwischen Kursniveau I und II höhere Standards anlegen als die Haupt- und Realschulen, so wird hier sichtbar, dass dies vor allem mit differenzierteren Kurszuweisungen einhergeht. Um nun für die integrierten Schulformen die allgemeine, also fächerübergreifende Fachleistung darstellen zu können, werden im Folgenden die Schülerinnen und Schüler, die höchstens in einem Fach auf dem höheren Kursniveau unterrichtet werden, dem unteren Stratum (IHR-H; Kursniveau II) zugerechnet, während diejenigen, die in mindestens zwei Fächern dem anspruchsvolleren Kursniveau angehören, dem oberen Stratum (IHR-R; Kursniveau I) zugeordnet werden. Dies entspricht auch den Festlegungen für die spätere Vergabe von Schulabschlüssen bzw. weiteren Zugangsberechtigungen.

Bei der Bildung eines Index für die allgemeine Fachleistung wird hier ähnlich verfahren wie schon im Bericht zur LAU 7: Es wird die durchschnittliche Testleistung für die Fächer Mathematik, Deutsch (zu gleichen Teilen gewichtet aus den Untertests *Leseverständnis*, *Sprache* und *Rechtschreibwissen*) und *Englisch* ermittelt und auf eine geeignete Skala mit einem Mittelwert von 50 und einer Standardabweichung von 10 projiziert.<sup>30</sup> Für die Schülerinnen und Schüler aus Latein- und Französischklassen umfasst die allgemeine Fachleistung nur den Durchschnitt aus den Fächern Deutsch und Mathematik.

Die Verteilung der so bestimmten allgemeinen Fachleistung der Schülerinnen und Schüler an Hamburger Schulen zum Ende der Klassenstufe 8 entspricht in recht guter Näherung einer Normalverteilung (mit einem durch den hohen Gymnasialanteil verursachten „Hauptgipfel“ im Bereich des gymnasialen Durchschnitts; ohne Abbildung). Die Unterschiede zwischen den Schulformen

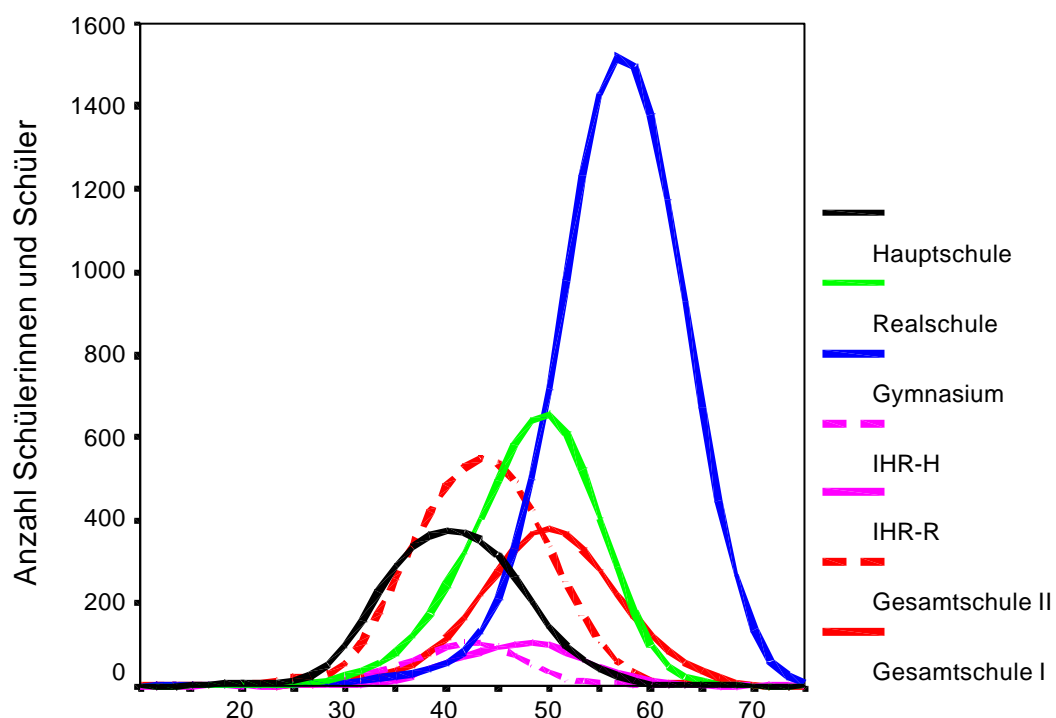
---

<sup>30</sup> Hierzu waren die Ausgangswerte zunächst über eine sog. „z-Transformation“ zu standardisieren. Dass sich die reale Gesamtstandardabweichung (8,8 Skalenpunkte) von der angegebenen Zielgröße (10 Skalenpunkte) unterscheidet, hängt mit fehlenden Werten für einzelne Gruppen zusammen, beispielsweise dem systematischen Ausfall von Englisch-Testwerten der Latein- und Französischklassen.

bestätigen das aus den fächerspezifischen Analysen bereits gewohnte Bild (vgl. Abbildung 15).

Wieder werden die großen Überschneidungsbereiche zwischen den Schulformen und Kursniveaus deutlich, die nicht nur zwischen gymnasialem und nichtgymnasialem Segment festzustellen sind, sondern auch zwischen Haupt- und Realschülerinnen bzw. -schülern, zwischen den Gruppen auf IHR-H- und IHR-R-Niveau und zwischen den (überwiegend) auf Kursniveau II unterrichteten Schülerinnen und Schülern der Gesamtschulen und ihren (überwiegend) dem Kursniveau I zugeordneten Mitschülerinnen und Mitschülern. So haben 26,3 Prozent der Hauptschülerinnen und Hauptschüler den Grenzwert zum *typischen Realschulbereich* ( $\geq 42$  Skalenpunkte) erreicht oder übertroffen, während 13,0 Prozent der Realschüler und Realschülerinnen dahinter zurückgeblieben sind. An den Integrierten Haupt- und Realschulen haben von den überwiegend auf dem Hauptschulniveau Unterrichteten 34,3 Prozent und auf dem Realschulniveau 87,0 Prozent den Schwellenwert dieser Schulform ( $\geq 44$  Skalenpunkte) erreicht oder überschritten.

Abbildung 15 Allgemeine Fachleistung: Verteilung der Lernstände am Ende der Klassenstufe 8 nach Schulform bzw. Kursniveau (absolut)



Allgemeine Fachleistung Ende Klassenstufe 8

An den Integrierten Gesamtschulen, wo einerseits wegen der höheren Standards beim Übergang zum Kursniveau I ( $\geq 48$  Skalenpunkte) ein kleinerer Überschneidungsbereich zu erwarten, andererseits jedoch ein erhöhter Anteil

an gemischten Zuweisungen festzustellen war, lagen 12,3 Prozent der überwiegend auf dem Kursniveau II und 62,1 Prozent der überwiegend auf dem Kursniveau I unterrichteten Schülerinnen und Schüler oberhalb des schulförmenspezifischen Grenzwerts. Bezogen auf den wie bisher definierten Grenzwert zum *typisch gymnasialen Bereich* ( $\geq 49$  Skalenpunkte) waren die Anteile derjenigen, die ihn erreicht oder übertroffen haben, 4,2 Prozent an den Hauptschulen, 44,0 Prozent an den Realschulen, 94,3 Prozent an den Gymnasien, 5,0 Prozent der überwiegend in H-Kursen bzw. 38,4 Prozent der überwiegend in R-Kursen an den Integrierten Haupt- und Realschulen unterrichteten Schülerinnen und Schüler sowie 9,3 Prozent der überwiegend in Kursen II und 55,1 Prozent der überwiegend in Kursen I an den Gesamtschulen unterrichteten Schülerinnen und Schüler.

Den gymnasialen Durchschnitt (57,4 Skalenpunkte) hat niemand an den Hauptschulen erreicht oder überschritten; an den Realschulen waren es 3,3 Prozent, an den Gymnasien 50,9 Prozent, von den überwiegend in H-Kursen an den Integrierten Haupt- und Realschulen unterrichteten Schülerinnen und Schülern waren es 0,3 Prozent, von den überwiegend in R-Kursen an den Integrierten Haupt- und Realschulen unterrichteten Schülerinnen und Schülern 1,6 Prozent, von den überwiegend in Kursen II an Gesamtschulen unterrichteten Schülerinnen und Schülern 0,4 Prozent und von den überwiegend in Kursen I unterrichteten Schülerinnen und Schülern 9,3 Prozent. Vergleicht man diese Werte mit den Anteilen derjenigen, die in der LAU 7 den gymnasialen Durchschnitt erreicht oder überschritten hatten (LEHMANN, GÄNSFUß & PEEK 1999, S. 67), so sind praktisch keine Veränderungen eingetreten: 1998 hatten 1,6 Prozent der Schülerinnen und Schüler aus den ehemaligen Beobachtungsstufen der Haupt- und Realschulen den gymnasialen Durchschnitt erreicht oder überschritten, 2000 waren es 1,8 Prozent. Ähnlich verhält es sich an den Gesamtschulen: 3,6 Prozent aus dem Jahre 1998 stehen im Jahre 2000 3,8 Prozent gegenüber. Dies gilt, obwohl sich der Anteil der Gymnasiastinnen und Gymnasiasten, die 2000 außerhalb des typisch gymnasialen Bereichs lagen, mit 5,8 Prozent gegenüber 1998 (11,9 Prozent) u. a. durch Schulformwechsel um mehr als die Hälfte vermindert hat.

Auch für die allgemeine Fachleistung ist es deshalb von Bedeutung, wie sich die jeweiligen Leistungsverteilungen zueinander verhalten. Analog zu den entsprechenden fachspezifischen Tabellen sind daher in Tabelle 23 die Mittelwerte, Standardabweichungen und Effektstärken zusammengestellt.

Tabelle 23 Allgemeine Fachleistung am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. Kursniveau

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	39,5	5,2	-1,18	1.296
Realschule	48,0	5,4	-0,21	2.035
Gymnasium	57,4	5,5	0,86	4.678
IHR Schule – H	40,5	4,9	-1,07	300
IHR Schule – R	47,2	5,1	-0,31	318
Gesamtschule II	42,2	5,3	-0,88	1.827
Gesamtschule I	49,7	6,3	-0,02	1.273
<i>insgesamt</i>	<i>49,9</i>	<i>8,8</i>	<i>---</i>	<i>11.727</i>

Die hier aufgelisteten Effektstärken *d* bezeichnen wieder, wie gewohnt, die relative Position der Mittelwerte für die Schulformen bzw. Kursniveaus, bezogen auf den Mittelwert und die Streuung der Gesamtheit aller untersuchten Schülerinnen und Schüler. Die Werte sind generell denen für die einzelnen Untertests dem Anschein nach nicht unähnlich. Wegen der Möglichkeit eines kumulativen Zusammenwirkens der einzelnen Untertests lohnt es sich aber, zum Vergleich auch das Bestimmtheitsmaß  $Eta^2$  anzugeben, das angibt, welche Varianzanteile mit der Schulform- bzw. Kursniveauezugehörigkeit verbunden sind, und so die Effektstärken zu einem einzigen Kennwert zusammenfasst. Für die allgemeine Fachleistung beträgt  $Eta^2 = 0,61$ , worin sich die vermuteten kumulativen Effekte der Untertests zeigen. Die Werte des Bestimmtheitsmaßes  $Eta^2$  sind für die einzelnen Testkomponenten entsprechend niedriger: für die Untertests *Mathematik*  $Eta^2 = 0,48$ , *Leseverständnis*  $Eta^2 = 0,32$ , *Sprache*  $Eta^2 = 0,44$ , *Rechtschreibwissen*  $Eta^2 = 0,39$  und *Englisch (C-Test)*  $Eta^2 = 0,46$ . Die entsprechenden Werte für das *Rechtschreibkönnen (HSP 5-9)* sind  $Eta^2 = 0,37$ , für die *Textproduktion*  $Eta^2 = 0,12$  und (zum Vergleich) für das *schlussfolgernde Denken (CFT 20)*  $Eta^2 = 0,21$ . Die kognitive Eingangsselektivität der verschiedenen Schulformen und die fachspezifischen Unterschiede zu Beginn der Sekundarstufe I sind wesentlich niedriger, wie sich aus dem Vergleich mit den entsprechenden Werten für  $Eta^2$  aus dem Bericht zur Lernausgangslagenuntersuchung LAU 5 ergibt (vgl. LEHMANN & PEEK 1997, S. 105). Die allgemeinen Leistungsunterschiede zwischen den Bildungsgängen nehmen also mit der Anzahl und Wirksamkeit äußerer Differenzierungen zu: Indem die bis zum Ende der Klassenstufe 6 einheitlich unterrichtete Schülerschaft der Haupt- und Realschulen nach Fachleistung in verschiedene Klassen aufgeteilt wird und in-

dem in den integrierten Schulformen, ebenfalls nach Fachleistung, unterschiedliche Kursniveaus eingerichtet werden, steigt der Anteil der damit verbundenen Leistungsdifferenzen. Es bleibt noch zu prüfen, ob gleichzeitig eine divergente Leistungsentwicklung zwischen den Schulformen bzw. Kursniveaus stattgefunden hat.

Zusätzliche Einsichten sind zu gewinnen, wenn man die wesentlichen statistischen Kennwerte für die Schulformen, die eine fachspezifische äußere Leistungsdifferenzierung kennen, separat danach zusammenstellt, in wie vielen Fächern eine Schülerin oder ein Schüler dem oberen Leistungsniveau zugeordnet ist (vgl. Tabelle 24).

Tabelle 24 Allgemeine Fachleistung an Schulen mit fachbezogener äußerer Leistungsdifferenzierung am Ende der Klassenstufe 8 (Mittelwert und Standardabweichung), nach Schulform und Kursniveau

	Mittelwert	Standardabweichung	N
<i>Integrierte Haupt- und Realschule</i>	44,0	6,0	618
kein R-Kurs	40,2	5,0	262
zwei H-Kurse, ein R-Kurs	42,3	3,9	38
ein H-Kurs, zwei R-Kurse	44,0	5,0	26
drei R-Kurse	47,5	5,0	292
<i>Gesamtschule</i>	45,3	6,9	3.100
kein Kurs I	41,2	5,1	1.254
zwei Kurse II, ein Kurs I	44,4	5,2	573
ein Kurs II, zwei Kurse I	47,3	5,3	454
drei Kurse I	51,1	6,5	819

Nach diesen Befunden sind die Schülerinnen und Schüler der Gesamtschulen denen der Integrierten Haupt- und Realschulen leistungsmäßig insgesamt, aber auch in jeder Untergruppe überlegen. Dabei erreichen diejenigen mit zwei Kursen I nicht ganz den Realschuldurchschnitt, diejenigen mit drei Kursen I übertreffen ihn um knapp eine halbe Standardabweichung ( $d = 0,47$ ), bleiben jedoch um fast eine Standardabweichung hinter dem Durchschnitt der Gymnasiastinnen und Gymnasiasten zurück ( $d = -0,98$ ). Angesichts der anders gelagerten Differenzierungsgrundsätze an den Integrierten Haupt- und Realschulen ist hier der Vergleich mit den traditionellen Haupt- und Realschulen vorrangig. Bei den beiden Extremgruppen (kein R-Kurs vs. drei R-Kurse) ist die Nähe zur Hauptschule bzw. zur Realschule offensichtlich.



Dass jedoch die Überlegenheit der einzelnen Differenzierungsebenen an Gesamtschulen und Integrierten Haupt- und Realschulen nicht allein (und dann möglicherweise nur mittelbar) an den unterschiedlichen Standards bei der Zuordnung zum höheren Kursniveau liegen kann, belegt der schon genannte höhere Gesamtmittelwert für die Gesamtschulen. Denkbar wäre indessen, dass hier auch unterschiedliche Lernvoraussetzungen eine Rolle spielen.

Um die Auswirkungen der äußeren Differenzierung aus einer anderen Perspektive zusätzlich zu charakterisieren, sind in Tabelle 25 noch einmal die Werte für das Bestimmtheitsmaß  $Eta^2$  aus den verschiedenen Phasen der Untersuchung zusammengestellt. Danach haben die Schulformunterschiede während der Klassenstufen 5 und 6 zunächst in allen untersuchten Leistungsbereichen zugenommen. Insbesondere an der teilweise sprunghaften Zunahme des Bestimmtheitsmaßes  $Eta^2$  beim Übergang vom Ende der Klassenstufe 6 zum Anfang der Klassenstufe 7 wird indessen nochmals ersichtlich, dass die gegenüber den Ergebnissen von 1998 (LAU 7) größeren Schulformunterschiede ganz überwiegend dadurch zu Stande gekommen sind, dass mit der Leistungsdifferenzierung an den Haupt- und Realschulen sowie an den Gesamtschulen zu Beginn der Klassenstufe 7 Niveauzuordnungen vorgenommen worden sind, die sich keineswegs durchgängig als Grundlage divergenter Lernentwicklungen darstellen. Nur im Fach Mathematik deuten sich differenzielle Lernentwicklungen an, während im Untertest *Leseverständnis* und, stärker noch, im *Englisch: C-Test* wegen der über die Zeit abnehmenden Werte für  $Eta^2$  sogar Konvergenzen vermutet werden könnten. Dem näher nachzugehen ist Aufgabe des Kapitels 4.

Tabelle 25 Unterschiede in den Fachleistungen zwischen den Schulformen bzw. Kursniveaus (Bestimmtheitsmaß  $Eta^2$ ) nach Untertests<sup>31</sup>

<i>(Unter-) Test</i>	<b>Anfang Klassen- stufe 5</b>	<b>Ende Klassen- stufe 6</b>	<b>Anfang Klassen- stufe 7</b>	<b>Ende Klassen- stufe 8</b>
Mathematik	0,22	0,36	0,43	0,48
Deutsch: Leseverständnis	0,24	0,27	0,37	0,32
Deutsch: Sprache	0,26	0,42	0,44	0,44
Rechtschreibwissen	0,23	---	---	0,39
Rechtschreibkönnen	0,23	0,26	0,37	0,37
Englisch: C-Test	---	0,27	0,61	0,46
allgemeine Fachleistung	0,34	0,54	0,64	0,61

### 3.3 Problemlösekompetenz

Neben den Tests zur Überprüfung der Fachleistungen wurde ein *Problemlösetest* eingesetzt, mit dem fächerübergreifende Kompetenzen der Schülerinnen und Schüler untersucht werden sollten. Den Autoren dieses Tests zufolge entscheidet beim Problemlösen nicht nur das abstrakte Kombinieren; bedeutsam ist auch, ob man – gestützt auf spezifisches Wissen – die konkrete Situation versteht und eine Folge von Lösungsschritten angemessen bearbeitet. Problemlöseprozesse sind also situations- und kontextspezifisch. Die Bearbeitung eines Problems erfolgt mehrstufig, wobei man zunächst das Problem erkennen (d. h. die Ausgangs- und die Zielsituation sowie deren Diskrepanzen erfassen), sodann den Lösungsweg als Abfolge von Denk- oder Handlungsschritten planen, schließlich die Lösung ausführen und das erreichte Ergebnis an dem angestrebten Ziel messen muss (zu Konzepten und Indikatoren solcher fächerübergreifender Kompetenzen vgl. ausführlich KLIEME, STANAT & ARTELT 2001). Bei der Aufgabenkonstruktion des *Problemlösetests* wurde darauf geachtet, dass für die Lösung der Aufgaben lediglich Alltagswissen und kein spezifisches Fachwissen vorausgesetzt wird.

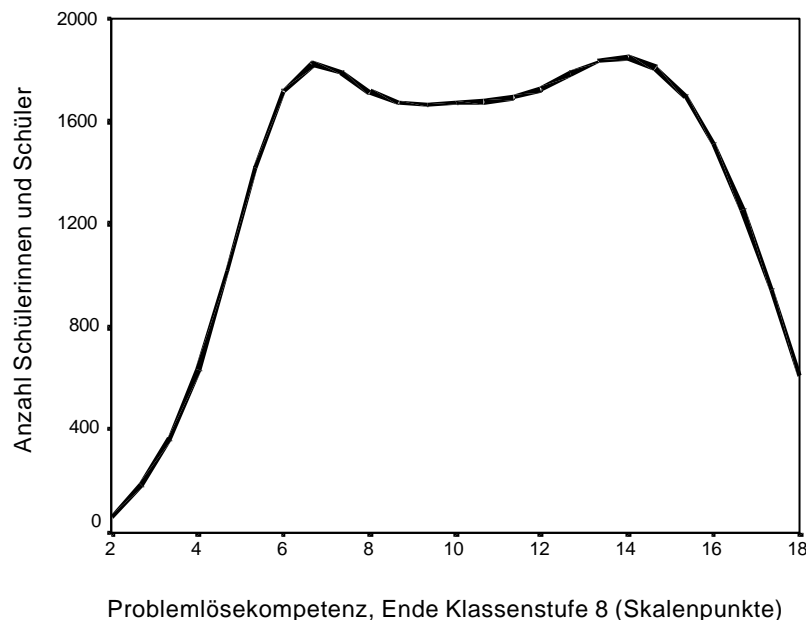
Die drei Projekte bzw. Komponenten des Untertests sollten von allen Schülerinnen und Schülern bearbeitet werden. Trotz der betonten Kontextspezifizität

<sup>31</sup> Um auch bei wechselnder Niveauezugehörigkeit Vergleichbarkeit zwischen der LAU 7 und der LAU 9 herzustellen, wurde (in einzelnen Fällen kontrafaktisch) für den Anfang der Klassenstufe 7 dasselbe Kursniveau unterstellt wie für das Ende der Klassenstufe 8. Für die integrierten Schulformen erfolgte die Gruppenbildung fachspezifisch. Eine Berücksichtigung der Textproduktion ist an dieser Stelle nicht sinnvoll, weil Vergleichswerte fehlen.

der einzelnen Aufgaben ergab sich eine durchschnittliche Interkorrelation von  $r = 0,64$  zwischen den Subtests, also doch ein gewisses Maß an Generalisierbarkeit der im Einzelfall beobachteten Problemlösekompetenz. Somit erschien es auch gerechtfertigt, die Resultate für die drei Aufgaben zu einem Summenscore zusammenzuziehen. Deren Zuverlässigkeit konnte mit Cronbachs Alpha = 0,83 bestimmt werden, also einem recht zufriedenstellenden Wert. Im Hinblick auf andere kognitive Leistungen sind die Testergebnisse am höchsten mit dem Sprachverständnis korreliert ( $r = 0,75$ ) und danach – in absteigender Reihenfolge – mit dem Leseverständnis ( $r = 0,72$ ), der Fachleistung Mathematik ( $r = 0,69$ ), der Fachleistung *Englisch: C-Test* ( $r = 0,68$ ) und der Fähigkeit zum schlussfolgernden Denken im *CFT 20* ( $r = 0,57$ ).

Die Testergebnisse sind, anders als in den meisten Fachleistungstests des *SL-HAM 8/9* wie auch hinsichtlich der Fähigkeit zum schlussfolgernden Denken (*CFT 20*), nicht angenähert normal oder zumindest eingipflig, sondern deutlich zweigipflig verteilt (vgl. Abbildung 16) – ein Hinweis darauf, dass die Ausbildung der im Test geforderten Fähigkeiten – es geht um die Gestaltung einer Veranstaltung, eines Arbeitsraums oder eines Energie- und Wassersparprogramms – in Zusammenhang mit den Schulformdifferenzierungen steht.

Abbildung 16 Problemlösekompetenz: Verteilung der Ergebnisse am Ende der Klassenstufe 8 (absolut)

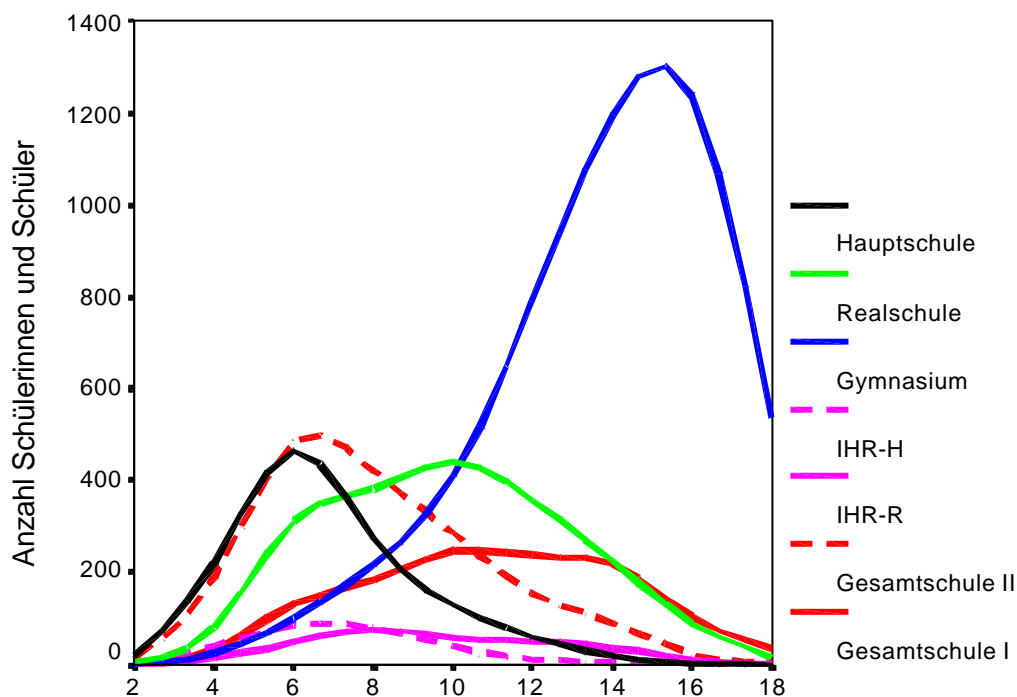


Die Struktur dieser Verteilung folgt nämlich im Wesentlichen der Unterscheidung zwischen den Schülerinnen und Schülern an Gymnasien einerseits und denjenigen anderer Schulformen bzw. Kursniveaus, ohne dass damit freilich schon geklärt wäre, durch welche Faktoren die Entwicklung solcher Kompetenzen begünstigt wird und wie sich dies auf die Schulformzuordnung aus-

wirkt. Daneben ist es (trotz eines leichten Deckeneffekts des Tests im gymnasialen Bereich) vor allem bemerkenswert, wie breit gestreut die Ergebnisse insgesamt sind.

In Abbildung 17 wird in der gewohnten Weise zwischen den Schulformen bzw. Kursniveaus unterschieden und über den Schnittpunkt der Verteilungskurven für die Realschulen und die Gymnasien (10,4 Skalenpunkte) der *typisch gymnasiale Leistungsbereich* bestimmt. Diesen Wert haben 9,0 Prozent der Schülerinnen und Schüler in Hauptschulklassen, 42,6 Prozent in Realschulklassen, 10,9 Prozent in den H-Kursen, 40,3 Prozent in den R-Kursen an den Integrierten Haupt- und Realschulen, 20,8 Prozent der überwiegend in Kursen II und 56,7 Prozent der überwiegend in Kursen I der Gesamtschulen erreicht oder überschritten.

Abbildung 17 Problemlösen: Verteilung der Ergebnisse am Ende der Klassenstufe 8 nach Schulformen bzw. überwiegender Kursniveauzugehörigkeit (absolut)



Problemlösekompetenz, Ende Klassenstufe 8 (Skalenpunkte)

Auffällig sind die relativ hohen Streuungen an den Gymnasien, auf den mittleren Niveaus (Realschulen, IHR-R) und vor allem auf dem Kursniveau I der Gesamtschulen (vgl. auch Tabelle 26). Die Effektstärken, die als Indikatoren für die relativen Unterschiede zwischen den Schulformen bzw. Kursniveaus dienen, sind ihrem absoluten Betrag nach schwächer ausgeprägt als diejenigen, die für den allgemeinen Fachleistungsindex bestimmt worden waren (vgl. Ab-

schnitt 3.2.6), größenordnungsmäßig aber den Verhältnissen in den einzelnen fachleistungsbezogenen Untertests gut vergleichbar, wie übrigens auch das Bestimmtheitsmaß zeigt, das hier  $Eta^2 = 0,44$  beträgt.

Tabelle 26 Ergebnisse im Problemlösetest am Ende der Klassenstufe 8: Mittelwerte, Standardabweichungen und Effektstärken, insgesamt und getrennt nach Schulform bzw. überwiegenden Kursniveaus

	Mittelwert	Standard- abweichung	Effektstärke <i>d</i>	<i>N</i>
Hauptschule	6,9	2,4	-1,02	1.195
Realschule	9,9	3,1	-0,27	1.913
Gymnasium	13,8	3,0	0,72	4.518
IHR Schule – H	7,4	2,6	-0,90	275
IHR Schule – R	9,6	3,1	-0,33	300
Gesamtschule II	8,1	2,9	-0,72	1.676
Gesamtschule I	11,0	3,3	0,01	1.192
<i>insgesamt</i>	<i>10,9</i>	<i>3,9</i>	<i>---</i>	<i>11.069</i>

### 3.4 Zusammenfassung

Die Analyse der Fachleistungen und der fächerübergreifenden Problemlösekompetenz von Schülerinnen und Schülern an Hamburger Schulen am Ende der Klassenstufe 8 hat vor allem drei Ergebnisse geliefert.

Erstens: Ausweislich der hohen Überschneidungsbereiche zwischen den Schulformen bzw. Kursniveaus gibt es zu Beginn der Klassenstufe 9 theoretisch ein beträchtliches Maß an Durchlässigkeit zwischen den Schulformen bzw. Kursniveaus. Die Fachleistungen der Schülerinnen und Schüler würden es in vielen Fällen erlauben, etwa von einer Realschule an ein Gymnasium zu wechseln. Umgekehrt findet sich in der Schülerschaft der Gymnasien ein nennenswerten Anteil solcher Jugendlicher, deren Fachleistungen eher zum Anforderungsniveau anderer Schulformen passen. Dies gilt, obwohl von den Schülerinnen und Schülern, die 1996 die Beobachtungsstufe des Gymnasiums besuchten, nur 65,3 Prozent in einem Gymnasium an den Erhebungen im Jahre 2000 teilgenommen haben, während 11,5 Prozent während oder nach der Beobachtungsstufe und weitere 8,9 Prozent (9,8 Prozent bezogen auf den Beginn der Klassenstufe 7) bis zum Beginn der Klassenstufe 9 das Gymnasium verlassen haben. Zieht man zusätzlich die Klassenwiederholungen in Betracht (vgl. dazu den Abschnitt 4.1), so ergibt sich daraus für das Spektrum der Fachleistungen – zumindest potenziell – ein beträchtlicher Homogenisierungseffekt,

der die Leistungsverteilung der Gymnasien deutlicher von denen der anderen Schulformen bzw. Kursniveaus hätte absetzen müssen.<sup>32</sup>

Zweitens: Während Klassenwiederholungen und Schulformwechsel die Streuung im Gesamtsystem Hamburger Schulen allenfalls indirekt beeinflussen, fällt bei der Untersuchung der Fachleistungen durchgängig die gegenüber den Befunden von 1998 reduzierte Varianz auf. Nun haben sich die mit der Schulform- bzw. Kursniveauzugehörigkeit assoziierten Unterschiede vom Beginn der Klassenstufe 7 bis zum Ende der Klassenstufe 8 nicht merklich verringert, wohl aber (mit Ausnahme der Hauptschulen) die Leistungsdifferenzen zwischen den Schulen innerhalb einer Schulform oder eines Kursniveaus. Es gibt also einen vorherrschenden Trend in Richtung auf Vereinheitlichung des Leistungsniveaus innerhalb der Schulformen bzw. Kursniveaus, und zwar insbesondere an den Gymnasien.

Drittens: Der Möglichkeit, der Herausforderung von Leistungsunterschieden in der Schülerschaft durch äußere Differenzierung zu begegnen, wird in Hamburg auf drei verschiedenen Wegen nachgegangen: Auch wenn in den Klassenstufen 9 und 10 noch Korrekturen möglich sind, befinden sich von der Schülerschaft, die die Beobachtungsstufe an traditionellen Haupt- und Realschulen absolviert hat, 60,7 Prozent am Ende der Klassenstufe 8 auf dem Wege zu einem mittleren Schulabschluss. An den Integrierten Haupt- und Realschulen sind es, wenn man das dominierende R-Kursniveau zum Maßstab nimmt, mindestens 50,9 Prozent und an den Gesamtschulen ausweislich der überwiegen- den Kurszugehörigkeit I mindestens 41,0 Prozent (in diesem Fall unter Ein- schluss der potenziellen Abiturientinnen und Abiturienten, jedoch unter Aus- schluss derer, die mit nur einem Kurs I den Realschulabschluss erwerben).

---

<sup>32</sup> Die bei LEHMANN, GÄNSFUB & PEEK 1999 berichtete Schulformwechslerquote vom Gymna- sium in Höhe von 6,2 Prozent bezog sich nur auf den Übergang nach Klasse 7; 5,3 Prozent haben also bereits während der Beobachtungsstufe das Gymnasium verlassen. Die im Ab- schnitt 4.1 berücksichtigten Quoten sind ebenfalls niedriger, weil dort nur diejenigen Fälle be- rücksichtigt werden konnten, für die auch Fachleistungsdaten vorhanden sind.

Somit werden an den traditionellen Haupt- und Realschulen potenziell mehr Schülerinnen und Schüler auf einen mittleren Schulabschluss, der ihnen u. U. auch den Weg in die gymnasiale Oberstufe öffnet, vorbereitet als in den integrierten Schulformen und speziell in den Gesamtschulen. Dem entspricht eine verglichen mit den Haupt- und Realschulen leistungsmäßige Überlegenheit der Kurse II und I der Gesamtschulen, die jedoch in engem Zusammenhang mit den unterschiedlichen Differenzierungsmustern zu sehen ist. Wollte man die 18,5 Prozent der Schülerschaft von Gesamtschulen, die nur einen Kurs I besuchen, denen zurechnen, die sich auf den mittleren Schulabschluss vorbereiten, so lägen die Testergebnisse dieser Gruppe trotz der darin enthaltenen potenziellen Abiturientinnen und Abiturienten mit durchschnittlich 48,1 Punkten auf dem Index für die allgemeine Fachleistung kaum noch oberhalb des Mittelwerts für die Realschulen (48,0 Punkte). Entsprechend verringerte sich auch die leistungsmäßige Überlegenheit der Kurse II (von 42,2 auf 41,2 gegenüber 39,5 Punkten an den Hauptschulen). Andererseits ist zu berücksichtigen, dass Klassenwiederholungen an Gesamtschulen selten sind. Im nachfolgenden Kapitel, das sich mit der Lernentwicklung in den Klassenstufen 7 und 8 beschäftigt, wird dem näher nachgegangen.

## **4 Entwicklung der Fachleistungen in den Klassenstufen 7 und 8**

### **4.1 Methodische Vorüberlegungen**

Im Grundsatz stehen den traditionellen Schulformen mehrere Möglichkeiten zu Gebote, das Leistungsspektrum der eigenen Schülerschaft zu homogenisieren bzw. – im Falle der Realschule und des Gymnasiums – auf relativ hohem Niveau homogen zu halten. Die wichtigsten sind – wie zum Beginn des Kapitels 3 bereits dargelegt wurde – der Schulformwechsel und die Klassenwiederholung. Der Schulwechsel innerhalb der Schulform kommt als weitere Möglichkeit hinzu, ist für die vorliegenden Fragestellungen aber weitgehend ohne Belang. Das „Springen“ erfolgt so selten, dass es hier ebenfalls vernachlässigt werden kann.

Klassenwiederholungen (und Überspringen einer Klassenstufe) haben – als Konsequenz der Anlage des Längsschnitts in der vorliegenden Untersuchung – dazu geführt, dass davon betroffene Schülerinnen und Schüler in den darauf folgenden Erhebungswellen nicht mehr erfasst werden konnten. Andererseits sind auf dem gleichen Wege auch neue Schülerinnen und Schüler in die Untersuchung gelangt. Trotzdem tritt durch Klassenwiederholungen – bezogen auf die Gesamtheit der 1996 untersuchten Schülerinnen und Schüler – insgesamt eine Überschätzung der jeweils erreichten Lernstände ein, die sich insbesondere auf die in diesem Kapitel zu diskutierenden Lernzuwachsrechnungen auswirkt. Der Grund dafür liegt darin, dass die Abgänge in der Regel aus dem unteren Bereich des Leistungsspektrums stammen, während für die Zugänge – zumindest innerhalb der ersten beiden Jahre – eine verbesserte Leistungsposition innerhalb der Klassenstufe und Schulform, ein „Erholungseffekt“, anzunehmen ist, was empirisch belegt werden kann.

Die Verzerrungen können näherungsweise so abgeschätzt werden, dass man statt der unvollständigen Lernstandswerte vom Ende der Klassenstufe 8 diejenigen vom Beginn der Klassenstufe 7 heranzieht, als die Wiederholerinnen und Wiederholer der Klassenstufen 7 und 8 noch Teil der Untersuchungsgruppe waren. Angenommen wird, dass die Nichtteilnehmer an der Erhebungswelle 2000 überwiegend durch Klassenwiederholung aus der Untersuchungsgruppe ausgeschieden sind. Dass die durch Klassenwiederholung erfolgten Neuzugänge dabei nicht berücksichtigt werden können, hat insofern nur begrenzten Einfluss, als deren Lernstände allenfalls teilweise den in den Klassenstufen 8 und 9 partiell neu konstituierten Lerngruppen zugeordnet werden können und in diesen Fällen von den Berechnungen zunächst ausge-



schlossen bleiben. Die Lernstandswerte vom Beginn der Klassenstufe 7 sind hoch mit den für den Anfang der Klassenstufe 9 (und damit auch für das Ende der Klassenstufe 8) erreichten Werten korreliert ( $r = 0,86$ ). Damit liefert dieser Ansatz mindestens annäherungsweise Anhaltspunkte für den Einfluss der Klassenwiederholung auf das Leistungsbild in den Ausgangslerngruppen der verschiedenen Schulformen. Tabelle 27 enthält die einschlägigen Befunde, unter Einschluss der integrierten Schulformen, jedoch ohne Rücksicht auf darin gegebene Kursniveaudifferenzierungen.

Tabelle 27 Allgemeine Fachleistung der Teilnehmerinnen und Teilnehmer an der Erhebung 1998 nach Schulform: Werte für den Beginn der Klassenstufe 7 im Vergleich zwischen Teilnehmern und Nichtteilnehmern an der Erhebung 2000

ursprüngliche Schulform-zugehörigkeit	Teilnahme an den Erhebungen 2000 (LAU 9)					Effektstärke $d$	Anteil Nichtteilnehmer in 2000
	nein		ja		insgesamt		
	$\bar{x}$	$s$	$\bar{x}$	$s$			
Hauptschule	29,6	8,3	32,6	9,8	31,8	0,09	32,8 %
Realschule	42,9	10,7	48,7	11,5	47,3	0,12	25,0 %
Gymnasium	64,5	12,3	71,7	11,5	70,8	0,08	13,1 %
IHR-Schule	36,4	10,5	40,4	12,1	39,4	0,08	26,2 %
Gesamtschule	41,0	14,5	44,6	13,8	44,4	0,02	10,5 %

Nach diesen Schätzungen, die – das sei wiederholt – grobe Annäherungen darstellen, weil sie Schulformwechsel, schulforminternen Kursniveauwechsel, Wegzug der Eltern, aber auch Neuzugänge durch Klassenwiederholungen und Zuzug nicht berücksichtigen, sind die angesprochenen Überschätzungen der „wahren“ Lernstände am Ende der Klassenstufe 8 in den Gesamtschulen fast zu vernachlässigen. In den traditionellen Schulformen sowie an den Integrierten Haupt- und Realschulen, die durch größere Wiederholerquoten charakterisiert sind, können sie hingegen nicht außer Betracht bleiben.

