



Petra Stanat
Hans Anand Pant
Katrín Böhme
Dirk Richter
(Hrsg.)



Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik

Ergebnisse des IQB-Ländervergleichs 2011

WAXMANN

Petra Stanat
Hans Anand Pant
Katrín Böhme
Dirk Richter (Hrsg.)

Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik

Ergebnisse des IQB-Ländervergleichs 2011



Waxmann 2012
Münster/New York/München/Berlin

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-8309-2777-8

© 2012, Waxmann Verlag GmbH,
Postfach 8603, D-48046 Münster

Waxmann Publishing Co.,
P. O. Box 1318, New York, NY 10028, USA.

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Christian Aeverbeck, Münster

Titelfoto: © Jacek Chabraszewski – Fotolia.com

Satz: Stoddart Satz- und Layoutservice, Münster

Druck: Hubert & Co., Göttingen

Gedruckt auf alterungsbeständigem Papier,
säurefrei gemäß ISO 9706



Printed in Germany

Alle Rechte vorbehalten. Nachdruck, auch auszugsweise, verboten.

Kein Teil dieses Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Inhalt

Vorwort des Präsidenten der Kultusministerkonferenz.....	9
-----------------------------------------------------------------	----------

Kapitel 1

Die länderübergreifenden Bildungsstandards in Deutschland.....	11
-----------------------------------------------------------------------	-----------

Katrin Böhme, Dirk Richter, Petra Stanat, Hans Anand Pant und Olaf Köller

1.1 Hintergrund der Einführung länderübergreifender Bildungsstandards.....	11
1.2 Charakterisierung der länderübergreifenden Bildungsstandards.....	13
1.3 Entwicklung von Testverfahren zur Überprüfung der länderübergreifenden Bildungsstandards im Primarbereich	15
Literatur.....	17

Kapitel 2

Die im Ländervergleich 2011 in den Fächern Deutsch und Mathematik untersuchten Kompetenzen	19
-------------------------------------------------------------------------------------------------------------	-----------

2.1 Beschreibung der im Fach Deutsch untersuchten Kompetenzen	19
Katrin Böhme und Albert Bremerich-Vos	
2.2 Beschreibung der im Fach Mathematik untersuchten Kompetenzen	34
Alexander Roppelt und Kristina Reiss	
Literatur.....	44

Kapitel 3

Kompetenzstufenmodelle für den Primarbereich	49
-----------------------------------------------------------	-----------

3.1 Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen	49
Hans Anand Pant, Katrin Böhme und Olaf Köller	
3.2 Kompetenzstufenmodelle im Fach Deutsch.....	56
Albert Bremerich-Vos, Katrin Böhme, Michael Krelle, Sebastian Weirich und Olaf Köller	
3.3 Kompetenzstufenmodelle im Fach Mathematik.....	72
Kristina Reiss, Alexander Roppelt, Nicole Haag, Hans Anand Pant und Olaf Köller	
Literatur.....	83

Kapitel 4 **Anlage und Durchführung des Ländervergleichs 85**

Dirk Richter, Maria Engelbert, Katrin Böhme, Nicole Haag, Jasmin Hannighofer, Heino Reimers,
Alexander Roppelt, Sebastian Weirich, Hans Anand Pant und Petra Stanat

4.1	Testinstrumente	85
4.2	Fragebogeninstrumente	87
4.3	Testdesign	89
4.4	Testablauf	90
4.5	Definition der Zielpopulation und Stichprobenziehung	91
4.6	Realisierte Stichprobe	93
4.7	Aufbereitung und Analyse der Daten	97
4.8	Qualitätssicherung	98
4.9	Beteiligte Institutionen und Personen	100
	Literatur	102

Kapitel 5 **Ländervergleich der in den Fächern Deutsch und Mathematik** **erzielten Kompetenzstände 103**

5.1	Der Ländervergleich im Fach Deutsch	103
	Katrin Böhme und Sebastian Weirich	
5.2	Der Ländervergleich im Fach Mathematik	117
	Nicole Haag und Alexander Roppelt	
5.3	Zusammenfassung zentraler Ergebnisse des Ländervergleichs in den Fächern Deutsch und Mathematik	128
	Katrin Böhme und Nicole Haag	
	Literatur	129

Kapitel 6 **Der Blick in die Länder 131**

Petra Stanat, Hans Anand Pant, Dirk Richter, Katrin Böhme, Maria Engelbert,
Nicole Haag, Jasmin Hannighofer, Josefine Prengel, Heino Reimers,
Alexander Roppelt und Sebastian Weirich

6.1	Baden-Württemberg	136
6.2	Bayern	138
6.3	Berlin	140
6.4	Brandenburg	143
6.5	Bremen	145
6.6	Hamburg	147
6.7	Hessen	150
6.8	Mecklenburg-Vorpommern	152
6.9	Niedersachsen	154
6.10	Nordrhein-Westfalen	156
6.11	Rheinland-Pfalz	158
6.12	Saarland	160
6.13	Sachsen	162

6.14 Sachsen-Anhalt	164
6.15 Schleswig-Holstein	167
6.16 Thüringen	169
6.17 Orthografiekompetenzen in der bundesweit repräsentativen Stichprobe.....	171
Literatur.....	172

Kapitel 7

Geschlechtsbezogene Disparitäten 173

Katrin Böhme und Alexander Roppelt

7.1 Geschlechtsbezogene Unterschiede im Bildungsbereich	173
7.2 Geschlechtsbezogene Kompetenzunterschiede im Ländervergleich 2011.....	180
7.3 Zusammenschau der Befunde und Fazit	187
Literatur.....	188

Kapitel 8

Soziale Disparitäten 191

Dirk Richter, Poldi Kuhl und Hans Anand Pant

8.1 Einführung	191
8.2 Indikatoren sozialer Disparitäten.....	193
8.3 Länderdifferenzen im sozioökonomischen Status.....	194
8.4 Kompetenzniveaus nach EGP-Klassen	201
8.5 Zusammenfassung und Diskussion	204
Literatur.....	206

Kapitel 9

Zuwanderungsbezogene Disparitäten 209

Nicole Haag, Katrin Böhme und Petra Stanat

9.1 Erfassung des Zuwanderungshintergrunds.....	211
9.2 Viertklässlerinnen und Viertklässler mit Zuwanderungshintergrund in den Ländern.....	212
9.3 Befunde für Kinder mit Zuwanderungshintergrund unterschiedlicher Herkunft in Deutschland	224
9.4 Zusammenfassung und Diskussion	231
Literatur.....	233

Kapitel 10

Aspekte der Aus- und Fortbildung von Lehrkräften in der Primarstufe 237

Dirk Richter, Poldi Kuhl, Heino Reimers und Hans Anand Pant

10.1 Demografische Angaben der Lehrkräfte im Ländervergleich.....	238
10.2 Fachfremdes Unterrichten im Primarbereich	239
10.3 Fortbildung von Lehrkräften	241
10.4 Zusammenfassung	248
Literatur.....	249

Kapitel 11

Sprach- und Leseförderung..... 251

Petra Stanat, Sebastian Weirich und Susanne Radmann

11.1 Analyse von Angeboten der Sprach- und Leseförderung und ihrer Nutzung	253
11.2 Diagnose von Sprachförderbedarf.....	255
11.3 Breite des Angebots von Sprach- und Leseförderung	257
11.4 Angebot systematischer Sprachförderung.....	261
11.5 Förderquoten	264
11.6 Differenzielle Förderquoten	267
11.7 Zusammenfassung und Diskussion	271
Literatur.....	274

Kapitel 12

Testdesign und Auswertung des Ländervergleichs: technische Grundlagen 277

Sebastian Weirich, Nicole Haag und Alexander Roppelt

12.1 Testdesign	277
12.2 Skalierung.....	281
Literatur.....	289

Kapitel 13

IQB-Ländervergleich 2011: Zusammenfassung und Einordnung der Befunde 291

Petra Stanat, Hans Anand Pant, Dirk Richter und Katrin Böhme

13.1 Mittleres Niveau und Streuung der erreichten Kompetenzen.....	292
13.2 Geschlechtsbezogene, soziale und zuwanderungsbezogene Disparitäten.....	294
13.3 Ausgewählte Bedingungen von Lehr-Lern-Prozessen: Aus- und Fortbildung von Lehrkräften und Maßnahmen der Sprach- und Leseförderung	295
13.4 Einordnung der Befunde anhand der Ergebnisse früherer Ländervergleichsstudien	297
Literatur.....	300

Dank

Viele haben zum Entstehen dieses Berichts beigetragen. Ein herzlicher Dank der Herausgeberinnen und Herausgeber gilt Juliane Dörk, Franziska Förster, Andrea Prater und Katharina Prestel für ihre Unterstützung bei der Zusammenstellung der Testmaterialien sowie Susanne Radmann und Jana Vogel für die Mitarbeit an den Analysen und deren Aufbereitung. Danken möchten wir auch Tobias Lewerenz für die Arbeit an der Dokumentation der Erhebungsinstrumente sowie Martha Mallwitz für die Mitarbeit am Bericht zum Qualitätsmonitoring. Ein besonderer Dank gilt ferner Josefine Pregel für die Erstellung der Grafiken und für redaktionelle Arbeiten an den Manuskripten. Für die Unterstützung bei den statistischen Analysen bedanken sich die Herausgeber bei Martin Hecht, Dr. Claudia Pöhlmann und Stefan Schipolowski. Auch allen Kolleginnen und Kollegen am IQB gilt ein herzlicher Dank, dass sie in den arbeitsintensiven Wochen der Berichtlegung die am Ländervergleich Beteiligten von anderen Arbeiten entlastet haben. Bei Beate Plugge und ihren Kolleginnen und Kollegen vom Waxmann Verlag bedanken wir uns für die produktive, freundliche und geduldige Zusammenarbeit bei der Erstellung dieses Berichts.

Vorwort des Präsidenten der Kultusministerkonferenz

Mit dem Ländervergleich für den Primarbereich veröffentlicht das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) zum zweiten Mal einen Bericht zur Überprüfung der Bildungsstandards, die seit 2004 eingeführt wurden. Sie bilden neben den internationalen Schulleistungsstudien PISA, IGLU und TIMSS einen bundesweiten Bezugsrahmen für das Bildungsmonitoring. Nach dem Ländervergleich 2009 zu den sprachlichen Kompetenzen in der Sekundarstufe I wird nun erstmalig ein Bericht über das Erreichen der fachbezogenen Kompetenzen zum Abschluss der Primarstufe vorgestellt. Der nächste Ländervergleich (Sekundarstufe I) wird 2013 veröffentlicht.

Bundesweit geltende Bildungsstandards gibt es gegenwärtig für die Fächer

- Deutsch und Mathematik für den Primarbereich (Jahrgangsstufe 4),
- Deutsch, Mathematik und erste Fremdsprache (Englisch, Französisch) für den Hauptschulabschluss (Jahrgangsstufe 9),
- Deutsch, Mathematik und erste Fremdsprache (Englisch, Französisch) für den Mittleren Schulabschluss (Jahrgangsstufe 10) sowie
- Biologie, Chemie und Physik für den Mittleren Schulabschluss (Jahrgangsstufe 10).

Das IQB hat den Auftrag, dazu empirisch fundierte Kompetenzstufenmodelle zu entwickeln, die Auskunft darüber geben, wie groß die Anteile der Schülerinnen und Schüler sind, die die Standards erreichen oder verfehlen.

Die zentrale Überprüfung des Erreichens der Bildungsstandards in Ländervergleichen hat mehrere Ziele: Sie soll

- für mehr Transparenz und eine größere Objektivität sorgen,
- einheitliche Leistungen in allen Ländern sicherstellen,
- die Rechenschaftslegung im Bildungsbereich verbessern und
- Anhaltspunkte liefern, an welchen Stellen das Bildungswesen und der Unterricht weiterzuentwickeln sind.

Die vom IQB durchgeführten Ländervergleiche lösen die Ländervergleiche im Rahmen von PISA (PISA-E) und PIRLS/IGLU (IGLU-E) ab. Im Gegensatz zu diesen Ergänzungen internationaler Studien orientieren sich die IQB-Ländervergleiche stärker an der Lehrplanwirklichkeit in den Schulen der Länder.

Die Kultusministerkonferenz hat entschieden, dass die Ergebnisse der Ländervergleiche stets *vor* den Ergebnissen der internationalen Schulleistungsstudien veröffentlicht werden, um der Politik und den Schulen eine rasche Rückmeldung

geben zu können. Auf diese Weise wird auch deutlich gemacht, dass die für alle Länder geltenden Bildungsstandards der Kultusministerkonferenz der entscheidende Referenzrahmen für unsere Schulen sind.

Die vorliegenden empirischen Daten sollen maßgeblich dazu beitragen, den Unterricht in allen Ländern weiterzuentwickeln. Dabei muss insbesondere der Anteil der Schülerinnen und Schüler, die die Schule ohne die für eine erfolgreiche Teilhabe am gesellschaftlichen und beruflichen Leben notwendigen Kompetenzen verlassen, weiter reduziert werden. Bereits im März 2010 hat die Kultusministerkonferenz in einer „Förderstrategie für leistungsschwächere Schülerinnen und Schüler“ die Maßnahmen der Länder zur nachhaltigen Förderung dieser Schülerinnen und Schüler zusammenfassend dargestellt und gemeinsame Leitlinien vereinbart.

Auf der Grundlage der jetzt vorliegenden Daten gilt es, aus den länderspezifischen Stärken und Schwächen Lehren zu ziehen. Die Länder haben mit den Bildungsstandards einen gemeinsamen Referenzrahmen geschaffen, der die Vergleichbarkeit von Anforderungen in den Kernbereichen zentraler Fächer gewährleistet. Mit der Verabschiedung von Bildungsstandards für die Allgemeine Hochschulreife und einem gemeinsamen Aufgabenpool wird dieser Weg konsequent fortgeführt.

Berlin, im Oktober 2012

Senator Ties Rabe
Präsident der Ständigen Konferenz der Kultusminister der Länder
in der Bundesrepublik Deutschland

Kapitel 1

Die länderübergreifenden Bildungsstandards in Deutschland

Katrin Böhme, Dirk Richter, Petra Stanat, Hans Anand Pant und Olaf Köller

1.1 Hintergrund der Einführung länderübergreifender Bildungsstandards

Die Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (*Kultusministerkonferenz, KMK*) leitete zu Beginn des neuen Jahrtausends einen weitreichenden Reformprozess in der deutschen Bildungspolitik ein. Die angestrebten Umgestaltungen basieren auf einem Paradigmenwechsel, der die bis dahin dominierende *Input-* und *Prozessorientierung* durch Elemente einer intensivierten *Outputorientierung* ergänzt. Damit tritt die Vergewisserung über das im Unterricht Erreichte stärker in den Vordergrund und der Stellenwert von Daten und Evaluationen in der Steuerung des deutschen Bildungssystems erhöht sich. Die Entwicklung und Sicherung der Bildungsqualität soll sich stärker als bisher an den Bildungserträgen der Schülerinnen und Schüler, der Schulen sowie des schulischen Bildungssystems insgesamt ausrichten. An die Stelle der früher vorrangigen Gestaltung von Lehrplänen sowie der Entwicklung und Erprobung didaktischer Modelle tritt der datengestützte Entwicklungskreislauf, bei dem Maßnahmen der Qualitätsentwicklung und -sicherung anhand von Bildungserträgen evaluiert werden.

Anlass für diese Reformschritte waren vor allem die ungünstigen Ergebnisse, die sich für Deutschland in den internationalen Vergleichsstudien TIMSS¹ 1995 (Baumert, Bos & Lehmann, 2000a, 2000b; Baumert et al., 1997) und PISA² 2000 (Baumert et al., 2002; Baumert et al., 2001) ergeben hatten. Zentraler Befund von TIMSS 1995 war, dass Schülerinnen und Schüler in Deutschland zum Ende der Pflichtschulzeit im mathematischen Bereich „ein bestenfalls mittleres Grundbildungsniveau“ erreichten (vgl. Baumert, Bos & Lehmann, 2000a, S. 163). Auch im internationalen Vergleich im Rahmen der PISA-Studie 2000 zeigten die in Deutschland unterrichteten Schülerinnen und Schüler in den Domänen Lesen, Mathematik und Naturwissenschaften im Mittel enttäuschende Leistungen (Baumert et al., 2002; Baumert et al., 2001). Ferner identifizierten die genannten Studien erhebliche Disparitäten zwischen den von verschiedenen Schülergruppen erreichten Kompetenzständen. So waren in Deutschland die Kompetenzunterschiede sowohl zwischen Jugendlichen aus unterschiedlichen sozialen Schichten als auch zwischen Jugendlichen aus zugewanderten Familien und Jugendlichen ohne Zuwanderungshintergrund besonders stark ausgeprägt (vgl. zusammenfassend Klieme et al., 2007). Zudem wurde deutlich,

1 Das Akronym TIMSS stand ursprünglich für *Third International Mathematics and Science Study*. Seit 2003 wird es in der Bedeutung *Trends in International Mathematics and Science Study* verwendet.

2 Das Akronym PISA steht für *Programme for International Student Assessment*.

dass die von den Schülerinnen und Schülern in den verschiedenen Ländern der Bundesrepublik Deutschland erreichten Kompetenzniveaus erheblich variierten (Baumert et al., 2002).

Da in Deutschland bis zum Beginn der 1990er Jahre kaum belastbare empirische Daten zu den Erträgen schulischer Bildungsprozesse vorlagen, waren diese Ergebnisse für viele überraschend und wurden in Politik, Schulpraxis, Wissenschaft und allgemeiner Öffentlichkeit ausführlich diskutiert. Die in den internationalen Vergleichsstudien aufgedeckten Problembereiche wurden zum Ausgangspunkt umfassender Bildungsreformen, die das Ziel verfolgten, die Qualität des deutschen Bildungssystems nachhaltig zu verbessern und langfristig zu sichern.

Als erste Reaktion auf das unerwartet schwache Abschneiden der Jugendlichen an deutschen Schulen im Rahmen von TIMSS 1995 erklärte die Kultusministerkonferenz bereits im Oktober 1997 mit dem sogenannten Konstanzer Beschluss die Sicherung der Qualität schulischer Bildung zur zentralen Zielstellung. Analysen der Bildungs- und Schulsysteme jener Staaten, die in internationalen Vergleichsstudien wiederholt besonders gute Ergebnisse erzielen konnten, ergaben Hinweise auf mögliche Ansatzpunkte für den angestrebten Reformprozess. Als wesentliche Erfolgsfaktoren wurden unter anderem die Etablierung von Bildungsstandards sowie die regelmäßige Durchführung von zentralen Vergleichsstudien identifiziert (Arbeitsgruppe „Internationale Vergleichsstudie“, 2007; van Ackeren, 2007).

Um für alle Länder in Deutschland eine gemeinsame Grundlage der Qualitätsentwicklung und des Bildungsmonitorings zu schaffen, erteilte die Kultusministerkonferenz zu Beginn des Jahrtausends den Auftrag, länderübergreifende Bildungsstandards für bestimmte Kernfächer und Schulabschlüsse zu entwickeln. Diese wurden für den Primarbereich und die Sekundarstufe I in den Jahren 2003 und 2004 von der Kultusministerkonferenz verabschiedet. Parallel hierzu wurde die Einrichtung des *Instituts zur Qualitätsentwicklung im Bildungswesen (IQB)* an der Humboldt-Universität zu Berlin angestoßen. Das IQB hat den Auftrag, die länderübergreifenden Bildungsstandards durch geeignete Aufgaben zu operationalisieren, diese zu normieren und ihr Erreichen regelmäßig im Rahmen von Ländervergleichsstudien zu überprüfen. Ferner ist das IQB mit der Entwicklung von Kompetenzstufenmodellen und der Weiterentwicklung der länderübergreifenden Bildungsstandards betraut.

Die empirische Überprüfung des Erreichens der Bildungsstandards, und somit auch der in diesem Bericht vorgestellte Ländervergleich in der Primarstufe 2011, sind Bestandteil einer Gesamtstrategie der Kultusministerkonferenz für das Monitoring des deutschen Bildungssystems (KMK, 2006). Diese Strategie umfasst die folgenden vier Säulen:

- die Beteiligung an stichprobenbasierten internationalen Schulleistungstudien: PISA in der Sekundarstufe I sowie PIRLS/IGLU³ und TIMSS in der Primarstufe;
- die stichprobenbasierte Überprüfung des Erreichens der länderübergreifenden Bildungsstandards für den Primarbereich in den Fächern Deutsch und Mathematik sowie in der Sekundarstufe I in den Fächern Deutsch, Mathematik, der ersten Fremdsprache (Englisch, Französisch) und den Natur-

3 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.

- wissenschaften in Ländervergleichsstudien, die in Anknüpfung an die internationalen Schulleistungsstudien durchgeführt werden;
- die regelmäßige Durchführung von flächendeckenden Vergleichsarbeiten (VERA, Kompetenztests bzw. Lernstandserhebungen) in der dritten und achten Jahrgangsstufe, die sich ebenfalls auf die länderübergreifenden Bildungsstandards beziehen und Impulse für die kompetenzorientierte Unterrichtsentwicklung sowie die Förderung der diagnostischen Kompetenz der Lehrkräfte liefern sollen;
- die gemeinsame Bildungsberichterstattung von Bund und Ländern.

Die ersten beiden Säulen der Gesamtstrategie verdeutlichen, dass die vom IQB seit dem Jahr 2009 durchgeführten Ländervergleichsstudien einen zentralen Stellenwert im Bildungsmonitoring einnehmen und die Teilnahme Deutschlands an internationalen Schulleistungsstudien ergänzen. Die Vergleiche der in den einzelnen Ländern von den Schülerinnen und Schülern erreichten Kompetenzstände erfolgen seit 2009 mit Hilfe von Testverfahren, die auf der Grundlage der länderübergreifenden Bildungsstandards entwickelt werden. Sie ersetzen die nationalen Ergänzungen der internationalen Schulleistungsstudien, wie beispielsweise die IGLU-E-Studie des Jahres 2006 (Bos et al., 2008). Die Ablösung der nationalen Ergänzungsstudien durch die IQB-Ländervergleiche und der damit verbundene Wechsel der Testsysteme erschwert vorübergehend Trendaussagen zur Entwicklung der in den deutschen Ländern erzielten Kompetenzausprägungen (vgl. Kapitel 5.1.4). Die Ländervergleichsstudien sind jedoch ebenfalls zyklisch angelegt, so dass zukünftig auch Aussagen darüber möglich sein werden, inwieweit der Anteil der Schülerinnen und Schüler, die die Bildungsstandards erreichen, in den einzelnen Ländern zu- oder abnimmt.

Die Entwicklung von Bildungsstandards und die regelmäßige Überprüfung ihres Erreichens sind zwar zentrale, aber nicht die einzigen Elemente der angestrebten Qualitätsoptimierung im deutschen Bildungssystem. Kompetenzorientierte Bildungsstandards und empirische Befunde zu den aktuellen Kompetenzständen der Schülerinnen und Schüler allein können ohne Implementationsbemühungen und die Unterstützung einer kompetenzorientierten Unterrichtsentwicklung keine Erfolge erzielen. Um positive Entwicklungsimpulse entfalten zu können, muss ein Bildungsmonitoring daher mit geeigneten Maßnahmen der Unterrichts- und Qualitätsentwicklung in den Schulen vor Ort verknüpft werden (vgl. Oelkers & Reusser, 2008).

1.2 Charakterisierung der länderübergreifenden Bildungsstandards

Für die Einführung der länderübergreifenden Bildungsstandards in Deutschland bildete die sogenannte Klieme-Expertise (Klieme et al., 2007) eine wichtige Grundlage. Sie wurde 2002 im Auftrag des Bundesministeriums für Bildung und Forschung erstellt und umfasst Überlegungen zur Konzeption und Funktion von Bildungsstandards, zu den Grundlagen ihrer Entwicklung, zu möglichen Konsequenzen für das Bildungssystem sowie zur Implementation von Bildungsstandards in Deutschland.

Nach der Klieme-Expertise beschreiben Bildungsstandards Ziele der pädagogischen Arbeit in Form von *Kompetenzerwartungen* und konkretisieren damit den Bildungsauftrag, den allgemeinbildende Schulen erfüllen sollen (Klieme et al., 2007, S. 19). Dabei werden Kompetenzen kognitiv definiert und als erlernbare,

kontextspezifische Leistungsdispositionen verstanden, die erforderlich sind, um Aufgaben oder Probleme eines bestimmten Inhaltsbereichs erfolgreich bewältigen zu können (vgl. bspw. Klieme & Hartig, 2007; Koeppen, Hartig, Klieme & Leutner, 2008).

Klieme und Kollegen (2007, S. 24ff.) sowie Köller (2010, S. 530f.) benennen wesentliche Merkmale von Bildungsstandards. Demzufolge zeichnen sich „gute“ Bildungsstandards neben ihrer Fachspezifität und Fokussierung auf die Kernbereiche des jeweiligen Faches auch dadurch aus, dass sie realistische Erwartungen benennen und kumulative Lernprozesse begünstigen. Ferner sollten sie die Möglichkeit zur Differenzierung verschiedener Kompetenzniveaus eröffnen und beispielsweise unterschiedliche Lernvoraussetzungen auf Seiten der Schülerinnen und Schüler berücksichtigen. Ein weiteres wichtiges Merkmal „guter“ Bildungsstandards ist außerdem, dass ihr Erreichen einer Überprüfung zugänglich gemacht werden kann. Die Standards müssen demnach so formuliert sein, dass sich Messinstrumente entwickeln lassen, mit denen die beschriebenen Kompetenzen erfasst werden können (vgl. Köller, 2010, S. 531).

Bezieht man diese Ausführungen auf die länderübergreifenden Bildungsstandards der Kultusministerkonferenz, so lassen sich diese als Konkretisierung fachspezifischer Kernziele schulischer Bildungsprozesse charakterisieren. Sie orientieren sich an den didaktischen Grundprinzipien des jeweiligen Unterrichtsfachs und beschreiben wesentliche fachbezogene Kompetenzen, die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt in ihrer Bildungsbiografie erworben haben sollen. Die Bildungsstandards formulieren die angestrebten Lernergebnisse in Form von Könnensbeschreibungen (*Can-do-Statements*), die spezifizieren, welche Anforderungen die Schülerinnen und Schüler zu bewältigen in der Lage sein sollen.

Nach einem im Oktober 2004 gefassten Beschluss der Kultusministerkonferenz bilden die Bildungsstandards für den Primarbereich in den Fächern Deutsch und Mathematik (KMK, 2005a, 2005b) seit Beginn des Schuljahres 2005/2006 die Grundlage für fachspezifische Anforderungen des Unterrichts in allen Ländern der Bundesrepublik Deutschland. Ferner verpflichteten sich die Länder dazu, die Bildungsstandards insbesondere über ihre Lehrpläne, Strategien der Schulentwicklung sowie die Lehreraus- und -fortbildung zu implementieren. Entsprechend wurden nach der verbindlichen Einführung der länderübergreifenden Bildungsstandards die Lehrpläne und Curricula in den Ländern überarbeitet, um sie auf die Bildungsstandards abzustimmen.

Die Bildungsstandards der Kultusministerkonferenz stellen somit eine länderübergreifend verbindliche Richtschnur dar, die wichtige Bewertungskriterien sowohl für externe als auch für interne Evaluationsprozesse liefert. Sie ermöglichen es, anhand von transparenten und anerkannten Erwartungen Stärken und Schwächen auf der Ebene des Bildungssystems und einzelner Schulen zu identifizieren und damit eine verlässliche Grundlage für Veränderungs- und Optimierungsprozesse zur Verfügung zu stellen. Entsprechend gewinnen großflächig angelegte Leistungserhebungen auf der Basis standardisierter Tests an Bedeutung. Diese Testungen dienen unter anderem der Überprüfung des Erreichens der Bildungsstandards und somit einem kontinuierlichen, länderübergreifenden Bildungsmonitoring. Die in solchen Studien gewonnenen empirischen Befunde bilden eine zentrale Komponente der ergebnisorientierten Systemsteuerung und damit einen wichtigen Ausgangspunkt der Qualitätsentwicklung und -sicherung im deutschen Bildungssystem (Böhme, 2012).

Es wird angenommen, dass das systematische Erheben von leistungsbezogenen Daten die Schul- und Unterrichtsentwicklung unterstützen kann, da diese Informationen wichtige Impulse für Schulentwicklungsprozesse liefern können (vgl. Köller & Pant, 2010). Mit dieser Intention werden auch die jährlich in der dritten Jahrgangsstufe auf der Basis der Bildungsstandards stattfindenden Vergleichsarbeiten durchgeführt. Die Lehrkräfte erhalten im Rahmen von Vergleichsarbeiten aus einer schulexternen Informationsquelle Hinweise auf mögliche Stärken und Schwächen ihrer Klassen und können ihre weitere Unterrichtsarbeit entsprechend ausrichten. Auf diese Weise sollen Vergleichsarbeiten zur Implementation der länderübergreifenden Bildungsstandards und zur Verbesserung der Leistungsfähigkeit des deutschen Bildungssystems beitragen.

Mit der Einführung der Bildungsstandards ist auch die Erwartung verbunden, dass die Lehrkräfte ihre Arbeit im Unterrichtsalltag zunehmend auf den Kompetenzaufbau der Schülerinnen und Schüler ausrichten. Das Innovationspotenzial der Bildungsstandards wird darin gesehen, dass sie durch ihren Bezug auf Kompetenzen die Inputorientierung reduzieren, bei der eine stofforientierte Unterrichtsgestaltung über Lehrpläne und Curricula vorherrscht (Köller, 2008). In den Fokus soll stattdessen der kumulative Kompetenzaufbau im jeweiligen Fach rücken. Durch das den länderübergreifenden Bildungsstandards zugrunde liegende Kompetenzkonzept und die stärkere Outputorientierung wird es möglich, Lehrkräfte bei ihrer Unterrichtsgestaltung weniger stark einzuschränken als bei einer Inputsteuerung, die mit detaillierten inhaltlichen Vorgaben verbunden ist. Dieser Aspekt des Potenzials kompetenzorientierter Bildungsstandards und ihrer Implementierung wird beispielsweise in der von der Kultusministerkonferenz in Zusammenarbeit mit dem IQB entwickelten Handreichung mit dem Titel „Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung“ (KMK, 2010) erläutert.

1.3 Entwicklung von Testverfahren zur Überprüfung der länderübergreifenden Bildungsstandards im Primarbereich

Um das Erreichen der länderübergreifenden Bildungsstandards in den Fächern Deutsch und Mathematik in der Primarstufe überprüfen zu können, war es zunächst erforderlich, auf ihrer Basis reliable und valide Messinstrumente zu entwickeln. Die Testkonstruktion für beide Fächer erfolgte unter Federführung des IQB (vgl. Granzer, 2009).

In Vorarbeiten des IQB wurde zunächst in enger Kooperation mit Expertinnen und Experten aus den Fachdidaktiken präzisiert, welche der in den Bildungsstandards beschriebenen Kompetenzaspekte im Rahmen eines *Large-Scale-Assessments* operationalisiert werden können und wie diese zu spezifizieren sind (vgl. Kapitel 2). Weiterhin wurden vom IQB unter Einbeziehung internationaler Standards Hinweise zur Entwicklung und Gestaltung von Testaufgaben (*Items*) erarbeitet. Mit Hilfe dieser Handreichungen wurden Aufgabenentwicklungsteams im Rahmen von intensiven Schulungen darüber informiert, wie gute Testaufgaben konstruiert sind und wie typische Fehler bei der Itementwicklung vermieden werden können. In praktischen Übungen erhielten die Aufgabenentwicklungsteams und -entwickler zudem die Möglichkeit, die in den Schulungen erworbenen Kenntnisse direkt anzuwenden.

Der Prozess der an den Bildungsstandards orientierten Testentwicklung erfolgte in beiden Fächern in den folgenden fünf Schritten:

1. Aufgabenentwicklung durch erfahrene Lehrkräfte

Die Entwicklung der Testaufgaben erfolgte in den Jahren 2005 bis 2007 durch Lehrkräfte aus allen 16 Ländern (vgl. Kapitel 4). Nach intensiven Schulungen arbeiteten diese in regionalen Arbeitsgruppen unter fachdidaktischer Anleitung an der Generierung und Optimierung der Testaufgaben für die Fächer Deutsch und Mathematik. Die Aufgaben wurden bereits in dieser frühen Entwicklungsphase durch die Lehrkräfte in ausgewählten Klassen im Rahmen sogenannter Präpilotierungen erprobt, um authentische Schülerlösungen als Grundlage für die Entwicklung von Auswertungsanleitungen und die Optimierung der Items zu erhalten.

2. Begutachtung der Items

Expertinnen und Experten aus den Bereichen Fachdidaktik, Bildungsforschung und Psychometrie beurteilten und kommentierten in einem zweiten Schritt die Items im Hinblick auf ihre fachdidaktische und testdiagnostische Qualität. Anhand dieser Rückmeldungen wurden die Aufgaben überarbeitet und optimiert.

3. Pilotierung – empirische Erprobung der entwickelten Aufgaben

Nach der Generierung eines umfassenden Itempools, der genügend Testaufgaben für alle einbezogenen Kompetenzbereiche der beiden Fächer umfasste, wurde im Jahr 2006 eine Pilotierungsstudie mit einer bundesweiten Stichprobe von $N = 15\,701$ Schülerinnen und Schülern der dritten und vierten Jahrgangsstufe durchgeführt (vgl. Winkelmann & Böhme, 2009). Die Festlegung der Stichprobengröße erfolgte dabei so, dass hinreichend robuste Schätzungen der Itemschwierigkeiten mit Modellen der *Item-Response-Theory* (vgl. Kapitel 12) möglich waren. Anhand entsprechender Auswertungen wurden statistisch problematische Items identifiziert und anschließend überarbeitet oder aus dem Itempool entfernt. Weiterhin wurden für Aufgaben mit offenen Antwortformaten Schülerlösungen aus der Pilotierungsstudie für die Optimierung von Auswertungsanleitungen genutzt. Dieser dritte Schritt der Testentwicklung resultierte in einem erprobten Aufgabenpool, der die Grundlage sowohl der nachfolgenden Normierungsstudie als auch des aktuellen Ländervergleichs bildete.

4. Normierung des Aufgabenpools

Im Jahr 2007 wurden die pilotierten Aufgaben und Items von einer bundesweit repräsentativen Stichprobe von $N = 10\,737$ Schülerinnen und Schülern der dritten und vierten Jahrgangsstufe bearbeitet. Auf der Basis dieser Daten wurden die statistischen Eigenschaften der Items (z.B. ihre empirische Schwierigkeit) wiederum auf der Grundlage von Modellen der *Item-Response-Theory* bestimmt. Anschließend erfolgte die Definition der Skalen, die den länderübergreifenden und bildungsstandardbasierten Maßstab für die von den Kindern erreichten Kompetenzstände in den unterschiedlichen Kompetenzbereichen der Fächer Deutsch und Mathematik in der Primarstufe bilden (vgl. Kapitel 3.1).

5. Entwicklung von Kompetenzstufenmodellen

Auf Grundlage der empirischen Befunde aus den Pilotierungs- und Normierungsstudien sowie unter Bezugnahme auf fachdidaktische Kompetenzstrukturmodelle wurden in einem abschließenden Entwicklungsschritt Kompetenzstufenmodelle definiert. Mit Hilfe solcher Modelle ist es möglich, sowohl die Testergebnisse von Schülerinnen und Schülern als auch spezifische Testaufgaben inhaltlich definierten Kompetenzniveaus zuzuordnen. So lassen sich Aussagen darüber treffen, welche durch das Aufgabenmaterial konkretisierten Anforderungen von den Schülerinnen und Schülern auf einer bestimmten Kompetenzstufe mit hoher Wahrscheinlichkeit bewältigt werden können.

Bei der Entwicklung von Kompetenzstufenmodellen kommen sogenannte *Standard-Setting-Verfahren* zum Einsatz (vgl. Pant, Rupp, Tiffin-Richards & Köller, 2009). Ziel eines Standard-Setting-Prozesses ist es, die für jeden Kompetenzbereich vorliegenden kontinuierlichen Kompetenzskalen in sinnvoll abgrenzbare Abschnitte (Kompetenzstufen oder Kompetenzniveaus) einzuteilen. Im Primarbereich wurden die Skalen aller Kompetenzbereiche in fünf gleichabständige Niveaustufen geteilt. Ferner wurde festgelegt, welche Stufen der Erfüllung von Mindest-, Regel- und Maximalstandards entsprechen.

Bei einem umfassenden Standard-Setting werden die Stufengrenzen in einem konsensuellen Verfahren definiert. Eine Gruppe von Expertinnen und Experten aus den Bereichen Fachdidaktik, Psychometrie und Bildungsadministration einigt sich auf die Festlegung der Grenzen zwischen den Stufen (*Cut-Scores*) und wählt diese so, dass sich die fachdidaktisch und kognitionspsychologisch begründeten Sprünge in den Aufgabenanforderungen, die zur Lösung der stufenspezifischen Testaufgaben notwendig sind, sinnvoll beschreiben lassen.

Die für den Ländervergleich 2011 relevanten Kompetenzstufenmodelle mit ihren genauen Grenzen und inhaltlichen Stufenbeschreibungen werden in Kapitel 3 dieses Berichts ausführlich dargestellt. Sie bilden die Grundlage für den „Blick in die Länder“ in Kapitel 6. Darin wird die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen dargestellt und überprüft, inwieweit in den einzelnen Ländern die Bildungsstandards erreicht worden sind.

Literatur

- Arbeitsgruppe „Internationale Vergleichsstudie“. (2007). *Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten*. Bonn: Bundesministerium für Bildung und Forschung.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Bos, W. & Lehmann, R. H. (Hrsg.). (2000a). *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Bd. 1., Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit)*. Opladen: Leske + Budrich.
- Baumert, J., Bos, W. & Lehmann, R. H. (Hrsg.). (2000b). *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Bd. 2., Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe)*. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillman, K.-J., & Weiß, M. (Hrsg.). (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.

- Baumert, J., Lehmann, R. H., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O. & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske + Budrich.
- Böhme, K. (2012). *Methodische und didaktische Überlegungen sowie empirische Befunde zur Erfassung sprachlicher Kompetenzen im Deutschen (Dissertation)*. Humboldt-Universität zu Berlin. Vorgesehen für eine elektronische Publikation unter <http://edoc.hu-berlin.de>. Berlin.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2008). *IGLU-E 2006: Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Granzer, D. (2009). Von Bildungsstandards zu ihrer Überprüfung: Grundlagen der Item- und Testentwicklung. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 21–30). Weinheim: Beltz.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (Hrsg.). (2007). *Zur Entwicklung nationaler Bildungsstandards. Expertise*. Bonn: Bundesministerium für Bildung und Forschung.
- Klieme, E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften.
- KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2005b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2006) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Wolters Kluwer.
- KMK (2010) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung*. Köln: Wolters Kluwer.
- Koepfen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current issues in competence modeling and assessment. *Zeitschrift für Psychologie/Journal of Psychology*, 216, 61–73.
- Köller, O. (2008). Bildungsstandards in Deutschland: Implikationen für die Qualitätssicherung und Unterrichtsqualität. In M. A. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), *Perspektiven der Didaktik* (S. 47–59). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O. (2010). Bildungsstandards. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch der Bildungsforschung* (S. 529–548). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O. & Pant, H. A. (2010). Die Rolle von Bildungsstandards in einem System der Qualitätssicherung und Qualitätsentwicklung. In B. Schaal & F. Huber (Hrsg.), *Qualitätssicherung im Bildungswesen*. (S. 55–67). Münster: Waxmann.
- Oelkers, J. & Reusser, K. (2008). *Expertise: Qualität entwickeln, Standards sichern, mit Differenz umgehen*. Bonn: Bundesministerium für Bildung und Forschung.
- Pant, H. A., Rupp, A. A., Tiffin-Richards, S. P. & Köller, O. (2009). Validity issues in standard-setting studies. *Studies in Educational Evaluation*, 35, 95–101.
- van Ackeren, I. (2007). *Nutzung großflächiger Tests für die Schulentwicklung. Exemplarische Analyse der Erfahrungen aus England, Frankreich und den Niederlanden*. Berlin: Bundesministerium für Bildung und Forschung.
- Winkelmann, H. & Böhme, K. (2009). Anlage und Durchführung der Pilotierung der Bildungsstandards. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 31–41). Weinheim: Beltz.

Kapitel 2

Die im Ländervergleich 2011 in den Fächern Deutsch und Mathematik untersuchten Kompetenzen

2.1 Beschreibung der im Fach Deutsch untersuchten Kompetenzen

Katrin Böhme und Albert Bremerich-Vos

Dieses Teilkapitel beschreibt die im Ländervergleich 2011 überprüften Kompetenzen im Fach Deutsch. Zunächst werden die zugrunde liegenden Bildungsstandards im Fach Deutsch für den Primarbereich vorgestellt (Abschnitt 2.1.1) und es wird kurz darauf eingegangen, wie die Struktur sprachlicher Kompetenzen im Deutschen konzipiert werden kann (Abschnitt 2.1.2). Daran schließen sich die Charakterisierungen der einzelnen Konstrukte und Erläuterungen zu ihrer Operationalisierbarkeit sowie eine Darstellung der im Ländervergleich 2011 jeweils eingesetzten Aufgaben zu den drei untersuchten Kompetenzbereichen *Zuhören* (Abschnitt 2.1.3), *Lesen* (Abschnitt 2.1.4) und *Orthografie* (Abschnitt 2.1.5) an.

2.1.1 Die Bildungsstandards im Fach Deutsch für den Primarbereich

Die Vermittlung und Förderung sprachlicher Kompetenzen ist ein zentraler Bestandteil des Bildungsauftrags im Primarbereich. Im schulischen Alltag dient Sprache in allen Fächern als Medium des Lernens und stellt eine wesentliche Voraussetzung für den schulischen Erfolg und die gesellschaftliche Teilhabe der Schülerinnen und Schüler dar. Somit kommt der sicheren Beherrschung der Instruktions- und Verkehrssprache Deutsch eine grundlegende Bedeutung für eine positive kognitive, emotionale und soziale Entwicklung der Kinder zu (vgl. KMK, 2005a, S. 6f.).

Die Bildungsstandards im Fach Deutsch für den Primarbereich definieren, welche Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Kernbereichen dieses Unterrichtsfachs in der Regel erreicht werden sollen (vgl. KMK, 2005a, S. 7). Sie fanden Eingang in die Curricula der Länder und geben somit vor, welche fachlichen Ziele in der Unterrichtsarbeit anzustreben sind. In den Bildungsstandards für das Fach Deutsch werden vier Kompetenzbereiche beschrieben, die in Abbildung 2.1 jeweils mit ihren zentralen Standards aufgeführt sind.

Abbildung 2.1: Kompetenzbereiche und zentrale Standards für das Fach Deutsch im Primarbereich (vgl. KMK, 2005a, S. 7)

<p>1) Sprechen und Zuhören</p> <ul style="list-style-type: none"> • zu anderen sprechen • verstehend zuhören • Gespräche führen • szenisch spielen • über Lernen sprechen 	<p>2) Schreiben</p> <ul style="list-style-type: none"> • über Schreibfertigkeiten verfügen • richtig schreiben • Texte planen • Texte schreiben • Texte überarbeiten 	<p>3) Lesen – mit Texten und Medien umgehen</p> <ul style="list-style-type: none"> • über Lesefähigkeiten verfügen • über Leseerfahrungen verfügen • Texte erschließen • Texte präsentieren
<p>Methoden und Arbeitstechniken</p> <p>Methoden und Arbeitstechniken werden jeweils in Zusammenhang mit den Inhalten jedes einzelnen Kompetenzbereichs erworben.</p>		
<p>4) Sprache und Sprachgebrauch untersuchen</p> <ul style="list-style-type: none"> • grundlegende sprachliche Strukturen und Begriffe kennen • sprachliche Verständigung untersuchen • an Wörtern, Sätzen, Texten arbeiten • Gemeinsamkeiten und Unterschiede von Sprachen entdecken 		

Die orthografische Kompetenz wird im Referenzdokument der Kultusministerkonferenz (KMK, 2005a) nicht als eigenständiger Kompetenzbereich aufgeführt, sondern in die Bereiche *Schreiben* sowie *Sprache und Sprachgebrauch untersuchen* eingeschlossen. Im Rahmen der Testentwicklung für die Überprüfung des Erreichens der Bildungsstandards wurde Orthografie jedoch als eigenständiger Kompetenzbereich behandelt und mit spezifischen Testinstrumenten operationalisiert. Die Gründe hierfür werden in Abschnitt 2.1.5 dargestellt.

In einem modernen, integrativ angelegten Deutschunterricht sind alle genannten Kompetenzbereiche aufeinander bezogen (KMK, 2005a, S. 8). Dies gilt insbesondere für die Kompetenz *Sprache und Sprachgebrauch untersuchen*, die eine besonders enge Anbindung an die anderen Kompetenzbereiche erfordert. Die angestrebte Verknüpfung wird in den Bildungsstandards im Fach Deutsch dadurch veranschaulicht, dass der Bereich *Sprache und Sprachgebrauch untersuchen* „quer“ zu den übrigen Kompetenzbereichen angeordnet ist (vgl. Abbildung 2.1).

Die in den Bildungsstandards für das Fach Deutsch beschriebenen Kompetenzerwartungen können kognitive Anforderungen und Operationen unterschiedlicher Komplexität umfassen. Um hierfür eine Orientierung zu geben, werden in den Bildungsstandards drei *Anforderungsbereiche* definiert (vgl. KMK, 2005a, S. 17). Bewältigen die Schülerinnen und Schüler Aufgaben, die dem Anforderungsbereich I „Wiedergeben“ zugeordnet sind, können sie bekannte Informationen wiedergeben und grundlegende Verfahren und Routinen anwenden. Im Anforderungsbereich II „Zusammenhänge herstellen“ bearbeiten die Schülerinnen und Schüler vertraute Sachverhalte, indem sie erworbenes Wissen und bekannte Methoden anwenden und miteinander verknüpfen. Anforderungsbereich III „Reflektieren und beurteilen“ bezieht sich auf die Bearbeitung neuartiger Problemstellungen, die von den Schülerinnen und Schülern eigenständige Beurteilungen und eigene Lösungsansätze erfordern (vgl. KMK, 2005a, S. 17).

Die Bildungsstandards im Fach Deutsch für den Primarbereich gehen von einer vorschulisch erworbenen Sprachhandlungskompetenz aus, die systematisch mit Hilfe von schulischen Lerngelegenheiten erweitert werden soll. Die angestrebten

Kompetenzen werden in den Bildungsstandards als Könnensbeschreibungen formuliert (*Can-do-Statements*). Explizite Definitionen der jeweiligen Kompetenzkonstrukte beinhalten sie jedoch nicht (vgl. Granzer, 2009). Um den Grad in dem die Bildungsstandards erreicht werden einer Überprüfung zugänglich machen zu können, war es daher vor der Überführung der einzelnen Standards in Testaufgaben erforderlich, Definitionen aller relevanten Kompetenzkonstrukte zu erarbeiten. Diese Arbeit wurde vom *Institut zur Qualitätsentwicklung im Bildungswesen* (IQB) in enger Kooperation mit Expertinnen und Experten aus der Deutschdidaktik geleistet. Dabei wurde sowohl auf die einschlägige internationale und nationale Forschung wie auch auf Vorarbeiten aus anderen etablierten *Large-Scale-Assessments* (insbesondere PISA¹ (Baumert et al., 2002; Baumert et al., 2001) sowie PIRLS/IGLU² (Bos et al., 2007; Bos, Lankes, Prenzel & Schwippert, 2003)) zurückgegriffen.

Nachfolgend werden für die drei im Ländervergleich 2011 überprüften Kompetenzbereiche Zuhören, Lesen und Orthografie jeweils die in den Bildungsstandards der Kultusministerkonferenz aufgeführten Kompetenzaspekte dargestellt. Ferner werden die Konstruktbeschreibungen, die der jeweiligen Testentwicklung zugrunde liegen, charakterisiert. Auch das Vorgehen bei der Überführung der Bildungsstandards in konkrete Testaufgaben (Operationalisierung) wird in den folgenden Abschnitten jeweils vertiefend erläutert. Zunächst soll allerdings aufgezeigt werden, warum im Fach Deutsch – anders als im Fach Mathematik – nicht von einer „globalen Deutschfähigkeit“ ausgegangen wird, sondern die untersuchten Kompetenzbereiche stets separat erfasst und analysiert werden.

2.1.2 Die Struktur sprachlicher Kompetenzen

Für die Überprüfung von Kompetenzen der Schülerinnen und Schüler im Fach Deutsch ist die Frage zentral, ob ein übergreifender Faktor sprachlicher Kompetenzen anzunehmen ist, auf den die Leistungen etwa beim Lesen und Zuhören zurückgeführt werden können (vgl. Oller, 1976) oder ob mehrere weitgehend unabhängige Teilkompetenzen zu unterscheiden sind (vgl. Böhme, Neumann & Schipolowski, 2010). Je nachdem, ob von einer allgemeinen sprachlichen Kompetenz oder von relativ unabhängigen Teilkompetenzen ausgegangen wird, unterscheiden sich Inhalt und Struktur der für die Überprüfung der Kompetenzstände eingesetzten Testinstrumente (vgl. Böhme, 2012; Jude, 2008).

Empirische Analysen zur Struktur der Bildungsstandards im Fach Deutsch für den Primarbereich wurden von Bremerich-Vos, Böhme und Robitzsch (2009) vorgelegt. Diese zeigen, dass die verschiedenen Kompetenzen zwar eng miteinander assoziiert sind, die Beschränkung auf einen Generalfaktor aber eine unzureichende Beschreibung der Kompetenzstrukturen liefert. Die theoretisch plausible Unterscheidung von produktiven und rezeptiven Fähigkeiten lässt sich im Wesentlichen auch empirisch bestätigen (Bremerich-Vos et al., 2009, S. 211ff.). Insbesondere das Zuhören ist von den übrigen Kompetenzbereichen im Fach Deutsch recht deutlich abgrenzbar.

1 Das Akronym *PISA* steht für *Programme for International Student Assessment*.

2 Das Akronym *PIRLS* steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.

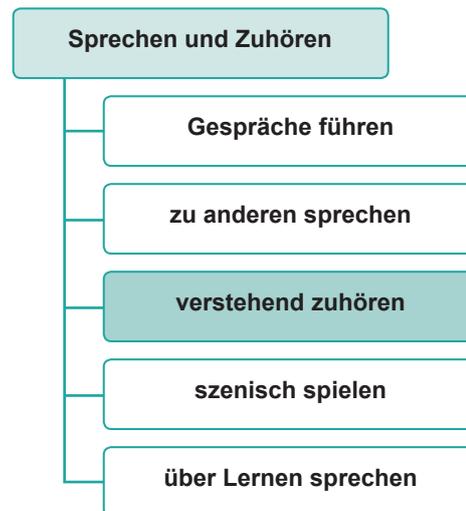
Insofern, wie auch aufgrund des theoretischen Diskussionsstandes in der Deutschdidaktik, liegt es nahe, die verschiedenen Bereiche separat zu modellieren. Aus diesem Grund wird im Fach Deutsch – anders als im Fach Mathematik (vgl. Kapitel 2.2) – auf die Verwendung eines Gesamtpunktwerts zur Erfassung einer globalen Deutschkompetenz verzichtet. In den nachfolgenden Kapiteln werden die Ergebnisse für die einzelnen Kompetenzbereiche (Lesen, Zuhören und Orthografie) also stets separat berichtet.

2.1.3 Beschreibung der Zuhörkompetenz und ihrer Operationalisierung

Die Zuhörkompetenz in den Bildungsstandards der KMK

In den Bildungsstandards umfasst der *Kompetenzbereich I: Sprechen und Zuhören* sowohl die produktive als auch die rezeptive mündliche Sprachkompetenz der Schülerinnen und Schüler. Dies folgt der Tradition des Deutschunterrichts, die beide Aspekte unter dem Begriff der Gesprächskompetenz zusammenfasst (vgl. Behrens & Eriksson, 2009). Der Schwerpunkt der schulischen Förderung liegt dabei auf dem Sprechen. Abbildung 2.2 stellt die für diesen Kompetenzbereich formulierten Standards dar.

Abbildung 2.2: Gruppierung der Bildungsstandards im Kompetenzbereich „Sprechen und Zuhören“ für den Primarbereich



Anmerkung. Hervorgehoben ist der Teilbereich „verstehend zuhören“, für den im Ländervergleich 2011 Testaufgaben entwickelt und eingesetzt wurden.

In den allgemeinen Erläuterungen der einzelnen Kompetenzbereiche heißt es zur Zuhörkompetenz der Schülerinnen und Schüler lediglich: „Sie [...] hören aufmerksam und genau zu, nehmen die Äußerungen anderer auf und setzen sich mit diesen konstruktiv auseinander“ (KMK, 2005a, S. 8). Die formulierten Standards beschränken sich auf das verstehende Zuhören, und zwar konkret darauf, Inhalte zuhörend zu verstehen, gezielt nachzufragen sowie Verstehen und Nicht-Verstehen zum Ausdruck zu bringen (KMK, 2005a, S. 10). Bei genauerer Betrachtung handelt es sich auch bei den beiden letztgenannten Teilkompetenzen um *produktive* Aspekte verbaler beziehungsweise nonverbaler Kommunikation, die keine Konstruktbestandteile des *rezeptiven* Zuhörens sind.

Zuhörkompetenz als Konstrukt

Die Zuhörkompetenz ist für das schulische Lernen im Primarbereich besonders wichtig. Belgrad, Eriksson, Pabst-Weinschenk und Vogt (2008) gehen davon aus, dass „Schüler(innen) bis zu zwei Drittel der Unterrichtszeit zuhören müssen“ (Belgrad et al., 2008, S. 20). Dieser Umstand verdeutlicht, dass ein Großteil des Lerninputs in Form von mündlicher Informationsvermittlung erfolgt. Daher ist es erforderlich, Schülerinnen und Schüler in die Lage zu versetzen, ihr Zuhörverhalten selbstständig und kompetent zu steuern.

Zuhören zu können umfasst mehr als eine gelingende akustische Wahrnehmung von Schallwellen. Im schulischen Kontext beinhaltet Zuhörkompetenz auch die Bereitschaft, anderen zuzuhören, die kognitive Verarbeitung des sprachlichen Inputs sowie die Reflexion von Sprecher- und Situationsmerkmalen vor dem Hintergrund gesellschaftlicher und persönlicher Erwartungen. Inwieweit für die Weiterentwicklung der Zuhörkompetenz eine explizite schulische Förderung erforderlich ist, ist nicht abschließend geklärt. Zwar beklagen Lehrkräfte, dass die Fähigkeit der Schülerinnen und Schüler, aufmerksam und konzentriert zuzuhören, ungenügend ausgeprägt ist und innerhalb der Schülerschaft eine (zu) große Varianz aufweist (Behrens, Böhme & Krelle, 2009), dennoch spielte der Bereich Zuhören in deutschdidaktischen Überlegungen lange Zeit nur eine untergeordnete Rolle (vgl. Krelle, 2010). Momentan findet in der Deutschdidaktik allerdings ein Wandel statt und dem Zuhören als Kompetenzbereich wird inzwischen ein deutlich höherer Stellenwert eingeräumt.

Wie für den Bereich Lesen liegen auch zum Zuhören kognitionspsychologisch und psycholinguistisch fundierte Überlegungen vor, die als Grundlage für die Erarbeitung eines theoretisch plausiblen Kompetenzmodells dienen können. Prominent sind insbesondere die Arbeiten von Imhof, die Zuhören als einen mehrstufigen Informationsverarbeitungsprozess versteht, bei dem in den Phasen der Intentionbildung, der Selektion, der Organisation und der Integration die Selbstregulation des Zuhörenden eine ausschlaggebende Rolle spielt (vgl. Imhof, 2003, 2010).

Wie im Kompetenzbereich Lesen können auch für das Zuhören unterschiedliche Verarbeitungsprozesse unterschieden werden. In der aktuellen Literatur wird betont, dass hierarchiehöhere Informationsverarbeitungsprozesse, wie etwa elaborative Schlüsse oder die Konstruktion eines mentalen Modells, beim Lese- und Hörverstehen sehr ähnlich ablaufen. Unterschiede beziehen sich primär auf die hierarchieniedrigen Ebenen des kognitiven Systems und sind auf modalitätsspezifische Besonderheiten der akustischen beziehungsweise visuellen Information zurückzuführen (vgl. z. B. Kürschner & Schnotz, 2008).

Operationalisierung und Überführung in Testaufgaben

Die Aufgabenentwicklung für den Ländervergleich 2011 bezog sich im Kompetenzbereich *Sprechen und Zuhören* zunächst nur auf den Standard *verstehend zuhören*, da das aktive Sprechen wie auch das szenische Spielen in Large-Scale-Assessments nicht mit realistischem Aufwand erhoben werden können. In den Bildungsstandards der Kultusministerkonferenz sind die für den Teilbereich Zuhören definierten Standards nur sehr grobkörnig beschrieben. Daher war es erforderlich, diesen Bereich als Grundlage für die empirischen Arbeiten zu den Bildungsstandards weiter auszudifferenzieren.

Wie bereits erwähnt, bestehen hinsichtlich der relevanten Informationsverarbeitungsprozesse beim Lesen und Zuhören zahlreiche Ähnlichkeiten; die wichtigsten unterscheidenden Merkmale beziehen sich auf modalitätsspezifische Besonderheiten wie beispielsweise die Flüchtigkeit gesprochener Sprache und die sprachlich-artikulatorische (prosodische) Gestaltung des Gesagten. In der gesprochenen Sprache werden Einheiten wie Wörter, Phrasen oder Sätze akzentuiert und durch Pausen unterschiedlicher Länge voneinander getrennt. Es gibt Veränderungen in der Tonhöhe, der Geschwindigkeit und der Lautstärke, womit unter anderem Gewichtung und emotionale Beteiligung markiert werden (vgl. z.B. Imhof, 2003). Diese prosodischen Merkmale können für die Bedeutungskonstruktion und Interpretation relevant sein.

Sind Hörtexte dialogisch gestaltet oder gibt es darin dialogische Abschnitte, müssen Aspekte der Gesprächsorganisation verstanden werden (vgl. z. B. Deppermann, 2008). Dabei geht es zum Beispiel um die Abfolge von Frage und Antwort oder Vorwurf und Rechtfertigung sowie um Regeln und Modalitäten des Sprecherwechsels. Für das Verständnis können darüber hinaus spezifische Merkmale gesprochener Sprache relevant sein, wie beispielsweise eine regional-sprachliche Lautung und syntaktische Konstruktionen, die in der geschriebenen Sprache nicht oder nur selten vorkommen.

In literarischen Hörtexten, die den Bildungsstandards entsprechend nicht vernachlässigt werden dürfen (KMK, 2005a, S. 9), spielen solche Merkmale allerdings kaum eine Rolle. Hier geht es in erster Linie um kognitiv-emotionale Aspekte des Zuhörens, wie beispielsweise die Hörästhetik (vgl. z. B. Gailberger, 2008).

In der internationalen Forschung, insbesondere im Rahmen des englischsprachigen *Language Testing*, haben die empirische Analyse und die kompetenzorientierte Diagnostik des Zuhörens beziehungsweise des Hörverstehens eine wesentlich längere Tradition als in relevanten deutschsprachigen Forschungszweigen (vgl. z.B. Buck, 2001). Für die Testentwicklung im Bereich Zuhören können aus dem derzeitigen Forschungsstand folgende Empfehlungen abgeleitet werden (Buck, 2001, S. 113):

- Das Hauptaugenmerk sollte auf der Erfassung solcher sprachlicher Kompetenzen liegen, die für das Hörverstehen einzigartig sind. Abgeprüft werden sollten also schnelle, automatisierte Verarbeitungsprozesse von Hörtexten, die den Besonderheiten gesprochener Sprache gerecht werden.
- Ferner sollte überprüft werden, ob der Zuhörende basale Informationen in einer Vielzahl verschiedenartiger Hörtexte zu unterschiedlichen Themen erfassen kann.
- Es sollten auch längere Texte und Gespräche Verwendung finden, da diese Strategien des Umgangs mit großen Informationsmengen erfordern und man dauerhaft aufmerksam sein muss.
- Es sollte nicht nur überprüft werden, ob den Hörtexten explizite Informationen abgewonnen werden können, sondern auch, ob Bedeutungen schlussfolgernd erfasst werden.

Beschreibung der im Ländervergleich 2011 eingesetzten Zuhöraufgaben

Zur Überprüfung der Zuhörkompetenz der Schülerinnen und Schüler wurden im Ländervergleich 2011 fünf verschiedene Aufgabenblöcke eingesetzt (vgl. Kapitel 12.1). Diese Blöcke beinhalteten jeweils eine Zuhöraufgabe, die aus einem

Hörtext (auditiver Stimulus) und mehreren schriftlich vorgegebenen und auch schriftlich zu bearbeitenden Items bestand. In der Summe handelte es sich um 51 Items in insgesamt fünf verschiedenen Zuhöraufgaben.

Die auditiven Stimulustexte entstammten unterschiedlichen Quellen. Um der großen Bandbreite möglicher Zuhörsituationen Rechnung zu tragen und den oben beschriebenen Empfehlungen von Buck (2001) zu folgen, wurden sowohl klassische Hörspiele als auch Kinderradiosendungen eingesetzt. Letztere zeichnen sich deutlich stärker als Hörspiele durch die Merkmale authentischer Mündlichkeit aus. Alle Hörtexte wurden durch eine kurze Situierung, die den Kindern vorgelesen wurde und auch zum Mitlesen in den Testheften abgedruckt war, eingeführt und einmal von CD vorgespielt. Im Anschluss hatten die Schülerinnen und Schüler Zeit, die zugehörigen Items zu bearbeiten. Längere Stimuli wurden geteilt und in Ausschnitten vorgespielt. Zwischen diesen Teilen wurden die Kinder gebeten, Items zu beantworten, die sich auf den jeweils gehörten Ausschnitt bezogen. Die Länge der Hörtexte beziehungsweise der aus diesen Texten vorgespielten Ausschnitte variierte zwischen zwei und viereinhalb Minuten. Zusätzlich wurde ein längerer Hörtext mit neun Minuten Dauer in die Testung einbezogen, um auch die Zuhörkompetenzen von Schülerinnen und Schülern im oberen Kompetenzbereich differenziert erfassen zu können. Die für die Beantwortung der Items jeweils vorgesehene Bearbeitungszeit war so bemessen, dass die Kinder nicht unter Zeitdruck arbeiten mussten, sondern die Gelegenheit hatten, alle Items in Ruhe durchzulesen und zu beantworten.

Bei den Items handelte es sich überwiegend um geschlossene Formate. Die 30 geschlossenen Items zeichneten sich dadurch aus, dass keine produktive Schreibleistung verlangt wurde, sondern aus vorgegebenen Antwortoptionen ausgewählt oder entschieden werden musste, ob bestimmte Aussagen korrekt sind oder nicht. Das gängigste der geschlossenen Itemformate ist das *Multiple-Choice-Format*, bei dem unter jeweils vier vorgegebenen Antwortoptionen die richtige beziehungsweise die am besten passende Option ausgewählt werden soll. Insgesamt 15 Items hatten ein halboffenes Format und verlangten von den Schülerinnen und Schülern eine Kurzantwort, also eine kurze schriftliche Äußerung im Umfang eines einzelnen Wortes oder einer Wortgruppe. Zusätzlich wurden sechs offene Items vorgelegt, für deren Lösung eine längere schriftliche Antwort beziehungsweise Begründung im Umfang von zwei bis drei Sätzen verfasst werden musste. Alle Itemformate sowie der Ablauf der Testung wurden den Schülerinnen und Schülern zu Beginn ausführlich erklärt und die Bearbeitung der verschiedenen Itemformate wurde anhand von Beispielitems illustriert.

2.1.4 Beschreibung der Lesekompetenz und ihrer Operationalisierung

Die Lesekompetenz in den Bildungsstandards der KMK

Lesekompetenz als die Fähigkeit, kontinuierliche, diskontinuierliche sowie multimediale Texte zu verstehen, ist eine Schlüsselqualifikation, die für eine gelingende Lebensführung unerlässlich ist. Sie ist nicht nur für den schulischen Wissenserwerb und das lebenslange, selbstbestimmte Lernen eine wichtige Voraussetzung, sondern spielt auch für die Teilnahme an der Kommunikation über gesellschaftlich relevante Themen eine zentrale Rolle (vgl. Arbeitsgruppe „Internationale Vergleichsstudie“, 2007; Bos et al., 2003). In den Bildungsstandards der Kultusministerkonferenz werden im Kompetenzbereich Lesen aber

nicht nur die Lesefertigkeit und das sinnverstehende Lesen, sondern auch Aspekte des genießenden Lesens, des Leseinteresses und der kritischen Auseinandersetzung mit Texten angesprochen (vgl. KMK, 2005a, S. 9).

Die Standards beziehen sich primär auf Kompetenzen, die durch den Umgang mit Texten erworben werden sollen und die für den Umgang mit Texten notwendig sind. Ferner werden einige wissensorientierte Standards aufgeführt, die sich unter anderem auf Textsortenkenntnisse beziehen (vgl. auch Köster & Rosebrock, 2009). Die insgesamt 26 Standards im Kompetenzbereich *Lesen – mit Texten und Medien umgehen* sind vier Kategorien zugeordnet (vgl. KMK, 2005a, S. 11f.):

- über Lesefähigkeiten verfügen
- über Leseerfahrungen verfügen
- Texte erschließen
- Texte präsentieren.

Aus dem breiten Verständnis der Lesekompetenz der Bildungsstandards ergeben sich eine Vielzahl von lese- und medienbezogenen Standards, die allerdings „nur vergleichsweise vage und generelle Vorgaben liefern“ (Köster & Rosebrock, 2009, S. 104) und nur sehr eingeschränkt voneinander abgegrenzt sind (vgl. auch Köster, 2008). Die hieraus resultierenden Konsequenzen für die Konstruktbeschreibung und die Überführung der Bildungsstandards in Testaufgaben werden in den folgenden Abschnitten näher erläutert.

Lesekompetenz als Konstrukt

Das *Literacy*-Konzept versteht Lesekompetenz als eine Fähigkeit, die „erforderlich ist für die Bewältigung der charakteristischen Kommunikations- und Handlungsanforderungen, denen ein [...] Gesellschaftsteilnehmer in seinem Alltag und Beruf begegnet“ (Hurrelmann, 2007, S. 21). Vor allem im Rahmen der Lesesozialisationsforschung und Literaturdidaktik werden darüber hinaus Aspekte von Lesekompetenz erörtert, die im *Literacy*-Konzept nicht oder nur am Rande berücksichtigt sind. Unter Bezug auf bildungstheoretische Diskurse steht dabei nicht nur die instrumentelle Funktion des Lesens, etwa als Mittel für den Zweck des Lernens, im Zentrum der Aufmerksamkeit, sondern es wird zudem der Beitrag des Lesens zur Persönlichkeitsbildung thematisiert. So setzt Hurrelmann (2007, S. 22f.) Lesen mit der Herausbildung von ästhetischer und sprachlicher Sensibilität, Moralentwicklung und Empathiefähigkeit, Fremdverstehen und Teilhabe am kulturellen Gedächtnis in Bezug. Neben kognitiven und motivationalen Aspekten wird damit auch die emotionale Beteiligung als Facette der Lesekompetenz angesehen sowie weiterhin die Bereitschaft und Fähigkeit zur Anschlusskommunikation, also zum Diskurs über das Gelesene, eingeschlossen.

Im Rahmen großer Schulleistungstudien wie PISA (Baumert et al., 2001) oder PIRLS/IGLU (Bos et al., 2007) wird jedoch zumeist ein pragmatisch-funktionales Verständnis von Lesekompetenz im Sinne der *Reading Literacy* zugrunde gelegt, wobei in der testdiagnostischen Überprüfung der Lesekompetenz die kognitive Dimension, also das Lese- oder Textverstehen, dominiert. Psychologische Theorien des Textverstehens beruhen auf der zentralen Annahme, dass der Verstehensprozess eine Konstruktionsleistung darstellt. Dies wurde auch im Rahmen der PISA-Studie des Jahres 2000 aufgegriffen: „Lesen ist keine passive Rezeption dessen, was im jeweiligen Text an Information enthalten ist, sondern aktive (Re-)Konstruktion der Textbedeutung. Die im Text enthaltenen Aus-

sagen werden aktiv mit dem Vor-, Welt- und Sprachwissen des Lesers verbunden. Die Auseinandersetzung mit dem Text lässt sich als ein Akt der Bedeutungsgenerierung verstehen, bei dem das Vorwissen der Leser und die objektive Textvorgabe interagieren“ (Artelt, Stanat, Schneider & Schiefele, 2001, S. 70f.).

Textverstehen kann als komplexe Interaktion textgeleiteter und (vor-)wissensgeleiteter kognitiver Informationsverarbeitung verstanden werden und umfasst Prozesse auf verschiedenen Ebenen. Hierarchieniedrig, und bei geübten Lesern automatisiert, sind Prozesse auf subsemantischer Ebene wie die Identifikation von Graphemen und Silben sowie ihre Zuordnung zu lautlichen Einheiten, aber auch das Dekodieren, also das Erfassen von Wortbedeutungen, sowie darüber hinaus die (Re-)Konstruktion der Satzbedeutung. Hierarchiehöhere Prozesse betreffen die Konstruktion eines umfassenden mentalen beziehungsweise Situationsmodells des Textinhalts. Häufig werden zudem drei Formen der mentalen Repräsentation von Bedeutung unterschieden, nämlich eine an der sprachlichen Oberfläche orientierte Repräsentation, eine propositionale Repräsentation, die Relationen von Prädikaten und Argumenten umfasst, sowie eine Repräsentation des Textinhalts als mentales oder Situationsmodell (vgl. Bremerich-Vos & Böhme, 2009).

Operationalisierung und Überführung in Testaufgaben

Wie bereits oben angesprochen, sind die Standards im Kompetenzbereich Lesen sehr heterogen und umfassen nicht nur Könnens-, sondern auch Kennensbeschreibungen. Zudem sind verschiedene Standards nicht klar voneinander abgrenzbar (vgl. Böhme, 2012; Köster & Rosebrock, 2009). Ferner sind zahlreiche Standards nicht oder nur mit erheblichem Aufwand in Testaufgaben überführbar.

Für einen Teil der Standards kann im Rahmen großer Schulleistungsstudien, wie des hier beschriebenen Ländervergleichs 2011, daher nicht überprüft werden, inwieweit die Schülerinnen und Schüler sie erreichen. So ist etwa die Mitwirkung bei Lesungen und Aufführungen (subsumiert unter *Texte präsentieren*) nicht mit einem standardisierten Testverfahren im Rahmen eines Large-Scale-Assessments prüfbar. Die Standards *Kinderliteratur kennen: Werke, Autoren und Autorinnen, Figuren, Handlungen* sowie *über Leseerfahrungen verfügen* ließen sich zwar durchaus mit Testaufgaben operationalisieren oder in Fragebögen erheben, dabei wäre es jedoch praktisch unmöglich, dem Gütekriterium der Testfairness gerecht zu werden: Teilgruppen der Schülerinnen und Schüler, die mit den ausgewählten Autoren oder Werken nicht vertraut sind, dafür aber andere Autoren und Werke kennen, wären benachteiligt. Im Fokus der Testung stünde ferner das (Vor-)Wissen und nicht die Lesekompetenz der Kinder. Als nicht testbar erscheint ferner der auf elaborative Prozesse zielende Standard *lebendige Vorstellungen beim Lesen und Hören literarischer Texte entwickeln*, da kaum objektiv zu entscheiden ist, wann Vorstellungen als lebendig zu werten sind. Zudem erscheint es fraglich, ob bildliche Vorstellungen mit verbalen Verfahren, womöglich sogar schriftlich, valide erfasst werden können.

Für die Umsetzung in Testaufgaben kommen daher ausschließlich solche Standards in Betracht, für die präzise Konstruktdefinitionen vorliegen oder auf der Basis vorhandener Theorien erarbeitet werden können und für die eine Operationalisierung gemäß den Gütekriterien für standardisierte Testungen möglich ist (vgl. z. B. Rost, 2004). Hierbei handelt es sich im Wesentlichen um Standards der Kategorie *Texte erschließen*; hinzu kommt der als *über Lesefähigkeiten verfügen*

bezeichnete, auf die Konstruktion eines Situationsmodells zielende zentrale Standard *altersgemäße Texte sinnverstehend lesen*.

Für die Operationalisierung dieser Standards kann auf die vielfältigen Erfahrungen aus großen internationalen Schulleistungsstudien wie PISA oder PIRLS/IGLU zurückgegriffen werden. In diesen Studien spielt die Überprüfung der Lesekompetenz stets eine zentrale Rolle, so dass umfangreiche Hinweise und Empfehlungen zur Entwicklung von Testaufgaben im Bereich Lesen vorliegen (vgl. z. B. Mullis, Martin & Kennedy, 2004), an die bei der Operationalisierung der Lesekompetenz für die Überprüfung des Erreichens der Bildungsstandards angeknüpft werden konnte (vgl. Bremerich-Vos & Böhme, 2009).

Beschreibung der im Ländervergleich 2011 eingesetzten Leseaufgaben

Zur Überprüfung der Lesekompetenz der Schülerinnen und Schüler wurden im Ländervergleich 2011 neun Aufgabenblöcke eingesetzt (vgl. Kapitel 12.1). Diese Blöcke beinhalteten jeweils ein bis zwei Leseaufgaben, die aus einem Stimulustext und mehreren schriftlich vorgegebenen und auch schriftlich zu bearbeitenden Items bestanden. Insgesamt kamen elf Leseaufgaben zum Einsatz, die in Summe 80 Items umfassten.

Im Test wurden sowohl kontinuierliche literarische Texte als auch kontinuierliche und diskontinuierliche Sachtexte eingesetzt. Die literarischen Texte bezogen sich auf ein breites Themenspektrum, das der Lebenswelt von Grundschulkindern entstammt und ihren Interessen entgegenkommt. Um verschiedene Textsorten zu berücksichtigen, befanden sich unter den sechs literarischen Texten ein Gedicht und eine Fabel. Die kontinuierlichen Sachtexte thematisierten überwiegend die Lebensweise verschiedener Tiere oder deren Erforschung. Ferner wurde ein diskontinuierlicher Sachtext eingesetzt, bei dem es sich um eine Tabelle handelte.

Wie im Kompetenzbereich Zuhören hatte auch mehr als die Hälfte der Items zu den Lesetexten ein geschlossenes Format. Zusätzlich wurden 15 halboffene Items eingesetzt, bei denen die Schülerinnen und Schüler um Kurzantworten oder Zu- beziehungsweise Umordnungen vorgegebener Sachverhalte gebeten wurden. Der Anteil an offenen Items, die eine längere, selbstständige Schülerantwort verlangten, war im Kompetenzbereich Lesen mit 22 Items sowohl absolut als auch im Verhältnis zur Gesamtzahl der vorgelegten Items höher als im Kompetenzbereich Zuhören.

2.1.5 Beschreibung der orthografischen Kompetenz und ihrer Operationalisierung

Die orthografische Kompetenz in den Bildungsstandards der KMK

Unter der Bezeichnung *richtig schreiben* ist neben der Fähigkeit zur Textproduktion auch die orthografische Kompetenz der Schülerinnen und Schüler Bestandteil des Kompetenzbereichs *Schreiben*. Die Zeichensetzung und die Groß- und Kleinschreibung sind ergänzend im Bereich *Sprache und Sprachgebrauch untersuchen* verortet.

So sinnvoll die Integration der Orthografie in den Bereich *Schreiben* für die unterrichtliche Praxis ist, so problematisch ist sie für die testdiagnostische Umsetzung (vgl. Böhme, 2012). Vor allem zwei Gründe sprechen dafür, ortho-

grafische Fähigkeiten separat zu erheben: Zum einen liegt empirische Evidenz dafür vor, dass sich die orthografische Kompetenz von Grundschulkindern als eigenständiges Konstrukt erfassen lässt (Bremerich-Vos et al., 2009, S. 210). Zum anderen unterscheiden sich die von Schülerinnen und Schülern in einer Testsituation verfassten Texte beträchtlich hinsichtlich ihrer Länge und ihrer sprachlichen Komplexität. Dies betrifft auch die Anzahl und Auswahl der in Texten realisierten orthografischen Phänomene. So ist die Wahrscheinlichkeit, dass ein Kind grammatikalische Fehler macht, größer, wenn es vergleichsweise lange Texte mit komplexeren Satzkonstruktionen schreibt, als wenn es sich auf das Schreiben eines kurzen Textes mit vergleichsweise knappen Hauptsätzen beschränkt. Kinder, die in ihren Texten gleich viele orthografische und grammatikalische Fehler machen, verfügen also nicht zwingend über dasselbe Maß an Kompetenz in diesem Bereich. Alle Fehler, die ein Kind macht, müssten anhand der Zahl und der Qualität der im jeweiligen Text *möglichen* Fehler relativiert werden. Dies ist im Rahmen von groß angelegten Schulleistungsstudien jedoch nicht umsetzbar.

Die orthografische Kompetenz der Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe wird in den Bildungsstandards wie folgt beschrieben:

Die Kinder verfügen über grundlegende Rechtschreibstrategien. Sie können lautentsprechend verschriften und berücksichtigen orthografische und morphematische Regelungen und grammatisches Wissen. Sie haben erste Einsichten in die Prinzipien der Rechtschreibung gewonnen. Sie erproben und vergleichen Schreibweisen und denken über sie nach. Sie gelangen durch Vergleichen, Nachschlagen im Wörterbuch und Anwenden von Regeln zur richtigen Schreibweise. Sie entwickeln Rechtschreibgespür und Selbstverantwortung ihren Texten gegenüber. (KMK, 2005a, S. 8)

Unter der Überschrift *richtig schreiben* werden die folgenden Kompetenzaspekte benannt:

- geübte, rechtschreibwichtige Wörter normgerecht schreiben
- Rechtschreibstrategien verwenden: Mitsprechen, Ableiten, Einprägen
- Zeichensetzung beachten: Punkt, Fragezeichen, Ausrufezeichen, Zeichen bei wörtlicher Rede
- über Fehlersensibilität und Rechtschreibgespür verfügen
- Rechtschreibhilfen verwenden: Wörterbuch nutzen, Rechtschreibhilfen des Computers kritisch nutzen
- Arbeitstechniken nutzen: methodisch sinnvoll abschreiben, Übungsformen selbstständig nutzen, Texte auf orthografische Richtigkeit überprüfen und korrigieren

Orthografische Kompetenz als Konstrukt

Im Rahmen der deutschsprachigen Orthografiedidaktik wird seit geraumer Zeit mit Modellen von Stufen³ des Erwerbs der orthografischen Kompetenz gearbeitet. Der Stufenbegriff bedeutet in diesem Zusammenhang aber nicht, dass im Erwerbsprozess klar abgrenzbare, strikt aufeinanderfolgende Entwicklungsstufen beobachtet werden können, sondern es handelt sich vielmehr um Zugriffsweisen

3 Der Begriff „Stufe“ bezieht sich hierbei auf Vorstellungen zur Entwicklung der orthografischen Kompetenz, nicht auf die Stufen eines Kompetenzstufenmodells.

beziehungsweise Strategien, die sich zumeist nach und nach entfalten, zum Teil aber auch parallel verfügbar sind. Für die Grundschule ist die Unterscheidung einer alphabetischen von einer orthografischen Stufe zentral⁴ (vgl. z. B. Böhme & Bremerich-Vos, 2009).

Nach Schulbeginn lernen die Kinder sehr schnell, den kontinuierlichen Lautstrom nicht nur in Silben, sondern auch in einzelne Laute (Phoneme) zu zerlegen und diese mit bestimmten Buchstaben (Graphemen) zu verknüpfen. Sind diese Laut-Buchstaben-Beziehungen bekannt, dann schreiben die Schülerinnen und Schüler so, wie sie (mehr oder weniger dialektal) sprechen. Man könnte sagen, dass sie nach der Maxime verfahren: „Beachte die gesprochene Lautfolge und schreibe für jeden Laut einen Buchstaben!“ (vgl. May, 2002, S. 29). Die Beherrschung von Laut-Buchstaben-Beziehungen auf der Ebene der Basisgrapheme ist kennzeichnend für die *alphabetische Stufe*.

Auf der *orthografischen Stufe* setzen sich die Kinder mit dem Phänomen der Dehnung (z. B. *Bohnen*), Schärfung (z. B. *Ball*), Umlautableitung (z. B. *Bäcker*), Auslautverhärtung (z. B. *gelb*) und mit „Merkelementen“ wie [v] bei *Vater*, [x] bei *Hexe* und so weiter auseinander (vgl. May, 2002, S. 30). Sie orientieren sich mehr und mehr am Stammprinzip (*Wand* wegen *Wände*). Auch die Großschreibung von Substantiven, insbesondere von Konkreta, gelingt in wachsendem Maße, ebenso die Abgrenzung von Sätzen, unter anderem im Zusammenhang mit der direkten Rede. Hier kommt syntaktisches, auf den Satzbau bezogenes Wissen ins Spiel.

Operationalisierung und Überführung in Testaufgaben

Bei der Überführung der Bildungsstandards in Testaufgaben wurde die Mehrheit der oben aufgeführten Standards einbezogen. Besonderer Wert wurde auf die Überprüfung der Fähigkeit gelegt, *rechtschreibwichtige Wörter normgerecht zu schreiben*. Welche Wörter hier im Einzelnen relevant sind, ist nicht leicht zu entscheiden. In den Ländern der Bundesrepublik Deutschland sind Grundwortschätze, auf die man sich aktuell beziehen könnte, rar. Für die Auswahl der Testwörter wurde dennoch auf einige, darunter auch ältere, Grundwortschätze zurückgegriffen. Weiterhin wurden Wörter aus bewährten standardisierten Tests sowie Wörter aus dem von Naumann (1999) zusammengestellten Orientierungswortschatz verwendet.

Die in den Standards angesprochenen *Rechtschreibstrategien* beziehen sich auf die Laut-Buchstaben- beziehungsweise Phonem-Graphem-Korrespondenz (Mitsprechen), auf Dehnung, Schärfung, Umlautung und Auslautverhärtung (Ableiten) und auf Lernwörter, die zum Beispiel spezielle Grapheme wie <v> oder <x> enthalten (Einprägen). Schreibt ein Kind beispielsweise *want* statt *Wand*, so kann ihm korrektes „Mitsprechen“ attestiert werden. Es schreibt zwar „alphabetisch“, aber noch nicht „orthografisch“, da es Großschreibung und Auslautverhärtung nicht berücksichtigt hat. Folgt man derzeit weithin akzeptierten Theorien des Rechtschreiberwerbs (vgl. u. a. May, 2002; Scheele, 2006; Thomé, 2003), dann sollten Schülerinnen und Schüler der vierten Jahrgangsstufe bereits weitgehend die Phonem-Graphem-Zuordnung bewältigen. Dem wurde insofern Rechnung getragen, als im Test Wörter überwiegen, bei denen orthogra-

4 Frühere Stufen – die präliterale-symbolische sowie die logographemische Stufe – fallen in den Elementarbereich und sind daher nicht Gegenstand der Überprüfung der Bildungsstandards für den Primarbereich.

fische Besonderheiten zu beachten sind. Was speziell die Auslautverhärtung und Umlautung angeht, so wurde anhand einiger Aufgaben zusätzlich überprüft, ob sie nicht nur prozedural beherrscht werden, sondern ob die jeweilige Schreibung auch explizit begründet werden kann.

Der Standard zu *Fehlersensibilität* kann anhand von Korrekturaufgaben überprüft werden. Die Schülerinnen und Schüler werden dabei aufgefordert, fehlerhafte Wörter zu markieren und die korrekte Version hinzuschreiben.

Die Fähigkeit der Kinder, *Wörterbücher und Computerhilfen* zu nutzen, konnte aus untersuchungspraktischen Gründen nicht überprüft werden. Auch die Nutzung der Arbeitstechniken *Abschreiben* und *selbstständiges Üben* wurde nicht überprüft. Das Abschreiben hätte zu viel Zeit gekostet und Übungsaufgaben sind keine Testaufgaben. Erhoben wurde aber die Fähigkeit der Kinder, eine weitere Arbeitstechnik anzuwenden, nämlich – wie bereits erwähnt – Texte auf orthografische Korrektheit hin zu überprüfen und zu korrigieren.

Da es stets viele Möglichkeiten gibt, ein Wort falsch zu schreiben, können Fehler als Indikatoren für spezifische kognitive Prozesse gedeutet und zum Gegenstand qualitativer Fehleranalysen gemacht werden. Im Rahmen einer solchen Analyse werden einzelne Grapheme beziehungsweise Graphemgruppen als Lupenstellen, das heißt als Repräsentanten bestimmter Problembereiche verstanden. Für die Testung wurde in Anlehnung an die Aachener Förderdiagnostische Rechtschreibfehler-Analyse (AFRA)⁵ eine Reihe von Fehlerkategorien verwendet, die sich unter anderem auf folgende Aspekte beziehen:⁶

1. Eine falsche *Graphem-Auswahl* (GA) liegt vor, wenn das relevante Graphem bezogen auf die Standardsprache keine lauttreue Verschriftung des entsprechenden Phonems darstellt (z. B. *schlüm* statt *schlimm*).
2. Ein Fehler in der *Graphemfolge* (GF) bezieht sich auf die Wortdurchgliederung und betrifft beispielsweise die Auslassung oder Hinzufügung eines Graphems sowie die Vertauschung von Graphemen (z. B. *sthen* statt *stehen*).
3. *Spezielle Grapheme* (SG) beziehen sich auf die Verwendung ungewöhnlicher Grapheme und auf einschlägige Fehler (z. B. bei der Schreibung des Lauts [ng] wie in *Bangk* oder bei der Schreibung von [kw] wie in *Kwelle* statt *Quelle*).
4. In der *Mehrheit* der Fälle wird die *Vokallänge* (VL+) nicht markiert, im Normalfall erfolgt eine Schreibung mit dem einfachen Vokalbuchstaben⁷. Falsch markiert ist die Länge zum Beispiel, wenn ein Dehnungs-h verwendet wurde (z. B. *Dahme* statt *Dame*).
5. In der *Minderheit* der Fälle muss die *Vokallänge* (VL-) markiert werden. Fehler sind hier Auslassungen des Dehnungs-h (z. B. *Sane* statt *Sahne*). Wörter mit doppeltem Vokalbuchstaben kamen im Test nicht vor.
6. Das Phänomen der *Vokalkürze* (VK) bezieht sich darauf, dass in der Regel dann, wenn im Stamm auf einen kurzen betonten Vokal ein einzelner Kon-

5 Aachener Förderdiagnostische Rechtschreibfehler-Analyse (AFRA, Herné & Naumann, 2002). Andere Instrumente wie DoSe (Löffler & Meyer-Schepers, 1992), der sprachsystematische Rechtschreibtest (Voss, Blatt & Kowalski, 2007) oder die Hamburger Schreibprobe (May, 2002) sind im Kern, wenn auch nicht in den Details, vergleichbar (vgl. Valtin, Badel, Löffler, Meyer-Schepers & Voss, 2003; Voss, Blatt & Kowalski, 2007).

6 Bei den hier aufgeführten Wörtern handelt es sich um Beispiele, die nur teilweise aus dem Testinstrument stammen, das zur Überprüfung des Erreichens der Bildungsstandards eingesetzt wurde.

7 Dies trifft nur auf die langen Vokale [a:], [e:], [o:] und [u:] zu; ein lang gesprochenes [i:] wird in der Mehrheit der Fälle als <ie> verschriftet.

- sonant folgt, der Konsonantenbuchstabe doppelt geschrieben wird (z. B. fälschlich *Waser* statt *Wasser*).⁸
7. Die *vokalische Ableitung* (VA) betrifft Fälle wie *Bäcker* (wegen *backen*), fälschlich *Becker* geschrieben.
 8. Die *konsonantische Ableitung* (KA) betrifft das Phänomen der Auslautverhärtung (z. B. fälschlich *liep* statt *lieb*).
 9. Die Kategorie der *häufigen Morpheme* (HM) bezieht sich auf die Schreibung geläufiger Prä- und Suffixe wie *ver-* und *-ig* (z. B. *ferarbeiten* statt *verarbeiten*, *sonnich* statt *sonnig*).
 10. Die Kategorie der *Morphemgrenze* (MG) bezieht sich auf morphologisches Strukturwissen beziehungsweise die Wortdurchgliederung (z. B. *Spieleute* statt *Spielleute*).
 11. Die Kategorie *Groß- und Kleinschreibung* (GK) bezieht sich auf den Satz-anfang und auf Konkreta (z. B. *tanne* statt *Tanne*, aber auch *Wir kommen Heute*).

Die genannten Fehlerkategorien können den oben dargestellten Stufen des Rechtschreiberwerbs wie folgt zugeordnet werden:

Fehlerelemente auf der alphabetischen Stufe: Fehler in der Graphemauswahl (GA) sowie in der Graphemfolge (GF) können als mangelnde Beherrschung von Laut-Buchstaben-Beziehungen interpretiert werden. Solche Fehler weisen darauf hin, dass die alphabetische Strategie noch nicht sicher gemeistert wird.

Fehlerelemente auf der orthografischen Stufe: Fehlerelemente, die der orthografischen Stufe zugeordnet werden können, sind jene Lupenstellen, die sich auf spezielle Grapheme (SG), auf die nicht notwendige (VL+) oder erforderliche Markierung der Vokallänge (VL-), die Vokalkürze (VK) sowie auf vokalische (VA) und konsonantische Ableitungen (KA) beziehen. Ferner lassen sich dieser Stufe Lupenstellen zuordnen, die sich auf die explizite Berücksichtigung morphologischer Aspekte beziehen. Dies sind Fehler in Bezug auf häufige Morpheme (HM) und Morphemgrenzen (MG). Hinzu kommen Fehler der Groß- und Kleinschreibung (GK) auch in Fällen, bei denen die Syntax beachtet werden muss.

Beschreibung der im Ländervergleich 2011 eingesetzten Orthografieaufgaben

Zur Überprüfung der orthografischen Kompetenz der Schülerinnen und Schüler wurden im Ländervergleich 2011 zwölf Aufgabenblöcke eingesetzt (vgl. Kapitel 12.1). Diese Blöcke beinhalteten jeweils zwei bis drei der zwölf Orthografieaufgaben. Insgesamt umfassten die Orthografieaufgaben 152 fast ausschließlich halboffene Items.

Die eingesetzten Aufgaben lassen sich drei verschiedenen Gruppen zuordnen. Etwas mehr als 40 Prozent der Items gehören zu *Lückensatzdiktaten*. Dabei handelt es sich um einzelne Sätze, in denen jeweils ein bis zwei Wörter fehlen. Diese Sätze wurden den Kindern vollständig vorgelesen und sie wurden gebeten, die fehlenden Wörter zu ergänzen. Lückensatzdiktate prüfen somit produktive bezie-

⁸ In der amtlichen Regelung der Orthografie heißt es, dass der Buchstabe für einen Konsonanten doppelt geschrieben wird, wenn im Stamm auf einen kurzen betonten Vokal nur dieser Konsonant folgt. In einer alternativen Regelformulierung ist der Begriff des Silbengelenks zentral. Mit der Kategorie „Vokalkürze“ wird auf die amtliche Regelung Bezug genommen.

hungsweise aktive Anteile der orthografischen Kompetenz. Dieser Aufgabentyp hat verschiedene Vorteile (vgl. Böhme & Bremerich-Vos, 2009). So ist bei dieser Diktatform der Schreibaufwand für die Kinder begrenzt und Unterschiede im Schreibtempo fallen weniger stark ins Gewicht als bei Textdiktaten. Das Schreiben von sicher beherrschten Wörtern wie Artikeln und Pronomen entfällt. Damit wird gewährleistet, dass die begrenzte Testzeit gezielt und effizient genutzt wird. Außerdem kann die Aufmerksamkeit der Schülerinnen und Schüler ungeteilt der Rechtschreibung gelten, da nicht parallel Wortgruppen oder Sätze erinnert werden müssen, wie es üblicherweise bei Textdiktaten der Fall ist.

Die zweite Gruppe von Aufgaben macht ebenfalls etwas mehr als 40 Prozent der Orthografieitems aus und bezieht sich auf die Korrektur vorgegebener Schreibungen. *Korrekturaufgaben* greifen auf passive Aspekte der orthografischen Kompetenz zurück und dienen der Erfassung von Fehlersensibilität sowie der Überprüfung, ob die Kinder Texte auf orthografische Richtigkeit kontrollieren und korrigieren können.

Bei den verbleibenden 15 Prozent der Items handelt es sich um unterschiedliche Aufgabentypen, die mehrheitlich auf die *Anwendung von Strategiewissen* abzielen, beispielsweise im Kontext konsonantischer Ableitungen.

2.2 Beschreibung der im Fach Mathematik untersuchten Kompetenzen

Alexander Roppelt und Kristina Reiss

Dieses Teilkapitel beschreibt die im Ländervergleich 2011 überprüften Kompetenzen im Fach Mathematik. Zunächst werden die Bildungsstandards im Fach Mathematik für den Primarbereich und das darin zugrunde gelegte Modell mathematischer Kompetenzen vorgestellt. Um das Erreichen der Bildungsstandards messbar zu machen, muss dieses beschreibende Modell in ein psychometrisches Modell überführt werden, das den Rahmenbedingungen und Zielen großer Schulleistungsstudien Rechnung trägt. Diese Aufgabe ist untrennbar mit der Frage verbunden, mit welcher dimensional Struktur mathematische Kompetenzen empirisch am besten beschrieben werden können. Im zweiten Teil dieses Teilkapitels folgt deshalb eine Darstellung konzeptueller und methodischer Überlegungen sowie empirischer Befunde zur Struktur der mathematischen Kompetenzen, aus denen sich das Vorgehen im Ländervergleich 2011 ableitet.

2.2.1 Die Bildungsstandards im Fach Mathematik für den Primarbereich

Bei den Bildungsstandards im Fach Mathematik für den Primarbereich (KMK, 2005b) handelt es sich formal um eine präskriptive beziehungsweise normative Setzung, die definiert, welche mathematischen Kompetenzen Schülerinnen und Schüler in Deutschland bis zur vierten Jahrgangsstufe erworben haben sollen. Gleichzeitig sollen die Bildungsstandards zur Weiterentwicklung des Mathematikunterrichts beitragen (z. B. Walther, Selter & Neubrand, 2007). Ihre Einführung erfolgte unter anderem mit der Zielstellung, das unterrichtliche Handeln von Lehrkräften stärker auf den Ausbau von Kompetenzen auszurichten (Kompetenzorientierung). Im Fach Mathematik bezieht sich dies insbesondere auf die allgemeinen mathematischen Kompetenzen, die mit den Bildungsstandards stärker als bislang in den Fokus rücken sollen. Sie haben einen inhaltsübergreifenden, prozesshaften Charakter. Die Aufmerksamkeit auf sie zu richten soll zur Folge haben, dass in der Schule nicht nur isolierte, kontextgebundene Kenntnisse und Fertigkeiten vermittelt werden, die lediglich zum Lösen eines schulischen Kanons von typischen Mathematikaufgaben ausreichen. Aufgebaut werden soll vielmehr „intelligentes Wissen“, also ein „wohlorganisiertes, disziplinär, interdisziplinär und lebenspraktisch vernetztes System von flexibel nutzbaren Fähigkeiten, Fertigkeiten, Kenntnissen und metakognitiven Kompetenzen“ (Weinert, 2000, zitiert nach Helmke, 2010, S. 43). Vor diesem Hintergrund ist die Förderung mathematischer Kompetenzen als ein Teil grundlegender Allgemeinbildung zu verstehen, die zu vermitteln Auftrag der Grundschule ist. Der Mathematikunterricht geht dabei von frühen mathematischen Alltagserfahrungen aus, die zu grundlegenden mathematischen Kompetenzen ausgebaut werden sollen. Diese bilden das Fundament für den anschließenden Kompetenzaufbau in weiterführenden Schulen.

Wie Walther, Selter und Neubrand (2007) erläutern, existiert „in der Mathematikdidaktik eine lange Tradition der Auseinandersetzung mit Bildungszielen und zentralen mathematischen Anforderungen“ (S. 25). In diese Diskussion reihen sich die Zielsetzungen ein, die in den Bildungsstandards formuliert sind. Die

bildungstheoretischen Grundsätze der Bildungsstandards für das Fach Mathematik fußen dabei insbesondere auf dem häufig zitierten Grundbildungskonzept von Winter (1995). Demnach soll es der Mathematikunterricht ermöglichen, die folgenden drei Gesichtspunkte der Mathematik als Grunderfahrung zu erleben:

1. Erscheinungen der Welt um uns [...] in einer spezifischen Art wahrzunehmen und zu verstehen,
2. mathematische Gegenstände und Sachverhalte [...] als eine deduktiv geordnete Welt eigener Art kennen zu lernen und zu begreifen,
3. in der Auseinandersetzung mit Aufgaben Problemlösefähigkeiten, die über die Mathematik hinaus gehen (heuristische Fähigkeiten), zu erwerben. (S. 38)

Diese drei Grunderfahrungen bezeichnet Blum (2006, S. 21) als „anwendungs-, struktur- und problemorientierten Aspekt der Mathematik“.

Die Bildungsstandards beschränken sich auf die Beschreibung fachlicher Kernziele des Mathematikunterrichts. Soziale oder personale Kompetenzen werden zwar nicht im Detail aufgeführt, aber dennoch ausdrücklich als Bestandteil von grundlegender Bildung genannt und sollen auch in der Auseinandersetzung mit fachlichen Inhalten erworben werden. Besonders der Förderung der allgemeinen mathematischen Kompetenzen wird zugeschrieben, den Schülerinnen und Schülern zugleich allgemeine Bildungsziele wie die genannten sozialen und personalen Kompetenzen näher zu bringen sowie eine positive Einstellung zum Fach zu begünstigen.

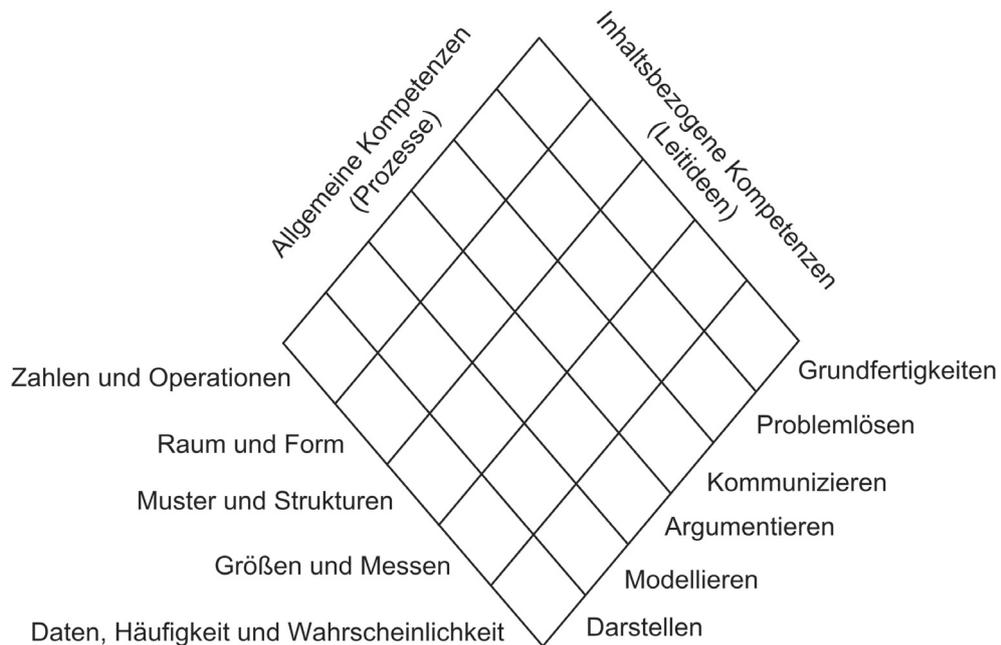
In den folgenden Abschnitten wird zunächst das theoretische Modell dargestellt, das die Bildungsstandards zur Beschreibung mathematischer Kompetenz heranziehen. Anschließend werden die einzelnen Komponenten des Modells genauer charakterisiert.

Kompetenzmodell der Bildungsstandards im Fach Mathematik

Die Bildungsstandards im Fach Mathematik beschreiben mathematische Kompetenzen, die sich anhand von drei „Klassen“ oder „Dimensionen“ kategorisieren lassen. Dabei handelt es sich um eine Prozess-, eine Inhalts- und eine Anspruchsdimension (vgl. Abbildung 2.3). Diese Unterscheidung basiert auf historischen Modellen kognitiver Fähigkeiten aus der Psychologie und wurde in ähnlicher Weise beispielsweise schon von Guilford (1967) vorgeschlagen. Als unmittelbares Vorbild für das Modell der Bildungsstandards der KMK im Fach Mathematik dienten die wegweisenden Standards des amerikanischen *National Council of Teachers of Mathematics* (NCTM, 2000) und das Modell der PISA-Studie (OECD, 2003). In den Bildungsstandards werden der Prozessdimension sechs *allgemeine mathematische Kompetenzen* zugeordnet. Die Inhaltsdimension wird durch fünf *Leitideen* oder *Kompetenzbereiche* beschrieben und die Anspruchsdimension unterscheidet drei sogenannte *Anforderungsbereiche*.

Alle drei Dimensionen sind untrennbar miteinander verwoben. Eine Mathematikaufgabe lässt sich stets in Bezug auf alle drei Dimensionen charakterisieren. Die Prozessdimension mit den allgemeinen mathematischen Kompetenzen soll zwar im Mittelpunkt des Mathematikunterrichts stehen, es ist jedoch keineswegs möglich, diese unabhängig von mathematischen Inhalten zu erwerben. Allerdings erlangen die allgemeinen Kompetenzen eine eigenständige Bedeutung, wenn

Abbildung 2.3: Modell mathematischer Kompetenz der Bildungsstandards im Fach Mathematik für den Primarbereich ohne Darstellung der Anforderungsbereiche



sie generalisiert werden, wenn sie also nicht mehr eng mit einzelnen konkreten Inhalten verbunden sind, sondern auch auf bislang unbekannte Inhalte angewendet werden können. Eine solche Generalisierung wird nur möglich, wenn die prozessbezogenen Kompetenzen bezogen auf ein breites inhaltliches Spektrum – einschließlich authentischer Anwendungen – erlernt und genutzt werden.

Der dreidimensionale Aufbau des Kompetenzmodells für den Primarbereich deckt sich mit der Struktur, die auch den Bildungsstandards für den Mittleren Schulabschluss und für den Hauptschulabschluss zugrunde liegt. Auch die Facetten der einzelnen Dimensionen werden im Primar- und im Sekundarbereich in ähnlicher Weise beschrieben (vgl. Blum, Drüke-Noe, Hartung & Köller, 2006). Die Modelle sind jedoch in etwas unterschiedlichen Traditionen verwurzelt und deshalb nicht vollständig deckungsgleich. So wird beispielsweise das Entnehmen von Informationen aus Sachtexten im Primarbereich der Kompetenz Modellieren, im Sekundarbereich jedoch dem Kommunizieren zugerechnet. Klassifiziert man Aufgaben nach den dabei beanspruchten allgemeinen und inhaltlichen Kompetenzen, können sich deshalb im Detail Unterschiede in Abhängigkeit davon ergeben, ob man vom Modell des Primarbereichs oder vom Modell des Sekundarbereichs ausgeht.

Allgemeine mathematische Kompetenzen (Prozesse)

Die allgemeinen mathematischen Kompetenzen sind kognitive Prozesse oder Operationen, die das mathematische Arbeiten über alle Inhaltsbereiche hinweg charakterisieren. Ihnen „ist eine herausragende Rolle bei der Entwicklung von auf Verständnis gegründeten inhaltlichen mathematischen Kompetenzen zugeordnet“ (Walther & Granzer, 2009, S. 116). Obwohl jede der in den Bildungsstandards beschriebenen allgemeinen Kompetenzen ihren eigenständigen Kern hat, las-

sen sie sich nicht trennscharf voneinander abgrenzen – die Übergänge sind fließend. Darüber hinaus werden bei den meisten mathematischen Tätigkeiten mehrere allgemeine Kompetenzen zugleich angesprochen. Beispielsweise sind zur Bearbeitung von Modellierungsaufgaben häufig auch Problemlöse- und Argumentationskompetenzen erforderlich. Eine Zuordnung von Testaufgaben zu nur *einer* allgemeinen Kompetenz ist deshalb höchstens bei sehr einfach strukturierten Aufgaben möglich. Dennoch lassen sich Testitems im Hinblick darauf unterscheiden, welche allgemeinen Kompetenzen damit schwerpunktmäßig angesprochen werden. Im Folgenden werden die allgemeinen mathematischen Kompetenzen auf Grundlage ihrer Charakterisierung in den Bildungsstandards für den Primarbereich (KMK, 2005b) genauer beschrieben.

Von *Problemlösen* wird gesprochen, wenn mathematische Kenntnisse, Fertigkeiten und Fähigkeiten genutzt werden, um Aufgaben zu lösen, deren Lösungsweg für das Kind nicht unmittelbar ersichtlich ist (vgl. auch Mayer & Hegarty, 1996). Um zu einer Lösung zu gelangen, müssen angemessene Strategien entwickelt und genutzt werden. Dies kann beispielsweise durch systematisches Probieren oder durch das Erkennen und Nutzen von Zusammenhängen geschehen.

Die allgemeine mathematische Kompetenz *Kommunizieren* betrifft nach dem Verständnis der Bildungsstandards für den Primarbereich vorrangig interaktive Lehr-Lern-Prozesse. Sie umfasst die Fähigkeit, das eigene Vorgehen und auch das Vorgehen der Lehrkraft oder der Mitschülerinnen und Mitschüler zu verstehen, zu beschreiben und zu reflektieren. Aufgaben sollen gemeinsam bearbeitet werden können, wobei hierfür erforderliche Vereinbarungen getroffen und eingehalten werden müssen. Darüber hinaus wird die sachgerechte Verwendung mathematischer Fachbegriffe und Zeichen als Aspekt des Kommunizierens verstanden.

Argumentieren bedeutet, dass mathematische Begründungen selbst gesucht oder die Aussagen anderer nachvollzogen, hinterfragt und auf Korrektheit geprüft werden. Hierfür ist es häufig erforderlich, mathematische Zusammenhänge zu erkennen. Dies birgt eine offenkundige Überschneidung mit der Kompetenz des Problemlösens. Während jedoch beim Problemlösen das Berücksichtigen von Zusammenhängen eher Mittel zum Zweck ist, um gegebene Probleme zu lösen, steht beim Argumentieren das Identifizieren von Zusammenhängen und das Anstellen von Vermutungen *an sich* im Zentrum der Tätigkeit, etwa um Begründungen für einen Sachverhalt angeben zu können.

Der instrumentelle, anwendungsbezogene Aspekt von Mathematik findet seinen unmittelbaren Niederschlag in der Kompetenz *Modellieren*. Sie ist gefordert, wenn reale Phänomene in die Sprache der Mathematik zu übersetzen sind. Häufig werden diese Phänomene in Texten und anderen Darstellungen der Lebenswirklichkeit präsentiert, aus denen die relevanten Informationen zunächst entnommen werden müssen. Ist eine mathematische Formulierung des Sachproblems gefunden, gilt es dieses innermathematisch zu lösen und die Ergebnisse schließlich wieder auf die Ausgangssituation zu beziehen. Darüber hinaus sollen die Schülerinnen und Schüler auch den umgekehrten Weg beschreiten und zu mathematischen Darstellungen wie Termen, Gleichungen oder Abbildungen Sachaufgaben formulieren können.

Eine besondere Herausforderung der Mathematik ist, dass sie zwar von den Phänomenen der realen Welt inspiriert ist, mathematische Objekte wie Quadrate, Zahlen oder Funktionen jedoch Gedankenkonstrukte sind und als solche nicht direkt beobachtet werden können. Obgleich sie sich darstellen lassen, sind derartige

Repräsentationen doch nicht identisch mit den Objekten selbst (Duval, 2006). Tatsächlich steht für dasselbe Objekt meist eine Reihe von teilweise sehr unterschiedlichen Darstellungen zur Verfügung. So kann eine natürliche Zahl beispielsweise durch die Anzahl der Elemente einer Menge, durch eine Markierung in einer Stellenwerttafel oder durch eine Ziffernfolge repräsentiert werden. Der Umgang mit und der Wechsel zwischen verschiedenen Darstellungen für dasselbe Objekt ist ein wichtiges Element mathematischen Arbeitens, das in den Bildungsstandards mit der Kompetenz *Darstellen* beschrieben wird. Je nach Problemstellung müssen geeignete Darstellungen entwickelt, ausgewählt und genutzt werden. Ferner sind unterschiedliche Repräsentationen miteinander zu vergleichen und zu bewerten.

Im Rahmen der Operationalisierung der Bildungsstandards durch das IQB wurde deutlich, dass sich Aufgaben, die ausschließlich mathematische Basisfertigkeiten erfordern, keiner der allgemeinen mathematischen Kompetenzen zuordnen lassen, die in den Bildungsstandards für den Primarbereich beschrieben sind. Basisfertigkeiten sind jedoch von großer Wichtigkeit, da „im Sinne eines kumulativen Wissenserwerbs nur von gesichertem, verfügbarem ‚Grundwissen‘ aus weiter gelernt werden kann“ (Walther & Granzer, 2009, S. 117). In Anlehnung an das Modell der Sekundarstufe wurde daher entschieden, die fünf ursprünglich formulierten prozessbezogenen Kompetenzen um *technische Grundfertigkeiten* zu ergänzen (IQB, 2008; Walther & Granzer, 2009; vgl. auch Winkelmann & Robitzsch, 2009, S. 171). Diese umfassen das Arbeiten und Rechnen mit Zahlen, Rechenausdrücken und geometrischen Elementen, den sinnvollen Gebrauch mathematischer Hilfsmittel wie Zirkel und Lineal sowie die verständige Nutzung der symbolischen und formalen Sprache in Arithmetik und Geometrie einschließlich ihrer Übersetzung in Worte der natürlichen Sprache und zurück.

Inhaltsbezogene mathematische Kompetenzen (Leitideen)

Die Bildungsstandards Mathematik gliedern die zu erwerbenden mathematischen Inhalte nach fünf Leitideen oder (inhaltlichen) Kompetenzbereichen. Diese Ideen ergeben sich aus den Phänomenen, die man in der Welt erkennen kann, wenn man sie aus der Perspektive der Mathematik betrachtet (Freudenthal, 1983). Man erkennt dort beispielsweise (An-)Zahlen und ihre Veränderungen (*Zahlen und Operationen*), ebene und räumliche Figuren und Gebilde (*Raum und Form*), Gesetzmäßigkeiten in geometrischen oder arithmetischen Mustern (*Muster und Strukturen*), vielfältige Quantifizierungen (*Größen und Messen*) sowie zufällige, aber in ihrer Gesamtheit doch Regeln folgende Ereignisse (*Daten, Häufigkeit und Wahrscheinlichkeit*). „Aus diesen Leitideen heraus haben sich die mathematischen Stoffgebiete [...] entwickelt. Leitideen und Stoffgebiete sind jedoch nicht identisch“ (Blum, 2006, S. 20). Die Gliederung der Inhalte nach Leitideen statt nach Stoffgebieten soll den vernetzten Charakter der Mathematik betonen und dazu beitragen, eine starre Unterteilung des Unterrichts und der Überprüfung des Lernerfolgs in Stoffgebiete zu überwinden. Die Leitideen sind relativ grobe Konzepte, deren Ränder unscharf sind. Es gibt Aspekte mathematischen Arbeitens, die Zugänge aus verschiedenen Leitideen umfassen und die sich folglich im Grenzbereich mehrerer Leitideen befinden. Werden derartige Aspekte pragmatisch einer dieser Leitideen zugerechnet, ist damit notwendigerweise eine gewisse Willkür verbunden. Im Gegensatz zu den allgemeinen Kompetenzen ist

dies bei den inhaltlichen Kompetenzen jedoch nicht die Regel. Sie lassen sich überwiegend einer einzigen, dominierenden Leitidee zuordnen.

Im Folgenden werden die fünf Leitideen jeweils knapp beschrieben. Ausführlichere Erläuterungen zu den Bestandteilen dieser inhaltlichen Kompetenzbereiche finden sich in den Bildungsstandards (KMK, 2005b) sowie in den einzelnen Kapiteln des Buches von Walther, van den Heuvel-Panhuizen, Granzer und Köller (2007), das auch zahlreiche Anregungen für den Unterricht enthält.

Die erste Leitidee ist mit *Zahlen und Operationen* (ZO) überschrieben. Sie nimmt insofern eine besondere Stellung unter den Inhaltsbereichen der Bildungsstandards ein, als die natürlichen Zahlen und das Operieren mit ihnen im Mittelpunkt des Mathematikunterrichts im Primarbereich stehen (Rasch & Schütte, 2007). Ein sicheres und verständiges Beherrschen der Arithmetik natürlicher Zahlen ist nicht nur Grundlage für praktisch alle weiteren Bereiche der Mathematik, sondern zugleich eine basale Kulturfertigkeit, die Voraussetzung für eine aktive gesellschaftliche Teilhabe ist. In den Bildungsstandards werden drei Teilbereiche der Leitidee *Zahlen und Operationen* beschrieben, die in enger Beziehung zueinander stehen. Der erste Teilbereich ist überschrieben mit „Zahldarstellungen und Zahlbeziehungen verstehen“ und zielt auf das Verstehen des dekadischen Stellenwertsystems ab. Der zweite Teilbereich „Rechenoperationen verstehen und beherrschen“ schließt nicht nur die Algorithmen arithmetischer Operationen und Hilfstechiken wie das Zerlegen von Zahlen ein, sondern er umfasst auch das gedächtnismäßige Beherrschen des kleinen Einspluseins und Einmaleins, deren Automatisierung eine Entlastung des Arbeitsgedächtnisses bei komplexeren Rechnungen bewirkt. Der dritte Teilbereich „in Kontexten rechnen“ unterstreicht die Bedeutung, die Sachbezüge für das Durchdringen der Grundrechenarten und ihrer Beziehungen untereinander besitzen (Rasch & Schütte, 2007). Er beinhaltet vornehmlich Aspekte, die das Modellieren als allgemeine mathematische Kompetenz mit den bereits genannten Bereichen der Leitidee *Zahlen und Operationen* verknüpfen. Zugleich wird erst durch das Rechnen mit Sachbezügen jene Verzahnung verwirklicht, die – der Kompetenzorientierung entsprechend – eine Anwendung der in der Schule erworbenen Fähigkeiten und Fertigkeiten in realen Kontexten ermöglicht.

Die Leitidee *Raum und Form* (RF) ist eng verbunden mit dem mathematischen Teilgebiet Geometrie, wobei Geometrie zu verstehen ist im Sinne Freudenthals als „grasping space“, also als ein „Fassen“ des Raumes (Freudenthal, 1973, S. 403). „Raum“ meint hier zunächst den realen Raum in dem wir leben und durch den wir uns bewegen. Diesen gilt es zu erforschen, zu durchdringen und zu verstehen. Die Mathematik unterstützt diesen Prozess und entwickelt daraus mathematische Modelle und Formalismen. Bezogen auf mathematische Objekte bezieht sich „Raum“ jedoch nicht nur auf den dreidimensionalen Anschauungsraum als Raum im engeren Sinne, sondern auch auf Ebenen, Geraden oder sogar Kugeloberflächen (Wollring & Rinkens, 2007, S. 119). Die Leitidee *Raum und Form* hat ihren Ursprung in Phänomenen der realen Welt, wie den Ähnlichkeiten und Unterschieden von Formen sowie den verschiedenen Repräsentationen oder Darstellungen, in denen dieselben Figuren auftreten. Den Raum zu fassen erfordert ein Verständnis der Eigenschaften von Objekten und deren relativer Position zueinander (OECD, 2003, S. 36). Die Bildungsstandards gliedern den Kompetenzbereich in vier Teilbereiche. Es sollen erstens geometrische Figuren und zweitens einfache geometrische Abbildungen wie Vergrößerungen oder Symmetrien erkannt, benannt und dargestellt werden können. Der dritte Teilbereich ist das Vergleichen und Messen von Flächen

und Volumina, wobei dies im Primarbereich noch stark orientiert an Tätigkeiten wie dem Zerlegen und Ausfüllen mit Einheitsflächen oder -würfeln geschieht. Während diese drei Teilbereiche eher die „klassische“ Geometrie widerspiegeln, verweist der vierte – „sich im Raum orientieren“ – auf die oben beschriebene erweiterte Charakterisierung als Fassen des Raumes.

Quer zu den klassischen mathematischen Teilbereichen liegt die Leitidee *Muster und Strukturen* (MS). Sie durchzieht und verbindet alle anderen inhaltlichen Kompetenzbereiche und bildet gewissermaßen den Kern mathematischen Arbeitens. So heben beispielsweise Wittmann und Müller (2007) hervor, dass „Mathematik als Wissenschaft von Mustern“ (S. 47) zu betrachten sei. Strukturen zu erkennen bedeutet, vom Einzelfall zu abstrahieren. Auf diese Weise wird Komplexität verringert, weil viele Einzelfälle durch eine allgemeinere Kategorie ersetzt werden können. Das Erkennen von Regelmäßigkeiten ist zugleich eng verbunden mit grundlegenden kognitiven Fähigkeiten. Insbesondere sind Kompetenzen im Bereich *Muster und Strukturen* verwandt mit quantitativem Schlussfolgern, aber auch mit dem allgemeinen schlussfolgernden Denken.⁹ Es liegt deshalb in der Natur mathematischen Arbeitens, dass einige Aufgabentypen im Bereich *Muster und Strukturen* in ähnlicher Form auch in Tests allgemeiner kognitiver Fähigkeiten vorkommen können.

Als Bestandteile der Leitidee *Muster und Strukturen* benennen die Bildungsstandards einerseits das allgemeine Erkennen, Beschreiben und Darstellen von Gesetzmäßigkeiten – beispielsweise in Zahlenfolgen, strukturierten Zahlendarstellungen oder geometrischen Mustern. Andererseits wird zusätzlich das Erkennen, Beschreiben und Darstellen von funktionalen Beziehungen als etwas speziellerer Teilbereich benannt und damit hervorgehoben. Entsprechend wird der Kompetenzbereich *Muster und Strukturen* auch als Vorläufer der Leitidee *Funktionaler Zusammenhang* ausgewiesen, wie sie in den Bildungsstandards der Sekundarstufe I beschrieben wird.

Im inhaltlichen Kompetenzbereich *Größen und Messen* (GM) werden Kenntnisse über das Vergleichen, Messen und Schätzen von Größen zusammengefasst. Der für diese Leitidee zentrale Prozess des Messens unterstützt das Begreifen der Welt, indem er sie der Berechnung zugänglich macht. Er übersetzt physikalische Quantitäten in die Sprache der Mathematik und ermöglicht auf diese Weise ihre Bearbeitung mit mathematischen Werkzeugen. Der Inhaltsbereich *Größen und Messen* hat somit einen direkten Bezug zur Lebenswelt. Innerhalb der Mathematik verbindet *Größen und Messen* die Gebiete Arithmetik und Geometrie. Im Einzelnen umfasst der Bereich den Umgang mit Größenbereichen der physikalischen Welt, wie etwa Längen, Volumina, Masse/Gewicht oder auch Zeitspannen, sowie mit Geldwerten. In Bezug auf diese Größenbereiche sollen neben den Standardeinheiten auch alltagsrelevante Repräsentanten sowie sachgerechte Messmethoden bekannt sein.

Bei vielen alltäglichen Prozessen lässt sich kein deterministischer Ausgang feststellen. Man spricht dann von zufälligen Ereignissen. Trotz der Unsicherheit, die mit dem Ausgang eines einzelnen solchen Ereignisses verbunden ist, las-

9 In der psychologischen Forschung hat sich als allgemeine Theorie kognitiver Kompetenzen in den letzten Jahren das sogenannte Cattell-Horn-Carroll-Modell durchgesetzt (McGrew & Hessler, 1995; McGrew, 2009). Darin werden drei hierarchische Ebenen oder Strata unterschieden: „Enge“ und „breite“ kognitive Fähigkeiten auf der ersten und zweiten Ebene sowie eine generelle kognitive Fähigkeit (g-Faktor) auf der höchsten Ebene. Die beiden „breiten“ Fähigkeiten, denen Kompetenzen im Bereich *Muster und Strukturen* besonders nahe stehen, werden in diesem Modell *Quantitative Reasoning* (Gq) und *Fluid Intelligence* (Gf) genannt.

sen sich häufig für eine Gesamtheit von Ereignissen Regelmäßigkeiten identifizieren, die sich mit mathematischen Mitteln beschreiben lassen (vgl. Hasemann, Mirwald & Hoffmann, 2007). Aus diesem Phänomen erwächst die Leitidee *Daten, Häufigkeit und Wahrscheinlichkeit* (DHW). Sie ist eng verbunden mit dem mathematischen Teilgebiet der Stochastik. Erst seit einigen Jahren nimmt dieses Teilgebiet vermehrt Raum in Lehrplänen und in der Unterrichtspraxis ein. Gleichwohl ist die Bedeutung eines verständigen Umgangs mit Daten und zufälligen Ereignissen für eine informierte Teilhabe an einer entwickelten Gesellschaft, in der die Verwendung von Statistiken eine Selbstverständlichkeit geworden ist, von zentraler Bedeutung. Im Primarbereich gliedern die Bildungsstandards diese Leitidee in zwei Teilbereiche, die der typischen Unterscheidung von Statistik und Wahrscheinlichkeitsrechnung entsprechen. Hinsichtlich des ersten Bereichs beschränkt man sich in der Primarstufe im Allgemeinen auf die Grundlagen der deskriptiven Statistik. Dies umfasst sowohl das aktive Beschreiben von Daten als auch das Entnehmen von Informationen aus Tabellen, Schaubildern und Diagrammen. Bezüglich der Wahrscheinlichkeitsrechnung wird in den Standards die Erwartung formuliert, dass die Kinder in der Lage sein sollen, die Wahrscheinlichkeiten von Ereignissen in Zufallsexperimenten zu vergleichen. Sie sollen Grundbegriffe wie „sicher“, „unmöglich“ oder „wahrscheinlich“ kennen und die Gewinnchancen bei einfachen Zufallsexperimenten einschätzen können. Eine formale Berechnung von Wahrscheinlichkeiten wird nicht erwartet und wäre aufgrund der nicht vorhandenen Vorkenntnisse, wie etwa dem Umgang mit Brüchen, im Allgemeinen auch nicht möglich.

Anforderungsbereiche

Die Anforderungsbereiche sollen auf theoretischer Ebene beschreiben, welche Qualität und Komplexität kognitive Leistungen aufweisen, die beim Lösen einer Aufgabe erbracht werden müssen. Die Bildungsstandards im Fach Mathematik unterscheiden dabei drei Ebenen des Anspruchs: *Reproduzieren*, *Zusammenhänge herstellen* sowie *Verallgemeinern und Reflektieren*. Der Anforderungsbereich einer Aufgabe kann nicht vollständig losgelöst von der Person beurteilt werden, welche diese Aufgabe bearbeitet, denn was bei dem einen Kind aufgrund mangelnder Erfahrung die Entwicklung einer Lösungsstrategie durch Verallgemeinern erfordert, kann für ein anderes Kind bereits zur reproduzierenden Routinetätigkeit geworden sein. Möchte man eine Aufgabe in einen der Anforderungsbereiche einordnen, muss man sich deshalb ein „typisches“ oder „durchschnittliches“ Kind vor Augen führen, das diese Aufgabe bearbeitet – bei aller Unbestimmtheit, die mit einer solchen Verallgemeinerung häufig verbunden ist. Trotz solcher Unschärfen sind die Anforderungsbereiche ein Werkzeug, das hilfreich sein kann, um die Variabilität von Mathematikaufgaben hinsichtlich ihres kognitiven Anspruchs zu beschreiben. Auch wenn sie eine nur grobe Orientierung geben, kann man sie nutzen, um einer „Verflachung“ von Unterricht und Tests aller Art entgegenzutreten und eine hinreichende Bandbreite von Anforderungen der eingesetzten Aufgaben sicherzustellen.

2.2.2 Struktur und Operationalisierung mathematischer Kompetenzen

Für den Ländervergleich 2011 stellt das im Teilkapitel 2.2.1 dargestellte Strukturmodell der Bildungsstandards den theoretisch-deskriptiven Ausgangspunkt dar. Es soll möglichst alle im Fach Mathematik relevanten Teilfähigkeiten integrieren. Daher ist das Modell hoch differenziert und umfasst sowohl Inhalts- als auch Prozessaspekte. Die verschiedenen allgemeinen und inhaltlichen Kompetenzen lassen sich dabei nicht vollständig trennscharf gegeneinander abgrenzen, sondern sie weisen an einigen Stellen Überlappungen und fließende Übergänge auf, wie es oben beispielsweise für Problemlösen und Argumentieren diskutiert wurde. Solche Überlappungen sind für die Verzahnung der verschiedenen Teilkompetenzen in der Unterrichtsarbeit sinnvoll und wünschenswert. Für eine psychometrische Messung, die eine trennscharfe, auf präzisen Definitionen basierende Operationalisierung der Kompetenzen erfordert, stellen sie jedoch eine Herausforderung dar. Während sich die inhaltlichen Kompetenzen noch relativ gut gegeneinander abgrenzen lassen, stellt sich dieses Problem in besonderer Weise für die allgemeinen mathematischen Kompetenzen, deren spezifische Messung gerade besonders wünschenswert wäre, weil sie im Mathematikunterricht eine Schlüsselrolle einnehmen sollen. Da diese prozessbezogenen Kompetenzen aber nicht nur große konzeptuelle Überlappungen aufweisen, sondern überdies meist in Kombinationen zum Einsatz kommen und dabei in komplexer Weise miteinander interagieren, ist eine separate Messung der einzelnen allgemeinen Kompetenzen praktisch nicht oder nur mit erheblichen Unschärfen möglich.

Für die Beschreibung der von den Schülerinnen und Schülern in den Ländern erzielten Kompetenzstände im Fach Mathematik ergibt sich somit die Frage, welche der in den Bildungsstandards unterschiedenen Aspekte auch empirisch voneinander differenziert werden können und sollten. Während es aus den eben genannten Gründen nicht realistisch erscheint, die allgemeinen mathematischen Kompetenzen separat zu erfassen, ist die getrennte Messung inhaltlicher Kompetenzen in großen Schulleistungsstudien üblich. Allerdings wird oft auch lediglich eine globale Mathematikfähigkeit berichtet (vgl. Blum et al., 2004, für die Sekundarstufe; Ufer, Reiss & Heinze, 2009, für den Primarbereich). Als Grundlage für die Entscheidung darüber, in welcher Weise der Ländervergleich 2011 mathematische Kompetenzen erfasst und ausweist, werden im Folgenden ausgewählte Ergebnisse empirischer Untersuchungen zur Dimensionalität mathematischer Kompetenzen dargestellt.

