

IQ:

Petra Stanat
Stefan Schipolowski
Camilla Rjosk
Sebastian Weirich
Nicole Haag
(Hrsg.)



IQB-Bildungstrend 2016

Kompetenzen in den Fächern Deutsch und
Mathematik am Ende der 4. Jahrgangsstufe
im zweiten Ländervergleich

WAXMANN

Petra Stanat, Stefan Schipolowski, Camilla Rjosk,
Sebastian Weirich, Nicole Haag (Hrsg.)

IQB-Bildungstrend 2016

Kompetenzen in den Fächern Deutsch und Mathematik am
Ende der 4. Jahrgangsstufe im zweiten Ländervergleich



Waxmann 2017
Münster · New York

Bibliografische Informationen der Deutschen Nationalbibliothek

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

ISBN 978-3-8309-3730-2

© 2017, Waxmann Verlag GmbH,
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Christian Averbeck, Münster
Umschlagfoto: © drubig-photo – fotolia.de
Satz: Stoddart Satz- und Layoutservice, Münster
Druck: Mediaprint, Paderborn

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

Danksagung	9
-------------------------	----------

Vorwort der Präsidentin der Kultusministerkonferenz	11
--	-----------

Kapitel 1

Konzeptuelle Grundlagen des IQB-Bildungstrends 2016	13
--	-----------

1.1 Die IQB-Bildungstrends als zentrales Element des Bildungsmonitorings in Deutschland	13
Petra Stanat, Stefan Schipolowski, Camilla Rjosk und Hans Anand Pant	
1.2 Beschreibung der im Fach Deutsch untersuchten Kompetenzen.....	20
Katrin Böhme, Maria Engelbert und Sebastian Weirich	
1.3 Beschreibung der im Fach Mathematik untersuchten Kompetenzen	31
Alexander Roppelt, Kristina Reiss und Nicole Haag	

Kapitel 2

Kompetenzstufenmodelle für den Primarbereich	45
---	-----------

2.1 Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen	45
Hans Anand Pant, Katrin Böhme, Petra Stanat, Stefan Schipolowski und Olaf Köller	
2.2 Kompetenzstufenmodelle im Fach Deutsch.....	53
Albert Bremerich-Vos, Katrin Böhme, Michael Krelle, Sebastian Weirich und Olaf Köller	
2.3 Kompetenzstufenmodelle im Fach Mathematik.....	71
Kristina Reiss, Alexander Roppelt, Nicole Haag, Hans Anand Pant und Olaf Köller	

Kapitel 3

Anlage, Durchführung und Auswertung des IQB-Bildungstrends 2016.....	83
---	-----------

3.1 Anlage und Durchführung.....	83
Camilla Rjosk, Maria Engelbert, Stefan Schipolowski und Pauline Kohrt	
3.2 Auswertung, Trendschätzung und Ergebnisdarstellung	106
Nicole Haag und Sebastian Weirich	

Kapitel 4**Kontextinformationen zu den Schulsystemen der Länder im Primarbereich..... 115**

Stefan Schipolowski, Petra Stanat und Camilla Rjosk

4.1	Einleitung.....	115
4.2	Ausgewählte Kontextmerkmale der Schulsysteme der Länder	116
4.3	Zusammenfassung.....	127

Kapitel 5**Kompetenzstufenbesetzungen im Ländervergleich..... 129**

5.1	Kompetenzstufenbesetzungen im Fach Deutsch.....	129
	Sebastian Weirich, Julia Wittig und Petra Stanat	
5.2	Kompetenzstufenbesetzungen im Fach Mathematik.....	140
	Pauline Kohrt, Nicole Haag und Petra Stanat	

Kapitel 6**Mittelwerte und Streuungen der erreichten Kompetenzen im Ländervergleich..... 153**

6.1	Mittelwerte und Streuungen der im Fach Deutsch erreichten Kompetenzen.....	153
	Julia Wittig und Sebastian Weirich	
6.2	Mittelwerte und Streuungen der im Fach Mathematik erreichten Kompetenzen.....	168
	Nicole Haag und Pauline Kohrt	

Kapitel 7**Geschlechtsbezogene Disparitäten..... 187**

Stefan Schipolowski, Julia Wittig, Sebastian Weirich und Katrin Böhme

7.1	Geschlechtsbezogene Unterschiede in schulischen Kompetenzen und motivationalen Merkmalen.....	188
7.2	Geschlechtsbezogene Kompetenzunterschiede im Jahr 2016.....	192
7.3	Geschlechtsbezogene Kompetenzunterschiede in den Jahren 2011 und 2016 im Vergleich.....	197
7.4	Geschlechtsbezogene Unterschiede in motivationalen Merkmalen.....	204
7.5	Zusammenfassung.....	208

Kapitel 8**Soziale Disparitäten..... 213**

Nicole Haag, Aleksander Kocaj, Malte Jansen und Poldi Kuhl

8.1	Einleitung.....	213
8.2	Indikatoren sozialer Disparitäten.....	214
8.3	Anmerkungen zur Ergebnisdarstellung	216
8.4	Verteilung des sozioökonomischen Status der Schülerinnen und Schüler	217

8.5	Soziale Gradienten.....	220
8.6	Kompetenzniveau nach EGP-Extremgruppen.....	224
8.7	Zusammenfassung und Diskussion	232

Kapitel 9

Zuwanderungsbezogene Disparitäten 237

Camilla Rjosk, Nicole Haag, Birgit Heppt und Petra Stanat

9.1	Analysen zuwanderungsbezogener Disparitäten als Gegenstand des Bildungsmonitorings	237
9.2	Methodisches Vorgehen.....	238
9.3	Kinder aus zugewanderten Familien und Kinder ohne Zuwanderungshintergrund in den Ländern	241
9.4	Erreichte Kompetenzen von Viertklässlerinnen und Viertklässlern.....	246
9.5	Merkmale der sozialen Integration und schulischen Motivation von Viertklässlerinnen und Viertklässlern nach Zuwanderungsstatus im Jahr 2016.....	265
9.6	Zusammenfassung und Diskussion	269

Kapitel 10

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Deutschland 277

10.1	Einleitung.....	277
	Cornelia Gresch, Camilla Rjosk, Aleksander Kocaj und Petra Stanat	
10.2	Amtlich festgestellter sonderpädagogischer Förderbedarf und sonderpädagogische Förderung	282
	Cornelia Gresch, Jenny Kölm und Aleksander Kocaj	
10.3	Hintergrundmerkmale von Schülerinnen und Schülern mit sonder- pädagogischem Förderbedarf an Förderschulen und an allgemeinen Schulen.....	291
	Jenny Kölm, Cornelia Gresch und Nicole Haag	
10.4	Schulische Kompetenzen und schulische Motivation von Kindern mit sonderpädagogischem Förderbedarf an Förderschulen und an allgemeinen Schulen....	302
	Aleksander Kocaj, Poldi Kuhl, Nicole Haag, Pauline Kohrt und Petra Stanat	

Kapitel 11

Leistungsstarke Schülerinnen und Schüler in Deutschland..... 317

Claudia Neuendorf, Poldi Kuhl und Malte Jansen

11.1	Einleitung.....	317
11.2	Leistungsstarke Viertklässlerinnen und Viertklässler in den Fächern Deutsch und Mathematik.....	318
11.3	Merkmale leistungsstarker Schülerinnen und Schüler.....	322
11.4	Schulische Umwelt.....	328
11.5	Fazit	331

Kapitel 12**Qualifikation von Lehrkräften und Einschätzungen zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf..... 335**

Camilla Rjosk, Lars Hoffmann, Dirk Richter, Alexandra Marx und Cornelia Gresch

12.1	Einleitung.....	335
12.2	Demografische Hintergrundmerkmale der Lehrkräfte	336
12.3	Berufliche Qualifikation von Lehrkräften und erreichte Kompetenzen von Schülerinnen und Schülern.....	340
12.4	Berufliche Qualifikation und Einschätzungen von Lehrkräften zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf.....	344
12.5	Zusammenfassung und Diskussion	348

Kapitel 13**Grundlagen der Kompetenzmessung im IQB-Bildungstrend 2016..... 355**

13.1	Testdesign und Auswertung des IQB-Bildungstrends 2016	355
	Sebastian Weirich, Nicole Haag und Karoline A. Sachse	
13.2	Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf.....	369
	Katrin Böhme, Pauline Kohrt, Nicole Haag und Sebastian Weirich	

Kapitel 14**Zusammenfassung und Einordnung der Befunde 387**

Petra Stanat, Stefan Schipolowski, Camilla Rjosk, Sebastian Weirich und Nicole Haag

14.1	Erreichen der Bildungsstandards in den Ländern im Jahr 2016.....	389
14.2	Erreichen der Bildungsstandards in den Ländern im Trend	394
14.3	Durchschnittliches Niveau der Kompetenzen im Jahr 2016 und im Trend	398
14.4	Geschlechtsbezogene, soziale und zuwanderungsbezogene Disparitäten.....	401
14.5	Umgang mit Heterogenität	405
14.6	Qualifikation von Lehrkräften und Einschätzungen zu gemeinsamem Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf.....	408
14.7	Fazit	409

Danksagung

Zur Erstellung des Berichts über den IQB-Bildungstrend 2016 haben viele Personen beigetragen. An dieser Stelle möchten wir uns bei allen bedanken, die uns bei diesem umfangreichen Projekt unterstützt haben. Personen, die durch inhaltliche und logistische Vorarbeiten die Durchführung der Studie ermöglicht haben, wird namentlich in Kapitel 3.1 gedankt.

Ein herzlicher Dank der Herausgeberinnen und Herausgeber gilt Jana Bastian-Wurzel, Simone Dubiel, Daniela Gensing, Friederike Keiderling, Katharina Krohmer, Ulrike Lippmann und Sophie Petershagen für ihre Unterstützung bei der Zusammenstellung und beim Layout der Testmaterialien.

Ein besonderer Dank geht auch an Johanna Busse, Aleksandr Merkulov, Felix Milles, Stefanie Pietz und Franziska Schwarzer für die Erstellung und Prüfung der zahlreichen Abbildungen und Tabellen. Weiterhin danken wir Eljesa Haxhiu, Ricarda Klein, Mischa Mangel, Juliane Neumann, Franziska Schwarzer und Semiha Yilmaz herzlich für redaktionelle Arbeiten an den Manuskripten.

Aber auch alle anderen Mitarbeiterinnen und Mitarbeiter des IQB waren direkt oder indirekt an den Arbeiten beteiligt, die im Rahmen des IQB-Bildungstrends 2016 zu leisten waren, und sind bereitwillig eingesprungen, wenn zusätzliche Unterstützung benötigt wurde. Dafür ganz herzlichen Dank.

Für die hilfreichen Rückmeldungen und Anregungen zu einzelnen Kapiteln bedanken sich die Herausgeberinnen und Herausgeber bei Werner Klein und Prof. Dr. Birgit Lütje-Klose. Dr. Alexander Robitzsch und Karoline Sachse danken wir für die kompetente Unterstützung in methodischen Fragen.

Dem Team vom Waxmann Verlag gilt unser herzlicher Dank für die professionelle, freundliche und geduldige Zusammenarbeit bei der Fertigstellung der Druckfassung dieses Berichts.

Schließlich möchten wir uns ganz besonders bei den Autorinnen und Autoren der vorausgegangenen Berichte zu den Ländervergleichsstudien und Bildungstrends des IQB bedanken, von deren Erfahrungen und Vorarbeiten wir im gesamten Verlauf der Berichtlegung profitiert haben. Beim vorliegenden Bericht handelt es sich um einen Band in einer fortlaufenden Reihe von Berichten über die IQB-Ländervergleichsstudien und IQB-Bildungstrends. Beschreibungen konstanter Sachverhalte wurden aus diesem Grund teilweise wörtlich aus früheren Berichtsbänden der Reihe übernommen, ohne dies im Einzelnen zu kennzeichnen.

Vorwort der Präsidentin der Kultusministerkonferenz

Mit dem IQB-Bildungstrend 2016 werden zum zweiten Mal die Kompetenzen der Schülerinnen und Schüler der 4. Jahrgangsstufe in den Fächern Deutsch und Mathematik auf Grundlage der Bildungsstandards untersucht.

Die bundesweit geltenden Bildungsstandards der Kultusministerkonferenz beschreiben Leistungserwartungen in Form fachlicher Kompetenzanforderungen, über die Schülerinnen und Schüler bis zum Ende der 4. Jahrgangsstufe, der Sekundarstufe I und der Sekundarstufe II verfügen sollen. Auf der Grundlage dieses für alle Länder verbindlichen Referenzrahmens können die Erträge des Bildungssystems regelmäßig überprüft werden.

Für den Primarbereich und die Sekundarstufe I wird im Rahmen der vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) durchgeführten Bildungstrends untersucht, inwieweit in den einzelnen Ländern die in den Bildungsstandards formulierten Kompetenzanforderungen vor Abschluss des jeweiligen Bildungsabschnitts tatsächlich erreicht werden. Die Studien des IQB richten sich damit stärker als internationale Schulleistungsuntersuchungen wie PISA, IGLU und TIMSS und die bis 2006 durchgeführten nationalen Erweiterungen der internationalen Vergleichsstudien (PISA-E und IGLU-E) an der Unterrichtspraxis in den Schulen aus.

Zur Überprüfung des Erreichens der Bildungsstandards hat das IQB fachdidaktisch und lernpsychologisch abgesicherte Kompetenzstufenmodelle entwickelt, die einen Vergleichsmaßstab bereitstellen, der neben den in den Bildungsstandards formulierten Regelstandards auch Mindeststandards und Optimalstandards festlegt.

Erstmalig wurde das Erreichen der Bildungsstandards für den Primarbereich im Jahr 2011 durch das IQB überprüft. Mit dem IQB-Bildungstrend 2016 liegen hierzu nun Daten im Zeitvergleich vor. Wir können damit differenzierte und evidenzbasierte Aussagen darüber machen, inwieweit sich das Kompetenzniveau der Schülerinnen und Schüler der 4. Jahrgangsstufe im Fach Deutsch in den Kompetenzbereichen Lesen, Zuhören und Orthografie sowie im Fach Mathematik im Vergleich zwischen den Jahren 2011 und 2016 verändert hat.

Im Fokus der Bildungstrends steht zudem stärker als in früheren Ländervergleichsstudien die Frage, in welchem Umfang die Schülerinnen und Schüler die verschiedenen Kompetenzstufen erreichen. So wird ersichtlich, wie hoch der Anteil der Viertklässlerinnen und Viertklässler ist, der in den genannten Kompetenzbereichen die Regel- oder sogar die Optimalstandards erreicht oder vielleicht auch die Mindeststandards verfehlt.

Die vorliegenden Ergebnisse der zweiten Überprüfung des Erreichens der Bildungsstandards für den Primarbereich im Ländervergleich liefern damit eine wertvolle Bilanz der Anstrengungen der vergangenen Jahre in Bildungspolitik

und Bildungspraxis: Wo stehen wir im Hinblick auf die mit den Bildungsstandards gemeinsam verabredeten Kompetenzerwartungen? Wo haben wir uns gegenüber den Ergebnissen aus dem Jahr 2011 verbessert? Wo besteht besonderer Handlungsbedarf? Wo werden neue Herausforderungen sichtbar?

Zwei ergänzende Analysen liefern uns darüber hinaus Forschungsbefunde zum Themenfeld „Umgang mit Heterogenität“, dem die Kultusministerkonferenz im Rahmen ihrer Gesamtstrategie zum Bildungsmonitoring besondere bildungspolitische Bedeutung zumisst. Eine Zusatzanalyse richtet ihr besonderes Augenmerk auf Aspekte der Lernsituation von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. Eine weitere Zusatzanalyse widmet sich Aspekten der Lernsituation leistungsstarker Schülerinnen und Schüler.

Der IQB-Bildungstrend 2016 reiht sich in die bisherigen und noch kommenden Ländervergleichsstudien des IQB ein und stellt ein Instrument des Bildungsmonitorings dar. Im vergangenen Jahr wurde mit der Veröffentlichung der Ergebnisse des IQB-Bildungstrends 2015, der sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe untersucht, der zweite Zyklus zur Überprüfung des Erreichens der Bildungsstandards eröffnet. Im Jahr 2018 wird der zweite Zyklus mit der Untersuchung der mathematischen und naturwissenschaftlichen Kompetenzen am Ende der 9. Jahrgangsstufe abgeschlossen. Im Jahr 2020 folgt abermals die Untersuchung der Kompetenzen in der 4. Jahrgangsstufe, womit dann der dritte Zyklus zur Überprüfung des Erreichens der Bildungsstandards eröffnet wird.

Dem IQB und seinen Mitarbeiterinnen und Mitarbeitern, die den aktuellen Bildungstrend in wissenschaftlicher Unabhängigkeit und „nach allen Regeln der Kunst“ erstellt haben, gebührt meine Anerkennung und mein Dank. Mit den Ergebnissen können die Entwicklungen in den Ländern beobachtet, Stärken und Schwächen analysiert und daraus Hinweise auf bildungspolitischen Handlungsbedarf datengestützt abgeleitet werden.

Berlin, im Oktober 2017

Ministerin Dr. Susanne Eisenmann
Präsidentin der Ständigen Konferenz der
Kultusminister der Länder in der Bundesrepublik Deutschland

Kapitel 1

Konzeptuelle Grundlagen des IQB-Bildungstrends 2016

1.1 Die IQB-Bildungstrends als zentrales Element des Bildungsmonitorings in Deutschland¹

Petra Stanat, Stefan Schipolowski, Camilla Rjosk und Hans Anand Pant

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, mit dem die bis dahin dominierende *Input-* und *Prozessorientierung* durch verschiedene Elemente einer *Output-*Steuerung ergänzt wurde (vgl. ausführlicher z. B. Böhme, Richter, Stanat, Pant & Köller, 2012; Klieme & Tippelt, 2008; Köller, 2010). Die Entwicklung und Sicherung der Bildungsqualität stützt sich seitdem stärker auf Erkenntnisse über die Bildungserträge der Schülerinnen und Schüler, der Schulen und des gesamten schulischen Bildungssystems.

Eine zentrale Grundlage für die Umsetzung dieses Paradigmenwechsels bilden in Deutschland die Bildungsstandards der KMK, die fächerspezifisch festlegen, welche Kompetenzen Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt in ihrer Schullaufbahn entwickelt haben sollen. Unter einer Kompetenz wird dabei die Fähigkeit verstanden, Wissen und Können in den jeweiligen Fächern zur Lösung von Problemen anzuwenden (vgl. Kapitel 2.1). Als normativ gesetzte Kompetenzerwartungen sind die Bildungsstandards das Resultat eines intensiven Diskussions- und Beratungsprozesses, an dem Vertreterinnen und Vertreter der Bildungspolitik und Bildungsadministration, der Fachdidaktiken und anderer Bildungswissenschaften sowie der Schulpraxis beteiligt waren. Die verbindliche Einführung der Bildungsstandards für den Primarbereich, deren Erreichen am Ende der 4. Jahrgangsstufe im IQB-Bildungstrend 2016 untersucht wird, erfolgte in allen Ländern in der Bundesrepublik Deutschland zum Schuljahresbeginn 2004/2005.

Die Bildungsstandards der KMK beziehen sich auf Bildungsabschlüsse beziehungsweise Bildungsabschnitte und legen fest, welche Anforderungen Schülerinnen und Schüler zu bewältigen in der Lage sein sollen, wenn sie die

¹ Die konzeptuellen Grundlagen der Untersuchungen des IQB zur Überprüfung des Erreichens der Bildungsstandards wurden bereits in den Berichten zu den IQB-Ländervergleichsstudien 2009, 2011 und 2012 umfassend dargestellt (Köller, Knigge & Tesch, 2010; Stanat, Pant, Böhme & Richter, 2012; Pant et al., 2013). Im Folgenden werden sie daher nicht noch einmal ausführlich beschrieben, sondern nur grob skizziert. Ferner wird auf relevante Weiterentwicklungen der konzeptuellen Grundlagen und der Berichtlegung eingegangen.

jeweilige Bildungsetappe abgeschlossen haben. Abweichend von den Empfehlungen der sogenannten Klieme-Expertise zur Entwicklung nationaler Bildungsstandards (Klieme et al., 2003) hat die KMK keine Mindeststandards, sondern Regelstandards formuliert. Die länderübergreifenden Vorgaben beschreiben also Kompetenzerwartungen, die Schülerinnen und Schüler am Ende der jeweiligen Bildungsetappe „in der Regel“ oder „im Durchschnitt“ erfüllen sollten (Klieme et al., 2003, S. 138). Mit Verabschiedung der Kompetenzstufenmodelle, die vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) auf der Grundlage der Bildungsstandards entwickelt wurden (vgl. Kapitel 2), hat die KMK jedoch später auch Mindestanforderungen festgelegt, die alle Schülerinnen und Schüler bis zum Abschluss der jeweiligen Bildungsetappe erreichen sollen.

Die Bildungsstandards spielen eine zentrale Rolle in der Gesamtstrategie der KMK zum Bildungsmonitoring, die den gemeinsamen Rahmen der Länder für eine an den Ergebnissen von Bildungsprozessen orientierte Steuerung des Bildungswesens bildet. Die zuerst im Jahr 2006 verabschiedete Gesamtstrategie (KMK, 2006) bündelte die verschiedenen Maßnahmen der datenbasierten Qualitätssicherung, die als Folge des sogenannten Konstanzer Beschlusses der KMK von 1997 (KMK, 1997) durch die Länder gemeinsam umgesetzt wurden, in vier Säulen:

1. Teilnahme Deutschlands an internationalen Schulleistungsuntersuchungen in der Primarstufe und Sekundarstufe I,
2. zentrale Überprüfungen des Erreichens der Bildungsstandards im Ländervergleich,
3. Durchführung von Vergleichsarbeiten in Anbindung oder Ankoppelung an die Bildungsstandards zur landesweiten oder länderübergreifenden Überprüfung der Leistungsfähigkeit aller Schulen sowie
4. gemeinsame Bildungsberichterstattung von Bund und Ländern.

Nachdem über einen Zeitraum von mehr als 10 Jahren Erfahrungen mit den verschiedenen Instrumenten des Bildungsmonitorings gesammelt werden konnten, wurde in den vergangenen Jahren der Erfolg der bis dahin umgesetzten Strategie überprüft. Dazu fand ein von der Kultusministerkonferenz getragener Diskussionsprozess statt, an dem auch Vertreterinnen und Vertreter der Bildungswissenschaften, Schülerinnen und Schüler, Eltern, Verbände, Gewerkschaften und Medien beteiligt waren. Als Resultat dieses Prozesses sah sich die KMK darin bestätigt, „dass die 2006 formulierten Ziele der Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring nach wie vor von hoher Relevanz sind“ (KMK, 2015, S. 5). Die Änderungen, die als Konsequenz dieses Diskussionsprozesses im Rahmen der Überarbeitung der Gesamtstrategie vorgenommen wurden, zielen daher vor allem auf eine Fortschreibung, Optimierung und Ergänzung des als grundsätzlich erfolgreich angesehenen Weges ab. In der überarbeiteten Gesamtstrategie der KMK zum Bildungsmonitoring, die im Juni 2015 verabschiedet wurde, werden

- die Umsetzung von Bildungsstandards (Implementation) stärker betont,
- die Vergleichsarbeiten (VERA) als Teil eines Bündels von Maßnahmen zur Qualitätssicherung auf Ebene der Schulen verortet,
- die im Jahr 2012 beschlossenen Bildungsstandards für die Allgemeine Hochschulreife und Einrichtung eines gemeinsamen Abituraufgabenpools der Länder berücksichtigt und
- der Bedarf nach mehr anwendungsbezogenem Wissen für Bildungspolitik und pädagogische Praxis hervorgehoben, indem unter anderem Themenfelder

für Forschungsfragen benannt werden, die von zentraler bildungspolitischer Bedeutung sind und sich auf praktische Schlüsselfragen der Schul- und Unterrichtsentwicklung beziehen.

Unter Einbeziehung dieser Modifikationen werden die vier Säulen der Gesamtstrategie in der Überarbeitung von 2015 wie folgt fortgeschrieben:

1. Teilnahme an internationalen Schulleistungsstudien (PIRLS/IGLU², TIMSS³-Grundschule, PISA⁴),
2. Überprüfung und Umsetzung von Bildungsstandards für die Primarstufe, die Sekundarstufe I und die Allgemeine Hochschulreife,
3. Verfahren zur Qualitätssicherung auf Ebene der Schulen sowie
4. gemeinsame Bildungsberichterstattung von Bund und Ländern.

Die erste Säule der Gesamtstrategie zielt darauf ab, die Leistungsfähigkeit des deutschen Bildungssystems im internationalen Vergleich festzustellen. Die Ergebnisse beziehen sich auf die in den internationalen Schulleistungsstudien getesteten Kompetenzbereiche und lassen Aussagen darüber zu, welche Leistungen die Schülerinnen und Schüler in Deutschland im Vergleich zu Schülerinnen und Schülern in anderen Staaten erzielen.

Bei der zweiten Säule geht es um die Frage, inwieweit innerhalb Deutschlands auf Ebene der Länder die mit den Bildungsstandards der KMK kriterial vorgegebenen Ziele für die Kompetenzentwicklung von Schülerinnen und Schülern in zentralen Bereichen erreicht werden. Dabei wird ein breiteres Spektrum an Fächern und Kompetenzbereichen untersucht als in den internationalen Schulleistungsstudien. So wird zum Beispiel im sprachlichen Bereich in den internationalen Studien ausschließlich die Lesekompetenz in den jeweiligen Landessprachen erfasst, während die bereits durchgeführten Studien zur Überprüfung der Bildungsstandards in Deutschland im Fach Deutsch zusätzlich die Kompetenzbereiche *Zuhören* und *Orthografie* einbeziehen.

Mit der Überprüfung des Erreichens der Bildungsstandards ist das Institut zur Qualitätsentwicklung im Bildungswesen betraut, das von der KMK im Jahr 2004 als An-Institut der Humboldt-Universität zu Berlin gegründet wurde. Unter Federführung des IQB und in enger Zusammenarbeit mit Fachdidaktikerinnen und Fachdidaktikern entwickeln Lehrkräfte Testaufgaben, mit denen die in den Bildungsstandards definierten Kompetenzen in zentralen Bereichen erfasst werden können. Weiterhin erarbeitet das IQB fachdidaktisch und lernpsychologisch fundierte Kompetenzstufenmodelle, die zur inhaltlichen Interpretation von Testwerten dienen. Anhand dieser Modelle lässt sich inhaltlich beschreiben, welche Anforderungen Schülerinnen und Schüler, die ein bestimmtes Testergebnis erzielt haben, bewältigen können.

Die Kompetenzstufenmodelle des IQB sind so konstruiert, dass sie ein breites Leistungsspektrum in wenige Stufen unterteilen, die sich sinnvoll abgrenzen lassen. Neben der Stufe für den von der KMK definierten Regelstandard werden in jedem Modell weitere Stufen ausgewiesen, die Leistungen unter oder über dem jeweiligen Regelstandard beschreiben. Die Kompetenzstufenmodelle für den

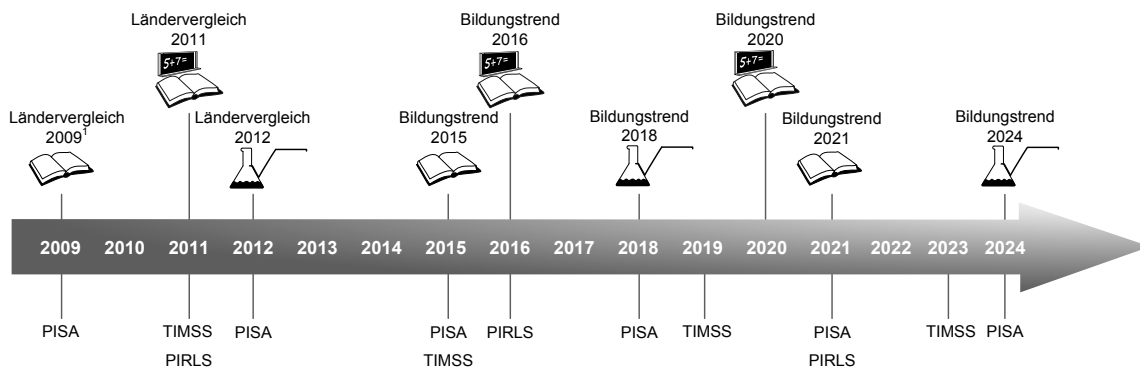
2 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.

3 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.




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

Primarbereich in den Fächern Deutsch und Mathematik werden in Kapitel 2 dieses Berichts beschrieben. Ausführungen zu den untersuchten Kompetenzen und deren Operationalisierung finden sich außerdem in den folgenden Kapiteln 1.2 und 1.3.

Abbildung 1.1: Zeitpunkte der Datenerhebungen der IQB-Ländervergleichsstudien bzw. IQB-Bildungstrends sowie der internationalen Schulleistungsstudien von 2009 bis 2024



IQB-Ländervergleichsstudien (bis zum Jahr 2012) bzw. IQB-Bildungstrends (ab dem Jahr 2015) auf Basis der Bildungsstandards in den folgenden Fächern und Schulstufen:

-  Deutsch, Englisch und Französisch in der Sekundarstufe I
-  Deutsch und Mathematik in der Primarstufe
-  Mathematik und Naturwissenschaften in der Sekundarstufe I

Anmerkung: ¹ Die Datenerhebung für das Fach Französisch fand bereits im Jahr 2008 statt.

Das IQB führt die Studien zur Überprüfung des Erreichens der Bildungsstandards im Auftrag der Kultusministerkonferenz in regelmäßigen Abständen durch. Die Erhebungen, die in zeitlicher Anknüpfung an die internationalen Schulleistungsstudien erfolgen, finden in der Primarstufe in den Fächern Deutsch und Mathematik in der Regel alle fünf Jahre, in der Sekundarstufe I alternierend in den Fächergruppen Deutsch, Englisch und Französisch einerseits sowie Mathematik, Biologie, Chemie und Physik andererseits alle drei Jahre statt (vgl. Abb. 1.1). Diesem Turnus folgend, war die Durchführung des nächsten IQB-Bildungstrends in der Primarstufe zunächst für das Jahr 2021 vorgesehen, in dem auch die Erhebung zum IQB-Bildungstrend in der Sekundarstufe I für die sprachlichen Fächer stattfinden soll. Da zeitgleiche Erhebungen zu zwei Bildungstrendstudien jedoch weder für das IQB noch für die Länder zu bewältigen wären, wird die dritte Studie im Primarbereich auf das Jahr 2020 vorgezogen.

Mit den Ländervergleichsstudien, die das IQB in den Jahren 2009 (Sekundarstufe I: Deutsch, Englisch, Französisch), 2011 (Primarstufe: Deutsch, Mathematik) und 2012 (Sekundarstufe I: Mathematik, Biologie, Chemie, Physik) durchgeführt hat, konnte der erste Zyklus der Überprüfungen des Erreichens der Bildungsstandards abgeschlossen werden. Der zweite Zyklus der Studien, die nunmehr als „IQB-Bildungstrends“ bezeichnet werden, begann mit der im Jahr 2015 durchgeführten Untersuchung zu den sprachlichen Fächern in der Sekundarstufe I und wird mit der im Jahr 2016 durchgeführten Untersuchung im Primarbereich fortgesetzt, über die der vorliegende Band berichtet. Die im IQB-Ländervergleich 2011 untersuchten Kompetenzen in den Fächern Deutsch und Mathematik wurden im Jahr 2016 erneut erfasst. Mit der wiederholten Messung

erweitert sich das Spektrum der Analysen erstmalig auch für den Primarbereich auf alle drei Vergleichsperspektiven, die bei der Bewertung von Ergebnissen eines Bildungsmonitorings herangezogen werden können (vgl. auch Stanat, Pant, Richter, Pöhlmann & Kuhl, 2013), wodurch der Informationsgehalt der Ergebnisse erheblich steigt:

1. Beim *sozialen Vergleich* werden die Ergebnisse der jeweiligen Untersuchungseinheiten miteinander verglichen, in den Studien des IQB also die Ergebnisse der Länder in der Bundesrepublik Deutschland. Bei dieser Vergleichsperspektive steht die Frage im Vordergrund, ob die Schülerinnen und Schüler in einem Land bessere oder schwächere Leistungen in den Kompetenztests erzielt haben als die Schülerinnen und Schüler in einem anderen Land oder in Deutschland insgesamt.
2. Beim *kriterialen Vergleich* bezieht sich die Bewertung der Ergebnisse auf vorab definierte Kriterien oder Zielvorgaben. In den Studien des IQB sind dies die Bildungsstandards der KMK und die darauf basierenden Kompetenzstufenmodelle, die zusätzlich zum Regelstandard einen Mindeststandard, einen Regelstandard plus und einen Optimalstandard definieren (vgl. Kapitel 2). Bei dieser Art von Vergleich steht beispielsweise die Frage im Mittelpunkt, welcher Anteil der Schülerinnen und Schüler in einem Land mindestens den Regelstandard im jeweiligen Kompetenzbereich erreicht hat.
3. Beim *ipsativen Vergleich* schließlich werden *Trends* als Veränderungen in den Ergebnissen über die Zeit beschrieben. In den IQB-Bildungstrends geht es dabei insbesondere um die Frage, inwieweit sich das von den Schülerinnen und Schülern erreichte Niveau der Kompetenzen in einem Land über die Zeit verändert hat. Dabei handelt es sich in den Studien des IQB nicht um individuelle Längsschnittanalysen, bei denen dieselben Schülerinnen und Schüler mehrfach getestet werden, sondern um Kohortenvergleiche. So hat das IQB sowohl im Jahr 2011 als auch im Jahr 2016 eine repräsentative Stichprobe von Schülerinnen und Schülern der 4. Jahrgangsstufe untersucht, sodass in der aktuellen Studie erstmalig Aussagen darüber getroffen werden können, ob es im Bereich der Primarstufe in den einzelnen Ländern über einen Zeitraum von fünf Jahren noch besser gelungen ist, die erfassten Kompetenzen zu fördern.

Die Erfahrungen der bisherigen Veröffentlichungen zeigen, dass der Fokus der medialen Berichterstattung über Schulleistungsstudien häufig allein auf den sozialen Vergleichen liegt, die in Form von Rangfolgen (*Rankings*) der Länder anhand der Kompetenzmittelwerte ihrer Schülerinnen und Schüler dargestellt werden. Aus Sicht des Bildungsmonitorings ist diese Art des Vergleichs jedoch die inhaltlich am wenigsten aufschlussreiche unter den drei genannten Vergleichsperspektiven und nicht selten auch problematisch. So werden häufig Unterschiede in Rangplätzen interpretiert, die statistisch nicht signifikant oder aufgrund ihrer geringen Größe praktisch nicht bedeutsam sind. Zudem werden bei einer ausschließlich sozialen Vergleichsperspektive Veränderungen nur dann sichtbar, wenn ein Land seinen Rangplatz gegenüber anderen Ländern verändert, andere Länder also gewissermaßen überholt hat oder überholt worden ist. Eine solche „Wettlaufperspektive“ ist für ein ländergemeinsames Bildungsmonitoring aber wenig aussagekräftig. Hier sollte es primär um die Zielsetzung gehen, das von den Schülerinnen und Schülern erreichte Kompetenzniveau insgesamt zu steigern und die Länderunterschiede auf hohem Niveau zu reduzieren.

Deutlich aufschlussreicher sind dagegen die kriterialen Vergleiche und die Trendaussagen, die daher auch im Fokus des vorliegenden Berichts stehen. So wird in Kapitel 5 dargestellt, wie sich die Viertklässlerinnen und Viertklässler im Jahr 2016 in den einzelnen Ländern auf die Kompetenzstufen verteilen. Dabei steht im Sinne des kriterialen Vergleichs die Frage im Vordergrund, welcher Anteil der Schülerinnen und Schüler mindestens die Regelstandards erreicht und welcher Anteil die Mindeststandards verfehlt hat. Zusätzliche Analysen nehmen darüber hinaus das Erreichen der Optimalstandards in den Blick. Anschließend wird die kriteriale Vergleichsperspektive mit der Trendanalyse verknüpft und geprüft, inwieweit sich die Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen in den einzelnen Ländern über die Zeit verändert haben. Dabei steht insbesondere im Blickpunkt, ob der Anteil der Kinder, deren Kompetenzen den Anforderungen der Mindeststandards nicht entsprechen, reduziert und der Anteil der Schülerinnen und Schüler, die mindestens die Regelstandards erreichen, erhöht werden konnte.

In Kapitel 6 wird zunächst unter der sozialen Vergleichsperspektive dargestellt, welches Kompetenzniveau die Schülerinnen und Schüler in den Ländern im Jahr 2016 im Durchschnitt entwickelt haben und wie groß die Streuung der erreichten Kompetenzen jeweils ausfällt. Anschließend wird auch hier der Fokus auf den Trend gerichtet und der Frage nachgegangen, inwieweit im Vergleich zum Jahr 2011 Veränderungen in den Mittelwerten und Streuungen zu verzeichnen sind.

Diese drei Vergleichsperspektiven kommen so weit wie möglich auch in den Analysen zu geschlechtsbezogenen, sozialen und zugewanderungsbezogenen Disparitäten im Bildungserfolg zur Anwendung (Kapitel 7–9), die ebenfalls einen wichtigen Bestandteil des nationalen Monitorings bilden. Dabei geht es um die Frage, inwieweit es in den Ländern gelingt, die mit diesen Hintergrundmerkmalen der Schülerinnen und Schüler verbundenen Unterschiede in den erreichten Kompetenzen zu verringern. Auch wenn nicht zu erwarten ist, dass Bildungssysteme ungünstige Eingangsvoraussetzungen vollständig kompensieren können, gilt es als wichtiges bildungspolitisches Ziel, die mit dem Geschlecht, der sozialen Herkunft und dem Zuwanderungshintergrund verbundenen Disparitäten so weit wie möglich zu reduzieren.

Ergänzt werden die in den Kapitel 5 bis 9 dargestellten Informationen zu den erreichten Kompetenzen der Viertklässlerinnen und Viertklässler um Zusatzanalysen zum Themenfeld „Umgang mit Heterogenität“, das die KMK als bildungspolitisch besonders relevant ausgewiesen hat. Insbesondere die Beschulung von Kindern und Jugendlichen mit sonderpädagogischem Förderbedarf stellt für die Bildungspolitik und die schulische Praxis in den Ländern aktuell eine zentrale Herausforderung dar. Vor diesem Hintergrund wird im vorliegenden Berichtsband das Potenzial einer bundesweit repräsentativen Datenerhebung mit großen Fallzahlen genutzt, um einen Beitrag zur Beschreibung der Bildungssituation von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf zu liefern (vgl. Kapitel 10). Darüber hinaus werden deskriptive Analysen zu besonders leistungsstarken Schülerinnen und Schülern (Kapitel 11) und zu verschiedenen Hintergrundmerkmalen und Einschätzungen von Lehrkräften (Kapitel 12) berichtet.

Mit den Ergebnissen des IQB-Bildungstrends 2016 erhalten die Länder Anhaltspunkte dafür, inwieweit die von ihnen umgesetzten Maßnahmen positive Entwicklungen ausgelöst haben und wo weiterer Handlungsbedarf besteht. Bei der Einordnung dieser Ergebnisse ist zu berücksichtigen, dass die in einem

Bildungssystem erreichten Erträge von vielfältigen Faktoren abhängen, zu denen neben bildungspolitischen Entscheidungen auch verschiedene, je nach Land teilweise unterschiedliche Rahmenbedingungen gehören, wie etwa die sozio-ökonomische Lage der Familien oder die Zusammensetzung der Schülerschaft. Veränderungen solcher Rahmenbedingungen können die Weiterentwicklung von Unterrichtsqualität befördern oder erschweren, sind aber durch bildungspolitische Maßnahmen allein nicht direkt zu beeinflussen. Bei sich ändernden Rahmenbedingungen kann es daher unter Umständen auch als Erfolg zu bewerten sein, wenn schulische Erträge stabil bleiben. In jedem Fall besteht eine wichtige bildungspolitische Aufgabe darin, auf sich ändernde Rahmenbedingungen rechtzeitig und adäquat zu reagieren.

Literatur

- Böhme, K., Richter, D., Stanat, P., Pant, H. A. & Köller, O. (2012). Die länderübergreifenden Bildungsstandards in Deutschland. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 11–18). Münster: Waxmann.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards – Expertise*. Berlin: BMBF (Bundesministerium für Bildung und Forschung).
- Klieme, E. & Tippelt, R. (2008). Qualitätssicherung im Bildungswesen. In E. Klieme & R. Tippelt (Hrsg.), *Qualitätssicherung im Bildungswesen. Eine aktuelle Zwischenbilanz*. 53. Beiheft der Zeitschrift für Pädagogik (S. 7–13). Weinheim: Beltz.
- KMK (1997) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1997). *Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland – Konstanzer Beschluss. Beschluss der Kultusministerkonferenz vom 24.10.1997*. Zugriff am 15.08.2016 unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf
- KMK (2006) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Beschluss der Kultusministerkonferenz vom 02.06.2006*. Zugriff am 15.08.2016 unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2006/2006_08_01-Gesamtstrategie-Bildungsmonitoring.pdf
- KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Beschluss der 350. Kultusministerkonferenz vom 11.06.2015*. Zugriff am 15.08.2016 unter https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf
- Köller, O. (2010). Bildungsstandards. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (3. Auflage, S. 529–548). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O., Knigge, M. & Tesch, B. (Hrsg.). (2010). *Sprachliche Kompetenzen im Ländervergleich*. Münster: Waxmann.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.). (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Stanat, P., Pant, H. A., Richter, D., Pöhlmann, C. & Kuhl, P. (2013). Was kann das IQB leisten? In S. Lin-Klitzing, D. DiFuccia & G. Müller-Frerich (Hrsg.), *Zur Vermessung von Schule: Empirische Bildungsforschung und Schulpraxis* (S. 125–152). Bad Heilbrunn: Klinkhardt.

1.2 Beschreibung der im Fach Deutsch untersuchten Kompetenzen

Katrin Böhme, Maria Engelbert und Sebastian Weirich

Im IQB-Bildungstrend 2016 werden fachliche Kompetenzen von Schülerinnen und Schülern am Ende der 4. Jahrgangsstufe untersucht. Die Basis dieser Untersuchung bilden die Bildungsstandards der Kultusministerkonferenz (KMK) für den Primarbereich (KMK, 2005). Die Vermittlung und Förderung sprachlicher Kompetenzen ist ein zentraler Bestandteil des Bildungsauftrags in diesem Bereich. Ferner dient Sprache im schulischen Alltag in allen Fächern als Medium des Lernens und stellt damit eine wesentliche Voraussetzung für den schulischen Erfolg und die gesellschaftliche Teilhabe der Schülerinnen und Schüler dar. Entsprechend ist davon auszugehen, dass der sicheren Beherrschung der Instruktions- und Verkehrssprache Deutsch eine „grundlegende Bedeutung für die kognitive, emotionale und soziale Entwicklung der Kinder“ zukommt (vgl. KMK, 2005, S. 6f).

Die Bildungsstandards im Fach Deutsch für den Primarbereich beschreiben sowohl produktive als auch rezeptive sprachliche Kompetenzen im mündlichen sowie schriftsprachlichen Bereich. Die gesamte Breite der in den Bildungsstandards thematisierten Kompetenzen lässt sich im Rahmen von standardisierten Schulleistungsstudien nicht abdecken. Einige der für die weitere Bildungslaufbahn der Schülerinnen und Schüler zentralen sprachlichen Kompetenzen konnten im IQB-Bildungstrend 2016 jedoch überprüft werden.

Im vorliegenden Kapitel werden die im IQB-Bildungstrend 2016 im Fach Deutsch untersuchten Kompetenzen und ihre Operationalisierung genauer beschrieben. Hierbei werden die drei relevanten Kompetenzbereiche beziehungsweise Teilbereiche *Lesen*, *Zuhören* und *Orthografie* nacheinander in den Blick genommen. Aussagen über die erreichten Kompetenzen der Schülerinnen und Schüler sind nur für diese drei Bereiche möglich; eine Verallgemeinerung der erzielten Ergebnisse auf andere Kompetenzbereiche oder auf eine globale „Deutschkompetenz“ ist nicht sinnvoll. Daher erfolgt auch die Darstellung der empirischen Befunde im Rahmen dieses Berichts für die getesteten Kompetenzbereiche stets separat.

1.2.1 Die Bildungsstandards im Fach Deutsch

Im Jahr 2004 hat die Kultusministerkonferenz verbindliche Bildungsstandards im Fach Deutsch für den Primarbereich verabschiedet (KMK, 2005). Diese Vorgaben gelten in allen Ländern in der Bundesrepublik Deutschland und legen fest, welche Kompetenzziele Schülerinnen und Schüler bis zum Ende der 4. Jahrgangsstufe, die in den meisten Ländern das Ende der Grundschulzeit markiert,¹ erreicht haben sollen. In jedem Land wurden Kernlehrpläne oder Curricula erarbeitet, in denen die Bildungsstandards konkretisiert und als verbindliche Lerngegenstände fixiert wurden.

¹ In den Ländern Berlin und Brandenburg erfolgt der Übergang von der Primarstufe in die Sekundarstufe I in der Regel erst nach der 6. Jahrgangsstufe.

Die Bildungsstandards für das Fach Deutsch im Primarbereich gehen von einer vorschulisch erworbenen Sprachhandlungskompetenz aus, die durch schulische Lerngelegenheiten systematisch erweitert werden soll. Die angestrebten Kompetenzen werden in den Bildungsstandards als Könnensbeschreibungen formuliert (*Can-do-Statements*). Explizite Definitionen der jeweiligen Kompetenzkonstrukte beinhalten die Bildungsstandards jedoch nicht (vgl. Granzer, 2009). Um den Grad des Erreichens der Bildungsstandards einer Überprüfung zugänglich machen zu können, war es daher vor der Operationalisierung der einzelnen Standards erforderlich, Definitionen aller relevanten Kompetenzkonstrukte zu erarbeiten. Dies erfolgte durch das IQB in enger Kooperation mit Expertinnen und Experten aus der Deutschdidaktik.

In der fachdidaktischen und sprachwissenschaftlichen Literatur werden sprachliche Kompetenzen häufig in verschiedene Teilbereiche unterteilt, zu meist in die Bereiche Lesen, Hören, Schreiben und Sprechen (vgl. Felder, 2003). Innerhalb dieser Bereiche können zum einen anhand der Sprachmodalität mündliche Kompetenzen (Sprechen und Hören beziehungsweise Zuhören) von schriftsprachlichen Kompetenzen (Schreiben und Lesen) unterschieden werden. Zum anderen lassen sich anhand der beteiligten sprachlichen Prozesse produktive Kompetenzen (Sprechen und Schreiben) von rezeptiven Kompetenzen (Hören und Lesen) differenzieren.

In den Bildungsstandards für den Primarbereich im Fach Deutsch sind die folgenden vier Kompetenzbereiche ausgewiesen, die diese Struktur sprachlicher Kompetenzen widerspiegeln (KMK, 2005, S. 7ff.):

- *Sprechen und Zuhören*: Gespräche führen, zu anderen sprechen, verstehend zuhören, szenisch spielen und über Lernen sprechen
- *Schreiben*: über Schreibfertigkeiten verfügen, richtig schreiben, Texte planen, schreiben und überarbeiten
- *Lesen – mit Texten und Medien umgehen*: über Lesefähigkeiten und Leseerfahrungen verfügen, Texte erschließen und präsentieren
- *Sprache und Sprachgebrauch untersuchen*: grundlegende sprachliche Strukturen und Begriffe kennen und verwenden, sprachliche Verständigung untersuchen, an Wörtern, Sätzen und Texten arbeiten und Gemeinsamkeiten und Unterschiede von Sprachen entdecken

Allen vier zentralen Kompetenzbereichen, die in Abbildung 1.2 noch einmal grafisch veranschaulicht sind, ist gemeinsam, dass sie als Stützkomponente Methoden und Arbeitstechniken beinhalten. Der Kompetenzbereich *Sprache und Sprachgebrauch untersuchen* nimmt zudem eine Sonderrolle ein. Im Rahmen eines integrativen Ansatzes ist dieser durchgängig mit den anderen Bereichen zu verbinden, da er Kompetenzen umfasst, die grundlegend für den Erwerb und den Ausbau anderer, sowohl rezeptiver (Zuhören, Lesen) als auch produktiver (Sprechen, Schreiben) sprachlicher Kompetenzen sind. Weiterhin zeigt Abbildung 1.2, dass *Sprechen und Zuhören* im Fach Deutsch einen gemeinsamen Kompetenzbereich bilden und *Orthografie* unter *Schreiben* subsumiert wird. Im Zuge fachdidaktischer Diskussionen und aufgrund von Befunden aus empirischen Studien des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) wurde jedoch sowohl für den Primarbereich als auch für die Sekundarstufe I deutlich, dass es sinnvoll ist, *Orthografie* in der Diagnostik als eigenen Teilbereich zu behandeln und separat zu testen. Diese Entscheidung wurde auch aus messtheoretischen Erwägungen getroffen, um die hier relevanten Teilkompetenzen durch den Einsatz spezifischer Aufgabenformate so präzise und passgenau wie mög-

lich erfassen zu können und eine Konfundierung mit anderen Fähigkeitsaspekten zu vermeiden. Ebenso wird *Zuhören* als separater Teilbereich von *Sprechen und Zuhören* getestet. Der Kompetenzbereich *Lesen – mit Texten und Medien umgehen* wird nachfolgend verkürzt als Kompetenzbereich *Lesen* bezeichnet.

Abbildung 1.2: Kompetenzbereiche in den Bildungsstandards der KMK für den Primarbereich im Fach Deutsch



Quelle: Abbildung in Anlehnung an KMK (2005, S. 7)

Für die drei im IQB-Bildungstrend 2016 untersuchten sprachlichen Kompetenzen im Fach Deutsch wurden basierend auf den Bildungsstandards der KMK Konstruktdefinitionen erarbeitet und Operationalisierungen abgeleitet. Beide Aspekte werden nachfolgend für die Kompetenzbereiche beziehungsweise Teilbereiche *Lesen*, *Zuhören* und *Orthografie* dargestellt.

1.2.2 Konstruktdefinitionen der Kompetenzbereiche im Fach Deutsch

Konstruktdefinition des Kompetenzbereichs *Lesen*

Die theoretischen Grundlagen für den Kompetenzbereich *Lesen* sind im Rahmen großer Schulleistungsstudien, wie etwa PISA², umfassend aufgearbeitet worden (vgl. z. B. Artelt, Stanat, Schneider & Schiefele, 2001; Kintsch, 1998). In diesen Studien wird Lesekompetenz als eine Schlüsselqualifikation verstanden und zumeist als die Fähigkeit definiert, kontinuierliche, diskontinuierliche und multimediale Texte zu verstehen.

Lesekompetenz ist nicht nur von zentraler Bedeutung für den Wissenserwerb und das lebenslange Lernen, sondern auch für die Teilnahme an der

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

Kommunikation über gesellschaftlich relevante Themen und für die Bewältigung von alltäglichen, beispielsweise beruflichen, Kommunikations- und Handlungsanforderungen (vgl. z. B. Bos et al., 2007; Hurrelmann, 2007).

Psychologische Theorien des Textverstehens beruhen auf der zentralen Annahme, dass der Verstehensprozess als Konstruktionsleistung zu fassen ist. Demnach ist Lesen keine passive Rezeption dessen, was im jeweiligen Text an Information enthalten ist, sondern aktive (Re-)Konstruktion der Textbedeutung. Die im Text enthaltenen Aussagen werden aktiv mit dem Vor-, Welt- und Sprachwissen des Lesers verbunden (Artelt et al., 2001, S. 70 f.). In großen Schulleistungsstudien wie PISA oder PIRLS/IGLU³ wird zumeist ein pragmatisch-funktionaler Begriff der *Reading Literacy* zugrunde gelegt und in der testdiagnostischen Überprüfung der Lesekompetenz die kognitive Dimension, also das Lese- oder Textverstehen, fokussiert.

Vor allem im Rahmen von 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 im Zentrum, sondern es wird auch der Beitrag des Lesens zur Persönlichkeitsbildung thematisiert (Böhme, 2012; Spinner, 2008).