Der Einfluss des Zugangs durch Wiederholung kann als Korrekturfaktor noch eingebracht werden, wenn man die Daten aus der Erhebung im Jahre 2000 heranzieht. Danach vermindert sich die Überschätzung an den Hauptschulen um  $d = 0,05$ , an den Realschulen um  $d = 0,02$  und an den Integrierten Haupt- und Realschulen um  $d = 0,03$ , während sie an den Gesamtschulen gänzlich verschwindet. Für die Gymnasien ist die Überschätzung um  $d = 0,03$  höher als zuvor anzunehmen.

Angesichts der zahlreichen ungeprüften Annahmen, auf denen diese Rechnungen beruhen, wird im Weiteren von einer Adjustierung der Lernstandswerte

für das Ende der Klassenstufe 8 und damit auch der Lernentwicklungsraten abgesehen. *Grundlage der Analysen ist deshalb auch in diesem Kapitel die Schulform- bzw. Kursniveauzugehörigkeit der Schülerinnen und Schüler nach dem Stand am Ende der Klassenstufe 8.* Es bleibt jedoch zu beachten, dass sich die relative Position der Gesamtschulen gegenüber den anderen Schulformen wegen der angesprochenen Verzerrungen mit hoher Wahrscheinlichkeit etwas günstiger darstellt, als es die manifesten Befunde nahe legen, und zwar auch und gerade, weil hier Klassenwiederholungen selten sind. Die Relationen zwischen den übrigen Schulformen bleiben hingegen von den Verzerrungen weitgehend unbeeinflusst.

Mit den Lernstandswerten für die Fächer Mathematik, Deutsch (Sprache und Leseverständnis)<sup>33</sup> und Englisch stehen Grundlagen für eine Einschätzung der Lernentwicklungen in den verschiedenen Schulformen bzw. Kursniveaus zur Verfügung. Für den Lateinunterricht ist die Datenbasis zu schmal, um vergleichbar gut begründete Erkenntnisse zu gewinnen. Die Fremdsprache Französisch wurde in der LAU 9 zum ersten Mal getestet.

In diesem Kapitel soll, mit der so begründeten Ausnahme der Fächer Latein und Französisch, jeweils in drei Schritten vorgegangen werden. Zunächst wird für jede Schulform bzw. für jedes Kursniveau der durchschnittliche Lernzuwachs bestimmt und vergleichenden Betrachtungen unterzogen. Sodann sollen – analog dem Bericht über die Lernentwicklung in den Klassenstufen 5 und 6 (vgl. LEHMANN, GÄNSFUß & PEEK 1999) – die differenziellen Lernzuwachsraten in den Klassenstufen 7 und 8 mit Bezug auf die Lernausgangslage 1998 untersucht werden. Analysen über die Lernentwicklung an einzelnen Schulen, bei denen – im Falle äußerer Leistungsdifferenzierung – auch die Charakteristik der Klassen- bzw. Kurszusammensetzung zu Beginn der Klassenstufe 7 berücksichtigt wird, schließen jeweils die Darstellungen ab. Hinsichtlich der Klassen- und Kurszusammensetzung ist anzumerken, dass insbesondere an den Gesamtschulen durch die systembedingte Mobilität zwischen den Kursniveaus nicht allen Verzweigungen der Auf- und Abstufung nachgegangen werden kann. Nicht in jedem Fall haben die Schülerinnen und Schüler, die sich am Ende der Klassenstufe 8 in einer bestimmten Lerngruppe auf dem Kursniveau I befunden haben, durchgehend dieser Gruppe angehört. Somit treten hier formal die gleichen Probleme auf wie hinsichtlich des Schulformwechsels; sie vervielfachen sich aber durch die Häufigkeit von Umstufungen sowie durch die Möglichkeit wiederholten Auf- und Abstiegs,

---

<sup>33</sup> Rechtschreibwissen wurde in der Untersuchung zur Klassenstufe 7 wegen der gerade eingeführten Neuregelung der Rechtschreibung / Orthografie nicht berücksichtigt.

wobei nur unvollständige Daten zu den Übergangspunkten vorliegen. Aussagen auf der Ebene einzelner Kurse werden deshalb hier vermieden. In geringerem Umfang gelten diese Vorbehalte auch für die Integrierten Haupt- und Realschulen. Auch aus diesem Grunde sollen sich die Analysen einheitlich auf die Schulform- bzw. Kursniveauzugehörigkeit am Ende der Klassenstufe 8 beschränken.

## **4.2 Entwicklung der Fachleistung Mathematik**

In der nachstehenden Tabelle 28 sind die statistischen Kennwerte für die Fachleistung Mathematik zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 sowie die daraus gebildeten Differenzwerte, die für eine Schulform oder ein Kursniveau die durchschnittlichen Lernfortschritte wiedergeben, zusammengefasst. Grundlage sind hier (wie in den nachfolgenden Analysen für andere Aspekte der Schulleistung) alle Schülerinnen und Schüler, für die Lernstandswerte sowohl aus der Erhebung des Jahres 1998 als auch aus derjenigen des Jahres 2000 zur Verfügung stehen.

Diesen Werten ist zunächst zu entnehmen, dass es in den beiden Schuljahren zwischen den Erhebungen von 1998 und 2000 über alle Schulformen hinweg einen Lernzuwachs gegeben hat, der mit 0,54 Standardabweichungen, d. h. rechnerisch 0,27 Standardabweichungen pro Jahr, unter dem der Klassenstufen 5 und 6 ( $d = 0,34/\text{Jahr}$ ), aber im Bereich dessen liegt, was bundesweit für die Klassenstufe 8 aus der TIMS-Studie bekannt ist ( $d = 0,29$ ; eigene Berechnung nach MARTIN & KELLY 1997, S. 95 f.). Allerdings verteilen sich die Zuwächse sehr ungleich auf die verschiedenen Schulformen bzw. Kursniveaus. Sie sind, absolut betrachtet, durchschnittlich am höchsten in den R-Kursen der Integrierten Haupt- und Realschulen, in den Realschulen und in den Kursen I der Gesamtschulen. Bemerkenswert groß sind die Lernzuwächse auch in den Kursen II der Gesamtschulen, insbesondere im Vergleich mit den Hauptschulen und den H-Kursen der Integrierten Haupt- und Realschulen. Demgegenüber verzeichnen die Gymnasien die niedrigsten absoluten Lernfortschritte.

Tabelle 28 Fachleistung Mathematik: Lernstandswerte zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 (Mittelwert, Standardabweichung) sowie Lernzuwachs (Mittelwert, Effektstärke und Varianzminderung), nach Schulform bzw. Kursniveau<sup>34</sup>

	Beginn Klassenstufe 7		Ende Klassenstufe 8		Lernzuwachs		
	$\bar{x}$	SD	$\bar{x}$	SD	Differenz	Effekt- stärke $d$	Varianz- minderung
Hauptschule	91,5	12,9	104,4	12,3	12,9	1,00	8,8 %
Realschule	104,6	16,9	120,1	14,3	15,4	0,91	28,4 %
Gymnasium	128,8	22,4	139,5	14,5	10,7	0,48	58,0 %
IHR Schule – H	93,3	15,2	104,6	12,3	11,3	0,74	34,6 %
IHR Schule – R	104,3	17,3	120,2	13,6	15,9	0,92	38,5 %
Gesamtschule II	92,9	13,8	107,7	13,0	14,8	1,08	10,4 %
Gesamtschule I	110,4	17,6	125,7	15,1	15,3	0,87	25,8 %
<i>insgesamt</i>	<i>111,9</i>	<i>24,1</i>	<i>124,9</i>	<i>19,6</i>	<i>13,0</i>	<i>0,54</i>	<i>33,6 %</i>

Fragt man nach den *relativen* Lernfortschritten, ergibt sich ein etwas anderes Bild. Unter Berücksichtigung der Veränderungen nach Maßgabe der Anfangsstreuung innerhalb der Schulform oder des Kursniveaus, die hier ebenso wie für die Gesamtheit über die Effektstärke  $d$  bestimmt sind, ist die Lernentwicklung in den Kursen II der Gesamtschulen und an den Hauptschulen besonders günstig verlaufen, gefolgt von den R-Kursen der Integrierten Haupt- und Realschulen und den Realschulen. Die zuletzt genannten Werte stehen, dem Berechnungsmodus entsprechend, in Abhängigkeit von der Homogenität der Fachleistung innerhalb der Schulform bzw. des Kursniveaus; man kann daraus also nicht ohne Weiteres auf einen Vorteil des Unterrichts mit

<sup>34</sup> Diese Werte stimmen nicht überall genau mit früher berichteten überein, weil die Zuordnung in der Tabelle über die Schulform- bzw. Kursniveauezugehörigkeit am Ende der Klassenstufe 8 erfolgt ist und weil hier nur diejenigen Schülerinnen und Schüler berücksichtigt sind, für die Testwerte aus beiden Erhebungen vorliegen. Daraus ergeben sich zwangsläufig Unschärfen. Für 9.388 Schülerinnen und Schüler konnten Differenzwerte berechnet werden, d. h. für 85,5 Prozent derjenigen, die an der LAU 7 teilgenommen haben, bzw. für 91,5 Prozent derjenigen, für die Daten aus der Erhebung im September 2000 vorliegen. Gründe für diese Fluktuation ergeben sich aus unklaren Schüleridentifikationen, aus der Wohnortmobilität der Eltern sowie aus Klassenwiederholungen und dem Überspringen von Klassenstufen. Auch Schulformwechsel kann wegen der dann häufiger auftretenden Schwierigkeit, einen Schüler oder eine Schülerin aus einer anderen Schule eindeutig zu identifizieren, zu den Abweichungen beigetragen haben.

leistungshomogenen Lerngruppen schließen; dies zeigt gerade das Beispiel der Hauptschulen, wo die absoluten Lernfortschritte knapp durchschnittlich sind, während der hohe Wert für die relativen Fortschritte ausschließlich aus der geringen Ausgangsstreuung (auf niedrigem Anfangsniveau) resultiert. Auffällig ist aber auch in dieser Perspektive des Bezugs zur Anfangsstreuung der niedrige Lernzuwachs an den Gymnasien. Da er sich in beiden Betrachtungsweisen so deutlich von dem der anderen Schulformen bzw. Kursniveaus abhebt, ist zu prüfen, ob hier ggf. ein Artefakt vorliegt.

Wie bereits angedeutet, kann für die gymnasiale Testversion ein Deckeneffekt ausgeschlossen werden. Lernstandswerte über 180 wurden bei einem Testumfang von 44 Items bereits mit 38 (das entspricht 86 Prozent) korrekt gelösten Aufgaben erzielt; der Anteil der Schülerinnen und Schüler in diesem Bereich besonders guter Fachleistungen beträgt an den Gymnasien jedoch nur 0,4 Prozent. Auch ein Zeitproblem käme für die Erklärung der unerwarteten Seltenheit von besonders hohen Lernstandswerten unter den Gymnasiastinnen und Gymnasiasten in Betracht. Nur eines der besonders selten korrekt gelösten Items befindet sich jedoch (mit der laufenden Nummer 40) in der Nähe des Endes dieser Testversion, so dass der negative Einfluss einer unbeabsichtigten „Speed“-Komponente nahezu ausgeschlossen werden kann. Eine dritte denkbare Erklärung für die unerwartet niedrigen Lernfortschritte könnte in der mangelnden curricularen Einschlägigkeit einiger Testaufgaben in der Gymnasialversion liegen. Die Lösungswahrscheinlichkeiten liegen indessen nur bei vier der 44 Aufgaben unterhalb der Ratewahrscheinlichkeit. Dabei handelt es sich um drei Algebraaufgaben und eine Aufgabe aus dem Bereich der Geometrie. Charakteristisch für diese Items ist weniger ein ungewohnter Inhalt als vielmehr ein ungewöhnliches Format (z. B. die Frage nach einem Parameter statt nach der Lösungsmenge), der Wechsel zwischen Dezimaldarstellung und Bruchdarstellung in der richtigen Antwortwahl oder die Zweischrittigkeit der Lösung bei der Geometrieaufgabe. Man kann in diesem Zusammenhang auf den Befund aus der TIMS-Studie verweisen, dass es selbst im hohen Leistungsbereich verhältnismäßig wenigen deutschen Schülerinnen und Schülern gelungen ist, mit der geforderten Flexibilität prinzipiell Bekanntes in einem ungewohnten Zusammenhang zielführend anzuwenden.

Für die Erklärung der vergleichsweise geringen Lernzuwächse an den Gymnasien käme prinzipiell auch ein Argument in Frage, das sich auf die Verzahnung der Testversionen von 1998 und 2000 bezieht. Wenn gezeigt werden könnte, dass die Ankeritems nicht auf die in den Klassenstufen 7 und 8 stattfindenden Lernprozesse reagieren, also bei der zweiten Erhebung keine höheren Lösungswahrscheinlichkeiten aufweisen als bei der ersten, wäre die

neue Gesamtmatrix angreifbar, und insbesondere käme es zu einer Unterschätzung der auf der Grundlage des neuen Tests ermittelten Lernstände. Sorgfältige Überprüfungen haben jedoch ergeben, dass dies nur für eines der zehn Ankeritems zutrifft. Die durchschnittliche Lösungswahrscheinlichkeit dieser Aufgaben konnte in der LAU 9 mit  $p = 0,57$  bestimmt werden, während sie in der LAU 7  $p = 0,42$  betrug. Hieraus und aus der Standardabweichung für die Ankeritems zum ersten Messzeitpunkt ergibt sich auch der Wert für den allgemeinen Lernzuwachs in den Klassenstufen 7 und 8 insgesamt ( $d = 0,53$ ), der demjenigen entspricht, der bereits berichtet wurde und wegen der guten Übereinstimmung mit den Ergebnissen aus der TIMS-Studie als plausibel gelten kann.<sup>35</sup>

---

<sup>35</sup> Weitere Kontrollrechnungen, die vor allem und exemplarisch das Ziel hatten, die unerwartet geringen Lernzuwächse an den Gymnasien gegen denkbare methodische Einwände abzusichern, können hier nur skizziert werden, weil es dabei z. T. um schwer eindeutig lösbare technische Fragen geht. So wurde z. B. geprüft, ob die Verwendung der Testrohwerte statt der Raschwerte zu anderen Ergebnissen führt. Dies konnte wegen der unterschiedlichen Testversionen in der LAU 9 nur innerhalb der Schulformen erfolgen. Gravierende Abweichungen hinsichtlich der korrelativen Zusammenhänge zwischen Eingangs- und Ausgangsmessung zeigten sich jedoch nicht (maximale Abweichung zwischen den Korrelationen an den Hauptschulen mit einer Differenz von 0,01). Somit kann ein Einfluss des Skalierungsverfahrens auf die hier zu prüfenden Befunde weitgehend ausgeschlossen werden. Auch eine Kovarianzanalyse mit Messwiederholung unter statistischer Kontrolle der kognitiven Eingangselektivität der Schulformen bzw. Kursniveaus bestätigt (mit Ausnahme der Hauptschule, die dann etwas niedrigere Zuwächse aufweist als das Gymnasium) die hier berichteten Befunde. Ein kovarianzanalytischer Überprüfungsansatz, bei dem der Einfluss der Schulform bzw. des Kursniveaus auf den Lernstand zum Ende der Klassenstufe 8 so geprüft wird, dass die Lernstände vom Anfang der Klassenstufe 7 als Kovariate verwendet werden und dann die Schulform bzw. das Kursniveau als Faktor behandelt wird, scheidet aus, weil die Residuen zwar definitionsgemäß unkorreliert sind mit den Prätestwerten, nicht aber mit den Posttestwerten. Dieser Zusammenhang beträgt modell- und messfehlerbedingt  $r = 0,69$  über alle Schulformen bzw. Kursniveaus hinweg und reicht *innerhalb* der Schulformen bzw. Kursniveaus von  $r = 0,55$  (Gymnasium) bis  $r = 0,80$  (Hauptschule). Innerhalb der Schulformen sind die Residuen außerdem negativ mit den Prätestwerten korreliert ( $-0,41 < r < -0,15$ ). Deshalb können die Mittelwerte der Residuen nicht als Indikatoren für den Lernzuwachs interpretiert werden; gemäß den Annahmen des Modells wäre dies nur dann zulässig, wenn innerhalb der Faktorausprägungen (hier: der Schulformen bzw. Kursniveaus) die Residuen unkorreliert mit den Prä- und Posttestwerten wären. Dies bedeutet insbesondere für die Gymnasien, dass mit diesem Ansatz der Lernzuwachs überschätzt würde, während er für die Schulformen und Kursniveaus mit niedrigeren Prätestwerten unterschätzt wäre. Besser begründbar wäre eine Analyse, die mit messfehlerbereinigten Posttestwerten arbeitet (vgl. dazu LEHMANN, HUSFELDT & PEEK 2001). In diesem Falle würden die Lernfortschritte an den

Weiterhin zeigen die Angaben der Tabelle 28 erneut, dass sich die Streuung der Mathematikleistungen im Hamburger Schulsystem, die bereits in den Klassenstufen 5 und 6 deutlich geschrumpft war, in den Klassenstufen 7 und 8 weiter vermindert hat, und zwar von  $SD = 24,1$  auf  $SD = 19,6$  Punkte, d. h. um 33,6 Prozent oder mehr als ein Drittel der Anfangsvarianz. Nachdem in keiner der eingesetzten Testversionen und in keiner der untersuchten Schulformen bzw. Kursniveaus ein ausgeprägter Decken- oder Bodeneffekt festzustellen war, insbesondere auch nicht in der gymnasialen Version, ist davon auszugehen, dass hinter diesen Werten tatsächlich – und entgegen den Erwartungen an ein mehrgliedriges Schulsystem – eine starke Homogenisierung der Leistungsverteilung steht. Wie in den Klassenstufen 5 und 6 bereits zu beobachten war, geht diese Varianzminderung in beträchtlichem Maße auf eine Streuungsminderung innerhalb der Schulformen bzw. Kursniveaus zurück, doch diese erfolgt in sehr ungleichem Maße. Während an den Hauptschulen mit 8,8 Prozent und in den Kursen II der Gesamtschulen mit 10,4 Prozent die Homogenisierung vergleichsweise gering ist, beträgt sie an den Gymnasien mit 58,0 Prozent mehr als die Hälfte der Anfangsvarianz. Daran ist bemerkenswert, dass es offenkundig einen engen Zusammenhang zwischen den relativen Wachstumsraten (den Effektstärken  $d$ ) und der Tendenz zur Homogenisierung der Fachleistungen innerhalb der Schulformen (der Varianzminderung) gibt: Je stärker sich die Fachleistungen innerhalb einer Schulform oder eines Kursniveaus einander angleichen, desto niedriger ist der relative – und im Falle

---

Gymnasien mit einem Differenzwert von 13,1 Punkten zwar höher als an den H-Kursen der Integrierten Haupt- und Realschulen (9,6), den Hauptschulen (10,5) und den Kursen II der Gesamtschulen (12,9) erscheinen, aber immer noch niedriger als an den Realschulen (14,5), den R-Kursen der Integrierten Haupt- und Realschulen (15,4) und den Kursen I der Gesamtschulen (15,7). Auch dabei würden jedoch die Lernfortschritte in dem Maße überschätzt werden, wie günstigere Anfangsbedingungen gegeben sind, weil die Messfehlerkorrektur auf dem Gesamtdatensatz beruht und so bei den für die Gymnasien typischen überdurchschnittlichen Prätestwerten höhere Differenzwerte erzeugt als bei unterdurchschnittlichen. Angesichts dieser hier letztlich nicht befriedigend zu lösenden Probleme wird es im vorliegenden Bericht vorgezogen, unmittelbar mit den Ausgangsdaten zu arbeiten, auch wenn dabei der Einfluss der sog. „Regression zur Mitte“ nicht explizit berücksichtigt wird. Hierfür wäre es nach dem oben Gesagten notwendig, zwischen entsprechenden Effekten innerhalb der Schulformen bzw. Kursniveaus und solchen zwischen ihnen analytisch zu trennen. In jedem Falle aber kann nach diesen Zusatzprüfungen festgestellt werden, dass mit den vorliegenden Daten die Annahme unvereinbar ist, die Lernfortschritte seien in den Klassenstufen 8 und 9 an den Gymnasien am größten. Eine noch weiter gehende Überprüfung der unerwartet niedrigen Lernzuwächse an den Gymnasien, etwa über die Analyse differenzieller Itemfunktionen, übersteigt die Möglichkeiten dieses Berichts.

der Gymnasien auch der absolute – Leistungszuwachs (die Korrelation zwischen Varianzminderung und der Effektstärke  $d$  für den relativen Lernzuwachs beträgt für die verschiedenen Schulformen bzw. Kursniveaus  $r = -0,90$ ). Gleichzeitig gibt es auf Schulform- bzw. Kursniveauebene einen Zusammenhang zwischen der Ausgangsvarianz und der Tendenz zur Homogenisierung ( $r = 0,90$ ): Je heterogener die Fachleistung Mathematik in einer Schulform oder auf einem Kursniveau zu Beginn der Klassenstufe 7 war (wobei das Gymnasium wiederum den Extremfall darstellt), umso stärker macht sich die Tendenz zum Ausgleich bemerkbar, so dass sich die Schulformen bzw. Kursniveaus am Ende der Klassenstufe 8 hinsichtlich der Streuung ihrer Fachleistungen deutlich weniger unterscheiden als zu Beginn der Klassenstufe 7. Es hat demnach den Anschein, als stelle ein auf Ausgleich gerichteter, kompensatorischer Unterricht vielfach eine Reaktion auf besonders heterogene mathematische Fachleistungen dar, die zugleich – zumindest auf Schulform- bzw. Kursniveauebene – mit vergleichsweise geringeren durchschnittlichen Lernfortschritten einhergeht. Im Fach Mathematik in den Klassenstufen 7 und 8 konvergieren also die Fachleistungen zwischen den Gymnasien und den anderen Schulformen. Dagegen entwickeln sie sich zwischen den Zweigen bzw. Kursniveaus an den Haupt- und Realschulen auseinander, und schließlich gibt es eine nahezu parallele Lernentwicklung auf den beiden Kursniveaus der Gesamtschulen. Dies entspricht im Wesentlichen der „Architektur“ des mehrgliedrigen Schulsystems in Hamburg, für die die Wahrung der Durchlässigkeit – insbesondere im Hinblick auf den Übergang in die gymnasiale Oberstufe – ein wichtiges Ziel darstellt. Dem steht entgegen, dass sich die mit der Schulform- bzw. Kursniveaueugehörigkeit verbundene Varianz in den Klassenstufen 7 und 8 von  $\eta^2 = 0,40$  auf  $\eta^2 = 0,49$  erhöht hat, was freilich auch mit der allgemeinen Varianzreduktion zusammenhängt. Es unterstreicht dies noch einmal die in allen Schulformen zu beobachtende Tendenz zur Verringerung des Leistungsspektrums, die durch Klassenwiederholung und Schulformwechsel noch verstärkt wird.<sup>36</sup>

Daneben ist aber auch danach zu fragen, welche Erkenntnisse sich hinsichtlich der Förderung der einzelnen Schülerinnen und Schüler, bezogen auf die Lernstände zu Beginn der Klassenstufe 7, gewinnen lassen. Abbildung 18

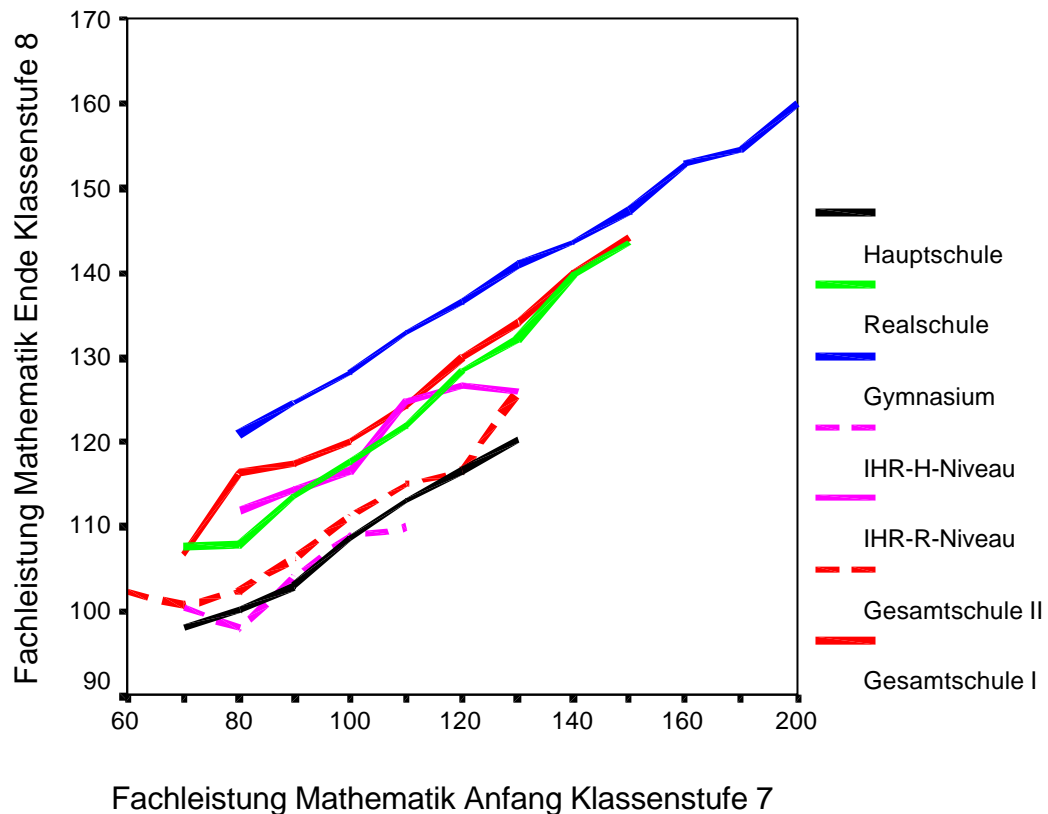
---

<sup>36</sup> Unbeschadet der oben dargelegten methodischen Probleme wurde diese Aussage mit dem extrem konservativen kovarianzanalytischen Ansatz geprüft: Unter Berücksichtigung der Lernstände zu Beginn der Klassenstufe 7 (Kovariate) gilt für alle Schulformen bzw. Kursniveaus, dass die Lernzuwächse in den anfangs leistungsschwächeren Gruppen höher sind als in den anfangs leistungsstarken (vgl. dazu auch LEHMANN, GÄNSFUß & PEEK 1999, S. 87 ff.).



stellt diese Zusammenhänge dar.<sup>37</sup>

Abbildung 18 Fachleistung Mathematik: Lernstände am Ende der Klassenstufe 8 nach Testleistung zu Beginn der Klassenstufe 7 und Schulform bzw. Kursniveau (Lernstandswerte)



In dieser Darstellung werden Zusammenhänge sichtbar, die bereits für die Lernentwicklung in den Klassenstufen 5 und 6 festgestellt worden waren (vgl. LEHMANN, GÄNSFUß & PEEK 1999, S. 114 ff.). So zeigt sich schon an den jeweiligen Wertebereichen, dass die verschiedenen Schulformen bzw. Kursniveaus auf der Grundlage sehr unterschiedlicher Lernausgangslagen zu Beginn der Klassenstufe 7 gearbeitet haben, was nicht ohne Einfluss auf die

<sup>37</sup> Gegen diese Form der Darstellung kann man den Einwand erheben, dass aus Gründen des Zusammenwirkens der Messfehler aus erster und zweiter Erhebung die „wahren“ Lernstände am Ende der Klassenstufe 8 am unteren Ende des Leistungsspektrums überschätzt und am oberen Ende unterschätzt werden (sog. „Regression zur Mitte“). Diese Erscheinung beeinflusst die Ordinate des Koordinatenkreuzes, betrifft aber alle Schulformen bzw. Kursniveaus, so dass sich an den Vergleichen qualitativ nichts ändert. Eine numerische Korrektur ist prinzipiell möglich, doch wird hier aus Gründen der Wahrung der Anschaulichkeit darauf verzichtet. Im Übrigen sind die Lernstandswerte für den Beginn der Klassenstufe 7 in Intervallen von je 10 Punkten gerundet und zusammengefasst. Gruppen mit einem Umfang von weniger als 10 Fällen sind in der Grafik nicht dargestellt.

Lernentwicklung bleiben konnte. Wie zu erwarten war, sind daher bei gleichem Anfangslernstand die Lernzuwächse umso größer, je anspruchsvoller die besuchte Schulform bzw. das Kursniveau ist, und zwar insbesondere im unteren Leistungsbereich. Dies betrifft auch und gerade diejenigen (wenigen) Gymnasiastinnen und Gymnasiasten, die sich zu Beginn der Klassenstufe 7 noch auf einem niedrigen Leistungsniveau befanden. Aber auch an den übrigen Schulformen, in denen eine Differenzierung entweder in Klassenverbände oder in fachspezifische Kursniveaus stattgefunden hat, gilt diese Beziehung ohne jede Ausnahme. Wenn dieser Effekt an den Gesamtschulen besonders stark ausgeprägt ist, zugleich aber die durchschnittlichen Lernzuwächse in den beiden Kursniveaus annähernd gleich sind, so hängt dies auch damit zusammen, dass sich hier die Verteilungen der Anfangswerte, dem vergleichsweise niedrigen Anteil der Schülerinnen und Schüler auf dem höheren Differenzierungsniveau entsprechend, zwischen den Niveaus stärker unterscheiden als an den Haupt- und Realschulen.

Der vermeintliche Widerspruch zwischen den – gemessen an der Lernausgangslage – hohen individuellen Lernzuwächsen der Gymnasiastinnen und Gymnasiasten und dem vergleichsweise niedrigen durchschnittlichen Lernfortschritt dieser Schulform liegt wiederum in dem Verhältnis der Anfangs- zur Endverteilung begründet. Die auffällige Schrumpfung der Leistungsvarianz am Gymnasium, die oben hervorgehoben wurde, geht *de facto* mit relativ geringen Zuwächsen schon im anteilmäßig starken mittleren, erst recht aber im oberen gymnasialen Leistungsbereich einher, so dass der Gesamtzuwachs deutlich hinter dem der anderen Schulformen bzw. Kursniveaus zurückbleibt. Demgegenüber sind vor allem in denjenigen Bildungsgängen, die mindestens auf einen mittleren Abschluss abzielen, die Zuwächse in den jeweiligen mittleren Leistungsbereichen erkennbar größer.

Auf das Argument, die besondere Situation der Gymnasien sei auch eine Folge von Klassenwiederholungen und Schulformwechsel, ist oben bereits eingegangen worden. Zwar würde sich einerseits der Vorsprung der Schülerinnen und Schüler des Gymnasiums vor denen der Real- und Gesamtschulen unter Berücksichtigung dieser Effekte bei der Bemessung der Lernzuwächse um durchschnittlich 0,8 bzw. 2,5 Punkte vermindern. Andererseits haben, wie sich zeigen lässt, die aufnehmenden Schulen von dem Schulformwechsel durchaus profitiert, weil die davon betroffenen Schülerinnen und Schüler an den Realschulen leistungsmäßig eine stark überdurchschnittliche Gruppe darstellen ( $d = 0,83$ ) und auch an den Gesamtschulen noch den Durchschnitt jener übertreffen, die in allen drei leistungsdifferenzierten Fächern auf dem Kursniveau I unterrichtet worden sind ( $d = 0,04$ ).

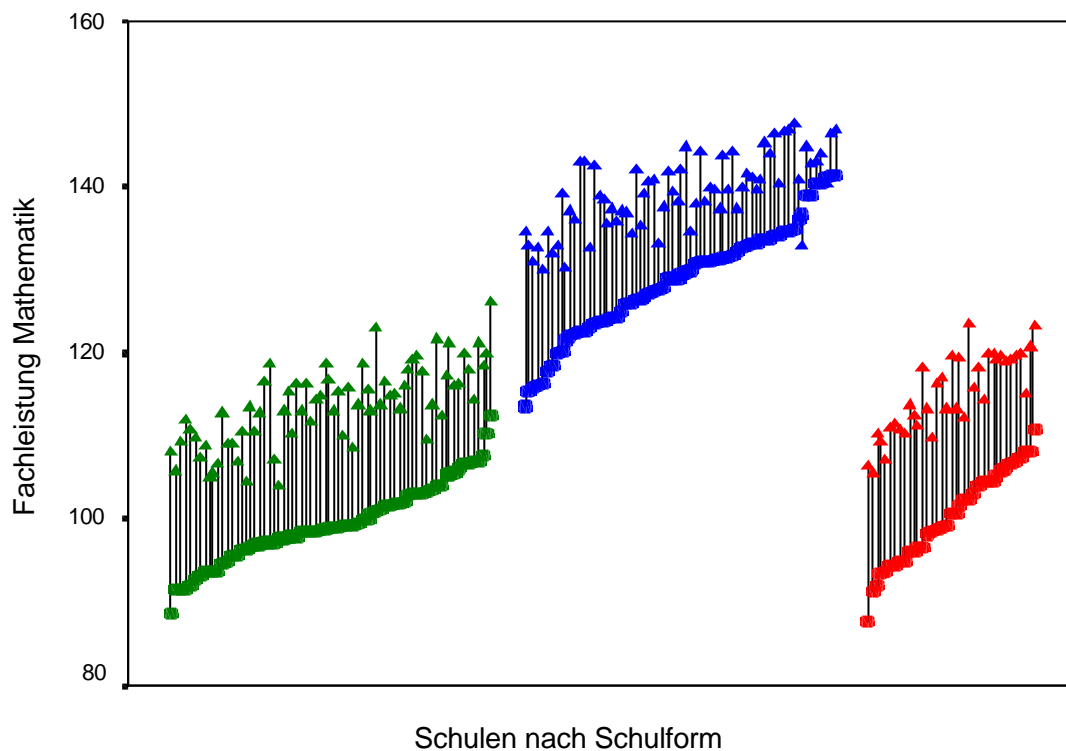
Schließlich ist zu prüfen, wie sich die durchschnittlichen Lernzuwächse zwischen den Einzelschulen unterscheiden. Klassenbezogene Analysen erweisen sich, zusätzlich zu dem bereits Gesagten, auch deshalb als schwierig, weil die Zuwachsraten von den Besonderheiten der Aufteilung der Schülerschaft in Schulzweige oder Kursniveaus abhängen oder jedenfalls abhängen können. Es war gezeigt worden, dass es in dieser Hinsicht große Unterschiede zwischen den Haupt- und Realschulen einerseits und den Gesamtschulen andererseits gibt. Dennoch liefert auch die Analyse auf Schulebene interessante Aufschlüsse. Im Hinblick auf die Herkunft der Schülerinnen und Schüler aus den noch undifferenzierten Organisationsformen der Beobachtungsstufen und der Gesamtschule wird dabei zunächst von einer Dreiteilung in Haupt- und Realschulen, Gymnasien und Gesamtschulen ausgegangen (vgl. Tabelle 29). Dabei sind nur diejenigen Schulen berücksichtigt, für die zu beiden Messzeitpunkten Lernstandswerte zur Verfügung stehen.

Tabelle 29 Fachleistung Mathematik zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 nach Schulform und Einzelschule

	<b>Beginn Klassenstufe 7</b>		<b>Ende Klassenstufe 8</b>		<b>Differenz</b>
	$\bar{x}$	SD	$\bar{x}$	SD	
Haupt- und Realschule	99,5	16,8	113,8	15,8	14,3
Gymnasium	128,8	22,6	139,5	14,5	10,7
Gesamtschule	101,0	17,9	116,0	16,6	15,1
<i>insgesamt</i>	<i>111,9</i>	<i>24,1</i>	<i>124,9</i>	<i>19,6</i>	<i>13,0</i>

Bei der Zusammenfassung der Differenzierungsniveaus zeigt sich in den Mittelwertdifferenzen, dass die Lernentwicklung in den Klassenstufen 7 und 8 an den Gesamtschulen insgesamt etwas günstiger verlaufen ist als an den (traditionellen und integrierten) Haupt- und Realschulen, obwohl dort der Anteil auf dem oberen Differenzierungsniveau geringer ist. An den auffallend niedrigen Zuwächsen der Gymnasien hat sich naturgemäß nichts geändert. Bei der Untersuchung der Lernstände und der Zuwächse auf Einzelschulebene werden freilich die großen Unterschiede sichtbar, die auch in den früheren Erhebungen zu erkennen waren (vgl. Abbildung 19).