Befunde zur dimensionalen Struktur mathematischer Kompetenzen berichten beispielsweise Winkelmann und Kollegen (Winkelmann, Robitzsch, Stanat & Köller, 2012; Winkelmann & Robitzsch, 2009). Sie gehen vom Kompetenzmodell der Bildungsstandards aus und untersuchen, inwieweit sich mathematische Kompetenzen durch einen einzigen Faktor repräsentieren lassen (Globalwert) oder ob ein Modell mit mehreren Teilkompetenzen empirisch angemessener ist. Bei der Arbeit handelt es sich, neben Analysen des Mathematiktests DEMAT (z. B. Roick, Gölitz & Hasselhorn, 2004), um eine der wenigen Studien zur Struktur mathematischer Kompetenzen im Grundschulalter. Sie zeichnet sich überdies dadurch aus, dass verschiedene Methoden diskutiert werden, wie der Herausforderung begegnet werden kann, dass zur Bearbeitung von Mathematikaufgaben häufig mehr als eine Teilkompetenz erforderlich ist (Mehrfachladungsstruktur). Die Autoren zeigen, dass die gemessene Stärke der korrelativen Zusammenhänge zwischen verschiedenen mathematischen Teil-

kompetenzen variiert, je nachdem, in welcher Weise Mehrfachladungen modelliert werden. In Modellen, in denen die Mehrfachladungsstruktur berücksichtigt wird, ergeben sich zwischen den inhaltlichen Kompetenzbereichen Korrelationen in der Spanne von .44 bis .68. Diese Korrelationen deuten auf substantielle Gemeinsamkeiten der Teilkompetenzen hin. Sie sind jedoch niedriger als jene in anderen, ähnlich angelegten Studien, die Korrelationen zwischen inhaltlichen Teilfähigkeiten von etwa .80 bis .90 berichten (Blum et al., 2004; Brunner, 2006; Klieme, 2000).¹⁰ Wie Winkelmann und Robitzsch (2009) ebenfalls zeigen, ist diese Differenz im Wesentlichen auf die Modellierung der Mehrfachladungsstruktur zurückzuführen. In einem alternativen Modell, in dem unterstellt wird, dass zur Lösung einer Aufgabe nur jeweils eine einzelne Teilkompetenz erforderlich ist (Einfachladungsstruktur), steigen die Korrelationskoeffizienten erheblich und erreichen in etwa die Werte der anderen Studien, die durchweg mit der Annahme einer Einfachladungsstruktur arbeiten. Dieser Effekt ist erwartungskonform, da bekannt ist, dass Modelle, die eine tatsächlich bestehende Mehrfachladungsstruktur vernachlässigen, zu einer Überschätzung der Korrelationen zwischen den Teilfähigkeiten führen (Robitzsch, 2009; Zhang, 2012). Umgekehrt kann eine fehlspezifizierte Mehrfachladungsstruktur allerdings auch zu einer Unterschätzung dieser Korrelationen führen, so dass keiner der beiden Ansätze *per se* zu bevorzugen ist.

In Übereinstimmung mit den Schlussfolgerungen von Klieme (2000) sowie von Winkelmann und Kollegen (2012), gehen wir davon aus, dass die auf den Bildungsstandards basierenden mathematischen Kompetenzen im Primarbereich eng miteinander verwoben sind. Zur Beschreibung interindividueller Unterschiede ist es deshalb in vielen Fällen ausreichend, einen „Globalwert“ zu betrachten, der die Kompetenzausprägung im Fach Mathematik mit einem einzelnen Wert erfasst. Zugleich zeigen sich jedoch innerhalb von bestimmten Gruppen von Aufgaben besonders hohe Korrelationen. Dies kann als Beleg für die Abgrenzbarkeit von Teilkompetenzen gewertet werden. Für die Bildungsstandards Mathematik im Primarbereich lassen sich in diesem Sinne die inhaltlichen Kompetenzen analytisch trennen und beschreiben. Ob es sinnvoll ist, diese getrennt zu messen und zu berichten, muss im Einzelfall in Abhängigkeit von den Zielen der Untersuchung und den Möglichkeiten der eingesetzten Messinstrumente entschieden werden.

Erfasst man die inhaltlichen mathematischen Kompetenzen getrennt voneinander, scheint es nicht ratsam zu sein, zusätzlich eine mögliche Mehrfachladungsstruktur explizit zu modellieren. Zwar stellt die Annahme einer Einfachladungsstruktur eine vereinfachte Betrachtungsweise dar, jedoch sind die zur Verfügung stehenden Modelle mit Mehrfachladungsstrukturen nicht notwendig

10 Darüber hinaus liegen einige Arbeiten vor, in denen komplexer strukturierte Modelle diskutiert werden. So findet Brunner (2006) eine gute Passung für ein Modell, in dem zur Vorhersage des Lösens von Mathematikaufgaben gleichzeitig spezifisch mathematische Teilfähigkeiten und ein Faktor allgemeiner kognitiver Fähigkeit herangezogen werden. Ein strukturell ähnliches Modell erweist sich auch bei Jasper (2009) als gut mit den Daten verträglich. Er berücksichtigt allerdings keine allgemeinen kognitiven Fähigkeiten. Stattdessen wird ein Globalfaktor mathematischer Kompetenz spezifiziert und ergänzt durch drei bis vier unkorrelierte Faktoren, die Abhängigkeiten zwischen Aufgaben modellieren, welche nicht durch den Globalfaktor erfasst werden. Bei beiden Ansätzen ist die inhaltliche Interpretation der modellierten Teilfähigkeiten jedoch vergleichsweise schwierig, weshalb sie für solche Untersuchungen wie den Ländervergleich 2011 weniger geeignet sind und hier nicht weiter betrachtet werden.

plausibler.¹¹ Darüber hinaus sind entsprechende Modelle wesentlich komplexer, schwieriger zu interpretieren und sie beinhalten im Detail theoretische und methodische Herausforderungen, die noch nicht vollständig überzeugend gelöst sind (vgl. Kapitel 12.2.2). Folglich ist ein Modell mit Einfachladungsstruktur als einfacheres Modell gemäß dem Prinzip der Sparsamkeit zu bevorzugen.

Die oben präsentierten Überlegungen und Befunde zur Dimensionalität mathematischer Kompetenzen legen nahe, die Berichtlegung im Ländervergleich 2011 analog zum Vorgehen bei früheren Schulleistungsstudien im Bereich der Sekundarstufe zu gestalten (z. B. Blum et al., 2004). Es werden also Ergebnisse für einen Globalwert berichtet, der die gesamte Breite des Fachs Mathematik abdeckt. In den zentralen Analysen in Kapitel 5 und für die Befunde zu Geschlechterdisparitäten in Kapitel 7 werden die Ergebnisse zusätzlich auch für die fünf inhaltlichen Kompetenzbereiche differenziert ausgewiesen. Eine explizite Modellierung der Mehrfachladungsstruktur erfolgt dabei nicht (vgl. Kapitel 12.2.2).

Literatur

- Arbeitsgruppe „Internationale Vergleichsstudie“. (2007). *Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten*. Bonn: Bundesministerium für Bildung und Forschung.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillman & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillman, K.-J. & Weiß, M. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Behrens, U., Böhme, K. & Krelle, M. (2009). Zuhören – Operationalisierung und fachdidaktische Implikationen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 357–375). Weinheim: Beltz.
- Behrens, U. & Eriksson, B. (2009). Sprechen und Zuhören. In A. Bremerich-Vos, D. Granzer, U. Behrens & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Deutsch konkret* (S. 43–47). Berlin: Cornelsen.
- Belgrad, J., Eriksson, B., Pabst-Weinschenk, M. & Vogt, R. (2008). Die Evaluation von Mündlichkeit. Kompetenzen in den Bereichen Sprechen, Zuhören und Szenisch Spielen. In Symposium Deutschdidaktik e.V. (Hrsg.), Sonderheft zum 16. Symposium „Deutschdidaktik Kompetenzen im Deutschunterricht“. *Didaktik Deutsch, 14*, 20–45.
- Blum, W. (2006). Die Bildungsstandards Mathematik. Einführung. In W. Blum, C. Drüke-Noe, R. Hartung & O. Köller (Hrsg.), *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen* (S. 14–32). Berlin: Cornelsen Scriptor.
- Blum, W., Drüke-Noe, C., Hartung, R. & Köller, O. (Hrsg.). (2006). *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. Berlin: Cornelsen Scriptor.

¹¹ Beispielsweise nutzt der Ansatz von Winkelmann und Kollegen (2012) ein rein additives Modell. Das heißt, die Wahrscheinlichkeit, eine Aufgabe zu lösen, bestimmt sich aus der Summe der Werte aller beteiligten Kompetenzen. Folglich kann ein Defizit in einer Kompetenz vollständig durch andere beteiligte Kompetenzen ausgeglichen werden. Ein derartiges, voll kompensatorisches Modell wird dem komplexen Zusammenspiel verschiedener Kompetenzfacetten vermutlich nicht in allen Fällen gerecht.

- Blum, W., Neubrand, M., Ehmke, T., Senkbeil, M., Jordan, A., Ulfig, F. & Carstensen, C. H. (2004). Mathematische Kompetenz. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. (S. 47–92). Münster: Waxmann.
- Böhme, K. (2012). *Methodische und didaktische Überlegungen sowie empirische Befunde zur Erfassung sprachlicher Kompetenzen im Deutschen (Dissertation)*. Humboldt-Universität zu Berlin. Vorgesehen für eine elektronische Publikation unter <http://edoc.hu-berlin.de>. Berlin.
- Böhme, K. & Bremerich-Vos, A. (2009). Diagnostik der Rechtschreibkompetenz in der Grundschule – Konstruktprüfung mittels Fehler- und Dimensionsanalysen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 330–356). Weinheim: Beltz.
- Böhme, K., Neumann, D. & Schipolowski, S. (2010). Beschreibung der im Ländervergleich im Fach Deutsch untersuchten Kompetenzen. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 19–25). Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2007). *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M. & Schwippert, K. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Bremerich-Vos, A., Böhme, K. & Robitzsch, A. (2009). Sprachliche Kompetenzen im Fach Deutsch. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 198–218). Weinheim: Beltz.
- Brunner, M. (2006). *Mathematische Schülerleistung. Struktur, Schulformunterschiede und Validität*. Dissertation, Humboldt-Universität zu Berlin, Berlin. Zugriff am 11.07.2012 unter <http://edoc.hu-berlin.de/dissertationen/brunner-martin-2006-02-08/PDF/brunner.pdf>
- Buck, G. (2001). *Assessing listening*. New York: Cambridge University Press.
- Deppermann, A. (2008). *Gespräche analysieren*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61, 103–131.
- Freudenthal, H. (1973). *Mathematics as an educational task*. Dordrecht: D. Reidel Publishing Company.
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Dordrecht: D. Reidel Publishing Company.
- Gailberger, S. (2008). Leseförderung durch Hörbücher. Eine verbal-auditive Leseförderungstheorie für den Deutschunterricht. In B. Lecke (Hrsg.), *Mediengeschichte, Intermedialität und Literaturdidaktik* (S. 395–446). Frankfurt, Main: Lang.
- Granzer, D. (2009). Von Bildungsstandards zu ihrer Überprüfung: Grundlagen der Item- und Testentwicklung. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 21–30). Weinheim: Beltz.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York, NY: McGraw-Hill.
- Hasemann, K., Mirwald, E. & Hoffmann, A. (2007). Daten, Häufigkeit, Wahrscheinlichkeit. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Mathematik konkret* (S. 141–161). Berlin: Cornelsen Scriptor.
- Helmke, A. (Hrsg.). (2010). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer.

- Herné, K.-L. & Naumann, C. L. (2002). *Aachener Förderdiagnostische Rechtschreibfehler-Analyse (AfRA Version 4)*. Systematische Einführung in die Praxis der Fehleranalyse mit Auswertungshilfen zu insgesamt 31 standardisierten Testverfahren als Kopier-vorlagen mit Beiträgen von Cordula Löffler. Aachen: Alfa Zentaurus.
- Hurrelmann, B. (2007). Modelle und Merkmale der Lesekompetenz. In A. Bertschi-Kaufmann (Hrsg.), *Lesekompetenz – Leseleistung – Leseförderung* (S. 18–28). Seelze-Velber: Kallmeyer.
- Imhof, M. (2003). *Zuhören. Psychologische Aspekte auditiver Informationsverarbeitung*. Göttingen: Vandenhoeck & Ruprecht.
- Imhof, M. (2010). Zuhören lernen und lehren. Psychologische Grundlagen zur Beschreibung und Förderung von Zuhörkompetenzen in Schule und Unterricht. In M. Imhof & V. Bernius (Hrsg.), *Zuhörkompetenz in Unterricht und Schule. Beiträge aus Wissenschaft und Praxis* (S. 15–30). Göttingen: Vandenhoeck & Ruprecht.
- IQB (2008) = Institut zur Qualitätsentwicklung im Bildungswesen (Hrsg.). *Kompetenzstufenmodell zu den Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4)*. Zugriff am 12.07.2012 unter http://www.iqb.hu-berlin.de/bista/dateien/Mathe_primar.pdf
- Jasper, F. (2009). *Zur Psychometrie der Mathematik am Ende der Sekundarstufe I*. Dissertation, Universität Mannheim, Mannheim. Zugriff am 12.07.2012 unter <http://ub-madoc.bib.uni-mannheim.de/2584/>
- Jude, N. (2008). *Zur Struktur von Sprachkompetenz*. Dissertation, Goethe-Universität Frankfurt am Main. Zugriff am 29.07.2012 unter http://publikationen.ub.uni-frankfurt.de/files/6694/Jude_Zur_Struktur_von_Sprachkompetenz.pdf
- Klieme, E. (2000). Fachleistungen im voruniversitären Mathematik- und Physikunterricht. Theoretische Grundlagen, Kompetenzstufen und Unterrichtsschwerpunkte. In J. Baumert (Hrsg.), *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Bd. 2., Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe)* (S. 57–128). Opladen: Leske + Budrich.
- KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2005b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- Köster, J. (2008). Lesekompetenz im Licht von Bildungsstandards und Kompetenzmodellen. In A. Bremerich-Vos, D. Granzer & O. Köller (Hrsg.), *Lernstandsbestimmung im Fach Deutsch. Gute Aufgaben für den Unterricht* (S. 162–183). Weinheim: Beltz.
- Köster, J. & Rosebrock, C. (2009). Lesen – mit Texten und Medien umgehen. In A. Bremerich-Vos, D. Granzer, U. Behrens & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Deutsch konkret* (S. 104–138). Berlin: Cornelsen.
- Krelle, M. (2010). Zuhördidaktik. Anmerkungen zur Förderung rezeptiver Fähigkeiten des mündlichen Sprachgebrauchs im Deutschunterricht. In V. Bernius & M. Imhof (Hrsg.), *Zuhörkompetenz in Unterricht und Schule. Beiträge aus Wissenschaft und Praxis* (S. 51–68). Göttingen: Vandenhoeck & Ruprecht.
- Kürschner, C. & Schnotz, W. (2008). Das Verhältnis gesprochener und geschriebener Sprache bei der Konstruktion mentaler Repräsentationen. *Psychologische Rundschau*, 59, 139–149.
- Loeffler, I. & Meyer-Schepers, U. (1992). *DoRA. Dortmunder Rechtschreibfehler-Analyse zur Ermittlung des Schriftsprachstatus rechtschreibschwacher Schüler*. Dortmund: ILT.
- May, P. (2002). *HSP 1-9. Diagnose orthographischer Kompetenz. Zur Erfassung grundlegender Rechtschreibstrategien mit der Hamburger Schreibprobe. Manual*. Hamburg: Verlag für Pädagogische Medien.
- Mayer, R. E. & Hegarty, M. (1996). The Process of Understanding Mathematical Problems. In R. J. Sternberg & T. Ben-Zeev (Hrsg.), *The Nature of Mathematical Thinking* (S. 29–53). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, 37, 1–10.

- McGrew, K. S. & Hessler, G. L. (1995). The relationship between the WJ-R Gf-Gc cognitive clusters and mathematics achievement across the life-span. *Journal of Psycho-educational Assessment*, 13, 21–38.
- Mullis, I. V. S., Martin, M. O. & Kennedy, A. M. (2004). *Item-Writing guidelines for the PIRLS 2006 field test*. Paper presented at the 2nd PIRLS 2006 NRC Meeting, Bratislava, Slovak Republic.
- Naumann, C. L. (1999). *Orientierungswortschatz. Die wichtigsten Wörter und Regeln für die Rechtschreibung Klassen 1-6*. Weinheim: Beltz.
- NCTM (2000) = National Council of Teachers of Mathematics. (Hrsg.). *Principles and standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- OECD. (Hrsg.). (2003). *The PISA 2003 Assessment Framework Mathematics, Reading, Science and Problem Solving Knowledge and Skills*. Paris: OECD.
- Oller, J. W. (1976). Evidence for a general language proficiency factor: An expectancy grammar. *Die Neueren Sprachen*, 75, 165–174.
- Rasch, R. & Schütte, S. (2007). Zahlen und Operationen. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Mathematik konkret* (S. 66–88). Berlin: Cornelsen Scriptor.
- Robitzsch, A. (2009). Methodische Herausforderungen bei der Kalibrierung von Leistungstests. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 42–106). Weinheim: Beltz.
- Roick, T., Gölitz, D. & Hasselhorn, M. (2004). *Deutscher Mathematiktest für dritte Klassen: DEMAT 3+. Manual*. Göttingen: Beltz.
- Rost, J. (2004). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- Scheele, V. (2006). *Entwicklung fortgeschrittener Rechtschreibfertigkeiten. Ein Beitrag zum Erwerb der „orthographischen“ Strategien*. Frankfurt, Main: Lang.
- Thomé, G. (2003). Entwicklung der basalen Rechtschreibkenntnisse. In U. Bredel, H. Günther, P. Klotz, J. Ossner & G. Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache: Ein Handbuch* (S. 369–379). Paderborn: Ferdinand Schöningh.
- Ufer, S., Reiss, K. & Heinze, A. (2009). BIGMATH – Ergebnisse zur Entwicklung mathematischer Kompetenz in der Primarstufe. In A. Heinze & M. Grüßing (Hrsg.), *Mathematiklernen vom Kindergarten bis zum Studium. Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (S. 61–85). Münster: Waxmann.
- Valtin, R., Badel, I., Löffler, I., Meyer-Schepers, U. & Voss, A. (2003). Orthographische Kompetenzen von Schülerinnen und Schülern der vierten Klasse. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 227–264). Münster: Waxmann.
- Voss, A., Blatt, I. & Kowalski, K. (2007). Zur Erfassung orthographischer Kompetenz in IGLU 2006. *Didaktik Deutsch*, 23, 15–33.
- Walther, G. & Granzer, D. (2009). Kompetenzmodell Mathematik. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 108–119). Weinheim: Beltz.
- Walther, G., Selzer, C. & Neubrand, J. (2007). Die Bildungsstandards Mathematik. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Mathematik konkret* (S. 16–41). Berlin: Cornelsen Scriptor.
- Walther, G., van den Heuvel-Panhuizen, M., Granzer, D. & Köller, O. (Hrsg.). (2007). *Bildungsstandards für die Grundschule: Mathematik konkret*. Berlin: Cornelsen Scriptor.
- Winkelmann, H. & Robitzsch, A. (2009). Modelle mathematischer Kompetenzen: Empirische Befunde zur Dimensionalität. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 169–196). Weinheim: Beltz.
- Winkelmann, H., Robitzsch, A., Stanat, P. & Köller, O. (2012). Mathematische Kompetenzen in der Grundschule: Struktur, Validierung und Zusammenspiel mit allgemeinen kognitiven Fähigkeiten. *Diagnostica*, 58, 15–30.
- Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61, 37–46.

- Wittmann, E. C. & Müller, G. N. (2007). Muster und Strukturen als fachliches Grundkonzept. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Mathematik konkret* (S. 42–65). Berlin: Cornelsen Scriptor.
- Wollring, B. & Rinkens, H.-D. (2007). Raum und Form. In G. Walther, M. van den Heuvel-Panhuizen, D. Granzer & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Mathematik konkret* (S. 118–140). Berlin: Cornelsen Scriptor.
- Zhang, J. (2012). Calibration of response data using MIRT models with simple and mixed structures. *Applied Psychological Measurement*, 36, 375–398.

Kapitel 3

Kompetenzstufenmodelle für den Primarbereich

In diesem Kapitel wird in Abschnitt 3.1 zunächst kurz das Kompetenzverständnis der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen im Allgemeinen erläutert. Anschließend werden in Abschnitt 3.2 für das Fach Deutsch und in Abschnitt 3.3 für das Fach Mathematik im Primarbereich die fach- und domänenspezifischen Schritte der Modellentwicklung beschrieben und die resultierenden Kompetenzstufenmodelle für die einzelnen Kompetenzbereiche im Detail vorgestellt.

3.1 Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen

Hans Anand Pant, Katrin Böhme und Olaf Köller

3.1.1 Zum Kompetenzbegriff

Kompetenzen werden in Erziehungswissenschaft, Psychologie und Fachdidaktik als theoretische Konstrukte betrachtet, die erst mit Hilfe von Messinstrumenten der Beobachtung zugänglich gemacht werden können. Das genaue theoretische Verständnis von Kompetenz ist jedoch je nach fachwissenschaftlicher Ausrichtung durchaus unterschiedlich (vgl. für einen Überblick Böhme, 2012; Klieme & Hartig, 2007). Das den Bildungsstandards zugrunde liegende Kompetenzkonzept spiegelt zwar ein breit akzeptiertes und häufig verwendetes, aber dennoch nur *ein* mögliches Verständnis des Kompetenzbegriffs wider.

Der Kompetenzbegriff der Bildungsstandards bezieht sich bewusst auf ein relativ pragmatisches Verständnis von Kompetenzen. Sie werden als Fähigkeiten und Fertigkeiten betrachtet, die sich in konkreten Anforderungssituationen als Können manifestieren. In den Bildungsstandards der Kultusministerkonferenz (KMK) zeigt sich dies in der durchgängigen Formulierung der Kompetenzerwartungen als Könnensbeschreibungen (*Can-do-Statements*), wie zum Beispiel Schülerinnen und Schüler können „zentrale Aussagen eines Textes erfassen und wiedergeben“ (KMK, 2005, S. 12).

Klieme und Leutner definieren Kompetenzen als das Ergebnis von Bildungsprozessen und als „*kontextspezifische kognitive Leistungsdispositionen*, die sich funktional auf Situationen und Anforderungen in einer bestimmten *Domäne* beziehen“ (Klieme & Leutner, 2006, S. 879, Hervorhebung i. O.). Kennzeichnend für dieses Kompetenzkonzept ist

- die Abgrenzung gegenüber Begabungskonzepten zugunsten einer Betonung der Erlernbarkeit und Förderbarkeit von Kompetenz;

- die Abgrenzung zu allgemeinen Fähigkeitskonstrukten wie Intelligenz zugunsten einer engeren Definition des Expertisebereichs („*kompetent wofür?*“);
- ein funktionaler Handlungsbezug („*Can-do*“-Aussagen);
- ein Bezug zum „wirklichen Leben“ in dem Sinne, dass Fähigkeiten gemeint sind, die in wechselnden Kontexten und unter situativen Unwägbarkeiten angewendet werden können und
- eine Fokussierung auf den kognitiven Aspekt, um das „Können“ getrennt von motivationalen (z. B. Fachinteresse) und affektiven (z. B. Leistungsangst) Einflussgrößen auf das Leistungsgeschehen zu betrachten.

3.1.2 Kompetenzmodelle: Struktur-, Stufen- und Entwicklungsmodelle

Bei der theoretischen Modellierung von Kompetenzen sind die Aspekte der *Struktur* eines Kompetenzbereichs, der *Graduierung* einzelner Kompetenzen in *Niveaus (Stufen)*¹ und der *Entwicklung von Kompetenzen* im Zeitverlauf zu unterscheiden. „*Kompetenzstrukturmodelle* befassen sich mit der Frage, welche und wie viele verschiedene Kompetenzdimensionen in einem spezifischen Bereich differenzierbar sind. Bei der Beschreibung von *Kompetenzniveaus* geht es darum, welche konkreten situativen Anforderungen Personen bei welcher Ausprägung einer Kompetenz bewältigen können“ (Klieme & Leutner, 2006, S. 883, Hervorhebung i. O.). Mit der *Kompetenzentwicklung* ist die Frage verbunden, wie sich Kompetenzdimensionen über die Zeit verändern.

Welche der Aspekte – Struktur, Niveaus oder Entwicklung – bei der Kompetenzmodellierung im Vordergrund stehen, hängt wesentlich von den Zielen ab, die mit dem empirischen Einsatz modellbasierter Kompetenzerfassung verknüpft sind. Will man primär individuelle Verläufe des Kompetenzaufbaus über einen bestimmten Zeitraum (z. B. die Primarschulzeit) verfolgen, dann sind theoretische Entwicklungsmodelle unabdingbar. So sind etwa für die Beantwortung solcher Fragen wie zum Beispiel *Welche kognitiven Vorläuferkompetenzen müssen notwendig gegeben sein, um später komplexere Kompetenzen ausbilden zu können?* oder *Differenzieren sich frühe Kompetenzen im Entwicklungsverlauf der Primarschule in verschiedene Teilkompetenzen aus?* Entwicklungsmodelle erforderlich (vgl. Reiss, Heinze & Pekrun, 2007). Die Überprüfung des Erreichens der Bildungsstandards im Ländervergleich fokussiert allerdings *nicht* auf Entwicklungsverläufe, sondern möchte zu einem bestimmten Zeitpunkt des Bildungsverlaufs (vierte Jahrgangsstufe der Primarschulzeit) bilanzierende Aussagen über das Spektrum und die Verteilung von Kompetenzen im Bildungssystem treffen. Zu diesem Zweck werden theoretisch ausgearbeitete Kompetenzstrukturmodelle und Kompetenzstufenmodelle benötigt, die für jede als relevant angesehene Teilkompetenz Graduierungen der gemessenen Fähigkeiten in Kompetenzniveaus erlauben.

Die fachdidaktisch und lernpsychologisch begründeten Strukturmodelle für die Fächer Deutsch und Mathematik im Primarbereich wurden in den Kapiteln 2.1 sowie 2.2 vorgestellt. Im Folgenden soll kurz das allgemeine, domänenunspezifische Vorgehen bei der Definition der Kompetenzniveaus beschrieben werden, um anschließend die fachbezogenen Kompetenzstufenmodelle darzustellen.

¹ Die Begriffe „Kompetenzniveau“ und „Kompetenzstufe“ werden in diesem Bericht synonym verwendet.

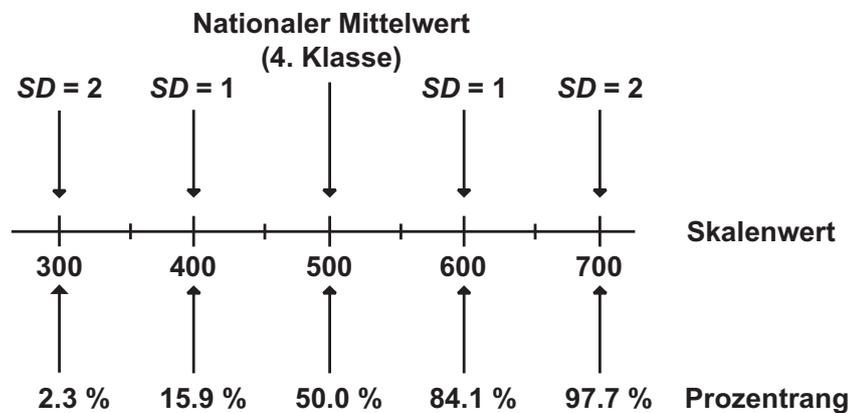
3.1.3 Standard-Setting: von Testwerten zu Kompetenzstufen²

Das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) hat in den Jahren 2006 und 2007 für die Fächer Deutsch und Mathematik auf der Basis national repräsentativer Stichproben der dritten und vierten Jahrgangsstufen Kalibrierungsbeziehungswise Normierungsstudien durchgeführt. Ziel dieser Studien war es, anhand empirisch gewonnener Daten länderübergreifend gültige Skalen zu definieren, auf denen sich die Schülerinnen und Schüler mit ihren Fähigkeiten und die jeweiligen (Teil-)Aufgaben mit ihren Schwierigkeiten verorten lassen. Damit liegen für die im Ländervergleich 2011 getesteten Kompetenzen Skalen vor, die es erlauben, das Erreichen der Bildungsstandards für den Primarbereich zu überprüfen.

Im Ländervergleich 2011 gilt für alle Teilkompetenzen, dass sie auf einer Skala abgebildet werden, die für alle Viertklässlerinnen und Viertklässler des allgemeinbildenden Schulsystems einen Mittelwert von $M = 500$ Kompetenzpunkten und eine Streuung von $SD = 100$ Kompetenzpunkten aufweisen. Diese Skalenbildung ist arbiträr, folgt damit aber den aus dem IQB-Ländervergleich sprachlicher Kompetenzen 2009 (Köller, Knigge & Tesch, 2010) und den internationalen Studien wie PIRLS/IGLU³ oder PISA⁴ bekannten Konventionen (vgl. bspw. Baumert et al., 2002; Bos et al., 2003).⁵

Unter der Annahme, dass die Daten normalverteilt sind, ergeben sich bei dieser Skalendefinition die in Abbildung 3.1 dargestellten Zusammenhänge zwischen der Standardabweichung (SD) und Prozenträngen.

Abbildung 3.1: Kontinuierliche Kompetenzskala ($M = 500$ und $SD = 100$) mit Prozenträngen unter der Annahme einer Normalverteilung



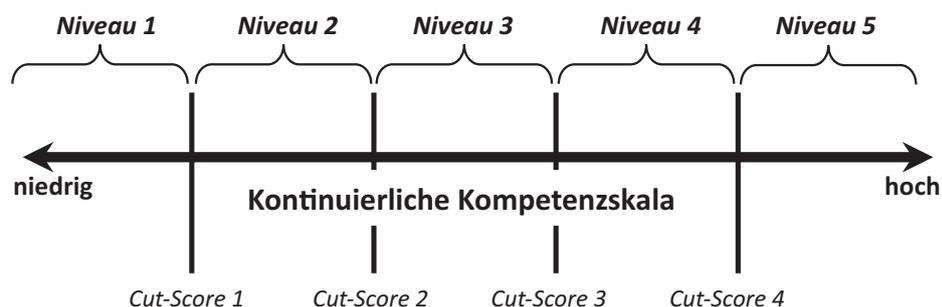
- 2 Teile dieses Abschnitts wurden in modifizierter Form aus Böhme und Köller (2010) übernommen.
- 3 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*, die deutsche Bezeichnung dieser Studie lautet *Internationale Grundschul-Lese-Untersuchung* (IGLU).
- 4 Das Akronym PISA steht für *Programme for International Student Assessment*.
- 5 Es sei an dieser Stelle ausdrücklich darauf hingewiesen, dass sich – ungeachtet der rein numerisch gleichen Skalenfestlegung – die Testergebnisse aus verschiedenen Untersuchungen nicht eins zu eins miteinander vergleichen lassen. Ein Mittelwert von 500 Kompetenzpunkten für Bundesland X im Lesetest des IQB-Ländervergleichs bedeutet also *nicht* das Gleiche wie ein Mittelwert von 500 Punkten für Land Y im IGLU-Lesetest (vgl. Kapitel 5.1.4). Um die Ergebnisse von unterschiedlichen Testverfahren miteinander vergleichen zu können, bedarf es zuvor einer inhaltlichen und statistischen Überprüfung der Äquivalenz (vgl. Pietsch, Böhme, Robitzsch & Stubbe, 2009).

Für den Mittelwert von $M = 500$ gilt somit theoretisch, dass jeweils 50 Prozent der Schülerinnen und Schüler darunter und darüber liegen. Für einen Schüler mit einem Fähigkeitswert von 600 Punkten (Mittelwert plus eine Standardabweichung) gilt, dass er mit seinem Wert über den Leistungen von rund 84 Prozent der Viertklässlerinnen und Viertklässler liegt, lediglich 16 Prozent erreichen einen noch höheren Wert. Für einen Kompetenzwert von 400 Punkten (Mittelwert minus eine Standardabweichung) gilt, dass ihn lediglich 16 Prozent der Schülerschaft unterschreiten.

Für die Interpretation von Skalenwerten ist bedeutsam, dass sich auch Teilaufgaben (*Items*) mit ihrer Schwierigkeit auf derselben Skala abbilden lassen. Sehr leichte Items liegen bei Schwierigkeitswerten von 400 und geringer, sehr schwierige Items bei Werten von 600 und höher. Hat ein Item eine Schwierigkeit von exakt 600 Punkten, so bedeutet dies, dass Personen mit einer Fähigkeit von 600 und mehr Punkten dieses Item mit einer hinreichenden Wahrscheinlichkeit lösen, Personen mit einem Fähigkeitswert unter 600 Punkten lösen es mit einer geringeren Wahrscheinlichkeit.⁶

Die Möglichkeit, Personen und Items auf einer gemeinsamen Skala verorten zu können, macht man sich bei der Definition von *Kompetenzstufen* zunutze. So kann ein Item, das beispielsweise einen Schwierigkeitswert von 600 aufweist, im Hinblick darauf analysiert werden, welche kognitiven Operationen zu seiner Lösung erforderlich sind. Dies wiederum lässt den Rückschluss zu, dass Personen, die einen Skalenwert von 600 oder höher erreicht haben, die zur Lösung dieses Items erforderlichen Operationen beherrschen. Erweitert man diese Idee, so kann man viele Items in ihrer Schwierigkeit betrachten und Punktwerte (Grenzen) auf der Skala definieren, bei denen sich die Items hinsichtlich ihrer kognitiven Anforderungen qualitativ verändern, also in inhaltlich beschreibbarer Weise komplexer werden. Die Definition solcher Kriterien erfolgt im Zuge einer Kompetenzstufensetzung, die im englischsprachigen Raum als *Standard-Setting* (Cizek, 2006) bezeichnet wird.

Abbildung 3.2: Schematische Darstellung eines Standard-Settings



6 Hinreichende Wahrscheinlichkeit bedeutet hier, dass die Lösungswahrscheinlichkeit bei $p = .625$ liegt (vgl. Köller et al., 2010). Diese Definition wurde in Anlehnung an die PISA-Studien vorgenommen (vgl. Frey, Carstensen, Walter, Rönnebeck & Gommelka, 2008).

Standard-Setting bezeichnet die Festlegung von Schwellenwerten (*Cut-Scores*), die eine Kompetenzskala in sukzessive Teilbereiche – sogenannte Kompetenzstufen oder Kompetenzniveaus – einteilt (Pant, Tiffin-Richards & Köller, 2010). Die Abbildung 3.2 zeigt schematisch, wie eine Cut-Score-Festlegung bei der Überprüfung des Erreichens von Bildungsstandards aussehen kann. Man legt auf dem Kontinuum der Kompetenzskala durch die Cut-Score-Setzungen fest, wo sich jeweils Übergänge zwischen Aufgabengruppen befinden, die abgrenzbare Anforderungen beinhalten und für deren sichere Lösung aufsteigende kognitive Leistungen notwendig sind. Die inhaltliche Beschreibung der Kompetenzniveaus erfolgt anhand der entsprechenden Könnensbeschreibungen.

Standard-Setting-Verfahren stützen sich auf inhaltliche Einschätzungen und Urteile von Expertinnen und Experten unterschiedlicher Expertisebereiche (hier: Fachdidaktik, Testentwicklung, Schulpraxis, Bildungsadministration). Das Setzen von Cut-Scores durch ein Expertenurteil stellt *per se* einen bewertenden Vorgang dar. In der psychometrischen Fachliteratur wird deshalb nicht davon ausgegangen, dass durch Standard-Setting-Prozeduren ein „wahrer“ Cut-Score gefunden werden kann (Kane, 2001).

Für die Festlegung der Kompetenzstufen stehen zahlreiche Verfahren zur Verfügung, die unterschiedliche Vor- und Nachteile aufweisen und teilweise unterschiedliche Zielstellungen verfolgen. Am prominentesten sind die *Bookmark*- sowie die *Angoff*-Methode (vgl. für einen Überblick Cizek, 2001; Cizek & Bunch, 2007; Kaftandjieva, 2010; Pant, Rupp, Tiffin-Richards & Köller, 2009). Für die im Kontext der Bildungsstandards gegebenen Bedingungen, also auf Standards basierende Kompetenztests mit ausschließlich *richtig-falsch* kodierten Aufgaben und der Notwendigkeit, mehrere Cut-Scores pro Kompetenzskala zu setzen, empfiehlt eine Synopse des *Educational Testing Service* (ETS; Morgan & Perie, 2004) die *Bookmark*-Methode (Mitzel, Lewis, Patz & Green, 2001). Diese wurde in den US-amerikanischen Large-Scale-Assessments der letzten Jahre mit Abstand am häufigsten angewendet. Bei der Festlegung von Stufen für die im Ländervergleich 2011 berücksichtigten Kompetenzen wurde daher die *Bookmark*-Methode in modifizierter Form verwendet.

Bei dieser Methode wird dem Expertenpanel ein „Buch“ vorgegeben, das alle Items nach ihrer empirisch ermittelten Schwierigkeit geordnet enthält. Die Aufgabe der Panelteilnehmer ist es, im wiederholten Abgleich mit den Kompetenzstufendeskriptoren an denjenigen Stellen im Item-Buch Markierungen (*bookmarks*) zu setzen, an denen ein für die jeweilige Stufe gerade kompetenter Schüler (*minimally competent learner*) die Teilaufgabe sicher, das heißt mit der hier festgesetzten Antwortwahrscheinlichkeit von knapp zwei Drittel, lösen würde.

Neben solchen Verfahren des Standard-Settings, die wesentlich auf Expertenurteilen beruhen, werden in neuerer Zeit auch stärker psychometrisch ausgerichtete Ansätze erprobt (vgl. Hartig & Frey, 2012; Hartig, Frey, Nold & Klieme, 2011). Da diese Modelle bei der Abgrenzung von Kompetenzstufen in größerem Maße berücksichtigen, was genau eine Aufgabe zu einer leichten beziehungsweise schwierigen macht, können sie für ein individualdiagnostisches Feedback eher geeignet sein. Für Zwecke des Large-Scale-Assessments wie den Ländervergleich erscheinen sie jedoch weniger geeignet, da Anzahl, Breite und inhaltliche Bedeutung der ermittelten Kompetenzstufen sehr sensitiv für den jeweils eingesetzten Aufgabenpool sein dürften. Die resultierenden Kompetenzstufenmodelle wären in diesem Falle also instabil, was für ein langfristig angelegtes Systemmonitoring, das Trends erfassen soll, problematisch wäre.

Weiterhin ist zu betonen, dass sich die Festlegung von Kompetenzstufen letztlich nie ausschließlich auf empirische Ergebnisse und fachdidaktische Erwägungen stützen kann. Stets spielen auch politische, normative und andere inhaltlich-sachliche Erwägungen eine Rolle. Camilli, Cizek und Lugg (2001, S. 455) formulieren pointiert: „Standard setting is less a scientific enterprise than a systematic and practical activity strongly influenced by social and policy perspectives“. Die Stufen werden also nicht als etwas mit den Daten bereits Gegebenes „gefunden“, sondern als etwas auf Grundlage der Daten zu Konstruierendes und mit politischen Vertretern Ausgehandeltes verstanden (vgl. bspw. AERA, APA, NCME, 1999; Cizek, 2001). Unabhängig von der Art des Vorgehens ist es wichtig, treffende Bezeichnungen und Beschreibungen für die Kompetenzstufen zu wählen, so dass nachvollziehbar wird, über welche konkreten Kompetenzen Schülerinnen und Schüler auf den jeweiligen Stufen verfügen.

3.1.4 Bildungsstandards als Verknüpfung von Kompetenzstufen mit normativen Erwartungen

Bei Bildungsstandards handelt es sich um normativ gesetzte Kompetenzerwartungen. Abweichend von den Empfehlungen der sogenannten Klieme-Expertise zur Entwicklung nationaler Bildungsstandards (Klieme et al., 2007) hat die Kultusministerkonferenz in ihren Beschlüssen keine *Mindeststandards* definiert, sondern *Regelstandards* formuliert. Die Bildungsstandards legen also Kompetenzerwartungen fest, die Schülerinnen und Schüler in der jeweiligen Bildungsetappe im Durchschnitt erfüllen sollten (Klieme et al., 2007, S. 138).

Bei der Erarbeitung der Kompetenzstufenmodelle für den Primarbereich war unter anderem festzulegen, bei welchem Niveau die Vorgaben des Regelstandards erfüllt werden. Darüber hinaus sollte ein Kompetenzstufenmodell im Sinne der von Klieme und Kollegen (2007) eingeforderten Differenzierung verschiedene Niveaus definieren, um auch Mindest- und Optimalstandards zu bestimmen. Da Regelstandards ein mittleres Erwartungsniveau beschreiben, wurden für das Kompetenzstufenmodell des Ländervergleichs fünf Kompetenzniveaus mit dem mittleren Niveau (Kompetenzstufe III) als Regelstandard abgegrenzt. Im Sinne dieser Differenzierung wurden die folgenden inhaltlichen Definitionen vorgenommen (vgl. Köller, 2010)⁷:

- *Mindeststandards* beziehen sich auf ein definiertes Minimum an Kompetenzen, das alle Schülerinnen und Schüler bis zu einem bestimmten Bildungsabschnitt erreicht haben sollten. Diese unterschreiten die in den Publikationen der KMK festgelegten Kompetenzerwartungen der Regelstandards. Sie beschreiben jedoch ein Bildungsminimum am Ende der Primarstufe, von dem angenommen werden kann, dass sich Schülerinnen und Schüler, die diesem Niveau zuzuordnen sind, bei entsprechender Unterstützung erfolgreich in die Sekundarstufe I integrieren werden.
- *Regelstandards* beziehen sich auf Kompetenzen, die im Durchschnitt von den Schülerinnen und Schülern bis zu einem bestimmten Bildungsabschnitt erreicht werden sollen und im Einklang mit den entsprechenden Veröffentlichungen der KMK stehen.

⁷ Die Kompetenzstufe I kennzeichnet Leistungen, die die Anforderungen der Mindeststandards noch unterschreiten. Für diese Stufe wurde keine eigene Bezeichnung eingeführt.

- Als *Regelstandard plus* wird ein Leistungsbereich definiert, der über den Regelstandards liegt und als Zielperspektive für die Weiterentwicklung von Unterricht angesehen werden kann.
- *Optimal-* beziehungsweise *Maximalstandards* beziehen sich auf Leistungserwartungen, die bei sehr guten oder ausgezeichneten individuellen Lernvoraussetzungen und der Bereitstellung besonders günstiger Lerngelegenheiten innerhalb und außerhalb der Schule erreicht werden können und bei weitem die Erwartungen der KMK-Bildungsstandards übertreffen.

Neben den genannten Kriterien müssen Festlegungen von Kompetenzstufen sowohl bildungspolitischen Erwartungen als auch pädagogischen Erfordernissen entsprechen. Kompetenzstufenmodelle sollten in dieser Hinsicht (vgl. Köller, 2010)

- herausfordernde und zugleich angemessene Leistungserwartungen konkretisieren, die der Leistungsstreuung innerhalb und zwischen den Ländern in angemessener Weise Rechnung tragen;
- trotz der zu erwartenden unterschiedlich hohen Anteile von Schülerinnen und Schülern, die den Regelanforderungen nicht entsprechen, für alle Länder ein Leistungsminimum beschreiben, das von allen Schülerinnen und Schülern mittelfristig erreicht wird;
- motivierende Leistungserwartungen formulieren, die Entwicklungsimpulse an den Schulen auslösen; hierfür dient insbesondere der *Regelstandard plus*;
- anschlussfähig an internationale Vorarbeiten sein, wie sie in den großen Schulleistungsstudien realisiert wurden;
- fünf annähernd gleich breite Kompetenzstufen zur einheitlichen, über mehrere Erhebungswellen hinweg stabilen kriterialen Kompetenzbeschreibung definieren sowie
- fachdidaktisch gut interpretierbare und vertretbare Grenzen zwischen den Kompetenzstufen aufweisen.

Nach dieser allgemeinen Darstellung, wie Kompetenzstufenmodelle entwickelt werden, folgt in den nächsten beiden Abschnitten eine genaue Beschreibung der fachspezifischen Kompetenzstufenmodelle, die vom IQB in enger Zusammenarbeit mit den fachdidaktischen Kooperationspartnern konstruiert worden sind und die die Grundlage für den Ländervergleich 2011 bilden.

3.2 Kompetenzstufenmodelle im Fach Deutsch

Albert Bremerich-Vos, Katrin Böhme, Michael Krelle,
Sebastian Weirich und Olaf Köller

Auf den Bildungsstandards der Kultusministerkonferenz (KMK) basierende Kompetenzstufenmodelle sollen herausfordernde und zugleich angemessene Kompetenzerwartungen beschreiben, die der Leistungsstreuung in den Ländern in geeigneter Weise Rechnung tragen. Zusätzlich zum „Regelniveau“ sollen sie ein „Mindestniveau“ definieren, das für alle Schülerinnen und Schüler anzustreben ist und letztlich auch erreicht werden soll.

Gute Kompetenzstufenmodelle müssen sowohl bildungspolitischen Erwartungen entsprechen als auch fachlichen, fachdidaktischen sowie testtheoretischen Anforderungen genügen. Es ist erforderlich, dass sich die Kompetenzstufenmodelle eng an die von der Kultusministerkonferenz verabschiedeten Bildungsstandards anlehnen und dabei nach Möglichkeit das gesamte Kompetenzspektrum berücksichtigen. Gleichzeitig ist eine Anbindung an nationale und internationale Vorarbeiten erstrebenswert.

In den Bildungsstandards im Fach Deutsch für den Primarbereich werden die Kompetenzbereiche *Sprechen und Zuhören*, *Schreiben*, *Lesen* sowie *Sprache und Sprachgebrauch untersuchen* unterschieden (KMK, 2005a). Orthografie ist dort nicht gesondert ausgewiesen, sondern wird unter Schreiben subsumiert. Im Rahmen des IQB-Ländervergleichs 2011 wurden nur Aufgaben zu den Domänen Zuhören, Lesen und Orthografie eingesetzt. Im Folgenden werden die Kompetenzstufenmodelle für diese Bereiche eingeführt und dargestellt.⁸

3.2.1 Das Kompetenzstufenmodell für den Teilbereich Zuhören

Das Kompetenzstufenmodell für den Teilbereich *Zuhören* wurde 2009 im Rahmen eines Standard-Settings unter Einbindung von Lehrkräften, Fachdidaktikerinnen und Fachdidaktikern sowie Expertinnen und Experten aus dem Bereich Psychometrie hauptsächlich auf Grundlage der Normierungsstudie für den Primarbereich (vgl. Kapitel 1) entwickelt und im Frühjahr 2010 von der Kultusministerkonferenz zustimmend zur Kenntnis genommen.

Als Referenzstichprobe für die in den nachfolgenden Kapiteln berichteten Befunde dient die vierte Jahrgangsstufe des 2011 durchgeführten Ländervergleichs. Dies bedeutet, dass die Kompetenzverteilung der im Rahmen der vorliegenden Studie getesteten Viertklässlerinnen und Viertklässler einen Mittelwert von genau 500 Punkten und eine Standardabweichung von genau 100 Punkten aufweist.

⁸ Die Beschreibungen der Kompetenzstufen basieren auf früheren, auf den Internetseiten des IQB publizierten Fassungen, die redaktionell überarbeitet worden sind. Da im vorliegenden Kapitel in den Abbildungen 3.4, 3.6 und 3.7 für jeden Bereich Beispielaufgaben präsentiert werden, die die Kompetenzstufen illustrieren, wurden aus den Stufenbeschreibungen Bezüge zu nicht publizierten Aufgaben entfernt. Allgemeinere Aussagen über die mit einer Stufe verbundenen Anforderungen haben wir jeweils an den Anfang der einzelnen Beschreibungen gestellt. Der Kern der Anforderungen, die Schülerinnen und Schüler auf den jeweiligen Kompetenzniveaus bewältigen können, wurde manchmal etwas deutlicher herausgearbeitet. Inhaltlich hat sich an der Definition der einzelnen Kompetenzstufen jedoch nichts geändert.

Das Modell umfasst fünf Kompetenzstufen mit einer Stufenbreite von jeweils 85 Punkten, wobei die Stufe I nach unten und die Stufe V nach oben offen ist. Konkret wurden folgende Kompetenzstufen bestimmt:

- Kompetenzstufe I: unter 360 Punkten
- Kompetenzstufe II: Werte von 360 bis 444 Punkten (Mindeststandard)
- Kompetenzstufe III: Werte von 445 bis 529 Punkten (Regelstandard)
- Kompetenzstufe IV: Werte von 530 bis 614 Punkten (Regelstandard plus)
- Kompetenzstufe V: Werte von 615 Punkten und mehr (Optimal- bzw. Maximalstandard)

Kompetenzstufe I (Punktwerte unter 360): prominente Einzelinformationen wiedererkennen

Auf Kompetenzstufe I können Schülerinnen und Schüler in der Hauptsache prominente Einzelinformationen wiedererkennen. Dabei sind die Anforderungen insofern erleichtert, als es um kurze Hörtexte geht und in der Regel nur eine Sprecherin oder ein Sprecher mit deutlicher Aussprache zu hören ist. Auch kommen keine Stör- oder Hintergrundgeräusche vor. Die Anforderungen werden vor allem dann bewältigt, wenn die Aufgaben ein *Multiple-Choice-Format* haben und die Distraktoren leicht auszuschließen sind. Im Rahmen solcher Bedingungen können Schülerinnen und Schüler beispielsweise Schauplätze oder Hauptfiguren wiedererkennen.

Die Anforderungen, die von Schülerinnen und Schülern auf Kompetenzstufe I bewältigt werden, liegen deutlich unter den Erwartungen der KMK-Bildungsstandards. Es ist davon auszugehen, dass bei Schülerinnen und Schülern auf dieser Kompetenzstufe der erfolgreiche Übergang in die Sekundarstufe I nur unter Einsatz intensiver Fördermaßnahmen gelingen wird.

Kompetenzstufe II (Punktwerte von 360 bis 444): benachbarte Informationen verknüpfen und weniger prominente Einzelinformationen wiedergeben

Schülerinnen und Schüler auf Kompetenzstufe II können weniger prominente Einzelinformationen wiedergeben und textnahe Schlüsse ziehen. Dabei sind die Zuhöranforderungen insofern erleichtert, als sie sich – wie auf Kompetenzstufe I – auf einfache Hörtexte beziehen. So ist häufig nur eine Sprecherin oder ein Sprecher zu hören. Wenn mehrere Sprecherinnen und Sprecher beteiligt sind, wird das Zuhören durch verständnisfördernde Pausen erleichtert; die Sprecherbeiträge überlappen sich nicht und die Aussprache ist deutlich. Stör- und Hintergrundgeräusche kommen selten vor. Lenkende Hinweise, wie sie geschlossene Formate beinhalten, werden nicht mehr durchgängig benötigt, um die Zuhöranforderungen zu bewältigen. Im Rahmen solcher Bedingungen können Schülerinnen und Schüler Einzelinformationen reproduzieren, auch wenn sie nicht explizit im Hörtext genannt wurden. Zudem können sie Informationen verknüpfen und einfache textnahe Schlüsse ziehen, wenn die Informationen im Hörtext kurz nacheinander gegeben wurden. Solche einfachen interpretatorischen Leistungen betreffen etwa kausale Beziehungen zwischen zentralen Handlungsaspekten oder Zusammenhänge zwischen Hauptfiguren und Schauplätzen, ohne dass damit ein Verstehen auf der Ebene des Textes insgesamt verbunden ist. Auch sind Schülerinnen und Schüler bereits auf dieser Kompetenzstufe in der Lage, Entscheidungen auf Basis einfacher Schlüsse zu begründen, sofern es nur um das Verstehen einzelner Textstellen geht.

Schülerinnen und Schüler auf Kompetenzstufe II sind damit in der Lage, Anforderungen im Bereich Zuhören zu bewältigen, die sich im Sinne von *Mindeststandards* interpretieren lassen.

Kompetenzstufe III (Punktwerte von 445 bis 529): verstreute Informationen miteinander verknüpfen und den Text ansatzweise als Ganzen erfassen

Auf dieser Kompetenzstufe können Schülerinnen und Schüler vor allem solche Anforderungen bewältigen, die sich auf die Verbindung voneinander entfernt platzierter Informationen beziehen. Zudem können sie den Text ansatzweise als Ganzen erfassen. Dabei sind die Anforderungen insofern anspruchsvoller als auf den vorherigen Stufen, als die Texte länger sind, Stör- und Hintergrundgeräusche das Verstehen erschweren und mehrere Sprecherinnen und Sprecher zu Wort kommen, deren Stimmen nicht immer klar zu unterscheiden sind. Lenkende Hinweise, wie sie in Multiple-Choice-Aufgaben enthalten sind, werden dabei nicht mehr durchgängig benötigt, um die Anforderungen zu bewältigen. Die Schülerinnen und Schüler können jetzt auch Fragen mit offenen Formaten beantworten. Im Rahmen solcher Bedingungen sind unter anderem Deutungen möglich, die den Text als Ganzen betreffen, etwa im Hinblick auf Motive oder Gefühle von Figuren, den Handlungsverlauf oder die Funktion des Textes. Schülerinnen und Schüler können auf dieser Stufe auch längere Erklärungen von Begriffen geben, die im Hörtext eine Rolle spielen, sowie Beschreibungen von Figuren, Orten oder Handlungskernen reproduzieren, die sie im Text gehört haben.

Insgesamt entsprechen die hier gezeigten Leistungen den Erwartungen der Bildungsstandards der KMK. Auf dieser Stufe können also die *Regelstandards* als erreicht gelten.

Kompetenzstufe IV (Punktwerte von 530 bis 614): auf der Ebene des Textes wesentliche Zusammenhänge erkennen und Details im Kontext verstehen

Schülerinnen und Schüler können auf dieser Kompetenzstufe zusätzlich Anforderungen bewältigen, die ein tieferes Verstehen des Gesamttextes oder von größeren Textabschnitten erfordern, und zwar selbst dann, wenn die Texte anspruchsvoll und lang sind. Dies gelingt auch bei Aufgaben mit offenen Formaten. Dabei kann es um die Integration von Informationen zu äußeren Merkmalen von Figuren und Objekten oder die Ausgestaltung von Schauplätzen und Handlungskernen gehen. Die Schülerinnen und Schüler bewältigen solche Anforderungen insbesondere auch dann, wenn die Informationen nicht an prominenter Stelle im Text vorkommen oder sprachlich schwieriger sind als auf den Stufen I bis III.

Insgesamt liegen die Leistungen auf dieser Kompetenzstufe bereits über den Erwartungen der KMK-Bildungsstandards und können so im Sinne des Erreichens des *Regelstandards plus* interpretiert werden.

Kompetenzstufe V (Punktwerte ab 615): auf zentrale Aspekte des Textes bezogene Aussagen eigenständig beurteilen und begründen

Auf der höchsten Kompetenzstufe sind Schülerinnen und Schüler in der Lage, Aussagen über zentrale Aspekte von Hörtexten eigenständig zu beurteilen und zu begründen. Dies gelingt auch dann, wenn es sich um Hörtexte handelt, die aus mehreren, teils längeren Sequenzen bestehen und eine große Anzahl von Sprecherinnen und Sprechern mit ähnlicher Sprechweise zu Wort kommt. Die Schülerinnen und Schüler bewältigen die Anforderungen dabei meist auch im Rahmen von komplexen offenen Aufgaben.

Im Rahmen solcher Bedingungen können Schülerinnen und Schüler Aussagen im Hinblick auf ihre Passung mit dem Hörtext prüfen, diese Aussagen erläutern oder begründet Stellung dazu nehmen. Dabei sind sie auch in der Lage, mehrere Aussagen über den Hörtext miteinander zu vergleichen und vorgegebene Begründungen zu untersuchen. Schülerinnen und Schüler können auf diesem Niveau zudem zentrale Unterschiede und Gemeinsamkeiten zwischen literarischen Hörtexten benennen.

Insgesamt bewältigen Schülerinnen und Schüler auf Kompetenzstufe V Anforderungen, die deutlich über den Vorgaben der KMK-Bildungsstandards liegen und nur bei optimalen schulischen und außerschulischen Lernbedingungen sowie günstigen individuellen Lernvoraussetzungen gemeistert werden. Diese Stufe kann daher als Erreichen von *Optimalstandards* gelten.

Zur Illustration der Kompetenzstufen I bis V finden sich nachfolgend in der Abbildung 3.3 die Transkripte der beiden Hörtexte der Zuhöraufgabe „Zwei Märchen“ und in der Abbildung 3.4 einige Beispielitems dieser Aufgabe.

Abbildung 3.3: Transkripte der Hörtexte zur Aufgabe „Zwei Märchen“

Der süße Brei

Du hörst gleich zwei Märchen. Blättere noch nicht um, sondern warte ab, bis wir das erste Märchen zu Ende gehört haben. Hör jetzt genau zu und pass gut auf.

Es war einmal ein armes, frommes Mädchen, das lebte mit seiner Mutter allein und sie hatten nichts mehr zu essen. Da ging das Kind hinaus in den Wald und dort begegnete ihm eine alte Frau, die wusste seinen Jammer schon und schenkte ihm ein Töpfchen. Zu dem sollte es sagen „Töpfchen, koche“, so kochte es guten, süßen Hirsebrei, und wenn es sagte „Töpfchen, steh“, so hörte es wieder auf zu kochen.

Das Mädchen brachte den Topf seiner Mutter heim, und nun waren sie ihrer Armut und ihres Hungers ledig und aßen süßen Brei, sooft sie wollten. Auf eine Zeit war das Mädchen ausgegangen, da sprach die Mutter „Töpfchen, koche“, da kocht es und sie isst sich satt. Nun will sie, dass das Töpfchen wieder aufhören soll, aber sie weiß das Wort nicht. Also kocht es fort, und der Brei steigt über den Rand hinaus und kocht immerzu, die Küche und das ganze Haus voll, und das zweite Haus und dann die Straße, als wollt's die ganze Welt satt machen, und ist die größte Not, und kein Mensch weiß sich da zu helfen. Endlich, wie nur noch ein einziges Haus übrig ist, da kommt das Kind heim und spricht nur „Töpfchen, steh.“ Da steht es und hört auf zu kochen. Und wer wieder in die Stadt wollte, der musste sich durchessen.

Blättere jetzt bitte um und beantworte die Fragen.

Das Märchen vom Salzwasser

Jetzt hören wir das zweite Märchen. Blättere noch nicht um, sondern warte ab, bis wir das zweite Märchen zu Ende gehört haben. Hör wieder genau zu und pass gut auf.

Es war einmal vor langer, langer Zeit, da lebten in einem kleinen Fischerdorf zwei Brüder. Der Ältere von ihnen hieß Chen. Er war böse, habgierig und verschlagen und nur auf seinen eigenen Vorteil bedacht. Lin aber, der Jüngere, war ein mutiger und redlicher Fischer, der mit dem zufrieden war, was er mit eigenen Händen erarbeitete. Jeden Morgen fuhr der fleißige Lin mit seinem kleinen Boot hinaus aufs Meer, um zu fischen. Aber seine Fischernetze waren sehr alt und er hatte kein Geld, um sich neue zu kaufen. Sobald sich Fische darin verfangen hatten, zerrissen seine Netze und der gesamte Fang entwichte ihm. Immer öfter kam Lin ohne einen einzigen Fisch nach Hause zu seiner Familie.

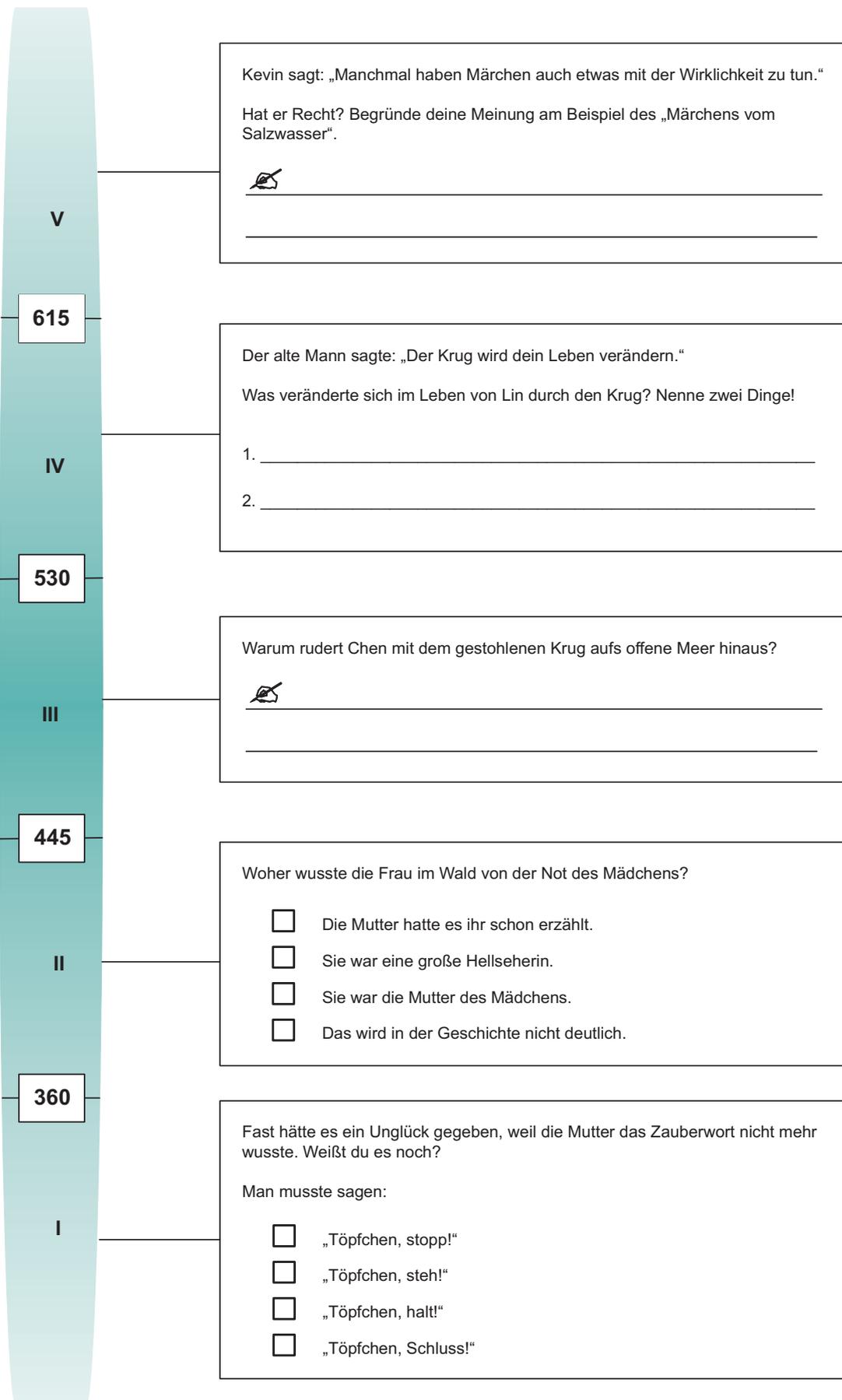
So ging es eine ganze Weile. Eines Abends war er aber so verzweifelt, dass er einfach nicht einschlafen konnte. Da stand mitten in der Nacht ein alter Mann vor ihm und sagte: „Lin, nun hast du Geduld und Mut bewiesen. Zur Belohnung schenke ich dir einen Krug. Er besitzt Zauberkräfte und wird dein Leben verändern. Aber gib Acht auf meine Worte! Du musst sagen: ‚Krug, gib Salz‘. Und er wird sich mit diesem wertvollen Gut füllen. Wenn du genug hast, sagst du zu ihm: ‚Halt ein, hab Dank‘. Und im selben Moment wird er sich wieder leeren.“ Lin bedankte sich und nahm das kostbare Geschenk an sich. Er tat, wie ihm gesagt. Jeden Tag besprach er von nun an den Zauberkrug. Dieser füllte sich mit Salz, sooft er es wollte.

Lin verkaufte es und wurde reich davon. Sein älterer Bruder Chen aber gönnte ihm sein Glück nicht. Er war schrecklich neidisch. Eines Morgens folgte er Lin und beobachtete ihn. Chen belauschte ihn, wie er mit folgenden Worten sprach: „Krug, gib Salz.“ Dann sah er, wie der Krug Salz spendete. Vor lauter Ungeduld hörte er aber nicht den zweiten Teil der Worte. Als Lin zum Markt ging, schlich er sich in dessen Haus und stahl den Zauberkrug. Sofort wollte er ihn ausprobieren. Er setzte sich in sein Boot und fuhr aufs offene Meer, damit ihn dabei niemand beobachtete. Kaum war er dort angekommen, sprach er schon: „Krug, gib mir Salz.“

Und wie schon bei seinem Bruder füllte sich der Krug mit Salz. Aber das Salz hörte nicht auf zu fließen. Es ergoss sich über seine Füße, über das gesamte Boot und wollte nicht mehr aufhören. Als Chen das bemerkte, schrie er in höchster Not: „Zu Hilfe, ich ertrinke.“ Er fluchte und bettelte, aber es half ihm nichts. Er kannte die Zauberformel nicht, die dem Salz Einhalt geboten hätte. Und so drückte das Gewicht des Salzes schließlich das gesamte Boot unter Wasser. Mit dem kleinen Fischerboot versanken auch Chen und der Krug in den Fluten. Seit diesem Tage liegt der Krug auf dem Meeresgrund. Und da niemand mehr zu ihm die Formel sagen kann, dass es aufhört, kommt das Salz bis zum heutigen Tage aus ihm heraus. Und deshalb ist das Meerwasser salzig.

Blättere jetzt bitte um und beantworte die Fragen.

Abbildung 3.4: Illustration des Kompetenzstufenmodells für den Teilbereich Zuhören



3.2.2 Das Kompetenzstufenmodell für den Bereich Lesen

Das Kompetenzstufenmodell für den Bereich *Lesen* wurde 2009 im Rahmen eines Standard-Settings unter Einbindung von Lehrkräften, Fachdidaktikerinnen und Fachdidaktikern sowie Expertinnen und Experten aus dem Bereich Psychometrie auf Grundlage der Pilotierungs- und Normierungsstudien für den Primarbereich (vgl. Kapitel 1) entwickelt und im Frühjahr 2010 von der Kultusministerkonferenz zustimmend zur Kenntnis genommen. Eine frühere Version des Kompetenzstufenmodells für den Bereich Lesen findet sich bei Bremerich-Vos und Böhme (2009).

Als Referenzstichprobe für die in den nachfolgenden Kapiteln berichteten Befunde dient die vierte Jahrgangsstufe des 2011 durchgeführten Ländervergleichs. Dies bedeutet, dass die Kompetenzverteilung der im Rahmen der vorliegenden Studie getesteten Viertklässlerinnen und Viertklässler einen Mittelwert von genau 500 Punkten und eine Standardabweichung von genau 100 Punkten aufweist.

Das Modell umfasst fünf Kompetenzstufen mit einer Stufenbreite von jeweils 75 Punkten, wobei die Stufe I nach unten und die Stufe V nach oben offen ist. Konkret wurden folgende Kompetenzstufen bestimmt:

- Kompetenzstufe I: unter 390 Punkten
- Kompetenzstufe II: Werte von 390 bis 464 Punkten (Mindeststandard)
- Kompetenzstufe III: Werte von 465 bis 539 Punkten (Regelstandard)
- Kompetenzstufe IV: Werte von 540 bis 614 Punkten (Regelstandard plus)
- Kompetenzstufe V: Werte von 615 Punkten und mehr (Optimal- bzw. Maximalstandard)

Kompetenzstufe I (Punktwerte unter 390): explizit angegebene Einzelinformationen identifizieren

Kinder auf der nach unten offenen Kompetenzstufe I sind in der Lage, konkrete Informationen zu identifizieren oder wiederzugeben, solange sich diese im kontinuierlichen beziehungsweise diskontinuierlichen Text an leicht auffindbaren Stellen befinden, wie etwa am Anfang eines längeren Absatzes, und wenn Teile der Formulierung der Aufgabe mit der Formulierung der gesuchten Informationen weitgehend identisch sind. Die Anforderungen sind dabei in der Regel nicht dadurch erschwert, dass der Text plausible Distraktoren enthält. Müssen bereits Schlüsse gezogen werden, dann sind diese textbasiert und beziehen sich fast durchgängig auf die Verknüpfung von Informationen auf lokaler Ebene, also aus benachbarten Sätzen.

Die Aufgaben haben in der Regel Multiple-Choice-Formate, enthalten also lenkende Hinweise. Es kommen auch halboffene und offene Items vor, wobei jedoch allenfalls ein Wort zu schreiben ist oder auch ein Satz, der zumeist aber wörtlich aus dem Text übernommen werden kann. Schülerinnen und Schüler auf dieser Kompetenzstufe bleiben deutlich hinter den Erwartungen der KMK-Bildungsstandards zurück. Es ist davon auszugehen, dass der erfolgreiche Übergang in die Sekundarstufe I nur unter Einsatz intensiver Fördermaßnahmen gelingen wird.

Kompetenzstufe II (Punktwerte von 390 bis 464): benachbarte Informationen miteinander verknüpfen

Kinder auf Kompetenzstufe II können in der Regel nicht nur einzelne konkrete Informationen erkennen und wiedergeben. Sie sind häufiger als Kinder auf Kompetenzstufe I in der Lage, Anforderungen zu bewältigen, bei denen eine gegebene Information mit einer anderen zu verknüpfen ist. Dabei sind diese Informationen in der Regel zwar im Text nicht weit voneinander entfernt, die Verknüpfung ist jedoch nicht explizit gegeben, sondern muss erschlossen werden. Die Kinder können einfache kausale Beziehungen, Motive von Handlungen und Ursachen von Vorgängen erkennen. Die Lokalisierung von Informationen ist im Vergleich mit Kompetenzstufe I manchmal dadurch erschwert, dass im Text oder in der Aufgabenstellung plausible Distraktoren vorkommen. Wie auf Kompetenzstufe I handelt es sich fast durchgängig um Multiple-Choice-Aufgaben und einfache halboffene oder offene Items.

Bei diskontinuierlichen Texten können die Schülerinnen und Schüler nicht nur einzelne Informationen in Texten lokalisieren, sondern sie können auch Verstehensanforderungen bewältigen, bei denen zyklisch operiert, eine Suchoperation im Text also mehrfach vollzogen werden muss. Insgesamt können Leistungen auf Kompetenzstufe II als Erreichen des *Mindeststandards* angesehen werden.

Kompetenzstufe III (Punktwerte von 465 bis 539): verstreute Informationen miteinander verknüpfen und den Text ansatzweise als Ganzen erfassen

Den Kern der Anforderungen, die Kinder auf Kompetenzstufe III bewältigen können, machen Aufgaben aus, bei denen es um die Verknüpfung beziehungsweise Integration von Informationen geht, die über den jeweiligen Text mehr oder weniger verstreut sind. Dabei handelt es sich beispielsweise um den Abgleich der Angaben von Tageszeiten zu Beginn und am Ende eines Textes, um die Begründung semantischer Ähnlichkeiten entfernt platzierter Sätze oder um die Angabe mehrerer Verhaltensmotive einer Figur. Fragen wie die, welches von mehreren genannten Textthemen wohl am ehesten zutrifft und welche Überschrift angemessener ist als andere, zielen ebenso auf ein Verständnis des gesamten Textes wie Aufgaben, bei denen eine rudimentäre Bewertung („*Wie findest du diese Geschichte?*“) in Form einer Kurzantwort – eines Adjektivs und eines begründenden Satzes – gefragt ist. Auch solche Anforderungen werden von Schülerinnen und Schülern auf Kompetenzstufe III weitgehend bewältigt.

Erst auf Kompetenzstufe III können die Kinder einen Text ansatzweise als Ganzen erfassen. Damit werden auf dieser Stufe Anforderungen mit hinreichender Sicherheit bewältigt, die dem *Regelstandard* entsprechen.

Kompetenzstufe IV (Punktwerte von 540 bis 614): für die Herstellung von Kohärenz auf der Ebene des Textes wesentliche Aspekte erfassen

Kinder auf Kompetenzstufe IV können über die bisher dargestellten Anforderungen hinaus auch Aufgaben lösen, bei denen vornehmlich im offenen Format nach textnahen Begründungen oder Motiven für Verhaltensweisen, Handlungen, Äußerungen und Ursachen von Emotionen gefragt wird. Dabei sind zur Herstellung von Kohärenz in der Regel Schlüsse zu ziehen, die sich nicht aus einer einzelnen Textstelle ergeben, sondern für die Informationen aus verschiedenen Textpassagen miteinander kombiniert werden müssen. Zu kombinieren und textnah zu schließen ist auch, wenn danach gefragt wird, welcher der Protagonisten eines Erzähltextes der Ich-Erzähler ist. Externes Wissen muss für die Antwort auf die Frage nach der Textsorte (z. B. Fabel) aktiviert werden. Diese

Anforderungen bewältigen die Schülerinnen und Schüler auch dann, wenn die Zuordnung der wesentlichen Informationen dadurch erschwert ist, dass im Text ähnliche Informationen vorkommen, die ablenkend wirken.

Kinder auf Kompetenzstufe IV übertreffen die in den Bildungsstandards der KMK formulierten Kompetenzerwartungen. Dementsprechend wird diese Stufe als *Regelstandard plus* bezeichnet.

Kompetenzstufe V (Punktwerte ab 615): auf zentrale Aspekte des Textes bezogene Aussagen selbstständig begründen

Kinder auf der höchsten Kompetenzstufe V können Textthemen und Handlungsmotive identifizieren, die im Text nicht explizit genannt sind. Zudem sind sie in der Lage, eigenständig über Texte nachzudenken und in den Aufgaben formulierte Behauptungen über Aspekte der Texte selbstständig zu begründen, auch wenn die verlangten Schlüsse komplexer sind oder wenn stärker auf Vorwissen zurückgegriffen werden muss. So können sie zum Beispiel vorgegebene Deutungshypothesen zum Zusammenhang von Emotionen und Verhaltensweisen durch die Angabe von Textstellen plausibel stützen.

Bei Aufgaben zu diskontinuierlichen Texten geht es durchgängig um zyklisches Lokalisieren, wobei der Suchraum jeweils das gesamte Dokument ist. Dabei hat die in komplexen Dokumenten zu lokalisierende Information mehrere Bedingungen zu erfüllen und ist von zahlreichen konkurrierenden Angaben nicht leicht zu unterscheiden. Kinder auf Kompetenzstufe V können deutlich anspruchsvollere Anforderungen bewältigen, als in der Regel zu erwarten ist. Sie entwickeln eigene Gedanken zu Texten und nehmen zu ihnen Stellung.

Leistungen auf der Kompetenzstufe V können daher als Erreichen eines *Optimalstandards* angesehen werden.

Zur Illustration der Kompetenzstufen I bis V finden sich nachfolgend in der Abbildung 3.5 der Stimulustext der Aufgabe „Der aufmerksame Beobachter“ und in der Abbildung 3.6 einige Beispielitems dieser Aufgabe.

Abbildung 3.5: Stimulustext der Aufgabe „Der aufmerksame Beobachter“

Der aufmerksame Beobachter

Ein kasachisches Märchen

Einmal kam einem Mann ein Kamel aus seiner Herde abhanden. Als er auszog, um es zu suchen, holte er in der Steppe einen Reiter ein. Sie begrüßten einander und steckten sich ihre Pfeifchen an.

„Ich hab ein Kamel verloren“, klagte der Mann. „Hast du es nicht gesehen?“

„Ist dein Kamel auf dem linken Auge blind, und fehlen ihm die Vorderzähne?“

„Jaja!“, rief der Mann froh. „Wo ist es denn?“

„Ich weiß nicht, wo dein Kamel ist, ich sah nur gestern seine Spuren.“



Der Besitzer des Kamels aber glaubte dem Reiter nicht, sondern beschuldigte ihn, es gestohlen zu haben, und führte ihn vor den Richter.

Der fremde Mann sagte zum Richter: „Ich kann noch mehr über das Kamel sagen und habe es doch nicht gesehen.“

„Nun, so sprich!“

„Auf der einen Seite trug es ein Fässchen mit Honig, auf der anderen einen prallen Sack Weizen.“

„Jaja, er ist der Dieb!“, rief der Besitzer des Kamels.

Sogar der Richter glaubte das jetzt, doch fragte er den Angeklagten lieber noch einmal: „Hast du das Kamel gesehen?“

„Nein.“

„Woher aber weißt du das alles?“

„Nun, dass das Kamel auf dem linken Auge blind ist, sah ich daran, dass nur rechts von seinem Weg Gras abgefressen war.“ – „Und woher weißt du, dass es keine Vorderzähne hat?“

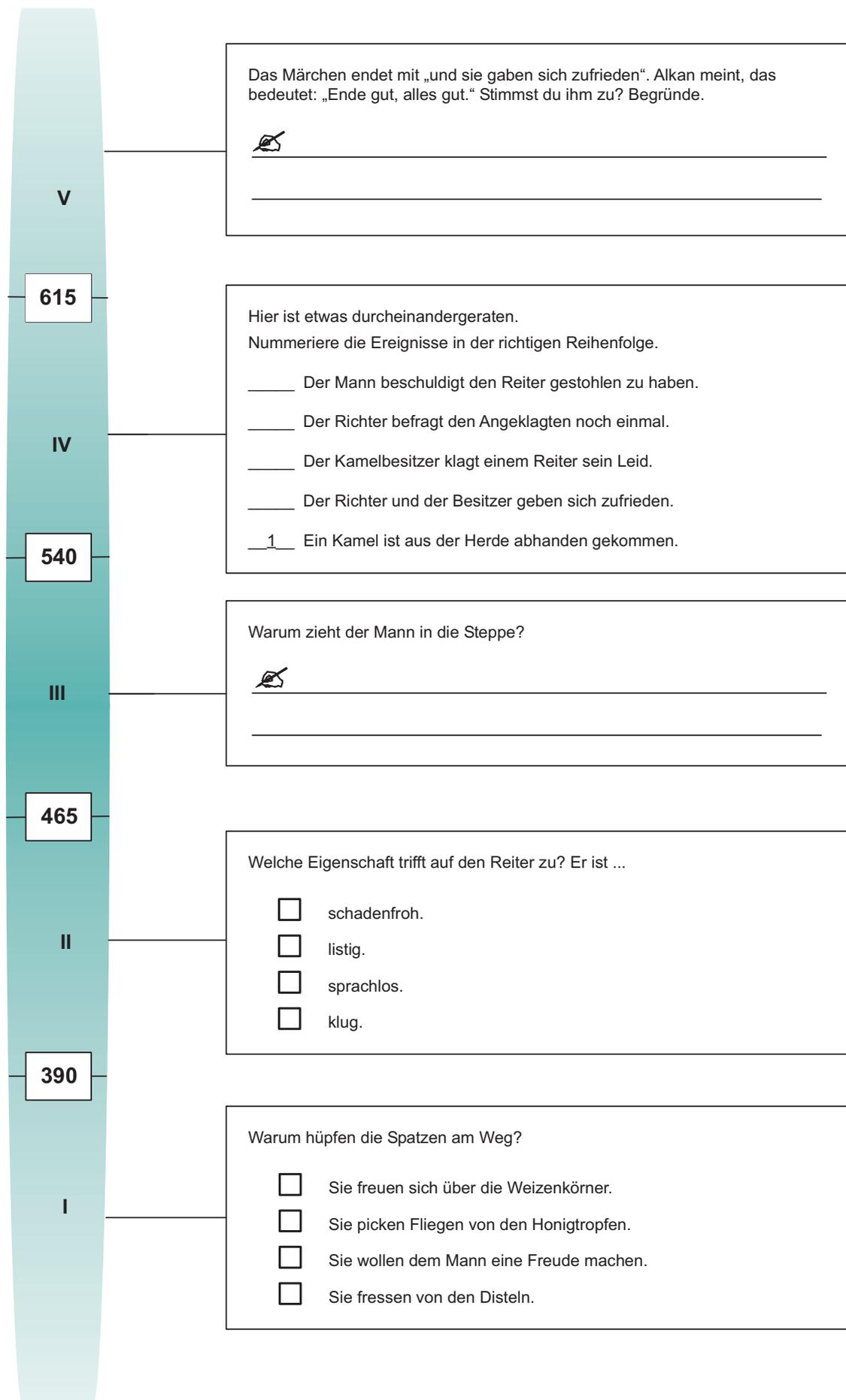
„Beim Grasens blieben in der Mitte immer einige Büschel der schmackhaften Disteln stehen.“

„So – und nun sag uns noch, woher du weißt, dass das Kamel Honig und Weizen trug?“

„Ganz einfach – auf der einen Seite des Weges saßen die Fliegen auf den Honigtropfen, und auf der anderen hüpfen die Spatzen und suchten Weizenkörner.“

„Ja, wenn das so war, dann glauben wir dir!“, riefen der Richter und der Kamelbesitzer, und sie gaben sich zufrieden.

Abbildung 3.6: Illustration des Kompetenzstufenmodells für den Bereich Lesen



3.2.3 Das Kompetenzstufenmodell für den Teilbereich Orthografie

Das Kompetenzstufenmodell für den Teilbereich *Orthografie* wurde 2010 im Rahmen eines Standard-Settings unter Einbindung von Lehrkräften, Fachdidaktikerinnen und Fachdidaktikern sowie Expertinnen und Experten aus dem Bereich Psychometrie primär auf Grundlage der Normierungsstudie für den Primarbereich (vgl. Kapitel 1) entwickelt und Ende des Jahres 2010 von der Kultusministerkonferenz zustimmend zur Kenntnis genommen.

Als Referenzstichprobe für die in den nachfolgenden Kapiteln berichteten Befunde dient die vierte Jahrgangsstufe des 2011 durchgeführten Ländervergleichs. Dies bedeutet, dass die Kompetenzverteilung der im Rahmen der vorliegenden Studie getesteten Viertklässlerinnen und Viertklässler einen Mittelwert von genau 500 Punkten und eine Standardabweichung von genau 100 Punkten aufweist.

Bei der Charakterisierung der Kompetenzstufen wird auf die in Kapitel 2.1.5 eingeführten Fehlerlupenstellen und ihre Zuordnung zu den verschiedenen Strategien im Erwerb der Rechtschreibkompetenz Bezug genommen. Dabei ist zu berücksichtigen, dass man beim Rechtschreiberwerb nicht von strikt aufeinander aufbauenden und sukzessive erreichten Entwicklungsstufen ausgehen darf. Während mitunter bei einigen Wörtern, die zum Beispiel gängig oder oft geübt worden sind, bereits Elemente fortgeschrittener Strategien beherrscht werden, können bei weniger gängigen oder ungeübten Wörtern parallel noch sehr basale Fehler auftreten.

Darüber hinaus kann, wie sich auch empirisch gezeigt hat, davon ausgegangen werden, dass die Anwendung automatisierter Wissensbestände beim eigenen Schreiben einfacher ist als der deklarative Einsatz dieses Wissens bei der orthografischen Korrektur fremder Texte.

Das Modell umfasst fünf Kompetenzstufen mit einer Stufenbreite von jeweils 75 Punkten, wobei die Stufe I nach unten und die Stufe V nach oben offen ist. Konkret wurden folgende Kompetenzstufen bestimmt:

- Kompetenzstufe I: unter 390 Punkten
- Kompetenzstufe II: Werte von 390 bis 464 Punkten (Mindeststandard)
- Kompetenzstufe III: Werte von 465 bis 539 Punkten (Regelstandard)
- Kompetenzstufe IV: Werte von 540 bis 614 Punkten (Regelstandard plus)
- Kompetenzstufe V: Werte von 615 Punkten und mehr (Optimal- bzw. Maximalstandard)

Kompetenzstufe I (Punktwerte unter 390)⁹

Kinder auf dieser Kompetenzstufe schreiben Wörter in der Regel so, dass ihre Lautstruktur erkennbar ist. Einige Schülerinnen und Schüler beherrschen aber die Laut-Buchstaben-Zuordnungen beziehungsweise die alphabetische Stufe noch nicht hinreichend.

Bei den Wörtern, die mit hoher Wahrscheinlichkeit von allen Schülerinnen und Schülern auf Kompetenzstufe I noch nicht richtig geschrieben werden, beträgt der Fehleranteil, der auf die basale Kategorie der Graphemauswahl (GA)

⁹ Anders als bei den Modellen für die Bereiche Zuhören und Lesen sind die Stufen des Modells für den Bereich Orthografie nicht mit Überschriften versehen. Da für die qualitative Fehleranalyse eine große Zahl an Kategorien genutzt wurde, fallen die Befunde entsprechend differenziert aus. Man hätte zwar für jede Stufe eine Kategorie als prototypisch ausweisen können. Ein solcher Prototyp würde die jeweilige Kompetenzstufe aber nur unzureichend abdecken und wäre daher irreführend.

entfällt, etwa 10 Prozent. Der Anteil der Graphemfolgefehler (FG) ist etwas höher.

Zu den Schreibungen, die sich auf die orthografische Stufe beziehen, lässt sich sagen, dass ein Drittel der Wörter mit <st> oder <sp> am Wort- beziehungsweise Morphemanfang richtig geschrieben wird. Auch die Schreibung einiger Wörter mit Dehnungs-h gelingt (VL-). Die Vokalkürze (VK) wird bereits häufig richtig markiert. Konsonantische Ableitungen (KA) meistern die Schülerinnen und Schüler vor allem in strukturell einfachen Wörtern wie *Hand*. Ebenfalls richtig schreiben die Kinder knapp die Hälfte der Wörter mit den Präfixen *ver-* und *vor-*. In Wörtern wie *Hausschuhe* und *Hexenmeister* wird die Morphemgrenze (MG) korrekt verschriftet. Was die Groß- beziehungsweise Kleinschreibung (GK) angeht, so werden einige Substantive mit gegenständlicher Bedeutung (Konkreta) wie *Gabel* richtig großgeschrieben.

Mit Korrekturaufgaben tun sich Kinder auf dieser Kompetenzstufe noch recht schwer. In einer Reihe von Fällen gelingt ihnen aber die orthografische Korrektur von alphabetischen Schreibungen wie *file* (statt *viele*). Auch Wörter wie *kempfer* (statt *Kämpfer*) und *Beume* (statt *Bäume*) können erfolgreich korrigiert werden, ebenso einige Fälle konsonantischer Ableitung wie bei *Berk* (statt *Berg*) und *Sant* (statt *Sand*).

Auf dieser Kompetenzstufe kann das für die vierte Jahrgangsstufe zu erwartende Kompetenzminimum im Sinne von *Mindeststandards* noch nicht als erreicht gelten.

Kompetenzstufe II (Punktwerte von 390 bis 464)

Auf dieser Kompetenzstufe werden die elementaren Laut-Buchstaben-Beziehungen beherrscht.

Die Kinder verschriften darüber hinaus mehr als drei Viertel der Wörter mit den speziellen Graphemen (SG) <st> oder <sp> am Wort- oder Morphemanfang richtig, ebenso Wörter mit <pf> beziehungsweise mit Konsonantenhäufungen wie *Pflanzen*, *pflügen* und *stumpf*. In der Hälfte der Fälle kann die Markierung der Vokallänge (VL-) geleistet werden und es werden auch mehrere Wörter mit silbenanlautendem *h* (wie *ruhig*) richtig geschrieben. Bewältigt werden ebenfalls circa 60 Prozent der Wörter, bei denen die Vokalkürze (VK) zu berücksichtigen ist. Nicht nur in einfachen, sondern auch in strukturell komplexeren Wörtern wie *gründlich* gelingt die konsonantische Ableitung (KA). Die Hälfte aller Lückenwörter, bei denen es unter anderem um vokalische Ableitung beziehungsweise Umlautung (VA) geht, wird auf dieser Kompetenzstufe richtig geschrieben (z. B. *nächste*). Bei den unmittelbar morphembezogenen Kategorien werden fast alle Wörter mit dem Suffix *-lich* korrekt verschriftet, ebenso drei Viertel der Wörter mit den Präfixen *ver-* beziehungsweise *vor-*. Beherrscht wird auch die Mehrzahl der Fälle, bei denen die Kategorie Morphemgrenze (MG) signiert ist, zum Beispiel bei *Fahrrad*. Korrekt großgeschrieben werden weitere Substantive mit gegenständlicher Bedeutung, vereinzelt auch ein Abstraktum (*Gesundheit*, *Verkehr*).

Bei den Wörtern, die von Kindern auf Kompetenzstufe II richtig korrigiert werden, geht es nicht mehr nur um alphabetische („lauttreue“) Schreibungen. In einem kurzen Text, in dem eine Reihe von Substantiven kleingeschrieben ist, können alle Falschschreibungen identifiziert werden. Die Aufgabe zu begründen, warum in Fällen wie *Kind* oder *Dieb* ein <d> oder zu schreiben ist, bewältigen die Kinder auf dieser Stufe ebenfalls mit hoher Wahrscheinlichkeit. Auch die Aufgabe, in einem kurzen Text fehlende Satzzeichen durch einen Strich zu mar-

kieren, meistern sie auf dieser Kompetenzstufe im Wesentlichen. Schließlich gelangen ihnen einige Sortieraufgaben¹⁰, bei denen die jeweils an zweiter Stelle stehenden Grapheme der zu reihenden Wörter unterschiedlich sind.

Die auf dieser Stufe erreichten Kompetenzen entsprechen noch nicht den Anforderungen, die von der Kultusministerkonferenz als Regelstandards definiert wurden, sie können aber als *Mindeststandard* interpretiert werden.

Kompetenzstufe III (Punktwerte von 465 bis 539)

Kinder auf dieser Kompetenzstufe schreiben mehr als 80 Prozent der Testwörter richtig.

Auf dieser Kompetenzstufe gelingt erstmals in größerem Umfang die Schreibung von Varianten des stimmlosen [s] (z.B. *gießen, blies, vermisst*). Auch Schreibungen des <qu> sind auf dieser Kompetenzstufe fast durchgängig korrekt (z.B. *quer*). Wenn es bei der Markierung der Vokallänge (VL-) um das Dehnungs-h geht, werden auf diesem Niveau vier Fünftel der Fälle gemeistert. Ebenso werden 80 Prozent der Wörter, bei denen die Vokalkürze (VK) zu markieren ist, bewältigt. Auch bei fast allen Beispielen konsonantischer Ableitungen (KA) ist die Schreibung korrekt. Charakteristisch für diese Kompetenzstufe ist, dass fast alle Beispiele vokalischer Ableitung beziehungsweise Umlautung (VA) richtig geschrieben werden können (z.B. *Gefängnis*). Hinsichtlich der unmittelbar morphembezogenen Kategorien werden alle Arten von Suffixen (HM) bewältigt, die im Korpus der Testwörter vorkommen (u. a. *-nis*). Richtig großgeschrieben (GK) werden neben Wörtern mit gegenständlicher Bedeutung jetzt auch Abstrakta, die als Substantive markiert sind (*Beschreibung, Ergebnis*).

Über die auf Kompetenzstufe II erfolgreich geleisteten Korrekturen hinaus werden nur wenige Schreibfehler erkannt, so zum Beispiel bei *Welenrauschen* (nicht markierte Vokalkürze). Allerdings können auf dieser Kompetenzstufe fast durchweg Begründungen für die Schreibung von Auslautverhärtungen beziehungsweise dafür gegeben werden, inwiefern es sich jeweils um konsonantische Ableitungen handelt (z.B. bei *sagt*).

Schließlich können fünf von sechs alphabetischen Sortieraufgaben mit relativ hoher Wahrscheinlichkeit gelöst werden. Bei den Aufgaben, die über die auf Kompetenzstufe II lösbaren hinaus gemeistert werden, geht es um die richtige Reihung von Wörtern, die nicht nur an zweiter Stelle, sondern zusätzlich an dritter Stelle identische Grapheme aufweisen.

Die Kompetenzstufe III lässt sich auf der Grundlage der Vorgaben der KMK als *Regelstandard* interpretieren.

Kompetenzstufe IV (Punktwerte von 540 bis 614)

Auf dieser Kompetenzstufe werden mehr als 90 Prozent der Wörter richtig geschrieben.

Es gelingt hier die Schreibung aller Wörter mit speziellen Graphemen (SG), insbesondere mit <qu>. Richtig geschrieben werden ebenfalls einige Wörter mit Dehnungs-h (VL-), die im aktiven Wortschatz der meisten Kinder selten sein dürften, wie *empfehlen*. Auch die Schreibung von Wörtern wie *kaputt* und *Schlittschuhläufer* wird gemeistert, bei denen die Markierung der Vokalkürze (VK) wegen der Betonungsverhältnisse beziehungsweise der Komplexität des Kompositums schwierig ist. Vokalisches und konsonantisches Ableitungen (VA, KA) bereiten fast durchgängig keine Schwierigkeiten mehr. Neben anderen

¹⁰ Sortieraufgaben wurden in den Pilotierungs- und Normierungsstudien eingesetzt, waren aber nicht Bestandteil des Ländervergleichs 2011.

Substantiven werden erstmals Nominalisierungen richtig großgeschrieben (GK; z. B. *Treffen*), was eine syntaktische Analyse voraussetzt.

Über die Korrektur von Fehlern in Bezug auf Vokalkürze (z. B. *Heuschrecken* statt *Heuschrecken*) hinaus gelingen auf dieser Kompetenzstufe auch Korrekturen von Fehlern im Feld der vokalischen Ableitung und es kann zusätzlich explizit angegeben werden, warum es sich um Fehler handelt (z. B. *Treume* ist falsch, *Träume* ist richtig, „weil es von ‚Traum‘ kommt“). Während eine Reihe von Wörtern mit stimmlosem [s] bereits auf Kompetenzstufe III richtig geschrieben werden kann, gelingt eine Korrekturaufgabe zum stimmlosen [s] in Form einer Richtig-Falsch-Batterie erst auf Kompetenzstufe IV.

Auf dieser Kompetenzstufe werden somit Anforderungen bewältigt, die bereits über den Vorgaben der KMK liegen, so dass sie als *Regelstandard plus* zu interpretieren ist.

Kompetenzstufe V (Punktwerte ab 615)

Auf dieser Kompetenzstufe stehen Wörter im Vordergrund, deren morphologische Struktur schwer zu erkennen ist oder für deren korrekte Schreibung auf syntaktisches Wissen zurückgegriffen werden muss.

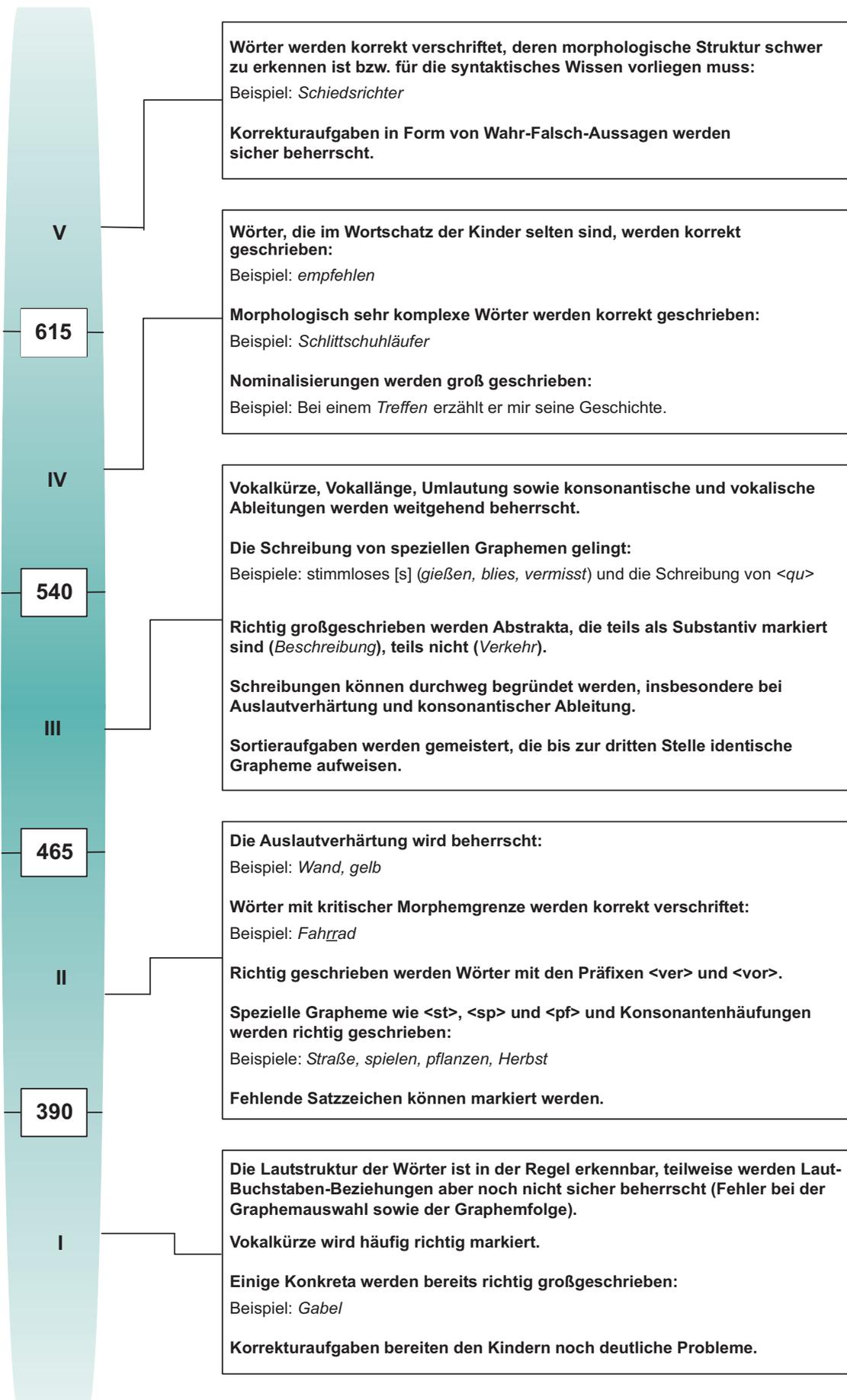
Hervorzuheben ist zum Beispiel das Wort *Schiedsrichter*, dessen Verwandtschaft mit *entscheiden* Kinder auf den unteren Niveaus vermutlich nicht bemerken. Was das syntaktische Wissen angeht, so gelingen die in den Testungen vorkommenden Nominalisierungen, wobei es sich im Test durchgängig um Verben handelt (GK, Großschreibung im Satzkontext).

Auf Kompetenzstufe V können alle Korrekturaufgaben in Form von Wahr-Falsch-Entscheidungen bewältigt werden. Schreiben die Kinder selbst, haben sie mit der <sp>- beziehungsweise <st>-Regel offensichtlich weniger Schwierigkeiten als mit der Beurteilung von Schreibungen wie *schpülen* und *vorschtrecken*.

Die Kompetenzstufe V lässt sich als *Optimalstandard* interpretieren.

Zur Illustration der Kompetenzstufen I bis V im Teilbereich Orthografie stellt Abbildung 3.7 exemplarisch dar, welche konkreten Anforderungen Schülerinnen und Schüler auf den verschiedenen Kompetenzstufen mit hoher Wahrscheinlichkeit bewältigen.

Abbildung 3.7: Illustration des Kompetenzstufenmodells für den Teilbereich Orthografie



3.3 Kompetenzstufenmodelle im Fach Mathematik

Kristina Reiss, Alexander Roppelt, Nicole Haag, Hans Anand Pant
und Olaf Köller

Für die Bildungsstandards im Fach Mathematik für den Primarbereich (KMK, 2005b) liegen Kompetenzstufenmodelle zu den fünf inhaltlichen Kompetenzbereichen sowie für das Fach insgesamt (Globalmodell) vor. Die im Folgenden vorgestellten Beschreibungen der Modelle wurden bereits in anderen Veröffentlichungen des IQB dargestellt (vgl. bspw. IQB, 2008; Reiss & Winkelmann, 2009), sie präzisieren jedoch einige Details. Im Einklang mit dem in Kapitel 3.1 dargestellten Vorgehen wurden die Berichtsskalen so gebildet, dass ihr Mittelwert (M) jeweils 500 und die Standardabweichung (SD) jeweils 100 Punkte für die Schülerinnen und Schüler der vierten Jahrgangsstufe beträgt. Die Stufengrenzen wurden für alle sechs Modelle einheitlich gesetzt. Der Wert 390 markiert jeweils den Übergang von der ersten zur zweiten Stufe, und alle weiteren Stufenübergänge folgen in einem Abstand von 70 Punkten.

Zunächst werden die Kompetenzstufen des Globalmodells dargestellt und mit Beispielaufgaben illustriert. Daran schließen sich Beschreibungen der einzelnen Kompetenzbereiche an, welche die Stufen des Globalmodells für den jeweiligen Inhaltsbereich detaillierter charakterisieren. Beispielaufgaben zu den einzelnen Kompetenzstufen finden sich in Abbildung 3.8.

3.3.1 Globalmodell

Kompetenzstufe I (Punktwerte unter 390): technische Grundlagen (Routineprozeduren auf Grundlage einfachen begrifflichen Wissens)

Schülerinnen und Schülern auf Kompetenzstufe I sind einfache mathematische Begriffe und Prozeduren bekannt und sie können diese in einem innermathematischen Kontext beziehungsweise in einem aus dem Alltag vertrauten oder gut geübten Kontext korrekt reproduzieren. Im Einzelnen werden die Grundaufgaben des kleinen Einpluseins und Einmaleins beherrscht und bei mündlichen, halbschriftlichen und schriftlichen Rechenverfahren genutzt, wenn die Aufgabenstellungen keine besonderen Schwierigkeiten aufweisen. Darüber hinaus werden sie auch in sehr einfachen Sachsituationen korrekt angewendet. Außerdem können Zahlen in Bezug auf ihre Größe verglichen und Zahldarstellungen in Stellentafeln insbesondere im Tausenderraum sicher gelesen werden. Grundlegende Begriffe der ebenen Geometrie (z. B. Kreis, Quadrat, Dreieck) werden bei prototypischen Darstellungen richtig verwendet. Sehr einfache Folgen und Muster können fortgesetzt werden. Gängige Größeneinheiten (z. B. m, km, kg) können gut vertrauten Repräsentanten zugeordnet werden. Auch einfache Größenvergleiche werden geleistet. Einfachen, klar strukturierten Diagrammen, Schaubildern und Tabellen mit Bezug zur Lebenswirklichkeit können unmittelbar ersichtliche Daten entnommen werden. Umgekehrt können einfache Informationen in eine Tabelle eingetragen werden. Es gelingt, sehr einfache und sehr anschauliche Zufallsexperimente in Bezug auf Gewinnchancen zu vergleichen.

Die auf diesem Niveau beschriebenen Kompetenzen sind als rudimentär zu werten; die eigentlichen Ziele des Mathematikunterrichts in der Grundschule werden weitgehend noch nicht erreicht.

Kompetenzstufe II (Punktwerte von 390 bis 459): einfache Anwendungen von Grundlagenwissen (Routineprozeduren in einem klar strukturierten Kontext)

Schülerinnen und Schüler mit Mathematikkompetenzen auf der Kompetenzstufe II können Grundlagenwissen in einfachen, klar strukturierten und bekannten Anwendungen nutzen. So werden Aufgaben zur Addition, Subtraktion und Multiplikation halbschriftlich und schriftlich durchgeführt, wobei auch Überschlagsrechnungen geleistet und Größenordnungen von Ergebnissen korrekt erkannt werden. Außerdem wird die Struktur des Dezimalsystems genutzt, Gesetzmäßigkeiten werden erkannt und bei der Fortsetzung einfacher Zahlenfolgen, beim strukturierten Zählen und systematischen Probieren berücksichtigt. Grundbegriffe der räumlichen Geometrie (z.B. Würfel) werden korrekt verwendet, wenn diese einen Bezug zum Alltag haben. Einfache räumliche Gebilde werden in ihren strukturellen Eigenschaften gesehen und räumliche Beziehungen zur Lösung einfacher Probleme genutzt. Dabei werden einfache Lagebeziehungen auch dann korrekt erkannt, wenn hierfür die Perspektive einer anderen Person einzunehmen ist. Auch komplexere geometrische Muster und Zahlenreihen werden in ihrer Struktur erkannt und fortgesetzt. Es können einfache Sachaufgaben gelöst werden. Aus dem Alltag vertraute proportionale Zuordnungen werden erkannt und angewendet. Bei einfachem Zahlenmaterial wird das Umwandeln von Größen in gegebene Einheiten auch bei gemischten Größenangaben beziehungsweise einfachen Kommazahlen durchgeführt. Klar strukturierten, auch umfangreicheren Diagrammen, Schaubildern und Tabellen können unterschiedliche Daten entnommen werden. Diese müssen aber in der Regel noch explizit gegeben sein. Umgekehrt können einfache Daten nach gewissen Vorgaben dort auch eingetragen werden. Wesentliche Grundbegriffe aus dem Umfeld von Zufall und Wahrscheinlichkeit wie „sicher“, „unmöglich“ und „wahrscheinlich“ werden korrekt verwendet. Einfache Zufallsexperimente können in Bezug auf Gewinnchancen eingeschätzt werden.

Schülerinnen und Schüler auf dieser Kompetenzstufe erfüllen noch nicht die in den Bildungsstandards der KMK für die vierte Jahrgangsstufe formulierten Erwartungen. Sie besitzen jedoch bereits grundlegende Fähigkeiten, die über die erste Kompetenzstufe hinausgehen. Die zweite Kompetenzstufe kann als Erreichen der *Mindeststandards* interpretiert werden.

Kompetenzstufe III (Punktwerte von 460 bis 529): Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext

Schülerinnen und Schüler, die Kompetenzstufe III erreichen, können das erlernte Wissen flexibel in unterschiedlichen Problemstellungen innerhalb eines vertrauten Kontextes nutzen. Insbesondere wird mit Zahlen und Operationen im curricularen Umfang sicher umgegangen und Überschlagsrechnungen werden auch bei großen Zahlen sicher durchgeführt. Strukturelle Aspekte werden zumindest bei gut geübten Inhalten gesehen und können kommuniziert werden. Dies gilt auch für Inhalte der Geometrie, wobei etwa zwischen verschiedenen Darstellungsformen einer Figur vermittelt werden kann. Beispielsweise werden beim Umgang mit Netzen ebene und räumliche Informationen aufeinander bezogen. Weiterhin sind

Grundlagen geometrischer Abbildungen verfügbar, sodass insbesondere einfache Achsenspiegelungen durchgeführt werden können. Zahlenfolgen, die nach komplexeren Regeln aufgebaut sind, können fortgesetzt und fehlerhafte Zahlen in überschaubaren Zahlenfolgen gegebenenfalls identifiziert werden. Einfache Sachsituationen werden modelliert und die damit verbundenen Problemstellungen gelöst. Proportionale Zuordnungen werden sicher genutzt. Der Umgang mit Größen ist flexibel und berücksichtigt – etwa bei Uhrzeiten – spezielle Eigenschaften der Größenbereiche. Hierbei wird auch die Kommaschreibweise korrekt verwendet. Darüber hinaus können in Größenbereichen den gängigen Einheiten geeignete Repräsentanten zugeordnet werden. Daten und Informationen können in bekanntem Kontext flexibel dargestellt werden. Dabei gelingt auch der Umgang mit großen Zahlen. Diagrammen können Informationen entnommen werden, die nicht direkt ablesbar sind, sondern erschlossen oder berechnet werden müssen. Der Informationsgehalt einfacher Diagramme kann explizit beurteilt werden. Bei nicht allzu komplexen Zufallsexperimenten werden Gewinnchancen korrekt eingeschätzt und begründet.

Schülerinnen und Schüler, die diese Kompetenzstufe oder eine höhere erreicht haben, erfüllen die in den Bildungsstandards beschriebenen Erwartungen, so dass die Zugehörigkeit zur Kompetenzstufe III bereits als Erreichen der *Regelstandards* gelten kann.

Kompetenzstufe IV (Punktwerte von 530 bis 599): sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang

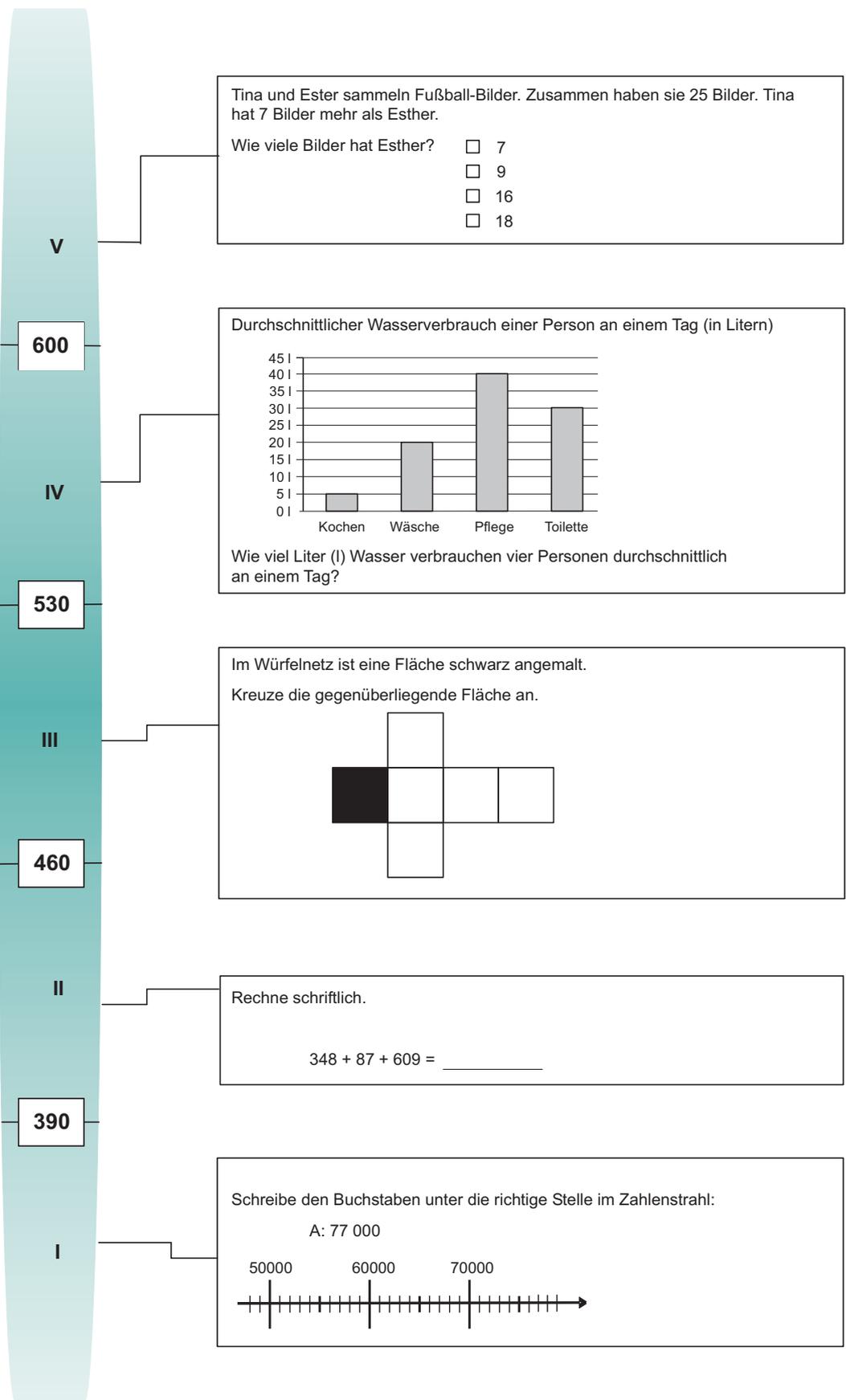
Schülerinnen und Schüler auf Kompetenzstufe IV wenden auch in einem wenig vertrauten Kontext mathematisches Wissen sicher an. Sie beschreiben eigene Vorgehensweisen korrekt, verstehen und reflektieren die Lösungswege anderer Kinder und beherrschen das Rechnen im curricularen Umfang in allen Varianten sicher. Zahldarstellungen in Stellenwerttafeln können auch bei sehr großen Zahlen (also im Zahlenraum bis zu einer Million) nach Vorschrift selbstständig manipuliert und systematisch verändert werden. Begriffe der ebenen und räumlichen Geometrie werden flexibel verwendet und geometrische Aussagen können hinterfragt und diskutiert werden. Auch bei komplexen Zahlenmustern wird die dahinterliegende Regel erkannt und das Muster korrekt fortgesetzt. Das Rechnen mit Größen ist sicher und flexibel und umfasst insbesondere auch Näherungs- und Überschlagsrechnungen. Modellierungsaufgaben werden selbst dann gelöst, wenn sie nicht unmittelbar auf einfachen Alltagserfahrungen basieren. Informationen aus unterschiedlichen Quellen können in einen Zusammenhang gestellt und in Modellierungsaufgaben selbstständig verwendet und manipuliert werden. Komplexere Zufallsexperimente werden angemessen beurteilt, mögliche Ergebnisse werden korrekt bestimmt.

Mit der Zugehörigkeit zur Kompetenzstufe IV ist ein Kompetenzniveau erreicht, das über den Erwartungen der Bildungsstandards liegt. Schülerinnen und Schüler, die sich mindestens auf dieser Stufe befinden, erreichen damit den *Regelstandard plus*.

Kompetenzstufe V (Punktwerte ab 600): Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien

Schülerinnen und Schüler auf der höchsten Kompetenzstufe bearbeiten mathematische Problemstellungen in allen Inhaltsbereichen auch in einem unbekanntem Kontext angemessen, sicher und flexibel. Dabei werden auf hohem

Abbildung 3.8: Globales Kompetenzstufenmodell und illustrierende Aufgaben im Fach Mathematik



Aus Platzgründen sind die Aufgaben in modifiziertem Layout dargestellt.

Niveau geeignete Strategien angewendet, sinnvolle Bewertungen abgegeben oder Verallgemeinerungen geleistet. Umfangreiches curricular verankertes Wissen wird auch in ungewohnten Situationen flexibel genutzt. Das Vorgehen kann nachvollziehbar kommuniziert und begründet werden. Mathematische Argumentationen werden angemessen bewertet. Komplexe Sachsituationen werden modelliert und bearbeitet, selbst wenn besondere Schwierigkeiten wie die Verwendung von Tabellen, der Umgang mit zusammengesetzten Größen oder das Rechnen mit Zahlen in Kommaschreibweise auftreten. Es können auch ungewohnte funktionale Zusammenhänge analysiert und genutzt werden. Die Lösung von Aufgaben gelingt auch dann, wenn sie ein hohes Maß an räumlichem Denken oder entsprechenden analytischen Fähigkeiten voraussetzt. Der Umgang mit Daten ist genauso wie die mathematische Aufbereitung von Zufallsexperimenten durch selbstständiges Arbeiten geprägt.

Mit der Kompetenzstufe V ist ein Leistungsniveau erreicht, das in der Regel nur bei sehr günstigen individuellen Voraussetzungen und optimalen schulischen und außerschulischen Lernangeboten erreicht werden kann. Es wird daher als *Optimalstandard* (Maximalstandard) bezeichnet.

Die Abbildung 3.8 illustriert die fünf Kompetenzstufen des Globalmodells mit Beispielaufgaben aus unterschiedlichen inhaltlichen Bereichen.

3.3.2 Zahlen und Operationen

Kompetenzstufe I (Punktwerte unter 390): technische Grundlagen (Routineprozeduren auf Grundlage einfachen begrifflichen Wissens)

Auf dieser Kompetenzstufe des Bereichs „Zahlen und Operationen“ sind die Grundlagen der Struktur des Dezimalsystems wie die Einteilung in Einer, Zehner, Hunderter usw. bekannt. Alle Grundaufgaben des kleinen Einspluseins und Einmaleins werden beherrscht. Kleinere Zahlen können halbschriftlich addiert und subtrahiert werden. Dabei können einfache Rechnungen kontrolliert und auf Korrektheit geprüft werden. Das Verfahren der schriftlichen Addition wird erfolgreich bei zwei Summanden verwendet. Die schriftliche Subtraktion wird ausgeführt, wenn kein Übertrag zu berücksichtigen ist. Der Zusammenhang zwischen Addition und Subtraktion wird bei einfachen Rechnungen genutzt. Bereits aus den ersten Schuljahren bekannte Strategien, wie etwa das Verdoppeln, werden auch in größeren Zahlenräumen angewendet. Einfache Zahlen insbesondere im Hunderter- oder Tausenderraum können am Zahlenstrahl bei gegebener Skalierung dargestellt und in Bezug auf ihre Größe verglichen werden.

Kompetenzstufe II (Punktwerte von 390 bis 459): einfache Anwendungen von Grundlagenwissen (Routineprozeduren in einem klar strukturierten Kontext)

Die Struktur des Dezimalsystems wird beim Umgang mit Zahlen in verschiedenen Darstellungen genutzt. Es werden Gesetzmäßigkeiten erkannt und bei der Fortsetzung einfacher Zahlenfolgen sowie beim strukturierten Zählen berücksichtigt. Einfache Aufgaben zu den Grundrechenarten werden im Kopf sowie halbschriftlich und schriftlich bearbeitet und gegebenenfalls auch durch systematisches Probieren gelöst. Dabei werden Überschlagsrechnungen durchgeführt und zur Bestimmung der Größenordnung von Ergebnissen herangezogen. In Bezug auf die Grundrechenarten werden grundlegende mathematische Begriffe (z. B.

„Summe“) korrekt benutzt. Die Grundrechenarten können zur Bearbeitung einfacher Sachaufgaben genutzt werden.

Kompetenzstufe III (Punktwerte von 460 bis 529): Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext

Zahlen werden im curricularen Umfang in unterschiedlichen Darstellungen (z.B. in einer Stellenwerttafel) sicher gelesen und geschrieben, wobei auch die Ziffer Null korrekt berücksichtigt wird. Alle halbschriftlichen und schriftlichen Rechenverfahren werden beherrscht, wobei die Division auf einstellige Divisoren beschränkt ist. Die Grundaufgaben des mündlichen Rechnens werden auch in nicht vertrautem Kontext angewendet. Einmaleinsaufgaben werden sinngemäß auf einen größeren Zahlenraum übertragen. Auch im Zusammenhang mit einem Sachkontext werden Überschlagsrechnungen selbst bei großen Zahlen sicher durchgeführt und Ergebnisse sinnvoll gerundet. Die Beziehungen zwischen Addition und Subtraktion beziehungsweise Multiplikation und Division werden erkannt. Einfache strukturelle Aspekte werden bei gut geübten Inhalten gesehen und können kommuniziert werden (z.B. im Zusammenhang mit den Einmaleinsreihen). Im curricular gegebenen Zahlenraum werden einfache Sachsituationen modelliert und die damit verbundenen Problemstellungen gelöst.

Kompetenzstufe IV (Punktwerte von 530 bis 599): sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang

Das Rechnen wird im curricularen Umfang bei allen Rechenarten und in allen Varianten sicher beherrscht. Dies schließt insbesondere die schriftliche Division ein. Dabei werden Eigenschaften des Dezimalsystems genauso systematisch genutzt wie die Beziehungen zwischen Operationen. Diese Kenntnisse werden auch bei der Untersuchung von Zahlenfolgen angewendet, in denen beispielsweise Fehler erkannt oder Vorgehensweisen begründet werden müssen. Rechenverfahren werden flexibel kombiniert, Ergebnisse werden angemessen geschätzt beziehungsweise gerundet. Lösungsstrategien wie das systematische Probieren werden auch bei komplexeren Rechenwegen und Gleichungen genutzt. Rechenregeln sind explizit bekannt und können sinnvoll angewendet werden. Komplexe Sachsituationen werden im geeigneter Weise modelliert, rechnerisch korrekt bearbeitet und ihre Lösungen werden angemessen dargestellt. Das begriffliche Wissen umfasst auch speziellere Fachbegriffe und kann sicher verwendet und kommuniziert werden.

Kompetenzstufe V (Punktwerte ab 600): Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien

Schwierigere mathematische Aufgaben können auf unterschiedlichen Wegen korrekt gelöst werden. Beziehungen zwischen Zahlen werden situationsangemessen erkannt. Mathematische Eigenschaften, wie etwa die Zerlegbarkeit natürlicher Zahlen in Faktoren, werden für Problemlösungen genutzt. Auf dieser Grundlage werden auch schwierige Gleichungen bearbeitet und beispielsweise durch systematisches Probieren gelöst. Besondere Aspekte wie das Rechnen mit Bruchzahlen oder das Rechnen mit Zahlen in Kommaschreibweise stellen keine Hürde dar. Unterschiedliche Lösungswege können nachvollzogen und beschrieben werden.

3.3.3 Raum und Form

Kompetenzstufe I (Punktwerte unter 390): technische Grundlagen (Routineprozeduren auf Grundlage einfachen begrifflichen Wissens)

Es werden in allen Bereichen Grundkenntnisse nachgewiesen, die reproduktiv eingesetzt werden können. Das begriffliche Wissen beschränkt sich auf einfache Formen der ebenen Geometrie mit klar trennbaren Eigenschaften und ist nur in einem eingeschränkten Kontext anwendbar. Spiegelbilder einfacher, geradlinig begrenzter Figuren werden korrekt gezeichnet. Mit geometrischen Formen kann auf Grundlage bildlicher Darstellungen beziehungsweise des Gitternetzes gearbeitet werden. Eine räumliche Orientierung ist in sehr einfachen Situationen möglich.

Kompetenzstufe II (Punktwerte von 390 bis 459): einfache Anwendungen von Grundlagenwissen (Routineprozeduren in einem klar strukturierten Kontext)

Einfache Grundbegriffe der ebenen und räumlichen Geometrie werden korrekt verwendet, wobei die räumlichen Begriffe in der Regel im Alltag verankert sind (z.B. Würfel). Einfache räumliche Gebilde werden in ihren strukturellen Eigenschaften gesehen und räumliche Beziehungen können zur Lösung einfacher Probleme genutzt werden. Dabei werden einfache Lagebeziehungen auch dann korrekt gesehen, wenn hierfür die Perspektive einer anderen Person einzunehmen ist. Komplexere geometrische Muster werden in ihrer Struktur erkannt und fortgesetzt.

Kompetenzstufe III (Punktwerte von 460 bis 529): Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext

Durch das Curriculum vorgegebene Begriffe der ebenen und räumlichen Geometrie werden korrekt verwendet. Räumliche Beziehungen werden zur Lösung komplexerer Aufgaben genutzt, wobei erweiterte Problemlösefähigkeiten gezeigt werden. Achsenspiegelungen werden in gut strukturierten Umgebungen (Gitternetz, Geobrett) ausgeführt. Vertraute ebene Figuren (z.B. Quadrat, Rechteck, Kreis) können im Hinblick auf ihre Symmetrieeigenschaften untersucht werden. Bei Körpern wird mit verschiedenen Darstellungen (z.B. Netz) sinnvoll gearbeitet, wobei Zusammenhänge und räumliche Orientierung genutzt werden.

Kompetenzstufe IV (Punktwerte von 530 bis 599): sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang

Begriffe der ebenen und der räumlichen Geometrie werden flexibel verwendet, wobei unterschiedliche Darstellungen sicher in ihren Zusammenhängen gesehen werden. Bedingungen etwa zur Lage und Größe von Quadraten werden bei ihrer Konstruktion berücksichtigt. Beispiele von Würfel- und Quadernetzen werden korrekt erkannt, mögliche Fehler können korrigiert werden. Maßstabsgerechtes Zeichnen wird bei gegebenen Teilfiguren geleistet. Es ist eine Vertrautheit mit grundlegenden mentalen Operationen im Raum gegeben. Begründungen für geometrische Zusammenhänge werden angemessen beurteilt. Insbesondere werden auch Aufgaben, die unterschiedliche allgemeine mathematische Kompetenzen erfordern, korrekt gelöst.

Kompetenzstufe V (Punktwerte ab 600): Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien

Umfangreiches curricular verankertes Wissen wird auch in ungewohnten Situationen flexibel genutzt. Begriffe und Methoden der ebenen und räumlichen Geometrie können in komplexen Situationen angewendet werden. Geometrische Zusammenhänge können korrekt begründet werden, Verallgemeinerungen werden geleistet. Es wird ein hohes Maß an räumlichem Denken, aber auch eine gute analytische Durchdringung von Sachverhalten gezeigt. Insbesondere wird räumliches Vorstellungsvermögen zur Lösung komplexer Probleme genutzt. Aufgabenstellungen und ihre Lösungen können kritisch reflektiert werden. Auch ohne Vorgaben kann maßstabsgerecht gezeichnet werden.

3.3.4 Muster und Strukturen

Kompetenzstufe I (Punktwerte unter 390): technische Grundlagen (Routineprozeduren auf Grundlage einfachen begrifflichen Wissens)

Es werden sehr einfache Gesetzmäßigkeiten in Mustern erkannt, die auf vertrauten numerischen Zusammenhängen („leichte“ Einmaleinsreihen, insbesondere Verdoppeln) beruhen. Sehr einfache geometrische Muster werden erkannt und fortgesetzt.

Kompetenzstufe II (Punktwerte von 390 bis 459): einfache Anwendungen von Grundlagenwissen (Routineprozeduren in einem klar strukturierten Kontext)

Einfache Zahlenfolgen werden nach einer gegebenen Vorschrift fortgesetzt, fehlerhafte Einträge werden in solchen Zahlenfolgen erkannt. Bei einfachen, bildlich oder numerisch gegebenen Folgen wird die Grundstruktur des Musters erkannt (z. B. fortgesetzte Addition einer kleinen Zahl). Zahlen im Tausenderraum werden in Stellenwerttafeln sicher dargestellt und nach Vorschrift variiert. Einfache proportionale Zuordnungen (wie etwa das Verdoppeln) werden erkannt und angewendet.

Kompetenzstufe III (Punktwerte von 460 bis 529): Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext

Gesetzmäßigkeiten werden auch bei komplexeren Mustern erkannt und Muster können entsprechend fortgesetzt werden. Dabei werden analytische Fähigkeiten genutzt, sodass die Bestimmung eines beliebigen Elements an einer vorgegebenen Stelle auch ohne eigene Anschauung oder konkretes Manipulieren möglich ist. Gesetzmäßigkeiten in Zahlenfolgen werden bei Zahlenmaterial im curricularen Umfang erkannt und begründet beziehungsweise kommuniziert. Zahlen werden in einer Stellentafel gezielt manipuliert und das Ergebnis wird sinnvoll interpretiert. Proportionale Zuordnungen werden erkannt und interpretiert. In einfachen Sachsituationen werden funktionale Beziehungen erkannt und interpretiert und es werden insbesondere proportionale Zuordnungen zur Modellierung und Lösung von Problemen genutzt.