Die Bildungsstandards der Kultusministerkonferenz umfassen für den Kompetenzbereich *Lesen* eine breite Palette von Teilkompetenzen. Diese beziehen sich nicht nur auf das Verstehen von Texten, sondern thematisieren auch explizit die Förderung von Leseinteresse und Lesefreude sowie die Ausbildung von Empathie und Fremdverstehen (KMK, 2005). Bezogen auf das Verstehen sollen die Schülerinnen und Schüler in die Lage versetzt werden, in Texten gezielt einzelne Informationen zu suchen und zentrale Aussagen eines Textes zu erfassen, diese wiederzugeben und mit Textstellen zu belegen. Ferner wird erwartet, dass die Schülerinnen und Schüler verschiedene Sorten von Sach- und Gebrauchstexten einerseits sowie Erzähltexten, lyrischen und szenischen Texten andererseits kennen und diese auch unterscheiden können. Darüber hinaus sollen die Kinder Texte begründet auswählen können, zu Texten eigene Gedanken entwickeln, zu ihnen Stellung nehmen und mit anderen über Texte sprechen können. Wesentlich ist auch der handelnde und gestaltende Umgang mit Texten (KMK, 2005).

Konstruktdefinition des Teilbereichs *Zuhören*

Auch der Kompetenzbereich *Sprechen und Zuhören* nimmt im schulischen Kontext einen zentralen Stellenwert ein, da er für die unterrichtliche Kommunikation und damit den Transport von Wissen in allen Fächern von entscheidender Bedeutung ist. Belgrad und Kollegen (2008) konstatieren, dass „Schüler(innen) bis zu zwei Drittel der Unterrichtszeit zuhören müssen“ (Belgrad, Eriksson, Pabst-Weinschenk & Vogt, 2008, S. 20). Ein Großteil des Lerninputs erfolgt also als mündliche Informationsvermittlung, weshalb Schülerinnen und Schüler in die Lage versetzt werden müssen, ihr Zuhörverhalten selbstständig und kompetent zu steuern (vgl. Böhme, 2012). Hierbei ist relevant, dass sich schulische Anforderungen an mündlichen Sprachgebrauch von den Anforderungen der Alltagskommunikation unterscheiden (vgl. Heppt, 2016). Das im institutionalisierten

3 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*. Das Akronym IGLU steht für *Internationale Grundschul-Lese-Untersuchung*.

sierten Kontext verwendete schulspezifische Sprachregister orientiert sich viel stärker als alltäglicher Sprachgebrauch an der schriftsprachlichen Norm und „unterscheidet sich auf lexikalischer, syntaktischer und pragmatischer Ebene deutlich von der Alltagssprache“ (Nauwerck, 2009, S. 260).

In den Bildungsstandards der Kultusministerkonferenz im Fach Deutsch für den Primarbereich (KMK, 2005) umfasst der Kompetenzbereich *Sprechen und Zuhören* sowohl die produktive als auch die rezeptive mündliche Sprachkompetenz der Schülerinnen und Schüler. Dies folgt der Tradition des Deutschunterrichts, in dem diese beiden Teilbereiche unter Gesprächskompetenz subsumiert werden (vgl. Krelle & Neumann, 2014). Die in den Bildungsstandards im Fach Deutsch genannten Bereiche mündlicher Sprache lassen sich wie folgt gruppieren (vgl. KMK, 2005, S. 9f.):

- sprechen (Gespräche führen, zu anderen sprechen, über Lernen sprechen),
- zuhören (verstehend zuhören) und
- szenisch spielen.

Aufgrund des interaktiven Charakters mündlicher Sprachverwendung stellt die hierbei vorgenommene Trennung eine starke Vereinfachung der realen Kommunikationssituation dar, da beispielsweise die in den Standards thematisierte Gesprächsführung stets sowohl Aspekte des Sprechens als auch des Zuhörens vereint. Die Standards legen den Fokus der Kompetenzbeschreibungen auf die in Schule und Alltag relevante Kommunikationsfunktion mündlicher Sprache. Es soll den Schülerinnen und Schülern ermöglicht werden, Gespräche unter Beachtung gemeinsam entwickelter Gesprächsregeln zu führen. Somit geht es um den für eine adäquate soziale Interaktion relevanten Kommunikationsaspekt. Die Abbildung dieser komplexen Kommunikationssituation in großen, standardisierten Schulleitungsstudien ist derzeit allerdings nicht ohne Weiteres möglich. Im Rahmen der IQB-Bildungstrends wird der Fokus der Kompetenzüberprüfung daher auf ein Element der mündlichen Kommunikation gelegt: auf den Teilbereich *Zuhören*. In den allgemeinen Erläuterungen der einzelnen Kompetenzbereiche heißt es zur Zuhörkompetenz der Schülerinnen und Schüler konkret: „Sie [...] hören aufmerksam und genau zu, nehmen die Äußerungen anderer auf und setzen sich mit diesen konstruktiv auseinander“ (KMK, 2005, S. 8). Die formulierten Standards beziehen sich auf das verstehende Zuhören, und zwar konkret darauf, Inhalte zuhörend zu verstehen, gezielt nachzufragen sowie Verstehen und Nichtverstehen zum Ausdruck zu bringen (KMK, 2005).

Konstruktdefinition des Teilbereichs *Orthografie*

Auch Rechtschreibfertigkeiten haben eine wichtige pädagogische und gesellschaftliche Bedeutung (Scheele, 2006). Die deutschsprachige und internationale Fachliteratur hierzu bezieht sich vorrangig auf die Entwicklung der orthografischen Kompetenz im Vor- und Grundschulalter (vgl. z. B. Thomé, 2003). Den Erwerb der orthografischen Kompetenz konzipiert die deutschsprachige Orthografiedidaktik seit geraumer Zeit als Stufenmodell.⁴ Hierbei handelt es sich allerdings nicht um klar abgrenzbare, strikt aufeinanderfolgende Entwicklungsstufen, sondern um Zugriffsweisen beziehungsweise Strategien, die sich zumeist nach und nach entfalten, zum Teil aber auch parallel verfügbar sind

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

(May, 2002). Für die Grundschule ist die Unterscheidung einer alphabetischen von einer orthografischen Stufe zentral⁵ (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* usw. auseinander (vgl. May, 2002, S. 30). Sie orientieren sich mehr und mehr am Stamprinzip (*Wand* wegen *Wände*). Auch die Großschreibung von Substantiven, insbesondere von Konkreta, gelingt in zunehmendem Maße, ebenso die Abgrenzung von Sätzen, unter anderem im Zusammenhang mit der direkten Rede. Hier kommt syntaktisches, auf den Satzbau bezogenes Wissen ins Spiel.

In den Bildungsstandards der KMK ist unter der Bezeichnung *richtig schreiben* die orthografische Kompetenz der Schülerinnen und Schüler neben der Fähigkeit zur Textproduktion Bestandteil des Kompetenzbereichs *Schreiben*. Die Zeichensetzung und die Groß- und Kleinschreibung hingegen sind im Kompetenzbereich *Sprache und Sprachgebrauch untersuchen* verortet. Die orthografische Kompetenz der Schülerinnen und Schüler zum Ende der 4. 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, 2005, S. 8)

1.2.3 Operationalisierung der Bildungsstandards im Fach Deutsch

Allgemeine Hinweise zur Operationalisierung

In den Kompetenztests des IQB für das Fach Deutsch werden Aufgaben eingesetzt, die in aller Regel aus einem Stimulus, also beispielsweise einem Lese- oder Hörtext, und einer Reihe von einzelnen Aufgabenstellungen, sogenannten Items, bestehen. Die Itemtypen sind über die Kompetenzbereiche beziehungsweise Teilbereiche *Lesen*, *Zuhören* und *Orthografie* hinweg sehr ähnlich und werden daher an dieser Stelle für alle drei Kompetenzen beschrieben. Für die Überprüfung der orthografischen Kompetenz werden darüber hinaus weite-

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

re, spezifische Itemtypen eingesetzt, die im Abschnitt zur Operationalisierung im Bereich *Orthografie* vorgestellt werden.

Die Formate der Items, die in den Kompetenztests des IQB zum Einsatz kommen, lassen sich nach dem Grad des von den Schülerinnen und Schülern eigenständig zu leistenden produktiven Gehalts in geschlossene, halboffene und offene Formate einteilen. Geschlossene Items bieten Antwortmöglichkeiten an, die angekreuzt, unterstrichen oder geordnet werden müssen. Bei halboffenen und offenen Antwortformaten werden den Schülerinnen und Schülern keine Antwortoptionen angeboten. Vielmehr verlangen die entsprechenden Items eine eigenständig formulierte Antwort, die aus nur einem Wort, einer Wortgruppe oder mehreren Sätzen bestehen kann. Die Übergänge zwischen halboffenen und offenen Formaten sind fließend. Im IQB-Bildungstrend 2016 wurden die folgenden Itemformate eingesetzt (vgl. IQB, 2013a, 2013b, 2013c):

- **Multiple-Choice-Items (MC-Items):** Bei diesen Items werden zu einer Aussage oder einer Frage (Itemstimulus) vier Antwortmöglichkeiten vorgegeben, von denen genau eine zutreffend ist (Attraktor); diese soll angekreuzt werden. Distraktoren können sich beispielsweise auf typische Fehlvorstellungen von Schülerinnen und Schülern beziehen.
- **Richtig-Falsch-Items:** Bei diesen Items werden mehrere Aussagen vorgegeben, für die jeweils entschieden werden muss, ob sie korrekt sind. Diese Entscheidungen werden zeilenweise durch Ankreuzen („richtig“ oder „falsch“) getroffen. Die Schwierigkeit dieses Itemtyps kann unter anderem dadurch variiert werden, dass das Item nur dann insgesamt als richtig bearbeitet bewertet wird, wenn eine Mindestzahl an korrekten Entscheidungen getroffen und keine unzutreffenden Antworten angekreuzt wurden.
- **Zuordnungsitems:** Bei dieser Art von Items müssen verschiedene Elemente aus zwei Gruppen – beispielsweise Problem und Lösung oder Frage und Antwort – einander zugeordnet werden. Aufgrund der Abhängigkeiten, die zwischen den einzelnen Entscheidungen bestehen, wird nur die richtige Zuordnung *aller* Elemente als korrekte Antwort bewertet. Die Schwierigkeit solcher Items kann somit durch die Anzahl der zuzuordnenden Elemente variiert werden.
- **Kurzantworten:** Bei diesen Items mit halboffenem Format müssen Schülerinnen und Schüler einzelne Informationen, wie etwa Zahlen, Daten, Eigennamen oder einzelne Wörter, in ein kurzes Freitextfeld eintragen.
- **Offene Items:** Bei dieser Art von Items müssen die Schülerinnen und Schüler ihre Antwort eigenständig formulieren. Solche Items sind im Allgemeinen schwieriger als Auswahlitems, weil hier keine Antwortoptionen vorliegen, die Hinweise auf die richtige Antwort geben könnten.

Für die Kompetenzmessung werden die einzelnen Aufgaben mit ihren Items in sogenannten Aufgabenblöcken gruppiert (vgl. Kapitel 3.1). Je nach Art und zeitlichem Umfang für die Aufgabenbearbeitung enthalten die Aufgabenblöcke im Fach Deutsch zwischen einer Aufgabe mit 7 Items (*Lesen*) und 14 Aufgaben mit 36 Items (*Orthografie*). Alle Aufgaben und Items eines Aufgabenblocks beziehen sich immer auf denselben Kompetenzbereich.

Operationalisierung im Kompetenzbereich *Lesen*

Nicht alle von der KMK zum Kompetenzbereich *Lesen* formulierten Standards lassen sich so operationalisieren und in Testaufgaben überführen, dass sie im Rahmen von großangelegten, standardisierten Leistungsmessungen erfassbar sind. So lassen sich beispielsweise die Standards „lebendige Vorstellungen beim Lesen und Hören literarischer Texte entwickeln“ (KMK, 2005, S. 11) oder „bei Lesungen und Aufführungen mitwirken“ (KMK, 2005, S. 13) kaum mit stift- und papierbasierten Tests im Klassenverband erfassen. Unter Berücksichtigung dieser Beschränkungen liegt der Schwerpunkt der Testung im Kompetenzbereich *Lesen* auf der Komponente *Texte erschließen*. Hinzu kommt der als *über Lesefähigkeiten verfügen* bezeichnete, auf die Konstruktion eines Situationsmodells zielende zentrale Standard *altersgemäße Texte sinnverstehend lesen*.

Für die Erhebung der Lesekompetenz werden in den Kompetenztests des IQB Aufgaben eingesetzt, die aus einem Lesetext und einer Reihe von einzelnen Aufgabenstellungen (Items) bestehen. Bei der Aufgabenentwicklung wurde darauf geachtet, dass die den Aufgaben zugrunde liegenden Stimuli zu ungefähr gleichen Anteilen sowohl literarische als auch nichtliterarische Texte umfassen. Ferner kamen neben mehrheitlich kontinuierlichen Texten auch zwei diskontinuierliche informierende Texte zum Einsatz.

Im IQB-Bildungstrend 2016 wurden für die Kompetenzmessung im Teilbereich *Lesen* 18 verschiedene Aufgabenblöcke eingesetzt, die insgesamt 21 Aufgaben mit insgesamt 155 Items enthielten. Diese Items haben zu etwa 55 Prozent ein geschlossenes Format und zu knapp 30 Prozent ein offenes Format. Die verbleibenden ca. 15 Prozent waren halboffene Items wie beispielsweise Kurzantworten. Für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf wurden angepasste Testaufgaben konstruiert. Der Prozess der Entwicklung dieser Aufgaben und die Art der dabei vorgenommenen Anpassungen werden in Kapitel 13.2 dargestellt.

Operationalisierung im Teilbereich *Zuhören*

Die im Kompetenzbereich *Sprechen und Zuhören* in großen Schulleistungsstudien überprüfbareren Kompetenzen des Teilbereichs *Zuhören* beziehen sich überwiegend auf das verstehende Zuhören, bei dem Gesprächsbeiträge anderer verfolgt und rezipiert, wesentliche Informationen verstanden, gesichert und wiedergegeben sowie auch nonverbale Äußerungen (zum Beispiel Stimmführung) verstanden werden sollen (KMK, 2005).

Für die Erhebung der Zuhörkompetenz werden in den Kompetenztests des IQB Aufgaben eingesetzt, die aus einem Hörstimulus und einer Reihe von einzelnen Aufgabenstellungen (Items) bestehen. Die Hörtexte, die als Aufgabenstimuli zum Einsatz kommen, werden in der Testsituation von CD abgespielt. Zuhörkompetenz kann sich sowohl auf gesprochene Sprache, die eher konzeptionell schriftlich ist (z. B. Hörbuch), beziehen als auch auf eher konzeptionell mündliche gesprochene Sprache (z. B. Podcast). Konzeptionell mündliche Sprache weist typischerweise ein höheres Maß an Redundanz auf und ist mit Interjektionen (*ach, aha, nanu* etc.) und Verzögerungsausdrücken (*äh, ähm*) durchsetzt (vgl. Buck, 2001). Für jede Art von Zuhörsituation sind auch paraverbale Merkmale (Tempo, Prosodie, emotionale Einfärbung) sowie Hintergrund-

geräusche relevant. In Hörspielen, Radiobeiträgen etc. wird zudem Hintergrundmusik eingesetzt.

Als Stimuli werden in den Kompetenztests des IQB sowohl fiktional-literarische als auch informierend-faktische Hörtexte, wie etwa Podcasts, Radiobeiträge oder Ausschnitte aus Hörspielen, verwendet. Die Stimuli decken somit einerseits eher konzeptionell mündliche, andererseits aber auch eher konzeptionell schriftliche Sprache ab. Sie werden durch eine mündliche und/oder schriftliche Instruktion gerahmt, um beispielsweise auf bestimmte Aspekte hinzuweisen, die für die Lösung der Aufgabe wichtig sind, oder um Begriffe, die den Schülerinnen und Schülern mit größerer Wahrscheinlichkeit unbekannt sind, zu erläutern. Dann wird der Stimulus den Schülerinnen und Schülern einmal vorgespielt und im Anschluss beantworten die Kinder die auf den Stimulus bezogenen Items.

Im IQB-Bildungstrend 2016 wurden für die Kompetenzmessung im Teilbereich *Zuhören* 12 Aufgabenblöcke eingesetzt, die 16 Zuhöraufgaben mit insgesamt 115 Items enthielten. Knapp 70 Prozent der Items sind geschlossen, etwa ein Fünftel offen und knapp 10 Prozent haben ein halboffenes Format. Auch im Teilbereich *Zuhören* wurden für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf angepasste Testaufgaben konstruiert (vgl. Kapitel 13.2).

Operationalisierung im Teilbereich *Orthografie*

Zur Überprüfung der Rechtschreibkompetenz der Schülerinnen und Schüler wurden im IQB-Bildungstrend 2016 vier verschiedene Aufgabentypen eingesetzt, die es erlauben, die Mehrheit der Bildungsstandards aus dem Bereich *richtig schreiben* in die Überprüfung einzubeziehen. Der Standard *rechtschreibwichtige Wörter normgerecht schreiben* wird überwiegend durch Lückensatzdiktate operationalisiert. Diese sind gut geeignet, um die richtige Schreibung von Wörtern im Kontext und die Anwendung orthografischer Regeln zu erheben (vgl. z. B. Böhme & Bremerich-Vos, 2009; Herné & Naumann, 2002). Bei den Lückensatzdiktaten handelt es sich um einzelne Sätze, in denen jeweils ein bis zwei Wörter fehlen. Die Sätze wurden den Kindern vollständig vorgelesen und sie wurden gebeten, die fehlenden Wörter zu ergänzen. Dadurch liegt der Fokus der Fähigkeitsüberprüfung auf Lückenwörtern, mit denen gezielt orthografische Schwerpunkte erfasst werden können. Dieser Aufgabentyp hat verschiedene Vorteile (vgl. Böhme & Bremerich-Vos, 2009). Bei dieser Variante von Diktaten ist der Schreibaufwand für die Kinder begrenzt, sodass Unterschiede im Schreibtempo weniger stark ins Gewicht fallen als bei Textdiktaten. Das Schreiben von sicher beherrschten Wörtern wie Artikeln und Pronomen entfällt. Damit ist 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. Zur Auswahl der Lückenwörter für den Test wurden Grundwortschätze herangezogen, beispielsweise der von Naumann (1999) zusammengestellte Orientierungswortschatz, sowie Wortbestände aus bewährten standardisierten Tests.

Wie im IQB-Ländervergleich 2011 wurde auch im IQB-Bildungstrend 2016 für die Lückensatzdiktate auf eine qualitative Fehleranalyse, also die Betrachtung von Einzelfehlern pro Wort, verzichtet. In den Vergleichsarbeiten (VERA), bei denen ebenfalls vom IQB entwickelte Testaufgaben auf Basis der

Bildungsstandards zum Einsatz kommen, werden entsprechende Auswertungen jedoch regelmäßig angeregt und erläutert (Bremerich-Vos et al., 2010).

Im IQB-Bildungstrend 2016 kamen 109 Items des Aufgabentyps Lückensatzdiktat zum Einsatz, die in etwa die Hälfte aller Items im Bereich *Orthografie* darstellen. Etwa ein Drittel der Items aus Lückensatzdiktaten wurde für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf so angepasst, dass diese ausschließlich sehr einfache Lückenwörter enthalten, die in der Regel in der Schuleingangsphase (Jahrgangsstufen 1 bis 3) erworben werden.

Die Fähigkeit, *über Fehlersensibilität und Rechtschreibgespür zu verfügen*, wurde mit Korrekturaufgaben erfasst. Die 71 Items dieses Aufgabentyps haben ein halboffenes Format. Die Schülerinnen und Schüler werden dabei aufgefordert, in kurzen Texten fehlerhafte Wörter zu markieren und die korrekte Schreibung zu notieren. Im Test liegt der Anteil fehlerhafter Wörter pro Text bei rund 15 bis 50 Prozent, wobei innerhalb eines fehlerhaften Wortes ein oder zwei verschiedene Einzelfehler enthalten sind. Die Fehler beziehen sich ausschließlich auf die Schreibung der Wörter (Groß- und Kleinschreibung, Dehnung, Schärfung etc.), nicht auf die Zeichensetzung. Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf haben diesen Aufgabentyp nicht bearbeitet.

Eine weitere Gruppe von Aufgaben überprüft die Anwendung von *Strategiewissen*. Bei diesen 12 Items im halboffenen und offenen Format sollen die Schülerinnen und Schüler durch Bildung von Wortformen bzw. Ableitungen Schreibungen finden, aus denen die Richtigschreibung hervorgeht. In einigen Fällen sollen die Kinder die zugrunde liegende Regel aufschreiben. Auch dieser Aufgabentyp wurde Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf nicht vorgelegt.

Die übrigen Items im Bereich *Orthografie* fordern die Schülerinnen und Schüler dazu auf, *Texte auf orthografische Richtigkeit zu überprüfen*, also in Texten Falschschreibungen zu erkennen und zu markieren. Ein Teil dieser 12 Items im geschlossenen Format wurde auch Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf vorgelegt.

Literatur

- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich.
- 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.
- Böhme, K. (2012). *Methodische und didaktische Überlegungen sowie empirische Befunde zur Erfassung sprachlicher Kompetenzen im Deutschen. Analysen zu den Bildungsstandards im Fach Deutsch für den Primarbereich*. Humboldt Universität zu Berlin, Philosophische Fakultät IV, 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.
- 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.

- Bremerich-Vos, A. Behrens, U., Böhme, K., Engelbert, M., Linkert, D. & Krelle, M. (2010). *Vergleichsarbeiten 2010. 3. Jahrgangsstufe (VERA-3). Deutsch – DIDAKTISCHE HANDREICHUNG zu Testheft II – Rechtschreibung*. Berlin: IQB.
- Buck, G. (2001). *Assessing Listening*. New York: Cambridge University Press.
- Felder, E. (2003). Sprache als Medium und Gegenstand des Unterrichts. In U. Bredel, H. Günther, P. Klotz, J. Ossner & G. Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache. Band 1* (S. 42–51). Paderborn: Ferdinand Schöningh.
- 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.
- Heppt, B. (2016). *Verständnis von Bildungssprache bei Kindern mit deutscher und nicht-deutscher Familiensprache* (Doktorarbeit). Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät. Zugriff am 08.08.2017 unter <https://edoc.hu-berlin.de/handle/18452/18186>
- Herné, K.-L. & Naumann, C. L. (2002). *Aachener Förderdiagnostische Rechtschreibfehler-Analyse (AFRA). Systematische Einführung in die Praxis der Fehleranalyse mit Auswertungshilfen zu insgesamt 33 standardisierten Testverfahren als Kopiervorlagen*. 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.
- IQB (2013a) = Institut zur Qualitätsentwicklung im Bildungswesen (Hrsg.). (2013). Kompetenzstufenmodell zu den Bildungsstandards für das Fach Deutsch im Kompetenzbereich *Lesen – mit Texten und Medien umgehen*. Primarbereich. Zugriff am 08.08.2017 unter https://www.iqb.hu-berlin.de/bista/ksm/KSM_GS_Deutsch_L_1.pdf.
- IQB (2013b) = Institut zur Qualitätsentwicklung im Bildungswesen (Hrsg.). (2013). Kompetenzstufenmodell zu den Bildungsstandards für das Fach Deutsch im Kompetenzbereich *Sprechen und Zuhören*, Teilbereich *Zuhören*. Primarbereich. Zugriff am 08.08.2017 unter https://www.iqb.hu-berlin.de/bista/ksm/KSM_GS_Deutsch_Z_2.pdf.
- IQB (2013c) = Institut zur Qualitätsentwicklung im Bildungswesen (Hrsg.). (2013). Kompetenzstufenmodell zu den Bildungsstandards für das Fach Deutsch im Kompetenzbereich *Schreiben*, Teilbereich *Rechtschreibung*. Primarbereich. Zugriff am 08.08.2017 unter https://www.iqb.hu-berlin.de/bista/ksm/KSM_GS_Deutsch_R_2.pdf.
- Kintsch, W. (1998). *Comprehension*. Cambridge: Cambridge University Press.
- KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- Krelle, M. & Neumann, D. (2014). Sprechen und Zuhören. In U. Behrens, A. Bremerich-Vos, M. Krelle, K. Böhme & S. Hunger (Hrsg.), *Bildungsstandards Deutsch: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen* (S. 14–45). Berlin: Cornelsen.
- May, P. (2002). *HSP 1–9. Diagnose orthographischer Kompetenz. Zur Erfassung der grundlegenden Rechtschreibstrategien mit der Hamburger Schreibprobe. Manual*. Hamburg: Verlag für Pädagogische Medien.
- Naumann, C. L. (1999). *Orientierungswortschatz. Die wichtigsten Wörter und Regeln für die Rechtschreibung Klassen 1–6*. Weinheim: Beltz.
- Nauwerck, P. (2009). Sprachstandsdiagnose und Sprachförderung im Übergang vom Kindergarten in die Grundschule: Förderbedarf erkennen und (kommunikative) Kompetenzen entwickeln. In M. Krelle & C. Spiegel (Hrsg.), *Sprechen und Kommunizieren. Entwicklungsperspektiven, Diagnosemöglichkeiten und Lernszenarien in Deutschunterricht und Deutschdidaktik* (S. 260–275). Baltmannsweiler: Schneider Verlag Hohengehren.
- Scheele, V. (2006). *Entwicklung fortgeschrittener Rechtschreibfertigkeiten. Ein Beitrag zum Erwerb der „orthographischen“ Strategien*. Frankfurt am Main: Peter Lang.
- Spinner, K. H. (2008). Bildungsstandards und Literaturunterricht. In M. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), *Perspektiven der Didaktik. Zeitschrift für Erziehungswissenschaft, Sonderheft 9/2008* (S. 313–323). Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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.

1.3 Beschreibung der im Fach Mathematik untersuchten Kompetenzen

Alexander Roppelt, Kristina Reiss und Nicole Haag

Dieses Kapitel beschreibt die im IQB-Bildungstrend 2016 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, wie mathematische Kompetenzen empirisch am besten beschrieben werden können. Im zweiten Teil dieses Kapitels folgt deshalb eine Darstellung zur Struktur und zur Operationalisierung der mathematischen Kompetenzen im IQB-Bildungstrend 2016.

1.3.1 Die Bildungsstandards im Fach Mathematik für den Primarbereich

Bei den Bildungsstandards im Fach Mathematik für den Primarbereich (KMK, 2005) handelt es sich formal um eine präskriptive beziehungsweise normative Setzung, die definiert, welche mathematischen Kompetenzen Schülerinnen und Schüler in Deutschland bis zur 4. 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 und Kollegen (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

füßen 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.

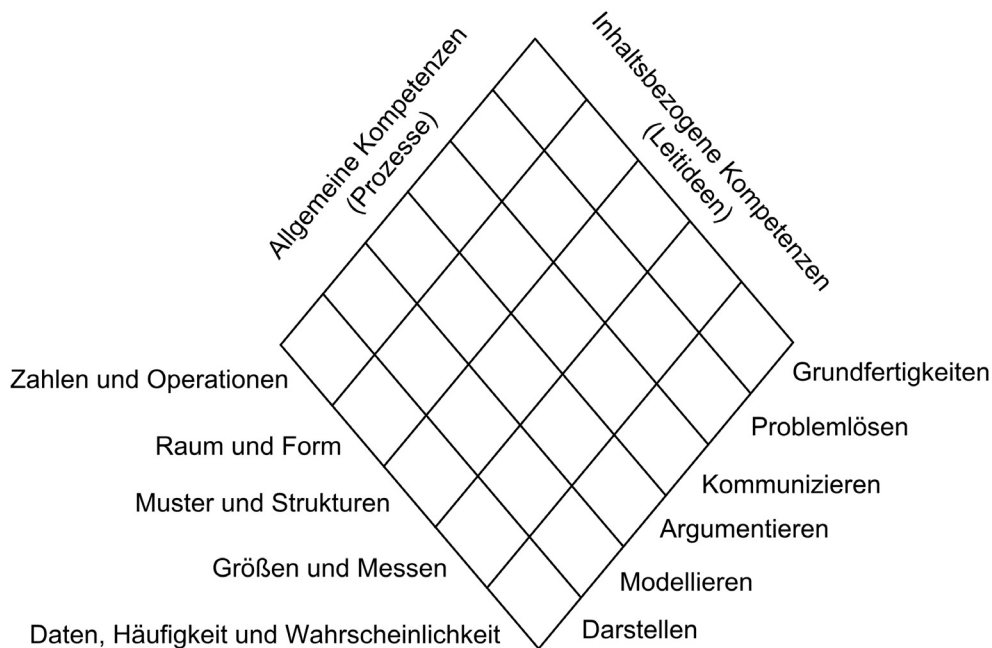
In den folgenden Abschnitten wird zunächst das theoretische Modell dargestellt, das die Bildungsstandards zur Beschreibung mathematischer Kompetenz heranziehen. Anschließend werden die einzelnen Komponenten des Modells genauer charakterisiert.

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. Abb. 1.3). 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 sie generalisiert werden, wenn sie also nicht mehr eng mit einzelnen konkre-

Abbildung 1.3: Modell mathematischer Kompetenz der Bildungsstandards im Fach Mathematik für den Primarbereich ohne Darstellung der Anforderungsbereiche



ten 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 sowie den Bildungsstandards für die Allgemeine Hochschulreife zugrunde liegt. Auch die Facetten der einzelnen Dimensionen werden im Primar- und im Sekundarbereich in ähnlicher Weise beschrieben (vgl. Blum, 2015; 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 zuge-dacht“ (Walther & Granzer, 2009, S. 116). Obwohl jede der in den Bildungsstandards beschriebenen allgemeinen Kompetenzen ihren eigenständigen Kern

hat, lassen 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, 2005) 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, 2013; 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, 2005) 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 Hilfstechniken 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.¹ 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-

1 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, die 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.

1.3.2 Struktur mathematischer Kompetenzen im IQB-Bildungstrend 2016

Wie bereits im IQB-Ländervergleich 2011 basiert die Erfassung der mathematischen Kompetenzen auch im IQB-Bildungstrend 2016 auf dem in Abschnitt 1.3.1 beziehungsweise Abbildung 1.3 dargestellten Strukturmodell der Bildungsstandards für den Primarbereich. Das Modell soll möglichst alle im Fach Mathematik relevanten Teilfähigkeiten integrieren und umfasst daher sowohl Inhalts- als auch Prozessaspekte. Die verschiedenen inhaltlichen und allgemeinen Kompetenzen lassen sich dabei nicht vollständig trennscharf gegeneinander abgrenzen, sondern sie weisen – wie bereits erwähnt – an einigen Stellen Überlappungen und fließende Übergänge auf. 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 (vgl. auch Weirich, Haag & Roppelt, 2012). 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 besonders wünschenswert wäre, weil sie im Mathematikunterricht eine Schlüsselrolle einnehmen sollen.

Für die Beschreibung der von den Schülerinnen und Schülern in den Ländern erreichten Kompetenzen im Fach Mathematik ergibt sich somit die Frage, welche der in den Bildungsstandards unterschiedenen Aspekte auch empirisch voneinander differenziert werden können. Die prozessbezogenen allgemeinen Kompetenzen weisen nicht nur große konzeptuelle Überlappungen auf, sondern kommen überdies meist in Kombinationen zum Einsatz und interagieren dabei in komplexer Weise miteinander. Eine separate Messung der einzelnen allgemeinen Kompetenzen ist daher praktisch nicht oder nur mit erheblichen Unschärfen möglich und wird aus diesem Grund in den Studien des IQB nicht vorgenommen. Die inhaltlichen Kompetenzen beziehungsweise Leitideen weisen zwar ebenfalls teilweise deutliche Überlappungen auf, können aber insgesamt besser voneinander abgegrenzt werden. Über drei Viertel der eingesetzten Testaufgaben lassen sich eindeutig einer dominierenden Leitidee zuordnen (siehe unten), sodass eine separate Messung der Leitideen möglich ist.

Im IQB-Ländervergleich 2011 wurden neben Ergebnissen für die einzelnen Leitideen auch Werte auf einer Globalskala der Mathematikkompetenz berichtet (Haag & Roppelt, 2012). Auf dieser Globalskala wurde die Kompetenzausprägung einer Schülerin beziehungsweise eines Schülers im Fach Mathematik anhand eines einzelnen Zahlenwerts zusammenfassend dargestellt. Die Betrachtung der globalen Mathematikkompetenz ist für die meisten Fragestellungen ausreichend aussagekräftig, da die auf die Leitideen bezogenen mathematischen Teilfähigkeiten stark miteinander zusammenhängen (vgl. Winkelmann, Robitzsch, Stanat & Köller, 2012). Sowohl die Darstellung der erreichten Kompetenzen auf einer Globalskala mathematischer Kompetenz als auch die differenziertere Betrachtung der in den einzelnen Leitideen erzielten Kompetenzwerte haben sich empirisch bewährt (z.B. Klieme, 2000; Winkelmann & Robitzsch, 2009; Winkelmann et al., 2012) und beides ist in großen Schulleistungsstudien üblich (vgl. Blum et al., 2004 sowie Roppelt, Penk, Pöhlmann & Pietsch, 2013 für die Sekundarstufe; Ufer, Reiss & Heinze, 2009 sowie Roppelt & Reiss, 2012 für den Primarbereich). Ob es sinnvoll ist, die ein-

zelen Leitideen getrennt zu berichten, ist in Abhängigkeit von den Zielen der Untersuchung und der Datenlage zu entscheiden.

Die im Bericht zum IQB-Ländervergleich 2011 ausführlich dargestellten Überlegungen und Befunde zur Dimensionalität mathematischer Kompetenzen (Roppelt & Reiss, 2012) legen nahe, die Berichtlegung im IQB-Bildungstrend 2016 analog zum Vorgehen bei früheren Schulleistungsstudien zu gestalten (z.B. Blum et al., 2004). Es werden also vornehmlich Ergebnisse für einen Globalwert berichtet, der die gesamte Breite der Kompetenzen im Fach Mathematik abdeckt. Bei den in Kapitel 6.2 dargestellten Analysen zu den Kompetenzwerten, die Schülerinnen und Schüler in den Ländern im Mittel erreicht haben, werden die Ergebnisse auch für die fünf Leitideen einzeln ausgewiesen. Für die Analysen zum Erreichen der Bildungsstandards und zu den Kompetenzstufenbesetzungen im Ländervergleich findet sich die differenzierte Darstellung nach Kompetenzbereichen für das Jahr 2016 in Kapitel 5.2. Die Ergebnisse der Trendanalysen für die einzelnen Leitideen können auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

1.3.3 Operationalisierung mathematischer Kompetenzen im IQB-Bildungstrend 2016

Die in den Ländern erreichten mathematischen Kompetenzen wurden mit Aufgaben aus dem IQB-Aufgabenpool für die Primarstufe erfasst. Ein Großteil der im IQB-Bildungstrend 2016 eingesetzten Aufgaben wurde bereits im Vorfeld des IQB-Ländervergleichs 2011 in einem aufwändigen Prozess entwickelt und erprobt (vgl. Granzer, 2009). Um den im IQB-Ländervergleich 2011 eingesetzten Itempool zu aktualisieren, wurden für die Testdurchführung im Jahr 2016 zusätzliche Aufgaben entwickelt und im Rahmen der Pilotierungen zu den Vergleichsarbeiten empirisch erprobt. Für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf wurden ebenfalls neue Aufgaben erstellt, deren Erprobung im Herbst 2015 erfolgte. Die Entwicklung dieser Aufgaben wird in Kapitel 13.2 dargestellt. Die im IQB-Bildungstrend 2016 eingesetzte Itemauswahl umfasst sowohl sehr leichte Aufgaben, um Kompetenzen im unteren Leistungsbereich abbilden zu können, als auch herausfordernde Aufgaben, um die Kompetenzen von leistungsstarken Schülerinnen und Schülern zu erfassen.

In den meisten Fällen bestehen die im IQB-Bildungstrend 2016 eingesetzten Aufgaben aus einem bis zu fünf Zeilen langen Aufgabentext (Stimulus), der die Schülerinnen und Schüler in den Aufgabenkontext einführt. Einige Stimuli enthalten zusätzlich zum Text Diagramme, Zeichnungen oder Illustrationen. Auf einen Stimulus folgen bis zu sieben darauf bezogene Teilaufgaben (Items), wobei der überwiegende Teil der Aufgaben im Fach Mathematik nur ein Item enthält. Im IQB-Bildungstrend 2016 wurden im Fach Mathematik 237 Aufgaben mit insgesamt 406 Items eingesetzt, von denen 56 Items zwei Leitideen zugeordnet waren. Ungefähr ein Viertel dieser Items hatten ein Mehrfachwahlformat (*Multiple-Choice*). Bei knapp drei Vierteln der Items mussten die Schülerinnen und Schüler eine kurze Antwort angeben, beispielsweise das Ergebnis einer Rechnung, oder eine Zeichnung ergänzen. Nur bei einigen wenigen Items wurde eine mehrere Sätze umfassende Antwort verlangt, in der beispielsweise ein Rechenweg beschrieben werden sollte. Bezogen auf die einzelnen Leitideen variiert die Anzahl der Items zwischen 120 im Bereich *Zahlen und Operationen* und 66 im Bereich *Muster und Strukturen*. Die zwischen den Leitideen unterschiedli-

chen Itemzahlen sind teilweise durch Unterschiede in der inhaltlichen Breite der Leitideen und teilweise durch den Anteil mehrfach zugeordneter Items bedingt. Mehrfachzuordnungen von Items zu Leitideen kommen dadurch zustande, dass für die Bearbeitung des jeweiligen Items verschiedene inhaltliche Kompetenzen benötigt werden. Die im IQB-Bildungstrend 2016 eingesetzten Items wurden so entwickelt, dass sie maximal zwei unterschiedlichen Leitideen zugeordnet werden können. Items, für deren Bearbeitung mehr als eine inhaltliche Kompetenz benötigt wird, werden beiden relevanten Leitideen zugeordnet und in der Zählung berücksichtigt. Eine detaillierte Aufstellung der Itemzahlen pro Leitidee findet sich in Kapitel 13.1.

Literatur

- 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. (2015). Zur Konzeption der Bildungsstandards Mathematik für die Allgemeine Hochschulreife. In W. Blum, S. Vogel, C. Drüke-Noe, & A. Roppelt (Hrsg.), *Bildungsstandards aktuell. Mathematik in der Sekundarstufe II* (S. 16–30). Braunschweig: Bildungshaus Schulbuchverlage.
- Blum, W., Drüke-Noe, C., Hartung, R. & Köller, O. (Hrsg.). (2006). *Bildungsstandards Mathematik: konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen*. Berlin: Cornelsen Scriptor.
- 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, C. 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.
- 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.
- 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.
- Haag, N. & Roppelt, A. (2012). Der Ländervergleich im Fach Mathematik. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 117–127). Münster: Waxmann.
- 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.
- IQB (2013) = Institut zur Qualitätsentwicklung im Bildungswesen. (Hrsg.). *Kompetenzstufenmodell zu den Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4)*. Zugriff am 07.09.2017 unter https://www.iqb.hu-berlin.de/bista/ksm/KSM_GS_Mathemati_4.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 Naturwissen-*

- schaftsstudie – 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 (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- 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 Psychoeducational Assessment*, 13, 21–38.
- NCTM (2000) = National Council of Teachers of Mathematics (Hrsg.). (2000). *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.
- 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.
- Roppelt, A., Penk, C., Pöhlmann, C. & Pietsch, E. (2013). Der Ländervergleich im Fach Mathematik. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 123–140) Münster: Waxmann.
- Roppelt, A. & Reiss, K. (2012). Beschreibung der im Fach Mathematik untersuchten Kompetenzen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 34–48). Münster: Waxmann.
- 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.
- 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., Selter, 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.
- Weirich, S., Haag, N. & Roppelt, A. (2012). Testdesign und Auswertung des Ländervergleichs: technische Grundlagen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 277–290). Münster: Waxmann.
- 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.

Kapitel 2

Kompetenzstufenmodelle für den Primarbereich

2.1 Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen

Hans Anand Pant, Katrin Böhme, Petra Stanat, Stefan Schipolowski und Olaf Köller

2.1.1 Das Kompetenzkonzept der Bildungsstandards

Bei Kompetenzen handelt es sich um theoretische Konstrukte, die nur mit Hilfe von geeigneten Messinstrumenten der Beobachtung zugänglich gemacht werden können (Köller, 2008). Das genaue theoretische Verständnis von Kompetenz ist dabei – je nach fachwissenschaftlicher Ausrichtung – durchaus unterschiedlich (für einen Überblick vgl. Klieme & Hartig, 2007). Den Bildungsstandards der Kultusministerkonferenz (KMK) liegt ein breit akzeptiertes und häufig verwendetes Kompetenzkonzept zugrunde, das Kompetenzen als Fähigkeiten und Fertigkeiten betrachtet, die sich in konkreten Anforderungssituationen als „Können“ manifestieren. Entsprechend werden die Kompetenzerwartungen der Bildungsstandards im Sinne von *Can-do-Statements* als Fähigkeiten der Schülerinnen und Schüler definiert, bestimmte Anforderungen bewältigen zu können, wie zum Beispiel:

„[Sie können] altersgemäße Texte sinnverstehend lesen, [...] gezielt einzelne Informationen suchen, [...] Aussagen mit Textstellen belegen, [...].“ (Kompetenzbereich *Lesen* – mit Texten und Medien umgehen; KMK, 2005, S. 11f.)

„[Sie können] geübte, rechtschreibwichtige Wörter normgerecht schreiben [und] Rechtschreibstrategien verwenden: Mitsprechen, Ableiten, Einprägen.“ (Kompetenzbereich *Schreiben*, Teilbereich „richtig schreiben“; KMK, 2005, S. 10)

Dieser handlungsorientierte Kompetenzbegriff der Bildungsstandards entspricht der Konzeption von Klieme und Leutner (2006). Sie 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; Hervorhebungen im Original). Kennzeichnend für dieses Kompetenzkonzept sind

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

- die Abgrenzung zu allgemeinen Fähigkeitskonstrukten wie etwa breiten Intelligenzfaktoren zugunsten einer engeren Definition des Expertisebereichs („*kompetent wofür?*“),
- der funktionale Handlungsbezug (*Can-do-Aussagen*),
- der Bezug auf Fähigkeiten, die situationsangemessen und in wechselnden Kontexten angewendet werden können, und
- die Fokussierung des kognitiven Aspekts, um das „Können“ getrennt von motivationalen (z.B. Fachinteresse) und affektiven (z.B. Leistungsangst) Einflussgrößen auf das Leistungsgeschehen zu betrachten.

2.1.2 Kompetenzstufenmodelle

Die oben dargelegte Definition beschreibt zunächst, was allgemein unter einer Kompetenz verstanden wird. Eine genauere theoretische Modellierung von Kompetenzen muss jedoch domänenspezifisch erfolgen, wobei die Aspekte der *Struktur* eines Kompetenzbereichs, der *Graduierung* einzelner Kompetenzen in Niveaus beziehungsweise Stufen und der *individuellen Entwicklung* von Kompetenzen im Zeitverlauf zu unterscheiden sind. Da die auf den Bildungsstandards basierenden Kompetenzmessungen primär für punktuelle Bestandsaufnahmen auf der Ebene des Systems (IQB-Bildungstrends) und auf der Ebene einzelner Schulen beziehungsweise Klassen (Vergleichsarbeiten) genutzt werden sollen, liegt der Fokus hier auf Fragen der Graduierung, wobei im Prozess der Aufgabenentwicklung auch Strukturaspekte berücksichtigt werden (Pant, Böhme & Köller, 2012). Die Modellierung individueller Verläufe, für die Entwicklungsmodelle erforderlich sind, ist für solche Bestandsaufnahmen dagegen weniger relevant.

Die Überprüfung des Erreichens der Bildungsstandards soll zu einem wichtigen Zeitpunkt des Bildungsverlaufs bilanzierende Aussagen über die Verteilung der von Schülerinnen und Schülern erreichten Kompetenzen in den Bildungssystemen der Länder treffen und zudem aufzeigen, wie sich diese Verteilungen über die Zeit verändern. Zu diesem Zweck werden theoretisch ausgearbeitete *Kompetenzstufenmodelle* benötigt. Kompetenzstufen- beziehungsweise Kompetenzniveaumodelle stellen dar, ab welcher in Stufen graduierten Ausprägung einer Kompetenz Schülerinnen und Schüler welche konkreten kognitiven Anforderungen bewältigen können (vgl. Klieme & Leutner, 2006). Die Modelle unterteilen also das kontinuierliche Fähigkeitsspektrum in Abschnitte. Die Beschreibung der Anforderungen, die auf den einzelnen Kompetenzstufen mit hoher Wahrscheinlichkeit bewältigt werden, basiert auf einer Analyse von Testaufgaben beziehungsweise -items, die aufgrund ihrer empirisch bestimmten Schwierigkeit auf der Kompetenzskala derselben Stufe zuzuordnen sind. Es wird dabei bestimmt, welche kognitiven Operationen zur Bewältigung der Aufgabenstellungen jeweils erforderlich sind.

Die fachdidaktisch und lernpsychologisch begründeten Kompetenzstufenmodelle für die Fächer Deutsch und Mathematik im Primarbereich werden in den Kapiteln 2.2 und 2.3 detailliert vorgestellt. Im Folgenden soll zunächst kurz das allgemeine, fächerübergreifende Vorgehen bei der Definition der Kompetenzstufen beschrieben werden.

Das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) hat in den Jahren 2006 und 2007 für die Fächer Deutsch und Mathematik Kalibrierungsbeziehungsweise Normierungsstudien durchgeführt (vgl. hierzu auch Granzer

et al., 2009). Ziel der Studien in der Primarstufe, die mit repräsentativen Stichproben von Schülerinnen und Schülern der Jahrgangsstufen 3 und 4 durchgeführt wurden, war es, anhand empirisch gewonnener Daten länderübergreifend gültige Skalen zu definieren, auf denen sich die von den Schülerinnen und Schülern erreichten Kompetenzen und die Schwierigkeiten der jeweiligen (Teil-)Aufgaben verorten lassen. Diese bildeten wiederum den Ausgangspunkt für die Entwicklung von Kompetenzstufenmodellen zu den Bildungsstandards im Primarbereich. Auf dieser Grundlage wurde im Jahr 2011 der erste IQB-Ländervergleich in der Primarstufe durchgeführt. Im Rahmen dieser Studie wurden die Kompetenzskalen zu den Kompetenzbereichen *Lesen* und *Zuhören* im Fach Deutsch sowie die Skalen im Fach Mathematik (*Globalskala* sowie die Skalen zu den fünf mathematischen Leitideen; vgl. Kapitel 1.3) so angepasst, dass sie in der Population der Viertklässlerinnen und Viertklässler des allgemeinbildenden Schulsystems im Jahr 2011 einen Mittelwert von $M = 500$ Punkten und eine Streuung von $SD = 100$ Punkten aufweisen. Diese Skalenbildung folgt den aus bisherigen IQB-Ländervergleichsstudien und IQB-Bildungstrends (vgl. bspw. Köller, Knigge & Tesch, 2010; Stanat, Böhme, Schipolowski & Haag, 2016; Stanat, Pant, Böhme & Richter, 2012) und aus internationalen Studien wie PIRLS/IGLU¹ und PISA² bekannten Konventionen (vgl. bspw. Baumert et al., 2002; Bos et al., 2003).³ Die Berichtlegung im IQB-Bildungstrend 2016 zu den genannten Kompetenzbereichen basiert ebenfalls auf diesen Festlegungen. Davon abweichend wurde die Skala im Kompetenzbereich *Orthografie* im IQB-Bildungstrend 2016 so angepasst, dass sich der Mittelwert von $M = 500$ Punkten und die Standardabweichung von $SD = 100$ Punkten auf die Population der Viertklässlerinnen und Viertklässler im Jahr 2016 bezieht. Dieses Vorgehen liegt darin begründet, dass in die Testungen zum IQB-Ländervergleich 2011 im Kompetenzbereich *Orthografie* – anders als im *Lesen* und im *Zuhören* – keine Schülerinnen und Schüler aus Förderschulen einbezogen wurden. Die Ergebnisse des IQB-Bildungstrends 2016 im Bereich *Orthografie* werden daher nicht auf der Metrik des Jahres 2011, sondern auf der Metrik des Jahres 2016 dargestellt, die auch Schülerinnen und Schüler an Förderschulen berücksichtigt und somit den Grundstein für zukünftige Trendschätzungen auf Länderebene in diesem Kompetenzbereich legt (siehe auch Kapitel 3.2).⁴

Unter der idealtypischen Annahme, dass die Daten normalverteilt sind, ergeben sich bei dieser Skalendefinition die in Abbildung 2.1 dargestellten Zusammenhänge zwischen der Standardabweichung (SD) und den Prozenträngen. Für den Kompetenzmittelwert von $M = 500$ gilt theoretisch, dass die Kompetenzen

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

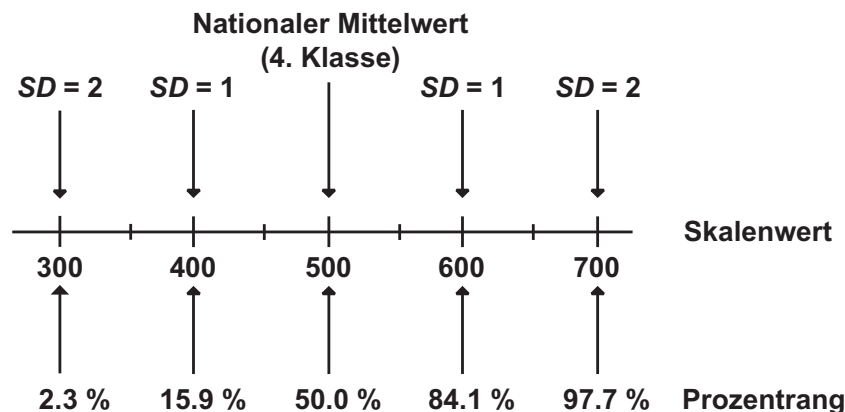
2 Das Akronym PISA steht für *Programme for International Student Assessment*.

3 Es sei an dieser Stelle ausdrücklich darauf hingewiesen, dass sich – ungeachtet der rein numerisch gleichen Skalenfestlegung – die Testergebnisse aus verschiedenen Untersuchungen nicht direkt miteinander vergleichen lassen. Ein Mittelwert von 500 Kompetenzpunkten für Land X im Lesekompetenztest des IQB-Bildungstrends bedeutet also nicht das Gleiche wie ein Mittelwert von 500 Punkten für Land Y im PISA-Lesekompetenztest. Um die Ergebnisse von unterschiedlichen Testverfahren miteinander vergleichen zu können, bedarf es zuvor einer inhaltlichen und statistischen Überprüfung der Äquivalenz (vgl. Böhme et al., 2014; Hartig & Frey, 2012; Pietsch, Böhme, Robitzsch & Stubbe, 2009).

4 Diese Skalenfestlegung weicht von der Metrik des bereits veröffentlichten Kompetenzstufenmodells des IQB für den Kompetenzbereich *Orthografie* ab. Daher wurde für die Berechnungen zum IQB-Bildungstrend 2016 eine rein nominelle Anpassung der Stufengrenzen vorgenommen, um sicherzustellen, dass die inhaltliche Bedeutung der Stufen für den Kompetenzbereich *Orthografie* unverändert bleibt.

von 50 Prozent der Schülerinnen und Schüler unter dem Mittelwert liegen und die Kompetenzen von 50 Prozent der Schülerinnen und Schüler darüber. Bei einem Wert von 600 Punkten (Mittelwert plus eine Standardabweichung) liegen die Leistungen von rund 84 Prozent der Grundschülerinnen und Grundschüler darunter und lediglich 16 Prozent erreichen einen noch höheren Wert. Umgekehrt gilt für einen Kompetenzwert von 400 Punkten (Mittelwert minus eine Standardabweichung), dass ihn lediglich 16 Prozent der Schülerschaft unterschreiten.