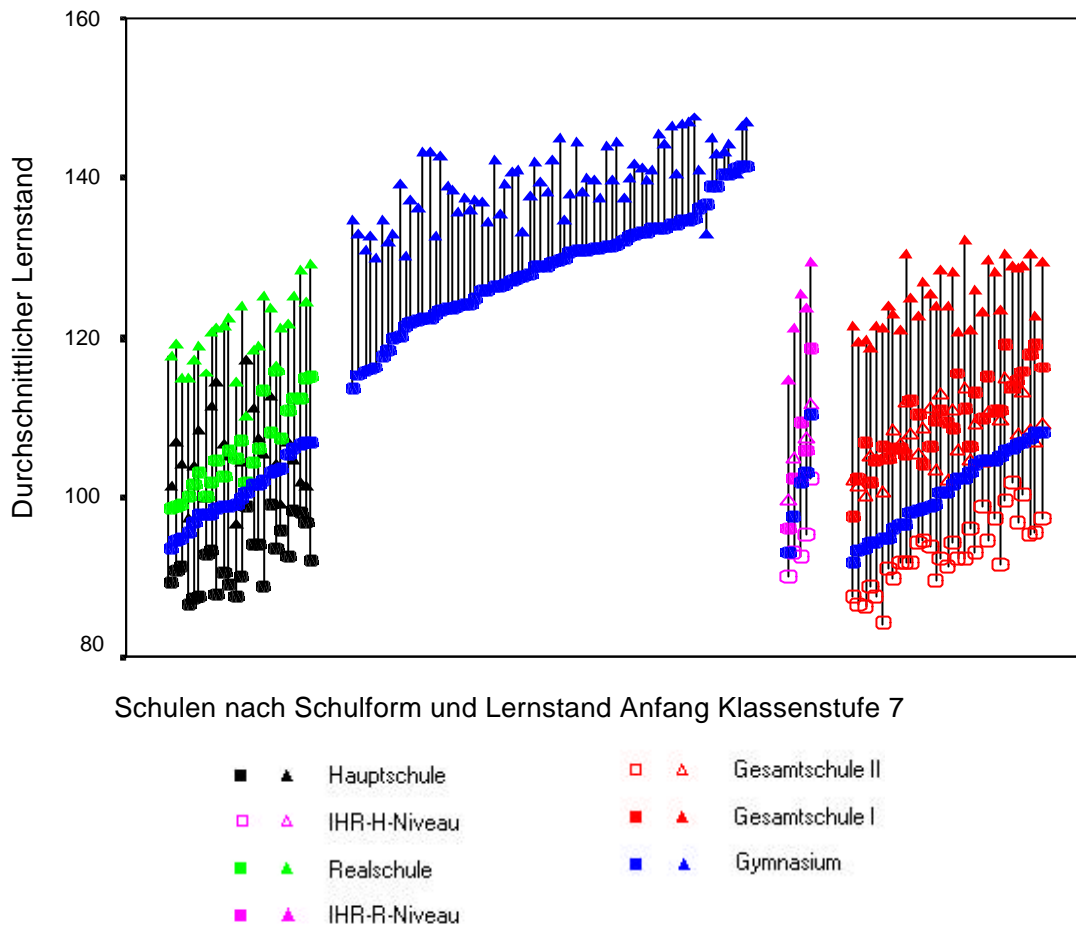
Abbildung 19 Fachleistung Mathematik: Lernentwicklung vom Beginn der Klassenstufe 7 bis zum Ende der Klassenstufe 8 nach Schulform und Einzelschule



In dieser Grafik sind die Lernzuwächse daran zu erkennen, dass jeweils ein Quadrat, das den durchschnittlichen Lernstand an einer Schule zu Beginn der Klassenstufe 7 als Lernausgangslage markiert, mit einem dazugehörigen Dreieck verbindet, welches den Lernstand am Ende der Klassenstufe 8 bezeichnet. Die verschiedenen Schulformen sind farblich abgesetzt: dunkelgrün die (traditionellen und integrierten) Haupt- und Realschulen, blau die Gymnasien, rot die integrierten Gesamtschulen. Innerhalb der jeweiligen Schulform sind die Einzelschulen entsprechend ihrer Lernausgangslage *vor Wirksamwerden äußerer Differenzierungsmaßnahmen* aufgereiht. Dabei wird sichtbar, dass es in allen Schulformen Schulen gibt, deren durchschnittlicher Lernstand am Ende der Klassenstufe 8 noch unterhalb der Lernausgangslage anderer Schulen zu Beginn der Klassenstufe 7 liegt. Die folgende Abbildung 20, in der nur diejenigen Schulen berücksichtigt sind, für die zu beiden Messzeitpunkten und ggf. auch für jeden Schulzweig bzw. jedes Kursniveau Lernstandswerte für mindestens 15 Schülerinnen und Schüler vorliegen, illustriert die komplexen Verhältnisse, die sich im Bezugsrahmen der Einzelschule aus der Überlagerung von Differenzierungsmaßnahmen und Lernentwicklung ergeben. Auch diese Grafik ist so angelegt, dass immer die Schulen einer Schulform zusammengestellt und nach der dort zu Beginn der Klassenstufe 7 erhobenen durchschnittlichen Fachleistung Mathematik geordnet sind (dargestellt durch blaue Quadrate). Weiterhin wird für diejenigen

Schulen, in denen eine äußere Differenzierung stattfindet, die spezifische Lernausgangslage des jeweiligen Differenzierungsniveaus markiert, an den traditionellen Haupt- und Realschulen z. B. durch schwarze (Hauptschule) und hellgrüne (Realschule) Quadrate. An den Integrierten Haupt- und Realschulen sind dies umrandete rosa Quadrate für die H-Kurse und ausgefüllte rosa Quadrate für die R-Kurse; eine analoge Darstellung (rot) wurde für die Gesamtschulen mit ihren Kursniveaus II (umrandet) und I (ausgefüllt) gewählt. Jeder Schulform bzw. jedem Differenzierungsniveau ist ein Dreieck gleicher Farbe für den Lernstand am Ende der Klassenstufe 8 zugeordnet. Demnach entspricht für jede Schule mit äußerer Fachleistungsdifferenzierung der Abstand der beiden Quadrate, die jeweils für den niveauspezifischen Lernstand zu Beginn der Klassenstufe 7 stehen, dem Grad, in dem fachleistungsspezifisch differenziert worden ist. Der Abstand zwischen den einander zugeordneten Quadraten und Dreiecken entspricht dem schul- bzw. kursniveauspezifischen Lernzuwachs. Für die Haupt- und Realschulen und für die Gesamtschulen wird nun sichtbar, wie nachhaltig die Differenzierung in Haupt- und Realschulzweige bzw. in Kursniveaus für das Fach Mathematik wirksam sein kann, dann nämlich, wenn z. B. eine Hauptschulklasse bis zum Ende der Klassenstufe 8 noch nicht den Durchschnitt der Lernausgangslage im Realschulzweig derselben Schule erreicht hat. In einzelnen Fällen ist am Ende der Klassenstufe 8 auf dem unteren Differenzierungsniveau nicht einmal der Gesamtmittelwert der Schule vom Anfang der Klassenstufe 7 erreicht worden. Zwar ist in stark negativ ausgelesenen Hauptschulklassen oder H-Kursen der Integrierten Haupt- und Realschulen – auch erkennbar am großen Abstand zwischen den Lernständen zu Beginn der Klassenstufe 7 vom Schulmittelwert – bis zum Ende der Klassenstufe 8 in der Regel ein kaum geringeres Leistungsniveau erreicht worden als in solchen, die sich zu Beginn der Klassenstufe 7 leistungsmäßig dichter an den Realschulklassen oder -kursen befanden (statistisch nicht signifikante Korrelation auf Schulebene:  $r = -0,15$  n.s.). Dem steht jedoch am Ende der Klassenstufe 8 ein ausnehmend hoher Leistungsstand in den besonders stark positiv ausgelesenen Realschulklassen bzw. -kursen gegenüber ( $r = 0,53$ ).

Abbildung 20 Fachleistung Mathematik: Lernzuwachs in den Klassenstufen 7 und 8 nach Schulform bzw. Kursniveau und Schulen



Allerdings ist zu beachten, dass die Neigung zu stark fachbezogener äußerer Differenzierung auf Schulebene ihrerseits mit dem Leistungsniveau zu Beginn der Klassenstufe 7 korreliert ( $r = 0,50$ ). Wenn man diese Tendenz statistisch berücksichtigt, halten sich die Vorteile vergleichsweise starker fachbezogener Differenzierung für die durchschnittliche Lernerentwicklung in Realschulklassen und -kursen einerseits und die negativen Konsequenzen für die Hauptschulklassen und -kurse andererseits die Waage.<sup>38</sup>

<sup>38</sup> Eine bestimmte Form der statistischen Berücksichtigung solcher Einflüsse von Drittvariablen auf den eigentlich interessierenden Zusammenhang bezeichnet man als „Auspartialisierung“; der dann zu berechnende Korrelationskoeffizient ist der Partialkorrelationskoeffizient. Im vorliegenden Fall beträgt die *partielle Korrelation* zwischen dem Differenzierungsgrad einer Schule, gemessen über die Leistungsdifferenz zwischen oberem und unterem Differenzierungsniveau am Anfang der Klassenstufe 7, und dem Leistungsstand am Ende der Klassenstufe 8 unter Berücksichtigung des Gesamtleistungsniveaus der Schule  $r(\text{partiell}) = 0,30$  für die Realschulklassen und  $r(\text{partiell}) = -0,30$  für die Hauptschulklassen.

An den Gesamtschulen sind andere Zusammenhänge zu beobachten. Generell wird hier im Fach Mathematik stärker fachleistungsbezogen differenziert als an Haupt- und Realschulen, was man nicht losgelöst davon sehen wird, dass an den Gesamtschulen der Anteil derer, die zum anspruchsvolleren Kurs zugelassen werden, wesentlich geringer ist. Auch an den Gesamtschulen findet zwar an Schulen mit höheren mathematischen Lernständen zu Beginn der Klassenstufe 7 auch eine konsequentere fachleistungsbezogene Differenzierung statt, doch die Korrelation ist mit  $r = 0,28$  n.s. merklich schwächer als an den Haupt- und Realschulen. Eine Tendenz der Art, dass durch eine stärker fachleistungsbezogene Differenzierung bzw. eine konsequentere Auslese Vorteile für die durchschnittliche Lernentwicklung in den Kursen I zu erzielen sind, lässt sich für die Gesamtschulen nicht belegen (partiell  $r = -0,12$  n.s.). Hingegen gibt es auch in den Gesamtschulen Anzeichen für Nachteile einer starken Auslese nach Fachleistung für die Lernentwicklung in den Kursen II (partiell  $r = -0,24$  n.s.).

Fasst man diese Befunde zusammen, so spricht manches dafür, dass im Fach Mathematik leistungsschwächere Schülerinnen und Schüler davon profitieren können, gemeinsam mit leistungsstärkeren unterrichtet zu werden. Ob für die Leistungsstärkeren durch positive Auslese ein höheres Lerntempo ermöglicht wird, bleibt zunächst eine offene Frage. Wie die Analysen der Zusammenhänge zwischen konsequenter Leistungsdifferenzierung und spezifischer Lernentwicklung auf den Differenzierungsniveaus (unter Berücksichtigung der gemeinsamen Lernausgangslage) gezeigt haben, scheint dies an den Haupt- und Realschulen tendenziell der Fall zu sein, dagegen eindeutig nicht an den Gesamtschulen. In beiden Schulformen kann aber nicht ausgeschlossen werden, dass auf dem jeweils unteren Differenzierungsniveau vergleichsweise geringere Lernfortschritte erzielt werden, wenn die leistungsfähigsten Schülerinnen und Schüler konsequent dem oberen Niveau zugewiesen werden. Ob durch einen anderen Mathematikunterricht trotz solcher systembedingter negativer Auslese auch auf dem unteren Differenzierungsniveau eine günstigere Lernentwicklung zu erzielen wäre, ist eine Frage, die auf der Basis der vorliegenden Daten nicht zu beantworten ist.

Das Problem der äußeren Differenzierung stellt sich an den Gymnasien nicht. Dafür aber legt die Abbildung 20 die Vermutung nahe, dass an denjenigen Gymnasien, an denen zu Beginn der Klassenstufe 7 relativ niedrige Lernstände zu verzeichnen waren, eine besonders günstige Lernentwicklung stattgefunden hat, so dass die oben als auffällig beschriebene Varianzmin- derung in dieser Schulform vor allem auf eine tendenzielle Angleichung der

mathematischen Fachleistung zwischen den Schulen zurückzuführen wäre.<sup>39</sup> Die letztgenannte Vermutung hält indessen einer näheren Überprüfung nur teilweise stand: Zwar korreliert der mittlere Lernzuwachs auf Schulebene negativ mit dem Lernstand zu Beginn der Klassenstufe 7 ( $r = -0,75$ ). Der Anteil der Zwischen-Schul-Varianz an der Gesamtstreuung ist jedoch mit 9,3 Prozent (Klassenstufe 7) bzw. 8,9 Prozent (Klassenstufe 9) kaum geschrumpft, und auf Klassenebene ist er mit 15,2 bzw. 15,1 Prozent praktisch gleich geblieben. Demnach muss sich die Homogenisierung der mathematischen Fachleistung an den Gymnasien auch auf Schulebene und innerhalb der Schulklassen vollzogen haben, ein Prozess, den man mit der schon für die Beobachtungsstufe festgestellten Tendenz im Zusammenhang sehen wird, eine gemeinsame Basis der meisten Schülerinnen und Schüler für künftige Lernprozesse im Fach Mathematik zu sichern. Schulformwechsel vermag jedenfalls die relativ großen Lernfortschritte an den weniger leistungsstarken Gymnasien nicht zu erklären; denn er kommt hier nicht nennenswert häufiger vor als an den schon anfangs leistungsstarken Schulen. Allerdings scheint es mit dieser Strategie einherzugehen, dass bei den Schülerinnen und Schülern mit hohem Leistungsstand und selbst an Gymnasien mit günstigen Ausgangsbedingungen in den Klassenstufen 7 und 8 vergleichsweise geringe Lernfortschritte erzielt werden, jedenfalls in dem Bereich der Mathematik, dessen Beherrschung durch den eingesetzten Test erfasst werden konnte.

### 4.3 Entwicklung der Fachleistung Deutsch

Wie bei der Fachleistung Mathematik sollen auch für die Fachleistung Deutsch – differenziert für *Leseverständnis* und *Sprache* – die Lernentwicklungen der Hamburger Schülerinnen und Schüler vom Beginn der Klassenstufe 7 bis zum Ende der Klassenstufe 8 aufgezeigt werden.

---

<sup>39</sup> Da es sich hier um Schulmittelwerte handelt, die jeweils auf einer recht großen Zahl von Einzelwerten beruhen, ist die Annahme, die niedrigen Werte am Ende der Klassenstufe 8 seien vor allem messfehlerbedingt, schwerlich plausibel. Eine (für den Einzelfall angemessene) und dann auf Schulebene aggregierte Messfehlerkorrektur würde die Lernstände am Ende der Klassenstufe 8 an Schulen mit niedrigem Anfangslernstand künstlich nach unten korrigieren und solchen mit hohen Anfangslernständen für das Ende der Klassenstufe 8 einen höheren Wert attestieren, als tatsächlich erreicht wurde. Aus diesem Grunde wurde auch hier auf die Möglichkeit einer solchen Adjustierung verzichtet.



## • Leseverständnis

Zunächst sollen für die Ergebnisse im Untertest *Leseverständnis* die statistischen Kennwerte für den Beginn der Klassenstufe 7 und das Ende der Klassenstufe 8 sowie die daraus gebildeten Differenzwerte als Indikatoren für die Lernfortschritte berichtet werden.

Tabelle 30 Fachleistung Leseverständnis: Lernstandswerte zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 (Mittelwert, Standardabweichung) sowie Lernzuwachs (Mittelwert, Effektstärke und Varianzminderung), nach Schulform bzw. Kursniveau

	Beginn Klassenstufe 7		Ende Klassenstufe 8		Lernzuwachs		
	$\bar{x}$	SD	$\bar{x}$	SD	Differenz	Effekt- stärke $d$	Varianz- minderung
Hauptschule	91,1	16,0	104,4	18,8	13,3	0,83	-38,0
Realschule	108,5	17,4	120,8	16,3	12,3	0,70	13,0
Gymnasium	126,0	15,5	134,2	13,0	8,2	0,53	29,9
IHR Schule – H	92,9	17,7	105,9	16,3	13,1	0,74	14,4
IHR Schule – R	105,7	18,2	119,4	15,9	13,8	0,76	23,1
Gesamtschule II	98,2	18,1	109,5	18,5	11,4	0,63	-5,3
Gesamtschule I	114,6	18,1	124,9	16,9	10,3	0,57	13,1
<i>insgesamt</i>	<i>112,2</i>	<i>21,2</i>	<i>122,6</i>	<i>19,4</i>	<i>10,4</i>	<i>0,49</i>	<i>16,5</i>

Hiernach ist für das Leseverständnis über alle Schulformen bzw. Kursniveaus hinweg eine Erhöhung der Lernstandswerte um 10,4 Punkte oder 0,5 Standardabweichungen festzustellen, was mit rechnerisch 0,25 Standardabweichungen pro Jahr annähernd dem Zuwachs in *Mathematik* entspricht. Allerdings sind die Unterschiede im Leseverständnis zwischen den Schulformen und Kursniveaus deutlich geringer ausgeprägt als im Fach *Mathematik*. Wiederum ist für die Gymnasien festzustellen, dass hier der Lernzuwachs wesentlich geringer ist als in den anderen Schulformen.

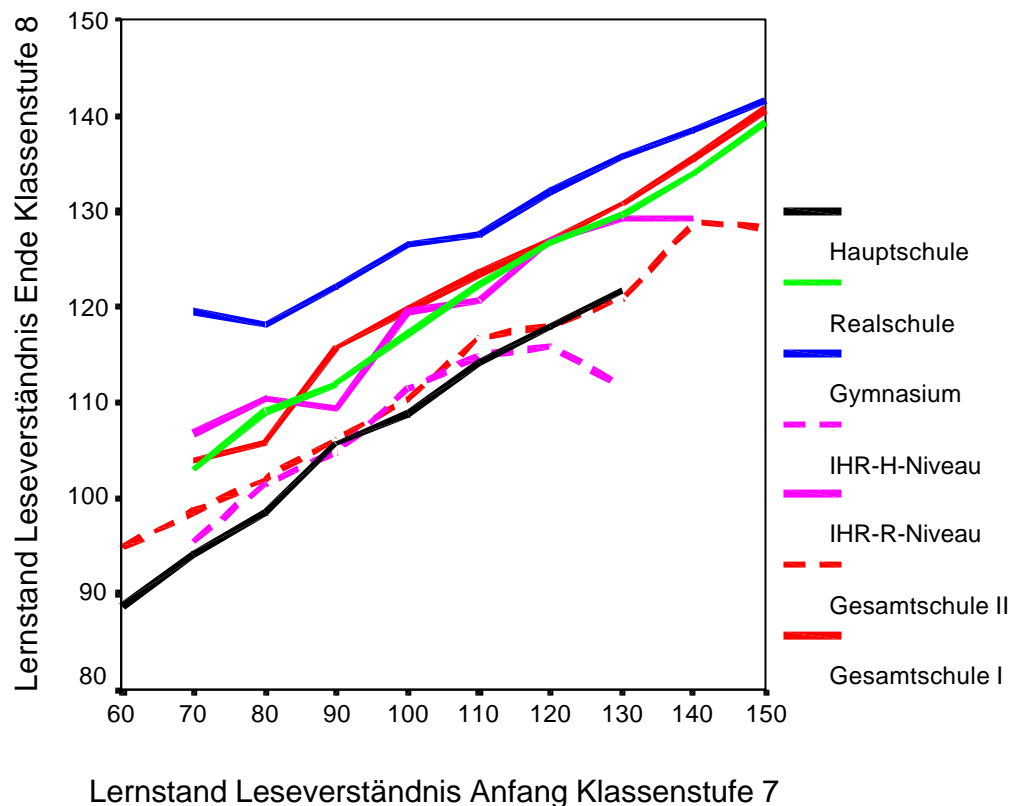
Weder ein Deckeneffekt des Tests noch ein Zeitfaktor sind dafür verantwortlich: Bei einem Umfang von 44 Items in der Gymnasialfassung des Leseverständnis-tests hat niemand alle Aufgaben gelöst, und lediglich 0,3 Prozent der Gymnasiastinnen und Gymnasiasten haben mehr als 40 richtige Lösungen erzielt. Darüber hinaus befand sich nur eine der besonders schwierigen Aufgaben in einer der Testversionen im hinteren Teil des Testhefts.

Gleichwohl ist denkbar, dass hinsichtlich des Leseverständnisses insofern andere Verhältnisse gegeben sind, als hier nicht unbedingt von einem linear-kumulativen Lernprozess ausgegangen werden kann, sondern vielleicht besser von allmählich verlangsamten Lernzuwächsen, sobald ein gewisses Niveau erreicht ist. Diese Annahme liegt u. a. deshalb nahe, weil auf dem Hauptschulniveau die größten Zuwächse zu verzeichnen sind, während neben den Gymnasien auch für die Kurse I der Gesamtschulen eine niedrigere Wachstumsrate festzustellen ist. Einen Sonderfall stellen die Kurse II der Gesamtschulen dar, in denen die Lernentwicklung absolut und relativ gesehen durchschnittlich langsamer verlaufen ist als auf dem sonst zumeist vergleichbaren Hauptschulniveau. Dennoch bleibt es auffällig, dass besonders gute Testergebnisse insgesamt selten sind. Hier bestätigt sich für Hamburg der Befund aus internationalen Leistungsstudien, demzufolge deutsche Schülerinnen und Schüler im oberen Leistungsbereich deutlich unterrepräsentiert sind (vgl. LEHMANN, PEEK, PIEPER & VON STRITZKY 1995; vgl. auch LEHMANN 2000).

Die allgemeine Verminderung der Varianz im Hamburger Gesamtsystem ist mit 16,5 Prozent im Falle des Leseverständnisses nur halb so stark ausgeprägt wie in der Fachleistung Mathematik (33,6 Prozent). Auch innerhalb der Schulformen bzw. Kursniveaus ist sie in der Regel schwächer ausgeprägt, und in den Kursen II der Gesamtschulen sowie an den Hauptschulen erhöht sich in den Klassenstufen 7 und 8 die Streuung sogar um 5,3 bzw. 38,0 Prozent. Die nahe liegende Vermutung, diese Besonderheit könnte auf den vergleichsweise hohen Ausländeranteil an den Hauptschulen bzw. in den Kursen II der Gesamtschulen zurückzuführen sein, trifft allerdings nicht zu. Die Überprüfung zeigt, dass dieser Anteil an den Hauptschulen mit 26,7 Prozent dem in den Kursen II der Gesamtschulen (26,6 Prozent) entspricht. Dagegen ist er in den H-Kursen der Integrierten Haupt- und Realschulen mit 36,7 Prozent deutlich höher, hier aber ist die von den anderen Schulformen und Kursniveaus her gewohnte Varianzminderung zu beobachten.

Angesichts dieses uneinheitlichen Bildes empfiehlt es sich wiederum, die Lernstände am Ende der Klassenstufe 8 in Abhängigkeit von der Schulform und den dort gegebenen Lernständen am Anfang der Klassenstufe 7 zu betrachten (vgl. Abbildung 21).

Abbildung 21 Fachleistung Leseverständnis: Lernstände am Ende der Klassenstufe 8 nach Testleistung zu Beginn der Klassenstufe 7 und Schulform bzw. Kursniveau (Lernstandswerte)



Auch hier gilt wieder, dass der individuelle Lernstand am Ende der Klassenstufe 8 umso höher ist, je günstiger die *Leistungsverteilung* zu Beginn der Klassenstufe 7 war, und zwar im Vergleich der Schulformen ebenso wie im Vergleich der Kursniveaus innerhalb einer Schulform. So ist bei Schülerinnen und Schülern gleicher Lesekompetenz am Anfang der Klassenstufe 7 der Lernstand am Ende der Klassenstufe 8 generell an den Gymnasien am günstigsten, und zwar um bis zu 14 Lernstandspunkte gegenüber der nächsten Schülergruppe<sup>40</sup>. Es folgen Schülerinnen und Schüler aus den Kursen I der Gesamtschulen, aus den Realschulen und den R-Kursen der Integrierten Haupt- und Realschulen; zwischen diesen drei Gruppen überschreiten die Abstände allerdings in keinem Bereich des Leistungsspektrums 5 Punkte. Demgegenüber ist die Lernentwicklung in den Kursen II der Gesamtschulen vor allem im unteren Leistungsbereich günstiger als an den Hauptschulen, und

<sup>40</sup> Dieser Effekt ist wahrscheinlich nicht völlig unbeeinflusst von Einflüssen der Klassenwiederholung und des Schulformwechsels. Man beachte aber, dass es sich bei den Schülerinnen und Schülern in den unteren Bereichen des Leistungsspektrums um solche handelt, die gerade *nicht* eine Klassenstufe wiederholen oder die Schulform wechseln mussten.

zwar um bis zu 6 Lernstandspunkte. Bei diesen Vergleichen ist erneut daran zu erinnern, dass die scheinbaren Widersprüche zu den in Tabelle 29 berichteten Befunden daraus resultieren, dass letztere die Ausgangsverteilung implizit durch Gewichtung der einzelnen Kompetenzniveaus berücksichtigen, während in der Abbildung 21 gerade von der Ausgangsverteilung abstrahiert wird. Dennoch ist es bemerkenswert, dass über das gesamte Leistungsspektrum hinweg betrachtet die Gymnasien, die Kurse I der Gesamtschule und die Realschulen in den Klassenstufen 7 und 8 konvergieren, auch wenn die Anteile der Schülerinnen und Schüler in den Gruppen mit besonders hoher Ausgangsleistung (d. h., mit Lernstandswerten über 135 zu Beginn der Klassenstufe 7) mit 35,6 Prozent an den Gymnasien, 16,1 Prozent in den Kursen I der Gesamtschulen und 8,1 Prozent an den Realschulen sehr verschiedenen Umfang haben. Bei der Interpretation dieser Befunde ist im Hinblick auf die nichtgymnasialen Schulformen insbesondere auch der mehrfach erwähnte unterschiedliche Anteil im anspruchsvolleren Differenzierungsniveau zu berücksichtigen. So ist es denkbar, dass sich restriktivere Formen der Fachleistungsdifferenzierung zu Gunsten der Lernentwicklung auf dem oberen Differenzierungsniveau auswirken.

Ehe auf die Lernentwicklung auf der Ebene der Einzelschulen eingegangen wird, werden in Tabelle 31 zunächst die Lernstände zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 berichtet, getrennt nach Haupt- und Realschulen, Gymnasien und Gesamtschulen.

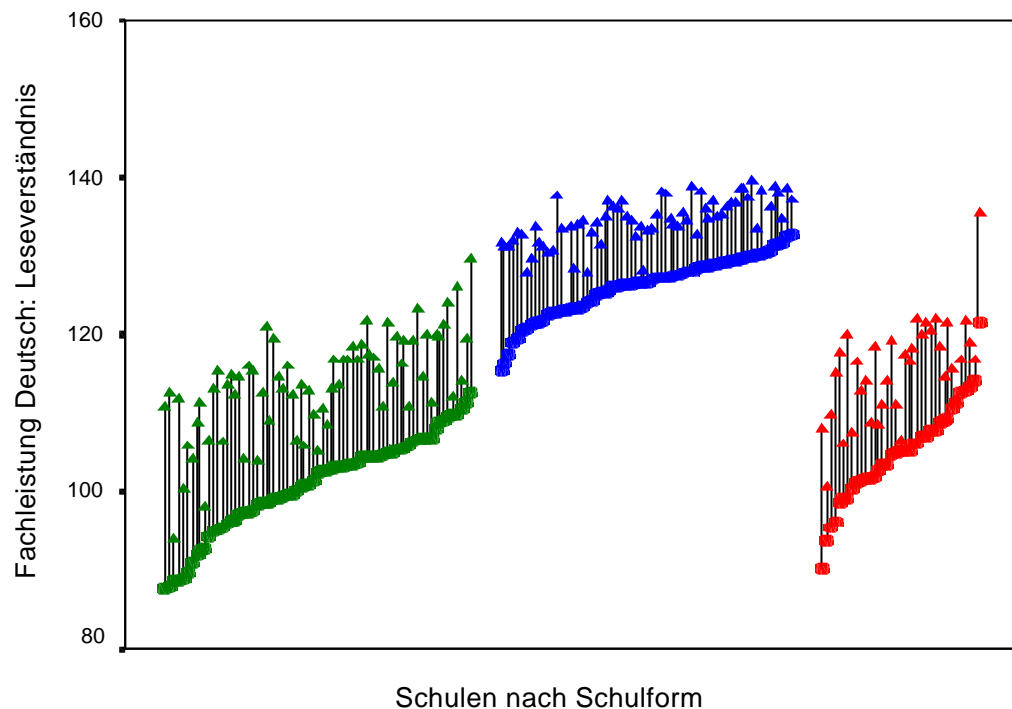
Tabelle 31 Fachleistung Leseverständnis zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8, nach Schulform und Einzelschule

	<b>Beginn Klassenstufe 7</b>		<b>Ende Klassenstufe 8</b>		Differenz
	$\bar{x}$	SD	$\bar{x}$	SD	
Haupt- und Realschule	101,2	19,0	114,1	18,8	12,8
Gymnasium	126,0	15,5	134,2	13,0	8,2
Gesamtschule	105,2	19,8	116,1	19,5	10,9
<i>insgesamt</i>	<i>112,2</i>	<i>21,2</i>	<i>122,7</i>	<i>19,4</i>	<i>10,4</i>

Im Bereich Leseverständnis sind demnach die Lernzuwächse in den Haupt- und Realschulen – hier zunächst verstanden als einheitliche Schulform, die an die gemeinsame Beobachtungsstufe anknüpft – am größten, gefolgt von den Gesamtschulen und den Gymnasien. Von Bedeutung ist diese Betrachtungsweise dann, wenn man wieder die Einzelschule in den Blick nimmt, innerhalb derer äußere Differenzierung und Lernentwicklung zusammenwirken. In

Abbildung 22 ist deshalb die Lernentwicklung im Bereich Leseverständnis innerhalb einzelner Schulen, getrennt nach Schulform und durchschnittlichem Lernstand zu Beginn der Klassenstufe 7, dargestellt, und zwar zunächst wieder in der vereinfachten Form ohne Berücksichtigung der Differenzierung ab Klassenstufe 7.

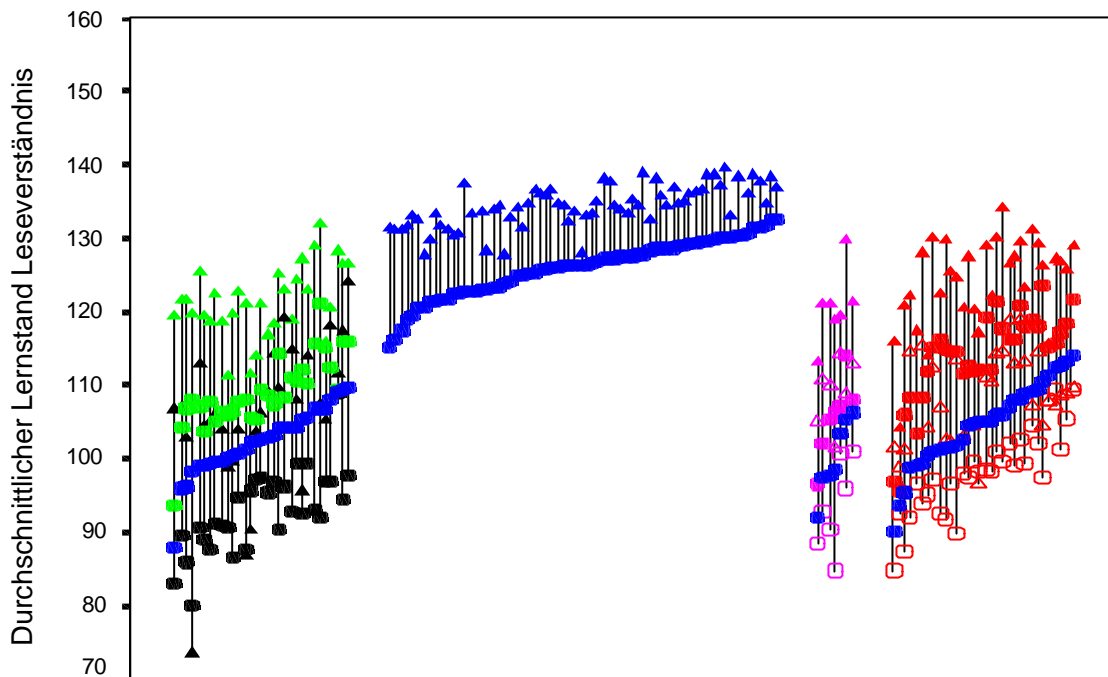
Abbildung 22 Fachleistung Leseverständnis: Lernentwicklung vom Beginn der Klassenstufe 7 bis zum Ende der Klassenstufe 8 nach Schulform und Einzelschule



Wie in der Abbildung 19 ist jede Schulform durch eine eigene Farbe gekennzeichnet; die schulspezifischen Lernausgangslagen werden durch Quadrate wiedergegeben, die mit Dreiecken verbunden sind, die dem durchschnittlich erzielten Lernstand am Ende der Klassenstufe 8 entsprechen. Die Länge der verbindenden Strecken repräsentiert den schulspezifischen Lernzuwachs. Dabei wird auch für den Bereich des Leseverständnisses deutlich, dass in allen Schulformen an einigen Schulen bis zum Ende der Klassenstufe 8 im Mittel das Niveau nicht erreicht wird, das für andere die Lernausgangslage zu Beginn der Klassenstufe 7 darstellt.

Auch hier ist es jedoch wieder notwendig, sich nicht auf die Einzelschule als administrative Einheit zu beschränken, sondern insbesondere dem Zusammenhang von äußerer Differenzierung und Lernentwicklung auf den jeweiligen Differenzierungsniveaus nachzugehen

Abbildung 23 Fachleistung Leseverständnis: Lernzuwachs in den Klassenstufen 7 und 8 nach Schulform bzw. Kursniveau und Schulen



Schulen nach Schulform und durchschnittlichem Lernstand

- |     |              |     |                 |
|-----|--------------|-----|-----------------|
| ■ ▲ | Hauptschule  | □ ▲ | Gesamtschule II |
| □ ▲ | IHR-H-Niveau | ■ ▲ | Gesamtschule I  |
| ■ ▲ | Realschule   | ■ ▲ | Gymnasium       |
| ■ ▲ | IHR-R-Niveau |     |                 |

Wie im Falle der Fachleistung Mathematik (vgl. die analoge Grafik) liegen angesichts der relativ hohen Übergangsquoten von Schülerinnen und Schülern aus der Beobachtungsstufe der Haupt- und Realschulen in Realschulklassen die Lernausgangswerte für die Realschulklassen deutlich dichter an den Schulmittelwerten, als dies etwa für die Kurse I an den Gesamtschulen der Fall ist. Umgekehrt sind die Hauptschulklassen leistungsmäßig stärker negativ ausgelesen als die Kurse II an den Gesamtschulen. Auch hier liegt deshalb die Frage nahe, ob es einen Zusammenhang zwischen dem Grad der Differenzierung (gemessen über die Unterschiede der Anfangsleistungen zwischen den beiden Differenzierungsniveaus) und der Leistungsentwicklung gibt.

An den Haupt- und Realschulen sind auf Schulebene die Hauptschulklassen offenbar durch eine konsequent leistungsbezogene Differenzierungspraxis beeinträchtigt – dadurch also, dass die leistungsstärksten Schülerinnen und Schüler nicht mehr im Klassenverband sind ( $r = -0,36$ ). Dies ist deshalb

besonders bemerkenswert, weil eine hohe Quote von Realschülerinnen und Realschülern für sich genommen keineswegs notwendig solche Folgen zu haben scheint ( $r = -0,11$  n.s.). Andererseits ist der Übergang der Realschülerinnen und Realschüler in eine homogenere, leistungsstärkere Lerngruppe offenbar in vielen Fällen vorteilhaft für deren Lernentwicklung ( $r = 0,32$ ), wobei der Realschulanteil an der Schülerschaft der Schule allein wiederum keinen Einfluss auf das Leistungsniveau in der Realschulgruppe hat ( $r = -0,09$  n.s.). Berücksichtigt man statistisch die Unterschiede im Anfangsniveau und die Übergangsquote auf die Realschule, so verstärken sich diese Zusammenhänge noch. Insgesamt kovariert an den Haupt- und Realschulen zwar die Übergangsquote in die Realschulklassen zu Beginn der Klassenstufe 7 mit dem durchschnittlichen Leseverständnis an der jeweiligen Schule ( $r = 0,45$ ), nicht aber der Grad, in dem die Differenzierung nach Maßgabe der tatsächlichen Fachleistung erfolgt ( $r = 0,08$  n.s.). Angesichts der leistungsmäßigen Vorteile solcher Differenzierung für die Realschülerinnen und Realschüler bei gleichzeitigen Nachteilen für die Schülerinnen und Schüler in Hauptschulklassen wird es verständlich, dass ein Zusammenhang zwischen der Differenzierungspraxis und dem Gesamtmittelwert für das an einer Haupt- und Realschule am Ende der Klassenstufe 8 erreichte Leseverständnis fehlt ( $r = -0,04$ ).

An den Gesamtschulen sind die Verhältnisse etwas anders. Bei dem dort üblichen Anteil von ca. 40 Prozent der Schülerschaft, der auf die Kurse I entfällt, scheinen aus einer konsequent leistungsorientierten Auslese bei der Bildung leistungsmäßig homogenerer Lerngruppen beide Kursniveaus Vorteile für die Lernentwicklung zu ziehen (Kurse II:  $r = 0,40$ ; Kurse I:  $r = 0,30$  n.s.); diese Effekte verschieben sich unter Konstanthaltung des Anfangsniveaus und des Anteils von Schülerinnen und Schülern auf dem höheren Kursniveau I zu dessen Gunsten (Kurse II: partielles  $r = 0,30$  n.s.; Kurse I: partielles  $r = 0,58$ ). Für die Kurse II sind mit einer konsequent leistungsorientierten Auslese also keine Nachteile für die Lernentwicklung, sondern tendenziell eher Vorteile verbunden; für die Kurse I fallen diese Vorteile sehr deutlich aus. Dabei hängt die Differenz zwischen dem oberen und dem unteren Kursniveau zu Beginn der Klassenstufe 7, also der Grad, in dem die Niveaudifferenzierung an die Fachleistung gekoppelt ist, nicht mit der durchschnittlichen Anfangsleistung an der Schule zusammen ( $r = 0,04$  n.s.), in relativ hohem Maße aber mit dem Schulmittelwert am Ende der Klassenstufe 8 ( $r = 0,40$ ).

Für die Gymnasien legt die Grafik wiederum den Eindruck nahe, der vergleichsweise geringe durchschnittliche Lernzuwachs der Gymnasien liege im Wesentlichen daran, dass an Schulen mit besonders günstigen Voraussetzungen zu Beginn der Klassenstufe 7 nur mehr geringe weitere Fortschritte erzielt wurden. Dieser Eindruck lässt sich auch quantitativ belegen: Der schulspezifische Zuwachs korreliert eindeutig negativ mit dem Schulmittelwert zu Beginn der Klassenstufe 7 ( $r = -0,60$ ), wonach die günstigere Lernentwicklung an Schulen mit ungünstiger Ausgangslage stattgefunden hat.

Gleichzeitig hat sich aber das Bestimmtheitsmaß  $Eta^2$  für die Schulzugehörigkeit innerhalb der Schulform in den beiden Jahren des Untersuchungszeitraums praktisch nicht verringert ( $Eta^2 = 0,06$ ). Daraus ist zu schließen, dass auch innerhalb der Schulen eine deutliche Reduktion der Leistungsvarianz stattgefunden hat, die übrigens im Durchschnitt mit  $SD = 14,9$  (Klassenstufe 7) zu  $SD = 12,2$  (Klassenstufe 9) recht genau den Verhältnissen in der gymnasialen Schülerschaft insgesamt entspricht. Zwischen den Schulklassen hingegen hat der spezifische Varianzanteil an den Gymnasien von  $Eta^2 = 0,11$  (Klassenstufe 7) auf  $Eta^2 = 0,14$  (Klassenstufe 9) zugenommen. Am deutlichsten ist auf Schulebene der Zusammenhang zwischen der Tendenz, die Binnenvarianz zu reduzieren, und dem durchschnittlichen Leistungszuwachs: je stärker die Tendenz, das Leistungsspektrum in der Klasse zu reduzieren, desto größer waren die durchschnittlichen Lerngewinne. Die Korrelation beträgt  $r = 0,63$  und verstärkt sich sogar noch, wenn man die Anfangsleistung berücksichtigt ( $r = 0,79$ ). M. a. W.: Im Hinblick auf das Leseverständnis verdankt sich der festgestellte Lernzuwachs in den Klassenstufen 7 und 8 an den Gymnasien also allem Anschein nach zum großen Teil den kompensatorischen Leistungen von Schulen mit vergleichsweise niedrigen Lernständen zu Beginn des Betrachtungszeitraums.

- **Sprache**

Wie bereits aufgezeigt wurde, lässt der Untertest *Sprache* ebenso wenig wie die anderen fachbezogenen Untertests einen Deckeneffekt erkennen, auch nicht für die Schülerinnen und Schüler an den Gymnasien. Dennoch sind die Lernzuwächse in diesem Untertest geringer als in den anderen untersuchten Leistungsbereichen, und eher noch als im Falle des Leseverständnisses könnte man vermuten, dass hier die Lernentwicklung nicht oder nicht mehr linear verlaufen ist, sondern sich mit abnehmenden Lernzuwächsen einem individuellen oder gruppenspezifischen Maximum annähert. In Tabelle 32 sind die Befunde, auf die sich diese Vermutung stützt, zusammengefasst.