Kompetenzstufe IV (Punktwerte von 530 bis 599): sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang

Komplexe Muster können analysiert und zeichnerisch oder numerisch fortgesetzt werden. Dabei werden insbesondere Gesetzmäßigkeiten auch dann erkannt, wenn Zahlenfolgen nicht durch die Addition einer bestimmten Zahl oder die Multiplikation mit einem bestimmten Faktor erzeugt wurden, beziehungsweise wenn die Glieder nicht in Form von Zahlen, sondern in Form von Rechentermen gegeben sind. Zusammenhänge zwischen verschiedenen Darstellungen (z. B. bildliche beziehungsweise numerische Darstellungen) werden auch bei schwierigeren Folgen oder Mustern gesehen. Zahldarstellungen in Stellenwerttafeln können auch bei sehr großen Zahlen nach Vorschrift selbstständig und systematisch verändert werden. Proportionale Zuordnungen werden zur Modellierung und Lösung von Sachaufgaben genutzt.

Kompetenzstufe V (Punktwerte ab 600): Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien

Der Umgang auch mit schwierigen Zahlenfolgen (z. B. unter Verwendung von Quadratzahlen oder mehreren Rechenarten) ist sicher. Gesetzmäßigkeiten werden auch erkannt, wenn verschiedene Operationen zu verbinden sind. Diese können auch begründet werden. Arithmetische Muster werden nach vorgegebenen Kriterien konstruiert, wobei eigenständig Lösungsstrategien entwickelt werden. Proportionale Zuordnungen können im Zusammenhang mit einfachen Bruchzahlen und Dezimalbrüchen angewendet und in Sachsituationen auch unter Verwendung von tabellarischen Darstellungen interpretiert werden. Selbst ungewohnte funktionale Zusammenhänge werden in Sachsituationen modelliert, analysiert und genutzt.

3.3.5 Größen und Messen**Kompetenzstufe I (Punktwerte unter 390): technische Grundlagen (Routineprozeduren auf Grundlage einfachen begrifflichen Wissens)**

Gängige standardisierte Einheiten für Geldwerte, Zeitspannen, Längen, Gewichte und Rauminhalte werden richtig verwendet und können gut vertrauten Repräsentanten zugeordnet werden. Größenvergleiche werden bei übereinstimmender Einheit beziehungsweise in Bezug auf Längen und Gewichte auch bei benachbarten Einheiten durchgeführt. Bei Geldwerten werden Eurobeträge in Cent umgewandelt, wobei in einfachen Fällen auch die Kommaschreibweise verwendet wird. Bei Längen und Gewichten kann in eine benachbarte Einheit ohne Verwendung der Kommaschreibweise umgewandelt werden.

Kompetenzstufe II (Punktwerte von 390 bis 459): einfache Anwendungen von Grundlagenwissen (Routineprozeduren in einem klar strukturierten Kontext)

Das Umwandeln von Größen in gegebene Einheiten wird auch bei gemischten Größenangaben durchgeführt, wenn das Zahlenmaterial einfach ist. Ein Vergleich von Größen ist auch bei Verwendung unterschiedlicher Einheiten möglich. Bei aus dem Alltag vertrauten Größen wird die Kommaschreibweise in einfachen Situationen richtig verwendet (Längen, Geldwerte). Größen können in einfachen Sachsituationen angewendet werden, wenn die dabei benutzten Einheiten einen deutlichen Bezug zum Alltag der Kinder haben.

Kompetenzstufe III (Punktwerte von 460 bis 529): Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext

Größen werden in jeweils benachbarte Einheiten umgewandelt, wobei gemischte Größen, aber auch komplexeres Zahlenmaterial (z.B. große Zahlen, einfache Brüche, Zahlen in Kommaschreibweise auch mit einer Null vor dem Komma) verwendet werden. Es werden für die verschiedenen Einheiten jeweils geeignete Repräsentanten verwendet, beziehungsweise Repräsentanten werden in Sachsituationen sinnvoll zugeordnet. Der Umgang mit Größen ist flexibel und berücksichtigt – etwa bei Uhrzeiten – spezielle Eigenschaften der Größenbereiche. Allen im Unterricht behandelten Größen können geeignete Repräsentanten zugeordnet werden.

Kompetenzstufe IV (Punktwerte von 530 bis 599): sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang

Das Rechnen mit Größen ist sicher und flexibel und umfasst insbesondere Näherungsrechnungen und Überschlagsrechnungen (z.B. mit Geldwerten), aber auch das angemessene Schätzen sowie den Umgang mit Hohlmaßen. Die Kenntnisse können auch in komplexen Sachsituationen angewendet werden, wobei sowohl das Entwickeln eines geeigneten Modells für eine Lösung als auch das rechnerische Lösen geleistet werden. Dabei können auch mehrere Größen auftreten und Zusammenhänge zwischen ihnen genutzt werden.

Kompetenzstufe V (Punktwerte ab 600): Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien

Es werden schwierige Umrechnungen insbesondere im Zusammenhang mit Zeitspannen und Gewichten geleistet, bei denen auch gemischte Größen und nicht benachbarte Einheiten auftreten. Das Zahlenmaterial umfasst den ganzen curricularen Umfang, wobei gleichzeitig sehr große und sehr kleine Zahlen auftreten, die Kommaschreibweise verwendet wird oder Brüche benutzt werden. Komplexe Sachsituationen, die nicht unbedingt dem eigenen Erfahrungsbereich entnommen sind, werden sicher modelliert und auch bei rechnerisch hohen Anforderungen gelöst. Dabei werden innerhalb einer Aufgabe Bezüge zwischen verschiedenen Größen hergestellt sowie mehrere Grundrechenarten und schwieriges Zahlenmaterial verwendet. Zusammenhänge zwischen Größen werden erkannt, interpretiert und genutzt. Insbesondere können Größen im Zusammenhang mit proportionalen Zuordnungen verwendet werden. Es werden logische Beziehungen auf der Grundlage eines umfassenden Verständnisses gegebener Größen analysiert. Informationen zu Größen werden dem jeweiligen Kontext angemessen dargestellt.

3.3.6 Daten, Häufigkeit und Wahrscheinlichkeit

Kompetenzstufe I (Punktwerte unter 390): technische Grundlagen (Routineprozeduren auf Grundlage einfachen begrifflichen Wissens)

Es wird mit einfachen Daten sinnvoll umgegangen, wobei insbesondere unmittelbar ersichtliche Informationen aus einfachen, klar strukturierten Diagrammen, Schaubildern und Tabellen entnommen werden. Einfache Daten können darüber hinaus in gegebenen Diagrammen, Schaubildern und Tabellen ergänzt

werden. Dabei handelt es sich um übersichtliche Daten, die keine besondere Strukturierung erfordern (etwa relativ kleine natürliche Zahlen). Sehr einfache und sehr anschauliche Zufallsexperimente können im Hinblick etwa auf Gewinnchancen verglichen werden („beim ersten Glücksrad hat man die besten Gewinnchancen“).

Kompetenzstufe II (Punktwerte von 390 bis 459): einfache Anwendungen von Grundlagenwissen (Routineprozeduren in einem klar strukturierten Kontext)

Der Umgang mit Daten umfasst Zahlen im curricular gegebenen Umfang. Diagramme, Schaubilder und Tabellen können dabei eine erweiterte Struktur haben, also beispielsweise mehrere Spalten umfassen. Ihnen werden Informationen entnommen, die in der Regel noch explizit gegeben sein müssen. Es werden wesentliche Grundbegriffe aus dem Umfeld von Zufall und Wahrscheinlichkeit wie „sicher“, „unmöglich“ oder „wahrscheinlich“ korrekt verwendet. Gewinnchancen können bei einfachen und alltagsnahen Zufallsexperimenten eingeschätzt werden, wobei jedoch keine explizite Begründung geleistet wird (z. B. bei gegebener Häufigkeitsverteilung die wahrscheinlichste Farbe für das Ziehen aus „Urnen“ benennen).

Kompetenzstufe III (Punktwerte von 460 bis 529): Erkennen und Nutzen von Zusammenhängen in einem vertrauten (mathematischen und sachbezogenen) Kontext

Daten und Informationen können auf verschiedene Arten dargestellt und entsprechenden (auch in Textform gegebenen) Darstellungen entnommen werden, wenn der Kontext der Lebenswelt und den Alltagserfahrungen der Kinder entstammt. Es werden Zusammenhänge zwischen den Darstellungsformen genutzt. Bei nicht allzu komplexen Zufallsexperimenten werden die relativen Gewinnchancen korrekt angegeben und die Einschätzung wird sinnvoll begründet. Einfache Zufallsexperimente mit gegebener Gewinnwahrscheinlichkeit werden korrekt modelliert.

Kompetenzstufe IV (Punktwerte von 530 bis 599): sicheres und flexibles Anwenden von begrifflichem Wissen und Prozeduren im curricularen Umfang

Die Nutzung von Informationen aus Diagrammen, Schaubildern und Tabellen erfolgt völlig flexibel und ist auch bei wenig strukturiertem Informationsmaterial möglich. Informationen aus unterschiedlichen Quellen und gegebenenfalls auch aus unterschiedlichen Darstellungsformen können in einen Zusammenhang gestellt werden. Die entnommenen Daten werden in Modellierungsaufgaben eigenständig verwendet. Mögliche Ausgänge eines zweistufigen Zufallsexperiments werden beschrieben. Auch schwierigere Zufallsexperimente werden korrekt beurteilt.

Kompetenzstufe V (Punktwerte ab 600): Modellierung komplexer Probleme unter selbstständiger Entwicklung geeigneter Strategien

Der Umgang mit Daten berücksichtigt auch Detailinformationen in dem jeweiligen Kontext in angemessener Weise und umfasst die selbstständige Entnahme sowie Aufbereitung der Daten. Diese können damit in unterschiedlichen Problemaufgaben angemessen verwendet werden. Insbesondere werden entsprechende Modellierungsaufgaben und anspruchsvolle Probleme sicher gelöst.

Komplexe Zufallsexperimente werden angemessen in Bezug auf die Wahrscheinlichkeit bestimmter Ereignisse beurteilt und können etwa in Bezug auf Gewinnchancen verglichen werden.

Literatur

- AERA, APA, NCME (1999) = American Educational Research Association (AERA), American Psychological Association (APA) & National Council on Measurement in Education (NCME). *Standards for Educational and Psychological Testing*. Washington, DC: APA.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Böhme, K. (2012). *Methodische und didaktische Überlegungen sowie empirische Befunde zur Erfassung sprachlicher Kompetenzen im Deutschen (Dissertation)*. Humboldt-Universität zu Berlin. Vorgesehen für eine elektronische Publikation unter <http://edoc.hu-berlin.de>
- Böhme, K. & Köller, O. (2010). Kompetenzstufenmodelle für den Mittleren Schulabschluss: Einleitung. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 33–35). Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R. & Walther, G. (2003). *Erste Ergebnisse aus IGLU: Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Camilli, G., Cizek, G. J. & Lugg, C. A. (2001). Psychometric theory and the validation of performance standards: History and future perspectives. In G. J. Cizek (Hrsg.), *Setting performance standards: Concepts, methods, and perspectives* (S. 445–475). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Cizek, G. J. (2001). Conjectures on the rise and fall of standard setting: An introduction to context and practice. In G. J. Cizek (Hrsg.), *Setting performance standards: Concepts, methods, and perspectives* (S. 3–17). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Cizek, G. J. & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications Ltd.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 375–397). Münster: Waxmann.
- Hartig, J. & Frey, A. (2012). Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten. *Psychologische Rundschau*, 63, 43–49.
- Hartig, J., Frey, A., Nold, G. & Klieme, E. (2011). An application of explanatory item response modeling for model-based proficiency scaling. *Educational and Psychological Measurement*, vor dem Druck online veröffentlicht am 29. Dezember 2011.
- IQB (2008) = Institut zur Qualitätsentwicklung im Bildungswesen. (Hrsg.). *Kompetenzstufenmodell zu den Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4)*. Zugriff am 12.07.2012 unter http://www.iqb.hu-berlin.de/bista/dateien/Mathe_primar.pdf
- Kaftandjewa, F. (2010). *Methods for setting cut scores in criterion-referenced achievement tests*. Arnheim: Cito.
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38, 319–342.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (Hrsg.). (2007). *Zur Entwick-*

lung nationaler Bildungsstandards. *Expertise*. Bonn: Bundesministerium für Bildung und Forschung.

- Klieme, E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52, 876–903.
- KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2005b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- Köller, O. (2010). Politische und inhaltliche Rahmenbedingungen bei der Setzung von Kompetenzstufen. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 35–37). Münster: Waxmann.
- Köller, O., Knigge, M. & Tesch, B. (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Mitzel, H. C., Lewis, D. M., Patz, R. J. & Green, D. R. (2001). The bookmark procedure: Psychological perspectives. In G. J. Cizek (Hrsg.), *Setting performance standards: Concepts, methods, and perspectives* (S. 249–281). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Morgan, D. L. & Perie, M. (2004). *Setting standards in education: Choosing the best method for your assessment and population*. ETS. Princeton. [Unpublished Paper].
- Pant, H. A., Rupp, A. A., Tiffin-Richards, S. P. & Köller, O. (2009). Validity issues in standard-setting studies. *Studies in Educational Evaluation*, 35, 95–101.
- Pant, H. A., Tiffin-Richards, S. P. & Köller, O. (2010). Standard-Setting für Kompetenztests im Large-Scale-Assessment. Projekt Standardsetting. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*. (S. 175–188). Weinheim: Beltz.
- Pietsch, M., Böhme, K., Robitzsch, A. & Stubbe, T. C. (2009). Das Stufenmodell zur Lesekompetenz der länderübergreifenden Bildungsstandards im Vergleich zu IGLU 2006. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik*. (S. 393–416). Weinheim: Beltz.
- Reiss, K., Heinze, A. & Pekrun, R. (2007). Mathematische Kompetenz und ihre Entwicklung in der Grundschule. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 107–127). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. BremerichVos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 120–141). Weinheim: Beltz.

Kapitel 4

Anlage und Durchführung des Ländervergleichs

Dirk Richter, Maria Engelbert, Katrin Böhme, Nicole Haag, Jasmin Hannighofer, Heino Reimers, Alexander Roppelt, Sebastian Weirich, Hans Anand Pant und Petra Stanat

Der Ländervergleich in der Primarstufe 2011 zielt darauf ab, das Erreichen der in den Bildungsstandards der Kultusministerkonferenz (KMK, 2005) beschriebenen Kompetenzen in den Fächern Deutsch und Mathematik am Ende der vierten Jahrgangsstufe zu überprüfen. Weiterhin sollen die Kompetenzstände der Schülerinnen und Schüler ländervergleichend untersucht werden. Zu diesem Zweck nahmen insgesamt 27 081 Schülerinnen und Schüler aus den 16 Ländern in der Bundesrepublik Deutschland an der Untersuchung teil.

Zeitgleich zur Erhebung des bundesweiten Ländervergleichs 2011 fanden die *Internationale Grundschul-Lese-Untersuchung* (PIRLS/IGLU¹) 2011 und die *Trends in International Mathematics and Science Study* (TIMSS²) 2011 statt, die vom Institut für Schulentwicklungsforschung (IfS) der Technischen Universität Dortmund verantwortet werden. Um Aussagen zur Äquivalenz der Ergebnisse der verschiedenen Studien zu ermöglichen, wurde eine sogenannte Validierungsuntersuchung durchgeführt, in der sowohl die Tests des Ländervergleichs 2011 als auch die Tests der beiden internationalen Studien zum Einsatz kamen. Die Ergebnisse dieser Analysen werden zu einem späteren Zeitpunkt veröffentlicht.

Das vorliegende Kapitel beschreibt die Vorbereitung und Durchführung des Ländervergleichs 2011. Dies umfasst die Testentwicklung, die eingesetzten Fragebogeninstrumente, das Testdesign, das Verfahren der Stichprobenziehung und die resultierende Stichprobe sowie die Maßnahmen zur Qualitätssicherung während der Datenerhebung. Weiterhin werden die wichtigsten Schritte der Datenanalyse und Datenauswertung skizziert sowie die an der Studie beteiligten Institutionen und Personen vorgestellt. Genauere Beschreibungen der technischen Details sind in Kapitel 12 zu finden.

4.1 Testinstrumente

Im Ländervergleich 2011 wurden auf der Grundlage der Bildungsstandards entwickelte Instrumente zur Erfassung einer Reihe von Kompetenzen in den Fächern Deutsch und Mathematik eingesetzt. Im Fach Deutsch wurden Kompetenzen in den Bereichen Lesen, Zuhören und Orthografie erfasst. Die Testung von Ortho-

-
- 1 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.
 - 2 Das Akronym TIMSS stand ursprünglich für *Third International Mathematics and Science Study*. Seit 2003 wird es in der Bedeutung *Trends in International Mathematics and Science Study* verwendet.

grafie diene jedoch ausschließlich der Überprüfung des Kompetenzstufenmodells für diesen Kompetenzbereich und erfolgte daher nur in einer für Deutschland repräsentativen Stichprobe. Innerhalb der Länder sind die Stichproben zu klein, um anhand der Daten einen Ländervergleich durchzuführen. Im Fach Mathematik wurden alle fünf inhaltsbezogenen mathematischen Kompetenzen geprüft: (1) Zahlen und Operationen, (2) Raum und Form, (3) Muster und Strukturen, (4) Größen und Messen sowie (5) Daten, Häufigkeit und Wahrscheinlichkeit.

Die eingesetzten Aufgaben entstammen dem Itempool des IQB, der von erfahrenen Lehrkräften entwickelt und anschließend in großen nationalen Erhebungen pilotiert und normiert worden ist (vgl. Kapitel 1 in diesem Band).

Die Leseaufgaben im Fach Deutsch umfassen jeweils einen Stimulustext von einer bis anderthalb Seiten, zu dem mehrere unterschiedlich komplexe Teilaufgaben (*Items*) gestellt werden. Es werden dabei sowohl kontinuierliche literarische Texte als auch kontinuierliche und diskontinuierliche Sachtexte verwendet. Bei den Zuhöraufgaben besteht der Stimulus aus einem Hörtext mit einer Dauer von zwei bis neun Minuten, zu dem ebenfalls jeweils mehrere Items mit unterschiedlichen Antwortformaten bearbeitet werden. Bei den Hörtexten handelt es sich um Hörspiele oder Radiosendungen für Kinder. Zur Erfassung der orthografischen Kompetenz wird eine Reihe von Aufgaben mit unterschiedlichen Formaten eingesetzt. Dazu gehören Lückensatzdiktate, Aufgaben zur Korrektur von Falschschreibungen sowie Aufgaben, bei denen die Kenntnis und Anwendung von orthografischen und morphematischen Regeln gezeigt werden soll. Weitere Informationen zu den Testinstrumenten für die drei getesteten Kompetenzbereiche im Fach Deutsch finden sich im Kapitel 2.1.

Die Aufgaben im Fach Mathematik bestehen jeweils aus einem kurzen Stimulustext (ca. 1–5 Zeilen), zu dem bis zu sechs Fragen (*Items*) gestellt werden. Es werden je nach inhaltlichem Kompetenzbereich verschiedene Arten von Aufgaben verwendet: Neben einfachen Berechnungen müssen die Schülerinnen und Schüler bei Aufgaben zum Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit* Informationen aus Diagrammen entnehmen und interpretieren oder die Wahrscheinlichkeit von Ereignissen einschätzen. Bei Aufgaben zum Kompetenzbereich *Größen und Messen* schätzen die Kinder Größen, wandeln sie um und ordnen sie Einheiten zu. Im Kompetenzbereich *Raum und Form* werden die Schülerinnen und Schüler aufgefordert, Figuren zu spiegeln oder Spiegelachsen einzuzeichnen sowie die Eigenschaften von Figuren oder Körpern zu erkennen und zu beschreiben. Aufgaben zum Kompetenzbereich *Muster und Strukturen* schließlich erfordern von den Kindern, Muster in Rechenkettens oder Figuren zu erkennen und fortzusetzen. Eine detaillierte Beschreibung der erfassten Kompetenzen im Fach Mathematik ist in Kapitel 2.2 zu finden.

In den Tests zu den Fächern Deutsch und Mathematik kommen sowohl geschlossene als auch offene und halboffene Aufgabenformate zur Anwendung. Zu den geschlossenen Aufgabenformaten gehören Multiple-Choice-Aufgaben und Richtig-Falsch-Aufgaben, bei denen jeweils aus verschiedenen Antwortalternativen die richtige angekreuzt werden muss, sowie Zuordnungsaufgaben, bei denen zusammengehörige Paare (z. B. Wortpaare) gebildet werden sollen. Bei halboffenen Aufgaben ergänzen die Schülerinnen und Schüler einzelne Wörter (z. B. in Lückensatzdiktaten) oder Wortgruppen, während sie bei offenen Aufgaben eigenständig kurze Texte im Umfang von mehreren Zeilen produzieren.

Die im Ländervergleich eingesetzten Aufgaben sollen das gesamte Leistungsspektrum von Schülerinnen und Schülern der vierten Jahrgangsstufe abdecken. Zur Erfassung von Kompetenzen im unteren Leistungsbereich wurden Aufgaben

eingesetzt, für deren Lösung Kompetenzen erforderlich sind, die bereits bis zum Ende der dritten Klasse erworben sein sollten. Die Kompetenzmessung im oberen Leistungsbereich erfolgte mit Aufgaben, deren Anforderungen am Ende der Primarstufe noch sehr herausfordernd sind. Mit diesem Vorgehen soll sichergestellt werden, dass die Kompetenzen der Schülerinnen und Schüler auch am oberen und unteren Ende des Leistungsspektrums ausreichend genau erfasst werden.

In Ergänzung zu den fachspezifischen Tests bearbeiteten die Kinder einen Test zur Erfassung kognitiver Grundfähigkeiten. Dieser besteht aus kurzen Subtests des Kognitiven Fähigkeitstests (KFT 4–12+R, Heller & Perleth, 2000) und des Salzburger Lese-Screenings (SLS 1–4, Mayringer & Wimmer, 2003). Aus dem Kognitiven Fähigkeitstest wurden die Subtests „Wortschatz“ und „Figurenanalogien“ eingesetzt, die jeweils 10 Minuten in Anspruch nehmen. Im Wortschatztest sollten die Kinder aus einer Reihe von Wörtern dasjenige identifizieren, welches am besten zu einem vorgegebenen Wort als Oberbegriff oder Synonym passt. Im Subtest „Figurenanalogien“ wurden die Kinder dazu aufgefordert, zu einem vorgegebenen Figurenpaar ein Äquivalent aus einer Reihe von Figuren zu wählen. Das Salzburger Lesescreening misst die Lesegeschwindigkeit mit Aufgaben, in denen eine Reihe von einfachen Sätzen als wahr oder falsch zu kennzeichnen ist. Die Bearbeitungszeit des Salzburger Lesescreenings umfasst insgesamt 3 Minuten. Die Ergebnisse des Kognitiven Fähigkeitstests und des Salzburger Lesescreenings dienen lediglich als Indikatoren verbaler und räumlicher kognitiver Grundfähigkeiten. Sie werden zur möglichst genauen statistischen Schätzung der fachbezogenen Kompetenzen benötigt.

4.2 Fragebogeninstrumente

In Ergänzung zu den Tests wurden die Schülerinnen und Schüler, Eltern, Lehrkräfte sowie Schulleiterinnen und Schulleiter jeweils gebeten, einen Fragebogen zu bearbeiten. Diese Instrumente dienen unter anderem der Erfassung individueller Hintergrundmerkmale von Schülerinnen und Schülern, um Disparitäten in den erworbenen Kompetenzen zu untersuchen, die mit dem Geschlecht, dem Zuwanderungshintergrund und der sozialen Herkunft verbunden sind. Weiterhin wurden in den Fragebögen verschiedene Merkmale des Unterrichts, der Schule und des familiären Umfelds erfasst, um Zusammenhänge zwischen den zur Verfügung stehenden Lerngelegenheiten und den erreichten Kompetenzen analysieren zu können. Die Inhalte der Fragebögen wurden im Vorfeld der Studie von den Kultusministerien der Länder insbesondere in Bezug auf die Einhaltung der datenschutzrechtlichen Bestimmungen geprüft. Im Folgenden wird beschrieben, wie die einzelnen Fragebogeninstrumente administriert wurden und welche Inhalte sie jeweils umfassten.

Schülerfragebogen

Der Schülerfragebogen wurde von den Schülerinnen und Schülern am ersten Testtag direkt im Anschluss an die Testung bearbeitet. Aufgrund von Unterschieden in den rechtlichen Regelungen war die Teilnahme in den Ländern Berlin, Brandenburg, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt und Thüringen für die Schülerinnen und Schüler verpflichtend, während in allen anderen Ländern die Eltern der Fragebogenbearbeitung ihrer Kinder zustimmen mussten. Mit dem Schülerfragebogen wurden Aspekte der Einwanderungsgeschichte der Familie, des Sprachgebrauchs und des Lese-

verhaltens der Kinder, ihr Selbstkonzept in den Fächern Deutsch und Mathematik sowie motivationale Merkmale in Bezug auf das Lernverhalten erfasst.

Elternfragebogen

Der Fragebogen für die Eltern wurde den Schülerinnen und Schülern am ersten Testtag mit nach Hause gegeben. Dieser enthielt Fragen zu Merkmalen der häuslichen Lernumgebung, des sozialen Hintergrunds und der Einwanderungsgeschichte. Hierzu zählen insbesondere Angaben zum Geburtsland der Familienmitglieder und zum Sprachgebrauch in der Familie, zur schulischen und beruflichen Ausbildung der Eltern und ihrer beruflichen Tätigkeit, zu den Bildungsaspirationen der Eltern, zu Merkmalen der vorschulischen Betreuung des Kindes und zur Nutzung von Förderangeboten innerhalb und außerhalb der Schule.

Lehrerfragebogen

An der Lehrerbefragung nahmen jeweils die in der Klasse unterrichtenden Lehrkräfte in den Fächern Deutsch und Mathematik teil. Aufgrund variierender Rechtsgrundlagen unterschieden sich die Regelungen für die Teilnahme an dieser Befragung in den Ländern. In Baden-Württemberg, Bayern, Hamburg und Sachsen war die Teilnahme vollständig freiwillig; in Bremen und dem Saarland waren berufs- und schulbezogene Fragen (z. B. zum Studium) verpflichtend, die anderen Fragen dagegen freiwillig zu beantworten. In allen anderen Ländern war die Beteiligung an der Fragebogenerhebung für Lehrkräfte verpflichtend.

Die Inhalte des Lehrerfragebogens bezogen sich auf die berufliche Qualifikation der Befragten, ihre Nutzung von Fort- und Weiterbildungsangeboten sowie Merkmale des Unterrichts in der am Ländervergleich 2011 teilnehmenden Klasse. Im Einzelnen wurden Fragen zum Beschäftigungsverhältnis und dem Beschäftigungsumfang der Lehrkräfte, zur Unterrichtstätigkeit in den Fächern Deutsch und Mathematik, zur Lehreraus- und -fortbildung, zum Einsatz von Schulbüchern in Deutsch beziehungsweise Mathematik, zur Kompetenzorientierung im Unterricht, zum Erstleseunterricht sowie zu Maßnahmen der Sprachförderung gestellt. Weiterhin wurden Angaben zu lerntheoretischen Überzeugungen, zur Zusammenarbeit im Kollegium, zum Wohlbefinden im Beruf und zu Einstellungen gegenüber Schulleistungstests erfasst.

Schulleiterfragebogen

Die Schulleiterinnen und Schulleiter wurden im Rahmen des Ländervergleichs 2011 ebenfalls schriftlich befragt, um schulische Rahmenbedingungen und die in den Schulen vorhandenen Unterstützungsangebote zu ermitteln. Die länderspezifischen Regelungen für die Teilnahme an der Schulleiterbefragung waren für die meisten Länder dieselben wie für die Lehrerbefragung. Nur in Hamburg war die Schulleiterbefragung im Gegensatz zur Lehrkräftebefragung für die schul- und berufsbezogenen Angaben verpflichtend. Zu den mit dem Fragebogen erfassten Merkmalen zählen jahrgangsübergreifender Unterricht, Ganztagsbetrieb, Verfügbarkeit schulischer Förderangebote, insbesondere der Sprach- und Leseförderung, Bereitstellung von Fortbildungsangeboten und der Umgang mit schulischen Leistungstests und deren Ergebnissen.

4.3 Testdesign

Im Ländervergleich 2011 wurde eine große Zahl von Aufgaben eingesetzt, um alle Facetten der getesteten Kompetenzbereiche hinreichend zu repräsentieren. Da diese Zahl von Aufgaben in ihrer Gesamtheit nicht von allen Schülerinnen und Schülern bearbeitet werden kann, wurden sie auf verschiedene Testhefte verteilt (Multi-Matrix-Design). Die Zusammenstellung der Testhefte erfolgte in einem mehrstufigen Verfahren, bei dem zunächst die Aufgaben Aufgabenblöcken und anschließend die Aufgabenblöcke Testheften zugeordnet wurden. Die Blöcke enthielten jeweils Aufgaben zu nur einem Kompetenzbereich (z. B. Deutsch – Lesen, Mathematik – Größen und Messen), die wiederum zu Deutsch- beziehungsweise Mathematiktestheften zusammengestellt wurden. Der Umfang des Aufgabenmaterials wurde dabei so gewählt, dass die Aufgabenblöcke in Deutsch jeweils 20 Minuten und in Mathematik jeweils 10 Minuten Bearbeitungszeit erfordern. Die unterschiedlichen Blocklängen für Deutsch und Mathematik sind auf die längeren Verarbeitungszeiten der Lese- und Zuhörstimuli im Fach Deutsch zurückzuführen. Die Zahl der Aufgaben innerhalb der Blöcke war so bemessen, dass alle Kinder diese ohne Zeitdruck bearbeiten konnten.

Die Zusammenstellung der Aufgabenblöcke und die Verteilung dieser Blöcke auf verschiedene Testhefte wurden in einem sogenannten Testdesign geplant. Bei der Erstellung des Testdesigns wurde darauf geachtet, dass in jedem Heft mehrere Kompetenzbereiche überprüft werden. Im Fach Deutsch wurden mit allen Testheften die Kompetenzbereiche Lesen und Zuhören geprüft, nur ein Teil der Testhefte enthielt auch Blöcke mit Orthografieaufgaben. In Mathematik zielte die Verteilung der Blöcke darauf ab, mit jedem Heft drei bis fünf Kompetenzbereiche zu testen. Besonders wichtig bei einem Multi-Matrix-Design ist, die Blöcke so auf die Testhefte zu verteilen, dass alle Aufgaben eines Fachs über die verschiedenen Testhefte bei der späteren Auswertung miteinander verknüpft werden können. Auf diese Weise ist es möglich, die Leistungen der Schülerinnen und Schüler auf einer einheitlichen Skala abzubilden, auch wenn sie unterschiedliche Testhefte und Aufgaben bearbeitet haben.

Insgesamt wurden im Fach Deutsch 40 Testhefte für Regelschulen und vier Testhefte für Förderschulen, im Fach Mathematik 35 Testhefte für Regelschulen und drei Testhefte für Förderschulen entwickelt. Die Testzeit betrug an Regelschulen 80 Minuten und an Förderschulen 40 Minuten.³ Die in Förderschulen eingesetzten Testhefte enthielten zudem überwiegend leichtere Aufgaben.

Bei der Verteilung der Testhefte auf die Klassen wurde in den Fächern Deutsch und Mathematik ein unterschiedliches Vorgehen gewählt. Im Fach Deutsch konnten die Testhefte nicht beliebig den einzelnen Schülerinnen und Schülern zugewiesen werden, da die Hörtexte und die Lückensatzdiktate im Bereich Orthografie innerhalb einer Klasse jeweils zur gleichen Zeit bearbeitet werden mussten. Aus diesem Grund wurden in Deutsch jeweils zwei Testheftversionen pro Klasse eingesetzt, die identische Zuhöraufgaben und Lückensatzdiktate, aber unterschiedliche Leseaufgaben enthielten. Diese zwei Versionen wurden den einzelnen Schülerinnen und Schülern einer Klasse zufällig zugewiesen. In Mathematik konnten alle Testhefte zufällig auf die Schülerinnen und Schüler einer Klasse verteilt werden.

³ Im Land Bremen erhielten auch Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, die an Regelschulen inklusiv unterrichtet werden, die verkürzten Testhefte.

Tabelle 4.1 zeigt einen Ausschnitt aus dem Testdesign für das Fach Deutsch. Hier wird deutlich, dass jeweils zwei aufeinanderfolgende Testheftversionen zu einer Testheftgruppe zusammengefasst wurden. In jeder Klasse kam eine dieser Testheftgruppen zum Einsatz. Anhand der Tabelle lässt sich auch die Idee des Multi-Matrix-Designs erläutern. In Testheftgruppe 2 wurden zum Beispiel die Aufgabenblöcke D-Zuhören03 und D-Zuhören04 gemeinsam administriert, während in Testheftgruppe 3 die Aufgabenblöcke D-Zuhören03 und D-Zuhören05 zum Einsatz kamen. Durch die Kombination der Blöcke D-Zuhören04 sowie D-Zuhören05 mit dem in beiden Testheftgruppen vorhandenen Block D-Zuhören03 ist es möglich, auch die Aufgaben des Blocks D-Zuhören04 mit denen in D-Zuhören05 in Beziehung zu setzen, ohne dass dieselben Schülerinnen und Schüler diese Blöcke bearbeitet haben.

In der Tabelle 4.1 wird ebenfalls erkennbar, dass die Positionen der Aufgabenblöcke zwischen den Testheften variieren. So befindet sich beispielsweise der Block D-Lesen01 im Testheft TH01 an der zweiten Blockposition, im Testheft TH03 an der ersten Blockposition und im Testheft TH05 an der dritten Blockposition. Eine Variation der Position von Aufgabenblöcken ist in Large-Scale-Assessments wie dem Ländervergleich erforderlich, da sich im Verlauf der Testung solche Faktoren wie die Anstrengungsbereitschaft, Motivation oder Konzentrationsfähigkeit der Schülerinnen und Schüler ändern können. Unter anderem können im Verlauf der Testdurchführung Ermüdungseffekte auftreten, die sich vor allem auf die Bearbeitung von Aufgabenblöcken am Ende eines Testheftes auswirken können (Robitzsch, 2009). Um solche Positionseffekte in den statistischen Analysen kontrollieren zu können, wurden alle Aufgabenblöcke in den verschiedenen Testheften auf unterschiedlichen Blockpositionen platziert.

Tabelle 4.1: Ausschnitt des Testdesigns im Ländervergleich 2011

Testheftgruppe	Testheft-nr.	Block 1	Block 2	Block 3	Block 4
1	TH01	D-Zuhören01	D-Lesen01	D-Zuhören02	D-Lesen32
	TH02	D-Zuhören01	D-Lesen31	D-Zuhören02	D-Lesen03
2	TH03	D-Lesen01	D-Zuhören04	D-Lesen02	D-Zuhören03
	TH04	D-Lesen21	D-Zuhören04	D-Lesen42	D-Zuhören03
3	TH05	D-Zuhören03	D-Lesen03	D-Lesen01	D-Zuhören05
	TH06	D-Zuhören03	D-Lesen11	D-Lesen21	D-Zuhören05

Anmerkungen. TH = Testheft. Die Bezeichnung der Blöcke setzt sich zusammen aus einer Abkürzung für das zu testende Fach (D = Deutsch), der Bezeichnung der operationalisierten Kompetenz und einer fortlaufenden Nummer für den Aufgabenblock zu dieser Kompetenz in diesem Fach.

4.4 Testablauf

Die Erhebungen für den Ländervergleich 2011 fanden im Zeitraum vom 23. Mai bis zum 22. Juli 2011 unter hoch standardisierten Bedingungen statt. Die Untersuchung in den Schulen erfolgte jeweils an zwei Testtagen, die in der Regel direkt aufeinander folgten. An einem der beiden Testtage bearbeiteten die Schülerinnen und Schüler ausschließlich Aufgaben im Fach Deutsch, am anderen Testtag ausschließlich Aufgaben im Fach Mathematik. Die Zuordnung der Fächer zu den Testtagen erfolgte per Zufall. Der Ablauf der Erhebungen an Regelschulen ist in Tabelle 4.2 dargestellt.

Tabelle 4.2: Ablauf des ersten und zweiten Testtages an Regelschulen

Dauer in Minuten	Tätigkeit
10	Beginn der Testsitzung, Verteilung der Testhefte, Instruktion der Schülerinnen und Schüler
40	Bearbeitung der Testaufgaben (Deutsch oder Mathematik)
10	Pause
40	Bearbeitung der Testaufgaben (Deutsch oder Mathematik)
10	Pause
5	Instruktion zur Bearbeitung des Schülerfragebogens (1. Testtag) beziehungsweise zum Test der kognitiven Grundfähigkeiten (2. Testtag)
20	Bearbeitung des Schülerfragebogens (1. Testtag) beziehungsweise des Tests der kognitiven Grundfähigkeiten (2. Testtag)
5	Abschluss der Testsitzung, Einsammeln der Materialien

Die Testungen führten schulexterne Testleiterinnen und Testleiter durch, die im Vorfeld vom *IEA Data Processing and Research Center* (DPC) gründlich geschult und mit den Testmaterialien vertraut gemacht wurden. Um einen standardisierten und über alle Testungen hinweg vergleichbaren Ablauf sicherzustellen, folgten die Testleiterinnen und Testleiter detaillierten Skripts und lasen die Instruktionen wörtlich vor. Anhand des Skripts erklärten sie den Schülerinnen und Schülern zunächst, wie die verschiedenen Aufgabenformate zu bearbeiten sind, und demonstrierten dies anhand von Beispielen. Die Testleiterinnen und Testleiter überwachten zudem die Bearbeitungszeit der Testaufgaben und des Schülerfragebogens beziehungsweise des Tests der kognitiven Grundfähigkeiten. Sie dokumentierten die Bearbeitungszeiten, etwaige Störungen, andere Vorkommnisse in einem Testsitzungsprotokoll.

Zur Sicherung einer hohen Teilnahmequote wurde, wie auch in internationalen Schulleistungsstudien üblich, eine Regelung zur Durchführung von Nachtests festgelegt. Danach wurde die Erhebung in einer Schule wiederholt, wenn sowohl mindestens 15 Prozent als auch mindestens drei der für den Test zugelassenen Schülerinnen und Schüler einer Klasse an beiden Testtagen abwesend waren. Da jedoch die Beteiligung an den Testungen in allen Schulen ausreichend hoch war, mussten im Ländervergleich 2011 keine Nachtestungen durchgeführt werden.

4.5 Definition der Zielpopulation und Stichprobenziehung

Im Ländervergleich 2011 sollte die Population der Schülerinnen und Schüler der vierten Jahrgangsstufe für jedes Land in der getesteten Stichprobe möglichst genau repräsentiert sein. Dies erfordert zunächst eine präzise Bestimmung der Zielpopulation. Die Grundlage hierfür bildete die Schulstatistik für die dritte Jahrgangsstufe im Schuljahr 2009/2010, die für die Prognose der Schülerzahlen in der vierten Jahrgangsstufe herangezogen wurde.

Tabelle 4.3 gibt einen Überblick über die Anzahl der Schülerinnen und Schüler der vierten Jahrgangsstufe pro Land und Schulart für das Schuljahr 2010/2011. Die Übersicht zeigt, dass sich die Anteile der Schülerinnen und Schüler in den Förderschulen der Länder zum Teil deutlich unterscheiden. Die Anteile variieren

zwischen 0.6 Prozent in Bremen und 7.8 Prozent in Mecklenburg-Vorpommern. Diese Differenzen sind in erster Linie auf länderspezifische Ansätze der Beschulung von Kindern mit sonderpädagogischem Förderbedarf zurückzuführen. Während Kinder mit sonderpädagogischem Förderbedarf in einigen Ländern überwiegend Regelschulen besuchen, werden sie in anderen Ländern mehrheitlich in Förderschulen unterrichtet. Für einen validen Ländervergleich ist es erforderlich, die Zielpopulation einheitlich zu definieren. Um sicherzustellen, dass Kinder mit sonderpädagogischem Förderbedarf in allen Ländern in vergleichbarem Maße repräsentiert sind, wurden in den Ländervergleich 2011 nicht nur Regelschulen, sondern auch Förderschulen einbezogen. Allerdings beschränkte sich die Einbeziehung auf Schulen mit solchen Förderschwerpunkten, bei denen davon auszugehen ist, dass die jeweiligen Schülerinnen und Schüler zur Bearbeitung des Tests grundsätzlich in der Lage sind. Dazu zählen Schulen mit den Schwerpunkten „Lernen“, „soziale und emotionale Entwicklung“ sowie „Sprache“. Schulen mit anderen Förderschwerpunkten (geistige Entwicklung, Hören, Sehen sowie körperliche und motorische Entwicklung) wurden nicht als Teil der Zielpopulation definiert.

Tabelle 4.3: Schülerinnen und Schüler der vierten Jahrgangsstufe im allgemeinbildenden Schulsystem nach Land und Schulart (prozentuale Angaben) im Schuljahr 2010/2011

Land	Insgesamt ¹	GS	IGS	FrW	FS
Baden-Württemberg ²	107 764	94.1	--	1.8	4.1
Bayern	123 663	95.3	--	0.6	4.1
Berlin	24 942	91.9	3.2	1.3	3.5
Brandenburg	19 534	95.5	0.3	0.5	3.7
Bremen	5 480	97.9	--	1.4	0.6
Hamburg	15 175	86.3	8.2	1.7	3.8
Hessen	56 651	94.7	1.3	0.7	3.3
Mecklenburg-Vorpommern	12 725	91.6	--	0.6	7.8
Niedersachsen	79 427	95.5	0.3	0.8	3.5
Nordrhein-Westfalen	179 039	95.3	--	0.7	3.9
Rheinland-Pfalz	38 476	97.0	--	0.5	2.5
Saarland	8 032	97.6	--	1.2	1.2
Sachsen	31 308	93.7	--	0.4	5.9
Sachsen-Anhalt	17 253	93.1	--	0.3	6.6
Schleswig-Holstein	28 159	97.3	--	1.4	1.3
Thüringen	16 928	94.7	0.8	0.7	3.9

Anmerkungen. GS = Grundschulen; IGS = Integrierte Gesamtschulen; FrW = Freie Waldorfschulen; FS = Förderschulen.

¹Absolute Häufigkeiten. ² Förderschülerinnen und Förderschüler nach Klassenstufe tlw. vom Statistischen Bundesamt geschätzt. Quelle: Fachserie 11 des Statistischen Bundesamtes. Schuljahr 2010/2011.

Alle Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf wurden darüber hinaus danach unterschieden, ob sie entsprechend dem (Rahmen-) Lehrplan für Regelschulen oder auf Grundlage eines spezifischen (Rahmen-) Lehrplans unterrichtet werden. Diese Zuordnung wird auch als *zielgleiches beziehungsweise zieldifferentes Unterrichten* bezeichnet. Die Entscheidung, welche Schülerinnen und Schüler zielgleich oder zieldifferent unterrichtet werden, hängt eng mit den jeweiligen Förderschwerpunkten zusammen und wird von jedem Land nach eigenen Regelungen festgelegt. Eine über alle Länder der Bundesrepublik Deutschland einheitliche Zuordnung von Kindern mit sonderpädagogischem Förderbedarf zu einer dieser beiden Gruppen ist somit nicht möglich. Daher werden im Ländervergleich sowohl die zielgleich als auch die zieldifferent unterrichteten Kinder in die Zielpopulation einbezogen. Dies betrifft die Ergebnisse, die in den Kapiteln 5 sowie 7 bis 11 berichtet werden. Die Darstellung der Kompetenzstufenverteilung in Kapitel 6 bezieht sich jedoch ausschließlich auf die zielgleich unterrichteten Schülerinnen und Schüler, da davon auszugehen ist, dass für zieldifferent unterrichtete Schülerinnen und Schüler andere Lernziele gelten als die mit den Bildungsstandards definierten.

Die Stichprobenziehung erfolgte in mehreren Schritten (Frey, Carstensen, Walter, Rönnebeck & Gomolka, 2008). Zunächst wurde innerhalb jedes Landes eine Zufallsstichprobe von Schulen aus der Gesamtheit aller gemeldeten Schulen gezogen. Die Anzahl der zu untersuchenden Schulen orientierte sich an der Leistungsstreuung, die aufgrund von Daten bereits durchgeführter Vergleichsuntersuchungen zu erwarten war. In der Untersuchung IGLU-E⁴ 2006 ergaben sich für die Stadtstaaten Berlin, Hamburg und Bremen vergleichsweise große Stichprobenfehler (Bos et al., 2008), so dass in diesen Ländern eine entsprechend größere Anzahl von Schulen gezogen wurde. In den Flächenländern wurde eine Stichprobengröße von jeweils 80 Schulen und in den Stadtstaaten eine Stichprobe von 120 (Berlin und Hamburg) beziehungsweise 100 Schulen (Bremen)⁵ unter proportionaler Einbeziehung der Anzahl von Förderschülerinnen und Förderschülern anvisiert.

Im zweiten Schritt wurde innerhalb jeder Schule eine Klasse der vierten Jahrgangsstufe zufällig zur Teilnahme an der Untersuchung ausgewählt. In diesen Klassen nahmen alle Schülerinnen und Schüler an der Testung teil. An Schulen mit jahrgangsübergreifendem Lernen bezog sich die Zufallsauswahl nicht auf eine Klasse, sondern auf eine Lerngruppe, in der sich Schülerinnen und Schüler der vierten Jahrgangsstufe befanden. In die Testung wurden dann jedoch nur diejenigen Kinder einbezogen, die aufgrund ihres Einschulungszeitpunkts dem vierten Jahrgang zuzuordnen waren.

4.6 Realisierte Stichprobe

Die Teilnahme am Ländervergleich war für alle Schulen in öffentlicher Trägerschaft verbindlich, Privatschulen konnten die Teilnahme dagegen ablehnen. Tabelle 4.4 gibt einen Überblick über die Anzahl der teilnehmenden Schulen pro Land und Schulart.

4 Als IGLU-E wird die nationale Ergänzungsstudie bezeichnet, die aufgrund einer zusätzlich gezogenen Stichprobe in Anbindung an PIRLS/IGLU einen Vergleich der in den einzelnen Ländern erreichten Kompetenzen ermöglichte.

5 In Bremen wurde die Stichprobe auf 100 Schulen beschränkt, da durch die parallele Durchführung von PIRLS/IGLU 2011 und TIMSS 2011 sowie aufgrund weiterer Erhebungen nahezu alle Bremer Schulen im Frühjahr 2011 an Studien beteiligt waren.

Tabelle 4.4: Realisierte Schulstichprobe nach Land und Schulart

Land	GS	FrW	FS	Gesamt
Baden-Württemberg	73	0	5	78
Bayern	75	0	3	78
Berlin	113	0	5	118
Brandenburg	77	0	2	79
Bremen	91	0	1	92
Hamburg	112	0	2	114
Hessen	76	1	2	79
Mecklenburg-Vorpommern	73	0	7	80
Niedersachsen	76	0	4	80
Nordrhein-Westfalen	76	0	3	79
Rheinland-Pfalz	77	0	2	79
Saarland	75	1	1	77
Sachsen	74	0	5	79
Sachsen-Anhalt	75	0	5	80
Schleswig-Holstein	74	1	2	77
Thüringen	78	0	2	80
Gesamt	1 295	3	51	1 349

Anmerkung. GS = Grundschulen, FrW = Freie Waldorfschulen, FS = Förderschulen.

Da die Ziehung von Ersatzschulen nur in Einzelfällen (z. B. bei Renovierungsarbeiten) möglich war, wurde die angestrebte Schulstichprobe von 1 380 Schulen mit 1 349 teilnehmenden Schulen nicht ganz erreicht. Gründe für den Ausfall lagen vor allem in der Schließung oder Zusammenlegung von Schulen und der kurzfristigen Absage von Privatschulen, die eine nachträgliche Rekrutierung von Ersatzschulen unmöglich machte. Internationale Schulen, in denen der Unterricht in einer Fremdsprache erteilt wird, wurden ebenfalls aus der Stichprobe ausgeschlossen. Alle gezogenen Schulen wurden bis zum Januar 2011 über ihre Auswahl zur Teilnahme an der Studie informiert.

Die Teilnahme am Ländervergleich war nicht nur für die öffentlichen Schulen, sondern auch für ihre Schülerinnen und Schüler verpflichtend. Nur an Privatschulen konnten die Eltern einer Teilnahme an der Testung widersprechen. In Anlehnung an die internationalen Grundschulstudien PIRLS/IGLU 2011 und TIMSS 2011 konnten Schülerinnen und Schüler ausgeschlossen werden, wenn auf sie mindestens eines der drei folgenden Kriterien zutraf:

1. Schülerinnen und Schüler mit einer dauerhaften körperlichen Beeinträchtigung, die ihnen die Teilnahme an der Testsitzung unmöglich machte.
2. Geistig beeinträchtigte Schülerinnen und Schüler, die nach professioneller Einschätzung der Schulleitung oder einer anderen qualifizierten Person des Kollegiums emotional oder geistig nicht in der Lage waren, auch nur den allgemeinen Anweisungen des Tests zu folgen, oder die durch die Testsituation in unzumutbarem Maße emotional belastet worden wären.

3. Schülerinnen und Schüler nicht deutscher Muttersprache, die weniger als ein Jahr in deutscher Sprache unterrichtet wurden, und die nicht in der Lage waren, Deutsch zu lesen oder zu sprechen.

Insgesamt wurden 171 Schülerinnen und Schüler aus einem dieser Gründe von der Teilnahme am Ländervergleich 2011 ausgeschlossen. Dies entspricht einem Anteil von weniger als einem Prozent aller für den Test ausgewählten Schülerinnen und Schüler. Die Entscheidung über den Ausschluss von Schülerinnen und Schülern nach diesen Kriterien traf die jeweilige Schulkoordinatorin beziehungsweise der jeweilige Schulkoordinator.

Die realisierte Gesamtstichprobe umfasst 27 081 Schülerinnen und Schüler der vierten Jahrgangsstufe. Diese Zahl entspricht einer gewichteten Teilnahmequote⁶ für die Tests von 98,3 Prozent und liegt damit etwas höher als bei den zuletzt in Deutschland durchgeführten Grundschuluntersuchungen IGLU 2006 (93,5 %, Hornberg, Bos, Buddeberg, Potthoff & Stubbe, 2007) und TIMSS 2007 (96,0 %, Bensen, Lintorf, Bos & Frey, 2008). Die Teilnahmequote für die Schülerfragebögen fällt mit 87,3 Prozent insgesamt etwas niedriger aus, da deren Bearbeitung in einigen Ländern freiwillig war. Der Rücklauf der Elternfragebögen liegt im Durchschnitt bei 81,4 Prozent und damit zwischen dem Niveau der Studie IGLU 2006 (87 %, Hornberg et al., 2007) und TIMSS 2007 (73,6 %, Lintorf, 2012).

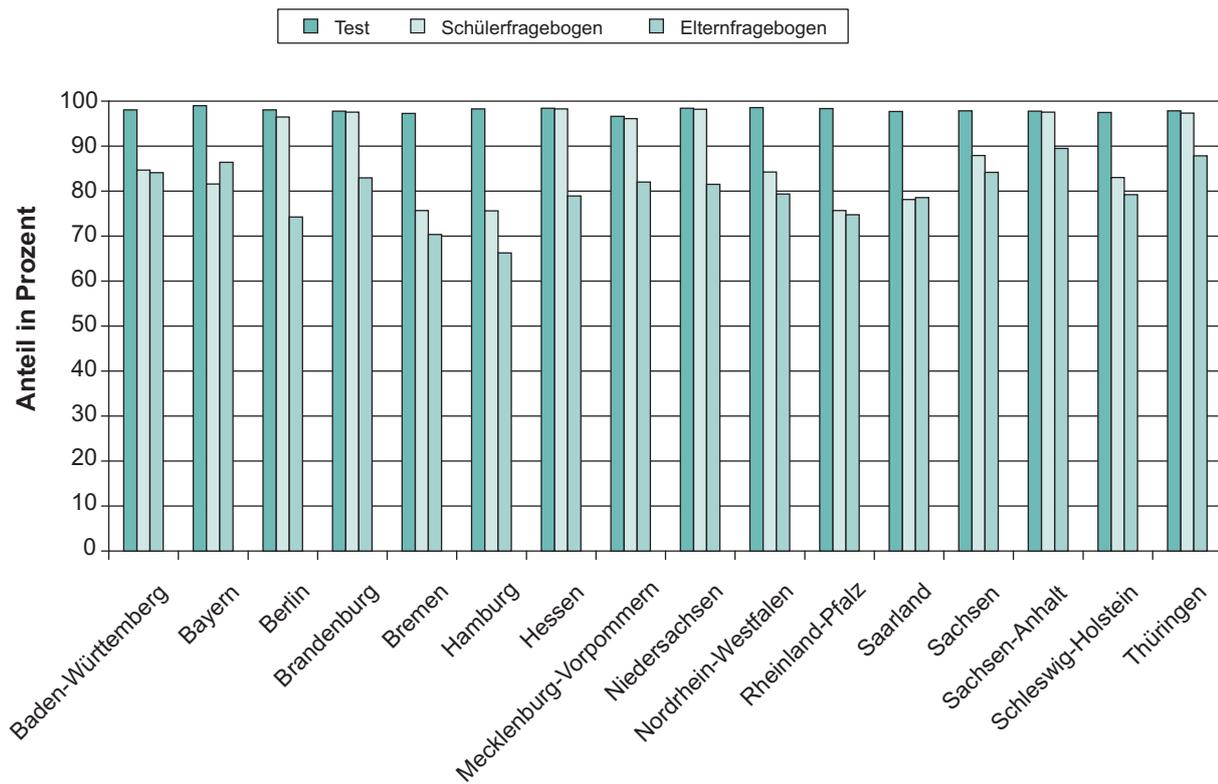
Auch innerhalb der einzelnen Länder liegt die Beteiligungsquote an den Tests mit jeweils über 96 Prozent auf konstant hohem Niveau. In Bezug auf die Quote der bearbeiteten Schülerfragebögen unterscheiden sich die Länder dagegen deutlich (vgl. Abbildung 4.1). Die Beteiligung variiert zwischen 75,6 Prozent in Hamburg und 98,3 Prozent in Hessen. Diese Differenzen dürften vor allem darauf zurückzuführen sein, dass die Teilnahme an der Schülerbefragung in einigen Ländern freiwillig, in anderen Ländern verbindlich war. Aber auch für die Elternbefragung ist eine große Spannweite der Teilnahmequoten zu beobachten. Die niedrigste Quote findet sich wiederum in Hamburg (66,3 %), die höchste in Sachsen-Anhalt (89,5 %).

An der Lehrerbefragung nahmen insgesamt 1 816 Lehrkräfte der Fächer Deutsch und Mathematik teil. Für das Fach Deutsch ist aus 88,5 Prozent der Klassen mindestens ein Lehrerfragebogen eingegangen, für das Fach Mathematik liegen Fragebögen aus 86,8 Prozent der Klassen vor. Deutlich höher fiel mit 94,3 Prozent die Beteiligung bei der Schulleiterbefragung aus. Dieser Wert entspricht in etwa der Teilnahmequote, die im Rahmen von IGLU 2006 für die Schulleiterbefragung erreicht wurde (Hornberg et al., 2007).

In der länderspezifischen Darstellung der Beteiligungsquoten für die Lehrkräfte- und Schulleiterbefragung in Abbildung 4.2 sind ebenfalls deutliche Unterschiede zu erkennen. Die geringste Beteiligung weist wiederum Hamburg auf. Teilnahmequoten mit über 90 Prozent für die Lehrkräfte- und Schulleiterbefragung finden sich in Bayern, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen-Anhalt und Thüringen.

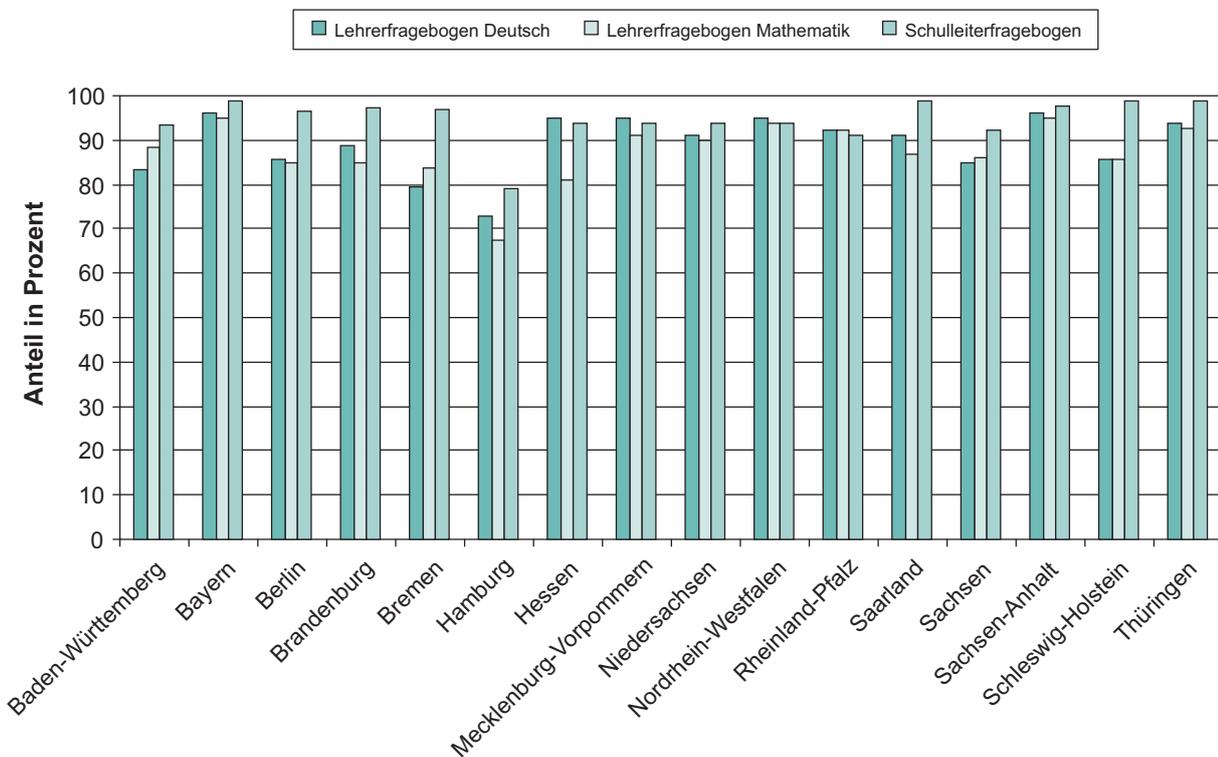
6 Zur Berechnung der Teilnahmequote werden in großen Schulleistungsstudien individuelle Schülegewichte verwendet. Die Gewichte geben an, wie viele Schülerinnen und Schüler der Population jede einzelne Schülerin / jeder einzelne Schüler in der Stichprobe repräsentiert. Die gewichtete Teilnahmequote kennzeichnet, welcher Anteil der Population durch die Stichprobe abgebildet wird. Eine gewichtete Teilnahmequote wurde für die Tests sowie für die Schüler- und Elternfragebögen berechnet. Die Berechnung der Teilnahmequoten der Lehrer- und Schulleiterbefragung erfolgte ungewichtet, da keine Informationen über die Population vorliegen.

Abbildung 4.1: Teilnahmequoten für den Test sowie den Schüler- und Elternfragebogen nach Land



Anmerkung. Für den Test sowie die Schüler- und Elternfragebögen werden gewichtete Teilnahmequoten berichtet.

Abbildung 4.2: Teilnahmequoten für den Lehrer- und Schulleiterfragebogen nach Land



Insgesamt liegen die Teilnahmequoten für die Tests und die Fragebogenerhebungen im Vergleich zu anderen internationalen und nationalen Large-Scale-Assessments auf sehr hohem Niveau (vgl. auch Köller, Knigge & Tesch, 2010). In den Stadtstaaten Bremen und Hamburg fielen jedoch die Beteiligungsquoten für alle Fragebögen vergleichsweise niedrig aus.

4.7 Aufbereitung und Analyse der Daten

Für die Skalierung der Testdaten wird das Rasch-Modell aus der Familie der probabilistischen Testmodelle herangezogen, bei dem für jedes Item ein Schwierigkeitsparameter geschätzt wird (Rasch, 1960). Die Kalibrierung erfolgt unter Einbeziehung aller zur Verfügung stehenden Testdaten pro Fach und Kompetenzbereich in jeweils eindimensionalen Modellen. Nach der Kalibrierung werden die Personenfähigkeiten unter Berücksichtigung individueller Hintergrundmerkmale (z.B. zuwanderungsbezogene und soziale Merkmale) als sogenannte *Plausible Values* geschätzt (von Davier, Gonzalez & Mislevy, 2009).

Dieses Verfahren gewährleistet, dass sowohl Zusammenhänge zwischen Individualmerkmalen und Leistungskennwerten als auch die Populationsvarianzen möglichst unverzerrt abgebildet werden (Mislevy, Beaton, Kaplan & Sheehan, 1992). Aus diesem Grund wurden die Merkmale, die in den vertiefenden Analysen von Daten des Ländervergleichs genauer betrachtet werden, in das Hintergrundmodell aufgenommen. Angaben, die für einzelne Schülerinnen und Schüler fehlten, wurden unter Nutzung des EM-Algorithmus imputiert (Schafer & Graham, 2002). Die Imputationen erfolgten separat für jedes Land, da die Zusammenhänge der Variablen zwischen den Ländern variieren können. Nach erfolgter Imputation wurde eine Hauptkomponentenanalyse durchgeführt, um die verfügbaren Informationen aus dem Hintergrundmodell in untereinander unkorrelierte Variablenbündel zusammenzufassen. Pro Land wurden die Hauptkomponenten in das Hintergrundmodell aufgenommen, die zusammen mindestens 95 Prozent der Variabilität in den Hintergrundvariablen erklären (vgl. Böhme et al., 2010).

Der üblichen Praxis in Schulleistungsstudien entsprechend, wurden die ermittelten Personenfähigkeiten für jeden Kompetenzbereich auf eine Skala mit dem Mittelwert 500 und der Standardabweichung 100 transformiert („Berichtsmetrik“).

Um aus den erhobenen Daten der Stichprobe Rückschlüsse auf die Population ziehen zu können, muss diese angemessen repräsentiert sein. In großen Schulleistungsstudien wie PISA, IGLU, TIMSS und dem Ländervergleich werden Schülergewichte in die Analysen einbezogen, um die Repräsentativität der Ergebnisse zu gewährleisten. In die Berechnung der Schülergewichte gehen Faktoren auf der Ebene der Schule, der Klasse und des Individuums ein (siehe auch OECD, 2012). Auf Schulebene wird berücksichtigt, dass nur eine Auswahl aller im Land vorhandenen Schulen an der Untersuchung teilnimmt, jedoch Aussagen für alle Schulen getroffen werden sollen. Auf Klassenebene wird dafür Rechnung getragen, dass sich Schulen in der Anzahl der Klassen in der vierten Jahrgangsstufe unterscheiden. In Schulen mit zwei Klassen in der vierten Jahrgangsstufe liegt die Ziehungswahrscheinlichkeit pro Klasse bei 50 Prozent, bei vier Klassen beträgt sie nur 25 Prozent. Auf Individualebene wird darüber hinaus berücksichtigt, dass einzelne Schülerinnen und Schüler am Testtag fehlen. Unter Beachtung aller genannten Faktoren erhält jede Schülerin und jeder Schüler ein Gewicht, welches

angibt, wie viele Schülerinnen und Schüler der Zielpopulation sie beziehungsweise er repräsentiert. Mit diesem Gewicht ist es möglich, einen für die Population unverzerrten Durchschnittswert für die erreichten Kompetenzen zu schätzen. Die Gewichte werden auch in den Analysen der Kapitel 5 bis 11 berücksichtigt.

Die Ergebnisse dieses Berichts sollen eine Aussage darüber erlauben, welches Kompetenzniveau die Zielpopulation der Schülerinnen und Schüler in der vierten Jahrgangsstufe in den Ländern erreicht und welcher Anteil die Anforderungen der Bildungsstandards erfüllt. Da jedoch aus ökonomischen Gründen nicht alle Schülerinnen und Schüler getestet wurden, können die anhand der Stichprobe ermittelten statistischen Kennwerte (z. B. der Mittelwert) nur mit einem bestimmten Maß an Sicherheit auf die Zielpopulation übertragen werden. Zur Abschätzung des Grades an Unsicherheit wird deshalb ein Standardfehler (*SE*) berechnet, der angibt, wie genau ein Kennwert die Zielpopulation beschreibt. Die Bestimmung des Standardfehlers erfolgt in Schulleistungsstudien dieser Größe mit dem sogenannten Jackknife-Verfahren, das die hierarchische Datenstruktur (Testung von Schülerinnen und Schülern innerhalb von Klassen) berücksichtigt (siehe Kapitel 12).

4.8 Qualitätssicherung

Im Ländervergleich 2011 wurde ein Qualitätsmonitoring durchgeführt, um stichprobenartig die Einhaltung der standardisierten Durchführungsbestimmungen zu überprüfen und um Aussagen zur Durchführungsobjektivität treffen zu können. Hierfür wurde ein Vorgehen gewählt, das sich bereits im Ländervergleich 2009 und in anderen großen Schulleistungsstudien bewährt hat. Beim Qualitätsmonitoring besuchen geschulte Beobachterinnen und Beobachter unangekündigt ausgewählte Teilnehmerschulen während des Tests und dokumentieren den Testablauf, ohne selbst Einfluss auf die Testung zu nehmen. Zusätzlich zur Beobachtung haben sie die Aufgabe, die Schulkoordinatorin oder den Schulkoordinator zum Ablauf der Vorbereitung und Durchführung der Studie zu befragen.

Für das Qualitätsmonitoring wurden insgesamt 100 Schulen aus der Gesamtzahl der Teilnehmerschulen per Zufall durch das DPC ausgewählt. Dies entspricht einem Anteil von 7,4 Prozent aller teilnehmenden Schulen. Die Zahl der während der Testdurchführung besuchten Schulen richtete sich dabei nach der Anzahl der im jeweiligen Land teilnehmenden Schulen. In Ländern mit größeren Schulstichproben nahmen auch mehr Schulen am Qualitätsmonitoring teil. Das Monitoring fand jeweils nur an einem von zwei Testtagen statt. Die Festlegung, an welchem der beiden Tage die Schule besucht wurde, erfolgte ebenfalls per Zufall, so dass die eine Hälfte der Schulen am ersten, die andere Hälfte der Schulen am zweiten Testtag beobachtet wurde.

Für die Erfassung des Ablaufs der Erhebungen erhielten die Beobachterinnen und Beobachter einen standardisierten Fragebogen, der sich vor allem auf die Vorbereitung des Testraumes und der Testmaterialien, die Einhaltung der vorgegebenen Instruktionen und Testzeiten, die Anwesenheit der Schülerinnen und Schüler und das Verhalten der Schülerinnen und Schüler in der Testsituation bezog. Die zentralen Ergebnisse des Qualitätsmonitorings werden im Folgenden kurz beschrieben. Genauere Informationen dazu finden sich im Bericht von Böhme und Mallwitz (2011).

Vorbereitung der Erhebung

Die Vorbereitung der Testung im Klassenzimmer verlief in fast allen Schulen entsprechend dem vorgeschriebenen Prozedere. Auch die von der Schule zur Verfügung gestellten Räume ermöglichten in den meisten Fällen ein störungsfreies Bearbeiten der Tests. Nur vereinzelt, in sieben Schulen, entstanden durch die Nähe zum Pausenhof oder zu anderen Räumen störende Geräusche. Die Anwesenheit der Schülerinnen und Schüler wurde den Vorgaben entsprechend zu Beginn der Testung und vor der Fragebogenerhebung von der Testleiterin beziehungsweise dem Testleiter überprüft.

Administration der Tests und Fragebögen

Die Durchführung der Testung und der Einsatz der Fragebögen verliefen nach Einschätzung der Beobachterinnen und Beobachter weitgehend problemlos. Die Instruktionen wurden in den meisten Fällen durch die Testleiterinnen und Testleiter wörtlich verlesen und nur in wenigen Ausnahmefällen geringfügig ergänzt oder verkürzt. Auch die vorgeschriebenen Test- und Pausenzeiten wurden von den Testleiterinnen und Testleitern eingehalten. In einzelnen Fällen wurde entsprechend den Vorgaben für die Bearbeitung des Schülerfragebogens mehr Zeit zur Verfügung gestellt, um es den Schülerinnen und Schülern zu ermöglichen, alle Fragen zu beantworten. Nahezu alle Testleiterinnen und Testleiter wirkten im Umgang mit den Testunterlagen vertraut. Die Schülerinnen und Schüler verhielten sich während der Testung in der Regel diszipliniert und kooperativ. Störungen, die mehr als eine Minute anhielten, wurden nur in einem Fall beobachtet. Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf stellten teilweise zusätzliche Rückfragen, bearbeiteten jedoch nach Angaben der Beobachterinnen und Beobachter in fast allen Fällen die Testaufgaben und Fragebögen vollständig.

Einschätzung der Schulkoordinatorinnen und -koordinatoren

Die Befragung der Schulkoordinatorinnen und Schulkoordinatoren ergab, dass sich eine große Mehrheit gut über den Ländervergleich informiert fühlte und der Studie gegenüber weitgehend positiv eingestellt war. Die Organisation der Testung an der Schule erfolgte nach ihrer Einschätzung in den meisten Fällen problemlos. Auch die Zusammenarbeit mit den Eltern und die Unterstützung der Lehrkräfte wurden von den Schulkoordinatorinnen und Schulkoordinatoren positiv eingeschätzt. Über die Hälfte der Befragten sprach sich jedoch für eine Testung zu einem früheren Zeitpunkt im Schuljahr aus.

Implikationen für die Interpretation der Daten

Insgesamt weisen die Ergebnisse des Qualitätsmonitorings darauf hin, dass die Testvorbereitung durch die Schulkoordinatorinnen und Schulkoordinatoren erfolgreich verlief und die Testleiterinnen und Testleiter die standardisierten Durchführungsbestimmungen, von wenigen Ausnahmen abgesehen, eingehalten haben. Somit sind bei der Testdurchführung offenbar keine gravierenden Probleme aufgetreten, die sich auf die Interpretierbarkeit der Daten des Ländervergleichs 2011 auswirken könnten.

4.9 Beteiligte Institutionen und Personen

Die wissenschaftliche Leitung des Ländervergleichs 2011 liegt bei den Direktoren des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin, Prof. Dr. Petra Stanat und Prof. Dr. Hans Anand Pant. Die Vorbereitung der Studie, die Aufbereitung und Analyse der Daten sowie die Berichterlegung erfolgte durch die folgenden Mitarbeiterinnen und Mitarbeiter des Arbeitsbereichs Grundschule im IQB:

Wissenschaftliche und administrative Koordination der Studie

Dr. Dirk Richter
Dr. Katrin Böhme
Dr. Heino Reimers

Mitarbeit im Fach Deutsch

Maria Engelbert
Josefine Prengel
Sebastian Weirich

Mitarbeit im Fach Mathematik

Nicole Haag
Jasmin Hannighofer
Alexander Roppelt

Mit der praktischen Durchführung der Studie wurde das DPC in Hamburg beauftragt. Die Aufgaben des DPC umfassten unter anderem die Stichprobenziehung, die Koordination der Datenerhebung, die Kodierung der Schülerleistungen sowie die Eingabe und Aufbereitung der Daten. Am DPC waren folgende Personen in den Ländervergleich 2011 eingebunden:

Julia Bockelmann	Guido Martin
Tina Ebert	Heiko Sibberns (Leitung)
Jens Gomolka	Anja Waschk

Aufgrund der großen Anzahl der teilnehmenden Schulen mussten wichtige Aspekte der Vorbereitung und Durchführung des Ländervergleichs von Projektkoordinatorinnen und -koordinatoren in den einzelnen Ländern übernommen werden. Zu deren Aufgaben gehörte die Organisation von Schulleiterinformationsveranstaltungen, in denen das IQB gemeinsam mit dem DPC über die Ziele, die Organisation und den Ablauf der Studie informierte. Weiterhin waren sie für die Rekrutierung und Einsatzplanung der Testleiterinnen und Testleiter, die Koordination der Testleiterschulungen, die Kommunikation mit den teilnehmenden Schulen sowie die Übermittlung der Klassenstichprobenziehung an das DPC verantwortlich. Bei den Projektkoordinatorinnen und -koordinatoren handelt es sich um Mitarbeiterinnen und Mitarbeiter der Ministerien, der Landesinstitute für Lehrerfortbildung oder der Qualitätsagenturen der Länder. Als Projektkoordinatorinnen und Projektkoordinatoren waren die folgenden Personen tätig:

Dr. Desiree Burba (SH)	Gerhard Krohne (NI)
Rüdiger Cwielong (SL)	Karin Lindenstruth (NW)
Dr. Grit Elsner (SN)	Johannes Miethner (RP)

Dr. Holger Gärtner (BE/BB)	Daniela Neumann, Maria Skejic (HE)
Oliver Groth (HH)	Steffen Pleßmann (BW)
Dietmar Kirchhoff, Karsten Wolf (HB)	Bianca Schmidt (BY)
Frank Kirchner (ST)	Maria Skejic (HE)
Astrid Koriller (MV)	Heiko Wontroba (TH)

Die Testungen in den Schulen wurden von Testleiterinnen und Testleitern durchgeführt, die vom DPC geschult wurden. Dabei handelte es sich in der Regel um Studierende, Referendarinnen und Referendare, Lehrkräfte oder Mitarbeiterinnen und Mitarbeiter des schulppsychologischen Dienstes.

An der Entwicklung der Aufgaben, die im Ländervergleich eingesetzt wurden, wirkten Fachlehrkräfte aus allen 16 Ländern mit. Die Aufgabenentwicklerinnen und Aufgabenentwickler arbeiteten in regionalen Gruppen zusammen und erhielten Unterstützung von fachdidaktischen Beraterinnen und Beratern. Die fachdidaktische Anleitung erfolgte im Fach Deutsch durch Prof. Dr. Albert Bremerich-Vos (Universität Duisburg-Essen) und im Fach Mathematik durch Prof. Dr. Gerd Walther (Universität Kiel) sowie Prof. Dr. Marja van den Heuvel-Panhuizen (Universität Utrecht). In den Teams arbeiteten die folgenden Aufgabenentwicklerinnen und -entwickler mit:

Aufgabenentwicklung im Fach Deutsch

Petra Bittins (BB)	Ulrike Potthoff (NW)
Waltraud Credé (HE)	Catrin Puschmann (NI)
Sabine Dörnhaus (SH)	Ilse Stork (BY)
Christine Förster (HE)	Elvira Throm (BW)
Waltraud Frölich (BW)	Elke Uthe (ST)
Christiane Hermes (MV)	Sylke Wiegand (TH)
Brigitte Heßler (SN)	Frauke Wietzke (SH)
Petra Klein (SL)	Uta Zimmermann (HH)
Karin Möcklinghoff (HH)	

Aufgabenentwicklung im Fach Mathematik

Ute Alsdorf (TH)	Hanna Haubold (MV)
Ute Baumann (BB)	Antje Hoffmann (NI)
Thomas Bongartz (NW)	Barbara Meyer-Wirth (SL)
Brigitte Dedekind (SH)	Kirsten Rähling (HH)
Rita Dürr (BW)	Charlotte Rechtsteiner-Merz (BW)
Ilse Eckhardt (HE)	Sabine Schmidt (ST)
Gerda Frommeyer (HB)	Charlotte Schorr-Brill (SL)
Hedwig Gasteiger (BY)	Christian Schuster (RP)
Astrid Gebert (BE)	

Die Einhaltung der Vorgaben zur Durchführung der Studie wurde im Rahmen eines Qualitätsmonitorings überprüft. Das Qualitätsmonitoring wurde durch Wissenschaftlerinnen und Wissenschaftler organisiert, die in der Region, für die sie jeweils zuständig waren, Beobachterinnen und Beobachter rekrutierten und deren unangekündigte Schulbesuche koordinierten. Als Koordinatorinnen und Koordinatoren für das Qualitätsmonitoring wirkten mit:

Dr. Anja Friedrich (BW, RP, SL)	Prof. Dr. Barbara Moschner (HB, NI)
Prof. Dr. Mareike Kunter (HE)	Christine Schmeißer (BY, TH)
Prof. Dr. Nele McElvany (NW)	Prof. Dr. Hans-Martin Süß (SN, ST)
Prof. Dr. Jens Möller (HH, SH)	Prof. Dr. Rainer Watermann (BE, BB, MV)

Literatur

- Böhme, K., Leucht, M., Schipolowski, S., Porsch, R., Knigge, M. & Köller, O. (2010). Anlage und Durchführung des Ländervergleichs. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 65–85). Münster: Waxmann.
- Böhme, K. & Mallwitz, M. (2011). Qualitätsmonitoring im Rahmen des Ländervergleichs in der Primarstufe 2011. Unveröffentlichter Bericht, Institut zur Qualitätsentwicklung im Bildungswesen.
- Bonsen, M., Lintorf, K., Bos, W. & Frey, K. A. (2008). TIMSS 2007 Grundschule – Eine Einführung in die Studie. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 19–48). Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2008). *IGLU-E 2006: Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 375–397). Münster: Waxmann.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz Test.
- Hornberg, S., Bos, W., Buddeberg, I., Potthoff, B. & Stubbe, T. C. (2007). Anlage und Durchführung von IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 21–79). Münster: Waxmann.
- KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards der Kultusministerkonferenz: Erläuterungen zur Konzeption und Entwicklung*. Neuwied: Luchterhand.
- Köller, O., Knigge, M. & Tesch, B. (Hrsg.). (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Lintorf, K. (2012). *Wie vorhersagbar sind Grundschulnoten? Prädiktionskraft individueller und kontextspezifischer Merkmale*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mayringer, H. & Wimmer, H. (2003). *Salzburger Lese-Screening für die Klassenstufen 1–4*. Bern: Verlag Hans Huber.
- Mislevy, R. J., Beaton, A. E., Kaplan, B. & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of items. *Journal of Educational Measurement*, 29, 133–164.
- OECD. (2012). *PISA 2009 Technical Report*. PISA, OECD Publishing, <http://dx.doi.org/10.1787/9789264167872-en>.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.
- Robitzsch, A. (2009). Methodische Herausforderungen bei der Kalibrierung von Leistungstests. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Wather (Hrsg.), *Bildungsstandards Deutsch und Mathematik: Leistungsmessung in der Grundschule* (S. 42–106). Weinheim: Beltz.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Statistisches Bundesamt. (2011). *Fachserie 11: Bildung und Kultur, Reihe 1: Allgemeinbildende Schulen*. Wiesbaden: Statistisches Bundesamt.
- von Davier, M., Gonzalez, E. & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series*, 4, 9–36.

Kapitel 5

Ländervergleich der in den Fächern Deutsch und Mathematik erzielten Kompetenzstände

5.1 Der Ländervergleich im Fach Deutsch

Katrin Böhme und Sebastian Weirich

Das vorliegende Teilkapitel beschreibt die von den Schülerinnen und Schülern in den einzelnen Ländern der Bundesrepublik Deutschland im Fach Deutsch erreichten Kompetenzstände. Der länderübergreifende Vergleich bezieht sich auf die Kompetenzbereiche Lesen und Zuhören, die in Kapitel 2.1 dieses Berichts konzeptuell und operational definiert worden sind.

Im Folgenden werden nun zunächst die im Mittel erreichten Kompetenzausprägungen dargestellt, und es wird erläutert, wie sich diese zwischen den Ländern unterscheiden (Abschnitt 5.1.1). Im Anschluss wird die Leistungsheterogenität innerhalb der einzelnen Länder genauer betrachtet (Abschnitt 5.1.2), und es wird der Frage nachgegangen, inwieweit die von Schülerinnen und Schülern erreichten Kompetenzen innerhalb und zwischen Schulen variieren (Abschnitt 5.1.3). Abschließend wird darauf eingegangen, wie sich die Ergebnisse des aktuellen Ländervergleichs 2011 für den Kompetenzbereich Lesen in die Befunde der nationalen Ergänzung der *Internationalen Grundschul-Lese-Untersuchung* IGLU-E 2006 einordnen lassen (Abschnitt 5.1.4).

5.1.1 Vergleich der im Fach Deutsch in den Ländern im Mittel erreichten Kompetenzstände

Zu den wichtigsten Indikatoren für die von den Grundschulkindern erreichten Kompetenzstände gehören die Kompetenzmittelwerte der 16 Länder. Diese sind für die beiden im Fach Deutsch untersuchten Kompetenzbereiche in Tabelle 5.1 dargestellt. Zusätzlich werden auch die Ergebnisse für Schülerinnen und Schüler in Großstädten aufgeführt. Diese Befunde stellen insbesondere für die drei Stadtstaaten Berlin, Bremen und Hamburg, die sich in ihrer Bevölkerungsstruktur teilweise deutlich von den Flächenländern unterscheiden, einen ergänzenden Vergleichsmaßstab zur Verfügung. In Anlehnung an die in PISA¹ 2000 verwendete Definition wurden dabei alle Städte mit einer Einwohnerzahl von mehr als 300 000 als Großstädte klassifiziert (Baumert et al., 2002).² Schülerinnen und

1 Das Akronym PISA steht für *Programme for International Student Assessment*.

2 Großstädte mit einer Einwohnerzahl von mehr als 300 000 sind Bielefeld, Bochum, Bonn, Dortmund, Dresden, Duisburg, Düsseldorf, Essen, Frankfurt am Main, Hannover, Köln, Leipzig, Mannheim, München, Nürnberg, Stuttgart und Wuppertal (Gebietsstand 31.12.2010, Quelle: Statistisches Bundesamt).

Schüler der Stadtstaaten Berlin, Bremen und Hamburg wurden dabei nicht in die Berechnung der Ergebnisse für Großstädte eingeschlossen, um einen sinnvollen Vergleich mit den drei Stadtstaaten zu ermöglichen.

Die in diesem Kapitel präsentierten Befunde basieren, wie in Kapitel 3.1 ausgeführt, auf einer Berichtsmetrik mit einem Mittelwert (M) von 500 und einer Standardabweichung (SD)³ von 100 Punkten. Diese gelten für die Population der Schülerinnen und Schüler, die 2011 in Deutschland die vierte Jahrgangsstufe besuchten.⁴

Tabelle 5.1: In den Ländern im Mittel erreichte Kompetenzstände im Fach Deutsch

Land	Lesen		Zuhören	
	M	(SE)	M	(SE)
Baden-Württemberg	505	(4.6)	509	(5.0)
Bayern	515	(5.3)	513	(6.0)
Berlin	467	(5.4)	472	(6.2)
Brandenburg	497	(3.8)	490	(4.7)
Bremen	463	(6.1)	467	(7.3)
Hamburg	478	(4.4)	486	(4.8)
Hessen	493	(5.4)	492	(6.4)
Mecklenburg-Vorpommern	496	(4.2)	495	(4.4)
Niedersachsen	503	(5.0)	507	(4.9)
Nordrhein-Westfalen	494	(5.0)	496	(6.0)
Rheinland-Pfalz	493	(5.7)	487	(5.1)
Saarland	495	(4.4)	484	(5.5)
Sachsen	513	(4.9)	499	(5.1)
Sachsen-Anhalt	511	(4.1)	487	(4.4)
Schleswig-Holstein	495	(5.2)	507	(5.5)
Thüringen	510	(4.7)	501	(4.8)
Großstädte ¹	493	(9.1)	490	(9.0)
Deutschland	500	(1.7)	500	(1.9)

Anmerkungen. Fett gedruckte Mittelwerte unterscheiden sich im jeweiligen Kompetenzbereich signifikant vom deutschen Mittelwert; M = Mittelwert; SE = Standardfehler. ¹ Ohne Berlin, Bremen und Hamburg.

Für jedes Land ist in Tabelle 5.1 neben den im Mittel erbrachten Leistungen (Mittelwert, M) zusätzlich der Standardfehler des jeweiligen Mittelwerts (*Standard Error*, SE) angegeben. Dieser Fehlerwert kennzeichnet die Genauigkeit der Schätzung der Populationsmittelwerte. Je kleiner der Standardfehler des Mittelwerts ausgeprägt ist, desto präziser bildet der Stichprobenmittelwert den tatsächlichen oder – in der Sprache der Inferenzstatistik – den „wahren“ Wert der Zielpopulation ab.

- 3 Die Standardabweichung ist ein Maß dafür, wie stark die Werte in einer Verteilung um den Mittelwert streuen.
- 4 Die Zielpopulation, und somit Grundlage der Ergebnisdarstellung in diesem Kapitel, bilden alle Schülerinnen und Schüler der vierten Jahrgangsstufe, die zum Zeitpunkt der Erhebung eine Grundschule oder Förderschule (Förderschwerpunkte Sprache, soziale und emotionale Entwicklung oder Lernen) in Deutschland besuchten, unabhängig davon, ob sie an dieser Schule zielgleich oder zieldifferent unterrichtet wurden (vgl. Kapitel 4).

Zur Quantifizierung der Unsicherheiten, mit denen die Schätzungen der Populationsmittelwerte behaftet sind, werden die Stichprobenkennwerte häufig durch die Angabe von Konfidenzintervallen ergänzt. Das Intervall von etwa zwei Standardfehlern ober- und unterhalb des jeweiligen Mittelwerts bildet den Wertebereich ab, in dem der wahre Mittelwert der Population mit einer Wahrscheinlichkeit von 95 Prozent liegt.⁵ Dies lässt sich anhand eines Beispiels wie folgt veranschaulichen: Die mittlere Kompetenz der Schülerinnen und Schüler im Bereich Zuhören liegt in der zufällig gezogenen Stichprobe des Landes Nordrhein-Westfalen bei 496 Punkten auf der Berichtsmetrik (vgl. Tabelle 5.1). Der Standardfehler dieser mittleren Kompetenz beträgt sechs Kompetenzpunkte. Berechnet man nun den Bereich, der zwischen zwei Standardfehlern oberhalb und zwei Standardfehlern unterhalb des Stichprobenmittelwertes von 496 liegt, ergibt sich ein Konfidenzintervall mit einer Breite von plus/minus 12 Punkten, also ein Bereich von 484 bis 508 Punkten.

Um einen Vergleich der in den einzelnen Ländern im Lesen und Zuhören erzielten Kompetenzausprägungen zu erleichtern, sind die Ergebnisse in den Abbildungen 5.1 und 5.2 zusätzlich grafisch dargestellt. In diesen Abbildungen sind die Länder absteigend nach ihrem jeweils erzielten Mittelwert angeordnet. Neben den Mittelwerten enthalten die Abbildungen ergänzend die für die Länder ermittelten Konfidenzintervalle, Standardabweichungen und verschiedene Perzentilwerte. Bei einem Perzentil handelt es sich um einen Wert auf der Berichtsskala, der einen bestimmten Bereich der Verteilung abschneidet. Der jeweilige Bereich wird dabei in Prozent angegeben. So finden sich beispielsweise für das Land Rheinland-Pfalz im Kompetenzbereich Lesen unterhalb des Wertes 320, der das 5. Perzentil markiert, die schwächsten 5 Prozent der Schülerinnen und Schüler in der Kompetenzverteilung dieses Landes. Der Wert 647 gibt für Rheinland-Pfalz das 95. Perzentil und damit denjenigen Wert auf der Berichtsskala an, unterhalb dessen sich 95 Prozent der Schülerinnen und Schüler befinden. Werte von 647 oder mehr Kompetenzpunkten erreichen im Lesen nur die 5 Prozent leistungsstärksten Schülerinnen und Schüler in Rheinland-Pfalz.

Vergleicht man für die beiden Kompetenzbereiche Lesen und Zuhören die in den einzelnen Ländern erzielten Kompetenzmittelwerte mit dem deutschen Mittelwert, so lassen sich jeweils drei Ländergruppen identifizieren. Im rechten Teil der Abbildungen 5.1 und 5.2 sind diese Gruppen jeweils durch unterschiedliche Umrandungen um die Perzentilbänder gekennzeichnet.

In der ersten Gruppe sind jene Länder zusammengefasst, in denen mittlere Kompetenzausprägungen erzielt wurden, die statistisch signifikant *über* dem deutschen Mittelwert von 500 Punkten liegen.⁶ Dies sind im Kompetenzbereich Lesen die Länder Bayern, Sachsen, Sachsen-Anhalt und Thüringen. Betrachtet man den Kompetenzbereich Zuhören, zeigt sich, dass es hier nur im Freistaat

5 Die Wahrscheinlichkeitsangabe von 95 Prozent lässt sich wie folgt interpretieren: Würde man die hier zugrunde liegende Untersuchung sehr oft, beispielsweise einhundertmal, mit zufällig gezogenen Stichproben wiederholen, dann kann davon ausgegangen werden, dass das für den jeweils erhaltenen Stichprobenmittelwert berechnete Konfidenzintervall in 95 dieser 100 Untersuchungen den wahren Mittelwert der Population einschließt.

6 Bei der Prüfung, ob sich ein Land in statistisch signifikantem Ausmaß vom deutschen Mittelwert unterscheidet, wurde – wie auch im Rahmen des Ländervergleichs 2009 (vgl. Köller, Knigge & Tesch, 2010) – ein Vorgehen gewählt, das in der Signifikanztestung die Standardfehler beider Gruppen (d. h. Deutschland insgesamt und das jeweilige Land) berücksichtigt.

Abbildung 5.1: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich Lesen

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Bayern	515	(5.3)	103	(2.0)	334	380	453	586	639	669	336
Sachsen	513	(4.9)	99	(2.3)	347	387	450	581	634	665	318
Sachsen-Anhalt	511	(4.1)	104	(2.7)	328	386	455	580	631	659	332
Thüringen	510	(4.7)	92	(2.5)	351	389	450	576	627	654	304
Baden-Württemberg	505	(4.6)	93	(2.1)	346	386	446	570	621	651	304
Niedersachsen	503	(5.0)	102	(3.8)	332	372	437	573	629	661	329
Deutschland	500	(1.7)	100	(1.0)	329	370	436	570	624	655	325
Brandenburg	497	(3.8)	98	(2.5)	326	369	433	566	620	650	324
Mecklenburg-Vorpommern	496	(4.2)	103	(3.5)	318	372	441	565	614	641	323
Schleswig-Holstein	495	(5.2)	99	(2.5)	324	363	429	564	620	651	327
Saarland	495	(4.4)	103	(3.5)	313	358	430	567	622	653	340
Nordrhein-Westfalen	494	(5.0)	97	(2.3)	332	368	428	562	619	651	319
Großstädte ¹	493	(9.1)	106	(6.0)	314	357	422	567	626	659	345
Hessen	493	(5.4)	103	(4.1)	319	364	430	563	617	649	330
Rheinland-Pfalz	493	(5.7)	100	(2.7)	320	362	428	563	617	647	327
Hamburg	478	(4.4)	102	(2.9)	303	345	410	550	604	634	331
Berlin	467	(5.4)	105	(2.6)	289	329	395	543	600	632	343
Bremen	463	(6.1)	101	(2.5)	298	333	393	536	594	625	327

☐ Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
 ☐ Länder liegen nicht signifikant vom deutschen Mittelwert ab.
 ☐ Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.

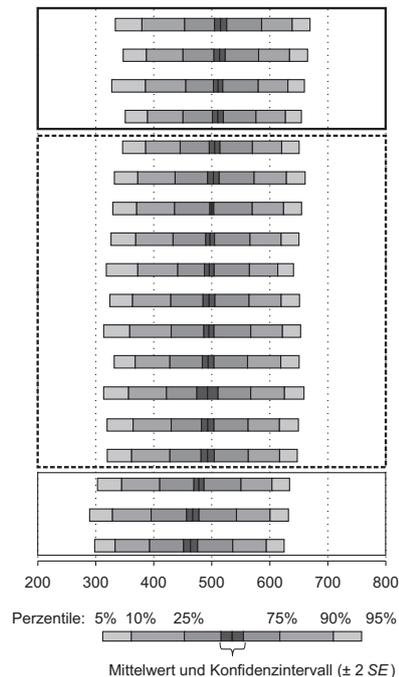
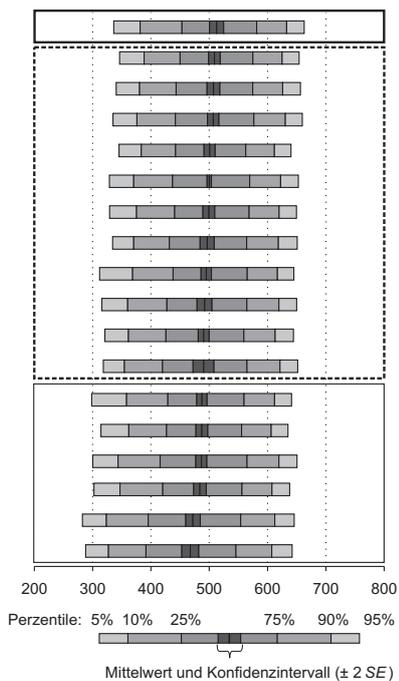


Abbildung 5.2: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich Zuhören

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Bayern	513	(6.0)	100	(2.4)	336	381	453	581	633	663	327
Baden-Württemberg	509	(5.0)	95	(2.6)	346	388	450	575	625	654	308
Schleswig-Holstein	507	(5.5)	97	(3.3)	340	381	443	574	626	656	316
Niedersachsen	507	(4.9)	100	(3.2)	335	376	442	577	631	659	324
Thüringen	501	(4.8)	89	(2.4)	345	383	442	563	612	640	295
Deutschland	500	(1.9)	100	(1.0)	329	371	437	570	622	653	324
Sachsen	499	(5.1)	102	(2.6)	329	375	441	568	620	650	321
Nordrhein-Westfalen	496	(6.0)	97	(2.5)	334	370	432	564	618	651	317
Mecklenburg-Vorpommern	495	(4.4)	105	(4.0)	312	368	438	565	617	645	333
Hessen	492	(6.4)	106	(4.8)	316	360	427	565	620	650	334
Brandenburg	490	(4.7)	98	(2.7)	321	362	426	559	613	645	324
Großstädte ¹	490	(9.0)	104	(5.5)	318	354	420	565	622	652	333
Sachsen-Anhalt	487	(4.4)	107	(3.6)	298	358	429	560	612	641	343
Rheinland-Pfalz	487	(5.1)	98	(3.1)	314	362	427	556	606	635	321
Hamburg	486	(4.8)	109	(3.8)	300	344	416	565	620	651	350
Saarland	484	(5.5)	101	(3.1)	302	347	420	556	608	638	336
Berlin	472	(6.2)	112	(2.6)	283	323	395	554	613	646	363
Bremen	467	(7.3)	109	(3.2)	288	327	391	546	607	642	354

☐ Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
 ☐ Länder liegen nicht signifikant vom deutschen Mittelwert ab.
 ☐ Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.



Bayern gelingt, einen signifikant überdurchschnittlichen Kompetenzstand zu erzielen.

Bei der zweiten Gruppe handelt es sich um jene Länder, in denen im Mittel Kompetenzausprägungen erreicht werden, die sich nicht statistisch signifikant vom deutschen Mittelwert unterscheiden. Die Mittelwerte dieser Länder liegen zwar selten bei einem Wert von genau 500 Punkten, weichen aber in so geringem Umfang nach oben oder unten vom deutschen Mittelwert ab, dass die Differenzen nicht statistisch bedeutsam sind. In diese zweite Gruppe fallen sowohl im Kompetenzbereich Lesen als auch im Kompetenzbereich Zuhören 9 der 16 Länder in der Bundesrepublik Deutschland.

In der dritten Ländergruppe werden jene Länder zusammengefasst, deren Ergebnisse in statistisch bedeutsamem Ausmaß *unterhalb* des deutschen Mittelwerts von 500 Punkten liegen. Im Kompetenzbereich Lesen handelt es sich hierbei um die drei Stadtstaaten Hamburg, Berlin und Bremen. Es bestätigt sich also erneut der sowohl für den Primarbereich als auch für die Sekundarstufe I in früheren Studien beobachtete Befund, dass die in den Stadtstaaten erreichten mittleren Lesekompetenzen recht deutlich unterhalb der in den Flächenländern erzielten Kompetenzstände liegen (Bos et al., 2008; Köller, Knigge & Tesch, 2010). Vergleicht man die in den drei Stadtstaaten erreichten Kompetenzniveaus untereinander, zeichnet sich ein leichter Vorsprung Hamburgs vor den Ländern Berlin und Bremen ab, wo im Mittel sehr ähnliche Kompetenzausprägungen erzielt werden. Der Vorsprung Hamburgs ist statistisch jedoch nicht signifikant. Ergänzend können die in den Stadtstaaten erreichten Kompetenzausprägungen auch mit dem Kompetenzmittelwert der Viertklässlerinnen und Viertklässler in anderen Großstädten verglichen werden. Dieser beträgt im Kompetenzbereich Lesen 493 Punkte und unterscheidet sich nicht statistisch bedeutsam vom deutschen Mittelwert. Im Vergleich liegt die mittlere Lesekompetenz von Schülerinnen und Schülern aus Hamburg 15 Punkte unter den in Großstädten im Mittel erreichten Kompetenzwerten. Für das Land Bremen beträgt diese Differenz 30 Punkte.

Etwas anders stellt sich die Befundlage für den Kompetenzbereich Zuhören dar. Hier befinden sich neben den drei Stadtstaaten auch die Flächenländer Rheinland-Pfalz, das Saarland sowie Sachsen-Anhalt – insgesamt also sechs Länder – in der Gruppe derjenigen Länder, in denen die mittlere Kompetenzausprägung statistisch signifikant unterhalb des deutschen Mittelwerts liegt. Im Zuhören erzielen also deutlich mehr Länder als im Lesen Kompetenzstände unterhalb der in Deutschland im Mittel erbrachten Leistung.

Als globales Maß der Kompetenzunterschiede zwischen den verschiedenen Ländern kann die Spannweite der Ländermittelwerte herangezogen werden. Diese Spannweite entspricht der Differenz zwischen dem höchsten und dem niedrigsten Landesmittelwert innerhalb eines Kompetenzbereichs. Die höchsten Mittelwerte werden sowohl im Lesen als auch im Zuhören in Bayern, die niedrigsten Mittelwerte in Bremen erzielt. Im Bereich Lesen beträgt die Spannweite zwischen diesen beiden Ländern 52 Kompetenzpunkte, im Bereich Zuhören 46 Punkte.

Um die Bedeutung dieser Unterschiede einzuschätzen, ist es hilfreich, als Maßstab die bei Grundschulkindern im Durchschnitt zu erwartenden Leistungszuwächse innerhalb eines Schuljahres heranzuziehen. Da bislang keine echten Längsschnittuntersuchungen der Lernentwicklung von Grundschulkindern auf Basis der länderübergreifenden Bildungsstandards durchgeführt worden sind, müssen hierfür querschnittliche Daten zu den Kompetenzständen von

Schülerinnen und Schülern verschiedener Jahrgangsstufen herangezogen werden. In den Studien des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) zur Pilotierung und Normierung der Bildungsstandards im Primarbereich (vgl. Granzer et al., 2009) lag der querschnittlich ermittelte Kompetenzunterschied zwischen der dritten und der vierten Jahrgangsstufe sowohl für das Lesen als auch für das Zuhören bei $d = 0.60$ (Bremerich-Vos & Böhme, 2009; Behrens, Böhme & Krelle, 2009). Demnach entspricht die Spannweite der Ländermittelwerte von 52 beziehungsweise 46 Punkten einem Kompetenzunterschied von jeweils etwa einem drei viertel Schuljahr.

Betrachtet man den *Zusammenhang* der im Mittel in den Ländern erreichten Kompetenzstände im Lesen und Zuhören, ergibt sich eine Korrelation von $r = .80$. Auf Länderebene gehen also zum Beispiel überdurchschnittliche Kompetenzen im Lesen tendenziell mit ebenfalls überdurchschnittlichen Kompetenzen im Zuhören einher. Dass im Einzelfall aber auch deutliche Unterschiede in den mittleren Kompetenzständen in diesen beiden Bereichen vorliegen können, belegt das Beispiel Sachsen-Anhalt. Während den Kindern dort offenbar im Bereich Lesen Lerngelegenheiten zur Verfügung stehen, die es ihnen gestatten, überdurchschnittliche Lesekompetenzen zu entwickeln, weist dieselbe Schülerschaft im Zuhören Kompetenzausprägungen auf, die statistisch signifikant unterhalb des deutschen Mittelwerts liegen.

5.1.2 Heterogenität der im Fach Deutsch in den Ländern erreichten Kompetenzstände

Während im letzten Abschnitt die in den einzelnen Ländern *im Mittel* erzielten Kompetenzstände in den Bereichen Lesen und Zuhören dargestellt und verglichen wurden, sollen in diesem Abschnitt die *Leistungsstreuungen* näher betrachtet werden. Die Streuung ist insofern relevant, als nicht nur das im Mittel erreichte Kompetenzniveau Auskunft über die Bildungsqualität eines Landes gibt, sondern gleichzeitig berücksichtigt werden sollte, inwieweit das schulische System allen Schülerinnen und Schülern gerecht wird. Im Idealfall gelingt es, durch schulische Lerngelegenheiten mittlere Kompetenzausprägungen auf hohem Niveau bei gleichzeitig geringer Streuung der Kompetenzstände zu gewährleisten.

Einen Hinweis darauf, wie stark die Unterschiede im Lesen und Zuhören *innerhalb* der einzelnen Länder ausgeprägt sind, liefert die in den Abbildungen 5.1 und 5.2 aufgeführte Standardabweichung (*SD*). Im Kompetenzbereich Lesen weisen die Länder Baden-Württemberg und Thüringen mit etwas mehr als 90 Punkten eine sehr geringe Streuung auf. Im Zuhören beträgt die Standardabweichung im Freistaat Thüringen sogar nur 89 Punkte. Eine vergleichsweise geringe Streuung der Kompetenzstände innerhalb eines Landes weist auf eine relativ zu den anderen Ländern große Homogenität der Kompetenzstände hin. Die erstrebenswerte Konstellation eines hohen mittleren Kompetenzniveaus bei gleichzeitig geringer Streuung findet sich somit tendenziell in beiden Kompetenzbereichen sowohl in Baden-Württemberg als auch in Thüringen.

Relativ große Streuungen mit Standardabweichungen von deutlich über 100 Punkten finden sich im Kompetenzbereich Lesen in den Ländern Sachsen-Anhalt ($SD = 104$) sowie Berlin ($SD = 105$) und damit sowohl in einem Land, das signifikant über dem deutschen Mittelwert liegt (Sachsen-Anhalt), als auch in einem Land, das ein signifikant unterdurchschnittliches Ergebnis erzielt hat

(Berlin). Im Kompetenzbereich Zuhören weisen deutlich mehr Länder als im Lesen eine große Streuung auf, die im Zuhören zusätzlich noch ausgeprägter ist als im Lesen. Vergleichsweise große Standardabweichungen von mehr als 105 Punkten finden sich in den Ländern Berlin ($SD = 112$), Bremen ($SD = 109$), Hamburg ($SD = 109$), Hessen ($SD = 106$) und Sachsen-Anhalt ($SD = 107$).

Zur Veranschaulichung der Leistungsstreuungen enthalten die Abbildungen 5.1 und 5.2 für die beiden Kompetenzbereiche Lesen und Zuhören ferner die Perzentile der Kompetenzverteilungen für jedes Land. Zusätzlich zur tabellarischen Darstellung der Perzentilwerte 5, 10, 25, 75, 90 und 95 werden diese Informationen jeweils auch in Form von Perzentilbändern grafisch dargeboten.

Perzentile sind nicht nur Indikatoren für die Spannweite der Leistungsverteilung, sondern sie ermöglichen auch einen Vergleich der sehr schwachen und sehr starken Schülerinnen und Schüler zwischen den einzelnen Ländern.⁷ Ein Vergleich der beiden Länder mit dem höchsten und dem geringsten Wert für das 5. Perzentil im Kompetenzbereich Lesen zeigt, dass die Kompetenzausprägung der leistungsschwächsten 5 Prozent der Schülerinnen und Schüler im Land Thüringen bei höchstens 351 Kompetenzpunkten, im Land Berlin hingegen bei nur maximal 289 Kompetenzpunkten liegt (vgl. Abbildung 5.1). Die 5 Prozent leistungsschwächsten Viertklässlerinnen und Viertklässler des Freistaats Thüringen haben also einen Kompetenzvorsprung von 62 Punkten vor den 5 Prozent leistungsschwächsten Schülerinnen und Schülern des Landes Berlin. Dies entspricht in etwa einem mittleren Kompetenzzuwachs von einem Schuljahr (vgl. Bremerich-Vos & Böhme, 2009). Bei der Interpretation dieser Differenz ist allerdings zu berücksichtigen, dass sich die Zusammensetzung der Schülerschaft in den Ländern Thüringen und Berlin beispielsweise im Hinblick auf den Anteil an Schülerinnen und Schüler mit Zuwanderungshintergrund deutlich unterscheidet.

Im Kompetenzbereich Zuhören erzielen die 5 Prozent leistungsschwächsten Schülerinnen und Schüler in Baden-Württemberg Kompetenzstände von maximal 346 Punkten, im Land Berlin liegt dieser Wert bei 283 Punkten. Hier beträgt der Abstand zwischen den 5 Prozent leistungsschwächsten Viertklässlerinnen und Viertklässlern im Vergleich der beiden Länder 63 Kompetenzpunkte. Auch diese Differenz entspricht ungefähr der Schätzung des durchschnittlichen Lernzuwachses, der im Kompetenzbereich Zuhören innerhalb eines Schuljahres erreicht wird (vgl. Behrens et al., 2009).

Am anderen Ende des Leistungsspektrums können die sehr kompetenzstarken Schülerinnen und Schüler verglichen werden, die in den Abbildungen 5.1 und 5.2 durch die Angabe des 95. Perzentils markiert sind. Die höchsten Kompetenzwerte in der Leistungsspitze finden sich sowohl für das Lesen als auch für das Zuhören im Freistaat Bayern. Hier erzielen die 5 Prozent leistungsstärksten Schülerinnen und Schüler im Lesen mindestens 669 Punkte und im Zuhören mindestens 663 Punkte. Im Vergleich dazu schneiden die leistungsstärksten 5 Prozent der Schülerinnen und Schüler im Kompetenzbereich Lesen im Land Bremen mit mindestens 625 Punkten am schlechtesten ab. Im Kompetenzbereich Zuhören erreichen die kompetenzstärksten 5 Prozent der Schülerschaft in Rheinland-Pfalz mindestens 635 Punkte und markieren somit das untere Ende der Kompetenzausprägungen in der Spitzengruppe. Die Differenz zwischen den von den Spitzengruppen in Bayern und Bremen im Lesen erzielten Kompetenzständen

⁷ Ein solcher Vergleich bezieht sich auf konkrete Perzentilwerte und vernachlässigt Unterschiede in den Verteilungsformen im untersten und obersten Kompetenzspektrum der betrachteten Länder.

beträgt 44 Punkte, im Kompetenzbereich Zuhören beträgt der entsprechende Abstand zwischen Bayern und Rheinland-Pfalz 28 Punkte. Im Vergleich liegen somit die besonders leistungsstarken Viertklässlerinnen und Viertklässler im Kompetenzbereich Zuhören bundesweit gesehen etwas näher beieinander als im Lesen.

Betrachtet man den Punktabstand zwischen dem 5. und dem 95. Perzentil, erhält man ein weiteres anschauliches Maß dafür, wie heterogen die Kompetenzausprägungen der Schülerschaft innerhalb eines Landes sind. Diese Angabe findet sich in der jeweils ganz rechten Spalte der Tabellenteile der Abbildungen 5.1 und 5.2. Je kleiner dieser Wert ist, desto homogener sind die in einem Land erreichten Kompetenzstände. Relativ hohe Werte weisen auf eine deutliche Leistungsdifferenz zwischen den stärksten und schwächsten 5 Prozent der Schülerinnen und Schüler innerhalb eines Landes hin. Dies ist vor dem Hintergrund angestrebter Bildungsgerechtigkeit vor allem dann problematisch, wenn ein großer Anteil der Schülerschaft nur über sehr gering ausgeprägte Kompetenzen verfügt. Für den Bereich Lesen (Abbildung 5.1) zeigt der Vergleich der Länder mit dem geringsten und dem größten Abstand zwischen dem 5. und dem 95. Perzentil, dass sich dieser in Thüringen und Baden-Württemberg auf 304, in Berlin hingegen auf 343 Punkte beläuft. Die Schülerschaft im Freistaat Thüringen sowie im Land Baden-Württemberg ist im Hinblick auf ihre Lesekompetenz also jeweils deutlich homogener als die im Stadtstaat Berlin. Im Kompetenzbereich Zuhören beträgt die Spannweite zwischen den 5 Prozent leistungsschwächsten und den 5 Prozent leistungsstärksten Schülerinnen und Schülern in Thüringen sogar weniger als 300 Punkte. Erneut zeigt sich im Stadtstaat Berlin mit einer Differenz von 363 Punkten eine besonders große Leistungsstreuung innerhalb der Schülerschaft.

Insgesamt zeichnet sich als Muster ab, dass die Stadtstaaten den anderen Großstädten zwar in der Heterogenität der erreichten Kompetenzstände ähnlich sind, sich von diesen aber durch ein insgesamt unterdurchschnittliches Kompetenzniveau unterscheiden. Unter den Flächenländern finden sich einerseits Länder, die relative Leistungshomogenität auf hohem Niveau aufweisen (z. B. Thüringen und Baden-Württemberg), andererseits aber auch Länder wie Bayern, die sich bei hohem mittlerem Niveau gleichzeitig mit einer erheblichen Leistungsheterogenität konfrontiert sehen.

5.1.3 Heterogenität in den Kompetenzständen innerhalb und zwischen Schulen

In allen Ländern in der Bundesrepublik Deutschland sind die von den Schülerinnen und Schülern erreichten Kompetenzstände heterogen. Diese Heterogenität kann jedoch unterschiedlich verteilt sein. Unterschiede in den erreichten Kompetenzausprägungen können zum einen zwischen Schülerinnen und Schülern bestehen, die *dieselbe* Schule besuchen, und zum anderen zwischen Schülerinnen und Schülern, die *unterschiedliche* Schulen besuchen (vgl. z. B. Baumert, Watermann & Stanat, 2006). Wie die Gesamtvarianz auf diese unterschiedlichen Quellen verteilt ist, kann sich von Land zu Land unterscheiden. Die Untersuchung der Frage, ob ein hoher Anteil der Varianz innerhalb oder zwischen Schulen liegt, ist relevant, da je nach Verteilungsmuster unterschiedliche Strategien der Qualitätsentwicklung sinnvoll sein können (Willms, 2006). So würde man beispielsweise bei einem erheblichen Varianzanteil *zwischen* Schulen primär solche

Unterstützungsstrategien ins Auge fassen, die auf gelingende Schulentwicklung im Sinne definierter Standards und Kriterien der Schul- und Unterrichtsqualität abzielen und dabei vor allem Schulen in den Blick nehmen, die besondere Herausforderungen zu bewältigen haben (Pietsch, 2010). Wäre dagegen die Leistungsvarianz *innerhalb* der Schulen sehr groß, wären eher Maßnahmen zum verbesserten Umgang mit individueller (Leistungs-)Heterogenität und zur individuellen Förderung, etwa durch diagnosebasierte Sprach- und Leseförderung, angezeigt. Beide Ansatzpunkte der Qualitätsentwicklung schließen einander selbstverständlich nicht aus.

Liegt ein Großteil der Varianz zwischen den Schulen eines Landes, können diese Unterschiede auf Schulebene durch verschiedene Faktoren bedingt sein, von denen nachfolgend drei mögliche erläutert werden.

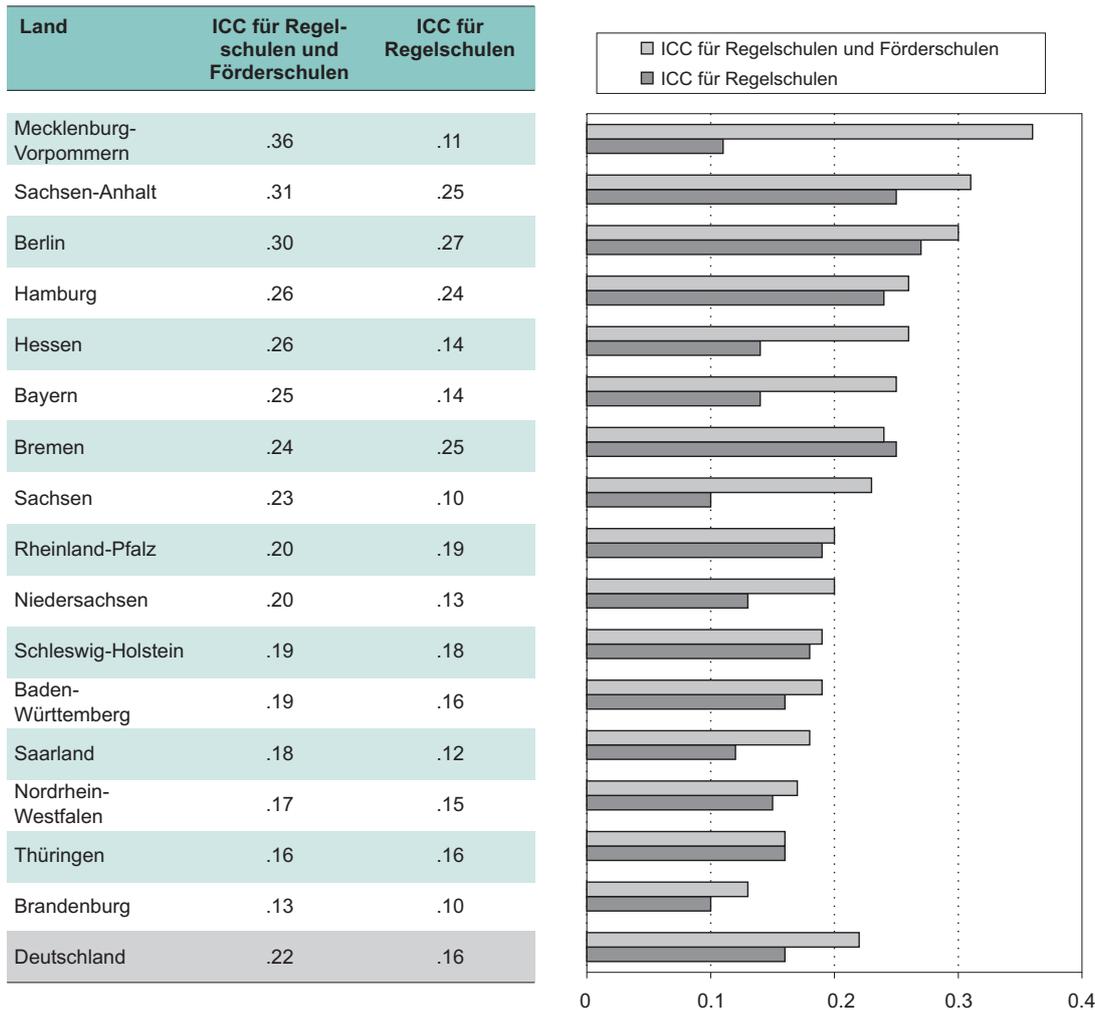
Ein wichtiger Faktor besteht darin, dass leistungsstärkere und leistungsschwächere Kinder mitunter von Beginn ihrer Schullaufbahn an nicht die gleichen Schulen besuchen. In diesem Fall verteilen sich Schülerinnen und Schüler mit unterschiedlichen Eingangsvoraussetzungen ungleichmäßig auf die Schulen eines Landes und die Unterschiedlichkeit zwischen Schulen basiert auf *Selektions-effekten*. Insbesondere in den Stadtstaaten und anderen Großstädten sind die Einzugsgebiete der untersuchten Schulen hinsichtlich ihrer sozialen Zusammensetzung oft sehr unterschiedlich, was dazu führt, dass auch die Zusammensetzung der Schülerschaft, etwa hinsichtlich des Anteils an Schülerinnen und Schülern mit nicht deutscher Herkunftssprache oder aus sozial schwachen Familien, stark variiert. Die Leistungsheterogenität zwischen Schulen kann in solchen Fällen also zu einem erheblichen Teil durch die unterschiedlich zusammengesetzte Schülerschaft bedingt sein, die diese Schulen besucht.

Auch die Entscheidung, ob Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Regelschulen oder in Förderschulen unterrichtet werden, trägt zu Leistungsunterschieden auf Schulebene bei. In den einzelnen Ländern werden in dieser Hinsicht zum Teil sehr unterschiedliche Strategien verfolgt. Daher variiert der Anteil der Kinder mit sonderpädagogischem Förderbedarf, die an Regelschulen unterrichtet werden, mitunter deutlich zwischen den einzelnen Ländern (vgl. Kapitel 4, Tabelle 4.2).

Eine dritte wichtige Ursache für Kompetenzunterschiede zwischen Schulen sind heterogene Lernbedingungen beziehungsweise Unterschiede in den zur Verfügung gestellten Lerngelegenheiten und damit in der Schul- und Unterrichtsqualität. Manchen Schulen gelingt es besser als anderen, den Unterricht so zu gestalten, dass der Kompetenzerwerb der Grundschul Kinder effektiv gefördert wird.

Um zu bestimmen, wie sich die Heterogenität der Kompetenzstände der Schülerinnen und Schüler in den einzelnen Ländern verteilt, wurden *Intra-klassenkorrelationen* (*Intra-Class-Correlation, ICC*) berechnet. Eine Intra-klassenkorrelation kann Werte zwischen 0 und 1 annehmen und gibt den auf die Kompetenzausprägung bezogenen Varianzanteil an, der innerhalb eines Landes zwischen einzelnen Schulen besteht. Eine hohe Intraklassenkorrelation impliziert somit, dass ein großer Teil der Unterschiede in den Kompetenzständen der Schülerinnen und Schüler *zwischen* den einzelnen Schulen dieses Landes liegt.

Wie bereits erwähnt, unterscheiden sich die Länder darin, inwieweit Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf primär an Regelschulen oder an Förderschulen unterrichtet werden. Dies beeinflusst, welche Varianzanteile innerhalb und zwischen Schulen liegen. Um auch zu bestimmen, wie die Heterogenität von Schülerinnen und Schülern im Bereich der Regelschulen verteilt ist, werden die Intraklassenkorrelationen in den Abbildungen 5.3 und 5.4

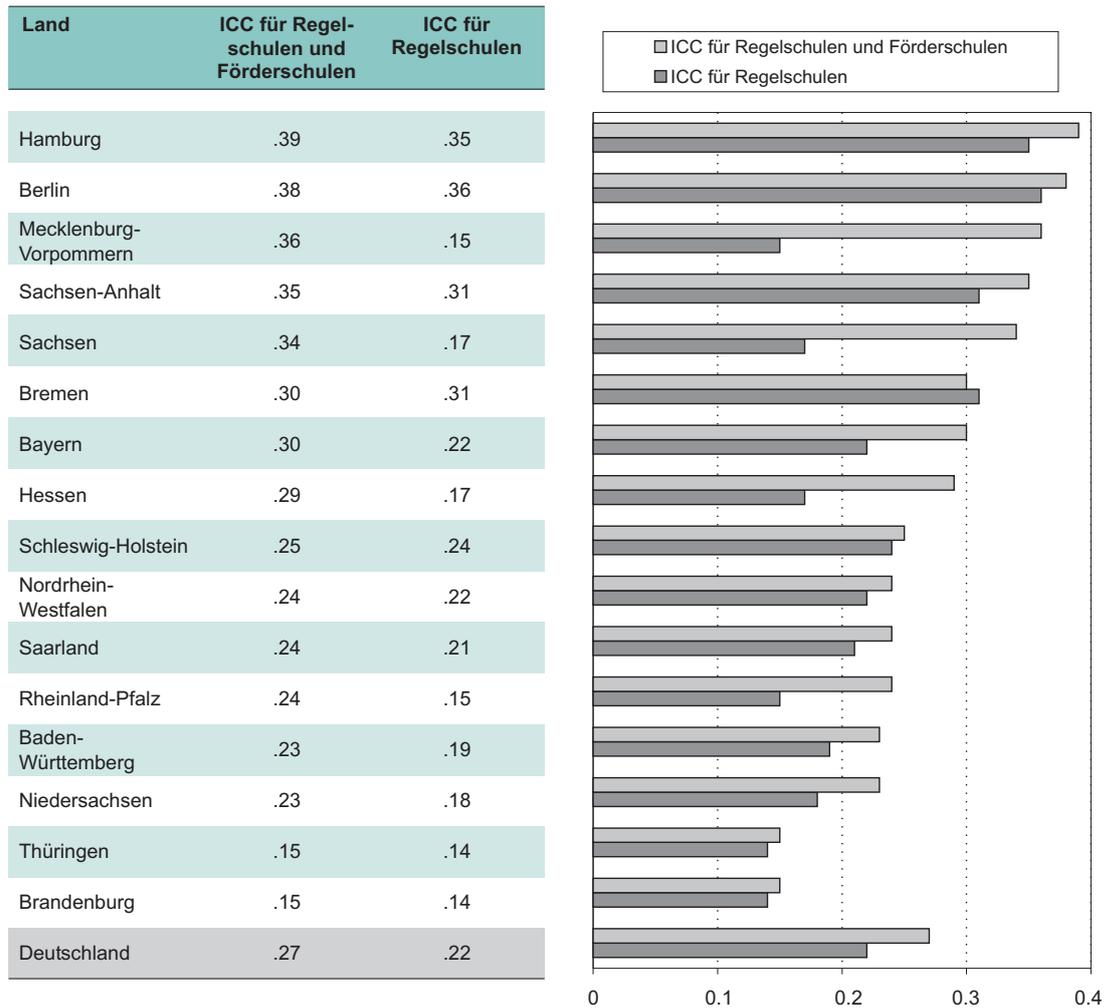
Abbildung 5.3: Intraklassenkorrelationen (ICC) für den Bereich Lesen insgesamt (Regel- und Förderschulen) sowie ausschließlich für Regelschulen

Anmerkung. Die Länder sind absteigend nach der ICC für Regelschulen und Förderschulen geordnet.

zum einen für alle Schülerinnen und Schüler und zum anderen nur für die an Regelschulen unterrichteten Schülerinnen und Schüler berichtet.

Für Deutschland insgesamt zeigt sich, dass 22 Prozent der Kompetenzunterschiede im Bereich Lesen zwischen den untersuchten Schulen liegen (vgl. Abbildung 5.3). Betrachtet man ausschließlich die Kompetenzstände von Schülerinnen und Schülern an Regelschulen, verringert sich dieser Anteil auf 16 Prozent. Im Kompetenzbereich Zuhören liegt der Anteil der Kompetenzunterschiede zwischen Schulen etwas höher und beträgt 27 Prozent für die gesamte Schülerschaft in der vierten Jahrgangsstufe und 22 Prozent für Schülerinnen und Schüler an Regelschulen (vgl. Abbildung 5.4). Für die beiden im Fach Deutsch untersuchten Kompetenzbereiche gilt folglich, dass der Großteil der Varianz nicht zwischen Schulen, sondern zwischen Schülerinnen und Schülern innerhalb von Schulen liegt.

Im Kompetenzbereich Lesen finden sich bei Berücksichtigung aller Schülerinnen und Schüler die höchsten Intraklassenkorrelationen in den Ländern Mecklenburg-Vorpommern, Sachsen-Anhalt und Berlin. Für Schülerinnen und Schüler an Regelschulen liegen in den drei Stadtstaaten sowie im Land Sachsen-Anhalt die größten Varianzanteile zwischen Schulen. Diese sind aber insgesamt et-

Abbildung 5.4: Intraklassenkorrelationen (ICC) für den Bereich Zuhören insgesamt (Regel- und Förderschulen) sowie ausschließlich für Regelschulen

Anmerkung. Die Länder sind absteigend nach der ICC für Regelschulen und Förderschulen geordnet.

was niedriger als die Intraklassenkorrelationen für die gesamte Schülerschaft. Besonders geringe Intraklassenkorrelationen finden sich sowohl für Schülerinnen und Schüler an Regelschulen als auch für die gesamte Schülerschaft in den Ländern Brandenburg und Thüringen. In diesen beiden Ländern sind sich also die Primarschulen in dem Sinne relativ ähnlich, dass die Heterogenität in den Kompetenzständen nur zu 10 bis circa 15 Prozent auf Unterschiede zwischen verschiedenen Schulen zurückzuführen ist. Hier scheinen also weder Selektivitätseffekte noch Unterschiede in der Schulqualität der beteiligten Schulen eine größere Rolle zu spielen. Letzteres ist insbesondere dann wünschenswert, wenn das im Mittel erreichte Kompetenzniveau hoch ist. Dies ist im Freistaat Thüringen eher der Fall als im Land Brandenburg.

Die Intraklassenkorrelationen, die für den Kompetenzbereich Zuhören gefunden wurden, liegen insgesamt etwas höher als die für den Bereich Lesen ermittelten Ergebnisse. Demnach ist die Heterogenität der Kompetenzstände im Bereich Zuhören also etwas stärker als im Lesen zwischen den Schulen angesiedelt. Dieser Befund dürfte unter anderem durch die Art und Weise der Testdurchführung im Zuhören begründet sein, da immer alle Schülerinnen und Schüler einer Klasse dieselben Zuhöraufgaben bearbeitet haben (vgl. Kapitel 12). Daher

sollten die Unterschiede zwischen den ICCs in den beiden Kompetenzbereichen nicht inhaltlich interpretiert werden.

Unter Einbeziehung aller Schülerinnen und Schüler finden sich auch im Kompetenzbereich Zuhören in den Ländern Berlin, Mecklenburg-Vorpommern und Sachsen-Anhalt vergleichsweise hohe ICCs. Ähnliches gilt für die Länder Hamburg und Sachsen. Besonders gering fallen die Intraklassenkorrelationen im Bereich Zuhören wiederum für die Länder Brandenburg und Thüringen aus.

Die Abbildungen 5.3 und 5.4 verdeutlichen, dass sich die für alle Schülerinnen und Schüler ermittelten Intraklassenkorrelationen oft nur geringfügig von denen unterscheiden, die ausschließlich für Kinder an Regelschulen identifiziert wurden. Abweichend hiervon fallen in einigen Ländern die Intraklassenkorrelationen für Kinder an Regelschulen jedoch deutlich kleiner aus als bei der Gesamtschülerschaft. Dies trifft insbesondere auf solche Länder zu, die einen vergleichsweise hohen Anteil der Kinder mit sonderpädagogischem Förderbedarf nicht in Regelschulen, sondern an Förderschulen unterrichten, so zum Beispiel die Länder Mecklenburg-Vorpommern oder Sachsen. Die Unterschiede zwischen Schulen bestehen in diesen Ländern also vorrangig zwischen Regel- und Förderschulen. In einigen anderen Ländern unterscheiden sich die Intraklassenkorrelationen für Schülerinnen und Schüler an Regelschulen von denen für die gesamte Schülerschaft, obwohl der Anteil der an Förderschulen unterrichteten Kinder nur im mittleren Bereich liegt. Dies gilt im Kompetenzbereich Lesen beispielsweise für das Land Niedersachsen und im Kompetenzbereich Zuhören für die Länder Rheinland-Pfalz und Hessen. Um den Ursachen dieses Befundes auf den Grund zu gehen, wären vertiefende Analysen erforderlich.