Abbildung 2.1: Kontinuierliche Kompetenzskala ($M = 500$ und $SD = 100$) mit Prozenträngen unter der Annahme einer Normalverteilung



Für die Interpretation von Skalenwerten ist es wichtig, dass sich auch die Aufgaben beziehungsweise Teilaufgaben (Items) mit ihrer Schwierigkeit auf dieser Skala abbilden lassen. Sehr leichte Items liegen bei Schwierigkeitswerten von 400 Punkten und geringer, sehr schwierige Items bei Werten von 600 Punkten und höher. Hat ein Item zum Beispiel eine Schwierigkeit von exakt 600 Punkten, so bedeutet dies, dass Schülerinnen und Schüler mit einer Fähigkeit von 600 und mehr Punkten dieses Item mit einer hinreichenden Wahrscheinlichkeit lösen.⁵ Schülerinnen und Schüler mit einem Fähigkeitswert unter 600 Punkten lösen das Item mit einer geringeren Wahrscheinlichkeit.

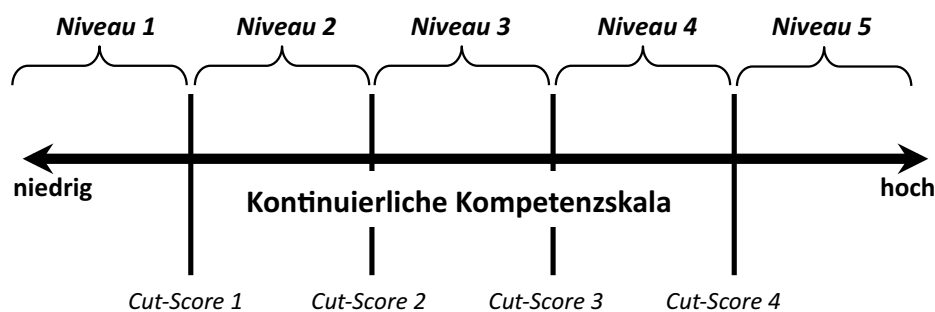
Die Möglichkeit, die Fähigkeit von Personen und die Schwierigkeit von Items auf einer gemeinsamen Skala verorten zu können, macht man sich bei der Definition von Kompetenzstufen zunutze. So kann ein Item, das beispielsweise einen Schwierigkeitswert von 600 Punkten aufweist, im Hinblick darauf analysiert werden, welche kognitiven Operationen zu seiner Lösung erforderlich sind. Dies wiederum lässt den Rückschluss zu, dass Schülerinnen und Schüler, die einen Skalenwert von 600 Punkten oder höher erreicht haben, die zur Lösung dieses Items erforderlichen Operationen beherrschen. Erweitert man diese Idee, so kann man viele Items nach ihrer Schwierigkeit ordnen 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. Dieses Vorgehen wird als *Standard-Setting* bezeichnet (Cizek & Bunch, 2007). Ein *Standard-Setting* zielt somit auf die Festlegung von Schwellenwerten (*Cut-Scores*) ab, die eine Kompetenzskala in sukzessive Teilbereiche – sogenannte Kompetenzstufen oder Kompetenzniveaus – einteilen (Pant, Tiffin-Richards & Köller, 2010).

5 „Hinreichende Wahrscheinlichkeit“ bedeutet hier, dass die Lösungswahrscheinlichkeit bei $p = .625$ bzw. 62.5 Prozent liegt (vgl. Köller et al., 2010). Diese Definition orientiert sich an den PISA-Studien (vgl. Frey, Carstensen, Walter, Rönnebeck & Gomolka, 2008).

Abbildung 2.2 zeigt schematisch, dass beim Standard-Setting auf dem Kontinuum der Kompetenzskala durch Cut-Scores festgelegt wird, wo sich jeweils Übergänge zwischen Aufgabengruppen befinden, die abgrenzbare Anforderungen beinhalten und zu deren Bewältigung zunehmend komplexe kognitive Fähigkeiten erforderlich sind. Die inhaltliche Beschreibung der Kompetenzniveaus erfolgt dann anhand der entsprechenden Fähigkeiten (sog. Kompetenzstufendeskriptoren), um zu verdeutlichen, über welche konkreten Kompetenzen Schülerinnen und Schüler auf den jeweiligen Stufen verfügen (z. B. „Schülerinnen und Schülern auf der untersten Kompetenzstufe gelingt es, Einzelinformationen im Text zu lokalisieren und zu verarbeiten ...“).

In die vom IQB durchgeführten Standard-Setting-Verfahren zu den Kompetenzstufenmodellen im Primarbereich wurden Expertinnen und Experten aus der Fachdidaktik, Testentwicklung, Schulpraxis und Bildungsadministration einbezogen, die anhand der empirischen Daten zur Aufgabenschwierigkeit aus den oben genannten Kalibrierungs- beziehungsweise Normierungsstudien die Stufengrenzen festgelegt haben.

Abbildung 2.2: Schematische Darstellung der in einem Standard-Setting festgelegten Stufengrenzen



2.1.3 Die 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., 2003) hat die KMK in ihren Beschlüssen keine *Mindeststandards*, sondern *Regelstandards* definiert. Die Bildungsstandards im Primarbereich legen also Kompetenzerwartungen fest, die Schülerinnen und Schüler zum Ende der 4. Jahrgangsstufe „in der Regel“ beziehungsweise „im Mittel“ erfüllen sollten. Entsprechend war bei der Erarbeitung der Kompetenzstufenmodelle unter anderem festzulegen, ab welchem Niveau die Vorgaben der jeweiligen Regelstandards erreicht sind. Darüber hinaus sollten die Kompetenzstufenmodelle im Sinne der von Klieme und Kollegen (2003) geforderten Differenzierung weitere Niveaus definieren, um auch das Erreichen von Mindest- und Optimalstandards bestimmen zu können. Entsprechend wurden für die Kompetenzstufenmodelle in der Primarstufe, die auch die Basis der Berichtlegung im IQB-Bildungstrend 2016 bilden, jeweils fünf Kompetenzstufen festgelegt, von denen die mittlere Stufe (Kompetenzstufe III) als Regelstandard interpretiert wird. Ferner wurden die folgenden inhaltlichen Definitionen vorgenommen (vgl. Köller, 2010):

- *Mindeststandards* (Kompetenzstufe II) beziehen sich auf ein definiertes Minimum an Kompetenzen, das alle Schülerinnen und Schüler bis zu einem bestimmten Bildungsabschnitt erreicht haben sollten. Dieses Minimum unterschreitet die in den Publikationen der KMK festgelegten Kompetenzerwartungen der Regelstandards.
- *Regelstandards* (Kompetenzstufe III) beziehen sich auf Kompetenzen, die im Durchschnitt von den Schülerinnen und Schülern bis zu einem bestimmten Bildungsabschnitt erreicht werden sollen und den Veröffentlichungen der KMK-Bildungsstandards entsprechen.
- Als *Regelstandard plus* (Kompetenzstufe IV) wird ein Leistungsbereich definiert, der über den Regelstandards liegt und als Zielperspektive für die Weiterentwicklung von Unterricht angesehen werden kann.
- *Optimalstandards* (Kompetenzstufe V) beziehen sich auf Kompetenzerwartungen, die bei sehr guten oder ausgezeichneten individuellen Lernvoraussetzungen und der Bereitstellung geeigneter Lerngelegenheiten innerhalb und außerhalb der Schule erreicht werden können und die 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 der Länder 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 Kompetenzminimum beschreiben, das von *allen* Schülerinnen und Schülern mittelfristig erreicht werden soll (*Mindeststandard*),
- motivierende Kompetenzerwartungen formulieren, die Entwicklungsimpulse an den Schulen auslösen; hierfür dient insbesondere der *Regelstandard plus*,
- anschlussfähig an internationale Vorarbeiten, für die Primarstufe vor allem aus PIRLS/IGLU und TIMSS⁶, sein,
- ungefähr gleich breite Kompetenzstufen zur einheitlichen, über mehrere Erhebungswellen hinweg stabilen kriterialen Kompetenzbeschreibung definieren sowie
- fachdidaktisch gut interpretierbare und vertretbare Grenzen zwischen den Kompetenzstufen aufweisen.

Diese inhaltlichen, bildungspolitischen und pädagogischen Anforderungen gelten für alle Kompetenzstufenmodelle des IQB.

In den nächsten beiden Kapiteln folgt eine Beschreibung der fachspezifischen Kompetenzstufenmodelle, die vom IQB in enger Zusammenarbeit mit fachdidaktischen Kooperationspartnern konstruiert worden sind und die Grundlage für den IQB-Bildungstrend 2016 bilden.

6 Das Akronym TIMSS steht für *Trends in International Mathematics and Science Study*.

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.
- Böhme, K., Richter, D., Weirich, S., Haag, N., Wendt, H., Bos, W., Pant, H. A. & Stanat, P. (2014). Messen wir dasselbe? Zur Vergleichbarkeit des IQB-Ländervergleichs 2011 mit den internationalen Studien IGLU und TIMSS 2011. *Unterrichtswissenschaft*, 42 (4), 342–365.
- 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.
- Cizek, G. J. & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. 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. 375–397). 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.
- Hartig, J. & Frey, A. (2012). Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten. *Psychologische Rundschau*, 63, 43–49.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards – Expertise*. Bonn, Berlin: 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 (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- Köller, O. (2008). Bildungsstandards – Verfahren und Kriterien bei der Entwicklung von Messinstrumenten. *Zeitschrift für Pädagogik*, 54, 163–173.
- 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.
- Pant, H. A., Böhme, K. & Köller, O. (2012). Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (S. 49–55). Münster: Waxmann.
- 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.

- Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Hrsg.). (2016). *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

2.2 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 der Bundesrepublik Deutschland 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 den von der Kultusministerkonferenz verabschiedeten Bildungsstandards orientieren und dabei nach Möglichkeit das gesamte Spektrum der Kompetenzausprägungen auf Seiten der Schülerinnen und Schüler 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 – Umgang mit Texten und Medien* sowie *Sprache und Sprachgebrauch untersuchen* unterschieden (KMK, 2005). Der Bereich *Orthografie* ist dort nicht gesondert ausgewiesen, sondern wird unter *Schreiben* subsumiert. Im Rahmen des IQB-Bildungstrends 2016 wurden – wie bereits im IQB-Ländervergleich 2011 – nur Aufgaben zu den Kompetenzbereichen beziehungsweise Teilbereichen *Lesen*, *Zuhören* und *Orthografie* eingesetzt (vgl. Kapitel 1.2). Im Folgenden werden die Kompetenzstufenmodelle für diese Bereiche vorgestellt und durch Beispielaufgaben illustriert.⁷

2.2.1 Das Kompetenzstufenmodell für den Bereich *Lesen*

Das Kompetenzstufenmodell für den Bereich *Lesen* wurde im Jahr 2009 im Rahmen eines *Standard-Settings* unter Einbindung von Lehrkräften, Fachdidaktikerinnen und Fachdidaktikern sowie Expertinnen und Experten aus der Psychometrie auf Grundlage der Kalibrierungs- beziehungsweise Normierungsstudien des IQB im Primarbereich entwickelt (vgl. Kapitel 2.1) und im Frühjahr 2010 von der Kultusministerkonferenz zustimmend zur Kenntnis genommen. Im Rahmen der Berichtlegung zum IQB-Ländervergleich 2011 wurde das Kompetenzstufenmodell redaktionell geringfügig überarbeitet. Zudem wurde die zugrunde liegende Kompetenzskala so angepasst, dass sie in der Population der Viertklässlerinnen und Viertklässler des allgemeinbildenden Schulsystems im Jahr 2011 einen Mittelwert von 500 Punkten und eine Standardabweichung von 100 Punkten aufweist. Auf diese Skala beziehen sich auch die Ergebnisse zum Kompetenzbereich *Lesen* im IQB-Bildungstrend 2016.

⁷ Die im Folgenden vorgestellten Beschreibungen der Modelle einschließlich weiterführender Erläuterungen sind auf der Webseite des IQB verfügbar unter <https://www.iqb.hu-berlin.de/bista/ksm>

Das Kompetenzstufenmodell 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 (Optimalstandard)

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 Abbildung 2.3 der Stimulustext der Aufgabe „Der aufmerksame Beobachter“ und in Abbildung 2.4 einige Beispielimens zu dieser Aufgabe.⁸

Die in Abbildung 2.4 gezeigten Beispielimens beziehen sich auf das Märchen „Der aufmerksame Beobachter“. Bei den abgedruckten Fragen geht es durchweg um Figurenaspekte beziehungsweise um Aspekte des Handlungsverlaufs. Damit zielen die Items auf einen zentralen Gegenstand des Deutschunterrichts in der Grundschule ab, der auch in den Bildungsstandards benannt wird: „Beim [...] Lesen literarischer Texte [...] setzen sich [die Kinder] identifizierend und abgrenzend mit literarischen Figuren auseinander.“ (KMK, 2005, S. 9) Das Beispielimens für Kompetenzstufe I bezieht sich auf die Handlungsmotive weniger zentraler Figuren (die Spatzen). Die richtige (erste) Option ist der Formulierung im Text ähnlich. Wer den Text als Ganzen nicht gelesen oder noch Probleme mit dem Dekodieren hat, mag auch die Distraktoren als plausibel ansehen. Kompetenzstufe II wird durch ein Item illustriert, bei dem nach den Eigenschaften des klugen Reiters gefragt wird. Für die richtige Lösung muss das Verhalten an mehreren Textstellen beziehungsweise im Gesamtkontext des Märchens erschlossen werden. Das Beispielimens auf Kompetenzstufe III zielt auf die Handlungsmotive einer zentralen Hauptfigur ab, und zwar auf den Mann, der sein Kamel verloren hat. Dieser macht sich auf den Weg, um sein Kamel zu suchen. Das Beispielimens zu Kompetenzstufe IV zielt hingegen darauf ab, auf Ebene des Textes wesentliche Aspekte zu erfassen, indem verschiedene Ereignisse in der Geschichte in eine Reihenfolge gebracht werden müssen. Kompetenzstufe V wird schließlich durch ein komplexes offenes Item illustriert, bei dem eine Behauptung selbstständig begründet werden muss. Dazu müssen die Kinder erkennen, dass der Text folgende Komplikation aufweist: Zwar geht die Gerichtsverhandlung gut aus; am Ende wird schließlich niemand verurteilt. Das eigentliche Ausgangsproblem ist jedoch nicht gelöst: Das Kamel bleibt verschwunden.

2.2.2 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 der Psychometrie hauptsächlich auf Grundlage der Normierungsstudie für den Primarbereich entwickelt (vgl. Kapitel 2.1) und im Frühjahr 2010 von der Kultusministerkonferenz zustimmend zur Kenntnis genommen. Im Rahmen der Berichtlegung zum

⁸ Für die Abschätzung von Trends ist es erforderlich, die Testaufgaben in mehreren aufeinanderfolgenden Bildungstrend-Studien unter vergleichbaren Bedingungen einzusetzen. Um dies zu gewährleisten, dürfen die Aufgaben nicht veröffentlicht werden. Bei der Aufgabe „Der aufmerksame Beobachter“ handelt es sich daher lediglich um ein illustrierendes Beispiel für das Kompetenzstufenmodell im Bereich *Lesen*, das den Schülerinnen und Schülern im IQB-Bildungstrend 2016 jedoch nicht zur Bearbeitung vorgelegt wurde.

Abbildung 2.3: 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?“

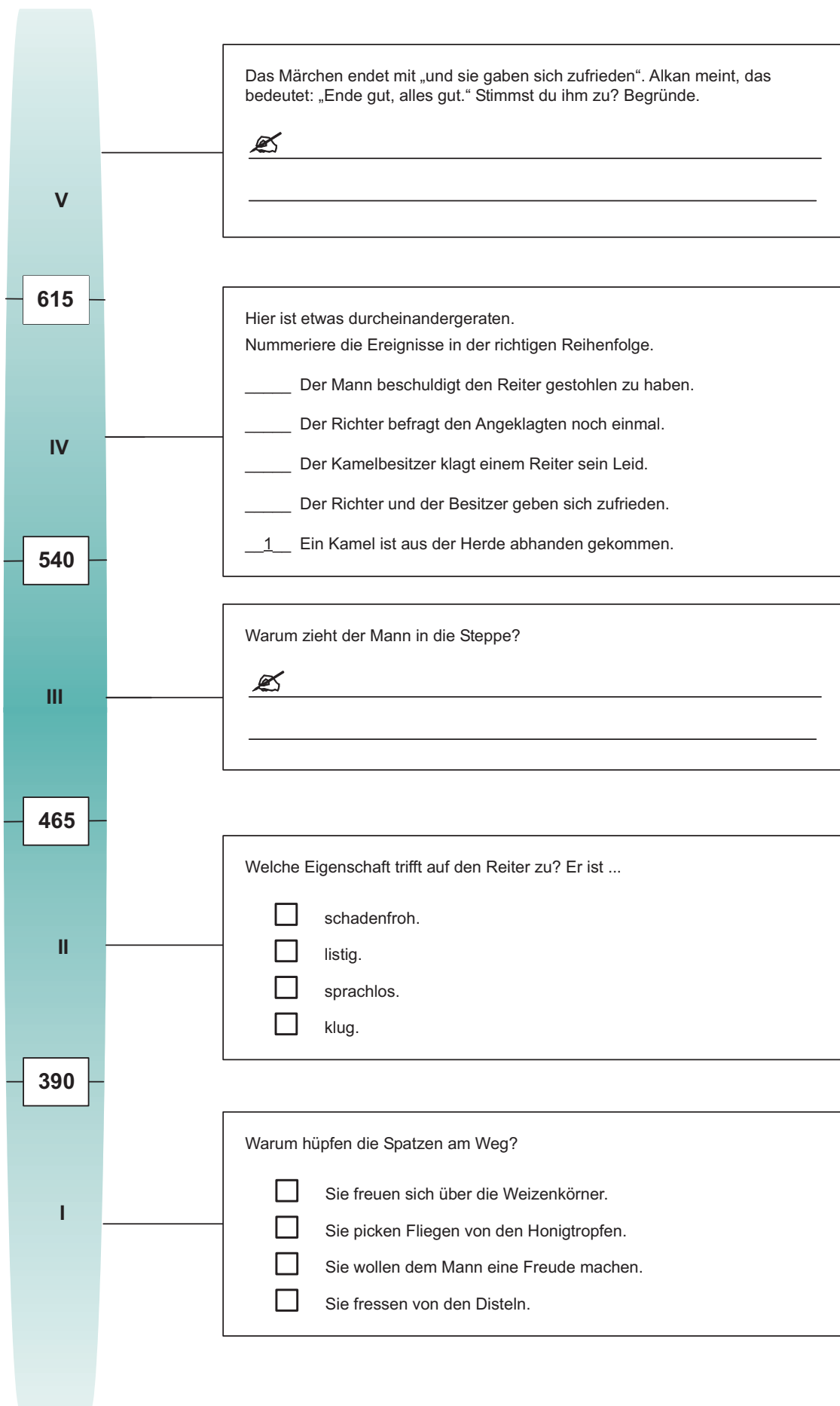
„Beim Grasens blieben in der Mitte immer einige Büschel der schmackhaften Disteln stehen.“

„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 2.4: Illustration des Kompetenzstufenmodells für den Bereich *Lesen*



V

615

IV

540

III

465

II

390

I

Das Märchen endet mit „und sie gaben sich zufrieden“. Alkan meint, das bedeutet: „Ende gut, alles gut.“ Stimmt du ihm zu? Begründe.



Hier ist etwas durcheinandergeraten.
Nummeriere die Ereignisse in der richtigen Reihenfolge.

- _____ Der Mann beschuldigt den Reiter gestohlen zu haben.
- _____ Der Richter befragt den Angeklagten noch einmal.
- _____ Der Kamelbesitzer klagt einem Reiter sein Leid.
- _____ Der Richter und der Besitzer geben sich zufrieden.
- __1__ Ein Kamel ist aus der Herde abhanden gekommen.

Warum zieht der Mann in die Steppe?



Welche Eigenschaft trifft auf den Reiter zu? Er ist ...

- schadenfroh.
- listig.
- sprachlos.
- klug.

Warum hüpfen die Spatzen am Weg?

- Sie freuen sich über die Weizenkörner.
- Sie picken Fliegen von den Honigtropfen.
- Sie wollen dem Mann eine Freude machen.
- Sie fressen von den Disteln.

IQB-Ländervergleich 2011 wurde auch das Kompetenzstufenmodell für den Teilbereich *Zuhören* redaktionell geringfügig überarbeitet und die zugrunde liegende Kompetenzskala so angepasst, dass sie in der Population der Viertklässlerinnen und Viertklässler des allgemeinbildenden Schulsystems im Jahr 2011 einen Mittelwert von 500 Punkten und eine Standardabweichung von 100 Punkten aufweist. Auf diese Skala beziehen sich auch die Ergebnisse zum Kompetenzbereich *Zuhören* im IQB-Bildungstrend 2016.

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 (Optimalstandard)

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 interpre-

tatorischen 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 des *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 kann also der *Regelstandard* 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 des *Optimalstandards* gelten.

Abbildung 2.5: 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.

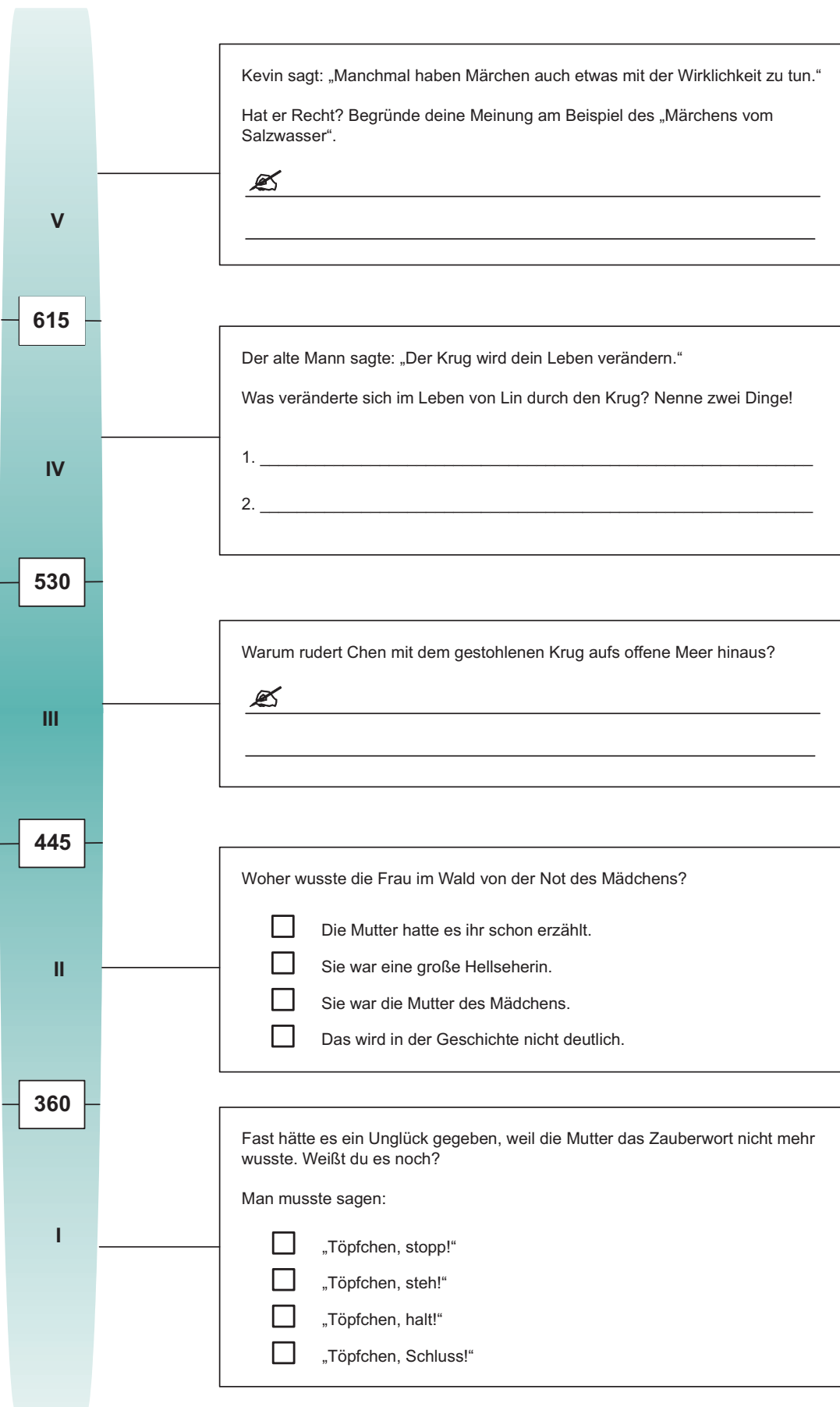
So ging es eine ganze Weile. Eines Abends war er aber so verzweifelt, dass er einfach nicht einschlafen konnte. Da stand mitten in der Nacht ein alter Mann vor ihm und sagte: „Lin, nun hast du Geduld und Mut bewiesen. Zur Belohnung schenke ich dir einen Krug. Er besitzt Zauberkräfte und wird dein Leben 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 2.6: Illustration des Kompetenzstufenmodells für den Teilbereich *Zuhören*



Zur Illustration der Kompetenzstufen I bis V finden sich in der Abbildung 2.5 die Transkripte der beiden Hörtexte der Zuhöraufgabe „Zwei Märchen“ und in der Abbildung 2.6 einige Beispielimens dieser Aufgabe.⁹

Die Zuhöraufgabe „Zwei Märchen“ besteht aus zwei literarischen Texten: dem Märchen „Der süße Brei“ sowie dem „Märchen vom Salzwasser“. Die abgedruckten Beispielimens sind auf jeweils einen der beiden Hörtexte bezogen. Es wurden allerdings auch Items zu Textvergleichen eingesetzt, die an dieser Stelle aus Platzgründen nicht abgedruckt sind. Das Beispielimens auf Kompetenzstufe I zielt auf die formelhafte Rede im Märchen „Der süße Brei“ ab. Hier muss eine prominente Information („Töpfchen, steh“) wiedergegeben werden, die für den Handlungsverlauf zentral ist und mehrmals im Hörtext explizit genannt wird. Kompetenzstufe II wird durch ein Item illustriert, das auf eine Unbestimmtheitsstelle im gleichen Text bezogen ist: Woher die Frau im Wald von der Not des Mädchens wusste, wird im Märchen nicht deutlich. Zur Lösung dieses Items sind einfache Schlüsse nötig, um Distraktoren der MC-Aufgabe ausschließen zu können. Das Beispielimens auf Kompetenzstufe III zielt auf ein zentrales Handlungsmotiv einer der Hauptfiguren des Märchens vom Salzwasser ab. Chen fährt mit dem Boot auf das Meer, um unbeobachtet zu sein. Dies muss von den Schülerinnen und Schülern in einer eigenen, freien Antwort kurz schriftlich formuliert werden. Auch die Beispielimens auf den Kompetenzstufen IV und V erfordern eigene kurze Antworten. Zur Lösung des Beispielimens zu Kompetenzstufe IV müssen zentrale Stellen im Text interpretiert werden: Die Schülerinnen und Schüler können eine mittelbare oder unmittelbare Folge für die Figur Lin (Salz, Reichtum) beziehungsweise die Figurenkonstellation (z. B. Chen ist neidisch auf Lin) oder den Schauplatz („Das Meer wird salzig“) benennen. Auf Kompetenzstufe V erfordert das Beispielimens einen Bezug des Textes auf die lebensweltlichen Erfahrungen der Kinder. Hierbei wird von den Schülerinnen und Schülern eine begründete Stellungnahme erwartet, beispielsweise „Ja, weil ...“. Ob die Kinder der vorgegebenen Behauptung zustimmen oder nicht, gilt hierbei nicht als Kriterium für eine korrekte Antwort. Entscheidend ist, dass sie konsistent ausfällt.

2.2.3 Das Kompetenzstufenmodell für den Teilbereich *Orthografie*

Das Kompetenzstufenmodell für den Teilbereich *Orthografie* wurde im Jahr 2010 im Rahmen eines Standard-Settings unter Einbindung von Lehrkräften, Fachdidaktikerinnen und Fachdidaktikern sowie Expertinnen und Experten aus der Psychometrie entwickelt. Datengrundlage des Modells waren zum einen die Normierungsstudie im Fach Deutsch für den Primarbereich und andererseits die Pilotierung des IQB für die Vergleichsarbeiten (VERA-3) des Jahres 2010 (Testzeitpunkt 2009). Das Kompetenzstufenmodell wurde Ende des Jahres 2010 von der Kultusministerkonferenz zustimmend zur Kenntnis genommen. Im Rahmen des IQB-Ländervergleichs 2011 wurde für den Kompetenzbereich *Orthografie* zunächst nur die länderübergreifende Verteilung der Viertklässlerinnen und Viertklässler an allgemeinen Schulen (ohne Schülerinnen und Schüler an Förderschulen) auf die Kompetenzstufen untersucht. Dabei erfolgte zudem eine geringfügige redaktionelle Überarbeitung des Modells. Im

9 Bei der Aufgabe „Zwei Märchen“ handelt es sich um ein illustrierendes Beispiel für das Kompetenzstufenmodell im Teilbereich *Zuhören*. Sie wurde den Schülerinnen und Schülern im IQB-Bildungstrend 2016 nicht zur Bearbeitung vorgelegt.

IQB-Bildungstrend 2016 liegt der Testung zur orthografischen Kompetenz somit erstmalig eine auf Ebene der 16 Länder repräsentative Stichprobe der Viertklässlerinnen und Viertklässler des allgemeinbildenden Schulsystems – einschließlich Schülerinnen und Schülern an Förderschulen – zugrunde. Da diese Datengrundlage die Ausgangsbasis für zukünftige Trendschätzungen auf Länderebene darstellt, wurde die Kompetenzskala für den Kompetenzbereich *Orthografie* für die Berichterlegung zum IQB-Bildungstrend 2016 so angepasst, dass sie in der Population der Viertklässlerinnen und Viertklässler im Jahr 2016 einen Mittelwert von 500 Punkten und eine Standardabweichung von 100 Punkten aufweist (vgl. auch Kapitel 2.1).

Bei der Charakterisierung der Kompetenzstufen wird auf die in Böhme und Bremerich-Vos (2012) 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 häufig genutzt werden 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 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 (Optimalstandard)

Kompetenzstufe I (Punktwerte unter 390)¹⁰

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) 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 häu-

¹⁰ Anders als bei den Modellen für die Kompetenzbereiche *Lesen* und *Zuhören* 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.

fig 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 *vielen*). 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 4. Jahrgangsstufe zu erwartende Kompetenzminimum im Sinne des *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 Morphemangang 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 Falschreibungen 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 markieren, meistern sie auf dieser Kompetenzstufe im Wesentlichen. Schließlich gelingen ihnen einige Sortieraufgaben, 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 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, sodass 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*.

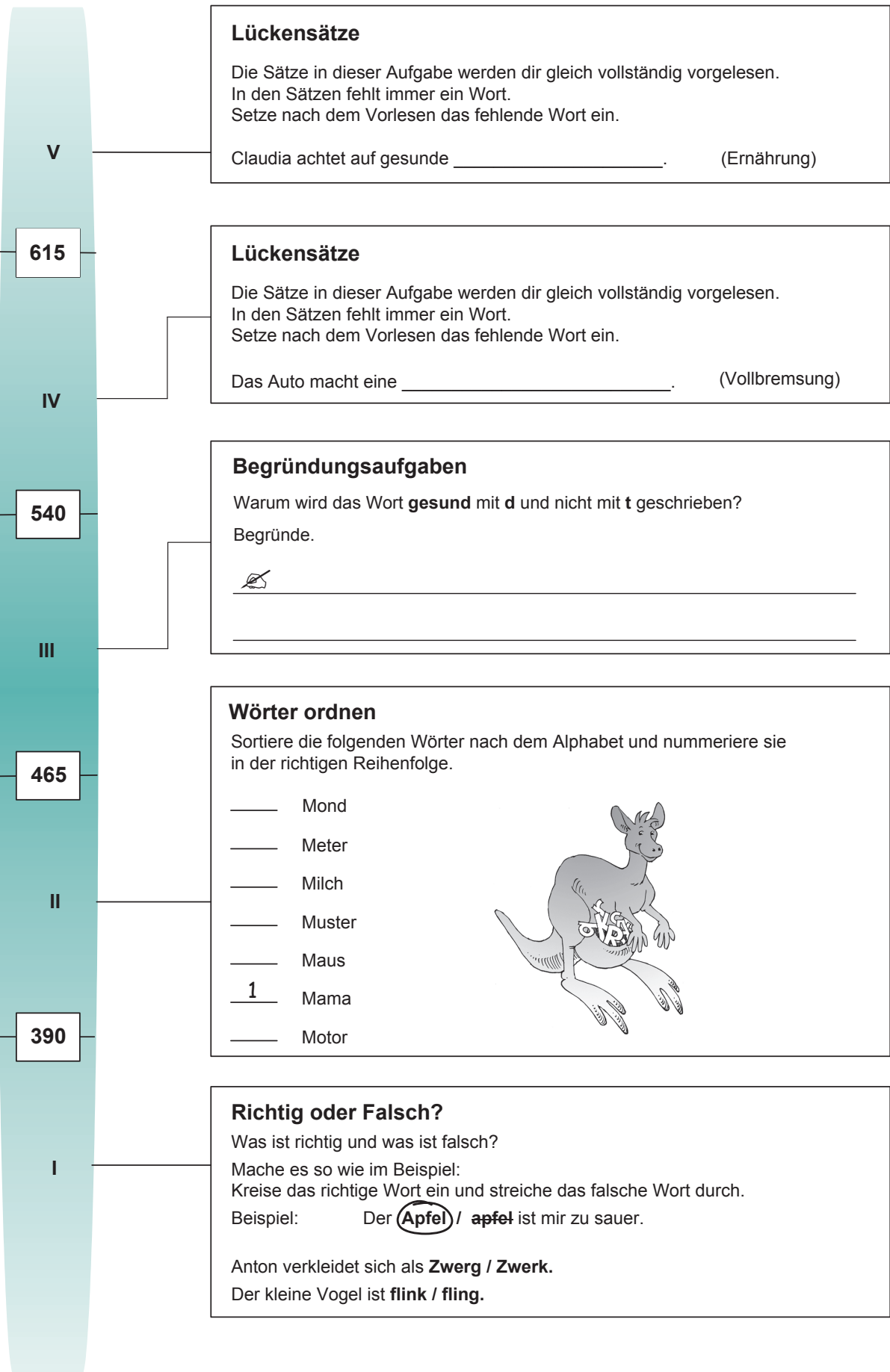
Die Kompetenzstufe V lässt sich als *Optimalstandard* interpretieren.

Zur Illustration der Kompetenzstufen I bis V im Teilbereich *Orthografie* finden sich nachfolgend in der Abbildung 2.7 Items, die beispielhaft die hier relevanten Kompetenzen und die in der Testung eingesetzten Aufgabenformate veranschaulichen. Die Beispiele stammen aus verschiedenen Studien, unter anderem aus den Vergleichsarbeiten in der 3. Jahrgangsstufe („VERA 3“).¹¹

Das abgedruckte Aufgabenbeispiel zu Kompetenzstufe I bezieht sich vor allem auf die Fehlerkategorie konsonantische Ableitung (KA) beziehungsweise auf die Auslautverhärtung: Um auf die richtige Schreibung des Konsonanten zu kommen, sind aus *Zwerg* und *flink* jeweils zweisilbige Wörter zu bilden (*Zwer-ge*, *flin-ke*). Bei der Sortieraufgabe, die Kompetenzstufe II illustriert, sind Wörter nach dem Alphabet zu ordnen. Solche Aufgaben dienen als Substitut für die Arbeit mit Wörterbüchern. Bei dieser Beispielaufgabe sind die jeweils an zweiter Stelle stehenden Grapheme der zu reihenden Wörter unterschiedlich, dies ist anspruchsvoller als Sortieraufträge zu Wörtern, die sich im ersten Graphem unterscheiden. Bei dem Aufgabenbeispiel für Kompetenzstufe III geht es für die Schülerinnen und Schüler darum, eine bestimmte Schreibung (von *gesund*) zu begründen, indem sie auf die Anwendung einer Strategie verweisen: Richtig sind Antworten, die auf eine Verlängerung des Wortes zielen, z. B. in Form von *gesunde*. Inhaltlich entspricht die Anwendung dieser Strategie der Anforderung, die auch bei den Schreibungen zu den konsonantischen Ableitungen auf Kompetenzstufe I angewendet werden muss. Hier muss diese Strategie aber nicht nur (implizit) zur Anwendung kommen, sondern auch aktiv und nachvollziehbar produziert werden. Kompetenzstufe IV wird durch eine Lückensatzaufgabe illustriert. Fehlerschwerpunkte sind hier auf die Markierung der Vokalkürze bezogen (*Volbremsung* statt *Vollbremsung*, Fehlerkategorie VK). Zudem spielt die Groß- und Kleinschreibung (GK) eine Rolle. Die erforderliche Großschreibung ist für die Schülerinnen und Schüler an dem Suffix *-ung* erkennbar (UM). Bei der Lückensatzaufgabe zu Kompetenzstufe V können Schülerinnen und Schüler

¹¹ Bei den dargestellten Items handelt es sich um illustrierende Beispiele für das Kompetenzstufenmodell im Teilbereich *Orthografie*. Sie wurden den Schülerinnen und Schülern im IQB-Bildungstrend 2016 nicht zur Bearbeitung vorgelegt.

Abbildung 2.7: Illustration des Kompetenzstufenmodells für den Teilbereich *Orthografie*



schließlich die Verwandtschaft von *Ernährung* mit *Nahrung* erkennen (VA) und bei diesem weniger geläufigen Abstraktum (GK) die Vokallänge durch ein Dehnungs-h markieren (VL).

Literatur

- Böhme, K. & Bremerich-Vos, A. (2012). Beschreibung der im Fach Deutsch untersuchten Kompetenzen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 19–33). Münster: Waxmann.
- KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.

2.3 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, 2005) liegen Kompetenzstufenmodelle zu den fünf inhaltlichen Kompetenzbereichen sowie für das Fach insgesamt (Globalmodell) vor.¹² Im Einklang mit dem in Kapitel 2.1 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 4. Jahrgangsstufe im Jahr 2011 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 2.8.

2.3.1 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.

¹² Die im Folgenden vorgestellten Beschreibungen der Modelle einschließlich weiterführender Erläuterungen sind auf der Webseite des IQB verfügbar unter <https://www.iqb.hu-berlin.de/bista/ksm>

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 4. Jahrgangsstufe formulierten Erwartungen. Sie besitzen jedoch bereits grundlegende Fähigkeiten, die über die erste Kompetenzstufe hinausgehen. Leistungen auf der zweiten Kompetenzstufe können 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, sodass 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

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 2.8 illustriert die fünf Kompetenzstufen des Globalmodells mit Beispielaufgaben aus unterschiedlichen inhaltlichen Bereichen.¹³

2.3.2 Zahlen und Operationen

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

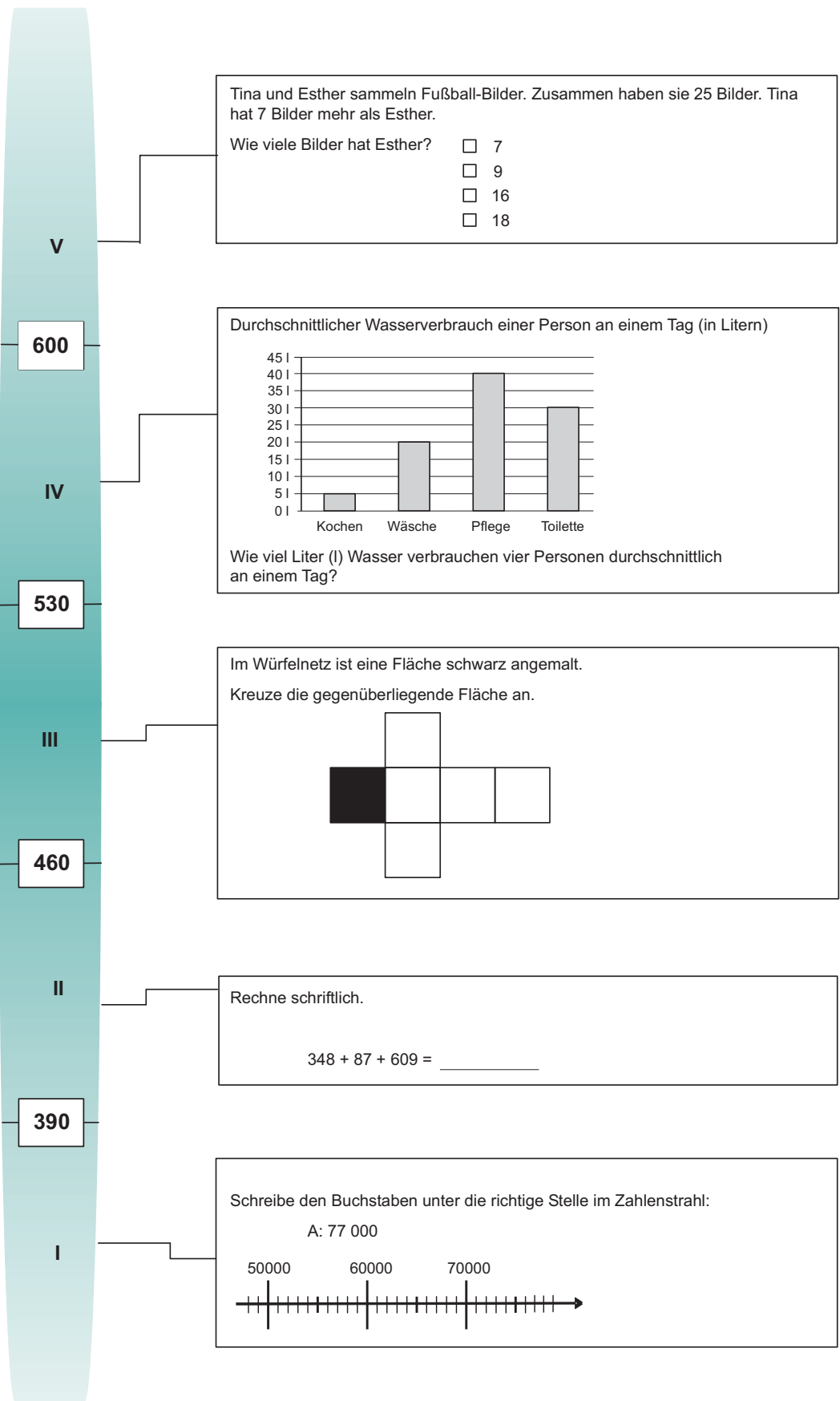
Auf dieser Kompetenzstufe des Bereichs *Zahlen und Operationen* sind die Grundlagen der Struktur des Dezimalsystems wie die Einteilung in Einer, Zehner, Hunderter usw. bekannt. Alle Grundaufgaben des kleinen Einspluseins und Einmaleins werden beherrscht. Kleinere Zahlen können halbschriftlich addiert 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 berücksich-

¹³ Für die Abschätzung von Trends ist es erforderlich, die Testaufgaben in mehreren aufeinanderfolgenden Bildungstrend-Studien unter vergleichbaren Bedingungen einzusetzen. Um dies zu gewährleisten, dürfen die Aufgaben nicht veröffentlicht werden. Bei den dargestellten Aufgaben handelt es sich daher lediglich um illustrierende Beispiele für das Kompetenzstufenmodell, die den Schülerinnen und Schülern im IQB-Bildungstrend 2016 jedoch nicht zur Bearbeitung vorgelegt wurden.

Abbildung 2.8: Globales Kompetenzstufenmodell und illustrierende Aufgaben im Fach Mathematik



Aus Platzgründen sind die Aufgaben in modifiziertem Layout dargestellt.

tigt. 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 in 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 syste-

matisches 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.

2.3.3 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.

2.3.4 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.

2.3.5 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.

2.3.6 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

KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.

Kapitel 3

Anlage, Durchführung und Auswertung des IQB-Bildungstrends 2016

3.1 Anlage und Durchführung

Camilla Rjosk, Maria Engelbert, Stefan Schipolowski
und Pauline Kohrt

Mit dem IQB-Bildungstrend 2016 wird das Ziel verfolgt, die von Schülerinnen und Schülern der 4. Jahrgangsstufe in den Ländern in der Bundesrepublik Deutschland erreichten Kompetenzen in den Fächern Deutsch und Mathematik zu erfassen. Insbesondere wird der Frage nachgegangen, welcher Anteil der Viertklässlerinnen und Viertklässler im Jahr 2016 die von der Kultusministerkonferenz (KMK, 2005a, 2005b) in den untersuchten Kompetenzbereichen verabschiedeten Regel- und Optimalstandards erreicht beziehungsweise die Mindeststandards verfehlt. Nach dem im Jahr 2011 durchgeführten IQB-Ländervergleich fand im Jahr 2016 zum zweiten Mal eine Überprüfung von Kompetenzen in der Primarstufe auf Basis der KMK-Bildungsstandards statt. Daher ist es erstmals möglich zu untersuchen, wie sich die Kompetenzen der Schülerinnen und Schüler über einen Zeitraum von fünf Jahren entwickelt haben. Einen weiteren Schwerpunkt des IQB-Bildungstrends 2016 bilden vertiefende Analysen zum Themenfeld „Umgang mit Heterogenität“ unter besonderer Berücksichtigung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) (vgl. Kapitel 10). Dieser Schwerpunkt wurde in Einklang mit dem in der überarbeiteten Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring beschriebenen Ziel gewählt, Schulleistungsstudien verstärkt zur Untersuchung von Schlüsselfragen der Schul- und Unterrichtsentwicklung zu nutzen (vgl. Kapitel 1.1).

Im vorliegenden Kapitel 3.1 werden Anlage und Durchführung des IQB-Bildungstrends 2016 vorgestellt. Dabei wird auch auf Aspekte eingegangen, die speziell die Gruppe der Schülerinnen und Schüler mit SPF betreffen. Zunächst erfolgt eine Beschreibung der eingesetzten Kompetenztests und Fragebogen, gefolgt von Erläuterungen zum Testdesign und zum Ablauf der Testungen an den Schulen. Im Anschluss wird auf die Definition der Zielpopulation, die Stichprobenziehung und die im IQB-Bildungstrend 2016 realisierte Stichprobe eingegangen. Das Kapitel schließt mit einer Übersicht aller an der Studie beteiligten Personen und Institutionen.

3.1.1 Kompetenztests

Im IQB-Bildungstrend 2016 wurden Tests zur Erfassung von in den Bildungsstandards für die Primarstufe beschriebenen Kompetenzen in den Fächern Deutsch und Mathematik sowie kognitiver Grundfähigkeiten eingesetzt. Um Trendaussagen zu ermöglichen, waren die in den Fächern Deutsch und Mathematik verwendeten Aufgaben zu einem Großteil mit den Aufgaben identisch, die bereits im IQB-Ländervergleich 2011 zum Einsatz gekommen waren. Alle Aufgaben wurden von erfahrenen Lehrkräften in Zusammenarbeit mit Wissenschaftlerinnen und Wissenschaftlern entwickelt und vor ihrem Einsatz im IQB-Bildungstrend 2016 in umfangreichen Erhebungen empirisch erprobt und normiert (Granzer et al., 2009).

Im Fach Deutsch wurden Kompetenzen im Bereich „Lesen – mit Texten und Medien umgehen“ (kurz: *Lesen*), im Teilbereich *Zuhören* des Bereichs „Sprechen und Zuhören“ sowie im Teilbereich *Orthografie* des Bereichs „Schreiben“ erfasst (vgl. Kapitel 1.2 und 2.2). Die Testaufgaben im Bereich *Lesen* ermöglichen es zu prüfen, ob Schülerinnen und Schüler die Inhalte von altersgemäßen Texten verstehen. Zur Erfassung dieser Kompetenz wurden den Schülerinnen und Schülern Lesetexte (Stimuli) von einer halben bis anderthalb Seiten vorgelegt, zu denen unterschiedlich komplexe Fragen zu beantworten waren (sog. Teilaufgaben oder Items). Als Lesetexte wurden sowohl kontinuierliche literarische Texte als auch kontinuierliche Sachtexte, zum Beispiel Sachbeschreibungen, und diskontinuierliche Sachtexte, zum Beispiel Tabellen und Diagramme, eingesetzt. Mit den Testaufgaben im Teilbereich *Zuhören* wurde überprüft, inwieweit Kinder auditiv präsentierte Inhalte verstehen können. Das Stimulusmaterial bestand unter anderem aus zwei bis neun Minuten langen Hörspielen oder Radiosendungen für Kinder, zu denen wiederum jeweils mehrere Items zu bearbeiten waren. Die Hörtexte wurden in der Testsituation von einer CD abgespielt. Mit den Aufgaben im Teilbereich *Orthografie* schließlich wurde erfasst, inwieweit die Kinder die korrekte Schreibweise von Wörtern beherrschen und ob sie Strategien kennen, um die Schreibweise von Wörtern abzuleiten. Die Items im Fach Deutsch umfassten unterschiedliche Itemformate, darunter geschlossene Formate, wie Mehrfachwahlaufgaben (*Multiple-Choice*) oder Zuordnungsaufgaben, sowie offene Formate, die schriftliche Kurzantworten auf Fragen oder das Ausfüllen von Lückensatzdiktaten erforderten. Aufgabenbeispiele für das Fach Deutsch sind in Kapitel 2.2 zu finden.

Im Fach Mathematik wurden Kompetenzen in den fünf Bereichen (1) *Zahlen und Operationen*, (2) *Raum und Form*, (3) *Muster und Strukturen*, (4) *Größen und Messen* sowie (5) *Daten, Häufigkeit und Wahrscheinlichkeit* erfasst (vgl. Kapitel 1.3 und 2.3). Das Stimulusmaterial bestand aus kurzen Anweisungen oder Problemstellungen, die bis zu fünf Zeilen umfassten. Abhängig vom Kompetenzbereich folgten jeweils unterschiedliche Aufgabenarten. Im Bereich *Zahlen und Operationen* erforderten Aufgaben unter anderem die Anwendung von Grundrechenarten. Im Bereich *Raum und Form* ging es zum Beispiel darum, grafische Muster zu vervollständigen oder Figuren zu spiegeln und im Bereich *Muster und Strukturen* sollten die Kinder unter anderem Rechenkettten vervollständigen oder mit Stellenwerttafeln umgehen. Aufgaben zum Bereich *Größen und Messen* erforderten das Einschätzen, Umwandeln und Zuordnen verschiedener Maßeinheiten und Größen und bei Aufgaben des Bereichs *Daten, Häufigkeit und Wahrscheinlichkeit* schließlich sollten die Kinder die Wahrscheinlichkeit von Ereignissen einschätzen sowie Informationen aus Diagrammen entnehmen und

interpretieren. Auch im Fach Mathematik kamen geschlossene (beispielsweise Multiple-Choice) und offene Aufgabenformate zum Einsatz. Beispielaufgaben für das Fach Mathematik sind Kapitel 2.3 zu entnehmen.

Zur Erfassung der *kognitiven Grundfähigkeiten* wurden zwei Tests eingesetzt. Zum einen der Subtest „Wortschatz“ des Kognitiven Fähigkeitstests (KFT 4–12+ R, Heller & Perleth, 2000) mit der Anforderung, zu vorgegebenen Zielwörtern je ein Synonym oder einen passenden Oberbegriff aus fünf möglichen Wörtern auszuwählen. Zum anderen wurden figurale (nonverbale) Aufgaben zum schlussfolgernden Denken vorgelegt (Schroeders, Schipolowski & Wilhelm, 2017; vgl. auch Wilhelm, Schroeders & Schipolowski, 2014). Dabei sollten die Kinder bei einer Reihe von Zeichnungen, die sich nach bestimmten Regeln ändern, die jeweiligen Regeln erkennen und angeben, wie die nächsten beiden Zeichnungen in der Reihe aussehen müssen.