Tabelle 32 Fachleistung Sprache: Lernstandswerte zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 (Mittelwert, Standardabweichung) sowie Lernzuwachs (Mittelwert, Effektstärke und Varianzminderung), nach Schulform bzw. Kursniveau

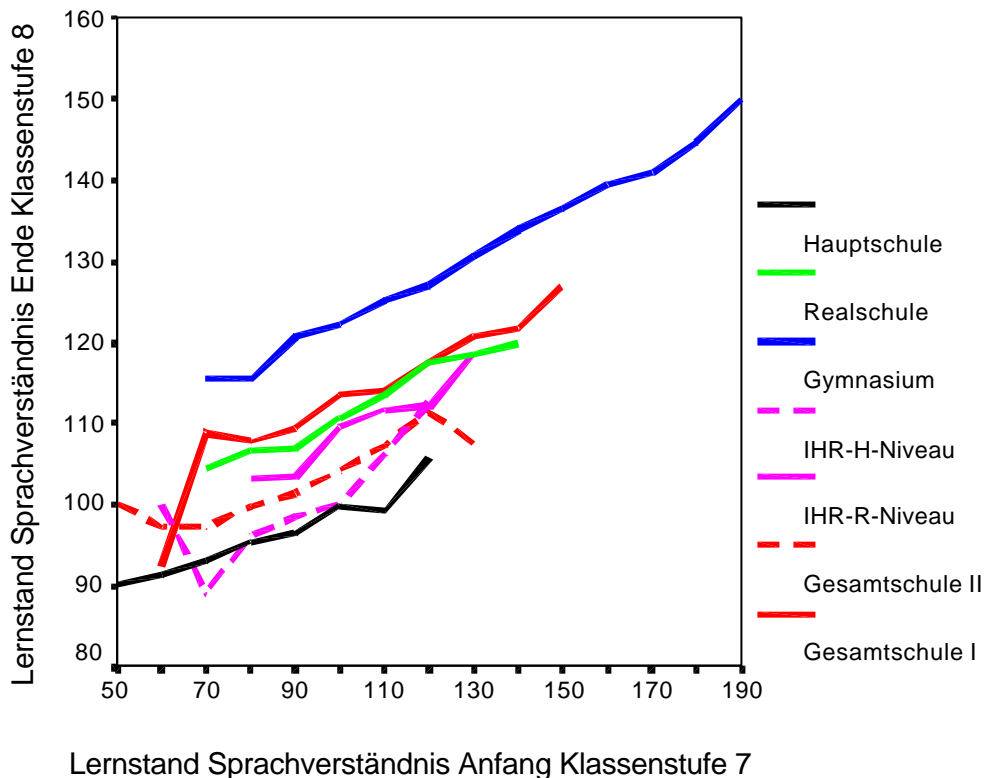
	Beginn Klassenstufe 7		Ende Klassenstufe 8		Lernzuwachs		
	$\bar{x}$	SD	$\bar{x}$	SD	Differenz	Effekt- stärke $d$	Varianz- minderung
Hauptschule	87,6	14,8	96,4	13,9	8,8	0,59	10,9
Realschule	103,8	15,6	111,8	13,0	8,0	0,51	30,5
Gymnasium	127,1	18,5	129,7	13,5	2,7	0,14	46,5
IHR Schule – H	89,8	15,6	98,7	14,6	8,9	0,57	13,5
IHR Schule – R	102,6	14,9	108,9	12,7	6,3	0,42	28,2
Gesamtschule II	92,9	16,0	102,7	14,3	9,9	0,62	20,6
Gesamtschule I	107,2	17,4	114,1	14,4	6,9	0,40	31,6
<i>insgesamt</i>	<i>109,8</i>	<i>22,9</i>	<i>115,9</i>	<i>18,6</i>	<i>6,1</i>	<i>0,26</i>	<i>34,3</i>

Bei einem Gesamtzuwachs von 6,1 Lernstandspunkten bzw. einer Effektstärke von  $d = 0,26$  (also rechnerisch 0,13 Standardabweichungen pro Jahr) ist die Leistungsentwicklung im Bereich des Sprachverständnisses nur etwa halb so stark ausgeprägt wie im Fach Mathematik und im Bereich Leseverständnis. Differenziert man zwischen den Schulformen und Kursniveaus, so zeigen sich ähnliche Relationen wie die schon zuvor beobachteten: Am höchsten sind – absolut wie relativ betrachtet – die Zuwächse in den Kursen II der Gesamtschulen, in den H-Kursen der Integrierten Haupt- und Realschulen sowie in den Hauptschulen und am geringsten in den Gymnasien. Die übrigen Schulformen bzw. Kursniveaus liegen dazwischen, wobei die Entwicklung – absolut und relativ – an den Realschulen etwas günstiger verläuft als in den Kursen I der Gesamtschulen und in den R-Kursen der Integrierten Haupt- und Realschulen. Dabei wird ein weiteres Mal deutlich, dass tendenziell die Lernzuwächse umso höher sind, je geringer die Tendenz zur Leistungshomogenisierung auf der Ebene der Schulform bzw. des Kursniveaus ausgeprägt ist (zum analogen Zusammenhang zwischen Effektstärke und Varianzminderung im Fach Mathematik vgl. Tabelle 28 und die daran angeschlossene Diskussion).<sup>41</sup> Auch bei der Differenzierung nach Schulform und Testleistung

<sup>41</sup> Zu erinnern ist in diesem Zusammenhang nochmals an die Überschätzungen, die sich aus häufiger Klassenwiederholung und häufigem Schulformwechsel ergeben, die sich also hier vor allem zu Gunsten der Realschulen auswirken. Da jedoch im Rahmen dieses Berichts ähnliche Berechnungen für den Kursniveauwechsel nicht durchgeführt werden konnten, bleibt das

zu Beginn der Klassenstufe 7 zeigen sich im Wesentlichen wieder die bereits aus den anderen Leistungsbereichen bekannten Zusammenhänge (vgl. Abbildung 24).

Abbildung 24 Fachleistung Sprache: Lernstände am Ende der Klassenstufe 8 nach Testleistung zu Beginn der Klassenstufe 7 und Schulform bzw. Kursniveau (Lernstandswerte)



Die Lernentwicklung verläuft in den Schulformen bzw. Kursniveaus ungefähr parallel, sofern man den Lernstand zu Beginn der Klassenstufe 7 berücksichtigt. Dabei ist bei gegebener Anfangsleistung der Lernstand am Ende der Klassenstufe 8 im Allgemeinen umso höher, je anspruchsvoller die besuchte Schulform oder das besuchte Kursniveau ist. Anders als im Bereich des Leseverständnisses ist im Bereich Sprache keine Konvergenz bei den oberen Leistungsgruppen zu erkennen. Unabhängig davon, welches Testergebnis zu Beginn der Klassenstufe 7 erzielt wurde, ist der Lernstand am Ende der Klassenstufe 8 an den Gymnasien durchgehend höher als in den anderen Gruppen, und innerhalb der Schulen mit äußerer Fachleistungsdifferenzierung sind bei gleicher Anfangsleistung jeweils in dem Zweig oder Kurs, der (mindestens) auf

---

genaue Ausmaß dieser Verzerrung quantitativ unbestimmt. In jedem Falle stehen die berichteten Unterschiede zwischen Schulformen und Kursniveaus unter dem Vorbehalt solcher Unsicherheiten.

den Erwerb eines mittleren Schulabschlusses zielt, die Lernstände am Ende der Klassenstufe 8 höher als auf dem unteren Differenzierungsniveau. Dass im Bereich des Sprachverständnisses der Anstieg generell niedriger ist als im Fach Mathematik oder beim Leseverständnis, wird zwar in der Abbildung 24 optisch nicht deutlich sichtbar, ergibt sich aber aus dem identischen Maßstab und aus der bereits diskutierten numerischen Darstellung in Tabelle 32. Um eine angemessene Interpretation der Befunde für die Einzelschulen zu ermöglichen, seien wieder die schulformspezifischen Kennwerte ohne Berücksichtigung der Differenzierungsniveaus angeführt (vgl. Tabelle 33).

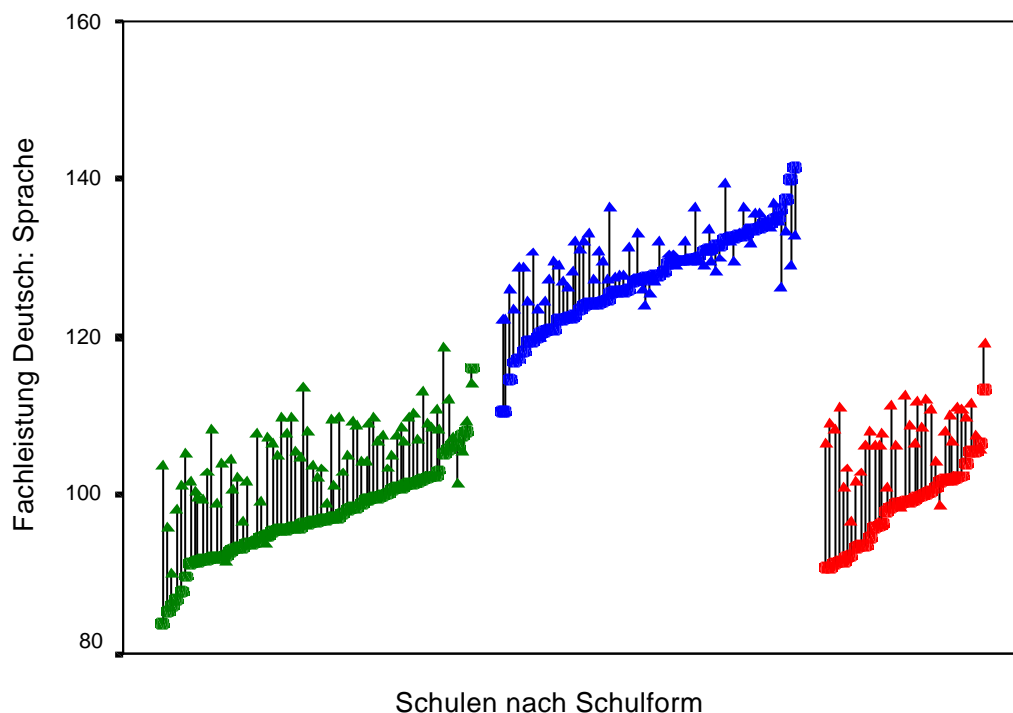
Tabelle 33 Fachleistung Sprache zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 nach Schulform und Einzelschule

	<b>Beginn</b>		<b>Ende</b>		Differenz
	<b>Klassenstufe 7</b>		<b>Klassenstufe 8</b>		
	$\bar{x}$	SD	$\bar{x}$	SD	
Haupt- und Realschule	97,3	17,1	105,5	15,2	8,2
Gymnasium	127,1	18,5	129,7	13,5	2,7
Gesamtschule	99,0	18,1	107,6	15,4	8,6
<i>insgesamt</i>	<i>109,9</i>	<i>22,9</i>	<i>115,9</i>	<i>18,6</i>	<i>6,1</i>

Anders als beim Leseverständnis haben die Gesamtschulen im Bereich Sprachverständnis eine etwas höhere durchschnittliche Lernentwicklungsrate zu verzeichnen als die Haupt- und Realschulen, während für die Gymnasien, wie bereits diskutiert, nur ein extrem geringer mittlerer Zuwachs festgestellt werden kann.

Bei der Differenzierung nach Schulformen und Schulen, geordnet nach dem durchschnittlichen Lernstand zu Beginn der Klassenstufe 7 (vgl. Abbildung 25), sind prinzipiell die gleichen Zusammenhänge zu beobachten wie in den analogen Abbildungen zur Fachleistung Mathematik und zum Leseverständnis: Die Lernausgangslagen unterscheiden sich sehr beträchtlich, und auch hinsichtlich der am Ende der Klassenstufe 8 erreichten Lernstände bestehen so große Unterschiede, dass an etlichen Schulen zu diesem Zeitpunkt die Anfangslernstände anderer Schulen derselben Schulform noch unterschritten sind. Auch stehen in vielen Fällen hohen Lernzuwächsen an Schulen mit eher niedrigen Durchschnittswerten am Anfang der Klassenstufe 7 geringere (gelegentlich auch fehlende) Zuwächse an ursprünglich sehr leistungsstarken Schulen gegenüber.

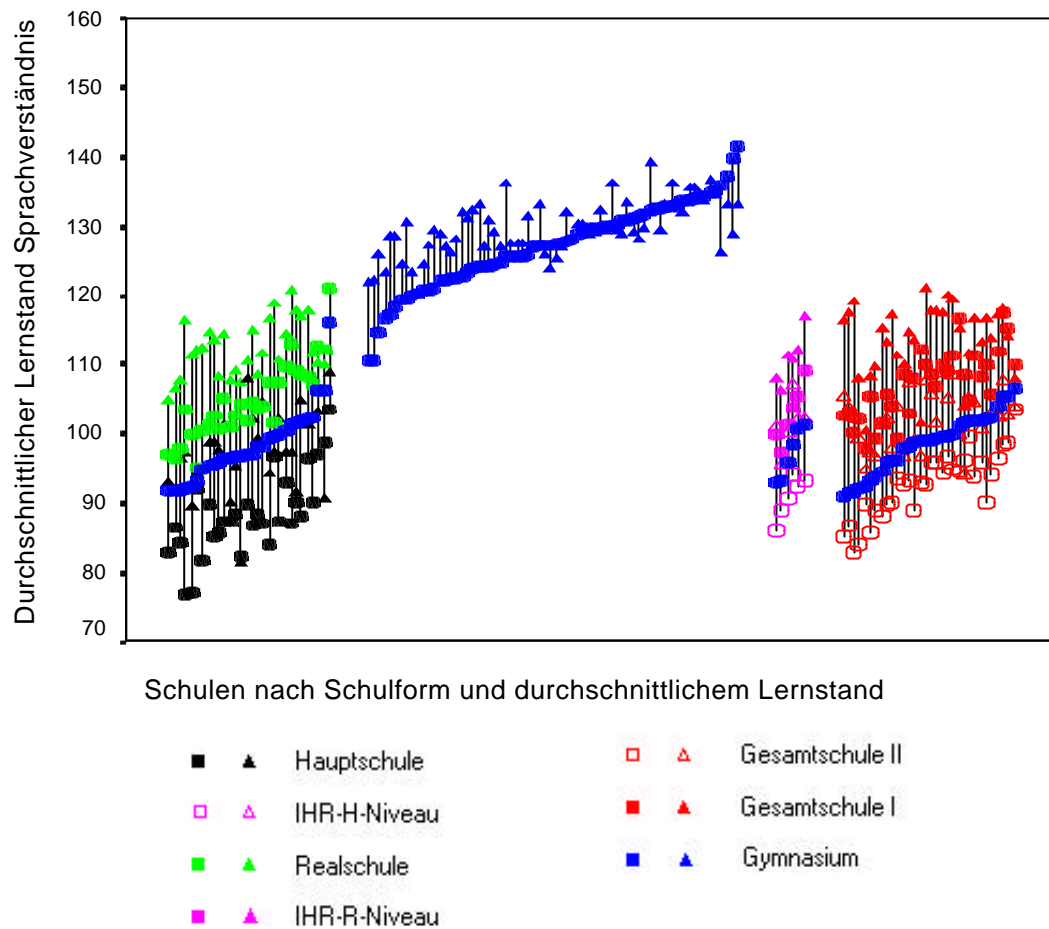
Abbildung 25 Fachleistung Sprache: Lernzuwachs in den Klassenstufen 7 und 8 nach Schulform und Einzelschule



Selbstverständlich wird man auch hier der komplexen Überlagerung von Differenzierungsmustern und niveauspezifischen Lernentwicklungen erst dann gerecht, wenn man für jede Schule gleichzeitig die Aufspaltung in unterschiedliche Lerngruppen (Schulzweige, Kursniveaus) und die jeweilige Lernentwicklung betrachtet (vgl. Abbildung 26).

Untersucht man nämlich auch für den Bereich Sprachverständnis bei den Schulen mit äußerer Fachleistungsdifferenzierung die Zusammenhänge zwischen Differenzierungsgrad (gemessen über die Differenz der mittleren Anfangsleistungen auf dem oberen bzw. unteren Differenzierungsniveau) und der Lernentwicklung, so sind bei einer stärker fachleistungsbezogenen Differenzierung an den Haupt- und Realschulen keine niedrigeren Lernstände in den Hauptschulklassen bzw. H-Kursen nachweisbar als dort, wo sich die beiden Differenzierungsniveaus anfangs relativ wenig unterscheiden ( $r = -0,10$  n.s.). Wohl aber sind in den Realschulklassen bzw. -kursen bei konsequent leistungsbezogener Bildung der Lerngruppen am Ende der Klassenstufe 8 zumindest tendenziell höhere Lernstände erreicht ( $r = 0,31$  n.s.).

Abbildung 26 Fachleistung Sprache: Lernzuwachs in den Klassenstufen 7 und 8 nach Schulform bzw. Kursniveau und Schulen



Dabei spielt der jeweilige Anteil der dem oberen Differenzierungsniveau zugewiesenen Schülerinnen und Schüler, also die Übergangsquote in die Realschulklassen bzw. -kurse, nur insofern eine Rolle, als bei statistischer Berücksichtigung dieses Einflussfaktors die leistungsmäßige Überlegenheit hoch ausgelesener Realschulklassen bzw. -kurse am Ende der Klassenstufe 8 noch deutlicher wird (partielles  $r = 0,44$ ). Der Grund hierfür liegt darin, dass der Realschulanteil zwar generell an leistungsstarken Schulen geringfügig höher ist ( $r = 0,24$  n.s.), dass sich dabei aber die möglicherweise günstigen Einflüsse der Bildung homogener Lerngruppen auf dem oberen Differenzierungsniveau nur dann zu entfalten vermögen, wenn die Lerngruppenbildung konsequent an Fachleistungsunterschiede gebunden ist (was gerade an den schon zu Anfang des Betrachtungszeitraums leistungsstarken Schulen der Fall ist:  $r = 0,50$ ).

Ähnlich sieht es in diesem Leistungsbereich an den Gesamtschulen aus. Während eine konsequent an der Fachleistung orientierte Praxis der Kurszuweisung allenfalls (aber zufallskritisch nicht gesichert) den Kursen I zugute kommt ( $r = 0,25$  n.s.), sind – wie auch bei anderen Leistungsaspekten für die

Kurse II der Gesamtschulen keine negativen Auswirkungen solchen Vorgehens auf die Lernentwicklung zu erkennen ( $r = -0,04$  n.s.). Auch hier ist der Anteil der den Kursen I zugewiesenen Schülerinnen und Schüler unerheblich. Anders jedoch als an den Haupt- und Realschulen verstärkt die Berücksichtigung dieses Merkmals nicht den Zusammenhang zwischen konsequent fachbezogener Differenzierung und Leistungsniveau auf dem Kursniveau I, weil zwar der relative Anteil der Kurse I an leistungsstarken Gesamtschulen etwas höher, dafür aber dieser Anteil praktisch unabhängig vom Grad der Bindung der Kurszuweisung an die Fachleistung ist ( $r = -0,14$  n.s.).

An den Gymnasien zeigt sich neben dem Sachverhalt, dass an den anfangs leistungsstärkeren Schulen auch die höheren Lernstände am Ende der Klassenstufe 8 erreicht worden sind ( $r = 0,62$ ), wieder die auffällige Erscheinung, dass in den Schulen mit ursprünglich niedrigeren Lernstandswerten höhere Zuwächse zu verzeichnen sind ( $r = 0,79$ ). Ein statistisch bedeutsamer Zusammenhang zwischen durchschnittlicher Lernentwicklung und der Tendenz, die Lernstände an der Schule zu homogenisieren, liegt dabei nicht vor ( $r = 0,11$  n.s.), doch diese ist an den Gymnasien mit anfangs niedrigerem Leistungsdurchschnitt geringer ausgeprägt ( $r = -0,27$ ). Generell gleichen die Befunde für die Gymnasien also weitgehend jenen für die bereits untersuchten Leistungsaspekte. Nur ist das Fehlen substantieller Zuwächse an leistungsstarken Gymnasien hier noch markanter als in den bisher schon untersuchten Testergebnissen.

#### **4.4 Entwicklung der Fachleistung im Bereich der Fremdsprachen**

Im Folgenden soll aufgezeigt werden, wie sich die Lernentwicklungen im Bereich der ersten Fremdsprache (Englisch bzw. Latein) vom Beginn der Klassenstufe 7 bis zum Ende der Klassenstufe 8 darstellen.

- **Englisch**

Für die Schulleistung im Fach Englisch konnten im Bericht zu den Befunden aus der Erhebung des Jahres 1998 (Klassenstufe 7) noch keine Lernzuwächse berechnet werden, da zu Beginn der Klassenstufe 5 der *KS-HAM 4/5* eingesetzt wurde, der dieses Fach noch nicht berücksichtigen konnte. Nun aber liegen Testergebnisse für C-Tests aus den Jahren 1998 und 2000 (Beginn Klassenstufe 7 bzw. Ende Klassenstufe 8) vor, die miteinander so verzahnt sind, dass mit einer gemeinsamen Skala („Metrik“) für die beiden Erhebungszeitpunkte gearbeitet werden kann. Als Parameter der Skala wurde ein

Mittelwert von 100 und eine Standardabweichung von 25 für den Beginn der Klassenstufe 7 gewählt.<sup>42</sup> Die so berechneten Lernstandswerte und weitere Kenngrößen sind in Tabelle 34 zusammengefasst.

Tabelle 34 Fachleistung Englisch (C-Test): Lernstandswerte zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 (Mittelwert, Standardabweichung) sowie Lernzuwachs (Mittelwert, Effektstärke und Varianzminderung), nach Schulform bzw. Kursniveau

	Beginn Klassenstufe 7		Ende Klassenstufe 8		Lernzuwachs		
	$\bar{x}$	SD	$\bar{x}$	SD	Differenz	Effekt- stärke $d$	Varianz- minderung
Hauptschule	75,6	14,9	103,2	15,6	27,6	1,85	-10,6 %
Realschule	96,2	15,2	128,5	16,6	32,3	2,12	-18,7 %
Gymnasium	123,7	16,8	153,4	14,4	29,7	1,77	26,8 %
IHR Schule – H	77,6	15,0	108,0	15,3	30,4	2,03	-3,8 %
IHR Schule – R	92,0	16,8	125,8	14,9	33,8	2,01	21,7 %
Gesamtschule II	82,1	14,1	109,8	16,0	27,7	1,96	-27,6 %
Gesamtschule I	99,7	16,4	133,9	17,0	34,2	2,08	-6,8 %
<i>insgesamt</i>	<i>103,2</i>	<i>24,3</i>	<i>133,5</i>	<i>24,4</i>	<i>30,4</i>	<i>1,24</i>	<i>-0,5 %</i>

Aus diesen Zahlen wird ersichtlich, dass die Lernentwicklung im Fach Englisch gänzlich anders verlaufen ist als in den Fächern Deutsch und Mathematik. Der durchschnittliche Lernfortschritt in den Klassenstufen 7 und 8 ist mit 1,24 Standardabweichung bzw. 0,62 Standardabweichungen pro Schuljahr ausgesprochen hoch. Der eigentliche Unterschied zu diesen Fächern liegt aber darin, dass im Englischen die anfängliche Leistungsstreuung im Gesamtsystem erhalten geblieben ist; deren geringfügige Erhöhung von Klassenstufe 7 nach 9 ist unbedeutend. Wenig verändert haben sich auch die Unterschiede zwischen den Schulformen: Lagen (um die Extreme zu benennen) die Mittelwerte für das

<sup>42</sup> Dass in der nachstehenden Tabelle der Mittelwert und die Standardabweichung zu Beginn der Klassenstufe 7 nicht insgesamt 100 und 25 Punkte betragen, sondern davon etwas abweichende Werte angegeben sind, geht darauf zurück, dass durch die bereits angesprochenen, auf Grund der Datenlage unvermeidlichen Unklarheiten in der Zuordnung zu Schulformen und Kursniveaus mit dem Stand am Ende der Klassenstufe 8 vor allem eher leistungsschwache Schülerinnen und Schüler von den vorliegenden Betrachtungen ausgeschlossen werden mussten, was zu einer Überschätzung der durchschnittlichen Englischleistung um etwa 0,12 Standardabweichungen und zu einer Reduktion der Varianz um ca. 5 Prozent geführt hat.

Gymnasium und die Hauptschule zu Beginn der Klassenstufe 7 um 1,98 Standardabweichungen auseinander, so betrug dieser Abstand am Ende der Klassenstufe 8  $d = 2,06$ .<sup>43</sup>

Untersucht man die Lernfortschritte differenziert nach Schulform bzw. Kursniveau, so gibt es weder absolut noch relativ betrachtet Unterschiede, die ähnlich stark ausgeprägt wären wie in den Fächern Deutsch und Mathematik. Besonders hohe Zuwächse sind in den Kursen I der Gesamtschulen, in den R-Kursen der Integrierten Haupt- und Realschulen und an den Realschulen zu verzeichnen, ohne dass die H-Kurse der letztgenannten Schulform weit dahinter zurückbleiben. Es folgen (absolut betrachtet) die Gymnasien, die Kurse II der Gesamtschulen und schließlich die Hauptschulen. Dass an den Gymnasien wiederum eine besonders niedrige *relative* Zunahme der Lernstandswerte stattgefunden hat, steht zumindest formal, aber möglicherweise auch inhaltlich im Zusammenhang mit der vergleichsweise hohen Anfangsstreuung der Englischleistungen an dieser Schulform.

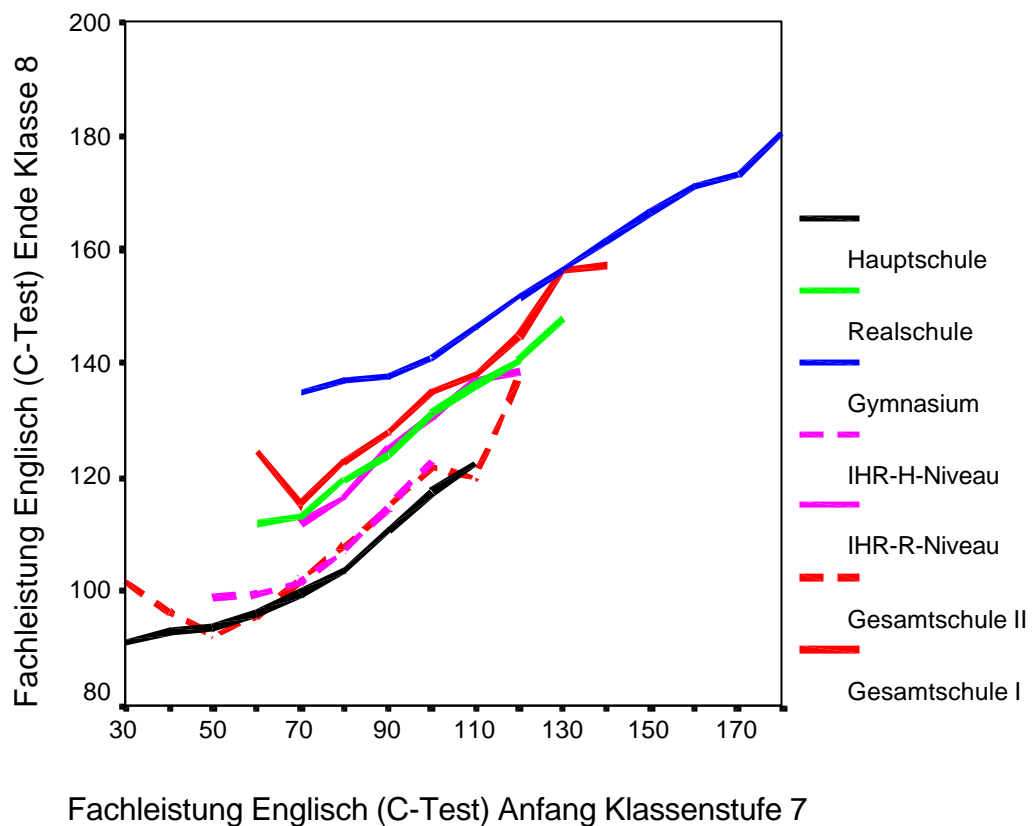
Auffällig ist der Umstand, dass sich – mit Ausnahme der Gymnasien und der R-Kurse der Integrierten Haupt- und Realschulen – innerhalb der übrigen Schulformen bzw. Kursniveaus keine Reduktion, sondern eine Erhöhung der Streuung zeigt. Im Fach Englisch sind demnach kompensatorische Prinzipien weniger maßgebend als in den beiden anderen untersuchten Fächern. Wenn die bisherigen Erkenntnisse zutreffen, nach denen besonders starke Lernzuwächse im unteren Bereich des Leistungsspektrums in der Regel mit niedrigeren allgemeinen Zuwachsraten einhergehen, dann liegt ein Zusammenhang zwischen der Varianzerhaltung bzw. -vermehrung im Fach Englisch und der dort durchgehend zu beobachtenden günstigen Lernerentwicklung nahe. Angesichts dieser speziellen Situation ist die Aufschlüsselung der Lernstände am Ende der Klassenstufe 8 nach Schulformen und Lernstand am Anfang der Klassenstufe 7 aufschlussreich (vgl. Abbildung 27).

---

<sup>43</sup> Für die Verzahnung mit der LAU 7 stand eine Anker-Textpassage mit 24 Items zur Verfügung. Die Wachstumssensitivität, d.h. die Verminderung der relativen Itemschwierigkeiten zwischen den beiden Messzeitpunkten, wurde überprüft. Diese ist für alle Items ebenso gegeben wie für eine aus den Ankeritems gebildete Summenskala, für die insgesamt ein Zuwachs über die Klassenstufen 7 und 8 von  $d = 0,86$  sowie schulformspezifische Zuwächse von  $d = 0,93$  (Gymnasium) bis  $d = 1,83$  (Hauptschule) ermittelt wurden.



Abbildung 27 Fachleistung Englisch (C-Test): Lernstände am Ende der Klassenstufe 8 nach Testleistung zu Beginn der Klassenstufe 7 und Schulform bzw. Kursniveau (Lernstandswerte)



Hier zeigt sich, dass die in allen Schulformen bzw. Kursniveaus, insbesondere aber in den Kursen I der Gesamtschulen und in den Realschulen festzustellenden hohen Lernfortschritte im Fach Englisch einhergehen mit einer deutlichen Förderung auch der leistungsstärkeren Gruppen der jeweiligen Schülerschaft. Unter anderem wird die Förderung leistungsstarker Schülerinnen und Schüler an der Konvergenz der Leistungsentwicklung zwischen den Schulformen bzw. Kursniveaus gerade im oberen Segment der Schülerschaften der Gesamtschulen und Realschulen sichtbar. Andererseits ist zu berücksichtigen, dass die Gymnasien auch im Fach Englisch – wie schon in den Fächern Deutsch und Mathematik – in diesem für sie mittleren, für die Gesamtschulen sowie die Haupt- und Realschulen aber oberen Leistungsbereich, unerwartet geringe Zuwächse aufweisen.

Für eine differenziertere Betrachtung der Lernentwicklung auf den Differenzierungsniveaus einzelner Schulen ist es nützlich, die Lernausgangslagen für die Schülerschaften, die aus den Beobachtungsstufen der Haupt- und Realschulen und der Gymnasien sowie aus der Klassenstufe 6 der Gesamtschulen hervorgegangen sind, tabellarisch zusammenzustellen (Tabelle 35).

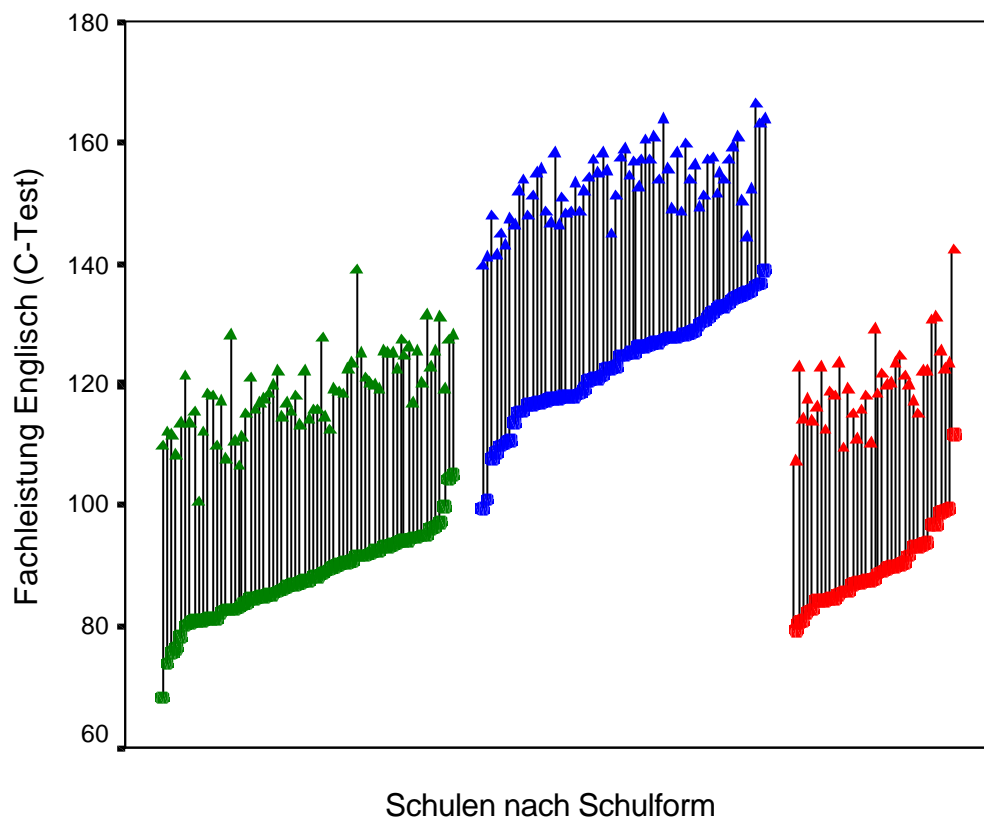
Tabelle 35 Fachleistung Englisch (C-Test) zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 nach Schulform und Einzelschule

	Beginn Klassenstufe 7		Ende Klassenstufe 8		Differenz
	$\bar{x}$	SD	$\bar{x}$	SD	
Haupt- und Realschule	88,2	18,0	119,0	19,0	30,9
Gymnasium	123,7	16,8	153,4	14,4	29,7
Gesamtschule	90,1	17,5	120,9	20,4	30,7
<i>Insgesamt</i>	<i>103,1</i>	<i>24,3</i>	<i>133,5</i>	<i>24,4</i>	<i>30,4</i>

Wie hier ersichtlich wird, sind die Lernzuwächse in allen drei Schulformen, insgesamt betrachtet, fast gleich. Wegen der unterschiedlichen Anfangsverteilungen steht dies nicht im Widerspruch zu den hohen individuellen Förderungserfolgen in den Gymnasien, belegt aber ein weiteres Mal die vergleichsweise geringen Lernzuwächse unter den Schülerinnen und Schülern mit relativ hohen Lernständen schon zu Beginn der Klassenstufe 7.

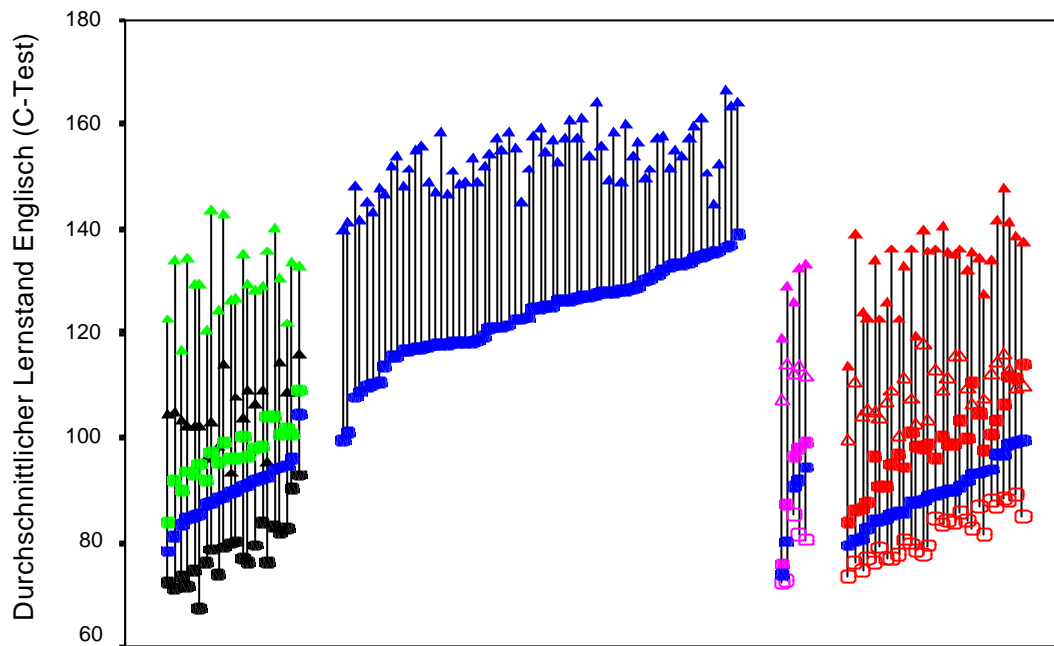
Auf der Ebene der Einzelschulen sind im Bereich Englisch deutlich gleichmäßigere Lernzuwächse zu verzeichnen als bei den anderen Untertests des *SL-HAM 8/9* (vgl. Abbildung 28). Zwar gibt es auch hier Beispiele dafür, dass an einer Schule, gemessen an der Lernausgangslage, eine unerwartet günstige oder auch eine unerwartet ungünstige Lernentwicklung stattgefunden hat. Insgesamt jedoch sind diese Fälle eher selten. Immerhin ist es bemerkenswert, wenn es an einer Haupt- und Realschule gelungen ist, ausgehend von einer für diese Schulform keineswegs untypischen Lernausgangslage *im Durchschnitt* den Mittelwert eines Gymnasiums zu erreichen. Im Gegensatz dazu lagen an der Gesamtschule, die ebenfalls im Mittel gymnasiales Niveau erreicht hat, von Anfang an besonders günstige und insoweit einigen Gymnasien vergleichbare Anfangsbedingungen vor. Umgekehrt ist z. B. an einem Gymnasium, das nach der Lernausgangslage zu den leistungsstärksten überhaupt gezählt hatte, für das Ende der Klassenstufe eines der niedrigsten Ergebnisse in dieser Schulform zu verzeichnen. Zu ermitteln, warum es zu solchen atypischen Entwicklungen gekommen ist, ist auf der Grundlage der vorliegenden Daten nicht zu leisten.

Abbildung 28 Fachleistung Englisch (C-Test): Lernentwicklung vom Beginn der Klassenstufe 7 bis zum Ende der Klassenstufe 8 nach Schulform und Einzelschule



Mit einer grafischen Darstellung der Lernzuwächse, geordnet nach Schulformen und durchschnittlichem Lernstand an den Einzelschulen zu Beginn der Klassenstufe 7, nun aber unter Einschluss der Differenzierungsbedingungen, rundet sich das Bild (vgl. Abbildung 29). Auch anhand dieser Darstellung ist leicht zu erkennen, dass in allen Schulformen bzw. Kursniveaus im Fach Englisch beträchtliche Lernzuwächse erzielt worden sind. Insbesondere haben die Hauptschulen, die H-Kurse an den Integrierten Haupt- und Realschulen und die Kurse II an den Gesamtschulen bis zum Ende der Klassenstufe 8 in der Regel den Anfangslernstand der Klassen bzw. Kurse des jeweils höheren Niveaus deutlich übertroffen, was so nicht für alle untersuchten Leistungsaspekte gilt. Weiterhin sind zwischen den Schulen derselben Schulform die erreichten Leistungsstände offenbar stärker abhängig von den Anfangslernständen, was quantitativ überprüft werden kann. Und schließlich bleibt auch hier zu fragen, in welchem Verhältnis die Klassen- bzw. Kurszusammensetzung am Beginn der Klassenstufe 7 zu der jeweiligen Lernentwicklung steht.