In den Stadtstaaten Berlin, Bremen und Hamburg in denen im Mittel unterdurchschnittliche Kompetenzstände ermittelt wurden, liegt ein relativ hoher Anteil der Varianz zwischen Schulen. Die hohen ICCs zwischen .25 und .40 dürften unter anderem auf deutliche Unterschiede in der Zusammensetzung der Schülerschaft an den verschiedenen Schulen zurückzuführen sein.

5.1.4 Einordnung der Befunde in die Ergebnisse bisheriger Ländervergleiche (IGLU-E 2006)

Im Ländervergleich 2011 werden die in den einzelnen Ländern erzielten Kompetenzstände der Schülerinnen und Schüler des Primarbereichs erstmals auf der Basis von Tests miteinander verglichen, die auf den Bildungsstandards basieren. Somit kann auch erstmalig überprüft werden, inwieweit die Schülerinnen und Schüler der Primarstufe die Bildungsstandards in den Fächern Deutsch und Mathematik erreichen (vgl. Kapitel 6).

Bislang erfolgten die Vergleiche der in den Ländern erreichten Kompetenzstände in nationalen Ergänzungen internationaler Vergleichsstudien wie der Internationalen Grundschul-Lese-Untersuchung (IGLU) der Jahre 2001 und 2006 (Bos et al., 2008; Bos, Lankes, Prenzel & Schwippert, 2003). Im Jahr 2001 nahmen an der als IGLU-E bezeichneten nationalen Erweiterung der IGLU-Studie lediglich 5 der 16 Länder der Bundesrepublik Deutschland teil. Im Jahr 2006 wurden alle Länder in die Erweiterung der Erhebung einbezogen. Die Ergebnisse der IGLU-Ergänzungsstudien sind jedoch mit den hier berichteten Befunden des IQB-Ländervergleichs 2011 nicht direkt vergleichbar, da zwischen beiden Studien im Hinblick auf die Definition der interessierenden Kompetenzkonstrukte, die Testkonstruktion und die Berichtsmetrik einige Unterschiede bestehen (vgl. hier-

zu Pietsch, Böhme, Robitzsch & Stubbe, 2009). Ferner ist die Datengrundlage im Ländervergleich 2011 robuster als in IGLU-E 2006. In IGLU-E wurden kleinere Stichproben gezogen, was im Vergleich zur aktuellen Studie zu etwas größeren Standardfehlern und einer größeren Unsicherheit der berichteten Ergebnisse führte (vgl. Valtin, Bos, Buddeberg, Goy & Potthoff, 2008, Abbildung IV.7, S. 61). Dies betrifft insbesondere die Spannweite und die absolute Höhe der größten ermittelten Standardfehler. Aufgrund der genannten Unterschiede ist keine Aussage darüber möglich, inwieweit sich die mittleren Kompetenzen der Schülerinnen und Schüler in Deutschland insgesamt oder in den einzelnen Ländern seit 2006 verändert haben. Solche Tendaussagen werden erst möglich sein, wenn 2016 der IQB-Ländervergleich für die Primarstufe wiederholt wird.

Möchte man dennoch eine Einordnung der hier vorgelegten Befunde in die Ergebnisse bisheriger Studien vornehmen, können die im Ländervergleich 2011 und die im Rahmen von IGLU-E 2006 für den Kompetenzbereich Lesen ermittelten Länderrangreihen gegenübergestellt werden. Da die Messwerte inhaltlich nicht vergleichbar sind, sollte sich der Vergleich auf die Frage beschränken, welche Länder 2006 und 2011 den drei oben erläuterten Gruppen angehörten und sich entweder positiv, negativ oder nicht statistisch bedeutsam vom deutschen Mittelwert unterschieden. Hierbei ist zu berücksichtigen, dass aufgrund der unterschiedlich großen Standardfehler die Wahrscheinlichkeit, dass Unterschiede als signifikant ausgewiesen werden, im IQB-Ländervergleich etwas größer ist, als sie es 2006 im Rahmen von IGLU-E war.

In IGLU-E 2006 wurden im Lesen nur im Freistaat Thüringen Kompetenzstände erzielt, die signifikant oberhalb des damaligen deutschen Mittelwerts lagen. Im Ländervergleich 2011 reißen sich im Kompetenzbereich Lesen zusätzlich die Länder Bayern, Sachsen und Sachsen-Anhalt in diese Spitzengruppe ein. Dieser Befund ist allerdings nicht überraschend, da die Länder Bayern, Sachsen und Sachsen-Anhalt schon 2006 Spitzenplätze in der Länderrangreihe belegten, sich damals aber noch nicht statistisch bedeutsam vom deutschen Mittelwert abheben konnten.

Die Gruppe derjenigen Länder, die Kompetenzausprägungen im mittleren Bereich aufweisen, war damals wie heute recht groß. In IGLU-E 2006 gehörten dieser Gruppe 13 der 16 Länder an, im Ländervergleich 2011 fallen 9 Länder in diese mittlere Kategorie. Da die Kompetenzausprägungen in diesen Ländern nicht nur sehr dicht am deutschen Mittelwert liegen, sondern sich auch untereinander oft nur geringfügig unterscheiden, gibt es in dieser zweiten Gruppe einige Vertauschungen in den Rangplätzen, die jedoch nicht bedeutsam sind. Festzuhalten ist hier aber der Befund, dass sich die Länder Baden-Württemberg, Brandenburg, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Schleswig-Holstein, Rheinland-Pfalz sowie das Saarland sowohl 2006 als auch 2011 nicht statistisch signifikant von den im Mittel in Deutschland erreichten Lesekompetenzwerten unterschieden.

Im Jahr 2006 wurde im Rahmen von IGLU-E ermittelt, dass die Lesekompetenz in den Stadtstaaten Bremen und Hamburg unterdurchschnittlich ausgeprägt war. Dieser Befund wurde im Ländervergleich 2011 repliziert. Zusätzlich erwiesen sich die im Land Berlin erreichten Lesekompetenzstände im Jahr 2011 als signifikant unterdurchschnittlich. Auch dieser Befund ist plausibel, da im Stadtstaat Berlin bereits im Jahr 2006 eher schwache Lesekompetenzen diagnostiziert wurden, die sich aber damals – möglicherweise aufgrund des großen Standardfehlers – nicht statistisch signifikant vom deutschen Mittelwert unterschieden.

Insgesamt ergibt der Vergleich der Länderrangreihen für den Kompetenzbereich Lesen aus den Jahren 2006 und 2011 also eine weitgehende Stabilität der Ergebnisse. Lediglich für die Länder, in denen die Schülerinnen und Schüler sehr gute und eher schwache Lesekompetenzen aufweisen, ergeben sich leichte Veränderungen in dem Sinne, dass sich die dort erzielten Ergebnisse im Ländervergleich 2011 für eine größere Zahl von Ländern signifikant vom deutschen Mittelwert unterscheiden. Dies dürfte jedoch auf die größeren Stichprobenumfänge im Ländervergleich 2011 und die hiermit einhergehenden geringeren Standardfehler zurückzuführen sein, die dazu führen, dass die Unterschiede zum deutschen Mittelwert eher als statistisch signifikant ausgewiesen werden.

5.2 Der Ländervergleich im Fach Mathematik

Nicole Haag und Alexander Roppelt

Dieses Teilkapitel beschreibt die Kompetenzstände, die in den Ländern der Bundesrepublik Deutschland im Fach Mathematik erreicht wurden. Zu ihrer Erfassung wurden Aufgaben zu allen in den Bildungsstandards (KMK, 2005) beschriebenen inhaltlichen und allgemeinen Kompetenzbereichen eingesetzt. Die Ergebnisse werden sowohl für die einzelnen inhaltlichen Kompetenzbereiche als auch für eine globale Mathematikskala dargestellt, die die Aufgaben aus allen Bereichen zusammenfasst. Eine genauere Beschreibung der Struktur der inhaltlichen Kompetenzbereiche sowie der globalen Mathematikskala ist in Kapitel 2.2 zu finden.

In Abschnitt 5.2.1 werden zunächst die durchschnittlich erreichten Kompetenzen der Schülerinnen und Schüler in den Ländern präsentiert, und es wird erläutert, wie sich diese zwischen den Ländern unterscheiden. Ferner werden die Unterschiede zwischen den Kompetenzständen auf der Globalskala und in den inhaltlichen Kompetenzbereichen für die einzelnen Länder dargestellt. Daran anschließend wird in Abschnitt 5.2.2 die Leistungsheterogenität innerhalb der Länder untersucht.

5.2.1 Vergleich der im Fach Mathematik in den Ländern im Mittel erreichten Kompetenzstände

Tabelle 5.2 enthält die Mittelwerte der von den Schülerinnen und Schülern in den 16 Ländern erreichten Kompetenzen im Fach Mathematik insgesamt (Globalskala) sowie in den fünf inhaltlichen Kompetenzbereichen. Zusätzlich weist die Tabelle die Standardfehler (SE) der Mittelwerte (M) aus. Die Standardfehler sind ein Maß für die Unsicherheit, die mit statistischen Kennwerten notwendigerweise verbunden ist, wenn sie aus einer stichprobenbasierten Untersuchung stammen (siehe auch Abschnitt 5.1.1 sowie Kapitel 12).

Die Mittelwerte werden auf einer Metrik berichtet, die für alle Viertklässlerinnen und Viertklässler in Deutschland einen Mittelwert von 500 und eine Standardabweichung von 100 Punkten aufweist (vgl. Kapitel 3.1). Länder, deren Mittelwert statistisch signifikant vom deutschen Mittelwert 500 abweicht, sind in der Tabelle fett gedruckt.

Um den Vergleich der in den einzelnen Ländern der Bundesrepublik Deutschland erzielten Kompetenzausprägungen sowohl auf der Globalskala als auch in den einzelnen untersuchten Kompetenzbereichen zu erleichtern, sind die Ergebnisse in den Abbildungen 5.5 bis 5.10 in Form von Perzentilbändern grafisch veranschaulicht. Diese enthalten neben den in Tabelle 5.2 dargestellten Mittelwerten und ihren Standardfehlern zusätzlich die für die Länder ermittelten Standardabweichungen und Perzentile. In diesen Abbildungen sind die Länder absteigend nach ihrem Mittelwert angeordnet. Verschiedene Umrandungen in der Grafik auf der rechten Seite geben an, ob die in den Ländern erreichten mittleren Kompetenzstände signifikant vom deutschen Mittelwert abweichen.

Für die Globalskala der mathematischen Kompetenz lässt sich eine eng zusammenliegende Gruppe von vier Ländern identifizieren, in denen die mittleren Kompetenzstände der Schülerinnen und Schüler signifikant über dem deutschen Mittelwert liegen (siehe Abbildung 5.5). Sie umfasst die Länder Bayern, Sachsen,

Tabelle 5.2: In den Ländern im Mittel erzielte Kompetenzstände im Fach Mathematik

Land	Global		Zahlen und Operationen		Raum und Form		Muster und Strukturen		Größen und Messen		Daten, Häufigkeit und Wahrscheinlichkeit	
	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>
Baden-Württemberg	512	(4.9)	510	(4.6)	510	(5.4)	506	(4.7)	518	(4.6)	511	(4.5)
Bayern	519	(5.1)	515	(5.1)	538	(4.8)	516	(5.0)	509	(5.3)	510	(4.8)
Berlin	451	(4.9)	451	(4.6)	459	(4.7)	455	(4.4)	448	(5.2)	461	(4.5)
Brandenburg	491	(3.8)	483	(3.5)	505	(3.7)	481	(3.7)	488	(3.8)	494	(4.0)
Bremen	452	(7.2)	454	(7.0)	464	(6.7)	459	(6.3)	448	(7.7)	459	(6.8)
Hamburg	470	(4.3)	474	(4.3)	476	(3.9)	481	(4.1)	464	(4.1)	475	(4.3)
Hessen	484	(4.4)	486	(4.1)	476	(5.1)	484	(4.1)	491	(4.3)	485	(4.5)
Mecklenburg-Vorpommern	494	(5.1)	489	(5.3)	500	(5.0)	482	(4.6)	497	(4.6)	494	(4.7)
Niedersachsen	496	(4.5)	495	(4.5)	494	(4.3)	497	(4.4)	498	(4.4)	498	(4.5)
Nordrhein-Westfalen	497	(5.0)	505	(4.8)	482	(4.9)	508	(4.8)	498	(5.0)	501	(4.6)
Rheinland-Pfalz	494	(5.2)	489	(5.3)	497	(5.2)	491	(4.6)	496	(5.3)	493	(4.8)
Saarland	492	(5.1)	495	(5.1)	470	(5.1)	492	(5.1)	509	(5.2)	493	(4.7)
Sachsen	517	(4.3)	511	(4.3)	528	(4.9)	506	(4.2)	513	(3.8)	513	(4.6)
Sachsen-Anhalt	517	(4.2)	511	(4.3)	524	(4.2)	500	(4.2)	517	(4.1)	516	(4.2)
Schleswig-Holstein	487	(4.3)	484	(4.2)	487	(4.8)	486	(4.5)	491	(4.3)	490	(4.1)
Thüringen	502	(5.4)	501	(5.2)	504	(5.1)	493	(5.6)	503	(5.6)	500	(5.1)
Großstädte ¹	495	(8.6)	503	(8.6)	493	(8.8)	507	(7.6)	493	(8.1)	496	(8.7)
Deutschland	500	(1.7)	500	(1.7)	500	(1.7)	500	(1.6)	500	(1.7)	500	(1.5)

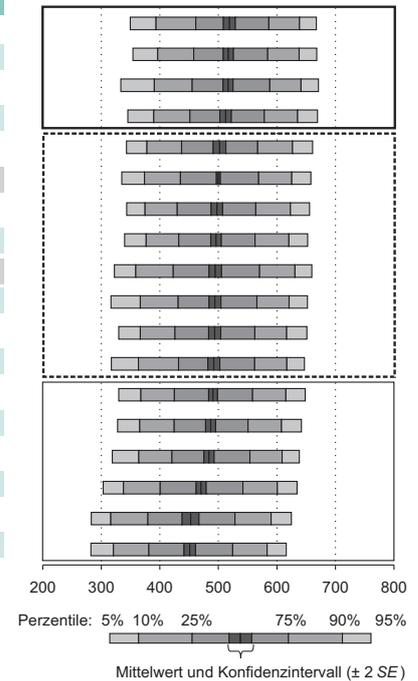
Anmerkungen. Fett gedruckte Mittelwerte unterscheiden sich im jeweiligen Kompetenzbereich signifikant vom deutschen Mittelwert; *M* = Mittelwert; *SE* = Standardfehler. ¹ Ohne Berlin, Bremen und Hamburg.

Sachsen-Anhalt und Baden-Württemberg. Mit gewissem Abstand folgt eine Mittelgruppe von sechs Ländern, deren Kompetenzmittelwerte sich nicht signifikant vom deutschen Mittelwert unterscheiden. Daran schließen sich die Länder Brandenburg, Schleswig-Holstein und Hessen sowie die Stadtstaaten Hamburg, Bremen und Berlin an, deren Ergebnisse statistisch bedeutsam unterhalb des deutschen Mittelwerts liegen. Auch im Vergleich zu den Schülerinnen und Schülern, die in anderen Großstädten zur Schule gehen, erreichen Schülerinnen und Schüler in den Stadtstaaten im Mittel signifikant niedrigere Kompetenzausprägungen. Es scheint in den Stadtstaaten also deutlich schwieriger zu sein, Lerngelegenheiten so zu gestalten, dass im Mittel ähnliche Kompetenzausprägungen erreicht werden wie in anderen Ländern und Großstädten. Innerhalb der Stadtstaaten besteht allerdings ein deutlicher Unterschied von 18 beziehungsweise 19 Punkten zwischen den in Hamburg erzielten Werten einerseits und den in Bremen sowie Berlin erzielten Werten andererseits.

Eine ländervergleichende Untersuchung wurde für die Mathematikkompetenz in der Grundschule bislang in Deutschland noch nicht durchgeführt. Jedoch decken sich die vergleichsweise niedrigen Kompetenzwerte in den Stadtstaaten im Bereich der mathematischen Kompetenz mit den Befunden von PISA 2006 in der Sekundarstufe I (Frey, Asseburg, Ehmke & Blum, 2008). In Abweichung zu PISA 2006 gehört Berlin jedoch im Ländervergleich 2011 nicht zur Durchschnittsgruppe, sondern zu den Ländern, in denen die Kompetenzstände signifikant unter dem deutschen Mittelwert liegen.

Abbildung 5.5: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Fach Mathematik (Globalskala)

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Bayern	519	(5.1)	99	(2.2)	349	393	461	586	639	667	318
Sachsen	517	(4.3)	100	(1.9)	354	396	458	584	638	668	314
Sachsen-Anhalt	517	(4.2)	106	(2.9)	333	391	455	588	641	671	338
Baden-Württemberg	512	(4.9)	99	(2.8)	345	389	451	578	635	669	324
Thüringen	502	(5.4)	97	(2.5)	343	378	437	567	627	661	318
Deutschland	500	(1.7)	100	(0.8)	335	374	435	569	625	658	323
Nordrhein-Westfalen	497	(5.0)	96	(2.1)	343	375	430	564	623	656	312
Niedersachsen	496	(4.5)	98	(4.0)	340	377	432	563	620	653	313
Großstädte ¹	495	(8.6)	106	(6.8)	322	359	423	570	631	660	337
Mecklenburg-Vorpommern	494	(5.1)	106	(3.6)	317	367	431	566	621	652	336
Rheinland-Pfalz	494	(5.2)	99	(3.1)	330	367	426	562	617	651	321
Saarland	492	(5.1)	105	(2.9)	317	363	432	562	617	647	330
Brandenburg	491	(3.8)	98	(3.0)	330	368	425	558	615	648	319
Schleswig-Holstein	487	(4.3)	95	(2.3)	328	366	424	551	608	642	314
Hessen	484	(4.4)	102	(2.6)	319	364	421	554	609	638	320
Hamburg	470	(4.3)	102	(2.7)	303	338	401	542	601	634	331
Bremen	452	(7.2)	106	(2.9)	283	316	379	528	590	625	342
Berlin	451	(4.9)	104	(2.5)	282	321	381	525	583	616	334



 Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
 Länder weichen nicht signifikant vom deutschen Mittelwert ab.
 Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.

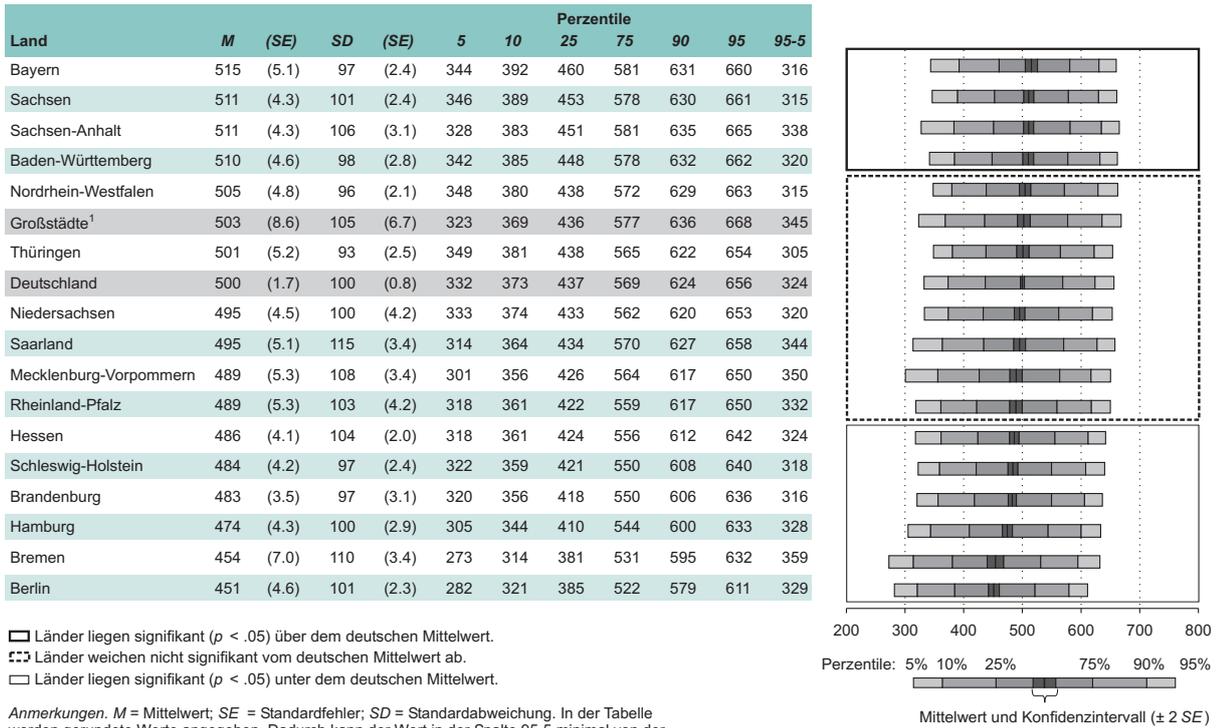
Inhaltliche Kompetenzbereiche

Für die einzelnen inhaltlichen Kompetenzbereiche ergibt sich insgesamt ein sehr ähnliches Bild wie für die Globalskala (vgl. Abbildung 5.6 bis 5.10). Dies zeigt sich auch in den hohen Korrelationen zwischen den in den Ländern erzielten mittleren Kompetenzniveaus auf der Globalskala einerseits und in den einzelnen Kompetenzbereichen andererseits. Die Korrelationen mit der Globalskala auf Länderebene reichen von $r = .88$ für *Raum und Form* bis $r = .99$ für *Daten, Häufigkeit und Wahrscheinlichkeit*. Auch die einzelnen Kompetenzbereiche untereinander zeigen auf Länderebene hohe Korrelationen (zwischen $r = .74$ für *Raum und Form – Größen und Messen* und $r = .97$ für *Zahlen und Operationen – Muster und Strukturen*). Dennoch ergeben sich für einige Kompetenzbereiche differenzielle Befundmuster, die auf relative Stärken und Schwächen innerhalb der Länder hinweisen. In der folgenden Beschreibung der Ergebnisse für die einzelnen Kompetenzbereiche werden diese zu den Befunden auf der Globalskala in Beziehung gesetzt, um auf solche spezifischen Stärken und Schwächen aufmerksam zu machen.

Im Kompetenzbereich *Zahlen und Operationen* sind im Vergleich zur Globalskala einige Plätze in der Rangreihe vertauscht (siehe Abbildung 5.6). Diese Verschiebungen sind jedoch meist auf sehr kleine, statistisch nicht bedeutende Unterschiede in den Mittelwerten zurückzuführen.

Ebenso wie auf der Globalskala werden auch im Kompetenzbereich *Raum und Form* in den Ländern Bayern, Sachsen und Sachsen-Anhalt statistisch überdurchschnittliche Kompetenzstände erreicht, wobei insbesondere Bayern durch weit überdurchschnittliche Werte hervorsteicht (siehe Abbildung 5.7). In Baden-Württemberg wird die Marke eines statistisch signifikant überdurch-

Abbildung 5.6: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich *Zahlen und Operationen*



schnittlichen Ergebnisses nur knapp verfehlt. Im Vergleich zur Globalskala sind die in Brandenburg, Nordrhein-Westfalen und dem Saarland erreichten Kompetenzstände auffällig. Während die Schülerinnen und Schüler in Brandenburg im Bereich *Raum und Form* relativ hohe Kompetenzausprägungen erzielen, fallen die Mittelwerte der anderen beiden Länder in diesem Bereich weniger günstig aus als für die Globalskala und für die anderen inhaltlichen Kompetenzbereiche. So befindet sich der im Saarland erzielte Mittelwert in diesem Kompetenzbereich in etwa auf dem Niveau der Stadtstaaten Hamburg und Bremen. Ein ähnlich niedriger Wert wurde außerdem in Hessen festgestellt. Innerhalb der Stadtstaaten unterscheiden sich die Ergebnisse für Hamburg im Gegensatz zu den Ergebnissen für die Globalskala nicht statistisch signifikant von den in Bremen und Berlin erreichten Kompetenzständen. Insgesamt weichen die Ergebnisse für den Kompetenzbereich *Raum und Form* am stärksten von denen für die Globalskala ab.

Im Kompetenzbereich *Muster und Strukturen* zeigen sich vor allem im oberen Leistungsbereich einige Verschiebungen (siehe Abbildung 5.8). Hier liegt nur der im Freistaat Bayern erzielte Kompetenzstand signifikant oberhalb des deutschen Mittelwerts. Die von den Schülerinnen und Schülern in Nordrhein-Westfalen erreichten Kompetenzen sind in diesem Bereich vergleichsweise hoch ausgeprägt und verfehlen im Vergleich mit dem deutschen Mittelwert nur knapp das Kriterium der statistischen Signifikanz. Weniger Veränderung ist in der Gruppe mit unterdurchschnittlichen Kompetenzständen zu verzeichnen, in der sich im Kompetenzbereich *Muster und Strukturen* zusätzlich Mecklenburg-Vorpommern befindet.

Bei der Leitidee *Größen und Messen* erreichen wiederum dieselben Länder statistisch unterdurchschnittliche Mittelwerte wie auf der Globalskala (siehe

Abbildung 5.7: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich *Raum und Form*

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Bayern	538	(4.8)	97	(2.5)	373	415	480	604	655	687	314
Sachsen	528	(4.9)	101	(2.6)	364	406	467	594	652	687	323
Sachsen-Anhalt	524	(4.2)	102	(2.9)	350	396	462	591	649	683	333
Baden-Württemberg	510	(5.4)	99	(2.4)	344	383	446	577	634	667	323
Brandenburg	505	(3.7)	94	(2.0)	351	383	441	570	626	659	307
Thüringen	504	(5.1)	99	(2.9)	341	378	439	571	626	662	321
Deutschland	500	(1.7)	100	(0.9)	335	374	435	568	625	659	324
Mecklenburg-Vorpommern	500	(5.0)	102	(3.2)	326	374	438	567	623	658	331
Rheinland-Pfalz	497	(5.2)	97	(2.8)	337	372	432	563	621	655	318
Niedersachsen	494	(4.3)	96	(2.7)	337	372	431	560	617	649	312
Großstädte ¹	493	(8.8)	105	(6.0)	318	358	423	564	626	662	344
Schleswig-Holstein	487	(4.8)	92	(2.4)	336	369	424	548	605	639	303
Nordrhein-Westfalen	482	(4.9)	95	(2.2)	327	362	419	546	602	637	311
Hessen	476	(5.1)	102	(5.0)	318	363	420	543	596	626	307
Hamburg	476	(3.9)	100	(2.5)	314	347	408	545	605	640	326
Saarland	470	(5.1)	99	(3.5)	309	350	411	535	590	624	315
Bremen	464	(6.7)	100	(2.8)	302	335	394	533	593	630	329
Berlin	459	(4.7)	102	(2.6)	294	331	391	528	591	629	335

- ☐ Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
- ☒ Länder weichen nicht signifikant vom deutschen Mittelwert ab.
- ☐ Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.

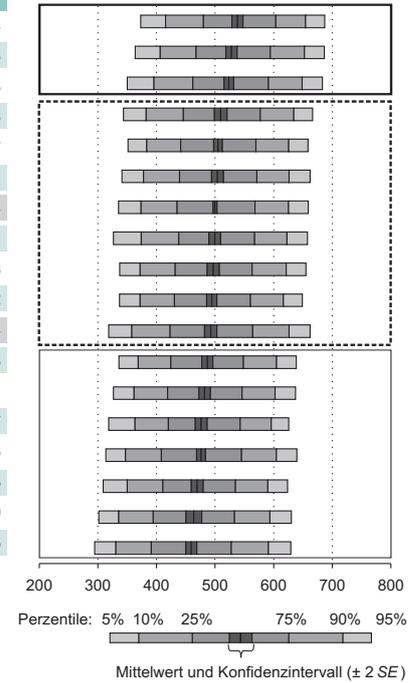


Abbildung 5.8: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich *Muster und Strukturen*

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Bayern	516	(5.0)	98	(2.1)	345	389	456	583	635	666	321
Nordrhein-Westfalen	508	(4.8)	96	(2.2)	347	383	442	576	632	663	316
Großstädte ¹	507	(7.6)	103	(5.4)	330	371	440	577	639	668	338
Baden-Württemberg	506	(4.7)	102	(2.9)	337	376	440	575	635	669	332
Sachsen	506	(4.2)	96	(2.2)	341	380	444	572	624	656	315
Deutschland	500	(1.6)	100	(0.9)	331	371	434	569	626	658	327
Sachsen-Anhalt	500	(4.2)	98	(2.3)	333	375	437	569	620	650	317
Niedersachsen	497	(4.4)	101	(2.9)	329	368	430	565	624	661	332
Thüringen	493	(5.6)	94	(3.0)	336	373	430	558	613	646	310
Saarland	492	(5.1)	105	(4.6)	316	362	428	563	619	650	334
Rheinland-Pfalz	491	(4.6)	99	(2.6)	327	363	423	559	618	652	326
Schleswig-Holstein	486	(4.5)	97	(2.2)	323	362	422	552	609	642	318
Hessen	484	(4.1)	102	(3.1)	315	358	420	555	611	642	327
Mecklenburg-Vorpommern	482	(4.6)	101	(2.5)	311	351	416	552	608	640	330
Hamburg	481	(4.1)	101	(2.5)	312	350	413	551	610	644	333
Brandenburg	481	(3.7)	98	(2.5)	318	354	414	549	605	638	320
Bremen	459	(6.3)	104	(3.0)	291	327	387	531	596	633	343
Berlin	455	(4.4)	103	(2.0)	285	323	385	527	588	622	338

- ☐ Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
- ☒ Länder weichen nicht signifikant vom deutschen Mittelwert ab.
- ☐ Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.

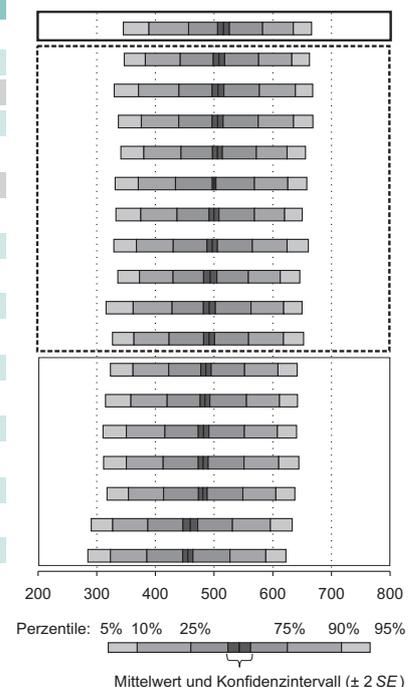


Abbildung 5.9: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich *Größen und Messen*

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Baden-Württemberg	518	(4.6)	98	(3.1)	346	393	460	585	637	666	320
Sachsen-Anhalt	517	(4.1)	103	(2.6)	333	386	459	587	638	667	335
Sachsen	513	(3.8)	97	(2.1)	350	393	458	579	626	655	306
Bayern	509	(5.3)	97	(2.3)	342	382	448	576	628	656	314
Saarland	509	(5.2)	110	(4.8)	318	371	447	584	637	668	349
Thüringen	503	(5.6)	101	(2.8)	330	370	437	573	631	664	335
Deutschland	500	(1.7)	100	(0.8)	329	371	437	569	623	653	324
Nordrhein-Westfalen	498	(5.0)	97	(2.3)	335	371	432	567	621	652	317
Niedersachsen	498	(4.4)	93	(3.7)	344	381	438	562	613	643	299
Mecklenburg-Vorpommern	497	(4.6)	102	(3.2)	321	369	439	566	617	649	328
Rheinland-Pfalz	496	(5.3)	101	(4.1)	330	368	433	565	619	652	321
Großstädte ¹	493	(8.1)	104	(5.8)	318	359	424	567	622	650	333
Schleswig-Holstein	491	(4.3)	98	(2.2)	322	361	426	558	613	646	324
Hessen	491	(4.3)	105	(2.2)	318	363	426	561	617	649	331
Brandenburg	488	(3.8)	96	(3.3)	326	365	426	553	607	639	313
Hamburg	464	(4.1)	107	(2.5)	284	324	392	539	599	635	351
Bremen	448	(7.7)	111	(3.1)	262	304	375	525	588	626	363
Berlin	448	(5.2)	110	(2.7)	261	304	374	525	587	623	362

- ☐ Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
- ☒ Länder weichen nicht signifikant vom deutschen Mittelwert ab.
- ☐ Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.

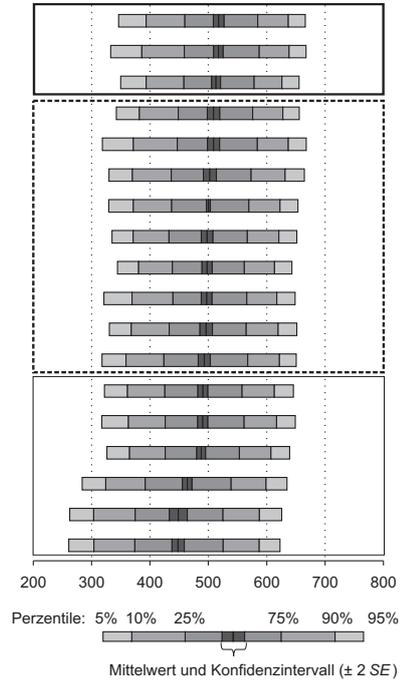


Abbildung 5.10: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Schülerinnen und Schülern der vierten Jahrgangsstufe erreichten Kompetenzstände im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit*

Land	M	(SE)	SD	(SE)	Perzentile						
					5	10	25	75	90	95-5	
Sachsen-Anhalt	516	(4.2)	108	(3.3)	330	385	454	588	642	673	343
Sachsen	513	(4.6)	99	(2.4)	344	388	454	580	632	664	321
Baden-Württemberg	511	(4.5)	96	(2.8)	345	386	451	576	631	663	318
Bayern	510	(4.8)	105	(2.5)	329	378	447	581	637	669	340
Nordrhein-Westfalen	501	(4.6)	92	(1.9)	347	383	440	565	618	648	301
Thüringen	500	(5.1)	97	(2.4)	335	373	437	565	623	659	323
Deutschland	500	(1.5)	100	(0.9)	330	372	436	569	624	656	326
Niedersachsen	498	(4.5)	102	(3.6)	327	368	433	567	625	659	332
Großstädte ¹	496	(8.7)	105	(7.3)	318	361	429	569	626	657	340
Mecklenburg-Vorpommern	494	(4.7)	105	(3.4)	316	367	433	565	618	650	334
Brandenburg	494	(4.0)	99	(2.9)	325	366	430	563	616	648	323
Rheinland-Pfalz	493	(4.8)	98	(2.7)	324	365	429	561	615	646	322
Saarland	493	(4.7)	108	(3.6)	312	359	429	566	623	655	343
Schleswig-Holstein	490	(4.1)	96	(2.4)	330	367	428	557	610	642	312
Hessen	485	(4.5)	104	(3.6)	310	353	419	558	613	645	335
Hamburg	475	(4.3)	106	(2.8)	299	337	404	550	608	642	343
Berlin	461	(4.5)	104	(2.2)	284	325	391	534	592	625	341
Bremen	459	(6.8)	109	(3.0)	275	318	385	536	597	632	357

- ☐ Länder liegen signifikant ($p < .05$) über dem deutschen Mittelwert.
- ☒ Länder weichen nicht signifikant vom deutschen Mittelwert ab.
- ☐ Länder liegen signifikant ($p < .05$) unter dem deutschen Mittelwert.

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. ¹Ohne Berlin, Bremen und Hamburg.

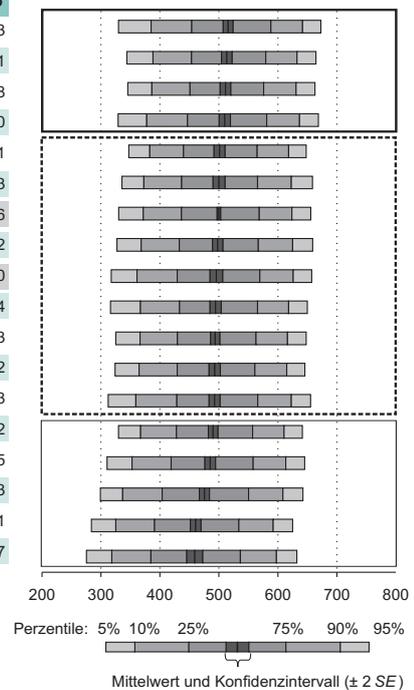


Abbildung 5.9). Im oberen Bereich der Verteilung sind die von den Schülerinnen und Schülern in Bayern erzielten Mittelwerte knapp nicht mehr signifikant überdurchschnittlich. Im Saarland, das in diesem Kompetenzbereich eine wesentlich günstigere Position einnimmt als in den anderen Kompetenzbereichen und auf der Globalskala, wird eine ebenso hohe mittlere Kompetenz erzielt wie im Freistaat Bayern.

Im Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit* schließlich ergeben sich nahezu die gleichen Gruppen wie schon für die Globalskala (siehe Abbildung 5.10). Die in Brandenburg erreichten Kompetenzstände in diesem Bereich sind allerdings als statistisch durchschnittlich einzuordnen. Obwohl der inhaltliche Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit* vergleichsweise neu in den Unterricht eingeführt wurde (vgl. Kapitel 2.2), decken sich die in diesem Bereich festgestellten Länderunterschiede weitgehend mit den für die Globalskala mathematischer Kompetenz ermittelten Werten.

Über die Kompetenzbereiche hinweg zeigt sich zusammenfassend folgendes Bild: Während die in Schleswig-Holstein, Hessen und den drei Stadtstaaten erreichten mittleren Kompetenzstände in sämtlichen Kompetenzbereichen signifikant unterhalb des deutschen Mittelwerts liegen, gibt es kein Land, in dem die Schülerinnen und Schüler in allen fünf inhaltlichen Kompetenzbereichen Mittelwerte erzielen, die statistisch signifikant höher ausfallen als der deutsche Mittelwert. Allerdings sind die Ergebnisse in Bayern, Sachsen und Sachsen-Anhalt in vier sowie in Baden-Württemberg in drei Inhaltsbereichen als klar überdurchschnittlich zu bezeichnen. Diese vier Länder können insofern als Spitzengruppe charakterisiert werden, als dass in keinem weiteren Land in einem der mathematischen Inhaltsbereiche ein signifikant überdurchschnittlicher Kompetenzstand zu verzeichnen ist.

Als globales Maß für die inhaltliche Bedeutsamkeit der Kompetenzunterschiede zwischen den verschiedenen Ländern kann, wie schon in Abschnitt 5.1.1 für die erreichten Kompetenzen im Fach Deutsch, die *Spannweite* der Ländermittelwerte herangezogen werden. Sie errechnet sich als Differenz zwischen dem höchsten und dem niedrigsten Ländermittelwert. Die Spannweite der Ländermittelwerte beträgt für die globale Mathematikkompetenz 68 Punkte; in den einzelnen inhaltlichen Kompetenzbereichen variiert sie zwischen 57 Punkten im Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit* und 79 Punkten im Kompetenzbereich *Raum und Form*. Diese Werte können wiederum mit geschätzten Leistungszuwächsen in Beziehung gesetzt werden, die in Mathematik in der Grundschule im Durchschnitt zu erwarten sind. In den Studien des IQB zur Pilotierung und Normierung der Bildungsstandards lag der querschnittlich ermittelte Kompetenzunterschied zwischen der dritten und vierten Jahrgangsstufe bei ungefähr 80 Punkten (Reiss & Winkelmann, 2009). Die Spannweite der Ländermittelwerte entspricht demnach einem erwarteten Leistungszuwachs von etwas weniger als einem Schuljahr.

Um die Kompetenzprofile der Schülerinnen und Schüler innerhalb der einzelnen Länder im Fach Mathematik genauer zu bestimmen, wurden pro Land die Abweichungen der in jedem Kompetenzbereich erzielten mittleren Kompetenzstände vom jeweiligen Landesmittelwert der Globalskala bestimmt. Hierbei lassen sich nur wenig ausgeprägte Kompetenzprofile einzelner Länder erkennen. Diese Profile zeigen sich zudem nur für einzelne Länder, nicht für Ländergruppen. Dies deutet darauf hin, dass die Abweichungen auf länderspezifische Schwerpunkte im Unterricht zurückzuführen sein könnten. Die größ-

te Differenz eines Kompetenzbereichs zum Landesmittelwert auf der Globalskala findet sich im Saarland für *Raum und Form*. In diesem Bereich sind die Kompetenzwerte der Schülerinnen und Schüler um 22 Punkte geringer als auf der Globalskala. Im Kompetenzbereich *Größen und Messen* hingegen weisen die Schülerinnen und Schüler im Saarland eine um 17 Punkte höhere mittlere Kompetenz im Vergleich zum Landesmittelwert auf der Globalskala auf. Nennenswerte Abweichungen nach oben relativ zum Globalwert bestehen sonst nur noch im Freistaat Bayern mit 19 Punkten und in Brandenburg mit 14 Punkten für den Kompetenzbereich *Raum und Form*. Im Vergleich zur Globalskala eher niedrige mittlere Werte erzielen die Schülerinnen und Schüler in den Ländern Sachsen-Anhalt im Kompetenzbereich *Muster und Strukturen* (17 Punkte) sowie Nordrhein-Westfalen im Kompetenzbereich *Raum und Form* (15 Punkte). Keine der beschriebenen Abweichungen der in den einzelnen Kompetenzbereichen erzielten mittleren Leistungen von den Werten auf der Globalskala erreicht allerdings das Kriterium für statistische Signifikanz.

5.2.2 Heterogenität der im Fach Mathematik in den Ländern erreichten Kompetenzstände

Neben den in den Ländern im Mittel erzielten Kompetenzständen ist außerdem von Interesse, wie stark die Unterschiede in der mathematischen Kompetenz *innerhalb* der einzelnen Länder ausgeprägt sind. Diese Information lässt sich an den in den Abbildungen 5.5 bis 5.10 dargestellten Standardabweichungen (*SD*) der Ergebnisse in den Ländern ablesen. Besonders große Streuungen auf der Globalskala weisen die Länder Sachsen-Anhalt, Mecklenburg-Vorpommern und Bremen mit Standardabweichungen von 106 Punkten auf (siehe Abbildung 5.5). Bemerkenswert ist, dass somit sowohl ein Land besonders große Heterogenität aufweist, das über dem deutschen Mittelwert liegt (Sachsen-Anhalt), als auch ein Land, das ein signifikant unterdurchschnittliches Ergebnis erzielt hat (Bremen). Dies weist erneut darauf hin, dass kein zwangsläufiger Zusammenhang zwischen erreichtem Kompetenzmittelwert und der Kompetenzstreuung besteht. Die geringste Standardabweichung wurde für Schleswig-Holstein festgestellt ($SD = 95$). Insgesamt sind die Unterschiede zwischen den Ländern in den Standardabweichungen für die Globalskala jedoch nicht sehr groß.

In den einzelnen inhaltlichen Kompetenzbereichen zeigen sich besonders große Standardabweichungen für das Saarland in den Bereichen *Zahlen und Operationen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* ($SD_{ZO} = 115$, $SD_{DHW} = 108$), für Bremen in den Bereichen *Größen und Messen, Zahlen und Operationen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* ($SD_{GM} = 111$, $SD_{ZO} = 110$, $SD_{DHW} = 109$), für Hamburg im Bereich *Größen und Messen* ($SD = 107$) sowie für Sachsen-Anhalt im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* ($SD = 108$). Vor allem in Bremen geht der im Durchschnitt niedrige Kompetenzstand also mit erheblichen interindividuellen Unterschieden einher. Aus der Spitzengruppe weist Bayern in drei von fünf Kompetenzbereichen die niedrigste Standardabweichung auf. Diese Art des Befundmusters – eine im Durchschnitt hoch ausgeprägte mathematische Kompetenz bei geringen interindividuellen Unterschieden – gilt allgemein als besonders wünschenswert.

Zusätzlich zur zentralen Tendenz und Streuung können bestimmte Punkte der Kompetenzverteilung, sogenannte Perzentile, zwischen den Ländern verglichen werden. Die Perzentilbänder stellen die Kompetenzverteilungen insge-

samt dar und verdeutlichen die Unterschiedlichkeit in der Lage einzelner Leistungssegmente. Beispielsweise befinden sich unterhalb des 5. Perzentils eines Kompetenzbereichs diejenigen 5 Prozent der Schülerinnen und Schüler des jeweiligen Landes, die die niedrigsten Kompetenzstände aufweisen. Der Wert für das 5. Perzentil gibt also den höchsten Kompetenzwert an, der von Schülerinnen und Schülern dieser Gruppe erreicht wurde. Ein hoher Wert für das 5. Perzentil indiziert, dass auch die leistungsschwächsten Schülerinnen und Schüler eines Landes noch relativ hohe Kompetenzwerte erzielen. Wird am anderen Ende des Kompetenzspektrums für das 95. Perzentil ein hoher Wert festgestellt, so ist dies ein Beleg für eine besonders kompetente Leistungsspitze. Unterhalb und oberhalb des Mittelwertes wurden jeweils drei Perzentile bestimmt: das 5., 10. und 25. beziehungsweise das 75., 90. und 95. Perzentil.

Wie Abbildung 5.5 zeigt, werden im oberen Leistungsbereich die insgesamt höchsten Kompetenzstände in Sachsen-Anhalt erzielt. Die leistungsstärksten 5 Prozent der Schülerinnen und Schüler erreichen hier Werte von mindestens 671 Punkten. Die leistungsstärksten Schülerinnen und Schüler in den Stadtstaaten erreichen ebenfalls Werte, die mehr als eine Standardabweichung über dem deutschen Mittelwert liegen (634 Punkte in Hamburg, 625 Punkte in Bremen, 616 Punkte in Berlin). Die Differenz des höchsten und niedrigsten Länderwerts für das 95. Perzentil beträgt 55 Punkte.

Für das 5. Perzentil werden die höchsten Werte in Sachsen erreicht (354 Punkte), die niedrigsten hingegen in Berlin (282 Punkte). Im untersten Kompetenzspektrum sind die Unterschiede zwischen den Ländern mit 72 Punkten deutlich größer als im obersten Kompetenzbereich.

Die Abbildung enthält außerdem Angaben zum Abstand zwischen dem 5. und dem 95. Perzentil. Dieser Wert markiert die Spanne, in der die Kompetenzen der mittleren 90 Prozent aller Schülerinnen und Schüler eines Landes liegen. Diese Spanne ist umso größer, je heterogener die von den Schülerinnen und Schülern innerhalb eines Landes erreichten Kompetenzen sind. In allen Ländern beträgt sie mehr als 300 Punkte. Vergleicht man die Differenz des 5. und 95. Perzentils zwischen den Ländern, ergeben sich vergleichsweise homogene Kompetenzwerte in den Ländern Nordrhein-Westfalen, Niedersachsen, Schleswig-Holstein und Sachsen, vergleichsweise heterogene Kompetenzwerte hingegen in Mecklenburg-Vorpommern, Sachsen-Anhalt und Bremen. Insgesamt sind die Differenzen zwischen den Ländern auf diesem Maß jedoch gering. Dies gilt in ähnlicher Weise auch für die Ergebnisse in den einzelnen Kompetenzbereichen (vgl. Abbildung 5.5 bis 5.10).

5.2.3 Heterogenität in den Kompetenzständen innerhalb und zwischen Schulen

Um zu überprüfen, inwieweit die in den Ländern erreichten Kompetenzstände zwischen Schulen oder innerhalb von Schulen variieren, wurden Intraklassenkorrelationen für alle Länder berechnet. Anhand dieser Größe lässt sich feststellen, wie stark sich die erreichten Kompetenzen von Schülerinnen und Schülern in verschiedenen Schulen unterscheiden. Mögliche Faktoren, die dieser Variation zugrunde liegen können, werden in Teilkapitel 5.1 näher erläutert. Wie bereits für die in Teilkapitel 5.1 berichteten Intraklassenkorrelationen für die Kompetenzbereiche Lesen und Zuhören im Fach Deutsch wurden die

Berechnungen pro Land einmal für alle Schulen und einmal nur für Regelschulen durchgeführt. Die Intraklassenkorrelationen sind in Abbildung 5.11 dargestellt.

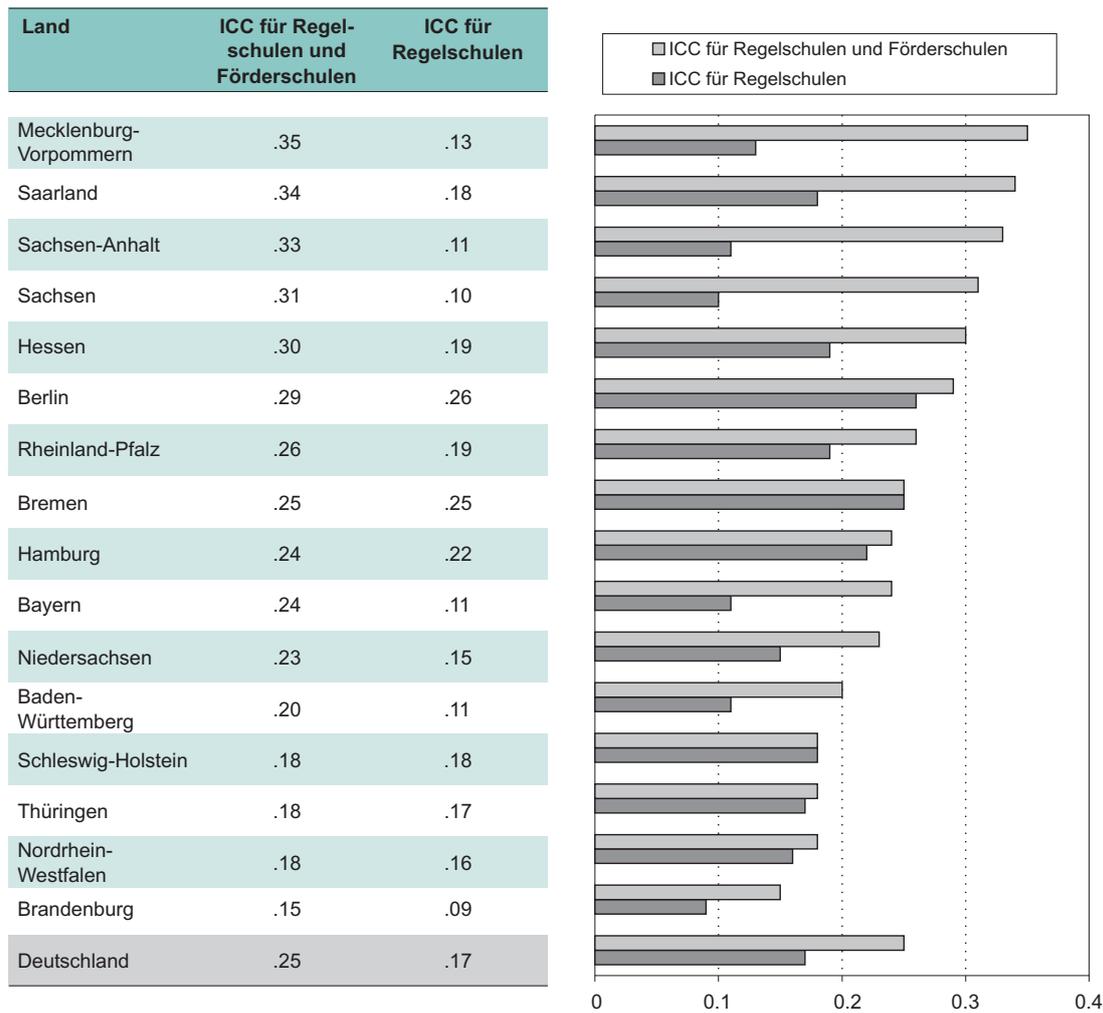
Für Deutschland zeigt sich, dass 25 Prozent der gesamten Varianz der Testergebnisse in Mathematik zwischen den untersuchten Schulen liegen. Werden nur die Regelschulen betrachtet, verringert sich der Anteil auf 17 Prozent. Betrachtet man Förder- und Regelschulen gemeinsam, weisen die Länder Mecklenburg-Vorpommern, Saarland, Sachsen-Anhalt, Sachsen, Hessen und Berlin vergleichsweise hohe, die Länder Baden-Württemberg, Schleswig-Holstein, Thüringen, Nordrhein-Westfalen und Brandenburg hingegen eher niedrigere Intraklassenkorrelationen auf. Bei den Intraklassenkorrelationen für Regelschulen zeigt sich hingegen ein anderes Bild; hier ergeben sich in den Stadtstaaten Berlin, Bremen und Hamburg die höchsten Intraklassenkorrelationen. In Brandenburg und Baden-Württemberg sind die Intraklassenkorrelationen in beiden Analysen im Vergleich zum Wert für Deutschland insgesamt eher klein.

In Abbildung 5.11 ist deutlich zu erkennen, dass die Intraklassenkorrelationen für Regelschulen in den meisten Ländern niedriger ausfallen als diejenigen für alle beteiligten Schulen. So sind die Intraklassenkorrelationen in Mecklenburg-Vorpommern, Sachsen-Anhalt und Sachsen um ungefähr .20 geringer, wenn man nur die Schülerinnen und Schüler in Regelschulen betrachtet, was jeweils einer Reduktion um über 60 Prozent entspricht. Allgemein ist der Unterschied innerhalb eines Landes zwischen der Intraklassenkorrelation für alle Schulen und derjenigen für Regelschulen umso höher, je größer in der Tendenz der Mittelwertunterschied zwischen Förder- und Regelschulen ist und je mehr Schülerinnen und Schüler sich in einer Förderschule befinden. Vor allem in Ländern, in denen ein vergleichsweise hoher Anteil von Schülerinnen und Schülern eine Förderschule besucht (vgl. Kapitel 4, Tabelle 4.2), ist der Unterschied daher stark ausgeprägt. Darauf dürften auch die vergleichsweise großen Differenzen zwischen den beiden Intraklassenkorrelationen in den Ländern Bayern, Mecklenburg-Vorpommern, Sachsen und Sachsen-Anhalt zumindest teilweise zurückzuführen sein. Die Stichprobenziehung im Ländervergleich 2011 ermöglicht es jedoch nicht, die im Mittel in Förderschulen erreichten Kompetenzen in jedem Land verlässlich separat zu schätzen.

Die Intraklassenkorrelationen zwischen den im Fach Deutsch untersuchten Kompetenzbereichen und denen für die Globalskala mathematischer Kompetenz sind in den meisten Ländern sehr ähnlich. Es zeigen sich jedoch Unterschiede in den ICCs einzelner Kompetenzbereiche in verschiedenen Ländern. Im Saarland sind die Intraklassenkorrelationen für Regel- und Förderschulen im Fach Mathematik ($ICC = .34$) höher als für die Bereiche Lesen ($ICC = .18$) und Zuhören ($ICC = .24$). Die Kompetenzunterschiede zwischen Schulen sind somit für Mathematik deutlich größer als im Lesen und im Zuhören. In Hamburg sind dagegen die Intraklassenkorrelationen im Bereich Zuhören ($ICC = .39$) stärker ausgeprägt als in Mathematik ($ICC = .24$). Dies weist darauf hin, dass sich die untersuchten Schulen deutlich stärker im Zuhören unterscheiden als in Mathematik. Werden nur die Regelschulen betrachtet, zeigen sich in Sachsen-Anhalt deutlich geringere Kompetenzunterschiede zwischen den untersuchten Schulen in Mathematik ($ICC = .11$) als in den beiden Kompetenzbereichen des Faches Deutsch ($ICC = .25$ für Lesen, $ICC = .31$ für Zuhören).

Setzt man die Werte für die Intraklassenkorrelation mit den in Abbildung 5.5 berichteten Ländermittelwerten auf der Globalskala in Beziehung, ergibt sich ein differenzierteres Bild. Für die Länder in der Spitzengruppe – Sachsen, Sachsen-Anhalt, Bayern und Baden-Württemberg – zeigen die vergleichsweise gerin-

Abbildung 5.11: Intraklassenkorrelationen (ICC) der mathematischen Kompetenz (Globalskala) für Regelschulen und Förderschulen gemeinsam sowie ausschließlich für Regelschulen



Anmerkung. Die Länder sind absteigend nach der ICC für Regelschulen und Förderschulen geordnet.

gen Intraklassenkorrelationen für Regelschulen, dass sich die Kompetenzen von Schülerinnen und Schülern in verschiedenen Schulen innerhalb dieser Länder nicht stark voneinander unterscheiden. Dies weist auf eine über die Schulen hinweg weitgehend vergleichbare Komposition der Schülerschaft und die Bereitstellung vergleichbar guter Lerngelegenheiten hin. Auch die niedrige Intraklassenkorrelation in Brandenburg im Zusammenhang mit dem unterdurchschnittlichen Abschneiden der Schülerinnen und Schüler deutet auf ähnliche Lehr-Lern-Bedingungen in den Schulen hin, die allerdings allgemein weniger günstig zu sein scheinen. In den Stadtstaaten Hamburg, Bremen und Berlin, in denen ebenfalls unterdurchschnittliche Kompetenzwerte erzielt wurden, sind hingegen die Intraklassenkorrelationen am höchsten ausgeprägt. Sie können hier unter anderem als Ausdruck der heterogenen Zusammensetzung der Schülerschaft unterschiedlicher Schulen in Bezug auf den sozialen Hintergrund und den Zuwanderungshintergrund gesehen werden.

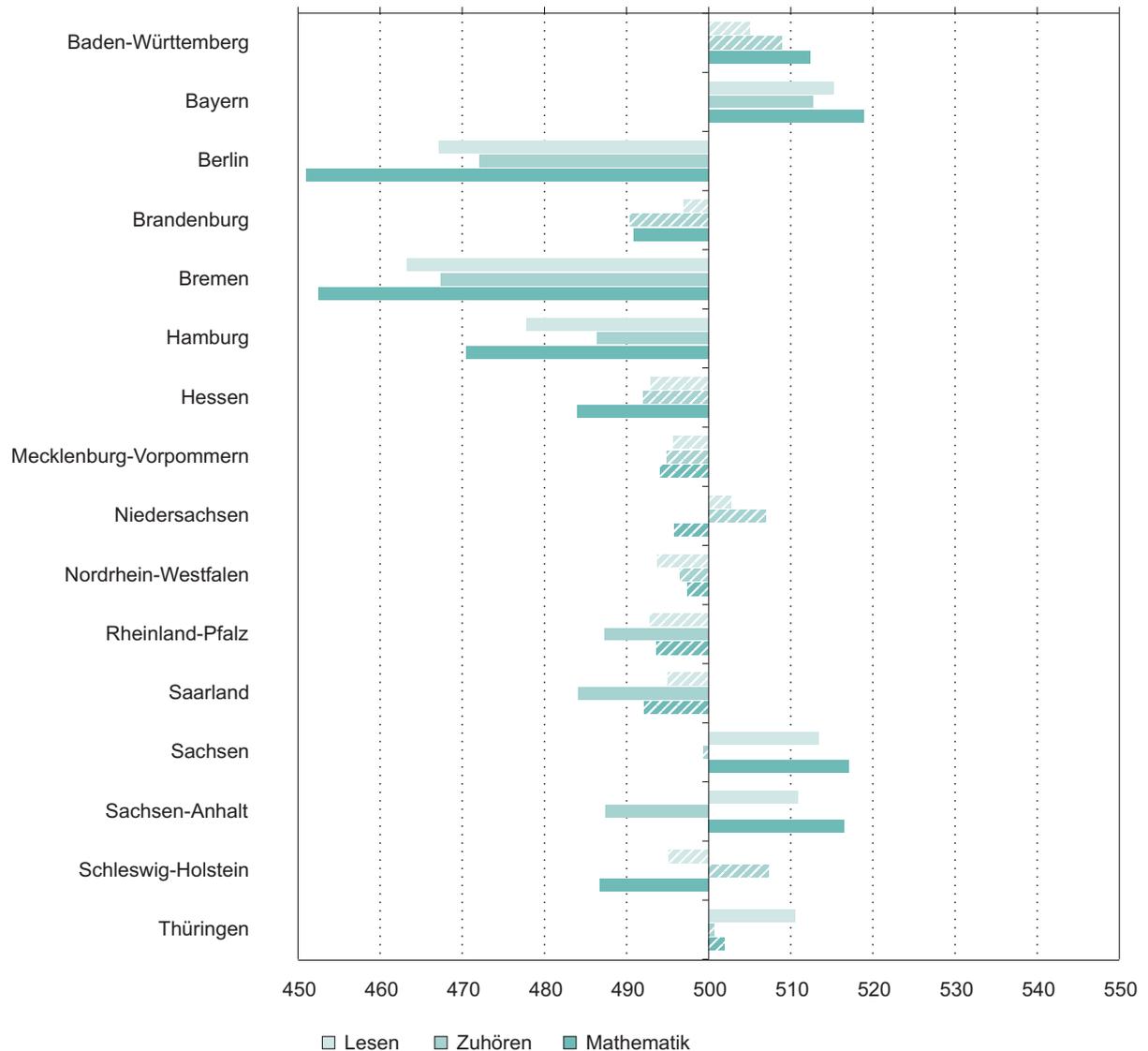
5.3 Zusammenfassung zentraler Ergebnisse des Ländervergleichs in den Fächern Deutsch und Mathematik

Katrin Böhme und Nicole Haag

In diesem letzten Teilkapitel soll eine Zusammenschau zentraler Befunde des Ländervergleichs 2011 vorgenommen werden. Die Abbildung 5.12 zeigt, inwieweit die in den Ländern erreichten Kompetenzstände in den drei Bereichen Lesen, Zuhören und Mathematik vom deutschen Mittelwert von 500 Punkten abweichen. Die Darstellung konzentriert sich auf die im Mittel erreichten Kompetenzwerte, berücksichtigt also nicht die jeweiligen Leistungsstreuungen innerhalb und zwischen den Ländern. In der Abbildung sind Differenzen, die sich nicht signifikant vom deutschen Mittelwert unterscheiden, als schraffierte Balken dargestellt. Abweichungen, die sich statistisch bedeutsam vom deutschen Mittelwert unterscheiden, sind dagegen vollfarbig abgebildet.

Es wird deutlich, dass in der Mehrzahl der Länder die Ergebnisse für alle drei Bereiche tendenziell in dieselbe Richtung weisen. Das heißt, die mittleren Kompetenzen weichen entweder konsistent nach unten oder nach oben von der in Deutschland im Mittel erzielten Leistung ab. Hierbei gelingt es nur Schülerinnen und Schülern im Freistaat Bayern sowohl im Lesen als auch im Zuhören sowie im Fach Mathematik Kompetenzen zu entwickeln, die signifikant oberhalb der mittleren Kompetenzausprägungen in Deutschland insgesamt liegen. In den drei Stadtstaaten Berlin, Bremen und Hamburg liegen die von den Schülerinnen und Schülern erreichten Kompetenzen hingegen in allen drei Bereichen statistisch bedeutsam unterhalb des deutschen Mittelwerts. In der überwiegenden Mehrzahl der Länder werden Kompetenzstände erreicht, die sich nicht oder nur für einzelne Bereiche signifikant von den jeweiligen deutschen Mittelwerten unterscheiden. Über alle drei Kompetenzbereiche hinweg konsistente, also von der Tendenz her identische Abweichungen von den mittleren deutschen Kompetenzständen finden sich in den Ländern Baden-Württemberg, Brandenburg, Hessen, Mecklenburg-Vorpommern, Nordrhein-Westfalen, Rheinland-Pfalz, im Saarland und dem Freistaat Thüringen.

In Niedersachsen liegen die Befunde für alle drei Kompetenzen sehr dicht am deutschen Mittelwert, weichen aber für Lesen und Zuhören geringfügig nach oben und für Mathematik geringfügig nach unten ab. In Sachsen, Sachsen-Anhalt und Schleswig-Holstein weisen die Abweichungen in den drei untersuchten Bereichen in unterschiedliche Richtungen. In den Ländern Sachsen und Sachsen-Anhalt werden in den Kompetenzbereichen Lesen und Mathematik Kompetenzstände erreicht, die oberhalb des deutschen Mittelwerts liegen, die im Bereich Zuhören erzielten Kompetenzen weichen hingegen in Sachsen-Anhalt signifikant und in Sachsen tendenziell nach unten ab. In Schleswig-Holstein findet sich hingegen der umgekehrte Befund, nämlich tendenziell überdurchschnittliche Kompetenzen im Zuhören bei gleichzeitig unterdurchschnittlichen Kompetenzen im Lesen und in Mathematik. Während also in vielen Ländern über die Kompetenzbereiche hinweg weitgehend einheitliche Ergebnisse erzielt werden, zeichnen sich in anderen Ländern einzelne Stärken und Schwächen ab.

Abbildung 5.12: Abweichungen der in den Ländern erreichten Kompetenzstände in den Bereichen Lesen, Zuhören und Mathematik vom deutschen Mittelwert

Anmerkung. Schraffierte Balken unterscheiden sich im jeweiligen Kompetenzbereich nicht signifikant vom deutschen Mittelwert.

Literatur

- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Behrens, U., Böhme, K. & Krelle, M. (2009). Zuhören – Operationalisierung und fachdidaktische Implikationen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 357–375). Weinheim: Beltz.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2008). *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.

- Bos, W., Lankes, E.-M., Prenzel, M. & Schwippert, K. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Frey, A., Asseburg, R., Ehmke, T. & Blum, W. (2008). Mathematische Kompetenz im Ländervergleich. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 127–147). Münster: Waxmann.
- Granzer, D., Köller, O., Bremerich-Vos, A., van den Heuvel-Panhuizen, M., Reiss, K. & Walther, G. (Hrsg.). (2009). *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule*. Weinheim: Beltz.
- KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- Köller, O., Knigge, M. & Tesch, B. (Hrsg.). (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Pietsch, M. (2010). Evaluation von Unterrichtsstandards. *Zeitschrift für Erziehungswissenschaft, 13*, 121–148.
- Pietsch, M., Böhme, K., Robitzsch, A. & Stubbe, T. C. (2009). Das Stufenmodell zur Lesekompetenz der länderübergreifenden Bildungsstandards im Vergleich zu IGLU 2006. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 393–416). Weinheim: Beltz.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 120–141). Weinheim: Beltz.
- Valtin, R., Bos, W., Buddeberg, I., Goy, M. & Potthoff, B. (2008). Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe im nationalen und internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 51–101). Münster: Waxmann.
- Willms, J. D. (2006). *Learning divides: Ten policy questions about the performance and equity of schools and schooling systems*. Montreal, Quebec: UNESCO Institute for Statistics.

Kapitel 6

Der Blick in die Länder

Petra Stanat, Hans Anand Pant, Dirk Richter, Katrin Böhme,
Maria Engelbert, Nicole Haag, Jasmin Hannighofer, Josefine Prengel,
Heino Reimers, Alexander Roppelt und Sebastian Weirich

Die von der Kultusministerkonferenz (KMK) 2006 verabschiedete Gesamtstrategie zum Bildungsmonitoring sieht vor, regelmäßig durch das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) in einem Ländervergleich überprüfen zu lassen, inwieweit die Bildungsstandards in den Ländern erreicht werden. Laut Beschluss der Kultusministerkonferenz sollen die Erhebungen im Primarbereich in der vierten Jahrgangsstufe durchgeführt und an die alle fünf beziehungsweise vier Jahre stattfindenden internationalen Schulleistungsstudien PIRLS/IGLU¹ und TIMSS² angebunden werden, um einen Bezug zu internationalen Vergleichsmaßstäben zu ermöglichen. Im Folgenden werden die Befunde der ersten zentralen Überprüfung des Erreichens der Bildungsstandards in der Primarstufe dargestellt. Wie bereits im Bericht zum 2009 durchgeführten Ländervergleich sprachlicher Kompetenzen in der Sekundarstufe I (Köller, Knigge & Tesch, 2010), werden die Ergebnisse jeweils getrennt für die einzelnen Länder beschrieben.

Im Vergleich zur Sekundarstufe I, die durch schulstrukturelle Heterogenität gekennzeichnet ist, stellt sich die Situation im Primarbereich in den Ländern deutlich homogener dar. Während in der Sekundarstufe I zwischen den Ländern erhebliche Unterschiede in der Anzahl und im Zuschnitt der angebotenen Schularten, in der Verteilung der Schülerinnen und Schüler auf diese Schularten und in den Quoten der erreichten Abschlüsse bestehen (vgl. Köller et al., 2010), besuchen Schülerinnen und Schüler der vierten Jahrgangsstufe in der Regel überall in Deutschland eine Grundschule. In der Primarstufe unterscheiden sich die Länder daher kaum in der Bildungsbeteiligung und in der Quote der Schülerinnen und Schüler, die diese Bildungsetappe abschließen. Wie bereits in Kapitel 4 dargestellt, bestehen jedoch in einem nicht unerheblichen Maße Unterschiede darin, wie groß der Anteil der Kinder mit einem sonderpädagogischen Förderbedarf ist und zu welchen Anteilen diese Kinder eine Förderschule oder eine Regelschule besuchen. Die entsprechenden Daten der amtlichen Statistik, die sich auf die Jahrgangsstufen 1–4 der Primarstufe beziehen, sind in Tabelle 6.1 dargestellt. In den Angaben sind alle Schülerinnen und Schüler enthalten, bei denen im Schuljahr 2010/2011 ein sonderpädagogischer Förderbedarf diagnostiziert wurde, unabhängig vom Förderschwerpunkt.

-
- 1 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.
 - 2 Das Akronym TIMSS steht für *Trends in International Mathematics and Science Study*.

Tabelle 6.1: Verteilung von Kindern mit sonderpädagogischem Förderbedarf (SFB) in den Jahrgangsstufen 1–4 im Schuljahr 2010/2011 nach Ländern

Land	Gesamt- schülerzahl	Schülerinnen und Schüler mit SFB		davon			
		Anzahl	Anteil in %	Schülerinnen und Schüler mit SFB in Regel- schulen	Schülerinnen und Schüler mit SFB in Förder- schulen ¹	Schülerinnen und Schüler mit SFB in Regel- schulen	Schülerinnen und Schüler mit SFB in Förder- schulen
				Anzahl	Anteil in %	Anzahl	Anteil in %
Baden- Württemberg	411 943	31 274	7.6	15 158	16 116	48.5	51.5
Bayern	470 907	32 146	6.8	9 110	23 036	28.3	71.7
Berlin	109 448	6 946	6.3	3 910	3 036	56.3	43.7
Brandenburg	78 939	4 742	6.0	2 613	2 129	55.1	44.9
Bremen	21 120	1 360	6.4	1 201	159	88.3	11.7
Hamburg	59 247	3 151	5.3	1 155	1 996	36.7	63.3
Hessen	220 290	8 571	3.9	2 429	6 142	28.3	71.7
Mecklenburg- Vorpommern	51 804	4 332	8.4	1 558	2 774	36.0	64.0
Niedersachsen ²	309 523	9 663	3.1	--	9 663	--	--
Nordrhein- Westfalen	691 833	35 462	5.1	11 560	23 902	32.6	67.4
Rheinland-Pfalz	150 004	4 981	3.3	1 406	3 575	28.2	71.8
Saarland ²	32 002	315	1.0	--	315	--	--
Sachsen	128 519	9 162	7.1	2 968	6 194	32.4	67.6
Sachsen-Anhalt	69 266	5 140	7.4	1 749	3 391	34.0	66.0
Schleswig- Holstein	108 758	4 603	4.2	3 569	1 034	77.5	22.5
Thüringen	68 935	3 353	4.9	1 338	2 015	39.9	60.1
Deutschland	2 982 538	165 201	5.5	59 724	105 477	36.2	63.8

Anmerkungen. ¹ Zusätzlich weist die Statistik des Statistischen Bundesamtes etwa 10 000 Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf an Förderschulen in den Klassenstufen 1–13 aus, die keiner bestimmten Jahrgangsstufe zugewiesen werden konnten. ² Die Angaben zu der Anzahl und dem Anteil von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf beziehen sich ausschließlich auf die Schülerinnen und Schüler in Förderschulen.

Quelle: Fachserie 11, Statistisches Bundesamt (2011)

Für die Überprüfung des Erreichens der Bildungsstandards ist vor allem die Frage von Bedeutung, ob für diese Schülerinnen und Schüler dieselben Lernziele maßgeblich sind wie für Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf. Da für *zieldifferent* unterrichtete Kinder eigene Anforderungen gelten, wäre es unabhängig vom Lernort unangemessen, die Bildungsstandards als Maßstab für die von ihnen erreichten Leistungen heranzuziehen. In die nachfolgend dargestellten Befunde zur Frage, inwieweit die Schülerinnen und Schüler in den einzelnen Ländern die Bildungsstandards erreichen, werden folglich nur solche Kinder einbezogen, die nach den Regelungen des jeweiligen Landes *zielgleich* unterrichtet werden.

Da die amtliche Statistik keine Angaben zur Anzahl der zielgleich und zieldifferent unterrichteten Schülerinnen und Schüler bereitstellt, müssen die entsprechenden Quoten anhand der Daten für die Stichprobe des Ländervergleichs 2011 geschätzt werden. Die Schulkoordinatorinnen und Schulkoordinatoren der teilnehmenden Schulen wurden gebeten, in einer dafür vorgesehenen Liste jeweils zu kennzeichnen, für welche Schülerinnen und Schüler andere Lehrpläne/Rahmenlehrpläne/Kerncurricula als die der Regelschulen gelten. Die anhand dieser Angaben ermittelten Quoten sind in

Tabelle 6.2 aufgeführt.³ Dabei wurden nur solche Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf einbezogen, die den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen zugeordnet wurden (vgl. Kapitel 4).

Tabelle 6.2: Anteil von zielgleich und zieldifferent unterrichteten Kindern mit sonderpädagogischem Förderbedarf in der vierten Jahrgangsstufe nach Ländern (Daten des Ländervergleichs 2011)

Land	Schülerinnen und Schüler mit SFB		Schülerinnen und Schüler mit SFB an Förderschulen		Schülerinnen und Schüler mit SFB an Regelschulen		Zielgleich unterrichtete Schülerinnen und Schüler mit SFB		Zieldifferent unterrichtete Schülerinnen und Schüler mit SFB	
	an Gesamt-population	an Gesamt-population	an SuS mit SFB	an Gesamt-population	an SuS mit SFB	an Gesamt-population	an SuS mit SFB ¹	an Gesamt-population	an SuS mit SFB ¹	
	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	
Baden-Württemberg	4.6	2.4	52.1	2.2	47.9	3.1	68.6	0.8	18.1	
Bayern	5.7	2.8	48.6	2.9	51.4	3.5	61.8	2.1	37.0	
Berlin	5.8	1.2	21.0	4.6	79.0	3.6	61.8	2.0	33.9	
Brandenburg	7.9	1.9	24.4	6.0	75.6	4.4	55.9	2.4	29.9	
Bremen	2.0	0.1	7.1	1.9	92.9	1.8	88.5	0.2	8.8	
Hamburg	5.2	1.1	20.8	4.1	79.2	3.2	61.0	1.9	36.3	
Hessen	4.1	1.7	40.4	2.4	59.6	2.0	48.4	2.1	50.1	
Mecklenburg-Vorpommern	11.3	4.5	39.3	6.9	60.7	6.5	57.7	4.1	36.2	
Niedersachsen	3.4	1.7	49.8	1.7	50.2	1.6	47.1	1.7	51.0	
Nordrhein-Westfalen	3.8	1.6	42.2	2.2	57.8	2.7	70.6	1.1	29.4	
Rheinland-Pfalz	3.2	1.5	47.3	1.7	52.7	1.0	31.1	2.2	68.9	
Saarland	4.4	1.8	40.3	2.6	59.7	1.5	35.1	2.9	64.9	
Sachsen	6.1	3.3	53.9	2.8	46.1	3.4	55.2	2.6	42.9	
Sachsen-Anhalt	6.6	4.1	61.9	2.5	38.1	4.5	67.7	2.1	32.3	
Schleswig-Holstein	4.7	1.0	21.7	3.7	78.3	1.8	38.6	2.9	61.4	
Thüringen	4.9	1.1	21.9	3.8	78.1	4.3	87.8	0.5	11.1	
Deutschland	4.7	2.0	43.2	2.7	56.8	2.8	60.4	1.7	36.1	

Anmerkungen. ¹ Die Anteile zielgleich und zieldifferent unterrichteter Schülerinnen und Schüler entsprechen wegen fehlender Werte in der Summe nicht immer 100 Prozent. SFB = Sonderpädagogischer Förderbedarf; SuS = Schülerinnen und Schüler.

3 Abweichungen zwischen Tabelle 6.1 und Tabelle 6.2 in der Anzahl von Kindern mit sonderpädagogischem Förderbedarf sind darauf zurückzuführen, dass sich die Zahlen in Tabelle 6.1 auf alle Förderschwerpunkte beziehen, in Tabelle 6.2 hingegen nur auf die Förderschwerpunkte Sprache, soziale und emotionale Entwicklung sowie Lernen. In Tabelle 6.1 sind zudem die Jahrgangsstufen 1–4 der Primarschule enthalten, wohingegen sich Tabelle 6.2 nur auf die Jahrgangsstufe 4 bezieht. Weiterhin basieren die Angaben in Tabelle 6.1 auf Populationsdaten, während es sich bei den Angaben in Tabelle 6.2 um Schätzungen anhand von Stichprobendaten handelt, die statistische Fehlertoleranzen aufweisen. Trotz dieser Unterschiede sind Abweichungen von drei Prozentpunkten oder mehr nur für Baden-Württemberg (3.0 %) und Bremen (4.4 %) zu beobachten. Die Abweichung für das Saarland von 3.4 Prozentpunkten dürfte darauf zurückzuführen sein, dass sich die Angabe der amtlichen Statistik nur auf Schülerinnen und Schüler in Förderschulen bezieht.

Wie aus Tabelle 6.2 hervorgeht, variieren auch die Quoten der zieldifferent unterrichteten Kinder mit sonderpädagogischem Förderbedarf zwischen den Ländern. Dies ist insbesondere dann der Fall, wenn der Anteil nicht bezogen auf alle Grundschul Kinder, sondern nur auf die Gruppe der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf berechnet wird. Während 2011 in Bremen nur knapp 9 Prozent der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf zieldifferent unterrichtet wurden, waren es in Rheinland-Pfalz fast 69 Prozent. Bezogen auf die Gesamtpopulation der Viertklässlerinnen und Viertklässler variiert die Quote der zieldifferent unterrichteten Kinder allerdings nur zwischen 0.2 Prozent in Bremen und etwa 4 Prozent in Mecklenburg-Vorpommern. Auch wenn der Anteil zieldifferent unterrichteter Schülerinnen und Schüler an der Gesamtpopulation der Viertklässlerinnen und Viertklässler in allen Ländern gering ist, sollte bei der Interpretation der in diesem Kapitel dargestellten Ergebnisse berücksichtigt werden, dass in den einzelnen Ländern geringfügig unterschiedliche Anteile von Kindern aus den Analysen und somit auch von den Verteilungen auf die Kompetenzstufen ausgeschlossen wurden. Entsprechend werden die Ergebnisse für jedes Land separat berichtet und interpretiert.

In den nachfolgenden Länderkapiteln wird zunächst noch einmal darauf eingegangen, welcher Anteil der Schülerinnen und Schüler nach der Diagnosepraxis des jeweiligen Landes einen sonderpädagogischen Förderbedarf aufweist, zu welchen Anteilen diese Schülerinnen und Schüler eine Förderschule oder eine Regelschule besuchen und ob sie zieldifferent oder zielgleich unterrichtet werden.⁴ Dabei bezieht sich der Förderbedarf immer nur auf die Bereiche Sprache, soziale und emotionale Entwicklung oder Lernen, da Schülerinnen und Schüler mit anderen Förderschwerpunkten aus der Populationsdefinition des Ländervergleichs 2011 von vornherein ausgeschlossen wurden (vgl. Kapitel 4). Aus Gründen der Lesbarkeit wird diese Beschränkung nicht jedes Mal ausdrücklich erwähnt.

Weitere Kontextinformationen, die in den Länderkapiteln dargestellt werden, beziehen sich auf Merkmale der Schullaufbahnen der Kinder im Primarbereich im jeweiligen Land. So hängt es von Stichtagsregelungen, Quoten der Zurückstellung von der Einschulung sowie Wiederholerquoten⁵ ab, welches Alter die Kinder zum Zeitpunkt der Testung in der vierten Jahrgangsstufe erreicht haben und in welchem Umfang sie bis zu diesem Zeitpunkt schulische Lerngelegenheiten nutzen konnten. Für den Umfang der Lerngelegenheiten ist vor allem auch die Anzahl der Unterrichtsstunden relevant. Daher wird berichtet, wie viele Stunden Unterricht in den Fächern Deutsch und Mathematik in der Primarstufe laut Stundentafel in einem Land mindestens erteilt werden

4 Dabei wird analog zum Nationalen Bildungsbericht 2012 (Autorengruppe Bildungsberichterstattung, 2012) keine Unterscheidung zwischen Integration beziehungsweise Inklusion von Kindern mit sonderpädagogischem Förderbedarf getroffen. Die Diskussion um inhaltliche Differenzen zwischen beiden Konzepten soll an dieser Stelle nicht aufgegriffen werden.

5 Da im Primarbereich oft jahrgangsübergreifend gelernt wird, kann nicht präzise bestimmt werden, ob eine Schülerin oder ein Schüler eine Klasse übersprungen hat. Daher werden nur die Quoten der Klassenwiederholungen berichtet.

sollen.^{6,7} Den Schätzungen der insgesamt in den ersten vier Schuljahren erteilten Unterrichtsstunden wurde ein 38 Unterrichtswochen umfassendes Schuljahr zugrunde gelegt.

Darüber hinaus wird berichtet, welcher Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe 2011 eine Ganztagschule besuchte. Diese Information basiert auf Angaben der Schulleiterinnen und Schulleiter zur Frage, ob ihre Schule im laufenden Schuljahr „Ganztagsbetrieb an mindestens drei Wochentagen über jeweils mindestens sieben Zeitstunden“ hatte. Weiterhin wurde gefragt, ob es sich um eine voll gebundene Form („alle Schülerinnen und Schüler sind verpflichtet, an mindestens drei Wochentagen für jeweils mindestens sieben Zeitstunden am Ganztagsbetrieb der Schule teilzunehmen“), um eine teilweise gebundene Form („ein Teil der Schülerinnen und Schüler verpflichtet sich, an mindestens drei Wochentagen für jeweils mindestens sieben Zeitstunden an den ganztägigen Angeboten der Schule teilzunehmen“) oder um eine offene Form („Aufenthalt verbunden mit einem Bildungs- und Betreuungsangebot in der Schule an mindestens drei Wochentagen von täglich mindestens sieben Zeitstunden für Schülerinnen und Schüler ist möglich“) des Ganztagsbetriebs handelte.

In der Darstellung dieser mit potenziellen Lerngelegenheiten verbundenen Kontextinformationen wird sowohl auf Daten der amtlichen Statistik als auch auf Auswertungen des hier vorgelegten Ländervergleichs 2011 zurückgegriffen. Die Datenbasis wird jeweils angegeben. Im Anschluss wird für jedes Land berichtet, wie sich die Schülerinnen und Schüler in den Fächern Deutsch (Lesen und Zuhören) sowie Mathematik (Globalskala) auf die Kompetenzstufen verteilen. Die diesen Analysen zugrunde gelegten Kompetenzstufenmodelle wurden in den Kapiteln 3.2 und 3.3 ausführlich beschrieben. Die Darstellungen der Länderergebnisse folgen durchgängig demselben Schema und schließen jeweils mit einer kurzen ländervergleichenden Einordnung auffälliger Befunde. Diese Einordnungen haben einen rein deskriptiv-resümierenden Charakter; eine statistische Signifikanzprüfung liegt ihnen nicht zu Grunde.

Nach den von der KMK verabschiedeten Kompetenzstufenmodellen bildet die Kompetenzstufe III den Regelstandard für den jeweiligen Bildungsabschnitt als „im Durchschnitt“ zu erreichendes Kompetenzniveau ab. Als Mindeststandard gilt die Kompetenzstufe II. Schülerinnen und Schüler, deren Leistungen lediglich der Kompetenzstufe I entsprechen, verfehlen die länderübergreifend festgelegten Minimalanforderungen. Es wird angenommen, dass sie in besonderem Maße gefährdet sind, in ihrer weiteren Bildungslaufbahn den Anschluss zu verlieren. Auf den Kompetenzstufen IV und V, die als „Regelstandard plus“ beziehungsweise „Optimalstandard“ bezeichnet werden, übertreffen die Schülerinnen und Schüler die Regelerwartungen der KMK.

6 In einigen Ländern weist die Stundentafel keine separaten Stundenzahlen für die Fächer Deutsch und Mathematik pro Schuljahr aus, sondern ist als sogenannte Kontingentstundentafel abgefasst. Die Kontingentstundentafel legt für jede Schulart fest, wie viele Jahresschulstunden insgesamt in den Schuljahren bis zum Abschluss des Bildungsgangs zu erteilen sind. Wie diese Jahresschulstunden jedoch auf die einzelnen Klassenstufen beziehungsweise Fächer verteilt werden, entscheiden die einzelnen Schulen. Sie erhalten damit pädagogischen Freiraum und können die Verteilung der Stunden nutzen, um Schwerpunkte zu setzen und die Schulkonzepte zu gestalten.

7 Die Herausgeber danken dem Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland für die Recherche und Bereitstellung von Länderinformationen zu den Stundentafeln. Der Dank gilt insbesondere Frau Andrea Schwermer.

6.1 Baden-Württemberg

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Baden-Württemberg, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei knapp 5 Prozent (vgl. Tabelle 6.2). Mehr als 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen etwa 52 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 48 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Baden-Württemberg an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Baden-Württemberg bezogen auf die gesamte Schülerschaft bei 0,8 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei etwa 18 Prozent.

Einschulungsalter

Der Stichtag für die Einschulung wurde in Baden-Württemberg zwischen 2005 und 2007 schrittweise um drei Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während vor 2005 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seit 2007 Kinder als schulpflichtig, wenn sie bis einschließlich 30. September sechs Jahre alt werden. Für die Zielpopulation des Ländervergleichs 2011 galt diese neue Stichtagsregelung bereits, das heißt, die Kinder wurden in der Regel eingeschult, wenn sie bis einschließlich 30.9.2007 ihr sechstes Lebensjahr vollendet hatten.

Von den Kindern, die in Baden-Württemberg 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik etwa 3 Prozent vorzeitig und etwas weniger als 9 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt knapp 7 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 4 Monate alt.

Unterrichtszeit

Die Stundentafeln Baden-Württembergs für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 26 Jahreswochenstunden Unterricht erteilt werden und im Fach Mathematik mindestens 19 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen). In den ersten vier Schuljahren sollten also Schülerinnen und Schüler in Baden-Württemberg kumulativ mindestens 988 Stunden Deutsch- und 722 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Baden-Württemberg insgesamt rund 38 Prozent der

Kinder eine Ganztagschule, wobei es sich in nur relativ wenigen Fällen um eine voll gebundene Form handelte (3.2 %). Eine teilweise gebundene Ganztagschule besuchten etwa 7 Prozent und eine offene Ganztagschule rund 24 Prozent der Schülerinnen und Schüler. Für knapp 4 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.3 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Baden-Württemberg auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 10 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei knapp 6 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Baden-Württemberg im Lesen 21 Prozent und im Zuhören 17 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau des KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Baden-Württemberg fast 69 Prozent der Kinder im Lesen und etwa 77 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen etwa 11 Prozent und im Zuhören knapp 13 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Baden-Württemberg, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei etwas weniger als 10 Prozent. Weitere knapp 18 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik fast 73 Prozent der Viertklässlerinnen und Viertklässler in Baden-Württemberg. Herausragende Leistungen (Kompetenzstufe V) erzielten knapp 19 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Baden-Württemberg mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Zuhören* gehört Baden-Württemberg zu den sieben Ländern, in denen der Anteil von Kindern, die den Mindeststandard verfehlten, mit unter 7 Prozent am niedrigsten ist. Der Anteil derjenigen Kinder, die im Zuhören den Regelstandard oder bessere Leistungen erzielen konnten, ist bundesweit in Baden-Württemberg am zweithöchsten. In den anderen Kompetenzbereichen bewegen sich die Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen in Baden-Württemberg im oberen Mittelfeld.

Tabelle 6.3: Kompetenzstufenverteilung in Baden-Württemberg für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	10.3	21.0	31.7	25.9	11.2
	Zuhören	5.8	17.0	33.5	31.1	12.6
Mathematik	Global	9.5	17.8	28.4	25.5	18.8

6.2 Bayern

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Bayern, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei knapp 6 Prozent (vgl. Tabelle 6.2). Etwa 3 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen fast 49 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 51 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Bayern an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Bayern bezogen auf die gesamte Schülerschaft bei rund 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei 37 Prozent.

Einschulungsalter

Der Stichtag für die Einschulung wurde in Bayern zwischen 2005 und 2009 schrittweise vom 30. Juni auf den 30. November, insgesamt also um fünf Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Die letzten zwei Schritte der Senkung des Einschulungsalters wurden 2010 jedoch wieder zurückgenommen. Seit 2010 gelten demnach Kinder als schulpflichtig, wenn sie vor dem oder am 30. September sechs Jahre alt werden. Auch für die Zielpopulation des Ländervergleichs 2011 galt bereits diese Stichtagsregelung, das heißt, die Kinder wurden in der Regel eingeschult, wenn sie bis einschließlich 30. September 2007 ihr sechstes Lebensjahr vollendet hatten.

Von den Kindern, die in Bayern 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik 3 Prozent vorzeitig und 8 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt knapp 5 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 5 Monate alt.

Unterrichtszeit

Die Stundentafeln Bayerns für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–2 hinweg in den Fächern Deutsch, Mathematik, Heimat- und Sachunterricht, Musikerziehung und Kunsterziehung mindestens 32 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die zwei Jahrgangsstufen) Unterricht erteilt werden. In den Jahrgangsstufen 3–4 erhalten bayerische Grundschul Kinder zusammen 12 Jahreswochenstunden Unterricht im Fach Deutsch und 10 Jahreswochenstunden im Fach Mathematik. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Bayern kumulativ mindestens 456 *plus X* Stunden (aus der zweijährigen Eingangsphase) Deutsch- und 380 *plus X* Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Bayern insgesamt knapp 24 Prozent der Kinder eine Ganztagschule, wobei es sich in 10 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten knapp 4 Prozent und eine offene Ganztagschule etwa 6 Prozent der Schülerinnen und Schüler. Für rund 4 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.4 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Bayern auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Knapp 10 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei weniger als 6 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Bayern im Lesen etwa 17 Prozent und im Zuhören rund 15 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Bayern etwa 74 Prozent der Kinder im Lesen und 79 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen knapp 16 Prozent und im Zuhören rund 15 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Bayern, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei knapp 8 Prozent. Rund 15 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik 77 Prozent der Viertklässlerinnen und Viertklässler in Bayern. Herausragende Leistungen (Kompetenzstufe V) erzielten fast 21 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Bayern mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Lesen* ist der zweithöchste Anteil von Schülerinnen und Schülern zu verzeichnen, die den Regelstandard erreichten. Die höchste Kompetenzstufe (Kompetenzstufe V) ist im Lesen in keinem Land stärker besetzt als in Bayern. Im Kompetenzbereich *Zuhören* gehört Bayern zu den sieben Ländern, in denen der Anteil von Kindern, die den Mindeststandard verfehlten,

Tabelle 6.4: Kompetenzstufenverteilung in Bayern für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	9.7	16.9	29.2	28.4	15.9
	Zuhören	5.8	15.2	32.1	32.2	14.7
Mathematik	Global	7.9	15.2	28.2	27.9	20.9

mit unter 7 Prozent am niedrigsten ist. Der Anteil der Schülerinnen und Schüler, die im Bereich Zuhören den Regelstandard erreichten, ist in Bayern am größten. Auch die höchste Kompetenzstufe erreichten im Kompetenzbereich Zuhören in keinem Land mehr Schülerinnen und Schüler als in Bayern. Im Fach *Mathematik* fällt die Verteilung der Kinder auf die Kompetenzstufen ebenfalls sehr günstig aus. Der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlten, ist nur in einem anderen Land etwas geringer ausgeprägt, während in keinem anderen Land mehr Kinder den Regelstandard erreichten als in Bayern. Auch die höchste Kompetenzstufe ist in Bayern (mit gleichem Anteilswert wie in Sachsen-Anhalt) stärker besetzt als in den anderen Ländern.

6.3 Berlin

Berlin nimmt zusammen mit Brandenburg eine Sonderstellung hinsichtlich der Regeldauer der Primarschulzeit ein. Anders als in den anderen 14 Ländern der Bundesrepublik Deutschland erfolgt in Berlin und Brandenburg der Übergang von der Primarschule auf eine weiterführende Schule in der Regel erst nach sechs und nicht schon nach vier Jahren (Lohauß, Rehkämper, Rockmann & Wendt, 2011).⁸ Dieser Umstand kann für die Interpretation der Kompetenzergebnisse insofern von Bedeutung sein, als in Berlin und Brandenburg zum Testzeitpunkt, anders als in den 14 anderen Ländern, keine positiven Lern- und Leistungseffekte zu erwarten sind, die mit der anstehenden Übergangsentscheidung („Endspurteffekte“) verbunden sein könnten.⁹ Die Quantifizierung des potenziellen Effekts einer vierjährigen gegenüber einer sechsjährigen Primarschulzeit auf die Kompetenzergebnisse am Ende der vierten Jahrgangsstufe ist anhand der vorliegenden Daten allerdings nicht möglich.

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Berlin, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei knapp 6 Prozent (vgl. Tabelle 6.2). Nur etwa ein Prozent *aller* Kinder der vierten Jahrgangsstufe besuchte eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen 21 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden 79 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Berlin an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Berlin bezogen auf die gesamte Schülerschaft bei 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei knapp 34 Prozent.

8 In beiden Ländern ist in Ausnahmefällen auch ein Übergang in eine weiterführende Schule bereits nach der vierten Klasse möglich. Dies erfolgt jedoch jeweils nur bei einer Minderheit (< 10 %) der Schülerinnen und Schüler

9 In einer Schweizer Studie zur Kompetenzentwicklung im Fach Französisch wurden motivationale Effekte auf den Lernzuwachs bedingt durch einen möglichen Wechsel der Schulart bei Erreichen eines bestimmten Notendurchschnitts diskutiert (Neumann et al., 2007).

Einschulungsalter

Anders als in den meisten anderen Ländern, die in den letzten sieben Jahren den Stichtag für die Einschulung vorgezogen haben, wurde das Einschulungsalter in Berlin nicht schrittweise, sondern einmalig um sechs Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während vor 2005 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seit 2005 Kinder als schulpflichtig, wenn sie vor dem oder am 31. Dezember sechs Jahre alt werden. Berlin hat damit derzeit unter allen Ländern den am weitesten vorgezogenen Stichtag für den Beginn der regulären Schulpflicht. Für die Zielpopulation des Ländervergleichs 2011 galt diese neue Stichtagsregelung bereits, das heißt, die Kinder wurden in der Regel eingeschult, wenn sie bis einschließlich 31. Dezember 2007 ihr sechstes Lebensjahr vollendet hatten.

Von den Kindern, die in Berlin 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik etwa 2 Prozent vorzeitig und 3 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt knapp 6 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 2 Monate alt. In Berlin war somit die untersuchte Population im Vergleich zu den anderen Ländern am jüngsten.

Unterrichtszeit

Die Stundentafeln Berlins für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 26 Jahreswochenstunden und im Fach Mathematik mindestens 20 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden.¹⁰ In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Berlin kumulativ mindestens 988 Stunden Deutsch- und 760 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Berlin insgesamt knapp 61 Prozent der Kinder eine Ganztagschule, wobei es sich in etwas weniger als 12 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten knapp 3 Prozent und eine offene Ganztagschule rund 42 Prozent der Schülerinnen und Schüler. Für etwa 4 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor. Insgesamt gehört Berlin zu den vier Ländern, in denen die Schulen von mehr als 60 Prozent der Viertklässlerinnen und Viertklässler angegeben haben, ein Ganztagsangebot zu haben.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.5 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Berlin auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 22 Prozent der zielgleich unterrichteten Viertklässlerinnen

¹⁰ Diese Angaben beziehen sich auf die für die Jahrgangsstufen 1 und 2 empfohlenen Richtwerte für Deutsch beziehungsweise Mathematik innerhalb des Gesamtkontingents für die Fächer Deutsch, Sachunterricht und Mathematik.

und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei über 15 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Berlin im Lesen etwa 24 Prozent und im Zuhören 23 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Berlin etwa 54 Prozent der Kinder im Lesen und rund 62 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen etwa 7 Prozent und im Zuhören rund 10 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Berlin, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei fast 27 Prozent. Knapp 26 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik rund 47 Prozent der Viertklässlerinnen und Viertklässler in Berlin. Herausragende Leistungen (Kompetenzstufe V) erzielten etwa 7 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Berlin mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Lesen* ist der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlten, in Berlin am größten. Auf der anderen Seite des Leistungsspektrums – das Erreichen des Regelstandards wie auch der Kompetenzstufe V – gibt es nur jeweils ein Land, das noch geringere Anteile aufweist. Auch im Bereich *Zuhören* verfehlte in Berlin ein größerer Anteil von Schülerinnen und Schülern den Mindeststandard als in allen anderen Ländern. Der Anteil der Kinder, deren Leistungen mindestens dem Regelstandard entsprechen, ist im Kompetenzbereich Zuhören wiederum in Berlin am zweitniedrigsten. Etwas positiver stellt sich im Bereich Zuhören hingegen die Situation am oberen Ende der Leistungsverteilung dar (Kompetenzstufe V). Hier ist der Anteil in Berlin etwas höher als in vier anderen Ländern. Im Kompetenzbereich *Mathematik* hingegen sind die Ergebnisse für Berlin durchgängig ungünstiger als in den anderen Ländern. Der Anteil der Kinder, die den Mindeststandard verfehlten, ist hier am höchsten, während besonders wenige Schülerinnen und Schüler den Regelstandard erreichten. Auch die oberste Kompetenzstufe ist in keinem anderen Land so dünn besetzt wie in Berlin.

Tabelle 6.5: Kompetenzstufenverteilung in Berlin für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	22.2	24.1	27.3	19.0	7.4
	Zuhören	15.4	23.0	28.5	23.3	9.7
Mathematik	Global	26.6	25.9	23.6	16.5	7.3

6.4 Brandenburg

Brandenburg nimmt zusammen mit Berlin eine Sonderstellung hinsichtlich der Regeldauer der Primarschulzeit ein. Anders als in den anderen 14 Ländern der Bundesrepublik Deutschland erfolgt in Brandenburg und Berlin der Übergang von der Primarschule auf eine weiterführende Schule in der Regel erst nach sechs und nicht schon nach vier Jahren (Lohauß et al., 2011).¹¹ Dieser Umstand kann für die Interpretation der Kompetenzergebnisse insofern von Bedeutung sein, als in Brandenburg und Berlin zum Testzeitpunkt, anders als in den 14 anderen Ländern, keine allgemein positiven motivationalen Lern- und Leistungseffekte zu erwarten sind, die mit der anstehenden Übergangentscheidung („Endspurteffekte“) verbunden sein könnten.¹² Die Quantifizierung des potenziellen Effekts einer vierjährigen gegenüber einer sechsjährigen Primarschulzeit auf die Kompetenzergebnisse am Ende der vierten Jahrgangsstufe ist anhand der vorliegenden Daten allerdings nicht möglich.

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Brandenburg, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei fast 8 Prozent (vgl. Tabelle 6.2). Knapp 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 24 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden fast 76 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Brandenburg an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Brandenburg bezogen auf die gesamte Schülerschaft bei gut 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei etwa 30 Prozent.

Einschulungsalter

Anders als in den meisten anderen Ländern, die in den letzten sieben Jahren den Stichtag für die Einschulung vorgezogen haben, wurde das Einschulungsalter in Brandenburg nicht schrittweise, sondern einmalig um drei Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während vor 2005 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seit 2005 Kinder als schulpflichtig, wenn sie vor dem oder am 30. September sechs Jahre alt werden. Für die Zielpopulation des Ländervergleichs 2011 galt diese neue Stichtagsregelung bereits, das heißt, die Kinder wurden in der Regel eingeschult, wenn sie bis einschließlich 30. September 2007 ihr sechstes Lebensjahr vollendet hatten.

11 In beiden Ländern ist in Ausnahmefällen auch ein Übergang in eine weiterführende Schule bereits nach der vierten Klasse möglich. Dies erfolgt jedoch jeweils nur bei einer Minderheit (< 10 %) der Schülerinnen und Schüler

12 In einer Schweizer Studie zur Kompetenzentwicklung im Fach Französisch wurden motivationale Effekte auf den Lernzuwachs bedingt durch einen möglichen Wechsel der Schulform bei Erreichen eines bestimmten Notendurchschnitts diskutiert (Neumann et al., 2007).

Von den Kindern, die in Brandenburg 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 3 Prozent vorzeitig und rund 13 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwa 8 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 5 Monate alt.

Unterrichtszeit

Die Stundentafeln Brandenburgs für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 25 Jahreswochenstunden und im Fach Mathematik mindestens 18 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Brandenburg kumulativ mindestens 950 Stunden Deutsch- und 684 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Brandenburg insgesamt etwa 40 Prozent der Kinder eine Ganztagschule, wobei es sich in etwa 7 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten 11 Prozent und eine offene Ganztagschule etwa 19 Prozent der Schülerinnen und Schüler. Für rund 2 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.¹³

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.6 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Brandenburg auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Rund 13 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei knapp 9 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Brandenburg sowohl im Lesen als auch im Zuhören etwa 21 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Brandenburg etwa 66 Prozent der Kinder im Lesen und fast 70 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen rund 11 Prozent und im Zuhören etwa 10 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Brandenburg, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei knapp 13 Prozent. Weitere rund 23 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik knapp 65 Prozent der Viertklässlerinnen und Viertklässler in Brandenburg. Herausragende Leistungen (Kompetenzstufe V) erzielten fast 14 Prozent der Schülerinnen und Schüler.

¹³ Aufgrund von Rundungsungenauigkeiten ergibt die Summe der Prozentangaben für die drei Ganztagsformen (inkl. des Anteils fehlender Angaben) nicht immer genau den Prozentsatz von Kindern, die insgesamt eine Ganztagschule besuchten.

Tabelle 6.6: Kompetenzstufenverteilung in Brandenburg für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	12.5	21.3	30.4	24.5	11.3
	Zuhören	8.8	21.3	33.2	26.9	9.8
Mathematik	Global	12.8	22.5	28.0	23.1	13.6

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Brandenburg mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Zuhören* ist Brandenburg das Flächenland mit dem zweithöchsten Anteil von Kindern, die den Mindeststandard verfehlten. Der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichten, ist im Zuhören in nur einem Flächenland geringer als in Brandenburg. In den anderen Kompetenzbereichen bewegen sich in Brandenburg die Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen weitgehend im Mittelfeld.

6.5 Bremen

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Bremen, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei 2 Prozent (vgl. Tabelle 6.2). Dies ist mit Abstand der niedrigste Wert unter den 16 Ländern. Nur 0.1 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 7 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurde mit fast 93 Prozent der weit überwiegende Anteil der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Bremen an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *ziel-different* unterrichtet wurden, lag in Bremen bezogen auf die gesamte Schülerschaft bei 0.2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei knapp 9 Prozent.

Einschulungsalter

Bremen gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Bremen eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die in Bremen 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 18 Prozent vorzeitig und 2 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt knapp 9 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 5 Monate alt.

Unterrichtszeit

Die Stundentafeln Bremens für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 23 Jahreswochenstunden und im Fach Mathematik mindestens 22 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Bremen kumulativ mindestens 874 Stunden Deutsch- und 836 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Bremen insgesamt knapp 24 Prozent der Kinder eine Ganztagschule, wobei es sich in etwas weniger als 17 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten rund 2 Prozent und eine offene Ganztagschule etwa 4 Prozent der Schülerinnen und Schüler. Für etwas mehr als ein Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.7 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Bremen auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Rund 21 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei etwas weniger als 14 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Bremen im Lesen etwa 27 Prozent und im Zuhören rund 26 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Bremen rund 52 Prozent der Kinder im Lesen und knapp 61 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen fast 7 Prozent und im Zuhören rund 9 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Bremen, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei 25 Prozent. Weitere rund 25 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichten oder übertrafen (Kompetenzstufe III, IV oder V) in Mathematik etwa 50 Prozent der Viertklässlerinnen und Viertklässler in Bremen. Herausragende Leistungen (Kompetenzstufe V) erzielten fast 9 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Bremen mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Lesen* ist der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlten, in Bremen am zweithöchsten. Auf der anderen Seite des Leistungsspektrums – das Erreichen des Regelstandards wie auch der Kompetenzstufe V – gibt es kein Land, das noch geringere Anteile aufweist. Auch im Bereich *Zuhören* verfehlte nur noch in einem Land ein größerer Anteil von Schülerinnen und Schülern den Mindeststandard. Der Anteil der Kinder, deren Leistungen mindestens dem Regelstandard entsprechen, ist im Kompetenzbereich *Zuhören* in Bremen wiederum am niedrigsten. Etwas positiver stellt sich im Bereich *Zuhören* hingegen die Situation am oberen Ende der Leistungsverteilung dar. Hier ist der Anteil in Bremen etwas höher als in zwei anderen Ländern. Im Kompetenzbereich *Mathematik* hingegen sind die Ergebnisse jeweils nur noch in einem Land ungünstiger als in Bremen. Der Anteil der Kinder, die den Mindeststandard verfehlten, ist hier am zweithöchsten, während in Bezug auf das Erreichen des Regelstandards und der Kompetenzstufe V jeweils die zweitniedrigste Quote zu verzeichnen ist.

Tabelle 6.7: Kompetenzstufenverteilung in Bremen für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	21.3	26.6	27.1	18.2	6.8
	Zuhören	13.7	25.8	29.0	22.2	9.3
Mathematik	Global	25.0	24.8	24.3	17.2	8.7

6.6 Hamburg

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Hamburg, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei etwa 5 Prozent (vgl. Tabelle 6.2). Etwa ein Prozent *aller* Kinder der vierten Jahrgangsstufe besuchte eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 21 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden rund 79 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Hamburg an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Hamburg bezogen auf die gesamte Schülerschaft bei etwa 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei rund 36 Prozent.

Einschulungsalter

Hamburg gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Hamburg eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die in Hamburg 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik etwas mehr als 13 Prozent vorzeitig und rund 3 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt knapp 7 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 5 Monate alt.

Unterrichtszeit

Die Stundentafeln Hamburgs für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch und Mathematik jeweils 18 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden.¹⁴ In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Hamburg kumulativ mindestens je 684 Stunden Deutsch- und Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Hamburg insgesamt etwa 18 Prozent der Kinder eine Ganztagschule, wobei es sich in knapp 7 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten knapp 3 Prozent und eine offene Ganztagschule 9 Prozent der Schülerinnen und Schüler.¹⁵

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.8 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Hamburg auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 18 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei knapp 12 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Hamburg im Lesen rund 23 Prozent und im Zuhören fast 21 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Hamburg knapp 59 Prozent der Kinder im Lesen und fast 68 Prozent im Zuhören. Die

14 Die Verordnung zur Einführung der Primarschule, der Stadtteilschule und des sechsstufigen Gymnasiums im Schuljahr 2010/2011 (VOE-PSG 2010/2011) sah vor, dass in der Primarstufe jeweils 27 Wochenstunden in Deutsch und Mathematik erteilt werden, von denen 9 Wochenstunden auf die Jahrgangsstufen 5 und 6 entfallen.

15 Aufgrund von Rundungsungenauigkeiten ergibt die Summe der Prozentangaben für die drei Ganztagsformen (inkl. des Anteils fehlender Angaben) nicht immer genau den Prozentsatz von Kindern, die insgesamt eine Ganztagschule besuchten.

Tabelle 6.8: Kompetenzstufenverteilung in Hamburg für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	18.3	23.1	29.1	21.3	8.2
	Zuhören	11.5	20.9	29.5	26.8	11.4
Mathematik	Global	20.5	23.5	26.3	19.3	10.4

höchste Kompetenzstufe erreichten im Lesen etwa 8 Prozent und im Zuhören etwas mehr als 11 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Hamburg, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei etwa 21 Prozent. Weitere knapp 24 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichten oder übertrafen (Kompetenzstufe III, IV oder V) in Mathematik 56 Prozent der Viertklässlerinnen und Viertklässler in Hamburg. Herausragende Leistungen (Kompetenzstufe V) erzielten etwas mehr als 10 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Hamburg mit denen der anderen 15 Länder, so fällt Folgendes auf:

Hamburg bildet in der Regel gemeinsam mit den anderen beiden Stadtstaaten die Gruppe mit den am wenigsten günstigen Ergebnissen. In einigen Bereichen setzt sich Hamburg innerhalb dieser Gruppe aber in positiver Richtung ab. So ist in Hamburg zwar der dritthöchste Anteil der Schülerinnen und Schüler zu finden, die im Kompetenzbereich *Lesen* nicht den Mindeststandard erreichten. Dennoch weist Hamburg etwa 4 Prozentpunkte Abstand zu dem Land mit dem höchsten Prozentsatz an Kindern unterhalb des Mindeststandards auf. Ebenso wird der Regelstandard nur in zwei Ländern seltener erreicht als in Hamburg, die Quote liegt hier jedoch um fast 7 Prozentpunkte höher als im Land mit dem geringsten Anteil. Auch die Gruppe von Kindern, die Kompetenzstufe V erreicht haben, ist etwas größer als in den beiden anderen Stadtstaaten.

Dieses Muster wiederholt sich in ähnlicher Weise in den Kompetenzbereichen *Zuhören* und *Mathematik*. Auch hier befindet sich Hamburg zumeist in der Gruppe der drei schwächsten Länder mit jedoch jeweils etwas günstigeren Ergebnissen als die anderen beiden Stadtstaaten. Eine Ausnahme bildet die Gruppe von Kindern mit herausragenden Ergebnissen im Zuhören. Hier gehört Hamburg zu den insgesamt sechs Ländern, in denen der höchste Niveaubereich (Optimalstandard) am stärksten besetzt ist.

6.7 Hessen

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Hessen, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei etwa 4 Prozent (vgl. Tabelle 6.2). Knapp 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen etwa 40 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden knapp 60 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Hessen an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Hessen bezogen auf die gesamte Schülerschaft bei etwa 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei rund 50 Prozent.

Einschulungsalter

Hessen gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Hessen eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die in Hessen 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 13 Prozent vorzeitig und 10 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwas mehr als 6 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 6 Monate alt.

Unterrichtszeit

Die Stundentafeln Hessens für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 22 Jahreswochenstunden und im Fach Mathematik mindestens 20 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden.¹⁶ In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Hessen kumulativ mindestens 836 Stunden Deutsch- und 760 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Hessen insgesamt knapp 29 Prozent der Kinder eine Ganztagschule, wobei es sich in keinem der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten etwa 3 Prozent und eine

¹⁶ Die Angaben beziehen sich auf die Verordnung über die Stundentafeln für die Primarstufe und die Sekundarstufe I vom 05. September 2011. Die Anzahl der Jahreswochenstunden im Schuljahr 2010/2011 entspricht den Angaben in dieser Verordnung.

offene Ganztagschule etwas mehr als 23 Prozent der Schülerinnen und Schüler. Für etwa 3 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.9 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Hessen auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 13 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei etwas über 8 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Hessen im Lesen etwas mehr als 22 Prozent und im Zuhören rund 21 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Hessen fast 65 Prozent der Kinder im Lesen und fast 71 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen und Zuhören jeweils rund 11 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Hessen, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei etwas mehr als 14 Prozent. Weitere knapp 24 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichten oder übertrafen (Kompetenzstufe III, IV oder V) in Mathematik etwa 62 Prozent der Viertklässlerinnen und Viertklässler in Hessen. Herausragende Leistungen (Kompetenzstufe V) erzielten rund 12 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Hessen mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Mathematik* ist der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlten, in keinem anderen Flächenland größer als in Hessen. Ebenso ist der Anteil der Kinder, die den Regelstandard erreichten, in Hessen kleiner als in den anderen Flächenländern. Der obere Leistungsbereich (Kompetenzstufe V) ist in nur einem Flächenland geringer besetzt als in Hessen. In den Kompetenzbereichen *Lesen* und *Zuhören* bewegen sich die Verteilungen der hessischen Schülerinnen und Schüler auf die Kompetenzstufen weitgehend im unteren Mittelfeld.

Tabelle 6.9: Kompetenzstufenverteilung in Hessen für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	13.1	22.4	30.5	23.5	10.6
	Zuhören	8.3	20.9	32.1	27.4	11.3
Mathematik	Global	14.4	23.5	27.7	22.1	12.3

6.8 Mecklenburg-Vorpommern

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Mecklenburg-Vorpommern, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei mehr als 11 Prozent (vgl. Tabelle 6.2). Dies ist mit Abstand der höchste Wert aller 16 Länder. Etwas weniger als 5 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 39 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden knapp 61 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Mecklenburg-Vorpommern an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Mecklenburg-Vorpommern bezogen auf die gesamte Schülerschaft bei etwa 4 Prozent, was ebenfalls höher ist als in allen anderen Ländern. Bezogen auf Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf liegt der Anteil *zieldifferent* unterrichteter Kinder bei rund 36 Prozent.

Einschulungsalter

Mecklenburg-Vorpommern gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Mecklenburg-Vorpommern eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die in Mecklenburg-Vorpommern 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 3 Prozent vorzeitig und etwa 6 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwas weniger als 6 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 7 Monate alt.

Unterrichtszeit

Die Stundentafeln Mecklenburg-Vorpommerns für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 26 bis 28 Jahreswochenstunden und im Fach Mathematik mindestens 22 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Mecklenburg-Vorpommern kumulativ mindestens 988 bis 1064 Stunden Deutsch- und 836 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Mecklenburg-Vorpommern insgesamt knapp 28 Prozent der Kinder eine Ganztagschule, wobei es sich in knapp 11 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten etwa ein Prozent und eine offene Ganztagschule etwas mehr als 15 Prozent der Schülerinnen und Schüler.¹⁷

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.10 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Mecklenburg-Vorpommern auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwas über 9 Prozent der zugleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei 6 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Mecklenburg-Vorpommern im Lesen knapp 21 Prozent und im Zuhören etwa 18 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Mecklenburg-Vorpommern etwa 70 Prozent der Kinder im Lesen und knapp 76 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen gut 10 Prozent und im Zuhören etwa 11 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt in Mecklenburg-Vorpommern der Anteil der Schülerinnen und Schüler, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei knapp 11 Prozent. Weitere rund 20 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik mehr als 69 Prozent der Viertklässlerinnen und Viertklässler in Mecklenburg-Vorpommern. Herausragende Leistungen (Kompetenzstufe V) erzielten knapp 16 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Mecklenburg-Vorpommern mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Lesen* gehört Mecklenburg-Vorpommern zu den drei Ländern, in denen der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlten, am geringsten ist. Gleichzeitig ist jedoch der Anteil der Schülerinnen und Schüler, die dem Optimalstandard entsprechende Leistungen erzielten (Kompetenzstufe V), in keinem anderen Flächenland geringer als in Mecklenburg-Vorpommern. Im Kompetenzbereich *Zuhören* gehört Mecklenburg-Vorpommern zu den sieben Ländern, in denen der Anteil von Kindern, die den Mindeststandard verfehlten, mit unter 7 Prozent am niedrigsten ist. Darüber hinaus nimmt Mecklenburg-Vorpommern den vierten Rang in Bezug auf den Anteil der Viertklässlerinnen und Viertklässler ein, die den Regelstandard im Zuhören erreichen oder übertreffen konnten. In *Mathematik* bewegt sich die Verteilung der Schülerinnen und Schüler in Mecklenburg-Vorpommern auf die Kompetenzstufen im oberen Mittelfeld.

¹⁷ Aufgrund von Rundungsungenauigkeiten ergibt die Summe der Prozentangaben für die drei Ganztagsformen (inkl. des Anteils fehlender Angaben) nicht immer genau den Prozentsatz von Kindern, die insgesamt eine Ganztagschule besuchten.

Tabelle 6.10: Kompetenzstufenverteilung in Mecklenburg-Vorpommern für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	9.1	20.6	33.2	27.0	10.1
	Zuhören	6.0	18.1	34.9	30.0	10.9
Mathematik	Global	10.8	20.2	28.8	24.8	15.5

6.9 Niedersachsen

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Niedersachsen, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei etwas mehr als 3 Prozent (vgl. Tabelle 6.2). Knapp 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen fast 50 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa die Hälfte der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Niedersachsen an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zielfferent* unterrichtet wurden, lag in Niedersachsen bezogen auf die gesamte Schülerschaft bei knapp 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei 51 Prozent.

Einschulungsalter

Der Stichtag für die Einschulung wurde in Niedersachsen zwischen 2010 und 2012 schrittweise um drei Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während vor 2010 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seit 2012 Kinder als schulpflichtig, wenn sie vor dem oder am 30. September sechs Jahre alt werden. Für die Zielpopulation des Ländervergleichs 2011 galt allerdings noch die alte Stichtagsregelung, wonach Kinder in der Regel eingeschult wurden, wenn sie bis einschließlich 30.6.2007 ihr sechstes Lebensjahr vollendet hatten.

Von den Kindern, die in Niedersachsen 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik rund 8 Prozent vorzeitig und knapp 6 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt rund 9 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 6 Monate alt.

Unterrichtszeit

Die Stundentafeln Niedersachsens für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 24 Jahreswochenstunden und im Fach Mathematik mindestens 20 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Niedersachsen kumulativ mindestens 912 Stunden Deutsch- und 760 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Niedersachsen insgesamt rund 16 Prozent der Kinder eine Ganztagschule, wobei es sich ausschließlich um offene Ganztagsangebote handelte. Für rund 3 Prozent der Kinder lagen allerdings keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.11 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Niedersachsen auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 12 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei knapp 7 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Niedersachsen im Lesen rund 21 Prozent und im Zuhören 18 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Niedersachsen rund 67 Prozent der Kinder im Lesen und etwa 75 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen etwa 13 Prozent und im Zuhören etwas mehr als 14 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Niedersachsen, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei etwas weniger als 12 Prozent. Weitere 23 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik knapp 66 Prozent der Viertklässlerinnen und Viertklässler in Niedersachsen. Herausragende Leistungen (Kompetenzstufe V) erzielten fast 15 Prozent der Schülerinnen und Schüler.

Tabelle 6.11: Kompetenzstufenverteilung in Niedersachsen für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	11.9	20.8	29.1	24.8	13.4
	Zuhören	6.7	18.0	31.2	29.9	14.1
Mathematik	Global	11.5	23.0	28.0	22.9	14.6

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Niedersachsen mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Zuhören* gehört Niedersachsen zu den sieben Ländern, in denen der Anteil von Kindern, die den Mindeststandard verfehlten, mit unter 7 Prozent am niedrigsten ist. Nur in einem anderen Land erreichten mehr Schülerinnen und Schüler im Zuhören die höchste Kompetenzstufe V. In den anderen Kompetenzbereichen bewegen sich die Verteilungen der Schülerinnen und Schüler in Niedersachsen auf die Kompetenzstufen im Mittelfeld, teilweise auch im oberen Mittelfeld.

6.10 Nordrhein-Westfalen

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Nordrhein-Westfalen, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei knapp 4 Prozent (vgl. Tabelle 6.2). Nicht ganz 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 42 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 58 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Nordrhein-Westfalen an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Nordrhein-Westfalen bezogen auf die gesamte Schülerschaft bei rund einem Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei etwas mehr als 29 Prozent.

Einschulungsalter

Der Stichtag für die Einschulung wurde in Nordrhein-Westfalen zwischen 2007 und 2011 schrittweise um 3 Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während vor 2007 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seit 2011 Kinder als schulpflichtig, wenn sie vor dem oder am 30. September sechs Jahre alt werden. Für die Zielpopulation des Ländervergleichs 2011 galt als Stichtag der 31. Juli, das heißt, die Kinder wurden in der Regel eingeschult, wenn sie bis einschließlich 31.7.2007 ihr sechstes Lebensjahr vollendet hatten.

Von den Kindern, die in Nordrhein-Westfalen 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 8 Prozent vorzeitig und weniger als ein Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwa 10 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 6 Monate alt.

Unterrichtszeit

Nach den Stundentafeln Nordrhein-Westfalens für das Schuljahr 2010/2011 und nach der neuen Ausbildungsordnung für die Grundschule werden in den Primarschulen über die Jahrgangsstufen 1–4 hinweg die Fächer Deutsch, Mathematik, Sachunterricht und Förderunterricht zusammen mit 53 bis 55 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) unterrichtet.¹⁸ Im Schulgesetz heißt es in § 3 dazu, dass der Unterricht fächerübergreifend auszurichten ist. Erfahrungswerte weisen darauf hin, dass in der Schuleingangsphase sowie den Jahrgangsstufen 3 und 4 insgesamt jeweils 18 Jahreswochenstunden in den Fächern Deutsch und Mathematik erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Nordrhein-Westfalen kumulativ jeweils 684 Stunden Deutsch- und Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Nordrhein-Westfalen insgesamt rund 74 Prozent der Kinder eine Ganztagschule. Dabei handelte es sich in nur wenigen Fällen um eine voll gebundene Form (0,5 %). Eine teilweise gebundene Ganztagschule besuchten etwa 4 Prozent und eine offene Ganztagschule etwas mehr als 67 Prozent der Schülerinnen und Schüler. Für knapp 3 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor. Insgesamt gehört Nordrhein-Westfalen zu den vier Ländern, in denen die Schulen von mehr als 60 Prozent der Viertklässlerinnen und Viertklässler angegeben haben, ein Ganztagsangebot zu haben.¹⁹

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.12 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Nordrhein-Westfalen auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Insgesamt rund 14 Prozent der zugleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei 8 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Nordrhein-Westfalen im Lesen etwa 23 Prozent und im Zuhören knapp 21 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Nordrhein-Westfalen knapp 63 Prozent der Kinder im Lesen und etwa 71 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen und im Zuhören jeweils knapp 11 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Nordrhein-Westfalen, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei 13 Prozent. Weitere knapp 22 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen

¹⁸ Diese Angaben beziehen sich auf die Verordnung über den Bildungsgang in der Grundschule (Ausbildungsordnung Grundschule – AO-GS) vom 23. März 2005, zuletzt geändert durch die Verordnung vom 17. Februar 2012.

¹⁹ Aufgrund von Rundungsungenauigkeiten ergibt die Summe der Prozentangaben für die drei Ganztagsformen (inkl. des Anteils fehlender Angaben) nicht immer genau den Prozentsatz von Kindern, die insgesamt eine Ganztagschule besuchten.

Tabelle 6.12: Kompetenzstufenverteilung in Nordrhein-Westfalen für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	14.2	23.2	29.7	22.2	10.7
	Zuhören	8.0	20.8	32.9	27.4	10.9
Mathematik	Global	13.0	21.7	28.1	22.1	15.2

(Kompetenzstufe III, IV oder V) konnten in Mathematik rund 65 Prozent der Viertklässlerinnen und Viertklässler in Nordrhein-Westfalen. Herausragende Leistungen (Kompetenzstufe V) erzielten etwa 15 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Nordrhein-Westfalen mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Lesen* ist der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlten, in nur einem Flächenland größer als in Nordrhein-Westfalen. Der Anteil der Kinder, die den Regelstandard erreichten oder übertrafen, ist im Vergleich zu den anderen Flächenländern am geringsten. In den anderen Kompetenzbereichen bewegen sich die Verteilungen der Schülerinnen und Schüler in Nordrhein-Westfalen auf die Kompetenzstufen weitgehend im Mittelfeld.

6.11 Rheinland-Pfalz

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Rheinland-Pfalz, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei etwa 3 Prozent (vgl. Tabelle 6.2). Weniger als 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 47 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 53 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Rheinland-Pfalz an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferenter* unterrichtet wurden, lag in Rheinland-Pfalz bezogen auf die gesamte Schülerschaft bei rund 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei fast 69 Prozent.

Einschulungsalter

Der Stichtag für die Einschulung wurde in Rheinland-Pfalz 2008 um zwei Monate vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während

vor 2008 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seit 2008 Kinder als schulpflichtig, wenn sie vor dem oder am 31. August sechs Jahre alt werden. Als Stichtag galt für die rheinland-pfälzische Zielpopulation des Ländervergleichs 2011, die 2007 ins schulpflichtige Alter kam, demnach noch der 30. Juni.

Von den Kindern, die in Rheinland-Pfalz 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 12 Prozent vorzeitig und etwas über 7 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwas weniger als 8 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt 10 Jahre und knapp 6 Monate alt.

Unterrichtszeit

Die Stundentafeln in Rheinland-Pfalz für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg in den Fächern Deutsch und Sachkunde mindestens 31 Jahreswochenstunden²⁰ und im Fach Mathematik mindestens 18 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Rheinland-Pfalz kumulativ mindestens 1 178 Stunden Unterricht in Deutsch und Sachkunde und 684 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Rheinland-Pfalz insgesamt 39 Prozent der Kinder eine Ganztagschule, wobei es sich in nur relativ wenigen Fällen um eine voll gebundene Form handelte (1,6 %). Eine teilweise gebundene Ganztagschule besuchten etwa 19 Prozent und eine offene Ganztagschule etwas mehr als 16 Prozent der Schülerinnen und Schüler. Für rund 2 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.13 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Rheinland-Pfalz auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwas über 14 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei etwa 8 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Rheinland-Pfalz im Lesen knapp 22 Prozent und im Zuhören ebenfalls rund 22 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Rheinland-Pfalz 64 Prozent der Kinder im Lesen und etwa 70 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen knapp 11 Prozent und im Zuhören etwa 8 Prozent der Schülerinnen und Schüler.

²⁰ In Rheinland-Pfalz umfasst eine Unterrichtsstunde 50 Minuten.

Tabelle 6.13: Kompetenzstufenverteilung in Rheinland-Pfalz für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	14.3	21.7	29.8	23.8	10.5
	Zuhören	8.1	21.5	34.5	27.5	8.4
Mathematik	Global	13.1	21.6	27.3	24.0	14.0

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Rheinland-Pfalz, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei rund 13 Prozent. Knapp 22 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichten oder übertrafen (Kompetenzstufe III, IV oder V) in Mathematik ungefähr 65 Prozent der Viertklässlerinnen und Viertklässler in Rheinland-Pfalz. Herausragende Leistungen (Kompetenzstufe V) erzielten 14 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Rheinland-Pfalz mit denen der anderen 15 Länder, so fällt Folgendes auf:

Der Anteil der Kinder, die im Kompetenzbereich *Lesen* den Mindeststandard verfehlten, ist unter allen Flächenländern der höchste. Auf der anderen Seite des Leistungsspektrums – das Erreichen der Regelstandards wie auch der Kompetenzstufe V – gibt es jeweils nur ein Flächenland, das im Lesen noch geringere Anteile aufweist. Im *Zuhören* ist insbesondere die höchste Kompetenzstufe in Rheinland-Pfalz schwächer besetzt als in allen anderen Ländern der Bundesrepublik Deutschland. In *Mathematik* bewegt sich die Verteilung der Schülerinnen und Schüler in Rheinland-Pfalz auf die Kompetenzstufen weitgehend im Mittelfeld.