Schülerinnen und Schüler mit SPF erhielten sowohl in den Fächern Deutsch und Mathematik als auch zur Erfassung der nonverbalen kognitiven Grundfähigkeiten im Schwierigkeitsniveau angepasstes Testmaterial. Ein Großteil der Aufgaben zur Prüfung der Kompetenzen in den Fächern Deutsch und Mathematik wurde in Kooperation mit Expertinnen und Experten der Inklusionspädagogik und der Fachdidaktiken eigens für diese Schülergruppe entwickelt und im Vorfeld im Rahmen einer Pilotierungsstudie erprobt. In Kapitel 13.2 wird dieser Aufgabenentwicklungsprozess beschrieben, Beispielimis vorgestellt und die empirische Güte der neu entwickelten Aufgaben sowie ihre Passung zu den Aufgaben aus dem IQB-Itempool erläutert. Die den Schülerinnen und Schülern mit SPF vorgelegte Testversion zu nonverbalen kognitiven Grundfähigkeiten bestand aus einer Auswahl besonders leichter Items, die ebenfalls im Vorfeld in einer Pilotierungsstudie erprobt wurden.

3.1.2 Fragebogen

Zusätzlich zu den Kompetenztests wurden verschiedene Fragebogen eingesetzt, die sich an Schülerinnen und Schüler, deren Eltern und Lehrkräfte sowie die Schulleitungen richteten. Dabei wurde die Anonymität der Befragten gewährleistet. In den Fragebogen ging es primär um die Erfassung von Hintergrundmerkmalen der Schülerinnen und Schüler, die für Analysen von geschlechtsbezogenen, sozialen und zuwanderungsbezogenen Disparitäten benötigt werden, und um die Erfassung von Informationen zu schulischen und außerschulischen Lerngelegenheiten. Die Inhalte wurden im Vorfeld von den Kultusministerien der Länder datenschutzrechtlich geprüft und genehmigt.

Zusätzlich zu den eingesetzten Fragebogen wurde vor der Testung erfragt, ob Schülerinnen und Schüler einen dokumentierten SPF aufwiesen und daher angepasstes Testmaterial erhalten sollten. In dieser Studie wurde länderübergreifend der Begriff „dokumentierter sonderpädagogischer Förderbedarf“ als Grundlage der Materialzuordnung genutzt. Dieser besteht, wenn in einem Feststellungsverfahren ein sonderpädagogischer Förderbedarf festgestellt wurde oder wenn die Schülerin beziehungsweise der Schüler infolge einer schulinternen Entscheidung auf angepasstem Anforderungsniveau beziehungsweise ziel-different unterrichtet wurde. Bei Schülerinnen und Schülern, die ausschließlich eine Teilleistungsschwäche (z. B. Lese-Rechtschreib-Schwäche/Legasthenie oder Rechenschwäche/Dyskalkulie) haben oder als hochbegabt gelten, liegt in diesem Sinne kein sonderpädagogischer Förderbedarf vor.

Im Gegensatz zur Teilnahme an den Kompetenztests, die an öffentlichen Schulen länderübergreifend für die Schülerinnen und Schüler verpflichtend war (vgl. Abschnitt 3.1.6), bestand nur in einigen Ländern eine allgemeine Verpflichtung zur Teilnahme an den Fragebogenerhebungen. Dafür maßgeblich sind die jeweiligen Landesschulgesetze, die Entscheidung des Schulträgers (öffentliche oder freie Trägerschaft) und die Datenschutzregelungen des Landes. Der Grad der Teilnahmeverpflichtung wird in Tabelle 3.1 für die unterschiedlichen Fragebogen dargestellt. Im Anschluss daran wird beschrieben, wie die einzelnen Fragebogeninstrumente administriert wurden und welche Inhalte sie jeweils umfassten.

Tabelle 3.1: Verpflichtungsgrad der Teilnahme an den Fragebogenerhebungen nach Land und Schulträgerschaft

Land	Fragebogen für Schülerinnen und Schüler ¹		Fragebogen für Lehrkräfte ² und Schulleitungen ²	
	öffentliche Schulen	Schulen in freier Trägerschaft	öffentliche Schulen	Schulen in freier Trägerschaft
Baden-Württemberg	○	○	○	○
Bayern	○	○	○	○
Berlin	◐ ³	○ ⁴	◐	○
Brandenburg	●	○	●	○
Bremen	●	●	◐	◐
Hamburg	○	○	◐ ⁵	◐ ⁵
Hessen	●	○	◐	○
Mecklenburg-Vorpommern	●	●	●	●
Niedersachsen	●	○	●	○
Nordrhein-Westfalen	○	○	●	○
Rheinland-Pfalz	○	○	◐	○
Saarland	○	○	◐	○
Sachsen	○	○	○	○
Sachsen-Anhalt	●	●	●	●
Schleswig-Holstein	○	○	○	○
Thüringen	●	Entscheidung Schule/Träger	●	Entscheidung Schule/Träger

● Verpflichtende Teilnahme

◐ Teilweise verpflichtende Teilnahme

○ Freiwillige Teilnahme

Anmerkungen. ¹ Bei freiwilliger Teilnahme zusätzlich Elterneinverständnis erforderlich. ² Bei teilweise verpflichtender Teilnahme schul- und unterrichtsbezogene Angaben verpflichtend, persönliche Angaben freiwillig. ³ Verpflichtend bis auf Fragen über Dritte. Die Beantwortung der Fragen über Dritte war freiwillig und erforderte eine Einverständniserklärung der Eltern. ⁴ Die Beantwortung von Fragen über Dritte erforderte eine zusätzliche Einverständniserklärung der Eltern. ⁵ Lehrkräftefragebogen freiwillig, Schulleitungsfragebogen teilverpflichtend.

Fragebogen für Schülerinnen und Schüler

Der Fragebogen für Schülerinnen und Schüler lag in drei verschiedenen Versionen vor, die in den Fragen zur Person und zum häuslichen Umfeld identisch waren, sich jedoch in weiterführenden Fragen zu schulischen Lerngelegenheiten und Merkmalen der Kinder, beispielsweise ihren fachbezogenen Interessen, unterschieden. Die Verwendung verschiedener Fragebogenversionen ermöglicht es, die zeitliche Belastung durch die Befragung gering zu halten und gleichzeitig Informationen zu einer größeren Anzahl relevanter Fragestellungen zu erhalten. Schülerinnen und Schüler mit SPF erhielten eine vierte, im Umfang reduzierte Fragebogenversion, bei der in besonderer Weise auf die sprachliche Verständlichkeit geachtet wurde. Ein Großteil der Fragen dieser Fragebogenversion wurde im Vorfeld im Rahmen einer Pilotierungsstudie mit Schülerinnen und Schülern mit SPF erprobt.

Insgesamt enthielten die Fragebogen Fragen zur Person (z.B. Geschlecht, Alter und Geburtsland), zum häuslichen Umfeld (z.B. Geburtsland der Eltern und Sprachgebrauch im Elternhaus), zu schulischen und außerschulischen Lernbedingungen (z.B. Merkmale des Unterrichts und des Lernklimas), zu schulischen Merkmalen der Schülerinnen und Schüler (z.B. fachbezogenes Interesse und Lernfreude), zur sozialen Eingebundenheit in der Klasse sowie zum Verhalten gegenüber den Mitschülerinnen und Mitschülern.

Sofern die Datenschutzbestimmungen in den Ländern eine Einverständniserklärung der Eltern zur Beantwortung des Fragebogens vorsahen (vgl. Tab. 3.1), wurde diese entsprechend eingeholt. Lag keine Einverständniserklärung der Eltern vor, wurde den entsprechenden Schülerinnen und Schülern der Fragebogen nicht zur Beantwortung vorgelegt beziehungsweise der einverständnispflichtige Fragebogenteil gestrichen. An Schulen in öffentlicher Trägerschaft in den Ländern Berlin, Brandenburg, Bremen, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt und Thüringen war das Ausfüllen des Fragebogens für Schülerinnen und Schüler verpflichtend, wobei in Berlin Fragen über Dritte, wie etwa zum Geburtsland der Eltern, von der Verpflichtung ausgenommen waren. In den übrigen Ländern war die Teilnahme freiwillig und erforderte das Einverständnis der Eltern.

Alle Kinder ohne SPF einer Klasse erhielten dieselbe Fragebogenversion, wobei die Versionen zufällig auf die Klassen verteilt wurden. Alle Schülerinnen und Schüler mit SPF an allgemeinen Schulen¹ und an Förderschulen erhielten die im Umfang angepasste Fragebogenversion für diese Schülergruppe. Die Bearbeitung des Fragebogens erfolgte am ersten von zwei Testtagen am Ende der Testsitzung im Klassenraum (vgl. Abschnitt 3.1.4).

Fragebogen für Eltern

Der Fragebogen für Eltern wurde den Schülerinnen und Schülern am ersten Testtag zusammen mit einem Begleitschreiben nach Hause mitgegeben. Die ausgefüllten Fragebogen wurden in verschlossenen Umschlägen in der Schule wieder eingesammelt. Der Bogen umfasste unter anderem Fragen zur Person (z.B. Geburtsland und Beruf der Eltern), zum häuslichen Umfeld (z. B. Sprachgebrauch

¹ Mit dem Begriff „allgemeine Schulen“ werden hier und in den folgenden Kapiteln alle allgemeinbildenden Schularten ohne die Förderschulen bezeichnet (vgl. Definitionenkatalog der KMK zur Schulstatistik; KMK, 2015).

in der Familie und Anzahl vorhandener Bücher), zur Schule und zur Nutzung von Förderangeboten (z.B. Teilnahme an schulischen Angeboten, wie etwa Zusatzunterricht für leistungsstarke Kinder oder Förderunterricht in bestimmten Fächern, und an außerschulischen Aktivitäten) sowie zum Bildungsweg und gegebenenfalls zum sonderpädagogischen Förderbedarf des Kindes. Die Teilnahme an der Elternbefragung war in allen Ländern freiwillig.

Fragebogen für Lehrkräfte

Um zentrale Rahmenbedingungen schulischer Bildungsprozesse beschreiben zu können, wurden die Lehrkräfte, die in den teilnehmenden Klassen die Fächer Deutsch und Mathematik im Schuljahr 2015/2016 unterrichteten, ebenfalls befragt. Dabei wurden unter anderem Fragen zur Person (z.B. Geschlecht und Alter), zur beruflichen Qualifikation (z.B. Lehramtsstudium und Studienfach), zu Unterrichtszielen und -gestaltung (z.B. Klassenführung und Methodenvielfalt im Unterricht) sowie zur Beschulung von Kindern mit SPF (z.B. Unterrichtsorganisation und spezifische Fördermaßnahmen) gestellt. Es kamen zwei Fragebogenversionen zum Einsatz: eine für Lehrkräfte an allgemeinen Schulen und eine für Lehrkräfte an Förderschulen.

Auch für die Teilnahme an der Lehrkräftebefragung unterschied sich der Grad der Verpflichtung zwischen den Ländern (vgl. Tab. 3.1). In Baden-Württemberg, Bayern, Hamburg, Sachsen und Schleswig-Holstein war die Teilnahme freiwillig, an öffentlichen Schulen in Brandenburg, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Sachsen-Anhalt und Thüringen hingegen vollständig verpflichtend. In den übrigen Ländern bestand zur Beantwortung schul- und unterrichtsbezogener Fragen an öffentlichen Schulen eine Verpflichtung, nicht aber zur Beantwortung personenbezogener Fragen.

Fragebogen für Schulleitungen

Auch die Schulleiterinnen und Schulleiter der teilnehmenden Schulen wurden im Rahmen des IQB-Bildungstrends 2016 schriftlich befragt, um die in den Schulen vorhandenen Unterstützungsangebote sowie schulische Rahmenbedingungen zu erfassen. Die Fragen bezogen sich einerseits auf Angaben zur Person, beispielsweise Geschlecht und Beschäftigungsdauer als Schulleitung, und andererseits auf Merkmale der Schule und Umgebung, wie etwa Trägerschaft, Größe der Schule, Ganztagsangebot, spezifische Förderangebote, pädagogisches Personal der Schule, Zusammensetzung der Schülerschaft und Beschulung von Schülerinnen und Schülern mit SPF. Auch hier kamen zwei Fragebogenversionen zum Einsatz: eine für Schulleitungen von allgemeinen Schulen und eine für Schulleitungen von Förderschulen.

Der Verpflichtungsgrad zur Teilnahme an der Schulleitungsbefragung entsprach in allen Ländern den Regelungen für die Lehrkräftebefragung, mit Ausnahme von Hamburg. Hier waren für Schulleitungen schulbezogene Angaben verpflichtend, personenbezogene Angaben jedoch freiwillig.

3.1.3 Testdesign

Im IQB-Bildungstrend 2016 wurde – analog zum IQB-Ländervergleich 2011 – eine große Zahl von Aufgaben eingesetzt, um alle Facetten der getesteten Kompetenzbereiche hinreichend abzudecken (vgl. Richter et al., 2012). Da die Gesamtheit der Aufgaben nicht von allen Schülerinnen und Schülern bearbeitet werden kann, wurden sie nach einem Multi-Matrix-Design (vgl. Gonzalez & Rutkowski, 2010) auf verschiedene Testhefte verteilt. Die Zusammenstellung der Testhefte erfolgte in einem mehrstufigen Verfahren, bei dem zunächst die Aufgaben Aufgabenblöcken und anschließend die Aufgabenblöcke Testheften zugeordnet wurden.

Aufgabenblöcke

Die Aufgabenblöcke enthielten jeweils Aufgaben zu nur einem Kompetenzbereich (z.B. *Lesen* im Fach Deutsch). Der Umfang des Aufgabenmaterials wurde so gewählt, dass die Aufgabenblöcke in Deutsch aufgrund der umfangreicheren Stimuli (Lese- bzw. Hörtexte) jeweils 20 Minuten und in Mathematik jeweils 10 Minuten Bearbeitungszeit erforderten. Dabei wurde darauf geachtet, dass die Kinder die Aufgaben in der Regel innerhalb der gegebenen Zeit vollständig bearbeiten konnten.

Die Zusammenstellung der Aufgabenblöcke und deren Verteilung auf verschiedene Testhefte (siehe folgender Abschnitt) zielten auf zuverlässige und valide Schätzungen (a) der Schülerkompetenzen im Jahr 2016, (b) von Veränderungen der Schülerkompetenzen zwischen den Jahren 2011 und 2016 (Trendaussagen) sowie (c) der Kompetenzen von Schülerinnen und Schülern mit und ohne SPF auf einer gemeinsamen Skala ab (vgl. Abschnitt 3.1.1 und Kapitel 13.2).

Zur zuverlässigen und validen Schätzung der Schülerkompetenzen wurde darauf geachtet, die verschiedenen Kompetenzbereiche durch ein breites Aufgabenspektrum und eine größere Anzahl von Aufgabenblöcken abzudecken. Um gesicherte Aussagen über Veränderungen der Schülerkompetenzen zwischen den Jahren 2011 und 2016 machen zu können, wurde ein Teil der Aufgabenblöcke aus dem IQB-Ländervergleich 2011 in identischer Form eingesetzt. Die für Schülerinnen und Schüler mit SPF im Schwierigkeitsniveau angepassten Testaufgaben (vgl. Abschnitt 3.1.1 und Kapitel 13.2) wurden spezifischen Aufgabenblöcken zugeordnet. Ein Überblick über die Anzahl der eingesetzten Aufgaben und Aufgabenblöcke findet sich in Kapitel 13.1.

Testhefte

Im nächsten Schritt wurden die Aufgabenblöcke zu Testheften zusammengestellt. Im Fach Deutsch wurden mit allen Testheften mindestens zwei Kompetenzteilbereiche geprüft, wobei in fast allen Testheften Blöcke zum *Lesen* und *Zuhören* enthalten waren. In Mathematik zielte die Verteilung der Blöcke darauf ab, mit jedem Heft drei bis fünf Kompetenzbereiche zu testen. Gemäß dem Multi-Matrix-Design wurden die Blöcke so auf die Testhefte verteilt, dass alle Aufgaben zu einem Fach über die verschiedenen Testhefte bei der späteren Auswertung miteinander verknüpft werden konnten, darunter auch

Aufgabenblöcke des IQB-Ländervergleichs 2011 sowie Aufgabenblöcke für Schülerinnen und Schüler mit und ohne SPF. So wurde sichergestellt, dass sich die Kompetenzen der Schülerinnen und Schüler auf einer einheitlichen Skala abbilden lassen, auch wenn sie unterschiedliche Testhefte und Aufgaben bearbeitet hatten. Dafür wurde unter anderem ein Teil der Aufgabenblöcke (im Fach Deutsch) beziehungsweise einzelne Aufgaben (im Fach Mathematik) sowohl bei Schülerinnen und Schülern ohne SPF als auch bei Schülerinnen und Schülern mit SPF eingesetzt.

Im Fach Deutsch gab es für die Zuordnung von Aufgabenblöcken zu Testheften drei Testdesigns: ein Teildesign für Schülerinnen und Schüler ohne SPF, ein Teildesign für Schülerinnen und Schüler mit SPF an allgemeinen Schulen und ein Teildesign für Schülerinnen und Schüler an Förderschulen. Im Fach Mathematik gab es zwei Teildesigns: eins für Schülerinnen und Schüler ohne SPF und eins für Schülerinnen und Schüler mit SPF, unabhängig von der besuchten Schulart. Weiterführende Erläuterungen zu methodischen Aspekten des Testdesigns, zur Anzahl der Testhefte pro Teildesign und zu Überschneidungen zwischen den Teildesigns sind in Kapitel 13.1 zu finden.

Insgesamt kamen im Fach Deutsch 76 Testhefte zum Einsatz. Davon waren 12 identisch mit den im IQB-Ländervergleich 2011 eingesetzten Testheften. Im Fach Mathematik gab es insgesamt 63 Testhefte, von denen 11 identisch mit den im IQB-Ländervergleich 2011 eingesetzten Testheften waren.

Die Verteilung der Testhefte auf die einzelnen Klassen (im Fach Deutsch) beziehungsweise die einzelnen Schülerinnen und Schüler einer Klasse (im Fach Mathematik) erfolgte im Rahmen der Teildesigns per Zufall. Für das Fach Deutsch musste darauf geachtet werden, dass Zuhöraufgaben und bestimmte Aufgaben zum Teilbereich *Orthografie* (Lückendiktate) von allen Kindern (mit SPF und ohne SPF) einer Klasse gleichzeitig bearbeitet werden konnten, da Zuhöraufgaben mit einem CD-Spieler abgespielt und einige Orthografieaufgaben diktiert wurden. Im Fach Mathematik erfolgte eine zufällige Zuteilung von Testheften zu Schülerinnen und Schülern sowohl innerhalb als auch zwischen Klassen mit der einzigen Bedingung, dass Kinder mit SPF eine für diese Testgruppe erstellte Testheftversion erhielten. Dies ist auch der Grund dafür, dass im Fach Deutsch drei Teildesigns zur stärkeren Differenzierung von Kompetenzniveaus, im Fach Mathematik hingegen nur zwei Teildesigns benötigt wurden.

3.1.4 Testablauf

Die Erhebungen zum IQB-Bildungstrend 2016 fanden im Zeitraum vom 7. Mai bis zum 15. Juli 2016 unter hoch standardisierten Bedingungen statt. Die Untersuchung in den Schulen erfolgte jeweils an zwei Testtagen, die in der Regel direkt aufeinanderfolgten. 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, sodass in jedem Land insgesamt etwa die Hälfte der Klassen mit Aufgaben im Fach Mathematik begann und die andere Hälfte mit Aufgaben im Fach Deutsch. Der Ablauf der Erhebungen, der an allgemeinen Schulen und Förderschulen identisch war, ist in Tabelle 3.2 dargestellt.

Tabelle 3.2: Ablauf des ersten und zweiten Testtages

Dauer in Minuten	Tätigkeit
15	Beginn der Testsitzung, Verteilung der Testhefte, Einweisung der Schülerinnen und Schüler
40	Bearbeitung der Testaufgaben (Deutsch oder Mathematik)
10	Pause
40	Bearbeitung der Testaufgaben (Deutsch oder Mathematik)
10	Pause
5	Einweisung zur Bearbeitung des Fragebogens für Schülerinnen und Schüler (1. Testtag) bzw. zu den kognitiven Grundfähigkeitstests (2. Testtag)
30	Bearbeitung des Fragebogens für Schülerinnen und Schüler (1. Testtag) bzw. der kognitiven Grundfähigkeitstests (2. Testtag)
5	Abschluss der Testsitzung, Einsammeln der Materialien

Der Testablauf entsprach dem Ablauf an allgemeinen Schulen im IQB-Ländervergleich 2011. Die Gesamtdauer der Testsitzung an einem Tag lag bei etwa zweieinhalb Stunden, wobei die reine Bearbeitungszeit für die Schülerinnen und Schüler insgesamt maximal 110 Minuten betrug. Die Schulen wurden gebeten, den Testbeginn nach Möglichkeit so einzurichten, dass in etwa die üblichen Pausenzeiten eingehalten werden konnten. Auch sollten die Schulen die Kinder frühzeitig über den Umstand informieren, dass es an den betreffenden Tagen zu zeitlichen Veränderungen gegenüber regulären Schultagen kommt.

Die Testungen führten schulexterne Testleiterinnen und Testleiter durch, die im Vorfeld von der *International Association for the Evaluation of Educational Achievement Hamburg (IEA Hamburg)* geschult und mit den Testmaterialien vertraut gemacht wurden (vgl. Abschnitt 3.1.7). Um einen standardisierten und über alle Testungen hinweg vergleichbaren Ablauf sicherzustellen, folgten die Testleiterinnen und Testleiter einem vorab detailliert festgelegten Ablaufplan und lasen den Schülerinnen und Schülern die einleitenden Instruktionen wortwörtlich vor. Sie erklärten den Schülerinnen und Schülern zunächst anhand eines Anleitungsposters, wie die verschiedenen Aufgabenformate zu bearbeiten sind, und demonstrierten dies mit Beispielen. Das Vorgehen beim Ausfüllen des Fragebogens für Schülerinnen und Schüler wurde ebenfalls anhand eines Posters durch die Testleiterin beziehungsweise den Testleiter erläutert. Ferner konnten die Testleiterinnen und Testleiter die Fragen des Fragebogens bei Bedarf laut vorlesen, um die Schülerinnen und Schüler bei der Bearbeitung zu unterstützen.²

Zur Sicherung einer hohen Teilnahmequote wurde, wie auch in internationalen Schulleistungsstudien üblich, eine Regelung zur Durchführung von Nachtests festgelegt. Die Erhebung in einer Schule wurde mit bei der Testung abwesenden Schülerinnen und Schülern nachgeholt, wenn sowohl mindestens 15 Prozent als auch mindestens drei der für den Test vorgesehenen Schülerinnen und Schüler einer Klasse an beiden Testtagen abwesend waren.

2 Die Entscheidung über das Vorlesen trafen die Testleiterinnen und Testleiter in Absprache mit der anwesenden Lehrkraft situationsabhängig, um sicherzustellen, dass alle Kinder die gestellten Fragen verstehen und zutreffend beantworten konnten.

3.1.5 Definition der Zielpopulation und Stichprobenziehung

Definition der Zielpopulation

Im IQB-Bildungstrend 2016 sollte die Population der Schülerinnen und Schüler der 4. Jahrgangsstufe innerhalb jedes Landes durch die getestete Stichprobe möglichst genau repräsentiert sein. Zur Vorbereitung der Stichprobenziehung im Sommer 2015 wurden daher als Grundlage zur präzisen Bestimmung der Zielpopulation Angaben der statistischen Landesämter für die 3. Jahrgangsstufe im Schuljahr 2013/2014 als Prognose der Schülerzahlen in der 4. Jahrgangsstufe herangezogen. In Einzelfällen musste auf Daten aus den Schuljahren 2012/2013 (im Saarland) beziehungsweise 2011/2012 (in Mecklenburg-Vorpommern) zurückgegriffen werden.

Tabelle 3.3 gibt einen Überblick über die Anzahl der Schülerinnen und Schüler pro Land und Schularart, die im Schuljahr der Erhebung (Schuljahr 2015/2016) nach Angaben der Fachserie 11 des Statistischen Bundesamts (2016) die 4. Jahrgangsstufe besuchten. Die Übersicht zeigt unter anderem, dass sich die Anteile der Schülerinnen und Schüler in den Förderschulen der Länder zum Teil deutlich unterscheiden; diese variieren zwischen 0,6 Prozent in Bremen sowie Schleswig-Holstein und 5,4 Prozent in Sachsen. Die Differenzen sind in erster Linie auf länderspezifische Regelungen zur Beschulung von Kindern mit SPF zurückzuführen (vgl. Kapitel 10): Während Kinder mit sonderpädagogischem Förderbedarf in einigen Ländern überwiegend allgemeine Schulen besuchen, werden sie in anderen Ländern mehrheitlich in Förderschulen unterrichtet. Um die Zielpopulation der Schülerinnen und Schüler als Voraussetzung für einen aussagekräftigen Ländervergleich einheitlich zu definieren, wurden in den IQB-Bildungstrend 2016 nicht nur allgemeine Schulen, sondern auch Förderschulen einbezogen. Auf diese Weise kann sichergestellt werden, dass Kinder mit sonderpädagogischem Förderbedarf in allen Ländern in vergleichbarem Maße repräsentiert sind und in vertiefenden Analysen zudem valide Aussagen über die Lernsituation von Schülerinnen und Schülern mit SPF in Deutschland gemacht werden können.

Die Zielpopulation des IQB-Bildungstrends 2016 umfasste somit die Schülerinnen und Schüler der 4. Jahrgangsstufe in Deutschland an allen Schularten. Nicht zur Zielpopulation gehörten lediglich Schülerinnen und Schüler mit SPF im Förderschwerpunkt „geistige Entwicklung“ sowie Schülerinnen und Schüler, die weniger als ein Jahr in deutscher Sprache unterrichtet wurden. Im nachfolgenden Schritt der Stichprobenziehung (siehe unten) konnten aus Gründen der Durchführbarkeit unter den gegebenen Testbedingungen weitere Personen – zum Beispiel Kinder mit dem Förderschwerpunkt „Sehen“, für die nach Einschätzung der Lehrkräfte ohne angepasstes Testmaterial keine Teilnahme möglich war – nicht in die Erhebung einbezogen werden. Dieses Vorgehen entspricht dem Standard großer internationaler Schulleistungsstudien, den früheren IQB-Ländervergleichsstudien und dem IQB-Bildungstrend 2015.

Bei Schülerinnen und Schülern mit SPF wurde für die Berichtlegung danach unterschieden, ob sie entsprechend dem Lehrplan beziehungsweise Rahmenlehrplan für allgemeine Schulen *zieltgleich* oder auf Grundlage eines abweichenden

Tabelle 3.3: Schülerinnen und Schüler der 4. Jahrgangsstufe im allgemeinbildenden Schulsystem nach Land und Schulart (prozentuale Angaben) im Schuljahr 2015/2016

Land	Insgesamt ^{1, 2}	GS	IGS ³	FrW	FS ⁴
Baden-Württemberg ⁵	97 205	82.4	11.7	1.9	4.0
Bayern	109 328	95.2	–	0.6	4.2
Berlin	27 888	91.3	5.4	1.3	2.0
Brandenburg	19 908	96.9	0.3	0.6	2.1
Bremen	5 464	97.9	–	1.5	0.6
Hamburg	15 026	88.8	7.1	1.9	2.2 ⁶
Hessen	53 559	94.6	1.6	0.7	3.0
Mecklenburg-Vorpommern	12 615	94.4	–	0.6	5.0 ⁶
Niedersachsen	74 212	96.1	0.4	0.7	2.8
Nordrhein-Westfalen	158 844	95.9	–	0.8	3.3 ⁶
Rheinland-Pfalz	34 001	96.8	–	0.6	2.6
Saarland	7 432	96.5	–	1.4	2.2 ⁶
Sachsen	32 157	94.2	–	0.5	5.4
Sachsen-Anhalt	16 443	95.5	–	0.4	4.1
Schleswig-Holstein	24 828	97.9	–	1.5	0.6
Thüringen	16 486	92.6	4.6	0.7	2.1
Deutschland	705 396	93.5	2.3	0.9	3.3

Anmerkungen. GS = Grundschulen; IGS = Integrierte Gesamtschulen; FrW = Freie Waldorfschulen; FS = Förderschulen. ¹ Absolute Häufigkeiten. ² Die Zahlen umfassen ausschließlich solche Schülerinnen und Schüler, die laut amtlicher Statistik eindeutig der 4. Jahrgangsstufe zugeordnet werden können. Kinder mit dem Förderschwerpunkt „geistige Entwicklung“, die der Grundstufe zugeordnet sind, bleiben unberücksichtigt. ³ In einigen Ländern existieren der amtlichen Statistik zufolge Integrierte Gesamtschulen, die auch die Jahrgangsstufen 1–4 umfassen. ⁴ Ohne Förderschulen mit dem Förderschwerpunkt „geistige Entwicklung“. ⁵ Förderschülerinnen und Förderschüler nach Jahrgangsstufe teilweise vom Statistischen Bundesamt geschätzt. ⁶ In vier Ländern weist die Fachserie 11 einen Teil der Schülerinnen und Schüler an Förderschulen nicht nach Jahrgangsstufen getrennt aus; diese Schülerinnen und Schüler sind hier nicht berücksichtigt. Der Anteil nicht nach Jahrgangsstufe ausgewiesener Schülerinnen und Schüler an allen Förderschülerinnen und -schülern beträgt im Saarland 44.7 %, in Nordrhein-Westfalen 5.8 %, in Hamburg 2.6 % und in Mecklenburg-Vorpommern weniger als 0.1 %.

Quelle: Statistisches Bundesamt (2016). Allgemeinbildende Schulen, Fachserie 11, Reihe 1 – Schuljahr 2015/2016.

den Lehrplans und somit *zielfifferent* unterrichtet wurden.³ Die Entscheidung, welche Schülerinnen und Schüler zielgleich oder zielfifferent unterrichtet werden, hängt eng mit den jeweiligen Förderschwerpunkten zusammen und wird von jedem Land nach eigenen Regelungen festgelegt. In den IQB-Bildungstrend wurden sowohl die zielgleich als auch die zielfifferent unterrichteten Kinder in die Zielpopulation einbezogen (vgl. Kapitel 4). Dies betrifft die Analysen, deren Ergebnisse in den Kapiteln 6 bis 13 berichtet werden. Die Darstellungen der Kompetenzstufenverteilungen in Kapitel 5 beziehen sich hingegen ausschließlich auf die zielgleich unterrichteten Schülerinnen und Schüler, da davon auszugehen ist, dass für zielfifferent unterrichtete Schülerinnen und Schüler andere Lernziele gelten als die in den Bildungsstandards definierten.

3 Zielgleich unterrichtete Schülerinnen und Schüler werden nach einem allgemeinen Lehrplan und somit entsprechend der Bildungsstandards unterrichtet. Die Lernziele zielfifferent unterrichteter Schülerinnen und Schüler sind hingegen nicht am allgemeingültigen Lehrplan ausgerichtet und basieren somit nicht zwangsläufig auf den Bildungsstandards.

Stichprobenziehung

Die Stichprobenziehung im Rahmen des IQB-Bildungstrends 2016 erfolgte, wie in großen Schulleistungsstudien üblich, in mehreren Schritten (Frey, Carstensen, Walter, Rönnebeck & Gomolka, 2008). In einem ersten Schritt wurden innerhalb jedes Landes alle vorhandenen Schulen einer von vier Ziehungsschichten zugeordnet, aus denen nachfolgend eine Zufallsstichprobe gezogen wurde: Förderschulen sowie allgemeine Schulen, wobei innerhalb der allgemeinen Schulen drei explizite Schichten differenziert wurden, die sich in der mittleren sozialen Zusammensetzung der Schülerschaft unterschieden. Bei Förderschulen beschränkte sich die Einbeziehung aus praktischen Gründen auf Schulen mit den Förderschwerpunkten „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“, bei denen davon auszugehen ist, dass die Schülerinnen und Schüler zur selbstständigen Bearbeitung des Tests grundsätzlich in der Lage sind. Schulen mit anderen Förderschwerpunkten („geistige Entwicklung“, „Hören“, „Sehen“, „körperliche und motorische Entwicklung“, „Unterricht kranker Schülerinnen und Schüler“ sowie „Unterricht von Kindern und Jugendlichen mit autistischem Verhalten“) wurden nicht in die Stichprobenziehung einbezogen. Dies führt zu einer geringfügigen, praktisch begründeten Einschränkung der Repräsentativität der Stichprobe für die Zielpopulation des IQB-Bildungstrends 2016.

Die Größe der Stichprobe pro Land wurde so gewählt, dass sie Aussagen mit einer vorab festgelegten Präzision (ermittelt anhand des sog. Stichproben- beziehungsweise Standardfehlers) ermöglicht. Dabei gilt als Grundregel: Je größer die Streuung der Schülerkompetenzen innerhalb eines Landes insgesamt und innerhalb von Schulen ist, desto größer muss die Anzahl der Schulen in der Stichprobe sein, um Kompetenzwerte präzise schätzen zu können. Die Anzahl der zu untersuchenden Schulen orientierte sich zum einen an der Kompetenzstreuung, die aufgrund von Daten des IQB-Ländervergleichs 2011 zu erwarten war. Diese war insbesondere in den Stadtstaaten Berlin, Hamburg und Bremen vergleichsweise groß (Böhme & Weirich, 2012; Haag & Roppelt, 2012), sodass in diesen Ländern eine im Verhältnis zur Populationsgröße relativ große Anzahl von Schulen gezogen werden musste. Zum anderen wurde berücksichtigt, dass Veränderungen des Schulsystems, wie beispielsweise durch Ausweitung der gemeinsamen Beschulung von Kindern mit und ohne sonderpädagogischen Förderbedarf an allgemeinen Schulen, zu Veränderungen in der Kompetenzstreuung innerhalb von Schulen führen können.

Insgesamt wurden 1559 Schulen für die Teilnahme am IQB-Bildungstrend 2016 ausgewählt. In den Flächenländern war eine Stichprobengröße von 80 bis 96, in den Stadtstaaten von 101 bis 132 allgemeinen Schulen vorgesehen. Für den Fall von Schulausfällen (vgl. Abschnitt 3.1.6) wurden jeder Schule zwei vergleichbare Ersatzschulen derselben Ziehungsschicht zugeordnet. Es wurde angestrebt, Förderschulen proportional zur Anzahl von Förderschülerinnen und Förderschülern im jeweiligen Land einzubeziehen, mindestens jedoch sechs Förderschulen pro Land, sofern vorhanden.

Im zweiten Schritt wurde innerhalb jeder Schule eine Klasse der 4. Jahrgangsstufe zufällig zur Teilnahme an der Untersuchung ausgewählt. In den ausgewählten Klassen sollten alle Schülerinnen und Schüler an der Testung teilnehmen (vgl. Abschnitt 3.1.6). An Schulen mit jahrgangsübergreifendem Lernen bezog sich die Zufallsauswahl nicht auf eine Klasse, sondern auf eine Lerngruppe. In die Testung wurden dann ausschließlich diejenigen Kinder einbezogen, die

aufgrund ihres Einschulungszeitpunkts der 4. Jahrgangsstufe zuzuordnen waren. An Förderschulen waren davon abweichend grundsätzlich alle Schülerinnen und Schüler der 4. Jahrgangsstufe mit mindestens einem der Förderschwerpunkte „Lernen“, „Sprache“ oder „emotionale und soziale Entwicklung“ in der jeweiligen Schule für die Teilnahme an der Erhebung vorgesehen.

In den Ländern Berlin, Bremen und Nordrhein-Westfalen wurde auf Wunsch der Kultusministerien zudem eine Erweiterungsstichprobe (sog. *Oversampling*) gezogen, um vertiefende Analysen zur Lernsituation von Schülerinnen und Schülern mit SPF in diesen Ländern, die über Fragestellungen dieses Berichtsbands hinausgehen, zu ermöglichen. In Berlin und in Nordrhein-Westfalen umfasste die Erweiterungsstichprobe jeweils 25 zusätzliche Schulen (20 allgemeine Schulen und 5 Förderschulen), in Bremen alle allgemeinen Schulen. In den allgemeinen Schulen der Erweiterungsstichprobe nahmen neben der gezogenen Klasse zusätzlich alle weiteren Schülerinnen und Schüler der 4. Jahrgangsstufe mit mindestens einem der oben genannten Förderschwerpunkte teil.⁴

3.1.6 Realisierte Stichprobe

Die Teilnahme am IQB-Bildungstrend 2016 war für alle ausgewählten Schulen in öffentlicher Trägerschaft verpflichtend. Schulen in freier Trägerschaft, wie beispielsweise Freie Waldorfschulen, hatten je nach landesspezifischer Gesetzgebung die Möglichkeit, die Teilnahme abzulehnen. Tabelle 3.4 gibt einen Überblick über die Anzahl der Schulen pro Land und Schulart, die an der Erhebung teilgenommen haben (realisierte Stichprobe).

Wenn an einer Schule ein reibungsloser Ablauf des Tests (z. B. aufgrund umfassender Renovierungsarbeiten) ausgeschlossen war oder eine Schule in freier Trägerschaft die Teilnahme ablehnte, wurde stattdessen eine Ersatzschule in die Stichprobe aufgenommen. Konnte hingegen eine gezogene Schule nicht an der Erhebung teilnehmen, weil sie zwischenzeitlich (nach Erstellung der für die Stichprobenziehung verwendeten Schullisten) geschlossen wurde, so erfolgte kein Ersatz, da die betroffenen Schulen nicht zur Zielpopulation gehörten.⁵ Fielen neben der ursprünglich gezogenen Schule auch beide Ersatzschulen aus, wurde ebenfalls nicht weiter ersetzt. Alle gezogenen Schulen wurden bis zum Januar 2016 über ihre Auswahl zur Teilnahme an der Studie informiert. Insgesamt nahmen einschließlich der Ersatzschulen 1 508 der 1 559 ausgewählten Schulen teil.

Die Teilnahme an den Kompetenztests zum IQB-Bildungstrend 2016 war, wie bereits oben erwähnt, bei öffentlichen Schulen sowohl auf Schulebene als auch auf Schülerebene verpflichtend. Nur an Schulen in freier Trägerschaft konnten die Eltern einer Teilnahme an der Testung in der Regel widersprechen. In Anlehnung an die Praxis der internationalen Schulleistungsstudien konnten zudem Schülerinnen und Schüler vom Test 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.

4 Die aufgrund der Erweiterungsstichprobe zusätzlich teilnehmenden Kinder mit sonderpädagogischem Förderbedarf führen durch die Verwendung von Fallgewichten in der Auswertung (vgl. Kapitel 3.2) zu keinen Verzerrungen der Ergebnisse.

5 Davon abweichend wurde bei Ausfall einer Förderschule in jedem Fall eine Ersatzschule herangezogen (sofern verfügbar), um für diese Schulen ausreichend große Fallzahlen zu erreichen.

Tabelle 3.4: Realisierte Schulstichprobe nach Land und Schulart

Land	Schülerinnen und Schüler insgesamt	Schulen insgesamt	Schulen nach Schulart			
			GS	IGS ¹	FrW	FS ²
Baden-Württemberg	1 555	87	76	2	0	9
Bayern	1 773	90	80	0	0	10
Berlin	3 048	146	130	6	0	10
Brandenburg	1 549	77	69	2	0	6
Bremen	1 859	94	94	0	0	0
Hamburg	2 325	118	112	0	0	6
Hessen	1 649	90	82	1	1	6
Mecklenburg-Vorpommern	1 842	92	88	0	0	4
Niedersachsen	1 549	84	80	0	0	4
Nordrhein-Westfalen	2 300	108	94	0	0	14
Rheinland-Pfalz	1 564	93	87	0	0	6
Saarland	1 750	88	82	0	0	6
Sachsen	1 796	89	83	0	0	6
Sachsen-Anhalt	1 381	78	72	0	0	6
Schleswig-Holstein	1 740	86	84	0	0	2
Thüringen	1 579	88	82	0	0	6
Deutschland	29 259	1 508	1 395	11	1	101

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

¹ In einigen Ländern existieren der amtlichen Statistik zufolge Integrierte Gesamtschulen, die auch die Jahrgangsstufen 1–4 umfassen.

² Förderschulen mit einem der Förderschwerpunkte „Lernen“, „Sprache“ oder „emotionale und soziale Entwicklung“.

2. Schülerinnen und Schüler, die nach professioneller Einschätzung der Schulleitung oder einer anderen qualifizierten Person des Kollegiums emotional oder geistig nicht in der Lage waren, den allgemeinen Anweisungen des Tests zu folgen, oder die durch die Testsituation in unzumutbarem Maße emotional belastet worden wären.
3. Schülerinnen und Schüler nichtdeutscher Herkunftssprache, die weniger als ein Jahr in deutscher Sprache unterrichtet wurden und die nicht in der Lage waren, Deutsch zu lesen oder zu sprechen.

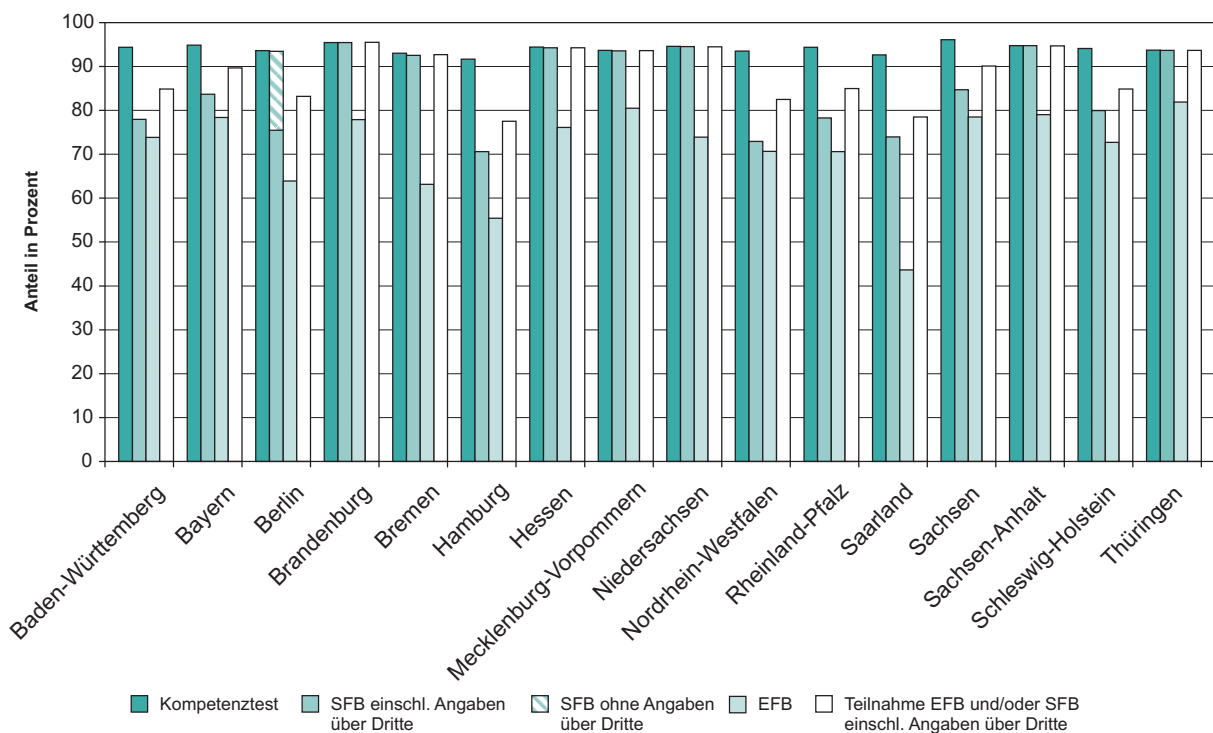
Insgesamt wurden 615 Schülerinnen und Schüler aus einem dieser Gründe von der Teilnahme am IQB-Bildungstrend 2016 ausgeschlossen, was einem Anteil von 2 Prozent aller für den Test ausgewählten Schülerinnen und Schüler entspricht.⁶ Die Entscheidung über den Ausschluss von Schülerinnen und Schülern nach diesen Kriterien traf die für den organisatorischen Ablauf an der Schule zuständige Lehrkraft (Schulkoordinatorin beziehungsweise Schulkoordinator) oder die Schulleitung.

6 Etwa die Hälfte dieser Schülerinnen und Schüler waren Kinder nichtdeutscher Herkunftssprache, die bislang weniger als ein Jahr in deutscher Sprache unterrichtet worden waren. Diese Schülerinnen und Schüler sind nach Angabe der Lehrkräfte zum Großteil im Jahr 2014 oder später als Geflüchtete nach Deutschland gekommen.

Die realisierte Gesamtstichprobe für den Test umfasst 29259 Schülerinnen und Schüler der 4. Jahrgangsstufe. Diese Zahl entspricht auf Schülerebene einer gewichteten Teilnahmequote⁷ von 94,3 Prozent. Damit liegt die Quote in etwa auf demselben Niveau wie in den zuletzt in Deutschland durchgeführten Grundschuluntersuchungen IGLU 2011 (95%, Tarelli, Wendt, Bos & Zylowski, 2012) und TIMSS 2015 (95%, Wendt et al., 2016), ist aber etwas niedriger als beim IQB-Ländervergleich 2011 (98,3%, Richter et al., 2012). Mit einer recht geringen Spannweite von 92 bis 96 Prozent ist die Beteiligungsquote an den Tests auch in den einzelnen Ländern insgesamt hoch (vgl. Abb. 3.1).

Im Rahmen der Datenerfassung und -auswertung waren vereinzelt weitere Fallausschlüsse erforderlich, etwa aufgrund einer Vertauschung von Testinstrumenten. Zudem wurden Schülerinnen und Schüler von der Auswertung ausgeschlossen, die trotz Anwesenheit am Testtag im jeweiligen Fach keine der Aufgaben bearbeitet hatten. Des Weiteren wurden Schülerinnen und Schüler aus den Analysen eines Fachs ausgeschlossen, wenn sie an einem der beiden Testtage abwesend waren und somit in diesem Fach keine Testaufgaben bearbeitet hatten. Insgesamt führten die genannten Gründe zum Ausschluss von 698 Kindern im Fach Deutsch und 742 Kindern im Fach Mathematik, was jeweils einem Anteil von etwa 2,5 Prozent der realisierten Gesamtstichprobe entspricht. Somit konnten im Fach Deutsch 28561 Schülerinnen und Schüler und im Fach Mathematik 28517 Schülerinnen und Schüler in die Analysen einbezogen werden.

Abbildung 3.1: Teilnahmequoten für den Kompetenztest und die Befragung der Schülerinnen und Schüler sowie Eltern nach Land (gewichtet) im IQB-Bildungstrend 2016



Anmerkung. SFB = Fragebogen für die Schülerinnen und Schüler; EFB = Elternfragebogen.

7 Zur Berechnung der Teilnahmequote werden in großen Schulleistungsstudien individuelle Schülergewichte verwendet. Die Gewichte geben an, wie viele Schülerinnen und Schüler der Population jede einzelne Schülerin beziehungsweise jeder einzelne Schüler in der Stichprobe repräsentiert. Die gewichtete Teilnahmequote kennzeichnet also, welcher Anteil der Population durch die Stichprobe abgebildet wird.

Neben der Beteiligung am Test ist für die Datenanalysen zum IQB-Bildungstrend 2016 eine hohe Beteiligungsquote der verschiedenen Befragungsinstrumente wichtig, die ebenfalls Abbildung 3.1 entnommen werden kann. In Deutschland insgesamt fällt für die Befragung der Schülerinnen und Schüler die Teilnahmequote mit 83.2 Prozent der Kinder, die für die Testung vorgesehen waren, niedriger aus als für den Kompetenztest. Diese ist zudem etwas geringer als die Teilnahmequote für die Befragung im IQB-Ländervergleich 2011 (87.3%). Im Gegensatz zur Quote für die Tests unterscheidet sich die Quote für die Befragung der Schülerinnen und Schüler zwischen den Ländern erheblich (vgl. Abb. 3.1). In den sieben Ländern, in denen die Teilnahme an der Befragung verpflichtend war (vgl. Tab. 3.1), sind die Beteiligungsquoten erwartungsgemäß besonders hoch (92.5% bis 95.5%). In den anderen acht Ländern, in denen die Teilnahme freiwillig und abhängig vom Einverständnis der Eltern war, liegt die Beteiligungsquote bei 70.6 Prozent (Hamburg) bis 84.7 Prozent (Sachsen). In Berlin war eine zusätzliche Einverständniserklärung der Eltern zu Fragebogenteilen erforderlich, die Angaben über Dritte (z. B. das Geburtsland der Eltern) enthielten. Dort liegen die Fragebogen inklusive Angaben über Dritte für 75.5 Prozent der Schülerinnen und Schüler vor.

Die Teilnahme an der Elternbefragung war in allen Ländern freiwillig. Die Quoten variieren zwischen 43.6 Prozent der Schülerinnen und Schüler, für die Angaben aus Elternfragebogen vorlagen, im Saarland⁸ und 81.9 Prozent in Thüringen. Für die in den Kapiteln 8 und 9 dargestellten Analysen zu sozialen und zuwanderungsbezogenen Disparitäten waren Angaben zum Beruf, zur Ausbildung und zum Geburtsland der Eltern der teilnehmenden Schülerinnen und Schüler erforderlich. Die Angaben zum Geburtsland der Eltern wurden sowohl von den Eltern als auch von den Schülerinnen und Schülern erfragt. Durch eine Kombination der Informationen von den Eltern und von den Schülerinnen und Schülern konnte in allen Ländern der Anteil fehlender Angaben zu den Geburtsländern reduziert werden (vgl. Abb. 3.1, Balken „Teilnahme EFB und/oder SFB“). Für die Analyse sozialer Disparitäten konnten hingegen ausschließlich Informationen des Elternfragebogens verwendet werden, da die Angaben zum Beruf der Eltern nicht von den Viertklässlerinnen und Viertklässlern erfragt werden konnten.

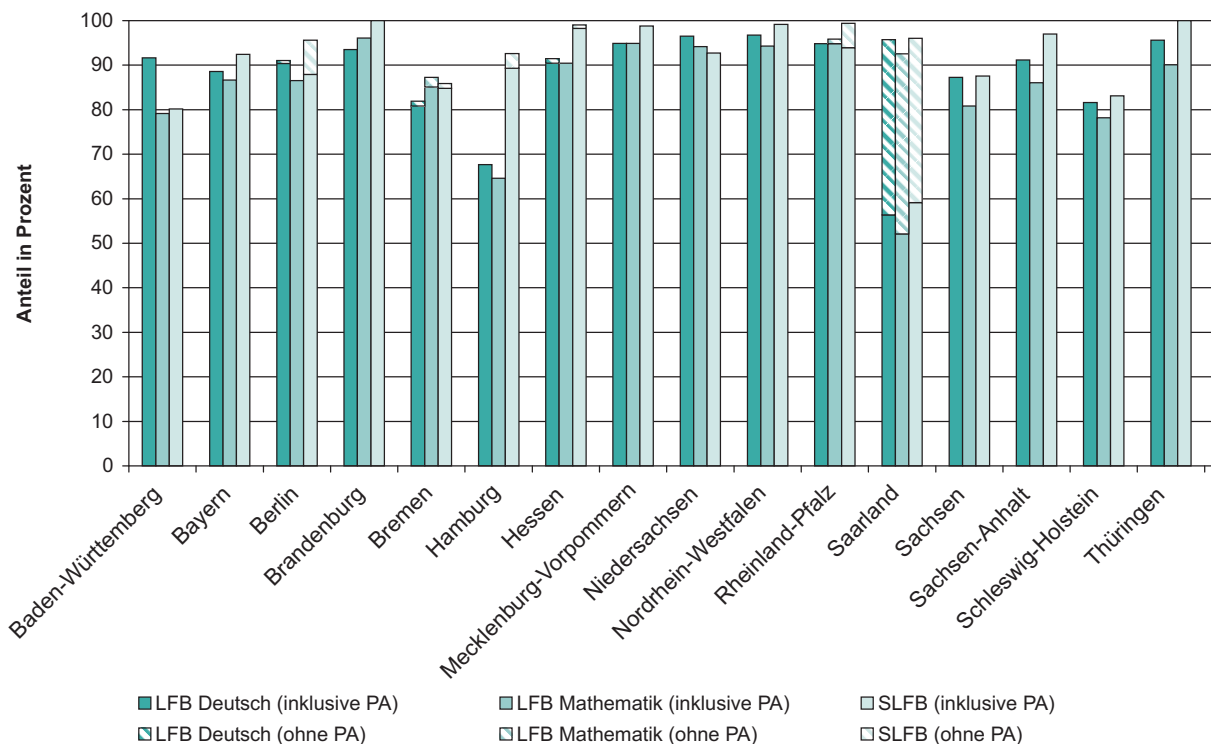
8 Im Saarland lagen Angaben der Eltern für 43.6 Prozent der Schülerinnen und Schüler vor. Weitere 20.9 Prozent der Eltern füllten den Elternfragebogen zwar aus, gaben jedoch nicht ihr Einverständnis für die Verwendung der gemachten Angaben in der Auswertung.