Abbildung 29 Fachleistung Englisch (C-Test): Lernzuwachs in den Klassenstufen 7 und 8 nach Schulform bzw. Kursniveau und Schulen



Schulen nach Schulform und durchschnittlichem Lernstand

- |     |              |     |                 |
|-----|--------------|-----|-----------------|
| ■ ▲ | Hauptschule  | □ ▲ | Gesamtschule II |
| □ ▲ | IHR-H-Niveau | ■ ▲ | Gesamtschule I  |
| ■ ▲ | Realschule   | ■ ▲ | Gymnasium       |
| ■ ▲ | IHR-R-Niveau |     |                 |

Generell hängen an den Haupt- und Realschulen Anfangs- und Endleistung im Fach Englisch relativ eng miteinander zusammen ( $r = 0,66$ ). Hier ergeben sich bei einer konsequent nach Fachleistung erfolgten Differenzierung der Lerngruppen nur in den Realschulklassen bzw. -kursen höhere Lernstände am Ende der Klassenstufe 8 ( $r = 0,46$ ), nicht aber in den Hauptschulklassen bzw. -kursen ( $r = -0,31$  n.s.). Dies entspricht weitgehend den bereits für die Fächer Deutsch und Mathematik herausgearbeiteten Befunden, und die Zusammenhänge sind bei statistischer Berücksichtigung der durchschnittlichen Anfangsleistung auf Schulebene sowie des Anteils der Schülerinnen und Schüler im Realschulzweig oder -kurs noch ausgeprägter (partielles  $r = 0,64$  auf Realschulniveau; partielles  $r = -0,39$  n.s. auf Hauptschulniveau). Auf dem Hauptschulniveau geht demnach die Abwesenheit der eher leistungsstarken Schülerinnen und Schüler anscheinend mit vergleichsweise geringeren Englischleistungen am Ende des Betrachtungszeitraums einher, auch wenn, wie beschrieben, immer noch beachtliche Lernfortschritte erzielt worden sind.<sup>44</sup>

<sup>44</sup> Die knapp verfehltete zufallskritische Absicherung ( $p = 0,05$ ) erscheint unbedenklich angesichts

Auch an den Gesamtschulen ist die Lernentwicklung in den Kursen des oberen Anforderungsniveaus bei einer konsequent fachbezogenen Leistungs-differenzierung besonders günstig ( $r = 0,39$ ), ohne dass hier leistungsmäßig negative Folgen für die Kurse II zu beobachten sind ( $r = 0,07$  n.s.). Dies entspricht ebenfalls den Befunden, die sich aus der Analyse anderer Leistungsaspekte ergeben hatten. Allerdings zeigen sich in der Tendenz negative Auswirkungen für die Kurse II dann, wenn man auf Schulebene das Anfangsleistungsniveau und die Übergangsquote in die Kurse I statistisch berücksichtigt. Weil die Fachleistungen Ende der Klassenstufe 8 sehr eng mit der Anfangsleistung zusammenhängen ( $r = 0,61$ ), deutet sich bei deren Einbeziehung ein negativer Einfluss konsequent fachleistungsbezogener Auslese auf die Leistungen im unteren Kursniveau an (partiell  $r = -0,38$  n.s.). Ein sog. „Creaming effect“ (Absahneffekt) lässt sich also auch für die Kurse II der Gesamtschulen beobachten, beschränkt sich hier aber auf Schulen mit einem insgesamt ungünstigen Leistungsstand.

An den Gymnasien korrelieren Anfangs- und Endleistung auf Schulebene im Fach Englisch mit  $r = 0,65$  und damit etwas enger als im Fach Deutsch (Leseverständnis:  $r = 0,61$ ; Sprachverständnis:  $r = 0,62$ ), jedoch weniger stark als im Fach Mathematik ( $r = 0,70$ ). Dabei ist die Lernentwicklung insgesamt – ähnlich, wie dies bereits in den Fächern Deutsch und Mathematik der Fall war – auch im Englischen an Schulen mit ursprünglich eher niedrigen Lernständen offenbar günstiger verlaufen als an solchen, die bereits zu Anfang des Untersuchungszeitraums hohe Lernstände zu verzeichnen hatten ( $r = -0,74$ ). So bestätigen sich für das Fach Englisch bei der differenziellen Analyse auf Schulebene im Wesentlichen die Zusammenhänge, die auch in den anderen Fächern sichtbar waren.

### • Latein

Die Lernentwicklung im Fach Latein lässt sich nur auf schmaler Datenbasis – einbezogen sind elf Hamburger Schulklassen, in denen Latein als erste Fremdsprache unterrichtet wurde – untersuchen. Auch sind die in den Klassenstufen 7 und 9 eingesetzten Lateintests nicht miteinander verzahnt, so dass sich keine übergreifende Metrik bilden ließ. Um trotzdem etwas über die relativen Lernfortschritte in diesem Fach aussagen zu können, wurden die Testrohwerte aus dem Untertest *Latein* des *SL-HAM 6/7* (kombiniert aus Sprachkenntnissen und Sachwissen „Römisches Leben“) analog zum Fach Englisch so skaliert, dass sich ein Mittelwert von 100 und eine Standardabweichung von 25 ergab.

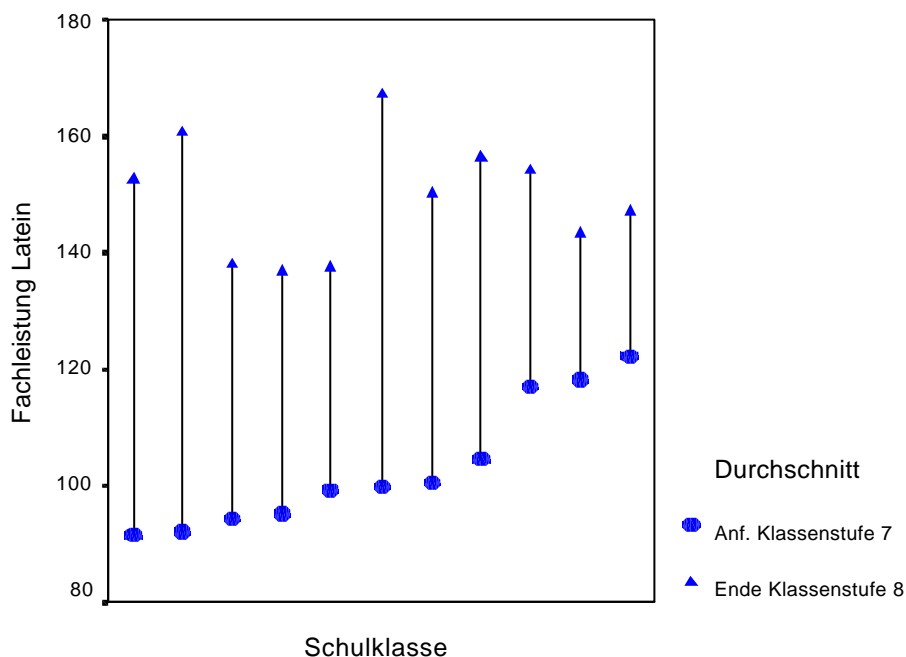
---

der Parallelität der Befunde in den anderen untersuchten Fächern.

Anders als im Fach Englisch handelt es sich dabei nur um eine lineare Transformation der Testrohwerte, nicht um das Resultat einer probabilistischen Testanalyse. Der zu Beginn der Jahrgangsstufe 9 durchgeführte Lateintest wurde ebenso verrechnet, nur dass hier ein Mittelwert von 150 gewählt wurde. Die Differenz zwischen diesen beiden Mittelwerten ist eine rein pragmatische Setzung. Sie entspricht in der Größenordnung den empirisch ermittelten Lernzuwächsen für die Klassenstufen 7 und 8 im Fach Englisch. Schulklassenvergleiche und Korrelationen bleiben von diesen Festsetzungen unberührt.

Auffällig ist, dass bereits auf Schülerebene die Lernstände zu Beginn der Klassenstufe 7 nur in geringem Umfang mit den Lernständen am Ende der Klassenstufe 8 zusammenhängen ( $r = 0,26$ ). Eine denkbare (Teil-) Erklärung hierfür liegt darin, dass in den verschiedenen Klassen sehr verschiedene durchschnittliche Lerngewinne erzielt wurden, und zwar ihrerseits gänzlich unabhängig vom Anfangslernstand ( $r = -0,03$  n.s.; vgl. Abbildung 30).<sup>45</sup>

Abbildung 30 Fachleistung Latein: Lernzuwachs in den Klassenstufen 7 und 8



Es ist in den ursprünglich leistungsstarken Klassen keineswegs durchgehend gelungen, diesen Vorsprung zu halten, während umgekehrt manche anfangs eher leistungsschwache Schulklassen merklich aufgeholt und am Ende der Klassenstufe 8 sogar Spitzenwerte erzielt haben. Es liegt nahe, diese Erschei-

<sup>45</sup> Wie in den anderen untersuchten Bereichen bleiben auch hier methodische Probleme der Lernzuwachsmessung ungelöst bestehen. Die Darstellung beruht auf den beobachteten, auf Schulklassenebene aggregierten Testergebnissen ohne Adjustierung.

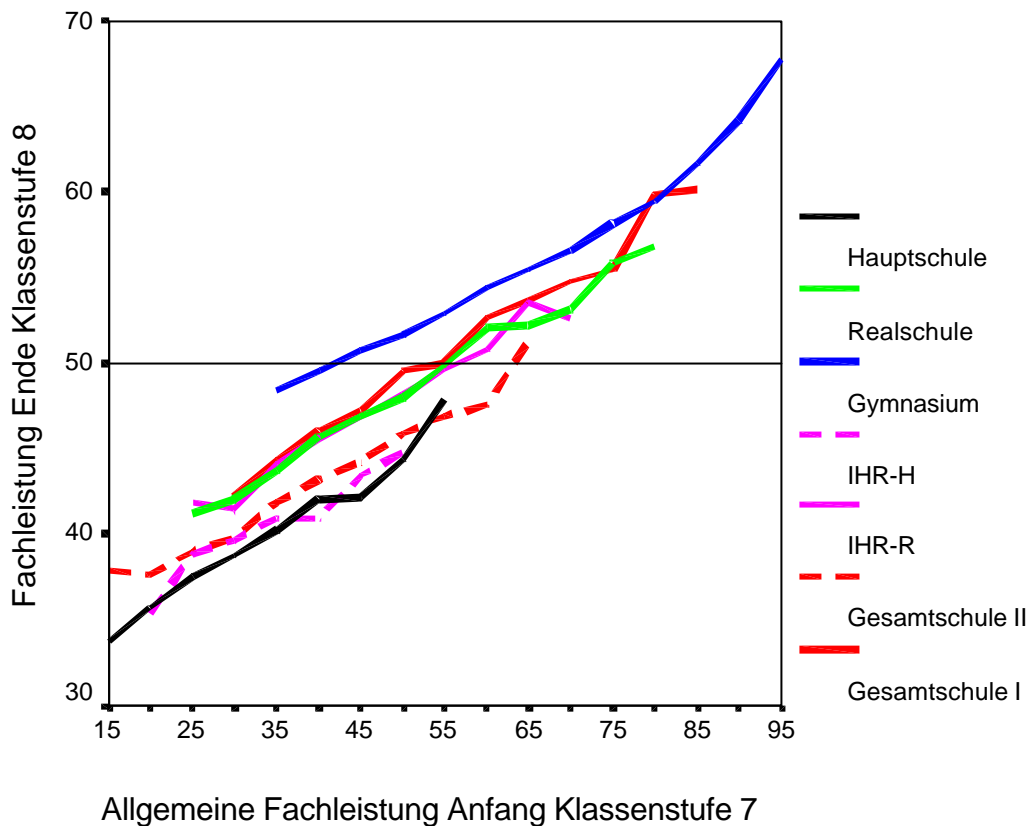
nung im Zusammenhang mit einer unterschiedlichen Ausrichtung oder Effizienz des Unterrichts zu sehen, doch es fehlen empirische Daten zu dessen Anlage und Durchführung.

## 4.5 Zusammenfassung

Die Befunde zur Lernentwicklung in den Klassenstufen 7 und 8 werden hier so zusammengefasst, dass zunächst der Rückbezug zu dem in Kapitel 3 gebildeten Index für die Allgemeine Fachleistung am Ende der Klassenstufe 8 und zu dem entsprechenden Index aus der Erhebung zu Beginn der Klassenstufe 7 hergestellt wird. Beide Indizes beruhen zu gleichen Teilen auf den Leistungen in den Fächern Mathematik, Deutsch und Englisch. Zwar liegen sie nicht auf derselben Skala, so dass Aussagen über allgemeine Lernfortschritte nicht gemacht werden können; sie erlauben jedoch allgemeine Einblicke in die relative Lernentwicklung von Schülerinnen und Schülern verschiedener Schulformen bzw. Kursniveaus und unterschiedliche Leistungsstände zu Beginn der Klassenstufe 7.

Der Fachleistungsindex für den Beginn der Klassenstufe 7 hat einen schulformübergreifenden Mittelwert von 53,3 Punkten und eine Standardabweichung von 18,8 Punkten. Um den Fachleistungsindex für das Ende der Klassenstufe 8 dazu in Beziehung setzen zu können, wurde hier pragmatisch ein Mittelwert von 50 Punkten und eine Standardabweichung von 10 Punkten festgesetzt. Unbedingt zu beachten ist, dass es sich hier um *zwei verschiedene Skalen* handelt, deren Werte nicht unmittelbar zueinander in Beziehung gesetzt werden dürfen. Die allgemeine Lernentwicklung lässt sich auf diese Weise nicht absolut, sondern nur relativ darstellen, dies jedoch auch in grafischer Form (vgl. Abbildung 31). Der Landesdurchschnitt für die allgemeine Fachleistung am Ende der Klassenstufe 8 ist durch eine waagerechte Linie gekennzeichnet. Unterhalb dieser Schwelle befanden sich zu diesem Zeitpunkt 7,9 Prozent der Gymnasiastinnen und Gymnasiasten, während zu Beginn der Klassenstufe 7 nur 6,5 Prozent unter dem Landesdurchschnitt lagen.

Abbildung 31 Allgemeine Fachleistung: Lernstände am Ende der Klassenstufe 8, nach Testleistung zu Beginn der Klassenstufe 7 und Schulform bzw. Kursniveau



Auch in den anderen Schulformen hat es leichte Verschiebungen zwischen den beiden Erhebungszeitpunkten gegeben. Während an den Hauptschulen der Anteil *über* dem Landesdurchschnitt mit 2,9 Prozent konstant geblieben ist, hat sich dieser Anteil der überwiegend in H-Kursen Unterrichteten an den Integrierten Haupt- und Realschulen von 4,2 auf 3,0 Prozent und von den überwiegend in Kursen II an den Gesamtschulen Unterrichteten von 8,5 auf 7,1 Prozent vermindert. Die Schülerinnen und Schüler an den Integrierten Haupt- und Realschulen, die überwiegend R-Kurse besucht haben, lagen zu Beginn der Klassenstufe 7 zu 31,8 Prozent und am Ende der Klassenstufe 8 zu 29,9 Prozent über dem Landesdurchschnitt. Erhöhungen dieser Quote sind an den Realschulen (von 36,4 auf 37,3 Prozent) und in den Kursen I der Gesamtschulen (von 47,9 auf 49,0 Prozent) festzustellen.

Für die integrierten Schulformen werden im Folgenden je nach Kurseinstufung die Anteile der Schülerschaft aufgeführt, deren allgemeine Fachleistung zu Beginn der Klassenstufe 7 und am Ende der Klassenstufe 8 oberhalb des jeweiligen Landesdurchschnitts lag (vgl. Tabelle 36).



Tabelle 36 Anteil der Schülerschaft von Schulen mit fachbezogener äußerer Leistungsdifferenzierung oberhalb des Landesdurchschnitts, nach Schulform und Differenzierungsniveau (in Prozent)

Differenzierungsniveau nach dem Stand Ende Klassenstufe 8	Anfang Klassenstufe 7	<i>N</i>	Ende Klassenstufe 8	<i>N</i>
<i>Integrierte Haupt- und Realschule</i>	19,0	510	16,8	618
kein R-Kurs	3,9	203	3,1	262
zwei H-Kurse, ein R-Kurs	6,1	33	2,6	38
ein H-Kurs, zwei R-Kurse	19,1	21	7,7	26
drei R-Kurse	32,8	253	31,9	292
<i>Gesamtschule</i>	25,4	2.617	24,3	3.100
kein Kurs I	4,3	1.013	3,4	1.254
zwei Kurse II, ein Kurs I	17,3	485	15,2	573
ein Kurs II, zwei Kurse I	27,2	397	29,5	454
drei Kurse I	59,3	722	59,8	819

Diese Befunde können insofern gut miteinander verglichen werden, als die Datenausfälle für die Integrierten Haupt- und Realschulen und die Gesamtschulen nahezu gleich sind. Auch ergibt sich kein nennenswert anderes Bild, wenn nur diejenigen berücksichtigt werden, die an beiden Erhebungen teilgenommen haben. Es wird sichtbar, dass sich die nach Übergangsquoten und Fachleistung stärker ausgelesenen Kurse I an den Gesamtschulen an anderen Zielsetzungen orientieren als die R-Kurse der Integrierten Haupt- und Realschulen. Die Leistungsverteilung der Schülerinnen und Schüler an Gesamtschulen mit zwei Kursen I entspricht annähernd derjenigen der Schülerinnen und Schüler an Integrierten Haupt- und Realschulen, die in allen drei Differenzierungsfächern R-Kurse besucht haben. Diejenigen Gesamtschülerinnen und Gesamtschüler hingegen, die in allen drei Fächern auf dem Kursniveau I unterrichtet worden sind, zeigen am Ende der Klassenstufe 8 mit 59,8 Prozent oberhalb des Landesdurchschnitts eine günstigere Leistungsverteilung als die Hamburger Schülerschaft insgesamt (50,8 Prozent), sind aber deutlich entfernt von dem gymnasialen Anteil oberhalb des Landesdurchschnitts (92,1 Prozent). Im Übrigen hat sich die relative Position der verschiedenen Schulformen bzw. Kursniveaus in den Klassenstufen 7 und 8 nur wenig verschoben. Außerdem wird aus der Abbildung 31 ersichtlich, was sich auch aus der Analyse der Ergebnisse der verschiedenen Untertests ergab: Hinsichtlich der Lernfortschritte in den Klassenstufen 7 und 8 verläuft die Entwicklung bei Berücksichtigung der Anfangsleistung zwischen den Schulformen etwa parallel. Am Gymnasium ist die *individuelle* Lernentwicklung (fast) über das gesamte Leistungsspektrum am günstigsten. In einem mittleren Bereich befinden sich

die Schülerinnen und Schüler der Kurse I an den Gesamtschulen, der Realschulen und der R-Kurse an den Integrierten Haupt- und Realschulen. Dabei macht der Anteil der Schülerinnen und Schüler der Kurse I an Gesamtschulen, die bei gleicher Lernausgangslage am Ende der Klassenstufe 8 einen dem Gymnasium entsprechenden Lernstand erreicht haben (d. h., die Gruppe derjenigen, die zu Beginn der Klassenstufe 7 einen Fachleistungsindex  $\geq 80$  erzielt hatten), 2,9 Prozent der Schülerschaft dieses Kursniveaus aus. Die geringsten individuellen Zuwächse sind in den Kursen II der Gesamtschulen, den H-Kursen der Integrierten Haupt- und Realschulen und schließlich in den Hauptschulen zu verzeichnen.

Obwohl an den Gymnasien in den Klassenstufen 7 und 8 im Durchschnitt deutlich geringere allgemeine Lernfortschritte als in den anderen Schulformen erzielt wurden und insbesondere die ursprünglich sehr leistungsstarken Gymnasien ihre Spitzenposition nicht immer zu halten, in keinem Falle aber deutlich auszubauen vermochten, erreichten auch in der Perspektive der allgemeinen Fachleistung Schülerinnen und Schüler mit einem eher unterdurchschnittlichen Ausgangslernstand an dieser Schulform die vergleichsweise höchsten Lernzuwächse. Dass Klassenwiederholung und Schulformwechsel, deren Auswirkungen zu Eingang dieses Kapitels nur global abgeschätzt werden konnten, als mögliche Risiken mit zu bedenken sind, ist freilich hinzuzufügen. An den Haupt- und Realschulen, in denen im Anschluss an die Beobachtungsstufe ca. 60 Prozent der Schülerschaft in Realschulklassen bzw. ca. 50 Prozent in R-Kurse übergegangen sind, macht sich nicht dieser Anteil als solcher, wohl aber eine besonders konsequente äußere (Fach-) Leistungsdifferenzierung bemerkbar. Diese geht in der Regel mit höheren Fachleistungen der Schülerinnen und Schüler der Realschulklassen bzw. -kurse einher, doch gleichzeitig sind negative Folgen in den Hauptschulklassen bzw. H-Kursen sehr wahrscheinlich („Creaming effect“). Je konsequenter die leistungsstärkeren Gruppen der Schülerschaft aus den ursprünglichen Klassenverbänden herausgelöst werden, umso günstiger verläuft in diesen Gruppen die Lernentwicklung. Dem steht entgegen, dass die auf homogen niedrigem Niveau gebildeten Klassen- bzw. Kursverbände (Hauptschulen bzw. H-Kurse) deutlich geringere Lernfortschritte erreichen als heterogenere Lerngruppen der gleichen Stufe. In den Kursen I an den Gesamtschulen, in denen knapp über 40 Prozent der Schülerschaft unterrichtet werden, sind vergleichsweise hohe Lernzuwächse festzustellen, und es ist denkbar, dass die leistungsmäßige Überlegenheit gegenüber den Realschulen und den R-Kursen der Integrierten Haupt- und Realschulen mit jener vergleichsweise geringeren Übergangsquote in Zusammenhang steht. Es ist indessen bemerkenswert und

konsistent gegenüber den gesamtschultypischen Fluktuationen zwischen den Kursniveaus, dass diese Selektivität außer im Fach Mathematik im Mittel nicht stärker auf die Fachleistung bezogen ist als an den Haupt- und Realschulen. Dazu passt, dass in den Fächern Mathematik und Englisch, nicht aber im Fach Deutsch, auch an den Gesamtschulen Ansätze zu einem „Creaming effect“ auf dem Kursniveau II zu erkennen sind. Hervorzuheben ist der Umstand, dass systembedingt der Anteil an Klassenwiederholungen in den Gesamtschulen sehr gering ist, was den Abstand des Kursniveaus I gegenüber dem Realschulniveau geringer erscheinen lässt, als er sich unter Berücksichtigung dieser Einflussgröße darstellen würde.

Als Fazit ergibt sich, dass anregungsreiche Lernmilieus – seien dies Schulformen oder Kursniveaus –, die durch die Anwesenheit leistungsstarker Schülerinnen und Schüler geprägt sind, in der Regel für alle dort Lernenden leistungsmäßig Vorteile bieten. Dem stehen vor allem für die weniger leistungsstarken Schülerinnen und Schüler dann Nachteile gegenüber, wenn auf diese Weise homogene Lerngruppen niedrigen Leistungsstands entstehen. Allerdings ist noch einmal zu betonen, dass auch in den positiv ausgelesenen Lerngruppen, insbesondere an den Gymnasien und dabei vor allem an den Schulen mit besonders günstiger Anfangsleistung, ein kompensatorischer Unterricht vorzuherrschen scheint, der es den besonders leistungsstarken Schülerinnen und Schülern offenbar nicht hinreichend ermöglicht, Lernfortschritte zu erzielen, die ihrem Potenzial entsprechen.

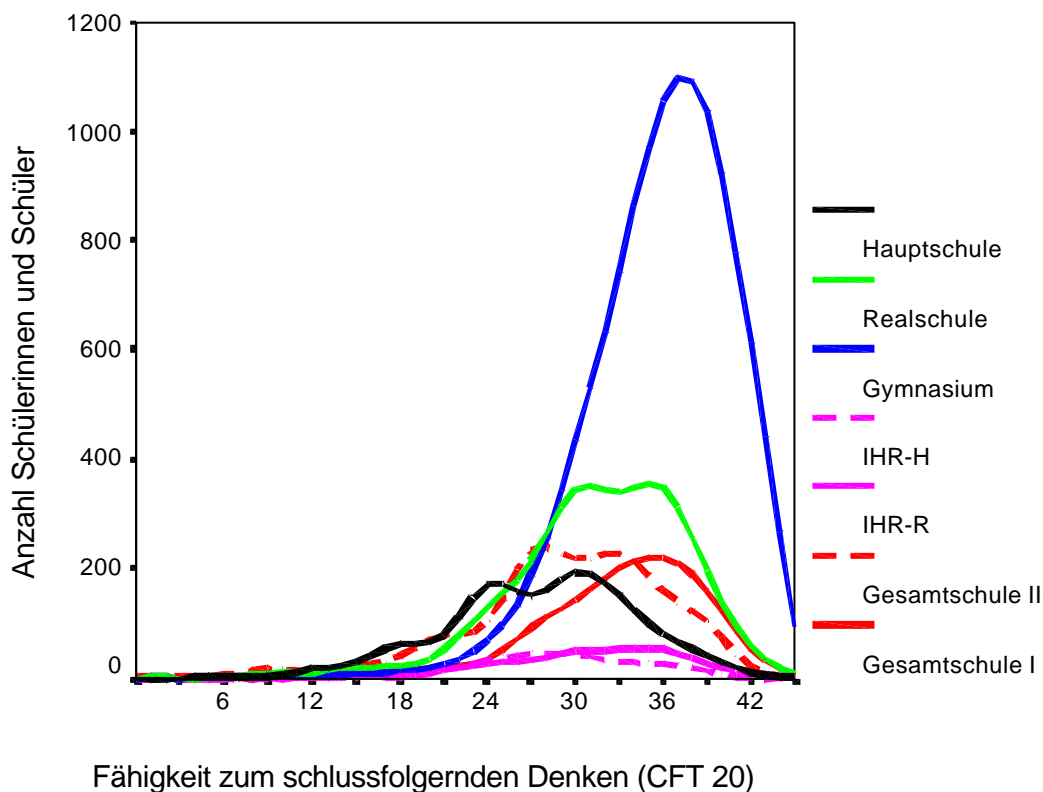
## 5 Individueller und sozialer Kontext der Schulleistungen

In diesem Kapitel werden einige Voraussetzungen und Bedingungen diskutiert, unter denen die für das Ende der Klassenstufe 8 festgestellten Schulleistungen zu Stande gekommen sind. Hierzu gehören kognitive Voraussetzungen – insbesondere die Fähigkeit zum schlussfolgernden Denken –, das Alter und Geschlecht der Schülerinnen und Schüler, Merkmale des familialen Umfeldes, ggf. ein Migrationshintergrund, und lern- bzw. schulbezogene Einstellungen der Schülerinnen und Schüler.

### 5.1 Kognitive Voraussetzungen

Zu den Daten, die im Jahre 2000 erhoben wurden, gehören die Ergebnisse im Test *CFT 20* (vgl. WEIB 1997), der mit nonverbalen Aufgaben die Fähigkeit zum schlussfolgernden Denken misst. Diese Fähigkeit unterscheidet sich praktisch nicht zwischen Jungen und Mädchen (Effektstärke  $d = 0,04$  zugunsten der Mädchen), wohl aber zwischen den Schulformen bzw. Kursniveaus. Dies wird in Abbildung 32 deutlich, die analog aufgebaut ist wie die entsprechenden Grafiken in Kapitel 4.

Abbildung 32 Fähigkeit zum schlussfolgernden Denken (Ergebnisse im *CFT 20*), nach Schulformen bzw. Kursniveaus



Die Ähnlichkeit dieser Verteilungen mit den Fachleistungsverteilungen sowie der Verteilung der Ergebnisse im Problemlösetest ist unverkennbar. Wegen des hohen Hamburger Gymnasialanteils dominiert in der Abbildung optisch wiederum diese Teilgruppe. Über alle Schulformen und Kursniveaus betrachtet, handelt es sich nicht um eine Normalverteilung: Verhältnismäßig viele Schülerinnen und Schüler haben durch hohe Testergebnisse eine asymmetrische („rechtssteile“) Verteilungsform entstehen lassen. Dies stimmt mit den von WEIß berichteten Ergebnissen seiner Eichstichprobe überein (1997, S. 53), wonach allerdings die Hamburger LAU-9-Resultate mit einem Mittelwert von 32,8 Punkten und einer Standardabweichung von 6,4 Punkten trotz niedrigerer Klassenstufe bei etwa gleicher Streuung generell günstiger sind. In der bundesweiten Eichstichprobe mit Schülerinnen und Schülern der Klassenstufen 9 und 10 liegt der Mittelwert bei 30,9 Punkten und die Standardabweichung bei 6,5.

Auffällig ist auch, dass an den Hauptschulen und Realschulen sowie bei der überwiegend auf dem Kursniveau II unterrichteten Schülerschaft der Gesamtschulen zweigipflige Verteilungsformen entstanden sind, die darauf hindeuten, dass sich hier auch Schülerinnen und Schüler befinden, deren kognitives Potenzial höher ist, als es ihre Schulform- bzw. Kursniveauzuordnung erwarten ließe. Diese Erscheinung, bei der die tatsächlichen Schulerfolge nicht dem kognitiven Potenzial entsprechen, wird in der Literatur zumeist mit dem englischen Fachausdruck „Underachievement“ bezeichnet und häufig mit motivationalen Defiziten in Verbindung gebracht. Im vorliegenden Kontext könnte man allerdings auch daran denken, dass es sich bei den „Underachievern“ vorzugsweise um Schülerinnen und Schüler mit Migrationshintergrund handelt, deren intelligenzunabhängige Sprachdefizite zu einer ungünstigen Schulform- oder Kursniveauzuweisung geführt haben. Diese Vermutung lässt sich für die hier untersuchte Schülerschaft eindeutig widerlegen: Definiert man als Underachiever für die genannten Schulformen bzw. Kursniveaus diejenigen, deren Ergebnisse im *CFT 20* oberhalb des Sattelpunkts der schulformspezifischen Werteverteilung liegen, so fallen 50,1 Prozent Deutsche ohne Migrationshintergrund, 47,7 Prozent Deutsche mit Migrationshintergrund und nur 37,5 Prozent Schülerinnen und Schüler ausländischer Nationalität unter diese Definition (zu dieser Unterscheidung vgl. den Abschnitt 5.4). Eine Benachteiligung der Gruppen mit einem Migrationshintergrund lässt sich also nicht belegen. Stattdessen ist Underachievement offenbar eher unter *deutschen* Schülern und Schülerinnen anzutreffen, und zwar deutlich häufiger unter den Jungen (52,2 Prozent gegenüber 47,5 Prozent bei den Mädchen). Dieser Geschlechterunterschied ist unter den Vergleichsgruppen sogar noch beträchtlich stärker ausgeprägt (Deut-

sche mit Migrationshintergrund: 52,3 Prozent bei den Jungen gegenüber 42,5 Prozent bei den Mädchen; Jugendliche ausländischer Nationalität: 41,4 Prozent bei den Jungen, 33,3 Prozent bei den Mädchen), doch dies belegt nur, dass Mädchen mit Migrationshintergrund, jedenfalls aus dieser Perspektive betrachtet, gerade *nicht* als eine besonders – gewissermaßen doppelt – benachteiligte Gruppe gelten können. Um die Relationen zwischen den Schulformen bzw. Kursniveaus auch numerisch greifbar zu machen, sind in Tabelle 37 die wichtigsten Kennzahlen zusammengefasst.

Tabelle 37 Fähigkeit zum schlussfolgernden Denken (Ergebnisse im *CFT 20*) am Ende der Klassenstufe 8, nach Schulform bzw. Kursniveau (Mittelwerte, Standardabweichungen, Effektstärken)

	$\bar{x}$	SD	Effektstärke <i>d</i>
Hauptschule	27,7	6,3	-0,81
Realschule	31,6	5,9	-0,19
Gymnasium	35,8	4,9	0,46
IHR Schule – H	28,4	5,9	-0,70
IHR Schule – R	31,1	5,9	-0,27
Gesamtschule II	29,4	6,7	-0,54
Gesamtschule I	33,1	5,8	0,05
<i>insgesamt</i>	<i>32,8</i>	<i>6,4</i>	<i>---</i>

Vergleicht man die hier berichteten Effektstärken mit denen für den Allgemeinen Fachleistungsindex (Tabelle 23), so fällt auf, dass die Zugehörigkeit zu einer Schulform oder zu einem Kursniveau weniger mit den kognitiven Lernvoraussetzungen zusammenhängt als mit der Fachleistung. Zwar könnte man argumentieren, dies sei insofern zu erwarten, als letztlich in der Tat die Fachleistung für die Einstufung maßgeblich sein müsse. Bekanntlich hängt diese aber nicht nur von den kognitiven Lernvoraussetzungen ab, sondern auch von anderen Faktoren, insbesondere auch den (hier nicht untersuchten) konkreten Lernangeboten einer Schulform oder eines Kursniveaus. Gleichwohl bleibt festzuhalten, dass es auf allen Ebenen der Untersuchung nicht wenige Schülerinnen und Schüler gibt, deren kognitive Lernvoraussetzungen den Besuch des Gymnasiums resp. des Kursniveaus I der Gesamtschulen gerechtfertigt hätten.

In diesem Kontext ist es von Interesse, in welchem Zusammenhang die Fähigkeit zum schlussfolgernden Denken zu den einzelnen Testleistungen steht (vgl. Tabelle 38).

Tabelle 38 Korrelation der Fähigkeit zum schlussfolgernden Denken (Er-

gebnisse im *CFT 20*) am Ende der Klassenstufe 8 mit Ergebnissen der Fachleistungstests und des Problemlösetests, nach Schulform bzw. Kursniveau

**SL-HAM 8/9, Untertest . . .**

	Mathe- matik	Lesever- ständnis	Sprache	Englisch (C-Test)	Allgemeine Fachleistung	Problem- lösen
Hauptschule	0,34	0,28	0,28	0,23	0,38	0,32
Realschule	0,40	0,36	0,35	0,24	0,43	0,39
Gymnasium	0,42	0,33	0,35	0,31	0,45	0,43
IHR Schule – H	0,41	0,29	0,26	0,19	0,38	0,37
IHR Schule – R	0,44	0,32	0,40	0,27	0,46	0,38
Gesamtschule II	0,43	0,36	0,28	0,10	0,40	0,37
Gesamtschule I	0,42	0,33	0,30	0,18	0,40	0,39
<i>insgesamt</i>	<i>0,58</i>	<i>0,51</i>	<i>0,52</i>	<i>0,48</i>	<i>0,59</i>	<i>0,57</i>

Den Werten für die Schülerschaft insgesamt ist zu entnehmen, dass die durch den praktisch sprachfreien *CFT 20* erfassten Fähigkeiten nahezu gleich hoch mit der Fachleistung Mathematik und den Ergebnissen im Problemlösetest korrelieren, ebenso mit dem Allgemeinen Fachleistungsindex. Erwartungsgemäß sind die Korrelationen mit den sprachlichen Untertests deutlich niedriger. Innerhalb der Schulformen sind die Zusammenhänge zwangsläufig abgeschwächt, weil erhebliche Varianzanteile auf die Unterschiede zwischen den Schulformen bzw. Kursniveaus entfallen. Dennoch sind in dieser Hinsicht einige Unterschiede auffällig. So stehen an den Hauptschulen die Fachleistungen generell in vergleichsweise geringem Zusammenhang mit den kognitiven Lernvoraussetzungen. Außer im Fach Mathematik ist dies auch bei den überwiegend auf Hauptschulniveau unterrichteten Schülerinnen und Schülern der Integrierten Haupt- und Realschulen der Fall. An den Gesamtschulen (beide Kursniveaus) zeigt sich diese Erscheinung vor allem im Bereich *Englisch*. Da in diesen Schulformen bzw. Kursniveaus die Ergebnisse im *CFT 20* z. T. eher überdurchschnittlich stark streuen und auch die Fachleistungen hier keineswegs durchgehend besonders homogen sind, gibt es keine erkennbaren formalstatistischen Gründe für eine solche Entkoppelung der Fachleistungen von den kognitiven Lernvoraussetzungen. So muss vorerst offen bleiben, warum in manchen der untersuchten Schulformen bzw. Kursniveaus und Testbereiche die Korrelationen zwischen der Fähigkeit zum schlussfolgernden Denken und den Fachleistungen besonders niedrig sind.

## 5.2 Alter und Geschlecht

Innerhalb einer Klassenstufe ist das Alter der Schülerinnen und Schüler gewöhnlich negativ mit der Fachleistung korreliert: Vorzeitige Einschulung und Springen sind selten, spätere Einschulung und Klassenwiederholungen hingegen gehen auch nach etlichen Schuljahren in der Regel mit niedrigeren Fachleistungen einher. Der untersuchte Schülerjahrgang macht in dieser Hinsicht keine Ausnahme: Die Korrelation zwischen allgemeiner Fachleistung und Lebensalter zum Testzeitpunkt beträgt  $r = -0,29$ .

Ein Teil dieser negativen Korrelation geht auch darauf zurück, dass sich die Schülerinnen und Schüler zwischen den Schulformen bzw. Kursniveaus in ihrem Durchschnittsalter unterscheiden. Dies wiederum ist zum Teil durch Unterschiede in der Häufigkeit von Klassenwiederholungen bedingt. In welchem Maße verspätete Einschulung und – im Falle von Zuwanderung aus dem Ausland – nicht unmittelbar altersentsprechende Klasseneinstufung beteiligt sind, lässt sich nur unvollständig rekonstruieren. Klassenwiederholungen in den Klassenstufen 5 bis 8 bzw. 9 konnten jedoch nahezu lückenlos erfasst werden. In der nachstehenden Tabelle 39 werden sie berichtet.

Die Altersunterschiede zwischen den einzelnen Schulformen bzw. Kursniveaus betragen maximal ein halbes Jahr, was mehr als einer dreiviertel Standardabweichung in der gesamten untersuchten Schülerschaft entspricht. Auch gibt es beträchtliche Unterschiede in den Streuungen der Schulformen bzw. Kursniveaus: Die Altersspannen sind z. B. an den Haupt- und Realschulen wesentlich größer als an den Gesamtschulen (Kursniveau I) und den Gymnasien. Selbstverständlich hängt beides unmittelbar mit den unterschiedlichen Anteilen an Klassenwiederholungen zusammen. An den Hauptschulen haben knapp 19 Prozent in den Klassenstufen 5 bis 9 eine Klassenstufe wiederholt, an den Realschulen sogar etwas mehr als 25 Prozent. An den Integrierten Haupt- und Realschulen sind die Wiederholerquoten trotz der sonst immer wieder deutlich gewordenen Vergleichbarkeit der Fachleistungen gegenüber den Hauptschulen auf der einen bzw. den Realschulen auf der anderen Seite wesentlich geringer, in der Größenordnung vergleichbar mit den Gymnasien. Systembedingt sind Klassenwiederholungen an den Gesamtschulen besonders selten.



Tabelle 39 Lebensalter der Schülerinnen und Schüler und Quote der Klassenwiederholungen seit Klassenstufe 5, nach Schulform bzw. Kursniveau (Mittelwerte, Standardabweichungen, Prozent der Wiederholungen und um Klassenwiederholung bereinigte Altersmittelwerte)

	Lebensalter		Klassenwiederholungen	adjustiertes Durchschnittsalter
	$\bar{x}$	SD		
Hauptschule	15,5	0,7	18,8 %	15,3
Realschule	15,3	0,7	25,5 %	15,1
Gymnasium	15,0	0,6	11,4 %	14,9
IHR Schule – H	15,3	0,7	11,0 %	15,2
IHR Schule – R	15,2	0,6	9,7 %	15,1
Gesamtschule II	15,2	0,6	4,2 %	15,2
Gesamtschule I	15,0	0,5	2,7 %	15,0
<i>insgesamt</i>	<i>15,2</i>	<i>0,6</i>	<i>12,4 %</i>	<i>15,0</i>

Berücksichtigt man die Klassenwiederholungen bei der Altersbestimmung – berechnet man also unter der (kontrafaktischen) Voraussetzung, dass die Schulformmittelwerte ausschließlich durch Klassenwiederholungen beeinflusst sind, das Durchschnittsalter gleichsam für das neunte Schulbesuchsjahr statt für die Klassenstufe 9 –, so zeigt sich eine andere Struktur. Es entsprechen dann die oben für die einzelnen Schulformen bzw. Kursniveaus berichteten aggregierten Durchschnittsleistungen in ihrer Rangfolge recht genau der umgekehrten Rangfolge der so adjustierten Altersmittelwerte: Je höher das allgemeine Leistungsniveau, umso jünger waren die Schülerinnen und Schüler in ihrem neunten Schulbesuchsjahr. Da jedoch die Mittelwertunterschiede auch dann noch erheblich sind, müssen noch andere Faktoren an ihrem Zustandekommen beteiligt sein. Vor allem spätere Einschulung, daneben auch Klassenwiederholungen in der Grundschule und nicht altersgemäße Klasseneinstufung bei Zuwanderern sind nahe liegende Möglichkeiten. Setzt man als theoretischen Altersdurchschnitt bei glattem Durchgang durch das Bildungssystem für den Testzeitpunkt 14,9 Jahre an, so ist im Mittel jede dritte Hauptschülerin bzw. jeder dritte Hauptschüler aus dem einen oder anderen dieser Gründe ein Jahr älter als vorgesehen, während sich beispielsweise an den Gymnasien Effekte durch frühe Einschulung oder Springen (0,2 Prozent an den Gymnasien) bzw. durch die in dieser Schülerschaft relativ seltenen Wiederholungen von Grundschulklassen die Waage halten (gerechnet *nach* Abzug des Einflusses von Klassenwiederholungen in der Sekundarstufe I).