6.12 Saarland

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe im Saarland, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei etwas über 4 Prozent (vgl. Tabelle 6.2). Knapp 2 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen rund 40 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden knapp 60 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 im Saarland an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag im Saarland bezogen auf die

gesamte Schülerschaft bei etwa 3 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei circa 65 Prozent.

Einschulungsalter

Das Saarland gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden im Saarland eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die im Saarland 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik 8 Prozent vorzeitig und knapp 6 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt 8 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 6 Monate alt.

Unterrichtszeit

Die Stundentafeln des Saarlands für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg sowohl im Fach Deutsch als auch im Fach Mathematik jeweils mindestens 20 Jahreswochenstunden Unterricht erteilt werden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen). In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler im Saarland kumulativ mindestens je 760 Stunden Deutsch- und Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten im Saarland insgesamt etwas weniger als 48 Prozent der Kinder eine Ganztagschule, wobei es sich in nur relativ wenigen Fällen um eine voll gebundene Form handelte (1,3 %). Eine teilweise gebundene Ganztagschule besuchten etwa 9 Prozent und eine offene Ganztagschule knapp 35 Prozent der Schülerinnen und Schüler. Für fast 3 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.14 gibt Auskunft darüber, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe im Saarland auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 13 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei etwas über 10 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten im Saarland im Lesen wie auch Zuhören jeweils knapp 21 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten im Saarland 66 Prozent der Kinder im Lesen und rund 69 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen etwa 12 Prozent und im Zuhören 9 Prozent der Schülerinnen und Schüler.

Tabelle 6.14: Kompetenzstufenverteilung im Saarland für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	13.2	20.8	30.3	23.8	11.9
	Zuhören	10.2	20.6	33.1	27.1	9.0
Mathematik	Global	11.8	19.6	30.0	24.2	14.4

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler im Saarland, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei fast 12 Prozent. Weitere knapp 20 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik knapp 69 Prozent der Viertklässlerinnen und Viertklässler im Saarland. Herausragende Leistungen (Kompetenzstufe V) erzielten mehr als 14 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern im Saarland mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im Kompetenzbereich *Zuhören* verfehlten nur die Kinder in den drei Stadtstaaten noch häufiger den Mindeststandard als die Viertklässlerinnen und Viertklässler im Saarland. Auch hinsichtlich des Anteils der Kinder, die im Zuhören den KMK-Regelstandard erreichten, schneidet das Saarland als schwächstes der Flächenländer ab. Insbesondere die höchste Kompetenzstufe ist im Zuhören im Saarland auffällig schwach besetzt; nur in einem anderen Land ist der Anteil niedriger. In den anderen Kompetenzbereichen bewegen sich die Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen im Saarland weitgehend im Mittelfeld.

6.13 Sachsen

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe im Freistaat Sachsen, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei etwa 6 Prozent (vgl. Tabelle 6.2). Etwas mehr als 3 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen fast 54 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 46 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Sachsen an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Sachsen bezogen auf die

gesamte Schülerschaft bei knapp 3 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei fast 43 Prozent.

Einschulungsalter

Sachsen gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Sachsen eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die im Freistaat Sachsen 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik weniger als ein Prozent vorzeitig und knapp 4 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwa 9 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt 10 Jahre und etwas über 7 Monate alt.

Unterrichtszeit

Die Stundentafeln im Freistaat Sachsen für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 27 Jahreswochenstunden und im Fach Mathematik mindestens 20 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Sachsen kumulativ mindestens 1 026 Stunden Deutsch- und 760 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Sachsen insgesamt etwas weniger als 66 Prozent der Kinder eine Ganztagschule, wobei es sich in nur ganz wenigen Fällen um eine voll gebundene Form handelte (0,4 %). Eine teilweise gebundene Ganztagschule besuchten etwa 8 Prozent und eine offene Ganztagschule knapp 55 Prozent der Schülerinnen und Schüler. Für rund 2 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor. Insgesamt gehört Sachsen zu den vier Ländern, in denen die Schulen von mehr als 60 Prozent der Viertklässlerinnen und Viertklässler angegeben haben, ein Ganztagsangebot vorzuhalten.²¹

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.15 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Sachsen auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Knapp 9 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei etwa 6 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard)

²¹ Aufgrund von Rundungsungenauigkeiten ergibt die Summe der Prozentangaben für die drei Ganztagsformen (inkl. des Anteils fehlender Angaben) nicht immer genau den Prozentsatz von Kindern, die insgesamt eine Ganztagschule besuchten.

Tabelle 6.15: Kompetenzstufenverteilung in Sachsen für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	8.5	19.1	29.8	27.8	14.8
	Zuhören	5.7	18.7	34.1	30.1	11.4
Mathematik	Global	6.6	17.1	28.4	27.1	20.8

erreichten in Sachsen im Lesen etwa 19 Prozent und im Zuhören ebenfalls rund 19 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Sachsen etwa 72 Prozent der Kinder im Lesen und rund 76 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen knapp 15 Prozent und im Zuhören etwas über 11 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Sachsen, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei knapp 7 Prozent. Rund 17 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichten oder übertrafen (Kompetenzstufe III, IV oder V) in Mathematik mehr als 76 Prozent der Viertklässlerinnen und Viertklässler in Sachsen. Herausragende Leistungen (Kompetenzstufe V) erzielten fast 21 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Sachsen mit denen der anderen 15 Länder, so fällt Folgendes auf:

Der Anteil der Kinder, die im Kompetenzbereich *Lesen* den Mindeststandard verfehlten, ist in keinem Land niedriger als in Sachsen. In Bezug auf das Erreichen des KMK-Regelstandards (mindestens Kompetenzstufe III) und der höchsten Kompetenzstufe im Lesen gehört Sachsen zu den drei Ländern mit den besten Ergebnissen. Im Kompetenzbereich *Zuhören* zählt Sachsen zu einer Gruppe von sieben Ländern, in denen der Anteil von Kindern, die den Mindeststandard verfehlten, unter 7 Prozent liegt. Im Kompetenzbereich *Mathematik* ist der Anteil der Viertklässlerinnen und Viertklässler, die den Mindeststandard verfehlten, der niedrigste aller 16 Länder. Sowohl beim Erreichen des KMK-Regelstandards (mindestens Kompetenzstufe III) als auch bei der Besetzung der höchsten Kompetenzstufe im Fach Mathematik gehört Sachsen wiederum zu den drei Ländern mit den besten Ergebnissen.

6.14 Sachsen-Anhalt

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Sachsen-Anhalt, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei knapp 7 Prozent

(vgl. Tabelle 6.2). Rund 4 Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen fast 62 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 38 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Sachsen-Anhalt an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zielfähig* unterrichtet wurden, lag in Sachsen-Anhalt bezogen auf die gesamte Schülerschaft bei etwa 2 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei etwa 32 Prozent.

Einschulungsalter

Sachsen-Anhalt gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Sachsen-Anhalt eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die in Sachsen-Anhalt 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik knapp 3 Prozent vorzeitig und weniger als ein Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt fast 12 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 7 Monate alt.

Unterrichtszeit

Die Stundentafeln in Sachsen-Anhalt für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 28 Jahreswochenstunden und im Fach Mathematik mindestens 22 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Sachsen-Anhalt kumulativ mindestens 1064 Stunden Deutsch- und 836 Stunden Mathematikunterricht erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Sachsen-Anhalt insgesamt nur etwa 9 Prozent der Kinder eine Ganztagschule, wobei es sich in keinem Fall um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten etwa ein Prozent und eine offene Ganztagschule rund 8 Prozent der Schülerinnen und Schüler.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

In Tabelle 6.16 wird dargestellt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Sachsen-Anhalt auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Insgesamt 9 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem

Tabelle 6.16: Kompetenzstufenverteilung in Sachsen-Anhalt für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	9.0	17.4	31.3	28.3	14.0
	Zuhören	8.7	20.0	34.5	27.1	9.7
Mathematik	Global	8.2	16.6	25.8	28.4	20.9

Mindeststandard ebenfalls bei knapp 9 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Sachsen-Anhalt im Lesen etwas mehr als 17 Prozent und im Zuhören 20 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Sachsen-Anhalt fast 74 Prozent der Kinder im Lesen und rund 71 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen 14 Prozent und im Zuhören rund 10 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Sachsen-Anhalt, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei etwa 8 Prozent. Weitere knapp 17 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik etwa 75 Prozent der Viertklässlerinnen und Viertklässler in Sachsen-Anhalt. Herausragende Leistungen (Kompetenzstufe V) erzielten knapp 21 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Sachsen-Anhalt mit denen der anderen 15 Länder, so fällt Folgendes auf:

Der Anteil der Kinder, die im Kompetenzbereich *Lesen* den Mindeststandard verfehlten, ist unter allen Ländern der zweitniedrigste. Für das Erreichen des KMK-Regelstandards ist in Sachsen-Anhalt im Lesen der höchste Anteil, für die Besetzung des Optimalstandards (Kompetenzstufe V) der dritthöchste Anteil zu verzeichnen. Während in den meisten Ländern die Verteilungen auf die Kompetenzstufen im *Zuhören* günstiger ausfallen als im Lesen, ist dies in Sachsen-Anhalt nicht der Fall. Im Kompetenzbereich *Mathematik* liegt der Anteil der Viertklässlerinnen und Viertklässler, die den Mindeststandard verfehlten, niedriger als in den meisten der 16 Länder (drittniedrigster Anteil). Sowohl in Bezug auf das Erreichen des KMK-Regelstandards als auch hinsichtlich der Besetzung der höchsten Kompetenzstufe im Fach Mathematik gehört Sachsen-Anhalt wiederum zu den drei Ländern mit den besten Ergebnissen.

6.15 Schleswig-Holstein

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe in Schleswig-Holstein, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei knapp 5 Prozent (vgl. Tabelle 6.2). Nur ein Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen fast 22 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden rund 78 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Schleswig-Holstein an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Schleswig-Holstein bezogen auf die gesamte Schülerschaft bei fast 3 Prozent, bezogen auf Kinder mit sonderpädagogischem Förderbedarf bei rund 61 Prozent.

Einschulungsalter

Schleswig-Holstein gehört zu den acht Ländern in der Bundesrepublik Deutschland, in denen sich das Einschulungsalter in den letzten Jahren nicht geändert hat. Kinder werden in Schleswig-Holstein eingeschult, wenn sie bis einschließlich 30. Juni das sechste Lebensjahr vollendet haben (Autorengruppe Bildungsberichterstattung, 2012). Diese Stichtagsregelung galt also auch 2007, als die Grundgesamtheit des Ländervergleichs 2011 das Einschulungsalter erreichte.

Von den Kindern, die in Schleswig-Holstein 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik etwas mehr als 7 Prozent vorzeitig und rund 5 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwas weniger als 9 Prozent der Schülerinnen und Schüler eine Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 6 Monate alt.

Unterrichtszeit

Die Stundentafeln Schleswig-Holsteins für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg im Fach Deutsch mindestens 24 Jahreswochenstunden und im Fach Mathematik mindestens 20 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden.²² In den ersten vier Schuljahren sollten somit Schülerinnen und Schüler in Schleswig-Holstein kumulativ mindestens 912 Stunden Deutsch- und 760 Stunden Mathematikunterricht erhalten haben.

22 Die Angaben beziehen sich auf den Runderlass des Ministeriums für Bildung und Kultur vom 1. August 2011. Die Anzahl der Jahreswochenstunden im Schuljahr 2010/2011 entspricht den Angaben in dieser Verordnung.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Schleswig-Holstein insgesamt 55 Prozent der Kinder eine Ganztagschule, wobei es sich in nur relativ wenigen Fällen um eine voll gebundene Form handelte (2.3 %). Eine teilweise gebundene Ganztagschule besuchten weniger als 2 Prozent und eine offene Ganztagschule 49 Prozent der Schülerinnen und Schüler. Für etwas über 2 Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.17 gibt Auskunft darüber, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe in Schleswig-Holstein auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Etwa 13 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei nur knapp 6 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Schleswig-Holstein im Lesen knapp 22 Prozent und im Zuhören rund 18 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Schleswig-Holstein rund 65 Prozent der Kinder im Lesen und etwa 76 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen rund 11 Prozent und im Zuhören knapp 13 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Schleswig-Holstein, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei 13 Prozent. Weitere knapp 23 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik etwa 64 Prozent der Viertklässlerinnen und Viertklässler in Schleswig-Holstein. Herausragende Leistungen (Kompetenzstufe V) erzielten etwa 12 Prozent der Schülerinnen und Schüler.

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Schleswig-Holstein mit denen der anderen 15 Länder, so fällt Folgendes auf:

Tabelle 6.17: Kompetenzstufenverteilung in Schleswig-Holstein für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	13.2	21.5	29.8	24.2	11.2
	Zuhören	5.8	17.8	32.3	31.3	12.8
Mathematik	Global	13.0	22.7	30.7	21.8	11.8

Anders als im Kompetenzbereich *Lesen*, wo für Schleswig-Holstein keine auffälligen Ergebnisse zu beobachten sind, zeichnet sich das Land durch eine Population von Viertklässlerinnen und Viertklässlern aus, die im Kompetenzbereich *Zuhören* relativ häufig den Regelstandard erreichen konnten (dritthöchster Anteil). Zudem weist Schleswig-Holstein den zweitniedrigsten Anteil von Kindern auf, die im Zuhören den Mindeststandard verfehlten. In Bezug auf die Besetzung der höchsten Kompetenzstufe gehört Schleswig-Holstein im Zuhören zu den drei Ländern mit den besten Ergebnissen. Im Kompetenzbereich *Mathematik* hingegen deuten die jeweils relativ niedrigen Anteile an Viertklässlerinnen und Viertklässlern, die mindestens den Regelstandard erreichten oder sogar Leistungen auf dem Niveau des Optimalstandards zeigten, auf eine relative Schwäche hin. Unter den Flächenländern nimmt Schleswig-Holstein hier den vorletzten beziehungsweise letzten Rangplatz ein.

6.16 Thüringen

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen

Der Anteil der Schülerinnen und Schüler der vierten Jahrgangsstufe im Freistaat Thüringen, bei denen ein sonderpädagogischer Förderbedarf in den Bereichen Sprache, soziale und emotionale Entwicklung oder Lernen diagnostiziert wurde, lag gemäß den Daten des Ländervergleichs 2011 bei fast 5 Prozent (vgl. Tabelle 6.2). Etwa ein Prozent *aller* Kinder der vierten Jahrgangsstufe besuchten eine Förderschule mit den Förderschwerpunkten Sprache, soziale und emotionale Entwicklung oder Lernen; bezogen auf Kinder mit sonderpädagogischem Förderbedarf gingen knapp 22 Prozent auf eine Förderschule mit einem dieser Förderschwerpunkte. Demzufolge wurden etwa 78 Prozent der Viertklässlerinnen und Viertklässler mit sonderpädagogischem Förderbedarf 2011 in Thüringen an Regelschulen integrativ beschult. Der Anteil der Schülerinnen und Schüler, die nach den Angaben der teilnehmenden Schulen in der vierten Jahrgangsstufe *zieldifferent* unterrichtet wurden, lag in Thüringen bezogen auf die gesamte Schülerschaft bei 0,5 Prozent; dies stellt bundesweit die zweitniedrigste Quote dar. Bezogen auf Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf liegt der Anteil *zieldifferent* unterrichteter Kinder bei etwa 11 Prozent.

Einschulungsalter

Der Stichtag für die Einschulung wurde in Thüringen 2003 um einen Monat vom 30. Juni auf den 31. Juli vorgezogen (Autorengruppe Bildungsberichterstattung, 2012). Während vor 2003 das Einschulungsalter erreicht war, wenn ein Kind bis einschließlich 30. Juni das sechste Lebensjahr vollendet hatte, gelten seitdem Kinder als schulpflichtig, wenn sie vor dem oder am 31. Juli sechs Jahre alt werden. Für die Zielpopulation des Ländervergleichs 2011 galt diese neue Stichtagsregelung bereits, das heißt, die Kinder wurden in der Regel eingeschult, wenn sie bis einschließlich 31. Juli ihr sechstes Lebensjahr vollendet hatten.

Von den Kindern, die in Thüringen 2007 das Einschulungsalter erreicht hatten, wurden nach Angaben der amtlichen Statistik 2 Prozent vorzeitig und rund 6 Prozent verspätet eingeschult (Autorengruppe Bildungsberichterstattung, 2012). Gegen Ende der vierten Jahrgangsstufe hatten nach den Daten des Ländervergleichs 2011 insgesamt etwas über 6 Prozent der Schülerinnen und Schüler eine

Klassenstufe wiederholt. Zum Zeitpunkt der Testung waren die Schülerinnen und Schüler im Durchschnitt etwa 10 Jahre und 6 Monate alt.

Unterrichtszeit

Die Stundentafeln in Thüringen für das Schuljahr 2010/2011 sahen vor, dass in den Primarschulen über die Jahrgangsstufen 1–4 hinweg in den Fächern Deutsch und Mathematik mindestens zwischen 42 und 46 Jahreswochenstunden (Summe der wöchentlichen Stundenzahl für die vier Jahrgangsstufen) Unterricht erteilt werden. In den ersten vier Schuljahren sollten Schülerinnen und Schüler in Thüringen kumulativ somit zwischen 1596 und 1748 Stunden Unterricht in beiden Fächern zusammen erhalten haben.

Besuch von Ganztagschulen

In der Population der Viertklässlerinnen und Viertklässler im Ländervergleich 2011 besuchten in Thüringen insgesamt 73 Prozent der Kinder eine Ganztagschule, wobei es sich in knapp 9 Prozent der Fälle um eine voll gebundene Form handelte. Eine teilweise gebundene Ganztagschule besuchten etwas unter 6 Prozent und eine offene Ganztagschule rund 57 Prozent der Schülerinnen und Schüler. Für etwas mehr als ein Prozent der Kinder lagen keine Angaben der Schulleiterinnen oder Schulleiter über die Form der besuchten Ganztagschule vor. Insgesamt gehört Thüringen zu den vier Ländern, in denen die Schulen von mehr als 60 Prozent der Viertklässlerinnen und Viertklässler angegeben haben, ein Ganztagsangebot zu haben.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Tabelle 6.18 zeigt, wie sich die Schülerinnen und Schüler der vierten Jahrgangsstufe im Freistaat Thüringen auf die Kompetenzstufen in denjenigen Kompetenzbereichen verteilen, die in den Fächern Deutsch und Mathematik im Ländervergleich 2011 getestet wurden. Knapp 10 Prozent der zielgleich unterrichteten Viertklässlerinnen und Viertklässler verfehlten im Bereich Lesen den Mindeststandard. Im Bereich Zuhören lag der Anteil der Kinder unter dem Mindeststandard bei etwas mehr als 6 Prozent. Lediglich die Kompetenzstufe II (Mindeststandard) erreichten in Thüringen im Lesen rund 20 Prozent und im Zuhören 19 Prozent der Viertklässlerinnen und Viertklässler. Leistungen auf dem Niveau der KMK-Regelstandards oder darüber (mindestens Kompetenzstufe III) erzielten in Thüringen etwa 71 Prozent der Kinder im Lesen und knapp 75 Prozent im Zuhören. Die höchste Kompetenzstufe erreichten im Lesen knapp 13 Prozent und im Zuhören etwa 9 Prozent der Schülerinnen und Schüler.

Im Bereich Mathematik liegt der Anteil der Schülerinnen und Schüler in Thüringen, die in der vierten Jahrgangsstufe den Mindeststandard verfehlten, bei etwas über 12 Prozent. Weitere gut 20 Prozent der Viertklässlerinnen und Viertklässler zeigten mathematische Kompetenzen, die lediglich dem Mindeststandard entsprechen. Den KMK-Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV oder V) konnten in Mathematik etwa 67 Prozent der Viertklässlerinnen und Viertklässler in Thüringen. Herausragende Leistungen (Kompetenzstufe V) erzielten knapp 16 Prozent der Schülerinnen und Schüler.

Tabelle 6.18: Kompetenzstufenverteilung in Thüringen für Schülerinnen und Schüler der vierten Jahrgangsstufe nach Kompetenzbereich (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Lesen	9.8	19.7	31.2	26.5	12.8
	Zuhören	6.3	19.0	35.9	29.3	9.4
Mathematik	Global	12.4	20.3	28.4	23.1	15.7

Vergleicht man die Befunde zur Erreichung der Bildungsstandards bei den Viertklässlerinnen und Viertklässlern in Thüringen mit denen der anderen 15 Länder, so fällt Folgendes auf:

Im *Lesen* liegen die Ergebnisse Thüringens sowohl hinsichtlich des Erreichens des KMK-Regelstandards als auch in Bezug auf die Besetzung der höchsten Kompetenzstufe V im oberen Mittelfeld. Im Kompetenzbereich *Zuhören* weist Thüringen eine auffällig auf die mittleren Kompetenzstufen konzentrierte Verteilung auf: Das Land zählt einerseits zu einer Gruppe von sieben Ländern, in denen der Anteil von Kindern, die den Mindeststandard verfehlten, unter 7 Prozent liegt. Andererseits ist die höchste Kompetenzstufe im Zuhören nur in drei Ländern schwächer besetzt als in Thüringen. Im Fach *Mathematik* liegt die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen bundesweit gesehen im Mittelfeld.

6.17 Orthografiekompetenzen in der bundesweit repräsentativen Stichprobe

Wie in den Kapiteln 2.1 und 3.2.3 dargelegt wurde, ist im Fach Deutsch nicht nur für die Teilkompetenzen Lesen und Zuhören, sondern auch für die orthografische Kompetenz ein Kompetenzstufenmodell entwickelt worden, um das Erreichen der Bildungsstandards in diesem Bereich überprüfen zu können. Im Ländervergleich 2011 sollte nach Beschluss der Kultusministerkonferenz vom 9. Dezember 2010 im Bereich Orthografie zunächst länderübergreifend die Verteilung der Viertklässlerinnen und Viertklässler auf die Kompetenzstufen untersucht werden. Für diese Teilkompetenz wurde daher eine bundesweit repräsentative Stichprobe in die Studie einbezogen, die jedoch, anders als bei den anderen drei erfassten Kompetenzbereichen, keine Auswertung für die einzelnen Länder zulässt.

Aus Tabelle 6.19 ist ersichtlich, dass sich die orthografische Kompetenz der Viertklässlerinnen und Viertklässler bundesweit in ähnlicher Weise auf die fünf Kompetenzstufen verteilt wie die in den anderen Bereichen überprüften Kompetenzen im Fach Deutsch. Insgesamt verfehlten in Deutschland fast 13 Prozent der Kinder in der vierten Jahrgangsstufe im Kompetenzbereich Orthografie den Mindeststandard der KMK. Weitere 23 Prozent erfüllten mit ihren Leistungen im Test diese Mindestwartungen. Kompetenzen auf dem Niveau des Regelstandards oder darüber wurden von fast zwei Dritteln (64,3 %) der Viertklässlerinnen und Viertklässler gezeigt. Die oberhalb des Regelstandards liegenden Kompetenzstufen IV und V erreichten knapp 35 Prozent der Kinder.

Tabelle 6.19: Länderübergreifende Kompetenzstufenverteilung für Schülerinnen und Schüler der vierten Jahrgangsstufe für die Kompetenzteilbereiche Orthografie, Lesen und Zuhören (prozentuale Angaben)

		Kompetenzstufen				
		I	II	III	IV	V
Deutsch	Orthografie	12.6	23.1	29.4	23.1	11.8
	Lesen	12.4	21.0	30.0	24.7	12.0
	Zuhören	7.4	18.9	32.6	29.1	11.9

Ein vergleichbares Befundmuster zeigte sich für Orthografie bereits in der 2007 durchgeführten Normierungsstudie des IQB, die zur Entwicklung der Kompetenzstufenmodelle im Primarbereich durchgeführt worden war.

Literatur

- Autorengruppe Bildungsberichterstattung (Hrsg.). (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann Verlag.
- KMK (2006) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Wolters Kluwer.
- Köller, O., Knigge, M. & Tesch, B. (Hrsg.). (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Lohauß, P., Rehkämper, K., Rockmann, U. & Wendt, W. (Hrsg.). (2011). *Bildung in Berlin und Brandenburg 2010. Ein indikatorengestützter Bericht zur Bildung im Lebenslauf*. Berlin, Potsdam: Institut für Schulqualität der Länder Berlin und Brandenburg und Amt für Statistik Berlin-Brandenburg.
- Neumann, M., Schnyder, I., Trautwein, U., Niggli, A., Lüdtke, O. & Cathomas, R. (2007). Schulformen als differenzielle Lernmilieus: Institutionelle und kompositionelle Effekte auf die Leistungsentwicklung im Fach Französisch. *Zeitschrift für Erziehungswissenschaft*, 10, 399–420.
- Statistisches Bundesamt. (2011). *Fachserie 11: Bildung und Kultur, Reihe 1: Allgemeinbildende Schulen*. Wiesbaden: Statistisches Bundesamt.

Kapitel 7

Geschlechtsbezogene Disparitäten

Katrin Böhme und Alexander Roppelt

7.1 Geschlechtsbezogene Unterschiede im Bildungsbereich

In der Diskussion über Fragen der Bildungsqualität und des Bildungserfolgs von Kindern und Jugendlichen spielen Geschlechterunterschiede noch immer eine zentrale Rolle. Geschlechtsbezogene Disparitäten bestehen sowohl in den Mustern der schulischen Bildungsbeteiligung und dem Erwerb von Bildungsabschlüssen als auch im schulischen Kompetenzerwerb (vgl. z.B. Hannover & Kessels, 2011).

Zur Erklärung dieser geschlechtsbezogenen Unterschiede im schulischen Kontext sind in der Literatur verschiedene Theorieansätze zu finden (für einen Überblick vgl. z.B. Benölken, 2011; Hannover, 2008; Stanat & Bergann, 2009). So gehen beispielsweise biologische und neurowissenschaftliche Erklärungsansätze davon aus, dass physiologische Unterschiede zwischen den Geschlechtern, wie etwa genetische und hormonelle Eigenschaften oder auch geschlechtsspezifische Merkmale der Gehirnorganisation, die Entwicklung unterschiedlicher Begabungen, Interessen, Vorlieben und somit Kompetenzen begünstigen. Inwieweit solche biologischen Faktoren bei der Entwicklung schulischer Disparitäten tatsächlich eine Rolle spielen, ist umstritten. So argumentiert etwa Budde (2011), dass die aktuellen Veränderungen in den Mustern der Bildungsbeteiligung und im Erwerb von Bildungsabschlüssen (vgl. Abschnitt 7.1.1) so nicht auftreten dürften, wenn die Unterschiede im Kompetenzerwerb von Jungen und Mädchen tatsächlich vorrangig biologische Ursachen hätten. Geschlechtsbezogene Disparitäten im schulischen Kontext werden daher im Allgemeinen primär als ein Ergebnis von Sozialisationseinflüssen betrachtet.

Sozialisationstheoretische Erklärungsansätze gehen davon aus, dass geschlechtstypische Eigenschaften, Einstellungen und Verhaltensweisen in Interaktion mit dem sozialen Umfeld erlernt werden (vgl. Hannover, 2008; Trautner, 2002). Durch Sozialisationsprozesse werden Unterschiede zwischen Jungen und Mädchen ausgeformt und modifiziert (vgl. z.B. Stanat & Bergann, 2009). Lerntheorien akzentuieren hierbei die Bedeutung von Rollenvorbildern und Erwartungen auf Seiten der Eltern, Lehrkräfte und Peers. Solche Erwartungen können sich beispielsweise in einer differenziellen Bekräftigung von Verhaltensweisen äußern, die eher männlich oder eher weiblich konnotiert sind. Auch (sozial-)psychologische Theorieansätze betonen, dass Geschlecht im sozialen Kontext konstruiert wird. Sie thematisieren die Entwicklung von Geschlechterstereotypen und untersuchen, welche situationalen und strukturellen Faktoren deren Entstehung und Aufrechterhaltung erklären können (vgl. Blossfeld et al., 2009).

Auf der Grundlage gegenwärtiger Erkenntnisse aus der soziologischen und psychologischen Forschung ist davon auszugehen, dass zwischen Jungen und Mädchen bereits beim Eintritt in das Bildungssystem Unterschiede bestehen, die primär als sozial determinierte Differenzen interpretiert werden können (vgl. Blossfeld et al., 2009; Hannover, 2008; Stanat & Bergann, 2009). Unterschiede bestehen vor allem bei bestimmten Aspekten des Sozialverhaltens und in Bezug auf Interessen, teilweise auch bei kognitiven Fähigkeiten, wie etwa dem räumlichen Denken oder verbalen Kompetenzen. Diese Differenzen, die insbesondere im kognitiven Bereich zu Beginn der Bildungslaufbahn noch sehr klein sind, formen sich im schulischen Kontext weiter aus und resultieren in geschlechtsbezogenen Disparitäten beim Kompetenzerwerb und in den Mustern der Bildungsbeteiligung.

Gleichzeitig liegen jedoch auch Hinweise darauf vor, dass sich geschlechtsbezogene Unterschiede durch gezielte Prävention und Intervention reduzieren lassen (vgl. für einen Überblick Hannover, 2008). Um für den Primarschulbereich festzustellen, inwieweit solche Bemühungen angezeigt sind, ist es wichtig, auf der Grundlage aktueller Daten zu bestimmen, ob und in welchem Ausmaß Kompetenzunterschiede zwischen Jungen und Mädchen in zentralen Kompetenzbereichen existieren. Dies ist Gegenstand des vorliegenden Kapitels.

7.1.1 Geschlechtsbezogene Unterschiede in der Bildungsbeteiligung

Geschlechtsbezogene Unterschiede wurden im schulischen Kontext in Deutschland bereits am Ende des 19. Jahrhunderts diskutiert (vgl. Stanat & Bergann, 2009). Zur damaligen Zeit und bis weit in die Phase der Bildungsexpansion in den 60er Jahren des letzten Jahrhunderts hinein, waren Mädchen im Bildungsbereich deutlich benachteiligt. Veranschaulicht wurde dies beispielsweise mit dem Bild vom „katholischen Arbeitermädchen vom Lande“ (Blossfeld et al., 2009, S. 40), das als Inbegriff der Bildungsbenachteiligung galt.

Gegenwärtig wird in der öffentlichen und wissenschaftlichen Diskussion allerdings verstärkt davon ausgegangen, dass nicht länger die Mädchen, sondern zunehmend die Jungen im schulischen Kontext benachteiligt sind (vgl. z.B. Diefenbach, 2010; Hannover & Kessels, 2011). Dies äußert sich vor allem darin, dass Jungen in höheren Bildungsgängen weniger stark vertreten sind und die entsprechenden höherwertigen Abschlüsse seltener als Mädchen erreichen. Beides soll nachfolgend anhand aktueller Daten des Statistischen Bundesamtes veranschaulicht werden. Hierfür sind zunächst in Tabelle 7.1 die Beteiligungsquoten für Jungen und Mädchen der neunten Jahrgangsstufe in den Schularten des Sekundarschulbereichs im Schuljahr 2010/2011 dargestellt.

Tabelle 7.1: Beteiligungsquoten für Jungen und Mädchen der neunten Jahrgangsstufe in den Schularten der Sekundarstufe I im Schuljahr 2010/2011

Schulart	Jungen		Mädchen	
	Anzahl	Anteil in %	Anzahl	Anteil in %
Hauptschulen	86 228	20.3	66 330	16.4
Schularten mit mehreren Bildungsgängen	35 446	8.3	30 732	7.6
Realschulen	105 216	24.8	102 889	25.5
Gymnasien	133 399	31.4	149 085	36.9
Integrierte Gesamtschulen	40 014	9.4	37 977	9.4
Freie Waldorfschulen	3 035	0.7	3 392	0.8
Förderschulen	21 468	5.1	13 268	3.3

Quelle: Auszug aus der Tabelle 3.4–2010 der Fachserie 11 für das Schuljahr 2010/2011 des Statistischen Bundesamtes (2011).

Die in Tabelle 7.1 zusammengestellten Informationen verdeutlichen, dass Jungen im Schuljahr 2010/2011 in der neunten Jahrgangsstufe häufiger als Mädchen solche Schularten besuchten, die zu einem geringerwertigen Schulabschluss hinführen. So sind etwa Jungen an Haupt- und Förderschulen, Mädchen hingegen an Gymnasien überrepräsentiert.

Unterschiede finden sich dementsprechend auch im Hinblick auf die von Jungen und Mädchen tatsächlich erreichten Bildungsabschlüsse. Um dies zu veranschaulichen, zeigt Tabelle 7.2 die Anzahl der vergebenen Abschlüsse und die prozentualen Anteile an Absolventen beziehungsweise Abgängern allgemeinbildender und beruflicher Schulen im jeweils typischen Abschlussalter im Jahr 2010 nach Abschlussarten und Geschlecht.

Tabelle 7.2: Von Jungen und Mädchen im Jahr 2010 in der Sekundarstufe erreichte Bildungsabschlüsse

Abschlussart	Jungen		Mädchen	
	Anzahl	Anteil in %	Anzahl	Anteil in %
Ohne Hauptschulabschluss	32 351	7.8	20 707	5.3
Hauptschulabschluss	125 442	30.3	88 863	22.6
Mittlerer Schulabschluss	223 516	52.2	222 555	54.7
Fachhochschulreife	74 336	15.4	68 297	14.8
Allgemeine Hochschulreife	142 238	29.4	173 985	37.8

Anmerkung. Die prozentualen Angaben beziehen sich auf den Durchschnitt der Wohnbevölkerung (am 31.12. des Vorjahres) im jeweils typischen Abschlussalter (ohne/mit Hauptschulabschluss: 15 bis unter 17 Jahre; Mittlerer Schulabschluss: 16 bis unter 18 Jahre; Fachhochschulreife und allgemeine Hochschulreife: 18 bis unter 21 Jahre). Dies führt dazu, dass sich die Anteile insgesamt zu mehr als 100 % addieren.

Quelle: Auszug aus der Tabelle D7-4web des Bildungsberichts 2012 (Autorengruppe Bildungsberichterstattung, 2012).

Zusammenfassend kann festgehalten werden, dass sich im Hinblick auf die Muster der schulischen Bildungsbeteiligung und die erreichten Bildungsabschlüsse deutliche Geschlechterunterschiede zugunsten der Mädchen abzeichnen. Neben den hier veranschaulichten Disparitäten finden sich weitere Unterschiede in der Bildungsbeteiligung. Während etwa Mädchen häufiger als Jungen vorzeitig eingeschult werden, sind Jungen häufiger von Zurückstellungen in der Schuleingangsphase betroffen, und auch bei den Übergangsempfehlungen von der Primarstufe in die Sekundarschule lassen sich Unterschiede zugunsten der

Mädchen feststellen. Entsprechend erreichen Mädchen proportional häufiger das Abitur, während relativ betrachtet mehr Jungen als Mädchen die Schule ohne Abschluss oder mit einem Hauptschulabschluss verlassen (vgl. Blossfeld et al., 2009; Budde, 2011; Hannover & Kessels, 2011; Stanat & Bergann, 2009).

Als Grund für den größeren Erfolg der Mädchen im Hinblick auf die Bildungsbeteiligung und den Erwerb von Bildungsabschlüssen wird diskutiert, dass Mädchen eher als Jungen solche Persönlichkeitseigenschaften und Verhaltensweisen zeigen, die im schulischen Kontext als angemessen und besser angepasst wahrgenommen werden (vgl. Hannover & Kessels, 2011). Beispielsweise konnten Anders, McElvany und Baumert (2010) empirisch belegen, dass Lehrkräfte Mädchen im Hinblick auf ihr Sozialverhalten sowie ihre Motivation und Lerntugenden positiver einschätzen als Jungen. Diese schulisch relevanten Aspekte überfachlicher Kompetenzen fließen auch in die Beurteilung und Benotung der Schülerinnen und Schüler ein. Hannover und Kessels (2011) gehen daher davon aus, dass die stärker ausgeprägten – beziehungsweise als stärker ausgeprägt wahrgenommenen – überfachlichen Kompetenzen der Mädchen zu ihren relativ besseren Noten und somit zu ihrer Überrepräsentation an höheren Bildungsgängen sowie ihren höher qualifizierenden Schulabschlüssen beitragen.

Eine andere Hypothese zu Ursachen differenzieller Muster der Bildungsbeteiligung von Jungen und Mädchen bezieht sich darauf, dass – insbesondere im Elementar- und im Primarschulbereich – der weit überwiegende Anteil der Betreuungspersonen und Lehrkräfte weiblichen Geschlechts ist (vgl. Blossfeld et al., 2009; Kapitel 10 in diesem Band) und es den Jungen auf diese Weise nicht nur an positiven Rollenbildern fehle, sondern der Unterricht auch durch „weibliche“ Interessen und Handlungsweisen geprägt sei. Belastbare empirische Evidenz für die These, dass Jungen durch den hohen Anteil von Frauen im Bildungssystem benachteiligt werden, liegt allerdings nicht vor (vgl. zusammenfassend Hannover & Kessels, 2011; Helbig, 2010).

7.1.2 Geschlechtsbezogene Kompetenzunterschiede

Geschlechtsbezogene Unterschiede finden sich nicht nur in der schulischen Bildungsbeteiligung und bei lehrerseitigen Leistungserwartungen und -einschätzungen, sondern auch in den von Jungen und Mädchen erreichten Kompetenzständen. Dies belegen sowohl die großen Schulleistungsstudien wie PIRLS/IGLU¹ oder PISA² als auch Meta-Analysen und zahlreiche kleinere empirische Untersuchungen (vgl. z. B. Hyde, 1981, 2005; Wieczerkowski & Prado, 1992).³

1 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.

2 Das Akronym PISA steht für *Programme for International Student Assessment*.

3 Da in der Literatur Geschlechterunterschiede häufig ausschließlich in Form der Effektstärke Cohens d (Cohen, 1988) angegeben werden, können auch im Folgenden teilweise nur d -Werte berichtet werden. Für die Bewertung der praktischen Bedeutsamkeit von Effekten schlägt Cohen (1988) folgende Klassifizierung vor: kleiner Effekt: $d = 0.20$, mittlerer Effekt: $d = 0.50$, großer Effekt: $d = 0.80$. Eine Interpretation von Effektstärken gemäß dieser Klassifizierung ist in Bildungsstudien jedoch nur bedingt aufschlussreich, da die gefundenen Effekte hier in der Regel geringer ausfallen als im experimentellen Kontext, in dem Cohen (1988) seine Kriterien entwickelt hat. Hilfreich ist vielmehr eine Orientierung an den innerhalb eines Schuljahres zu erwartenden Lernzuwächsen.

Geschlechtsbezogene Unterschiede in sprachlichen Kompetenzen

Bereits im Elementarbereich weisen Mädchen Vorsprünge beim Erwerb sprachlicher Kompetenzen auf und sind seltener von Störungen der Sprachentwicklung wie Stottern betroffen (Halpern, 2000). Zu Beginn der Grundschulphase sind die geschlechtsbezogenen Kompetenzunterschiede zwischen Jungen und Mädchen zwar nur gering ausgeprägt, die Disparitäten zugunsten der Mädchen vergrößern sich jedoch bis zum Ende der Grundschulzeit und insbesondere im Verlauf der Sekundarstufe deutlich (vgl. Stanat & Bergann, 2009).

Für den Bereich der *Lesekompetenz* wurde im Rahmen der großen internationalen Schulleistungsstudien PIRLS/IGLU und PISA wiederholt nachgewiesen, dass Mädchen im Mittel höhere Kompetenzstände als Jungen erreichen. Dies gilt sowohl international als auch innerhalb Deutschlands und sowohl im Primar- als auch im Sekundarbereich (z. B. Baumert et al., 2001, 2002; Bos et al., 2007, 2008). In der Grundschulzeit sind die Unterschiede zwischen den Lesekompetenzen von Jungen und Mädchen jedoch eher gering (vgl. Blossfeld et al., 2009). In IGLU-E 2006 betrug der Vorteil der Mädchen auf der Gesamtskala Lesen nur 7 Punkte (vgl. Valtin, Bos, Buddeberg, Goy & Potthoff, 2008, S. 77), wobei der Kompetenzvorsprung beim Lesen literarischer Texte am stärksten ausgeprägt war. Für Schülerinnen und Schüler der neunten Jahrgangsstufe wurde im Rahmen des bildungsstandardbasierten Ländervergleichs sprachlicher Kompetenzen, der 2009 vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) durchgeführt wurde (vgl. Köller, Knigge & Tesch, 2010), ein mittlerer Kompetenzvorsprung der Mädchen im Bereich Lesen in Höhe von 17 Punkten ermittelt, was einem Effekt von $d = 0.19$ entspricht (vgl. Winkelmann & Groeneveld, 2010, S. 180). Der im Rahmen der PISA-Studien der Jahre 2000 bis 2009 jeweils berichtete Unterschied in den Lesekompetenzen der Schülerinnen und Schüler in Deutschland bewegte sich stets um die 40 Punkte (vgl. Naumann, Artelt, Schneider & Stanat, 2010, S. 54) und war somit größer als der im Ländervergleich sprachlicher Kompetenzen 2009 ermittelte Unterschied.⁴

Für das deutschsprachige *Zuhören* liegen bislang nur wenige Befunde zu geschlechtsbezogenen Unterschieden vor. Ein Ergebnis für den Primarbereich entstammt der 2006 durchgeführten Studie des IQB zur Pilotierung der bildungsstandardbasierten Testaufgaben. Anhand dieser Daten konnten Behrens, Böhme und Krelle (2009, S. 368) für das Zuhören im Fach Deutsch keine Unterschiede in den Kompetenzständen zwischen Jungen und Mädchen der dritten und vierten Jahrgangsstufe identifizieren. Für die Sekundarstufe I wurde anhand von Daten des IQB-Ländervergleichs 2009 ein Kompetenzvorsprung der Mädchen im Umfang von 18 Punkten ($d = 0.18$) gefunden, der somit ähnlich groß ist wie im Kompetenzbereich Lesen (Winkelmann & Groeneveld, 2010, S. 180).

Für die *orthografische Kompetenz* von Schülerinnen und Schülern im Primarbereich berichten Böhme und Bremerich-Vos (2009) im Rahmen der vom IQB durchgeführten Normierung des bildungsstandardbasierten Aufgabenpools einen Kompetenzvorsprung der Mädchen in Höhe von $d = 0.22$ (Böhme & Bremerich-Vos, 2009, S. 346). Relativiert man diesen geschlechtsbezogenen Unterschied an dem zum Ende der Grundschulzeit zu erwartenden Kompetenzzuwachs (vgl. Böhme & Bremerich-Vos, 2009), ergibt dies eine Differenz von

4 Dass der Unterschied im Rahmen von PISA größer ausfiel als im IQB-Ländervergleich 2009 könnte unter anderem darauf zurückzuführen sein, dass in die PISA-Erhebungen auch Schülerinnen und Schüler an Förderschulen einbezogen wurden, in denen Jungen überrepräsentiert sind.

etwa einem Viertel Schuljahr. Der von Böhme und Bremerich-Vos (2009) berichtete Geschlechterunterschied ist mit dem im Rahmen von IGLU-E 2006 für den Bereich der Rechtschreibung gefundenen Effekt von $d = 0.20$ (Kowalski, Voss, Valtin & Bos, 2010, S. 36) praktisch identisch. Dieser Befund kann somit als gut gesichert gelten. Für Schülerinnen und Schüler der neunten Jahrgangsstufe wurde im Rahmen des IQB-Ländervergleichs sprachlicher Kompetenzen 2009 für den Bereich der Orthografie ein Effekt zugunsten der Mädchen in Höhe von 51 Punkten ($d = 0.53$) ermittelt (vgl. Winkelmann & Groeneveld, 2010, S. 179). Ebenso wie in den Kompetenzbereichen Lesen und Zuhören vergrößert sich also auch im Teilbereich Orthografie der Geschlechterunterschied von der Grundschule bis zur neunten Jahrgangsstufe.

Geschlechtsbezogene Unterschiede in mathematischen Kompetenzen

Auch für den Kompetenzerwerb in den Bereichen Mathematik, Informatik, Naturwissenschaft und Technik (MINT) zeigen sich geschlechtsbezogene Unterschiede – hier allerdings zugunsten der Jungen. Meta-Analysen zur mathematischen Kompetenz von Jungen und Mädchen sowie nationale und internationale Schulleistungsstudien belegen, dass bereits am Ende der Grundschulzeit Unterschiede zwischen Schülerinnen und Schülern bestehen und dass auch hier die Differenzen zwischen den beiden Geschlechtergruppen mit dem Alter beziehungsweise mit der untersuchten Schulstufe tendenziell zunehmen (vgl. Hyde, Fennema & Lamon, 1990; Rohe & Quaiser-Pohl, 2010). Auf Grundlage der Daten aus den Studien FIMS⁵, SIMS⁶ und TIMSS⁷ konnte Hanna (2000) nachweisen, dass sich die Unterschiede in den von Jungen und Mädchen erreichten mathematischen Kompetenzen in den letzten fünf Jahrzehnten verringert haben und sich die Kompetenzstände der Geschlechtergruppen historisch betrachtet annähern. Winkelmann und van den Heuvel-Panhuizen (2009) gehen aufgrund ihrer eigenen Befunde jedoch davon aus, dass dieser Trend gegenwärtig keine Fortsetzung findet.

Inwieweit sich in unterschiedlichen mathematischen Bereichen die Kompetenzstände von Jungen und Mädchen differenziell unterscheiden, ist empirisch nicht eindeutig belegt. Als gut gesichert gilt der Befund, dass Jungen Vorteile im räumlichen Denken aufweisen und zwar insbesondere dann, wenn mentale Rotationen oder räumliches Vorstellungsvermögen gefordert sind (vgl. Hyde, 2005). Anhand von Daten der IQB-Studien zur Pilotierung und Normierung der bildungsstandardbasierten Testaufgaben identifizierten Winkelmann und van den Heuvel-Panhuizen (2009) für den Primarbereich signifikant höhere Kompetenzstände der Jungen in allen inhaltsbezogenen mathematischen Kompetenzen. Diese Vorsprünge variieren jedoch in Abhängigkeit vom jeweiligen Kompetenzbereich und bewegen sich zwischen $d = 0.14$ für *Raum und Form* und $d = 0.36$ für *Größen und Messen* (Winkelmann & van den Heuvel-Panhuizen, 2009, S. 147f.). Relativiert man diese Unterschiede an dem innerhalb eines Schuljahres zu erwartenden Kompetenzzuwachs, der am Ende der Grundschulzeit im Fach Mathematik etwa 80 Punkte beträgt (Reiss & Winkel-

5 Das Akronym FIMS steht für *First International Mathematics Study*.

6 Das Akronym SIMS steht für *Second International Mathematics Study* und gilt gemeinsam mit FIMS als Vorläuferstudie von TIMSS.

7 Das Akronym TIMSS stand ursprünglich für *Third International Mathematics and Science Study*. Seit 2003 wird es in der Bedeutung *Trends in International Mathematics and Science Study* verwendet.

mann, 2009, S. 139), entsprechen die gefundenen Differenzen etwa einem Fünftel bis zu maximal der Hälfte eines Schuljahres.

Hinsichtlich der geschlechtsbezogenen Leistungsvariabilität wurde bereits vor mehreren Jahrzehnten die Hypothese aufgestellt, dass die Kompetenzstände von Jungen eine größere Varianz aufweisen als die der Mädchen (vgl. z.B. Feingold, 1992). Für den Bereich der mathematischen Kompetenz konnte diese Hypothese von Hyde und Kollegen (2008) bestätigt werden. Insgesamt werden in empirischen Studien aber primär die im Mittel von Jungen und Mädchen erreichten Kompetenzen untersucht, während Unterschiede in den jeweiligen Kompetenzverteilungen nur selten thematisiert werden. Daher ist wenig darüber bekannt, inwieweit sich die Verteilungen der Mädchen und Jungen in verschiedenen Kompetenzbereichen unterscheiden.

7.1.3 Fragestellungen

Wie der aktuelle Forschungsstand zeigt, bestehen zwischen den beiden Geschlechtergruppen bildungsbezogene Ungleichheiten. Diese betreffen zum einen die Bildungsbeteiligung und den Erwerb von Bildungsabschlüssen. Hier besteht mittlerweile ein Vorteil zugunsten der Mädchen. Betrachtet man den Kompetenzerwerb von Jungen und Mädchen, so zeigen sich im Hinblick auf sprachliche Kompetenzen, und hier insbesondere im Lesen und in der Orthografie, Vorsprünge der Mädchen; im Bereich der mathematischen Kompetenz finden sich hingegen Vorteile für die Jungen.

In der weit überwiegenden Mehrzahl der vorliegenden empirischen Befunde handelt es sich um Vergleiche der *mittleren* Kompetenzstände von Jungen und Mädchen, über Unterschiede in den *Leistungsverteilungen* werden hingegen nur selten Aussagen gemacht. Bei den untersuchten Kompetenzen liegt der Fokus vorwiegend auf der Lesekompetenz sowie auf allgemeinen mathematischen Kompetenzen. Weitere sprachliche Kompetenzen oder spezifische mathematische Inhaltsbereiche wurden bislang nur vereinzelt im Rahmen großer Schulleistungsstudien des Primarbereichs untersucht.

Von Interesse ist daher im Rahmen des Ländervergleichs 2011 erstens die Frage, inwieweit sich differenzielle Unterschiede in den Kompetenzständen von Jungen und Mädchen für die untersuchten Kompetenzbereiche der Fächer Deutsch und Mathematik identifizieren lassen. Zweitens sollen zusätzlich zu den mittleren Kompetenzunterschieden der Geschlechtergruppen auch die Kompetenzverteilungen von Jungen und Mädchen vertiefend betrachtet werden. Nachdem die Befunde zunächst für die bundesweite Gesamtstichprobe dargestellt werden, soll in einem dritten Schritt schließlich geprüft werden, inwieweit die Geschlechterunterschiede zwischen den Ländern variieren.

7.2 Geschlechtsbezogene Kompetenzunterschiede im Ländervergleich 2011

7.2.1 Geschlechtsbezogene Unterschiede in den Kompetenzmittelwerten

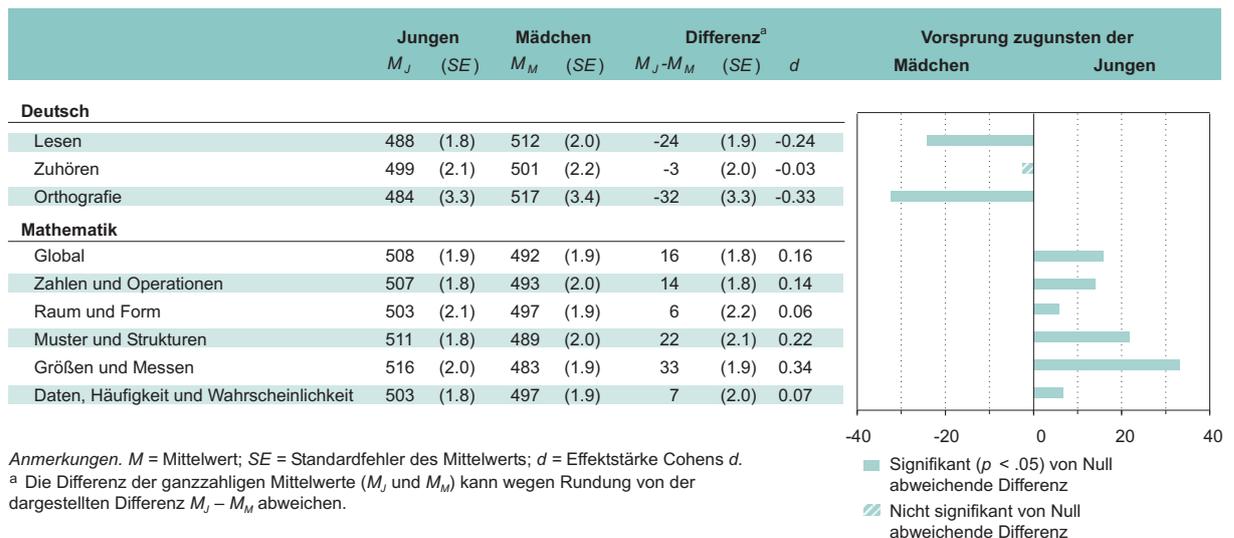
Zunächst werden die deutschlandweit ermittelten Kompetenzunterschiede zwischen Jungen und Mädchen in allen im Ländervergleich 2011 untersuchten Kompetenzbereichen vergleichend dargestellt. Abbildung 7.1 zeigt die jeweils von Jungen (M_J) und Mädchen (M_M) in Deutschland im Mittel erreichten Kompetenzstände sowie deren Differenz ($M_J - M_M$) und die zugehörigen Standardfehler (SE).⁸ Negative Differenzwerte markieren einen Kompetenzvorsprung der Mädchen, positive Differenzwerte hingegen einen Kompetenzvorsprung der Jungen. Die Kompetenzunterschiede zwischen den beiden Geschlechtergruppen sind zum einen als Punkteabstand zwischen den mittleren Leistungen der Jungen und Mädchen auf der Berichtsmetrik (vgl. Kapitel 3.1) und zum anderen als standardisierte Mittelwertsdifferenz (d) dargestellt. Die Betrachtung von Gruppenunterschieden in Punkten auf der Berichtsmetrik hat den Vorteil, dass sich diese anschaulich mit den erwarteten Lernzuwächsen in Beziehung setzen lassen. Da aber in der einschlägigen Literatur häufig d -Werte als Differenzmaße verwendet werden (vgl. Fußnote 3), werden diese für eine bessere Vergleichbarkeit unserer Befunde mit Ergebnissen anderer Studien zusätzlich berichtet. Um die Kompetenzunterschiede zwischen Jungen und Mädchen grafisch zu verdeutlichen, sind diese in Abbildung 7.1 zusätzlich als Balkendiagramm wiedergegeben, wobei die Abweichungen als Punkte auf der Berichtsmetrik dargestellt sind. Balken, die nach links weisen, zeigen einen Vorsprung zugunsten der Mädchen; Balken, die nach rechts weisen, einen Vorsprung der Jungen im jeweiligen Kompetenzbereich. Farbige ausgefüllte Balken markieren Unterschiede, die statistisch signifikant sind, schraffierte Balken kennzeichnen Gruppenunterschiede, die das Signifikanzkriterium verfehlen.

Erwartungsgemäß ergeben sich für die Kompetenzbereiche des Faches Deutsch Vorteile zugunsten der Mädchen, für die Kompetenzbereiche im Fach Mathematik dagegen Vorteile zugunsten der Jungen. Interessant ist hierbei der Befund, dass die geschlechtsbezogenen Disparitäten für verschiedene Kompetenzbereiche *innerhalb* der beiden Fächer stark variieren.

Im Fach Deutsch beträgt der Kompetenzvorsprung der Mädchen im Bereich *Lesen* 24 Punkte und liegt somit im Vergleich zu den in der Literatur berichteten Befunden in einer mittleren Größenordnung. Im Lesen wurde in querschnittlichen Analysen zwischen der dritten und vierten Jahrgangsstufe ein Kompetenzunterschied von etwa 60 Punkten ermittelt (Bremerich-Vos & Böhme, 2009, S. 240). Damit entspricht der Kompetenzvorsprung der Mädchen im Lesen also in etwa dem Lernzuwachs, der in einem Drittel bis einem halben Schuljahr erreicht wird.

Im Kompetenzbereich *Zuhören* beträgt der Unterschied zwischen Jungen und Mädchen lediglich drei Punkte auf der Berichtsmetrik. Diese Differenz ist statistisch nicht signifikant und somit zu vernachlässigen. Der Befund repliziert das Ergebnis von Behrens, Böhme und Krelle (2009), die auf Basis von Daten

⁸ Erklärungen zum Standardfehler und anderen hier verwendeten statistischen Begriffen finden sich in Kapitel 5.1

Abbildung 7.1: Kompetenzunterschiede zwischen Jungen und Mädchen der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik

der Pilotierung der bildungsstandardbasierten Testaufgaben im Bereich Zuhören ebenfalls keine signifikanten Unterschiede zwischen Jungen und Mädchen beobachten konnten.

Am größten ist der Vorsprung der Mädchen im Kompetenzbereich *Orthografie* mit 32 Punkten.⁹ Da in diesem Teilbereich der zu erwartende Kompetenzzuwachs von Jahrgangsstufe drei nach vier mit circa 70 Punkten etwas größer als im Lesen und Zuhören ist (vgl. Böhme & Bremerich-Vos, 2009, S. 345), entspricht der Vorteil der Mädchen hier etwa einem halben Schuljahr.

Im Fach Mathematik findet sich für die *globale mathematische Kompetenz* der Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe kein sehr großer, aber statistisch signifikanter Vorsprung zugunsten der Jungen im Umfang von 16 Punkten. Relativiert man diesen Kompetenzunterschied anhand der innerhalb eines Schuljahres zu erwartenden Kompetenzsteigerung, die sich nach Reiss und Winkelmann (2009, S. 139) auf ungefähr 80 Punkte beläuft, entspricht dieser Unterschied etwa dem Fünftel eines Schuljahres. Der Kompetenzvorteil der Jungen hinsichtlich der globalen mathematischen Kompetenz ist also deutlich geringer als der Vorsprung der Mädchen in den Kompetenzbereichen Lesen und Orthografie.

Betrachtet man zusätzlich zur Globalskala der mathematischen Kompetenz auch die Ergebnisse für die einzelnen inhaltlichen Leitideen (vgl. Kapitel 2.2), ergeben sich differenzielle Befunde. In den Bereichen *Raum und Form* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* beträgt der Kompetenzvorsprung der Jungen vor den Mädchen weniger als 10 Punkte und ist damit zwar statistisch signifikant, aber sehr klein. Im Kompetenzbereich *Zahlen und Operationen* schneiden die Jungen um 14 Punkte besser ab als die Mädchen, was in etwa dem Geschlechterunterschied auf der Globalskala mathematischer Kompetenz

9 Wie in Kapitel 12 erläutert, weicht die Stichprobe derjenigen Schülerinnen und Schüler, die Aufgaben des Kompetenzbereichs Orthografie bearbeitet haben, von der Stichprobe der anderen Kompetenzbereiche ab, da für Orthografie kein Ländervergleich angestrebt wurde. In die Orthografie-Stichprobe wurden keine Schülerinnen und Schüler an Förderschulen einbezogen. Da Jungen an Förderschulen in der Regel überrepräsentiert und dort vermutlich eher dem unteren Kompetenzspektrum zuzuordnen sind, dürften die hier berichteten geschlechtsbezogenen Unterschiede für den Bereich Orthografie wahrscheinlich eine leichte Unterschätzung der tatsächlichen Disparitäten darstellen.

entspricht. Größer sind die Differenzen dagegen mit mehr als 20 Punkten im Bereich *Muster und Strukturen* und mit mehr als 30 Punkten im Bereich *Größen und Messen*. Diese Kompetenzunterschiede entsprechen einer Lernzeit von einem Drittel bis knapp einem halben Schuljahr. Bereits Winkelmann und van den Heuvel-Panhuizen (2009, S. 147f.) identifizierten differenzielle Kompetenzunterschiede zwischen Jungen und Mädchen in Abhängigkeit von der jeweiligen inhaltlichen Leitidee. Die von diesen Autoren gefundenen Kompetenzvorsprünge waren insgesamt etwas größer als die hier ermittelten Differenzen und auch die relative Höhe der von ihnen berichteten Kompetenzunterschiede weicht von den hier für die einzelnen Leitideen identifizierten Unterschieden leicht ab. Dennoch ergibt sich hier wie dort konsistent der Befund, dass die geschlechtsbezogenen Unterschiede in den Bereichen *Raum und Form* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* eher geringer und im Bereich *Größen und Messen* mit Abstand am stärksten ausgeprägt sind.

Insgesamt kann festgehalten werden, dass sich erwartungskonform im Fach Deutsch in den Bereichen Lesen und Orthografie statistisch bedeutsame Kompetenzvorsprünge zugunsten der Mädchen identifizieren lassen, während in Mathematik für alle inhaltlichen Leitideen und für die globale mathematische Kompetenz signifikante Vorteile der Jungen ermittelt werden. Einzig im Kompetenzbereich Zuhören finden sich keine statistisch bedeutsamen Geschlechterunterschiede.

Trotz bestehender Mittelwertsdifferenzen sollte allerdings berücksichtigt werden, dass sich die Kompetenzverteilungen von Jungen und Mädchen in weiten Bereichen überschneiden und dass die Unterschiede in den Kompetenzständen *innerhalb* der Geschlechtergruppen deutlich ausgeprägter sind, als die Differenzen *zwischen* beiden Geschlechtern (vgl. z. B. Stanat & Bergann, 2009).

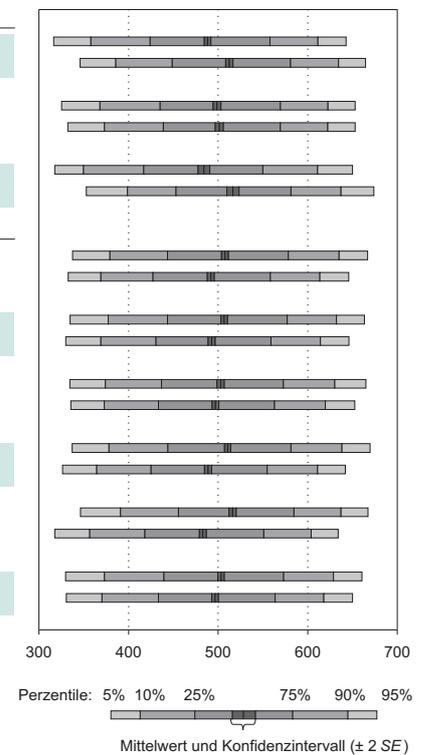
7.2.2 Geschlechtsbezogene Unterschiede in den Kompetenzverteilungen

Um Aufschluss darüber zu erhalten, ob sich neben den Geschlechterunterschieden in den im Mittel erreichten Kompetenzen auch Geschlechterunterschiede in den Leistungsverteilungen und damit relative Vor- oder Nachteile in bestimmten Leistungssegmenten ergeben, werden in diesem Abschnitt die *Leistungsstreuungen* von Jungen und Mädchen in Deutschland insgesamt untersucht. Hierfür sind in Abbildung 7.2 ergänzend zu den bereits beschriebenen Kompetenzmittelwerten verschiedene Maße der Leistungsvarianz dargestellt. Dies sind zum einen die Standardabweichungen (*SD*) und zum anderen verschiedene Perzentile. Letztere sind in Abbildung 7.2 zusätzlich als Perzentilbänder grafisch veranschaulicht. Sie stellen die Kompetenzverteilungen insgesamt dar und verdeutlichen die Unterschiedlichkeit einzelner Leistungssegmente. Beispielsweise befinden sich unterhalb des 5. Perzentils eines Kompetenzbereichs diejenigen 5 Prozent der Schülerinnen und Schüler, die die niedrigsten Kompetenzstände aufweisen. Der Wert für das 5. Perzentil gibt also den höchsten Kompetenzwert an, der von den 5 Prozent leistungsschwächsten Jungen und Mädchen erreicht wurde. Ein hoher Wert für das 5. Perzentil weist darauf hin, dass auch leistungsschwächere Kinder noch relativ hohe Kompetenzwerte erzielen. Wird am anderen Ende des Kompetenzspektrums für das 95. Perzentil ein hoher Wert festgestellt, so ist dies ein Beleg für ausgeprägte Kompetenzen im obersten Leistungsbereich. Unterhalb und oberhalb des Mittelwerts wurden jeweils drei Perzentile bestimmt: das 5., 10. und 25. sowie das 75., 90. und 95. Perzentil.

Abbildung 7.2: Mittelwerte, Streuungen, Perzentile und Perzentilbänder der von Jungen und Mädchen erreichten Kompetenzstände

		<i>M</i>	(<i>SE</i>)	<i>SD</i>	5	10	25	75	90	95
Deutsch										
Lesen	Jungen	488	(1.8)	101	317	358	424	558	611	643
	Mädchen	512	(2.0)	98	346	386	449	581	635	665
Zuhören	Jungen	499	(2.1)	102	325	368	435	570	623	653
	Mädchen	501	(2.2)	98	332	373	439	570	622	653
Orthografie	Jungen	484	(3.3)	101	318	350	417	550	611	650
	Mädchen	517	(3.4)	96	353	399	453	581	637	674
Mathematik										
Global	Jungen	508	(1.9)	103	337	379	443	578	635	667
	Mädchen	492	(1.9)	96	333	369	427	558	614	646
Zahlen und Operationen	Jungen	507	(1.8)	102	335	377	444	577	632	663
	Mädchen	493	(2.0)	97	330	369	431	559	614	646
Raum und Form	Jungen	503	(2.1)	103	335	374	437	573	630	665
	Mädchen	497	(1.9)	97	336	373	433	563	620	653
Muster und Strukturen	Jungen	511	(1.8)	102	337	378	444	581	638	670
	Mädchen	489	(2.0)	97	326	365	425	555	611	642
Größen und Messen	Jungen	516	(2.0)	100	346	391	456	585	637	667
	Mädchen	483	(1.9)	97	318	357	418	551	604	634
Daten, Häufigkeit und Wahrscheinlichkeit	Jungen	503	(1.8)	102	330	373	439	573	629	660
	Mädchen	497	(1.9)	98	331	370	433	564	618	650

Anmerkung. *M* = Mittelwert; *SE* = Standardfehler des Mittelwerts; *SD* = Standardabweichung.



Insgesamt zeigt sich über alle Kompetenzbereiche der Fächer Deutsch und Mathematik hinweg, dass die Kompetenzstände der Mädchen etwas homogener sind als die der Jungen. Dies ist zum einen an den unterschiedlichen Standardabweichungen beider Gruppen zu erkennen: Während diese für die Mädchen immer unterhalb von 100 Punkten liegen, werden für die Jungen stets Standardabweichungen von 100 oder mehr Punkten ermittelt. Die größten Unterschiede finden sich im Fach Deutsch für den Kompetenzbereich Orthografie ($SD_M = 96$, $SD_J = 101$) und für die globale Mathematikkompetenz ($SD_M = 96$, $SD_J = 103$).

Auch ein Vergleich unterschiedlicher Perzentile und damit unterschiedlicher Segmente der Kompetenzverteilungen kann wichtige Informationen zu differenziellen Kompetenzständen von Jungen und Mädchen liefern. Im Fach Deutsch sind sich die Verteilungen im Kompetenzbereich Zuhören am ähnlichsten. Im unteren Leistungsspektrum, markiert durch das 5. Perzentil, beträgt der Kompetenzvorsprung der Mädchen vor den Jungen 7 Punkte und ist somit als gering einzuschätzen. Im oberen Kompetenzbereich liegen die Werte des 95. Perzentils für Jungen und Mädchen bei jeweils 653 Punkten und sind somit identisch. Der von Jungen und Mädchen im Bereich Zuhören erreichte Stand der Kompetenzentwicklung ist also am Ende der vierten Jahrgangsstufe sowohl hinsichtlich der mittleren Kompetenzausprägung als auch in Bezug auf die Kompetenzverteilung sehr homogen. Es finden sich keine nennenswerten geschlechtsbezogenen Disparitäten.

In den Kompetenzbereichen Lesen und Orthografie weisen die Kompetenzverteilungen von Jungen und Mädchen jedoch unterschiedliche Verteilungsformen auf. Betrachtet man jeweils den unteren Leistungsbereich, ergeben

sich im Lesen für die Mädchen 346 und für die Jungen 317 Punkte für das 5. Perzentil. Dies entspricht einem Vorsprung der Mädchen von fast 30 Punkten im unteren Kompetenzspektrum. Im oberen Leistungsbereich erzielen Mädchen Werte von 665 und Jungen von 643 Punkten für das 95. Perzentil. Der Vorsprung der Mädchen ist im oberen Kompetenzsegment mit 22 Punkten also etwas geringer als im unteren Leistungsbereich. Geschlechtsbezogene Nachteile für Jungen mit sehr geringer Lesekompetenz sind zum Ende der Grundschulzeit also etwas stärker ausgeprägt als für Jungen, die sehr hohe Lesekompetenzwerte erzielen. Ein ganz ähnliches Befundmuster ergibt sich für den Teilbereich Orthografie. Hier beträgt der Abstand zwischen Jungen und Mädchen im unteren Kompetenzspektrum (5. Perzentil) 35 Punkte, im obersten Leistungssegment (95. Perzentil) jedoch nur 24 Punkte. Auch hier sind somit die geschlechtsbezogenen Disparitäten im unteren Leistungsbereich größer als bei Jungen und Mädchen mit hohen Kompetenzwerten.

Für das Fach Mathematik liefert die Analyse der Kompetenzverteilungen von Jungen und Mädchen ein etwas anderes Bild. So ergeben sich für die globale mathematische Kompetenz und für die inhaltlichen Kompetenzbereiche *Zahlen und Operationen*, *Raum und Form* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* sehr ähnliche Resultate. In diesen Bereichen bestehen im unteren Leistungssegment, markiert durch das 5. Perzentil, nur geringe Kompetenzunterschiede zwischen Jungen und Mädchen im Umfang von 5 Punkten oder weniger. Im oberen Kompetenzspektrum (95. Perzentil) ist der Kompetenzvorsprung der Jungen dagegen deutlich größer und bewegt sich zwischen einem Vorteil von 10 Punkten im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* und einer Differenz von 21 Punkten für die globale mathematische Kompetenz. Die geschlechtsbezogenen Unterschiede sind in den genannten Bereichen also im unteren Leistungssegment sehr gering und steigen mit zunehmendem Kompetenzniveau an.

Etwas abweichend hiervon sind die Kompetenzverteilungen für Jungen und Mädchen in den Bereichen *Muster und Strukturen* sowie *Größen und Messen* über das gesamte Kompetenzspektrum hinweg gegeneinander verschoben, so dass die Jungen in allen Segmenten der Verteilung recht deutlich überlegen sind. Auch hier sind jedoch die Kompetenzunterschiede im obersten Bereich der Verteilung etwas größer. So beträgt die Differenz der Werte zugunsten der Jungen für das 5. Perzentil im Bereich *Größen und Messen* 28 Punkte und für das 95. Perzentil 33 Punkte. Im Kompetenzbereich *Muster und Strukturen* vergrößert sich der Kompetenzunterschied zwischen Jungen und Mädchen von 11 Punkten für das 5. Perzentil auf 28 Punkte im oberen Kompetenzsegment (95. Perzentil). In diesem Kompetenzbereich der Mathematik steigen also die geschlechtsbezogenen Disparitäten am deutlichsten mit dem Kompetenzniveau an.

Insgesamt zeigen sich im Fach Mathematik im Mittel etwas geringere geschlechtsbezogene Kompetenzunterschiede als im Lesen und in der Orthografie. Deutlich wird ferner, dass im Fach Mathematik die Differenzen in den Kompetenzständen zwischen Jungen und Mädchen insbesondere das obere Leistungssegment betreffen. Hier sind die Vorsprünge der Jungen noch ausgeprägter als bei Kindern mit geringeren mathematischen Kompetenzen.

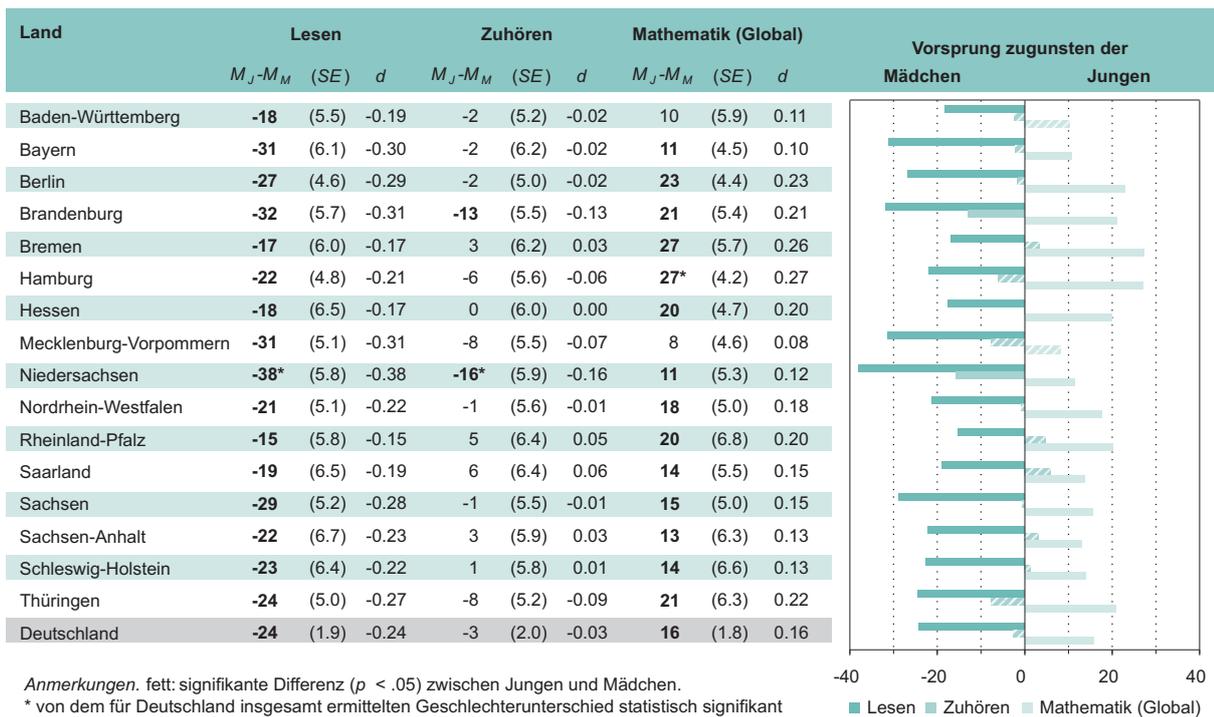
7.2.3 Geschlechtsbezogene Unterschiede in den Ländern

Bislang standen die für Deutschland insgesamt ermittelten geschlechtsbezogenen Unterschiede im Fokus. In diesem letzten Teilkapitel soll auf die Unterschiede in den einzelnen Ländern eingegangen werden. Wie auch in den anderen Kapiteln dieses Berichts beziehen sich die Analysen der Länderunterschiede ausschließlich auf die Kompetenzbereiche Lesen, Zuhören und die globale mathematische Kompetenz.

Nachfolgend sollen ländervergleichend zwei Fragen beantwortet werden. Zum einen wird bestimmt, ob die Geschlechterunterschiede in den verschiedenen Kompetenzbereichen innerhalb der einzelnen Länder signifikant sind. Zum anderen soll untersucht werden, ob die für die einzelnen Länder identifizierten Geschlechterunterschiede sich signifikant von den für Deutschland insgesamt ermittelten Differenzen unterscheiden.

Abbildung 7.3 gibt einen Überblick der Kompetenzunterschiede zwischen Jungen und Mädchen in den Bereichen Lesen, Zuhören und Mathematik (Globalskala). Im Tabellenteil der Abbildung sind für Lesen, Zuhören und die mathematische Kompetenz die Mittelwertsdifferenzen ($M_J - M_M$) als Punkte auf der Berichtsmetrik mit dem zugehörigen Standardfehler (SE) ausgewiesen. Ferner finden sich in Abbildung 7.3 für jeden Kompetenzbereich die entsprechenden standardisierten Mittelwertsdifferenzen (d). Negative Werte kennzeichnen wiederum Vorsprünge der Mädchen und positive Werte Vorteile der Jungen. Im rechten Teil der Abbildung 7.3 sind die Mittelwertsdifferenzen als Kompetenzpunkte zusätzlich grafisch dargestellt, wobei Balken, die nach links weisen, einen Vorsprung zugunsten der Mädchen und Balken, die nach rechts weisen, einen Vorsprung zugunsten der Jungen markieren.

Für den Kompetenzbereich Lesen ergeben sich in allen Ländern der Bundesrepublik Deutschland statistisch bedeutsame Vorsprünge zugunsten der Mädchen. Der für Deutschland insgesamt ermittelte Kompetenzunterschied beträgt dabei 24 Punkte. In vielen Ländern finden sich im Kompetenzbereich Lesen geschlechtsbezogene Differenzen, die sich in einer sehr ähnlichen Größenordnung bewegen. Es gibt aber auch eine Reihe von Ländern, in denen die Disparitäten zwischen Jungen und Mädchen wesentlich geringer oder wesentlich stärker ausgeprägt sind. Tendenziell kleinere Unterschiede als für Deutschland insgesamt finden sich im Lesen in Baden-Württemberg und Hessen mit einem Kompetenzvorsprung der Mädchen von 18 Punkten, in Bremen mit 17 Punkten sowie in Rheinland-Pfalz mit lediglich 15 Punkten. Obwohl die Kompetenzunterschiede zwischen Jungen und Mädchen in diesen Ländern etwas geringer als in Deutschland insgesamt ausfallen, sind diese Abweichungen von dem für Deutschland insgesamt berechneten Geschlechterunterschied nicht statistisch signifikant. Besonders große geschlechtsbezogene Unterschiede ergeben sich für den Bereich Lesen hingegen in Bayern sowie Mecklenburg-Vorpommern mit einem Vorteil der Mädchen von je 31 Punkten sowie in Brandenburg mit 32 Punkten. Der mit Abstand größte Kompetenzvorsprung der Mädchen vor den Jungen zeigt sich im Lesen in Niedersachsen. Hier beträgt die Differenz zwischen den Geschlechtergruppen 38 Punkte und entspricht damit dem Lernzuwachs von etwa zwei Dritteln eines Schuljahres (vgl. Bremerich-Vos & Böhme, 2009). Nur der für Niedersachsen ermittelte Kompetenzvorteil der Mädchen weicht statistisch signifikant von der für Deutschland insgesamt errechneten Differenz im Bereich Lesen ab.

Abbildung 7.3: Geschlechtsbezogene Kompetenzunterschiede von Viertklässlerinnen und Viertklässlern in den Ländern in den Fächern Deutsch und Mathematik

Anmerkungen. fett: signifikante Differenz ($p < .05$) zwischen Jungen und Mädchen.

* von dem für Deutschland insgesamt ermittelten Geschlechterunterschied statistisch signifikant abweichende Differenz. Schraffierte Balken markieren nicht statistisch signifikante Differenzen zwischen Jungen und Mädchen.

Wie oben erläutert, findet sich im Kompetenzbereich Zuhören in Deutschland insgesamt kein signifikanter Kompetenzunterschied zwischen Jungen und Mädchen. Dies ist auch innerhalb der Länder überwiegend der Fall. In vielen Ländern sind geringe Vorteile der Mädchen zu verzeichnen, mitunter zeigen sich auch tendenzielle, allerdings nie statistisch signifikante Kompetenzvorsprünge der Jungen, so beispielsweise in Rheinland-Pfalz oder im Saarland. Die größten Kompetenzvorsprünge zugunsten der Mädchen zeigen sich in Brandenburg mit 13 Punkten und in Niedersachsen mit 16 Punkten. Nur in diesen beiden Ländern sind die geschlechtsbezogenen Disparitäten im Kompetenzbereich Zuhören statistisch signifikant. Zusätzlich ist der Vorsprung der Mädchen in Niedersachsen signifikant größer als in Deutschland insgesamt. Damit ist Niedersachsen das einzige Land, in dem sowohl für das Lesen als auch für das Zuhören geschlechtsbezogene Disparitäten gefunden wurden, die in statistisch bedeutsamem Umfang über die für Deutschland insgesamt ermittelten Differenzen hinausgehen.

Im Fach Mathematik liegt der Kompetenzvorteil der Jungen in Deutschland insgesamt bei 16 Punkten. Auch in den einzelnen Ländern erzielten Jungen durchgehend höhere mittlere Kompetenzstände als Mädchen. Besonders groß sind die Unterschiede zugunsten der Jungen in den drei Stadtstaaten. Sie belaufen sich in Berlin auf 23 und in Hamburg sowie Bremen auf 27 Punkte. In Hamburg ist der geschlechtsbezogene Kompetenzunterschied statistisch signifikant größer als in Deutschland insgesamt. Trotz eines identischen Differenzwertes ist dies für Bremen nicht der Fall, was auf den größeren Standardfehler der für Bremen berechneten Mittelwertsdifferenz zwischen Jungen und Mädchen zurückzuführen ist. Sehr gering und statistisch nicht signifikant sind die Kompetenzunterschiede zwischen Jungen und Mädchen in den Ländern Baden-Württemberg (10 Punkte) und Mecklenburg-Vorpommern (8 Punkte), diese sind aber nicht signifikant geringer als in Deutschland insgesamt.

7.3 Zusammenschau der Befunde und Fazit

Das vorliegende Kapitel berichtet über geschlechtsbezogene Kompetenzunterschiede im Ländervergleich 2011. Erwartungskonform ergeben sich für Deutschland insgesamt in den Kompetenzbereichen Lesen und Orthografie im Mittel signifikante Kompetenzvorteile zugunsten der Mädchen und im Fach Mathematik Kompetenzvorteile zugunsten der Jungen. Lediglich für den Kompetenzbereich Zuhören werden keine geschlechtsbezogenen Disparitäten identifiziert.

Eine vergleichende Analyse der Kompetenzverteilungen ergibt durchgehend etwas geringere Leistungsstreuungen für Mädchen als für Jungen. Im Hinblick auf Unterschiede zwischen Jungen und Mädchen in verschiedenen Leistungssegmenten zeigt sich, dass im Lesen und im Teilbereich Orthografie die Kompetenzvorsprünge der Mädchen im unteren Leistungsbereich tendenziell stärker ausgeprägt sind als im oberen Kompetenzspektrum. Im Fach Mathematik zeichnet sich insofern das gegenteilige Befundmuster ab, als geschlechtsbezogene Disparitäten hier insbesondere das obere Kompetenzniveau betreffen und die Vorsprünge der Jungen im oberen Leistungsbereich deutlich ausgeprägter sind als im unteren Kompetenzspektrum. Trotz der beschriebenen Differenzen ist jedoch festzuhalten, dass die Leistungsstreuungen innerhalb der Geschlechtergruppen wesentlich größer sind als die mittleren Kompetenzunterschiede zwischen Jungen und Mädchen.

Relevant sind Analysen von Disparitäten und somit auch die hier berichteten Befunde, um Benachteiligungen im schulischen Kontext, die mit der Zugehörigkeit der Kinder und Jugendlichen zu einer bestimmten Gruppe verbunden sind, zu identifizieren und diese gegebenenfalls durch gezielte Förderung reduzieren zu können. Dies betrifft mit Merkmalen der sozialen Herkunft und des Zuwanderungshintergrunds assoziierte Unterschiede (vgl. Kapitel 8 und 9), aber auch Geschlechterunterschiede.

In Übereinstimmung mit den oben berichteten Befunden aus der einschlägigen Literatur weisen die in diesem Kapitel dargestellten Befunde darauf hin, dass gezielte Förderung im Primarschulbereich zur Reduktion geschlechtsbezogener Kompetenzunterschiede nach wie vor erforderlich ist. Diese sollte sich im Fach Deutsch auf die Stärkung der Lesekompetenz und die Unterstützung der Entwicklung der orthografischen Kompetenz der Jungen richten. Im Kompetenzbereich Zuhören bestehen dagegen keine geschlechtsbezogenen Disparitäten, weshalb für diesen Bereich eine geschlechtsspezifische Förderung nicht angezeigt ist. Im Fach Mathematik weisen die Ergebnisse darauf hin, dass eine gezielte Förderung der Mädchen insbesondere die Bereiche *Muster und Strukturen* sowie *Größen und Messen* in den Blick nehmen sollte.

Im Sinne von Budde (2011) ist allerdings zu betonen, dass es in der unterrichtlichen Praxis vermutlich nicht zielführend wäre, geschlechtsbezogene Unterschiede in der Unterrichtsgestaltung oder durch spezifische Förderprogramme zu sehr in den Vordergrund zu rücken. Dies könnte zu einer verstärkten Stereotypisierung *der* Mädchen und *der* Jungen führen, die der tatsächlichen Heterogenität der Schülerinnen und Schüler nicht gerecht wird. Denn Geschlechtergerechtigkeit bedeutet nicht nur, Jungen wie Mädchen im schulischen Kontext gleiche Chancen einzuräumen und ihnen die bestmögliche Förderung zukommen zu lassen, sondern auch, Stereotype abzubauen.

Literatur

- Anders, Y., McElvany, N. & Baumert, J. (2010). Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule. Wie differenziert urteilen Lehrkräfte? In K. Maaz, J. Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 313–330). Bonn: Bundesministerium für Bildung und Forschung.
- Autorengruppe Bildungsberichterstattung. (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann Verlag.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillman, K.-J. & Weiß, M. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Behrens, U., Böhme, K. & Krelle, M. (2009). Zuhören – Operationalisierung und fachdidaktische Implikationen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 357–375). Weinheim: Beltz.
- Benölken, R. (2011). *Mathematisch begabte Mädchen. Untersuchungen zu geschlechts- und begabungsspezifischen Besonderheiten im Grundschulalter*. Münster: WTM Verlag für wissenschaftliche Texte und Medien.
- Blossfeld, H.-P., Bos, W., Hannover, B., Lenzen, D., Müller-Böling, D., Prenzel, M. & Wößmann, L. (2009). *Geschlechterdifferenzen im Bildungssystem. Jahresgutachten 2009*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Böhme, K. & Bremerich-Vos, A. (2009). Diagnostik der Rechtschreibkompetenz in der Grundschule – Konstruktprüfung mittels Fehler- und Dimensionsanalysen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 330–356). Weinheim: Beltz.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2007). *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2008). *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Budde, J. (2011). Geschlechtersensible Schule. In H. Faulstich-Wieland (Hrsg.), *Umgang mit Heterogenität und Differenz* (S. 99–119). Baltmannsweiler: Schneider Verlag Hohengehren.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Diefenbach, H. (2010). Jungen – die „neuen“ Bildungsverlierer. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 245–271). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research*, 62, 61–84.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hanna, G. (2000). Declining gender differences from FIMS to TIMSS. *Zentralblatt für Didaktik der Mathematik / International reviews on mathematical education*, 32, 11–17.
- Hannover, B. (2008). Vom biologischen zum psychologischen Geschlecht: Die Entwicklung von Geschlechtsunterschieden. In A. Renkl (Hrsg.), *Lehrbuch Pädagogische Psychologie* (S. 339–388). Bern: Huber.

- Hannover, B. & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25, 89–103.
- Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? [Are female teachers responsible for the school performance gap between boys and girls?]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 93–111.
- Hyde, J. S. (1981). How large are cognitive gender differences? *American Psychologist*, 36, 892–901.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581–592.
- Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107, 139–155.
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B. & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321, 494–495.
- Köller, O., Knigge, M. & Tesch, B. (Hrsg.). (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Kowalski, K., Voss, A., Valtin, R. & Bos, W. (2010). Erhebungen zur Orthographie in IGLU 2001 und IGLU 2006: Haben sich die Rechtschreibleistungen verbessert? In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert, I. Tarelli & R. Valtin (Hrsg.), *IGLU 2006 – die Grundschule auf dem Prüfstand. Vertiefende Analysen zu Rahmenbedingungen schulischen Lernens* (S. 33–42). Münster: Waxmann.
- Naumann, J., Artelt, C., Schneider, W. & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23–71). Münster: Waxmann.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 120–141). Weinheim: Beltz.
- Rohe, A. M. & Quaiser-Pohl, C. (2010). Prädiktoren für mathematische Kompetenzen zu Beginn der Grundschule – Gibt es Unterschiede zwischen Mädchen und Jungen? In C. Quaiser-Pohl & M. Endepohls-Ulpe (Hrsg.), *Bildungsprozesse im MINT-Bereich. Interesse, Partizipation und Leistungen von Mädchen und Jungen* (S. 13–27). Münster: Waxmann.
- Stanat, P. & Bergann, S. (2009). Geschlechtsbezogene Disparitäten in der Bildung. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 513–527). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Statistisches Bundesamt. (Hrsg.). (2011). *Fachserie 11: Bildung und Kultur, Reihe 1: Allgemeinbildende Schulen*. Wiesbaden: Statistisches Bundesamt. Ausschließlich im Internet veröffentlichte Tabellen zugänglich unter https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/Allgemein_bildende_Schulen.html
- Trautner, H. M. (2002). Entwicklung der Geschlechtsidentität. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 648–674). Weinheim: Beltz.
- Valtin, R., Bos, W., Buddeberg, I., Goy, M. & Potthoff, B. (2008). Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe im nationalen und internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 51–101). Münster: Waxmann.
- Wieczerkowski, W. & Prado, T. M. (1992). Begabung und Geschlecht. In E. A. Hany & H. Nickel (Hrsg.), *Begabung und Hochbegabung. Theoretische Konzepte, empirische Befunde, praktische Konsequenzen* (S. 39–57). Bern: Huber.
- Winkelmann, H. & Groeneveld, I. (2010). Geschlechterdisparitäten. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 177–184). Münster: Waxmann.
- Winkelmann, H. & van den Heuvel-Panhuizen, M. (2009). Geschlechtsspezifische mathematische Kompetenz. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 142–156). Weinheim: Beltz.

Kapitel 8

Soziale Disparitäten

Dirk Richter, Poldi Kuhl und Hans Anand Pant

8.1 Einführung

Große Aufmerksamkeit erregten in den vergangenen Jahren die Befunde internationaler Schulleistungsstudien, die darauf hinwiesen, dass die Kompetenzen von Schülerinnen und Schülern in Deutschland in ganz besonders starkem Maße mit der sozialen Herkunft der Familie zusammenhängen. Für den Bereich der Grundschule zeigten die Untersuchungen einen signifikant positiven Zusammenhang zwischen der sozialen Herkunft und den erreichten Kompetenzen im Lesen (IGLU¹ 2006, Bos et al., 2007) sowie in Mathematik und den Naturwissenschaften (TIMSS² 2007, Bos et al., 2008). Schülerinnen und Schüler aus Familien mit einem niedrigen sozioökonomischen Status wiesen im Vergleich zu Kindern aus Familien mit höherem sozioökonomischen Status deutlich niedrigere Kompetenzwerte auf. Ähnliche Befundmuster zeigten sich auch in der PISA-Studie³ für die Sekundarstufe I (z. B. Ehmke & Baumert, 2008; Ehmke & Jude, 2010). Im Vergleich zu anderen Nationen bestand in Deutschland dabei eine besonders enge Kopplung zwischen sozialer Herkunft und den gemessenen Kompetenzen, die sich in den vergangenen Jahren, zumindest für die Sekundarstufe I, leicht abschwächte (Ehmke & Jude, 2010).

Sozial bedingte Bildungsungleichheiten, so fassen Maaz, Baumert und Trautwein (2010) zusammen, können an ganz verschiedenen Stellen ihren Ausgang nehmen. In der Bildungsforschung wurde ihre Entstehung und Entwicklung sowohl bei Bildungsübergängen als auch innerhalb von Bildungsinstitutionen, zwischen Bildungsinstitutionen und außerhalb des Bildungssystems untersucht (vgl. Krüger, Rabe-Kleberg, Kramer & Budde, 2010). Innerhalb des Systems Schule wurden vielfach soziale Disparitäten in der Sekundarstufe I sowie beim Übergang von der Primarstufe in die Sekundarstufe I betrachtet. Insbesondere der Übergang gilt als ein zentraler Zeitpunkt in der Bildungskarriere von Schülerinnen und Schülern, an dem Bildungsungleichheiten verstärkt werden (Maaz, Baumert, Gresch & McElvany, 2010). So konnten Arnold, Bos, Richter und Stubbe (2007) auf der Grundlage von Daten der IGLU-2006-Studie nachweisen, dass die Wahrscheinlichkeit einer Gymnasialempfehlung für Kinder oberer Sozialschichten bei gleicher Leistung und kognitiver Grundfähigkeit größer ist als für Kinder niedrigerer Sozialschichten.

Zur Erklärung sozialer Bildungsungleichheiten werden vielfach soziologische Theorien von Bourdieu (Bourdieu, Passeron & Nice, 1990; vgl. Kramer

1 Das Akronym IGLU steht für *Internationale Grundschul-Lese-Untersuchung*.

2 Das Akronym TIMSS steht für *Trends in International Mathematics and Science Study*.

3 Das Akronym PISA steht für *Programme for International Student Assessment*.

& Helsper, 2010) und Boudon (1986) herangezogen, die auf die Bedeutsamkeit sozialschichtspezifischer Struktur- und Prozessmerkmale für die Entstehung von Bildungsungleichheiten hinweisen. Die Bedeutung solcher Struktur- und Prozessmerkmale als Erklärungsfaktoren herkunftsbezogener Leistungsunterschiede in der Lesekompetenz untersuchten McElvany, Becker und Lüdtke (2009) für den Primarbereich. Die Autoren konnten zeigen, dass familiäre Strukturmerkmale, wie die Bildung der Eltern und der sozioökonomische Status der Familie, über familiäre Prozessmerkmale, wie beispielsweise kulturelle Ressourcen und Praxen, auf individuelle Merkmale der Schülerinnen und Schüler (Lesemotivation, Leseverhalten, Wortschatz) vermittelt werden, die wiederum den Erwerb der Lesekompetenz begünstigen oder behindern (vgl. auch Baumert, Watermann & Schümer, 2003).

Die ländervergleichenden Analysen großer nationaler Schulleistungsstudien haben darüber hinaus darauf hingewiesen, dass deutliche Unterschiede im Grad der Kopplung zwischen sozialer Herkunft und den erreichten Kompetenzständen zwischen den Ländern bestehen (PISA-E 2006: Ehmke & Baumert, 2008; Ländervergleich 2009: Knigge & Leucht, 2010; IGLU-E 2006: Stubbe, Bos & Hornberg, 2008). Ein Vergleich der Länder in der Bundesrepublik Deutschland in Bezug auf die sozialen Disparitäten wurde in der Grundschule bislang ausschließlich für die Lesekompetenz in der deutschen Ergänzungsstudie zu IGLU 2006 (IGLU-E 2006) durchgeführt (Stubbe et al., 2008). Dort zeigten sich große Länderunterschiede für den Zusammenhang zwischen der Leseleistung und dem kulturellen Kapital der Familie, gemessen an der Anzahl der im Haushalt verfügbaren Bücher. In den Stadtstaaten Berlin und Hamburg unterschieden sich die Leseleistungen von Schülerinnen und Schülern aus Familien mit mehr als 100 Büchern von denen mit weniger als 100 Büchern im Mittel um 60 bis 70 Punkte. Für das Land Bayern fand sich für die gleichen Vergleichsgruppen eine Leistungsdifferenz von lediglich 25 Punkten. Diese Länderunterschiede bestätigten sich auch in einer Zusatzanalyse für den Bildungsbericht 2010, die den Zusammenhang zwischen dem höchsten sozioökonomischen Status der Familie und den erreichten Leseleistungen untersuchte (Autorengruppe Bildungsberichterstattung, 2010).

Die Befunde des Ländervergleichs Sprachen, der in der Sekundarstufe I im Jahr 2009 durchgeführt wurde, weisen ebenfalls darauf hin, dass die soziale Herkunft, gemessen an der beruflichen Stellung der Eltern, mit den schulischen Leistungen in Deutsch und auch in Englisch in Beziehung steht (Knigge & Leucht, 2010). Schülerinnen und Schüler, deren Eltern eine höhere berufliche Qualifikation erworben hatten, erreichten deutlich höhere Kompetenzniveaus als Schülerinnen und Schüler, deren Eltern nur über geringe berufliche Qualifikationen verfügten. Im Ländervergleich 2009 standen die Kompetenzen von Schülerinnen und Schülern im Fach Deutsch (Lesen und Zuhören) in Berlin in einem besonders hohen Ausmaß mit der sozialen Herkunft der Eltern in Beziehung. In Brandenburg, Sachsen und Schleswig-Holstein hingegen war der Zusammenhang im Vergleich zum Bundeswert signifikant geringer.

In diesem Kapitel werden auf Länderebene Zusammenhänge zwischen der sozialen Herkunft und Schülerkompetenzen erstmals seit IGLU 2006 wieder für den Primarbereich berichtet. Zusätzlich werden die Ergebnisse um weitere Indikatoren sozialer Herkunft ergänzt und im Vergleich zu den Ergebnissen von 2006 betrachtet. Darüber hinaus ermöglichen die Analysen auch Vergleiche zu Ergebnissen des Ländervergleichs Sprachen, der im Jahr 2009 in der Sekundarstufe I durchgeführt wurde (Knigge & Leucht, 2010). Betrachtet wer-

den Beziehungen zwischen den Indikatoren der sozialen Herkunft und den im Ländervergleich 2011 erfassten Kompetenzen im Lesen, im Zuhören und in Mathematik. Auch in diesem Kapitel werden neben den Ergebnissen der Länder Befunde für eine Vergleichsgruppe deutscher Großstädte⁴ berichtet, die mehr als 300 000 Einwohner haben (ohne Berlin, Bremen und Hamburg). Dies schließt an Analysen aus der ersten PISA-E-Studie an, die für die Sekundarstufe I nachweisen konnten, dass in urbanen Zentren Leistung und soziale Herkunft in einer besonders engen Beziehung zueinander stehen (Baumert & Schümer, 2002).

8.2 Indikatoren sozialer Disparitäten

Die soziale Herkunft von Schülerinnen und Schülern wird in der Bildungsforschung häufig anhand der beruflichen Qualifikation und der beruflichen Tätigkeit der Eltern erfasst. In nationalen wie internationalen Schulleistungsstudien werden dafür die Angaben zum Beruf der Eltern auf der Grundlage der *International Standard Classification of Occupation* (ISCO) kategorisiert und zur Bildung weiterer Indizes verwendet. Bei PISA 2009 und im Ländervergleich Sprachen 2009 erfolgte die Einordnung der elterlichen Berufe auf Grundlage des sogenannten ISCO-88, eines Index aus dem Jahr 1988 (vgl. International Labour Office, 1990). Im vorliegenden Ländervergleich kam eine aktualisierte Fassung dieses Index aus dem Jahr 2008 zum Einsatz (ISCO-08), der auch die Klassifikation von inzwischen neu entstandenen Berufsgruppen (z. B. im IT-Bereich) ermöglicht (International Labour Office, 2012). Auf Grundlage des ISCO lassen sich zwei Indikatoren des sozialen Status der Eltern der getesteten Schülerinnen und Schüler erstellen: der sozioökonomische Index (International Socio-Economic Index, ISEI; Ganzeboom, de Graaf, Treiman & de Leeuw, 1992) und die EGP-Klassifikation⁵ (Erikson, Goldthorpe & Portocarero, 1979). Beide sollen im Folgenden kurz beschrieben werden.

Der sozioökonomische Index ISEI (Ganzeboom et al., 1992) stellt ein Maß für den Status der beruflichen Tätigkeit der Eltern unter Berücksichtigung des Einkommens und des Bildungsniveaus dar. Mit dem ISEI wird eine Skala gebildet, die die Berufe hierarchisch ordnet und den sozioökonomischen Status repräsentiert. Die ISEI-Skala basierend auf dem ISCO-08 kann Werte zwischen 10 und 89 Punkten annehmen. Geringere Werte stehen dabei für Berufe mit einem niedrigen sozioökonomischen Status (z. B. Kellner: 28 Punkte), während hohe Werte einen hohen sozioökonomischen Status indizieren (z. B. Arzt: 89 Punkte). In den Analysen dieses Kapitels wird zur Beschreibung des sozioökonomischen Status der Familie der höchste ISEI der Eltern herangezogen, der als HISEI bezeichnet wird (*Highest ISEI*).

Der HISEI wird häufig zur Bestimmung des *sozialen Gradienten* verwendet, der angibt, wie stark der sozioökonomische Status der Eltern mit den erreichten Kompetenzen der Schülerinnen und Schüler zusammenhängt. Er beschreibt mittels einer statistischen Regressionsgeraden die Enge des Zusammenhangs zwischen dem HISEI und den Ergebnissen der Kompetenzmessung. Der soziale Gradient wird in diesem Ländervergleich für jeden der Kompetenzbereiche

4 Großstädte mit einer Einwohnerzahl von mehr als 300 000 sind Bielefeld, Bochum, Bonn, Dortmund, Dresden, Duisburg, Düsseldorf, Essen, Frankfurt am Main, Hannover, Köln, Leipzig, Mannheim, München, Nürnberg, Stuttgart und Wuppertal (Gebietsstand 31.12.2010, Quelle: Statistisches Bundesamt).

5 Die Abkürzung besteht aus den ersten Buchstaben der Nachnamen der Urheber des Index, Erikson, Goldthorpe und Portocarero.

Lesen, Zuhören und Mathematik berechnet. Für die Analysen wird der HISEI so transformiert, dass er einen Mittelwert von 0 und eine Standardabweichung von 1 aufweist. Durch diese Transformation lässt sich an der Steigung des sozialen Gradienten (b) direkt ablesen, wie viele Kompetenzpunkte eine Schülerin oder ein Schüler durchschnittlich mehr erzielen würde, wenn der HISEI der Eltern um eine Standardabweichung höher läge.

Der zweite Indikator der sozialen Herkunft, die EGP-Klassifikation, stellt die verschiedenen Merkmale der beruflichen Tätigkeiten heraus (Erikson & Goldthorpe, 2002; Erikson et al., 1979). Die EGP-Klassifikation teilt die Berufe nach der Art der Tätigkeit (manuell, nicht manuell, landwirtschaftlich), der Stellung im Beruf (selbstständig, abhängig beschäftigt), der Weisungsbefugnis (keine, geringe, große) und den erforderlichen Qualifikationen (keine, niedrige, hohe) ein. Anhand dieser Merkmale lassen sich nach dem Modell von Erikson und Kollegen (1979) insgesamt sieben gesellschaftliche Schichten von der „obersten Dienstklasse“ (Klasse I) bis zu „un- und angelernten Arbeitern“ (Klasse VII) unterscheiden. Für die Analysen in diesem Ländervergleich wurde, analog zum Vorgehen in PISA-E 2000, PISA-E 2006 und im IQB-Ländervergleich Sprachen 2009, jedem Kind zunächst die EGP-Klasse zugeordnet, die dem höchsten EGP-Wert der Eltern entspricht (vgl. Baumert & Schümer, 2002; Ehmke & Baumert, 2008; Knigge & Leucht, 2010). Anschließend wurden, ebenfalls wie in den aufgeführten Studien, zwei Extremgruppen gebildet, indem zum einen die Klassen I und II und zum anderen die Klassen V bis VII zusammengefasst wurden. Die EGP-Klassen I und II umfassen beispielsweise Schülerinnen und Schüler, deren Eltern im mittleren Management, als Beamte im mittleren, gehobenen und höheren Dienst sowie als Hochschul- und Gymnasiallehrkräfte beschäftigt sind. Der Gruppe mit den EGP-Klassen V bis VII sind unter anderem die Kinder von Vorarbeitern, Meistern und Technikern, die in manuelle Arbeitsprozesse eingebunden sind, sowie Kinder von un- und angelernten Arbeiterinnen und Arbeitern zugehörig. Die Bildung dieser EGP-Extremgruppen macht es möglich, Kompetenzstände von Kindern zu vergleichen, deren Eltern hoch beziehungsweise niedrig qualifizierten Berufsgruppen zugeordnet sind.

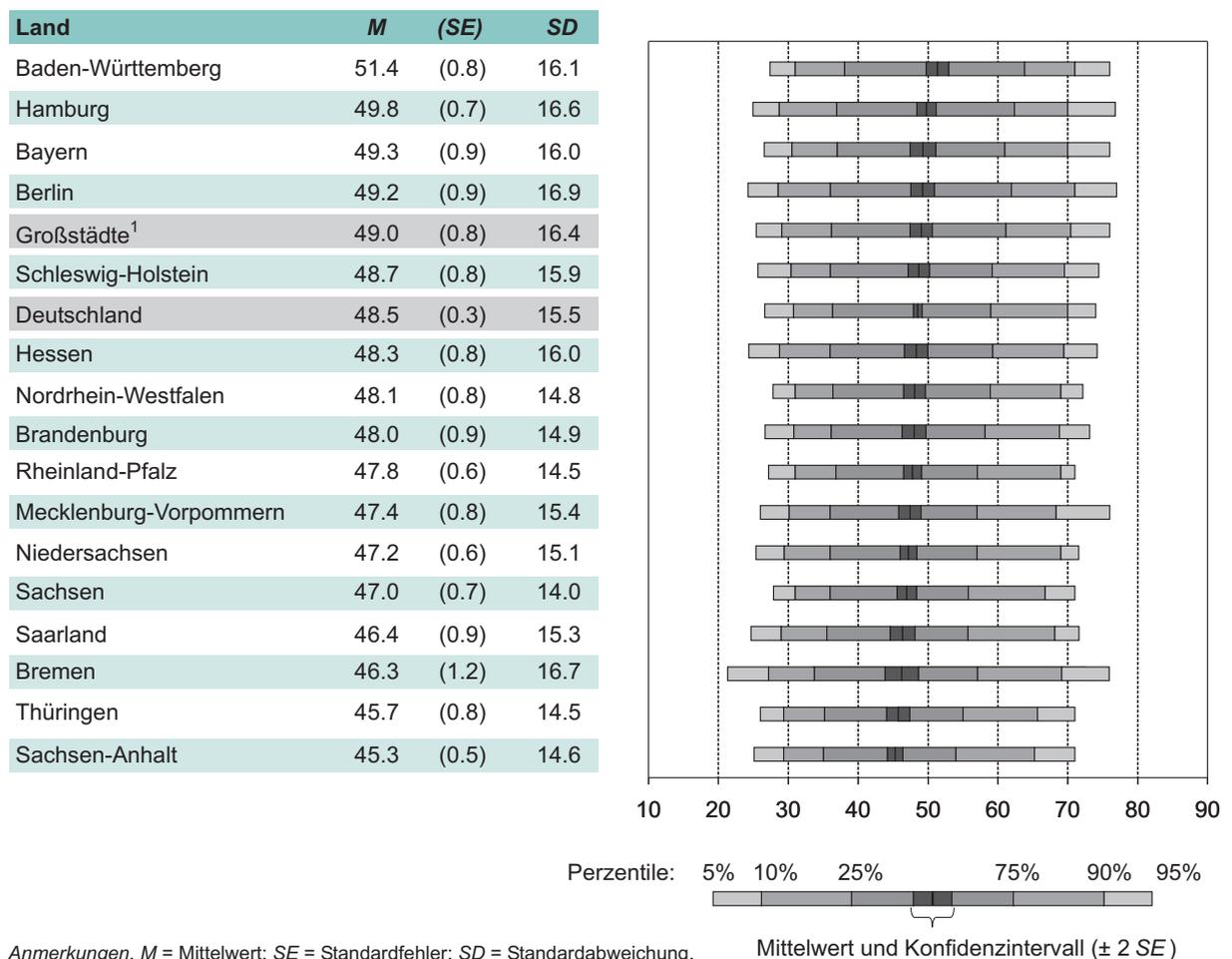
8.3 Länderdifferenzen im sozioökonomischen Status

In Abbildung 8.1 wird die Verteilung des sozioökonomischen Status der Eltern aller teilnehmenden Schülerinnen und Schüler anhand des Mittelwerts (M) und der Streuung (SD) des HISEI wiedergegeben. Die mittleren HISEI-Werte unterscheiden sich zwischen den Ländern nur geringfügig und bewegen sich im Bereich von 45.3 in Sachsen-Anhalt und 51.4 in Baden-Württemberg. Auch die Streuung (SD) der HISEI-Werte, die ein Maß für die soziale Heterogenität in einem Land darstellt, weist auf nur geringe Unterschiede von Land zu Land hin. Die Darstellung der Ergebnisse in Form von Perzentilbändern im rechten Teil der Abbildung 8.1 erlaubt darüber hinaus weitere Aussagen zur Verteilung und Spannweite der HISEI-Werte der Länder. Die Perzentile geben an, welcher Anteil der Schülerinnen und Schüler in bestimmten Bereichen des HISEI liegt. Beim 5. Perzentil handelt es sich beispielsweise um denjenigen HISEI-Wert, unter dem die fünf Prozent der Schülerschaft aus den sozial schwächsten Familien liegen. Das 95. Perzentil gibt den Wert des HISEI an, oberhalb dessen 5 Prozent aller Schülerinnen und Schüler zu verorten sind. Die Differenz zwischen dem 5. und 95. Perzentil der HISEI-Verteilung ist ein Maß für den Grad der sozialen

Heterogenität in den einzelnen Ländern beziehungsweise in Deutschland insgesamt. Diese Spanne gibt Auskunft darüber, in welchem Wertebereich die HISEI-Werte der mittleren 90 Prozent der Schülerinnen und Schüler liegen.

Die schmalste und somit homogenste Verteilung des HISEI findet sich in Sachsen mit einer Spannweite von 43 HISEI-Punkten zwischen dem 5. und dem 95. Perzentil. Für Sachsen zeigt sich also, dass sich die Schülerschaft hinsichtlich der sozialen Herkunft vergleichsweise wenig unterscheidet. Die größte Spannweite hinsichtlich des HISEI weist das Land Bremen mit einer Differenz von 55 HISEI-Punkten zwischen dem 5. und 95. Perzentil auf. In Bremen unterscheiden sich Schülerinnen und Schüler hinsichtlich ihrer sozialen Herkunft also am deutlichsten, aber auch in den beiden anderen Stadtstaaten, Berlin und Hamburg, ist jeweils ein deutliches soziales Gefälle zu konstatieren. Eine sehr heterogene Schülerschaft hinsichtlich der sozialen Lage findet sich in den Stadtstaaten in ähnlicher Form auch für den Primarbereich in IGLU-E 2006 (Stubbe et al., 2008) sowie im Ländervergleich Sprachen, an dem Schülerinnen und Schüler der neunten Jahrgangsstufe im Jahr 2009 teilnahmen (Knigge & Leucht, 2010). Die übereinstimmenden Ergebnisse dieser Studien deuten darauf hin, dass sich die länderspezifische soziale Zusammensetzung der Schülerschaft, gemessen am mittleren HISEI, in Deutschland über die Jahre kaum verändert hat.

Abbildung 8.1: Kennwerte der Verteilung des sozioökonomischen Status (*Highest International Socio-Economic Index, HISEI*) in den Ländern



Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. Fehlende Werte für den HISEI wurden durch multiple Imputation ersetzt.

¹ Ohne Berlin, Bremen und Hamburg.

8.3.1 Soziale Gradienten im Fach Deutsch

In der Tabelle 8.1 sind die Ergebnisse der Analysen zum Zusammenhang zwischen dem HISEI und den erreichten Kompetenzen im Fach Deutsch in den Bereichen Lesen und Zuhören für jedes Land aufgeführt. Die Tabelle beinhaltet für jedes Land sowohl Angaben zum Achsenabschnitt, der die mittlere Leistung für Schülerinnen und Schüler mit einem durchschnittlichem HISEI repräsentiert (s. Abbildung 8.1), als auch die Steigung des sozialen Gradienten (b), die ein Maß für die Stärke des Zusammenhangs zwischen der sozialen Herkunft und den Kompetenzen darstellt (vgl. Abschnitt 8.2). Die zusätzlich berichtete Varianzaufklärung (R^2) gibt an, welcher Anteil der Leistungsunterschiede in Prozent durch den HISEI als Maß für den sozioökonomischen Status erklärt werden kann, das heißt, wie wichtig die soziale Herkunft für die Erklärung von Kompetenzunterschieden von Schülerinnen und Schülern in einem Land ist. Bei der Interpretation muss berücksichtigt werden, dass die Ergebnisse dieser Analysen ausschließlich den Zusammenhang von Kompetenzausprägung und sozialem Status betrachten, mögliche weitere Einflussfaktoren auf die Leistungsergebnisse also außer Acht lassen und somit nicht kausal interpretiert werden können.

Im Lesen wie auch im Zuhören finden sich in allen Ländern statistisch signifikante Zusammenhänge zwischen dem HISEI und den erreichten Kompetenzen. Die Kompetenzen von Schülerinnen und Schülern variieren also in Abhängigkeit von der sozialen Herkunft, wobei sich für Deutschland insgesamt sowie für alle Länder zeigt, dass mit höheren HISEI-Werten auch höhere Kompetenzwerte einhergehen. Für Deutschland insgesamt liegt die Steigung des sozialen Gradienten bei 36 Punkten im Lesen und bei 37 Punkten im Zuhören. Durchschnittlich weichen also die Kompetenzen von Schülerinnen und Schülern, die sich hinsichtlich des HISEI um eine Standardabweichung unterscheiden, im Lesen um 36 und im Zuhören um 37 Punkte ab. Hinsichtlich der Varianzaufklärung wird deutlich, dass bundesweit etwa 13 Prozent der Unterschiede in der erzielten Lesekompetenz und knapp 14 Prozent in der Zuhörkompetenz durch den sozioökonomischen Status der Eltern erklärt werden können. Der sozioökonomische Status spielt somit bei der Erklärung der Kompetenzunterschiede von Schülerinnen und Schülern eine substantielle Rolle.

Der soziale Gradient⁶ variiert in den einzelnen Ländern und den Großstädten im Lesen zwischen 29 und 46 Punkten und im Zuhören zwischen 33 und 48 Punkten. Der höchste Sozialgradient findet sich im Lesen mit 46 Punkten für Schülerinnen und Schüler in Großstädten. Dort lässt sich in etwa ein Fünftel der Unterschiede in den erreichten Kompetenzen durch den sozialen Status der Eltern erklären. Vergleichsweise geringe soziale Disparitäten im Lesen finden sich mit einem sozialen Gradienten von 29 Punkten in Sachsen, wobei dieser Wert als einziger statistisch signifikant vom bundesweiten Wert ($b = 36$) abweicht. In Sachsen ist somit die Kopplung zwischen der sozialen Herkunft und den erreichten Lesekompetenzen deutlich niedriger als in Deutschland insgesamt. Im Kompetenzbereich Zuhören weisen die drei Stadtstaaten Berlin, Bremen und Hamburg und die Großstädte mit über 40 Punkten jeweils einen besonders hohen sozialen Gradienten auf, der für die drei Stadtstaaten signifikant über dem Bundeswert liegt. Dieser Befund deutet darauf hin, dass die Kompetenzen im

6 Die Bezeichnung „sozialer Gradient“ wird im Folgenden synonym mit der Steigung des sozialen Gradienten verwendet.

Tabelle 8.1: Soziale Gradienten für Lesen und Zuhören im Fach Deutsch nach Land (absteigend geordnet nach dem sozialen Gradienten)

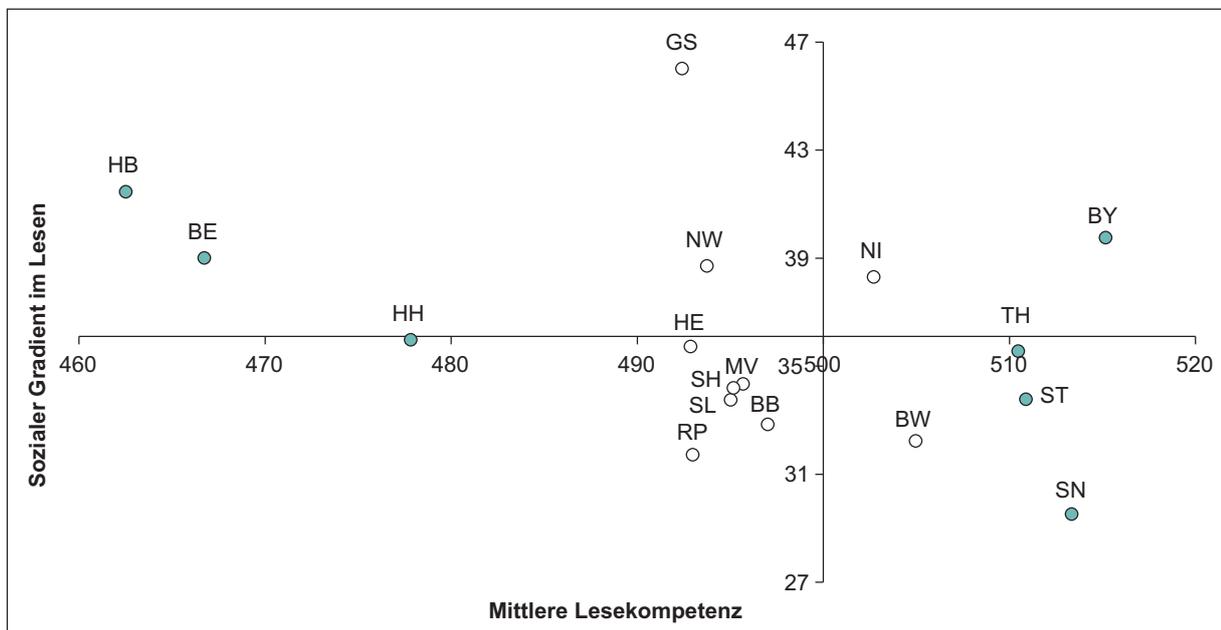
Land	Achsenabschnitt		Steigung des sozialen Gradienten		Varianzaufklärung R^2
		(SE)	b	(SE)	
Deutsch Lesen					
Großstädte ¹	492	(7.5)	46	(5.8)	20.5
Bremen	469	(4.1)	41	(3.7)	19.5
Bayern	513	(4.2)	40	(2.9)	15.7
Berlin	465	(4.4)	39	(2.7)	16.2
Nordrhein-Westfalen	495	(4.3)	39	(3.1)	14.6
Niedersachsen	506	(5.0)	38	(4.0)	13.3
Hamburg	475	(3.9)	36	(2.8)	14.4
Hessen	494	(4.1)	36	(4.5)	12.9
Thüringen	517	(3.7)	36	(2.9)	12.9
Mecklenburg-Vorpommern	498	(3.4)	34	(2.8)	11.1
Schleswig-Holstein	495	(4.7)	34	(3.3)	12.5
Sachsen-Anhalt	518	(4.3)	34	(4.1)	9.5
Saarland	500	(4.2)	34	(8.8)	10.6
Brandenburg	498	(3.4)	33	(3.6)	10.3
Baden-Württemberg	499	(4.2)	32	(2.6)	12.8
Rheinland-Pfalz	494	(5.2)	32	(3.4)	8.9
Sachsen	516	(4.4)	29	(3.0)	7.2
Deutschland	500	(1.4)	36	(1.2)	13.1
Deutsch Zuhören					
Berlin	470	(4.6)	48	(3.1)	21.6
Großstädte ¹	489	(7.2)	46	(5.8)	21.3
Bremen	474	(5.0)	46	(4.1)	20.5
Hamburg	483	(4.1)	45	(3.0)	19.6
Niedersachsen	510	(4.7)	40	(3.6)	14.9
Hessen	493	(5.1)	38	(4.2)	13.8
Bayern	511	(5.1)	38	(2.9)	15.1
Saarland	489	(4.4)	37	(7.5)	13.0
Nordrhein-Westfalen	498	(5.2)	36	(3.2)	13.0
Sachsen-Anhalt	495	(4.2)	36	(4.1)	10.2
Baden-Württemberg	502	(4.3)	36	(3.1)	15.4
Mecklenburg-Vorpommern	497	(3.7)	36	(3.1)	11.5
Thüringen	507	(3.6)	35	(3.1)	13.4
Sachsen	503	(4.4)	34	(3.3)	9.2
Rheinland-Pfalz	489	(4.6)	33	(3.5)	9.9
Schleswig-Holstein	507	(4.7)	33	(3.6)	12.0
Brandenburg	492	(4.3)	33	(3.1)	10.2
Deutschland	500	(1.6)	37	(1.2)	13.9

Anmerkungen. Die Steigung des sozialen Gradienten ist im Lesen und im Zuhören für jedes Land, die Großstädte und für Deutschland insgesamt signifikant von 0 verschieden. Fettgedruckte b unterscheiden sich signifikant ($p < .05$) vom mittleren b in Deutschland. Fehlende Werte für den HISEI wurden durch multiple Imputation ersetzt. b = unstandardisierter Regressionskoeffizient; SE = Standardfehler; R^2 = Determinationskoeffizient. ¹ Ohne Berlin, Bremen und Hamburg.

Zuhören in diesen Ländern in relativ starkem Maße an die soziale Herkunft der Schülerinnen und Schüler geknüpft sind.

Nachdem der Zusammenhang zwischen dem sozioökonomischen Status der Eltern und den erzielten Kompetenzen der Viertklässlerinnen und Viertklässler innerhalb der Länder und für Deutschland insgesamt auf Individualebene betrachtet wurde, soll dieser Zusammenhang nun aggregiert auf Länderebene untersucht werden. Mit dieser Aggregation lässt sich feststellen, inwiefern die durchschnittlich erreichten Kompetenzen in den Ländern mit dem Ausmaß sozialer Disparitäten einhergehen. Die Ergebnisse sind pro Kompetenzbereich in einem Koordinatensystem abgetragen, in dem die X-Achse das Kompetenzniveau und die Y-Achse den sozialen Gradienten in den Ländern repräsentiert, wobei sich die Achsen im jeweiligen Wert für Deutschland insgesamt schneiden. In Abbildung 8.2 ist der Zusammenhang für die Lesekompetenz dargestellt. Im oberen linken Quadranten sind mit Berlin, Bremen, Nordrhein-Westfalen und den Großstädten die Länder beziehungsweise Regionen verortet, für die sich bei vergleichsweise niedrigen Kompetenzständen zum Teil relativ ausgeprägte soziale Gradienten zeigen. Im unteren rechten Quadranten befinden sich hingegen die Länder, die bei eher niedriger Kopplung zwischen sozialer Herkunft und Testleistungen Kompetenzstände erreichen, die über dem deutschen Mittelwert liegen. Wie bereits bei den Darstellungen zum sozialen Gradienten deutlich wurde, gelingt es insbesondere im Land Sachsen, überdurchschnittliche Lesekompetenzen zu erzielen, ohne dass die Leistungen in besonders starkem Maße an den sozioökonomischen Status der Eltern geknüpft sind.

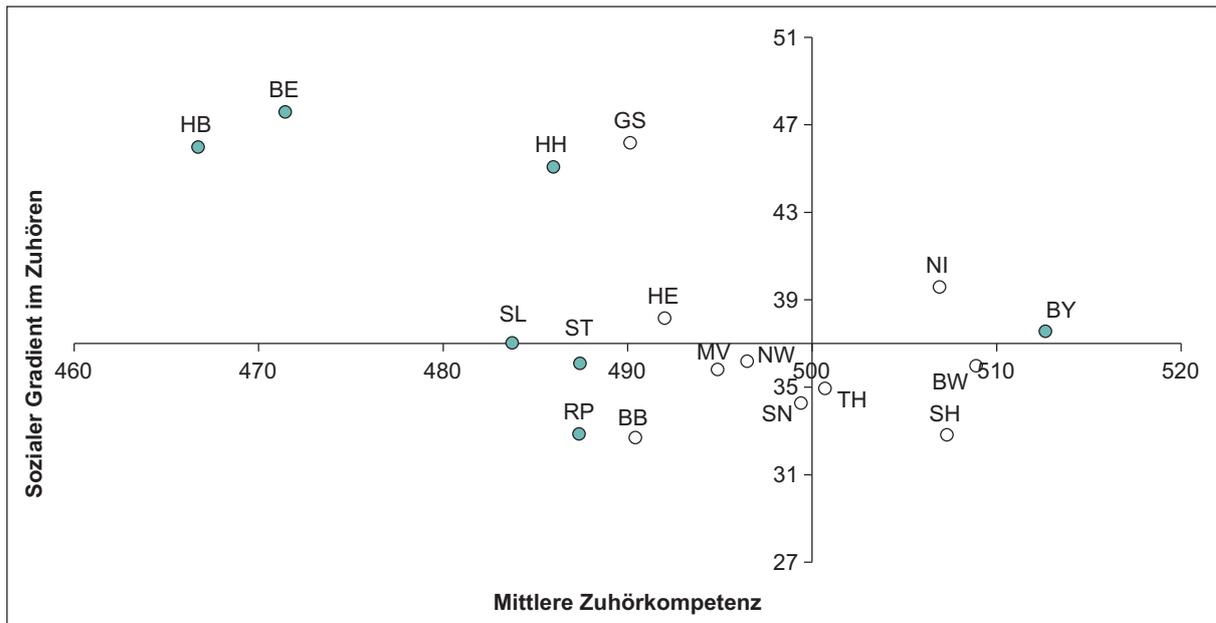
Abbildung 8.2: Zusammenhang zwischen dem sozialen Gradienten und dem mittleren Kompetenzstand im Lesen auf Länderebene



Anmerkungen. Länder mit ausgefüllten Markierungen unterscheiden sich in der Lesekompetenz signifikant vom deutschen Mittelwert. BB=Brandenburg, BE=Berlin, BW=Baden-Württemberg, BY=Bayern, GS=Großstädte, HB=Bremen, HE=Hessen, HH=Hamburg, MV=Mecklenburg-Vorpommern, NI=Niedersachsen, NW=Nordrhein-Westfalen, RP=Rheinland-Pfalz, SH=Schleswig-Holstein, SL=Saarland, SN=Sachsen, ST=Sachsen-Anhalt, TH=Thüringen.

Wie in Abbildung 8.3 zu erkennen ist, finden sich für alle Stadtstaaten überdurchschnittlich hohe Sozialgradienten bei gleichzeitig unterdurchschnittlichen Kompetenzständen im Zuhören. Für die Großstädte zeigt sich ein ähnlich hoher Wert für den Sozialgradienten, dieser lässt sich jedoch nicht zufallskritisch absichern. Alle anderen Länder weisen trotz sehr unterschiedlicher mittlerer Kompetenzstände im Zuhören nur geringfügige und nicht statistisch signifikante Abweichungen vom sozialen Gradienten in Deutschland auf.

Abbildung 8.3: Zusammenhang zwischen dem sozialen Gradienten und dem mittleren Kompetenzstand im Zuhören auf Länderebene



Anmerkungen. Länder mit ausgefüllten Markierungen unterscheiden sich in der Zuhörkompetenz signifikant vom deutschen Mittelwert. BB=Brandenburg, BE=Berlin, BW=Baden-Württemberg, BY=Bayern, GS=Großstädte, HB=Bremen, HE=Hessen, HH=Hamburg, MV=Mecklenburg-Vorpommern, NI=Niedersachsen, NW=Nordrhein-Westfalen, RP=Rheinland-Pfalz, SH=Schleswig-Holstein, SL=Saarland, SN=Sachsen, ST=Sachsen-Anhalt, TH=Thüringen.

8.3.2 Soziale Gradienten im Fach Mathematik

Der soziale Gradient in der bundesweiten Stichprobe liegt in Mathematik mit 35 Punkten auf einem ähnlichen Niveau wie im Lesen und im Zuhören. Grundschulkindern aus Familien mit einem niedrigen sozioökonomischen Status erzielen in jedem Land niedrigere Kompetenzwerte als Kinder aus Familien mit einem hohen sozioökonomischen Status. Die Spannweite der länderspezifischen Gradienten reicht von 28 bis zu 43 Punkten und ist damit ähnlich groß wie im Lesen und Zuhören. In Mathematik zeigen sich für die drei Stadtstaaten und die Großstädte erneut die höchsten sozialen Gradienten, jedoch weicht nur Bremen mit 43 Punkten statistisch signifikant vom bundesdeutschen Wert ab. Somit zeigt sich auch für Mathematik, dass in Bremen die Kompetenzen der Schülerinnen und Schüler in besonders hohem Maße an die soziale Herkunft der Familien gekoppelt sind.