Anteil fehlender Angaben und Konsequenzen für den Ergebnisbericht

Der Anteil fehlender Angaben zu Hintergrundmerkmalen der Schülerinnen und Schüler und ihrer Familien hat direkte Konsequenzen für die Analysen im IQB-Bildungstrend 2016 und den Bericht von länderspezifischen Ergebnissen zu sozialen und zuwanderungsbezogenen Disparitäten. Bei einem hohen Anteil fehlender Werte kann die Zuverlässigkeit von Analysen eingeschränkt sein. Verzerrungen können vor allem dann entstehen, wenn Kinder, für die keine Angaben zu den betrachteten Hintergrundmerkmalen vorliegen, tendenziell geringere Kompetenzwerte aufweisen. Da dies im IQB-Bildungstrend 2016 der Fall ist, können die Ergebnisse zu sozialen und zuwanderungsbezogenen Disparitäten (Kapitel 8 und 9) für einige Länder nicht oder nur unter Vorbehalt berichtet werden. Gemäß einem von der Amtschefscommission „Qualitätssicherung in Schulen“ in ihrer 74. Sitzung beschlossenen Vorgehen werden Ergebnisse unter Vorbehalt berichtet, wenn die jeweiligen Angaben für weniger als 80 Prozent, jedoch für mindestens 70 Prozent der Schülerinnen und Schüler eines Landes vorliegen. Liegen für weniger als 70 Prozent der Schülerinnen und Schüler Angaben vor, so werden für die betreffenden Länder keine Ergebnisse berichtet. Dies hat für die Berichterlegung zu den sozialen und zuwanderungsbezogenen Disparitäten folgende Konsequenzen:

Ergebnisse zu den **sozialen Disparitäten** (Kapitel 8) werden für die Länder Baden-Württemberg, Bayern, Brandenburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen, Sachsen-Anhalt sowie Schleswig-Holstein *unter Vorbehalt* berichtet. *Keine Ergebnisse* zu sozialen Disparitäten werden für die Länder Berlin, Bremen, Hamburg und das Saarland berichtet.

Die Darstellung der Ergebnisse zu **zuwanderungsbezogenen Disparitäten** (Kapitel 9) erfolgt *unter Vorbehalt* für Hamburg und das Saarland. Da in allen Ländern für mindestens 70 Prozent der Schülerinnen und Schüler Informationen zum Zuwanderungshintergrund vorliegen, werden für alle Länder Ergebnisse berichtet.

Abbildung 3.2: Teilnahmequoten¹ für die Befragung der Lehrkräfte (ungewichtet) und Schulleitungen (gewichtet) nach Land im IQB-Bildungstrend 2016

Anmerkungen. LFB = Fragebogen für die Lehrkräfte; SLFB = Fragebogen für die Schulleitungen; PA = personenbezogene Angaben.

¹ Bei Ländern, in denen die schulbezogenen Angaben verpflichtend, personenbezogene Angaben jedoch freiwillig waren, wurden die entsprechenden Quoten getrennt ausgewiesen.

An der Lehrkräftebefragung nahmen insgesamt 2094 Lehrerinnen und Lehrer der Fächer Deutsch und Mathematik teil. Für das Fach Deutsch liegt aus 89.6 Prozent der Klassen, für das Fach Mathematik aus 86.9 Prozent der Klassen mindestens ein Lehrkräftefragebogen vor⁹. Diese Anteile sind mit den Quoten aus dem Jahr 2011 mit Angaben der Lehrkräfte für 88.5 Prozent der Klassen für das Fach Deutsch und 86.8 Prozent der Klassen für das Fach Mathematik vergleichbar. Etwas höher fiel mit 92.9 Prozent die Beteiligung an der Schulleitungsbefragung aus. Dieser Wert entspricht in etwa der Teilnahmequote, die im Rahmen des IQB-Ländervergleichs 2011 für die Schulleitungsbefragung erreicht wurde (94.3%). Die länderspezifischen Beteiligungsquoten für die Lehrkräfte- und Schulleitungsbefragungen unterscheiden sich deutlich (vgl. Abb. 3.2). Die geringste Beteiligung an der Befragung der Lehrkräfte weist Hamburg (Mathematiklehrkräfte: 64.6% der Klassen, Deutschlehrkräfte: 67.7% der Klassen) und an der Schulleitungsbefragung mit 80.2 Prozent Baden-Württemberg auf. Teilnahmequoten mit über 90 Prozent für die Lehrkräfte- und

9 Eine gewichtete Teilnahmequote kann nur bezogen auf Schülerinnen und Schüler sowie Schulleitungen, nicht aber für Lehrkräfte berichtet werden, da der Lehrkräftefragebogen allen Lehrerinnen und Lehrern vorgelegt wurde, die in den teilnehmenden Klassen Deutsch und/oder Mathematik unterrichteten. Dabei kam es vor, dass Schülerinnen und Schüler von mehreren Lehrkräften im Fach Deutsch oder Mathematik unterrichtet wurden oder Teilungsunterricht stattfand; zudem wurden an Förderschulen oft Schülerinnen und Schüler aus mehreren Klassen in die Testung einbezogen. Die Zahl der Lehrerinnen und Lehrer, die in die Befragung einzubeziehen war, variierte somit zwischen den Schulen. Dies hat zur Folge, dass die Zielpopulation nicht eindeutig definiert werden kann, was für die Berechnung einer gewichteten Teilnahmequote jedoch erforderlich wäre.

Schulleitungsbefragung finden sich in Brandenburg, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz und Thüringen sowie im Saarland bezogen auf den Fragebogenteil ohne personenbezogene Angaben.

3.1.7 Beteiligte Personen und Institutionen

Die wissenschaftliche Leitung des IQB-Bildungstrends 2016 lag beim wissenschaftlichen Vorstand des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin, Prof. Dr. Petra Stanat. Die Vorbereitung der Studie, die Aufbereitung und Analyse der Daten sowie die Berichterlegung im IQB erfolgte durch die folgenden Mitarbeiterinnen und Mitarbeiter (in alphabetischer Reihenfolge):

Dr. Camilla Rjosk (wissenschaftliche Koordination IQB-Bildungstrend 2016)
Dr. Stefan Schipolowski (wissenschaftlicher Leiter IQB-Bildungstrends)

Maria Engelbert
Dr. Nicole Haag
Pauline Kohrt

Dr. Heino Reimers
Dr. Sebastian Weirich
Julia Wittig

An der Vorbereitung der Studie waren zudem Prof. Dr. Katrin Böhme (bis zum 31.03.2017 stellvertretende wissenschaftliche Leiterin des IQB), Prof. Dr. Dirk Richter, Dr. Cornelia Gresch sowie Prof. Dr. Poldi Kuhl beteiligt.

Mit der praktischen Durchführung der Studie wurde die IEA Hamburg beauftragt. Die Aufgaben der IEA Hamburg umfassten unter anderem die Stichprobenziehung, die Schulung aller Testleiterinnen und Testleiter, die Koordination der Datenerhebung, die Kodierung der Schülerantworten, die Eingabe und Aufbereitung der Test- und Fragebogendaten sowie die Erstellung der Fallgewichte. Bei der IEA Hamburg waren folgende Personen in die Studie eingebunden:

Heiko Sibberns (Leitung)
Viktoria Böhm
Tina Ebert
Jens Gomolka
Dr. Miriam Hellrung
Nina Hugh
Philipp Köhne
Guido Martin

Maren Meyer-Everdt
Karsten Penon
Katrin Rütbling
Sabine Tieck
Gleb Turezkiy
Anja Waschke
Sabine Weber

Aufgrund der großen Anzahl teilnehmender Schulen konnten wichtige Aspekte der Vorbereitung und Durchführung des IQB-Bildungstrends 2016 nicht zentral geleistet werden und wurden daher von Projektkoordinatorinnen und -koordinatoren in den einzelnen Ländern übernommen. Zu ihren Aufgaben gehörte die Organisation von Schulleiterinformationsveranstaltungen, in denen das IQB gemeinsam mit der IEA Hamburg 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 und die Kommunikation mit den teilnehmenden Schulen verantwortlich. Bei den Projektkoordinatorinnen und -koordinato-

ren handelte es sich um Mitarbeiterinnen und Mitarbeiter der Ministerien, der Landesinstitute für Lehrerfortbildung oder der Qualitätsagenturen der Länder. Als Projektkoordinatorinnen und -koordinatoren waren folgende Personen tätig:

Soraya Akkoub (SH)	Dietmar Kirchhoff (HB)
Dr. Stephan Blank (BW)	Annemarie Klemm (HB)
Volker Bock (ST)	Astrid Koriller (MV)
Gabriele Colell (HE)	Anja Kurpiers (NW)
Mathias Dewald (HH)	Karin Lindenstruth (NW)
Dr. Grit Elsner (SN)	Johannes Miethner (RP)
Christiana Fahrnbach (RP)	Eric Möller (MV)
Martina Fischer (HE)	Michaela Müller (TH)
Carola Frenzel (SN)	Sabine Ocklitz-Sichermann (NI)
Daniela Gill (NI)	Steffen Pleßmann (BW)
Oliver Groth (HH)	Marc-Oliver Richter (SL)
Claudia Hansen (SH)	Bianca Schmidt (BY)
Tobias Haring (SH)	Antje Springstubbe (NI)
Manuela Hennig (MV)	Heiko Wontroba (TH)
Tina Holz (BE/BB)	Alexander Zapff (HB)

Die Durchführung der Testungen in den Schulen erfolgte durch Testleiterinnen und Testleiter, die von der IEA Hamburg geschult wurden. Dabei handelte es sich in der Regel um Studierende, Referendarinnen und Referendare, Lehrkräfte oder Mitarbeiterinnen und Mitarbeiter des schulpsychologischen Dienstes.

An der vom IQB geleiteten Entwicklung der Aufgaben, die im Bildungstrend eingesetzt wurden, wirkten Fachlehrkräfte aus allen 16 Ländern mit. Die Aufgabenentwicklerinnen und Aufgabenentwickler arbeiteten in regionalen Gruppen zusammen und erhielten Unterstützung von Expertinnen und Experten aus den Bereichen Fachdidaktik und Empirische Bildungsforschung. In den Teams arbeiteten die folgenden Personen mit:¹⁰

Im Fach Deutsch

Dr. Ulrike Behrens (Universität Duisburg-Essen)
 Prof. Dr. Albert Bremerich-Vos (Universität Duisburg-Essen)
 Prof. Dr. Wolfgang Eichler (Carl von Ossietzky Universität Oldenburg)
 Prof. Dr. Michael Krelle (Technische Universität Chemnitz)
 Susanne Maria Lang (Universität Paderborn)
 Dr. Christopher Sappok (Universität zu Köln)
 Prof. Dr. Angelika Speck-Hamdan (Ludwig-Maximilians-Universität München)

¹⁰ Da ein wesentlicher Teil der im IQB-Bildungstrend 2016 eingesetzten Aufgaben bereits im Vorfeld des IQB-Ländervergleichs 2011 entwickelt und in beiden Studien eingesetzt wurde, umfassen die folgenden Auflistungen auch Entwicklerinnen und Entwickler sowie Expertinnen und Experten, die seinerzeit an der Aufgabenerstellung zum IQB-Ländervergleich 2011 mitgewirkt haben.

Petra Bittins (BB)	Ulrike Potthoff (NW)
Cathrin Coerdts (RP)	Catrin Puschmann (NI)
Waltraud Credé (HE)	Anke Schumacher (NW)
Sabine Dörnhaus (SH)	Dr. Rositta Seidel (SN)
Christine Förster (HE)	Ilse Stork (BY)
Waltraud Frölich (BW)	Lars Strömel (RP)
Iris Grünack (BB)	Elvira Throm (BW)
Christiane Hermes (MV)	Elke Uthe (ST)
Brigitte Heßler (SN)	Sylke Wiegand (TH)
Friedrun Kemmner (BW)	Frauke Wietzke (SH)
Petra Klein (SL)	Uta Zimmermann (HH)
Karin Möcklinghoff (HH)	

Im Fach Mathematik

Prof. Dr. Hedwig Gasteiger (Universität Osnabrück)
 Daniela Götz (Universität Osnabrück)
 Prof. Dr. Jens Holger Lorenz (Pädagogische Hochschule Heidelberg, im
 Ruhestand)
 Gabriele Moll (Hochschule Karlsruhe)
 Prof. Dr. Kristina Reiss (Technische Universität München)
 Prof. Dr. Marja van den Heuvel-Panhuizen (Universität Utrecht)
 Prof. Dr. Gerd Walther (Universität Kiel)
 Prof. Dr. Bernd Wollring (Universität Kassel)

Regine Affeldt (HH)	Dr. Antje Hoffmann (NI)
Ute Alsdorf (TH)	Dietmar Leder (BW)
Ute Baumann (BB)	Birgit Löhr (HE)
Karin Behring (HB)	Dr. Gabriele Loibl (BY)
Thomas Bongartz (NW)	Barbara Meyer-Wirth (SL)
Michaela Dannenberger (MV)	Hennrika Nehls (MV)
Brigitte Dedekind (SH)	Dr. Raphaela Porsch (NW)
Renate Dittmar (SN)	Kirsten Rätthling (HH)
Rita Dürr (BW)	Dr. Charlotte Rechtsteiner-Merz (BW)
Ilse Eckhardt (HE)	Sabine Schmidt (ST)
Reinhard Forthaus (NW)	Dirk Schnitzler (NW)
Heike Frohmann (SH)	Charlotte Schorr-Brill (SL)
Gerda Frommeyer (HB)	Christian Schuster (RP)
Astrid Gebert (BE)	Maria Skejic (HE)
Dr. Hanna Haubold (MV)	

Entwicklung und Erprobung angepasster Aufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

Für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf wurden unter der Leitung des IQB in Kooperation mit der Universität Potsdam spezielle Deutsch- und Mathematikaufgaben entwickelt. Folgende Personen waren an der Entwicklung und Erprobung dieser Aufgaben beteiligt:

Miriam Balt (Universität Potsdam)
 Prof. Dr. Katrin Böhme (bis März 2017 IQB, danach Universität Potsdam)
 Moritz Börnert-Ringleb (Universität Potsdam)
 Prof. Dr. Antje Ehlert (Universität Potsdam)
 Prof. Dr. Hedwig Gasteiger (Universität Osnabrück)
 Theresa Gerlach (Universität Potsdam)
 Prof. Dr. Michael Grosche (Universität Wuppertal)
 Dr. Nicole Haag (IQB)
 Pauline Kohrt (IQB)
 Prof. Dr. Michael Krelle (Universität Chemnitz)
 Prof. Dr. Dirk Richter (Universität Potsdam)
 Dr. Sebastian Weirich (IQB)
 Prof. Dr. Jürgen Wilbert (Universität Potsdam)

Literatur

- Böhme, K. & Weirich, S. (2012). Der Ländervergleich im Fach Deutsch. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 103–116). Münster: Waxmann.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. 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. 375–397). Münster: Waxmann.
- Gonzalez, E. & Rutkowski, L. (2010). Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. *IERI Monograph Series*, 3, 125–156.
- 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.
- Haag, N. & Roppelt, A. (2012). Der Ländervergleich im Fach Mathematik. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 117–127). Münster: Waxmann.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+R*. Göttingen: Beltz Test.
- KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005a). *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. (2005b). *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Definitionenkatalog zur Schulstatistik 2015 (2)*. Zugriff am 20.07.2016 unter http://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Defkat2015_2.pdf
- Richter, D., Engelbert, M., Böhme, K., Haag, N., Hannighofer, J., Reimers, H., Roppelt, A., Weirich, S., Pant, H. A., Stanat, P. (2012). Anlage und Durchführung des Ländervergleichs. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 85–102). Münster: Waxmann.
- Schroeders, U., Schipolowski, S. & Wilhelm, O. (2017). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 3. und 4. Jahrgangsstufe*. Manuskript in Vorbereitung.
- Tarelli, I., Wendt, H., Bos, W. & Zylowski, A. (2012). Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2011). In W. Bos, I. Tarelli,

- A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–67). Münster: Waxmann.
- Wendt, H., Bos, W., Kasper, D., Walzebug, A., Goy, M. & Jusufi, D. (2016). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2015). In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 31–77). Münster: Waxmann.
- Wilhelm, O., Schroeders, U. & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8–10)*. Göttingen: Hogrefe.

3.2 Auswertung, Trendschätzung und Ergebnisdarstellung

Nicole Haag und Sebastian Weirich

Im IQB-Bildungstrend 2016 kann bezogen auf die Bildungsstandards der Kultusministerkonferenz erstmalig untersucht werden, wie sich die in der 4. Jahrgangsstufe erreichten Kompetenzen in den Fächern Deutsch und Mathematik seit dem Jahr 2011 in den Ländern entwickelt haben. Dieses Kapitel gibt einen Überblick über die Datenaufbereitung sowie die Herausforderungen, die mit der Datenauswertung und der Interpretation der Ergebnisse verbunden sind. Ferner werden die wichtigsten im IQB-Bildungstrend 2016 verwendeten Formen der Ergebnisdarstellung erläutert. Auf technische Details zur Datenauswertung geht Kapitel 13.1 genauer ein.

3.2.1 Aufbereitung und Auswertung der Daten

In einem ersten Schritt der Datenaufbereitung wurden die von den Schülerinnen und Schülern bearbeiteten Testhefte durch die IEA eingescannt. Anschließend wurden die nun digital vorliegenden Antworten der Schülerinnen und Schüler kodiert, das heißt konkreten Antwortkategorien zugeordnet. Während Aufgaben mit geschlossenen Antwortformaten, zum Beispiel Multiple-Choice-Items, maschinell kodierbar waren, erfolgte die Kategorisierung von Kurzantworten und offenen Antworten durch geschulte Kodiererinnen und Kodierer. Die Kodieranweisungen für diese Itemtypen wurden zu einem Großteil bereits für den IQB-Ländervergleich 2011 erstellt und schon vor ihrem Einsatz im IQB-Ländervergleich 2011 fachdidaktisch und empirisch geprüft. Sie kamen im IQB-Bildungstrend 2016 weitgehend unverändert zum Einsatz, um eine vergleichbare Kodierung für beide Testungen zu gewährleisten. In seltenen Fällen waren Überarbeitungen der Kodieranweisungen für offene Antworten erforderlich. Diese dienten im Wesentlichen dazu, die Kategorisierung der Schülerantworten für die Kodiererinnen und Kodierer zu vereinfachen, indem beispielsweise ähnliche Arten von richtigen Lösungen zu einer Kategorie zusammengefasst wurden. In die Trendschätzung wurden jedoch nur Aufgaben einbezogen, deren Kodierung zwischen den Jahren 2011 und 2016 inhaltlich unverändert blieb. Die Kodieranweisungen der für den IQB-Bildungstrend 2016 neu entwickelten Aufgaben wurden ebenfalls fachdidaktisch geprüft und in Pilotierungsstudien empirisch erprobt.

Die kodierten Schülerantworten wurden anhand der in einer Aufgabendatenbank hinterlegten Vorschriften als richtig oder falsch bewertet und – sofern es das Aufgabenformat erforderte – zusammengefasst. Die aufbereiteten Testdaten wurden anschließend skaliert, um für jedes Item einen Schwierigkeitsparameter zu generieren. Die Schätzung der Itemparameter basierte auf den Testdaten aller Schülerinnen und Schüler der jeweiligen Zielpopulation und schloss somit Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf ein (vgl. Kapitel 3.1). Dabei kam ein Modell aus der probabilistischen Testtheorie zur Anwendung. Aufgrund der speziellen Messeigenschaften der neu entwickelten Items für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (vgl. Kapitel 13) wurde ein Modell gewählt, das leicht vom bisher in IQB-Ländervergleichsstudien und IQB-Bildungstrends zur Kalibrierung ein-

gesetzten Modell abweicht. Dieses Modell erlaubt es, für die bereits im IQB-Ländervergleich 2011 eingesetzten Items Schwierigkeitsparameter zu bestimmen, die für Trendanalysen geeignet sind, und dabei gleichzeitig den besonderen Eigenschaften der für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf neu entwickelten Items Rechnung zu tragen. Die Gewinnung der Itemparameter erfolgte, wie bereits im IQB-Ländervergleich 2011, separat pro Fach und Kompetenzbereich unter Anwendung eindimensionaler Modelle.

Die so gewonnenen Itemparameter wurden im nächsten Schritt zur Ermittlung der Kompetenzwerte der Schülerinnen und Schüler verwendet. Dabei mussten zwei Zielstellungen beachtet werden: Zum einen sollten die Kompetenzen aller Schülerinnen und Schüler auf einer gemeinsamen Skala abgebildet werden, obwohl jede Schülerin und jeder Schüler nur eine Teilmenge der insgesamt eingesetzten Aufgaben bearbeitet hatte (vgl. Kapitel 3.1). Zum anderen wurde angestrebt, möglichst präzise Aussagen über Populationen (z. B. alle Viertklässlerinnen und Viertklässler eines Landes) und genau definierte Subpopulationen (z. B. Viertklässlerinnen und Viertklässler eines Landes, die zielgleich unterrichtet werden) zu treffen.

Dementsprechend wurde ein Verfahren gewählt, das verlässliche Schätzungen der Kompetenzverteilungen in Populationen und Subpopulationen ermöglicht. Neben einer unverzerrten Schätzung von Mittelwerten und Varianzen sollten auch Zusammenhänge der Kompetenzwerte mit weiteren Merkmalen der Schülerinnen und Schüler verzerrungsfrei geschätzt werden können. Solche Zusammenhänge sind beispielsweise für die Bestimmung von Disparitäten relevant, die mit Geschlecht, sozialer Herkunft oder Zuwanderungshintergrund assoziiert sind. Der in den Studien des IQB und auch in internationalen Schulleistungsstudien üblicherweise verwendete *Plausible-Value*-Ansatz erlaubt es, Kompetenzwerte zu gewinnen, die diesen Anforderungen gerecht werden. Bei der Schätzung der Werte wurden daher außer den Testdaten im jeweiligen Fach unter anderem das Geschlecht der Schülerinnen und Schüler, Indikatoren ihres sozialen Hintergrunds und Zuwanderungshintergrunds sowie motivationale Merkmale berücksichtigt, die in den vertiefenden Analysen des Bildungstrends genauer untersucht werden.

Um anhand der Daten des IQB-Bildungstrends 2016 repräsentative Aussagen über die Populationen der Viertklässlerinnen und Viertklässler in den Ländern treffen zu können, wurden die Analysen unter Verwendung von Fallgewichten durchgeführt. In ihre Berechnung gingen Faktoren auf der Ebene der Schule, der Klasse und des Individuums ein (vgl. OECD, 2012). Dabei wurde auf Schulebene berücksichtigt, dass nur eine Auswahl aller im jeweiligen Land vorhandenen Schulen an der Untersuchung teilnimmt, die Aussagen sich aber auf die Gesamtheit aller Schulen beziehen sollen. Auf Klassenebene wurde der Tatsache Rechnung getragen, dass sich die Anzahl der Klassen in der 4. Jahrgangsstufe zwischen den Schulen unterscheidet. Auf Individualebene schließlich wurde berücksichtigt, dass einzelne Schülerinnen und Schüler am Testtag beispielsweise aufgrund von Erkrankung fehlten. Unter Beachtung aller genannten Faktoren erhielt jede Schülerin und jeder Schüler ein Gewicht, welches angibt, wie viele Kinder der Zielpopulation diese eine Schülerin beziehungsweise dieser eine Schüler repräsentiert. Dadurch war es möglich, die erreichten Kompetenzen für die jeweilige Zielpopulation unverzerrt zu schätzen.

3.2.2 Trendschätzung und Definition der Metrik für den Bildungstrend

Um Veränderungen der Kompetenzwerte und der Kompetenzstufenbesetzungen zwischen dem IQB-Ländervergleich 2011 und dem IQB-Bildungstrend 2016 beschreiben zu können, mussten die in den beiden Studien jeweils gewonnenen Kompetenzwerte auf einer gemeinsamen Metrik abgebildet werden. Für die Kompetenzbereiche *Lesen* und *Zuhören* im Fach Deutsch sowie für die Globalskala und die Skalen zu allen fünf Leitideen im Fach Mathematik wurden hierfür die Berichtsmetriken des IQB-Ländervergleichs 2011 gewählt. Diese haben bezogen auf die Population der Viertklässlerinnen und Viertklässler des allgemeinbildenden Schulsystems im Jahr 2011 einen Mittelwert von $M = 500$ Punkten und eine Streuung von $SD = 100$ Punkten.

Für den Kompetenzbereich *Orthografie* im Fach Deutsch musste ein anderes Vorgehen zur Festlegung der Berichtsmetrik gewählt werden, da die orthografischen Fähigkeiten im IQB-Ländervergleich 2011 nur in einer Teilstichprobe erhoben wurden, die ausschließlich Schülerinnen und Schüler an allgemeinen Schulen umfasste. Im IQB-Bildungstrend 2016 umfasste die Testung hingegen auch Schülerinnen und Schüler an Förderschulen (vgl. Kapitel 3.1 und Kapitel 13.1).

Um auch in zukünftigen IQB-Bildungstrends die orthografischen Kompetenzen *aller* Schülerinnen und Schüler unabhängig von der besuchten Schulart auf der Berichtsmetrik abbilden und sinnvoll interpretieren zu können, wurde im Bereich *Orthografie* eine neue Berichtsmetrik für den IQB-Bildungstrend 2016 definiert. Das Vorgehen orientierte sich an der Neudefinition der Berichtsmetriken für die Fächer Deutsch und Englisch im IQB-Bildungstrend 2015. Der Mittelwert (M) von 500 Punkten und die Standardabweichung (SD) von 100 Punkten beziehen sich im Bereich *Orthografie* demnach auf die im Jahr 2016 getestete Population, die entsprechenden Kennwerte in den anderen Kompetenzbereichen hingegen auf die im Jahr 2011 getestete Population.

Um den Trend pro Fach und Kompetenzbereich bestimmen zu können, wurden sowohl die im IQB-Bildungstrend 2016 ermittelten Kompetenzwerte der Schülerinnen und Schüler als auch die aus dem IQB-Ländervergleich 2011 stammenden Kompetenzwerte jeweils auf eine gemeinsame Berichtsmetrik umgerechnet. Die Umrechnungsvorschrift für die in beiden Jahren ermittelten Kompetenzwerte ließ sich anhand der Veränderungen in den Schwierigkeiten der Items bestimmen, die in beiden Studien in unveränderter Form zum Einsatz kamen. Hierfür wurde eine spezielle Eigenschaft der zur Auswertung verwendeten statistischen Modelle genutzt, wonach die Schwierigkeitswerte der Items und die Kompetenzwerte der Personen auf derselben Skala darstellbar sind. Durch sogenannte *Linking*-Verfahren lässt sich aus der mittleren Veränderung der Itemschwierigkeiten ein Wert ableiten, der angibt, wie stark sich die mittleren Kompetenzen der Schülerinnen und Schüler zwischen dem IQB-Ländervergleich 2011 und dem IQB-Bildungstrend 2016 verändert haben. Sinkt die mittlere Schwierigkeit der Items zwischen den beiden Messungen beispielsweise um 15 Punkte, so kann daraus geschlossen werden, dass die mittlere Kompetenz der Schülerinnen und Schüler um 15 Punkte gestiegen ist. Technische Details zur Trendschätzung sowie zur Abschätzung der damit verbundenen Unsicherheit werden in Kapitel 13.1 dargestellt.

Aufgrund der oben beschriebenen Bildung der neuen Metrik im Bereich *Orthografie*, die sich auf das Jahr 2016 bezieht, können die im Bericht zum IQB-Ländervergleich 2011 veröffentlichten Ergebnisse für den Bereich *Orthografie* nicht direkt mit den Ergebnissen im Bericht zum IQB-Bildungstrend 2016 verglichen werden, sondern müssen auf die neue Metrik aus dem Jahr 2016 transformiert werden.

3.2.3 Darstellung der Ergebnisse

Nachfolgend werden die wichtigsten Varianten der Ergebnisdarstellungen erläutert. Neben Mittelwerten, Streuungen und Kompetenzstufenverteilungen werden statistische Maßzahlen berichtet, die sich auf bestimmte Unterschiede beziehen. Hierbei kann es sich um Unterschiede zwischen den Erhebungszeitpunkten 2011 und 2016 oder um Unterschiede zwischen Subpopulationen (z. B. Jungen und Mädchen) zu einem der beiden Erhebungszeitpunkte handeln. Zusätzlich wird im Berichtsband dargestellt, wie sich die Heterogenität der Schülerinnen und Schüler sowohl zwischen Schulen als auch innerhalb von Schulen verteilt. Der Anteil der Gesamtheterogenität, der auf Unterschiede zwischen Schulen zurückgeht, wird als Intraklassenkorrelation (*ICC*) bezeichnet.

Mittelwerte und Standardabweichungen

Mittelwerte geben die in einer bestimmten Population oder Subpopulation im Durchschnitt erreichten Kompetenzwerte an. Im Gegensatz dazu zeigen Standardabweichungen, wie stark die Kompetenzen innerhalb einer Population streuen, das heißt wie heterogen (bzw. homogen) die Verteilung der Kompetenzen ist. Hohe Mittelwerte in Verbindung mit niedrigen Streuungen deuten dabei auf das anzustrebende Ziel homogen leistungsstarker Populationen oder Subpopulationen hin. Bei normalverteilten Kompetenzen liegen die Werte von etwa zwei Dritteln (68.2 %) der Viertklässlerinnen und Viertklässler im Bereich von ± 1 Standardabweichung um den Mittelwert. Im Bereich von ± 2 Standardabweichungen um den Mittelwert liegen 95.4 Prozent der Kompetenzwerte.

Perzentile und Perzentilbänder

Perzentile können dazu genutzt werden, die von unterschiedlich leistungsstarken Gruppen erzielten Kompetenzwerte zwischen den Ländern zu vergleichen, da sie eine Verteilung in unterschiedliche Anteile zerlegen. Vergleicht man beispielsweise das 95. Perzentil zwischen den Ländern, so lässt sich daraus ableiten, welche Kompetenzen die leistungsstärksten 5 Prozent der Schülerinnen und Schüler innerhalb jedes Landes mindestens erreicht haben und welche Kompetenzen die restlichen 95 Prozent der Schülerinnen und Schüler dieses Landes höchstens erreicht haben. Perzentilwerte können mit sogenannten Perzentilbändern grafisch veranschaulicht werden, die auch Auskunft über die Breite der Kompetenzverteilung geben (siehe dazu insbesondere Kapitel 6).

Kompetenzstufen

Während die bisher erläuterten Formen der Ergebnisdarstellung die Kompetenzverteilungen der Schülerinnen und Schüler lediglich statistisch beschreiben, lassen sie sich anhand von Kompetenzstufen auch inhaltlich interpretieren. Kompetenzstufen sind zudem besser dazu geeignet, die Anteile von besonders schwachen oder besonders starken Schülerinnen und Schülern in den Ländern zu bestimmen. Jede Kompetenzstufe umfasst einen gewissen Bereich von Testpunkten und eine inhaltliche Beschreibung der Kompetenzanforderungen, die Schülerinnen und Schüler mit Testwerten innerhalb dieses Bereichs typischerweise bewältigen können. Dazu werden die von den Schülerinnen und Schülern erreichten Testwerte auf die Kompetenzstufenmetrik umgerechnet (vgl. Kapitel 2.1).

Der Bezug zu den Bildungsstandards wird durch eine Zuordnung der einzelnen Kompetenzstufen zu den abschlussbezogenen Mindest-, Regel- und Optimalstandards hergestellt. Das Erreichen dieser Standards wird in Kapitel 5 dargestellt. Die im IQB-Bildungstrend verwendeten Kompetenzstufenmodelle werden in Kapitel 2 ausführlich beschrieben.

Standardfehler und Signifikanztests

Ziel des IQB-Bildungstrends ist es, von Ergebnissen, die für Stichproben ermittelt wurden, auf Kennwerte der gesamten Zielpopulation zu schließen. Diese Art von Schlüssen ist immer mit einer gewissen Unsicherheit verbunden, deren Ausmaß mit dem Standardfehler (*SE*) angegeben wird. Je kleiner der Standardfehler eines statistischen Kennwerts ist, desto präziser bildet der anhand der Stichprobe geschätzte Wert den tatsächlichen Wert in der Zielpopulation ab. Geschätzte Kennwerte mit einem großen Standardfehler sind dementsprechend mit einer größeren Unsicherheit behaftet. Die Standardfehler von Kennwerten können auch dazu genutzt werden, die statistische Signifikanz von Unterschieden und Trends zu ermitteln. Auf mögliche Quellen von Unsicherheiten, die beim Vergleich von statistischen Kennwerten zwischen Gruppen oder über die Zeit berücksichtigt werden müssen, wird in Kapitel 13.1 näher eingegangen.

Praktische Relevanz und standardisierte Mittelwertsdifferenzen

Neben der statistischen Signifikanz werden soweit wie möglich auch Indikatoren für die praktische Relevanz von Unterschieden angegeben, da die statistische Signifikanz alleine darüber nichts aussagt. Die statistische Signifikanz eines Unterschiedes hängt von der Unsicherheit der Schätzungen ab, die wiederum von der Stichprobengröße beeinflusst wird. So ist bei kleineren Stichproben die Unsicherheit der Schätzungen höher und bei größeren entsprechend kleiner. Werden zwei kleine Gruppen, zum Beispiel Schülerinnen und Schüler, deren Familien aus verschiedenen Herkunftsländern zugewandert sind, miteinander verglichen, so könnte ein bestimmter Mittelwertsunterschied als statistisch nicht signifikant ausgewiesen werden, wohingegen ein Mittelwertsunterschied in vergleichbarer Höhe zwischen zwei größeren Gruppen, zum Beispiel Schülerinnen und Schülern aus verschiedenen Ländern in Deutschland, sich als statistisch signifikant erweisen würde.

Zur Beurteilung der praktischen Relevanz lassen sich verschiedene Werte heranziehen. Zum einen können standardisierte Mittelwertsdifferenzen oder Effektstärken (d) berechnet werden. Diese relativieren die zwischen zwei Gruppen festgestellten Unterschiede an ihrer gemeinsamen Standardabweichung (Cohen, 1988). Anhand dieses Indikators lassen sich bestimmte Gruppenunterschiede direkt miteinander vergleichen, auch wenn die Streuung der Werte in den Gruppen variiert. Zum anderen kann für Unterschiede in den erreichten Kompetenzen angegeben werden, wie groß der Unterschied im Vergleich zum erwarteten Lernzuwachs in einem Schuljahr in der Primarstufe ausfällt. Die entsprechenden Schätzungen der erwarteten Lernzuwächse werden für das Fach Deutsch in Kapitel 6.1 und für das Fach Mathematik in Kapitel 6.2 dargestellt.

Bestimmung von Merkmalsunterschieden innerhalb und zwischen Schulen

Neben der Streuung von Merkmalen, beispielsweise der erreichten Kompetenzen oder des sozioökonomischen Status, innerhalb eines Landes ist auch von Interesse, ob Unterschiede in Bezug auf dieses Merkmal eher zwischen Schulen oder eher innerhalb von Schulen bestehen. Um diese Unterschiede zu bestimmen, wird die Varianz des Merkmals zerlegt in einen Teil, der zwischen einzelnen Schulen besteht, und einen Teil, der innerhalb von Schulen besteht. Die Größe dieser beiden Teile hängt auch von der Gesamtheterogenität des Merkmals im Land ab und lässt sich daher nur schwer zwischen Ländern vergleichen. Die Varianz zwischen Schulen und die Varianz innerhalb von Schulen werden daher jeweils durch die Gesamtvarianz des Landes geteilt. Da diese Anteile von der Gesamtheterogenität unabhängig sind, können sie einfacher zwischen den Ländern verglichen werden.

Die Intraklassenkorrelation (ICC) ist ein deskriptives Maß für den Anteil der Varianz zwischen Schulen an der Gesamtvarianz des Landes und kann Werte zwischen 0 und 1 annehmen. In Bezug auf die erreichten Kompetenzen würde ein Wert von 0 bedeuten, dass an allen Schulen eines Landes die gleiche mittlere Kompetenz erreicht wird und sich die Kompetenzwerte der Schülerinnen und Schüler ausschließlich innerhalb von Schulen unterscheiden. Bei einer Intraklassenkorrelation von 1 würde es sich genau umgekehrt verhalten, das heißt, die Unterschiede wären allein auf die Varianz zwischen Schulen zurückzuführen, wohingegen alle Schülerinnen und Schüler innerhalb einer Schule den gleichen Kompetenzwert erreichen würden.

Multipliziert man die Intraklassenkorrelation mit 100, ergibt dies den prozentualen Anteil der Varianz zwischen Schulen an der Gesamtvarianz des Landes. Dieser Anteil wird, gemeinsam mit dem Anteil der Varianz innerhalb von Schulen, in Kapitel 6 für die erreichten Kompetenzen und in Kapitel 8 für den sozioökonomischen Status berichtet.

3.2.4 Hinweise zur Interpretation der Trends

Die im IQB-Bildungstrend 2016 ermittelten Trendschätzungen werden als Veränderungen der im Mittel erreichten Kompetenzen im Vergleich zum Jahr 2011 interpretiert. Dafür müssen sich die Schätzungen der Kompetenzmittelwerte für beide Zeitpunkte auf eine vergleichbare Population beziehen. Für die Kompetenzbereiche *Lesen* und *Zuhören* im Fach Deutsch sowie für alle unter-

suchten Kompetenzbereiche im Fach Mathematik wurden in beiden Erhebungen Stichproben getestet, die sowohl bundesweit als auch auf Länderebene für die Gesamtpopulation der Viertklässlerinnen und Viertklässler im allgemeinbildenden Schulsystem (einschließlich Förderschulen) repräsentativ sind. Die im IQB-Bildungstrend 2016 ermittelten und auf die Berichtsmetrik umgerechneten Kompetenzwerte können in diesen Kompetenzbereichen also direkt mit den im IQB-Ländervergleich 2011 festgestellten Kompetenzwerten verglichen werden.

Die im IQB-Bildungstrend 2016 ermittelten Kompetenzwerte im Bereich *Orthografie* sind jedoch, wie oben erläutert, nicht ohne weiteres mit den im IQB-Ländervergleich 2011 ermittelten Kompetenzwerten vergleichbar, da im Jahr 2011 keine Förderschulen in die Erhebung einbezogen wurden. Im Jahr 2016 hingegen nahmen auch Schülerinnen und Schüler an Förderschulen an der Studie teil (vgl. Kapitel 3.1), sodass nur die Stichprobe des IQB-Bildungstrends 2016 repräsentativ für alle Schülerinnen und Schüler der 4. Jahrgangsstufe an allgemeinbildenden Schulen einschließlich Förderschulen¹ ist. Für die Trendberechnung war es im Bereich *Orthografie* daher erforderlich, die von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf erzielten Kompetenzwerte auszuschließen, um systematisch verzerrte Schätzungen der Veränderungen zu vermeiden. In diesem Kompetenzbereich werden die Ergebnisse des IQB-Bildungstrends 2016 daher für das Jahr 2016 bezogen auf alle Schülerinnen und Schüler der 4. Jahrgangsstufe, inklusive der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, und zur Beschreibung von Veränderungen gegenüber dem Jahr 2011 ausschließlich für Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf berechnet. In den Tabellen und Abbildungen der einzelnen Ergebniskapitel ist gekennzeichnet, auf welche Population sich die Ergebnisse jeweils beziehen.

Der Kompetenzbereich *Orthografie* wurde im IQB-Ländervergleich 2011 außerdem nur in einer kleineren Teilstichprobe erhoben, die nur für Schülerinnen und Schüler an allgemeinen Schulen in Deutschland insgesamt repräsentativ ist, jedoch keine Analysen auf Länderebene erlaubt. Dementsprechend können Veränderungen im Bereich *Orthografie* im Zeitraum von 2011 bis 2016 im Rahmen der Trendanalysen ebenfalls nur für Deutschland bestimmt werden. Die Schätzung von Veränderungen auf Länderebene wird erst im nächsten IQB-Bildungstrend in der Primarstufe möglich sein, der im Jahr 2020 durchgeführt wird.

Darüber hinaus sollten insbesondere bei der Interpretation von Veränderungen in den Kompetenzstufenbesetzungen innerhalb eines Landes im Vergleich der Jahre 2011 und 2016 folgende Hinweise berücksichtigt werden:

1. Die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen hängt auch mit der Streuung der erreichten Kompetenzwerte zusammen. Bei gleichem Mittelwert sind in einem Land mit einer hohen Streuung der Kompetenzwerte die Stufen an den Rändern der Kompetenzstufenverteilung stärker besetzt als in einem Land mit geringerer Streuung. Eine Veränderung an den Rändern der Kompetenzstufenverteilung weist daher nicht zwangsläufig darauf hin, dass sich das erreichte Kompetenzniveau der Schülerinnen und Schüler insgesamt erhöht oder verringert hat, sondern verweist zunächst auf eine Veränderung der Streuung. So ist beispielsweise ein steigender Anteil

¹ Die Repräsentativität für Schülerinnen und Schüler an Förderschulen ist geringfügig eingeschränkt, da aus praktischen Gründen nur Schulen mit den Förderschwerpunkten „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“ in die Stichprobenziehung einbezogen wurden (vgl. Kapitel 3.1).

von Schülerinnen und Schülern, die die Optimalstandards erreichen, vor allem dann eine wünschenswerte Entwicklung, wenn sich gleichzeitig der Anteil an Schülerinnen und Schülern verringert oder zumindest nicht erhöht hat, die die Mindeststandards verfehlen.

2. Veränderungen der Kompetenzstufenbesetzungen sollten ins Verhältnis zum Ausgangsniveau gesetzt werden. Ein um 10 Prozentpunkte höherer Anteil der Schülerinnen und Schüler, die die Regelstandards erreichen, dürfte zum Beispiel bei einem Ausgangsniveau von 25 Prozent anders zu bewerten sein als bei einem Ausgangsniveau von 70 Prozent.
3. Die identifizierten Differenzen können auch von sogenannten Boden- oder Deckeneffekten beeinflusst werden. Diese treten dann auf, wenn die Ränder der Kompetenzstufenverteilung bereits nahezu minimal oder maximal besetzt sind. Verfehlten beispielsweise im Jahr 2011 nur 3 Prozent der Schülerinnen und Schüler in einem Land die Mindeststandards, ist eine weitere Reduktion dieses Anteils kaum noch möglich.

Literatur

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- OECD (Hrsg.). (2012). *PISA 2009 technical report*. Paris: OECD.

Kapitel 4

Kontextinformationen zu den Schulsystemen der Länder im Primarbereich

Stefan Schipolowski, Petra Stanat und Camilla Rjosk

4.1 Einleitung

Bevor in den folgenden Kapiteln die von den Viertklässlerinnen und Viertklässlern in den Ländern erreichten Kompetenzen beschrieben werden, erfolgt an dieser Stelle zunächst eine überblicksartige Darstellung grundlegender Merkmale der Schulsysteme der Länder im Primarbereich, die für die Interpretation der Ergebnisse bedeutsam sein können. Sowohl aus Platzgründen als auch aufgrund der Tatsache, dass nicht zu allen potenziell relevanten Kontextfaktoren belastbare Daten vorliegen, kann dabei nur auf ausgewählte Systemmerkmale in der Primarstufe eingegangen werden. Diese werden zusammenfassend beschrieben, wobei im Ländervergleich auf Gemeinsamkeiten und Unterschiede aufmerksam gemacht wird.

Bei den im Folgenden dargestellten Informationen handelt es sich um Merkmale der Schullaufbahnen der Kinder im Primarbereich, die allgemeine Lernbedingungen und potenzielle Lerngelegenheiten kennzeichnen. Sie basieren überwiegend auf amtlichen Schulstatistiken und Vorgaben der Kultusbehörden der Länder. Konkret wird darauf eingegangen,

- welcher Anteil der Viertklässlerinnen und Viertklässler in den Ländern sonderpädagogischen Förderbedarf aufweist, inwieweit diese Kinder in Förderschulen oder allgemeinen Schulen unterrichtet werden und inwieweit für diese Kinder die Bildungspläne der allgemeinen Schule maßgeblich sind,
- wie viele Stunden Unterricht für die Fächer Deutsch und Mathematik in der Primarstufe mindestens vorgesehen sind,
- welche Stichtagsregelungen zur Einschulung bestehen und inwieweit Einschulungen verspätet beziehungsweise vorzeitig erfolgen und
- welcher Anteil der Schülerinnen und Schüler am Ganztagsbetrieb teilnimmt.

Entsprechend der in den Ergebniskapiteln beschriebenen Trends in den erreichten Kompetenzen, die sich auf Veränderungen zwischen den Jahren 2011 und 2016 beziehen, erfolgt auch die Darstellung der Kontextbedingungen für die Populationen der Viertklässlerinnen und Viertklässler des Jahres 2011 und des Jahres 2016, um mögliche Veränderungen in diesen Merkmalen abzubilden. Darüber hinaus wird auch innerhalb der einzelnen Ergebniskapitel über die Zusammensetzung der Schülerschaft in den Ländern informiert. Dies betrifft insbesondere den sozioökonomischen Hintergrund der Schülerinnen und Schüler (vgl. Kapitel 8) sowie den Anteil der Viertklässlerinnen und Viertklässler mit Zuwanderungshintergrund (vgl. Kapitel 9).

4.2 Ausgewählte Kontextmerkmale der Schulsysteme der Länder

4.2.1 Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

Ein wesentlicher schulstruktureller Aspekt, in dem sich die Länder im Primarbereich unterscheiden, betrifft den Anteil und den Beschulungsort von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF). In den Tabellen 4.1 und 4.2 ist für das Schuljahr 2010/2011 beziehungsweise 2015/2016 anhand von Daten der statistischen Ämter des Bundes und der Länder angegeben, wie groß der Anteil der Schülerinnen und Schüler der 4. Jahrgangsstufe ist, bei denen ein SPF vorliegt.¹ Bezieht man Viertklässlerinnen und Viertklässler aller Förderschwerpunkte sowohl an Förderschulen als auch an allgemeinen Schulen ein, so liegt dieser Anteil in Deutschland im Schuljahr 2015/2016 bei insgesamt rund 7 Prozent, wobei er auf Länderebene zwischen gut 5 Prozent in Rheinland-Pfalz² und fast 11 Prozent in Mecklenburg-Vorpommern erheblich variiert (vgl. Tab. 4.2). Darüber hinaus unterscheiden sich die Länder in der Primarstufe zum Teil deutlich im Hinblick darauf, welche Anteile der Schülerinnen und Schüler mit SPF an allgemeinen Schulen (d. h. *inklusiv* beziehungsweise *integrativ*)³ einerseits und an Förderschulen andererseits unterrichtet werden. Die entsprechenden Daten der statistischen Landesämter für die 4. Jahrgangsstufe, die sich auf alle von der Kultusministerkonferenz (KMK) definierten Förderschwerpunkte beziehen, sind ebenfalls in den Tabellen 4.1 und 4.2 dargestellt. Von allen Viertklässlerinnen und Viertklässlern mit SPF besuchen demnach im Schuljahr 2015/2016 je nach Land zwischen knapp 16 Prozent (Baden-Württemberg) und über 90 Prozent (Bremen) eine allgemeine Schule.⁴ Für Deutschland insgesamt liegt dieser Anteil bei rund 44 Prozent.

Ein Vergleich der Angaben für das Schuljahr 2015/2016 mit dem Schuljahr 2010/2011 zeigt, dass der Anteil der Viertklässlerinnen und Viertklässler mit SPF an allen Schülerinnen und Schülern der 4. Jahrgangsstufe über den Zeitraum von fünf Jahren nur geringfügig gestiegen ist (+1 Prozentpunkt). Eine deutlich größere Veränderung ist hingegen für die Verteilung auf die Schularten festzustellen. So liegt der Anteil der Schülerinnen und Schüler mit SPF, die eine allgemeine Schule besuchen, in der 4. Jahrgangsstufe im Schuljahr 2015/2016 bundesweit bei gut 44 Prozent und damit um etwa 17 Prozentpunkte höher als im Schuljahr 2010/2011.

1 Hierbei ist zu beachten, dass sich die Verfahren zur Feststellung sonderpädagogischen Förderbedarfs und somit auch die Grundlage für die Angaben teilweise zwischen den Ländern unterscheiden (vgl. Gresch & Piezunka, 2015; Malecki, 2013; Sälzer, Gebhardt, Müller & Pauly, 2015).

2 Die in Tabelle 4.2 angegebenen Quoten für die Länder Niedersachsen und Saarland sind noch geringer, sie beziehen sich jedoch ausschließlich auf Schülerinnen und Schüler an Förderschulen und wurden daher an dieser Stelle nicht berücksichtigt.

3 Dabei wird – wie schon in den IQB-Ländervergleichen 2011 und 2012 sowie im IQB-Bildungstrend 2015 (Stanat, Pant, Böhme & Richter, 2012; Pant et al., 2013; Stanat, Böhme, Schipolowski & Haag, 2016) – keine Unterscheidung zwischen Integration beziehungsweise Inklusion von Kindern mit sonderpädagogischem Förderbedarf getroffen. Die Diskussion zu inhaltlichen Differenzen zwischen beiden Konzepten (siehe z. B. Grosche, 2015) soll in diesem Bericht nicht aufgegriffen werden.

4 Den unterschiedlichen Anteilen von Schülerinnen und Schülern mit SPF an allgemeinen Schulen und an Förderschulen in den Ländern wird in den Ländervergleichsstudien und Bildungstrends des IQB seit dem Jahr 2011 dadurch Rechnung getragen, dass sowohl Schülerinnen und Schüler mit SPF an allgemeinen Schulen als auch Schülerinnen und Schüler an Förderschulen in die Erhebungen einbezogen werden. Hierdurch wird die länderübergreifende Vergleichbarkeit der Ergebnisse optimiert.