Mädchen und Jungen waren von Klassenwiederholungen in unterschiedlichem Maße betroffen. In den Klassenstufen 5 bis 9 haben 10,9 Prozent der Mädchen, aber 13,8 Prozent der Jungen (mindestens) eine Klassenstufe wiederholt. Dabei finden sich zum Teil erhebliche Unterschiede zwischen den Schulformen bzw. Kursniveaus: Die Mädchen sind an den Hauptschulen überdurchschnittlich häufig von Klassenwiederholungen betroffen (20,6 Prozent gegenüber 17,5 Prozent Jungen), geringfügig auch in den Kursen II der Gesamtschulen (4,4 Prozent gegenüber 4,0 Prozent Jungen). In den übrigen Schulformen aber gilt dies für die Jungen, am stärksten in den Gymnasien (14,8 Prozent gegenüber 8,3 Prozent Mädchen) und in den H-Kursen der Integrierten Haupt- und Realschulen (13,5 Prozent gegenüber 7,6 Prozent Mädchen).

Außer in den H-Kursen der Integrierten Haupt- und Realschulen sind die Mädchen überall leistungsmäßig stärker (Allgemeiner Fachleistungsindex:  $d = 0,15$ ), so dass die erhöhte Wiederholerquote der Mädchen an den Hauptschulen und in den Kursen II der Gesamtschulen erklärungsbedürftig scheint. Aber auch umgekehrt ist zu prüfen, ob die häufigeren Klassenwiederholungen der Jungen in den übrigen Gruppen vom Leistungsbild her gerechtfertigt sind. Dabei hat sich ergeben, dass an den Hauptschulen und in den Kursen II der Gesamtschulen mehr Mädchen als Jungen gerade erst die Klassenstufe 9 wiederholt hatten und in der Folge vergleichsweise hohe Fachleistungen aufwiesen. Auch sonst lässt sich zeigen, dass zwar schulform- bzw. kursniveauübergreifend die allgemeine Fachleistung von Jungen, die mindestens eine Klassenstufe wiederholt hatten, geringfügig besser war als die der Mädchen mit gleichem Hintergrund ( $d = 0,04$ ), dass aber innerhalb der Schulformen und Kursniveaus mit Ausnahme der H-Kurse an den Integrierten Haupt- und Realschulen die Fachleistung der Wiederholerinnen etwas höher war als die der Wiederholer. Somit gibt es keinen Anhaltspunkt für die Vermutung, dass die Jungen bei gleicher Fachleistung grundsätzlich von einem höheren Versetzungsrisiko bedroht seien als die Mädchen.

Der allgemeineren Frage, ob und ggf. wie sich die Schulleistungen der Jungen und Mädchen in den einzelnen Schulfächern innerhalb der Sekundarstufe I bis zum Beginn der Klassenstufe 9 differenziell entwickelt haben, soll hier so nachgegangen werden, dass für die vier zentralen Untersuchungsbereiche (Mathematik, Leseverständnis, Sprache, Englisch) die Leistungsentwicklung von der ersten Datenerhebung (Klassenstufe 5) bis zur vorerst letzten (Klassenstufe 9) nachgezeichnet wird. Dies soll jeweils zunächst in tabellarischer Form geschehen und dann auch grafisch veranschaulicht werden. Tabelle 40 enthält die entsprechenden Angaben für das Fach Mathematik.

Tabelle 40 Fachleistung *Mathematik* nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform bzw. Kursniveau (Mittelwerte; Effektstärke  $d$  für das Merkmal Geschlecht am Ende der Klassenstufe 8 bzw. zu Beginn der Klassenstufe 9<sup>46</sup>)

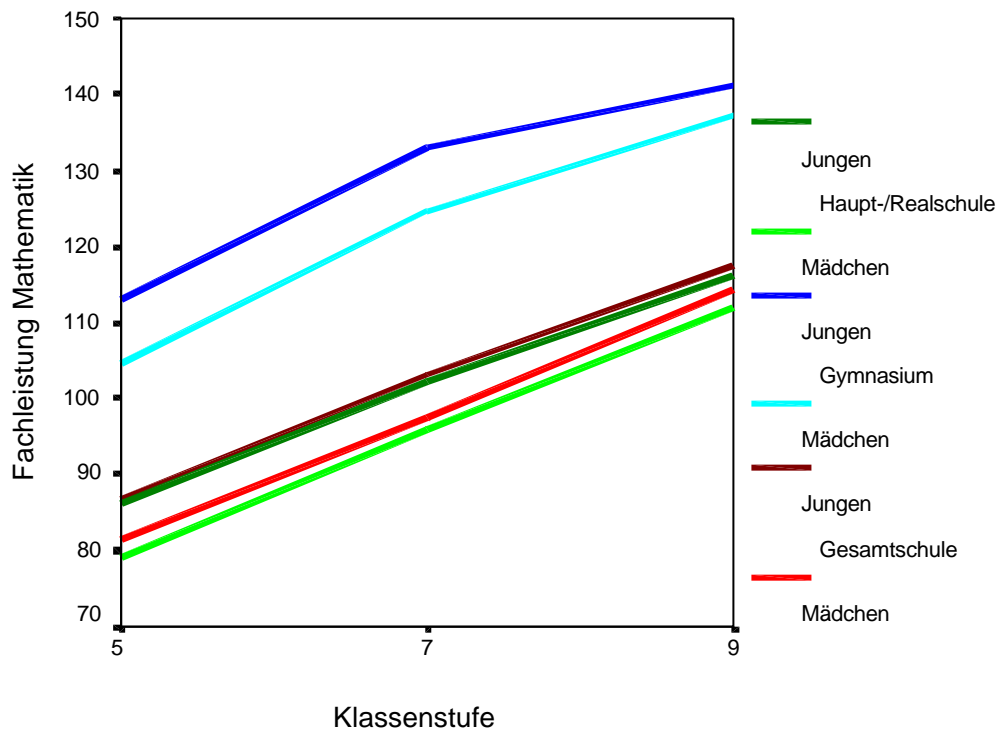
	Klassenstufe 5		Klassenstufe 7		Klassenstufe 9		Effektstärke $d$
	Jungen	Mädchen	Jungen	Mädchen	Jungen	Mädchen	
Hauptschule	77,7	71,3	93,9	88,1	105,9	102,8	-0,24
Realschule	94,0	84,9	108,7	100,6	123,2	117,5	-0,40
Gymnasium	113,0	104,5	132,9	124,8	141,2	137,2	-0,28
IHR Schule – H	74,3	69,6	94,6	89,8	106,9	101,3	-0,44
IHR Schule – R	89,6	81,6	107,2	99,9	123,6	117,6	-0,43
Gesamtschule II	79,9	74,7	95,2	89,5	109,3	106,2	-0,24
Gesamtschule I	94,7	88,8	112,6	106,7	127,2	124,2	-0,19
<i>insgesamt</i>	<i>96,7</i>	<i>90,9</i>	<i>113,8</i>	<i>108,3</i>	<i>126,2</i>	<i>123,3</i>	<i>-0,15</i>

Diesen Befunden zufolge ist der Unterschied zwischen den Mathematikleistungen der Jungen und Mädchen im Hamburger Gesamtsystem bis zum Beginn der Klassenstufe 9 ( $d = -0,15$ ) dem Betrage nach unter den Stand am Ende der Klassenstufe 6 ( $d = -0,22$ ) sowie der Grundschulzeit ( $d = -0,22$ ) gefallen. Innerhalb der Schulformen bzw. Kursniveaus (die hier allerdings anders aufgeschlüsselt sind als in den früheren Berichten) ist indessen keineswegs überall eine entsprechende Angleichung der Leistungen von Jungen und Mädchen festzustellen. Für die beiden Differenzierungsniveaus der Integrierten Haupt- und Realschulen gilt dies beispielsweise nicht, während sich an den Gymnasien der am Ende der Klassenstufe 6 bzw. Anfang der Klassenstufe 7 noch bestehende Unterschied zu Gunsten der Jungen von  $d = -0,36$  auf  $d = -0,28$  verringert hat. Zudem können die Modalitäten der äußeren Differenzierung dort, wo diese stattfindet, ihrerseits die Leistungsentwicklung von Jungen und Mädchen differenziell beeinflussen (vgl. die hohen Unterschiede zwischen den Geschlechtern auf den beiden Kursniveaus der Integrierten Haupt- und Realschulen mit den auffällig geringen auf den beiden Differenzie-

<sup>46</sup> Die Schulform- bzw. Kursniveauzuordnung ist hier in derselben Weise vorgenommen worden wie in den beiden vorangegangenen Kapiteln, bezieht sich also eigentlich auf das Ende der Klassenstufe 8 und nicht auf den Beginn der Klassenstufe 9. Auf diese Weise wird die Lernentwicklung der entsprechenden Schülergruppen *unabhängig von ihrer Gruppenzugehörigkeit zu früheren Erhebungszeitpunkten* ablesbar. Vergleichswerte für die Schulformen (bzw. Kursniveaus zu Beginn der Klassenstufe 7) zu den früheren Erhebungszeitpunkten können den bereits vorliegenden Berichten entnommen werden.

rungsstufen der Gesamtschulen). Um in dieser Hinsicht eine gewisse Vergleichbarkeit (und zugleich Übersichtlichkeit) zu erhalten, sind deshalb in der nachfolgenden Abbildung 33 jeweils alle Schülerinnen bzw. alle Schüler der Haupt- und Realschulen, der Gymnasien und der Gesamtschulen ohne Rücksicht auf ihre Niveaueinstufung am Ende der Klassenstufe 8 zusammengefasst.

Abbildung 33 Fachleistung Mathematik nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform (Mittelwerte)



In dieser Abbildung wird erkennbar, dass die Leistungsentwicklung der Jungen und Mädchen an den Haupt- und Realschulen insgesamt und – etwas stärker noch – an den Gesamtschulen in den Klassenstufen 5 bis Ende 8 tendenziell konvergent verlaufen ist. Es besteht demnach wenig Grund für die Annahme, dass der Unterricht in diesen Schulformen gewollt oder ungewollt geschlechtsbezogene Leistungsunterschiede im Fach Mathematik erzeuge oder verstärke. Das Gegenteil scheint der Fall zu sein: Unbeschadet teilweise erheblicher Differenzen innerhalb dieser Schulformen und ihrer Kursniveaus und unbeschadet der Ausnahme der Integrierten Haupt- und Realschulen, dominieren Effekte, die mit einer Verminderung der Geschlechtsunterschiede einhergehen, wobei sich insgesamt die Gesamtschule in dieser Hinsicht bei schon anfangs ausgeglicheneren Verhältnissen als recht wirksam erweist, erkennbar an der relativ geringen Effektstärke von  $d = -0,19$  zum Ende der Klassenstufe 8. Bei den Gymnasien fällt vor allem die diskontinuierliche Lernentwicklung auf, die unter Bezugnahme auf die Klassenstufen 7 und 8 bereits im Kapitel 4 dis-

kutiert worden ist. Daneben wird deutlich, dass diese Abschwächung kumulativen Lernens am Gymnasium bei den Jungen viel stärker ausgeprägt ist als bei den Mädchen. Somit ist die auch hier zu beobachtende Verringerung der Geschlechterdifferenzen im Fach Mathematik vor allem mit verminderten Lernfortschritten der Jungen verbunden, wobei allerdings die Mädchen den seit dem Übergang in die Sekundarstufe I festzustellenden Abstand zu den Jungen nicht ganz haben aufholen können. Danach ergeben sich keine Anhaltspunkte für besondere Schwierigkeiten oder Benachteiligungen der Mädchen in den Klassenstufen 7 und 8, die sich an größeren Lernzuwächsen bei den Jungen zeigen müssten.

In Tabelle 41 sind die Ergebnisse im Leseverständnistest, aufgeschlüsselt nach Erhebungszeitpunkt (Klassenstufe), Schulform bzw. Kursniveau und Geschlecht, zusammengefasst. Vergleicht man diese Werte wieder mit den Ergebnissen aus den beiden früheren Untersuchungen, so ist festzustellen, dass sich die Geschlechterunterschiede im Leseverständnis im Hamburger Bildungswesen insgesamt von einer Effektstärke  $d = 0,11$  (Klassenstufe 5) über  $d = 0,16$  (Klassenstufe 7) auf  $d = 0,29$  (Klassenstufe 9) zu Gunsten der Mädchen vergrößert haben.

Tabelle 41 Fachleistung *Leseverständnis* nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform bzw. Kursniveau (Mittelwerte; Effektstärke  $d$  für das Merkmal Geschlecht am Ende der Klassenstufe 8 bzw. zu Beginn der Klassenstufe 9)

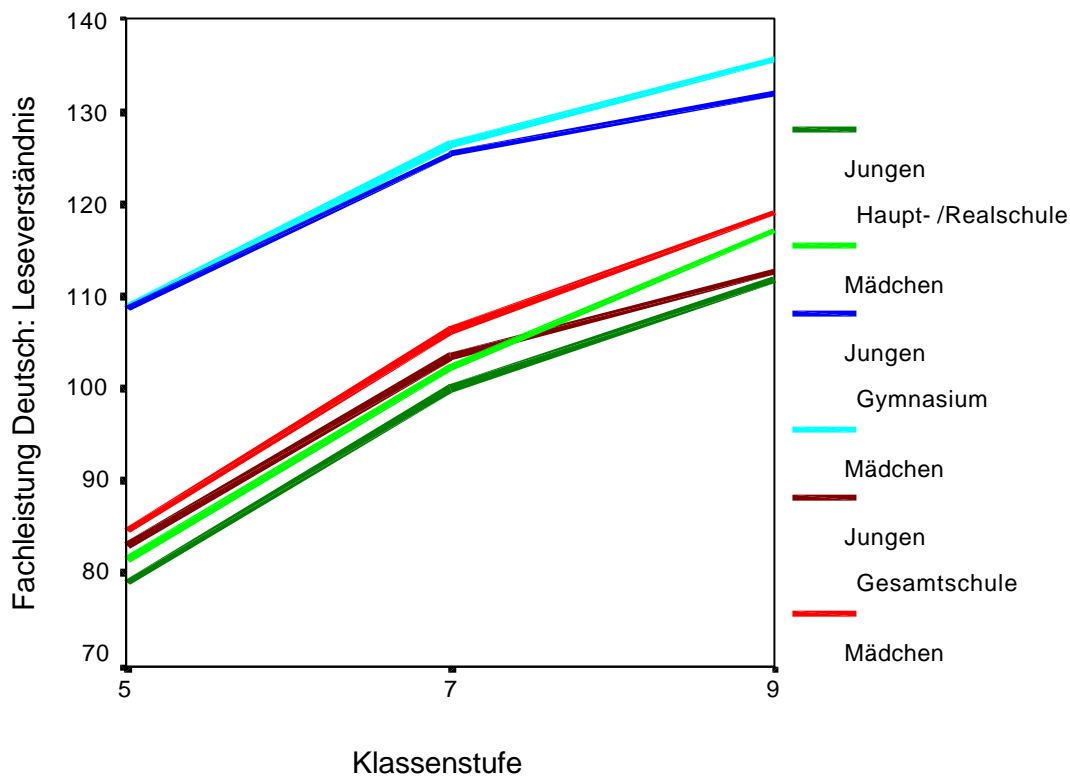
	Klassenstufe 5		Klassenstufe 7		Klassenstufe 9		Effektstärke $d$
	Jungen	Mädchen	Jungen	Mädchen	Jungen	Mädchen	
Hauptschule	71,1	70,6	91,0	91,5	101,6	108,4	0,35
Realschule	87,2	88,6	107,5	109,5	119,2	122,3	0,19
Gymnasium	108,7	108,9	125,4	126,4	131,9	135,7	0,28
IHR Schule – H	70,1	69,7	92,4	92,4	104,5	108,9	0,27
IHR Schule – R	8,8	85,2	104,2	106,5	117,7	121,6	0,24
Gesamtschule II	76,8	75,3	98,3	96,9	107,7	112,1	0,23
Gesamtschule I	94,9	92,8	113,7	115,0	123,3	125,9	0,15
<i>insgesamt</i>	<i>91,6</i>	<i>94,2</i>	<i>110,3</i>	<i>113,4</i>	<i>119,6</i>	<i>125,3</i>	<i>0,29</i>

Mit Ausnahme der Hauptschulen, wo diese Differenz nunmehr besonders hoch ausfällt, sowie der Kurse I der Gesamtschulen, wo sie fast auf dem Wert vom Beginn der Klassenstufe 7 ( $d = 0,16$ ) stehen geblieben und damit besonders niedrig ist, kann man innerhalb der Schulformen bzw. Kursniveaus eine recht

gleichmäßige Überlegenheit der Mädchen hinsichtlich ihres Leseverständnisses beobachten.

Auch hier ist wieder eine grafische Veranschaulichung aufschlussreich, die auf der Schulformzugehörigkeit basiert und die Niveauezugehörigkeit innerhalb der Schulformen mit äußerer Differenzierung außer Acht lässt (Abbildung 34).

Abbildung 34 Fachleistung *Leseverständnis* nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform (Mittelwerte)



Wie im Fach Mathematik erweisen sich hinsichtlich des Leseverständnisses die Jungen aus dieser Perspektive als diejenige Gruppe, in der sich die kumulativen Lernfortschritte besonders verlangsamt haben. Dem steht in allen Schulformen – vor allem aber an den Haupt- und Realschulen – eine stetigere Lernentwicklung der Mädchen gegenüber. Diese Beobachtung ist zugleich deswegen wichtig, weil sie eine Teilerklärung für die in Kapitel 4 berichteten unerwartet geringen Lernfortschritte in den Klassenstufen 7 und 8 liefert: Diese sind offenbar in beträchtlichem Maße mit der vergleichsweise ungünstigen Lernentwicklung der Jungen verbunden. Dabei wiederum ist es bedeutsam, dass es sich nicht um ein Methodenartefakt handeln kann, weil sich diese Erscheinung auch dann zeigt, wenn die Jungen die leistungsschwächere Gruppe darstellen. Im Übrigen ist die an den Gesamtschulen besonders deutliche Divergenz der Lernentwicklung von Mädchen und Jungen (Klassenstufe 9:  $d = 0,32$ ; zum Vergleich: Haupt- und Realschulen  $d = 0,27$ , Gymnasien  $d = 0,28$ ) mit den relativ geringen

gen Geschlechterdifferenzen auf der Ebene der Kursniveaus nur dadurch in Einklang zu bringen, dass die Jungen in den Kursen II ihrem Anteil, wenn auch nicht ihrer Fachleistung nach überrepräsentiert sind. Dies ist in der Tat der Fall: An den Gesamtschulen befinden sich in den Kursen II im Fach Deutsch 58,9 Prozent Jungen, in den Kursen I hingegen nur 39,7 Prozent.

Im Untertest *Sprache* haben sich unter der Frage nach Geschlechterdifferenzen insgesamt ähnliche Befunde ergeben wie beim Leseverständnis. Allerdings sind beim Sprachverständnis in den Schulformen bzw. Kursniveaus mit niedrigerem Anforderungsniveau, d. h. an den Hauptschulen, in den H-Kursen der Integrierten Haupt- und Realschulen und in den Kursen II der Gesamtschulen, keine oder nur sehr geringe Leistungsvorteile der Mädchen zu erkennen (vgl. Tabelle 42).

Tabelle 42 Fachleistung *Sprache* nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform bzw. Kursniveau (Mittelwerte; Effektstärke  $d$  für das Merkmal Geschlecht am Ende der Klassenstufe 8 bzw. zu Beginn der Klassenstufe 9)

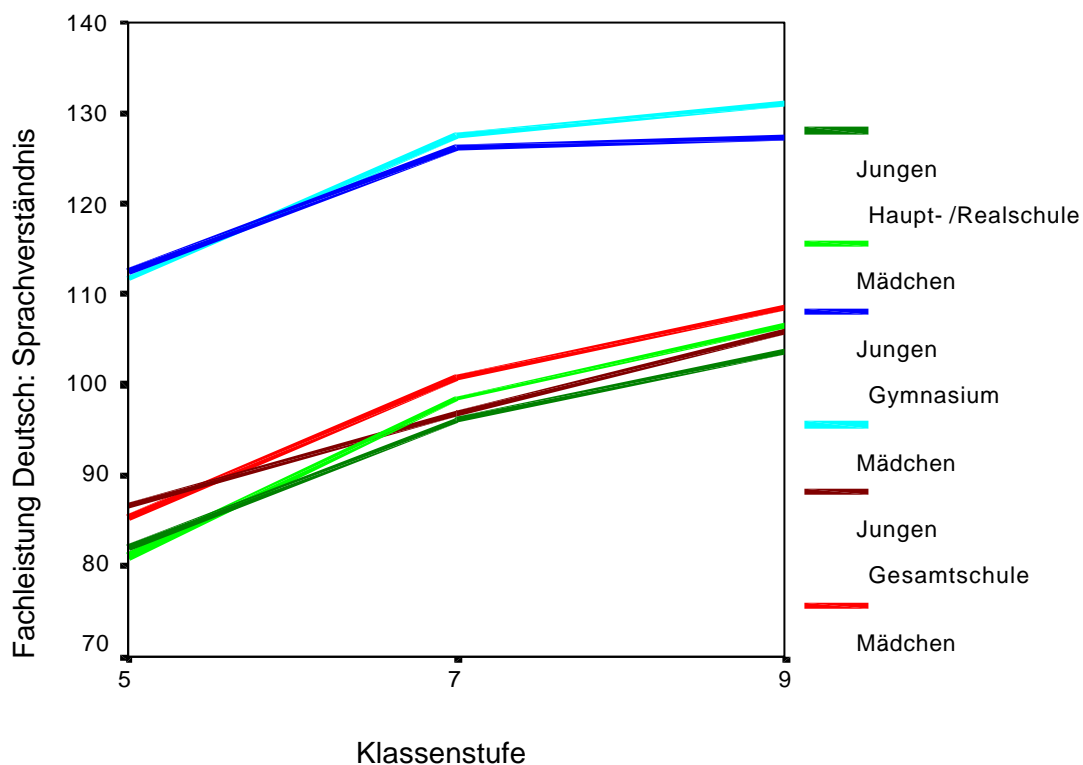
	Klassenstufe 5		Klassenstufe 7		Klassenstufe 9		Effektstärke $d$
	Jungen	Mädchen	Jungen	Mädchen	Jungen	Mädchen	
Hauptschule	73,0	72,5	87,2	88,2	95,4	96,7	0,08
Realschule	89,5	87,1	102,3	105,2	109,7	112,7	0,22
Gymnasium	112,6	111,7	126,2	127,6	127,3	131,0	0,27
IHR Schule – H	74,9	70,2	91,0	88,7	98,0	97,9	-0,00
IHR Schule – R	83,1	83,2	102,0	102,9	108,0	110,2	0,18
Gesamtschule II	80,5	75,2	92,4	92,6	102,5	102,4	-0,00
Gesamtschule I	97,9	94,3	105,7	108,1	112,7	114,7	0,14
<i>insgesamt</i>	<i>95,3</i>	<i>95,7</i>	<i>107,6</i>	<i>111,3</i>	<i>113,2</i>	<i>117,3</i>	<i>0,22</i>

Bei der Interpretation dieser Werte ist – wie schon zuvor – zu bedenken, dass die Einteilung der Schülerinnen und Schüler nach dem Stand der Schulform bzw. Kursniveauezugehörigkeit am Ende der Klassenstufe 8 erfolgt ist. So waren z. B. die Jungen, die im Schuljahr 1999/2000 im Fach Deutsch in den Kursen I der Gesamtschulen unterrichtet wurden und zu dieser Zeit einen etwas niedrigeren Lernstand als die Mädchen aufwiesen, diesen zu Beginn der Klassenstufe 5 *überlegen* ( $d = -0,16$ ), was zunächst nur etwas über die relative Leistungsentwicklung dieser speziellen Gruppe aussagt. Die Lernentwicklung ist jedoch überall analog verlaufen: Wo die Mädchengruppe anfangs hinter den Jungen im Hinblick auf das Sprachverständnis zurückstand, hat sie zumindest

gleichgezogen, besonders in den H-Kursen der Integrierten Haupt- und Realschulen und in den Kursen II der Gesamtschulen. Wo der Anfangsunterschied gering war, hat sich nunmehr ein Vorsprung der Mädchen herausgebildet, der vor allem an den Gymnasien auffällig ist. Im Hamburger Bildungssystem insgesamt ist die Effektstärke von  $d = 0,02$  auf  $d = 0,22$  angewachsen.

Eine Grafik, in der die ursprünglichen Schulformen ohne Berücksichtigung der Differenzierungsniveaus zusammengefasst sind, verdeutlicht, dass die Entwicklung des Sprachverständnisses in den Klassenstufen 5 bis 8 bei Jungen und Mädchen in den verschiedenen Schulformen recht unterschiedlich und anders als beim Leseverständnis verlaufen ist (vgl. Abbildung 35).

Abbildung 35 Fachleistung *Sprache* nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform (Mittelwerte)





Hier zeigt sich in allen Schulformen bei den Mädchen eine Verlangsamung der Lernfortschritte in den Klassenstufen 7 und 8, allerdings vor allem an den Gymnasien, wo der Abstand der Mädchen gegenüber den Schülerinnen der beiden anderen Schulformen abgenommen hat. Die Jungen an den Haupt- und Realschulen folgen ab Klassenstufe 7 etwa dieser allgemeinen Zuwachsrate der Mädchen, während die Jungen an den Gesamtschulen in den Klassenstufen 5 und 6 auffallend geringe Lernzuwächse erzielt, danach aber diese Zuwachsrate beibehalten haben. So hat sich ihr Abstand zu den Mädchen seither zumindest nicht vergrößert. Als besonders auffällige Gruppe stellt sich hier die männliche Schülerschaft der Gymnasien dar, in der in den Klassenstufen 7 und 8 praktisch kein Zuwachs mehr nachweisbar ist (1,1 Lernstandspunkte). Dass es sich hierbei nicht um ein Artefakt handelt, ergibt sich schon daraus, dass bei den Gymnasiastinnen auf höherem Niveau mit 3,4 Lernstandspunkten noch Lerngewinne festgestellt werden konnten.

Schließlich sollen die Befunde für das Fach *Englisch* vorgestellt werden, wobei allerdings aus der Klassenstufe 5 keine Testergebnisse zur Verfügung stehen (vgl. Tabelle 43).

Tabelle 43 Fachleistung Englisch (C-Test) nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform bzw. Kursniveau (Mittelwerte; Effektstärke  $d$  für das Merkmal Geschlecht am Ende der Klassenstufe 8 bzw. zu Beginn der Klassenstufe 9)

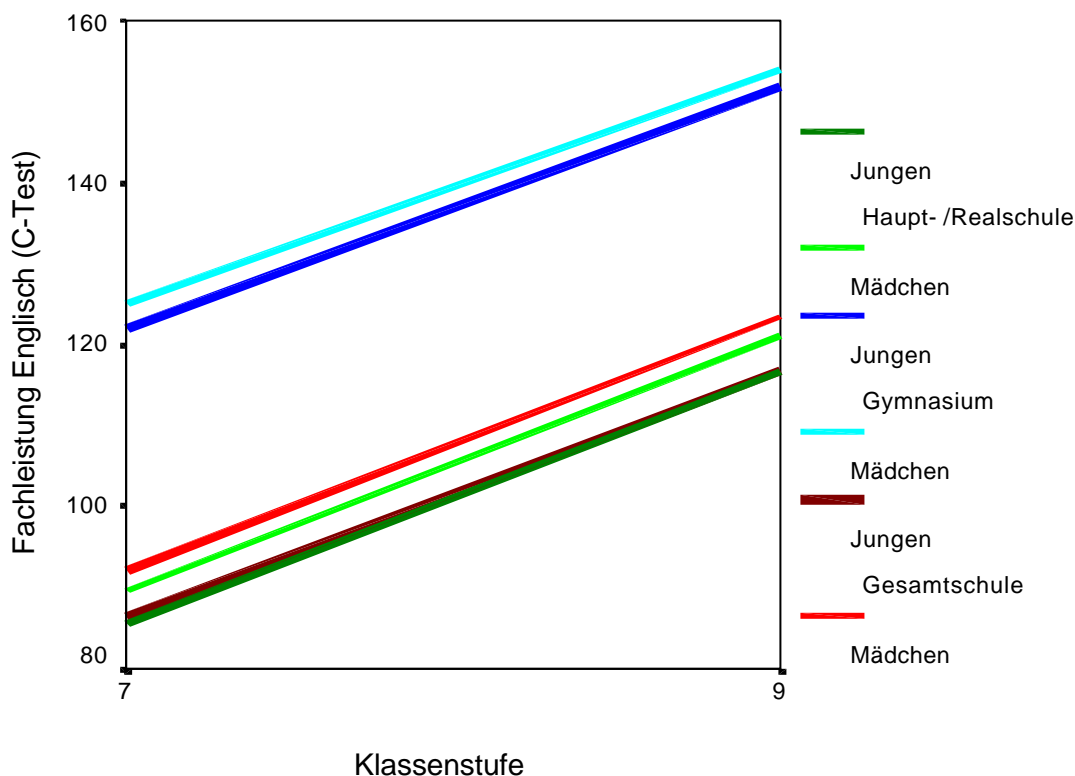
	Klassenstufe 7		Klassenstufe 9		Effektstärke $d$
	Jungen	Mädchen	Jungen	Mädchen	
Hauptschule	73,0	77,5	101,5	105,8	0,25
Realschule	94,7	97,4	126,9	130,0	0,19
Gymnasium	122,0	125,2	151,8	154,0	0,15
IHR Schule – H	76,0	78,6	106,4	108,8	0,15
IHR Schule – R	89,9	92,5	123,4	128,6	0,35
Gesamtschule II	79,5	83,3	107,7	111,6	0,24
Gesamtschule I	96,9	101,4	131,2	134,9	0,21
<i>insgesamt</i>	98,9	104,8	130,2	135,5	0,21

Bezogen auf das Hamburger Schulwesen insgesamt ist der Leistungsunterschied der Geschlechter im Fach Englisch richtungsgleich und etwa so groß wie im Untertest *Sprache*, jedoch weniger stark ausgeprägt als im Leseverständnis. Dagegen ist er deutlich größer als die leistungsmäßige Überlegenheit der Jungen im Fach Mathematik ( $d = -0,15$ , s. o.). Zudem hat sich die Ge-

schlechterdifferenz im Fach Englisch seit Beginn der Klassenstufe 7 ( $d = 0,24$ ) nur wenig verändert. Nimmt man die relativ kleine Gruppe der Schüler aus R-Kursen der Integrierten Haupt- und Realschulen aus, so sind die Differenzen trotz der unterschiedlichen Leistungsniveaus in allen Schulformen und Kursniveaus vergleichbar stark ausgeprägt.

Um die geschlechterspezifische Lernentwicklung – analog zu den Fächern Mathematik und Deutsch – auch für das Fach Englisch grafisch darzustellen, werden in der nachstehenden Abbildung 36 wieder die jeweiligen Mittelwerte für die verschiedenen Erhebungszeitpunkte dargestellt, und zwar in diesem Fall nur für die Klassenstufen 7 und 9, da für den Beginn der Klassenstufe 5 auf eine Lernstandserhebung verzichtet wurde. Der Gesamtmittelwert für die Klassenstufe 7 wurde im Prozess der Skalierung aus pragmatischen Gründen auf 100 festgesetzt.<sup>47</sup>

Abbildung 36 Fachleistung Englisch (C-Test) nach Erhebungszeitpunkt (Klassenstufe), Geschlecht und Schulform (Mittelwerte)



<sup>47</sup> Die geringfügige Abweichung von diesem Wert ergibt sich daraus, dass nicht für alle Schülerinnen und Schüler das Kursniveau zum Ende der Klassenstufe 8 eindeutig festgestellt werden konnte.

Diese Abbildung vermittelt drei wichtige Erkenntnisse: So muss es, erstens, an den Gymnasien in den Klassenstufen 5 und 6 im Fach Englisch höhere Lernfortschritte gegeben haben als an den anderen Schulformen, soweit nicht die bereits zu Beginn der Klassenstufe 5 vorhandenen Englischkenntnisse mit der späteren Schulformzugehörigkeit korreliert waren. Zweitens ist es den Gymnasien, trotz des erkennbaren hohen Leistungsvorsprungs zu Beginn der Klassenstufe 7, offenbar nicht gelungen, diesen weiter auszubauen, und zwar weder bei den Mädchen noch bei den Jungen. Drittens hat sich die zu Beginn der Klassenstufe 7 festgestellte Überlegenheit der Gymnasiastinnen gegenüber den Gymnasiasten ( $d = 0,20$ ) bis zum Ende der Klassenstufe 8 auf den in der Tabelle aufgeführten, vergleichsweise geringen Wert von  $d = 0,15$  vermindert. An den Haupt- und Realschulen sowie an den Gesamtschulen sind die bereits zu Beginn der Klassenstufe 7 bestehenden Leistungsunterschiede zwischen Jungen und Mädchen weitgehend erhalten geblieben.

Zusammenfassend ergibt sich für die untersuchten Leistungsaspekte, dass die teilweise unerwartet geringen Lernfortschritte in den Klassenstufen 7 und 8 eng mit einer vergleichsweise langsameren Lernentwicklung der Jungen in den Fächern Mathematik und Deutsch zusammenhängen: Vom Beginn der Klassenstufe 7 an gerechnet ist die Überlegenheit der Mädchen hinsichtlich der allgemeinen Fachleistung von  $d = 0,07$  auf etwa das Doppelte ( $d = 0,15$ ) angestiegen. Im Fach Mathematik geht dies mit einer Verringerung der Abstände zwischen den Geschlechtern in den Klassenstufen 7 und 8 einher, im Fach Deutsch dagegen mit einer umso bedeutsameren Vergrößerung. Im Fach Englisch schließlich bleiben die Leistungsunterschiede, die sich bis zum Beginn der Klassenstufe 7 eingestellt haben, weitgehend erhalten.

### **5.3 Außerschulischer Hintergrund**

Zum Kernbestand bildungssoziologischer Erkenntnisse gehört es, dass der familiäre Hintergrund – insbesondere die sog. „Bildungsnähe“ des Elternhauses – ihren Niederschlag in der Zuordnung von Kindern und Jugendlichen zu bestimmten Bildungskarrieren und darüber hinaus in den erreichten Leistungsständen findet. Dass Hamburger Schulen bzw. Lehrkräfte solche Effekte nicht abmildern, sondern eher verstärken, ist am Beispiel der Grundschulempfehlungen in einem früheren Bericht gezeigt worden (vgl. LEHMANN & PEEK 1997, S. 86 ff.). Zu prüfen ist hier, ob solche Chancenungleichheiten auch am Ende der Klassenstufe 8 noch nachweisbar sind. Als Kriterium der Bildungsnähe wird dabei der höchste erreichte Schulabschluss der Mutter oder des Vaters gewählt. Diese Variable korreliert über alle Schulformen und Kurs-

niveaus hinweg mit der allgemeinen Fachleistung in Höhe von  $r = 0,41$ : Schülerinnen und Schüler, deren Eltern über einen höheren Schulabschluss verfügen, haben erkennbar höhere Testleistungen erbracht als solche, deren Eltern die Schule nach der Pflichtschulzeit verlassen haben, und zwar gilt dies auch innerhalb der Schulformen (möglicher oder wahrscheinlicher Schulabschluss konstant gehalten: partielles  $r = 0,13$ ).

Ein häufig gewählter Ansatz zur Beantwortung der Frage nach dem Verhältnis von Schule und gesellschaftlicher Schichtung ist der, von Chancenungleichheit dann zu sprechen, wenn die Quoten des Schichtungsmerkmals (hier: der Schulabschlüsse) insgesamt abweichen von den Quoten in den Schulformen bzw. Kursniveaus. Tabelle 44 zeigt, dass dies in Hamburg am Ende der Klassenstufe 8 in der Tat der Fall ist.

Tabelle 44 Verteilung der Schulabschlüsse der Eltern bzw. der / des Erziehungsberechtigten nach Schulform bzw. Kursniveau (in Prozent)

	<i>Eltern bzw. Erziehungsberechtigte(r) ...</i>						<i>insgesamt</i>
	ohne Abschluss	POS 8 <sup>48</sup>	Hauptschule	Realschule/ POS 10 <sup>49</sup>	Fachhochschulreife	Abitur	
Hauptschule	21,9	15,7	21,2	10,1	6,6	2,9	9,5
Realschule	21,3	30,4	23,5	23,3	18,8	9,2	17,5
Gymnasium	10,6	14,7	15,2	34,7	48,5	70,5	45,1
IHR Schule – H	9,6	4,9	5,1	2,2	1,0	0,7	2,3
IHR Schule – R	4,3	3,9	4,2	3,7	3,2	1,1	2,8
Gesamtschule II	20,9	22,5	22,8	15,4	9,9	6,2	13,0
Gesamtschule I	11,3	7,8	8,0	10,6	12,0	9,5	9,8
<i>insgesamt (N)</i>	<i>(301)</i>	<i>(102)</i>	<i>(1.613)</i>	<i>(2.761)</i>	<i>(784)</i>	<i>(3.462)</i>	<i>(9.023)</i>

Nach diesem Maßstab sind Schülerinnen und Schüler, bei denen mindestens ein Elternteil die Allgemeine Hochschulreife erlangt hat, an Haupt- und Realschulen sowie in den Kursen II (nicht aber in den Kursen I) der Gesamtschulen deutlich unterrepräsentiert, wie die Prozentangaben in der Tabelle 35 belegen. Umgekehrt sind von den Schülerinnen und Schülern, bei denen kein Elternteil einen Schulabschluss erlangt hat, an den Hauptschulen und in den H-Kursen der Integrierten Haupt- und Realschulen um mehr als das Doppelte und in den

<sup>48</sup> Abgang aus der Polytechnischen Oberschule (POS) der ehemaligen DDR nach der Klassenstufe 8

<sup>49</sup> Abgang aus der POS der ehemaligen DDR nach der Klassenstufe 10

Kursen II der Gesamtschulen um mehr als 60 Prozent überrepräsentiert. Ihre Chance, ein Gymnasium zu besuchen, beträgt weniger als ein Viertel der Wahrscheinlichkeit für die Gesamtheit aller Schülerinnen und Schüler und nur ein Siebtel derjenigen, die dann gilt, wenn mindestens ein Elternteil das Abitur erworben hat.

Dieses „Quotenkriterium“ ist problematisch, weil dabei vorausgesetzt wird, dass Lernvoraussetzungen und Fachleistungen in allen Subgruppen gleich verteilt sind. Dies ist aber nicht der Fall. So sind, um wieder die beiden Extrembeispiele zu nehmen, die kognitiven Lernvoraussetzungen bei den Schülerinnen und Schülern, deren Eltern keinen Schulabschluss erlangt haben, deutlich – nämlich um  $d = 0,91$  Standardabweichungen – ungünstiger ausgeprägt als bei denjenigen, bei denen mindestens ein Elternteil das Abitur hat. Bei der allgemeinen Fachleistung am Anfang der Klassenstufe 7 beträgt der Unterschied sogar  $d = 1,30$  Standardabweichungen, was durch die differenziellen Übergangsquoten am Ende der Grundschulzeit mitbedingt ist.

Eine denkbare Alternative ist es, sich an den möglichen bzw. wahrscheinlichen Schulabschlüssen zu orientieren, so wie sie am Ende der Klassenstufe 8 aus der Schulform- bzw. Kursniveauezugehörigkeit abzuleiten sind. Dazu lässt sich ein dreistufiger Index gewinnen, der der Höhe der Schulabschlüsse entspricht.<sup>50</sup> Für die vorliegenden Zwecke wurden dazu folgende Festlegungen getroffen:

- *Stufe 1 (Hauptschulabschluss möglich oder wahrscheinlich)*: Schülerinnen und Schüler der Hauptschulen, der Integrierten Haupt- und Realschulen, sofern sie höchstens einen R-Kurs besuchen, und der Gesamtschulen, sofern sie durchweg dem Kursniveau II zugeordnet sind;
- *Stufe 2 (Realschulabschluss möglich oder wahrscheinlich)*: Schülerinnen und Schüler der Realschulen, der Integrierten Haupt- und Realschulen, sofern sie mindestens zwei R-Kurse besuchen, und der Gesamtschulen, sofern sie einem oder zwei Kursen des Kursniveaus I zugeordnet sind;
- *Stufe 3 (Abitur möglich oder wahrscheinlich)*: Schülerinnen und Schüler der Gymnasien und der Gesamtschulen, sofern sie durchgängig den Kursen des Kursniveaus I zugeordnet sind.