In Abbildung 8.4 wird für alle 16 Länder der Zusammenhang zwischen dem sozialen Gradienten und den mittleren Kompetenzständen in Mathematik dargestellt. Auch in dieser Grafik schneiden sich die Achsen bei den bundeswei-

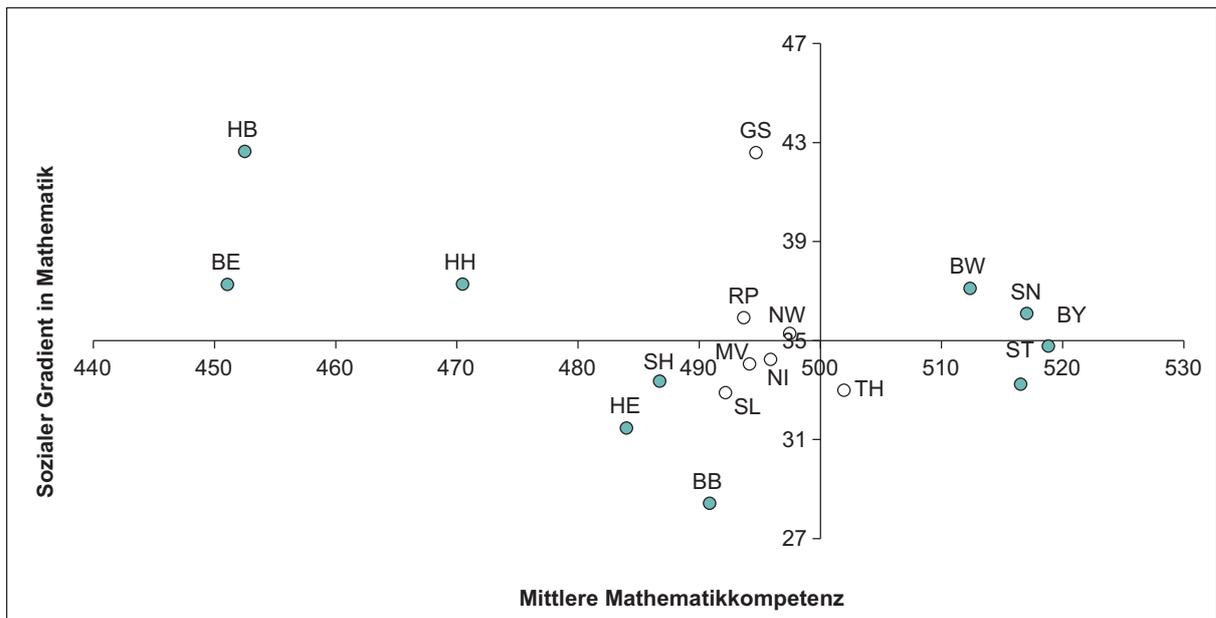
Tabelle 8.2: Soziale Gradienten für Mathematik pro Land (absteigend geordnet nach dem sozialen Gradienten)

Land	Achsenabschnitt		Steigung des sozialen Gradienten		Varianzaufklärung
		(SE)	<i>b</i>	(SE)	<i>R</i> ²
Bremen	458	(5.1)	43	(3.6)	18.8
Großstädte ¹	494	(6.9)	43	(7.0)	17.5
Hamburg	467	(3.9)	37	(2.9)	15.4
Berlin	449	(3.8)	37	(2.8)	15.3
Baden-Württemberg	505	(4.2)	37	(2.7)	15.2
Sachsen	521	(3.7)	36	(3.0)	10.7
Rheinland-Pfalz	495	(4.6)	36	(3.5)	11.4
Nordrhein-Westfalen	498	(4.1)	35	(3.4)	12.4
Bayern	517	(4.3)	35	(3.3)	13.3
Niedersachsen	499	(4.3)	34	(3.4)	11.6
Mecklenburg-Vorpommern	497	(4.5)	34	(3.4)	10.1
Schleswig-Holstein	486	(3.7)	33	(2.8)	13.0
Sachsen-Anhalt	523	(4.2)	33	(3.5)	8.7
Thüringen	508	(5.0)	33	(3.6)	10.1
Saarland	496	(4.8)	33	(10.3)	9.5
Hessen	484	(3.9)	31	(6.0)	10.1
Brandenburg	492	(3.2)	28	(3.7)	7.8
Deutschland	500	(1.4)	35	(1.4)	12.2

Anmerkungen. Die Steigung des sozialen Gradienten ist in Mathematik für jedes Land, die Großstädte und für Deutschland insgesamt signifikant von 0 verschieden. Fettgedruckte *b* unterscheiden sich signifikant ($p < .05$) vom mittleren *b* in Deutschland. Fehlende Werte für den HISEI wurden durch multiple Imputation ersetzt. *b* = unstandardisierter Regressionskoeffizient; SE = Standardfehler; *R*² = Determinationskoeffizient. ¹ Ohne Berlin, Bremen und Hamburg.

ten Mittelwerten für die Mathematikkompetenz ($M=500$) sowie für den sozialen Gradienten der deutschen Gesamtstichprobe in Mathematik ($b=35$). Die Länder verteilen sich, wie auch in den Bereichen Lesen und Zuhören, auf alle vier Quadranten, wobei mehrere Länder eher durchschnittliche oder unterdurchschnittliche Kompetenzen bei durchschnittlichen oder vergleichsweise niedrigen sozialen Disparitäten aufweisen (unterer linker Quadrant). Alle Stadtstaaten finden sich, wie schon beim Zuhören und zum Teil auch beim Lesen, im oberen linken Quadranten, jedoch zeichnet sich ausschließlich Bremen durch statistisch signifikant überdurchschnittliche soziale Disparitäten bei gleichzeitig signifikant unterdurchschnittlichen Mathematikkompetenzen aus. Darüber hinaus zeigt sich, dass die Länder, die in der Mathematikkompetenz statistisch signifikant über dem deutschen Mittelwert liegen (Bayern, Sachsen, Sachsen-Anhalt und Baden-Württemberg), hinsichtlich des sozialen Gradienten durchschnittliche Werte aufweisen. Die in Mathematik besonders erfolgreichen Länder sind demnach weder durch eine besonders starke noch eine besonders geringe Kopplung der Leistung an die soziale Herkunft gekennzeichnet.

Abbildung 8.4: Zusammenhang zwischen dem sozialen Gradienten und dem mittleren Kompetenzstand in Mathematik auf Länderebene



Anmerkungen. Länder mit ausgefüllten Markierungen unterscheiden sich in der Mathematikkompetenz signifikant vom deutschen Mittelwert. BB=Brandenburg, BE=Berlin, BW=Baden-Württemberg, BY=Bayern, GS=Großstädte, HB=Bremen, HE=Hessen, HH=Hamburg, MV=Mecklenburg-Vorpommern, NI=Niedersachsen, NW=Nordrhein-Westfalen, RP=Rheinland-Pfalz, SH=Schleswig-Holstein, SL=Saarland, SN=Sachsen, ST=Sachsen-Anhalt, TH=Thüringen.

8.4 Kompetenzniveaus nach EGP-Klassen

In diesem Abschnitt wird mit den EGP-Klassen ein weiterer Indikator für die soziale Herkunft herangezogen und mit den erzielten Kompetenzen in Zusammenhang gebracht. Dabei werden nicht alle EGP-Klassen miteinander verglichen, sondern zwei Extremgruppen sozialer Herkunft gebildet und kontrastiert (vgl. Abschnitt 8.2).

In Tabelle 8.3 werden die durchschnittlichen Kompetenzwerte für die Schülerinnen und Schüler beider EGP-Extremgruppen sowie die sich daraus ergebenden Kompetenzunterschiede berichtet. Im bundesweiten Vergleich zeigt sich zwischen den Schülergruppen niedriger und hoher sozialer Herkunft eine durchschnittliche Differenz von 81 Punkten im Lesen und 86 Punkten im Zuhören. Bezogen auf durchschnittlich in einem Jahr erreichte Kompetenzzuwächse im Lesen und im Zuhören entspricht dies einem Vorsprung der Schülerinnen und Schüler aus den EGP-Klassen I und II im Vergleich zu den Kindern aus den Klassen V bis VII von mehr als einem Schuljahr (Lesen: Bremerich-Vos & Böhme, 2009; Zuhören: Behrens, Böhme & Krelle, 2009). In ländervergleichender Perspektive zeigen sich starke Unterschiede im Kompetenzgefälle zwischen den beiden EGP-Gruppen. Sie reichen im Lesen von 52 bis zu 106 Punkten und im Zuhören von 60 bis zu 116 Punkten. Hohe, aber statistisch nicht signifikant vom deutschen Wert abweichende Differenzen zwischen den EGP-Extremgruppen finden sich im Lesen für die Großstädte und Bayern. Nur in Sachsen unterscheidet sich die Differenz der beiden Extremgruppen hinsichtlich der Lesekompetenz in statistisch signifikantem Maße vom bundesdeutschen Wert. Mit 52 Punkten Differenz unterscheiden sich die Lesekompetenzen von Schülerinnen und Schülern aus beiden Gruppen hier deutlich weniger stark

Tabelle 8.3: Durchschnittliche Kompetenzwerte der EGP¹-Extremgruppen und deren Differenz für Lesen und Zuhören im Fach Deutsch pro Land (absteigend geordnet nach der Differenz der Kompetenzmittelwerte)

Land	EGP-Klasse I-II		EGP-Klasse V-VII		Differenz	
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)	ΔM	(<i>SE</i>)
Deutsch Lesen						
Großstädte ²	534	(8.5)	428	(14.3)	106	(15.6)
Bayern	544	(5.1)	446	(8.6)	98	(9.4)
Bremen	504	(6.6)	414	(7.2)	91	(9.3)
Hessen	523	(5.0)	434	(10.9)	90	(11.9)
Hamburg	510	(4.2)	421	(9.9)	90	(10.7)
Sachsen-Anhalt	539	(5.6)	453	(8.2)	86	(10.0)
Berlin	498	(5.5)	413	(8.5)	85	(8.7)
Saarland	527	(7.2)	444	(12.5)	84	(16.6)
Baden-Württemberg	528	(5.0)	447	(8.9)	81	(9.5)
Niedersachsen	531	(4.8)	451	(10.5)	80	(10.0)
Nordrhein-Westfalen	523	(5.3)	444	(7.8)	79	(8.2)
Schleswig-Holstein	522	(5.8)	445	(8.4)	76	(9.7)
Thüringen	539	(4.7)	467	(7.0)	72	(7.8)
Rheinland-Pfalz	516	(5.9)	449	(7.9)	67	(8.0)
Mecklenburg-Vorpommern	520	(5.0)	454	(11.0)	66	(11.4)
Brandenburg	518	(4.4)	455	(10.4)	63	(11.2)
Sachsen	531	(5.8)	479	(10.7)	52	(11.3)
Deutschland	527	(1.8)	446	(3.2)	81	(3.5)
Deutsch Zuhören						
Hamburg	526	(4.8)	410	(9.5)	116	(11.0)
Berlin	510	(6.2)	401	(9.5)	109	(9.9)
Bremen	514	(7.6)	409	(7.9)	105	(10.3)
Großstädte ²	531	(8.5)	428	(14.1)	103	(14.9)
Hessen	526	(5.3)	430	(11.4)	96	(11.6)
Baden-Württemberg	535	(5.2)	440	(10.3)	96	(11.2)
Saarland	519	(6.6)	425	(11.5)	94	(14.4)
Bayern	540	(5.4)	448	(9.6)	92	(9.0)
Niedersachsen	537	(4.9)	447	(10.2)	91	(9.5)
Sachsen-Anhalt	515	(5.1)	427	(11.0)	88	(12.2)
Schleswig-Holstein	533	(6.3)	456	(9.5)	77	(10.7)
Nordrhein-Westfalen	525	(6.1)	449	(8.8)	77	(8.7)
Rheinland-Pfalz	511	(5.6)	436	(8.2)	75	(8.8)
Mecklenburg-Vorpommern	521	(5.1)	453	(10.5)	68	(11.2)
Thüringen	527	(4.9)	460	(7.1)	67	(7.7)
Brandenburg	511	(5.2)	449	(10.2)	62	(10.3)
Sachsen	521	(6.3)	461	(10.1)	60	(11.4)
Deutschland	529	(2.0)	443	(3.7)	86	(3.8)

Anmerkungen. Der Differenzwert zwischen den EGP-Extremgruppen (ΔM) ist im Lesen und im Zuhören für jedes Land, die Großstädte und für Deutschland insgesamt signifikant von 0 verschieden. Fettgedruckte Differenzwerte (ΔM) unterscheiden sich signifikant ($p < .05$) von der mittleren Differenz in Deutschland. Fehlende Werte für die EGP-Klasse wurden durch multiple Imputation ersetzt. *M* = Mittelwert; *SE* = Standardfehler. ¹ Abkürzung für Erikson, Goldthorpe und Portocarero. ² Ohne Berlin, Bremen und Hamburg.

als bundesweit. Im Zuhören sind besonders große Leistungsdifferenzen zwischen beiden EGP-Extremgruppen für die Stadtstaaten Hamburg und Berlin zu beobachten. Im Vergleich mit der bundesdeutschen Differenz zeigen sich deutlich geringere Kompetenzunterschiede im Zuhören in Brandenburg, Sachsen und Thüringen.

In Tabelle 8.4 sind die durchschnittlich erreichten Leistungen für die beiden EGP-Extremgruppen sowie deren Differenzwerte für das Fach Mathematik dargestellt. Bundesweit betrachtet entspricht die Differenz von 79 Punkten einem Leistungsunterschied von etwa einem Schuljahr (Reiss & Winkelmann, 2009). Die länderspezifischen Differenzwerte bewegen sich im Bereich von 58 bis 101 Punkten, jedoch unterscheidet sich keiner dieser Werte in statistisch bedeutsamer Weise vom deutschen Differenzwert.

Tabelle 8.4: Durchschnittliche Kompetenzwerte der EGP¹-Extremgruppen und deren Differenz für Mathematik pro Land (absteigend geordnet nach der Differenz der Kompetenzmittelwerte)

Land	EGP-Klasse I-II		EGP-Klasse V-VII		Differenz	
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)	ΔM	(<i>SE</i>)
Großstädte ²	536	(8.6)	436	(16.3)	101	(18.4)
Baden-Württemberg	540	(4.8)	445	(10.3)	96	(10.8)
Bremen	494	(6.9)	401	(8.2)	93	(9.9)
Saarland	526	(8.2)	438	(15.2)	88	(19.3)
Sachsen-Anhalt	543	(4.8)	461	(10.8)	83	(11.1)
Berlin	481	(5.4)	399	(7.8)	83	(8.4)
Niedersachsen	525	(4.2)	443	(9.7)	82	(9.4)
Bayern	544	(4.2)	463	(9.8)	81	(10.1)
Hamburg	502	(4.4)	422	(8.6)	80	(9.3)
Hessen	512	(6.3)	433	(9.1)	78	(12.6)
Rheinland-Pfalz	519	(6.1)	441	(8.1)	78	(8.4)
Nordrhein-Westfalen	525	(5.5)	452	(6.8)	73	(7.7)
Mecklenburg-Vorpommern	520	(6.1)	448	(9.8)	72	(10.8)
Schleswig-Holstein	511	(5.1)	443	(7.5)	68	(9.2)
Thüringen	526	(6.2)	462	(7.4)	64	(7.9)
Sachsen	541	(5.2)	479	(9.0)	62	(10.4)
Brandenburg	510	(4.7)	452	(9.3)	58	(10.8)
Deutschland	527	(1.8)	449	(3.3)	79	(3.6)

Anmerkungen. Der Differenzwert zwischen den EGP-Extremgruppen (ΔM) ist in Mathematik für jedes Land, die Großstädte und für Deutschland insgesamt signifikant von 0 verschieden. Es gibt jedoch keine statistisch signifikanten Unterschiede ($p < .05$) zwischen den Differenzwerten in den einzelnen Ländern und dem Differenzwert in Deutschland. Fehlende Werte für die EGP-Klasse wurden durch multiple Imputation ersetzt. *M* = Mittelwert; *SE* = Standardfehler. ¹ Abkürzung für Erikson, Goldthorpe und Portocarero. ² Ohne Berlin, Bremen und Hamburg.

8.5 Zusammenfassung und Diskussion

In diesem Kapitel wurden die sozialen Disparitäten von Schülerinnen und Schülern der vierten Jahrgangsstufe in den Ländern der Bundesrepublik Deutschland betrachtet, um Aussagen darüber treffen zu können, in welchem Ausmaß in den Ländern die von Schülerinnen und Schülern erreichten Kompetenzen an ihre soziale Herkunft gekoppelt sind. Hierfür wurden die erreichten Kompetenzen im Lesen, im Zuhören und in Mathematik in Beziehung zu Indikatoren der sozialen Herkunft gesetzt.

In allen Ländern Deutschlands besteht ein substanzieller und gleichgerichteter Zusammenhang zwischen der sozialen Herkunft und den erreichten Kompetenzen von Schülerinnen und Schülern. Je niedriger der soziale Status, desto geringer sind die gezeigten Kompetenzen. Allerdings unterscheiden sich die Länder in der Stärke dieses Zusammenhangs. Eine enge Kopplung findet sich sowohl für den sozialen Gradienten als auch für die EGP-Klassen-Vergleiche in der Tendenz in den Stadtstaaten und Großstädten. Ebenfalls auffällig ist in Bayern der relativ starke, wenn auch nicht signifikante Zusammenhang der Leseleistungen mit der sozialen Herkunft, der in diesem Ausmaß nicht für die Leistungen im Zuhören und in der Mathematik zu beobachten ist. Eine vergleichsweise geringe Kopplung der schulischen Leistung mit der sozialen Herkunft findet sich in Sachsen (Lesen und Zuhören) sowie in Brandenburg und Thüringen (Zuhören). Hier sind die Leistungsunterschiede zwischen Kindern von hoch und niedrig qualifizierten Berufsgruppen geringer als bundesweit betrachtet.

Ein ähnliches Befundmuster für die Verteilung von Ländern mit stark und weniger stark ausgeprägten sozialen Disparitäten fand sich auch bei IGLU-E 2006 (Autorengruppe Bildungsberichterstattung, 2010; Stubbe et al., 2008). Direkte Vergleiche zu den Ergebnissen von IGLU-E 2006 sind jedoch nur eingeschränkt möglich, da sich einerseits die Inhalte und Konzeptionen der eingesetzten Tests unterscheiden und andererseits verschiedene Maße zur Erfassung der sozialen Herkunft verwendet wurden (vgl. Kapitel 5 in diesem Band). Dennoch lässt sich feststellen, dass die Länder Berlin und Hamburg, die in IGLU-E 2006 noch die höchsten sozialen Gradienten im Lesen aufwiesen (Autorengruppe Bildungsberichterstattung, 2010), im Ländervergleich 2011 im Lesen und auch in Mathematik durchschnittliche soziale Gradienten hatten. Im Kompetenzbereich Zuhören hingegen zeigten sich für Berlin und Hamburg im Ländervergleich 2011 Werte für den sozialen Gradienten, die statistisch signifikant über dem deutschen Wert lagen. Im Gegensatz dazu wies Bremen in IGLU-E 2006 einen durchschnittlichen sozialen Gradienten auf, wohingegen im Ländervergleich 2011 überdurchschnittliche und statistisch signifikant vom deutschen Wert abweichende soziale Gradienten für die Kompetenzbereiche Zuhören und Mathematik beobachtet wurden. Die geringsten sozialen Gradienten im Lesen zeigten sich bei IGLU-E 2006 für die Länder Bayern, Sachsen-Anhalt und Thüringen. Im Ländervergleich 2011 wiesen diese Länder durchschnittliche Werte im sozialen Gradienten für die Kompetenzbereiche Lesen, Zuhören und Mathematik auf. Inwiefern dies als bedeutsame Veränderung interpretiert werden kann, lässt sich aufgrund der erwähnten Unterschiede zwischen den Studien nicht mit Sicherheit feststellen.

Eine Einordnung der in diesem Kapitel berichteten Ergebnisse lässt sich auch im Vergleich zu den Befunden des Ländervergleichs 2009 vornehmen, da in dieser Studie die gleichen Indizes der sozialen Herkunft für Schülerinnen und Schüler der neunten Jahrgangsstufe gebildet wurden. Da der Fokus im Ländervergleich 2009 auf den sprachlichen Kompetenzen lag, wer-

den in die vergleichende Betrachtung der Befunde des vorliegenden Ländervergleichs nur die Kompetenzbereiche Lesen und Zuhören einbezogen. Für beide Kompetenzbereiche wurde auch im Ländervergleich 2009 ein starker Zusammenhang zwischen den Testleistungen und dem sozioökonomischen Status der Eltern in Berlin festgestellt (Knigge & Leucht, 2010). Eine vergleichsweise geringe Kopplung fand sich 2009 in Brandenburg, Sachsen und in Schleswig-Holstein. Dieses Ergebnis deckt sich in der Tendenz mit dem Ergebnis aus dem Ländervergleich in der Primarstufe, was darauf hindeutet, dass sich relativ ähnliche Befundmuster für Indikatoren der sozialen Ungleichheit in einem Bildungssystem schulstufenübergreifend zeigen.

Neben der relativen Position der einzelnen Länder lässt sich auch die Ausprägung des sozialen Gradienten sowie die Differenz zwischen den EGP-Extremgruppen zwischen den Studien vergleichen. Beide betrachteten Kennwerte fallen im Ländervergleich 2011 für Lesen und Zuhören geringfügig höher aus als im Ländervergleich 2009. Erklärt werden kann dieser Unterschied unter anderem durch Spezifika der jeweils zugrunde liegenden Schülerpopulation. Während im Ländervergleich 2009 ausschließlich Schülerinnen und Schüler an Regelschulen in die Erhebung einbezogen wurden, nahmen 2011 auch Kinder aus Förderschulen teil. Diese weiter gefasste Definition der Population führt zu einer größeren Heterogenität in den gemessenen Kompetenzen und damit zu Veränderungen im sozialen Gradienten beziehungsweise in den Kompetenzunterschieden zwischen den EGP-Extremgruppen.

Das vorliegende Kapitel ergänzt die wenigen vorliegenden Untersuchungen zu den sozialen Disparitäten in der Primarstufe insbesondere durch die erstmalige ländervergleichende Betrachtung im Fach Mathematik. Im Gegensatz zur Lesekompetenz, von der angenommen wird, dass sie insbesondere durch das familiäre Anregungspotenzial und die sozialen Lebensverhältnisse geprägt wird, wird die Vermittlung mathematischer Kompetenzen in erster Linie der Schule zugeschrieben (Maaz, Baumert & Cortina, 2008). Dennoch zeigen sich im Hinblick auf die sozialen Disparitäten ähnliche Befundmuster für beide Fächer (vgl. auch TIMSS 2007, Bonsen, Frey & Bos, 2008). Ferner kann die Studie bestätigen, dass sozial bedingte Disparitäten nicht nur eine spezifische Herausforderung der drei Stadtstaaten, sondern allgemein von urbanen Ballungsräumen darstellen (vgl. Baumert & Schümer, 2002). Darüber hinaus machen die Ergebnisse erneut deutlich, dass bereits in der Grundschule erhebliche sozial bedingte Kompetenzunterschiede bestehen, die für den weiteren Bildungsweg und den Übergang auf die weiterführenden Schulen bedeutsam werden können. Daher sollte das Ziel aller Bemühungen sein, insbesondere in Ländern mit ausgeprägten Disparitäten stärker als bisher Mindeststandards zu sichern und zusätzlich den Anteil der Schülerinnen und Schüler, die die Regelstandards erreichen, zu erhöhen. Es bedarf somit auch zukünftig weiterer Anstrengungen innerhalb der Länder, um die Zusammenhänge zwischen sozialer Herkunft und Kompetenzerwerb zu entkoppeln.

Literatur

- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T. C. (2007). Schullaufbahnpräferenzen am Ende der vierten Klassenstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 271–297). Münster: Waxmann.
- Autorengruppe Bildungsberichterstattung. (2010). *Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demografischen Wandel*. Bielefeld: Bertelsmann Verlag.
- Baumert, J. & Schümer, G. (2002). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich* (S. 159–202). Opladen: Leske + Budrich.
- Baumert, J., Watermann, R. & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und individuelles Mediationsmodell. *Zeitschrift für Erziehungswissenschaft*, 6, 46–71.
- Behrens, U., Böhme, K. & Krelle, M. (2009). Zuhören – Operationalisierung und fachdidaktische Implikationen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 357–375). Weinheim: Beltz.
- Bonsen, M., Frey, K. A. & Bos, W. (2008). Soziale Herkunft. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 141–156). Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2007). *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, S. & Walther, G. (Hrsg.). (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Boudon, R. (1986). Education, mobility, and sociological theory. In J. G. Richardson (Hrsg.), *Handbook of theory and research for the sociology of education* (S. 261–274). New York: Greenwood.
- Bourdieu, P., Passeron, J.-C. & Nice, R. (1990). *Reproduction in education, society and culture*. Thousand Oaks, CA, US: Sage Publications, Inc.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 219–249). Weinheim: Beltz.
- Ehmke, T. & Baumert, J. (2008). Soziale Disparitäten des Kompetenzerwerbs und der Bildungsbeteiligung in den Ländern: Vergleiche zwischen PISA 2000 und 2006. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 319–342). Münster: Waxmann.
- Ehmke, T. & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, M. Schneider & P. Stanat (Hrsg.), *PISA 2009: Bilanz nach einem Jahrzehnt*. (S. 231–254). Münster: Waxmann.
- Erikson, R. & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives*, 16(3), 31–44.
- Erikson, R., Goldthorpe, J. H. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30, 341–415.
- Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D., J. & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1–56.
- International Labour Office. (1990). *International Standard Classification of Occupations (ISCO-88)*. Genf: ILO.
- International Labour Office. (2012). *International Standard Classification of Occupations (ISCO-08)*. Genf: ILO.

- Knigge, M. & Leucht, M. (2010). Soziale Disparitäten im Spracherwerb. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 185–201). Münster: Waxmann.
- Kramer, R.-T. & Helsper, W. (2010). Kulturelle Passung und Bildungsungleichheit – Potenziale einer an Bourdieu orientierten Analyse der Bildungsungleichheit. In H.-H. Krüger, U. Rabe-Kleberg, R.-T. Kramer & J. Budde (Hrsg.), *Bildungsungleichheit revisited. Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule* (S. 103–125). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krüger, H.-H., Rabe-Kleberg, U., Kramer, R.-T. & Budde, J. (Hrsg.). (2010). *Bildungsungleichheit revisited. Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maaz, K., Baumert, J. & Cortina, K. S. (2008). Soziale und regionale Ungleichheit im deutschen Bildungssystem. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 205–243). Reinbek bei Hamburg: Rowohlt.
- Maaz, K., Baumert, J., Gresch, C. & McElvany, N. (Hrsg.). (2010). *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Berlin: Bundesministerium für Bildung und Forschung.
- Maaz, K., Baumert, J. & Trautwein, U. (2010). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In K. Maaz, J. Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 27–63). Berlin: Bundesministerium für Bildung und Forschung.
- McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41, 121–131.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 120–141). Weinheim: Beltz.
- Stubbe, T. C., Bos, W. & Hornberg, S. (2008). Soziale und kulturelle Disparitäten der Schülerleistungen in den Ländern der Bundesrepublik Deutschland. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 103–109). Münster: Waxmann.

Kapitel 9

Zuwanderungsbezogene Disparitäten

Nicole Haag, Katrin Böhme und Petra Stanat

Über die Situation von Kindern und Jugendlichen aus zugewanderten Familien an deutschen Schulen war vor der Veröffentlichung von Befunden der internationalen Schulleistungsstudien TIMSS¹ (Baumert, Bos & Lehmann, 2000), PISA² (Baumert et al., 2002) und PIRLS/IGLU³ (Bos, Lankes, Prenzel & Schwippert, 2003) wenig bekannt. Dies lag unter anderem daran, dass der Zuwanderungshintergrund von Personen in der amtlichen Statistik ausschließlich anhand der Staatsbürgerschaft registriert wurde. Daher ließ sich lange Zeit nicht einmal feststellen, wie viele Kinder und Jugendliche mit Zuwanderungshintergrund eine Schule in Deutschland besuchten. Nach den Daten der amtlichen Statistik, in denen (Spät-)Aussiedler und eingebürgerte Ausländer nicht identifiziert werden konnten, lag der Anteil ausländischer Schülerinnen und Schüler in der Sekundarstufe I im Jahr 2000 bei etwa 8 Prozent (Statistisches Bundesamt, 2001). Im Bericht über die Ergebnisse von PISA 2000 zeigte sich jedoch, dass im selben Jahr der Anteil von 15-Jährigen mit mindestens einem im Ausland geborenen Elternteil 22 Prozent betrug und somit deutlich höher war (Baumert & Schümer, 2001). Die Zahl der Kinder und Jugendlichen mit Zuwanderungshintergrund, die in Deutschland eine Schule besuchen, ist in den folgenden Jahren weiter gestiegen und lag in der Gruppe der 15-Jährigen im Jahr 2009 bei knapp 26 Prozent (Stanat, Rauch & Segeritz, 2010). Für die Primarstufe ergaben die Befunde aus IGLU 2001 (Schwippert, Bos & Lankes, 2003), IGLU-E 2006 (Schwippert, Hornberg & Goy, 2008) und TIMSS 2007 (Bonsen, Kummer & Bos, 2008) ähnliche beziehungsweise zum Teil noch etwas höhere Anteile von Kindern mit mindestens einem im Ausland geborenen Elternteil von 22 Prozent, 26 Prozent und 29 Prozent.

Neben dem Fehlen relevanter demografischer Daten lag ein zweiter Grund dafür, dass die Bildungssituation von Heranwachsenden aus zugewanderten Familien nicht angemessen untersucht werden konnte, darin, dass in Deutschland lange Zeit darauf verzichtet wurde, die Erträge schulischer Bildungsprozesse mit Hilfe von geeigneten Kompetenztests zu bestimmen (vgl. Kapitel 1). Als Indikatoren für Bildungserfolg ließen sich daher höchstens (und auch dies nur in sehr eingeschränktem Maße) Merkmale von Bildungslaufbahnen, wie etwa die Muster der Bildungsbeteiligung oder Bildungsabschlüsse, untersuchen. So ließ sich zwar zum Beispiel anhand von Daten des Sozio-oekonomischen Panels (vgl. Schupp, 2009) feststellen, dass Schülerinnen und Schüler aus zu-

1 Das Akronym TIMSS steht für *Trends in International Mathematics and Science Study*.

2 Das Akronym PISA steht für *Programme for International Student Assessment*.

3 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.

gewanderten Familien an Hauptschulen erheblich überrepräsentiert und an Gymnasien deutlich unterrepräsentiert waren (vgl. Alba, Handl & Müller, 1994; Diefenbach, 2002). Ob diese Disparitäten auf Unterschiede in den bis zum Übergang in die Sekundarstufe I erreichten schulischen Kompetenzen oder auf andere Faktoren, wie etwa differenzielle Übergangsempfehlungen von Lehrkräften (Gomolla & Radtke, 2002), zurückzuführen waren, konnte aufgrund des Mangels an Kompetenzdaten jedoch nicht geklärt werden. Dies änderte sich mit der Teilnahme Deutschlands an internationalen Schulleistungsstudien und den sich daran anschließenden nationalen Untersuchungen der empirischen Bildungsforschung.

In diesen Studien zeigte sich, dass schon in der Grundschule teils erhebliche Nachteile für Kinder mit Zuwanderungshintergrund im Kompetenzerwerb festzustellen sind. Kinder aus zugewanderten Familien erzielten allgemein deutlich geringere Kompetenzniveaus in den Bereichen Lesen, Mathematik und Naturwissenschaften als Gleichaltrige ohne Zuwanderungshintergrund (Bonsen et al., 2008; Schwippert et al., 2003). Dieser Unterschied war in Deutschland deutlich größer als in den meisten anderen Ländern. Der in IGLU 2006 beobachtete Kompetenzrückstand im Lesen betrug für Schülerinnen und Schüler, deren Elternteile beide im Ausland geboren sind, im Mittel 48 Punkte (Schwippert, Hornberg, Freiberg & Stubbe, 2007), was dem im Rahmen derselben Studie geschätzten Leistungszuwachs von einem Schuljahr entspricht. Ein ähnlich großer Kompetenzrückstand von 46 Punkten zeigte sich in TIMSS 2007 auch für Mathematik (Bonsen et al., 2008).

Die mit dem Zuwanderungshintergrund verbundenen Disparitäten im Bildungserfolg lassen sich zu einem großen Teil auf Merkmale der sozialen Herkunft zurückführen. Nach Kontrolle des sozioökonomischen Status der Familie und des Bildungshintergrunds der Eltern reduzieren sich die Kompetenzunterschiede zwischen Heranwachsenden mit und Heranwachsenden ohne Zuwanderungshintergrund in der Regel erheblich (vgl. z.B. Segeritz, Walter & Stanat, 2010). Entsprechend ist davon auszugehen, dass der Entstehung zuwanderungsbezogener Disparitäten zu einem erheblichen Teil dieselben Mechanismen zugrunde liegen wie der Entstehung sozialer Ungleichheit (vgl. zusammenfassend Maaz, 2006; vgl. auch Kapitel 8). Nach Kontrolle der sozialen Herkunft bleiben jedoch in der Regel weitere zuwanderungsbezogene Disparitäten bestehen, die zum Teil auf die Umgangssprache in der Familie zurückgeführt werden können. So erreichten Jugendliche, die in der Familie eine andere Sprache als Deutsch sprechen, auch nach Kontrolle von ökonomischen und kulturellen Ressourcen der Familie (sozioökonomische Stellung, Bildungshintergrund, Kulturgüter) in PISA 2009 etwa 24 Punkte weniger auf der Lesekompetenzskala als Jugendliche mit deutscher Familiensprache (Stanat, Rauch & Segeritz, 2010). Dieser Effekt, der auch schon in der Grundschule nachweisbar ist (vgl. z.B. Kristen, 2008), kann als Indikator für einen Mangel an Lerngelegenheiten zum Kompetenzerwerb in der Verkehrs- und Instruktionssprache Deutsch interpretiert werden, der durch institutionelle Lerngelegenheiten ausgeglichen werden muss. Gestützt wird diese Annahme durch Ergebnisse von Studien mit Kleinkindern. Diese zeigen, dass Kinder aus zugewanderten Familien bereits beim Eintritt in die Kindertagesstätte deutlich geringere Kompetenzen in den Bereichen Wortschatz und Grammatik der deutschen Sprache aufweisen als Kinder ohne Zuwanderungshintergrund (vgl. z.B. Dubowy, Ebert, von Maurice & Weinert, 2008). Auch in Einschulungsuntersuchungen wird bei Kindern aus zugewanderten Familien ein im Durchschnitt geringeres sprachliches Kompetenzniveau beobachtet (vgl. z.B. Mengerling, 2005).

Die Situation von Heranwachsenden aus zugewanderten Familien ist jedoch nicht homogen, sondern unterscheidet sich zwischen verschiedenen Herkunftsgruppen teilweise erheblich. Während im IQB-Ländervergleich 2009 für alle betrachteten Herkunftsgruppen signifikante Kompetenznachteile bestanden (Böhme, Tiffin-Richards, Schipolowski & Leucht, 2010), waren in PISA 2009 für in Deutschland geborene Jugendliche, deren Eltern im Gebiet der ehemaligen Sowjetunion geboren sind (zweite Generation), keine Kompetenznachteile zu verzeichnen. In anderen Herkunftsgruppen waren die Kompetenznachteile nach Kontrolle von ökonomischen und kulturellen Ressourcen der Familie nicht mehr signifikant. Dies verweist nochmals sehr deutlich darauf, dass die Nachteile von Kindern und Jugendlichen aus zugewanderten Familien zu großen Teilen mit der sozialen Stellung der Familien verbunden sind.

Um genauer festzustellen, welche Gruppen von Schülerinnen und Schülern mit Zuwanderungshintergrund besonderer Förderung bedürfen, werden im vorliegenden Kapitel die Kompetenzstände von Kindern aus zugewanderten Familien nicht nur insgesamt, sondern auch für verschiedene Teilgruppen dargestellt. Nach einer kurzen Erläuterung, wie der Zuwanderungshintergrund der Schülerinnen und Schüler im Rahmen der verschiedenen Analysen definiert wird, werden zunächst die im Durchschnitt erreichten Kompetenzen in den Fächern Deutsch und Mathematik in Abhängigkeit vom Zuwanderungshintergrund berichtet und zwischen den Ländern verglichen. Im Anschluss werden diese Unterschiede zu Merkmalen des familiären Hintergrunds, wie beispielsweise das Herkunftsland und die Familiensprache, in Bezug gesetzt.

9.1 Erfassung des Zuwanderungshintergrunds

Im Ländervergleich 2011 wurden der Bestimmung des Zuwanderungshintergrunds Angaben über das Geburtsland der Schülerinnen und Schüler sowie der Geburtsländer ihrer Eltern zugrunde gelegt. Dabei wurden bevorzugt die Elternangaben verwendet. Lagen von den Eltern keine Informationen vor, wurden die Angaben der Schülerinnen und Schüler herangezogen. Die Klassifikation des Zuwanderungsstatus im Ländervergleich 2011 orientiert sich an der internationalen Literatur, in der im Allgemeinen zwischen verschiedenen Zuwanderungsgenerationen unterschieden wird (vgl. Stanat & Christensen, 2006; Walter & Taskinen, 2007). Insgesamt beziehen die in diesem Kapitel dargestellten Analysen die folgenden Gruppen ein:⁴

- *Schülerinnen und Schüler ohne Zuwanderungshintergrund*: beide Eltern sind in Deutschland geboren;
- *Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil*: ein Elternteil ist in Deutschland, der andere Elternteil ist im Ausland geboren;
- *Schülerinnen und Schüler der zweiten Generation*: beide Elternteile sind im Ausland geboren, die Schülerin oder der Schüler selbst ist in Deutschland geboren sowie
- *Schülerinnen und Schüler der ersten Generation*: sowohl beide Elternteile als auch die Schülerin oder der Schüler selbst sind im Ausland geboren.

4 Die Teilstichprobe der Schülerinnen und Schüler der dritten Generation (beide Elternteile sowie die Schülerin beziehungsweise der Schüler sind in Deutschland geboren; die Großeltern sind im Ausland geboren) ist zu klein, um separat analysiert werden zu können. Wie auch in anderen Large-Scale-Assessments (z.B. Stanat, Rauch & Segeritz, 2010) wurden sie der Gruppe von Schülerinnen und Schülern ohne Zuwanderungshintergrund zugeordnet.

In der Literatur werden Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil unterschiedlichen Gruppen zugeordnet. Im Rahmen des 2009 durchgeführten IQB-Ländervergleichs sprachlicher Kompetenzen in der Sekundarstufe I wurden sie zu den Jugendlichen ohne Zuwanderungshintergrund gezählt (Böhme et al., 2010). Der Bericht zu IGLU-E 2006 (Schwippert et al., 2008) und der nationale Bericht zu PISA 2009 (Stanat, Rauch & Segeritz, 2010) wiesen die Ergebnisse dieser Teilgruppe hingegen entweder getrennt aus oder ordneten sie den Jugendlichen mit Zuwanderungshintergrund zu. Da sich gezeigt hat, dass zumindest in einigen Zuwanderungsgruppen auch Heranwachsende mit nur einem im Ausland geborenen Elternteil benachteiligt sind, berichten wir die Ergebnisse für diese Gruppe im Ländervergleich 2011 separat.

Zusätzlich zur Zuwanderungsgeneration werden anhand der Geburtsländer der Eltern verschiedene Herkunftsgruppen unterschieden. Die Zuordnungen beziehen sich auf Familien, in denen entweder beide Eltern außerhalb Deutschlands geboren sind oder ein Elternteil im Ausland und der andere Elternteil in Deutschland geboren ist. Die Herkunftsgruppen sind wie folgt definiert:

- *Türkei*: mindestens ein Elternteil ist in der Türkei geboren;
- *ehemalige Sowjetunion*: mindestens ein Elternteil ist in Russland oder in einer der ehemaligen Sowjetrepubliken geboren;
- *Polen*: mindestens ein Elternteil ist in Polen geboren;
- *ehemaliges Jugoslawien*: mindestens ein Elternteil ist in Bosnien und Herzegowina, dem Kosovo, Kroatien, Mazedonien, Montenegro, Serbien oder Slowenien geboren sowie
- *anderes Land*: mindestens ein Elternteil ist in einem hier nicht genannten Land geboren oder die Eltern sind in zwei unterschiedlichen Ländern im Ausland geboren.

Auch beim Umgang mit fehlenden Werten wurden in früheren Studien unterschiedliche Strategien verfolgt. Im Ländervergleich 2011 wird in Anlehnung an den nationalen Bericht von PISA 2009 sowohl für den Zuwanderungshintergrund als auch für die Herkunftsländer jeweils eine zusätzliche Gruppe gebildet, die als „nicht zuzuordnen“ bezeichnet wird (vgl. Stanat, Rauch & Segeritz, 2010). Sie umfasst diejenigen Schülerinnen und Schüler, für die nicht alle notwendigen Angaben zur Bestimmung des Zuwanderungshintergrunds vorliegen.

9.2 Viertklässlerinnen und Viertklässler mit Zuwanderungshintergrund in den Ländern

In diesem Teilkapitel wird zunächst berichtet, welche Anteile von Schülerinnen und Schülern der vierten Jahrgangsstufe in den Ländern aus zugewanderten Familien stammen. Daran anschließend werden die erreichten Kompetenzen der Schülerinnen und Schüler aus zugewanderten Familien in den einzelnen Ländern berichtet und mit den Kompetenzständen der Schülerinnen und Schüler ohne Zuwanderungshintergrund verglichen. Schließlich werden verschiedene Hintergrundmerkmale der Familien differenziert nach dem Zuwanderungshintergrund der Kinder dargestellt.

9.2.1 Prozentuale Anteile von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund

In Tabelle 9.1 sind die Anteile der verschiedenen Teilgruppen von Kindern aus zugewanderten Familien und von Kindern ohne Zuwanderungshintergrund sowohl für Deutschland insgesamt als auch für die einzelnen Länder dargestellt. Außerdem ist angegeben, welcher Anteil der Schülerinnen und Schüler sich jeweils aufgrund fehlender Angaben zum Zuwanderungshintergrund keiner der Gruppen eindeutig zuordnen lässt. Wie bereits in Kapitel 5 werden auch hier ergänzend die Werte von Schülerinnen und Schülern aufgeführt, die in einer Großstadt zur Schule gehen. Als Großstädte gelten dabei in Anlehnung an die in PISA 2000 verwendete Definition alle Städte mit mehr als 300 000 Einwohnern (Baumert et al., 2001; siehe auch Kapitel 5). Schülerinnen und Schüler der Stadtstaaten Berlin, Bremen und Hamburg wurden nicht in die Berechnung der Ergebnisse für Großstädte eingeschlossen, da diese als zusätzlicher Vergleichsmaßstab vor allem für die von den Schülerinnen und Schülern in den Stadtstaaten erreichten Kompetenzstände dienen sollen.

Tabelle 9.1: Prozentuale Anteile von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund in den Ländern

Land	ohne Zuwanderungshintergrund		mit Zuwanderungshintergrund gesamt		ein Elternteil im Ausland geboren		zweite Generation		erste Generation		nicht zuzuordnen %
	gültige %	(SE)	gültige %	(SE)	gültige %	(SE)	gültige %	(SE)	gültige %	(SE)	
Baden-Württemberg	70.9	(2.0)	29.1	(2.0)	12.8	(0.8)	14.2	(1.5)	2.1	(0.5)	9.3
Bayern	75.4	(2.1)	24.6	(2.1)	10.9	(0.9)	11.6	(1.7)	2.1	(0.4)	9.4
Berlin	64.4	(1.9)	35.6	(1.9)	15.7	(1.0)	16.5	(1.2)	3.4	(0.5)	15.9
Brandenburg	94.5	(0.7)	5.5	(0.7)	3.4	(0.5)	1.3	(0.3)	0.8	(0.3)	2.5
Bremen	58.4	(2.9)	41.6	(2.9)	13.5	(1.0)	24.5	(2.1)	3.7	(0.7)	16.9
Hamburg	56.2	(1.5)	43.8	(1.5)	16.8	(0.9)	22.8	(1.4)	4.2	(0.5)	19.2
Hessen	67.8	(2.1)	32.2	(2.1)	12.2	(0.6)	16.9	(1.5)	3.1	(0.6)	4.4
Mecklenburg-Vorpommern	95.5	(0.5)	4.5	(0.5)	2.5	(0.4)	1.4	(0.3)	0.6	(0.2)	2.0
Niedersachsen	79.1	(1.7)	20.9	(1.7)	6.8	(0.8)	11.9	(1.2)	2.3	(0.4)	3.2
Nordrhein-Westfalen	69.8	(2.3)	30.2	(2.3)	11.9	(1.1)	16.4	(1.7)	2.0	(0.4)	10.3
Rheinland-Pfalz	72.7	(1.7)	27.3	(1.7)	11.8	(0.9)	13.5	(1.4)	2.0	(0.4)	15.4
Saarland	76.3	(2.0)	23.7	(2.0)	9.3	(0.9)	12.7	(1.5)	1.8	(0.5)	12.5
Sachsen	93.4	(1.0)	6.6	(1.0)	3.4	(0.5)	2.0	(0.5)	1.2	(0.4)	6.4
Sachsen-Anhalt	93.6	(1.0)	6.4	(1.0)	3.1	(0.5)	1.9	(0.4)	1.5	(0.5)	3.0
Schleswig-Holstein	85.2	(1.3)	14.8	(1.3)	7.4	(0.8)	5.6	(0.9)	1.9	(0.4)	9.8
Thüringen	94.2	(0.9)	5.8	(0.9)	2.9	(0.4)	1.9	(0.6)	1.0	(0.3)	1.7
Großstädte ¹	56.7	(3.7)	43.3	(3.7)	14.8	(1.8)	25.5	(2.6)	3.1	(0.8)	12.3
Deutschland	75.3	(0.8)	24.7	(0.8)	10.2	(0.4)	12.5	(0.6)	2.1	(0.2)	8.7

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. gültige %: Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind.
¹ Ohne Berlin, Bremen und Hamburg. SE = Standardfehler.

In Deutschland liegt der Anteil der Schülerinnen und Schüler mit Zuwanderungshintergrund in der vierten Jahrgangsstufe insgesamt bei knapp 25 Prozent. Die niedrigsten Anteile sind mit weniger als 10 Prozent in den ostdeutschen Ländern zu finden. Vor allem in den Stadtstaaten Berlin, Bremen und Hamburg ist der Anteil der Schülerinnen und Schüler mit Zuwanderungshintergrund mit 36, 42 und 44 Prozent deutlich höher als in Deutschland insgesamt. Ein ähnlich hoher Anteil von Grundschulkindern mit Zuwanderungshintergrund findet sich auch in der Vergleichsgruppe der Großstädte. In Hessen und Nordrhein-Westfalen stammen ebenfalls mehr als 30 Prozent der Schülerinnen und Schüler in der vierten Jahrgangsstufe aus einer zugewanderten Familie.

Innerhalb der Gruppe der Schülerinnen und Schüler mit Zuwanderungshintergrund ist ein Großteil der Kinder in Deutschland geboren (zweite Generation). Der Anteil der selbst zugewanderten Schülerinnen und Schüler (erste Generation) ist hingegen in allen Ländern sehr klein. Außer in Hamburg macht die erste Generation in keinem Land mehr als 4 Prozent aus. Der Anteil der zweiten Generation ist im Allgemeinen deutlich größer und in den Stadtstaaten (Berlin: 16.5 %, Bremen: 24.5 %, Hamburg: 22.8 %), Hessen (16.9 %) und Nordrhein-Westfalen (16.4 %) besonders hoch. In diesen Ländern sowie zusätzlich auch in Baden-Württemberg und Rheinland-Pfalz ist zudem der Anteil an Schülerinnen und Schülern mit einem im Ausland geborenen Elternteil relativ groß. Aufgrund der geringen Gruppengrößen der selbst zugewanderten Schülerinnen und Schüler werden in den folgenden Länderanalysen die erste und zweite Generation zu einer Gruppe zusammengefasst, die als „beide Eltern im Ausland geboren“ bezeichnet wird.

Der Anteil von Schülerinnen und Schülern, die sich aufgrund von fehlenden Angaben zum Geburtsland der Familie keiner Gruppe zuordnen lassen, variiert ebenfalls über die Länder. Diese Heterogenität ist vor allem auf die unterschiedlichen Rücklaufquoten der Eltern- und Schülerfragebögen in den einzelnen Ländern zurückzuführen (vgl. Abbildung 4.2 in Kapitel 4). Da die relevanten Informationen sowohl im Elternfragebogen als auch im Schülerfragebogen erhoben wurden, fällt jedoch der Anteil eindeutig zuzuordnender Schülerinnen und Schüler etwas höher aus als die Rücklaufquoten der einzelnen Fragebögen.

9.2.2 Kompetenzstände von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund

In diesem Abschnitt werden die von Kindern mit unterschiedlichem Zuwanderungshintergrund bis zur vierten Jahrgangsstufe erworbenen Kompetenzen für die Bereiche Lesen und Zuhören im Fach Deutsch sowie für die Globalskala im Fach Mathematik in den Ländern verglichen. Hierbei werden sowohl die im Mittel erreichten Kompetenzstände als auch die Streuungen innerhalb der Gruppen betrachtet. Um ausreichend präzise Schätzungen der Leistungsdifferenzen zu gewährleisten, werden die Ergebnisse der beiden Zuwanderungsgruppen nur für Länder berichtet, in denen der Anteil von mindestens einer der beiden Gruppen mehr als 10 Prozent beträgt (siehe Böhme et al., 2010; Frey, Carstensen, Walter, Rönnebeck & Gomolka, 2008).

Da sich die Streuungen der Kompetenzwerte in den insgesamt drei untersuchten Gruppen sowohl innerhalb als auch zwischen den Ländern teilweise erheblich unterscheiden, sind in den Abbildungen 9.1 bis 9.3 zusätzlich zu den Mittelwerten (M), Standardfehlern (SE), Standardabweichungen (SD) und

Perzentilbändern auch standardisierte Mittelwertsdifferenzen (d) angegeben. Die standardisierte Mittelwertsdifferenz relativiert die zwischen zwei Gruppen festgestellten Unterschiede an ihrer gemeinsamen Standardabweichung und ist somit von der Streuung des Merkmals in den Gruppen unabhängig. Dadurch wird die Interpretation von Vergleichen der Disparitäten zwischen den Ländern erleichtert. Die Betrachtung der Mittelwertsunterschiede zwischen den Gruppen auf der Berichtsmetrik hingegen hat den Vorteil, dass diese sich anschaulich mit den zu erwartenden Lernzuwächsen in einem Schuljahr in Beziehung setzen lassen. Daher werden beide Kennwerte berichtet. Die standardisierten Mittelwertsdifferenzen beziehen sich jeweils auf den Unterschied zu Kindern ohne Zuwanderungshintergrund, das heißt positive d -Werte kennzeichnen den Kompetenzvorteil dieser Gruppe im Vergleich zur jeweiligen Gruppe von Kindern aus zugewanderten Familien.

Abbildung 9.1 zeigt die Perzentilbänder für den Kompetenzbereich Lesen für Deutschland insgesamt, für die einzelnen Länder und für Kinder, die in einer Großstadt zur Schule gehen. Nach diesen Ergebnissen erreichen bundesweit betrachtet Schülerinnen und Schüler, deren Eltern in Deutschland geboren sind, im Kompetenzbereich Lesen im Mittel 514 Punkte. Die Lesekompetenz von Kindern mit einem im Ausland geborenen Elternteil liegt bei 489 Punkten und damit um 25 Punkte niedriger ($d = 0.26$). Kinder mit zwei im Ausland geborenen Elternteilen erreichen im Mittel einen Kompetenzstand von 460 Punkten, was 29 Punkte unter dem Wert der Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil liegt. Der Abstand von 54 Punkten zwischen Kindern ohne Zuwanderungshintergrund und Kindern mit zwei im Ausland geborenen Elternteilen ($d = 0.56$) entspricht dem erwarteten Leistungszuwachs von fast einem Schuljahr (61 Punkte; Bremerich-Vos & Böhme, 2009, S. 240). Die Standardabweichungen hingegen sind in allen drei Gruppen sehr ähnlich ($SD_{oZH} = 96$, $SD_{1Elternteil} = 99$, $SD_{2Elternteile} = 98$).

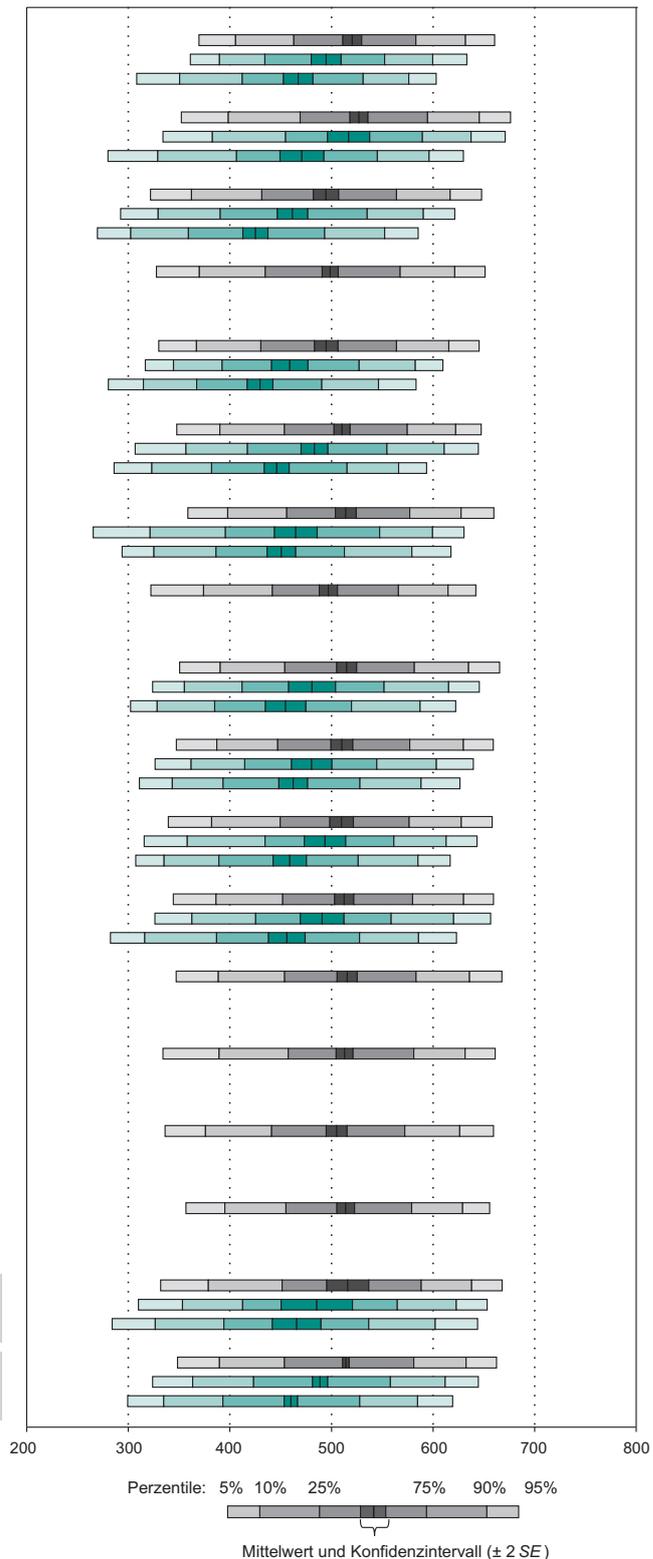
Die Mittelwertsunterschiede zwischen Kindern mit zwei im Ausland geborenen Elternteilen und Kindern ohne Zuwanderungshintergrund sind in allen Ländern statistisch signifikant, sie variieren jedoch teilweise erheblich. Die größte Differenz findet sich mit 70 Punkten in Berlin ($d = 0.70$), die kleinste mit 48 Punkten in Nordrhein-Westfalen ($d = 0.51$). Insgesamt bestehen in den Stadtstaaten besonders große Unterschiede zwischen Kindern mit zwei im Ausland geborenen Elternteilen und Kindern ohne Zuwanderungshintergrund. Diese sind wesentlich größer als in anderen Großstädten mit über 300 000 Einwohnern.⁵ Relativ ausgeprägte Disparitäten sind zudem mit 63 Punkten in Hessen zu verzeichnen ($d = 0.67$).

Zwischen Kindern mit einem im Ausland geborenen Elternteil und Kindern ohne Zuwanderungshintergrund zeigen sich in einigen Ländern keine signifikanten Unterschiede in den im Mittel erreichten Kompetenzständen. Dies ist in den Ländern Bayern, Rheinland-Pfalz und dem Saarland der Fall. Die insgesamt geringste Mittelwertsdifferenz für diese Gruppe findet sich mit 10 Punkten in Bayern ($d = 0.10$). In Hamburg und Hessen ist zudem die Streuung in der Gruppe der Kinder mit einem im Ausland geborenen Elternteil im Vergleich zu Kindern ohne Zuwanderungshintergrund deutlich größer. Hier scheinen die er-

5 Großstädte mit einer Einwohnerzahl von mehr als 300 000 sind Bielefeld, Bochum, Bonn, Dortmund, Dresden, Duisburg, Düsseldorf, Essen, Frankfurt am Main, Hannover, Köln, Leipzig, Mannheim, München, Nürnberg, Stuttgart, Wuppertal (Gebietsstand 31.12.2010, Quelle: Statistisches Bundesamt).

Abbildung 9.1: Mittelwerte, Streuungen und Perzentilbänder der Kompetenzverteilungen im Bereich Lesen nach Zuwanderungshintergrund und Land

Land	gültige %	M	(SE)	SD	d
Baden-Württemberg	70.8	520	(4.6)	90	
	12.9	495	(7.3)	83	0.29
	16.3	467	(7.2)	90	0.59
Bayern	75.4	527	(4.4)	100	
	10.9	517	(10.4)	104	0.10
	13.8	471	(10.8)	106	0.56
Berlin	64.4	495	(6.2)	100	
	15.7	462	(7.5)	101	0.33
	20.0	425	(6.2)	96	0.70
Brandenburg	94.6	499	4	98	
	--	--	--	--	
	--	--	--	--	
Bremen	58.5	495	(5.8)	97	
	13.3	459	(9.0)	93	0.37
	28.2	430	(6.3)	91	0.68
Hamburg	55.7	510	(3.9)	92	
	17.1	483	(6.6)	102	0.29
	27.2	446	(6.1)	95	0.69
Hessen	67.8	514	(5.2)	94	
	12.1	465	(10.5)	116	0.50
	20.1	451	(7.0)	98	0.67
Mecklenburg-Vorpommern	95.5	497	(4.5)	102	
	--	--	--	--	
	--	--	--	--	
Niedersachsen	79.1	515	(4.9)	98	
	6.8	481	(11.6)	102	0.35
	14.1	455	(10.0)	100	0.61
Nordrhein-Westfalen	69.7	510	(5.4)	94	
	12.1	481	(9.9)	95	0.31
	18.2	462	(7.1)	96	0.51
Rheinland-Pfalz	72.9	510	(5.8)	96	
	11.6	494	(10.2)	97	0.17
	15.5	459	(8.2)	97	0.53
Saarland	76.5	513	(4.7)	95	
	9.1	491	(10.7)	101	0.23
	14.4	456	(9.0)	103	0.59
Sachsen	93.5	515	(4.9)	99	
	--	--	--	--	
	--	--	--	--	
Sachsen-Anhalt	93.6	513	(4.2)	102	
	--	--	--	--	
	--	--	--	--	
Schleswig-Holstein	85.1	505	(5.1)	97	
	--	--	--	--	
	--	--	--	--	
Thüringen	94.2	514	(4.3)	91	
	--	--	--	--	
	--	--	--	--	
Großstädte ¹	56.8	516	(10.3)	102	
	14.8	486	(17.5)	111	0.29
	28.4	466	(12.0)	108	0.48
Deutschland	75.3	514	(1.6)	96	
	10.2	489	(3.7)	99	0.26
	14.6	460	(3.3)	98	0.56



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. gültige %: Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind.

¹ Ohne Berlin, Bremen und Hamburg.

fett: signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund

grau: Kinder ohne Zuwanderungshintergrund (beide Elternteile sind in Deutschland geboren)

grün (oberes Band): Kinder mit einem im Ausland geborenen Elternteil

grün (unteres Band): Kinder mit zwei im Ausland geborenen Elternteilen

Für Länder, deren Anteil an Zuwanderern in beiden Gruppen unter 10 Prozent liegt, werden nur die Ergebnisse für Kinder ohne Zuwanderungshintergrund berichtet. M = Mittelwert, SE = Standardfehler, SD = Standardabweichung, d = standardisierte Mittelwertsdifferenz.

reichten Kompetenzstände von Kindern mit einem zugewanderten Elternteil also heterogener zu sein.

An den Perzentilbändern ist außerdem zu erkennen, dass die Unterschiede zwischen Kindern mit einem im Ausland geborenen Elternteil und Kindern ohne Zuwanderungshintergrund nicht in allen Bereichen der Verteilung gleich ausgeprägt sind. In Hamburg etwa beträgt der Kompetenzrückstand von Schülerinnen und Schülern mit einem zugewanderten Elternteil im Mittel 27 Punkte. Betrachtet man hingegen den Abstand der jeweils leistungsschwächsten Schülerinnen und Schüler in beiden Gruppen, also das 5. Perzentil, ist ein Abstand von 41 Punkten zu verzeichnen, während sich die leistungsstärksten Schülerinnen und Schüler (95. Perzentil) nur um 3 Punkte unterscheiden. Dies deutet darauf hin, dass in Hamburg vor allem im unteren Kompetenzspektrum deutliche Kompetenzrückstände für Schülerinnen und Schüler mit einem zugewanderten Elternteil bestehen. Im Saarland, in Hessen und in Bayern bestehen ebenfalls besonders ausgeprägte Kompetenznachteile im unteren Kompetenzspektrum, wobei in Hessen jedoch der Kompetenzrückstand auch im obersten Bereich der Verteilung (95. Perzentil) mit 29 Punkten noch substantiell ist. Im Gegensatz dazu sind in Baden-Württemberg und Bremen die Unterschiede zwischen den beiden Gruppen für die leistungsschwächsten Schülerinnen und Schüler (Baden-Württemberg: 8 Punkte, Bremen: 13 Punkte) kleiner als die Unterschiede für die leistungsstärksten Schülerinnen und Schüler (Baden-Württemberg: 26 Punkte, Bremen: 36 Punkte). In diesen Ländern bestehen im unteren Kompetenzspektrum demnach deutlich geringere zuwanderungsbezogene Kompetenzrückstände.

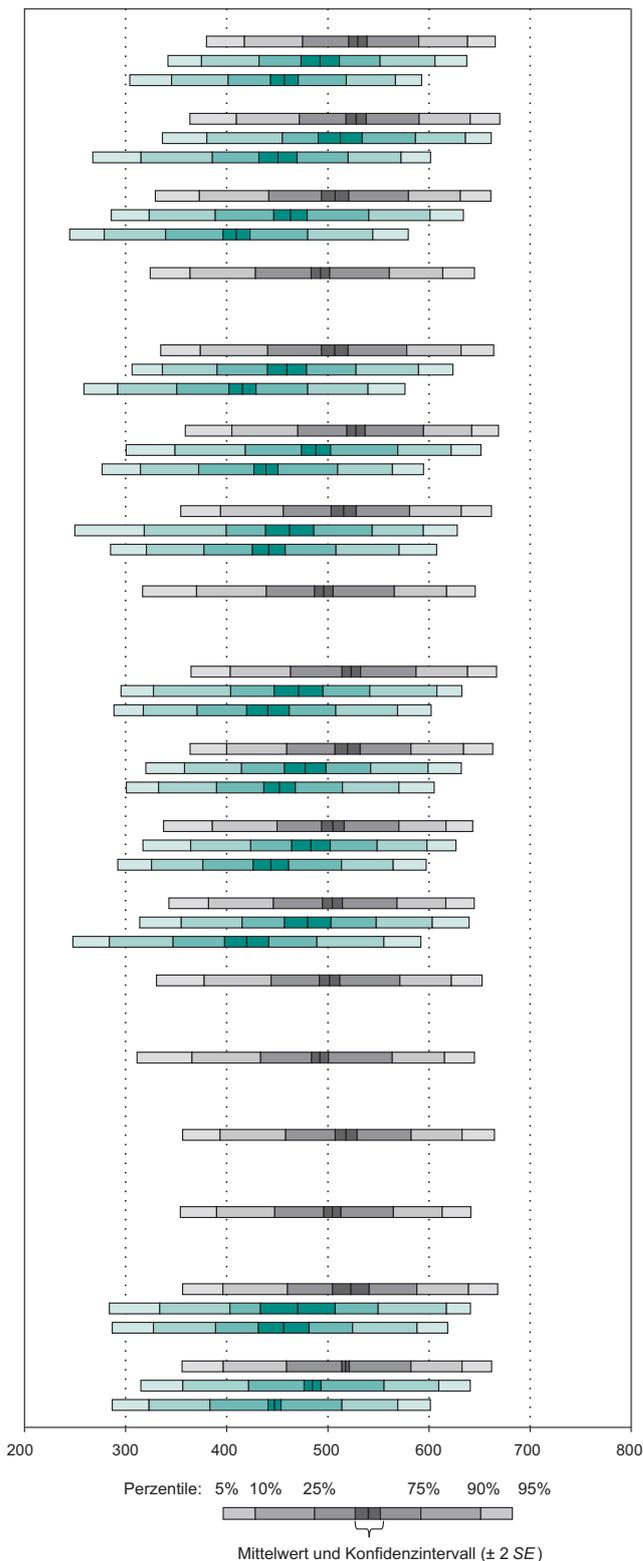
Die Ergebnisse für den Kompetenzbereich Zuhören sind in Abbildung 9.2 dargestellt. Auch in diesem Bereich erzielten Schülerinnen und Schüler, deren Eltern in Deutschland geboren sind, mit 518 Punkten bundesweit im Mittel die höchsten Kompetenzstände. Allgemein sind in diesem Kompetenzbereich die Mittelwertsdifferenzen etwas größer als im Kompetenzbereich Lesen. Dies ist insbesondere bei Schülerinnen und Schülern mit zwei im Ausland geborenen Elternteilen der Fall, die im Mittel 447 Punkte erreichen und somit um 71 Punkte ($d = 0.74$) hinter Schülerinnen und Schülern ohne Zuwanderungshintergrund zurückliegen. Der Kompetenzrückstand der Kinder mit zwei im Ausland geborenen Elternteilen ist im Bereich Zuhören somit sogar etwas größer als der zu erwartende Leistungszuwachs zwischen der dritten und vierten Jahrgangsstufe, der im Kompetenzbereich Zuhören bei etwa 60 Punkten liegt (Behrens, Böhme & Krelle, 2009, S. 368).

Kinder mit einem im Ausland geborenen Elternteil erreichen hingegen im Mittel einen Kompetenzstand von 485 Punkten, so dass die Differenz zwischen diesen Kindern und Kindern ohne Zuwanderungshintergrund 33 Punkte ($d = 0.34$) beträgt. Die Gruppe der Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil weist zudem, verglichen mit Kindern ohne Zuwanderungshintergrund, eine höhere Streuung auf ($SD_{\text{Elternteil}} = 101$; $SD_{\text{oZH}} = 95$).

Die Unterschiede zwischen Kindern mit zwei im Ausland geborenen Elternteilen und Kindern ohne Zuwanderungshintergrund sind im Bereich Zuhören in allen Ländern statistisch signifikant. Wie im Bereich Lesen sind auch im Bereich Zuhören die Disparitäten für diese Gruppe in den Stadtstaaten am größten und etwas stärker ausgeprägt als in anderen Großstädten. Sie variieren zwischen den Ländern noch etwas stärker als im Lesen und liegen zwischen 61 Punkten ($d = 0.64$) in Rheinland-Pfalz und 97 Punkten ($d = 0.96$) in Berlin. Der Unterschied zwischen Kindern mit einem im Ausland geborenen Elternteil und

Abbildung 9.2: Mittelwerte, Streuungen und Perzentilbänder der Kompetenzverteilungen im Bereich Zuhören nach Zuwanderungshintergrund und Land

Land	gültige %	M	(SE)	SD	d
Baden-Württemberg	70.8	530	(4.5)	89	
	12.9	492	(9.5)	89	0.42
	16.3	457	(6.8)	88	0.82
Bayern	75.4	528	(5.0)	93	
	10.9	512	(10.9)	103	0.17
	13.8	451	(9.4)	103	0.82
Berlin	64.4	507	(6.7)	102	
	15.7	463	(8.2)	107	0.43
	20.0	410	(6.7)	102	0.96
Brandenburg	94.6	493	(4.6)	98	
	--	--	--	--	
	--	--	--	--	
Bremen	58.5	507	(6.6)	101	
	13.3	460	(9.6)	99	0.47
	28.2	416	(6.7)	96	0.92
Hamburg	55.7	528	(4.5)	96	
	17.1	488	(7.3)	111	0.40
	27.2	439	(5.9)	98	0.93
Hessen	67.8	516	(6.2)	95	
	12.1	462	(11.9)	124	0.53
	20.1	442	(8.1)	101	0.77
Mecklenburg-Vorpommern	95.5	496	(4.6)	105	
	--	--	--	--	
	--	--	--	--	
Niedersachsen	79.1	523	(4.6)	93	
	6.8	471	(12.1)	102	0.55
	14.1	441	(10.5)	96	0.88
Nordrhein-Westfalen	69.7	520	(6.2)	91	
	12.1	478	(10.3)	94	0.46
	18.2	452	(7.8)	93	0.74
Rheinland-Pfalz	72.9	505	(5.6)	95	
	11.6	483	(9.5)	93	0.23
	15.5	444	(8.7)	98	0.64
Saarland	76.5	505	(5.0)	91	
	9.1	480	(11.5)	98	0.27
	14.4	420	(10.9)	103	0.91
Sachsen	93.5	502	(5.0)	104	
	--	--	--	--	
	--	--	--	--	
Sachsen-Anhalt	93.6	492	(4.2)	105	
	--	--	--	--	
	--	--	--	--	
Schleswig-Holstein	85.1	518	(5.4)	94	
	--	--	--	--	
	--	--	--	--	
Thüringen	94.2	504	(4.3)	87	
	--	--	--	--	
	--	--	--	--	
Großstädte ¹	56.8	523	(9.0)	94	
	14.8	470	(18.4)	114	0.53
	28.4	456	(12.5)	104	0.68
Deutschland	75.3	518	(1.8)	95	
	10.2	485	(4.2)	101	0.34
	14.6	447	(3.2)	97	0.74



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. gültige %: Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind.

¹ Ohne Berlin, Bremen und Hamburg.

fett: signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund
 grau: Kinder ohne Zuwanderungshintergrund (beide Elternteile sind in Deutschland geboren)
 grün (oberes Band): Kinder mit einem im Ausland geborenen Elternteil
 grün (unteres Band): Kinder mit zwei im Ausland geborenen Elternteilen

Für Länder, deren Anteil an Zuwanderern in beiden Gruppen unter 10 Prozent liegt, werden nur die Ergebnisse für Kinder ohne Zuwanderungshintergrund berichtet. M = Mittelwert, SE = Standardfehler; SD = Standardabweichung, d = standardisierte Mittelwertsdifferenz.

Kindern ohne Zuwanderungshintergrund ist in Bayern am kleinsten (16 Punkte; $d = 0.17$).

Auch im Kompetenzbereich Zuhören können die zuwanderungsbezogenen Disparitäten in unterschiedlichen Leistungssegmenten miteinander verglichen werden. Im Zuhören erreichen die leistungsstärksten Kinder (95. Perzentil) mit einem zugewanderten Elternteil im Saarland Kompetenzwerte, die denen der leistungsstärksten Kinder ohne Zuwanderungshintergrund nahezu entsprechen. Der Abstand zwischen den 5. Perzentilen der beiden Gruppen beträgt hier 29 Punkte, der Abstand zwischen den 95. Perzentilen hingegen nur fünf Punkte. Ein ähnliches Muster findet sich in Hamburg, Bayern und Hessen, wobei in Hessen auch im Bereich Zuhören mit einem Unterschied von 34 Punkten substantielle Disparitäten für Schülerinnen und Schüler im obersten Kompetenzspektrum (95. Perzentile) bestehen. Ein stärkerer Kompetenzunterschied im obersten Kompetenzbereich besteht ansatzweise in Bremen. Dort bleiben die erreichten Kompetenzstände der leistungsstärksten Kinder mit einem zugewanderten Elternteil hinter denen der leistungsstärksten Kinder ohne Zuwanderungshintergrund um 40 Punkte zurück, wohingegen der Unterschied für die jeweils leistungsschwächsten Schülerinnen und Schüler nur 28 Punkte beträgt.

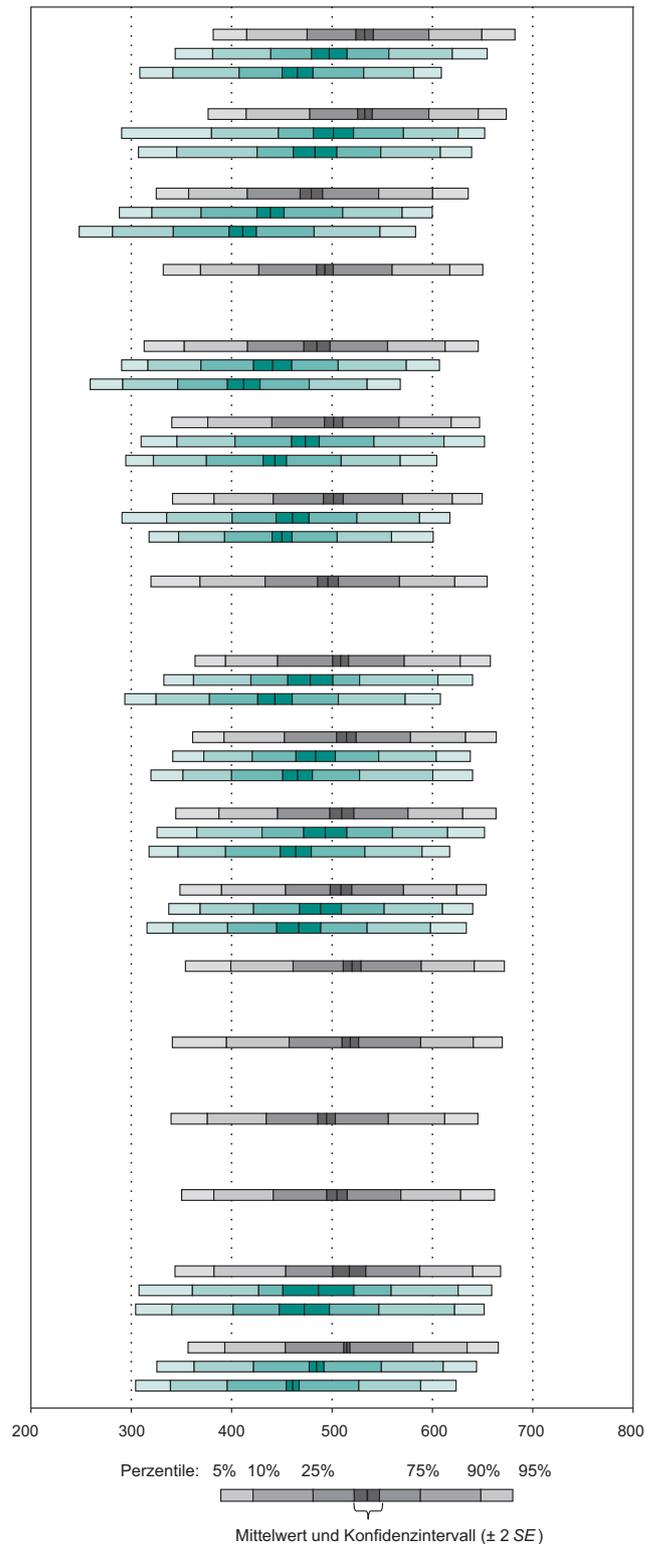
Die Ergebnisse für das Fach Mathematik (Globalskala) sind in Abbildung 9.3 dargestellt. In diesem Fach erreichen Schülerinnen und Schüler ohne Zuwanderungshintergrund 515 Punkte auf der Kompetenzskala. In Mathematik sind die mit einem Zuwanderungshintergrund verbundenen Disparitäten in der Bundesrepublik Deutschland insgesamt mit denen im Kompetenzbereich Lesen vergleichbar. Für Kinder mit zwei im Ausland geborenen Elternteilen ergibt sich ein Kompetenzrückstand von 54 Punkten ($d = 0.56$), der wesentlich stärker ausgeprägt ist als der Rückstand von Kindern mit nur einem im Ausland geborenen Elternteil (30 Punkte, $d = 0.31$). Im Vergleich zum Fach Deutsch sind in Mathematik mit circa 80 Punkten jedoch größere Lernzuwächse zwischen der dritten und vierten Jahrgangsstufe zu erwarten (Reiss & Winkelmann, 2009), was die Rückstände im Vergleich zum Kompetenzbereich Lesen etwas relativiert.

Die größten Disparitäten für Kinder mit zwei im Ausland geborenen Elternteilen in den Ländern finden sich wiederum in den Stadtstaaten sowie in Baden-Württemberg und Niedersachsen. Dies ist insbesondere deshalb bemerkenswert, da Baden-Württemberg in Mathematik zur Gruppe der Länder gehört, in denen im Vergleich zum deutschen Mittelwert insgesamt signifikant höhere Kompetenzstände erreicht werden (vgl. Kapitel 5.2). Dieses hohe Durchschnittsniveau wird von Schülerinnen und Schülern mit zwei zugewanderten Elternteilen offenbar nicht annähernd erreicht. Im Fach Mathematik sind die geringsten Disparitäten in Rheinland-Pfalz und im Saarland zu finden.

Auch im Bereich Mathematik bestehen für Schülerinnen und Schüler mit einem zugewanderten Elternteil in einigen Ländern differenzielle Kompetenzrückstände im unteren und oberen Kompetenzspektrum. Dies ist insbesondere in Bayern der Fall. Dort unterscheiden sich die leistungsstärksten Schülerinnen und Schüler (95. Perzentil) mit einem zugewanderten Elternteil um 22 Punkte von den leistungsstärksten Schülerinnen und Schülern ohne Zuwanderungshintergrund, wohingegen der Unterschied der jeweils leistungsschwächsten Schülerinnen und Schüler mit 86 Punkten ungefähr viermal so groß ist. In Hamburg erreichen die leistungsstärksten 5 Prozent der Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil sogar um 5 Punkte höhere Kompetenzwerte als die leistungsstärksten 5 Prozent der Kinder ohne Zuwanderungshintergrund;

Abbildung 9.3: Mittelwerte, Streuungen und Perzentilbänder der Kompetenzverteilungen im Fach Mathematik (Globalskala) nach Zuwanderungshintergrund und Land

Land	gültige %	M	(SE)	SD	d
Baden-Württemberg	71.0	533	(4.3)	94	
	12.7	497	(8.9)	92	0.37
	16.3	466	(7.7)	92	0.72
Bayern	75.4	533	(3.6)	93	
	10.9	501	(10.0)	104	0.33
	13.6	483	(10.8)	104	0.53
Berlin	64.2	479	(5.6)	97	
	15.7	439	(6.8)	98	0.42
	20.1	411	(6.8)	101	0.70
Brandenburg	94.5	493	(4.1)	98	
	--	--	--	--	
	--	--	--	--	
Bremen	58.3	485	(6.5)	102	
	13.6	441	(9.6)	98	0.44
	28.1	412	(8.1)	96	0.73
Hamburg	56.3	502	(4.5)	93	
	16.9	474	(6.9)	104	0.29
	26.7	443	(5.8)	95	0.63
Hessen	67.9	501	(5.0)	101	
	12.1	461	(8.2)	100	0.40
	19.9	450	(5.0)	88	0.52
Mecklenburg-Vorpommern	95.8	496	(5.1)	106	
	--	--	--	--	
	--	--	--	--	
Niedersachsen	79.2	509	(4.0)	93	
	6.7	478	(11.2)	93	0.32
	14.2	443	(8.6)	98	0.70
Nordrhein-Westfalen	69.9	514	(4.9)	92	
	11.8	484	(9.8)	90	0.33
	18.4	466	(7.4)	96	0.52
Rheinland-Pfalz	72.6	510	(6.0)	98	
	11.8	493	(10.7)	99	0.17
	15.6	464	(7.7)	96	0.47
Saarland	76.6	509	(5.3)	94	
	9.2	489	(10.4)	93	0.22
	14.1	467	(11.0)	96	0.45
Sachsen	93.5	520	(4.4)	101	
	--	--	--	--	
	--	--	--	--	
Sachsen-Anhalt	93.5	518	(4.1)	104	
	--	--	--	--	
	--	--	--	--	
Schleswig-Holstein	85.0	495	(4.3)	92	
	--	--	--	--	
	--	--	--	--	
Thüringen	94.2	505	(5.1)	95	
	--	--	--	--	
	--	--	--	--	
Großstädte ¹	56.7	517	(8.3)	99	
	14.7	486	(17.7)	110	0.30
	28.6	472	(12.5)	108	0.44
Deutschland	75.3	515	(1.5)	96	
	10.1	485	(3.6)	97	0.31
	14.6	461	(3.2)	98	0.56



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. gültige %: Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind.

¹ Ohne Berlin, Bremen und Hamburg.

fett: signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund

grau: Kinder ohne Zuwanderungshintergrund (beide Elternteile sind in Deutschland geboren)

grün (oberes Band): Kinder mit einem im Ausland geborenen Elternteil

grün (unteres Band): Kinder mit zwei im Ausland geborenen Elternteilen

Für Länder, deren Anteil an Zuwanderern in beiden Gruppen unter 10 Prozent liegt, werden nur die Ergebnisse für Kinder ohne Zuwanderungshintergrund berichtet. M = Mittelwert, SE = Standardfehler; SD = Standardabweichung, d = standardisierte Mittelwertsdifferenz.

im unteren Kompetenzspektrum ist dagegen mit 31 Punkten ein deutlicher Kompetenzrückstand zu beobachten. Im Gegensatz zu den Bereichen Lesen und Zuhören ist im Saarland für den Bereich Mathematik kein ausgeprägter Nachteil der Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil im untersten Kompetenzspektrum zu identifizieren.

Zusammenfassend kann festgehalten werden, dass die zuwanderungsbezogenen Disparitäten für Schülerinnen und Schüler mit zwei im Ausland geborenen Elternteilen in allen Ländern und in allen Kompetenzbereichen am stärksten ausgeprägt sind. Verglichen mit den im Mittel von der dritten bis zur vierten Jahrgangsstufe zu erwartenden Leistungszuwächsen entsprechen sie in der Bundesrepublik Deutschland insgesamt im Lesen etwa einem Schuljahr, im Zuhören etwas mehr als einem Schuljahr und in Mathematik etwa zwei Drittel eines Schuljahres. Die Kompetenzrückstände von Schülerinnen und Schülern mit einem im Ausland und einem in Deutschland geborenen Elternteil sind jeweils nur etwa halb so groß und nicht in allen Ländern signifikant.

9.2.3 Merkmale des familiären Hintergrunds von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund

Zur Interpretation der mit einem Zuwanderungshintergrund assoziierten Disparitäten ist es aufschlussreich, weitere Hintergrundmerkmale zwischen den Gruppen zu vergleichen, die für den Kompetenzerwerb relevant sein können. In diesem Abschnitt werden daher deskriptive Analysen für den sozioökonomischen Status, die in der Familie gesprochene Sprache und das Bildungsniveau der Eltern dargestellt.

Der sozioökonomische Status wird mit dem *Highest International Socio-Economic Index of Occupational Status* (HISEI; Ganzeboom, de Graaf, Treiman & de Leeuw, 1992) abgebildet. Eine genauere Beschreibung dieses Indikators ist in Kapitel 8 zu finden. Für die hier dargestellten Analysen wurden, wie in Kapitel 8, fehlende HISEI-Werte durch fünf multiple Imputationen ersetzt (vgl. Graham, 2009; Schafer & Graham, 2002). Die Ermittlung des Bildungsniveaus der Familie erfolgte anhand der Angaben der Eltern zu ihrem Bildungsabschluss. Hierbei wurde der höchste in der Familie erzielte Bildungsabschluss nach der *International Standard Classification of Education* (ISCED; OECD, 1999) klassifiziert und anschließend auf der Grundlage dieser Information die Anzahl der Bildungsjahre geschätzt (OECD, 2009). Zur Erfassung der Familiensprache wurden die Schülerinnen und Schüler gefragt, wie häufig in der Familie Deutsch gesprochen wird, wobei zur Beantwortung die Kategorien „immer“, „manchmal“ und „nie“ vorgegeben waren. Wenn von den Schülerinnen und Schülern keine Angaben vorlagen, wurden die Angaben der Eltern herangezogen, um den Anteil fehlender Werte auf dieser Variablen gering zu halten. In Tabelle 9.2 sind die Mittelwerte des sozioökonomischen Status und des Bildungsniveaus der Eltern sowie Informationen zur Familiensprache Deutsch zusammengestellt.

Insgesamt ist der sozioökonomische Status von zugewanderten Familien in Deutschland niedriger als von Familien ohne Zuwanderungshintergrund. Am ungünstigsten ist die soziale Lage von Kindern mit zwei im Ausland geborenen Elternteilen. Der Unterschied zu Kindern ohne Zuwanderungshintergrund ist für Kinder mit zwei zugewanderten Elternteilen mehr als drei Mal so groß wie für Kinder mit einem zugewanderten Elternteil. Die Disparitäten in der sozioökonomischen Lage zwischen Schülerinnen und Schülern aus zugewanderten Familien

Tabelle 9.2: Merkmale des familiären Hintergrunds von Viertklässlerinnen und Viertklässlern nach Zuwanderungshintergrund und Land

	sozio-ökonomischer Status		Bildungsniveau der Eltern ¹		Sprachgebrauch in der Familie					
					immer Deutsch		manchmal Deutsch		nie Deutsch ²	
	M	(SE)	M	(SE)	gültige %	(SE)	gültige %	(SE)	gültige %	(SE)
Baden-Württemberg										
ohne Zuwanderungshintergrund	54.6	(0.7)	15.1	(0.1)	96.1	(0.6)	3.9	(0.6)	0.0	(0.0)
ein Elternteil im Ausland geboren	50.1	(1.4)	14.5	(0.3)	61.2	(4.2)	38.4	(4.1)	0.4	(0.4)
beide Elternteile im Ausland geboren	41.2	(1.2)	13.2	(0.3)	34.4	(2.9)	59.4	(2.9)	6.1	(1.5)
Bayern										
ohne Zuwanderungshintergrund	51.6	(1.0)	14.2	(0.2)	97.0	(0.6)	2.6	(0.5)	0.4	(0.3)
ein Elternteil im Ausland geboren	49.5	(1.8)	14.3	(0.3)	68.4	(3.3)	30.2	(3.1)	1.4	(1.0)
beide Elternteile im Ausland geboren	40.3	(1.2)	13.5	(0.3)	38.1	(2.9)	56.6	(2.9)	5.3	(1.5)
Berlin										
ohne Zuwanderungshintergrund	54.2	(0.8)	15.3	(0.2)	93.1	(0.9)	6.8	(0.9)	0.1	(0.1)
ein Elternteil im Ausland geboren	47.3	(1.3)	14.6	(0.3)	54.8	(2.8)	44.4	(2.7)	0.8	(0.5)
beide Elternteile im Ausland geboren	40.7	(1.2)	13.4	(0.2)	18.4	(1.6)	74.8	(2.2)	6.8	(1.5)
Brandenburg										
ohne Zuwanderungshintergrund	48.2	(0.9)	14.5	(0.1)	96.0	(0.5)	4.0	(0.5)	0.0	(0.0)
ein Elternteil im Ausland geboren	--	--	--	--	--	--	--	--	--	--
beide Elternteile im Ausland geboren	--	--	--	--	--	--	--	--	--	--
Bremen										
ohne Zuwanderungshintergrund	51.7	(1.2)	14.6	(0.2)	95.0	(0.8)	4.9	(0.8)	0.1	(0.1)
ein Elternteil im Ausland geboren	43.1	(1.7)	13.8	(0.4)	56.6	(4.2)	43.4	(4.2)	0.0	(0.0)
beide Elternteile im Ausland geboren	39.2	(1.0)	13.3	(0.3)	29.7	(2.1)	67.2	(2.0)	3.2	(1.0)
Hamburg										
ohne Zuwanderungshintergrund	55.4	(0.8)	15.2	(0.2)	96.3	(0.7)	3.5	(0.7)	0.2	(0.1)
ein Elternteil im Ausland geboren	49.9	(1.2)	14.8	(0.2)	64.2	(3.0)	34.6	(2.9)	1.2	(0.6)
beide Elternteile im Ausland geboren	41.6	(1.0)	13.4	(0.3)	31.4	(2.1)	63.6	(2.1)	5.0	(1.2)
Hessen										
ohne Zuwanderungshintergrund	51.8	(0.9)	14.6	(0.2)	96.4	(0.7)	3.5	(0.6)	0.1	(0.1)
ein Elternteil im Ausland geboren	46.7	(1.6)	13.8	(0.3)	59.1	(3.4)	40.9	(3.4)	0.0	(0.0)
beide Elternteile im Ausland geboren	39.1	(1.0)	13.1	(0.3)	30.8	(2.2)	64.7	(2.1)	4.5	(1.2)
Mecklenburg-Vorpommern										
ohne Zuwanderungshintergrund	47.5	(0.8)	14.0	(0.1)	96.0	(0.6)	4.0	(0.6)	0.0	(0.0)
ein Elternteil im Ausland geboren	--	--	--	--	--	--	--	--	--	--
beide Elternteile im Ausland geboren	--	--	--	--	--	--	--	--	--	--
Niedersachsen										
ohne Zuwanderungshintergrund	49.1	(0.6)	14.4	(0.1)	97.1	(0.5)	2.9	(0.5)	0.0	(0.0)
ein Elternteil im Ausland geboren	46.1	(1.7)	14.4	(0.3)	62.5	(4.8)	35.8	(4.6)	1.8	(1.2)
beide Elternteile im Ausland geboren	38.6	(1.4)	13.1	(0.3)	33.1	(3.7)	62.0	(3.5)	4.9	(1.5)

	sozio- ökonomischer Status		Bildungs- niveau der Eltern ¹		Sprachgebrauch in der Familie					
					immer Deutsch		manchmal Deutsch		nie Deutsch ²	
	<i>M</i>	<i>(SE)</i>	<i>M</i>	<i>(SE)</i>	gültige %	<i>(SE)</i>	gültige %	<i>(SE)</i>	gültige %	<i>(SE)</i>
Nordrhein-Westfalen										
ohne Zuwanderungshintergrund	51.1	(0.7)	14.7	(0.1)	95.9	(0.6)	3.9	(0.5)	0.2	(0.1)
ein Elternteil im Ausland geboren	45.7	(1.8)	14.1	(0.3)	57.3	(4.5)	40.9	(4.5)	1.8	(1.0)
beide Elternteile im Ausland geboren	40.7	(1.2)	13.3	(0.3)	35.5	(3.5)	61.3	(3.5)	3.2	(1.3)
Rheinland-Pfalz										
ohne Zuwanderungshintergrund	50.2	(0.7)	14.4	(0.1)	96.4	(0.7)	3.4	(0.7)	0.2	(0.1)
ein Elternteil im Ausland geboren	46.8	(1.5)	14.0	(0.3)	66.2	(3.3)	33.2	(3.3)	0.6	(0.6)
beide Elternteile im Ausland geboren	41.6	(1.1)	13.1	(0.3)	46.7	(3.9)	50.9	(4.0)	2.4	(1.0)
Saarland										
ohne Zuwanderungshintergrund	48.8	(0.8)	14.3	(0.1)	97.1	(0.5)	2.7	(0.5)	0.1	(0.1)
ein Elternteil im Ausland geboren	42.8	(1.5)	13.9	(0.3)	64.5	(3.5)	35.5	(3.5)	0.0	(0.0)
beide Elternteile im Ausland geboren	38.2	(1.6)	12.6	(0.3)	32.1	(3.7)	65.7	(3.8)	2.2	(1.1)
Sachsen										
ohne Zuwanderungshintergrund	47.4	(0.6)	14.1	(0.1)	96.5	(0.7)	3.4	(0.7)	0.1	(0.1)
ein Elternteil im Ausland geboren	--	--	--	--	--	--	--	--	--	--
beide Elternteile im Ausland geboren	--	--	--	--	--	--	--	--	--	--
Sachsen-Anhalt										
ohne Zuwanderungshintergrund	45.5	(0.5)	13.9	(0.1)	96.7	(0.5)	3.0	(0.5)	0.3	(0.2)
ein Elternteil im Ausland geboren	--	--	--	--	--	--	--	--	--	--
beide Elternteile im Ausland geboren	--	--	--	--	--	--	--	--	--	--
Schleswig-Holstein										
ohne Zuwanderungshintergrund	49.9	(0.8)	14.3	(0.2)	97.0	(0.5)	3.0	(0.5)	0.0	(0.0)
ein Elternteil im Ausland geboren	--	--	--	--	--	--	--	--	--	--
beide Elternteile im Ausland geboren	--	--	--	--	--	--	--	--	--	--
Thüringen										
ohne Zuwanderungshintergrund	46.1	(0.8)	13.7	(0.1)	97.0	(0.5)	2.8	(0.5)	0.2	(0.1)
ein Elternteil im Ausland geboren	--	--	--	--	--	--	--	--	--	--
beide Elternteile im Ausland geboren	--	--	--	--	--	--	--	--	--	--
Großstädte										
ohne Zuwanderungshintergrund	53.0	(1.3)	14.8	(0.2)	94.4	(1.5)	4.4	(1.3)	1.2	(1.0)
ein Elternteil im Ausland geboren	48.2	(2.9)	14.4	(0.5)	48.9	(6.4)	49.3	(6.2)	1.9	(1.9)
beide Elternteile im Ausland geboren	43.2	(1.9)	14.0	(0.3)	27.5	(3.6)	63.8	(3.3)	8.6	(2.5)
Deutschland										
ohne Zuwanderungshintergrund	50.8	(0.3)	14.5	(0.1)	96.4	(0.2)	3.5	(0.2)	0.2	(0.1)
ein Elternteil im Ausland geboren	47.5	(0.7)	14.3	(0.1)	61.6	(1.7)	37.3	(1.7)	1.1	(0.3)
beide Elternteile im Ausland geboren	40.4	(0.5)	13.3	(0.1)	34.1	(1.3)	61.3	(1.3)	4.6	(0.6)

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. Gültige %: Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind. fett: signifikante Differenz ($p < .05$) zu Schülerinnen und Schülern ohne Zuwanderungshintergrund. *M* = Mittelwert; *SE* = Standardfehler. ¹ In Ausbildungsjahren (vgl. OECD, 2009). ² Insgesamt 20 Schülerinnen und Schüler ohne Zuwanderungshintergrund gaben an, „nie Deutsch“ in der Familie zu sprechen. Dies kann unter anderem darin begründet sein, dass Schülerinnen und Schüler der dritten Generation weiterhin die Herkunftssprache ihrer Familie sprechen.

und Schülerinnen und Schülern ohne Zuwanderungshintergrund sind in den Ländern insgesamt relativ ähnlich ausgeprägt.

Im Vergleich zum sozioökonomischen Status zeigen sich in allen Ländern eher geringe Disparitäten im elterlichen Bildungsniveau. Dennoch bestehen relativ zu Familien, in denen beide Elternteile in Deutschland geboren sind, in allen Ländern signifikante Nachteile im Bildungsniveau für Familien, in denen beide Elternteile zugewandert sind. Diese Gruppenunterschiede sind ebenfalls über die Länder hinweg relativ ähnlich. In den Ländern Berlin, Bremen und Hessen weisen auch Familien mit einem im Ausland geborenen Elternteil ein signifikant niedrigeres elterliches Bildungsniveau auf als Familien ohne Zuwanderungshintergrund.

Im Hinblick auf die Familiensprache Deutsch zeigen sich deutliche Unterschiede zwischen Kindern mit einem im Ausland geborenen Elternteil und Kindern mit zwei im Ausland geborenen Elternteilen. Während in Deutschland insgesamt in fast 62 Prozent der Familien mit nur einem im Ausland geborenen Elternteil „immer Deutsch“ gesprochen wird, liegt der entsprechende Anteil in Familien mit zwei im Ausland geborenen Elternteilen bei nur 34 Prozent. Auch dieses Muster ist über die Länder hinweg relativ ähnlich, wobei Berlin eine Ausnahme bildet. Dort geben Kinder aus Familien mit zwei im Ausland geborenen Elternteilen seltener als in anderen Ländern an, „immer Deutsch“ in der Familie zu sprechen (18.4 %).

Vergleichsweise niedrig ist selbst in der Gruppe der Schülerinnen und Schüler mit zwei im Ausland geborenen Elternteilen der Anteil derjenigen Familien, die in der Familie „nie Deutsch“ sprechen (4.6 %).

Insgesamt ist das Muster der Unterschiede in den untersuchten Hintergrundmerkmalen zwischen Schülerinnen und Schülern mit unterschiedlichem Zuwanderungshintergrund und Schülerinnen und Schülern ohne Zuwanderungshintergrund über die Länder hinweg weitgehend ähnlich. Die Länderunterschiede in den mit Zuwanderung verbundenen Disparitäten in den erreichten Kompetenzen dürften also nur bedingt auf Länderunterschiede in der sozialen Lage, im Bildungsniveau der Eltern oder in der Familiensprache der Zuwanderergruppen zurückgehen. Inwieweit sich die Disparitäten insgesamt auf diese Merkmale zurückführen lassen, wird in Abschnitt 9.3.2 mit Hilfe von Regressionsanalysen untersucht.

9.3 Befunde für Kinder mit Zuwanderungshintergrund unterschiedlicher Herkunft in Deutschland

Frühere Untersuchungen haben teilweise erhebliche Unterschiede in den erreichten Kompetenzen bei Kindern unterschiedlicher Herkunftsgruppen festgestellt (vgl. z. B. Ditton & Krüsken, 2006; Segeritz et al., 2010; Stanat, Rauch & Segeritz, 2010). Entsprechend sollen auch in diesem Kapitel die Kompetenzstände von verschiedenen Herkunftsgruppen differenziert analysiert werden, um Hinweise darauf zu erhalten, bei welchen Kindern besonderer Förderbedarf besteht. Da die Stichprobengrößen der einzelnen Teilgruppen in den einzelnen Ländern oft zu gering sind, werden die Berechnungen nur für die Bundesrepublik Deutschland insgesamt durchgeführt. Ferner wird untersucht, inwieweit die identifizierten Kompetenznachteile der Herkunftsgruppen auf Unterschiede in Merkmalen des familiären Hintergrunds zurückgeführt werden können.

Tabelle 9.3: Prozentuale Anteile und Kompetenzstände der Viertklässlerinnen und Viertklässler unterschiedlicher Herkunftsgruppen in der Bundesrepublik Deutschland

Herkunftsgruppe	Deutsch									Mathematik				
	<i>N</i>	%	Lesen			Zuhören			<i>N</i>	%				
			<i>M</i>	(<i>SE</i>)	<i>SD</i>	<i>M</i>	(<i>SE</i>)	<i>SD</i>			<i>M</i>	(<i>SE</i>)	<i>SD</i>	
ohne Zuwanderungshintergrund	18 546	68.8	514	(1.6)	96	518	(1.8)	95	18 548	69.0	515	(1.5)	96	
Türkei														
ein Elternteil im Ausland geboren	509	2.0	443	(7.3)	93	433	(8.6)	96	512	2.1	448	(6.7)	89	
beide Elternteile im Ausland geboren	811	3.3	421	(5.1)	91	408	(5.5)	91	807	3.4	423	(5.0)	86	
ehemalige Sowjetunion														
ein Elternteil im Ausland geboren	153	0.6	482	(11.9)	105	485	(13.7)	114	151	0.6	472	(11.3)	95	
beide Elternteile im Ausland geboren	933	3.8	481	(4.9)	91	472	(5.3)	90	936	3.8	485	(4.5)	90	
Polen														
ein Elternteil im Ausland geboren	242	1.0	494	(10.7)	93	493	(11.1)	94	246	1.0	494	(10.5)	99	
beide Elternteile im Ausland geboren	223	1.1	484	(10.3)	102	468	(9.5)	96	222	1.0	490	(11.3)	102	
ehemaliges Jugoslawien														
ein Elternteil im Ausland geboren	198	1.0	481	(10.8)	106	479	(11.3)	103	192	1.0	478	(10.7)	101	
beide Elternteile im Ausland geboren	255	1.1	445	(8.6)	89	435	(8.3)	91	256	1.1	448	(9.7)	100	
anderes Land														
ein Elternteil im Ausland geboren	1 156	4.7	510	(5.0)	94	507	(5.3)	94	1 159	4.7	502	(5.1)	95	
beide Elternteile im Ausland geboren	1 048	4.0	471	(6.5)	101	455	(5.9)	98	1 041	4.1	466	(5.8)	103	
nicht zuzuordnen	2 448	8.5	460	(4.1)	102	457	(4.1)	102	2 413	8.4	457	(4.2)	104	

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. fett: signifikante Differenz ($p < .05$) zu Schülerinnen und Schüler ohne Zuwanderungshintergrund. *N* = Fallzahl; *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung.

In Tabelle 9.3 sind zusammenfassend die Ergebnisse für Kinder, deren Familien aus unterschiedlichen Ländern zugewandert sind, dargestellt. Neben den Anteilen der Schülerinnen und Schüler unterschiedlicher Herkunftsgruppen sind die Mittelwerte (*M*), Standardfehler (*SE*) und Standardabweichungen (*SD*) der Kompetenzwerte aufgeführt.

Insgesamt kommen die Familien von etwa 5 Prozent der Kinder aus der Türkei, von etwa 4 Prozent aus der ehemaligen Sowjetunion sowie von jeweils etwa 2 Prozent aus Polen und dem ehemaligen Jugoslawien. Die größte Gruppe bilden jedoch mit einem Anteil von knapp 9 Prozent der Kinder solche Familien, die aus einem anderen Herkunftsland stammen. Dies verweist nochmals auf die erhebliche Heterogenität der Menschen mit Zuwanderungshintergrund in Deutschland, die aufgrund der starken Fokussierung der öffentlichen Diskussion auf einzelne Herkunftsgruppen häufig übersehen wird (vgl. auch Stanat, Rauch & Segeritz, 2010). Knapp 9 Prozent der Kinder konnten aufgrund fehlender Angaben zum Zuwanderungsstatus keiner der Gruppen zugeordnet werden.

In den im Mittel erreichten Kompetenzen sind deutliche Unterschiede zwischen Kindern unterschiedlicher Herkunftsländer zu verzeichnen. Innerhalb der Gruppe der Kinder mit Zuwanderungshintergrund erreichen Schülerinnen und Schüler türkischer Herkunft die niedrigsten Werte in den getesteten Kompetenzbereichen. Ihr Mittelwert unterscheidet sich statistisch signifikant von den Mittelwerten aller anderen Herkunftsgruppen. Auch die Kompetenzen der Schülerinnen und Schüler, für die keine ausreichenden Angaben zum Zuwanderungshintergrund vorliegen, liegen deutlich unter dem Mittelwert der Schülerinnen und Schüler deutscher Herkunft. Kinder polnischer Herkunft weisen im Lesen, im Zuhören und in Mathematik die höchsten Kompetenzstände aller betrachteten Zuwanderungsgruppen auf.

9.3.1 Merkmale des familiären Hintergrunds von Viertklässlerinnen und Viertklässlern unterschiedlicher Herkunftsgruppen

Auch für die verschiedenen Herkunftsgruppen soll der Frage nachgegangen werden, inwieweit sie sich in Bezug auf familiäre Hintergrundmerkmale voneinander unterscheiden. Der mittlere sozioökonomische Status, das mittlere Bildungsniveau der Eltern sowie die prozentualen Anteile der Häufigkeit, mit der Deutsch in der Familie gesprochen wird, sind in Tabelle 9.4 dargestellt.

In Tabelle 9.4 ist deutlich zu erkennen, dass Familien, in denen beide Elternteile zugewandert sind, in allen untersuchten Herkunftsgruppen über einen deutlich niedrigeren sozioökonomischen Status verfügen als Familien ohne Zuwanderungsgeschichte. Auch Familien, in denen nur ein Elternteil in der Türkei oder im ehemaligen Jugoslawien geboren ist, weisen einen besonders geringen sozioökonomischen Status auf. Ein ähnliches Muster ist für das Bildungsniveau der Eltern zu finden. Für Familien polnischer Herkunft sind hier jedoch keine signifikanten Unterschiede zu Familien ohne Zuwanderungshintergrund zu beobachten.

In Bezug auf die Familiensprache zeigt sich, dass Kinder mit einem im Ausland geborenen Elternteil in den meisten Herkunftsgruppen überwiegend angeben, zu Hause „immer Deutsch“ zu sprechen, wohingegen Kinder mit zwei im Ausland geborenen Elternteilen überwiegend angeben, zu Hause „manchmal Deutsch“ zu sprechen. In Familien, in denen ein Elternteil in der Türkei geboren ist, wird zu Hause überwiegend „manchmal Deutsch“ gesprochen. In allen untersuchten Herkunftsgruppen liegt der Anteil der Kinder, die zu Hause „nie Deutsch“ sprechen, bei unter 10 Prozent. Bei den untersuchten Merkmalen des familiären Hintergrunds sind durchgängig die Unterschiede für aus der Türkei zugewanderte Familien am größten.

Tabelle 9.4: Merkmale des familiären Hintergrunds von Viertklässlerinnen und Viertklässlern unterschiedlicher Herkunftsgruppen

Herkunftsgruppe	sozioökonomischer Status		Bildungsniveau der Eltern ¹		Sprachgebrauch in der Familie					
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)	immer Deutsch		manchmal Deutsch		nie Deutsch	
					gültige %	(<i>SE</i>)	gültige %	(<i>SE</i>)	gültige %	(<i>SE</i>)
ohne Zuwanderungshintergrund	50.8	(0.3)	14.5	(0.1)	96.4	(0.2)	3.5	(0.2)	0.2	(0.1)
Türkei										
ein Elternteil im Ausland geboren	39.3	(1.2)	12.8	(0.2)	42.7	(4.0)	55.3	(4.1)	2.0	(1.2)
beide Elternteile im Ausland geboren	36.8	(0.8)	11.8	(0.2)	27.0	(2.2)	70.1	(2.5)	3.0	(0.9)
ehemalige Sowjetunion										
ein Elternteil im Ausland geboren	49.6	(2.2)	15.0	(0.4)	66.6	(5.7)	33.0	(5.8)	0.4	(0.3)
beide Elternteile im Ausland geboren	41.6	(0.7)	14.0	(0.1)	42.8	(2.3)	52.7	(2.4)	4.5	(1.0)
Polen										
ein Elternteil im Ausland geboren	49.1	(1.6)	14.6	(0.3)	76.1	(3.9)	23.0	(3.9)	0.9	(0.7)
beide Elternteile im Ausland geboren	41.2	(1.3)	14.2	(0.3)	31.8	(4.6)	62.5	(4.6)	5.7	(2.0)
ehemaliges Jugoslawien										
ein Elternteil im Ausland geboren	43.4	(1.4)	13.7	(0.3)	59.8	(5.1)	38.1	(5.2)	2.1	(1.6)
beide Elternteile im Ausland geboren	38.8	(1.3)	12.7	(0.4)	27.2	(4.3)	66.2	(4.3)	6.6	(2.2)
anderes Land										
ein Elternteil im Ausland geboren	51.4	(0.9)	14.9	(0.2)	66.4	(2.5)	33.0	(2.4)	0.6	(0.3)
beide Elternteile im Ausland geboren	42.6	(1.0)	13.7	(0.2)	34.6	(2.2)	60.2	(2.3)	5.3	(1.0)
nicht zuzuordnen	44.2	(0.8)	13.5	(0.2)	67.9	(2.2)	30.8	(2.2)	1.3	(0.6)

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. Gültige %: Prozentangaben beruhen nur auf Angaben derjenigen Schülerinnen und Schüler, für die gültige Werte vorliegen. fett: signifikante Differenz ($p < .05$) zu Schülerinnen und Schülern ohne Zuwanderungshintergrund. *M* = Mittelwert; *SE* = Standardfehler. ¹ In Ausbildungsjahren (vgl. OECD, 2009).

9.3.2 Zusammenhänge zwischen zugewanderungsbezogenen Disparitäten und familiären Hintergrundmerkmalen

In diesem Abschnitt soll mit einem multivariaten Regressionsansatz abschließend untersucht werden, inwieweit sich die Kompetenznachteile von Kindern mit Zuwanderungshintergrund auf die beschriebenen familiären Hintergrundmerkmale zurückführen lassen. In die Analysen wurden nur diejenigen Kinder einbezogen, für die vollständige Daten zu allen Variablen vorlagen, die in die Regressionsmodelle eingehen. Aufgrund geringer Fallzahlen in den einzelnen Gruppen wurden auch hier die erste und zweite Generation zur Gruppe „beide Eltern im Ausland geboren“ zusammengefasst. In den Tabellen 9.5 bis 9.7 sind die Befunde für die Bereiche Lesen und Zuhören sowie für das Fach Mathematik dargestellt.

Tabelle 9.5: Regressionsmodelle zur Schätzung von Disparitäten im Bereich Lesen zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund

	Modell 1		Modell 2		Modell 3	
	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)
ohne Zuwanderungshintergrund	517	(1.6)	512	(1.5)	513	(1.5)
Türkei						
ein Elternteil im Ausland geboren	-66	(8.5)	-40	(8.4)	-32	(8.7)
beide Elternteile im Ausland geboren	-89	(6.7)	-55	(6.8)	-45	(7.0)
ehemalige Sowjetunion						
ein Elternteil im Ausland geboren	-31	(12.3)	-32	(12.3)	-27	(12.2)
beide Elternteile im Ausland geboren	-29	(5.5)	-12	(5.4)	-5	(5.6)
Polen						
ein Elternteil im Ausland geboren	-17	(11.3)	-17	(11.1)	-14	(11.3)
beide Elternteile im Ausland geboren	-27	(11.4)	-10	(11.2)	-1	(11.5)
ehemaliges Jugoslawien						
ein Elternteil im Ausland geboren	-34	(11.8)	-19	(11.5)	-14	(11.6)
beide Elternteile im Ausland geboren	-68	(10.3)	-41	(10.0)	-32	(10.0)
anderes Land						
ein Elternteil im Ausland geboren	-2	(5.5)	-5	(5.0)	-1	(5.2)
beide Elternteile im Ausland geboren	-34	(7.7)	-19	(6.4)	-11	(6.9)
nicht zuzuordnen	-45	(7.2)	-24	(7.0)	-21	(7.1)
sozialer Hintergrund						
HISEI ¹			25	(1.4)	25	(1.4)
Bildungsniveau der Eltern ¹			12	(1.5)	12	(1.5)
Familiensprache²						
manchmal Deutsch					-15	(4.0)
nie Deutsch					-10	(12.4)
<i>N</i>	21058		21058		21058	
<i>R</i> ²	.04		.15		.15	

Anmerkungen. Die Referenzgruppe sind Schülerinnen und Schüler ohne Zuwanderungshintergrund.

fett: signifikante Regressionskoeffizienten ($p < .05$). *b* = unstandardisierter Regressionskoeffizient; *SE* = Standardfehler.

¹ z-standardisiert. ² Referenzgruppe: immer Deutsch als Familiensprache.

Den Ergebnissen der deskriptiven Analysen entsprechend, zeigt sich für den Kompetenzbereich Lesen in Modell 1, dass die größten Kompetenznachteile für Kinder bestehen, deren Familien aus der Türkei zugewandert sind (vgl. Tabelle 9.5). Diese Differenz ist bei Kindern mit zwei im Ausland geborenen Elternteilen besonders ausgeprägt, aber auch bei Kindern mit einem im Ausland geborenen Elternteil noch erheblich.

Für Kinder, deren Eltern beide in der ehemaligen Sowjetunion, im ehemaligen Jugoslawien oder in einem anderen Land geboren sind, sind ebenfalls signifikante Kompetenznachteile zu verzeichnen. Hingegen weisen Kinder mit nur einem aus diesen Ländern zugewanderten Elternteil keine signifikanten Kompetenznachteile auf. Dasselbe gilt für Kinder, deren Familien aus Polen zugewandert sind, und zwar unabhängig davon, ob nur ein Elternteil oder ob beide Elternteile aus Polen stammen.

Werden in der Analyse das Bildungsniveau der Eltern und der sozioökonomische Status der Familien statistisch kontrolliert (Modell 2 in Tabelle 9.5), reduzieren sich die Disparitäten insbesondere bei Kindern mit zwei im Ausland geborenen Elternteilen in allen Gruppen deutlich. Für türkischstämmige Schülerinnen

Tabelle 9.6: Regressionsmodelle zur Schätzung von Disparitäten im Bereich Zuhören zwischen Kindern mit unterschiedlichem Zuwanderungshintergrund

	Modell 1		Modell 2		Modell 3	
	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)
ohne Zuwanderungshintergrund	520	(1.8)	515	(1.7)	516	(1.6)
Türkei						
ein Elternteil im Ausland geboren	-80	(9.8)	-54	(9.8)	-44	(9.8)
beide Elternteile im Ausland geboren	-110	(6.6)	-75	(6.7)	-62	(6.8)
ehemalige Sowjetunion						
ein Elternteil im Ausland geboren	-29	(10.9)	-30	(11.0)	-24	(10.8)
beide Elternteile im Ausland geboren	-43	(5.4)	-27	(5.2)	-17	(5.4)
Polen						
ein Elternteil im Ausland geboren	-20	(11.8)	-20	(11.8)	-16	(12.0)
beide Elternteile im Ausland geboren	-46	(10.8)	-30	(10.7)	-18	(11.0)
ehemaliges Jugoslawien						
ein Elternteil im Ausland geboren	-40	(12.2)	-25	(11.8)	-18	(12.1)
beide Elternteile im Ausland geboren	-80	(9.8)	-53	(9.3)	-40	(9.8)
anderes Land						
ein Elternteil im Ausland geboren	-7	(5.6)	-11	(5.2)	-5	(5.2)
beide Elternteile im Ausland geboren	-57	(6.8)	-42	(5.8)	-30	(6.2)
nicht zuzuordnen	-50	(7.4)	-30	(7.1)	-25	(7.3)
sozialer Hintergrund						
HISEI ¹			23	(1.3)	23	(1.3)
Bildungsniveau der Eltern ¹			15	(1.4)	15	(1.4)
Familiensprache²						
manchmal Deutsch					-20	(3.5)
nie Deutsch					-14	(13.7)
<i>N</i>	21 058		21 058		21 058	
<i>R</i> ²	.07		.17		.17	

Anmerkungen. Die Referenzgruppe sind Schülerinnen und Schüler ohne Zuwanderungshintergrund. fett: signifikante Regressionskoeffizienten ($p < .05$). *b* = unstandardisierter Regressionskoeffizient; SE = Standardfehler. ¹ z-standardisiert. ² Referenzgruppe: immer Deutsch als Familiensprache.

und Schüler sowie für Kinder, deren Eltern aus dem ehemaligen Jugoslawien stammen, sind sie jedoch weiterhin groß und statistisch bedeutsam.

Kontrolliert man zusätzlich noch die in der Familie gesprochene Sprache (Modell 3 in Tabelle 9.5), sind nur noch die Unterschiede für türkischstämmige Kinder, deren Eltern beide in der Türkei geboren sind, statistisch signifikant. Die Kompetenznachteile einiger anderer Teilgruppen sind ebenfalls groß, sie lassen sich jedoch aufgrund der geringen Fallzahlen nicht gegen den Zufall absichern.

Im Bereich Zuhören sind die Disparitäten in allen Gruppen etwas stärker ausgeprägt als im Bereich Lesen (vgl. Tabelle 9.6). Auch hier sind die Nachteile der türkischstämmigen Kinder am größten und selbst nach Kontrolle aller einbezogenen Merkmale des familiären Hintergrunds noch statistisch signifikant. Demnach lassen sich die Nachteile türkischstämmiger Kinder im Bereich Zuhören nicht allein auf die hier analysierten familiären Merkmale zurückführen. Der in Modell 3 für Kinder mit zwei in der Türkei geborenen Elternteilen geschätzte Kompetenznachteil von 63 Punkten entspricht dem erwarteten Lernzuwachs von etwa einem Schuljahr.

Tabelle 9.7: Regressionsmodelle zur Schätzung von Disparitäten im Fach Mathematik (Globalskala) zwischen Kindern mit unterschiedlichem Zuwanderungshintergrund

	Modell 1		Modell 2		Modell 3	
	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)
ohne Zuwanderungshintergrund	518	(1.6)	513	(1.4)	514	(1.4)
Türkei						
ein Elternteil im Ausland geboren	-65	(7.8)	-39	(7.3)	-29	(7.7)
beide Elternteile im Ausland geboren	-89	(6.5)	-56	(6.3)	-43	(6.8)
ehemalige Sowjetunion						
ein Elternteil im Ausland geboren	-42	(10.6)	-42	(10.9)	-36	(10.8)
beide Elternteile im Ausland geboren	-27	(4.6)	-11	(4.4)	-1	(5.0)
Polen						
ein Elternteil im Ausland geboren	-19	(12.3)	-19	(12.2)	-15	(12.4)
beide Elternteile im Ausland geboren	-18	(11.2)	-2	(10.8)	9	(11.5)
ehemaliges Jugoslawien						
ein Elternteil im Ausland geboren	-35	(11.6)	-21	(11.3)	-15	(11.7)
beide Elternteile im Ausland geboren	-63	(11.2)	-37	(10.1)	-24	(9.9)
anderes Land						
ein Elternteil im Ausland geboren	-11	(5.1)	-14	(4.6)	-9	(4.7)
beide Elternteile im Ausland geboren	-41	(7.0)	-27	(5.9)	-15	(6.5)
nicht zuzuordnen	-40	(7.1)	-20	(6.8)	-15	(6.9)
sozialer Hintergrund						
HISEI ¹			22	(1.4)	22	(1.4)
Bildungsniveau der Eltern ¹			15	(1.3)	15	(1.3)
Familiensprache²						
manchmal Deutsch					-19	(3.6)
nie Deutsch					-18	(14.8)
<i>N</i>	21 057		21 057		21 057	
<i>R</i> ²	.04		.14		.15	

Anmerkungen. Die Referenzgruppe sind Schülerinnen und Schüler ohne Zuwanderungshintergrund.

fett: signifikante Regressionskoeffizienten ($p < .05$). *b* = unstandardisierter Regressionskoeffizient; *SE* = Standardfehler.

¹ z-standardisiert. ² Referenzgruppe: immer Deutsch als Familiensprache.

Für Kinder, deren Eltern beide aus der ehemaligen Sowjetunion zugewandert sind, zeigen sich im Zuhören signifikante Kompetenznachteile, die auch nach Kontrolle des sozioökonomischen Status und des Bildungsniveaus der Eltern bestehen bleiben. Diese verschwinden jedoch, wenn zusätzlich die Familiensprache berücksichtigt wird (Modell 3 in Tabelle 9.6). Anders als im Bereich Lesen bestehen auch für Schülerinnen und Schüler, deren Eltern beide in Polen geboren sind, im Bereich Zuhören signifikante Kompetenznachteile, die sich jedoch nach Kontrolle der in die Analyse einbezogenen Merkmale des familiären Hintergrunds deutlich reduzieren.

Die Kompetenznachteile im Fach Mathematik sind etwas größer als im Kompetenzbereich Lesen, allerdings etwas geringer als im Zuhören (vgl. Tabelle 9.7). Auch in Mathematik ergeben sich die größten Nachteile für Kinder türkischer Herkunft. Diese reduzieren sich nach Kontrolle familiärer Hintergrundmerkmale um etwa die Hälfte, bleiben für Kinder mit zwei zugewanderten Elternteilen aber weiterhin statistisch signifikant. Alle weiteren zugewanderungsbezogenen

Disparitäten lassen sich in Mathematik auf die Merkmale des familiären Hintergrunds zurückführen.

In allen Regressionsanalysen sind die einbezogenen Merkmale des familiären Hintergrunds mit den jeweils untersuchten Kompetenzen assoziiert. Dies gilt durchgängig für das Bildungsniveau der Eltern und für den HISEI als Indikator des sozioökonomischen Status. Auch für die in der Familie gesprochene Sprache sind Zusammenhänge mit den erreichten Kompetenzen zu verzeichnen. Diese lassen sich jedoch nur zum Teil gegen den Zufall absichern. So erzielten Schülerinnen und Schüler, die in der Familie „manchmal Deutsch“ sprechen, im Zuhören 20 Punkte und in Mathematik 19 Punkte weniger als Schülerinnen und Schüler mit ausschließlich deutscher Familiensprache. Dies weist wiederum darauf hin, dass es erforderlich ist, Kindern nicht deutscher Herkunftssprache im Bildungssystem angemessene Lerngelegenheiten für den Erwerb der deutschen Sprache zur Verfügung zu stellen.

9.4 Zusammenfassung und Diskussion

In diesem Kapitel wurden die mit einem Zuwanderungshintergrund verbundenen Disparitäten in den Bereichen Lesen, Zuhören und Mathematik für Viertklässlerinnen und Viertklässler in den Ländern der Bundesrepublik Deutschland untersucht. Dabei ergaben sich für alle untersuchten Bereiche und in allen Ländern deutliche Nachteile für Schülerinnen und Schüler mit Zuwanderungshintergrund. Diese waren jeweils bei Kindern mit zwei im Ausland geborenen Elternteilen stärker ausgeprägt als bei Kindern mit einem im Ausland und einem in Deutschland geborenen Elternteil. Allgemein fielen die Disparitäten im Zuhören in allen Ländern größer aus als in den anderen untersuchten Bereichen.

Besonders ausgeprägte Disparitäten wurden in den Stadtstaaten Berlin, Bremen und Hamburg identifiziert. Vergleichsweise große Gruppenunterschiede waren außerdem im Bereich Lesen in Hessen, im Bereich Zuhören für Kinder mit zwei zugewanderten Elternteilen im Saarland sowie im Bereich Mathematik in Baden-Württemberg und Niedersachsen zu verzeichnen. Eher geringe Disparitäten zeigten sich für alle getesteten Bereiche in Rheinland-Pfalz sowie für Kinder mit einem zugewanderten Elternteil in Bayern und im Saarland. Für Kinder mit zwei zugewanderten Elternteilen wurden außerdem in Nordrhein-Westfalen für die Bereiche Lesen und Mathematik vergleichsweise geringe Gruppenunterschiede festgestellt.

Ähnliche Befundmuster zeigten sich für den Bereich Lesen bereits auch in IGLU-E 2006 (Schwippert et al., 2008). In IGLU-E 2006 waren ebenfalls in den Stadtstaaten und in den westdeutschen Flächenländern signifikante Unterschiede zwischen Kindern mit zwei zugewanderten Elternteilen und Kindern ohne Zuwanderungshintergrund zu verzeichnen. Während allerdings in IGLU-E 2006 nur in zwei Ländern signifikante Disparitäten zwischen Kindern mit einem zugewanderten Elternteil und Kindern ohne Zuwanderungshintergrund bestanden, ist dies im Ländervergleich 2011 in sieben Ländern der Fall. Dies könnte allerdings auf Unterschiede in der Höhe der Standardfehler zwischen IGLU-E und dem Ländervergleich 2011 zurückzuführen sein. Zudem lassen sich – wie bereits in Kapitel 5 erläutert – die erreichten Kompetenzstände der untersuchten Schülerinnen und Schüler des Ländervergleichs 2011 und die Ergebnisse von IGLU-E 2006 nur eingeschränkt vergleichen, da sich die eingesetzten Tests unterscheiden. Zu den Bereichen Zuhören und Mathematik lagen für die Primarstufe

vor dem IQB-Ländervergleich 2011 noch keine ländervergleichenden Analysen vor.

Im zweiten Teil dieses Kapitels wurden die Kompetenzstände verschiedener Herkunftsgruppen differenziert untersucht. Dabei ergaben sich erhebliche Gruppenunterschiede, wobei türkischstämmige Kinder im Mittel die geringsten Kompetenzstände erzielten. Ähnlich ausgeprägte Disparitäten wurden für die Sekundarstufe I im Bereich Lesen in PISA 2009 (Stanat, Rauch & Segeritz, 2010) sowie in den Bereichen Lesen und Zuhören im IQB-Ländervergleich sprachlicher Kompetenzen 2009 (Böhme et al., 2010) festgestellt. Die vorliegenden Analysen weisen darauf hin, dass die für die Sekundarstufe I beschriebenen zuwanderungsbezogenen Kompetenznachteile bereits in der Grundschule bestehen.

Die Ergebnisse von Regressionsanalysen zeigten, dass sich die zuwanderungsbezogenen Disparitäten teilweise auf den sozialen Status, das Bildungsniveau der Eltern und die Häufigkeit, mit der in der Familie Deutsch gesprochen wird, zurückführen lassen. Im Bereich Zuhören bestanden jedoch auch nach Kontrolle des familiären Hintergrunds für Kinder mit zwei zugewanderten Elternteilen in den meisten Herkunftsgruppen noch substanzielle Kompetenznachteile. Ähnliche Zusammenhänge zwischen zuwanderungsbezogenen Disparitäten und Merkmalen des familiären Hintergrunds ergaben sich für den Bereich Lesen in IGLU-E 2006 und für Mathematik in TIMSS 2007, wobei jedoch die Herkunftsländer der Kinder mit Zuwanderungsgeschichte in diesen Analysen nicht berücksichtigt wurden.

Nach Kontrolle des sozioökonomischen Status der Familien und des Bildungshintergrunds der Eltern reduzierten sich die Kompetenznachteile der Schülerinnen und Schüler mit Zuwanderungshintergrund der verschiedenen Herkunftsgruppen in den vorliegenden Analysen deutlich. Eine weitere Reduktion der Disparitäten war mit der zu Hause gesprochenen Sprache verbunden. Selbst nach Kontrolle dieser Merkmale des familiären Hintergrunds ergaben sich jedoch für Schülerinnen und Schüler türkischer Herkunft deutliche Kompetenznachteile. Wie bereits an anderer Stelle ausführlich diskutiert (Stanat, Rauch & Segeritz, 2010), ist unklar, wie dieser Befund, der auch in anderen Studien gefunden wurde (vgl. zusammenfassend Dollmann, 2010), zu erklären ist. An mangelnder Motivation kann es nicht liegen, da gerade auch Familien türkischer Herkunft ausgesprochen hohe Bildungsaspirationen für ihre Kinder haben (vgl. z. B. Stanat, Segeritz & Christensen, 2010). Ihre Aspirationen sind allerdings zum Teil so hoch, dass es nur sehr schwer möglich sein dürfte, sie einzulösen. Dies könnte einen Mangel an Information der Eltern darüber widerspiegeln, welche Voraussetzungen erfüllt sein müssen, um bestimmte Ziele im deutschen Bildungssystem zu erreichen (Kristen, 2002). Ein solcher Informationsmangel wiederum dürfte es ihnen erschweren, die Bildungslaufbahn ihrer Kinder zielgerichtet zu unterstützen.