Tabelle 4.1: Verteilung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in der 4. Jahrgangsstufe im Schuljahr 2010/2011 in den Ländern

Land	Gesamt-schülerzahl in der 4. Jahrgangsstufe ¹		SuS mit SPF ²		SuS mit SPF in den Förder-schwerpunkten LSE		SuS mit SPF in Förderschulen ²		SuS mit SPF in Allgemeinen Schulen		SuS mit SPF in den Förderschulen		SuS mit SPF in den Allgemeinen Schulen	
	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil an SuS mit SPF in %	Anzahl	Anteil an SuS mit SPF in %	Anzahl	Anteil an SuS mit SPF in %	Anzahl	Anteil an SuS mit SPF in %
Baden-Württemberg ^{3,4}	108 382	4.7	3 706	3.4	5 146	—	—	—	—	—	—	—	—	—
Bayern	124 589	7.0	5 359	4.3	6 012	68.9	2 719	31.1	2 859	53.3	2 500	46.7	2 500	46.7
Berlin	25 112	8.3	1 584	6.3	965	46.4	1 117	53.6	656	41.4	928	58.6	928	58.6
Brandenburg	19 552	8.2	1 401	7.2	719	44.8	886	55.2	655	46.8	746	53.2	746	53.2
Bremen	5 540	9.8	356	6.4	98	18.1	444	81.9	12	3.4	344	96.6	344	96.6
Hamburg	15 277	8.41	549	3.6	590	70.2	251	29.8	409	74.5	140	25.5	140	25.5
Hessen	57 079	5.1	2 057	3.6	2 285	78.6	621	21.4	1 523	74.0	534	26.0	534	26.0
Mecklenburg-Vorpommern	12 859	12.5	1 299	10.1	1 103	68.8	501	31.2	853	65.7	446	34.3	446	34.3
Niedersachsen ⁴	79 427	2 745	2 378	3.0	2 745	—	—	—	2 378	—	—	—	—	—
Nordrhein-Westfalen	180 562	11 721	6.5	8 710	4.8	8 589	73.3	3 132	26.7	69.4	2 665	30.6	2 665	30.6
Rheinland-Pfalz	38 694	1 697	4.4	1 256	3.2	1 236	72.8	461	27.2	67.4	410	32.6	410	32.6
Saarland ⁴	8 088	314	3.9	199	2.5	314	—	—	199	—	—	—	—	—
Sachsen	31 319	2 515	8.0	2 253	7.2	1 846	73.4	669	26.6	76.5	530	23.5	530	23.5
Sachsen-Anhalt ⁵	17 435	1 695	9.7	1 327	7.6	1 311	77.3	384	22.7	73.9	347	26.1	347	26.1
Schleswig-Holstein	28 484	1 759	6.2	1 250	4.4	654	37.2	1 105	62.8	25.4	932	74.6	932	74.6
Thüringen	17 132	1 259	7.3	960	5.6	841	66.8	418	33.2	64.3	343	35.7	343	35.7
Deutschland	769 531	47 162	6.1	34 644	4.5	34 455	73.1	12 708	26.9	68.6	10 865	31.4	10 865	31.4

Anmerkungen. SuS = Schülerinnen und Schüler; SPF = sonderpädagogischer Förderbedarf; LSE = Förderschwerpunkte „Lernen“, „Sprache“, „emotionale und soziale Entwicklung“. ¹ Zusätzlich weist die Statistik des Statistischen Bundesamtes 9 471 Schülerinnen und Schüler an Förderschulen sowie 1 100 weitere Schülerinnen und Schüler im Primarbereich aus, die keiner bestimmten Jahrgangsstufe zugeordnet werden konnten. Diese sind hier nicht berücksichtigt. ² In den Ländern Brandenburg, Bremen, Hamburg, Niedersachsen, Rheinland-Pfalz, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen ohne Schülerinnen und Schüler an Förderschulen im Förderschwerpunkt „Unterricht kranker Schülerinnen und Schüler“. ³ Die Angaben stellen Schätzungen dar, da in den statistischen Daten nicht ausreichend zwischen Jahrgangsstufe 4 und Grundstufe unterschieden wird. ⁴ Die Angaben zu Anzahl und Anteil von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf beziehen sich ausschließlich auf die Schülerinnen und Schüler in Förderschulen. ⁵ Angaben zu Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an allgemeinen Schulen liegen nur für Schulen in öffentlicher Trägerschaft vor.

Quelle: Statistische Ämter des Bundes und der Länder.

Tabelle 4.2: Verteilung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in der 4. Jahrgangsstufe im Schuljahr 2015/2016 in den Ländern

Land	Gesamt-schülerzahl in der 4. Jahrgangsstufe ¹		SuS mit SPF ²		SuS mit SPF in den Förder-schwerpunkten LSE		SuS mit SPF in Förderschulen ²		Allgemeinen Schulen		Förderschulen		SuS mit SPF in den Schwerpunkten LSE in Allgemeinen Schulen	
	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil in %	Anzahl	Anteil an SuS mit SPF in %	Anzahl	Anteil an SuS mit SPF in %	Anzahl	Anteil an SuS mit SPF in %	Anzahl	Anteil an SuS mit SPF in %
Baden-Württemberg	97 884	5,8	4 072	4,2	4 772	84,1	900	15,9	3 326	81,7	746	18,3		
Bayern	110 202	8,5	5 552	5,0	5 449	58,3	3 891	41,7	1 978	35,6	3 574	64,4		
Berlin	28 172	9,2	1 770	6,3	684	26,5	1 899	73,5	363	20,5	1 407	79,5		
Brandenburg	19 948	7,0	1 169	5,9	422	30,3	969	69,7	372	31,8	797	68,2		
Bremen	5 533	8,0	335	6,1	34	7,7	408	92,3	14	4,2	321	95,8		
Hamburg	15 148	10,4	1 226	8,1	375	23,8	1 201	76,2	236	19,2	990	80,8		
Hessen	53 982	6,0	2 416	4,5	1 979	60,9	1 271	39,1	1 305	54,0	1 111	46,0		
Mecklenburg-Vorpommern	12 728	10,8	1 121	8,8	647	46,9	732	53,1	484	43,2	637	56,8		
Niedersachsen ³	74 212	2,8	1 803	2,4	2 077	—	—	—	1 803	—	—	—		
Nordrhein-Westfalen	160 273	8,2	9 922	6,2	6 484	49,5	6 605	50,5	4 253	42,9	5 669	57,1		
Rheinland-Pfalz	34 243	5,4	1 359	4,0	1 198	65,1	642	34,9	770	56,7	589	43,3		
Saarland ³	7 474	289	3,9	166	2,2	289	—	—	166	—	—	—		
Sachsen	32 180	2 844	8,8	2 513	7,8	1 727	60,7	1 117	39,3	1 597	63,5	916	36,5	
Sachsen-Anhalt	16 622	1 572	9,5	1 221	7,3	849	54,0	723	46,0	578	47,3	643	52,7	
Schleswig-Holstein	25 187	1 661	6,6	1 034	4,1	416	25,1	1 245	74,9	98	9,5	936	90,5	
Thüringen	16 727	1 082	6,5	720	4,3	553	51,1	529	48,9	318	44,2	402	55,8	
Deutschland	710 516	50 089	7,0	36 399	5,1	27 956	55,8	22 133	44,2	17 661	48,5	18 738	51,5	

Anmerkungen. SuS = Schülerinnen und Schüler; SPF = sonderpädagogischer Förderbedarf; LSE = Förderschwerpunkte „Lernen“, „Sprache“, „emotionale und soziale Entwicklung“. ¹ Zusätzlich weist die Statistik des Statistischen Bundesamtes 4 890 Schülerinnen und Schüler an Förderschulen sowie 3 269 weitere Schülerinnen und Schüler im Primarbereich aus, die keiner bestimmten Jahrgangsstufe zugeordnet werden konnten. Diese sind hier nicht berücksichtigt. ² In den Ländern Brandenburg, Bremen, Hamburg, Niedersachsen, Rheinland-Pfalz, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und Thüringen ohne Schülerinnen und Schüler an Förderschulen im Förderschwerpunkt „Unterricht kranker Schülerinnen und Schüler“. ³ Die Angaben zu Anzahl und Anteil von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf beziehen sich ausschließlich auf die Schülerinnen und Schüler in Förderschulen.

Quelle: Statistische Ämter des Bundes und der Länder.

Wie in Kapitel 3.1 erläutert, wurden im IQB-Bildungstrend 2016 – ebenso wie bereits im IQB-Ländervergleich 2011 – nur Förderschulen in die Erhebungen einbezogen, die den Förderschwerpunkten „Lernen“, „Sprache“ oder „emotionale und soziale Entwicklung“ (LSE) zuzuordnen sind. Daher werden die zuvor beschriebenen Statistiken in den Tabellen 4.1 und 4.2 zusätzlich auch für Viertklässlerinnen und Viertklässler in diesen Förderschwerpunkten ausgewiesen. Da Schülerinnen und Schüler mit SPF im Primarbereich überwiegend einem oder mehreren dieser drei Förderschwerpunkte zugeordnet sind, ergeben sich für diese Teilpopulation erwartungsgemäß in den meisten Ländern ähnliche Verteilungen wie für die Gesamtgruppe aller Kinder mit SPF in der 4. Jahrgangsstufe.

Für die Überprüfung des Erreichens der Bildungsstandards ist die Frage von Bedeutung, ob für die Schülerinnen und Schüler mit SPF dieselben Bildungspläne maßgeblich sind wie für Schülerinnen und Schüler ohne SPF, ob sie also *zielgleich* oder *zieldifferent* unterrichtet werden. Da für zieldifferent unterrichtete Kinder eigene Anforderungen gelten, die nicht zwangsläufig auf den Bildungsstandards basieren, wäre es unangebracht, die Bildungsstandards als Maßstab für die von ihnen erreichten Kompetenzen heranzuziehen. Bei den Analysen zur Frage, inwieweit die Schülerinnen und Schüler in den einzelnen Ländern die Bildungsstandards erreichen (vgl. Kapitel 5), wurden daher nur jene Viertklässlerinnen und Viertklässler einbezogen, die nach den Regelungen des jeweiligen Landes zielgleich und somit auf Grundlage der Bildungsstandards unterrichtet werden. Da die amtliche Statistik keine Angaben dazu enthält, welcher Anteil der Schülerinnen und Schüler zieldifferent unterrichtet wird, müssen die entsprechenden Quoten anhand der Daten des IQB-Ländervergleichs 2011 und des IQB-Bildungstrends 2016 geschätzt werden. Die für den organisatorischen Ablauf der Studie an der Schule zuständigen Lehrkräfte (Schulkoordinatorinnen und Schulkoordinatoren) wurden deshalb gebeten, in einer dafür vorgesehenen Liste jeweils zu kennzeichnen, für welche Schülerinnen und Schüler andere Bildungspläne (bzw. Lehrpläne/Rahmenlehrpläne/Kerncurricula) als die der allgemeinen Schule gelten. Die anhand dieser Angaben ermittelten Quoten sind für die untersuchten Schülerpopulationen der Jahre 2011 und 2016 in Tabelle 4.3 aufgeführt. Diese beziehen sich, wie bereits erläutert, überwiegend auf die Förderschwerpunkte LSE.⁵ Zu beachten ist bei der Interpretation dieser Angaben, dass weder der IQB-Ländervergleich 2011 noch der IQB-Bildungstrend 2016 darauf angelegt waren, die Anteile der zielgleich und zieldifferent unterrichteten Schülerinnen und Schüler zu schätzen. Daher stellen die angegebenen Quoten keine präzisen Populationsbeschreibungen dar; sie sollen vor allem transparent machen, welcher Anteil von Schülerinnen und Schülern aus den Analysen der Kompetenzstufenverteilungen in den einzelnen Ländern (vgl. Kapitel 5) ausgeschlossen wurde.

Wie aus Tabelle 4.3 hervorgeht, ist der Anteil der zieldifferent unterrichteten Viertklässlerinnen und Viertklässler mit SPF an der Gesamtpopulation der Schülerinnen und Schüler der 4. Jahrgangsstufe mit rund 3 Prozent in

5 Abweichungen zwischen Tabelle 4.1/4.2 und Tabelle 4.3 im Anteil der Schülerinnen und Schüler mit SPF an der Gesamtpopulation der Viertklässlerinnen und Viertklässler sind insbesondere darauf zurückzuführen, dass die Angaben in den Tabellen 4.1 und 4.2 auf Populationsdaten der statistischen Ämter beruhen, während es sich bei den Angaben in Tabelle 4.3 um Schätzungen anhand von Stichprobendaten handelt, die statistische Unsicherheiten aufweisen. Zudem wurden im IQB-Bildungstrend 2016 nur Förderschulen mit den Förderschwerpunkten LSE einbezogen, während an allgemeinen Schulen auch Schülerinnen und Schüler mit anderen Förderschwerpunkten an den Testungen teilnahmen (vgl. Kapitel 3.1).

Tabelle 4.3: Anteile zielgleich und zieldifferent unterrichteter Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf nach Fach in der 4. Jahrgangsstufe in den Ländern

Land	2011				2016			
	zielgleich unterrichtete SuS mit SPF		zieldifferent unterrichtete SuS mit SPF		zielgleich unterrichtete SuS mit SPF		zieldifferent unterrichtete SuS mit SPF	
	Deutsch	Mathematik	Deutsch	Mathematik	Deutsch	Mathematik	Deutsch	Mathematik
	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %	Anteil in %
Baden-Württemberg	4	4	< 1	< 1	3	3	2	2
Bayern	4	4	2	2	4	4	3	3
Berlin	5	4	2	2	5	5	3	3
Brandenburg	5	5	2	2	3	3	3	3
Bremen	3	3	5	5	3	3	6	6
Hamburg	4	4	2	2	4	4	4	4
Hessen	2	2	2	2	2	2	2	2
Mecklenburg-Vorpommern	7	7	5	4	4	3	4	4
Niedersachsen	2	2	2	2	3	3	3	3
Nordrhein-Westfalen	3	3	1	1	3	3	4	4
Rheinland-Pfalz	1	1	2	2	1	2	3	3
Saarland	2	2	3	3	4	5	3	3
Sachsen	4	3	3	3	4	4	4	4
Sachsen-Anhalt	4	4	2	2	4	4	3	3
Schleswig-Holstein	3	3	3	3	1	1	4	4
Thüringen	5	5	< 1	< 1	2	3	3	3
Deutschland	3	3	2	2	3	3	3	3

Anmerkungen. SuS = Schülerinnen und Schüler; SPF = sonderpädagogischer Förderbedarf. Alle Werte sind ganzzahlig gerundet. Die Angaben beziehen sich überwiegend auf Schülerinnen und Schüler der Förderschwerpunkte „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ (vgl. Kapitel 3.1) und geben den Anteil dieser Schülerinnen und Schüler an allen Viertklässlerinnen und Viertklässlern in den Ländern wieder.

Quelle: Daten des IQB-Ländervergleichs 2011 und des IQB-Bildungstrends 2016.

Deutschland insgesamt klein und hat sich gegenüber dem ersten Testzeitpunkt im Schuljahr 2010/2011 nur geringfügig erhöht (+1 Prozentpunkt). Auf Länderebene variiert der Anteil zum Testzeitpunkt im Jahr 2016 weitgehend zwischen 3 und 4 Prozent; Ausnahmen bilden Baden-Württemberg und Hessen mit einem Anteil von je etwa 2 Prozent und Bremen mit rund 6 Prozent. Auch wenn der Anteil zieldifferent unterrichteter Schülerinnen und Schüler in allen Ländern gering ist, sollte bei der Interpretation der Ergebnisse zum Erreichen der Bildungsstandards berücksichtigt werden, dass in den einzelnen Ländern geringfügig unterschiedliche Anteile von Kindern aus den Analysen ausgeschlossen wurden.

4.2.2 Unterrichtszeit

Der Umfang schulischer Lerngelegenheiten der Schülerinnen und Schüler in der Primarstufe hängt insbesondere von den Vorgaben der Länder zur Anzahl der Unterrichtsstunden ab. In Tabelle 4.4 wird dargestellt, wie viele Stunden Unterricht in den Fächern Deutsch und Mathematik gemäß den Stundentafeln der Länder an allgemeinen Schulen in den Jahrgangsstufen 1 bis 4 mindestens erteilt werden sollen.⁶ Die Angaben in Tabelle 4.4 zeigen, dass die Regelungen zur Unterrichtszeit in der Primarstufe zwischen den Ländern erheblich variieren, sich jedoch seit dem Schuljahr 2010/2011 nicht verändert haben. Eine Ausnahme bildet hierbei Hamburg mit einer Erhöhung der Jahreswochenstundenzahl um 5 Stunden im Fach Deutsch und um 3 Stunden im Fach Mathematik. Zu berücksichtigen ist, dass die Übersicht zu den Stundenvorgaben – sofern nicht anders gekennzeichnet – nur die vorgegebenen Mindeststundenzahlen wiedergibt. Unberücksichtigt bleibt dabei, dass die Schulen in einigen Ländern innerhalb eines bestimmten Rahmens zusätzliche Stunden für bestimmte Fächer vorsehen können.⁷

Nicht in Tabelle 4.4 enthalten sind Angaben zu Förderschulen. Hier gilt in der Mehrzahl der Länder die Maßgabe, dass für Schülerinnen und Schüler an Förderschulen mit dem Schwerpunkt „Sprache“ oder „emotionale und soziale Entwicklung“, die zielgleich unterrichtet werden, in den Jahrgangsstufen 1 bis 4 die Stundentafel der Grundschule gilt. Für Schülerinnen und Schüler an Förderschulen mit dem Schwerpunkt „Lernen“, die zieldifferent unterrichtet werden, gibt es hingegen in zahlreichen Ländern von der Stundentafel der allgemeinen Schule abweichende Regelungen. An diesen Förderschulen ist im Vergleich zu den allgemeinen Schulen im Primarbereich je nach Land meist weniger Unterrichtszeit für die Fächer Deutsch und Mathematik vorgegeben und/oder die Schulen verfügen über einen größeren Gestaltungsspielraum bei der Festlegung der Stundentafel. Zudem ist an Förderschulen im Schwerpunkt „Lernen“ häufiger fächerübergreifender Unterricht vorgesehen, sodass keine fachspezifischen Vorgaben gemacht werden.

6 Die Herausgeberinnen und Herausgeber danken dem Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland für die aufwändige Recherche und Zusammenstellung von Länderinformationen zu den Stundentafeln. Unser Dank gilt insbesondere Frau Dr. Andrea Schwermer.

7 In einigen Ländern weist die Stundentafel keine separaten Stundenzahlen für die einzelnen Fächer pro Schuljahr aus, sondern ist als sogenannte Kontingentstundentafel abgefasst. Die Kontingentstundentafel legt fest, wie viele Jahreswochenstunden insgesamt innerhalb mehrerer Schuljahre zu erteilen sind. Wie diese Jahreswochenstunden auf die einzelnen Jahrgangsstufen beziehungsweise Fächer verteilt werden, entscheiden die einzelnen Schulen innerhalb eines bestimmten Rahmens selbst. Sie erhalten damit pädagogischen Freiraum und können die Verteilung der Stunden nutzen, um Schwerpunkte zu setzen und Schulkonzepte zu gestalten.

Tabelle 4.4: Summe der vorgesehenen Unterrichtsstunden in den Jahrgangsstufen 1 bis 4 an allgemeinen Schulen in den Schuljahren 2010/2011 und 2015/2016 in den Fächern Deutsch und Mathematik nach Land (Angaben in Jahreswochenstunden)

Land	Schuljahr 2010/2011		Schuljahr 2015/2016	
	Deutsch	Mathematik	Deutsch	Mathematik
Baden-Württemberg	26	19	26	19
Bayern	23 ¹	19 ¹	23 ¹	19 ¹
Berlin ²	27	20	27	20
Brandenburg ²	25	18	25	18
Bremen	23	22	23	22
Hamburg	18	18	23	21
Hessen	22	20	22	20
Mecklenburg-Vorpommern	26	22	26	22
Niedersachsen	24	21	24	21
Nordrhein-Westfalen	53 ³		53 ³	
Rheinland-Pfalz ⁴	31 ⁵	18	31 ⁵	18
Saarland	20	20	20	20
Sachsen	27	20	27	20
Sachsen-Anhalt	24 ⁶	20 ⁶	24 ⁶	20 ⁶
Schleswig-Holstein	24	20	24	20
Thüringen	42 ⁷		42 ⁷	

Anmerkungen. Sofern nicht anders vermerkt, handelt es sich bei den Angaben um die vorgesehenen Mindeststundenzahlen. ¹In den Jahrgangsstufen 1 und 2 werden die Fächer Deutsch und Mathematik im Rahmen des Grundlegenden Unterrichts erteilt. Dieser umfasst insgesamt 16 Wochenstunden und beinhaltet die Fächer Deutsch, Mathematik, Heimat- und Sachunterricht, Kunst und Musik. Die Stunden für das Fach Deutsch bzw. Mathematik für die Jahrgangsstufen 1 und 2 wurden jeweils anteilig bestimmt und in die Stundenzahl für die Jahrgangsstufen 1 bis 4 eingerechnet. ²Die Primarstufe umfasst in der Regel die Jahrgangsstufen 1 bis 6. Es wurden hier jedoch, wie in den anderen Ländern, nur die in den Jahrgangsstufen 1 bis 4 vorgesehenen Stunden berücksichtigt. ³Die Fächer Deutsch, Sachunterricht, Mathematik und Förderunterricht werden fächerübergreifend erteilt; fachbezogene Vorgaben zu den Unterrichtsstunden sind nicht vorgesehen. Die angegebene Stundenzahl bezieht sich daher auf alle genannten Fächer zusammen. ⁴In der Grundschule beträgt die Dauer einer Unterrichtsstunde 50 Minuten. ⁵Die angegebenen Stunden bilden den Lernbereich Deutsch/Sachunterricht ab. ⁶Die Vorgaben der Stundentafel beziehen sich auf den Fächerkomplex Deutsch/Mathematik/Sachkunde (im Schuljahr 2015/2016) bzw. Deutsch/Sachkunde (im Schuljahr 2010/2011). Die Stunden für das Fach Deutsch bzw. das Fach Mathematik wurden jeweils anteilig bestimmt. ⁷Die Vorgaben der Stundentafel beziehen sich auf den Fächerkomplex Deutsch/Mathematik. Die angegebene Stundenzahl bezieht sich daher auf beide Fächer zusammen.

Quelle: Recherche des Sekretariats der KMK anhand des in den Stundentafeln der Länder ausgewiesenen Unterrichts an Schulen des Primarbereichs.

4.2.3 Einschulungsalter und Alter zum Testzeitpunkt

Weitere Kontextinformationen, die sich auf das Alter der Kinder zum Zeitpunkt der Einschulung sowie zum Zeitpunkt der Testung in der 4. Jahrgangsstufe beziehen oder damit in Zusammenhang stehen, werden in Tabelle 4.5 dargestellt. So sind für die Einschulung in den Ländern teilweise unterschiedliche Stichtage maßgeblich. Für die Zielpopulation des IQB-Ländervergleichs 2011 waren dafür die im Jahr 2007 geltenden Regelungen entscheidend, für die Zielpopulation des IQB-Bildungstrends 2016 die Regelungen im Jahr 2012. Die im Jahr 2012 geltenden Regelungen sahen in acht Ländern den Stichtag 30. Juni vor. In diesen Ländern wurden also Kinder in der Regel eingeschult, wenn sie bis einschließlich 30. Juni ihr sechstes Lebensjahr vollendet hatten. In den übrigen Ländern galten im Jahr 2012 spätere Stichtage. Da in Berlin sowohl im Jahr 2007 als auch im Jahr 2012 der im Vergleich mit den anderen Ländern späteste Stichtag

Tabelle 4.5: Stichtage zur Einschulung, Anteile vorzeitiger und verspäteter Einschulungen sowie mittleres Alter zum Testzeitpunkt bezogen auf die in den Jahren 2011 und 2016 getesteten Schülerpopulationen nach Land

Land	Stichtag Einschulung		vorzeitige Einschulungen ¹		verspätete Einschulungen ¹		Alter zum Testzeitpunkt ²	
	2007	2012	2007	2012	2007	2012	2011	2016
			Anteil in %		Anteil in %		Mittelwert	
Baden-Württemberg	30.09.	30.09.	3.1	2.4	8.5	9.8	10;4	10;5
Bayern	30.09.	30.09.	3.0	2.3	7.9	11.6	10;5	10;4
Berlin	31.12.	31.12.	2.4	1.0	3.0	8.2	10;2	10;3
Brandenburg	30.09.	30.09.	2.8	1.8	12.5	11.0	10;5	10;5
Bremen	30.06.	30.06.	17.5	15.0	2.3	1.9	10;5	10;5
Hamburg	30.06.	30.06.	13.3	9.9	3.4	2.2	10;5	10;5
Hessen	30.06.	30.06.	12.6	7.2	10.1	9.8	10;6	10;7
Mecklenburg- Vorpommern	30.06.	30.06.	2.8	2.5	6.3	6.2	10;7	10;8
Niedersachsen	30.06.	30.09.	7.9	1.7	5.6	5.7	10;6	10;5
Nordrhein-Westfalen	31.07.	30.09.	7.9	2.9	0.8	2.4	10;6	10;5
Rheinland-Pfalz	30.06.	31.08.	11.6	4.6	7.3	4.7	10;6	10;5
Saarland	30.06.	30.06.	7.9	6.9	5.6	3.0	10;6	10;6
Sachsen	30.06.	30.06.	0.7	0.2	3.9	5.8	10;7	10;7
Sachsen-Anhalt	30.06.	30.06.	2.7	1.6	0.7	2.0	10;7	10;8
Schleswig-Holstein	30.06.	30.06.	7.2	5.8	4.9	1.6	10;6	10;7
Thüringen	31.07.	31.07.	2.0	1.8	6.4	6.9	10;7	10;7
Deutschland			6.2	3.1	5.5	6.6	10;6	10;5

Anmerkungen. ¹ Anteil der vorzeitigen bzw. verspäteten Einschulungen an allen Einschulungen des Jahres 2007 bzw. 2012.

² Mittleres Alter angegeben als „Jahr;Monat“ basierend auf den Daten der Viertklässlerinnen und Viertklässler im IQB-Ländervergleich 2011 bzw. im IQB-Bildungstrend 2016.

Quelle: Nationaler Bildungsbericht 2016 (Autorengruppe Bildungsberichterstattung, 2016), Angaben des Sekretariats der KMK sowie Daten des IQB-Ländervergleichs 2011 und des IQB-Bildungstrends 2016.

(31. Dezember) galt, war der Anteil der Erstklässlerinnen und Erstklässler, die zum Zeitpunkt der Einschulung das sechste Lebensjahr noch nicht vollendet hatten, in Berlin jeweils besonders hoch. Zwischen dem Jahr 2007 und dem Jahr 2012 haben sich nur in Niedersachsen, Nordrhein-Westfalen und Rheinland-Pfalz die Regelungen geändert. In diesen Ländern wurde der Stichtag für die Einschulung auf einen späteren Termin verlegt, sodass im Jahr 2012 die Schülerinnen und Schüler zum Zeitpunkt der Einschulung im Durchschnitt etwas jünger waren.

Obgleich die Einschulung in der Regel gemäß der oben genannten Stichtagsregelungen erfolgt, sind je nach Land unter bestimmten Bedingungen sowohl frühere als auch spätere Einschulungen möglich. In Tabelle 4.5 ist daher auch der Anteil der vorzeitigen und der verspäteten Einschulungen an allen Einschulungen des Jahres 2007 beziehungsweise 2012 angegeben. *Ceteris paribus* hat ein hoher Anteil vorzeitiger Einschulungen ein im Mittel geringeres Lebensalter und ein hoher Anteil verspäteter Einschulungen ein im Mittel höheres Lebensalter zum Testzeitpunkt zur Folge. Wie Tabelle 4.5 zeigt, variieren die Anteile vorzeitiger und verspäteter Einschulungen je nach Land teilweise erheblich. So erfolgten im Jahr 2012 rund 15 Prozent der Einschulungen im Land Bremen vorzeitig, wäh-

rend die entsprechende Quote in Sachsen, wo ebenfalls der Stichtag 30. Juni gilt, bei unter einem Prozent lag. Auch der Anteil verspäteter Einschulungen variierte im Jahr 2012 zwischen fast 12 Prozent in Bayern und weniger als 2 Prozent in Bremen sowie Schleswig-Holstein, wobei in Bayern ein späterer Stichtag galt. Stellt man die Anteile vorzeitiger und verspäteter Einschulungen im Jahr 2012 denen des Jahres 2007 gegenüber, wird deutlich, dass sich die Anteile in den Ländern teilweise verändert haben, wobei kein einheitliches Muster zu erkennen ist.

Da die vorgenannten Regelungen und Unterschiede zwischen den Ländern Auswirkungen auf das Alter der Kinder zum Testzeitpunkt gegen Ende der 4. Jahrgangsstufe haben, werden in Tabelle 4.5 zusätzlich die anhand der jeweiligen Datensätze berechneten Altersmittelwerte für die Zielpopulationen im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 angegeben.⁸ In keinem Land unterscheidet sich der Altersmittelwert der Schülerpopulation des Jahres 2016 von der des Jahres 2011 zum Testzeitpunkt um mehr als einen Monat. Allerdings variieren die Altersmittelwerte teilweise geringfügig zwischen den Ländern.

4.2.4 Besuch von Ganztagschulen

Abschließend soll darauf eingegangen werden, inwieweit in den Ländern Ganztagsangebote bereitgestellt werden und welche Anteile der Grundschülerinnen und Grundschüler in den Ländern am Ganztagsbetrieb teilnehmen. Bei Ganztagschulen handelt es sich nach der Definition der Kultusministerkonferenz um „Schulen, bei denen [...]

- an mindestens drei Tagen in der Woche ein ganztägiges Angebot für die Schülerinnen und Schüler bereitgestellt wird, das täglich mindestens sieben Zeitstunden umfasst;
- an allen Tagen des Ganztagschulbetriebs den teilnehmenden Schülerinnen und Schülern ein Mittagessen bereitgestellt wird;
- die Ganztagsangebote unter der Aufsicht und Verantwortung der Schulleitung organisiert und in enger Kooperation mit der Schulleitung durchgeführt werden sowie in einem konzeptionellen Zusammenhang mit dem Unterricht stehen.“ (KMK, 2017, S. 4)

Unabhängig von dieser Definition werden je nach Land teilweise verschiedene Bezeichnungen für Ganztagschulen verwendet. Zudem bestehen Unterschiede in weiteren organisatorischen und inhaltlichen Aspekten, die an dieser Stelle nicht im Detail dargestellt werden können (vgl. hierzu etwa KMK, 2015). Eine wesentliche Differenzierung, die hier aufgegriffen werden soll, betrifft allerdings die Frage, ob die Teilnahme am Ganztagsbetrieb für die Kinder freiwillig oder verpflichtend ist.

⁸ Berechnungsgrundlage waren Angaben der Schulen zu Geburtsmonat und -jahr der Kinder. Für die Auswertung wurde der Testzeitpunkt länderübergreifend auf Juni 2011 bzw. Juni 2016 festgelegt. Der tatsächliche Testzeitpunkt variierte aus organisatorischen Gründen geringfügig innerhalb und zwischen den Ländern (vgl. Kapitel 3.1).

Tabelle 4.6: Anteil der Ganztagschulen an allen Grundschulen in den Schuljahren 2010/2011 und 2015/2016 in den Ländern

Land	Schuljahr 2010/2011				Schuljahr 2015/2016			
	Anteil der Ganztagschulen (in %)				Anteil der Ganztagschulen (in %)			
	insgesamt	in voll gebundener Form	in teilweise gebundener Form	in offener Form	insgesamt	in voll gebundener Form	in teilweise gebundener Form	in offener Form
Baden-Württemberg	12.7	2.2	3.3	7.1	21.9	2.7	2.7	16.5
Bayern	23.5	0.7	8.6	14.2	25.4	0.9	15.2	9.3
Berlin	97.2	10.9	5.1	81.3	98.4	11.3	5.9	81.2
Brandenburg	47.4	–	–	47.4	48.8	–	–	48.8
Bremen	24.8	20.8	–	4.0	50.0	27.6	–	22.4
Hamburg	26.6	10.6	3.2	12.8	97.7	14.1	9.1	74.5
Hessen ¹	22.2	0.5	–	21.7	46.1	0.7	–	45.4
Mecklenburg-Vorpommern	4.0	2.2	0.3	1.5	26.1	12.7	4.0	9.3
Niedersachsen ¹	24.5	0.4	0.4	23.7	53.1	0.1	1.0	52.0
Nordrhein-Westfalen	83.5	0.5	–	83.0	92.2	0.5	–	91.7
Rheinland-Pfalz	56.0	0.7	26.4	28.9	69.2	0.9	32.7	35.5
Saarland	98.1	0.6	1.2	96.3	98.1	4.9	–	93.2
Sachsen	99.3	3.2	18.9	77.2	98.9	4.5	26.2	68.3
Sachsen-Anhalt ^{1,2}	3.5	0.6	–	2.9	4.0	0.7	–	3.3
Schleswig-Holstein	35.2	1.1	0.7	33.5	56.1	1.2	0.8	54.1
Thüringen	100.0	6.6	1.9	91.5	100.0	5.3	1.6	93.1
Deutschland	44.7	1.7	4.7	38.2	56.1	2.2	6.8	47.0

Anmerkungen. Da gerundete Werte angegeben werden, kann die Summe der Anteile für die unterschiedlichen Formen der Ganztagschule geringfügig vom Anteil der Ganztagschulen insgesamt abweichen. Zur Vermeidung von Doppelzählungen werden Schulen, die sowohl ein Ganztagsangebot in teilweise gebundener als auch in offener Form anbieten, ausschließlich bei den Ganztagschulen in teilweise gebundener Form gezählt. ¹Für Hessen, Niedersachsen und Sachsen-Anhalt liegen keine Angaben über Ganztagsangebote an Schulen in freier Trägerschaft vor. Die angegebenen Anteile beziehen sich auf Schulen in öffentlicher Trägerschaft. ²In Sachsen-Anhalt bestehen an nahezu allen Grundschulstandorten schulbezogene Ganztagsangebote, zumeist durch Kooperationen mit Horteinrichtungen.

Quelle: KMK (2016, 2017).

Es werden dazu drei Formen der Ganztagschule unterschieden:

- „In der *voll gebundenen Form* sind *alle* Schülerinnen und Schüler verpflichtet, an mindestens drei Wochentagen für jeweils mindestens sieben Zeitstunden an den ganztägigen Angeboten der Schule teilzunehmen.
- In der *teilweise gebundenen Form* verpflichtet sich *ein Teil* der Schülerinnen und Schüler (z. B. einzelne Klassen oder Klassenstufen), an mindestens drei Wochentagen für jeweils mindestens sieben Zeitstunden an den ganztägigen Angeboten der Schule teilzunehmen.
- In der *offenen Form* können *einzelne* Schülerinnen und Schüler auf Wunsch an den ganztägigen Angeboten [...] teilnehmen. Für die Schülerinnen und Schüler ist ein Aufenthalt, verbunden mit einem Bildungs- und Betreuungsangebot in der Schule, an mindestens drei Wochentagen im Umfang von täglich mindestens sieben Zeitstunden möglich.“ (KMK, 2017, S. 5; Hervorhebungen im Original)

In Tabelle 4.6 ist zunächst basierend auf der Statistik der KMK für die einzelnen Länder angegeben, welcher Anteil der Grundschulen⁹ gemäß der oben genannten Definition der KMK allgemeinbildende Schulen in Ganztagsform sind. Im Schuljahr 2015/2016 handelt es sich demnach bei rund 56 Prozent der Grundschulen in Deutschland um Ganztagschulen. Dieser Anteil variiert auf Länderebene allerdings erheblich zwischen knapp 22 Prozent in Baden-Württemberg¹⁰ und 100 Prozent in Thüringen. Im Vergleich zum Schuljahr 2010/2011 hat sich der Anteil der Ganztagschulen im Primarbereich in Deutschland insgesamt deutlich erhöht (+11 Prozentpunkte). Dies gilt auch für viele der 16 Länder. In keinem Land ist der Anteil der Ganztagschulen nennenswert zurückgegangen.

Wird nach der Form der Ganztagschule differenziert, so zeigt sich gemäß Tabelle 4.6, dass die Grundschulen meist einen Ganztagsbetrieb in offener (freiwilliger) Form anbieten. Dies gilt für beide der hier betrachteten Schuljahre. Wiederum sind jedoch Unterschiede zwischen den Ländern zu verzeichnen. In einigen Ländern weisen weniger als ein Prozent der Grundschulen einen Ganztagsbetrieb in voll oder teilweise gebundener Form auf und verpflichten somit alle oder einen Teil der Schülerinnen und Schüler zur Teilnahme am Ganztagsbetrieb. In anderen Ländern ist dies dagegen bei mehr als 20 Prozent der Grundschulen insgesamt der Fall und bei der Mehrzahl der Grundschulen mit Ganztagsangebot.

Da die Nutzung des Ganztagsangebots an vielen Schulen freiwillig ist, stellt sich die Frage, welcher Anteil der Grundschülerinnen und Grundschüler in der Primarstufe am Ganztagsbetrieb teilnimmt. Entsprechende Angaben, die ebenfalls auf der Statistik der KMK basieren, können Tabelle 4.7 entnommen werden, wobei analog zur Darstellung auf Schulebene auch nach der Form des Ganztagsbetriebs (offen versus gebunden) differenziert wird. Im Schuljahr 2015/2016 nahmen demnach in Deutschland insgesamt fast 35 Prozent der Schülerinnen und Schüler in Grundschulen am Ganztagsbetrieb teil. Dies stellt gegenüber dem Schuljahr 2010/2011 eine deutliche Steigerung dar (+12 Prozentpunkte). Ähnlich wie für die Ergebnisse auf Schulebene zeigen sich für beide dargestellten Schuljahre auch auf Schülerebene erhebliche Unterschiede zwischen den Ländern.

9 Integrierte Gesamtschulen bleiben unberücksichtigt, da zu dieser Schulart für den Primarbereich die entsprechenden Daten erst ab dem Schuljahr 2012/2013 vorliegen. Unberücksichtigt bleiben zudem die Förderschulen, da in den statistischen Daten für diese Schulart keine Differenzierung zwischen Primar- und Sekundarstufe vorgenommen wird.

10 Eine noch geringere Quote weist die Statistik der KMK für Sachsen-Anhalt aus. Diese wurde hier jedoch nicht berücksichtigt, da in Sachsen-Anhalt die Besonderheit vorliegt, dass an nahezu allen Grundschulstandorten schulbezogene Ganztagsangebote, zumeist in Kooperation mit Horteinrichtungen, bestehen (KMK, 2015). Zudem bezieht die Statistik für Sachsen-Anhalt nur Schulen in öffentlicher Trägerschaft ein.

Tabelle 4.7: Anteil der Schülerinnen und Schüler im Ganztagsbetrieb in der Grundschule in den Schuljahren 2010/2011 und 2015/2016 nach Land

Land	Schuljahr 2010/2011			Schuljahr 2015/2016		
	Anteil SuS im Ganztagsbetrieb (in %)			Anteil SuS im Ganztagsbetrieb (in %)		
	insgesamt	in gebundener Form	in offener Form	insgesamt	in gebundener Form	in offener Form
Baden-Württemberg	7.8	4.1	3.7	14.2	4.9	9.3
Bayern	6.3	2.5	3.8	11.1	6.6	4.4
Berlin	74.9	15.4	59.5	77.2	17.9	59.3
Brandenburg	42.5	–	42.5	41.0	–	41.0
Bremen	23.2	21.5	1.6	41.4	29.4	12.0
Hamburg	29.1	15.0	14.2	98.1	22.6	75.5
Hessen ¹	27.6	0.8	26.9	27.5	1.1	26.4
Mecklenburg-Vorpommern	2.4	1.5	0.9	2.6	2.0	0.6
Niedersachsen ¹	27.5	0.8	26.7	30.5	1.4	29.1
Nordrhein-Westfalen	28.7	0.5	28.2	42.4	0.5	41.9
Rheinland-Pfalz	21.6	15.4	6.2	33.7	22.2	11.5
Saarland	31.2	1.9	29.3	46.9	4.5	42.5
Sachsen	71.3	20.7	50.6	85.6	31.1	54.5
Sachsen-Anhalt ^{1,2}	4.1	1.0	3.1	4.1	1.0	3.1
Schleswig-Holstein	14.0	1.1	12.9	19.8	2.3	17.5
Thüringen	80.9	6.0	74.9	85.3	6.2	79.1
Deutschland	22.8	4.1	18.7	34.5	6.4	28.1

Anmerkungen. SuS = Schülerinnen und Schüler. Da gerundete Werte angegeben werden, kann die Summe der Anteile für die unterschiedlichen Formen des Ganztagsbetriebs geringfügig vom Anteil der insgesamt am Ganztagsbetrieb teilnehmenden Kinder abweichen. ¹Für Hessen, Niedersachsen und Sachsen-Anhalt liegen keine Angaben für Kinder an Schulen in freier Trägerschaft vor. Die angegebenen Anteile beziehen sich auf Kinder an Schulen in öffentlicher Trägerschaft. ²In Sachsen-Anhalt bestehen an nahezu allen Grundschulstandorten schulbezogene Ganztagsangebote, zumeist durch Kooperationen mit Horteinrichtungen. Die Angebote wurden im Schuljahr 2015/16 von etwa 68 Prozent aller Schülerinnen und Schüler der öffentlichen Grundschulen genutzt.

Quelle: KMK (2016, 2017).

4.3 Zusammenfassung

Die im vorliegenden Kapitel dargestellten Kontextinformationen zu den Schulsystemen der Länder im Primarbereich verdeutlichen, dass zwischen den Ländern zum Teil erhebliche Unterschiede im Anteil von Kindern mit sonderpädagogischem Förderbedarf und in deren Verteilung auf die Schularten (allgemeine Schule oder Förderschule) bestehen. Der Anteil der Kinder, die zieldifferent unterrichtet werden, ist – bezogen auf die Gesamtpopulation der Viertklässlerinnen und Viertklässler – in allen Ländern recht gering. Länderunterschiede bestehen ferner in der Anzahl der vorgesehenen Unterrichtsstunden in den Fächern Deutsch und Mathematik im Primarbereich, in den Regelungen zur Einschulung, im mittleren Alter der Viertklässlerinnen und Viertklässler zum Testzeitpunkt sowie im Ganztagsangebot und in der Teilnahme der Schülerinnen und Schüler am Ganztagsbetrieb.

Für einige dieser Merkmale sind Veränderungen zwischen dem Schuljahr 2010/2011 und dem Schuljahr 2015/2016 festzustellen, in dem die Erhebung zum IQB-Ländervergleich 2011 beziehungsweise IQB-Bildungstrend 2016 stattfand. Dies betrifft insbesondere den in Deutschland insgesamt deutlich gestiegenen Anteil an Viertklässlerinnen und Viertklässlern mit SPF, die eine allgemeine

Schule besuchen, sowie die in den meisten Ländern zunehmende Nutzung von Ganztagsangeboten im Primarbereich.

Insgesamt wird deutlich, dass sich die Schulsysteme der Länder im Primarbereich teilweise erheblich unterscheiden und in verschiedener Weise weiterentwickeln. Dies sollte bei der Interpretation der in den folgenden Kapiteln berichteten Ergebnisse berücksichtigt werden. Inwieweit die hier dargestellten oder auch weitere, hier nicht untersuchte Systemmerkmale für die Unterschiede in den Kompetenzen verantwortlich sind, die von den Schülerinnen und Schülern in den Ländern erreicht werden, lässt sich anhand der vorliegenden Daten allerdings nicht bestimmen.

Literatur

- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: W. Bertelsmann Verlag.
- Gresch, C. & Piezunka, A. (2015). Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich „Lernen“) an Regelschulen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 181–220). Wiesbaden: Springer VS.
- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer VS.
- KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Ganztagssschulen in Deutschland. Bericht der Kultusministerkonferenz vom 03.12.2015*. Zugriff am 14.08.2017 unter https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015-12-03-Ganztagsschulbericht.pdf
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Allgemeinbildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland. Statistik 2010 bis 2014*. Zugriff am 27.06.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/GTS_2014_Bericht_Text.pdf
- KMK (2017) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017). *Allgemeinbildende Schulen in Ganztagsform in den Ländern in der Bundesrepublik Deutschland. Statistik 2011 bis 2015*. Zugriff am 27.06.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/GTS_2015_Bericht.pdf
- Malecki, A. (2013). Sonderpädagogische Förderung in Deutschland – eine Analyse der Datenlage in der Schulstatistik. *Wirtschaft und Statistik*, 5, 356–365.
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (2013). *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.
- Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 129–152). Wiesbaden: Springer VS.
- Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Hrsg.). (2016). *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

Kapitel 5

Kompetenzstufenbesetzungen im Ländervergleich

5.1 Kompetenzstufenbesetzungen im Fach Deutsch

Sebastian Weirich, Julia Wittig und Petra Stanat

Im IQB-Bildungstrend 2016 wurden zur Erfassung sprachlicher Kompetenzen im Fach Deutsch Aufgaben zu den folgenden drei Kompetenzbereichen der Bildungsstandards für den Primarbereich eingesetzt (KMK, 2005):

- 1) *Lesen* – mit Texten und Medien umgehen
- 2) Sprechen und Zuhören, Teilbereich *Zuhören*
- 3) Schreiben, Teilbereich *Orthografie*

In diesem Kapitel wird erläutert, wie sich die Schülerinnen und Schüler in diesen drei Bereichen in den einzelnen Ländern auf die jeweiligen Kompetenzstufen verteilen. Die Kompetenzstufenmodelle, die diesen Analysen zugrunde liegen, werden in Kapitel 2.2 ausführlich dargestellt. Eine detaillierte Beschreibung der drei Kompetenzbereiche findet sich in Kapitel 1.2.

Im Folgenden werden zunächst die Kompetenzstufenbesetzungen in den Ländern für das Jahr 2016 berichtet, wobei insbesondere auf Unterschiede zwischen den Ländern eingegangen wird (Abschnitt 5.1.1). Im Anschluss wird der Frage nachgegangen, wie sich die Kompetenzstufenbesetzungen in den Ländern zwischen den Jahren 2011 und 2016 verändert haben (Abschnitt 5.1.2). Das Kapitel schließt mit einer Zusammenfassung der zentralen Befunde (Abschnitt 5.1.3).

5.1.1 Kompetenzstufenbesetzungen in den Ländern im Jahr 2016

Die zu den Bildungsstandards der Kultusministerkonferenz (KMK) entwickelten Kompetenzstufenmodelle des IQB für den Primarbereich umfassen jeweils fünf Kompetenzstufen als Grundlage für die Definition von Mindest-, Regel- und Optimalstandards. Auf Kompetenzstufe III, dem *Regelstandard*, werden Kompetenzen beschrieben, über die Schülerinnen und Schüler am Ende der 4. Jahrgangsstufe in der Regel verfügen sollen. Die Kompetenzstufe II entspricht hingegen dem *Mindeststandard*; Schülerinnen und Schüler, die Stufe II nicht erreichen, verfehlen demnach die länderübergreifend festgelegten Minimalanforderungen im Primarbereich. Schülerinnen und Schüler, die dagegen die Kompetenzstufe IV, den *Regelstandard plus*, oder die Kompetenzstufe V, den

Optimalstandard, erreichen, übertreffen die Regelanforderungen der KMK für die Primarstufe zum Teil erheblich.

Im Folgenden wird für alle drei Kompetenzbereiche und alle Länder sowie für Deutschland insgesamt jeweils der Anteil der Schülerinnen und Schüler berichtet, die

- a) den Mindeststandard verfehlen,
- b) den Regelstandard erreichen oder übertreffen beziehungsweise
- c) den Optimalstandard erreichen.¹

Zur Interpretation von Länderunterschieden können Angaben zu den in einem Schuljahr zu erwartenden Veränderungen in den Kompetenzstufenverteilungen herangezogen werden. Entsprechende Schätzungen lassen sich aus einer repräsentativen Normierungsstudie des IQB aus dem Jahr 2007 mit Schülerinnen und Schülern der 3. und 4. Jahrgangsstufe ableiten (Granzer et al., 2009). Die im Vergleich der beiden Jahrgangsstufen festgestellten Unterschiede in den Kompetenzstufenbesetzungen variieren zwischen den einzelnen Kompetenzbereichen. So liegt der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, in der 4. Jahrgangsstufe im *Lesen* und im *Zuhören* um etwa 20 Prozentpunkte und in der *Orthografie* um etwa 40 Prozentpunkte höher als in der 3. Jahrgangsstufe.

Die Besetzung der Kompetenzstufen im *Lesen* für das Jahr 2016 ist in Abbildung 5.1 dargestellt. Die Population, auf die sich die in dieser Abbildung dargestellten Werte beziehen, umfasst alle Viertklässlerinnen und Viertklässler in Deutschland, die zielgleich auf Grundlage der Bildungsstandards unterrichtet werden. Nicht berücksichtigt sind zieldifferent unterrichtete Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, da für sie andere Bildungsziele (Lehrpläne, Rahmenlehrpläne oder Kerncurricula) und somit auch andere als die in den Bildungsstandards der KMK definierten Anforderungen gelten.

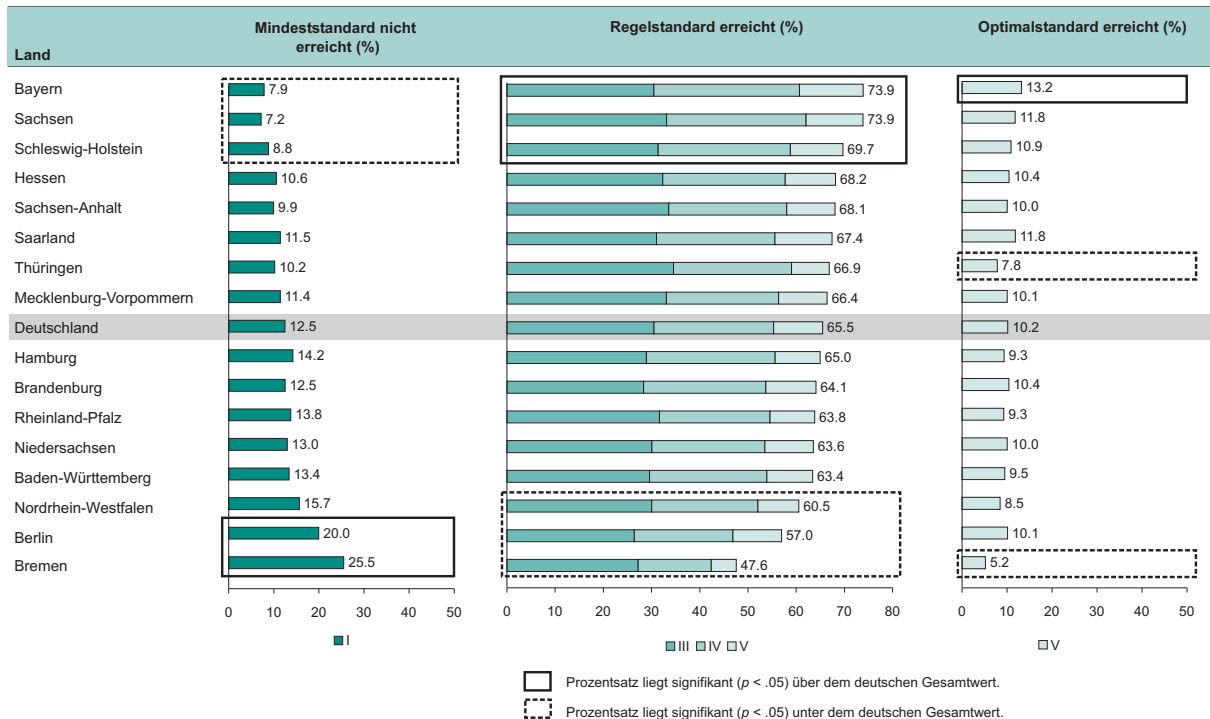
Abbildung 5.1 enthält drei Spalten mit Balkendiagrammen, die für die einzelnen Länder sowie für Deutschland insgesamt den Anteil der Schülerinnen und Schüler angeben, die den Mindeststandard verfehlen (Kompetenzstufe I), den Regelstandard erreichen oder übertreffen (Kompetenzstufe III, IV und V) beziehungsweise den Optimalstandard erreichen (Kompetenzstufe V). In den Balken zum Erreichen des Regelstandards sind auch die prozentualen Anteile für die einzelnen Kompetenzstufen gekennzeichnet, die sich zu der Gruppe „mindestens Regelstandard erreicht“ (Kompetenzstufen III, IV und V) summieren. Das Diagramm weist ferner aus, für welche Länder der jeweilige Prozentanteil statistisch signifikant ($p < .05$) vom Anteil für Deutschland insgesamt² abweicht. Liegt der Anteil eines Landes signifikant über dem deutschen Gesamtwert, ist dies durch einen Rahmen mit durchgezogener Linie gekennzeichnet. Länder, deren Prozentanteile signifikant unter dem Gesamtwert liegen, sind durch einen Rahmen mit gestrichelter Linie kenntlich gemacht.