Es ergeben sich auf diesem Wege Prozentwerte, die – gemessen an den Ab-

---

<sup>50</sup> Streng genommen handelt es sich dabei um eine ordinalskalierte Variable; sie wird aus Gründen des leichteren Nachvollzugs in einigen der nachstehenden Analysen wie eine intervallskalierte Variable behandelt.

solventen- und Übergangsquoten in Hamburg – nicht unplausibel sind: 24,9 Prozent Schülerinnen und Schüler der Stufe 1, 28,9 Prozent der Stufe 2 und 46,3 Prozent der Stufe 3.

Statistisch kann man nun zeigen, dass der Bildungshintergrund der Eltern bzw. der oder des Erziehungsberechtigten unabhängig von der allgemeinen Fachleistung zu Beginn der Klassenstufe 7 einen deutlichen eigenständigen Einfluss auf die mutmaßliche künftige Bildungskarriere ausübt, und zwar stärker noch als die kognitiven Lernvoraussetzungen. Hierbei wird für jede nach dem elterlichen Schulabschluss gebildete Gruppe der tatsächliche Mittelwert auf der von 1 bis 3 reichenden Skala des möglichen oder wahrscheinlichen Schulabschlusses bestimmt. Außerdem wird ein bereinigter („adjustierter“) Gruppenmittelwert berechnet, der die Unterschiede zwischen den Gruppen in den kognitiven Lernvoraussetzungen und im allgemeinen Lernstand am Anfang der Klassenstufe 7 berücksichtigt. Wenn die soziale Herkunft nach Berücksichtigung der kognitiven Lernvoraussetzungen und der Fachleistung keine Rolle mehr spielte, wären die adjustierten Werte identisch mit dem allgemeinen Mittelwert, der 2,36 beträgt (vgl. Tabelle 45).

Tabelle 45 Möglicher oder wahrscheinlicher Schulabschluss der Jugendlichen: tatsächlicher und um Unterschiede in den kognitiven Lernvoraussetzungen sowie in der Fachleistung zu Beginn der Klassenstufe 7 bereinigter („adjustierter“) Wert, nach Schulabschluss der Eltern oder der bzw. des Erziehungsberechtigten<sup>51</sup>

möglicher oder wahrscheinlicher Schulabschluss	Eltern bzw. Erziehungsberechtigte(r)...					
	ohne Abschluss	POS 8	Hauptschule	Realschule/ POS 10	Fachhochschulreife	Abitur
tatsächlicher Wert	1,74	1,94	1,83	2,24	2,47	2,73
adjustierter Wert	2,20	2,31	2,14	2,33	2,41	2,48
<i>Differenz zwischen adjustiertem Wert und Durchschnitt (= 2,36)</i>	-0,16	-0,05	-0,22	-0,03	0,05	0,12

An den Differenzen zwischen dem um den Einfluss der Lernvoraussetzungen und des Lernstands am Beginn der Klassenstufe 7 bereinigten Wert für den

<sup>51</sup> Es handelt sich hier um Ergebnisse einer Kovarianzanalyse mit dem dreistufigen Index für den möglichen oder wahrscheinlichen Schulabschluss als abhängiger Variable und den Ergebnissen im CFT 20 und dem Allgemeinen Fachleistungsindex aus der LAU-7-Untersuchung als Kovariaten.

möglichen (oder wahrscheinlichen) Schulabschluss der Jugendlichen und dem allgemeinen Durchschnitt wird deutlich, dass es in der Tat an den Hamburger Schulen eine Benachteiligung bestimmter Schülergruppen auf Grund ihrer sozialen Herkunft gibt. Am stärksten ist dieser Effekt in der Gruppe derjenigen Schülerinnen und Schüler, deren Eltern bzw. Erziehungsberechtigte (höchstens) den Hauptschulabschluss erreicht haben. Auf der anderen Seite werden die Schülerinnen und Schüler, bei denen mindestens ein Elternteil über die Allgemeine Hochschulreife verfügt, hinsichtlich ihrer Zuordnung zu einer Schulform (das ist in 70,5 Prozent dieser Gruppe das Gymnasium) oder eines Kursniveaus eindeutig bevorzugt. Somit sind – jedenfalls unter der Voraussetzung, dass kognitives Lernpotenzial und Fachleistung für die Festlegung eines Bildungsgangs maßgeblich sein sollten – auch am Ende der Klassenstufe 8 immer noch beträchtliche Bildungsbarrieren für die Kinder bzw. Jugendlichen aus bildungsferneren Schichten der Bevölkerung festzustellen. Umgekehrt werden so die nachhaltigen, nicht mit den unterschiedlichen Lernvoraussetzungen begründbaren Vorteile derjenigen sichtbar, bei denen mindestens ein Elternteil das Abitur hat. Dass diese Gegensätze schwerlich allein durch differenzielle Lernförderung in den verschiedenen sozialen Schichten bedingt sind, legen vor allem die eingangs erwähnten Ergebnisse aus der Untersuchung von 1996 nahe, nach denen schon beim Übertritt in die weiterführenden Schulen soziale Barrieren wirksam sind.

Hier schließt sich die Frage an, ob aus den vorliegenden Hintergrunddaten Hinweise auf Gründe für die privilegierte Position einiger Schülergruppen gewonnen werden können. Eine häufig geäußerte Vermutung bezieht sich auf den Nachhilfeunterricht und damit auf eine Bevorzugung von Schülerinnen und Schülern aus finanziell besser gestellten Familien. Rund 30 Prozent der Schülerinnen und Schüler haben angegeben, in mindestens einem Schulfach zumindest gelegentlich Nachhilfeunterricht zu erhalten, wobei Mathematik und Fremdsprache(n) mit etwa gleicher Häufigkeit dominieren. Die übliche Form ist der private Nachhilfeunterricht, genannt werden aber auch Nachhilfeschulen. Die Überprüfung des vermuteten Zusammenhangs ergibt jedoch, dass die (in der Regel besser gestellten) Eltern mit Abitur keineswegs am häufigsten für Nachhilfeunterricht sorgen und aufkommen, sondern mit 28,2 Prozent etwa ebenso häufig wie die Eltern ohne Schulabschluss (27,9 Prozent) und die Eltern, die nach der Klassenstufe 8 die Polytechnische Oberschule der ehemaligen DDR verlassen haben (28,0 Prozent). Die Werte für die Eltern mit Hauptschulabschluss (30,9 Prozent), Realschulabschluss bzw. Abschluss der zehnklassigen Polytechnischen Oberschule (31,2 Prozent) und Fachhochschulreife (31,1 Prozent) sind etwas höher. Zwar sind diese Unterschiede nicht sehr groß,

aber sie widerlegen die genannte Vermutung.

Ebenfalls bemerkenswert ist es, dass Nachhilfeunterricht keine typisch gymnasiale Erscheinung ist. Die Gymnasiastinnen und Gymnasiasten nehmen mit 29,8 Prozent eine durchschnittliche Position ein, während die Schülerinnen und Schüler der Integrierten Haupt- und Realschulen, die überwiegend H-Kurse besuchen (35,2 Prozent), und die Schülerschaft der Realschulen (33,4 Prozent) stärker vom Nachhilfeunterricht betroffen sind. Der Grund ist wohl darin zu suchen, dass Nachhilfeunterricht vor allem durch drohende Klassenwiederholung veranlasst wird, weshalb beispielsweise die wiederholungsintensiven Realschulen, aber auch die jeweils unteren Kursniveaus an Integrierten Haupt- und Realschulen sowie an Gesamtschulen relativ hohe Nachhilfesüherquoten haben. Stellt man im Übrigen die negative Korrelation zwischen der allgemeinen Fachleistung und dem Nachhilfeeinsatz statistisch in Rechnung, so bleibt nur ein äußerst geringer, praktisch irrelevanter Zusammenhang mit dem Bildungsniveau der Eltern bzw. der oder des Erziehungsberechtigten bestehen ( $\beta = 0,04$ ).

Ein weiterer Erklärungsansatz für die relativ privilegierte bzw. unterprivilegierte Position bestimmter Schülergruppen nimmt auf das (fehlende) Vorhandensein bildungsrelevanter Ressourcen in den Elternhäusern Bezug. Hierzu zählen Bücher (speziell auch die eigenen), Lexika, Computer, Kassettenrekorder und CD-Abspielgeräte. Interessanterweise zählen ausweislich der empirischen Befunde der eigene Fernseher und der eigene Videorekorder nicht zu dieser Gruppe. Man kann diese bildungsrelevanten Ressourcen statistisch zu einem Faktor zusammenfassen und prüfen, in welchem Zusammenhang er mit der allgemeinen Fachleistung und auch mit dem möglichen bzw. zu erwartenden Schulabschluss steht. Nach den vorliegenden Daten ist der Ansatz bei den bildungsrelevanten Ressourcen aussagekräftiger als der Bezug auf den (vor allem finanziell konnotierten) Nachhilfeunterricht: 78,8 Prozent der Varianz der allgemeinen Fachleistung am Ende der Klassenstufe 8 lässt sich aufklären, wenn man die Fachleistung zu Beginn der Klassenstufe 7 ( $\beta = 0,75$ ) und die kognitiven Lernvoraussetzungen ( $\beta = 0,16$ ) berücksichtigt und dann den eigenständigen Einfluss bildungsrelevanter Ressourcen ( $\beta = 0,06$ ) sowie der Schulabschlüsse der Eltern ( $\beta = 0,05$ ) untersucht.<sup>52</sup> Ein

---

<sup>52</sup> Die  $\beta$ -Koeffizienten beziehen sich stets auf den *eigenständigen* Erklärungsbeitrag einer Variable. Da die bildungsrelevanten Ressourcen und die Schulabschlüsse der Eltern aus vielen Gründen deutlich mit der Fachleistung zu Beginn der Klassenstufe 7 korreliert sind, ist ihr davon unabhängiger Einfluss hier gering. Hingegen stand in der LAU 5 kein Prädiktor für die Fachleistung zu einem früheren Zeitpunkt zur Verfügung, weshalb dort der Einfluss der sozialen Hintergrundvariablen viel größer erschien.



davon unabhängiger positiver Einfluss von Nachhilfeunterricht lässt sich auf diesem Wege nicht belegen. Selbstverständlich erfassen diese Faktoren auch wesentliche Facetten der Lerngeschichte *vor* dem Eintritt in die Klassenstufe 7, nicht zuletzt vermittelt durch die Schulform- bzw. Kursniveaueugehörigkeit. Zumindest so viel wird aber sichtbar, dass es in der Tat unterstützende materielle und wohl auch immaterielle Leistungen bildungsnaher Elternhäuser gibt, mit denen die Lernentwicklung auch in den Klassenstufen 7 und 8 unterstützt wird.

Kehrt man zurück zu der Frage, welche Hintergründe die Überrepräsentanz von Schülerinnen und Schülern aus bildungsnahen Elternhäusern in den aussichtsreicheren Bildungsgängen haben könnte, so gibt es nunmehr deutliche Hinweise auf die daran beteiligten Faktoren. Mit den genannten Prädiktoren lassen sich immerhin 59,7 Prozent der Varianz der Variable „möglicher bzw. wahrscheinlicher Schulabschluss“ aufklären. An erster Stelle steht wieder die allgemeine Fachleistung am Ende der Beobachtungsstufe resp. am Ende der Klassenstufe 6 an den Gesamtschulen ( $\beta = 0,64$ ), gefolgt vom höchsten Schulabschluss der Eltern bzw. der oder des Erziehungsberechtigten ( $\beta = 0,14$ ) und von den bildungsrelevanten Ressourcen im Elternhaus ( $\beta = 0,07$ ). Die kognitiven Lernvoraussetzungen stehen bei dieser Formulierung der Fragestellung mit ebenfalls  $\beta = 0,07$  an letzter Stelle. Somit bestätigt sich noch einmal die begünstigende Rolle eines hohen Bildungsabschlusses der Eltern bzw. der oder des Erziehungsberechtigten, die durch die unterstützende Funktion der bildungsrelevanten Ressourcen im Elternhaus ergänzt wird.

Somit kann die Unterrepräsentanz von Kindern bzw. Jugendlichen aus bildungsfernen Schichten in den aussichtsreicheren Bildungsgängen keineswegs allein unter Berufung auf die Ungleichverteilung kognitiver Lernvoraussetzungen und bereits gezeigter Fachleistung „erklärt“ werden. Benachteiligungen auf Grund des niedrigen Bildungsstatus mancher Eltern und des damit häufig einhergehenden Mangels an bildungsrelevanten Ressourcen sind mit hoher Wahrscheinlichkeit mitbeteiligt.

## 5.4 Migrationshintergrund

Zu den wichtigsten Aufgaben der Hamburger Schulen zählt die Integration von Schülerinnen und Schülern, die einen Migrationshintergrund haben. Hierzu gehört nicht nur die Schaffung sozialer Akzeptanz für diese Gruppe, sondern auch die Eröffnung von Lebenschancen, die an die von der Schule vermittelten Kompetenzen geknüpft sind.

Für die Untersuchung der Frage, in welchem Maße es bis zum Ende der Klassenstufe 8 gelungen ist, Schülerinnen und Schüler mit Migrationshintergrund trotz der damit oftmals verbundenen sprachlichen und kulturellen Barrieren an das allgemeine Leistungsniveau heranzuführen, ist es nicht unerheblich, wie diese Gruppe formal umschrieben wird. Am eindeutigsten ist der Rekurs auf die *Staatsangehörigkeit*, die in den Schülerakten aufgenommen ist. Nach diesem Kriterium sind 20,6 Prozent der Schülerinnen und Schüler Ausländer. Um die Zusammensetzung dieser Schülerschaft zu charakterisieren, wird in der Tabelle 46 die Häufigkeit der unterschiedlichen Staatsangehörigkeiten mitgeteilt.

Tabelle 46 Staatsangehörigkeit der ausländischen Schülerinnen und Schüler, teilweise zusammengefasst zu Regionen (in Prozent)

Türkei	32,4 %	Afrika (Subsahara)	3,4 %
Afghanistan	11,6 %	sonstiges Europa	3,3 %
ehemaliges Jugoslawien	9,3 %	Ost- und Südostasien	2,6 %
Europäische Union	7,3 %	Südasien (Indien und Pakistan)	1,7 %
Iran	7,0 %	Nordafrika und Naher Osten	0,8 %
Polen	6,2 %	<i>sonstige Staatsangehörigkeit</i>	3,1 %
Russische Föderation	4,7 %	<i>doppelte (inkl. deutsche) Staatsangehörigkeit</i>	6,5 %

Da Informationen über die Staatsangehörigkeit aus allen drei bisherigen Erhebungswellen vorliegen, ist es aufschlussreich zu prüfen, ob sich – beispielsweise durch Zuwanderung oder Schulformwechsel<sup>53</sup> – Verschiebungen in der Zusammensetzung der Schülerschaft ergeben haben. Tabelle 47 bietet die entsprechenden Angaben, wobei für die früheren Erhebungen der Sache nach keine Differenzierung nach Kurszugehörigkeit erfolgen kann.

<sup>53</sup> Unschärfen ergeben sich allerdings durch Klassenwiederholungen, die die Zusammensetzung der untersuchten Schülerschaft insgesamt beeinflusst haben.

Die Veränderungen in den Ausländeranteilen an den verschiedenen Schulformen sind überwiegend geringfügig. Vorstellbar ist, dass die Ausländerquote seit der Klassenstufe 7 insgesamt trotz einiger später Zugänge durch überzufällig häufige Schulabgänge ohne Abschluss leicht zurückgegangen ist; dafür sprechen vor allem die Rückgänge in den Hauptschulen, aber auch in den Integrierten Haupt- und Realschulen. An den Gymnasien ist bereits während der oder im Anschluss an die Beobachtungsstufe ein dem allgemeinen Trend zuwider laufender Rückgang zu beobachten, was vermutlich mit Schulleistungsproblemen und nachfolgendem Schulformwechsel zusammenhängt. Danach allerdings bleibt der Ausländeranteil an den Gymnasien ebenso stabil wie an den Gesamtschulen. Im Übrigen wird deutlich, dass sich in den Schulformen mit äußerer Differenzierung die ausländischen Schülerinnen und Schüler jeweils in den unteren Kursen konzentrieren.

Tabelle 47 Anteil von Schülerinnen und Schülern mit ausländischer Staatsangehörigkeit nach Schulform und ggf. Kursniveau (in Prozent)<sup>54</sup>

	<b>Klassenstufe 5</b>	<b>Klassenstufe 7</b>	<b>Klassenstufe 9</b>
Hauptschule	28,8	32,2	30,0
Realschule		23,2	23,0
Gymnasium	13,8	12,9	12,9
IHR Schule – H	31,7	32,8	34,2
IHR Schule – R			25,9
Gesamtschule II			27,1
Gesamtschule I	22,2	23,3	17,5
<i>insgesamt</i>	<i>20,7</i>	<i>20,9</i>	<i>20,2<sup>55</sup></i>

Ausländische Schülerinnen und Schüler waren, wie die (nicht immer ganz eindeutigen aus der Erhebung von 1996 stammenden) Angaben zum Einschulungsjahr im Vergleich mit den Daten zur Sekundarstufe I nahe legen, häufiger von Klassenwiederholungen schon in der Grundschulzeit betroffen. Klassenwiederholungen im Sekundarbereich waren in der deutschen und ausländischen Schülerschaft fast exakt gleich häufig. Eine Einschulung nach Beginn der Pflichtschulzeit kann kein sehr häufiges Phänomen sein, da sich 1996 nur bei 4,1 Prozent der ausländischen Schülerinnen und Schüler Hinweise darauf fan-

<sup>54</sup> Die mögliche Überrepräsentanz von Ausländern in den Sonderschulen muss hier außer Betracht bleiben (vgl. dazu Powell & Wagner 2001).

<sup>55</sup> Diese Zahl weicht von den oben genannten 20,6 Prozent ab, weil für einen Teil der Schülerschaft die Kursniveauzugehörigkeit nicht eindeutig bestimmt werden konnte.

den. Die ausländische Schülerschaft war zum Erhebungszeitpunkt im Durchschnitt ein Vierteljahr älter als die deutsche, was demnach vor allem auf die Klassenwiederholungen in der Grundschulzeit zurückzuführen ist.

Zu berücksichtigen ist allerdings, dass die Staatsbürgerschaft kein hinreichend aussagekräftiges Kriterium ist, wenn es um die besondere Lage der Kinder und Jugendlichen mit Migrationshintergrund geht, die unter Umständen schon in der zweiten oder dritten Generation in Deutschland leben und einen deutschen Pass besitzen oder die neu zugewandert sind und als Aussiedler Deutsche sind, dabei aber nicht immer über gute Kenntnisse der deutschen Sprache verfügen. Deshalb sollen hier auch solche Kriterien in Betracht gezogen werden, die auf den Angaben der Schülerinnen und Schüler zur eigenen Muttersprache und auf den Angaben der Eltern bzw. Erziehungsberechtigten zu der in der Familie gesprochenen Sprache beruhen.

Bezieht man sich auf die Angaben zur *Muttersprache*, so haben 28,4 Prozent der Schülerinnen und Schüler einen Migrationshintergrund; stützt man sich auf die Angaben der Eltern zur Herkunft und *Familiensprache*, so sind es sogar 33,6 Prozent. Diese Schätzungen sind nicht unbeeinflusst von den unterschiedlichen Rücklaufquoten der eingesetzten Tests, Schüler- und Elternfragebögen. Generell gilt, dass dabei die Schüler- oder Elternfragebögen seltener bei Leistungsstarken fehlen als bei Leistungsschwachen, dies jedoch stärker bei den deutschen Familien als bei den Migranten. So erscheinen die Leistungsunterschiede zwischen den beiden Schülergruppen als etwas geringer, wenn man die Elternangaben zur Definition des Migrantenstatus heranzieht. Gleichzeitig ermöglicht die Kombination von Elternangaben und Kenntnis der Staatsbürgerschaft, differenziertere Aussagen über die Situation der verschiedenen Gruppen zu machen. Somit stehen drei mögliche Kategorisierungen zur Verfügung:

1. Migrationshintergrund, definiert über die *Staatsbürgerschaft*: Deutsche vs. Nichtdeutsche;
2. Migrationshintergrund, definiert über die Schülerangaben zur *Muttersprache*: Schülerinnen und Schüler mit deutscher Muttersprache vs. solche mit einer oder maximal zwei nichtdeutschen Muttersprachen;<sup>56</sup>
3. Migrationshintergrund, definiert über die Elternangaben zur *Familien-*

---

<sup>56</sup> Deutsch war nicht vorgegeben. Die schulischen Fremdsprachen Englisch, Französisch und Spanisch wurden nur dann berücksichtigt, wenn sie im Abgleich mit der Staatsbürgerschaft als mögliche Muttersprachen erschienen. Dagegen kommen in Familien mit gemischtem ethnischen Hintergrund auch mehrere Sprachen vor.

*sprache*: Schülerinnen und Schüler, in deren Familien eine fremde Sprache gesprochen wird<sup>57</sup>, vs. solche, bei denen dies nicht der Fall ist.

In der nachfolgenden Tabelle 48 werden diese Kategorisierungen miteinander verglichen, indem die entsprechenden Quoten nach Schulform bzw. Kursniveau aufgeschlüsselt werden.

Tabelle 48 Migrationshintergrund, definiert über die Kriterien Staatsbürgerschaft, Muttersprache und Familiensprache, nach Schulform bzw. Kursniveau (Migrantenanteil in der jeweiligen Schulform und Bildungsbeteiligung von Migranten, in Prozent)

	<i>Migranten nach ...</i>						<i>deutsche Staatsbürger ohne Migrationshintergrund</i>
	<b>Staatsbürgerschaft</b>		<b>Muttersprache</b>		<b>Familiensprache</b>		
	<b>Anteil</b>	<b>Bildungsbeteilig.</b>	<b>Anteil</b>	<b>Bildungsbeteilig.</b>	<b>Anteil</b>	<b>Bildungsbeteilig.</b>	<b>Bildungsbeteiligung</b>
Hauptschule	30,0	16,2	38,5	14,0	44,1	12,8	8,3
Realschule	23,0	19,2	36,1	23,0	40,1	20,9	15,8
Gymnasium	12,9	24,8	19,9	30,6	25,9	34,4	50,0
IHR Schule – H	34,2	4,4	42,5	3,8	43,3	3,0	2,0
IHR Schule – R	25,9	3,4	36,6	3,7	40,9	3,3	2,4
Gesamtschule II	27,1	22,1	34,2	15,8	41,3	16,0	11,6
Gesamtschule I	17,5	9,9	27,0	9,2	32,6	9,5	9,9
<i>insgesamt</i>	<i>20,2</i>	<i>100,0</i>	<i>28,4</i>	<i>100,0</i>	<i>33,6</i>	<i>100,0</i>	<i>100,0</i>
<i>(N)</i>	<i>(11.965)</i>		<i>(9.851)</i>		<i>(9.183)</i>		

Ergebnis dieses Vergleichs ist es, dass man zu sehr unterschiedlichen Anteilen von Schülerinnen und Schülern mit Migrationshintergrund gelangt, je nachdem, welches Kriterium herangezogen wird. Je weiter der Kreis derer gezogen wird, die einen Migrationshintergrund haben, desto mehr gleichen sich die relativen Anteile an die Verteilung der deutschen Schülerinnen und Schüler ohne Migrationshintergrund auf die verschiedenen Schulformen an. In jedem Falle aber konzentrieren sich die Schülerinnen und Schüler mit einem Migrationshintergrund, selbst im weitesten Sinne gefasst, in den Bildungsgängen mit Haupt- oder Realschulabschluss, während sie an den Gymnasien unterrepräsentiert sind. Nur in den Kursen I der Gesamtschulen sind sie ungefähr ihrem Anteil

<sup>57</sup> Hier galten die gleichen Einschränkungen wie bei dem Kriterium der Muttersprache.

entsprechend vertreten. Somit sind die geringeren Bildungsaussichten für die Migranten auch in dieser Schulform erkennbar, und zwar hier an den unterschiedlichen Quoten im Gefolge der Kursniveaudifferenzierung. Es ist jedoch bemerkenswert, dass in Hamburg auch das Gymnasium von einer Schülerschaft besucht wird, die zu mehr als einem Viertel, sprachlich gesehen, einen Migrationshintergrund hat.

Nun gilt für die Frage der Chancengerechtigkeit im Hamburger Schulwesen – bezogen auf die Migranten – das Gleiche wie im Hinblick auf die soziale Schichtung. Von einer echten Benachteiligung wird man nur dann sprechen, wenn die Schülerinnen und Schüler mit Migrationshintergrund auch nach angemessener Berücksichtigung der kognitiven Lernvoraussetzungen und der Fachleistungen in den anspruchsvolleren Schulformen bzw. Kursniveaus unterrepräsentiert sind. Es bietet sich an, diese Frage methodisch in derselben Weise zu untersuchen wie im vorangegangenen Abschnitt und hierzu den Index des möglichen oder wahrscheinlichen Schulabschlusses heranzuziehen. Um Einflüsse unterschiedlicher Lerngelegenheiten in den Klassenstufen 7 und 8 angemessen zu berücksichtigen, ist zur Kontrolle der Fachleistungen wieder der allgemeine Lernstand zu Beginn der Klassenstufe 7 heranzuziehen. Bei der Klassifikation des Migrationshintergrundes wird eine Kombination des Kriteriums der Staatsbürgerschaft mit dem der Familiensprache gewählt. Demnach wird unterschieden zwischen

- Schülerinnen und Schülern mit deutscher Staatsangehörigkeit ohne einen an der Familiensprache erkennbaren Migrationshintergrund (66,2 Prozent),
- solchen mit deutscher Staatsangehörigkeit, jedoch mit mindestens einer nichtdeutschen Sprache in der Familie (15,6 Prozent) und
- solchen ohne deutsche Staatsangehörigkeit (18,3 Prozent).

Ein Vergleich der tatsächlichen Bildungschancen im oben definierten Sinne und der Chancen, die entsprechend den kognitiven Lernvoraussetzungen und den Fachleistungen adjustiert sind, findet sich für diese drei Gruppen in Tabelle 49. Bei einer unter dem Gesichtspunkt der Lernvoraussetzungen und Fachleistungen „gerechten“ Zuordnung der Schülerinnen und Schüler auf die verschiedenen Bildungsgänge sollten dabei keine Abweichungen zwischen dem Gesamtmittelwert für die möglichen oder wahrscheinlichen Schulabschlüsse (Skalenwert 2,35) und den adjustierten Werten auftreten.

Das Ergebnis der Analyse wird vielleicht überraschen, doch es stimmt mit einem Befund aus der Untersuchung zu Beginn der Klassenstufe 5 überein, demzufolge Migranten (damals lediglich nach der Staatsbürgerschaft definiert)

etwas *günstigere* Chancen beim Übertritt in die weiterführenden Schulen hatten, als es ihren Lernvoraussetzungen und ihrem Lernstand entsprach. Alles spricht dafür, dass in Hamburg Deutsche mit Migrationshintergrund ebenso wie Ausländer durchaus in dem Maße in den anspruchsvolleren Bildungsgängen vertreten sind, wie es ihre Lernvoraussetzungen und ihre tatsächlichen Fachleistungen nahe legen. Streng genommen sind sie nach Maßgabe der Lernvoraussetzungen und Fachleistungen sogar in den aussichtsreicheren Bildungsgängen geringfügig überrepräsentiert.

Tabelle 49 Möglicher oder wahrscheinlicher Schulabschluss der Jugendlichen: tatsächlicher und um Unterschiede in den kognitiven Lernvoraussetzungen sowie in der Fachleistung zu Beginn der Klassenstufe 7 bereinigter Wert, nach Migrationshintergrund<sup>58</sup>

möglicher oder wahrscheinlicher Schulabschluss	Deutsche ohne Migrationshintergrund	Deutsche mit Migrationshintergrund	Ausländer
tatsächlicher Wert	2,42	2,29	2,11
adjustierter Wert	2,34	2,37	2,38
<i>Differenz zwischen adjustiertem Wert und Durchschnitt (= 2,35)</i>	-0,01	0,02	0,03

Dieser Befund steht im Zusammenhang mit anderen, denen zufolge die kognitiven Lernvoraussetzungen und die Fachleistungen in den Migrantengruppen merklich geringer sind als in der Vergleichsgruppe mit deutscher Staatsangehörigkeit. Daher ist es wichtig, die tatsächlichen Leistungsstände in den Gruppen miteinander zu vergleichen. Tabelle 50 benennt die allgemeinen Lernstände (Index Allgemeine Fachleistung) und gibt zugleich die Abstände vom jeweiligen Gruppenmittelwert als Effektstärken an.

<sup>58</sup> Es handelt sich hier wieder um Ergebnisse einer Kovarianzanalyse mit dem dreistufigen Index für den möglichen oder wahrscheinlichen Schulabschluss als abhängiger Variable und den Ergebnissen im *CFT 20* und dem Allgemeinen Fachleistungsindex aus der LAU-7-Untersuchung als Kovariaten.

Tabelle 50 Allgemeine Fachleistung nach Migrationshintergrund und Schulform bzw. Kursniveau (Mittelwerte und Effektstärken  $d$ )<sup>59</sup>

	Deutsche ohne Migrationshintergrund		Deutsche mit Migrationshintergrund		Ausländer	
	$\bar{x}$	Effektstärke $d$	$\bar{x}$	Effektstärke $d$	$\bar{x}$	Effektstärke $d$
Hauptschule	40,3	0,10	39,8	-0,01	38,8	-0,21
Realschule	48,9	0,15	47,8	-0,06	46,2	-0,35
Gymnasium	58,2	0,10	56,9	-0,11	55,0	-0,47
IHR Schule – H	41,3	0,11	41,0	0,04	39,9	-0,20
IHR Schule – R	48,1	0,10	46,7	-0,15	46,8	-0,15
Gesamtschule II	43,4	0,15	42,5	-0,03	40,8	-0,35
Gesamtschule I	50,7	0,12	49,9	-0,03	50,1	-0,41
<i>insgesamt</i>	<i>52,2</i>	<i>0,16</i>	<i>49,9</i>	<i>-0,11</i>	<i>46,7</i>	<i>-0,49</i>

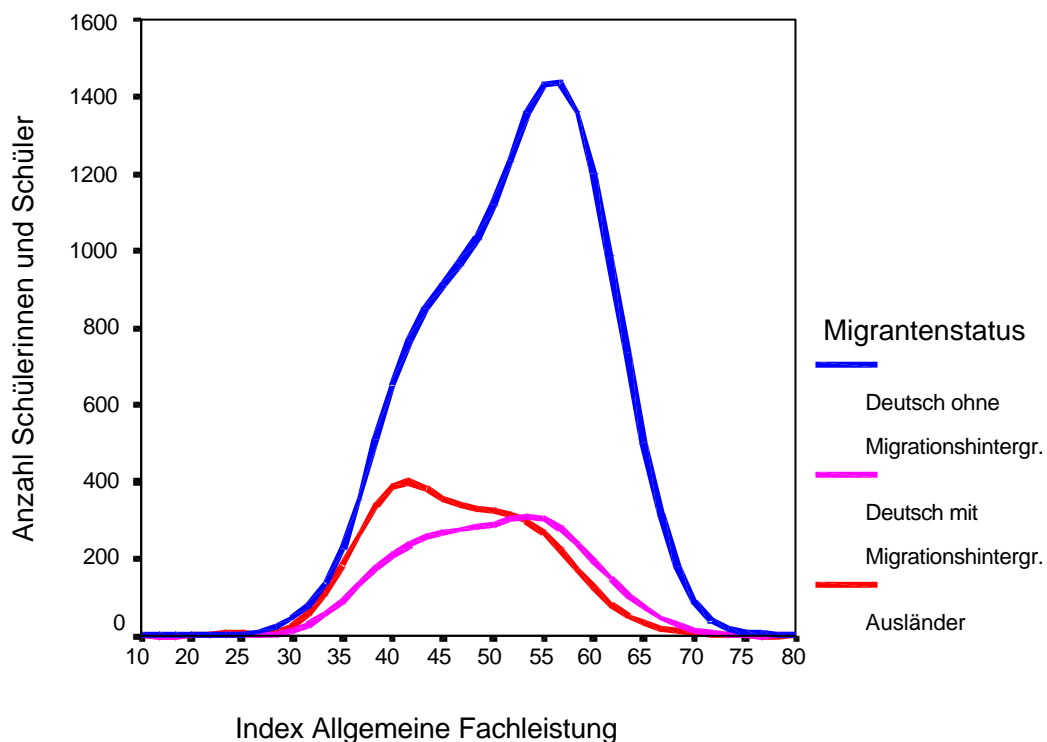
Der Leistungsabstand zwischen den drei Gruppen ist am stärksten im Gesamtsystem ausgeprägt, weil sich hier die nach Schulform bzw. Kursniveau unterschiedlichen Anteile von Schülerinnen und Schülern mit Migrationshintergrund zusätzlich bemerkbar machen. Insgesamt ist deutlich, dass die Schülerinnen und Schüler mit ausländischer Staatsbürgerschaft in der ungünstigsten Position sind, während die Deutschen mit Migrationshintergrund (also Deutsche mit nichtdeutscher Familiensprache, z. B. Aussiedlerkinder, Kinder von Eingebürgerten oder Jugendliche mit einem deutschen Elternteil, auf den die Staatsangehörigkeit zurückgeht) eine Mittelstellung einnehmen, die sich nicht sehr stark vom allgemeinen Mittelwert unterscheidet. Der Abstand der Schülerinnen und Schüler mit ausländischer Staatsangehörigkeit vom Gesamtdurchschnitt erscheint mit knapp einer halben Standardabweichung geringer als in den früheren Untersuchungen. Nimmt man jedoch das dort verwendete Kriterium der Staatsangehörigkeit allein zum Maßstab, so ergibt sich eine Fachleistungsdifferenz zwischen deutschen und ausländischen Schülerinnen bzw. Schülern in Höhe von  $d = 0,59$ , ein Wert, der zwar deutlich geringer ist als zu Beginn der Klassenstufe 5 ( $d = 0,82$ ; vgl. LEHMANN & PEEK 1997, S. 69), jedoch nicht niedriger als zum Ende der Klassenstufe 6 ( $d = 0,55$ ; vgl. LEHMANN, GÄNSFUß & PEEK 1999, S. 137). Der Prozess der Heranführung von Schülerinnen und Schülern ausländischer Nationalität an das allgemeine Fachleistungsniveau ist also in den Klassenstufen 5 und 6 bzw. in den Beob-

<sup>59</sup> Maßgeblich (auch rückwirkend) ist wieder die Schulform- bzw. Kursniveaueugehörigkeit vom Ende der Klassenstufe 8.



achtungsstufen recht effektiv verlaufen, hat sich aber in den Klassenstufen 7 und 8 so nicht fortgesetzt. Um die Leistungsverteilungen in den drei Gruppen anschaulich vor Augen zu führen, werden die Testergebnisse (Index Allgemeine Fachleistung) in der Abbildung 37 grafisch dargestellt. Hiernach sind die deutschen Schülerinnen und Schüler in allen Leistungsgruppen am zahlreichsten vertreten, wobei die auch bei ihnen zu beobachtende Abweichung von einer symmetrischen („normalen“) Leistungsverteilung mit der Unterscheidung zwischen dem Gymnasium und den anderen Schulformen in Verbindung stehen dürfte. Deutlich ist aber auch die Konzentration von Schülerinnen und Schülern ausländischer Nationalität im Bereich, der etwa 10 Punkte (entsprechend einer Standardabweichung) unterhalb des Gesamtmittelwerts liegt. Eine umgekehrte Asymmetrie ist bei den Deutschen mit Migrationshintergrund zu erkennen, die – trotz großer Streuung – sogar am zahlreichsten im leicht überdurchschnittlichen Bereich vertreten sind.

Abbildung 37 Allgemeine Fachleistung nach Migrationshintergrund



Sucht man nach Gründen für die geringeren Fachleistungen der Migranten, so ist zunächst die ungünstigere soziale Stellung zu nennen. Die Schulabschlüsse der Eltern sind bei den Deutschen mit Migrationshintergrund etwas höher als bei den Deutschen ohne Migrationshintergrund ( $d = 0,08$ ); eindeutig niedriger sind sie aber in der Gruppe mit ausländischer Nationalität ( $d = -0,35$ ). Bildungsrelevante Ressourcen sind sowohl bei den Migranten mit deutscher Staatsangehörigkeit ( $d = -0,32$ ) als auch bei Migranten mit auslän-

discher Staatsangehörigkeit ( $d = -0,77$ ) in geringerem Umfang vorhanden. Aber auch dann, wenn man diese beiden Faktoren statistisch berücksichtigt, bleibt ein leistungsmäßiger Rückstand nachweisbar ( $\beta = -0,08$  für deutsche Schülerinnen und Schüler mit Migrationshintergrund;  $\beta = -0,12$  für Schülerinnen und Schüler mit ausländischer Staatsbürgerschaft). Eine gewisse Rolle spielt unter bestimmten Bedingungen auch der Gebrauch der deutschen Sprache außerhalb der Schule. Ihr regelmäßiger oder zumindest gelegentlicher Gebrauch in der Familie leistet – unter Berücksichtigung der beiden anderen Faktoren – in der Gruppe der Deutschen mit Migrationshintergrund, und nur bei ihnen, offenbar einen eigenständigen Beitrag zur Erreichung günstiger Fachleistungen ( $\beta = 0,10$ ).

Weiterhin liegt die Frage nahe, ob es innerhalb der Gruppe der Schülerinnen und Schüler mit ausländischer Staatsangehörigkeit Unterschiede in den Fachleistungen je nach Herkunftsland bzw. Herkunftsregion gibt. Kovarianzanalytisch lassen sich solche Differenzen in der Tat nachweisen, wobei Unterschiede hinsichtlich der Schulformzugehörigkeit, des Bildungsstatus der Eltern und der bildungsrelevanten Ressourcen in Rechnung zu stellen sind. Der maßgebliche Gesamtmittelwert für diese Schülergruppe beträgt 47,52 Punkte auf der Skala für die Allgemeine Fachleistung. Tabelle 51 enthält die Ergebnisse dieser Analyse.

Die Herkunftsländer sind entsprechend ihrem Anteil in der ausländischen Schülerschaft aufgeführt. Dabei werden manifeste Leistungsunterschiede zwischen den Herkunftsländern bzw. -regionen sichtbar, die von 0,32 Standardabweichungen unter dem Durchschnitt für die ausländischen Schülerinnen und Schüler (Türkei) bis zu 0,30 Standardabweichungen darüber (Russische Föderation und „Sonstige“) liegen. Die Position der afghanischen Jugendlichen ähnelt der der türkischen, die der polnischen der der russischen. Auffallend hohe Testergebnisse haben auch die iranischen Schülerinnen und Schüler erzielt. Diejenigen, deren Familien aus einem Land der Europäischen Union stammen (und zwar zu 88,8 Prozent aus den „klassischen Anwerbeländern“ Griechenland, Italien, Portugal und Spanien), liegen ebenfalls über dem Leistungsdurchschnitt der Schülerinnen und Schüler ausländischer Nationalität.