Aber auch mangelnde Wertschätzung und Akzeptanz türkischstämmiger Zuwanderer in Deutschland könnten dazu beitragen, dass ihre Kinder weniger erfolgreich sind als Schülerinnen und Schüler ohne Zuwanderungshintergrund (Stanat, Rauch & Segeritz, 2010). Wie Studien in der Sozialpsychologie wiederholt gezeigt haben, sind gesellschaftlich verbreitete Stereotype sehr mächtige Einflussfaktoren, die die Wahrnehmung und das Verhalten von Menschen auch dann lenken können, wenn sie die stereotypen Annahmen nicht teilen (Nelson, 2002). Es kann daher nicht ausgeschlossen werden, dass negative Stereotype über türkischstämmige Zuwanderer, wie sie in aktuellen Debatten über Zuwanderung und Integration teilweise deutlich zu erkennen sind, dazu beitragen, dass es

Kinder und Jugendliche dieser Herkunftsgruppe schwerer haben, im deutschen Bildungssystem erfolgreich zu sein. Die Frage, inwieweit diese oder andere Faktoren für den beschriebenen Befund verantwortlich sind, muss jedoch an dieser Stelle aufgrund fehlender empirischer Studien offen bleiben.

Aber auch die Lage der Schülerinnen und Schüler aus anderen Herkunftsgruppen kann nicht zufrieden stellen, wenn ihre Kompetenznachteile nach Kontrolle des sozioökonomischen Hintergrunds und des Bildungsniveaus ihrer Eltern verschwinden. Dies bedeutet lediglich, dass Kinder dieser Herkunftsgruppen im Bildungssystem ebenso wenig erfolgreich sind wie ihre Mitschülerinnen und Mitschüler aus sozial schwachen Familien ohne Zuwanderungshintergrund, die ebenfalls in hohem Maße benachteiligt sind (vgl. Kapitel 8). Allgemeines Ziel muss daher sein, die Lerngelegenheiten für alle Schülerinnen und Schülern so zu gestalten, dass sie unbedingt die Mindeststandards, möglichst aber auch die Regelstandards erreichen und den Übergang zunächst von der Primarstufe in die Sekundarstufe sowie anschließend in die Berufsausbildung oder die Hochschule erfolgreich bewältigen. Dies erfordert unter anderem eine effektive Förderung sprachlicher Kompetenzen, die bereits im vorschulischen Bereich einsetzen und in der Schule fortgeführt werden muss.

Literatur

- Alba, R. D., Handl, J. & Müller, W. (1994). Ethnische Ungleichheit im deutschen Bildungssystem. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 46, 209–237.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2002). *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich*. Opladen: Leske + Budrich.
- Baumert, J., Bos, W. & Lehmann, R. H. (Hrsg.). (2000). *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn (Bd. 1., Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit)*. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillman, K.-J. & Weiß, M. (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–410). Opladen: Leske + Budrich.
- Behrens, U., Böhme, K. & Krelle, M. (2009). Zuhören – Operationalisierung und fachdidaktische Implikationen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 357–375). Weinheim: Beltz.
- Böhme, K., Tiffin-Richards, S. P., Schipolowski, S. & Leucht, M. (2010). Migrationsbedingte Disparitäten bei sprachlichen Kompetenzen. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 203–225). Münster: Waxmann.
- Bonsen, M., Kummer, N. & Bos, W. (2008). Schülerinnen und Schüler mit Migrationshintergrund. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich* (S. 156–175). Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M. & Schwippert, K. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bremerich-Vos, A. & Böhme, K. (2009). Lesekompetenzdiagnostik – die Entwicklung eines Kompetenzstufenmodells für den Bereich Lesen. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.),

- Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 219–249). Weinheim: Beltz.
- Diefenbach, H. (2002). Bildungsbeteiligung und Berufseinmündung von Kindern und Jugendlichen aus Migrantenfamilien: Eine Fortschreibung der Daten des Sozio-ökonomischen Panels. In Sachverständigenkommission Elfter Kinder- und Jugendbericht (Hrsg.), *Migration und die europäische Integration: Herausforderungen für Kinder- und Jugendhilfe* (S. 9–70). München: Deutsches Jugendinstitut.
- Ditton, H. & Krüsken, J. (2006). Der Übergang von der Grundschule in die Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 9, 348–372.
- Dollmann, J. (2010). *Türkischstämmige Kinder am ersten Bildungsübergang. Primäre und sekundäre Herkunftseffekte*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dubowy, M., Ebert, S., von Maurice, J. & Weinert, S. (2008). Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40, 124–134.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 375–397). Münster: Waxmann.
- Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D., J. & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1–56.
- Gomolla, M. & Radtke, F.-O. (2002). *Institutionelle Diskriminierung: Die Herstellung ethnischer Differenz in der Schule*. Opladen: Leske + Budrich.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576.
- Kristen, C. (2002). Hauptschule, Realschule oder Gymnasium? Ethnische Unterschiede am ersten Bildungsübergang. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 54, 534–552.
- Kristen, C. (2008). Schulische Leistungen von Kindern aus türkischen Familien am Ende der Grundschulzeit. Befunde aus der IGLU-Studie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 48, 230–251.
- Maaz, K. (2006). *Soziale Herkunft und Hochschulzugang. Effekte institutioneller Öffnung im Bildungssystem*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mengering, F. (2005). Bärenstark – Empirische Ergebnisse der Berliner Sprachstandserhebung an Kindern im Vorschulalter. *Zeitschrift für Erziehungswissenschaft*, 8, 241–262.
- Nelson, T. D. (2002). *The psychology of prejudice*. Needham Heights, MA: Allyn & Bacon.
- OECD. (1999). *Classifying educational programs. Manual for ISCED-97 implementation in OECD countries*. Paris: OECD.
- OECD. (2009). *PISA 2006 technical report*. Paris: OECD.
- Reiss, K. & Winkelmann, H. (2009). Kompetenzstufenmodelle für das Fach Mathematik im Primarbereich. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 120–141). Weinheim: Beltz.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schupp, J. (2009). 25 Jahre Sozio-oekonomisches Panel – Ein Infrastrukturprojekt der empirischen Sozial- und Wirtschaftsforschung in Deutschland. *Zeitschrift für Soziologie*, 38, 350–357.
- Schwippert, K., Bos, W. & Lankes, E.-M. (2003). Heterogenität und Chancengleichheit am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 265–302). Münster: Waxmann.
- Schwippert, K., Hornberg, S., Freiberg, M. & Stubbe, T. C. (2007). Lesekompetenz von Kindern mit Migrationshintergrund im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 249–269). Münster: Waxmann.
- Schwippert, K., Hornberg, S. & Goy, M. (2008). Lesekompetenzen von Kindern mit Migrationshintergrund im internationalen Vergleich. In W. Bos, S. Hornberg,

- K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 111–125). Münster: Waxmann.
- Segeritz, M., Walter, O. & Stanat, P. (2010). Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 113–138.
- Stanat, P. & Christensen, G. S. (2006). *Where immigrant students succeed: A comparative review of performance and engagement in PISA 2003*. Paris: OECD.
- Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 200–230). Münster: Waxmann.
- Stanat, P., Segeritz, M. & Christensen, G. S. (2010). Schulbezogene Motivation und Aspiration von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (S. 31–57). Münster: Waxmann.
- Statistisches Bundesamt. (2001). *Fachserie 11: Bildung und Kultur, Reihe 1: Allgemeinbildende Schulen, Schuljahr 1999/2000*. Stuttgart: Metzler-Poeschel.
- Walter, O. & Taskinen, P. (2007). Kompetenzen und bildungsrelevante Einstellungen von Jugendlichen mit Migrationshintergrund in Deutschland: Ein Vergleich mit OECD-Staaten. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 337–366). Münster: Waxmann.

Kapitel 10

Aspekte der Aus- und Fortbildung von Lehrkräften in der Primarstufe

Dirk Richter, Poldi Kuhl, Heino Reimers und Hans Anand Pant

Die professionelle Kompetenz von Lehrkräften gilt als eine zentrale Voraussetzung für gelingende Lehr-Lern-Prozesse im Unterricht und damit für den schulischen Lernerfolg von Schülerinnen und Schülern. In Deutschland widmeten sich in den vergangenen zehn Jahren mehrere Studien der Modellierung und Erfassung professioneller Kompetenz von Lehrkräften (Blömeke, Kaiser & Lehmann, 2010a, 2010b; Borowski et al., 2010; Kunter et al., 2011). Mit den Ergebnissen der COACTIV-Studie (*Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz*; Kunter et al., 2011) konnte empirisch gezeigt werden, dass sich Lehrkräfte in ihrem professionellen Wissen deutlich unterscheiden und dass vor allem das Ausmaß fachdidaktischen Wissens von Lehrkräften bedeutsam mit dem Grad der kognitiven Aktivierung und der individuellen Unterstützung im Unterricht zusammenhängt (Baumert et al., 2010). Diese beiden Merkmale guten Unterrichts wiederum sagen den beobachteten Lernzuwachs im Zeitraum eines Schuljahres vorher. Zudem zeigt eine zusammenfassende Auswertung vieler Studien (Meta-Analyse) zu den Determinanten von Schülerleistungen, dass die Art, wie Lehrkräfte Unterricht gestalten, zu den wichtigsten Faktoren im Lernprozess gehört (Hattie, 2009). Vor dem Hintergrund der Bedeutung, die Lehrkräften für den Lernerfolg von Schülerinnen und Schülern beigemessen werden kann, werden im Ländervergleich 2011 ausgewählte Ergebnisse der begleitenden Lehrkräftebefragung berichtet.

Das Kapitel zur Aus- und Fortbildung von Lehrkräften in der Primarstufe gliedert sich in drei Abschnitte: Im ersten Teil wird die Gruppe der teilnehmenden Deutsch- und Mathematiklehrkräfte hinsichtlich ihrer demografischen Merkmale charakterisiert. Im zweiten Teil wird die fachliche Qualifikation der Lehrkräfte für die unterrichteten Fächer in den Blick genommen. Insbesondere wird der Frage nachgegangen, welcher Anteil von Lehrkräften in jedem Land ein Fachstudium in den getesteten Fächern (Deutsch und Mathematik) absolviert hat und wie dies, länderübergreifend betrachtet, in Beziehung zu den erreichten Kompetenzen der Schülerinnen und Schüler in diesen Fächern steht. Im dritten Abschnitt des Kapitels stehen schließlich die Fortbildungsaktivitäten von Lehrkräften im Mittelpunkt. Nach einer kurzen Beschreibung der organisatorischen Rahmenbedingungen von Fortbildungen in den Ländern werden inhaltsbezogene und quantitative Angaben zur Fortbildungsteilnahme dargestellt. Darüber hinaus wird berichtet, in welchen Themenbereichen Fortbildungsbedarfe geäußert werden und welche Faktoren die Teilnahme an Fortbildungen aus Sicht der Befragten erschweren oder sogar verhindern können.

Die Datengrundlage dieses Kapitels bildet die Lehrkräftebefragung, die im Rahmen des Ländervergleichs 2011 in den beteiligten Schulen stattfand. Befragt

wurden jeweils die Fachlehrkräfte der Fächer Deutsch und Mathematik der am Ländervergleich teilnehmenden Klassen. Insgesamt beteiligten sich 1 816 Lehrkräfte, von denen 582 nur das Fach Deutsch, 555 nur das Fach Mathematik und 645 beide Fächer unterrichteten.¹ Angaben über die Teilnahmequoten der Lehrkräftebefragung innerhalb der Länder finden sich in Kapitel 4 dieses Bandes.

10.1 Demografische Angaben der Lehrkräfte im Ländervergleich

Vor dem Hintergrund der öffentlichen Diskussion über einen Mangel an männlichen Lehrkräften im Primarbereich und auch einer potenziellen Überalterung der deutschen Lehrerkollegien werden in diesem Abschnitt die an der Studie teilnehmenden Lehrkräfte zunächst bezüglich ihrer Geschlechter- und Altersverteilung charakterisiert. So werden zum einen Unterschiede zwischen den Lehrerkollegien der Länder deutlich, zum anderen wird sichergestellt, dass mit einer für das Land repräsentativen Stichprobe belastbare Aussagen auf Länderebene hinsichtlich der Aus- und Fortbildung getroffen werden können.²

Die in Tabelle 10.1 dargestellten Ergebnisse machen deutlich, dass der Anteil weiblicher Lehrkräfte in den Grundschulen durchgehend hoch ist. Über die Länder hinweg variiert dieser Anteil zwischen knapp 82 Prozent in Baden-Württemberg und circa 93 Prozent in Sachsen-Anhalt.

Im Unterschied zum durchgängig hohen Anteil weiblicher Lehrkräfte unterscheidet sich die Lehrerschaft hinsichtlich der Altersstruktur zwischen den Ländern deutlich. Hier fallen insbesondere die Länder auf, in denen mehr als die Hälfte der Lehrerschaft 50 Jahre und älter ist; dies sind die Länder Sachsen-Anhalt (59.2 %), Sachsen (57.6 %), Baden-Württemberg (57.4 %), Berlin (55.3 %), Thüringen (51.4 %) und Bremen (50.4 %). Deutschlandweit befindet sich knapp die Hälfte der Lehrerschaft (49.0 %) in dem Alterssegment der über 50-Jährigen. Im Saarland hingegen ist nur etwa jede dritte Lehrkraft im Primarbereich über 50 Jahre alt.

Auch im Hinblick auf den Anteil der Nachwuchslehrkräfte nimmt das Saarland eine Sonderstellung ein. Hier ist der Anteil sehr junger Lehrkräfte (bis 29 Jahre) mit knapp 20 Prozent vergleichsweise hoch. Ebenfalls relativ hohe Anteile an jungen Lehrkräften finden sich in den Ländern Hessen (12.0 %), Nordrhein-Westfalen (13.3 %) und Rheinland-Pfalz (11.8 %). In den ostdeutschen Ländern fällt der Anteil der Nachwuchslehrkräfte durchweg erheblich niedriger aus und variiert zwischen etwa einem Prozent in Mecklenburg-Vorpommern und Sachsen-Anhalt sowie knapp 5 Prozent in Thüringen.

Der hohe Altersdurchschnitt der Lehrkräfte lässt sich in den ostdeutschen Flächenländern in erster Linie mit den stark rückläufigen Schülerzahlen seit der Wiedervereinigung und den damit einhergehenden niedrigen Neueinstellungszahlen junger Lehrerinnen und Lehrer erklären (Statistisches Bundesamt, 2011). Trotz der immer noch vergleichsweise niedrigen Schülerzahlen ist auf-

- 1 Von 34 Lehrkräften liegen keine Angaben darüber vor, welches Fach sie in der getesteten Klasse unterrichteten.
- 2 Die von den am Ländervergleich teilnehmenden Lehrkräften berichteten Angaben wurden mit Daten der Schulstatistik des Statistischen Bundesamtes beziehungsweise der Statistischen Landesämter abgeglichen, um mögliche Verzerrungen der Stichprobe zu prüfen (Statistisches Bundesamt, 2011). Der Vergleich der Angaben zur Geschlechter- und Altersverteilung zeigte keine statistisch signifikanten Abweichungen von der Grundgesamtheit aller Lehrkräfte der Grundschulen. Somit kann davon ausgegangen werden, dass die vorliegende Stichprobe die länderspezifische Verteilung der Lehrkräfte in Bezug auf das Geschlecht und das Alter in unverzerrter Form wiedergibt.

Tabelle 10.1: Geschlecht und Alter der teilnehmenden Deutsch- und Mathematiklehrkräfte nach Ländern

Land	N	Weiblich (in %)	Alter in Jahren (in %)				
			bis 29	30–39	40–49	50–59	ab 60
Baden-Württemberg	94	81.9	8.5	25.5	8.5	43.6	13.8
Bayern	82	86.4	6.3	18.8	30.0	33.8	11.3
Berlin	166	86.0	3.7	11.2	29.8	39.8	15.5
Brandenburg	111	91.9	2.8	11.0	38.5	45.0	2.8
Bremen	129	90.9	5.0	23.5	21.0	37.8	12.6
Hamburg	140	84.8	5.8	24.1	24.1	34.3	11.7
Hessen	100	89.0	12.0	22.0	21.0	32.0	13.0
Mecklenburg-Vorpommern	105	92.4	1.0	3.8	48.6	41.9	4.8
Niedersachsen	132	85.5	5.4	20.0	26.9	31.5	16.2
Nordrhein-Westfalen	107	90.7	13.3	26.7	18.1	30.5	11.4
Rheinland-Pfalz	86	82.6	11.8	23.5	23.5	34.1	7.1
Saarland	83	87.3	19.7	35.5	7.9	18.4	18.4
Sachsen	120	90.0	3.4	7.6	31.4	43.2	14.4
Sachsen-Anhalt	124	92.7	0.8	1.7	38.3	50.0	9.2
Schleswig-Holstein	132	90.1	6.3	22.8	24.4	34.6	11.8
Thüringen	105	88.6	4.9	9.7	34.0	42.7	8.7
Deutschland	1816	88.3	6.4	17.4	27.2	37.5	11.5

grund steigender Pensionierungszahlen jedoch in den kommenden Jahren mit einem größeren Bedarf an Neueinstellungen auch in den ostdeutschen Ländern zu rechnen (KMK, 2011).

10.2 Fachfremdes Unterrichten im Primarbereich

Neben der Geschlechterverteilung und Alterszusammensetzung der Lehrerschaft wird häufig auch die Frage der fachlichen Qualifikation von Lehrkräften diskutiert (Blömeke et al., 2010a, 2010b; Kunter et al., 2011). In der Regel erwerben Lehrkräfte im Studium und im anschließenden Referendariat eine Lehrbefähigung für bestimmte Fächer, jedoch werden sie aufgrund schulspezifischer Erfordernisse teilweise auch in Fächern eingesetzt, in denen sie kein Studium absolviert haben. In der Grundschule ist darüber hinaus das Klassenlehrerprinzip stark verbreitet, nach dem die Klassenlehrkraft mehrere Fächer in ihrer Klasse unterrichtet (Schorch, 2009). In der Praxis bedeutet dies, dass Lehrkräfte nicht nur ihre studierten Fächer unterrichten, sondern auch Unterricht in Fächern erteilen, für die sie keine berufliche Qualifikation im Studium und Referendariat erworben haben. In welchem Umfang Lehrkräfte der Primarstufe fachfremd eingesetzt werden und ob fachfremder Unterricht im Zusammenhang mit den erreichten Kompetenzen von Schülerinnen und Schülern steht, ist bislang in Deutschland nur wenig erforscht.

Im Ländervergleich 2011 gaben knapp 17 Prozent der befragten Deutschlehrkräfte und etwa 27 Prozent der befragten Mathematiklehrkräfte an, die betreffenden Unterrichtsfächer zu unterrichten, diese aber nicht studiert zu haben (vgl. Tabelle 10.2). Die länderspezifische Aufschlüsselung weist da-

bei auf große Unterschiede hin. Der geringste Anteil fachfremd unterrichtender Lehrkräfte findet sich in Thüringen (Deutsch: 0 %; Mathematik: 1.3 %). Hohe Anteile fachfremd unterrichtender Lehrkräfte sind hingegen in Hamburg zu verzeichnen, wo circa 34 Prozent der Deutsch- beziehungsweise 48 Prozent der Mathematiklehrkräfte unterrichten, ohne im jeweiligen Fach ein Studium absolviert zu haben. Im Vergleich der Länder zeigt sich, dass vor allem die ostdeutschen Länder und das Land Bayern geringe Anteile fachfremd unterrichtender Lehrkräfte aufweisen. Dieser Befund lässt sich zumindest in den ostdeutschen Ländern darauf zurückführen, dass die Lehrkräfte dort deutlich älter sind als in den westdeutschen Ländern und ihre Ausbildung somit zum Großteil noch in der DDR absolviert haben. Im Ausbildungssystem der DDR wurden Unterstufenlehrkräfte sowohl in Deutsch als auch in Mathematik und einem weiteren Wahlpflichtfach ausgebildet (Fuchs, 2007). In der Bundesrepublik Deutschland legen die einzelnen Länder selbst fest, welche Fächer im Studium des Primarstufenlehramtes verpflichtend belegt werden müssen. Diese Vorgaben der Kultusministerien tragen dazu bei, dass sich die Lehrkräfte in den Ländern hinsichtlich ihrer fachlichen Qualifikationen in den von ihnen unterrichteten Fächern zum Teil unterscheiden.

Tabelle 10.2: Anteil der fachfremd unterrichtenden Lehrkräfte in Deutsch und Mathematik pro Land und bundesweit

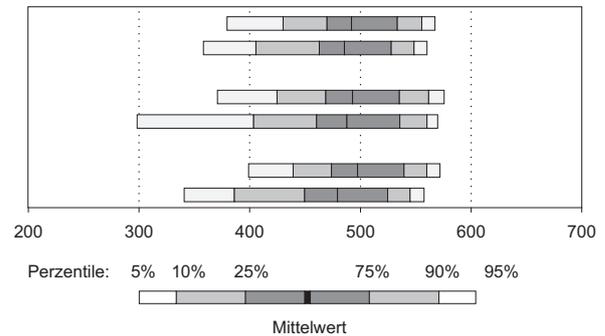
Land	Deutsch		Mathematik	
	N	Anteil fachfremd unterrichtende Lehrkräfte in %	N	Anteil fachfremd unterrichtende Lehrkräfte in %
Baden-Württemberg	66	34.8	69	44.9
Bayern	79	2.5	76	15.8
Berlin	103	27.2	101	37.6
Brandenburg	71	14.1	68	14.7
Bremen	78	20.5	84	46.4
Hamburg	86	33.7	79	48.1
Hessen	75	21.3	65	40.0
Mecklenburg-Vorpommern	79	6.3	75	13.3
Niedersachsen	75	33.3	73	39.7
Nordrhein-Westfalen	76	19.7	77	27.3
Rheinland-Pfalz	73	15.1	73	24.7
Saarland	72	23.6	69	33.3
Sachsen	68	2.9	69	2.9
Sachsen-Anhalt	79	5.1	78	5.1
Schleswig-Holstein	69	5.8	68	38.2
Thüringen	78	0.0	76	1.3
Deutschland	1227	16.9	1200	27.3

Wie hängen nun der Anteil fachfremd unterrichtender Lehrkräfte und die durchschnittlich erreichten Kompetenzen von Schülerinnen und Schülern im Lesen, im Zuhören und in Mathematik zusammen? In Abbildung 10.1 sind in Form von Perzentilbändern die durchschnittlich erreichten Kompetenzstände in Klassen mit fachfremd beziehungsweise nicht fachfremd unterrichtenden Lehrkräften dargestellt. Im Lesen schneiden Klassen, die durch eine Lehrkraft mit einer Lehrbefähigung für Deutsch unterrichtet wurden, im Durchschnitt um 7 Punkte

Abbildung 10.1: Gewichtete Verteilung der Kompetenzstände auf Klassenebene für fachfremd und nicht fachfremd unterrichtende Lehrkräfte nach Kompetenzbereich

Kompetenzbereich	Lehrkräfte	Anteil in %	M	SD	d
Deutsch – Lesen	nicht fachfremd	81.6	492	69	.09
	fachfremd	18.4	485	65	
Deutsch – Zuhören	nicht fachfremd	81.6	493	71	.07
	fachfremd	18.4	488	73	
Mathematik	nicht fachfremd	70.2	497	68	.26
	fachfremd	29.8	479	70	

Anmerkung. In dieser Analyse wurden die Daten auf Klassenebene aggregiert. Unterschiede zu den Anteilen fachfremd unterrichtender Lehrkräfte in Tabelle 10.2 lassen sich darauf zurückführen, dass Gewichte für die Berechnung der Kompetenzstände berücksichtigt und Klassen zum Teil von zwei Lehrkräften in Deutsch bzw. Mathematik unterrichtet wurden.



oberes Band: mittlere Klassenleistung bei nicht fachfremden Lehrkräften
unteres Band: mittlere Klassenleistung bei fachfremden Lehrkräften

und im Zuhören um 5 Punkte besser ab. In Mathematik ist die Differenz mit 18 Punkten deutlich größer.

Besonders deutliche Kompetenzunterschiede zwischen Klassen, die von Lehrkräften in ihrem studierten Fach beziehungsweise fachfremd unterrichtet wurden, zeigen sich dann, wenn man die fünf Prozent der Leistungsschwächsten gesondert betrachtet (5. Perzentil). Im Kompetenzbereich Lesen beträgt die Differenz in dieser Gruppe 21 Punkte, im Zuhören 73 Punkte und in Mathematik 58 Punkte.

10.3 Fortbildung von Lehrkräften

Neben der Absolvierung eines Fachstudiums sind auch spätere Fortbildungen wichtig für die professionelle Kompetenz von Lehrerinnen und Lehrern (Lipowsky, 2011; Timperley, Wilson, Barrar & Fung, 2007; Yoon, Duncan, Lee, Scarloss & Shapley, 2007). Bevor jedoch die Fortbildungsaktivitäten der befragten Lehrkräfte analysiert werden, sollen zunächst die länderspezifischen Rahmenbedingungen für Fortbildungen für das Schuljahr 2010/2011 dargestellt werden (vgl. Tabelle 10.3). Dies umfasst die vorgeschriebene Mindestanzahl der zu besuchenden Fortbildungsstunden pro Schuljahr, die Möglichkeit, für Fortbildungen Bildungsurlaub zu beanspruchen oder diese innerhalb der Unterrichtszeit besuchen zu können und die Organisation schulinterner Fortbildungen.

Obwohl die kontinuierliche Fortbildung für Lehrkräfte in allen Ländern prinzipiell verpflichtend ist (Avenarius & Heckel, 2000), gibt es nur in vier Ländern konkrete Vorgaben zum Umfang der zu besuchenden Fortbildungen. In Hamburg und Bremen wird erwartet, dass Lehrkräfte mindestens 30 Stunden Fortbildung innerhalb eines Schuljahres besuchen, während in Bayern zwölf Fortbildungstage über einen Zeitraum von insgesamt vier Schuljahren erbracht werden sollen.³ In Hessen bestand bis zum 31.12.2010 die Vorgabe, innerhalb von drei Schuljahren 150 Leistungspunkte zu sammeln, die durch verschiedene Aktivitäten, unter anderem die Teilnahme an Fortbildungen, die Mitarbeit in der Lehrerausbildung oder

3 Bremen: Verordnung über die Fortbildung der Lehrkräfte und Lehrer in besonderer Funktion, 2005; Hamburg: Verordnung über die Arbeitszeit der Lehrerinnen und Lehrer an staatlichen Schulen, 2003; Bayern: Lehrerfortbildung in Bayern, 2002.

Tabelle 10.3: Rechtliche Rahmenbedingungen der Lehrerfortbildung in den Ländern im Schuljahr 2010/2011

Land	Mindestteilnahme	Umfang	Anspruch auf Bildungsurlaub	Anspruch auf Freistellung während der Unterrichtszeit	Verpflichtung zur Erstellung eines Fortbildungsplans	Schuleigenes Budget	Höhe des Budgets ¹
Baden-Württemberg	○		– ²	● ³	●	●	Nach Maßgabe des Staatshaushaltsplans
Bayern	●	12 Tage in 4 Jahren	○ ⁴	○	●	○ ⁵	
Berlin	○		●	○	●	○ ⁶	
Brandenburg	○		●	○	●	○ ⁷	
Bremen	●	30 Stunden im Jahr	●	● ⁸	●	●	
Hamburg	●	30 Stunden im Jahr an allg.-bild. Schulen	●	○	●	● ⁹	
Hessen	● ¹⁰	150 Punkte in 3 Jahren	●	● ³	●	●	Nach Maßgabe des Haushaltsgesetzes
Mecklenburg-Vorpommern	○		●	● ¹¹	●	●	150 € pro Schule und 70 € pro Lehrkraft
Niedersachsen	○		●	● ¹¹	●	●	Je nach Anzahl der Lehrkraft pro Schule
Nordrhein-Westfalen	○		●	● ¹²	●	●	45 € pro Lehrkraft, mindestens jedoch 800 € pro Schule
Rheinland-Pfalz	○		●	● ¹³	●	▶ ¹⁴	
Saarland	○		●	●	●	●	
Sachsen	○		○	● ¹¹	●	●	25 € pro Lehrkraft pro Schuljahr für schulinterne Lehrerfortbildung
Sachsen-Anhalt	○		●	●	●	○	
Schleswig-Holstein	○		●	● ¹¹	●	○	
Thüringen	○		○	● ³	●	●	500 € pro Schule und Haushaltsjahr

Anmerkungen. ○ nicht vorhanden ▶ teilweise vorhanden ● vorhanden. ¹ Wenn keine Informationen zum schulinternen Budget aufgeführt sind, lagen keine konkreten Zahlen vor. ² Keine Information verfügbar. ³ Anspruch auf Freistellung besteht, wenn die Veranstaltung als dienstliche Fortbildung anerkannt ist. ⁴ Für Personalratsmitglieder für bis zu 5 Kalendertage möglich. ⁵ Es besteht die Möglichkeit, einen Zuschuss z. B. für Referentenhonorare und Sachkosten bei der zuständigen Bewilligungsbehörde zu beantragen. Die Höhe des Zuschusses ist abhängig von den zur Verfügung stehenden Haushaltsmitteln. ⁶ Finanzierung der Fortbildungen durch externe Anbieter möglich. ⁷ Die Schulen können in Abstimmung mit dem zuständigen Schulamt Mittel für die Durchführung von Fortbildungsveranstaltungen erhalten. ⁸ Sobald die Veranstaltungen während der Unterrichtszeit stattfinden, müssen diese Fortbildungen von der Schulleitung genehmigt werden. In diesem Fall sind die grundsätzlichen Regelungen zur Vermeidung von Unterrichtsausfall zu berücksichtigen. ⁹ Die Schulen verfügen über ein Budget, dass sie u. a. auch für die Finanzierung von schulinternen Fortbildungen einsetzen können. ¹⁰ Im Jahr 2011 wurde diese Vorgabe in Hessen aufgehoben. ¹¹ Anspruch besteht, sofern die Unterrichtsversorgung gesichert ist. ¹² Freistellung kann durch die Schulleitung gewährt werden. Die Genehmigung von Fortbildungen während der Unterrichtszeit setzt in der Regel voraus, dass eine Vertretung gesichert ist oder der Unterricht vorgezogen oder nachgeholt oder Unterrichtsausfall auf andere Weise vermieden wird. ¹³ Anspruch besteht, wenn im Einzelfall die Veranstaltung vom Kultusministerium zum Dienst erklärt wird und die Schulleitung festgestellt hat, dass die Teilnahme für die jeweilige Lehrkraft von dienstlichem Nutzen ist und dienstliche Belange nicht entgegenstehen. ¹⁴ Im Rahmen einiger Projekte stehen Budgets für Fortbildung zur Verfügung.

Quelle: Kultusministerien der Länder und öffentlich verfügbare Rechtsvorschriften zur Regelung der Lehrerfortbildung.

die Übernahme administrativer schulischer Tätigkeiten erbracht werden konnten. Diese Regelung wurde ab dem Jahr 2011 durch eine allgemeine Verpflichtung zur Teilnahme an Fortbildungsmaßnahmen abgelöst, ohne dass Umfang oder Art spezifiziert wurden.⁴

Die Angaben zur Mindestteilnahme in Tabelle 10.3 machen deutlich, dass Lehrkräfte in den meisten Ländern selbst bestimmen können, in welchem Umfang sie in jedem Schuljahr an Fortbildungen partizipieren. In den meisten Ländern haben Lehrkräfte auch die Möglichkeit, für Fortbildungen Bildungsurlaub in Anspruch zu nehmen und innerhalb der Unterrichtszeit freigestellt zu werden. Ferner sind die Schulen in allen Ländern dazu angehalten, eine Fortbildungsplanung vorzunehmen, um regelmäßige Fortbildungsangebote zu realisieren, die den Bedürfnissen der Schule entsprechen. Zur Ausgestaltung dieser Pläne stehen mit Ausnahme der Länder Bayern, Berlin, Brandenburg, Sachsen-Anhalt und Schleswig-Holstein in allen Ländern schulinterne Budgets zur Verfügung. Diese Budgets werden in Thüringen als Festbetrag, andernorts in Abhängigkeit von der Lehrerschaft zur Verfügung gestellt (Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen und Sachsen).

10.3.1 Anzahl und Umfang der besuchten Lehrerfortbildungen

Bislang wurden nur wenige repräsentative Lehrkräftebefragungen durchgeführt, die Aufschluss über die Quantität der besuchten Fortbildungsveranstaltungen in Deutschland geben. Im Rahmen der PISA-Studie⁵ aus dem Jahr 2006 beispielsweise wurden Lehrkräfte der Mathematik und der Naturwissenschaften in der Sekundarstufe I zu ihren Fortbildungsaktivitäten befragt (Frey et al., 2009). In dieser Erhebung gaben etwa 82 Prozent der Lehrkräfte an, im letzten Jahr eine Fortbildungsveranstaltung besucht zu haben. Eine ähnlich hohe Beteiligungsquote von circa 81 Prozent findet sich auch in der COACTIV-Studie, in der Mathematiklehrkräfte der Sekundarstufe I nach ihren Fortbildungsaktivitäten innerhalb von zwei Schuljahren befragt wurden (Richter et al., 2010).

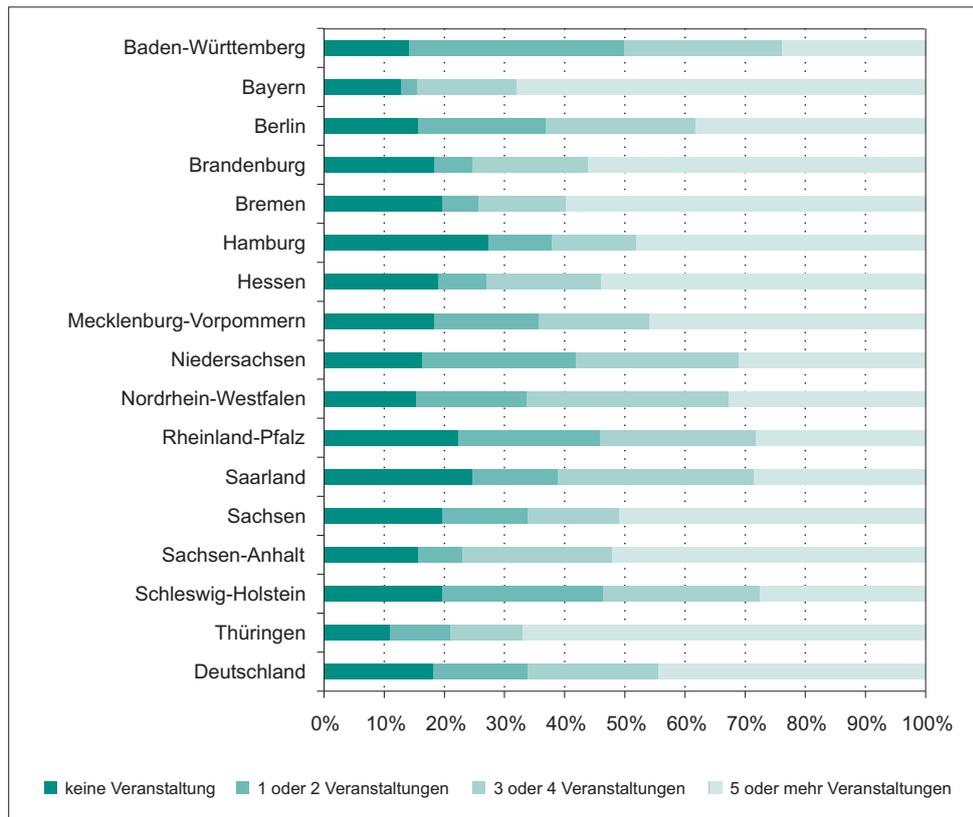
Zur Erfassung der Fortbildungsaktivitäten in der Primarstufe wurden im Ländervergleich 2011 die Deutsch- und Mathematiklehrkräfte der teilnehmenden Klassen gebeten, alle besuchten Fortbildungsveranstaltungen in den Schuljahren 2009/2010 und 2010/2011 zu benennen und darüber hinaus auch Angaben zum Themenbereich, zum genauen Fortbildungsthema, zum Stundenumfang und zur Veranstaltungsform zu machen.

Die Ergebnisse in Abbildung 10.2 zeigen, dass die Anzahl besuchter Fortbildungsveranstaltungen zwischen den Ländern deutlich variiert. In Thüringen haben lediglich 11 Prozent der Lehrkräfte innerhalb des Berichtszeitraums keine Fortbildungsveranstaltung besucht, während der Anteil in Hamburg mit etwa 27 Prozent mehr als doppelt so hoch war. Die Gruppe der fortbildungsaktivsten Lehrkräfte (fünf oder mehr Veranstaltungen) ist in Bayern und Thüringen mit über 60 Prozent der Lehrkräfte am größten und in Baden-Württemberg mit etwas weniger als einem Viertel der Lehrkräfte vergleichsweise klein.

Vor dem Hintergrund der konkreten Fortbildungsverpflichtung in Bayern, Bremen, Hamburg und Hessen wäre zu vermuten, dass alle Lehrkräfte die-

4 Die konkreten Vorgaben zur Fortbildungsteilnahme wurden in der Verordnung zur Umsetzung des Hessischen Lehrerbildungsgesetzes aus dem Jahr 2005 festgeschrieben. Diese Regelung hat das Hessische Lehrerbildungsgesetz im Jahr 2011 abgelöst.

5 Das Akronym PISA steht für *Programme for International Student Assessment*.

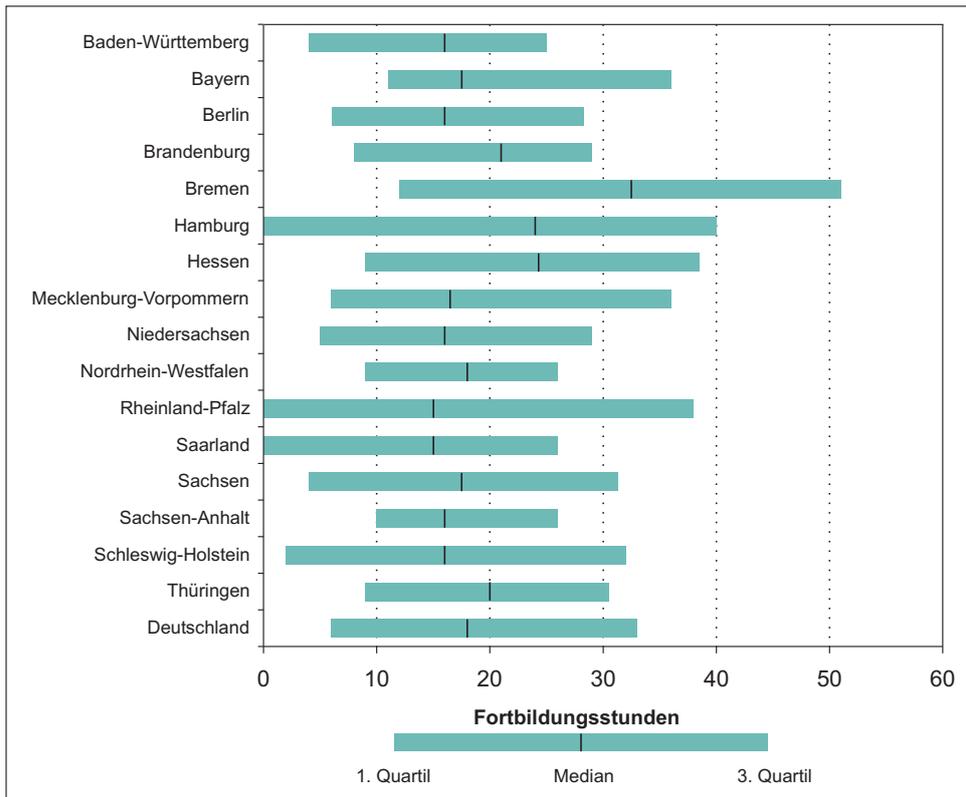
Abbildung 10.2: Prozentuale Verteilung der Anzahl von Lehrkräften besuchter Fortbildungsveranstaltungen in den Schuljahren 2009/2010 und 2010/2011 nach Ländern

ser Länder an Fortbildungen teilgenommen haben. Entgegen dieser Erwartung zeigt sich jedoch, dass auch in diesen Ländern ein substantieller Anteil der Lehrkräfte angibt, in den letzten beiden Schuljahren keine Fortbildungen besucht zu haben. Von den Ländern ohne konkrete Fortbildungsverpflichtung fallen die ostdeutschen Länder mit einer allgemein hohen Fortbildungsaktivität auf. Insbesondere Thüringen weist ähnlich hohe Teilnehmeraten wie Bayern auf, obwohl es dort keine spezifische Vorgabe gibt. In den westdeutschen Ländern ohne Mindestvorgaben berichten Lehrkräfte weitgehend vergleichbare Teilnehmeraten: Der Anteil der Lehrkräfte mit fünf oder mehr Veranstaltungen bewegt sich dort zwischen knapp 24 Prozent in Baden-Württemberg und etwa 33 Prozent in Nordrhein-Westfalen.

Bei der Betrachtung von Fortbildungen spielt nicht nur die Anzahl der Veranstaltungen eine Rolle, sondern auch die insgesamt investierte Zeit. Aus den Stundenangaben lassen sich Hinweise darauf gewinnen, ob länderspezifische zeitliche Nutzungsmuster vorliegen, ob also beispielsweise eher mehrstündige Nachmittagsangebote oder ganztägige Veranstaltungen beziehungsweise Veranstaltungsreihen besucht werden.

Abbildung 10.3 gibt für jedes Land den Median und die Quartile für die insgesamt besuchten Fortbildungsstunden an.⁶ Über alle Länder hinweg liegt der

6 Der Median teilt die berichteten Fortbildungsstunden in zwei Hälften, sodass 50 Prozent der Werte unterhalb und 50 Prozent der Werte oberhalb des Medians liegen. Darüber hinaus geben die Quartile ein Maß für die Variabilität der angegebenen Fortbildungsstunden an. Das untere Quartil gibt die Stundenzahl an, unter der 25 Prozent aller berichteten Angaben liegen, während das obere Quartil den Wert darstellt, unter dem 75 Prozent aller berichteten Angaben liegen.

Abbildung 10.3: Median und Quartile der von Lehrkräften aufgewendeten Fortbildungsstunden in den Schuljahren 2009/2010 und 2010/2011 nach Ländern

Median für die in den zwei Schuljahren besuchten Fortbildungen bei 18 Stunden. Deutlich oberhalb dieses Wertes befinden sich mit Bremen (32.5 Stunden), Hessen (24.3 Stunden) und Hamburg (24 Stunden) nur solche Länder, die im Schuljahr der Befragung verbindliche Vorgaben zur Teilnahme an Fortbildung machten. Die Mediane aller anderen Länder variieren zwischen 15 Stunden (Saarland und Rheinland-Pfalz) und 21 Stunden (Brandenburg). Auffällig ist ebenfalls die unterschiedlich große Spannweite zwischen dem unteren und dem oberen Quartil. Eine vergleichsweise große Spanne an Fortbildungsstunden zeigt sich in Bremen (12 bis 51 Stunden), Hamburg (0 bis 40 Stunden) und Rheinland-Pfalz (0 bis 38 Stunden). Dieses Ergebnis weist darauf hin, dass sich die Fortbildungsaktivität in diesen Ländern in hohem Maße zwischen Lehrkräften unterscheidet, und dies, obwohl in den erstgenannten Ländern eine verbindliche Verpflichtung von 30 Stunden pro Schuljahr besteht. Nur geringe Differenzen zwischen den Quartilen zeigen sich hingegen in Nordrhein-Westfalen (9 bis 26 Stunden) und Sachsen-Anhalt (10 bis 26 Stunden). In diesen Ländern investiert somit ein Großteil der Primarstufenlehrkräfte einen vergleichsweise ähnlichen Stundenumfang in Fortbildungen.

Werden die besuchten Veranstaltungen (Abbildung 10.2) den investierten Fortbildungsstunden (Abbildung 10.3) gegenübergestellt, fällt auf, dass in Ländern, in denen Lehrkräfte viele Fortbildungsveranstaltungen besuchen, nicht immer auch hohe Stundenzahlen dafür aufgewendet werden. In Bayern und Thüringen, in denen über 60 Prozent der Lehrkräfte fünf oder mehr Veranstaltungen innerhalb von zwei Jahren besuchen, beträgt der Median der besuchten Fortbildungsstunden 17.5 beziehungsweise 20 Stunden. Dieser Befund lässt darauf schließen, dass die besuchten Veranstaltungen in diesen beiden Ländern durchschnittlich eher kürzer sind als in anderen Ländern.

10.3.2 Themenbereiche von Fortbildungen: Inhalte und subjektive Bedarfe

In den internationalen Schulleistungsstudien wie PISA, TIMSS⁷ und PIRLS/IGLU⁸ wurden häufig nur Informationen über die Teilnahme an Fortbildungen zu wenigen und eng umgrenzten Themengebieten (z. B. Fachdidaktik, effektives Unterrichten) erhoben (Frey et al., 2009). Im Rahmen der Lehrkräftebefragung im IQB-Ländervergleich 2011 wurden die thematischen Schwerpunkte besuchter Fortbildungsveranstaltungen erstmals differenzierter erfasst. Die befragten Lehrkräfte ordneten die von ihnen ausgewählten Veranstaltungen selbst einer vorgegebenen Liste von Themenbereichen zu (siehe Tabelle 10.4). Diese Liste wurde aus dem Lehrerfragebogen der internationalen TALIS-Befragung (*Teaching and Learning International Survey*) im Jahr 2008 übernommen (OECD, 2009) und für die vorliegende Studie angepasst und erweitert. Darüber hinaus wurden die Lehrkräfte auch zu ihrem subjektiven Fortbildungsbedarf in dem jeweiligen Themengebiet befragt. Für jedes der Themengebiete gaben die Lehrkräfte auf einer vierstufigen Skala von (1) „überhaupt kein Bedarf“ bis (4) „sehr hoher Bedarf“ ihre Einschätzung ab.

In Tabelle 10.4 wird ersichtlich, zu welchem Anteil die Lehrkräfte zu den aufgeführten Themenbereichen mindestens eine Fortbildungsveranstaltung innerhalb der letzten beiden Schuljahre besucht haben. Fast die Hälfte der befragten Lehrkräfte besuchte Veranstaltungen zu fachdidaktischen Themen (44.2 %) und etwa ein Viertel zu Unterrichtsformen und -methoden (26.8 %). Eine größere Gruppe von Lehrkräften besuchte auch Fortbildungen zu Fragen der individuellen Förderung, zum Beispiel zur Förderung lernschwacher Schülerinnen und Schüler (17.5 %) und zur Binnendifferenzierung (13.9 %). Vergleichsweise wenige Lehrkräfte nahmen Veranstaltungen im Bereich der interkulturellen Pädagogik (1.6 %) und zu Ganztagsangeboten (2.3 %) wahr.

Tabelle 10.4 gibt für jedes der aufgelisteten Fortbildungsthemengebiete auch den Anteil der Lehrkräfte an, die für den jeweiligen Themenbereich einen hohen oder sehr hohen subjektiven Fortbildungsbedarf berichten. Mehr als zwei Drittel der Lehrkräfte gab an, einen besonderen Bedarf an Fortbildungen zur Förderung lernschwacher Schülerinnen und Schüler sowie zur Integration von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf zu haben. Insbesondere die vor dem Hintergrund der zunehmenden Inklusionsmaßnahmen der Länder relevante Integration von Förderschülerinnen und Förderschülern scheint bislang nur selten Inhalt besuchter Fortbildungsveranstaltungen gewesen zu sein. Hier steht dem hohen subjektiven Bedarf, der von fast 70 Prozent der befragten Lehrkräfte geäußert wurde, nur etwa ein Zehntel der Lehrkräfte gegenüber, die an mindestens einer Veranstaltung zu diesem Thema teilgenommen hat. Ein hoher oder sehr hoher subjektiver Fortbildungsbedarf wurde von über 60 Prozent der Lehrkräfte auch zu Fragen der Binnendifferenzierung und der individuellen Förderung geäußert. Auch in diesem Bereich ist der Anteil der Lehrkräfte, die einen subjektiven Bedarf in diesem Bereich geäußert hat, deutlich größer als der Anteil der Lehrkräfte, die solche Fortbildungsveranstaltungen besucht hat. Eine deutliche Kluft zwischen Fortbildungsbesuch und -bedarf zeigte sich zudem für

7 Das Akronym TIMSS stand ursprünglich für *Third International Mathematics and Science Study*. Seit 2003 wird es in der Bedeutung *Trends in International Mathematics and Science Study* verwendet.

8 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.

die Themenbereiche Sprachförderung und jahrgangübergreifendes Lernen, bei denen einem hohen Bedarf nur geringe Teilnahmequoten gegenüberstehen.

10.3.3 Subjektive Barrieren der Fortbildungsteilnahme

In den Analysen des zeitlichen Fortbildungsaufwands ist deutlich geworden, dass in den Ländern zwischen 11 Prozent und etwa 27 Prozent der Lehrkräfte in den letzten beiden Schuljahren nach eigenen Angaben gar keine Fortbildung besucht haben (vgl. Abschnitt 10.3.1). Da Fortbildungen eine wichtige Rolle für die fortlaufende Professionalisierung von Lehrerinnen und Lehrern spielen, sind auch die subjektiv empfundenen Barrieren von Interesse. Um diesen nachzugehen, wurde allen Lehrkräften eine Liste von möglichen Gründen vorgelegt, die eine Teilnahme an Fortbildungsveranstaltungen potenziell erschweren oder verhindern (vgl. Tabelle 10.5). Die Lehrkräfte gaben zu jedem Grund auf einer vierstufigen Skala von (1) „gar nicht“ bis (4) „sehr stark“ an, in welchem Ausmaß er die Teilnahme an Fortbildungsveranstaltungen für sie erschwert oder verhindert hat.

Tabelle 10.4: Teilnahme an Fortbildungsveranstaltungen (2009/2010; 2010/2011) und Anteil der Lehrkräfte mit einem hohen beziehungsweise sehr hohen subjektiven Fortbildungsbedarf nach Themenbereichen (absteigend nach Anteil der Teilnahme geordnet)

Themenbereich	Teilnahme		Fortbildungsbedarf	
	N	Anteil ¹ in %	N	Anteil ² in %
Fachdidaktische Themen	1 744	44.2	1 654	37.1
Unterrichtsformen und -methoden (fächerübergreifend)	1 744	26.8	1 675	44.9
Leistungsdiagnostik und -beurteilung	1 744	17.7	1 675	52.8
Förderung von lernschwachen Schülerinnen und Schülern	1 744	17.5	1 695	69.4
Nutzung von Medien im Unterricht	1 744	17.4	1 677	36.0
Curricula (Lehrpläne, Rahmenpläne)	1 744	16.7	1 663	19.8
Bildungsstandards der Kultusministerkonferenz	1 744	16.0	1 649	26.0
Fachliche Themen ohne Bezug zum Unterricht	1 744	15.1	1 647	9.5
Schulorganisation / Schulentwicklung	1 744	14.4	1 651	24.2
Gewaltprävention	1 744	13.9	1 680	38.9
Binnendifferenzierung / individuelle Förderung	1 744	13.9	1 670	61.5
Umgang mit Störungen im Unterricht	1 744	12.1	1 691	41.3
Beratung von Eltern oder Schülerinnen und Schülern	1 744	10.6	1 677	31.7
Leseförderung	1 744	10.4	1 668	36.1
Integration / Inklusion von Schülerinnen und Schülern mit sonderpäd. Förderbedarf	1 744	9.5	1 681	68.0
Sprachförderung	1 744	8.9	1 668	45.7
Jahrgangübergreifendes Lernen	1 744	7.1	1 677	44.8
Ganztagsangebote	1 744	2.3	1 648	19.8
Interkulturelle Pädagogik	1 744	1.6	1 659	24.8
Andere	1 744	27.4	–	–

Anmerkungen. Lehrkräfte ohne Angaben bei der Fortbildungsteilnahme und dem Fortbildungsbedarf wurden in diese Analyse nicht einbezogen ($N = 72$). Die Fallzahl (N) unterscheidet sich für den Fortbildungsbedarf zwischen den Themenbereichen, da zum Teil Lehrkräfte bei einzelnen Themenbereichen keine Angaben gemacht haben. ¹ Lehrkräfte mit mindestens einer Fortbildung in diesem Themenbereich in zwei Schuljahren. ² Zusammengefasster Anteil der Lehrkräfte mit hohem oder sehr hohem subjektiven Fortbildungsbedarf.

Tabelle 10.5: Subjektive Hinderungsgründe für die Teilnahme an Fortbildungen (absteigend nach Anteil geordnet)

Hinderungsgründe	starke oder sehr starke Erschwernis	
	N	Anteil in %
Hoher Zeitaufwand für Unterrichtsvorbereitungen	1 720	61.4
Hohes Unterrichtsdeputat	1 686	56.6
Dienstliche Aufgaben an der Schule (Konferenzen etc.)	1 701	56.3
Ungünstiger Ort der Fortbildung	1 715	51.5
Ungünstiger Termin der Fortbildung	1 706	49.5
Mangelnde Passung von Fortbildungsangebot und Fortbildungsbedarf	1 691	43.1
Familiäre Verpflichtungen	1 713	32.0
Ungenügende Qualität	1 670	28.0
Geringes Angebot	1 696	27.7
Probleme mit der Finanzierung	1 698	15.0
Probleme mit der Freistellung von der Schule	1 710	10.9

Anmerkung. Die Fallzahl (N) unterscheidet sich zwischen den Hinderungsgründen, da zum Teil Lehrkräfte bei einzelnen Hinderungsgründen keine Angaben gemacht haben.

Am häufigsten nannten die befragten Lehrkräfte als starken oder sehr starken Hinderungsgrund ihre hohe berufliche Belastung, die mit der Unterrichtsvorbereitung, dem Unterrichten an sich oder den dienstlichen Aufgaben in der Schule verbunden ist. Als weitere wichtige Hinderungsgründe für die Teilnahme an Fortbildungen wurden von etwa der Hälfte der Lehrkräfte Merkmale der Veranstaltungsorganisation, wie zum Beispiel Ort und Termin der Fortbildung, genannt. Fehlende Angebote oder die ungenügende Qualität vorhandener Angebote wurden nur von jeweils einem Viertel der Lehrkräfte als Erschwernis angeführt. Auch die fehlende Freistellung von der Schule wurde nur von einer Minderheit der Lehrkräfte als problematisch erlebt. Insgesamt zeigt sich, dass vor allem die umfangreichen beruflichen Verpflichtungen der Lehrkräfte und weniger die mangelnde Qualität beziehungsweise das fehlende Angebot als Hinderungsgrund gesehen wurden.

10.4 Zusammenfassung

In diesem Kapitel wurden verschiedene Hintergrundmerkmale von Lehrkräften im Primarschulbereich, ihre fachadäquate Ausbildung und ihre Fortbildungsaktivitäten beschrieben. Dabei zeigte sich, dass die in der Primarstufe tätigen Lehrkräfte in allen Ländern ganz überwiegend weiblich sind. Der Blick auf die Alterszusammensetzung der Kollegien wies auf Gemeinsamkeiten und Unterschiede zwischen den Ländern hin. In fast allen Ländern sind viele Lehrerinnen und Lehrer bereits mindestens 50 Jahre alt. Der Anteil junger Lehrkräfte unter 30 Jahren hingegen ist in fast allen Ländern sehr gering. Diese Altersstruktur stellt eine große Herausforderung hinsichtlich der bevorstehenden Pensionierungswelle und der dadurch notwendigen Neuanstellung qualifizierter Junglehrerinnen und -lehrer dar.

Zur Frage des fachfremd erteilten Unterrichts zeigte sich, dass der Anteil der Lehrkräfte ohne Fachstudium zwischen den Ländern erheblich variiert und

sich auch zwischen den Fächern Deutsch und Mathematik unterscheidet. Ein Vergleich der mittleren Kompetenzstände in Klassen von Lehrkräften mit und ohne Fachstudium machte deutlich, dass diese sich insbesondere in Mathematik unterscheiden. Demnach werden im Fach Mathematik im Durchschnitt bessere Ergebnisse erzielt, wenn die Lehrkraft ein Studium in diesem Fach abgeschlossen hat. Dieses Ergebnis ergänzt die Befunde der Studie TEDS-M (*Teacher Education and Development Study – Learning to Teach Mathematics*, Blömeke et al., 2010b), die zeigte, dass angehende Primarstufenlehrkräfte mit Mathematik als Schwerpunktfach über ein umfangreicheres mathematisches und mathematikdidaktisches Wissen verfügen als angehende Lehrkräfte ohne Mathematikstudium. Es lässt sich demnach begründet annehmen, dass Primarstufenlehrkräfte durch ein Studium in Mathematik fachliche Kompetenzen erwerben, die sich messbar in den Lernergebnissen der Schülerinnen und Schüler im Fach Mathematik niederschlagen. Zur Prüfung dieser Hypothese bedarf es jedoch weiterer Studien, die der Frage nachgehen, welche Bedeutung das Fachstudium im Primarbereich für die Gestaltung von Lernprozessen und für die Lernergebnisse von Schülerinnen und Schülern relativ zu anderen möglichen Einflussgrößen besitzt.

Hinsichtlich der Teilnahme an Fortbildungen bestehen länderspezifische Regelungen. Obwohl alle Länder ihre Lehrkräfte zu Fortbildungsaktivitäten anhalten, können diese in der Mehrzahl der Länder selbst darüber entscheiden, wie viel Zeit sie pro Jahr in Fortbildungen investieren. Dies spiegelt sich in sehr heterogenen Angaben zur investierten Zeit für Fortbildungen innerhalb und zwischen den Ländern wider. Die Angaben zu den Themenbereichen der besuchten Veranstaltungen verdeutlichen, dass Lehrkräfte sich zu einem breiten Spektrum von Inhalten weiterqualifizieren. Ein Fortbildungsschwerpunkt lässt sich jedoch im Bereich der fachdidaktischen und methodischen Veranstaltungen ausmachen. Die Auskünfte nach den subjektiven Fortbildungsbedarfen geben Hinweise darauf, in welchen Themenbereichen sich Lehrkräfte weiterqualifizieren möchten. Ein Missverhältnis von Bedarf und Teilnahmeverhalten scheint in besonderer Weise für Fortbildungen zur Integration beziehungsweise Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf zu bestehen. Als bedeutsame Hinderungsgründe zur Teilnahme an (weiteren) Fortbildungen werden von Lehrkräften vor allem die zeitlichen Anforderungen des Lehrerberufs genannt und nur zu einem geringen Anteil die wahrgenommene Qualität oder Quantität des Fortbildungsangebots.

Zusammenfassend ermöglichte die Lehrerbefragung im Rahmen des Ländervergleichs 2011 einen Einblick in einige aktuelle Aspekte der beruflichen Arbeit von Deutsch- und Mathematiklehrkräften in der Primarstufe. Anhand einer für die Länder Deutschlands repräsentativen Stichprobe konnte dargestellt werden, wer mit welchem Ausbildungsgrad in der vierten Jahrgangsstufe unterrichtet. Darüber hinaus bieten die Ergebnisse zu den Fortbildungsaktivitäten und -bedarfen Anknüpfungspunkte für die Diskussion über die Lehreraus- und -fortbildung in den Ländern.

Literatur

- Avenarius, H. & Heckel, H. (2000). *Schulrechtskunde: Ein Handbuch für Praxis, Rechtsprechung und Wissenschaft*. Neuwied: Luchterhand.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y. M. (2010). Teachers' mathematical knowledge, cogni-

- tive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133–180.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010a). *TEDS-M 2008. Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010b). *TEDS-M 2008: Professionelle Kompetenz und Lerngelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich*. Münster: Waxmann.
- Borowski, A., Neuhaus, B. J., Tepner, O., Wirth, J., Fischer, H. E., Leutner, D., Sandmann, A. & Sumfleth, E. (2010). ProwiN: Das Professionswissen von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 341–349.
- Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (2009). *PISA 2006 Skalenhandbuch. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Fuchs, H.-W. (2007). *Bildung und Wissenschaft in der SBZ/DDR 1945-1989*. Hamburg: Universität der Bundeswehr.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to student achievement*. London: Routledge.
- Hessisches Lehrerbildungsgesetz vom 28.09.2011. HLbG. (GVBl. I S. 590).
- KMK (2011) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. *Lehrereinstellungsbedarf und Lehrereinstellungsangebot in der Bundesrepublik Deutschland: Modellrechnung 2010–2020*. Zugriff am 08.08.2012 unter http://www.kmk.org/fileadmin/pdf/Statistik/BERICHT_MODELL_RECHNUNG_online.pdf.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Lehrerfortbildung in Bayern. Bekanntmachung des Bayerischen Staatsministeriums für Unterricht und Kultus vom 30.08.2002. (KWMBL. I 2002 S. 260).
- Lipowsky, F. (2011). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und Weiterbildung. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zur Lehrerbildung* (S. 398–417). Münster: Waxmann.
- OECD. (2009). *Creating effective teaching and learning environments: First results from TALIS*. Paris: OECD.
- Richter, D., Kunter, M., Anders, Y., Klusmann, U., Lüdtke, O. & Baumert, J. (2010). Inhalte und Prädiktoren beruflicher Fortbildung von Mathematiklehrkräften. *Empirische Pädagogik*, 24, 151–168.
- Schorch, G. (2009). Grundschule. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtasik & W. Sacher (Hrsg.), *Handbuch Schule: Theorie – Organisation – Entwicklung* (S. 228–235). Bad Heilbrunn: Klinkhardt Verlag.
- Statistisches Bundesamt. (2011). *Fachserie II: Bildung und Kultur, Reihe I: Allgemeinbildende Schulen*. Wiesbaden: Statistisches Bundesamt.
- Timperley, H., Wilson, A., Barrar, H. & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Wellington, New Zealand: Ministry of Education.
- Verordnung über die Arbeitszeit der Lehrerinnen und Lehrer an staatlichen Schulen vom 1.7.2003. (HmbGVBl. S. 197).
- Verordnung über die Fortbildung der Lehrkräfte und Lehrer in besonderer Funktion an öffentlichen Schulen (Lehrerfortbildungsverordnung) vom 02.08.2005. LehrerFBVO. (Brem.GBl. S. 386).
- Verordnung zur Umsetzung des Hessischen Lehrerbildungsgesetzes vom 16.03.2005. HLbG-UVO.
- Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B. & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement (Issues & Answers Report, REL 2007-No. 033)*. Zugriff am 08.08.2012 unter http://ies.ed.gov/ncee/edlabs/regions/southwest/pdf/REL_2007033.pdf. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest.

Kapitel 11

Sprach- und Leseförderung

Petra Stanat, Sebastian Weirich und Susanne Radmann

Sprach- und Lesekompetenzen spielen für gelingende Bildungsprozesse eine Schlüsselrolle. Um von den Lerngelegenheiten zu profitieren, die der Unterricht bietet, müssen Schülerinnen und Schüler die in der Schule verwendete Instruktionssprache verstehen und aktiv nutzen können. Dabei steigen im Bildungsverlauf die sprachlichen Anforderungen, die sich Kindern und Jugendlichen stellen, kontinuierlich an. Obwohl systematische Analysen des Sprachgebrauchs im Unterricht weitgehend fehlen, kann mit einiger Sicherheit davon ausgegangen werden, dass er zunehmend durch Merkmale gekennzeichnet ist, die mit dem Konzept der Bildungssprache assoziiert werden (vgl. z. B. Bailey, Butler, Stevens & Lord, 2007; Cummins, 2000; Gogolin & Lange, 2011; Heppt, Dragon, Berendes, Stanat & Weinert, im Druck; Redder et al., 2011; Schleppegrell, 2004). Demnach sind mit bildungssprachlich geprägter Kommunikation im Vergleich zur Alltagskommunikation anspruchsvollere Anforderungen in den Bereichen Wortschatz, Grammatik und Organisation sprachlicher Ausführungen verbunden. Schülerinnen und Schüler müssen in die Lage versetzt werden, diese Anforderungen zu bewältigen, um in der Schule erfolgreich sein zu können.

Im Bereich Lesen erwerben Kinder in der Primarstufe zunächst basale Lesefertigkeiten, die es ihnen erlauben, Wörter zu rekodieren (Lautierung) und zu dekodieren (Sinnentnahme). Spätestens am Ende der Grundschulzeit wird dann von ihnen erwartet, dass sie ganze Texte flüssig lesen und verstehen können. Im weiteren Bildungsverlauf spielt Leseverständnis eine zentrale Rolle, da die Arbeit mit Texten in schulischen Lehr-Lern-Prozessen zunehmend an Bedeutung gewinnt. Schülerinnen und Schüler, die kaum in der Lage sind, Texte für den eigenen Wissenserwerb zu nutzen, haben Schwierigkeiten, ihre Bildungslaufbahn erfolgreich zu absolvieren.

Sprach- und Lesekompetenzen werden jedoch in besonderem Maße nicht nur in der Schule, sondern auch außerschulisch, vor allem im familiären Kontext erworben. Der Spracherwerb beginnt spätestens mit der Geburt und wird sowohl durch die Quantität als auch durch die Qualität der Lerngelegenheiten in der Umwelt beeinflusst. So konnten Hart und Risley (1995) in einer Längsschnittstudie mit 42 Familien unterschiedlicher sozialer Lage zeigen, dass Kinder aus Familien mit einem niedrigen sozioökonomischen Status (in der Literatur meist mit „SES“ für *socio-economic status* abgekürzt) im Alter von 13 Monaten bis drei Jahren deutlich anregungsärmere Interaktionen mit ihren Müttern erfuhren als Kinder aus Familien mit einem hohen sozioökonomischen Status. Mütter mit hohem SES benutzten pro Stunde beinahe viermal so viele Wörter, wenn sie mit ihren Kindern interagierten, als Mütter mit niedrigem SES (2 100 vs. 600 Wörter). Zudem war in Familien mit hohem SES ein größerer Variantenreichtum in den verwendeten Wörtern, Sätzen und Redewendungen zu beobachten. Diese

Merkmale der Interaktionen korrelierten mit den sprachlichen Kompetenzen der Kinder. Entsprechend konnten bereits im Alter von zwei Jahren bei Kindern aus Familien mit hohem SES deutlich weiterentwickelte Kompetenzen im Bereich des Wortschatzes und der grammatischen Strukturen nachgewiesen werden als bei Kindern aus sozial schwachen Familien (Hart & Risley, 1995; vgl. auch Hoff-Ginsberg, 2000).

Auch für die Entwicklung von Lesekompetenz spielt das familiäre Umfeld eine wichtige Rolle (vgl. z.B. Bus, Van IJzendoorn & Pellegrini, 1995; McElvany, 2008). So konnte in einer deutschen Längsschnittstudie gezeigt werden, dass solche lesebezogenen Prozessmerkmale der Familie, wie etwa Gespräche über Bücher, Anregungen zum Lesen oder die selbst eingeschätzte Fähigkeit der Eltern, ihre Kinder beim Lesen unterstützen zu können, die Entwicklung der Fähigkeit von Kindern, verstehend zu lesen, beeinflussen (McElvany, Becker & Lüdtke, 2009). Diese Prozessmerkmale waren zudem mit dem sozioökonomischen Status der Familie, dem Schulabschluss der Eltern und der zu Hause gesprochenen Sprache korreliert.

Die Entwicklung von Sprach- und Lesekompetenzen ist also abhängig von Lerngelegenheiten, die in sozial schwachen Familien oft eingeschränkt sind. Da Schülerinnen und Schüler mit Zuwanderungshintergrund häufig aus Familien mit einem niedrigen sozioökonomischen Status stammen (vgl. z.B. Segeritz, Walter & Stanat, 2010), sind auch bei ihnen die außerschulischen Lernbedingungen schon aus diesem Grund oft weniger günstig. Zusätzlich können bei diesen Kindern die Möglichkeiten des Zugangs zur deutschen Sprache stark begrenzt sein, etwa wenn im familiären Umfeld kaum Deutsch gesprochen wird und der Eintritt in die Kita erst spät oder gar nicht erfolgt. Diese Einschränkungen in den Gelegenheiten für den Erwerb von Sprach- und Lesekompetenzen in der schulischen Instruktionssprache, die mit einem niedrigen sozioökonomischen Status und einem Zuwanderungshintergrund verbunden sein können, tragen in erheblichem Maße zur Entwicklung von zugewanderungsbezogenen und sozialen Disparitäten im Bildungserfolg bei (z.B. Dubowy, Ebert, von Maurice & Weinert, 2008; vgl. auch Kapitel 8 und 9). Will man diese Disparitäten reduzieren, so müssen institutionelle Lerngelegenheiten geschaffen und optimiert werden, die es auch Kindern und Jugendlichen aus Familien mit Zuwanderungshintergrund und mit niedrigem SES ermöglichen, die für einen erfolgreichen Bildungsweg erforderlichen Sprach- und Lesekompetenzen zu entwickeln. Dies ist eine notwendige Voraussetzung, um Mindeststandards des schulischen Kompetenzerwerbs für alle Schülerinnen und Schüler in allen Fächern zu sichern.

Dieses Erfordernis ist in Deutschland spätestens nach der Publikation der Ergebnisse von PISA 2000 (Baumert et al., 2001) erkannt worden. Zu den sieben Handlungsfeldern, in denen sich die Kultusministerkonferenz (KMK) nach Erscheinen des Berichts über PISA 2000 vornahm, vorrangig aktiv zu werden, gehörten unter anderem „Maßnahmen zur Verbesserung der Sprachkompetenz bereits im vorschulischen Bereich“ sowie „Maßnahmen zur Verbesserung der Grundschulbildung und durchgängige Verbesserung der Lesekompetenz [...]“ (KMK, 2002). In den gemeinsamen Empfehlungen der Kultusministerkonferenz und des Bundesministeriums für Bildung und Forschung zu den Ergebnissen der internationalen Vergleiche im Rahmen von PIRLS/IGLU 2006 und PISA 2006 wurde die Optimierung der Förderung von Sprach- und Lesekompetenz erneut bekräftigt. Darin heißt es: „In der Sekundarstufe I werden die vorhandenen Ansätze für eine durchgehende Sprachförderung weiter ausgebaut und durch möglichst flächendeckende diagnostische Verfahren ergänzt. Um einen guten

Start zu gewährleisten, werden im Elementar- und Primarbereich die Aktivitäten zur Sprach- und Leseförderung verstärkt und mit einer gezielten Einbindung der Eltern verbunden.“ (KMK & BMBF, 2008, S. 7)

Tatsächlich ist in Deutschland im Bereich der Sprach- und Leseförderung in den letzten zehn Jahren sehr viel unternommen worden, wie verschiedene Bestandsaufnahmen von Verfahren der Sprachstandsdiagnostik und Maßnahmen der Sprachförderung der Länder zeigen (Bundesamt für Migration und Flüchtlinge, 2007; Lisker, 2011; Redder et al., 2011). Viele dieser Aktivitäten zielen darauf ab, sprachliche Kompetenzen bereits in der Kita intensiv zu fördern, um den Kindern einen erfolgreichen Übergang in die Schule zu ermöglichen. Aber auch in der Primarstufe wurden die Bemühungen, schwache Schülerinnen und Schüler in ihrer sprachlichen und schriftsprachlichen Entwicklung zu unterstützen, teilweise deutlich intensiviert.

Die vorliegenden Übersichten der verschiedenen Aktivitäten basieren in der Regel auf Angaben der Länder und beschreiben die zentralen Elemente der jeweiligen Diagnose- und Förderstrategien konzeptuell und strukturell. Sie verdeutlichen die große Vielfalt an Anstrengungen, die im Bereich der Sprach- und Leseförderung unternommen werden. Zusätzlich zur überwiegend ungeklärten Frage nach der Wirksamkeit dieser Maßnahmen (Darsow, Paetsch, Stanat & Felbrich, 2012; Paetsch, Wolf, Stanat & Darsow, im Druck; Redder et al., 2011) bleibt dabei jedoch offen, inwieweit die jeweiligen Maßnahmen zielgenau diejenigen Schülerinnen und Schüler erreichen, die Förderung benötigen. Dieser Frage soll im vorliegenden Kapitel nachgegangen werden.

11.1 Analyse von Angeboten der Sprach- und Leseförderung und ihrer Nutzung

Ziel dieses Kapitels ist es, im Sinne des Bildungsmonitorings den Ländern eine Rückmeldung darüber zu geben, inwieweit Kinder, die besondere Förderung im Bereich der Sprach- und Lesekompetenz¹ benötigen, in ihren Schulen ein entsprechendes Angebot vorfinden. Dabei geht es nicht darum, die Länder untereinander zu vergleichen oder gar Aussagen darüber zu treffen, welche Länder im Bereich der Sprach- und Leseförderung besonders erfolgreich sind. Da bislang nur wenige Studien zur Effektivität der verschiedenen Maßnahmen vorliegen, ist unklar, welche dieser Ansätze mehr oder weniger wirksam sind (vgl. z. B. Darsow et al., 2012; Paetsch et al., im Druck; Redder et al., 2011; Stanat, Becker, Baumert, Lüdtke & Eckhardt, 2012). Entsprechend kann aus den im Folgenden dargestellten Befunden keine Schlussfolgerung darüber abgeleitet werden, inwieweit die Schwerpunkte, die in den Ländern im Bereich der Sprach- und Leseförderung gesetzt werden, zielführend sind – diese Frage lässt sich nur mit Hilfe von systematischen Interventions- und Evaluationsstudien beantworten (Felbrich, Stanat, Paetsch & Darsow, 2012). Ausgehend von Befunden der Kompetenzmessungen, die im Fach Deutsch im Ländervergleich 2011 durchgeführt worden sind, soll in diesem Kapitel im ersten Schritt analysiert werden, ob diejenigen Schülerinnen und Schüler, die offenbar Unterstützung benötigen, diese zumindest potenziell erhalten könnten, da ihre Schulen entsprechende Angebote bereithalten. Es geht also im Folgenden zunächst um den Aspekt der bedarfsgerechten Versorgung.

¹ Je nach Fragestellung und verfügbarer Information werden Maßnahmen der Sprach- und Leseförderung im Folgenden teilweise gemeinsam, teilweise differenziert betrachtet. Dies wird jeweils entsprechend gekennzeichnet.

Im zweiten Schritt wird der Frage nachgegangen, inwieweit Kinder, die im Schuljahr 2010/2011 zusätzliche Sprach- oder Leseförderung erhielten, diese auch tatsächlich benötigten. Drittens wird untersucht, in welchem Maße es möglicherweise vom Geschlecht, der sozialen Herkunft oder dem Zuwanderungshintergrund abhängt, ob Kinder *bei gleichem Bedarf* zusätzliche Förderung erhalten. Diese beiden Fragen betreffen den Aspekt einer validen Diagnosepraxis. Wie die vorliegenden Bestandsaufnahmen zu Aktivitäten der Länder im Bereich der Sprach- und Leseförderung zeigen, werden sehr unterschiedliche Verfahren eingesetzt, um den Förderbedarf der Kinder und Jugendlichen zu bestimmen (vgl. z.B. Redder et al., 2011). Da im Primarbereich kaum standardisierte und validierte Instrumente zur Erfassung von Sprachförderbedarf zur Verfügung stehen, muss zumeist auf andere Verfahren zurückgegriffen werden. Dabei kommt der Beobachtung der Lehrkräfte eine zentrale Rolle zu (s. u.). Verschiedene Studien haben gezeigt, dass es Lehrkräften im Allgemeinen sehr gut gelingt, die Schülerinnen und Schüler in ihren Klassen in eine Leistungsrangreihe zu bringen (z.B. Südkamp, Kaiser & Möller, 2012), sie scheinen jedoch größere Probleme damit zu haben, das absolute Leistungsniveau von einzelnen Schülerinnen und Schülern nach kriterialen Maßstäben genau zu bestimmen (vgl. z.B. Artelt, Stanat, Schneider & Schiefele, 2001; Brunner, Anders, Hachfeld & Krauss, 2011). Zudem finden sich in der Literatur Hinweise darauf, dass Lehrerurteile durch Hintergrundmerkmale, wie etwa das Geschlecht der Kinder, den ethnischen Hintergrund der Familie oder den sozioökonomischen Status der Eltern, beeinflusst sein können (vgl. z.B. Ready & Wright, 2011). Die Befundlage zu solchen Effekten ist allerdings gemischt (vgl. z.B. Karing, Matthäi & Artelt, 2011; Rjosk, McElvany, Anders & Becker, 2011) und bezieht sich in der Regel nicht auf die Identifikation von Förderbedarf. Anhand von Daten des Ländervergleichs soll daher untersucht werden, inwieweit bestimmte Hintergrundmerkmale beeinflussen, ob Kinder Sprach- beziehungsweise Leseförderung erhalten oder nicht. Da die Stichprobengrößen für manche Teilgruppen, die dabei betrachtet werden sollen, innerhalb der Länder oft sehr klein sind, werden diese Analysen nicht für die einzelnen Länder getrennt, sondern für die länderübergreifende Gesamtstichprobe durchgeführt.

Dem Vorgehen in Kapitel 6 folgend, werden im vorliegenden Kapitel nur Schülerinnen und Schüler einbezogen, die zielgleich unterrichtet werden, für die die Bildungsstandards also gelten. Dabei werden je nach Analyse maximal drei Gruppen von Schülerinnen und Schülern unterschieden, die anhand der in Kapitel 3.2 beschriebenen Kompetenzstufenmodelle (KSM) gebildet wurden („Kompetenzstufengruppen“). In diesen Modellen markiert die Kompetenzstufe II den von der KMK definierten Mindeststandard. Bei der ersten Kompetenzstufengruppe, die für die folgenden Analysen gebildet wurde, handelt es sich um Kinder, die in mindestens einem der beiden Kompetenzbereiche im Fach Deutsch (Zuhören bzw. Lesen) den Mindeststandard verfehlen und im jeweils anderen Bereich maximal den Mindeststandard erreichen („Mindeststandard verfehlt“ bzw. „KS I“; 11.9 % der Schülerinnen und Schüler). Anhand der in Kapitel 3.2 dargestellten Beschreibungen der Kompetenzstufen ist davon auszugehen, dass diesen Kindern der erfolgreiche Übergang in die Sekundarstufe I nur unter Einsatz intensiver Fördermaßnahmen gelingen wird. Die zweite Gruppe von Kindern, die in den Analysen betrachtet wird, hat in beiden Kompetenzbereichen im Fach Deutsch lediglich die Mindeststandards erreicht („Mindeststandards erreicht“ bzw. „KS II“; 7.5 % der Schülerinnen und Schüler). Geht man davon aus, dass auch für diese Schülerinnen und Schüler das Erreichen der Regelstandards

angestrebt wird, wäre für sie ebenfalls verstärkte Förderung wünschenswert. Die dritte Gruppe schließlich umfasst Schülerinnen und Schüler, die sowohl im Bereich Zuhören als auch im Bereich Lesen mindestens den Anforderungen der Regelstandards entsprechen konnten („Regelstandards erreicht“ bzw. „KS III oder höher“; 59.7 % der Schülerinnen und Schüler). Kinder, die diesen drei Gruppen nicht eindeutig zugeordnet werden können, wurden aus den Analysen ausgeschlossen (20.9 % der Schülerinnen und Schüler)².

In Analysen, die der Frage nachgehen, welche Schülerinnen und Schüler Sprach- oder Leseförderung erhalten, werden die Ergebnisse für alle drei Kompetenzstufengruppen präsentiert (Abschnitte 11.5 und 11.6). Bei Analysen, die darauf abzielen, das Angebot der Länder im Bereich der Sprach- und Leseförderung zu charakterisieren (Abschnitte 11.2 bis 11.4), würde dieser Grad der Differenzierung jedoch in sehr unübersichtlichen Ergebnisdarstellungen resultieren. Daher werden hier jeweils die beiden schwächeren Kompetenzstufengruppen („KS I“ und „KS II“) in einer Gruppe zusammengefasst, die als „höchstens Mindeststandards erreicht“ beziehungsweise „höchstens KS II“ bezeichnet wird. Da die Befundmuster der Analysen für beide Differenzierungsgrade sehr ähnlich sind, ist der Informationsverlust, der mit dieser Zusammenfassung der Gruppen verbunden ist, gering.

11.2 Diagnose von Sprachförderbedarf

Im ersten Schritt soll der Frage nachgegangen werden, ob und in welcher Weise in den Ländern der Sprachförderbedarf von Kindern diagnostiziert wird. Hierzu wurde den Schulleiterinnen und Schulleitern der am Ländervergleich 2011 teilnehmenden Schulen folgende Frage gestellt: „Wie wird bestimmt, ob eine Schülerin / ein Schüler einen sprachlichen Förderbedarf hat?“ Bei dieser Frage waren Mehrfachnennungen möglich. Wie Tabelle 11.1 zeigt, besuchen nahezu alle Schülerinnen und Schüler der vierten Jahrgangsstufe eine Schule, in der in unterschiedlicher Weise ein möglicher Sprachförderbedarf bestimmt wird. Der Anteil der Kinder, bei denen dies nicht der Fall ist, beträgt in jedem der Länder weniger als 2 Prozent und ist somit zu vernachlässigen.

Nahezu an allen Schulen werden zur Entscheidung, ob Förderbedarf besteht, Beobachtungen der Lehrkräfte herangezogen. Bundesweit besuchen etwa 92 Prozent der Schülerinnen und Schüler eine Schule, in der dies der Fall ist. Auch Informationen von anderen Pädagoginnen und Pädagogen (76.1 %), Informationen aus schulärztlichen Untersuchungen (67.4 %) sowie Informationen der Eltern (63.6 %) werden von Primarschulen häufig genutzt, um Sprachförderbedarf festzustellen. Schwache Noten werden als Indikator für Sprachförderbedarf zwar seltener, absolut gesehen jedoch ebenfalls recht häufig verwendet (49.7 %).

Die Länder unterscheiden sich zum Teil darin, welche dieser einzelnen Informationen schwerpunktmäßig genutzt werden. So scheint etwa die schulärztliche Untersuchung für die Bestimmung sprachlichen Förderbedarfs insbesondere in Brandenburg, Schleswig-Holstein und Thüringen in nahezu allen Schulen eine sehr wichtige Rolle zu spielen, während in Baden-Württemberg, Bayern und

2 Ein Großteil der ausgeschlossenen Schülerinnen und Schüler (12.6 %) hat in einem der beiden Kompetenzbereiche die Stufe III erreicht und im jeweils anderen Kompetenzbereich die Stufe II. Die im Folgenden dargestellten Befundmuster ändern sich kaum, wenn man diese Schülerinnen und Schüler in die Analysen einbezieht und der Gruppe „Regelstandards erreicht“ zuordnet.

Tabelle 11.1: Anteile der Schülerinnen und Schüler in der vierten Jahrgangsstufe, die eine Schule besuchen, in der Sprachförderbedarf mit verschiedenen Verfahren bestimmt wird
 „Wie wird bestimmt, ob eine Schülerin / ein Schüler Ihrer Schule einen sprachlichen Förderbedarf hat? (Mehrfachnennungen möglich)“ (Frage im Schulleiterfragebogen)

Land	gar nicht		eigene Beobachtung (der Lehrkraft)		Information anderer Pädagogen		Informationen aus schülärztlicher Untersuchung		Informationen der Eltern		schwache Noten		standardisierte Tests		unstandardisierte Verfahren		keine Angabe	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Baden-Württemberg	1.2	(1.2)	91.0	(3.4)	62.1	(5.1)	49.5	(5.4)	58.4	(6.5)	56.4	(5.8)	72.9	(5.7)	9.8	(2.9)	4.4	
Bayern	1.4	(1.4)	95.9	(2.3)	80.4	(5.3)	42.9	(5.6)	53.7	(6.5)	66.8	(5.3)	55.0	(6.0)	40.5	(5.5)	1.4	
Berlin	0.0	(0.0)	90.9	(3.1)	78.8	(4.1)	68.4	(4.1)	58.9	(5.1)	54.4	(5.0)	86.5	(3.2)	32.6	(4.8)	3.6	
Brandenburg	0.0	(0.0)	93.7	(2.8)	84.3	(4.5)	90.0	(3.6)	82.2	(4.4)	39.4	(5.4)	57.0	(5.8)	24.2	(4.8)	2.6	
Bremen	0.0	(0.0)	92.5	(3.0)	82.5	(4.3)	85.8	(4.0)	56.4	(6.2)	14.5	(4.3)	80.4	(4.9)	27.4	(4.5)	3.6	
Hamburg	0.0	(0.0)	67.8	(4.1)	39.9	(4.5)	41.8	(4.7)	35.8	(4.8)	31.7	(4.0)	79.4	(3.7)	19.7	(3.5)	20.6	
Hessen	0.0	(0.0)	91.0	(3.5)	78.8	(5.4)	80.1	(4.7)	67.2	(5.5)	48.3	(6.7)	67.7	(5.8)	17.8	(5.0)	6.9	
Mecklenburg-Vorpommern	0.0	(0.0)	90.6	(3.7)	71.6	(5.6)	85.4	(4.2)	76.2	(5.5)	51.1	(6.4)	27.3	(4.4)	9.9	(3.7)	5.9	
Niedersachsen	0.0	(0.0)	90.5	(3.6)	83.9	(4.6)	72.5	(5.5)	60.1	(5.4)	34.1	(4.9)	36.7	(5.1)	28.3	(4.9)	5.4	
Nordrhein-Westfalen	0.0	(0.0)	93.0	(3.2)	74.2	(5.2)	75.1	(5.4)	66.6	(5.6)	48.0	(6.1)	69.3	(5.6)	31.6	(5.3)	7.0	
Rheinland-Pfalz	0.0	(0.0)	87.9	(5.7)	81.1	(6.0)	73.9	(6.5)	72.6	(6.6)	44.0	(6.6)	25.4	(4.6)	21.9	(5.1)	12.1	
Saarland	0.0	(0.0)	97.2	(2.0)	67.2	(6.0)	84.4	(4.6)	60.8	(6.5)	52.7	(5.4)	48.2	(6.3)	6.9	(3.1)	0.0	
Sachsen	0.0	(0.0)	90.9	(3.2)	74.6	(5.0)	83.9	(4.5)	69.0	(4.8)	44.5	(6.1)	32.0	(5.2)	22.0	(4.9)	8.0	
Sachsen-Anhalt	0.0	(0.0)	97.4	(1.9)	85.5	(4.4)	74.2	(5.2)	73.5	(5.0)	47.2	(4.9)	42.8	(5.4)	33.3	(5.5)	2.6	
Schleswig-Holstein	0.0	(0.0)	90.7	(3.6)	89.1	(3.9)	91.9	(2.5)	74.0	(4.0)	32.4	(6.3)	38.1	(5.6)	38.0	(5.4)	1.6	
Thüringen	1.0	(1.0)	97.4	(1.8)	93.8	(2.8)	90.2	(3.0)	84.8	(4.1)	56.3	(6.0)	34.1	(5.8)	25.3	(5.4)	1.2	
Deutschland	0.4	(0.3)	92.1	(1.1)	76.1	(1.8)	67.4	(2.0)	63.6	(2.1)	49.7	(2.1)	57.4	(2.0)	26.8	(1.6)	5.3	

Anmerkungen. Die Frage wurde von den Schulleiterinnen und Schulleitern der beteiligten Schulen beantwortet. Im Wortlaut konnten sie die folgenden Antwortalternativen ankreuzen, wobei Mehrfachnennungen möglich waren: a) gar nicht, b) eigene Beobachtung (der Lehrkraft), c) Informationen anderer Pädagogen (Hort, Kolleginnen und Kollegen, Erzieherinnen und Erzieher), d) Informationen aus schülärztlicher Untersuchung, e) Informationen der Eltern, f) schwache Noten, g) standardisierte Tests (z. B. Hamburger Schreibprobe, MSVK, HASE, BISC), h) unstandardisierte Verfahren (z. B. Beobachtungsverfahren SISMIK, SELDAK, Lernausgangslagenanalysen).

Hamburg nicht einmal die Hälfte der Kinder eine Schule besucht, in der diese Informationsquelle herangezogen wird.

Inwieweit standardisierte Tests zur Bestimmung des Sprachförderbedarfs eingesetzt werden, variiert zwischen den Ländern besonders stark. Während etwa in Mecklenburg-Vorpommern und in Rheinland-Pfalz weniger als 30 Prozent der Schülerinnen und Schüler eine Schule besuchen, in denen der Sprachförderbedarf mit einem standardisierten Verfahren bestimmt wird, ist dieses Vorgehen insbesondere in Berlin mit fast 87 Prozent, in Bremen mit etwa 80 Prozent und in Hamburg mit rund 79 Prozent deutlich weiter verbreitet. Besonders häufig wurden dabei die Hamburger Schreibprobe (May, 1998) und der Stolperwörtertest (Metze, 2003) genannt.

Unstandardisierte Verfahren hingegen werden zur Bestimmung des Sprachförderbedarfs insgesamt erheblich seltener herangezogen. Am weitesten verbreitet ist diese Art des Vorgehens in Bayern. Hier besuchen etwa 41 Prozent der Schülerinnen und Schüler eine Schule, in denen Informationen aus unstandardisierten Verfahren verwendet werden. Dabei handelt es sich nach Angabe der Schulleiterinnen und Schulleiter vor allem um die Beobachtungsverfahren SISMIK³ und SELDAK⁴, die in Kitas eingesetzt werden, also auf Beobachtungen von Erzieherinnen und Erziehern basieren (Ulich & Mayr, 2003, 2006).

Insgesamt scheinen in Primarschulen in Deutschland also unterschiedliche Informationsquellen herangezogen zu werden, um den sprachlichen Förderbedarf von Schülerinnen und Schülern zu bestimmen. In der Regel wird mehr als eine Informationsquelle genutzt. In nur zwei Ländern (Mecklenburg-Vorpommern und Rheinland-Pfalz) besucht mehr als die Hälfte der Schülerinnen und Schüler eine Schule, in der weder standardisierte Tests noch unstandardisierte Verfahren der Sprachstandsfeststellung eingesetzt werden. Es ist allerdings noch weitgehend ungeklärt, welche der verschiedenen Verfahren am besten geeignet sind, um möglichst präzise und effizient zu entscheiden, ob ein Kind im sprachlichen Bereich einer besonderen Unterstützung bedarf. Entsprechend lässt sich anhand der hier dargestellten deskriptiven Analysen auch keine Aussage darüber treffen, inwieweit die von den einzelnen Ländern gewählten Strategien zu validen Ergebnissen führen.

11.3 Breite des Angebots von Sprach- und Leseförderung

Welche Arten von Angeboten der Sprach- und Leseförderung in Primarschulen zur Verfügung stehen, wurde im Rahmen des Ländervergleichs ebenfalls mit dem Schulleiterfragebogen erfasst. Anknüpfend an frühere Studien, wie etwa PIRLS/IGLU oder PISA, enthielt der Fragebogen eine Liste der gängigsten Maßnahmen, und die Schulleiterinnen und Schulleiter wurden gebeten, alle in ihrer Schule umgesetzten Angebote anzugeben. Diese Angaben wurden anschließend zu vier Kategorien zusammengefasst (Förderkurse/Zusatzunterricht, spezielle Förderprogramme Lesen, Förderung Lesehäufigkeit, Büchereien). Unterhalb der Tabelle 11.2 ist angegeben, auf welche Angebote sich die einzelnen Kategorien jeweils beziehen.

3 Das Akronym SISMIK steht für *Sprachverhalten und Interesse an Sprache bei Migrantenkindern in Kindertageseinrichtungen*.

4 Das Akronym SELDAK steht für *Sprachentwicklung und Literacy bei deutschsprachig aufwachsenden Kindern*.

Tabelle 11.2: Anteile der Schülerinnen und Schüler in der vierten Jahrgangsstufe, die eine Schule mit verschiedenen Angeboten der Deutsch-, Sprach- und Leseförderung besuchen, nach Kompetenzstufengruppen
 „Gibt es an Ihrer Schule die folgenden Angebote für Schülerinnen und Schüler?“ / „Welche zusätzlichen Maßnahmen zur Lese- und Sprachförderung werden an Ihrer Schule umgesetzt?“

Land	Förderkurse / Zusatzunterricht			Spezielle Förderprogramme Lesen			höchstens KS II			Förderung Lesehäufigkeit			Büchereien		
	%	(SE)	% k. A.	%	(SE)	% k. A.	%	(SE)	% k. A.	%	(SE)	% k. A.	%	(SE)	% k. A.
Baden-Württemberg	92.3	(3.7)	3.5	8.0	(3.7)	3.5	41.8	(7.3)	3.5	93.5	(3.4)	3.5	93.5	(3.4)	3.5
Bayern	92.9	(4.0)	0.4	26.9	(7.0)	0.4	56.3	(8.3)	0.4	99.6	(0.6)	0.4	99.6	(0.6)	0.4
Berlin	93.1	(3.4)	6.9	29.6	(5.6)	6.9	50.8	(5.7)	6.9	93.1	(3.4)	6.9	93.1	(3.4)	6.9
Brandenburg	94.3	(3.1)	1.0	29.1	(6.1)	1.0	45.4	(6.8)	1.0	98.3	(1.2)	1.0	98.3	(1.2)	1.0
Bremen	92.7	(3.8)	5.7	36.6	(6.9)	5.7	93.7	(3.6)	5.7	94.3	(3.5)	5.7	94.3	(3.5)	5.7
Hamburg	71.3*	(6.3)	27.9	29.1*	(5.0)	27.9	29.4*	(5.3)	27.9	72.1*	(6.3)	27.9	72.1*	(6.3)	27.9
Hessen	94.8	(2.8)	5.2	37.1	(7.7)	5.2	38.8	(7.0)	5.2	94.8	(2.8)	5.2	94.8	(2.8)	5.2
Mecklenburg-Vorpommern	92.5	(4.9)	7.5	31.9	(7.2)	7.5	47.1	(7.5)	7.5	91.5	(5.0)	7.5	91.5	(5.0)	7.5
Niedersachsen	93.8	(3.3)	4.0	41.8	(6.6)	4.0	55.7	(6.8)	4.0	96.0	(2.5)	4.0	96.0	(2.5)	4.0
Nordrhein-Westfalen	87.9	(5.1)	10.6	42.6	(7.6)	10.6	46.3	(7.3)	10.6	89.4	(5.2)	10.6	89.4	(5.2)	10.6
Rheinland-Pfalz	82.1	(8.4)	14.8	25.6	(6.3)	14.8	44.7	(8.1)	14.8	85.2	(8.5)	14.8	85.2	(8.5)	14.8
Saarland	94.5	(4.4)	0.0	23.1	(6.6)	0.0	37.7	(7.3)	0.0	99.9	(0.2)	0.0	99.9	(0.2)	0.0
Sachsen	93.0	(3.1)	5.8	18.0	(5.7)	5.8	34.3	(7.1)	5.8	92.6	(2.6)	5.8	92.6	(2.6)	5.8
Sachsen-Anhalt	96.6	(2.4)	2.5	14.6	(4.0)	2.5	46.0	(6.4)	2.5	97.5	(2.1)	2.5	97.5	(2.1)	2.5
Schleswig-Holstein	92.2	(3.9)	0.5	42.4	(7.1)	0.5	23.2	(8.0)	0.5	93.3	(1.3)	0.5	93.3	(1.3)	0.5
Thüringen	98.0	(1.6)	0.9	25.1	(6.7)	0.9	44.2	(7.6)	0.9	96.1	(2.6)	0.9	96.1	(2.6)	0.9
Deutschland	90.8	(1.8)	6.5	31.3	(2.7)	6.5	46.2	(2.6)	6.5	92.8	(1.7)	6.5	92.8	(1.7)	6.5

Land	Förderkurse / Zusatzunterricht			Spezielle Förderprogramme Lesen			KS III oder höher			Förderung Lesefähigkeit			Büchereien		
	%	(SE)	% k. A.	%	(SE)	% k. A.	%	(SE)	% k. A.	%	(SE)	% k. A.	%	(SE)	% k. A.
Baden-Württemberg	91.7	(3.4)	4.3	11.8	(4.2)	4.3	43.5	(6.1)	4.3	93.4	(3.0)	4.3			
Bayern	91.9	(3.4)	1.7	24.1	(4.7)	1.7	49.9	(6.3)	1.7	98.3	(1.7)	1.7			
Berlin	98.5	(0.9)	1.5	34.8	(5.3)	1.5	56.5	(5.8)	1.5	98.5	(0.9)	1.5			
Brandenburg	93.1	(3.2)	3.4	29.2	(5.3)	3.4	40.3	(5.7)	3.4	94.7	(3.1)	3.4			
Bremen	96.6	(1.9)	2.0	44.9	(6.5)	2.0	96.6	(1.9)	2.0	98.0	(1.5)	2.0			
Hamburg	82.5	(3.9)	16.9	27.7	(5.0)	16.9	28.9	(5.2)	16.9	83.1	(3.8)	16.9			
Hessen	92.3	(3.6)	7.7	32.7	(6.6)	7.7	31.6	(5.0)	7.7	92.3	(3.6)	7.7			
Mecklenburg-Vorpommern	94.4	(2.9)	5.6	31.7	(5.5)	5.6	47.3	(5.8)	5.6	92.8	(3.3)	5.6			
Niedersachsen	93.2	(3.2)	6.0	46.6	(6.2)	6.0	51.6	(5.6)	6.0	94.0	(3.1)	6.0			
Nordrhein-Westfalen	91.8	(3.1)	5.0	33.5	(5.7)	5.0	42.1	(6.3)	5.0	95.0	(2.4)	5.0			
Rheinland-Pfalz	87.1	(4.6)	10.5	34.1	(5.5)	10.5	46.7	(6.1)	10.5	89.5	(4.5)	10.5			
Saarland	96.3	(2.4)	0.0	16.0	(4.7)	0.0	32.4	(6.5)	0.0	98.0	(2.0)	0.0			
Sachsen	89.5	(3.8)	9.0	23.4	(5.5)	9.0	34.5	(6.2)	9.0	89.8	(3.1)	9.0			
Sachsen-Anhalt	95.2	(2.8)	2.9	16.6	(4.8)	2.9	53.8	(6.0)	2.9	97.1	(2.1)	2.9			
Schleswig-Holstein	90.2	(3.8)	2.1	38.6	(5.6)	2.1	31.6	(5.5)	2.1	96.7	(2.2)	2.1			
Thüringen	97.0	(2.0)	1.5	23.3	(5.1)	1.5	39.1	(6.5)	1.5	96.1	(2.3)	1.5			
Deutschland	92.0	(1.3)	4.8	28.6	(2.0)	4.8	44.1	(2.1)	4.8	94.6	(1.0)	4.8			

Anmerkungen. Die Fragen wurden von den Schulleiterinnen und Schulleitern der beteiligten Schulen beantwortet. Die Kategorien umfassen jeweils mindestens eins der folgenden Angebote: „Förderkurse / Zusatzunterricht“ = Förderangebote für leistungsschwache Schülerinnen und Schüler im Fach Deutsch, Förderkurse für lese- und rechtschreibschwache Schülerinnen und Schüler und/oder Zusatzunterricht in Deutsch;

„Spezielle Förderprogramme Lesen“ = Spezielle Programme zur Förderung des Leseverständnisses, der Lesefähigkeit und/oder der phonologischen Bewusstheit;

„Förderung Lesefähigkeit“ = AG „Lesen“/„Literatur“, Einrichtung von Lesecubs, Bücherzeitung und/oder Projekt „Lesekoffer“;

„Büchereien“ = Schulbibliothek, Einrichtung von Klassenbüchereien und/oder Kooperation mit Bibliotheken.

KS = Kompetenzstufe; SE = Standardfehler; k. A. = „keine Angabe“.

*Die Befunde stehen aufgrund fehlender Angaben von über 25 Prozent unter Vorbehalt.

In Tabelle 11.2 ist dargestellt, welcher Anteil von Schülerinnen und Schülern in der vierten Jahrgangsstufe eine Schule besucht, in der mindestens eine der Maßnahmen angeboten wird, die der jeweiligen Kategorie zugeordnet ist. Um zu prüfen, ob gerade auch die schwächeren Schülerinnen und Schüler Zugang zu solchen Angeboten haben, werden die Ergebnisse pro Land für die oben beschriebenen zwei Gruppen differenziert, die anhand der im Zuhören und im Lesen erreichten Kompetenzstufen gebildet worden sind.

Wie die Ergebnisse deutlich zeigen, haben fast alle Schülerinnen und Schüler in Deutschland – zumindest potenziell – in ihren Schulen Zugang zu Förder- oder Zusatzkursen, die allgemein im Fach Deutsch oder speziell für lese- und rechtschreibschwache Kinder angeboten werden.⁵ Dies gilt für alle Kinder unabhängig vom erreichten Kompetenzniveau. So besuchen bundesweit etwa 91 Prozent der Kinder, die sich unter oder auf Kompetenzstufe II befinden, und 92 Prozent der Kinder, die mindestens Kompetenzstufe III erreichen, eine Schule mit einem solchen Angebot.

Nicht nur insgesamt deutlich weniger verbreitet, sondern auch über die Länder hinweg deutlich weniger gleichförmig verbreitet als Förder- oder Zusatzkurse im Fach Deutsch beziehungsweise für lese- und rechtschreibschwache Kinder sind spezielle Programme zur Förderung der Lesekompetenz (Förderung des Leseverständnisses, der Leseflüssigkeit oder der phonologischen Bewusstheit). In der Gruppe der Kinder, die höchstens die Mindeststandards erreichen, liegt der Verbreitungsgrad bundesweit bei ungefähr 31 Prozent und variiert zwischen 8 Prozent in Baden-Württemberg und mehr als 40 Prozent in Niedersachsen, Nordrhein-Westfalen sowie Schleswig-Holstein. Auch diese Art des Angebots finden Schülerinnen und Schüler, die höchstens Kompetenzstufe II erreichen, und Schülerinnen und Schüler, die die Anforderungen der Regelstandards erfüllen, in ihren Schulen ähnlich häufig vor.

In vielen Ländern sind Maßnahmen zur Förderung der Lesehäufigkeit deutlich weiter verbreitet als Programme zur Förderung der Lesekompetenz. In der Gruppe der Schülerinnen und Schüler, die höchstens die Mindeststandards erreichen, besuchen bundesweit 46 Prozent der Kinder eine Schule, die angegeben hat, mindestens eine der aufgeführten Maßnahmen zur Förderung der Lesehäufigkeit umzusetzen. Differenziert nach Ländern variiert der Anteil in dieser Kompetenzstufengruppe zwischen rund 23 Prozent in Schleswig-Holstein und 94 Prozent in Bremen. In der leistungsstärkeren Gruppe (KS III oder höher) haben zwischen etwa 29 Prozent (Hamburg) und 97 Prozent (Bremen) der Kinder in ihren Schulen potenziell Zugang zu dieser Art der Förderung.

Der Zugang zu Bibliotheken oder Klassenbüchereien schließlich scheint in fast allen Ländern nahezu flächendeckend gewährleistet zu sein. Auch zu dieser Art des Angebots haben alle Schülerinnen und Schüler weitgehend unabhängig von den im Fach Deutsch erreichten Kompetenzniveaus in ähnlichem Maße Zugang.

5 In denjenigen Ländern, in denen weniger als 90 Prozent der Schülerinnen und Schüler, die im Fach Deutsch höchstens die Mindeststandards erreicht haben, potenziell Zugang zu Förder- oder Zusatzkursen haben (Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz), sind die entsprechenden Fragen von den Schulleiterinnen und Schulleitern relativ häufig unbeantwortet geblieben. Da nicht festgestellt werden kann, inwieweit Schulen mit fehlenden Werten solche Kurse anbieten, muss offen bleiben, ob die Reichweite des Angebots in diesen Ländern tatsächlich geringer ist.

11.4 Angebot systematischer Sprachförderung

Um etwas genauer zu bestimmen, in welcher Form Schülerinnen und Schüler im Primarbereich Sprachförderung erhalten, wurden die Schulleiterinnen und Schulleiter zusätzlich gefragt, ob in ihren Schulen eine systematische Sprachförderung⁶ angeboten wird und ob diese im Regelunterricht oder zusätzlich zum Regelunterricht erfolgt. Dabei wurden vielfach beide Kategorien ausgewählt, so dass sich auch die Angaben in Tabelle 11.3 zeilenweise zu mehr als 100 Prozent addieren.

Die Ergebnisse zeigen, dass in acht Ländern (Bayern, Brandenburg, Mecklenburg-Vorpommern, Niedersachsen, Sachsen, Sachsen-Anhalt, Schleswig-Holstein, Thüringen) mehr als 25 Prozent der Schülerinnen und Schüler, die im Fach Deutsch höchstens die Kompetenzstufe II erreichen, eine Schule besuchen, in der keine systematische Sprachförderung durchgeführt wird. In den Schulen dieser Kinder wird möglicherweise weitgehend darauf vertraut, dass mit dem regulären Deutschunterricht eine ausreichende Unterstützung sprachlicher Entwicklung gewährleistet werden kann.

In sieben Ländern (Berlin, Brandenburg, Bremen, Nordrhein-Westfalen, Saarland, Sachsen-Anhalt und Thüringen) hat hingegen mindestens die Hälfte der Kinder, die im Fach Deutsch höchstens die Mindeststandards erreichen, in ihrer Schule Zugang zu systematischer Sprachförderung, die im Regelunterricht stattfindet. In nur wenigen Ländern liegt die entsprechende Quote bei unter 40 Prozent (Bayern und Niedersachsen).

Deutlich stärker variiert der Verbreitungsgrad von systematischer Sprachförderung, die zusätzlich zum Regelunterricht angeboten wird. In der Gruppe der schwächeren Schülerinnen und Schüler (höchstens KS II) besuchen in sieben Ländern (Brandenburg, Mecklenburg-Vorpommern, Rheinland-Pfalz, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen) jeweils weniger als 30 Prozent der Schülerinnen und Schüler eine Schule mit entsprechendem Angebot. In den Stadtstaaten Berlin, Bremen und Hamburg sowie in Hessen hingegen haben mehr als 50 Prozent der Schülerinnen und Schüler Zugang zu dieser Art des Angebots.⁷ Dies korrespondiert mit dem insgesamt besonders hohen Anteil von Schülerinnen und Schülern, die in den Stadtstaaten die Mindeststandards nicht erreichen.

In dieser Analyse ist zu beobachten, dass der potenzielle Zugang zu systematischer Förderung mit dem Kompetenzniveau der Schülerinnen und Schüler kovariiert. Dies ist vor allem für den Anteil der Schülerinnen und Schüler der Fall, deren Schulen überhaupt keine systematische Sprachförderung anbieten. Bundesweit liegt dieser Anteil in der Gruppe der Kinder, die die

6 Da keine allgemein akzeptierte Definition von systematischer Sprachförderung existiert, wurde offen gelassen, was unter diesem Begriff zu verstehen ist. Die Ergebnisse lassen also keine Rückschlüsse darauf zu, wie die jeweilige Sprachförderung angelegt ist, die Schulleiterinnen und Schulleiter als systematisch einstufen. Aufschlussreich ist jedoch, wie viele Schülerinnen und Schüler eine Schule besuchen, die nach Angaben der Schulleiterinnen und Schulleiter *keine* systematische Sprachförderung betreibt. Es dürfte kaum vorgekommen sein, dass Schulen, die ein solches Angebot bereithalten, dieses im Fragebogen nicht angegeben haben.

7 Differenziert man diese Ergebnisse für die drei Kompetenzstufengruppen, so ergibt sich das folgende Befundmuster: Innerhalb der Gruppe der Schülerinnen und Schüler, die im Fach Deutsch nicht die Mindeststandards erreichen (KS I), besuchen in den Ländern Berlin, Bremen, Hamburg, Hessen und Niedersachsen mehr als 50 Prozent der Kinder eine Schule, die Sprachförderung zusätzlich zum Regelunterricht anbietet; innerhalb der Gruppe der Schülerinnen und Schüler, die genau die Mindeststandards erreichen (KS II), ist dies in Berlin, Bremen, Hamburg und Hessen der Fall.

Tabelle 11.3: Anteile der Schülerinnen und Schüler in der vierten Jahrgangsstufe, die eine Schule mit systematischer Sprachförderung besuchen, nach Kompetenzstufengruppen „Findet an Ihrer Schule systematische Sprachförderung im Regelunterricht oder zusätzlich zum Regelunterricht statt?“ (Frage im Schulleiterfragebogen)

Land	im Regelunterricht			höchstens KS II			KS III oder höher							
	%	(SE)	%	zusätzlich zum Regelunterricht	keine systematische Sprachförderung	keine Angabe	im Regelunterricht	zusätzlich zum Regelunterricht	keine systematische Sprachförderung	keine Angabe				
	%	(SE)	%	(SE)	(SE)	%	(SE)	(SE)	(SE)	%				
Baden-Württemberg	42.1	(7.7)	45.3	(7.2)	21.3	(5.0)	3.5	36.5	(6.1)	39.2	(5.4)	32.9	(5.5)	4.3
Bayern	38.9	(8.9)	39.2	(7.9)	28.6	(7.9)	0.4	36.8	(6.6)	31.1	(5.5)	32.4	(6.2)	1.7
Berlin	63.1	(6.0)	52.6	(6.6)	8.7	(4.0)	6.9	54.2	(5.7)	47.0	(5.1)	21.7	(5.0)	1.5
Brandenburg	59.6	(7.0)	23.9	(5.8)	27.2	(6.4)	1.0	60.0	(6.0)	29.5	(5.0)	21.9	(4.4)	3.4
Bremen	63.6	(7.6)	60.8	(7.5)	4.7	(2.5)	5.7	61.7	(6.8)	58.8	(7.4)	11.1	(4.9)	2.0
Hamburg	49.2*	(6.7)	66.5*	(6.0)	3.6*	(2.3)	27.9	53.4	(5.5)	61.5	(5.0)	4.9	(2.4)	16.9
Hessen	44.3	(7.1)	54.1	(6.7)	15.4	(5.0)	5.2	48.3	(5.4)	41.8	(5.9)	20.2	(4.6)	7.7
Mecklenburg-Vorpommern	47.4	(6.8)	29.3	(6.5)	27.6	(5.7)	7.5	43.7	(5.6)	25.2	(4.9)	38.3	(5.7)	5.6
Niedersachsen	33.8	(7.8)	49.3	(7.9)	29.0	(7.3)	4.0	21.9	(4.5)	41.8	(6.3)	40.5	(6.1)	6.0
Nordrhein-Westfalen	55.1	(7.5)	47.7	(7.5)	14.2	(4.6)	10.6	46.7	(6.7)	36.4	(5.6)	28.1	(6.4)	5.0
Rheinland-Pfalz	48.1	(7.9)	28.9	(6.0)	22.7	(6.5)	14.8	42.7	(6.3)	20.9	(3.8)	35.7	(5.9)	10.5
Saarland	57.5	(7.5)	46.9	(7.1)	11.7	(4.2)	0.0	52.3	(5.9)	39.4	(5.5)	21.8	(5.3)	0.0
Sachsen	42.5	(7.1)	26.0	(6.0)	27.9	(6.9)	5.8	36.0	(5.8)	22.8	(4.9)	32.9	(6.6)	9.0
Sachsen-Anhalt	57.4	(6.0)	25.7	(8.5)	30.7	(6.0)	2.5	49.1	(6.3)	12.5	(4.0)	39.5	(6.5)	2.9
Schleswig-Holstein	49.2	(8.4)	26.3	(7.1)	32.6	(6.6)	0.5	34.0	(6.0)	23.0	(5.2)	47.1	(5.6)	2.1
Thüringen	58.5	(6.8)	24.5	(6.0)	28.6	(6.1)	0.9	54.2	(6.0)	17.0	(5.0)	35.0	(5.5)	1.5
Deutschland	48.1	(2.8)	43.5	(2.7)	20.4	(2.0)	6.5	40.8	(2.4)	34.4	(2.0)	31.3	(2.2)	4.8

Anmerkungen. KS = Kompetenzstufe; SE = Standardfehler. *Die Befunde stehen aufgrund fehlender Angaben von über 25 Prozent unter Vorbehalt.

Tabelle 11.4: Anteile der Schülerinnen und Schüler in der vierten Jahrgangsstufe, die eine Schule besuchen, in der ein konkretes Sprachförderkonzept verfolgt wird, nach Kompetenzstufengruppen

„Wird in Ihrer Schule ein konkretes Sprachförderkonzept verfolgt?“ (Frage im Schulleiterfragebogen)

Land	höchstens KS II				KS III oder höher			
	vorgegebenes Konzept	selbst entwickeltes Konzept	keine systematische Sprachförderung oder kein konkretes Konzept	keine systematische Sprachförderung oder kein konkretes Konzept	vorgegebenes Konzept	selbst entwickeltes Konzept	keine systematische Sprachförderung oder kein konkretes Konzept	keine systematische Sprachförderung oder kein konkretes Konzept
	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)
Baden-Württemberg	3.9 (2.9)	31.4 (7.0)	46.3 (7.7)	18.3	5.0 (3.0)	26.1 (5.1)	56.3 (6.4)	12.6
Bayern	8.2 (4.3)	19.5 (7.2)	66.0 (7.7)	6.2	4.0 (2.1)	14.1 (3.9)	76.4 (4.8)	5.5
Berlin	2.5 (1.7)	55.4 (6.2)	29.8 (5.5)	12.3	4.6 (2.5)	42.6 (5.2)	45.7 (5.5)	7.2
Brandenburg	2.2 (1.7)	14.1 (5.0)	76.3 (5.3)	7.4	3.0 (2.0)	13.4 (4.4)	73.8 (4.7)	9.8
Bremen	12.5 (4.9)	31.3 (6.6)	44.0 (6.7)	12.1	8.3 (3.4)	33.5 (6.4)	50.5 (6.2)	7.7
Hamburg	2.1* (1.6)	58.7* (6.9)	3.9*	35.4	1.6 (1.2)	60.4 (5.3)	16.1 (4.4)	21.8
Hessen	4.2 (2.9)	27.5 (6.3)	53.0 (7.2)	15.3	3.7 (2.2)	23.0 (5.1)	57.5 (6.1)	15.7
Mecklenburg-Vorpommern	4.8 (2.8)	14.7 (6.0)	66.4 (7.8)	14.1	8.0 (3.5)	10.5 (3.8)	68.9 (6.0)	12.6
Niedersachsen	7.0 (4.7)	28.5 (7.1)	55.5 (6.9)	9.1	2.3 (1.6)	27.9 (5.8)	58.6 (6.4)	11.2
Nordrhein-Westfalen	11.8 (4.6)	38.8 (7.3)	34.3 (6.2)	15.2	12.0 (3.8)	23.8 (4.9)	57.5 (5.9)	6.7
Rheinland-Pfalz	13.9 (4.8)	10.9 (3.5)	52.7 (8.1)	22.6	8.2 (2.7)	12.7 (4.1)	60.2 (6.2)	18.9
Saarland	19.8 (5.6)	10.9 (5.0)	58.2 (6.7)	11.1	10.5 (3.8)	10.6 (4.1)	63.4 (6.2)	15.4
Sachsen	3.1 (3.2)	7.2 (3.3)	79.8 (5.2)	9.9	0.7 (0.7)	5.3 (2.8)	79.6 (5.0)	14.4
Sachsen-Anhalt	0.0 (0.0)	5.2 (2.8)	81.8 (5.1)	13.0	0.0 (0.0)	6.4 (2.9)	84.7 (4.4)	8.9
Schleswig-Holstein	2.0 (1.7)	17.7 (6.3)	74.5 (6.4)	5.8	2.3 (1.7)	13.1 (4.2)	78.6 (4.4)	6.0
Thüringen	10.6 (6.1)	6.5 (3.6)	80.1 (6.7)	2.9	3.0 (2.0)	5.6 (2.6)	88.8 (3.4)	2.5
Deutschland	7.5 (1.6)	28.8 (2.6)	50.1 (2.6)	13.6	5.7 (1.1)	20.7 (1.6)	63.8 (1.9)	9.9

Anmerkungen. KS = Kompetenzstufe; SE = Standardfehler. *Die Befunde stehen aufgrund fehlender Angaben von über 25 Prozent unter Vorbehalt.

Regelstandards erreichen (KS III oder höher) bei 31 Prozent und damit um ungefähr 10 Prozentpunkte höher als in der Gruppe der Kinder, die höchstens die Mindeststandards erreichen (höchstens KS II). Besonders deutlich ist diese Differenzierung in Baden-Württemberg, Berlin, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz und Schleswig-Holstein, wo der entsprechende Unterschied mindestens 11 Prozentpunkte beträgt. Auch der potenzielle Zugang zu systematischer Sprachförderung, die zusätzlich zum Regelunterricht angeboten wird, kovariiert mit dem Kompetenzniveau der Schülerinnen und Schüler. In Deutschland besuchen insgesamt fast 44 Prozent der Kinder, die im Fach Deutsch höchstens die Mindeststandards erreichen, aber nur etwa 34 Prozent der Kinder, die den Anforderungen der Regelstandards entsprechen, eine Schule, die zusätzliche Sprachförderung anbietet.

Diejenigen Schulleiterinnen und Schulleiter, die angegeben haben, dass in ihrer Schule systematische Sprachförderung stattfindet, wurden im Schulleiterfragebogen zudem gefragt, ob dabei ein konkretes Sprachförderkonzept verfolgt wird. Die Antwortalternativen lauteten: „ja, wir verfolgen ein vorgegebenes Sprachförderkonzept“, „ja, wir verfolgen ein selbst entwickeltes Sprachförderkonzept“ und „nein“. Die Verteilungen der Antworten auf diese Frage sind in Tabelle 11.4 dargestellt. Dabei wurden Schulen ohne Sprachförderangebot und solche ohne konkretes Sprachförderkonzept zusammengefasst. Anhand der Ergebnisse wird deutlich, dass in Primarschulen nur relativ selten extern entwickelte Konzepte umgesetzt werden. Wenn ein konkretes Sprachförderkonzept verfolgt wird, dann wurde dieses von der Schule selbst entwickelt. Hierzu dürften die Schulleiterinnen und Schulleiter auch solche Konzepte gezählt haben, für die im jeweiligen Land zwar ein gewisser Rahmen vorgegeben wird, die aber von den Schulen konkretisiert und angepasst werden müssen.

Inwieweit Schulen konkrete Förderkonzepte verfolgen, scheint zwischen den Ländern allerdings stark zu variieren. Innerhalb der Gruppe von Schülerinnen und Schülern, die höchstens Kompetenzstufe II erreichen, liegt der Anteil derjenigen, die eine Schule mit vorgegebenem oder selbst entwickeltem Sprachförderkonzept besuchen, in Berlin, Bremen, Hamburg und Nordrhein-Westfalen bei mehr als 40 Prozent. In Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen hingegen sind die entsprechenden Anteile mit maximal rund 20 Prozent deutlich kleiner.

In insgesamt 11 Ländern hat mehr als die Hälfte der Kinder, die im Fach Deutsch schwache Leistungen erzielen (höchstens KS II), entweder keinen Zugang zu systematischer Sprachförderung oder die Sprachförderung, die in der Schule umgesetzt wird, folgt keinem konkreten Konzept. Allerdings hat bei dieser Frage in einigen Ländern insgesamt ein vergleichsweise hoher Anteil von Schulleiterinnen und Schulleitern keine Angabe gemacht.

11.5 Förderquoten

In den ersten Abschnitten dieses Kapitels wurde der Frage nachgegangen, inwieweit die Schulen der im Rahmen des Ländervergleichs 2011 getesteten Kinder bestimmte Angebote im Bereich der Lese- und Sprachförderung vorhalten. Der Versorgungsgrad mit schulischen Lerngelegenheiten lässt jedoch keine Rückschlüsse darüber zu, inwieweit diese auch tatsächlich genutzt werden (vgl. Helmke, 2006). Hierzu ist es erforderlich, für die einzelnen Schülerinnen und Schüler zu erfassen, ob sie eine bestimmte Förderung erhalten oder nicht. In

der vierten Jahrgangsstufe ist es allerdings schwierig, diese Information bei den Schülerinnen und Schülern selbst zu erfragen, da sie vermutlich häufig nicht wissen, welche ihrer schulischen Aktivitäten als Sprach- oder Leseförderung einzustufen sind. Ebenso ist unsicher, inwieweit die Eltern immer darüber informiert sind, ob und welche Art der Förderung ihre Kinder in den Schulen erhalten. Daher wurden im Rahmen des Ländervergleichs 2011 die Schulkoordinatorinnen und Schulkoordinatoren gebeten, auf der Schülerteilnahmeliste zu vermerken, ob die an der Studie teilnehmenden Kinder jeweils Sprach- oder Leseförderung erhalten. Um den Aufwand, zu einzelnen Kindern Informationen einzuholen und anzugeben, in Grenzen zu halten, haben wir uns bei dieser Erhebung auf eine Frage beschränkt. Diese bezog sich ausschließlich auf *zusätzliche* Sprach- oder Leseförderung, die außerhalb des regulären Unterrichts in der Schule erteilt wird. Es wurde also nicht erfasst, ob Schülerinnen und Schüler innerhalb des Regelunterrichts Sprach- oder Leseförderung erhalten. Da kaum einheitliche Vorstellungen darüber bestehen, was in den Regelunterricht integrierte Sprach- und Leseförderung umfasst, wäre diese Information auch nur bedingt aufschlussreich gewesen. Bei der Interpretation der im Folgenden präsentierten Ergebnisse muss jedoch berücksichtigt werden, dass der Anteil der Schülerinnen und Schüler, die gezielte Unterstützung erhalten, möglicherweise unterschätzt wird.

Da sich die Länder darin unterscheiden können, inwieweit in ihren Schulen der Schwerpunkt auf zusätzlicher Sprachförderung oder auf in den Regelunterricht integrierte Sprachförderung gelegt wird, ist es zudem nur bedingt informativ, die Länder untereinander zu vergleichen. So wäre es durchaus denkbar, dass ein relativ hoher Anteil von Kindern mit Förderbedarf in einem Land zwar keine zusätzliche Förderung außerhalb des Regelunterrichts erhält, innerhalb des Regelunterrichts jedoch durchaus effektiv unterstützt wird.⁸ Dennoch ist es aufschlussreich zu bestimmen, inwieweit die zusätzliche Förderung, die in den einzelnen Ländern angeboten wird, diejenigen Kinder erreicht, die im Fach Deutsch besondere Schwierigkeiten haben. Dieser Frage soll im Folgenden nachgegangen werden.

Tabelle 11.5 zeigt, welcher Anteil von Schülerinnen und Schülern in den drei differenzierten Kompetenzstufengruppen im Fach Deutsch jeweils zusätzliche Sprach- oder Leseförderung außerhalb des regulären Unterrichts erhält. Wie aufgrund der unterschiedlichen Förderstrategien zu erwarten ist, variieren die Anteile zwischen den Ländern erheblich. Innerhalb der Länder ist aber sehr klar zu erkennen, dass diejenigen Schülerinnen und Schüler, die nach den Daten des Ländervergleichs 2011 im Fach Deutsch ein besonders geringes Kompetenzniveau erreicht haben, deutlich häufiger zusätzliche Förderung erhalten als Schülerinnen und Schüler, die den Anforderungen der Regelstandards entsprechen. In der Gruppe der Schülerinnen und Schüler, die mindestens Kompetenzstufe III erreicht haben, ist der Anteil derjenigen, die zusätzliche Förderung erhalten, relativ klein (bundesweit 6.5 %). Dieses Muster spricht dafür, dass es insgesamt vergleichsweise gut gelingt, diejenigen Kinder zu identifizieren, die *keine* Förderung benötigen, wobei der Grad der Differenzierung zwischen den Ländern allerdings variiert. Insbesondere in Hamburg scheint die Identifizierung von Kindern mit unterschiedlich ausgeprägtem Förderbedarf gut zu gelingen. Während in der Gruppe der Schülerinnen und Schüler, die im

8 Auch zur Frage, inwieweit diese beiden Formen der Umsetzung von Sprach- und Leseförderung effektiv sein können, liegen derzeit noch keine gesicherten Erkenntnisse vor (vgl. Darsow et al., 2012; Paetsch, Wolf & Stanat, 2010; Paetsch et al., im Druck; Redder et al., 2011).

Tabelle 11.5: Anteile der Schülerinnen und Schüler in der vierten Jahrgangsstufe, die zusätzliche Sprach- oder Leseförderung erhalten, nach Kompetenzstufengruppen
 „Erhält das Kind zusätzliche Sprach- oder Leseförderung außerhalb des regulären Unterrichts innerhalb der Schule?“ (Frage in der Schülereteilnahmeliste)

Land	KS I			KS II			KS III oder höher		
	% Förde- rung	(SE)	% k. A.	% Förde- rung	(SE)	% k. A.	% Förde- rung	(SE)	% k. A.
Baden-Württemberg	41.6	(6.4)	0.0	33.5	(8.1)	0.0	7.8	(2.0)	0.0
Bayern	29.9	(6.2)	0.0	21.0	(6.2)	0.0	6.1	(2.0)	0.0
Berlin	26.2	(4.0)	3.4	14.7	(3.7)	2.7	4.8	(1.1)	1.2
Brandenburg	21.3	(4.4)	0.0	14.8	(4.6)	0.0	4.1	(1.0)	0.0
Bremen	40.0	(5.0)	0.6	26.6	(5.4)	1.0	8.8	(1.9)	1.4
Hamburg	49.9	(3.5)	0.9	26.8	(4.4)	1.5	5.4	(0.9)	2.4
Hessen	48.3	(6.2)	0.5	34.8	(6.3)	1.2	8.5	(1.3)	1.5
Mecklenburg- Vorpommern	26.8	(5.9)	1.3	20.4	(6.1)	1.8	7.0	(1.6)	0.8
Niedersachsen	39.0	(4.7)	0.0	22.9	(5.3)	0.0	3.8	(0.8)	0.1
Nordrhein-Westfalen	42.0	(6.1)	2.8	29.4	(6.7)	3.8	8.9	(2.4)	3.6
Rheinland-Pfalz	15.5	(3.4)	0.4	11.2	(3.6)	1.6	1.8	(0.6)	3.1
Saarland	25.8	(5.2)	0.0	17.1	(4.9)	0.0	3.9	(0.8)	0.0
Sachsen	18.0	(6.1)	0.0	14.0	(5.2)	0.0	5.6	(1.9)	2.3
Sachsen-Anhalt	32.3	(10.7)	0.0	15.0	(5.6)	0.0	4.6	(1.7)	0.0
Schleswig-Holstein	23.4	(4.8)	0.0	19.6	(5.7)	0.0	8.0	(1.3)	0.0
Thüringen	24.4	(4.6)	0.0	16.5	(5.1)	0.0	1.6	(0.5)	0.0
Deutschland	35.5	(2.2)	1.1	24.8	(2.4)	1.4	6.5	(0.7)	1.2

Anmerkung. KS = Kompetenzstufe; SE = Standardfehler; k. A. = „keine Angabe“.