Dabei sollte der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, möglichst hoch ausfallen, das heißt über dem Gesamtwert für Deutschland liegen. Analog gilt dies auch für das Erreichen des Optimalstandards. Demgegenüber ist es wünschenswert, dass der Anteil derjeni-

1 Die vollständigen Verteilungen auf alle fünf Kompetenzstufen können für die drei Kompetenzbereiche auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

2 Bei der nachfolgenden Ergebnisdarstellung im Text wird dieser Anteil als „(deutscher) Gesamtwert“ bezeichnet.

Abbildung 5.1: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Lesen* im Fach Deutsch den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen



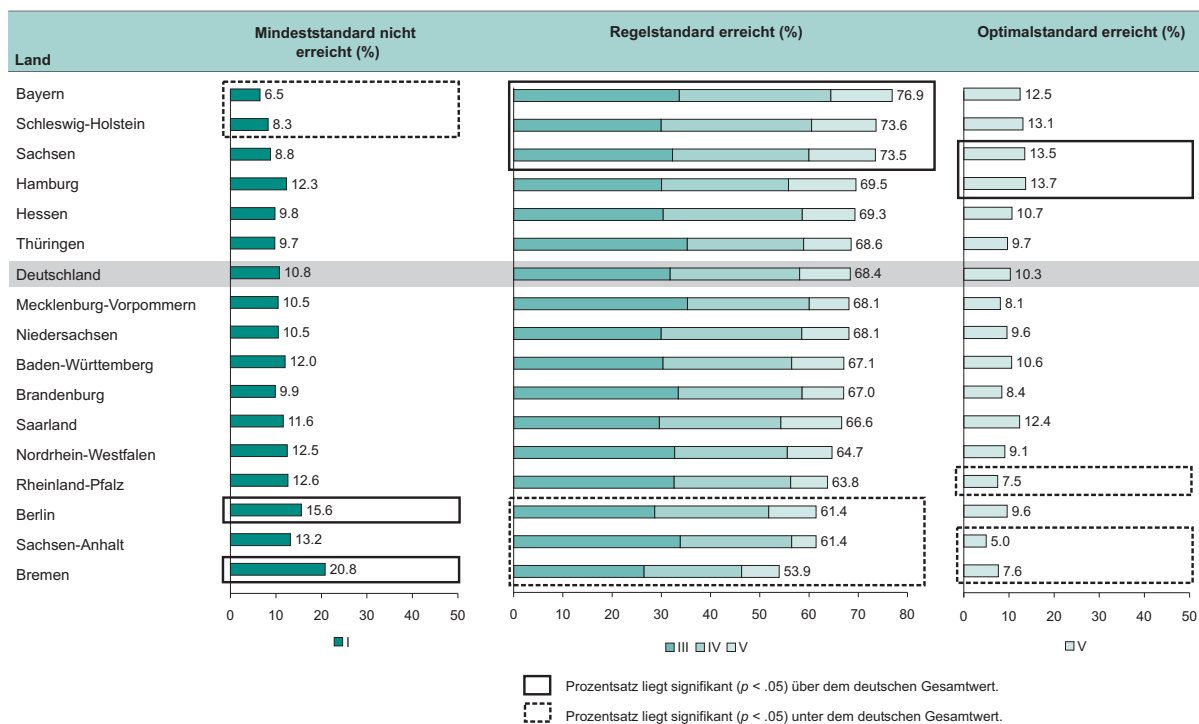
gen Schülerinnen und Schüler, die den Mindeststandard verfehlen, so gering wie möglich ist, das heißt unter dem Gesamtwert liegt.

Aus Abbildung 5.1 geht hervor, dass die prozentualen Anteile der Viertklässlerinnen und Viertklässler, die im Kompetenzbereich *Lesen* den Mindeststandard verfehlen, mindestens den Regelstandard erreichen beziehungsweise den Optimalstandard erreichen, zwischen den Ländern in der Bundesrepublik Deutschland stark variieren. Für den Regelstandard beträgt die Differenz zwischen Bayern, dem Land mit dem höchsten Anteil, und Bremen, dem Land mit dem geringsten Anteil, etwa 26 Prozentpunkte. In den Ländern Bayern, Sachsen und Schleswig-Holstein fällt der Anteil der Schülerinnen und Schüler, die im *Lesen* mindestens den Regelstandard erreichen, signifikant höher aus als der deutsche Gesamtwert (66 %). Signifikant unter dem Gesamtwert liegen hingegen die Anteile für die Länder Nordrhein-Westfalen, Berlin und Bremen.

Die Ergebnisse zum Mindeststandard zeigen, dass in Deutschland insgesamt die Lesekompetenz von fast 13 Prozent der Viertklässlerinnen und Viertklässler nicht den Minimalanforderungen für den Primarbereich entspricht. Die Differenz zwischen Sachsen, dem Land mit dem niedrigsten Anteil (7 %) und Bremen, dem Land mit dem höchsten Anteil (26 %) beträgt rund 19 Prozentpunkte. Signifikant höher als der deutsche Gesamtwert fällt der Anteil der Schülerinnen und Schüler, die im *Lesen* den Mindeststandard verfehlen, in den Ländern Berlin und Bremen aus. Signifikant niedriger als der Gesamtwert sind die für die Länder Bayern, Sachsen und Schleswig-Holstein festgestellten Ergebnisse.

Den Optimalstandard im Kompetenzbereich *Lesen* erreichen in Bayern 13 Prozent der Viertklässlerinnen und Viertklässler, jedoch nur 5 Prozent in Bremen, was einer Differenz von 8 Prozentpunkten entspricht. Signifikant höher als der deutsche Gesamtwert von 10 Prozent fällt der Anteil nur in Bayern aus. Signifikant niedriger als der Gesamtwert ist der Anteil in Thüringen und Bremen.

Abbildung 5.2: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Zuhören* im Fach Deutsch den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen

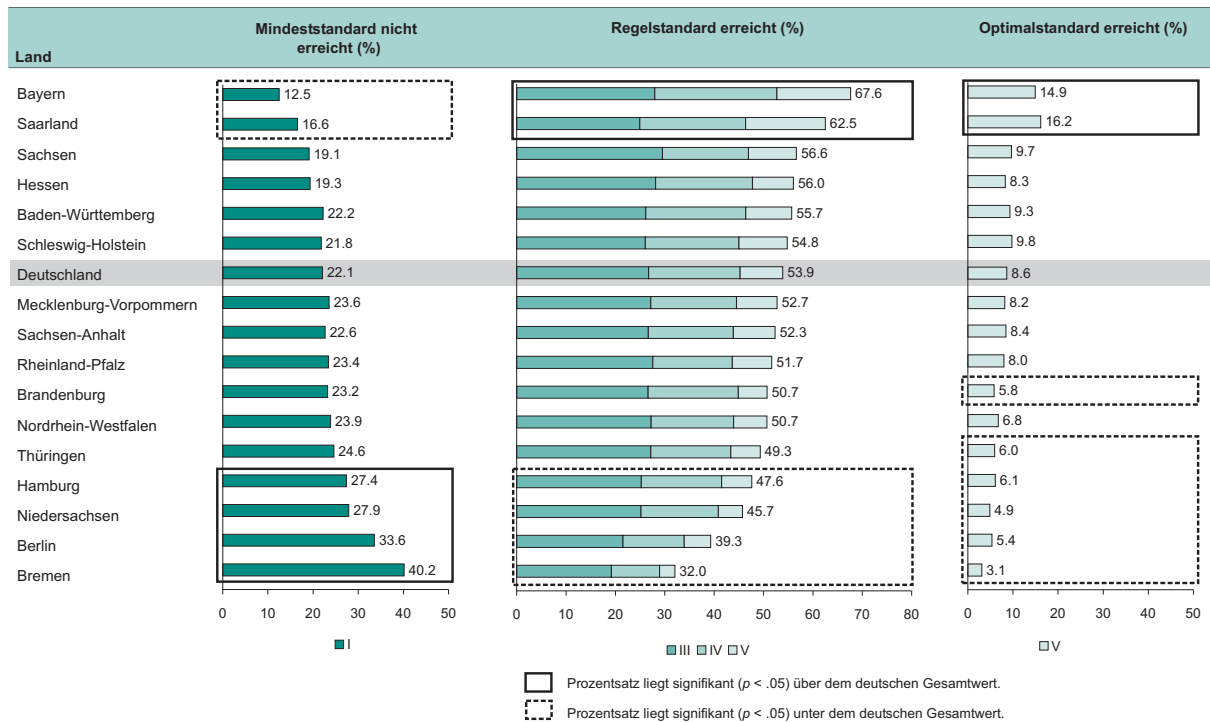


Angaben zur Kompetenzstufenbesetzung im Bereich *Zuhören* für das Jahr 2016 finden sich in Abbildung 5.2, die analog zur Ergebnisdarstellung für den Kompetenzbereich *Lesen* strukturiert ist. Auch hier beziehen sich die berichteten Anteile auf alle Schülerinnen und Schüler der 4. Jahrgangsstufe, die zielgleich unterrichtet werden. Die Länderunterschiede im *Zuhören* fallen ähnlich groß aus wie für den Kompetenzbereich *Lesen*. Bei den für das Erreichen oder Übertreffen des Regelstandards ermittelten Anteilen beträgt die Differenz zwischen dem höchsten Landeswert (Bayern) und dem niedrigsten Landeswert (Bremen) 23 Prozentpunkte. In Deutschland insgesamt ist der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard im *Zuhören* erreichen, mit insgesamt 68 Prozent etwas höher als im *Lesen*. Signifikant über dem deutschen Gesamtwert liegen die prozentualen Anteile in den Ländern Bayern, Schleswig-Holstein und Sachsen; signifikant darunter die Anteile in Berlin, Sachsen-Anhalt und Bremen.

Betrachtet man die prozentualen Anteile der Viertklässlerinnen und Viertklässler, die die Mindestanforderungen im *Zuhören* in der Primarstufe verfehlen, beträgt der Unterschied zwischen dem geringsten Wert in Bayern (7 %) und dem höchsten Wert in Bremen (21 %) rund 14 Prozentpunkte. Insgesamt fällt der Anteil der Schülerinnen und Schüler, die im *Zuhören* den Mindeststandard verfehlen, mit 11 Prozent etwas niedriger aus als für den Bereich *Lesen*. In Berlin und Bremen ist dieser Anteil signifikant höher als der Gesamtwert für Deutschland, in Bayern und Schleswig-Holstein signifikant geringer.

Der Anteil der Schülerinnen und Schüler, die im *Zuhören* den Optimalstandard erreichen, beträgt für Deutschland insgesamt 10 Prozent. Signifikant über diesem Wert liegen die Anteile in Sachsen und Hamburg; signifikant darunter die Anteile in Rheinland-Pfalz, Sachsen-Anhalt und Bremen.

Abbildung 5.3: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Orthografie* im Fach Deutsch den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen



Die Ergebnisse zur Kompetenzstufenbesetzung im Bereich *Orthografie* können der Abbildung 5.3 entnommen werden. Die Spannweite der für die einzelnen Länder ermittelten Anteile ist hier besonders groß. So beträgt die Differenz zwischen dem Land mit dem höchsten (Bayern) und dem Land mit dem geringsten (Bremen) Anteil an Schülerinnen und Schülern, die mindestens den Regelstandard erreichen, 36 Prozentpunkte. Mit 54 Prozent fällt der für Deutschland insgesamt ermittelte Anteil der Schülerinnen und Schüler, die den Regelstandard erreichen oder übertreffen, im Kompetenzbereich *Orthografie* geringer aus als in den Kompetenzbereichen *Lesen* und *Zuhören*.³ Signifikant über dem deutschen Gesamtwert liegen die für Bayern und das Saarland ermittelten Anteile. Signifikant unter dem deutschen Gesamtwert liegen die prozentualen Anteile in Hamburg, Niedersachsen, Berlin und Bremen.

Ähnlich groß wie für das Erreichen oder Übertreffen des Regelstandards fallen die Länderunterscheide für das Verfehlen des Mindeststandards aus. Hier zeigt sich zwischen dem geringsten (Bayern) und dem höchsten Landeswert (Bremen) ein Unterschied von 28 Prozentpunkten. Während in Bayern und im Saarland der prozentuale Anteil der Schülerinnen und Schüler, deren orthografische Kompetenz nicht den Mindestanforderungen entspricht, signifikant unter dem deutschen Gesamtwert von 22 Prozent liegt, ist der Anteil dieser Schülergruppe in Hamburg, Niedersachsen, Berlin und Bremen signifikant höher als in Deutschland insgesamt.

3 Dieser Befund allein kann jedoch nicht dahingehend interpretiert werden, dass die Ergebnisse für *Orthografie* pauschal ungünstiger ausfallen als etwa für den Kompetenzbereich *Lesen*. Die Anforderungen, die das Kriterium „Regelstandard erfüllt“ an die Kompetenzen der Schülerinnen und Schüler stellt, unterscheiden sich qualitativ zwischen Kompetenzbereichen und können nicht direkt zueinander in Bezug gesetzt werden.

Die Differenz der Anteile der Viertklässlerinnen und Viertklässler, die den Optimalstandard in *Orthografie* erreichen, beträgt zwischen dem höchsten und dem geringsten Landeswert 13 Prozentpunkte. Signifikant höher als der gesamtdeutsche Wert von 9 Prozentpunkten fallen die Anteile in Bayern und dem Saarland aus. Signifikant unter dem deutschen Gesamtwert liegen die Anteile in Brandenburg, Thüringen, Hamburg, Niedersachsen, Berlin und Bremen.

5.1.2 Kompetenzstufenbesetzungen in den Ländern für die Jahre 2011 und 2016 im Vergleich

Im folgenden Abschnitt wird berichtet, wie sich in den Ländern die Kompetenzstufenbesetzungen in den Bereichen *Lesen*, *Zuhören* und *Orthografie* im zeitlichen Verlauf von fünf Jahren, also von 2011 bis 2016 bezogen auf das Erreichen der Regelstandards, das Verfehlen der Mindeststandards und das Erreichen der Optimalstandards verändert haben. Die entsprechenden Ergebnisse sind in den Abbildungen 5.4, 5.5 und 5.6 zu finden.⁴

Für den Kompetenzbereich *Orthografie* ist zu berücksichtigen, dass die Kompetenzen der Schülerinnen und Schüler im Jahr 2011 nur in einer bundesweit repräsentativen Teilstichprobe erhoben wurden, die weder eine Auswertung auf Länderebene erlaubte noch Schülerinnen und Schüler in Förderschulen einschloss. Deshalb bezieht sich der Trend in den Abbildungen 5.4, 5.5 und 5.6 im Kompetenzbereich *Orthografie* lediglich auf Deutschland insgesamt und darüber hinaus nur auf Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf.

Im Tabellenteil von Abbildung 5.4 ist sowohl für Deutschland insgesamt als auch für jedes einzelne Land und – jeweils in einer separaten Zeile – für jeden der zwei Kompetenzbereiche *Lesen* und *Zuhören* aufgeführt, wie hoch der Anteil der Schülerinnen und Schüler ausfällt, die in den Jahren 2011 und 2016 mindestens den Regelstandard erreichen. Zu diesen Anteilen werden auch die zugehörigen Standardfehler (*SE*)⁵ berichtet. In der rechten Spalte der Tabelle sind die Differenzen zwischen den Prozentangaben für 2016 und 2011 sowie die Standardfehler dieser Differenzen ausgewiesen. Positive Differenzwerte (Balkendiagramm weist nach rechts) kennzeichnen eine günstige Entwicklung, da die Regelstandards im Jahr 2016 von mehr Schülerinnen und Schülern erreicht werden als im Jahr 2011. Negative Werte (Balkendiagramm weist nach links) bilden dagegen eine ungünstige Entwicklung ab, da die Regelstandards im Jahr 2016 von weniger Schülerinnen und Schülern erreicht werden als im Jahr 2011.

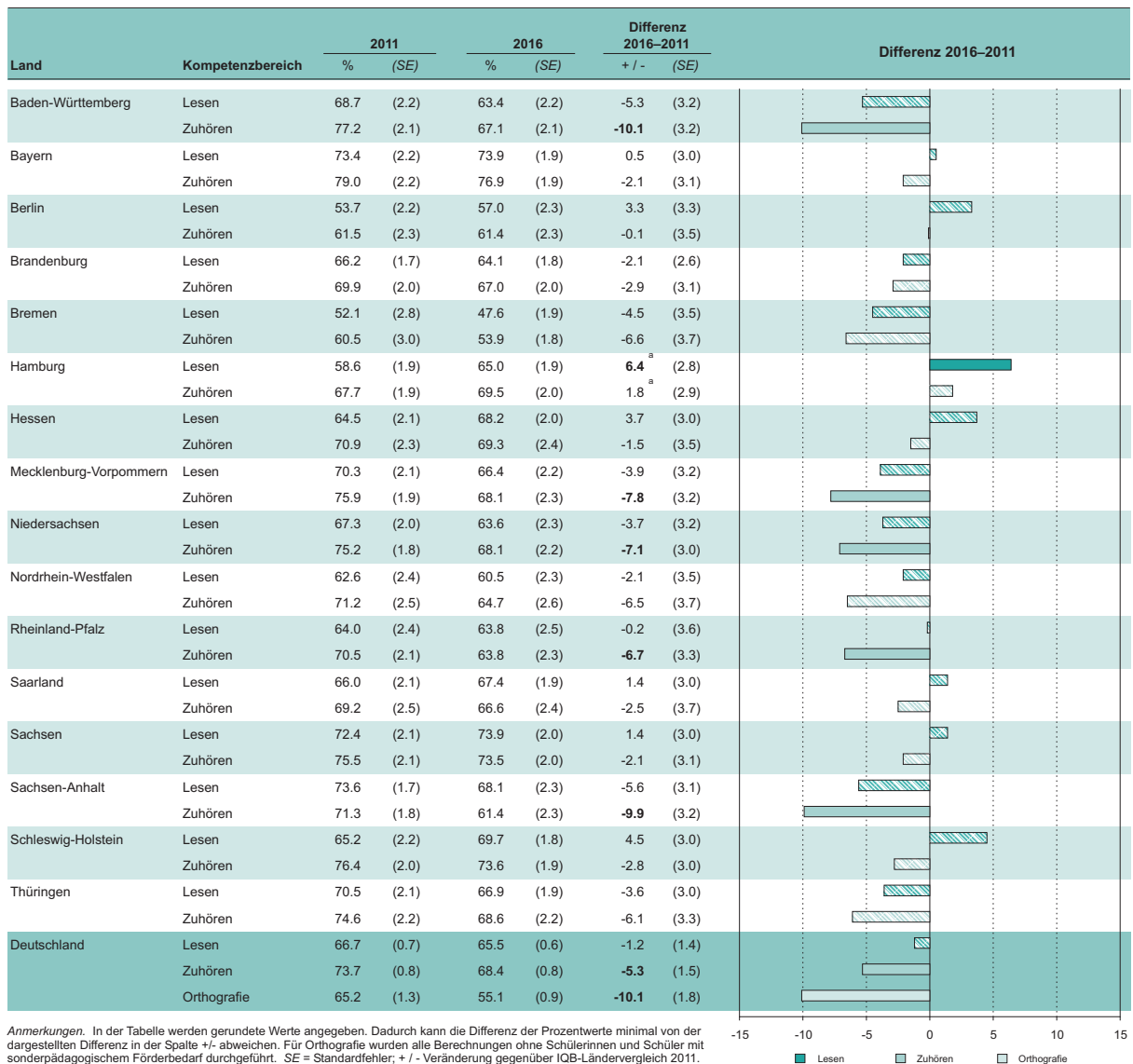
Das Balkendiagramm im rechten Teil der Abbildung kennzeichnet darüber hinaus, für welches Land die jeweils ermittelten Differenzen zwischen den Jahren 2011 und 2016 statistisch signifikant ($p < .05$) von Null abweichen. Wo das der Fall ist, wird dies durch einen vollständig ausgefüllten Balken markiert. Schraffierte Balken kennzeichnen nicht signifikante Veränderungen.

Darüber hinaus kann verglichen werden, ob sich die für ein spezifisches Land ermittelte Veränderung des jeweiligen Anteils signifikant von der für Deutschland insgesamt ermittelten Veränderung unterscheidet. Gekennzeichnet ist dies durch das Fußnotenzeichen „a“ in der Spalte „Differenz 2016–2011“.

4 Trendanalysen für die vollständigen Verteilungen auf alle fünf Kompetenzstufen können auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

5 Die Berechnung der Standardfehler wird in Kapitel 13.1 genauer beschrieben. Bedeutung und Interpretation der Standardfehler werden in Kapitel 3.2 erläutert.

Abbildung 5.4: Vergleich der für die Jahre 2011 und 2016 ermittelten prozentualen Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die in den Kompetenzbereichen *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch den Regelstandard erreichen oder übertreffen



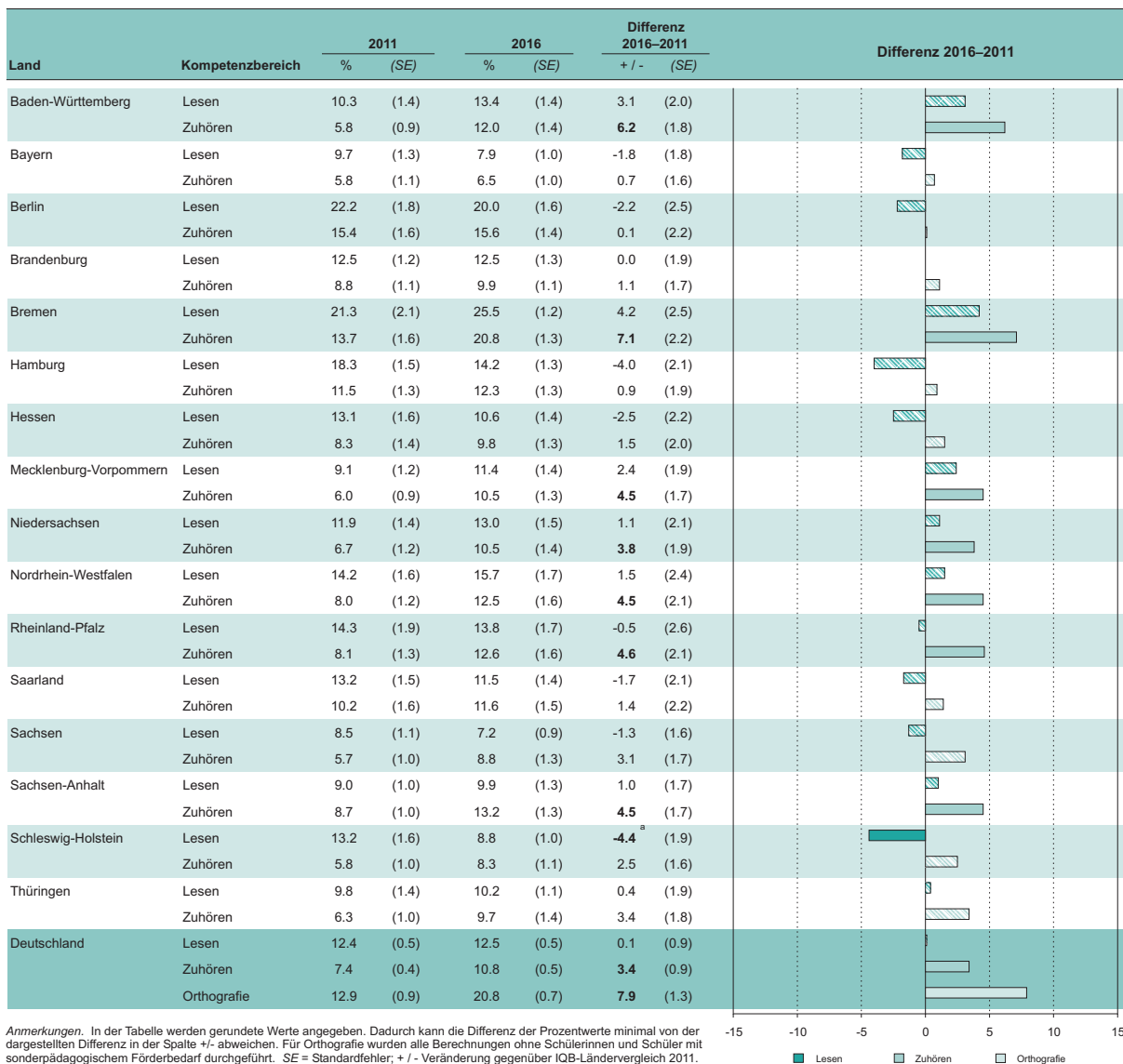
Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Prozentwerte minimal von der dargestellten Differenz in der Spalte +/- abweichen. Für Orthografie wurden alle Berechnungen ohne Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf durchgeführt. SE = Standardfehler; + / - Veränderung gegenüber IQB-Ländervergleich 2011.
^a statistisch signifikante Differenz ($p < .05$) zu Deutschland insgesamt.
 Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.
 Fettgedruckte Werte kennzeichnen eine statistisch signifikante Differenz zwischen 2011 und 2016.

Abbildung 5.4 zeigt zunächst, dass in Deutschland insgesamt der Anteil der Schülerinnen und Schüler, die im *Zuhören* und in der *Orthografie* mindestens den Regelstandard erreichen, im Jahr 2016 jeweils signifikant geringer ausfällt als im Jahr 2011 (*Zuhören*: –5 Prozentpunkte; *Orthografie*: –10 Prozentpunkte). Demgegenüber hat sich der Anteil der Schülerinnen und Schüler, die im *Lesen* mindestens den Regelstandard erreichen, nicht statistisch bedeutsam verändert (–1 Prozentpunkt).

Auf Länderebene zeigt sich für den Kompetenzbereich *Lesen* ein signifikant positiver Trend einzig für Hamburg (+6 Prozentpunkte). Ein signifikant negativer Trend ist hingegen für kein Land zu beobachten. Hamburg ist auch das einzige Land, dessen Veränderung sich signifikant von der Veränderung in Deutschland insgesamt unterscheidet.

Im Kompetenzbereich *Zuhören* zeigt sich für kein Land ein signifikant positiver Trend. Signifikant negativ fällt der Trend hingegen für Baden-Württemberg

Abbildung 5.5: Vergleich der für die Jahre 2011 und 2016 ermittelten prozentualen Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die in den Kompetenzbereichen *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch den Mindeststandard verfehlen



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Prozentwerte minimal von der dargestellten Differenz in der Spalte +/- abweichen. Für Orthografie wurden alle Berechnungen ohne Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf durchgeführt. SE = Standardfehler; + / - Veränderung gegenüber IQB-Ländervergleich 2011.

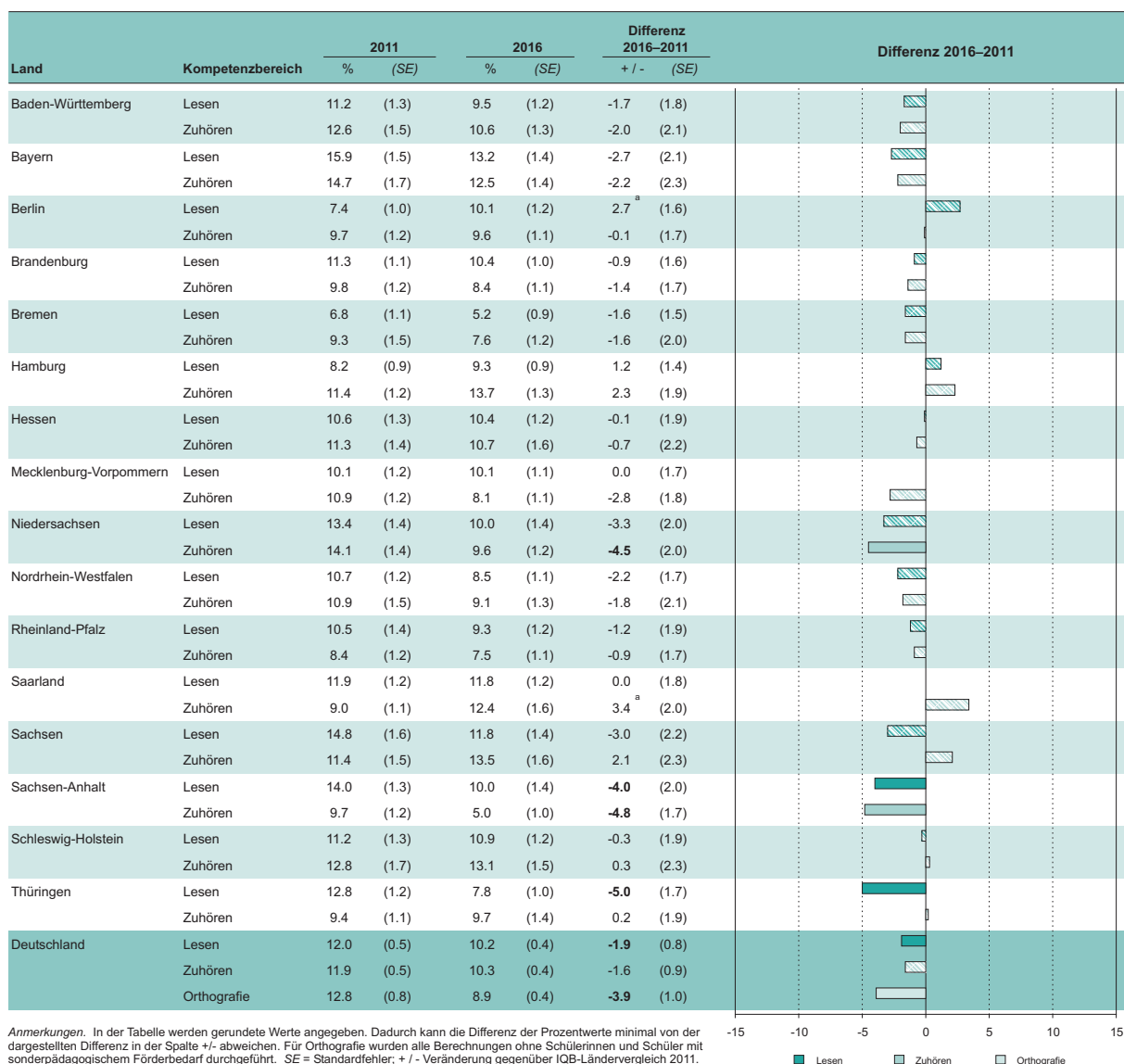
^a statistisch signifikante Differenz ($p < .05$) zu Deutschland insgesamt. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Fettgedruckte Werte kennzeichnen eine statistisch signifikante Differenz zwischen 2011 und 2016.

(-10 Prozentpunkte), Mecklenburg-Vorpommern (-8 Prozentpunkte), Niedersachsen (-7 Prozentpunkte), Rheinland-Pfalz (-7 Prozentpunkte) und Sachsen-Anhalt (-10 Prozentpunkte) aus. Vergleicht man die für die einzelnen Länder ermittelten Veränderungen in den Anteilen mit der Veränderung für Deutschland insgesamt, so findet sich wiederum lediglich für Hamburg eine signifikante Abweichung: Die leicht positive Entwicklung in Hamburg (+2 Prozentpunkte) unterscheidet sich signifikant von der negativen Veränderung für Deutschland insgesamt (-5 Prozentpunkte).

Abbildung 5.5 zeigt, dass der Anteil der Schülerinnen und Schüler, die im Bereich *Zuhören* und im Bereich *Orthografie* den Mindeststandard nicht erreichen, in Deutschland im Jahr 2016 jeweils signifikant höher ausfällt als im Jahr 2011 (*Zuhören*: +3 Prozentpunkte; *Orthografie*: +8 Prozentpunkte). Demgegenüber hat sich der Anteil der Schülerinnen und Schüler, die im Bereich *Lesen* den Mindeststandard verfehlen, nicht statistisch bedeutsam verändert.

Abbildung 5.6: Vergleich der für die Jahre 2011 und 2016 ermittelten prozentualen Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die in den Kompetenzbereichen *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch den Optimalstandard erreichen



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Prozentwerte minimal von der dargestellten Differenz in der Spalte +/- abweichen. Für Orthografie wurden alle Berechnungen ohne Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf durchgeführt. SE = Standardfehler; +/- Veränderung gegenüber IOB-Ländervergleich 2011.

^a statistisch signifikante Differenz ($p < .05$) zu Deutschland insgesamt.

Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Fettgedruckte Werte kennzeichnen eine statistisch signifikante Differenz zwischen 2011 und 2016.

Auf Ebene der Länder findet sich für den Kompetenzbereich *Lesen* ein signifikant positiver Trend (im Sinne einer Verringerung des Anteils an Schülerinnen und Schülern, die den Mindeststandard nicht erreichen) einzig für Schleswig-Holstein (-4 Prozentpunkte). Ein signifikant negativer Trend ist hingegen für kein Land zu beobachten. Schleswig-Holstein ist auch das einzige Land, dessen (positive) Veränderung sich signifikant von der Veränderung in Deutschland insgesamt unterscheidet.

Im Kompetenzbereich *Zuhören* ist für kein Land ein signifikant positiver Trend (also eine Verringerung des Anteils an Schülerinnen und Schülern, die den Mindeststandard nicht erreichen) zu beobachten. Signifikant negativ fällt dagegen der Trend in Baden-Württemberg (+6 Prozentpunkte), Bremen (+7 Prozentpunkte), Mecklenburg-Vorpommern (+5 Prozentpunkte), Niedersachsen (+4 Prozentpunkte), Nordrhein-Westfalen (+5 Prozentpunkte), Rheinland-Pfalz (+5 Prozentpunkte) und Sachsen-Anhalt (+5 Prozentpunkte) aus. In kei-

nem Land unterscheidet sich die Veränderung signifikant von der Veränderung in Deutschland insgesamt (+3 Prozentpunkte).

Abbildung 5.6 zeigt, dass der Anteil der Schülerinnen und Schüler, die in den Bereichen *Lesen* und *Orthografie* den Optimalstandard erreichen, in Deutschland im Jahr 2016 signifikant geringer ausfällt als im Jahr 2011 (*Lesen*: –2 Prozentpunkte; *Orthografie*: –4 Prozentpunkte). Demgegenüber hat sich der Anteil der Schülerinnen und Schüler, die im Bereich *Zuhören* den Optimalstandard erreichen, gegenüber 2011 nicht statistisch bedeutsam verändert (–2 Prozentpunkte).

Auf Länderebene ist für den Kompetenzbereich *Lesen* für kein Land ein signifikant positiver Trend zu verzeichnen, während der Trend für Sachsen-Anhalt (–4 Prozentpunkte) und Thüringen (–5 Prozentpunkte) signifikant negativ ausfällt. Vergleicht man die Entwicklungen in den Ländern mit der Entwicklung für Deutschland insgesamt, so ist Berlin das einzige Land, dessen (positive) Veränderung (+3 Prozentpunkte) sich signifikant von der negativen Veränderung (–2 Prozentpunkte) in Deutschland insgesamt unterscheidet.

Im Kompetenzbereich *Zuhören* zeigt sich ebenfalls für kein Land ein signifikant positiver Trend, während der Trend für Niedersachsen (–5 Prozentpunkte) und Sachsen-Anhalt (–5 Prozentpunkte) signifikant negativ ausfällt. Nur die positive Veränderung (+3 Prozentpunkte) im Saarland weicht signifikant von der Veränderung für Deutschland insgesamt (–2 Prozentpunkte) ab.

5.1.3 Zusammenfassung

Für das Jahr 2016 fällt zunächst auf, dass in Deutschland insgesamt die Verteilungen in den Bereichen *Lesen* und *Zuhören* recht ähnlich sind, während die Ergebnisse für *Orthografie* im Vergleich dazu weniger günstig ausfallen. So erreichen bundesweit in den Bereichen *Lesen* und *Zuhören* etwa zwei Drittel, im Bereich *Orthografie* aber nur etwas über die Hälfte der Schülerinnen und Schüler mindestens den Regelstandard. Der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlen, ist in *Orthografie* sogar fast doppelt so groß wie im *Lesen* und im *Zuhören*. Dieser Befund ist für sich allein genommen jedoch noch kein Beleg dafür, dass die Ergebnisse für *Orthografie* pauschal ungünstiger ausfallen als etwa für den Kompetenzbereich *Lesen*, da sich die Anforderungen, die die Kompetenzstufenmodelle etwa an die Erfüllung des Regelstandards knüpfen, zwischen den Kompetenzbereichen qualitativ unterscheiden. Im Trend – zunächst wiederum nur für Deutschland insgesamt – zeigt sich allerdings ein ähnliches Bild: Für alle drei Kriterien (mindestens Regelstandard erreicht, Mindeststandard verfehlt, Optimalstandard erreicht) fällt der Trend für *Orthografie* ungünstiger aus als für *Lesen* oder *Zuhören*.

Zwischen den Ländern unterscheiden sich die für das Jahr 2016 ermittelten Ergebnisse zum Erreichen der Bildungsstandards in allen drei untersuchten Kompetenzbereichen erheblich, wobei die Unterschiede in der *Orthografie* besonders ausgeprägt sind. So variiert der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, zwischen dem Land mit dem höchsten und dem Land mit dem niedrigsten Anteil im *Lesen* um 26 und im *Zuhören* um 23 Prozentpunkte, in der *Orthografie* jedoch um fast 36 Prozentpunkte.

Im Vergleich der Jahre 2011 und 2016 sind die Ergebnisse zum Erreichen der Regelstandards in Deutschland insgesamt für den Kompetenzbereich *Lesen* nahezu unverändert. Für die Kompetenzbereiche *Zuhören* und *Orthografie* ist hingegen eine signifikant negative Veränderung zu verzeichnen, die allerdings in

der *Orthografie* (–10 Prozentpunkte) größer ausfällt als im *Zuhören* (–5 Prozentpunkte).

Innerhalb der Länder sind nur im Kompetenzbereich *Lesen* vereinzelt positive Trends zu verzeichnen. In Hamburg konnte der Anteil der Schülerinnen und Schüler, die den Regelstandard erreichen oder übertreffen, signifikant gesteigert (+6 Prozentpunkte) und in Schleswig-Holstein konnte der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlen, signifikant reduziert werden (–4 Prozentpunkte). Negative Veränderungen zwischen den Jahren 2011 und 2016 sind für den Bereich *Lesen* nur für zwei Länder (Sachsen-Anhalt und Thüringen) und nur bezogen auf das Erreichen des Optimalstandards zu verzeichnen.

Im Bereich *Zuhören*, in dem für kein Land ein positiver Trend nachgewiesen werden konnte, ist dagegen die Anzahl der negativen Trends deutlich größer. In fünf Ländern (Baden-Württemberg, Mecklenburg-Vorpommern, Niedersachsen, Rheinland-Pfalz und Sachsen-Anhalt) ist zwischen den Jahren 2011 und 2016 der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, signifikant gesunken und der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlen, signifikant gestiegen. Negative Trends im *Zuhören* sind zudem in Bremen bezogen auf den Mindeststandard sowie in Niedersachsen und Sachsen-Anhalt bezogen auf den Optimalstandard zu verzeichnen.

Insgesamt sind die Kompetenzstufenverteilungen im Bereich *Lesen* deutschlandweit zwischen den Jahren 2011 und 2016 weitgehend stabil geblieben. Im Bereich *Zuhören* haben sich die Kompetenzverteilungen in einigen Ländern dagegen eher ungünstig entwickelt. Vergleichsweise stark ausgeprägte negative Trends liegen für Deutschland insgesamt im Kompetenzbereich *Orthografie* vor.

Literatur

- 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. (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.

5.2 Kompetenzstufenbesetzungen im Fach Mathematik

Pauline Kohrt, Nicole Haag und Petra Stanat

Im IQB-Bildungstrend 2016 wurden – wie auch schon im IQB-Ländervergleich 2011 – die mathematischen Kompetenzen der Schülerinnen und Schüler zu allen fünf in den Bildungsstandards der Kultusministerkonferenz (KMK, 2005) beschriebenen inhaltlichen Kompetenzbereichen erfasst: *Zahlen und Operationen*, *Raum und Form*, *Muster und Strukturen*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit*. Die Aufgaben zu diesen Kompetenzbereichen lassen sich zum einen separat auswerten und zum anderen auf einer Globalskala mathematischer Kompetenz verorten, sodass auch eine zusammenfassende Auswertung möglich ist (vgl. Kapitel 1.3).

Dieses Kapitel stellt dar, wie sich die mathematischen Kompetenzen der Schülerinnen und Schüler in den einzelnen Ländern auf die jeweiligen Kompetenzstufen der vom IQB entwickelten Kompetenzstufenmodelle (vgl. Kapitel 2.3) verteilen. Im Folgenden werden zunächst die Kompetenzstufenbesetzungen in den Ländern für das Jahr 2016 berichtet, wobei insbesondere auf Unterschiede zwischen den Ländern eingegangen wird. Dabei wird sowohl für die Globalskala mathematischer Kompetenz als auch für die einzelnen Leitideen dargestellt, welche prozentualen Anteile der Schülerinnen und Schüler in den Ländern die Bildungsstandards der Kultusministerkonferenz (KMK) erreichen (Abschnitt 5.2.1). Neben den Regelstandards werden auch die Mindest- und Optimalstandards in den Blick genommen. Im Anschluss wird für die globale mathematische Kompetenz der Frage nachgegangen, inwieweit sich die Kompetenzstufenbesetzungen in den Ländern zwischen den Jahren 2011 und 2016 verändert haben (Abschnitt 5.2.2). Die entsprechenden Ergebnisse für die einzelnen Leitideen können auf der Webseite des IQB als Zusatzmaterial abgerufen werden. Den Abschluss des Kapitels bildet eine Zusammenschau der zentralen Befunde (Abschnitt 5.2.3).

5.2.1 Kompetenzstufenbesetzungen in den Ländern im Jahr 2016

Für die mathematischen Kompetenzen im Primarbereich wurden für jede Leitidee und für die Globalskala fünf Kompetenzstufen festgelegt, welche die Grundlage für die Definition von *Mindeststandards*, *Regelstandards*, *Regelstandards plus* und *Optimalstandards* bilden. Die Kompetenzstufe III entspricht hierbei dem *Regelstandard* und kennzeichnet die Kompetenzerwartungen, die Schülerinnen und Schüler zum Ende der 4. Jahrgangsstufe in der Regel erreicht haben sollten. Die Kompetenzstufe II beschreibt den *Mindeststandard*, das heißt, unterhalb dieser Stufe werden die für den Primarbereich festgelegten Minimalanforderungen noch nicht erreicht. Auf den als *Regelstandard plus* beziehungsweise *Optimalstandard* bezeichneten Kompetenzstufen IV und V werden die Regelanforderungen der KMK für die Primarstufe zum Teil erheblich übertroffen.

Im Folgenden wird für die Globalskala mathematischer Kompetenz und für die fünf inhaltlichen Kompetenzbereiche der Anteil der Schülerinnen und Schüler berichtet, die jeweils innerhalb der Länder und in Deutschland insgesamt

- a) den Mindeststandard verfehlen,
- b) den Regelstandard erreichen oder übertreffen beziehungsweise
- c) den Optimalstandard erreichen.¹

Wie bereits in Kapitel 3.1 erläutert, wurden in allen Analysen dieses Kapitels nur die zielgleich auf Grundlage der Bildungsstandards unterrichteten Viertklässlerinnen und Viertklässler einbezogen. Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, die zieldifferent unterrichtet werden, bleiben daher unberücksichtigt.

Zur besseren Interpretation der Länderunterschiede in den Anteilen von Schülerinnen und Schülern, die den Regelstandard erreichen oder übertreffen, können diese zu den Veränderungen in den Kompetenzstufenverteilungen ins Verhältnis gesetzt werden, die in einem Schuljahr am Ende der Grundschulzeit zu erwarten sind. Entsprechende Schätzungen für auf den Bildungsstandards basierende Kompetenzmessungen lassen sich anhand von Daten aus Normierungsstudien des IQB aus den Jahren 2006 und 2007 bestimmen. Hierbei handelt es sich um eine in der 3. und 4. Jahrgangsstufe durchgeführte, bundesweit repräsentative Erhebung in den Fächern Deutsch und Mathematik. Ein Vergleich der von Schülerinnen und Schülern der 3. und der 4. Jahrgangsstufe erreichten Kompetenzstufen zeigt für die Globalskala mathematischer Kompetenz, dass der Anteil der Kinder, die den Regelstandard erreichen, in der 4. Jahrgangsstufe um etwa 35 Prozentpunkte höher liegt als in der 3. Jahrgangsstufe.

In den Abbildungen 5.7 bis 5.12 ist für die Globalskala mathematischer Kompetenzen und für die fünf Leitideen dargestellt, welcher Anteil der Schülerinnen und Schüler im Fach Mathematik im Jahr 2016 jeweils die Mindeststandards nicht erreicht, die Regelstandards erreicht oder übertrifft beziehungsweise die Optimalstandards erreicht. Die Abbildungen enthalten die entsprechenden Prozentangaben sowohl für die einzelnen Länder als auch für Deutschland insgesamt.² Durch Einrahmungen ist zudem gekennzeichnet, ob die Anteile in einzelnen Ländern statistisch signifikant ($p < .05$) höher (durchgezogene Linie) oder geringer (gestrichelte Linie) ausfallen als in Deutschland insgesamt. Für die Interpretation der Ergebnisse gilt, dass der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard oder den Optimalstandard erreichen, möglichst hoch und der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlen, möglichst gering ausfallen sollte.

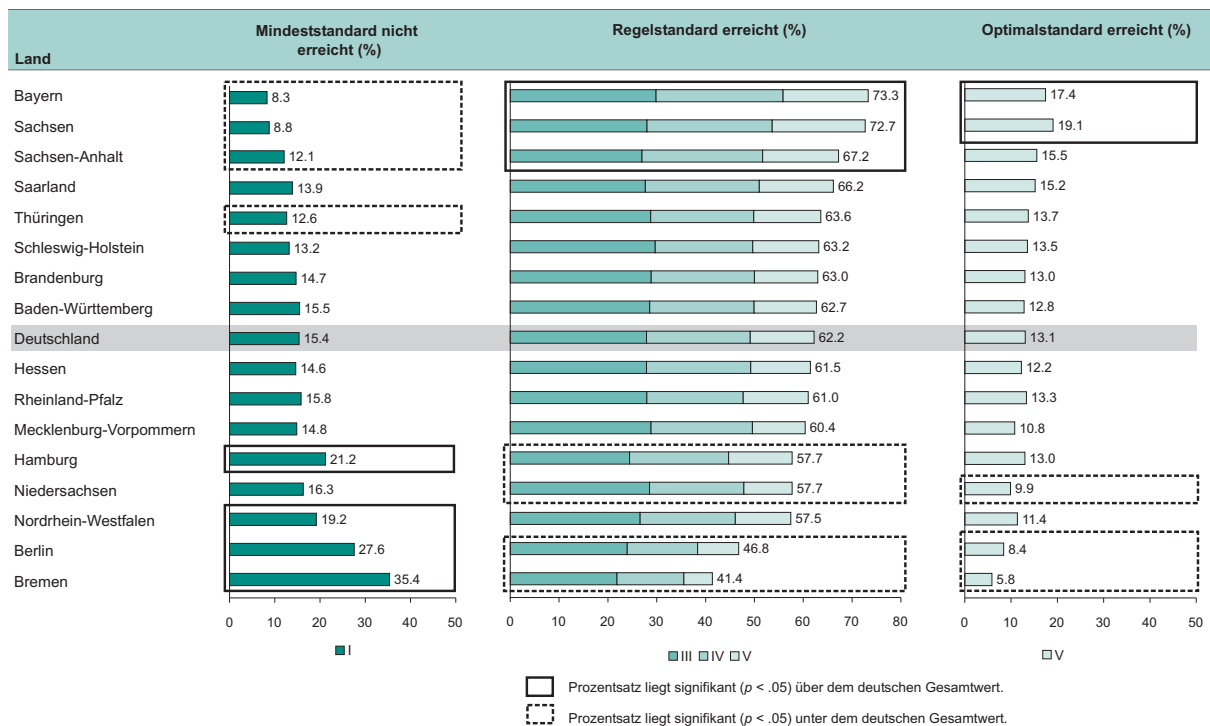
In Abbildung 5.7 ist zu sehen, dass die Anteile der Schülerinnen und Schüler, die am Ende der 4. Jahrgangsstufe im Fach Mathematik auf der Globalskala den Mindeststandard verfehlen, mindestens den Regelstandard erreichen beziehungsweise den Optimalstandard erreichen, zwischen den Ländern stark variieren.

Deutschlandweit erreichen oder übertreffen etwa zwei Drittel der Schülerinnen und Schüler den Regelstandard (62 %). Zwischen Bayern, dem Land mit dem höchsten Anteil (73 %), und Bremen, dem Land mit dem geringsten Anteil (41 %), besteht eine Differenz von etwa 32 Prozentpunkten. In den Ländern Bayern, Sachsen und Sachsen-Anhalt liegt der ermittelte Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, signifikant über dem

1 Die vollständigen Verteilungen auf alle fünf Kompetenzstufen können für alle Kompetenzbereiche und für die Globalskala auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

2 Bei der nachfolgenden Ergebnisdarstellung im Text wird dieser Anteil als „(deutscher) Gesamtwert“ bezeichnet.

Abbildung 5.7: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Fach Mathematik (*Globalskala*) den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen



deutschen Gesamtwert, in den Stadtstaaten Berlin, Bremen und Hamburg sowie in Niedersachsen signifikant darunter.

Für den Anteil der Schülerinnen und Schüler, die den Mindeststandard noch nicht erreichen, zeigt sich ebenfalls eine große Spannweite von 27 Prozentpunkten zwischen dem niedrigsten Anteil in Bayern (8 %) und dem höchsten Anteil in Bremen (35 %). Zudem fällt dieser Anteil in Bayern sowie in Sachsen, Sachsen-Anhalt und Thüringen signifikant geringer aus als in Deutschland insgesamt (15 %). Signifikant höhere Anteile als bundesweit finden sich hingegen in Berlin, Bremen, Hamburg und Nordrhein-Westfalen.

Betrachtet man nun die Schülerinnen und Schüler, die den Optimalstandard erreichen, so beträgt der Unterschied zwischen dem Land mit dem höchsten Anteil (Sachsen, 19 %) und dem Land mit dem geringsten Anteil (Bremen, 6 %) 13 Prozentpunkte. In Bayern und Sachsen wird der Optimalstandard im Vergleich zum deutschen Gesamtwert (13 %) überdurchschnittlich häufig erreicht, in Berlin, Bremen und Niedersachsen hingegen signifikant seltener.

Im Folgenden werden die Ergebnisse zu den Kompetenzbereichen *Zahlen und Operationen*, *Raum und Form*, *Muster und Strukturen*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* dargestellt. Die Befundmuster für die prozentualen Anteile der Viertklässlerinnen und Viertklässler, die den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen beziehungsweise den Optimalstandard erreichen, gleichen dabei weitgehend den für die Globalskala mathematischer Kompetenzen bestimmten Anteilen.

Die Kompetenzstufenbesetzung für den Bereich *Zahlen und Operationen* findet sich in Abbildung 5.8. Der Anteil der Schülerinnen und Schüler, die in diesem Kompetenzbereich mindestens den Regelstandard erreichen, liegt in Deutschland

Abbildung 5.8: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Zahlen und Operationen* im Fach Mathematik den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen

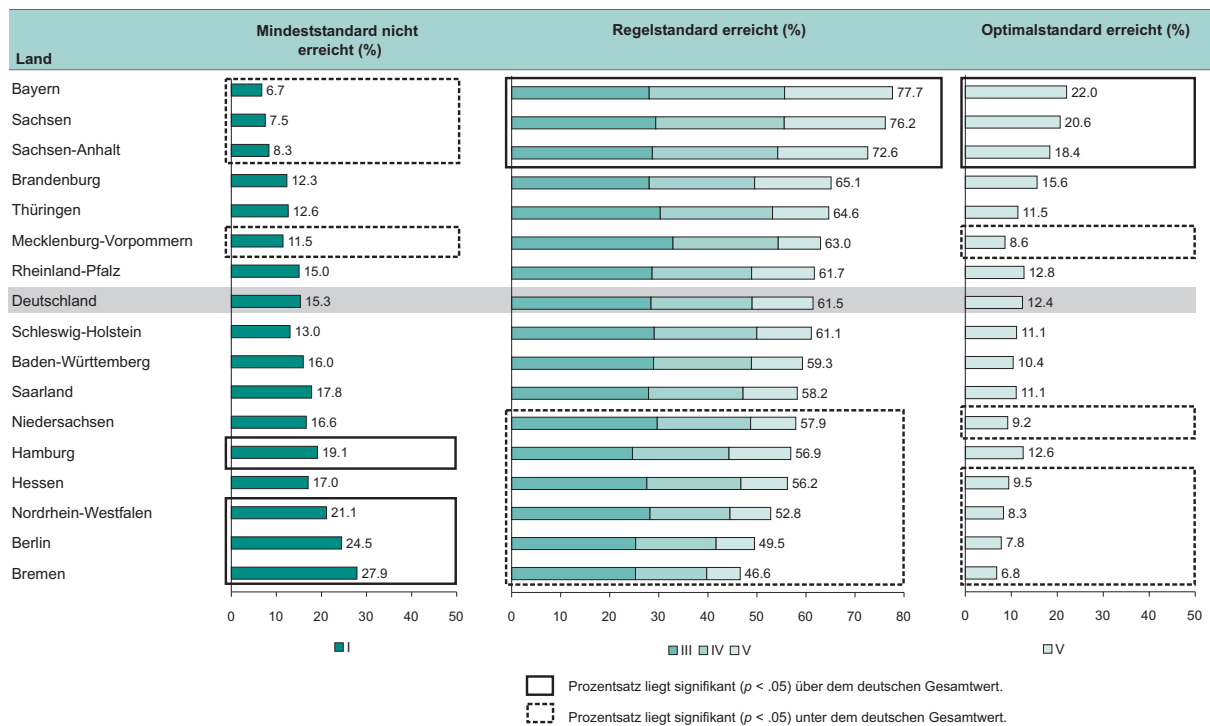


insgesamt bei knapp 63 Prozent. Die Differenz zwischen dem höchsten (Bayern, 72 %) und dem geringsten Anteil (Bremen, 41 %) beträgt 31 Prozentpunkte. In Bayern, Sachsen und dem Saarland liegt der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen oder übertreffen, signifikant über dem deutschen Gesamtwert, in den Ländern Berlin, Bremen, Hamburg und Niedersachsen signifikant darunter.