Tabelle 51 Allgemeine Fachleistung ausländischer Schülerinnen und Schüler nach Herkunftsland bzw. Herkunftsregion: tatsächlicher und um Unterschiede in der Verteilung auf die Schulformen bzw. Kursniveaus, im Bildungsstatus der oder des Erziehungsberechtigten sowie in der Verfügbarkeit bildungsrelevanter Ressourcen bereinigter Wert<sup>60</sup>

	<i>Herkunftsland</i>							
	Türkei	Afghanistan	ehem. Jugoslawien	Europ. Union	Iran	Polen	Russland	Sonstige
tatsächlicher Wert	45,00	45,29	47,09	48,30	49,28	49,81	49,88	49,94
Effektstärke <i>d</i>	-0,32	-0,28	-0,05	0,10	0,22	0,29	0,30	0,30
adjustierter Wert	47,02	46,78	47,48	48,85	46,70	48,22	47,38	48,05
<i>Differenz zwischen adjustiertem Wert und Durchschnitt (= 47,52 Punkte)</i>	-0,50	-0,74	-0,04	0,33	-0,82	0,70	-0,14	0,53

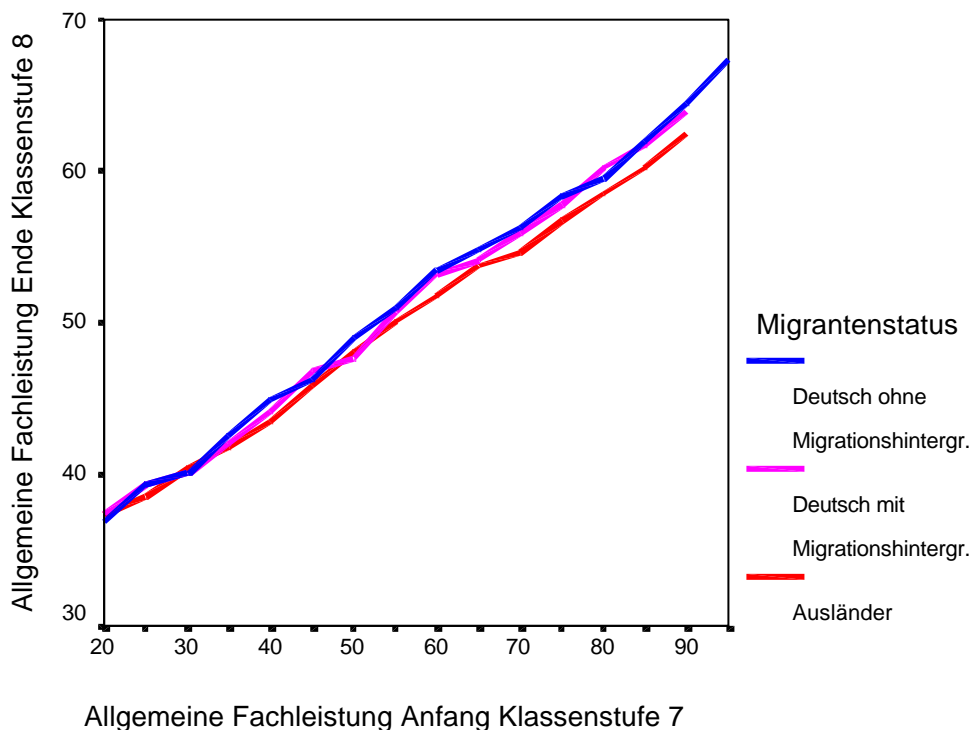
Berücksichtigt man nun die unterschiedliche Verteilung der Herkunftsgruppen auf die Schulformen, den Bildungsstatus der Eltern und die Verfügbarkeit bildungsrelevanter Ressourcen, so zeigt sich ein anderes Bild. Trotz ihres relativ hohen Lernstandes haben die iranischen Schülerinnen und Schüler offenbar mit Schwierigkeiten zu kämpfen; denn angesichts ihrer hohen Bildungsbeteiligung und des in vielen Fällen offenbar recht „bildungsnahe“ sozialen Hintergrunds wären eigentlich noch bessere Fachleistungen zu erwarten. In der afghanischen, auch in der türkischen Gruppe, die beide vergleichsweise geringe Fachleistungen erbracht haben, hätte man nach Berücksichtigung der Bildungsbeteiligung und der Hintergrundfaktoren gleichfalls höhere Lernstände erwartet. In allen drei Fällen müssen also zusätzliche, erfolgshemmende Einflussgrößen wirksam sein, die in der Rechnung nicht berücksichtigt wurden. Auffällig erfolgreich, unter den jeweils gegebenen individuellen Umständen gute Fachleistungen zu erbringen, sind die polnischen Schülerinnen und Schüler, die Gruppe der

<sup>60</sup> Ergebnisse einer Kovarianzanalyse mit dem Index für die allgemeine Fachleistung als abhängiger Variable, einer Einteilung der ausländischen Schülerinnen und Schüler in die sieben meistvertretenen Gruppen sowie eine Restkategorie als unabhängiger Variable und dem möglichen oder wahrscheinlichen Schulabschluss, dem höchsten Bildungsstatus der Eltern oder der bzw. des Erziehungsberechtigten und den bildungsrelevanten Ressourcen als Kovariaten. Die tatsächlichen Werte beruhen auf denselben Fallzahlen ( $N = 1.437$ ) wie die adjustierten. Zur besseren Einschätzung der Bedeutung der Unterschiede in den tatsächlichen Werten sind hier die Effektstärken *d* (berechnet innerhalb der Analysegruppe) ergänzt.

„Sonstigen“ und die Jugendlichen aus Ländern der Europäischen Union.

Von Bedeutung ist es ferner, ob die Lernentwicklung der ausländischen Schülerinnen und Schüler je nach Lernstand zu Beginn der Klassenstufe 7 differenziell anders verlaufen ist als bei den deutschen mit bzw. ohne Migrationshintergrund. Abbildung 38 enthält die entsprechenden Informationen. Wie in dieser Abbildung zu erkennen ist, unterscheiden sich unter dieser Perspektive die deutschen Schülerinnen und Schüler ohne Migrationshintergrund von denen, die die deutsche Staatsbürgerschaft besitzen, aber einen Migrationshintergrund erkennen lassen, kaum und nur unsystematisch; lediglich im vergleichsweise dichter besetzten mittleren Leistungsbereich (d. h., bei einem Lernstand am Anfang der Klassenstufe 7 von etwa 50 Punkten) ist bei letzteren ein leichter Rückstand zu erkennen. Anders verhält es sich bei den ausländischen Schülerinnen und Schülern: Hier sind vor allem im oberen Bereich des Leistungsspektrums bei gleichen Anfangswerten am Ende der Klassenstufe 8 eindeutig niedrigere Fachleistungen als in den beiden Vergleichsgruppen zu verzeichnen.

Abbildung 38 Allgemeine Fachleistung am Ende der Klassenstufe 8, nach Fachleistung zu Anfang der Klassenstufe 7 und Migrationshintergrund



Abschließend soll im Einzelnen untersucht werden, in welchem Maße sich in den Schulfächern Mathematik, Deutsch und Englisch der Einfluss des Migrationshintergrundes bemerkbar gemacht hat, wobei auch hier die Unterscheidung von Deutschen ohne bzw. mit Migrationshintergrund und Schülerinnen und Schülern ausländischer Nationalität maßgeblich sein soll (vgl. Tabelle 52).

Tabelle 52 Fachleistungen in den Bereichen *Mathematik, Deutsch und Englisch* nach Migrationshintergrund (Mittelwerte, Effektstärken  $d$ )

	Deutsche ohne Migrationshintergrund		Deutsche mit Migrationshintergrund		Ausländer	
	$\bar{x}$	Effektstärke $d$	$\bar{x}$	Effektstärke $d$	$\bar{x}$	Effektstärke $d$
Mathematik	128,88	0,13	125,02	-0,08	118,55	-0,42
Deutsch <sup>61</sup>	131,40	0,19	124,32	-0,14	114,89	-0,57
Leseverständnis	127,48	0,18	122,09	-0,11	113,61	-0,56
Sprache	120,66	0,28	114,14	-0,15	105,67	-0,61
Englisch	137,48	0,10	133,62	-0,06	127,38	-0,32
Allgemeine Fachleistung	52,20	0,16	49,93	-0,11	46,68	-0,49

Wie erkennbar wird, sind die Einflüsse des Migrationshintergrundes am deutlichsten im Bereich Deutsch ausgeprägt, und hier im Untertest *Sprachverständnis* noch stärker als im Untertest *Leseverständnis*. Im Untertest *Mathematik* und vor allem im C-Test *Englisch* sind die Unterschiede geringer. Aufschlussreich ist es, diese Werte mit den 1998 zu Beginn der Jahrgangsstufe 7 bzw. am Ende der Klassenstufe 6 gemessenen zu vergleichen (vgl. LEHMANN, GÄNSFUß & PEEK 1999, S. 138), wobei man wieder auf die Unterscheidung zwischen Deutschen und Ausländern nach der Staatsbürgerschaft zurückgreifen muss. Danach haben sich die Unterschiede zwischen diesen beiden Gruppen im Untertest *Mathematik* etwas erhöht (1998:  $d = 0,47$ ; 2000:  $d = 0,51$ ), während sie im C-Test *Englisch* praktisch gleich geblieben sind (1998:  $d = 0,40$ ; 2000:  $d = 0,39$ ). Im Deutschen gibt es ebenfalls kaum eine Veränderung hinsichtlich des Leseverständnisses (1998:  $d = 0,71$ ; 2000:  $d = 0,68$ ), wohl aber im Sprachverständnis. Hier sind die Unterschiede drastisch gestiegen (1998:  $d = 0,34$ ; 2000:  $d = 0,72$ ). Damit werden einerseits Befunde einer internationalen Vergleichsuntersuchung der IEA bestätigt, nach denen der Abstand zwischen vierzehnjährigen Schülerinnen und Schülern deutscher Mutter-

<sup>61</sup> Index, gebildet aus den Untertests *Leseverständnis* und *Sprache*

sprache und anderen bei einem insgesamt schwierigeren Test im Bereich von einer Standardabweichung ( $d = 1,0$ ) lag, was einen internationalen Spitzenwert darstellte (vgl. ELLEY 1992, S. 61). Andererseits sind die vergleichsweise geringen Leistungsrückstände der ausländischen Schülerinnen und Schüler im eingesetzten C-Test *Englisch* auffällig. Es gibt Hinweise darauf, dass sich hier positive Transfereffekte bemerkbar machen könnten; denn bei statistischer Kontrolle der kognitiven Lernvoraussetzungen, des (deutschen) Sprachverständnisses, der Schulformzugehörigkeit und des Bildungsstatus der Eltern bzw. der oder des Erziehungsberechtigten zeigt sich ein eigenständiger, leicht positiver Einfluss ( $\beta = 0,05$ ), der mit einer ausländischen Nationalität verknüpft ist.

Wenn man davon ausgeht, dass mit den in der vorliegenden Untersuchung eingesetzten Tests für spätere Qualifikationsprozesse wesentliche Kompetenzen überprüft werden, dann geben – jenseits der Frage der Chancengerechtigkeit für Migranten – vor allem die bei ihnen festgestellten Leistungsrückstände allen Anlass zur Suche nach Möglichkeiten, diese auf das in Ländern mit vergleichbarer Bevölkerungsstruktur bereits erreichte, also offenbar auch erreichbare Maß zu verringern.

## **5.5 Lern- und schulbezogene Einstellungen der Schülerinnen und Schüler**

Im Schülerfragebogen sollten die Schülerinnen und Schüler Angaben zu ihrer *Schulzufriedenheit*, ihrem *leistungsbezogenen Selbstkonzept* und zu ihrer *Anstrengungsbereitschaft* machen. Bei einem Wertebereich, der bei den zusammengefassten Einzelaussagen von 1 (für positiv formulierte Items: „trifft überhaupt nicht zu“) bis 4 (für positiv formulierte Items: „trifft völlig zu“) reicht, sind die Mittelwerte für alle Skalen und Schulformen eindeutig größer als 2,5 (theoretischer Skalenmittelwert). Insoweit äußern die Schülerinnen und Schüler überwiegend positive lern- und schulbezogene Einstellungen.

### **Schulzufriedenheit**

Die erste Dimension der erfragten Einstellungen drückt sich in der Bewertung von insgesamt 10 Aussagen aus wie z. B.: „Ich gehe gern zur Schule“, „In unserer Schule fühle ich mich wohl“ und „In unserer Schule macht das Lernen Spaß“<sup>62</sup>. Für die unterschiedlichen Schul- bzw. Kursformen sind die Mittelwerte und die Streuungen um den jeweiligen Mittelwert in den einzelnen Vergleichsgruppen sehr ähnlich (vgl. Tabelle 53).

---

<sup>62</sup> Die interne Konsistenz der Skala „Schulzufriedenheit“ liegt bei  $\alpha = 0,84$ .

Tabelle 53 Statistische Kennwerte der Skala „Schulzufriedenheit“

	Mittelwert*	Standardabweichung	N
Hauptschule	2,79	0,58	1.041
Realschule	2,73	0,57	1.818
Gymnasium	2,78	0,56	4.344
IHR Schule – H	2,84	0,57	244
IHR Schule – R	2,83	0,55	285
Gesamtschule II	2,74	0,60	1.306
Gesamtschule I	2,77	0,57	960
<i>insgesamt</i>	<i>2,77</i>	<i>0,57</i>	<i>9.998</i>

\* Ausprägungen hier und in den Tabellen 54 und 55: 1 = sehr gering bis 4 = sehr hoch

Damit liegen die Werte der Hamburger Schülerinnen und Schüler für Schulzufriedenheit zu Beginn der Klassenstufe 9 etwas niedriger als in der LAU-7-Untersuchung (vgl. LEHMANN, GÄNSFUß & PEEK 1999, S. 81ff.), jedoch so hoch, wie sie sich z. B. bei den Neuntklässlerinnen und Neuntklässlern in der QuaSUM-Untersuchung in Brandenburg gezeigt haben (vgl. LEHMANN ET AL. 2000). Der zumindest zeitweilige Rückgang der Schulzufriedenheit in der Adoleszenz ist eine Erscheinung, die auch in anderen Untersuchungen gefunden wurde (vgl. PLATH 1997, S. 102); für die Klassenstufe 8 an Gymnasien wurden für mehrere Bundesländer Werte berichtet, die den hier vorgelegten vergleichbar sind (vgl. BÖHM-KASPER ET AL. 2001, S. 172 f.).

In allen hier unterschiedenen Gruppen äußern die Mädchen mit einem durchschnittlichen Skalenwert von 2,84 ein höheres Maß an Schulzufriedenheit als die Jungen (2,70), wobei diese Unterschiede nicht überall gleich sind: Die Schülerinnen und Schüler in den Kursen I der Gesamtschulen ( $d = 0,33$ ), aus Hauptschulklassen ( $d = 0,32$ ) und aus Gymnasialklassen ( $d = 0,29$ ) zeigen hier die deutlichsten Geschlechterdifferenzen.

Schülerinnen und Schüler ausländischer Nationalität haben eine etwas höhere Schulzufriedenheit ( $d = 0,16$ ) geäußert als deutsche (ohne Migrationshintergrund:  $d = -0,05$ ; mit Migrationshintergrund:  $d = 0,03$ ). Ähnliches hatte sich auch schon in den Untersuchungen zu Beginn der Klassenstufen 5 und 7 gezeigt (vgl. LEHMANN, GÄNSFUß & PEEK 1999, S. 140). Mit der Schulform- oder Kursniveauezugehörigkeit hängt diese Erscheinung anscheinend nur insofern zusammen, als sie an den Gymnasien auffällig gering ausgeprägt ist ( $d = 0,04$ ), dafür aber besonders stark in den Kursen II der Gesamtschulen ( $d = 0,27$ ).

### Leistungsbezogenes Selbstkonzept

Die zweite Dimension der erfragten Einstellungen der Schülerinnen und Schüler bezieht sich auf insgesamt 12 Aussagen wie z. B.: „Ich halte mich für eine erfolgreiche Schülerin bzw. einen erfolgreichen Schüler“ und „Ich weiß genau, wo meine Stärken liegen“<sup>63</sup>. Die Ausprägungen sind in Tabelle 54 aufgeführt. Aus den hier mitgeteilten Werten wird deutlich, dass die Schülerinnen und Schüler der Hauptschulkurse und -klassen sowie des Kursniveaus II der Gesamtschulen über deutlich niedrigere Selbsteinschätzungen verfügen als diejenigen aus Realschulkursen der IHR-Schulen, aus den Kursen I der Gesamtschulen und aus Gymnasien. Zu betonen ist die von HORSTKEMPER (1987) berichtete und auch in der QuaSUM-Untersuchung (LEHMANN ET AL. 2000) festgestellte Differenz im Selbstvertrauen der Jungen und Mädchen.

Tabelle 54 Statistische Kennwerte der Skala „leistungsbezogenes Selbstkonzept“

	Mittelwert*	Standardabweichung	N
Hauptschule	2,87	0,49	1.041
Realschule	2,93	0,48	1.820
Gymnasium	3,02	0,50	4.343
IHR Schule – H	2,67	0,47	244
IHR Schule – R	3,06	0,47	285
Gesamtschule II	2,76	0,49	1.306
Gesamtschule I	3,04	0,47	960
<i>insgesamt</i>	<i>2,95</i>	<i>0,50</i>	<i>9.999</i>

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

Die Jungen sind über die Schulformen und Kursniveaus hinweg selbstbewusster ( $d = 0,18$ ), wobei sich die deutlich ausgeprägtesten Unterschiede bei den Schülerinnen und Schülern zeigen, die in H-Kursen der Integrierten Haupt- und Realschulen ( $d = 0,54$ ), überwiegend in Kursen II der Gesamtschulen ( $d = 0,39$ ) und in Hauptschulen ( $d = 0,32$ ) unterrichtet wurden. Dieses Ergebnis steht in Übereinstimmung mit anderen Befunden (vgl. z. B. HORSTKEMPER 1987), nach denen Jungen in weniger anspruchsvollen Lerngruppen ein unerwartet hohes Selbstbewusstsein haben, das sich offenbar nicht aus schulischen Lernerfolgen speist.

<sup>63</sup> Die interne Konsistenz der Skala liegt bei  $\alpha = 0,87$ .



## Anstrengungsbereitschaft

Die dritte Dimension der erfragten Einstellungen der Schülerinnen und Schüler bezieht sich auf insgesamt fünf Aussagen wie z. B.: „Lernen ist nicht meine Sache“ (negativ) und „Meine Noten werden durch Anstrengung auch nicht besser“ (negativ)<sup>64</sup>. Die entsprechenden Ergebnisse werden in Tabelle 55 mitgeteilt. Dass schulischer Erfolg in starkem Maße (auch) von der persönlichen Anstrengung abhängt, gehört nach diesen Ergebnissen zu den Grundüberzeugungen der Schülerinnen und Schüler, und zwar an den Schulen mit äußerer Leistungsdifferenzierung in den anspruchsvolleren Bildungsgängen jeweils etwas stärker ausgeprägt als in den weniger anspruchsvollen. Insgesamt zeigen sich hier keine nennenswerten geschlechtstypischen Unterschiede, wenn auch in den beiden integrierten Schulformen auf dem unteren Differenzierungsniveau jeweils die Jungen ihre Anstrengungsbereitschaft stärker betonen, während dies auf dem oberen für die Mädchen gilt.

Tabelle 55 Statistische Kennwerte der Skala „Anstrengungsbereitschaft“

	Mittelwert*	Standardabweichung	N
Hauptschule	3,09	0,52	1.041
Realschule	3,10	0,52	1.818
Gymnasium	3,07	0,53	4.344
IHR Schule – H	2,98	0,51	244
IHR Schule – R	3,20	0,49	285
Gesamtschule II	3,00	0,54	1.304
Gesamtschule I	3,13	0,53	960
<i>insgesamt</i>	<i>3,08</i>	<i>0,53</i>	<i>9.996</i>

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

Die erfragten Aspekte der lern- und schulbezogenen Einstellungen der Schülerinnen und Schüler sind nicht voneinander unabhängig. Hohe Werte auf einer Skala gehen in der Regel mit relativ hohen auf einer anderen einher (vgl. Tabelle 56). Dabei zeigt sich eine über alle Schulformen hin stabile Struktur: Die Schülerinnen und Schüler mit einem ausgeprägten leistungsbezogenen Selbstkonzept betonen die eigene Anstrengungsbereitschaft; auch sind diejenigen mit hohen Werten auf diesen beiden Skalen zufriedener mit der Schule.

<sup>64</sup> Die interne Konsistenz der Skala liegt bei  $\alpha = 0,65$ , was durch die Kürze der Skala zumindest mitbedingt ist.

Tabelle 56 Lern- und schulbezogene Einstellungen bzw. Selbsteinschätzungen: Interkorrelationen

	leistungsbezogenes Selbstkonzept	Anstrengungsbereitschaft
Schulzufriedenheit	0,28	0,40
leistungsbezogenes Selbstkonzept	---	0,68

In der Tabelle 57 werden die lern- und schulbezogenen Einstellungen zu den allgemeinen Fachleistungen in Beziehung gesetzt. Es zeigen sich – insgesamt und schulform- bzw. kursniveaubezogen – nur geringe Zusammenhänge zwischen der allgemeinen Schulzufriedenheit der Schülerinnen und Schüler und ihren allgemeinen Fachleistungen.

Tabelle 57 Lern- und schulbezogene Einstellungen bzw. Selbsteinschätzungen: Korrelationen mit der allgemeinen Fachleistung, nach Schulform bzw. Kursniveau

	Schul- zufriedenheit	leistungsbezogenes Selbstkonzept	Anstrengungs- bereitschaft
Hauptschule	0,03	0,14	0,11
Realschule	0,02	0,18	0,06
Gymnasium	0,09	0,31	0,15
IHR Schule – H	-0,08	0,11	-0,00
IHR Schule – R	-0,05	0,24	0,07
Gesamtschule II	-0,02	0,15	0,07
Gesamtschule I	0,08	0,30	0,16
<i>insgesamt</i>	<i>0,04</i>	<i>0,28</i>	<i>0,08</i>

Hingegen sind die Zusammenhänge zwischen einem positiven leistungsbezogenen Selbstkonzept und den allgemeinen Fachleistungen deutlich enger und bei den Schülerinnen und Schülern, die in Gymnasialklassen oder überwiegend in Kursen I der Gesamtschulen unterrichtet werden, am stärksten ausgeprägt. Gleichzeitig betonen in diesen beiden Schülergruppen die leistungsstärkeren Schülerinnen und Schüler auch verstärkt die eigene Anstrengungsbereitschaft.

## 6 Schlussbetrachtung

Der Ertrag der Untersuchung *Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9* umfasst eine Reihe von Erkenntnissen, die nicht sämtlich neu sind und gleichwohl Ansatzpunkte für Verbesserungen der Bildungsangebote von Hamburger Sekundarschulen bieten können. Einige Befunde waren – zumindest in dieser Weise – nicht zu erwarten und bieten verstärkt Anlass, den damit aufgeworfenen Fragen weiter nachzugehen. Nachdem in den vorangegangenen Kapiteln von den Ergebnissen dieser dritten Erhebung der ursprünglich so genannten „Hamburger Lernausgangslagenuntersuchung“ (LAU) oftmals mit starker Konzentration auf Details berichtet worden ist, sollen nun einige zentrale Befunde noch einmal in gedrängter Form dargelegt werden.

1. Das Hamburger Sekundarschulwesen ist, gemessen an den Fachleistungen der Schülerinnen und Schüler, auch am Ende der Klassenstufe 8 noch durch ein hohes Maß an potenzieller Durchlässigkeit charakterisiert. In den Realschulklassen und auf den oberen Differenzierungsniveaus der integrierten Schulformen (Integrierte Haupt- und Realschule und Gesamtschule) befinden sich viele, deren Fachleistungen denen eines großen Anteils der Gymnasiastinnen und Gymnasiasten nicht nachstehen. Obwohl ein fundiertes Urteil zur Bedeutung dieses Überschneidungsbereichs für die im Hamburger Bildungssystem vorgesehenen Möglichkeiten zum Übertritt in die gymnasiale Oberstufe im Anschluss an die Klassenstufe 10 jetzt noch nicht gefällt werden kann, liegen hier offenkundig Chancen, junge Menschen für Bildungsgänge zu gewinnen, die ihren Fähigkeiten entsprechen.
2. Dass es unter der Perspektive der Fachleistungen beträchtliche Überschneidungsbereiche zwischen den Schulformen bzw. Kursniveaus gibt, ist auch aus anderen Untersuchungen bekannt. Auch wenn eine unzweideutige Zuordnung von Fachleistungen zu Schulformen bzw. Kursniveaus wohl unerreichbar und auch nicht in jedem Einzelfall begründet und wünschenswert erscheint, bleibt die Frage nach der Chancengerechtigkeit im Bildungssystem unabweisbar. Die vorliegende Untersuchung hat erneut gravierende Hinweise darauf ergeben, dass an den Hamburger Schulen Bildungsbarrieren für diejenigen fortbestehen, die aus bildungsfernen Schichten stammen, während Schülerinnen und Schüler, deren Eltern einen hohen Schulabschluss erworben haben, hinsichtlich ihres schulischen Fortkommens nach wie vor unbegründete Vorteile genießen.
3. Ein weiterer Aspekt lässt die leistungsmäßigen Überschneidungen zwischen den verschiedenen Hamburger Schulformen bzw. Kursniveaus eher pro-

blematisch erscheinen. Von einem gegliederten Schulwesen ist zu erwarten, dass sich das Leistungsspektrum (die sog. „Varianz“) über die Klassenstufen hinweg immer mehr erweitert, weil in den verschiedenen Bildungsgängen unterschiedlich schnell gelernt werden soll. Untersuchungen in anderen Bundesländern belegen diesen Effekt (vgl. z. B. SCHNABEL, KÖLLER & BAUMERT 2001, S. 213). Eine allmähliche Verringerung der Überschneidungsbereiche wäre die Folge. In Hamburg scheint sich aber in wichtigen Bereichen (*Mathematik, Deutsch: Leseverständnis* und *Sprache*) die Leistungsstreuung eher zu vermindern, während sie sich im Bereich *Englisch* zwar nicht verringert, aber eben auch nicht zunimmt. Innerhalb der Schulformen bzw. Kursniveaus zeigt sich ein ähnliches Bild. Generell ist die Lernentwicklung an den Hamburger Gymnasien im Durchschnitt gerade *nicht* durch größere Lernfortschritte gekennzeichnet als diejenige in anderen Schulformen und ihren Differenzierungsniveaus. Diese Besonderheit wird nicht unabhängig sein von dem ungewöhnlich hohen Gymnasialanteil in Hamburg (36,8 Prozent zu Beginn der Klassenstufe 9).

4. Innerhalb einer Schulform oder eines Kursniveaus gibt es verschiedene Möglichkeiten, eine Erweiterung des Leistungsspektrums zu verhindern oder dieses sogar zu verringern. Dazu gehören die Verfügung einer Klassenwiederholung oder eines Schulformwechsels, ggf. auch die Abschulung nach Beendigung der Vollzeitschulpflicht. Die Anlage der vorliegenden Untersuchung hat es erlaubt, den Einfluss solcher Maßnahmen abzuschätzen, ohne ihn quantitativ genau bestimmen zu können. Hervorzuheben ist jedenfalls, dass die traditionellen Schulformen vergleichsweise häufig davon Gebrauch machen und so ein etwas günstigeres Leistungsbild bieten, als es ihrer ‚wahren Produktivität‘ entspricht. Die Gesamtschulen haben in dieser Hinsicht eine Sonderstellung, weil Klassenwiederholungen und Schulformwechsel hier besonders selten sind.
5. Einer der Gründe, weshalb die Lernentwicklung in den Klassenstufen 7 und 8 an Hamburger Schulen in vielen Fällen nicht so erfolgreich verläuft, wie man es vielleicht erwarten würde, liegt in einer deutlichen Verlangsamung der Lernfortschritte bei den Jungen, verglichen mit den Mädchen. Diese betrifft alle untersuchten Fachleistungen, soweit hier Längsschnittvergleiche möglich waren. So hat sich der ursprüngliche Vorsprung der Jungen im Fach Mathematik in den Klassenstufen 7 und 8 eindeutig verringert, während die Leistungsvorteile für die Mädchen im sprachlichen Bereich im gleichen Zeitraum ebenso eindeutig gewachsen sind. Letztere sind zudem erheblich größer als die Geschlechterdifferenz beim Untertest *Mathematik*. Zusammenhänge mit den erhobenen schulbezogenen Einstellungen sind

möglich: Allgemein ist die Schulzufriedenheit der Mädchen größer, während das leistungsbezogene Selbstkonzept, das besonders eng an die Mathematikleistung gekoppelt ist, bei den Jungen stärker ausgeprägt ist.

6. Berücksichtigt man nicht nur die Staatsangehörigkeit der Schülerinnen und Schüler, sondern auch die in der Familie gesprochene Mutter- oder Herkunftssprache, so wird bei etwa einem Drittel der Hamburger Schülerschaft der Klassenstufe 9 ein Migrationshintergrund erkennbar. Wie auch die früheren Untersuchungen gezeigt haben, kann man in Hamburg von keiner Benachteiligung dieser Schülerinnen und Schüler sprechen, jedenfalls dann nicht, wenn man deren kognitiven Lernvoraussetzungen und die tatsächlich erreichten Fachleistungen in Rechnung stellt. Auch äußern die Schülerinnen und Schüler mit ausländischer Staatsangehörigkeit ein höheres Maß an Schulzufriedenheit als diejenigen deutscher Nationalität. Gleichwohl lassen sich (nach ihrer Staatsangehörigkeit definierte) Gruppen erkennen, denen der Anschluss an den Leistungsstand der deutschen Schülerinnen und Schüler offenbar schwerer fällt als anderen, so wie umgekehrt Schülerinnen und Schüler mittel- und osteuropäischer Herkunft bemerkenswert erfolgreich sind.

Der Reichtum der in den vorliegenden Datensätzen enthaltenen Informationen verlangt und ermöglicht zusätzliche Untersuchungen, die weit über die hier berichteten Befunde hinausgehen. Es muss künftigen Veröffentlichungen vorbehalten bleiben, die dann gewonnenen Erkenntnisse zur Diskussion zu stellen.

## Literaturverzeichnis

- BAUMERT, J. / BOS, W. / LEHMANN, R.H. (Hrsg.) (2000): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band I: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen.
- BEHÖRDE FÜR SCHULE, JUGEND UND BERUFSBILDUNG, AMT FÜR SCHULE, HAMBURG (Hrsg.) (1998): Der Hamburger Schulleistungstest für sechste und siebte Klassen – SL-HAM 6/7. Hamburg.
- BEHÖRDE FÜR SCHULE, JUGEND UND BERUFSBILDUNG, AMT FÜR SCHULE, HAMBURG (Hrsg.) (2000): Hamburger Schulleistungstest für achte und neunte Klassen – SL-HAM 8/9. Hamburg.
- BÖHM-KASPER, O. / BOS, W. / KÖRNER, S. / WEISHAUPT, H. (2001): Sind 12 Schuljahre stressiger? Belastung und Beanspruchung von Lehrern und Schülern am Gymnasium. Weinheim und München.
- ELLEY, W.B. (1992): How in the World Do Students Read? IEA Study of Reading Literacy. Hamburg.
- FISCHER, G.H. / MOLENAAR, I.H. (1995): Rasch Models – Foundations, Recent Developments, and Applications. New York.
- GROTJAHN, R. (1995): Der C-Test. State of the Art. In: Zeitschrift für Fremdsprachenforschung. 6. Jahrgang. Heft 2. S. 37-60.
- HAMBLETON, R.K. / SWAMINATHAN, H. / ROGERS, H.J. (Hrsg.) (1991): Fundamentals of Item Response Theory. Newbury Park, CA.
- HORSTKEMPER, M. (1987): Schule, Geschlecht und Selbstvertrauen. Eine Längsschnittstudie über Mädchensozialisation in der Schule. Weinheim und München.
- KLIEME, E. / STANAT, P. / ARTELT, C. (2001): Fächerübergreifende Kompetenzen: Konzepte und Indikatoren. In: F.E. Weinert (Hrsg.). Leistungsmessungen in Schulen. Weinheim und Basel. S. 203-218.
- LEHMANN, R.H. (2000): Anregungen für bildungspolitische Konsequenzen der empirischen Forschung. In: W. Stark / T. Fitzner / C. Schubert (Hrsg.): Von der Alphabetisierung zur Leseförderung. Eine Fachtagung. Stuttgart. S. 153-165.

- LEHMANN, R.H. (2001): Systembeobachtung: Lernausgangslage und Lernentwicklung in der Sekundarstufe I. In: K.J. Tillmann / W. Vollstädt (Hrsg.): Politikberatung durch Bildungsforschung. Das Beispiel: Schulentwicklung in Hamburg. Opladen. S. 61-74.
- LEHMANN, R.H. / GÄNSFUß, R. / PEEK, R. (1999): Aspekte der Lernauslage und der Lernentwicklung von Schülerinnen und Schülern an Hamburger Schulen – Klassenstufe 7. Bericht über die Untersuchung im September 1998. Hamburg.
- LEHMANN, R.H. / HUSFELDT, V. / PEEK, R. (2001): Lernstände und Lernentwicklungen im Fach Mathematik – Ergebnisse der Hamburger Untersuchung (LAU) in den Klassenstufen 5 und 6. In: G. Kaiser / N. Koche / D. Lind / W. Zillmer (Hrsg.): Leistungsvergleiche im Mathematikunterricht – ein Überblick über aktuelle nationale Studien. Hildesheim und Berlin. S. 29-52.
- 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. Hamburg.
- LEHMANN, R.H. / PEEK, R. / PIEPER, I. / VON STRITZKY, R. (1995): Leseverständnis und Lesegewohnheiten deutscher Schülerinnen und Schüler. Weinheim und Basel.
- LEHMANN, R.H. / PEEK, R. / GÄNSFUß, R. / LUTKAT, S. / MÜCKE, S. / BARTH, I. (2000): QuaSUM. Qualitätsuntersuchung an Schulen zum Unterricht in Mathematik. Ergebnisse einer repräsentativen Untersuchung im Land Brandenburg (Reihe Schulforschung in Brandenburg, Heft 1). Potsdam.
- MARTIN, M.O. / KELLY, D.L. (Hrsg.) (1997): Third International Mathematics and Science Study. Technical Report. Vol. II: Implementation and Analysis. Primary and Middle School Years (Population 1 and Population 2). Chestnut Hill.
- MAY, P. (1994): HSP zur Erfassung der grundlegenden Rechtschreibstrategien. Handbuch für die Hamburger Schreibprobe (HSP). Hamburg.
- MIETZEL, G. / WILLENBERG, H. (2000): Hamburger Schulleistungstest für 4. und 5. Klassen (HST 4/5). Göttingen.

- PLATH, M. (1997): Wie wohl fühlen sich Schüler? Ein Beitrag zur Diskussion um die innere Schulreform. In: P. Zedler / H. Weishaupt (Hrsg.): Kontinuität und Wandel. Weinheim. S. 91-114.
- POWELL, J. / WAGNER, S. (2001): Daten und Fakten zu Migrantenjugendlichen an Sonderschulen in der Bundesrepublik Deutschland. Selbstständige Nachwuchsgruppe Working Paper 1/2001. Berlin.
- RAATZ, U. / KLEIN-BRALEY, C. (1983): Ein neuer Ansatz zur Messung der Sprachleistung. Der C-Test: Theorie und Praxis. In: R. Horn / K. Ingenkamp / R.S. Jäger (Hrsg.): Tests und Trends 3. Jahrbuch der Pädagogischen Diagnostik. Weinheim und Basel. S. 107-137.
- SCHNABEL, K. / KÖLLER, O. / BAUMERT, J. (2001): Stimmungsindikatoren und Leistungsängstlichkeit als Prädiktoren für schulische Leistungsentwicklung: Eine längsschnittliche Vergleichsstudie. In: C. Finkbeiner / G.W. Schnaitmann (Hrsg.): Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik. Donauwörth. S. 208-221.
- WEIB, R. (1997): Grundintelligenztest Skala 2 (CFT 20) mit Wortschatztest (WS) und Zahlenfolgentest (ZF). 4. überarbeitete Auflage. Göttingen.



## Glossar

### Alpha

Alpha (genauer: „Cronbachs Alpha“) ist eine Kennzahl für die interne Konsistenz einer  $\Rightarrow$  Skala. Werden inhaltlich zusammenpassende Aufgaben oder Fragen zu einer Skala zusammengefasst, z. B. 10 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

$\Rightarrow$  Mittelwert.

### Beta(-Gewicht)

$\Rightarrow$  Regressionsanalyse.

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

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

### erklärte Varianz $\Rightarrow$ Kovarianzanalyse,

$\Rightarrow$  Regressionsanalyse,  $\Rightarrow$  Varianzanalyse

### *Eta*<sup>2</sup>, Bestimmtheitsmaß *Eta*<sup>2</sup>

ist die Maßzahl für die erklärte Varianz bei der  $\Rightarrow$  Kovarianzanalyse und der  $\Rightarrow$  Varianzanalyse. *Eta*<sup>2</sup> 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 *Eta*<sup>2</sup> für ein Leistungsmerkmal Auskunft darüber, welchen Anteil Schulformunterschiede (Mittelwertdifferenzen) an den Leistungsunterschieden zwischen den Schülerinnen und Schülern überhaupt haben.

### interne Konsistenz

$\Rightarrow$  Alpha.

### intervallskalierte Variable

Wenn Gleichheit der Abstände zwischen zwei benachbarten Werten einer

Variable angenommen werden darf, spricht man von einer intervallskalierten Variable. Ist dies nicht der Fall, liegt eine nominaleskalierte (keine Rangfolge der Werte) oder eine ordinalskalierte (Rangfolge der Werte, aber keine Abstandsgleichheit) Variable vor.

### **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.

### **Kovarianzanalyse**

Die Kovarianzanalyse ist eine Mischform aus ⇒ Regressionsanalyse und ⇒ Varianzanalyse. Die abhängige Variable wird durch unabhängige Variablen erklärt, von denen mindestens eine wie bei der Varianzanalyse nominaleskaliert ist (der sog. „Faktor“), andere dagegen wie bei der Regressionsanalyse intervallskaliert (Kovariaten). Die Maßzahl für die insgesamt erklärte Varianz ist ⇒  $R^2$ ; für den Faktor allein kann auch ein ⇒  $\text{Eta}^2$  bestimmt werden.

### **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  $\Rightarrow$  intervallskalierte Daten vorliegen. Er ist die Summe der Einzelwerte aller Fälle dividiert durch die Fallzahl.

## **multiple Regressionsanalyse**

$\Rightarrow$  Regressionsanalyse.

## **multiple $R$ , multiple $R^2$**

$\Rightarrow$  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 Auswerterers sind.

## **Prädiktor(variable)**

unabhängige Variable,  $\Rightarrow$  Regressionsanalyse.

$r$  Abkürzung für Pearsons  $r$ ,  $\Rightarrow$  Korrelation.

## **Regressionsanalyse**

Die (multiple) Regressionsanalyse ist ein Analyseverfahren, das den Zusammenhang zwischen einer  $\Rightarrow$  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  $\Rightarrow$  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  $\Rightarrow$  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  $\Rightarrow$  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.  $\Rightarrow$  Cronbachs Alpha.

### **Skala**

1. Kurzbezeichnung für die Ausprägungen einer Einschätzskala (Ratingskala). Bei der Vorgabe der Antwortmöglichkeiten von 1 = „trifft völlig zu“ bis 4 = „trifft überhaupt nicht zu“ im Schülerfragebogen spricht man z. B. von einer vierstufigen Skala.
2. Inhaltlich zusammenfassende 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 19 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**

$\Rightarrow$  Standardabweichung.

### **Validität**

Validität ist ein Gütekriterium 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 Aufgaben der fachbezogenen Leistungstests mit den Curricula der jeweiligen Schulform und Klassenstufe validieren (curriculare Validität).

### **Variable**

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

### **Varianz**

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

### **Varianzanalyse**

Dieses Analyseverfahren ist mit der  $\Rightarrow$  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  $\Rightarrow$  intervallskaliert zu sein braucht. Die Maßzahl für die erklärte Varianz heißt  $\Rightarrow$   $\eta^2$ .