Fach Deutsch die Mindeststandards verfehlen, etwa die Hälfte der Kinder zusätzliche Sprach- oder Leseförderung erhält, liegt der entsprechende Anteil in der Gruppe der Schülerinnen und Schüler, die zumindest die Mindeststandards erreichen, bei nur knapp 27 Prozent und in der Gruppe der Schülerinnen und Schüler, die den Regelstandards entsprechende Kompetenzen aufweisen, bei nur ungefähr 5 Prozent. Mehr als 30 Prozentpunkte Unterschied in der Förderquote für Kinder, die die Regelstandards erreichen einerseits und für Kinder, die die Mindeststandards verfehlen andererseits, sind auch in Baden-Württemberg, Bremen, Hessen, Niedersachsen und Nordrhein-Westfalen zu beobachten.

Die Sprach- oder Leseförderung, die in den Ländern zusätzlich zum Regelunterricht angeboten wird, erreicht also offenbar vor allem diejenigen Schülerinnen und Schüler, die eine solche Förderung tatsächlich benötigen. In den Begriffen der sogenannten ROC-Analyse der Signalentdeckungstheorie, die in der pädagogisch-psychologischen Diagnostik häufig verwendet wird (vgl. z. B. Goldhammer & Hartig, 2007), bedeutet dies, dass die Anteile sowohl der „richtig positiven“ Entscheidungen (es besteht Förderbedarf und es wird Förderung erteilt) als auch der „richtig negativen“ Entscheidungen (es besteht kein Förderbedarf und es wird keine Förderung erteilt) relativ hoch sind und zugleich relativ wenige „falsch positive“ Entscheidungen (es besteht kein Förderbedarf aber es wird Förderung erteilt) getroffen werden (vgl. Tabelle 11.6). Auf Basis der im Ländervergleich 2011 erhobenen Informationen kann jedoch nicht be-

Tabelle 11.6: Mögliche fehlerhafte und korrekte Diagnosen von Förderbedarf

	Tatsächlich besteht Förderbedarf	Tatsächlich besteht kein Förderbedarf
Kind erhält Förderung	„richtig positiv“	„falsch positiv“
Kind erhält keine Förderung	„falsch negativ“	„richtig negativ“

stimmt werden, inwieweit bei entsprechendem Angebot zusätzlicher Sprach- und Leseförderung Schülerinnen und Schüler, die diese Förderung benötigen würden, diese nicht erhalten („falsch negative“ Entscheidungen). So weist Tabelle 11.5 darauf hin, dass in keinem Land auch nur annähernd alle Schülerinnen und Schüler, die die Mindeststandards im Fach Deutsch verfehlen, zusätzliche Sprach- oder Leseförderung erhalten. Da wir nicht wissen, ob diese Kinder in anderer Weise systematisch gefördert werden (also z. B. mit in den Regelunterricht integrierter Förderung), muss an dieser Stelle offen bleiben, wie dieser Befund zu beurteilen ist.

11.6 Differenzielle Förderquoten

Ob ein Schüler oder eine Schülerin gezielte Förderung erhält, sollte im Idealfall allein davon abhängen, inwieweit er oder sie im jeweiligen Bereich ein geringes Kompetenzniveau erreicht und somit besonderen Förderbedarf aufweist. Wie bereits oben beschrieben, finden sich in der Literatur jedoch Hinweise darauf, dass in der Praxis Leistungsdiagnosen von Merkmalen der Schülerinnen und Schüler beeinflusst sein können. Dies kann dann entsprechend zu verzerrten Urteilen führen (vgl. z. B. Ready & Wright, 2011). Im Folgenden soll untersucht werden, ob bestimmte Gruppen von Schülerinnen und Schülern bei gleichem Kompetenzniveau eine größere beziehungsweise geringere Chance haben, Sprach- und Leseförderung zu erhalten. Da die Stichprobengrößen für einige dieser Teilgruppen innerhalb der Länder oft zu klein sind, um separat betrachtet zu werden, werden die Analysen länderübergreifend durchgeführt.

In Tabelle 11.7 ist für verschiedene Gruppen von Schülerinnen und Schülern dargestellt, ob sie eine zusätzliche Sprach- oder Leseförderung außerhalb des Regelunterrichts in der Schule erhalten. Dabei wird deutlich, dass sich die Förderquoten zwischen den Gruppen teilweise erheblich unterscheiden. Betrachtet man zunächst nur die Schülerinnen und Schüler, die im Fach Deutsch nicht die Mindeststandards erreichen (KS I), so erhalten Jungen häufiger Sprach- oder Leseförderung (38.3 %) als Mädchen (32.1 %). Noch größer sind die Unterschiede zwischen Kindern mit unterschiedlichem Zuwanderungshintergrund. Innerhalb der Gruppe der Kinder, die im Fach Deutsch die Mindeststandards verfehlen (KS I), erhalten fast 51 Prozent der selbst zugewanderten Kinder (erste Generation) Sprach- und Leseförderung, aber nur 43 Prozent der in Deutschland geborenen Kinder von Eltern, die beide im Ausland geboren sind (zweite Generation), und rund 41 Prozent der Kinder mit nur einem im Ausland geborenen Elternteil. Mit lediglich etwa 27 Prozent ist die Förderquote bei Kindern ohne Zuwanderungshintergrund in dieser Kompetenzstufengruppe besonders gering.

Ob ein Kind zusätzliche Sprach- oder Leseförderung erhält, scheint auch mit seinem sozioökonomischen Hintergrund zu kovariieren. Schülerinnen und

Tabelle 11.7: Anteile der Schülerinnen und Schüler in der vierten Jahrgangsstufe, die zusätzliche Sprach- oder Leseförderung erhalten, nach Kompetenzstufengruppe und verschiedenen Hintergrundmerkmalen

	KS I			KS II			KS III oder höher		
	ja % (SE)	nein % (SE)	k. A. %	ja % (SE)	nein % (SE)	k. A. %	ja % (SE)	nein % (SE)	k. A. %
Geschlecht									
Jungen	38.3 (2.7)	60.7 (2.6)	1.0	27.2 (3.2)	71.4 (3.2)	1.4	7.5 (0.8)	91.4 (1.0)	1.2
Mädchen	32.1 (2.6)	66.8 (2.5)	1.0	22.3 (2.9)	76.7 (2.9)	1.1	5.6 (0.8)	93.2 (0.9)	1.2
Zuwanderungshintergrund									
erste Generation	50.6 (6.7)	48.2 (6.8)	1.2	40.3 (12.2)	58.0 (12.3)	1.6	15.3 (4.6)	83.0 (4.8)	1.8
zweite Generation	43.0 (4.0)	55.5 (4.0)	1.5	34.0 (5.2)	63.5 (5.2)	2.4	11.9 (2.2)	87.1 (2.3)	1.0
ein Elternteil im Ausland geboren	41.4 (5.5)	58.1 (5.5)	0.5	29.6 (7.2)	69.2 (7.2)	1.3	8.2 (1.7)	90.2 (2.0)	1.5
kein Zuwanderungshintergrund	26.7 (2.3)	73.0 (2.3)	0.3	18.8 (2.4)	80.4 (2.4)	0.7	5.4 (0.7)	93.4 (0.8)	1.2
sozioökonomischer Hintergrund									
EGP V-VII	39.1 (3.5)	60.1 (3.4)	0.8	27.6 (5.0)	70.8 (5.0)	1.6	11.3 (2.2)	87.4 (2.4)	1.2
EGP III-IV	31.3 (3.7)	68.4 (3.7)	0.3	19.8 (4.1)	79.2 (4.0)	1.0	6.8 (1.1)	92.4 (1.2)	0.9
EGP I-II	29.9 (3.4)	69.7 (3.4)	0.4	21.2 (3.7)	77.9 (3.8)	0.9	4.9 (0.7)	93.8 (0.8)	1.3

Anmerkungen. EGP = Abkürzung für Erikson, Goldthorpe und Portocarero. EGP V-VII = z. B. Vorarbeiter, Meister, Techniker, die in manuelle Arbeitsprozesse eingebunden sind; un- und angelernte Arbeiter; EGP III-IV = z. B. klassische Büro- und Verwaltungsbetriebe mit Routinetätigkeiten; Verkaufs- und Servicetätigkeiten; Selbstständige aus manuellen Berufen; EGP I-II = z. B. mittleres Management; Beamte im mittleren, gehobenen und höheren Dienst; Hochschul- und Gymnasiallehrer; KS = Kompetenzstufe; SE = Standardfehler; k. A. = „keine Angabe“.

Schüler, die im Fach Deutsch nicht die Mindeststandards erreichen (KS I), erhalten etwas häufiger Förderung, wenn sie aus Familien mit einem vergleichsweise niedrigen sozioökonomischen Status stammen (EGP V-VII), als wenn ihre Familien den mittleren (EGP III-IV) oder höheren (EGP I-II) sozialen Schichten zuzuordnen sind (für eine genauere Definition der EGP-Klassen vgl. Kapitel 8). Während die Förderquote in der erstgenannten Gruppe bei etwa 39 Prozent liegt, beträgt sie in den anderen beiden Gruppen jeweils rund 30 Prozent.

Dieses Befundmuster wiederholt sich für Schülerinnen und Schüler, die die Mindeststandards erreicht haben (KS II), wie auch für Schülerinnen und Schüler, die die Anforderungen der Regelstandards erfüllen konnten (KS III oder höher), in ähnlicher Weise, allerdings sind die Förderquoten hier insgesamt etwas geringer. Auch in diesen Gruppen erhalten Jungen tendenziell häufiger als Mädchen, Kinder aus zugewanderten Familien deutlich häufiger als Kinder ohne Zuwanderungshintergrund sowie Schülerinnen und Schüler aus den niedrigeren Sozialschichten etwas häufiger als Schülerinnen und Schüler aus den mittleren und höheren Sozialschichten zusätzliche Sprach- und Leseförderung.

Nun sind aber auch innerhalb der drei Kompetenzstufengruppen die von den Kindern erreichten Leistungen natürlich nicht homogen. Entsprechend wäre es denkbar, dass zum Beispiel Jungen oder Kinder aus zugewanderten Familien innerhalb der Gruppen geringere Kompetenzniveaus erreicht haben als Mädchen oder Kinder ohne Zuwanderungshintergrund. In diesem Fall wäre es möglich, dass die differenziellen Muster in den Förderquoten gerechtfertigt sind. Zudem sind Zuwanderungshintergrund und sozioökonomischer Status der Familien nicht unabhängig voneinander, so dass die Betrachtung von Einzelzusammenhängen keine Aussage darüber zulässt, inwieweit tatsächlich *beide* Faktoren einen Einfluss darauf haben, ob ein Kind Förderung erhält oder nicht. Um diese Einschränkungen der Interpretierbarkeit der Befundmuster auszuräumen, wurden die Zusammenhänge zwischen den Hintergrundmerkmalen und der Teilnahme an zusätzlicher Förderung simultan mit logistischen Regressionsanalysen untersucht.

Mit den im Folgenden dargestellten Analysen wird der Frage nachgegangen, inwieweit die relative Chance, dass eine Schülerin oder ein Schüler zusätzliche Sprach- oder Leseförderung außerhalb des Unterrichts erhält, zum einen vom erreichten Kompetenzniveau und zum anderen darüber hinaus von den drei Hintergrundmerkmalen Geschlecht, Zuwanderungshintergrund und soziale Stellung der Familie bestimmt wird. Die Koeffizienten dieser Analysen sind in Tabelle 11.8 dargestellt. In den Modellen 2 bis 4 kennzeichnen die *Odds Ratios*, ob ein Kind bei gleichem Kompetenzniveau im Zuhören beziehungsweise im Lesen eine höhere (*Odds Ratio* größer 1) oder eine geringere (*Odds Ratio* kleiner 1) Chance hat, zusätzliche Sprach- oder Leseförderung zu erhalten, als ein Kind in der jeweiligen Referenzgruppe.

Modell 1a enthält zunächst nur die Zugehörigkeit der Schülerinnen und Schüler zu den drei Kompetenzstufengruppen, wie sie in den deskriptiven Analysen differenziert wurden. Die *Odds Ratios* zeigen, dass Schülerinnen und Schüler, die weder im Zuhören noch im Lesen die Mindeststandards erreichen (KS I), eine 7.93-mal so hohe Chance haben, Sprach- oder Leseförderung zu erhalten als Schülerinnen und Schüler, die in beiden Kompetenzbereichen die Anforderungen der Regelstandards erfüllen (KS III oder höher). Auch bei Kindern, die nur die Mindeststandards erreichen (KS II) ist die Chance, eine Sprach- oder Leseförderung zu erhalten, relativ zu den Kindern, die mindestens den Anforderungen der Regelstandards entsprechen (KS III oder höher), erhöht, allerdings nur um das 4.75-fache.

Tabelle 11.8: Vorhersage des Erhalts zusätzlicher Sprach- oder Leseförderung anhand von erreichten Kompetenzen im Zuhören und Lesen, Geschlecht, Zuwanderungshintergrund (ZH) und sozioökonomischem Hintergrund. Ergebnisse logistischer Regressionen

	Modell 1a			Modell 1b			Modell 2			Modell 3			Modell 4		
	b	(SE)	Odds Ratio												
Intercept	-2.65	(0.12)	0.07	-2.00	(0.09)	0.14	-2.13	(0.10)	0.12	-2.37	(0.11)	0.09	-2.49	(0.14)	0.08
Kompetenzstufengruppe (Referenz: KS III oder höher)															
KS I	2.07	(0.13)	7.93												
KS II	1.56	(0.14)	4.75												
Erreichte Kompetenzen															
Lesen				-0.64	(0.07)	0.53	-0.61	(0.07)	0.54	-0.64	(0.07)	0.53	-0.63	(0.09)	0.53
Zuhören				-0.37	(0.07)	0.69	-0.40	(0.08)	0.67	-0.26	(0.08)	0.77	-0.27	(0.10)	0.76
Geschlecht (Referenz: Mädchen)															
Jungen				0.24	(0.07)	1.28	0.25	(0.08)	1.28	0.25	(0.08)	1.28	0.28	(0.08)	1.32
Zuwanderungshintergrund (Referenz: kein ZH)															
erste Generation							1.02	(0.20)	2.77	0.70	(0.25)	2.01	0.66	(0.15)	1.94
zweite Generation							0.76	(0.12)	2.15	0.76	(0.15)	1.68	0.43	(0.13)	1.54
ein Elternteil im Ausland geboren							0.52	(0.15)	1.68	0.52	(0.15)	1.68	0.43	(0.13)	1.54
sozioökonomischer Hintergrund (Referenz: EGP I-II)															
EGP III-IV													0.06	(0.09)	1.06
EGP V-VII													0.33	(0.11)	1.39
R ²							.12		.12			.12			.12

Anmerkungen. EGP V-VII = z. B. Vorarbeiter, Meister, Techniker, die in manuelle Arbeitsprozesse eingebunden sind; un- und angelernte Arbeiter; EGP III-IV = z. B. klassische Büro- und Verwaltungsberufe mit Routinetätigkeiten; Verkauf- und Servicetätigkeiten; Selbstständige aus manuellen Berufen; EGP I-II = z. B. mittleres Management; Beamte im mittleren, gehobenen und höheren Dienst; Hochschul- und Gymnasiallehrer; EGP = Abkürzung für Erikson, Goldthorpe und Portocarero; b = unstandardisierter Regressionskoeffizient; SE = Standardfehler; R² = Determinationskoeffizient. fett: signifikante Regressionskoeffizienten ($p < .05$).

In den weiteren Analysen wird das von den Schülerinnen und Schülern erreichte Kompetenzniveau, wie in Modell 1b dargestellt, als kontinuierliche Variable kontrolliert, um den spezifischen Einfluss, der mit den einzelnen Hintergrundmerkmalen verbunden ist, möglichst präzise schätzen zu können. In Modell 2 wird zunächst die Rolle des Geschlechts analysiert. Der entsprechende Koeffizient zeigt, dass bei gleichen Kompetenzniveaus in den Bereichen Zuhören und Lesen Jungen eine 1.28-mal höhere Chance haben, Sprach- oder Leseförderung zu erhalten als Mädchen.

Auch mit dem Zuwanderungshintergrund sind zusätzliche Effekte verbunden (vgl. Modell 3). Bei gleichen Kompetenzniveaus im Fach Deutsch ist die relative Chance eines Kindes, eine Sprach- oder Leseförderung zu erhalten, um das 2.77-fache höher, wenn es selbst nach Deutschland zugewandert ist (erste Generation) als wenn es aus einer Familie ohne Zuwanderungsgeschichte stammt. Auch Schülerinnen und Schüler, die als Kinder zugewanderter Eltern in Deutschland geboren sind (zweite Generation) oder die nur einen im Ausland geborenen Elternteil haben, erhalten bei vergleichbaren Kompetenzniveaus eher Sprach- oder Leseförderung als Kinder aus Familien ohne Zuwanderungsgeschichte (*Odds Ratio* = 2.15 und *Odds Ratio* = 1.68).

Der Effekt des Zuwanderungshintergrunds reduziert sich allerdings, wenn – wie in Modell 4 geschehen – der sozioökonomische Hintergrund der Schülerinnen und Schüler kontrolliert wird, die Koeffizienten sind jedoch weiterhin substantiell (*Odds Ratio* = 2.01 für die erste Generation; *Odds Ratio* = 1.94 für die zweite Generation; *Odds Ratio* = 1.54 für einen im Ausland geborenen Elternteil). Mit dem sozioökonomischen Hintergrund ist zudem ein eigenständiger Effekt verbunden, allerdings nur für den Vergleich der niedrigsten Sozialschichtgruppe mit der höchsten Sozialschichtgruppe. So ist selbst bei vergleichbaren Kompetenzniveaus im Zuhören und Lesen die Chance eines Kindes, eine Sprach- oder Leseförderung zu erhalten, um das 1.39-fache höher, wenn der soziale Status seiner Eltern niedrig ist (EGP-Klassen V-VII), als wenn es aus einer Familie mit hohem Sozialstatus stammt (EGP-Klassen I-II).

11.7 Zusammenfassung und Diskussion

In diesem Kapitel wurde der Frage nachgegangen, inwieweit insbesondere Schülerinnen und Schüler, die im Fach Deutsch ein vergleichsweise niedriges Kompetenzniveau erreichen, in ihren Schulen Zugang zu verschiedenen Maßnahmen der Sprach- und Leseförderung haben. Die Befunde verweisen erneut darauf, dass in den Ländern sowohl in Bezug auf die Diagnose von Förderbedarf als auch in Bezug auf die Förderaktivitäten teilweise sehr unterschiedliche Strategien verfolgt werden (vgl. auch Bundesamt für Migration und Flüchtlinge, 2007; Redder et al., 2011).

Zur Feststellung von sprachlichem Förderbedarf werden in so gut wie allen Ländern nahezu flächendeckend Beobachtungen der Lehrkräfte herangezogen. Auch andere Informationsquellen (andere Pädagoginnen und Pädagogen, schulärztliche Untersuchungen, Eltern, Noten) werden relativ häufig verwendet, um zu entscheiden, ob bei einem Kind sprachlicher Förderbedarf besteht oder nicht. Inwieweit standardisierte Tests oder unstandardisierte Verfahren zur Ermittlung des Förderbedarfs eingesetzt werden, variiert dagegen erheblich zwischen den Ländern. Insgesamt besuchen jedoch in nur zwei Ländern mehr als die Hälfte der

Schülerinnen und Schüler eine Schule, in der weder standardisierte Tests noch unstandardisierte Verfahren der Sprachstandsfeststellung eingesetzt werden.

Insgesamt stellen Schulen ein vielfältiges Angebot der Sprach- und Leseförderung zur Verfügung. Fast alle Schülerinnen und Schüler besuchen Schulen, die Zusatzkurse im Fach Deutsch beziehungsweise Förderkurse für lese- und rechtschreibschwache Kinder anbieten. Auch der Zugang zu Büchereien ist nahezu für alle Schülerinnen und Schüler gewährleistet. Das Angebot von Maßnahmen zur Förderung der Lesehäufigkeit und vor allem von speziellen Förderprogrammen im Bereich Lesen variiert dagegen stärker zwischen den Ländern.

Inwieweit und in welcher Form in Primarschulen systematische Sprachförderung stattfindet, variiert ebenfalls zwischen den Ländern. In einigen Ländern besucht mehr als ein Viertel der Schülerinnen und Schüler, die höchstens die Mindeststandards erreichen, eine Schule, die nach eigener Angabe keine systematische Sprachförderung anbietet. In den anderen Ländern erfolgt eine systematische Förderung häufig im Regelunterricht, oft aber auch zusätzlich zum Regelunterricht, wobei vor allem die Reichweite des Angebots zusätzlicher Förderung zwischen den Ländern deutlich variiert. Insgesamt haben aber nach Angaben der Schulleiterinnen und Schulleiter bundesweit mindestens 73 Prozent der Kinder, die im Fach Deutsch höchstens die Mindeststandards erreichen, Zugang zu systematischer Sprachförderung, entweder im Regelunterricht oder zusätzlich zum Regelunterricht oder auch beides.

Weniger verbreitet sind dagegen Sprachförderangebote, die ein konkretes Konzept verfolgen. In elf Ländern besucht mehr als die Hälfte der Schülerinnen und Schüler, die im Fach Deutsch höchstens die Mindeststandards erreichen, eine Schule, in der entweder keine systematische Sprachförderung angeboten wird oder eine Sprachförderung, die sich an keinem konkreten Konzept orientiert. Da wenig Einigkeit darüber bestehen dürfte, was genau unter einem Sprachförderkonzept zu verstehen ist, sollte dieser Befund vorsichtig interpretiert werden. Es erscheint allerdings wenig plausibel, dass eine Schule tatsächlich ein konkretes Konzept (ob vorgegeben oder selbst entwickelt) verfolgt, dieses jedoch nicht als solches erkennt und im Fragebogen daher fälschlicherweise angegeben hat, keine an einem Konzept orientierte Sprachförderung umzusetzen. Da der umgekehrte Fall (eine Schule gibt an, ein Konzept zu verfolgen, das aber nur oberflächlich ein solches darstellt) vermutlich deutlich häufiger vorkommt, dürfte es sich bei der ermittelten Quote der Schülerinnen und Schüler, die eine Schule ohne konkretes Sprachförderkonzept besuchen, eher um eine Unterschätzung handeln.

Schülerinnen und Schüler, die tatsächlich eine zusätzliche Sprach- oder Leseförderung erhalten, haben in der Regel auch einen entsprechenden Bedarf, da sie im Fach Deutsch in den Bereichen Zuhören und Lesen höchstens die Mindeststandards erreichen. Umgekehrt nehmen nur relativ wenige Schülerinnen und Schüler, die im Fach Deutsch den Anforderungen der Regelstandards entsprechen, Sprach- oder Leseförderung in Anspruch. In fast allen Ländern erhält jedoch ein erheblicher Anteil von Schülerinnen und Schülern, die die Mindeststandards entweder nicht oder gerade erreichen (KS I bzw. KS II), keine zusätzliche Sprachförderung. Dies kann vielfältige Gründe haben. So wäre es möglich, dass die Schulen dieser Kinder keine Sprachförderung oder ausschließlich in den Regelunterricht integrierte Sprachförderung zur Verfügung stellen. Weiterhin wäre es denkbar, dass der Sprachförderbedarf der Schülerinnen und Schüler bislang unerkannt geblieben ist oder sie beziehungsweise ihre Eltern das

Angebot einer zusätzlichen Sprachförderung nicht nutzen wollen. Welche dieser Faktoren für den hohen Anteil der Schülerinnen und Schüler, die trotz eines entsprechenden Bedarfs keine zusätzliche Sprach- oder Leseförderung erhalten, entscheidend sind, lässt sich anhand der vorliegenden Daten nicht beantworten.

Die Frage, ob die Teilnahme eines Kindes an Sprach- oder Leseförderung nicht nur vom Kompetenzniveau abhängt, das es im Fach Deutsch erreicht hat, sondern darüber hinaus auch von seinem Geschlecht, seinem Zuwanderungshintergrund oder der sozioökonomischen Stellung seiner Familie beeinflusst wird, ist zu bejahen. Bei gleichem Kompetenzniveau in den Bereichen Zuhören und Lesen erhalten Jungen eher als Mädchen, Kinder aus zugewanderten Familien eher als Kinder aus Familien ohne Zuwanderungsgeschichte und Kinder aus sozial schwachen Familien eher als Kinder aus sozioökonomisch besser gestellten Familien zusätzliche Sprach- oder Leseförderung. Besonders ausgeprägt ist der Zusammenhang des Zuwanderungshintergrunds mit der relativen Chance der Schülerinnen und Schüler, zusätzliche Sprach- und Leseförderung zu erhalten. Bei gleichem Kompetenzniveau im Fach Deutsch ist diese Chance bei Kindern ohne Zuwanderungshintergrund relativ zu Kindern aus zugewanderten Familien erheblich geringer. Auch diesen Befundmustern können mehrere Ursachen zugrunde liegen. Eine Ursache könnte darin bestehen, dass Lehrkräfte mit den Hintergrundmerkmalen der Schülerinnen und Schüler bestimmte Erwartungen verbinden, die beeinflussen, inwieweit sie einen gegebenen Förderbedarf wahrnehmen (z. B. Ready & Wright, 2011). Da die Beobachtung durch Lehrkräfte bei der Bestimmung des Förderbedarfs eine zentrale Rolle spielt, könnte dies die differenziellen Muster erklären. Denkbar wäre aber auch, dass Mädchen, Kinder ohne Zuwanderungshintergrund und Kinder aus sozial besser gestellten Familien, die einen Förderbedarf aufweisen, das Angebot zusätzlicher Förderung seltener wahrnehmen. Die geringere Förderquote bei Kindern ohne Zuwanderungshintergrund schließlich könnte zudem darauf zurückzuführen sein, dass für die Förderung von Kompetenzen in Deutsch als Muttersprache weniger Konzepte zur Verfügung stehen als für die Förderung von Deutsch als Zweitsprache.

Insgesamt ergeben die in diesem Kapitel dargestellten Befunde ein Bild, das auf erhebliche Anstrengungen der Schulen im Bereich der Sprach- und Leseförderung hinweist. Sie bestätigen, dass in diesem Bereich sehr viel unternommen wird. Der Ausbau von sprachlicher Förderung, den sich die Länder nach Veröffentlichung der Ergebnisse von PISA 2000 beziehungsweise von PIRLS/IGLU 2006 und PISA 2006 zur Aufgabe gemacht haben (KMK, 2002; KMK & BMBF, 2008), ist offenbar erfolgt. Gleichzeitig deuten die Ergebnisse aber auch darauf hin, dass es häufig an systematischer Förderung fehlt, die sich an konkreten Konzepten orientiert. Bundesweit besucht ein relativ hoher Anteil von Kindern Schulen, die kein konkretes Sprachförderkonzept verfolgen. Dies wirft die Frage auf, inwieweit Schulen in der Lage sind, ein solches Konzept selbstständig zu entwickeln und umzusetzen. Die Förderung von Sprach- und Lesekompetenz ist eine hoch anspruchsvolle Aufgabe, auf die Lehrkräfte oft nicht ausreichend vorbereitet zu sein scheinen. So zeigen die Ergebnisse einer kürzlich veröffentlichten Umfrage, dass sich nur etwas mehr als die Hälfte der Lehrkräfte „für den Umgang mit sprachlichen Problemen von Schülern ausreichend qualifiziert“ fühlt (Becker-Mrotzek, Hentschel, Hippmann & Linnemann, 2012, S. 7). Entsprechend wünschen sich 48 Prozent der Lehrerinnen und Lehrer Unterstützung durch Kräfte, die in diesem Bereich speziell qualifiziert sind. Auch die Lehrkräftebefragung, die im Rahmen des Ländervergleichs 2011 durchgeführt worden ist, weist darauf hin, dass im Bereich der Sprachförderung besonde-

rer Qualifizierungsbedarf besteht (vgl. Kapitel 10). Fast 46 Prozent der befragten Lehrkräfte gaben an, zum Themenbereich der Sprachförderung einen hohen oder sehr hohen Fortbildungsbedarf zu haben.

Angesichts des in den meisten Ländern nicht unerheblichen Anteils von Schülerinnen und Schülern, die im Fach Deutsch die Mindeststandards verfehlen oder sie gerade erreichen (vgl. Kapitel 6), und angesichts der demografischen Entwicklung, die eher eine Zunahme des sprachlichen Förderbedarfs erwarten lässt (Expertenrat „Herkunft und Bildungserfolg“, 2011), erscheint es wichtig, die Aktivitäten im Bereich der Sprach- und Leseförderung kontinuierlich zu optimieren. Hierzu ist es erforderlich, die bestehenden Konzepte auf der Grundlage von Praxiserfahrungen und von wissenschaftlichen Erkenntnissen weiterzuentwickeln und die Effektivität verschiedener Ansätze im Rahmen von längsschnittlichen Interventions- und Evaluationsstudien zu überprüfen. Zudem bedarf es einer systematischen Qualifizierung des pädagogischen Personals, um dieses in die Lage zu versetzen, die Entwicklung der Sprach- und Lesekompetenzen ihrer Schülerinnen und Schüler gezielt und effektiv zu unterstützen.

Literatur

- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillman & M. Weiß (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich.
- Bailey, A. L., Butler, F. A., Stevens, R. & Lord, C. (2007). Further specifying the language demands of school. In A. L. Bailey (Hrsg.), *The language demands of school. Putting academic English to the test* (S. 103–156). New Haven, CT: Yale University Press.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillman, K.-J. & Weiß, M. (Hrsg.). (2001). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Becker-Mrotzek, M., Hentschel, B., Hippmann, K. & Linnemann, M. (2012). Sprachförderung in deutschen Schulen – die Sicht der Lehrerinnen und Lehrer. Ergebnisse einer Umfrage unter Lehrerinnen und Lehrern. Köln: Universität zu Köln. Zugriff am 30.07.2012 unter http://www.stiftung-mercator.de/fileadmin/user_upload/DAZ/Lehrerumfrage_Langfassung_final_30_05_12.pdf
- Brunner, M., Anders, Y., Hachfeld, A. & Krauss, S. (2011). Diagnostische Fähigkeiten von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 215–234). Münster: Waxmann.
- Bundesamt für Migration und Flüchtlinge. (2007). *Bundesweites Integrationsprogramm (§ 45 Aufenthaltsgesetz): Feststellung der Sprachförderangebote des Bundes und der Länder. Dokumentation*. Nürnberg: Bundesamt für Migration und Flüchtlinge (BaMF).
- Bus, A. G., Van IJzendoorn, M. H. & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65, 1–21.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon: Multilingual Matters.
- Darsow, A., Paetsch, J., Stanat, P. & Felbrich, A. (2012). Konzepte der Zweitsprachförderung: Eine Systematisierung. *Unterrichtswissenschaft*, 40, 64–82.
- Dubowy, M., Ebert, S., von Maurice, J. & Weinert, S. (2008). Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten: Ein Vergleich von Kindern mit und ohne Migrationshintergrund. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40, 124–134.
- Expertenrat „Herkunft und Bildungserfolg“. (2011). *Empfehlungen für Bildungspolitische Weichenstellungen in der Perspektive auf das Jahr 2020 (BW 2020)*. Baden-Württemberg: Ministerium für Kultur, Jugend und Sport.

- Felbrich, A., Stanat, P., Paetsch, J. & Darsow, A. (2012). Das Erkenntnispotential experimenteller Studien zur Untersuchung der Wirksamkeit von Sprachfördermaßnahmen. In B. Ahrenholz (Hrsg.), *Einblicke in die Zweitspracherwerbsforschung und ihre methodischen Verfahren. Reihe „DaZ-Forschung. Deutsch als Zweitsprache, Mehrsprachigkeit und Migration“*, Bd. 1 (S. 145–172). Berlin: De Gruyter.
- Gogolin, I. & Lange, I. (2011). Bildungssprache und Durchgängige Sprachbildung. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit* (S. 107–128). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Goldhammer, F. & Hartig, J. (2007). Testwertinterpretation. In H. Moosbrugger & A. Kelava (Hrsg.), *Test- und Fragebogenkonstruktion* (S. 165–192). Berlin: Springer.
- Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes.
- Helmke, A. (2006). *Unterrichtsqualität: Erfassen, Bewerten, Verbessern*. Seelze-Velber: Kallmeyer.
- Heppt, B., Dragon, N., Berendes, K., Stanat, P. & Weinert, S. (im Druck). Aspekte bildungssprachlicher Kompetenzen bei Kindern im Grundschulalter. *Diskurs Kindheits- und Jugendforschung*.
- Hoff-Ginsberg, E. (2000). Soziale Umwelt und Sprachlernen. In H. Grimm (Hrsg.), *Sprachentwicklung. Enzyklopädie der Psychologie* (C, III, 3, S. 463–494). Göttingen: Hogrefe.
- Karing, C., Matthäi, J. & Artelt, C. (2011). Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I – Eine Frage der Spezifität? *Zeitschrift für Pädagogische Psychologie*, 25, 159–172.
- KMK (2002) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. PISA 2000 – Zentrale Handlungsfelder. Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern. Beschluss der 299. Kultusministerkonferenz vom 17./18.10.2002. Zugriff am 30.07.2012 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2002/massnahmen.pdf>
- KMK & BMBF (2008) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland & Bundesministerium für Bildung und Forschung. Ergebnisse von PIRLS/IGLU 2006-I und PISA 2006-I: Gemeinsame Empfehlungen der Kultusministerkonferenz und des Bundesministeriums für Bildung und Forschung. Zugriff am 30.07.2012 unter http://www.bmbf.de/pubRD/gemeinsame_empfehlungen_iglu_und_pisa_06.pdf
- Lisker, A. (2011). *Additive Maßnahmen zur Sprachförderung im Kindergarten – Eine Bestandsaufnahme in den Bundesländern. Expertise im Auftrag des Deutschen Jugendinstituts*. München: Deutsches Jugendinstitut.
- May, P. (1998). *Diagnose orthographischer Kompetenz. Zur Erfassung der grundlegenden Rechtschreibstrategien mit der Hamburger Schreibprobe. Unter Mitarbeit von Ulrich Vieluf und Volkmar Malitzky*. Hamburg: Verlag für pädagogische Medien.
- McElvany, N. (2008). *Förderung von Lesekompetenz im Kontext der Familie*. Münster: Waxmann.
- McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41, 121–131.
- Metze, W. (2003). *Der Stolperwörter-Lesetest*. Zugriff am 30.07.2012 unter <http://www.wilfriedmetze.de/html/stolper.html>
- Paetsch, J., Wolf, K. M. & Stanat, P. (2010). *Förderung von Kindern und Jugendlichen aus Zuwandererfamilien*. Unveröffentlichte Expertise für den Expertenrat „Herkunft und Bildungserfolg“ im Auftrag des Ministeriums für Kultus, Jugend und Sport des Landes Baden-Württemberg. Berlin: AB Empirische Bildungsforschung, Freie Universität Berlin.
- Paetsch, J., Wolf, K. M., Stanat, P. & Darsow, A. (im Druck). Förderung von Kindern und Jugendlichen aus Zuwandererfamilien. *Zeitschrift für Erziehungswissenschaft, Sonderheft*.
- Ready, D. D. & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, 48, 335–360.
- Redder, A., Schwippert, K., Hasselhorn, M., Forschner, S., Fickermann, D. & Ehlich, K. (2011). *Bilanz und Konzeptualisierung von strukturierter Forschung zu „Sprach-*

- diagnostik und Sprachförderung*“. ZUSE-Berichte Band 2. Hamburg: Hamburger Zentrum zur Unterstützung der wissenschaftlichen Begleitung und Erforschung schulischer Entwicklungsprozesse (ZUSE). Zugriff am 30.07.2012 unter http://www.zuse.uni-hamburg.de/501publikation/zuse_berichte_02.pdf
- Rjosk, C., McElvany, N., Anders, Y. & Becker, M. (2011). Diagnostische Fähigkeiten von Lehrkräften bei der Einschätzung der basalen Lesefähigkeit ihrer Schülerinnen und Schüler. *Psychologie in Erziehung und Unterricht*, 58, 92–105.
- Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Mahwah, NJ: Routledge.
- Segeritz, M., Walter, O. & Stanat, P. (2010). Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 113–138.
- Stanat, P., Becker, M., Baumert, J., Lüdtke, O. & Eckhardt, A. G. (2012). Improving second language skills of immigrant students: A field trial study evaluating the effects of a summer learning program. *Learning and Instruction*, 22, 159–170.
- Südkamp, A., Kaiser, J. & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104, 743–762.
- Ulich, M. & Mayr, T. (2003). *sismik. Sprachverhalten und Interesse an Sprache bei Migrantenkindern in Kindertageseinrichtungen (Beobachtungsbogen und Begleitheft)*. Freiburg: Herder.
- Ulich, M. & Mayr, T. (2006). *seldak. Sprachentwicklung und Literacy bei deutschsprachig aufwachsenden Kindern (Beobachtungsbogen und Begleitheft)*. Freiburg: Herder.

Kapitel 12

Testdesign und Auswertung des Ländervergleichs: technische Grundlagen

Sebastian Weirich, Nicole Haag und Alexander Roppelt

Das folgende Kapitel beschreibt zentrale methodische Aspekte der Durchführung und statistischen Auswertung des Ländervergleichs 2011. Die hier verwendeten Verfahren sind eng an die Methoden angelehnt, die sich in ähnlichen großen Schulleistungsstudien wie PISA¹ (OECD, 2003), TIMSS² (Bonsen, Lintorf, Bos & Frey, 2008), PIRLS/IGLU³ (Hornberg, Bos, Buddeberg, Pothoff & Stubbe, 2007) oder NAEP⁴ (z.B. Allen, Donoghue & Schoeps, 2001) bewährt haben. Ohne die statistischen und wahrscheinlichkeitstheoretischen Grundlagen ausführlich darzustellen, soll hier beschrieben werden, wie diese Verfahren im Ländervergleich 2011 umgesetzt wurden. An verschiedenen Stellen wird auf weiterführende Literatur verwiesen.

12.1 Testdesign

Im Ländervergleich 2011 sollten sowohl möglichst alle Facetten der getesteten Kompetenzbereiche in den Fächern Deutsch und Mathematik als auch das gesamte Leistungsspektrum von Schülerinnen und Schülern in diesen Fächern erfasst werden. Um diese Ziele zu erreichen, ist es erforderlich, eine große Menge an Aufgaben einzusetzen, die eine Schülerin oder ein Schüler nicht in ihrer Gesamtheit bearbeiten kann. Daher wurde im Ländervergleich 2011, wie auch in PISA, TIMSS und PIRLS/IGLU, ein Multi-Matrix-Design umgesetzt (Böhme et al., 2010; Bonsen et al., 2008; Hornberg et al., 2007; siehe auch Kolen, 2006). In einem solchen Design bekommt jedes Kind ein Testheft, das nur eine Teilmenge der insgesamt zur Verfügung stehenden Aufgaben enthält. Dies bedeutet zugleich, dass jedes Item nur von einer Teilmenge aller am Ländervergleich teilnehmenden Kinder bearbeitet wird.

Die Erstellung der Testhefte für den Ländervergleich 2011 umfasste mehrere Teilschritte. Zunächst wurden Aufgaben zu Aufgabenblöcken zusammengefasst; im nächsten Schritt wurden mit diesen Aufgabenblöcken Testhefte erstellt. Bei der Testdurchführung erfolgte die Verteilung der Testhefte in den Klassen nach vorab festgelegten Kriterien. Nachfolgend sollen die genannten Teilschritte näher erläutert werden.

- 1 Das Akronym PISA steht für *Programme for International Student Assessment*.
- 2 Das Akronym TIMSS stand ursprünglich für *Third International Mathematics and Science Study*. Seit 2003 wird es in der Bedeutung *Trends in International Mathematics and Science Study* verwendet.
- 3 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie aber im Allgemeinen als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.
- 4 Das Akronym NAEP steht für *National Assessment of Educational Progress*.

12.1.1 Gruppierung der Aufgaben zu Aufgabenblöcken

Im Ländervergleich 2011 umfasst ein Aufgabenblock im Fach Deutsch jeweils 20 Minuten und in Mathematik jeweils 10 Minuten. In den Kompetenzbereichen Lesen und Zuhören besteht ein Aufgabenblock in der Regel aus nur einer Aufgabe, da die Auseinandersetzung mit dem Lese- beziehungsweise Hörstimulus bereits eine längere Bearbeitungszeit beansprucht. In Mathematik und Orthografie umfasst ein Block in der Regel mehrere Aufgaben, da die Rezeption der verwendeten Stimuli weniger zeitaufwändig ist. Während in den Kompetenzbereichen Lesen und Zuhören sowie in allen Kompetenzbereichen des Fachs Mathematik eine Aufgabe immer genau einem Block zugeordnet wird, kann eine Aufgabe im Kompetenzbereich Orthografie in mehreren Blöcken vorkommen. Der Grund hierfür ist, dass bestimmte Orthografieaufgaben, sogenannte Lückensatzdiktate, häufiger in Testheften eingesetzt werden sollten als andere Orthografieaufgaben (zu den Gründen siehe Kapitel 12.1). Tabelle 12.1 gibt einen Überblick über die Anzahl der im Ländervergleich 2011 eingesetzten Aufgabenblöcke, Aufgaben und Items pro Kompetenzbereich.

Tabelle 12.1: Anzahl der im Ländervergleich 2011 eingesetzten Aufgabenblöcke, Aufgaben und Items nach Kompetenzbereich

Fach	Kompetenzbereich	Anzahl Aufgabenblöcke	Anzahl Aufgaben	Anzahl Items
Deutsch	Lesen	9	11	80
	Zuhören	5	5	51
	Orthografie	12	12	152
Mathematik	Zahlen und Operationen	6	55	88
	Raum und Form	8	42	60
	Muster und Strukturen	6	27	48
	Größen und Messen	7	37	68
	Daten, Häufigkeit und Wahrscheinlichkeit	8	40	66

12.1.2 Zusammenstellung von Testheften

Im nächsten Schritt wurden aus den Aufgabenblöcken verschiedene Testhefte zusammengestellt. Dies erfolgte getrennt für die Fächer Deutsch und Mathematik. Ferner wurden innerhalb der beiden Fächer separate Teildesigns für Regel- und Förderschulen⁵ erstellt. Jedes einzelne Teildesign sollte dabei mehrere Kriterien erfüllen.

Die Testhefte sollten so zusammengestellt werden, dass jeder Aufgabenblock über alle Testhefte eines Teildesigns hinweg an jeder Position ungefähr gleich häufig vorkommt. Die Erstellung eines solchen balancierten Blockdesigns ist notwendig, da bekannt ist, dass die Aufmerksamkeit oder Motivation der Schülerinnen und Schüler im Verlauf der Testbearbeitung nachlässt (Robitzsch, 2009). Mit diesem Design wird angestrebt, Effekte der Itemposition im Testheft auszugleichen und diese in der Analyse statistisch zu kontrollieren.

⁵ Im Land Bremen erhielten auch Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, die an Regelschulen integrativ unterrichtet wurden, die für den Einsatz an Förderschulen entwickelten Testhefte.

Wie bereits erwähnt, wird jede Aufgabe nur von einer Teilmenge aller Kinder bearbeitet. Die Leistungsstreuung soll dabei in allen diesen Substichproben ähnlich groß sein, um homogene Schätzungen für die Itemparameter zu erhalten. Dies ließe sich am einfachsten realisieren, wenn die Testhefte zufällig auf die Schülerinnen und Schüler einer Testgruppe verteilt würden. Diese zufällige Verteilung ist jedoch nicht immer möglich, da Aufgaben zu Hörtexten und Lückensatzdiktate immer der gesamten Klasse gemeinsam vorgelegt werden müssen, da sie entweder von einer CD abgespielt oder durch die Testleiter vorgelesen werden. Anders als Leseaufgaben können sie daher nicht innerhalb einer Klasse zufällig verteilt werden. Dies würde bedeuten, dass die Substichproben, die die einzelnen Aufgaben in den Bereichen Zuhören und Orthografie bearbeitet haben, aus weniger Klassen stammen als die Substichproben für die einzelnen Leseaufgaben. Damit würde der auf Unterschiede zwischen Klassen zurückzuführende Anteil der Populationsvarianz für Höraufgaben und Lückensatzdiktate schlechter abgebildet als etwa für Leseaufgaben (Luke, 2004). Um trotz der unterschiedlichen Administrationsbedingungen eine vergleichbare Präzision bei der Schätzung der Populationsvarianz für Lese-, Zuhör- und Lückensatzaufgaben zu erreichen, wurden in den Deutsch-Testheften für die Regelschulen die Zuhör- und Lückensatzdiktataufgaben häufiger eingesetzt als die Leseaufgaben und die anderen Aufgaben des Kompetenzbereichs Orthografie.

Die Testhefte sollten des Weiteren so miteinander „verknüpft“ oder „verlinkt“ sein, dass die Kompetenzwerte aller Schülerinnen und Schüler ohne zusätzliche Annahmen auf einer gemeinsamen Skala abgebildet werden können, auch wenn sie unterschiedliche Testhefte bearbeitet haben (Kolen, 2006; Robitzsch, 2009). Die Verlinkung der Testhefte erfolgt direkt oder indirekt über gemeinsame Aufgabenblöcke. Sind wiederum alle Aufgabenblöcke direkt oder indirekt miteinander verknüpft, so sind es auch die Testhefte. Eine direkte Verbindung zweier Aufgabenblöcke liegt vor, wenn sie gemeinsam im selben Testheft auftreten. Eine indirekte Verbindung ist gegeben, wenn jeder dieser Aufgabenblöcke gemeinsam mit einem identischen dritten Aufgabenblock in einem zusätzlichen Testheft auftritt. Dabei müssen sowohl die einzelnen Aufgabenblöcke innerhalb eines Kompetenzbereichs miteinander verbunden werden als auch die Aufgabenblöcke über Kompetenzbereiche hinweg.

Teildesigns

Im Ländervergleich 2011 wurden insgesamt fünf Teildesigns verwendet. Tabelle 12.2 gibt die Anzahl und die Länge der eingesetzten Testhefte in jedem Teildesign an.

Tabelle 12.2: Anzahl und Länge der Testhefte nach Teildesign pro Fach im Ländervergleich 2011

Fach	Teildesign	Eingesetzt in	Kompetenzbereiche	Anzahl Blöcke je Testheft	Anzahl Testhefte
Deutsch	1. Teildesign	Regelschulen	Lesen, Zuhören	4 à 20 Minuten	20
	2. Teildesign	Regelschulen	Lesen, Zuhören, Orthografie	4 à 20 Minuten	20
	3. Teildesign	Förderschulen	Lesen, Zuhören	2 à 20 Minuten	4
Mathematik	1. Teildesign	Regelschulen	Alle	8 à 10 Minuten	35
	2. Teildesign	Förderschulen	Alle	4 à 10 Minuten	3

Im Fach Deutsch wurden drei Teildesigns entwickelt, die sich zum einen in den getesteten Kompetenzbereichen und zum anderen in den Schularten, in denen sie zum Einsatz kamen, unterscheiden. Die ersten beiden Teildesigns wurden für den Einsatz in Regelschulen entwickelt und umfassten jeweils 20 Testhefte. Das erste Teildesign enthielt sämtliche in Tabelle 12.1 aufgeführten Aufgabenblöcke der Kompetenzbereiche Lesen und Zuhören. In diesem Teildesign wurden pro Testheft zwei Blöcke aus dem Kompetenzbereich Lesen und zwei Blöcke aus dem Kompetenzbereich Zuhören eingesetzt. In einem zweiten Teildesign waren zusätzlich Aufgaben des Kompetenzbereichs Orthografie enthalten. Für beide Teildesigns betrug die pro Testheft insgesamt veranschlagte Bearbeitungszeit 80 Minuten. Im Teildesign für Förderschulen kamen die zwei leichtesten Blöcke des Kompetenzbereichs Lesen und die zwei leichtesten Blöcke des Kompetenzbereichs Zuhören zum Einsatz. Sie wurden in vier Testheften mit jeweils zwei Blöcken je Testheft verwendet, was einer vorgesehenen Bearbeitungszeit von 40 Minuten pro Testheft entspricht.

Im Fach Mathematik wurden zwei Teildesigns entwickelt, die jeweils Aufgaben aller fünf inhaltlichen Kompetenzbereiche (Zahlen und Operationen, Raum und Form, Muster und Strukturen, Größen und Messen sowie Daten, Häufigkeit und Wahrscheinlichkeit) enthielten. Das erste Teildesign umfasste 35 Testhefte mit einer Bearbeitungszeit von 80 Minuten, die in Regelschulen eingesetzt wurden. Sie enthielten Aufgaben zu jeweils drei, vier oder fünf unterschiedlichen inhaltlichen Kompetenzbereichen. Im Teildesign für Förderschulen wurde für jeden der fünf Kompetenzbereiche der leichteste Block genutzt. Diese Blöcke wurden in drei Testheften mit jeweils vier Blöcken pro Testheft eingesetzt. Ebenso wie im Fach Deutsch umfasste die Bearbeitungszeit auch hier 40 Minuten.

Verteilung der Testhefte in den Klassen

Im letzten Teilschritt wurden die im Testdesign definierten Testheftversionen auf die Schülerinnen und Schüler der getesteten Klassen verteilt. Diese Verteilung erfolgte in den Fächern Deutsch und Mathematik auf unterschiedliche Weise. Wie oben beschrieben, konnten im Fach Deutsch nicht alle Testheftversionen zufällig den Schülerinnen und Schülern in einer Klasse zugewiesen werden. Daher wurden jeweils zwei Testheftversionen zu einer Testheftgruppe zusammengefasst. Die beiden Testhefte einer Testheftgruppe enthielten identische Zuhör- und Lückensatzdiktataufgaben, jedoch unterschiedliche Lese- und übrige Orthografieaufgaben. Die Testheftgruppen wurden in gleicher Häufigkeit über alle am Ländervergleich 2011 teilnehmenden Klassen verteilt. Da in Mathematik keine Aufgaben eingesetzt wurden, die von allen Schülerinnen und Schülern einer Klasse gleichzeitig bearbeitet werden mussten, konnten hier alle Testheftversionen eines Teildesigns in jeder Klasse eingesetzt werden. Die Verteilung der Testhefte auf die Schülerinnen und Schüler einer Klasse erfolgte daher zufällig und insgesamt in gleicher Häufigkeit über alle am Ländervergleich 2011 teilnehmenden Schülerinnen und Schüler.

12.2 Skalierung

Wie in anderen vergleichbaren Schulleistungsstudien beruht die Bestimmung der Kompetenzwerte auf Modellen der probabilistischen Testtheorie oder *Item Response Theory* (IRT). Ziel dieser Modelle ist es, von den beobachteten Antworten der Schülerinnen und Schüler auf ihre latenten, also nicht beobachtbaren, Kompetenzstände zu schließen. Ausführliche Darstellungen finden sich dazu etwa in Embretson und Reise (2000) oder in Rost (2004). Jeder Kompetenzbereich der beiden Fächer Deutsch und Mathematik definiert jeweils eine Skala, auf der für jede Person modellbasiert ein bestimmter Leistungswert angenommen wird. Das dabei zugrunde gelegte Rasch-Modell aus der Familie der IRT-Modelle soll folgende Anforderungen erfüllen:

- Die Kompetenzwerte der Personen sollen ohne systematische Verzerrungen mit der größtmöglichen Genauigkeit geschätzt werden. Wie in Studien im Bereich des Systemmonitorings üblich, steht dabei nicht die Genauigkeit individueller Kompetenzwerte im Vordergrund, sondern die Genauigkeit auf Gruppenebene aggregierter Statistiken, wie zum Beispiel Mittelwerte und Standardabweichungen der Gesamtpopulation oder von Teilpopulationen.
- Parameter der Population werden anhand von Kennwerten der Stichprobe geschätzt. Die durch Zufallsprozesse bedingte Unsicherheit der Stichprobenkennwerte soll ohne systematische Verzerrungen geschätzt werden.
- Ebenso wie die Kompetenzausprägung unbeobachtet ist und daher latent modelliert werden muss, wird auch die Korrelation der Kompetenzbereiche latent modelliert. Die Korrelation der individuellen Kompetenzwerte soll die „wahre“ Korrelation der Kompetenzbereiche in der Population unverzerrt schätzen.

Um dies zu gewährleisten, wird die Bestimmung der Kompetenzwerte in einem mehrstufigen Verfahren realisiert:

1. Kalibrierung der Items zur Gewinnung von Itemschwierigkeitsparametern;
2. Bestimmung von Personenparametern (*Plausible Values*);
3. Bestimmung einer Transformationsvorschrift für die Berichtsmetrik;
4. Bestimmung der Standardfehler für die relevanten Statistiken.

Für die Analysen wurden verschiedene Computerprogramme verwendet. ACER ConQuest (Wu, Adams, Wilson & Haldane, 2007) kam bei den Schritten 1 und 2 zum Einsatz, die Software R (R Development Core Team, 2011) bei den Schritten 1, 2 und 3 sowie das Programm WesVar (Westat, 2000) bei der Bestimmung der Standardfehler in Schritt 4. Die genannten Analyseschritte werden im Folgenden näher erläutert.

Zur Qualitätssicherung der in diesem Band berichteten Befunde wurden alle Auswertungsschritte von zwei Personen unabhängig voneinander durchgeführt und die jeweiligen Ergebnisse der Analysen miteinander abgeglichen.

12.2.1 Kalibrierung der Items

Kalibrierung bezeichnet in der IRT die modellbasierte Bestimmung von Itemschwierigkeitsparametern. Bei der Kalibrierung der Items im Ländervergleich 2011 wurde für jeden Kompetenzbereich ein separates eindimensionales Modell zugrunde gelegt. Die Itemparameter eines Kompetenzbereichs definieren dabei

die jeweilige Skala des Kompetenzbereichs. Alle Schülerinnen und Schüler, die mindestens ein Item im jeweiligen Kompetenzbereich beantwortet hatten, wurden in die Kalibrierung einbezogen.

Im Fach Mathematik wurden einzelne Items mehreren Kompetenzbereichen zugeordnet, da für die Lösung dieser Items Kompetenzen aus unterschiedlichen Bereichen erforderlich sind (genauere Ausführungen zu dieser Datenstruktur folgen in Kapitel 12.2.2). In der Kalibrierung wurden diese Items mehrfach verwendet, also jeweils für alle Kompetenzbereiche, denen sie zugeordnet sind. Wurde ein Item beispielsweise den Kompetenzbereichen „Zahlen und Operationen“ und „Größen und Messen“ zugeordnet, werden für dieses Item im Zuge der Kalibrierung zwei Schwierigkeitsparameter geschätzt, die sich geringfügig unterscheiden können: ein Parameter, der seine Schwierigkeit auf der Skala des Kompetenzbereichs „Zahlen und Operationen“ beschreibt, und ein Parameter, der seine Schwierigkeit auf der Skala des Kompetenzbereichs „Größen und Messen“ beschreibt. Um Aussagen zur globalen mathematischen Fähigkeit der Schülerinnen und Schüler treffen zu können, wurden zusätzlich alle Items in einem eindimensionalen Modell gemeinsam kalibriert. Hierbei wurde für jedes Item ein weiterer Parameter bestimmt, der seine Schwierigkeit auf der globalen Skala, die durch alle Mathematik-Items definiert ist, beschreibt. Von einer gemeinsamen mehrdimensionalen Kalibrierung der Items wurde abgesehen, da dies für Items mit Mehrfachladungsstruktur zu einer problematischen Interpretation der Itemparameter führen kann (siehe dazu ausführlicher der Abschnitt „Berücksichtigung der Ladungs- und Korrelationsstruktur der Kompetenzbereiche“ auf S. 285).

12.2.2 Bestimmung von Personenparametern

Die Bestimmung der Personenparameter basiert auf einer zentralen Modellvorstellung der IRT: Die Kompetenz eines Kindes, das heißt seine latente Fähigkeit, kann nicht direkt beobachtet werden. Sie beeinflusst jedoch die Wahrscheinlichkeit, wie viele Items das Kind richtig löst. Umgekehrt kann auf Basis des Prozentsatzes gelöster Items auf den Kompetenzwert des Kindes, das diese Items bearbeitet hat, geschlossen werden. So wäre ein hoher Kompetenzwert bei einem Kind, das nur fünf Prozent aller Items beantwortet hat, theoretisch möglich, aber sehr unwahrscheinlich. Wahrscheinlicher wäre hier ein geringer Kompetenzwert. Ist die Schwierigkeit der Items bekannt, lässt sich bestimmen, wie wahrscheinlich es ist, dass ein Kind mit einem bestimmten, hypothetisch angenommenen Kompetenzwert den empirisch ermittelten Anteil an Items gelöst hat. Es ist naheliegend, für das Kind denjenigen Kompetenzwert anzunehmen, für den der beobachtete Anteil gelöster Items am wahrscheinlichsten ist.

Trägt man für verschiedene hypothetische Kompetenzwerte die entsprechenden bedingten Wahrscheinlichkeiten ab, erhält man eine sogenannte intraindividuelle Verteilung von Fähigkeitswerten. Um beruhend auf dieser Verteilung Personenparameter zu bestimmen, wurde im Ländervergleich 2011 der Plausible-Values-Ansatz verwendet (Mislevy, Beaton, Kaplan & Sheehan, 1992; von Davier, Gonzalez & Mislevy, 2009). Unter Fixierung der Itemparameter auf die Werte der Kalibrierung wird dabei für jede Person eine solche intraindividuelle Verteilung von Fähigkeitswerten bestimmt. Aus dieser Verteilung können dann beliebig viele „plausible“ Werte gezogen werden. Jeder Wert ist eine mögliche Repräsentation des Kompetenzwerts dieser Person. Das Ziehen mehrerer Werte

berücksichtigt die Tatsache, dass die individuelle Kompetenzausprägung unbeobachtbar und die Schätzung dieser Ausprägung mit einer Unsicherheit verbunden ist. Für eine bestimmte Person repräsentiert die Abweichung der Plausible Values untereinander die Unsicherheit der individuellen Fähigkeitsschätzung. Prinzipiell entspricht dieses Verfahren dem einer multiplen Imputation (vgl. Rubin, 1987), bei der ebenfalls für unbeobachtete Variablen eine Reihe plausibler Werte beruhend auf beobachteten Variablen bestimmt wird. Die beobachteten Variablen sind in diesem Fall die Items des Kompetenztests. Üblicherweise werden je Person fünf Plausible Values gezogen (von Davier et al., 2009). In aktuellen Arbeiten wird diskutiert, inwiefern eine größere Genauigkeit gewonnen werden kann, wenn die Anzahl der Imputationen erhöht wird (Graham, 2009). Im Ländervergleich 2011 wurden 15 Plausible Values je Person gezogen.

Bei Plausible Values handelt es sich nicht um individuelle Personenfähigkeitsschätzer, da sie dem Erwartungswert der Fähigkeit einer einzelnen Person nicht entsprechen. Ist man jedoch, wie im Ländervergleich 2011, vorwiegend an Statistiken für Gruppen interessiert, besitzen sie verschiedene günstige Eigenschaften. Anders als individuelle Personenparameter probabilistischer Testmodelle repräsentieren die über alle Personen aggregierten Plausible Values die Eigenschaften der Population in konsistenter Weise – sie konvergieren zu Populationsparametern (Mislevy et al., 1992). Die Unsicherheit der Fähigkeitsschätzung, die durch die Breite der intraindividuellen Verteilung gekennzeichnet ist, lässt sich reduzieren, wenn man zusätzlich Hintergrundmerkmale in die Schätzung einbezieht (Mislevy et al., 1992; von Davier et al., 2009). Die Idee dabei ist: Wenn Hintergrundmerkmale, wie etwa das Geschlecht oder der sozioökonomische Status der Eltern, mit Kompetenzwerten assoziiert sind, können diese Informationen für eine genauere Schätzung der Kompetenzwerte verwendet werden. Man nutzt also neben den Antworten auf die eigentlichen Testaufgaben weitere beobachtete Variablen für eine präzisere Schätzung der latenten Kompetenzausprägung. Die Einbeziehung von Hintergrundmerkmalen erfolgt in einem Populationsmodell, das einen linearen Zusammenhang von Hintergrundmerkmalen, wie dem Geschlecht der Person oder ihrem sozioökonomischen Status, mit ihren Kompetenzwerten annimmt. Die Berücksichtigung eines Merkmals im Populationsmodell ist erforderlich, um anschließend eine unverzerrte Schätzung des Zusammenhangs dieses Merkmals mit der gemessenen Kompetenz zu erhalten. Deshalb müssen alle Hintergrundmerkmale in die Personenparameterschätzung einbezogen werden, für deren Analyse die Plausible Values später genutzt werden sollen (Frey, Carstensen, Walter, Rönnebeck & Gomolka, 2008). Sollen beispielsweise Kompetenzunterschiede zwischen Jungen und Mädchen in Abhängigkeit vom Zuwanderungshintergrund differenziert werden, müssen sowohl das Geschlecht als auch der Zuwanderungshintergrund und deren Interaktion ins Populationsmodell einbezogen werden. Für die eigentliche Analyse kann man dann auf einfacher zu spezifizierende Regressions- oder varianzanalytische Modelle im Rahmen des Allgemeinen Linearen Modells zurückgreifen.

Um das oben beschriebene Verfahren technisch zu realisieren, muss zunächst das Hintergrundmodell spezifiziert werden. Hierfür ist es erforderlich, fehlende Werte auf den Hintergrundvariablen zu imputieren und anschließend die Hintergrundvariablen zu Hauptkomponenten zusammenzufassen. Weiterhin muss die Ladungs- und Korrelationsstruktur der Kompetenzbereiche bei der Ziehung der Plausible Values berücksichtigt werden. Im Folgenden werden die genannten Punkte näher beschrieben.

Imputation fehlender Werte auf den Hintergrundvariablen und Faktorisierung

Die in der Software ConQuest realisierte lineare Regressionsanalyse, die dem Hintergrundmodell zugrunde lag, stellt zwei Anforderungen an die Hintergrundvariablen: Sie sollen zum einen vollständig beobachtet sein, also keine fehlenden Werte aufweisen, und sie dürfen zum anderen nicht multikollinear abhängig sein. Unter Multikollinearität versteht man eine hohe Korrelation verschiedener unabhängiger Variablen oder deren Linearkombinationen (vgl. etwa Bortz & Weber, 2005).

Da in den vorliegenden Daten nicht alle Angaben für Personen vollständig waren, wurden im Ländervergleich 2011 fehlende Werte auf den Hintergrundvariablen durch ein Maximum-Likelihood-Verfahren unter Verwendung des EM-Algorithmus (z. B. Graham, 2009; Schafer & Graham, 2002) mithilfe des Pakets NORM in R imputiert (vgl. Schafer, 1997). Für jedes Land fand die Imputation separat in einem Modell mit je 56 Variablen statt. Folgende Merkmale wurden bei der Imputation berücksichtigt: Alter, Geschlecht, Schulart (Grundschule vs. Förderschule), Deutsch- und Mathematiknote, sonderpädagogischer Förderbedarf, zielgleicher beziehungsweise zieldifferenter Unterricht, Zuwanderungshintergrund (Geburtsland der Mutter, des Vaters und des Kindes), die in der Familie gesprochene Sprache, sozialer Hintergrund sowie die kognitive Grundfähigkeit der Schülerinnen und Schüler. Der Unterschiedlichkeit der Benotung in Regel- und Förderschulen wurde durch die Aufnahme eines Interaktionsterms zwischen Schulart und Note Rechnung getragen. Um die Beziehungen der Hintergrundvariablen zu den Kompetenzwerten abbilden zu können, wurden außerdem Kompetenzwerte als *Weighted Likelihood Estimates* für die Kompetenzbereiche Lesen, Zuhören und Mathematik in das Imputationsmodell aufgenommen. Die Berücksichtigung der Mehrebenenstruktur der Daten erfolgte durch die Aufnahme der Klassenmittelwerte dieser Kompetenzwerte und der kognitiven Grundfähigkeit. Für jede verwendete Variable wurde ein Missingindikator in das Modell aufgenommen, der angibt, ob es sich bei dem betreffenden Wert um einen tatsächlich beobachteten oder imputierten Wert handelt. Vor der Imputation wurden aus technischen Gründen in den einzelnen Ländern konstante Variablen entfernt. Korrelierten zwei Variablen extrem hoch miteinander ($|r| > .98$), ging nur eine Variable in die Imputation mit ein. Variablen mit geringer Varianz – es handelte sich dabei stets um Missingindikatoren für Variablen mit nur sehr geringen Missinganteilen – gingen zwar in die Imputation mit ein, nicht jedoch in die anschließende Faktorisierung. Um den Einfluss einzelner imputierter Werte auf die folgenden Schritte nicht durch Ausreißer zu verzerren, wurden die imputierten Werte auf den Wertebereich der entsprechenden Variablen beschränkt.

Nach der Imputation wurden die Hintergrundvariablen in einer Hauptkomponentenanalyse zu orthogonalen Faktoren zusammengefasst. Zuvor wurden die individuellen Kompetenzwerte sowie ihre Missingindikatoren aus den imputierten Hintergrunddaten ausgeschlossen, um die Unabhängigkeit von Hintergrundfaktoren und den Ergebnissen des Kompetenztests zu gewährleisten. Die Hauptkomponenten sollten mindestens 95 Prozent der Gesamtvarianz der Hintergrundvariablen erklären. Die Anzahl der extrahierten Hauptkomponenten variiert zwischen den Ländern von 24 bis 29. Diese Hauptkomponenten bildeten das Hintergrundmodell, das für alle fächerbezogenen und länderbezogenen Skalierungen mit ConQuest verwendet wurde.

Berücksichtigung der Ladungs- und Korrelationsstruktur der Kompetenzbereiche

Wie bereits erwähnt, soll die Korrelation der Plausible Values die „wahre“ Korrelationsstruktur der Kompetenzbereiche unverzerrt schätzen. Im Fach Deutsch wurden dazu für jedes Land Plausible Values in einem zweidimensionalen Modell für die beiden Kompetenzbereiche Lesen und Zuhören unter Einbeziehung der für dieses Land extrahierten Hauptkomponenten des Hintergrundmodells bestimmt. Dabei wurde jede Testaufgabe dem (und nur dem) Kompetenzbereich zugerechnet, für dessen Messung sie konstruiert worden ist. Die Modellvorstellung ist, dass die Kompetenzbereiche klar trennbare, wenn auch miteinander korrelierte Konstrukte darstellen.

Dieses Vorgehen lässt sich für das Fach Mathematik nicht in gleicher Weise umsetzen. Wie in Kapitel 2.2 erläutert, sind dort die Kompetenzbereiche nicht immer klar voneinander abzugrenzen und es ist üblich – und so intendiert –, dass Items Aspekte mehrerer Leitideen ansprechen und deshalb auf mehr als einer Dimension „laden“. Man spricht in einem solchen Fall von *Within-Item-Dimensionality* (Wu & Adams, 2006; Zhang, 2004, 2005) oder einer Mehrfachladungsstruktur (*mixed structure*, vgl. Zhang, 2012). Da die Lösungswahrscheinlichkeit eines Items nun nicht mehr nur von einer, sondern von mindestens zwei latenten Teilkompetenzen abhängt, muss die Art dieser Abhängigkeit spezifiziert werden. Hierfür ist eine Vielzahl verschiedener Möglichkeiten denkbar, von denen im Folgenden zwei Ansätze vorgestellt werden. Ein Überblick möglicher Messmodelle für Mehrfachladungsstrukturen findet sich zum Beispiel in DiBello, Roussos und Stout (2007), eine Diskussion ihrer Einsatzmöglichkeiten in Kunina-Habenicht, Wilhelm, Matthes und Rupp (2010).

Die Lösungswahrscheinlichkeiten der Items, die mehreren Kompetenzbereichen zugeordnet sind, können entweder kompensatorisch oder nicht kompensatorisch (konjunktiv) modelliert werden. In einem kompensatorischen Modell kann eine geringe Fähigkeitsausprägung in einem Kompetenzbereich durch eine hohe Fähigkeitsausprägung in einem anderen Kompetenzbereich ausgeglichen (kompensiert) werden. In nicht kompensatorischen Modellen müssen Kompetenzen aller Bereiche beherrscht werden, um das Item zu lösen. Ein additives Modell ist das einfachste Beispiel dafür, wie zwei Personenfähigkeiten θ_1 und θ_2 kompensatorisch zusammenwirken können. Ein solches Modell wird beispielsweise in der Software ConQuest im Regelfall verwendet. Die Wahrscheinlichkeit, ein Item mit Schwierigkeit σ zu lösen, wird dabei durch den Term $\theta_1 + \theta_2 - \sigma$ beschrieben. Gehen die beiden Fähigkeiten als allgemeine Linearkombination mit unterschiedlichen Gewichten a_1 und a_2 ein, entspricht dies dem in linearen Strukturgleichungsmodellen üblichen Vorgehen beziehungsweise dem 2-PL-Modell (vgl. Embretson & Reise, 2000). Die Lösungswahrscheinlichkeit ist dann eine Funktion des Terms $a_1\theta_1 + a_2\theta_2 - \sigma$. Lässt man die Gewichtung der Fähigkeiten nicht als Parameter schätzen, sondern bestimmt sie auf Basis theoretischer Überlegungen und gibt sie fixiert in das Modell hinein, lässt sich auch dieses Modell im Rahmen des allgemeinen Messmodells von ConQuest spezifizieren.

Die Messmodelle für Mehrfachladungsstrukturen sind mit einer Reihe von theoretischen wie mathematisch-statistischen Problemen behaftet, von denen im Folgenden einige exemplarisch dargestellt werden. So können beispielsweise kompensatorische Modelle zu dem paradoxen Befund führen, dass für Personen, die eine größere Anzahl Items gelöst haben, ein geringerer Kompetenzwert ge-

schätzt wird als für Personen, die eine geringere Anzahl Items gelöst haben (Hooker, Finkelman & Schwartzman, 2009).

Ein anderes Beispiel betrifft die Interpretation der Fähigkeitsskala. Es ist einer der größten Vorteile eindimensionaler IRT-Modelle, dass Personenfähigkeiten und Itemschwierigkeiten auf derselben Skala gemessen werden und deshalb direkt aufeinander bezogen werden können. Dies ermöglicht eine inhaltliche Beschreibung verschiedener Abschnitte der Fähigkeitsskala, wie es etwa ein Kompetenzstufenmodell leistet (vgl. Kapitel 3.1). In einer gemischten Struktur, in der manche Items auf einem und andere auf mehreren Kompetenzbereichen laden, ist dies nicht ohne weiteres möglich. Dies soll anhand des additiven Messmodells beispielhaft verdeutlicht werden. Betrachtet werden ein Item, von dem man annimmt, dass es ausschließlich Kompetenzen aus dem Bereich Zahlen und Operationen (ZO) erfordert, sowie ein zweites Item, zu dessen Lösung Fähigkeiten sowohl aus dem Bereich ZO als auch aus dem Bereich Muster und Strukturen (MS) benötigt werden. Für das erste Item gilt das eindimensionale Rasch-Modell, und seine Lösungswahrscheinlichkeit ergibt sich bei einer Fähigkeitsausprägung θ_1 im Bereich ZO aus der Differenz $\theta_1 - \sigma_1$, wobei σ_1 die Schwierigkeit des Items ist. Für das zweite Item gilt entsprechend $\theta_1 + \theta_2 - \sigma_{12}$. Hierbei beschreibt σ_{12} die Schwierigkeit des zweiten Items, das auf beiden Kompetenzbereichen lädt. θ_1 beziehungsweise θ_2 bezeichnen die Fähigkeitsausprägungen in diesen beiden Bereichen (ZO beziehungsweise MS). Für das erste Item liegen Fähigkeit und Schwierigkeit auf derselben Skala, und die Differenz $\theta_1 - \sigma_1$ lässt sich als Abstand auf dieser Skala interpretieren. Die Schwierigkeit des zweiten Items lässt sich hingegen auf keiner der beiden Fähigkeitsskalen verorten. Es repräsentiert vielmehr eine eigene Dimension, die sich aus einer „Mischung“ der ZO- und der MS-Dimension ergibt. Dies wird auch deutlich, wenn man konkrete Zahlen für alle Parameter einsetzt. Nehmen diese beispielsweise alle den Wert 1 an (also $\theta_1 = \theta_2 = \sigma_1 = \sigma_{12} = 1$), so ist für das erste Item $\theta_1 - \sigma_1 = 0$ und für das zweite $\theta_1 + \theta_2 - \sigma_{12} = 1$. Das bedeutet, dass eine Person mit einer ZO-Fähigkeit von 1 für die zwei Items mit der numerisch gleichen Schwierigkeit unterschiedliche Lösungswahrscheinlichkeiten hat, nämlich 50 Prozent für das erste und 73 Prozent für das zweite Item.

Eine große Herausforderung bei der Modellierung der Mehrfachladungsstruktur im Bereich Mathematik ist die theoretische Begründung der Wahl eines Messmodells, das für alle Items gültig ist. Denn es scheint plausibel, dass mit der Mannigfaltigkeit von Mathematikaufgaben auch eine große Zahl von Möglichkeiten einhergeht, wie die jeweils beteiligten Kompetenzen zusammenspielen. Ein Modell, das dieser Vielfalt gerecht wird und gleichzeitig technisch umsetzbar ist, liegt bislang nicht vor. Das einzige Modell, das derzeit als etabliert betrachtet werden kann, ist das oben besprochene lineare Messmodell und dessen Spezialfall, das additive Messmodell (vgl. Kunina-Habenicht et al., 2010). Es kann jedoch nur für manche Mathematikaufgaben als adäquat betrachtet werden. In zahlreichen anderen ist es nicht möglich, eine geforderte Kompetenz vollständig durch eine andere zu kompensieren. Vielmehr zeichnen sich eine Vielzahl von Mathematikaufgaben gerade dadurch aus, dass es erforderlich ist, mehrere Kompetenzen zugleich einzusetzen. Folglich wäre für jede einzelne Aufgabe theoretisch zu bestimmen, in welcher Weise kompensatorische oder konjunktive Wechselwirkungen zwischen den beteiligten Fähigkeiten anzunehmen sind. In Anbetracht der Herausforderung, die bereits das Identifizieren der beteiligten Fähigkeiten in vergleichsweise einfachen Aufgaben darstellt (Kunina-Habenicht et al., 2010), stellt dies ein Desiderat für zukünftige Forschungsarbeiten dar.

Aufgrund dieser durch die Mehrfachladungen bedingten methodischen Herausforderungen wurde zur Skalierung der mathematischen Kompetenzen ein Vorgehen gewählt, das einerseits die Mehrfachladungen der Items berücksichtigt, zugleich aber auch kohärent zum Vorgehen in anderen Schulleistungsstudien ist. Die Plausible Values im Fach Mathematik wurden daher in mehreren eindimensionalen Modellen bestimmt, ohne die Mehrfachladungsstruktur explizit zu modellieren. Dabei wurden, wie bereits im Schritt der Kalibrierung, jeweils alle Items einbezogen, die gemäß einem von drei Fachdidaktikerinnen und Fachdidaktikern durchgeführten Expertenrating Kompetenzen des jeweiligen Inhaltsbereichs erfordern. Insgesamt wurden im Fach Mathematik pro Land sechs eindimensionale Modelle zur Personenparameterschätzung berechnet: ein Globalmodell mit allen Items und fünf kompetenzbereichsspezifische Modelle. In jedem Modell wurden für sämtliche Personen Plausible Values geschätzt. Die mit diesem Vorgehen verbundene geringfügige Unterschätzung der latenten Korrelationen zwischen den einzelnen Kompetenzbereichen wurde in Kauf genommen, um die andernfalls zu erwartenden, potenziell erheblichen Verschätzungen der Personenfähigkeiten auszuschließen (vgl. Hooker et al., 2009).

12.2.3 Bestimmung einer Transformationsvorschrift für die Berichtsmetrik

Die Kompetenzwerte der untersuchten Schülerinnen und Schüler befinden sich nach der Skalierung zunächst auf einer Skala, die als Logitmetrik bezeichnet wird. Diese weist sowohl positive als auch negative Werte auf und ist daher für die Ergebnisdarstellung wenig anschaulich. Um besser interpretierbare Werte zu erhalten, werden sie in eine Metrik transformiert, die auch in anderen Schulleistungsstudien (z. B. PISA) verwendet wird. Diese ist so definiert, dass die für einen Kompetenzbereich maßgebliche Referenzpopulation zum Zeitpunkt der Messung einen Mittelwert von 500 und eine Standardabweichung von 100 hat. Bei der Referenzpopulation handelt es sich hierbei um Primarschülerinnen und -schüler der vierten Jahrgangsstufe in Deutschland.

Für die Transformation wird je Kompetenzbereich der gewichtete Mittelwert M_g und die gewichtete Standardabweichung SD_g über die in den jeweiligen Ländern gezogenen Plausible Values bestimmt. Das Fallgewicht einer Person beschreibt, wie viele Personen in der Population diese Person der Stichprobe repräsentiert. Der gewichtete Mittelwert berechnet sich für jeden Kompetenzbereich folgendermaßen:

$$M_g = \frac{1}{I} \sum_{i=1}^I \sum_{n=1}^N \frac{w_n}{W} \cdot p_{v_{n,i}}$$

Dabei ist I die Anzahl der gezogenen Plausible Values je Person, N ist die Gesamtzahl der Personen in der Stichprobe, w_n ist das Fallgewicht der n -ten Person in der Stichprobe, W ist die Summe aller Gewichte, und $pv_{n,i}$ ist der i -te Plausible Value der n -ten Person. Analog dazu berechnet sich die gewichtete Standardabweichung für jeden Kompetenzbereich wie folgt:

$$SD_g = \frac{1}{I} \sum_{i=1}^I \sqrt{\sum_{n=1}^N \frac{w_n}{W-1} \cdot (pv_{n,i} - M_{g,i})^2}$$

Hier ist $M_{g,i}$ der gewichtete Mittelwert über das i -te Set von Plausible Values. Die so gewonnenen Werte werden nun benutzt, um jeden einzelnen Personenparameter von der Logitmetrik auf die Berichtsmetrik zu übertragen. Die inhaltliche Bedeutung der Kompetenzskalen wird durch die Transformation nicht verändert. Die Transformationsvorschrift ist für jeden Kompetenzbereich folgendermaßen definiert:

$$pv_{LV} = (pv_{Logit} - M_g) \cdot \frac{100}{SD_g} + 500$$

12.2.4 Bestimmung der Standardfehler

Ziel des Ländervergleichs ist es, gemittelte Fähigkeitsschätzungen aus einer repräsentativen Stichprobe auf die Grundgesamtheit, also die Population, zu übertragen. Da auch jede repräsentative Stichprobe in ihren Kennwerten jedoch aufgrund von Zufallsfehlern von den Parametern in der Population abweicht, ist es wichtig, den zu erwartenden Grad der Abweichung zu bestimmen. Im Falle des Ländervergleichs bedeutet dies beispielsweise, dass der aus der Stichprobe geschätzte durchschnittliche Fähigkeitswert der Schülerinnen und Schüler eines Landes nicht exakt dem tatsächlichen Mittelwert aller Schülerinnen und Schüler des Landes entsprechen muss, sondern eine Unsicherheit aufweist. Diese Unsicherheit wird durch den Standardfehler oder Standardschätzfehler (*standard error*, *SE*) quantifiziert. Der Fehler lässt sich in einer vollständig zufällig gezogenen Stichprobe bestimmen, indem man die Varianz des Merkmals in der Stichprobe ins Verhältnis zur Größe der Stichprobe setzt. Voraussetzung hierfür ist, dass die Beobachtungen – im vorliegenden Fall die Antworten der Schülerinnen und Schüler – voneinander unabhängig sind. Im Ländervergleich werden aus erhebungsökonomischen Gründen jedoch keine einzelnen Schülerinnen und Schüler, sondern ganze Klassen pro Land gezogen und alle Kinder einer Klasse getestet. Man spricht in einem solchen Fall von einer stratifizierten Clusterstichprobe, da die Grundgesamtheit in mehrere Gruppierungen oder Schichten, sogenannte Strata (hier: Land) unterteilt wird, aus denen separat Cluster (hier: Schulen und Klassen) gezogen werden. Da sich Kinder einer Klasse jedoch hinsichtlich ihrer Messwerte ähnlicher sind als Kinder verschiedener Klassen, ist eine solche Stichprobe in sich homogener als eine Stichprobe gleicher Größe, die nicht durch die Ziehung von Clustern, sondern durch Ziehung individueller Schülerinnen und Schüler zustande kommt. Ihre geringere Streuung muss berücksichtigt werden, um die Unsicherheit der

Mittelwertsschätzungen nicht systematisch zu unterschätzen. Aus diesem Grund werden die Standardfehler bei großen Schulleistungsstudien mit komplexen statistischen Verfahren bestimmt. Beim Ländervergleich 2011 kam dabei das *Jackknife-Verfahren* (vgl. Foy, Galia & Li, 2008; Wolter, 1985) zum Einsatz.

Beim Jackknife-Verfahren werden aus der bestehenden Stichprobe künstlich mehrere neue Stichproben (Replikationen) gebildet. In stratifizierten Stichproben werden dabei in jedem Stratum (Land) Paare der Ziehungseinheiten (Klassen) gebildet. Befinden sich in einem Stratum 100 Klassen, so werden daraus 50 Paare gebildet. Für eine Stichprobenreplikation wird nun eine Klasse aus einem Paar durch die andere Klasse des entsprechenden Paares ersetzt. Jede Stichprobenreplikation ist daher bezüglich der Klassen geringfügig homogener als die Gesamtstichprobe. Der interessierende Kennwert (z. B. Mittelwert der Schülerinnen und Schüler eines Landes im Lesen) wird nun mit dieser Replikation bestimmt. Das Verfahren wird in analoger Weise mit einer weiteren Replikation wiederholt. Je stärker die berechneten Kennwerte zwischen den Stichprobenreplikationen voneinander abweichen, umso größer ist die Variabilität von Kindern verschiedener Klassen verglichen mit der Variabilität von Kindern gleicher Klassen. Kennt man diese Differenz, lässt sich abschätzen, wie viel Variabilität man in einer ungeclusterten Stichprobe gleicher Größe erwarten würde. Diese erwartete Variabilität wird benutzt, um in jedem Stratum unverfälschte Stichprobenfehler zu bestimmen.

Literatur

- Allen, N. L., Donoghue, J. R. & Schoeps, T. L. (2001). *The NAEP 1998 technical report*. Washington: National Center for Educational Statistics.
- Böhme, K., Leucht, M., Schipolowski, S., Porsch, R., Knigge, M. & Köller, O. (2010). Anlage und Durchführung des Ländervergleichs. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 65–85). Münster: Waxmann.
- Bonsen, M., Lintorf, K., Bos, W. & Frey, K. A. (2008). TIMSS 2007 Grundschule – Eine Einführung in die Studie. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 19–48). Münster: Waxmann.
- Bortz, J. & Weber, R. (2005). *Statistik für Human- und Sozialwissenschaftler: mit 242 Tabellen*. Heidelberg: Springer Medizin.
- DiBello, L. V., Roussos, L. A. & Stout, W. (2007). Review of cognitively diagnostic assessment and a summary of psychometric models. In C. R. Rao & S. Sinharay (Hrsg.), *Handbook of Statistics* (S. 979–1030). New York: Elsevier.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Foy, P., Galia, J. & Li, I. (2008). Scaling the data from the TIMSS 2007 mathematics and science assessment. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Hrsg.), *TIMSS 2007 Technical Report* (S. 225–280). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2006 in Deutschland: Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 375–397). Münster: Waxmann.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576.
- Hooker, G., Finkelman, M. & Schwartzman, A. (2009). Paradoxical results in multidimensional Item Response Theory. *Psychometrika*, 74, 419–442.
- Hornberg, S., Bos, W., Buddeberg, I., Potthoff, B. & Stubbe, T. C. (2007). Anlage und Durchführung von IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust,

- L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 21–79). Münster: Waxmann.
- Kolen, M. J. (2006). Scaling and norming. In R. L. Brennan (Hrsg.), *Educational Measurement* (4th ed., S. 155–186). Westport, CT: American Council on Education.
- Kunina-Habenicht, O., Wilhelm, O., Matthes, F. & Rupp, A. A. (2010). Kognitive Diagnosemodelle: Theoretisches Potential und methodische Probleme. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes* (S. 75–85). Weinheim: Beltz.
- Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage.
- Mislevy, R. J., Beaton, A. E., Kaplan, B. & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29, 133–161.
- OECD. (2003). *The PISA 2003 assessment framework mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD.
- R Development Core Team. (2011). R: A language and environment for statistical computing (Version 2.13.2). Vienna: R Foundation for Statistical Computing. Zugriff am 12.08.2012 unter <http://www.R-project.org/>
- Robitzsch, A. (2009). Methodische Herausforderungen bei der Kalibrierung von Leistungstests. In D. Granzer, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik* (S. 42–106). Weinheim: Beltz.
- Rost, J. (2004). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Huber.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- von Davier, M., Gonzalez, E. & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series*, 4, 9–36.
- Westat. (2000). *WesVar*. Rockville, MD: Westat.
- Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer.
- Wu, M. & Adams, R. J. (2006). Modelling mathematics problem solving item responses using a multidimensional IRT model. *Mathematics Education Research Journal*, 18, 93–113.
- Wu, M., Adams, R. J., Wilson, M. R. & Haldane, S. (2007). *ACER ConQuest 2.0 – Generalised item response modelling software*. Camberwell: ACER.
- Zhang, J. (2004). Comparison of unidimensional and multidimensional approaches to IRT parameter estimation. *ETS Research Report, RR-04-44*. Princeton, NJ: Educational Testing Service.
- Zhang, J. (2005). Estimating multidimensional item response models with mixed structure. *ETS Research Report, RR-05-44*. Princeton, NJ: Educational Testing Service.
- Zhang, J. (2012). Calibration of response data using MIRT models with simple and mixed structures. *Applied Psychological Measurement*, 36, 375–398.

Kapitel 13

IQB-Ländervergleich 2011: Zusammenfassung und Einordnung der Befunde

Petra Stanat, Hans Anand Pant, Dirk Richter und Katrin Böhme

In diesem Bericht wurden die Ergebnisse der ersten Studie zur Überprüfung von Erträgen schulischer Lehr-Lern-Prozesse im Primarbereich dargestellt, die auf Grundlage der von den Kultusministern der Länder in der Bundesrepublik Deutschland beschlossenen Bildungsstandards durchgeführt worden ist (vgl. Kapitel 1). Die Analysen bezogen sich auf Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Im Fach Deutsch wurden die Kompetenzbereiche Lesen und Zuhören, im Fach Mathematik sowohl die globale mathematische Kompetenz als auch die fünf inhaltsbezogenen mathematischen Kompetenzen (Zahlen und Operationen; Raum und Form; Muster und Strukturen; Größen und Messen; Daten, Häufigkeit und Wahrscheinlichkeit) untersucht. Die Auswertungen basieren auf repräsentativen Stichproben von Schülerinnen und Schülern, die in allen 16 Ländern nach einem Zufallsverfahren ausgewählt worden sind. Für diese Kompetenzbereiche konnten daher Aussagen darüber getroffen werden, inwieweit sich die Ergebnisse zwischen den Ländern unterscheiden. Der Kompetenzbereich Orthografie wurde ebenfalls erfasst, allerdings nur in einer für Deutschland insgesamt repräsentativen Stichprobe, so dass sich für diesen Bereich kein Ländervergleich durchführen ließ.

Wie in Schulleistungsstudien allgemein üblich, wurden verschiedene Indikatoren für die Erträge schulischer Lehr-Lern-Prozesse ermittelt und im Bericht dargestellt. Diese beziehen sich zunächst auf das Niveau der von den Schülerinnen und Schülern in den Ländern insgesamt erreichten Kompetenzen, das anhand von Ländermittelwerten dargestellt wurde. Gleichzeitig wurde in den Analysen der Frage nachgegangen, wie stark die Leistungen der Schülerinnen und Schüler innerhalb der Länder streuen (Kapitel 5). Dabei gelten Muster, die auf ein hohes mittleres Kompetenzniveau bei gleichzeitig geringer Streuung hinweisen, als erstrebenswert.

Die in den Ländern jeweils erreichten Kompetenzniveaus wurden jedoch nicht nur in Form von abstrakten Punktwerten, sondern auch anhand von Verteilungen auf inhaltlich interpretierbare Kompetenzstufen dargestellt (Kapitel 6). Eine wichtige Neuerung besteht dabei darin, dass die Kompetenzstufen auf den Bildungsstandards basieren mit denen vorab kriterial festgelegt wurde, welche Kompetenzniveaus als mehr oder weniger zufriedenstellend zu bewerten sind. Anhand der Analysen wird also sichtbar, in welchem Ausmaß es den Ländern gelingt, Mindeststandards zu sichern und zu gewährleisten, dass ein möglichst hoher Anteil von Schülerinnen und Schülern die Regelstandards der Kultusministerkonferenz erreicht.

Eine zweite Gruppe von Indikatoren bezieht sich auf die Frage, inwieweit Schülerinnen und Schüler mit bestimmten Hintergrundmerkmalen ein hohes

Kompetenzniveau erreichen. Untersucht wurden Kompetenzunterschiede zwischen Mädchen und Jungen (Kapitel 7), Zusammenhänge zwischen Merkmalen der sozialen Herkunft und erreichten Kompetenzen (Kapitel 8) sowie Leistungsdifferenzen zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund (Kapitel 9). Auch wenn es in Bezug auf den familiären Hintergrund kaum möglich sein wird, die damit verbundenen Unterschiede in den Eingangsvoraussetzungen vollständig auszugleichen, so gilt es doch als erstrebenswert, Disparitäten so weit wie möglich zu reduzieren. Anhand der Daten des Ländervergleichs 2011 wird erkennbar, inwieweit dies, bezogen auf die drei genannten Hintergrundmerkmale, gelingt.

Mit den Ergebnissen dieser Analysen erhalten die Länder eine auf den *Output* bezogene Rückmeldung über Stärken und Schwächen ihres Bildungssystems im Primarbereich. Ein Bildungsmonitoring, das für Steuerung relevante Evidenz generieren will, sollte jedoch nicht nur die Ergebnisse von Lehr-Lern-Prozessen in den Blick nehmen, sondern auch zentrale Bedingungen dieser Prozesse einbeziehen. Im vorliegenden Band wurden zwei dieser Aspekte genauer analysiert: Merkmale der Lehrerschaft und Umsetzung von Maßnahmen der Sprach- und Leseförderung im Primarbereich. Zum einen wurde untersucht, wie sich die Lehrerschaft in Schulen der Primarstufe in den Ländern jeweils zusammensetzt und inwieweit Lehrkräfte ihrer Qualifikation entsprechend in den Fächern Deutsch und Mathematik eingesetzt werden. Weiterhin wurde berichtet, welche Angebote der Fortbildung Lehrerinnen und Lehrer nutzen, in welchen Bereichen sie nach eigenen Angaben einen Fortbildungsbedarf sehen und welche Faktoren sie an der Nutzung von Fortbildungsangeboten hindern (Kapitel 10). Zum anderen wurde der Frage nachgegangen, welche Maßnahmen der Sprach- und Leseförderung in Primarschulen umgesetzt werden und inwieweit diese Maßnahmen diejenigen Schülerinnen und Schüler erreichen, die gemessen an ihren Testergebnissen im Fach Deutsch einen entsprechenden Bedarf aufweisen.

Im Folgenden sollen die Ergebnisse der beschriebenen Analysen knapp zusammengefasst werden. Dabei wird nicht auf die genauen Befundmuster für die einzelnen Länder eingegangen, sondern es werden zentrale Tendenzen skizziert und besondere Herausforderungen hervorgehoben. In Bezug auf das Fach Mathematik beschränken sich die Ausführungen auf die Ergebnisse für die globale mathematische Kompetenz. Abschließend folgt eine Einordnung der Befunde in Bezug auf Ergebnisse früherer Ländervergleichsstudien, die im Primarbereich (IGLU-E 2006) und in der Sekundarstufe I (Ländervergleich 2009) durchgeführt worden sind.

13.1 Mittleres Niveau und Streuung der erreichten Kompetenzen

Ein vergleichsweise hohes *Niveau* erreichen in allen erfassten Kompetenzbereichen die Schülerinnen und Schüler in Bayern. Sowohl im Lesen und im Zuhören als auch in Mathematik liegen die Kompetenzmittelwerte in Bayern signifikant über dem deutschen Gesamtmittelwert. Das Gegenteil ist der Fall in den Stadtstaaten Berlin, Bremen und Hamburg. Hier erreichen die Schülerinnen und Schülern in allen drei Bereichen im Durchschnitt signifikant geringere Kompetenzwerte als Schülerinnen und Schüler in Deutschland insgesamt. Allerdings ist die Situation über die drei Stadtstaaten hinweg keineswegs vollständig homogen, da sich Hamburg innerhalb dieser Gruppe durch insgesamt etwas günstigere Ergebnisse auszeichnet. So weist Hamburg im Vergleich zum

Land mit den jeweils schwächsten Ergebnissen einen Vorsprung von 15 Punkten im Lesen und von 19 Punkten sowohl im Zuhören als auch in Mathematik auf.

Wie in Bayern einerseits und in den Stadtstaaten andererseits weisen auch in den meisten anderen Ländern die Ergebnisse für die verschiedenen Kompetenzbereiche jeweils in eine einheitliche Richtung, wobei sie sich jedoch nur zum Teil signifikant vom deutschen Mittelwert unterscheiden. Signifikant höhere Kompetenzwerte erzielen die Schülerinnen und Schüler in einzelnen Kompetenzbereichen in Baden-Württemberg (Mathematik), in Sachsen und Sachsen-Anhalt (Lesen und Mathematik) sowie in Thüringen (Lesen). Signifikant unterhalb des deutschen Mittelwerts liegen die erreichten Kompetenzen der Schülerinnen und Schüler zusätzlich zu den Stadtstaaten in Brandenburg, Hessen und Schleswig-Holstein (Mathematik) sowie in Rheinland-Pfalz, im Saarland und in Sachsen-Anhalt (Zuhören). Solche Ergebnismuster können als Hinweis auf differenzielle Herausforderungen für die schulische Förderung in den jeweiligen Ländern gewertet werden.

Die *Streuung* der Kompetenzwerte variiert ebenfalls zwischen den Ländern, wobei sich für die drei Kompetenzbereiche allerdings nur bedingt einheitliche Muster ergeben. In allen drei Bereichen ist die Streuung in Sachsen-Anhalt vergleichsweise groß. Besonders stark gehen die Leistungen der Schülerinnen und Schüler zudem in Berlin (Lesen und Zuhören), Bremen (Zuhören und Mathematik), Hamburg (Zuhören) und Mecklenburg-Vorpommern (Mathematik) auseinander. Vergleichsweise geringe Streuungen sind dagegen in Thüringen und Baden-Württemberg (Lesen und Zuhören) sowie in Nordrhein-Westfalen und Schleswig-Holstein (Mathematik) zu verzeichnen. Dabei hebt sich jedoch kaum ein Land dadurch von den anderen ab, dass es in allen überprüften Bereichen das erstrebenswerte Muster eines hohen mittleren Kompetenzniveaus bei gleichzeitig geringer Streuung aufweist. Dieses Muster ist relativ klar nur im Bereich Lesen für Thüringen sowie tendenziell auch für Baden-Württemberg zu erkennen. Die problematische Kopplung eines geringen Kompetenzniveaus mit einer relativ großen Streuung besteht vor allem in Berlin und Bremen.

Länderspezifische Herausforderungen für Bildungspolitik und Schulpraxis werden besonders deutlich, wenn man die Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen betrachtet, die anhand inhaltlicher Kriterien definiert wurden (vgl. Kapitel 3 und 6). Dabei ist davon auszugehen, dass es Schülerinnen und Schüler, die in der vierten Jahrgangsstufe die Kompetenzstufe II und damit den Mindeststandard verfehlen, sehr schwer haben werden, in der Sekundarstufe I den Anschluss zu finden.

Im Kompetenzbereich Zuhören verfehlen in keinem Land mehr als 16 Prozent der Schülerinnen und Schüler die Mindeststandards, wobei der Anteil in den Stadtstaaten und im Saarland mehr als 10 Prozent beträgt. In den Kompetenzbereichen Lesen und Mathematik hingegen gibt es Länder, in denen mehr als ein Fünftel der Schülerinnen und Schüler die Mindeststandards nicht erreicht. Dies ist im Lesen in den Ländern Berlin (22 %) und Bremen (21 %) der Fall sowie in Mathematik in Berlin (27 %), Bremen (25 %) und Hamburg (21 %). Besonders geringe Anteile von Schülerinnen und Schülern, die die Mindeststandards verfehlen, sind dagegen in Bayern, Sachsen und Sachsen-Anhalt (Lesen und Mathematik), in Baden-Württemberg (Mathematik) sowie in Mecklenburg-Vorpommern und Thüringen (Lesen) mit jeweils unter 10 Prozent zu finden.

Auf der anderen Seite des Leistungsspektrums, im Bereich der Optimalstandards (Kompetenzstufe V), sind die Unterschiede zwischen den Ländern

nicht nur im Zuhören, sondern auch im Lesen geringer ausgeprägt. Weniger als 10 Prozent sehr gute Leserinnen und Leser sind nur in Berlin (7 %), Bremen (7 %) und Hamburg (8 %) zu finden. Im Zuhören gehören neben den Stadtstaaten Berlin und Bremen auch Brandenburg, Rheinland-Pfalz, Saarland, Sachsen-Anhalt und Thüringen zu der Ländergruppe, in denen weniger als 10 Prozent der Schülerinnen und Schüler diese Kompetenzstufe erreichen. Deutlich größere Länderunterschiede sind im oberen Leistungsbereich dagegen im Fach Mathematik zu verzeichnen. Hier erreichen in Bayern, Sachsen und Sachsen-Anhalt mehr als 20 Prozent der Schülerinnen und Schüler die Optimalstandards, während in Berlin und Bremen nicht einmal 10 Prozent der Schülerinnen und Schüler auf diesem sehr hohen Kompetenzniveau zu finden sind (7 % bzw. 9 %). Ganz ähnlich stellt sich die Situation dar, wenn man die Anteile der Schülerinnen und Schüler betrachtet, die im Fach Mathematik die Regelstandards erreichen. Diese liegen in Bayern, Sachsen und Sachsen-Anhalt bei mehr als 75 Prozent, in Berlin und Bremen jedoch nur bei 47 beziehungsweise bei 50 Prozent. Damit geht die Schere zwischen den Ländern im Fach Mathematik am weitesten auseinander.

13.2 Geschlechtsbezogene, soziale und zugwanderungsbezogene Disparitäten

In den Analysen der Kompetenzunterschiede zwischen Jungen und Mädchen findet sich das bekannte Muster eines deutlichen Vorsprungs der Mädchen im Lesen (bundesweit: 24 Punkte) und eines etwas weniger stark ausgeprägten Vorsprungs der Jungen in Mathematik (bundesweit: 16 Punkte). Im Kompetenzbereich Zuhören hingegen ist in Deutschland insgesamt kein signifikanter Geschlechterunterschied zu verzeichnen. Dieses Muster wiederholt sich in den meisten Ländern in etwas unterschiedlich ausgeprägtem Maße. Besonders markante Geschlechterunterschiede, die signifikant größer sind als die Geschlechterdifferenzen in Deutschland insgesamt, sind nur in Niedersachsen für die Bereiche Lesen und Zuhören (38 Punkte beziehungsweise 16 Punkte zugunsten der Mädchen) sowie in Hamburg für die Globalskala der mathematischen Kompetenz (27 Punkte zugunsten der Jungen) zu verzeichnen. Aber auch in den anderen Ländern sind die Geschlechterunterschiede im Lesen und in Mathematik überwiegend statistisch signifikant. Lediglich in Baden-Württemberg und in Mecklenburg-Vorpommern sind die Leistungsdifferenzen zwischen Jungen und Mädchen in Mathematik so gering, dass sie sich nicht zufallskritisch absichern lassen und somit vernachlässigt werden können.

In allen Ländern Deutschlands besteht ein substanzieller und gleichgerichteter Zusammenhang zwischen der sozialen Herkunft und den erreichten Kompetenzen von Schülerinnen und Schülern; je niedriger der soziale Status ist, desto schwächer ausgeprägt sind die erreichten Kompetenzen. Allerdings unterscheiden sich die Länder in der Stärke dieses Zusammenhangs teilweise deutlich. Eine im Vergleich zum deutschen Gesamtwert signifikant engere Kopplung zwischen der schulischen Leistung und dem sozioökonomischen Status der Eltern besteht für den Bereich Zuhören in allen Stadtstaaten und für den Bereich Mathematik ausschließlich in Bremen. Zusammenhänge, die signifikant kleiner sind als der deutsche Gesamtwert, finden sich nur für den Bereich Lesen in Sachsen. Kompetenzunterschiede zwischen hoch und niedrig qualifizierten Berufsgruppen zeigen sich ebenfalls in allen Ländern, jedoch unterscheiden sich diese nur in

wenigen Ländern in den Bereichen Lesen und Zuhören statistisch signifikant vom deutschen Mittelwert. Eine Leistungsdifferenz von über 100 Punkten, die signifikant größer ist als die Differenz für Deutschland insgesamt, zeigt sich im Zuhören für Berlin und Hamburg. Sehr niedrige Differenzwerte sind im Zuhören in Brandenburg (62 Punkte), Sachsen (60 Punkte) und Thüringen (67 Punkte) zu finden. Im Bereich Lesen weist nur Sachsen mit 52 Punkten eine vom deutschen Wert signifikant abweichende Leistungsdifferenz auf, die etwa halb so groß ist wie die größten Differenzwerte in Bayern (98 Punkte) und in den Großstädten (106 Punkte).

Zuwanderungsbezogene Disparitäten wurden differenziert nur für diejenigen Länder analysiert, in denen der Anteil von Kindern aus zugewanderten Familien ausreichend groß ist. In diesen insgesamt 10 Ländern ergaben sich für alle untersuchten Kompetenzbereiche deutliche Nachteile für Schülerinnen und Schüler mit Zuwanderungshintergrund. Diese waren jeweils bei Kindern mit zwei im Ausland geborenen Elternteilen stärker ausgeprägt als bei Kindern mit einem im Ausland und einem in Deutschland geborenen Elternteil. Die Disparitäten im Zuhören fielen in allen Ländern größer aus als in den anderen untersuchten Bereichen.

Besonders große Leistungsnachteile von Kindern aus zugewanderten Familien im Vergleich zu Kindern ohne Zuwanderungshintergrund wurden für die Stadtstaaten Berlin, Bremen und Hamburg identifiziert (z. B. Kompetenzbereich Lesen, beide Elternteile im Ausland geboren: $Diff=70$ Punkte, 65 Punkte und 64 Punkte). In den Großstädten der Flächenländer sind die Unterschiede dagegen geringer ausgeprägt. Vergleichsweise große Gruppenunterschiede sind ferner im Bereich Lesen in Hessen, im Bereich Zuhören im Saarland und im Bereich Mathematik in Baden-Württemberg sowie Niedersachsen zu verzeichnen. In Nordrhein-Westfalen und Rheinland-Pfalz hingegen ergaben sich in allen getesteten Bereichen geringere zuwanderungsbezogene Disparitäten (z. B. Kompetenzbereich Lesen, beide Elternteile im Ausland geboren: $Diff=48$ Punkte und 51 Punkte).

Die Ergebnisse von Regressionsanalysen weisen darauf hin, dass sich die zuwanderungsbezogenen Disparitäten teilweise auf den sozialen Status, das Bildungsniveau der Eltern und die Häufigkeit, mit der in der Familie Deutsch gesprochen wird, zurückführen lassen. Im Bereich Zuhören bestehen jedoch auch nach Kontrolle dieser Merkmale des familiären Hintergrunds für Kinder mit zwei zugewanderten Elternteilen aus den meisten Herkunftsländern noch substanzielle Kompetenznachteile. Der Effekt der Familiensprache, der im Bereich Zuhören und in Mathematik auch noch nach Kontrolle des sozioökonomischen Status der Familien und dem Bildungshintergrund der Eltern statistisch signifikant bleibt, kann als Hinweis darauf gewertet werden, dass es erforderlich ist, Kindern mit nicht deutscher Familiensprache effektive institutionelle Lerngelegenheiten zum Erwerb der deutschen Sprache zur Verfügung zu stellen.

13.3 Ausgewählte Bedingungen von Lehr-Lern-Prozessen: Aus- und Fortbildung von Lehrkräften und Maßnahmen der Sprach- und Leseförderung

Die Analysen zur Zusammensetzung der Lehrerschaft in der Primarstufe zeigten, dass in allen Ländern die Lehrkräfte überwiegend weiblich sind und deutschlandweit etwa die Hälfte der Lehrkräfte 50 Jahre und älter ist. Ein substanzieller Anteil der Lehrkräfte, die in der Primarstufe die Fächer Deutsch und Mathematik

unterrichten, hat kein entsprechendes Fachstudium absolviert. Ein Vergleich der durchschnittlich erreichten Kompetenzen zwischen Klassen mit und Klassen ohne Fachlehrkraft ergab im Fach Mathematik einen deutlichen Unterschied. In Klassen, in denen Lehrkräfte mit einem mathematischen Fachstudium unterrichten, lagen die im Mittel erreichten Kompetenzen 18 Punkte höher als in Klassen mit einer Lehrkraft ohne ein entsprechendes Fachstudium.

Ein Vergleich der Rahmenbedingungen der Lehrerfortbildung hat gezeigt, dass sich die Länder im Grad der Verpflichtung und in den Unterstützungsangeboten (z. B. Bereitstellung schulinterner Fortbildungsbudgets) deutlich unterscheiden. Die Angaben über die besuchten Fortbildungen der Lehrkräfte lassen jedoch erkennen, dass in allen Ländern die Mehrzahl der Lehrkräfte Angebote der Lehrerfortbildung wahrnimmt. Über 80 Prozent aller deutschlandweit befragten Lehrkräfte geben an, innerhalb der letzten beiden Schuljahre mindestens eine Fortbildungsveranstaltung besucht zu haben, und etwas mehr als 40 Prozent berichten, an fünf und mehr Veranstaltungen teilgenommen zu haben. Im Vergleich der Länder fällt der Anteil von Lehrkräften mit dieser hohen Zahl an Fortbildungen in Bayern und Thüringen besonders hoch und in Baden-Württemberg vergleichsweise niedrig aus. Auf die Frage, welche Faktoren ihnen die Teilnahme an Fortbildungen erschweren, nennen die Lehrkräfte insbesondere die hohen zeitlichen Anforderungen des Lehrerberufs und nur zu einem vergleichsweise geringen Anteil die mangelnde Qualität beziehungsweise das fehlende Angebot an Lehrerfortbildungen.

Die von Lehrkräften häufig gewählten Fortbildungen beziehen sich insbesondere auf fachdidaktische Inhalte sowie die Nutzung von spezifischen Unterrichtsformen und -methoden. Einen hohen Fortbildungsbedarf benennen die Befragten vor allem zu Fragen der Förderung lernschwacher Schülerinnen und Schüler, zu Binnendifferenzierung und individueller Förderung sowie zur Integration beziehungsweise Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. Dies zeigt, dass sich die bestehende kulturelle, soziale und leistungsbezogene Heterogenität im Klassenzimmer auch in den Fortbildungswünschen der Lehrkräfte widerspiegelt und diese offenbar einen starken Bedarf haben, sich mit diesen Themen aktiv auseinanderzusetzen.

Die Analysen zu Maßnahmen der Sprach- und Leseförderung, die in den Ländern umgesetzt werden, ergeben ein Bild, das erhebliche Anstrengungen der Schulen in diesem Bereich belegt. Die Schulen stellen insgesamt ein vielfältiges Angebot der Sprach- und Leseförderung zur Verfügung. So besuchen fast alle Schülerinnen und Schüler Schulen, die Zusatzkurse im Fach Deutsch beziehungsweise Förderkurse für lese- und rechtschreibschwache Kinder anbieten. Auch der Zugang zu Büchereien ist nahezu für alle Schülerinnen und Schüler gewährleistet. Das Angebot von Maßnahmen zur Förderung der Lesehäufigkeit und vor allem von speziellen Förderprogrammen im Bereich Lesen variiert dagegen stärker zwischen den Ländern.

Insgesamt haben nach Angaben der Schulleiterinnen und Schulleiter in allen Ländern mehr als 70 Prozent der Kinder, die im Fach Deutsch höchstens die Mindeststandards erreichen, Zugang zu systematischer Sprachförderung, entweder innerhalb des Regelunterrichts, zusätzlich zum Regelunterricht oder auch beides. Weniger verbreitet scheinen dagegen Sprachförderangebote zu sein, die ein konkretes Konzept verfolgen. In elf Ländern besuchen mehr als die Hälfte der Schülerinnen und Schüler, die im Fach Deutsch höchstens die Mindeststandards erreichen, eine Schule, in der entweder keine systematische Sprachförderung an-

geboten wird oder lediglich eine Sprachförderung, die sich an keinem konkreten Konzept orientiert.

Schülerinnen und Schüler, die eine zusätzliche Sprach- oder Leseförderung erhalten, haben in der Regel auch einen entsprechenden Bedarf, da sie im Fach Deutsch in den Bereichen Zuhören und Lesen höchstens die Mindeststandards erreichen. In fast allen Ländern erhält jedoch ein erheblicher Anteil von Schülerinnen und Schülern, die maximal die Mindeststandards erreichen, keine zusätzliche Sprachförderung (Kompetenzstufe I: von 50 % in Hamburg bis 85 % in Rheinland-Pfalz; Kompetenzstufe II: von 65 % in Hessen bis 89 % in Rheinland-Pfalz). Welche der möglichen Ursachen diesem Befund zugrunde liegen (z. B. die Kinder erhalten keine Sprachförderung, weil ihr Förderbedarf bislang unerkannt blieb oder weil die Eltern ein vorhandenes Sprachförderangebot nicht nutzen wollen, oder die Kinder erhalten ausschließlich in den Regelunterricht integrierte Sprachförderung) lässt sich anhand der vorliegenden Daten nicht prüfen.

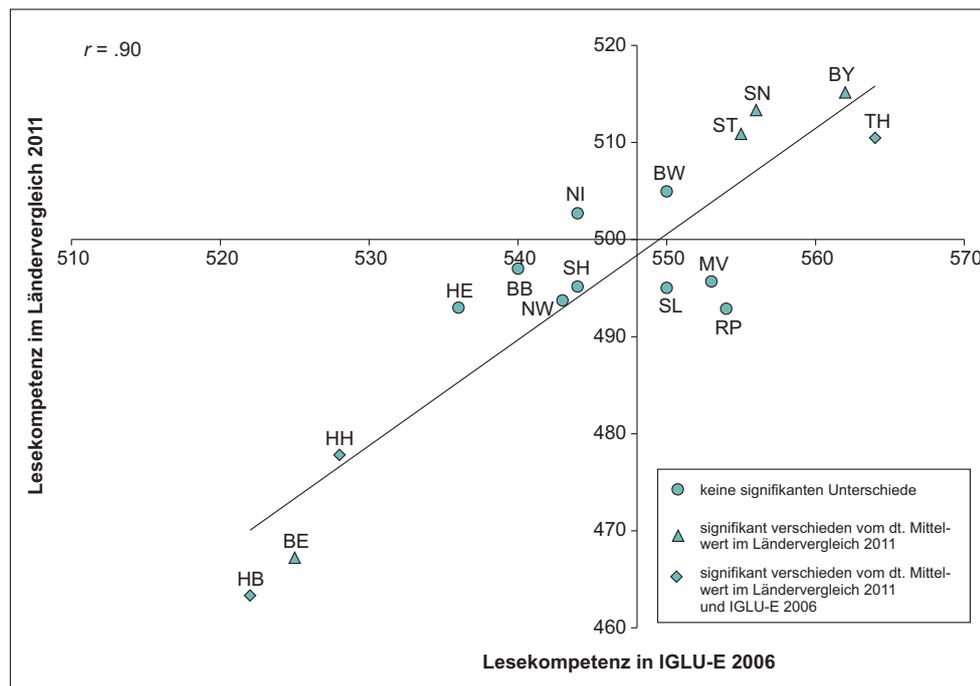
Ob ein Kind Sprach- oder Leseförderung erhält, hängt nicht nur von seinem Kompetenzniveau, sondern darüber hinaus auch von verschiedenen Hintergrundmerkmalen ab. Bei jeweils gleichem Kompetenzniveau in den Bereichen Lesen und Zuhören erhalten Jungen eher als Mädchen, Kinder aus zugewanderten Familien eher als Kinder aus Familien ohne Zuwanderungsgeschichte und Kinder aus sozial schwachen Familien eher als Kinder aus sozioökonomisch besser gestellten Familien zusätzliche Sprach- oder Leseförderung. Ob ein Kind eine solche Förderung erhält oder nicht, hängt dabei insbesondere mit dem Zuwanderungshintergrund zusammen. Demnach haben Kinder ohne Zuwanderungshintergrund bei gleichem Kompetenzniveau im Fach Deutsch eine geringere Chance, Sprach- oder Leseförderung zu erhalten, als Kinder aus zugewanderten Familien.

13.4 Einordnung der Befunde anhand der Ergebnisse früherer Ländervergleichsstudien

Der erste in Deutschland durchgeführte Ländervergleich im Primarbereich, an dem alle Länder beteiligt waren, fand 2006 im Rahmen der internationalen Schulleistungsstudie PIRLS/IGLU statt (IGLU-E 2006, Bos et al., 2008). Wie in Kapitel 5 ausführlich beschrieben, ist ein direkter Vergleich der Ergebnisse dieser früheren Untersuchung mit den Ergebnissen des Ländervergleichs 2011 nicht sinnvoll, da sich die Studien unter anderem in der Definition der Kompetenzkonstrukte, der Testkonstruktion und der Berichtsmetrik unterscheiden (vgl. hierzu Pietsch, Böhme, Robitzsch & Stubbe, 2009). Es ist daher keine Aussage darüber möglich, inwieweit sich die mittleren Kompetenzen der Schülerinnen und Schüler in Deutschland insgesamt oder in den einzelnen Ländern zwischen 2006 und 2011 verändert haben. Solche Trendaussagen werden erst möglich sein, wenn 2016 der IQB-Ländervergleich für die Primarstufe wiederholt wird.

Der Frage hingegen, inwieweit die relativen Positionen der Länder im IQB-Ländervergleich 2011 und in IGLU-E 2006 ähnlich sind, kann für den Kompetenzbereich Lesen anhand der vorliegenden Daten nachgegangen werden. Wie in Abbildung 13.1 deutlich zu erkennen ist, sind die Ergebnisse in dieser Hinsicht weitgehend stabil. Die Korrelation zwischen den Ländermittelwerten in den beiden Studien liegt bei $r = .90$. Damit besteht eine deutlich erkennbare Tendenz, dass Länder mit überdurchschnittlichen Mittelwerten in IGLU-E 2006

Abbildung 13.1: Mittelwerte der Länder im Kompetenzbereich Lesen in der Studie IGLU-E 2006 und im Ländervergleich 2011

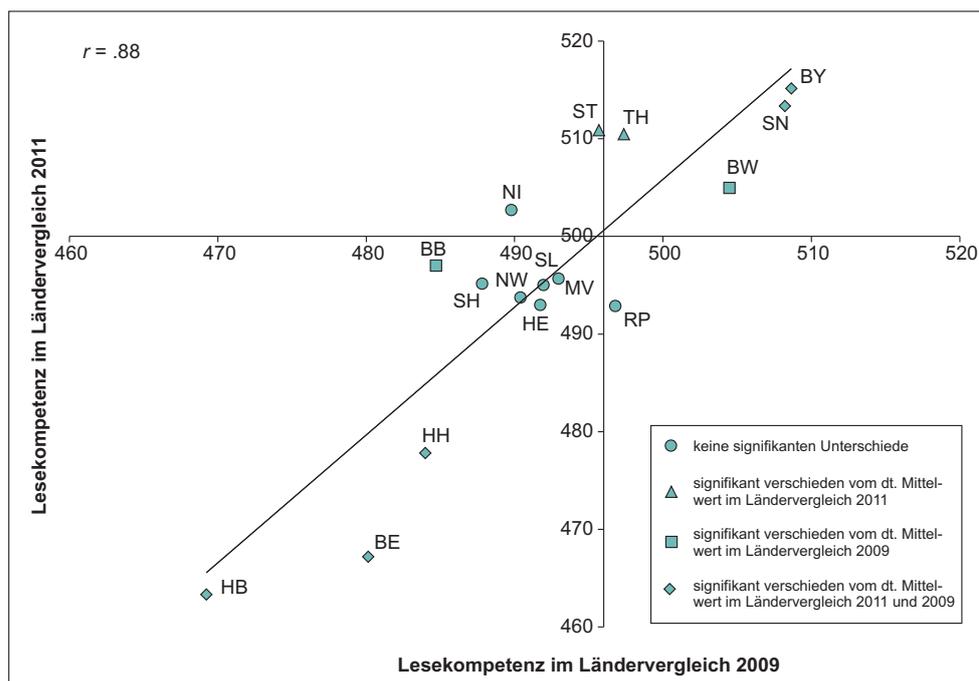


Anmerkung. BB=Brandenburg, BE=Berlin, BW=Baden-Württemberg, BY=Bayern, HB=Bremen, HE=Hessen, HH=Hamburg, MV=Mecklenburg-Vorpommern, NI=Niedersachsen, NW=Nordrhein-Westfalen, RP=Rheinland-Pfalz, SH=Schleswig-Holstein, SL=Saarland, SN=Sachsen, ST=Sachsen-Anhalt, TH=Thüringen.

auch im Ländervergleich 2011 im Lesen über dem deutschen Durchschnitt liegen und sich Länder mit unterdurchschnittlichen Mittelwerten in IGLU-E 2006 auch im Ländervergleich 2011 im unteren Leistungsspektrum bewegen. Größere Abweichungen in den relativen Positionen sind lediglich für Brandenburg und für Rheinland-Pfalz zu verzeichnen, wobei im Ländervergleich 2011 die Position Brandenburgs etwas günstiger, die Position von Rheinland-Pfalz hingegen etwas ungünstiger ausfällt als in IGLU-E 2006. Diese Unterschiede sollten aufgrund der genannten Einschränkungen der Vergleichbarkeit beider Studien jedoch nicht überinterpretiert werden.

Weiterhin stellt sich die Frage, inwieweit die in diesem Bericht dargestellten Befunde des Ländervergleichs 2011 für den Primarbereich ein ähnliches Muster aufweisen wie diejenigen des Ländervergleichs 2009 für die Sekundarstufe I (Köller, Knigge & Tesch, 2010). Unterschiede in der relativen Position der Länder in den beiden Studien könnten auf eine differenzielle Effektivität der schulischen Systeme in den beiden Schulstufen hinweisen. Die Ländermittelwerte der zwei Ländervergleichsstudien sind in Abbildung 13.2 für den Kompetenzbereich Lesen und in Abbildung 13.3 für den Kompetenzbereich Zuhören dargestellt. Auch hier gilt, dass die Kompetenzwerte an sich nicht vergleichbar sind und daher nur die relativen Positionen der Länder in den beiden Studien sinnvoll zueinander in Beziehung gesetzt werden können. Dabei ergibt sich für das Lesen mit einer Korrelation von $r = .88$ ein ähnlich enger Zusammenhang zwischen den Ländermittelwerten wie im Vergleich der Ergebnisse des Ländervergleichs 2011 mit den Ergebnissen aus IGLU-E 2006. Lediglich die relativen Positionen von Rheinland-Pfalz, Niedersachsen und Sachsen-Anhalt sind etwas verschoben. Während die Mittelwerte Niedersachsens und Sachsen-Anhalts in der Primarstufe relativ zu denen anderer Länder etwas

Abbildung 13.2: Mittelwerte der Länder im Kompetenzbereich Lesen im Ländervergleich 2009 (Sekundarstufe I) und im Ländervergleich 2011 (Primarstufe)



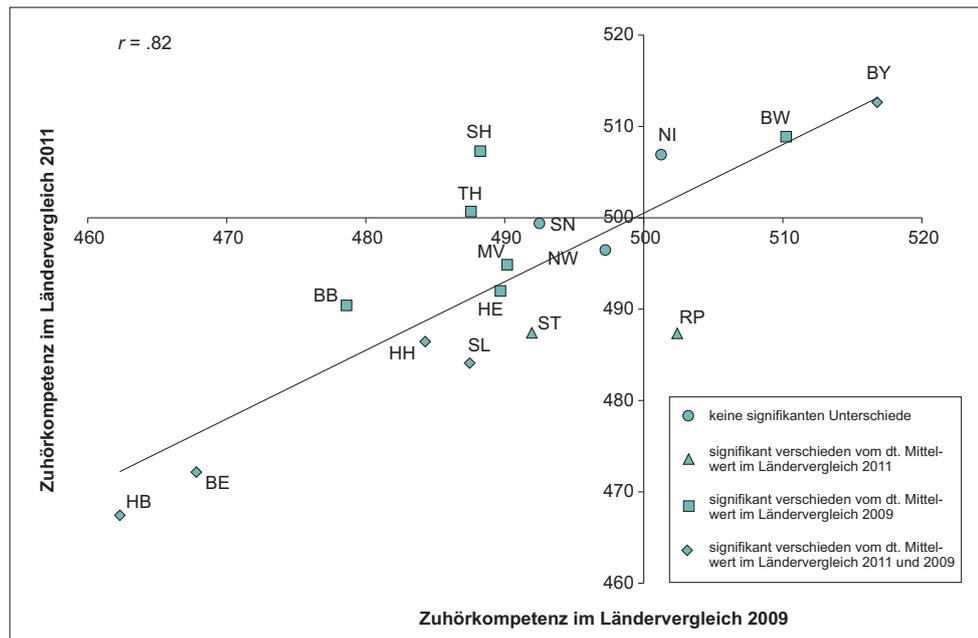
Anmerkung. BB=Brandenburg, BE=Berlin, BW=Baden-Württemberg, BY=Bayern, HB=Bremen, HE=Hessen, HH=Hamburg, MV=Mecklenburg-Vorpommern, NI=Niedersachsen, NW=Nordrhein-Westfalen, RP=Rheinland-Pfalz, SH=Schleswig-Holstein, SL=Saarland, SN=Sachsen, ST=Sachsen-Anhalt, TH=Thüringen.

besser ausfallen als in der Sekundarstufe I, ist in Rheinland-Pfalz das umgekehrte Muster zu verzeichnen. Absolut betrachtet ist der für Rheinland-Pfalz identifizierte Unterschied jedoch klein. Im Ländervergleich 2009 lag der Mittelwert des Landes einen Punkt über dem deutschen Durchschnitt, im Ländervergleich 2011 sieben Punkte darunter. Diese Differenz dürfte für Fragen der Steuerung kaum relevant sein.

Für den Kompetenzbereich Zuhören sind die Ergebnismuster des Ländervergleichs 2009 und des Ländervergleichs 2011 ebenfalls recht ähnlich, wenn auch die Korrelation zwischen den Mittelwerten etwas kleiner ausfällt ($r = .82$). Wiederum ist die relative Position von Rheinland-Pfalz in der Primarstufe etwas weniger günstig als in der Sekundarstufe I, wobei hier der Unterschied auch absolut betrachtet etwas größer ist als im Lesen. Schleswig-Holstein und Thüringen nehmen hingegen in der Primarstufe eine etwas bessere Position ein als in der Sekundarstufe I.

Allgemein weist die relative Konsistenz der Befunde darauf hin, dass in Ländern, in denen eher ungünstige Ergebnismuster zu beobachten sind, versucht werden sollte, die Qualität schulischer Lehr-Lern-Prozesse insgesamt zu optimieren. Dies gilt insbesondere für die Stadtstaaten und innerhalb dieser Gruppe vor allem für Berlin und Bremen. Dass die überwiegend schwachen Ergebnisse dieser Länder nicht nur auf die soziodemografische Zusammensetzung der Schülerschaft, wie etwa einen besonders hohen Anteil von Heranwachsenden aus sozial schwachen und zugewanderten Familien, zurückzuführen sind, wird anhand der Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen deutlich. So verfehlen in Berlin und Bremen nicht nur besonders viele Kinder der vierten Jahrgangsstufe die Mindeststandards, sondern es erreichen auch besonders wenige Kinder Optimalstandards (Kompetenzstufe V). Letzteres dürfte kaum auf ei-

Abbildung 13.3: Mittelwerte der Länder im Kompetenzbereich Zuhören im Ländervergleich 2009 (Sekundarstufe I) und im Ländervergleich 2011 (Primarstufe)



Anmerkung. BB=Brandenburg, BE=Berlin, BW=Baden-Württemberg, BY=Bayern, HB=Bremen, HE=Hessen, HH=Hamburg, MV=Mecklenburg-Vorpommern, NI=Niedersachsen, NW=Nordrhein-Westfalen, RP=Rheinland-Pfalz, SH=Schleswig-Holstein, SL=Saarland, SN=Sachsen, ST=Sachsen-Anhalt, TH=Thüringen.

nen Mangel an Schülerinnen und Schülern zurückzuführen sein, die über sehr gute Lernvoraussetzungen verfügen. So ist der sozioökonomische Status der 25 Prozent privilegiertesten Familien in den Stadtstaaten keineswegs schwächer, in Berlin sogar tendenziell etwas höher ausgeprägt als in anderen Ländern.

Einige Befunde des Ländervergleichs 2011 und früherer Ländervergleichsstudien weisen aber auch auf spezifische Stärken und Schwächen in den einzelnen Ländern hin. Hier wird zu prüfen sein, inwieweit die Schwächen so ausgeprägt sind, dass besondere Maßnahmen zu ihrer Verringerung erforderlich sein könnten. In den meisten Fällen dürfte es jedoch sinnvoller sein, die Strategien der Qualitätsentwicklung und Qualitätssicherung, die bereits auf den Weg gebracht worden sind, weiter zu optimieren und die einzelnen Bestandteile so aufeinander zu beziehen, dass ein kohärentes und transparentes System entsteht. Die Ergebnisse der PISA-Studien weisen darauf hin, dass sich die bisherigen Bemühungen der Entwicklung von Schul- und Unterrichtsqualität ausgezahlt haben und es daher sinnvoll erscheint, diese Strategie weiter zu verfolgen.

Literatur

- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2008). *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Köller, O., Knigge, M. & Tesch, B. (Hrsg.). (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Pietsch, M., Böhme, K., Robitzsch, A. & Stubbe, T. C. (2009). Das Stufenmodell zur Lesekompetenz der länderübergreifenden Bildungsstandards im Vergleich zu IGLU 2006. In D. Granzler, O. Köller, A. Bremerich-Vos, M. van den Heuvel-Panhuizen, K. Reiss & G. Walther (Hrsg.), *Bildungsstandards Deutsch und Mathematik. Leistungsmessung in der Grundschule* (S. 393–416). Weinheim: Beltz.