Der Anteil der Schülerinnen und Schüler, die im Bereich *Zahlen und Operationen* den Mindeststandard nicht erreichen, liegt in Deutschland insgesamt bei gut 15 Prozent mit einer Spannweite von 25 Prozentpunkten zwischen Bayern (8 %), dem Land mit dem geringsten Anteil, und Bremen (33 %), dem Land mit dem höchsten Anteil. Signifikant geringere Anteile verglichen mit dem deutschen Gesamtwert und somit günstigere Ergebnisse werden in den Ländern Bayern, Sachsen und Saarland erzielt, in den Stadtstaaten Berlin, Bremen und Hamburg verhält es sich umgekehrt: Dort sind die entsprechenden Anteile signifikant höher.

Der Optimalstandard im Bereich *Zahlen und Operationen* wird bundesweit von 13 Prozent der Schülerinnen und Schüler erreicht. Hierbei ergibt sich ein Abstand von 11 Prozentpunkten zwischen dem Land mit dem höchsten Anteil (Sachsen, 18 %) und dem Land mit dem geringsten Anteil (Bremen, 7 %). Sachsen liegt mit einem Anteil von 18 Prozent als einziges Land signifikant über dem deutschen Gesamtwert, Berlin, Bremen und Niedersachsen liegen signifikant darunter.

Abbildung 5.9: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Raum und Form* im Fach Mathematik den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen



In Abbildung 5.9 ist die Kompetenzstufenbesetzung in den Ländern für den Bereich *Raum und Form* dargestellt. In Deutschland insgesamt liegt der Anteil der Schülerinnen und Schüler, die den Regelstandard erreichen oder übertreffen, bei knapp 62 Prozent. Der Abstand zwischen dem höchsten Landeswert (Bayern, 78 %) und dem geringsten Landeswert (Bremen, 47 %) liegt hier bei 31 Prozentpunkten. Schülerinnen und Schüler in Bayern, Sachsen und Sachsen-Anhalt erreichen signifikant häufiger mindestens den Regelstandard als Schülerinnen und Schüler in Deutschland insgesamt, genau umgekehrt verhält es sich in Berlin, Bremen, Hamburg, Hessen, Niedersachsen und Nordrhein-Westfalen.

Im Bereich *Raum und Form* erreichen deutschlandweit etwa 15 Prozent der Schülerinnen und Schüler nicht den Mindeststandard. Die Differenz zwischen dem Land mit dem geringsten (Bayern, 7 %) und dem höchsten (Bremen, 28 %) Anteil beträgt 21 Prozentpunkte. Signifikant geringere Anteile von Schülerinnen und Schülern, die den Mindeststandard verfehlen, als in Deutschland insgesamt bestehen in Bayern, Mecklenburg-Vorpommern, Sachsen und Sachsen-Anhalt, höhere Anteile in Berlin, Bremen, Hamburg und Nordrhein-Westfalen.

Betrachtet man nun die Befunde zum Erreichen des Optimalstandards für *Raum und Form*, der in Deutschland insgesamt von gut 12 Prozent der Schülerinnen und Schüler erreicht wird, zeigt sich eine recht große Spannweite von 15 Prozentpunkten zwischen den Ländern mit dem höchsten (Bayern, 22 %) und dem geringsten (Bremen, 7 %) Anteil. In Bayern, Sachsen und Sachsen-Anhalt liegen die Anteile signifikant über dem deutschen Gesamtwert, in Berlin, Bremen, Hessen, Mecklenburg-Vorpommern, Niedersachsen und Nordrhein-Westfalen signifikant darunter.

Abbildung 5.10: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Muster und Strukturen* im Fach Mathematik den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen

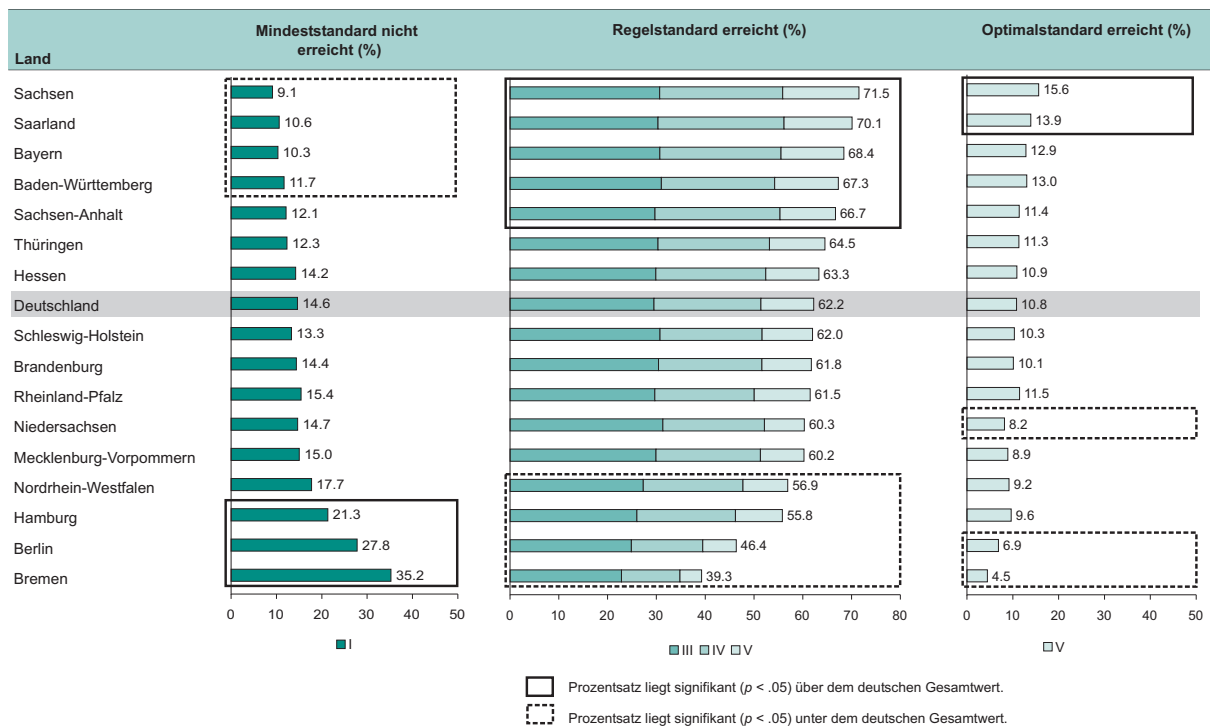


Abbildung 5.10 zeigt die Kompetenzstufenverteilung für den Bereich *Muster und Strukturen*. In Deutschland insgesamt liegt der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, bei gut 63 Prozent, wobei die Spannweite zwischen den Ländern mit dem höchsten (Bayern, 73 %) und niedrigsten (Bremen, 44 %) Anteilen 29 Prozentpunkte beträgt. Auch in diesem Kompetenzbereich werden in Bayern und Sachsen signifikant günstigere, in den Stadtstaaten Berlin und Bremen dagegen signifikant ungünstigere Ergebnisse erzielt als in Deutschland insgesamt.

Die Ergebnisse zum Mindeststandard, der bundesweit von knapp 15 Prozent der Schülerinnen und Schüler verfehlt wird, ergeben für *Muster und Strukturen* eine Differenz von 22 Prozentpunkten zwischen dem Land mit dem geringsten und dem Land mit dem höchsten prozentualen Anteil (Bayern, 8 % und Bremen, 30 %). In den Ländern Bayern und Sachsen wird der Mindeststandard im Vergleich zum deutschen Gesamtwert signifikant seltener verfehlt, in Berlin, Bremen und Hamburg dagegen signifikant häufiger.

Den Optimalstandard im Kompetenzbereich *Muster und Strukturen* erreichen bundesweit etwa 12 Prozent der Schülerinnen und Schüler. Hier zeigt sich eine Abweichung von 9 Prozentpunkten zwischen Bayern (15 %) und Bremen (6 %), den Ländern mit dem höchsten bzw. dem geringsten prozentualen Anteil. Signifikant höhere Anteile im Vergleich zum deutschen Gesamtwert ergeben sich für Bayern und Sachsen, signifikant geringere in Berlin, Brandenburg, Bremen, Mecklenburg-Vorpommern und Niedersachsen.

Abbildung 5.11: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Größen und Messen* im Fach Mathematik den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen

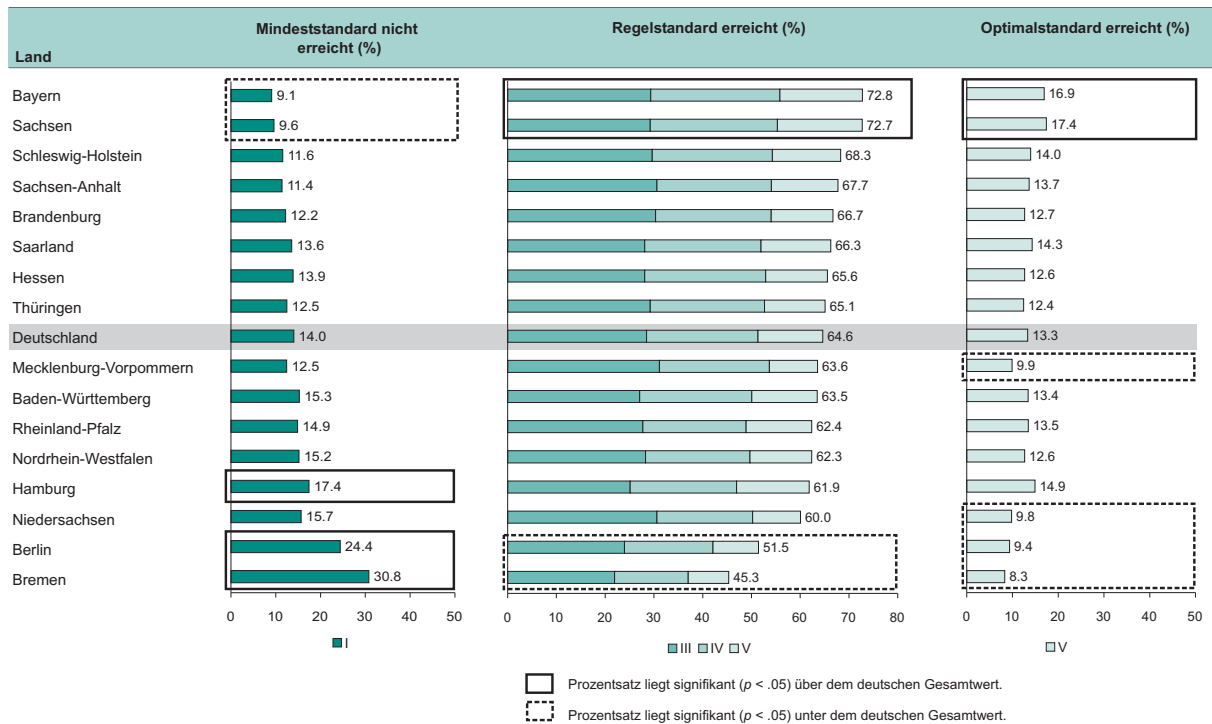


In Abbildung 5.11 sind die Kompetenzstufenverteilungen für den Bereich *Größen und Messen* dargestellt. Der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, liegt in Deutschland insgesamt bei etwa 62 Prozent und ist im Vergleich der Länder in Sachsen (72 %) am größten. Die Differenz zu Bremen (39 %), dem Land mit dem geringsten prozentualen Anteil, beträgt 33 Prozentpunkte. Signifikant über dem deutschen Gesamtwert liegen die Anteile der Schülerinnen und Schüler, die im Bereich *Größen und Messen* den Regelstandard erreichen oder übertreffen, in Baden-Württemberg, Bayern, Saarland, Sachsen und Sachsen-Anhalt, signifikant darunter in Berlin, Bremen, Hamburg und Nordrhein-Westfalen.

Deutschlandweit erreichen knapp 15 Prozent der Schülerinnen und Schüler im Bereich *Größen und Messen* nicht den Mindeststandard. Die Spannweite zwischen den Ländern mit dem niedrigsten (Sachsen, 9 %) und dem höchsten (Bremen, 35 %) Anteil beträgt hier 26 Prozentpunkte. Signifikant unter dem deutschen Gesamtwert liegen die festgestellten Anteile in Baden-Württemberg, Bayern, Saarland und Sachsen, signifikant darüber in den Stadtstaaten Berlin, Bremen und Hamburg.

Für das Erreichen des Optimalstandards ergibt sich im Bereich *Größen und Messen* bundesweit ein Anteil von knapp 11 Prozent der Schülerinnen und Schüler mit einer Differenz von 11 Prozentpunkten zwischen dem Land mit dem höchsten (Sachsen, 16 %) und dem geringsten (Bremen, 5 %) prozentualen Anteil. Verglichen mit dem deutschen Gesamtwert weisen auch hier die Länder Saarland und Sachsen signifikant günstigere Werte auf. Signifikant seltener als bundesweit wird der Optimalstandard in Berlin, Bremen und Niedersachsen erreicht.

Abbildung 5.12: Prozentuale Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit* im Fach Mathematik den Mindeststandard nicht erreichen, den Regelstandard erreichen oder übertreffen bzw. den Optimalstandard erreichen



Die Ergebnisse zu den Kompetenzstufenbesetzungen im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* sind in Abbildung 5.12 zu finden. Der für Deutschland insgesamt ermittelte Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, liegt bei knapp 65 Prozentpunkten. Der höchste Anteil wurde für Bayern (73 %) festgestellt, der geringste Anteil für Bremen (45 %). Die Differenz zwischen diesen beiden Landeswerten beträgt 28 Prozentpunkte. Auch hier ergeben sich für Bayern und Sachsen signifikant höhere und für Berlin und Bremen signifikant geringere Anteile als für Deutschland insgesamt.

Der Anteil der Schülerinnen und Schüler, die den Mindeststandard nicht erreichen, liegt im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* deutschlandweit bei 14 Prozent, bei einer Spannweite von 22 Prozentpunkten zwischen dem Land mit dem niedrigsten Anteil (Bayern, 9 %) und dem Land mit dem höchsten Anteil (Bremen, 31 %). Signifikant unter dem deutschen Gesamtwert liegen die für Bayern und Sachsen festgestellten Anteile, signifikant darüber in Berlin, Bremen und Hamburg.

Den Optimalstandard im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* erreichen gut 13 Prozent der Schülerinnen und Schüler bundesweit; zwischen den Ländern (höchster Anteil in Sachsen, 17 %; geringster Anteil in Bremen, 8 %) besteht eine Spannweite von 9 Prozentpunkten. Auch in diesem Kompetenzbereich liegt der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, in Bayern und Sachsen signifikant über dem deutschen Gesamtwert, während dieser in Berlin, Bremen, Mecklenburg-Vorpommern und Niedersachsen signifikant geringer ist.

5.2.2 Kompetenzstufenbesetzungen in den Ländern für die Jahre 2011 und 2016 im Vergleich

Im folgenden Abschnitt wird berichtet, wie sich die Kompetenzstufenbesetzungen im Fach Mathematik zwischen den Jahren 2011 und 2016 verändert haben. Wie bereits für die auf das Jahr 2016 bezogenen Ergebnisse werden auch für die Trends die Anteile der zielgleich unterrichteten Schülerinnen und Schüler betrachtet, die auf der Globalskala den Regelstandard erreichen oder übertreffen (Abb. 5.13), den Mindeststandard verfehlen (Abb. 5.14) und den Optimalstandard erreichen (Abb. 5.15).³ Da die Ergebnismuster für die Kompetenzbereiche *Zahlen und Operationen*, *Raum und Form*, *Muster und Strukturen*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* den für die globalen mathematischen Kompetenzen ermittelten Ergebnissen insgesamt recht ähnlich sind, wird im Text nicht weiter darauf eingegangen. Die entsprechenden Abbildungen können auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

Anhand der Abbildungen 5.13, 5.14 und 5.15 ist sowohl für Deutschland insgesamt als auch für die einzelnen Länder ersichtlich, inwieweit sich zwischen den Jahren 2011 und 2016 Veränderungen im Anteil der Viertklässlerinnen und Viertklässler ergeben haben, die den Regelstandard im Fach Mathematik (*Globalskala*) erreichen oder übertreffen, den Mindeststandard verfehlen, beziehungsweise den Optimalstandard erreichen. Die Reihenfolge der Länder in der Abbildung ist alphabetisch und stellt somit keine Rangfolge dar.

Neben den prozentualen Anteilen für die Jahre 2011 und 2016 und der Differenz dieser Werte sind auch die Standardfehler der Schätzungen (*SE*)⁴ aufgeführt. Positive Differenzwerte zeigen, dass der entsprechende Anteil im Jahr 2016 höher ausfällt als im Jahr 2011; bei negativen Differenzwerten ist hingegen der Anteil im Jahr 2016 geringer. Signifikant vom deutschen Gesamtwert abweichende Veränderungen in den Ländern sind durch ein hochgestelltes „a“ gekennzeichnet. Die Differenz der prozentualen Anteile zwischen den Jahren 2016 und 2011 wird zusätzlich in Form eines Balkendiagramms grafisch veranschaulicht. Vollständig gefüllte Balken zeigen, dass die Veränderung des jeweiligen Prozentsatzes statistisch signifikant ($p < .05$) von Null abweichen. Bei schraffiert gefüllten Balken ist diese Differenz dagegen statistisch nicht signifikant.

Abbildung 5.13 zeigt, dass der Anteil der Schülerinnen und Schüler, die im Fach Mathematik im Jahr 2016 mindestens den Regelstandard erreichen, in Deutschland insgesamt signifikant geringer ausfällt als im Jahr 2011 (–6 Prozentpunkte). Dies gilt auch für Baden-Württemberg (–10 Prozentpunkte), Bremen (–9 Prozentpunkte), Mecklenburg-Vorpommern (–9 Prozentpunkte), Niedersachsen (–8 Prozentpunkte), Nordrhein-Westfalen (–8 Prozentpunkte) und Sachsen-Anhalt (–8 Prozentpunkte). Ein signifikanter Anstieg dieses Anteils konnte für kein Land ermittelt werden. Lediglich Hamburg weicht mit der festgestellten leichten Zunahme (+2 Prozentpunkte) signifikant von der negativen Veränderung für Deutschland insgesamt ab. Hamburg kann damit den Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, signifikant stabiler halten als Deutschland.

3 Trendanalysen für die vollständigen Verteilungen auf alle fünf Kompetenzstufen können für alle Kompetenzbereiche auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

4 Die Berechnung der Standardfehler wird in Kapitel 13.1 genauer beschrieben. Bedeutung und Interpretation der Standardfehler werden in Kapitel 3.2 erläutert.

Abbildung 5.13: Vergleich der für die Jahre 2011 und 2016 ermittelten prozentualen Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Fach Mathematik (*Globalskala*) den Regelstandard erreichen oder übertreffen

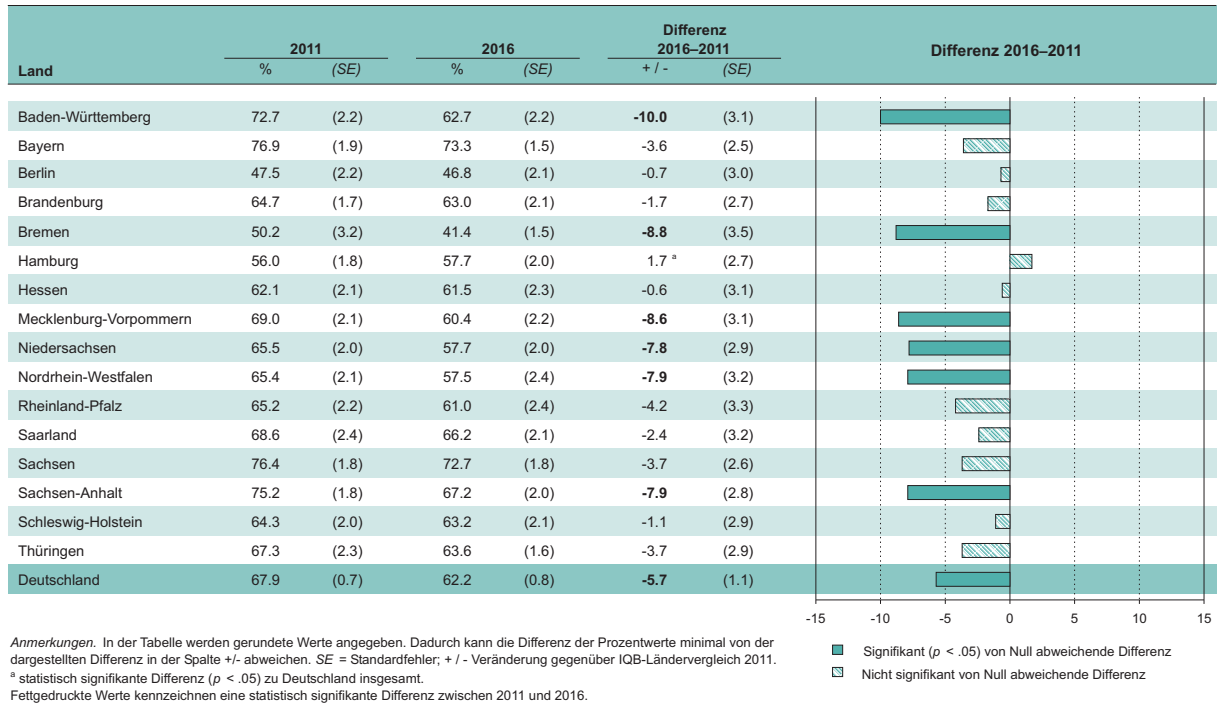
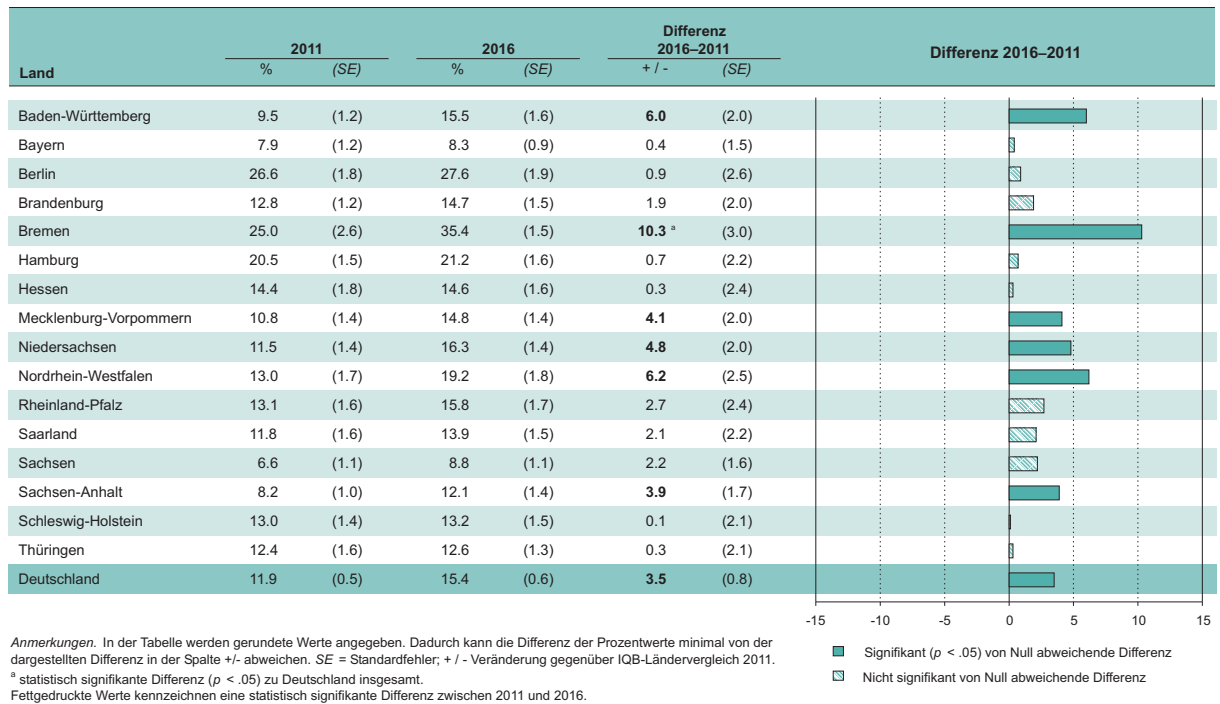


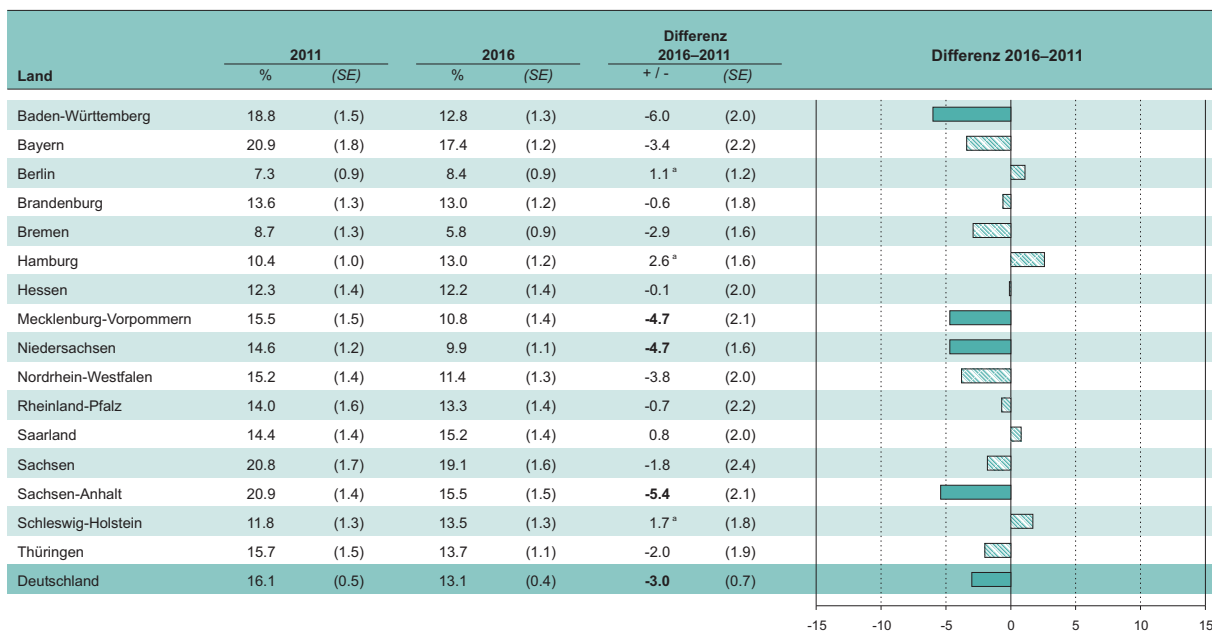
Abbildung 5.14: Vergleich der für die Jahre 2011 und 2016 ermittelten prozentualen Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Fach Mathematik (*Globalskala*) den Mindeststandard verfehlen



In Abbildung 5.14 ist der Trend für den Anteil der Schülerinnen und Schüler, die den Mindeststandard im Fach Mathematik nicht erreichen, dargestellt. Für Deutschland insgesamt zeigt sich ein signifikanter Anstieg dieses Anteils um etwa 4 Prozentpunkte. Damit wird der Mindeststandard im Jahr 2016 insgesamt häufiger verfehlt als im Jahr 2011. Diese ungünstige Entwicklung findet sich in unterschiedlicher Ausprägung in allen Ländern; eine Verringerung des Anteils der Schülerinnen und Schüler, die in Mathematik noch nicht den Mindeststandard erreichen, konnte in keinem Land festgestellt werden. Statistisch signifikant sind die Veränderungen in Baden-Württemberg (+6 Prozentpunkte), Bremen (+10 Prozentpunkte), Mecklenburg-Vorpommern (+4 Prozentpunkte), Niedersachsen (+5 Prozentpunkte), Nordrhein-Westfalen (+6 Prozentpunkte) und Sachsen-Anhalt (+4 Prozentpunkte). Im Vergleich mit der Veränderung für Deutschland insgesamt zeigt sich für Bremen zwischen 2011 und 2016 ein signifikant höherer Anstieg des Anteils der Schülerinnen und Schüler, die den Mindeststandard verfehlen (+10 Prozentpunkte). Signifikant günstigere Ergebnisse in Bezug auf das Verfehlen des Mindeststandards konnten im Vergleich zu Deutschland insgesamt für kein Land identifiziert werden.

Abbildung 5.15 kann entnommen werden, inwieweit sich der Anteil der Viertklässlerinnen und Viertklässler, die im Fach Mathematik den Optimalstandard erreichen, zwischen den Jahren 2011 und 2016 verändert hat. Für Deutschland zeigt sich insgesamt eine signifikante Abnahme dieses Anteils von 3 Prozentpunkten; der Optimalstandard wird im Jahr 2016 also von etwas weniger Schülerinnen und Schülern erreicht als im Jahr 2011. Auf Länderebene bestehen in Baden-Württemberg (-6 Prozentpunkte), Mecklenburg-Vorpommern (-5 Prozentpunkte), Niedersachsen (-5 Prozentpunkte) und Sachsen-Anhalt (-5 Prozentpunkte) signifikant negative Trends. Signifikant positive Trends konnten auch hier für kein Land festgestellt werden, auch wenn sich in Berlin, Hamburg und Schleswig-

Abbildung 5.15: Vergleich der für die Jahre 2011 und 2016 ermittelten prozentualen Anteile der Schülerinnen und Schüler der 4. Jahrgangsstufe, die im Fach Mathematik (*Globalskala*) den Optimalstandard erreichen



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Prozentwerte minimal von der dargestellten Differenz in der Spalte +/- abweichen. SE = Standardfehler; + / - Veränderung gegenüber IQB-Ländervergleich 2011.
^a statistisch signifikante Differenz ($p < .05$) zu Deutschland insgesamt.
 Fettgedruckte Werte kennzeichnen eine statistisch signifikante Differenz zwischen 2011 und 2016.

Signifikant ($p < .05$) von Null abweichende Differenz
 Nicht signifikant von Null abweichende Differenz

Holstein etwas günstigere Entwicklungen abzeichnen als bundesweit. Die in diesen Ländern leicht positiven Entwicklungen im Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, unterscheiden sich signifikant vom für Deutschland insgesamt ermittelten Trend.

5.2.3 Zusammenfassung

Für das Jahr 2016 kann zusammenfassend festgestellt werden, dass die Ergebnisse zu den Kompetenzstufenbesetzungen im Fach Mathematik für die Global-skala und für die fünf mathematischen Leitideen sehr ähnlich ausfallen. Der Regelstandard wird jeweils von etwa zwei Dritteln der Schülerinnen und Schüler in Deutschland insgesamt erreicht oder übertroffen, der Mindeststandard jeweils von etwa 15 Prozent der Schülerinnen und Schüler verfehlt und der Optimalstandard jeweils von ungefähr 15 Prozent der Schülerinnen und Schüler erreicht. Die Ergebnisse einzelner Länder weichen jedoch teilweise erheblich von diesen Werten ab. Die Anteile der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, variieren zwischen dem Land mit dem höchsten Anteil und dem Land mit dem niedrigsten Anteil in allen Leitideen um ungefähr 30 Prozentpunkte. In Bayern und in Sachsen werden mit Bezug auf nahezu alle Standards und Kompetenzbereiche besonders günstige Ergebnisse erreicht. Besonders ungünstige Ergebnisse hingegen ergeben sich durchgängig in den Ländern Berlin und Bremen.

Die Trendanalysen zeigen, dass sich die Kompetenzstufenverteilungen im Fach Mathematik deutschlandweit zwischen den Jahren 2011 und 2016 eher ungünstig entwickelt haben. Die mit den Bildungsstandards definierten Kompetenzziele werden im Jahr 2016 in den Ländern demnach von weniger Schülerinnen und Schülern erreicht als im Jahr 2011. Diese ungünstige Entwicklung ist sowohl für das Erreichen bzw. Übertreffen des Regelstandards, das Verfehlen des Mindeststandards als auch für das Erreichen des Optimalstandards zu beobachten. So hat sich der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, in Deutschland insgesamt signifikant verringert (–6 Prozentpunkte). Gleichzeitig verfehlt ein größerer Anteil der Schülerinnen und Schüler den Mindeststandard (+4 Prozentpunkte) und ein geringerer Anteil erreicht den Optimalstandard (–3 Prozentpunkte).

In den Ländern zeigen sich weitgehend ähnliche Befundmuster. In sechs Ländern (Baden-Württemberg, Bremen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen und Sachsen-Anhalt) wird im Jahr 2016 der Regelstandard signifikant seltener erreicht bzw. übertroffen und der Mindeststandard signifikant häufiger verfehlt als im Jahr 2011. In vier Ländern (Baden-Württemberg, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt) wird zudem im Jahr 2016 der Optimalstandard signifikant seltener erreicht als im Jahr 2011.

Während für kein Land signifikant positive Trends festgestellt werden können, gelingt es einigen Ländern zumindest, sich vom negativen Trend für Deutschland insgesamt abzusetzen. In Hamburg ist sowohl der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen, als auch der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, signifikant stabiler als deutschlandweit. Für die Anteile der Schülerinnen und Schüler, die den Optimalstandard erreichen, gilt dies auch für Berlin und Schleswig-Holstein. Bezogen auf den Mindeststandard kann für kein Land ein signifikant stabileres

Ergebnis festgestellt werden. In Bremen ist jedoch der Anteil der Schülerinnen und Schüler, die den Mindeststandard verfehlen, signifikant stärker angestiegen als bundesweit.

Literatur

KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.

Kapitel 6

Mittelwerte und Streuungen der erreichten Kompetenzen im Ländervergleich

6.1 Mittelwerte und Streuungen der im Fach Deutsch erreichten Kompetenzen

Julia Wittig und Sebastian Weirich

In Kapitel 6.1 werden die im Fach Deutsch von Schülerinnen und Schülern der 4. Jahrgangsstufe im Mittel erreichten Kompetenzen sowie ihre Heterogenität in den einzelnen Ländern berichtet. Dabei werden die Ergebnisse getrennt für die Kompetenzbereiche *Lesen*, *Zuhören* und *Orthografie* dargestellt. Die zur Erfassung dieser Domänen eingesetzten Testaufgaben werden in Kapitel 1.2 beschrieben.

Für die Kompetenzbereiche *Lesen* und *Zuhören* werden die erreichten Kompetenzen auf der Berichtsmetrik dargestellt, die auch im IQB-Ländervergleich 2011 für Schülerinnen und Schüler der 4. Jahrgangsstufe in Deutschland verwendet wurde. In diesen beiden Domänen wurde der Mittelwert im Jahr 2011 auf $M = 500$ und die Standardabweichung auf $SD = 100$ festgelegt (Pant, Böhme & Köller, 2012). Für den Kompetenzbereich *Orthografie* ist die Berichtsmetrik dagegen so definiert, dass sich der Mittelwert von $M = 500$ und die Standardabweichung von $SD = 100$ auf das Jahr 2016 beziehen. Maßgeblich hierfür ist, dass *Orthografie* im Jahr 2011 nur in einer Teilstichprobe erhoben wurde, die keine Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf einschloss und demnach nicht der Populationsdefinition des Jahres 2016 entspricht.

Unterschiede in den erreichten Kompetenzständen werden als Punktwertdifferenzen auf der Berichtsmetrik dargestellt (vgl. Kapitel 3.2). Um die ländervergleichende Interpretation dieser Differenzen zu veranschaulichen, können diese zu Lernzuwächsen ins Verhältnis gesetzt werden, die im Mittel am Ende der Primarstufe zu erwarten sind. Eine auf den Bildungsstandards basierende Normierungsstudie des IQB aus dem Jahr 2006 weist darauf hin, dass die Punktwertdifferenz zwischen der 3. und 4. Jahrgangsstufe in den Kompetenzbereichen *Lesen* und *Zuhören* auf der Berichtsmetrik jeweils etwa 60 Punkte (Behrens, Böhme & Krelle, 2009; Böhme & Bremerich-Vos, 2009), im Kompetenzbereich *Orthografie* etwa 100 Punkte beträgt.

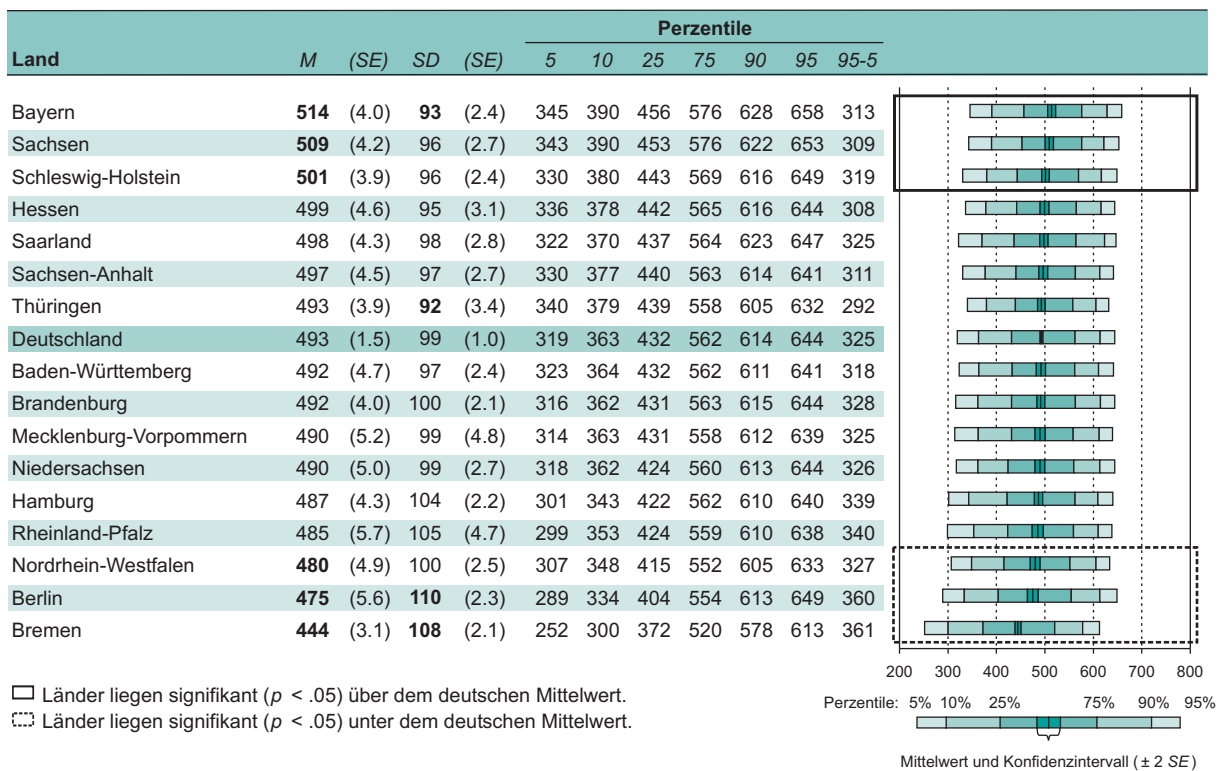
In Abschnitt 6.1.1 wird zunächst für das Erhebungsjahr 2016 dargestellt, welches Kompetenzniveau die Viertklässlerinnen und Viertklässler in den einzelnen Ländern durchschnittlich erreichen. Weiterhin wird die Streuung der im Mittel erreichten Kompetenzwerte innerhalb der Länder berichtet (Abschnitt 6.1.2). Abschnitt 6.1.3 fasst die Ergebnisse für das Jahr 2016 zusammen. Abschnitt 6.1.4 geht der Frage nach, wie sich das durchschnittliche Kompetenzniveau zwi-

schen den Jahren 2011 und 2016 im Trend verändert hat. In Abschnitt 6.1.5 werden Veränderungen in der Heterogenität der erreichten Kompetenzen dargestellt. Dabei wird sowohl auf die Varianz eingegangen, die innerhalb von Schulen besteht, als auch auf die Varianz, die zwischen Schulen liegt (Abschnitt 6.1.6). Abschnitt 6.1.7 fasst die zentralen Ergebnisse des Kapitels zusammen.

6.1.1 Im Mittel erreichte Kompetenzen im Jahr 2016

Neben dem Erreichen der Mindest-, Regel-, und Optimalstandards, die in Kapitel 5 dargestellt wurden, werden im Rahmen des nationalen Bildungsmonitorings auch die Kompetenzmittelwerte der 16 Länder als Indikatoren für das Erreichen bildungspolitischer Ziele herangezogen. Im Folgenden werden die Ergebnisse separat für die Kompetenzbereiche *Lesen*, *Zuhören* und *Orthografie* dargestellt (Abb. 6.1 bis 6.3). Sortiert sind die Länder im Tabellenteil der Abbildungen absteigend nach den jeweils erreichten Kompetenzmittelwerten. Neben dem Mittelwert (*M*) sind in den Tabellen auch der Standardfehler des Mittelwerts (*SE*), die Standardabweichung (*SD*) und ihr Standardfehler sowie verschiedene Perzentile angegeben. Wie diese statistischen Kennwerte zu interpretieren sind, wird in Kapitel 3.2 erläutert. Welche Ländermittelwerte sich signifikant vom Gesamtmittelwert für Deutschland unterscheiden, ist im Abbildungsteil durch Einrahmungen markiert. Werte, die signifikant über dem Mittelwert für Deutschland liegen, sind durch einen durchgezogenen Rahmen hervorgehoben. Werte, die signifikant unter dem Mittelwert für Deutschland liegen, sind durch einen durchgezogenen Rahmen hervorgehoben.

Abbildung 6.1: Mittelwerte, Streuungen und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Lesen*



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

Signifikant unterhalb des deutschen Mittelwerts liegende Werte sind durch einen gestrichelten Rahmen gekennzeichnet.

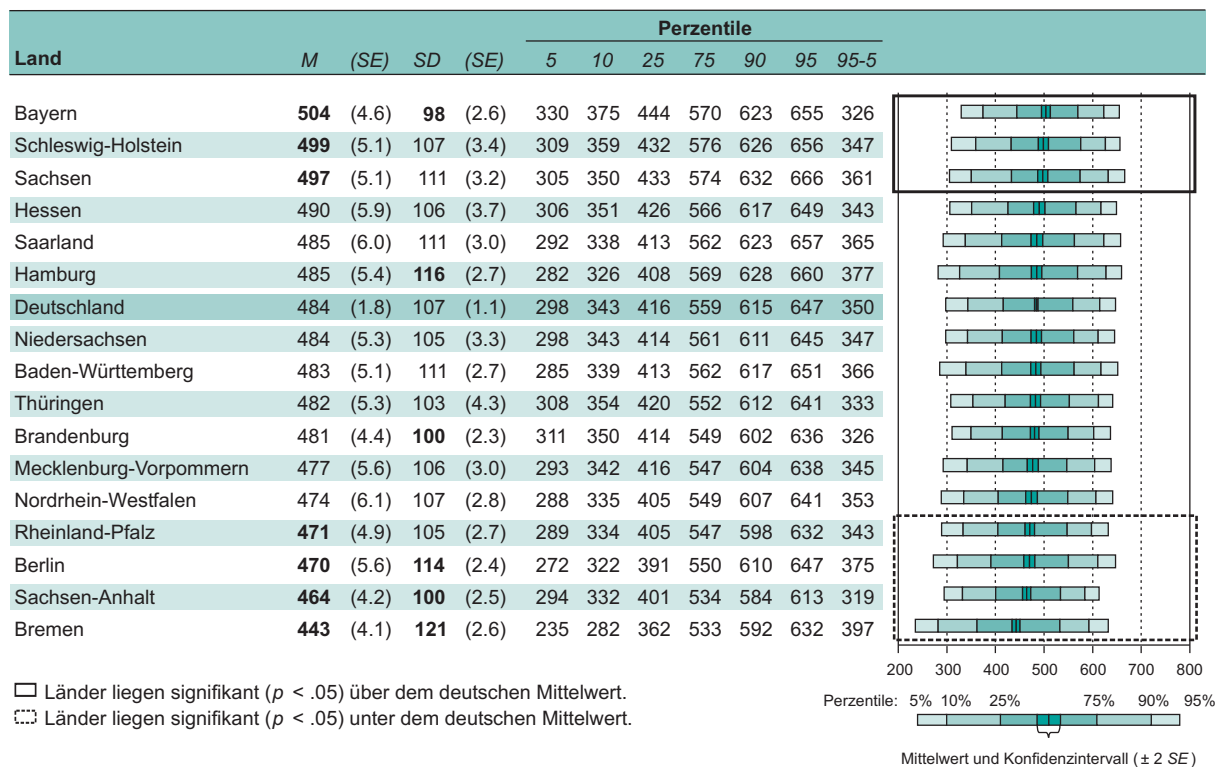
Die Analysen dieses Kapitels beziehen sich auf alle Schülerinnen und Schüler, das heißt auch auf Viertklässlerinnen und Viertklässler, die einen sonderpädagogischen Förderbedarf haben und nicht nach dem allgemeinen Lehrplan, sondern zieldifferent unterrichtet werden (siehe Kapitel 3.1).

Im Kompetenzbereich *Lesen* beträgt der Mittelwert im Jahr 2016 für Deutschland insgesamt 493 Punkte (Abb. 6.1). Die Werte der einzelnen Länder bewegen sich zwischen 514 Punkten in Bayern und 444 Punkten in Bremen. Zwischen dem Land mit dem höchsten und dem Land mit dem niedrigsten Mittelwert liegen somit 70 Punkte auf der Berichtsmetrik, was einem Lernzuwachs von etwas mehr als einem Schuljahr entspricht.

Durchschnittliche Kompetenzwerte, die signifikant über dem Mittelwert für Deutschland insgesamt liegen, erreichen im *Lesen* die Länder Bayern, Sachsen und Schleswig-Holstein. Signifikant unterhalb des bundesdeutschen Durchschnitts liegen die mittleren Kompetenzen, die in den Ländern Nordrhein-Westfalen, Berlin und Bremen erreicht werden.

Im Kompetenzbereich *Zuhören* erreichen Viertklässlerinnen und Viertklässler in Deutschland im Mittel 484 Punkte auf der Berichtsmetrik (Abb. 6.2). Zwischen Bayern, dem Land mit dem höchsten Wert von 504 Punkten, und Bremen, dem Land mit dem niedrigsten Wert von 443 Punkten, liegen 61 Punkte und damit eine etwas geringere Distanz als im Kompetenzbereich *Lesen*. Im *Zuhören* entspricht dies einem erwarteten Lernzuwachs von einem Schuljahr.

Abbildung 6.2: Mittelwerte, Streuungen und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Zuhören*



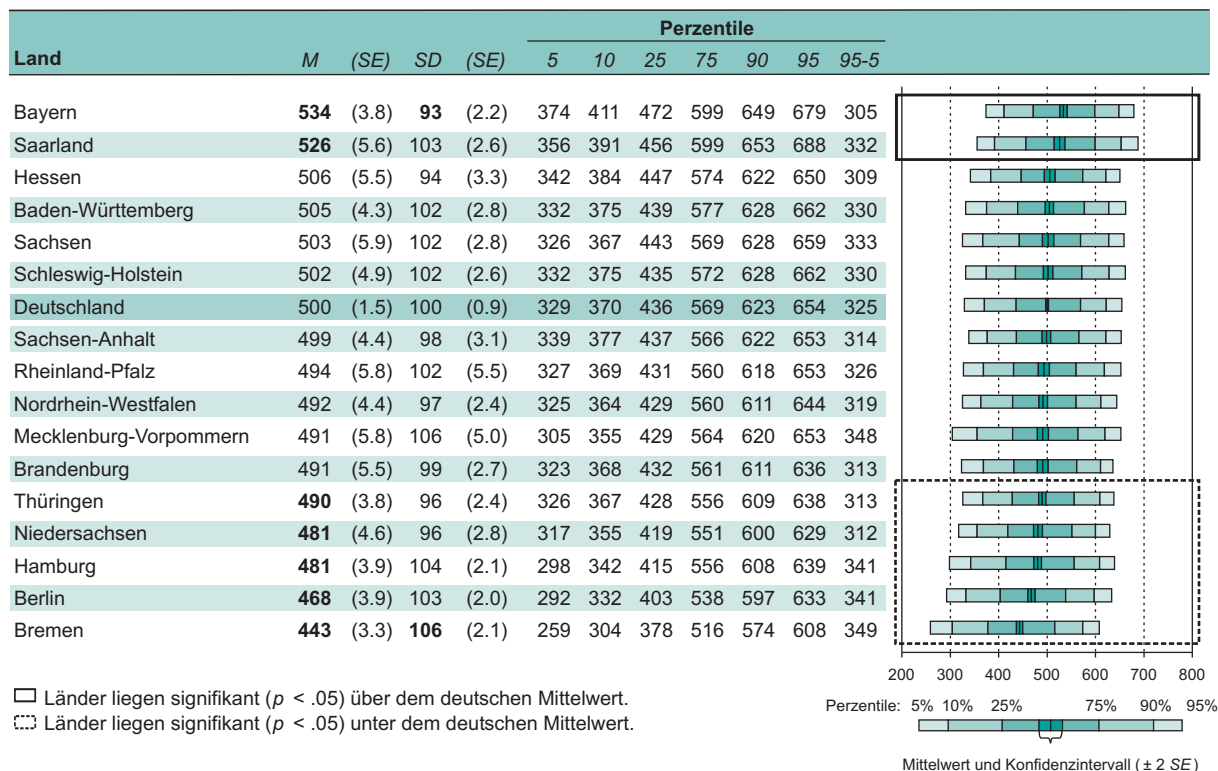
Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

Wie bereits im *Lesen* liegen auch im *Zuhören* die in den Ländern Bayern, Schleswig-Holstein und Sachsen im Durchschnitt erreichten Kompetenzen signifikant über dem deutschen Mittelwert. Mittlere Kompetenzwerte, die signifikant geringer sind als der bundesdeutsche Durchschnitt, werden hingegen in Rheinland-Pfalz, Berlin, Sachsen-Anhalt und Bremen erreicht.

Im Kompetenzbereich *Orthografie* beträgt der Gesamtmittelwert definitionsgemäß 500 Punkte (Abb. 6.3). Zwischen Bayern (534 Punkte), dem Land mit dem höchsten Mittelwert, und Bremen (443 Punkte), dem Land mit dem niedrigsten Mittelwert, liegen 91 Punkte. Im Vergleich zu den Kompetenzbereichen *Lesen* und *Zuhören* sind die Unterschiede in den Kompetenzständen zwischen den Ländern hier numerisch deutlich größer. Da der zu erwartende Kompetenzzuwachs zwischen der 3. und der 4. Jahrgangsstufe für *Orthografie* etwa 100 Punkte beträgt, entspricht auch diese Differenz, wie im *Zuhören*, inhaltlich einem Kompetenzzuwachs von etwa einem Schuljahr.

Signifikant über dem deutschen Durchschnitt liegen die Mittelwerte Bayerns und des Saarlands. Signifikant geringere Kompetenzmittelwerte werden in Thüringen, Niedersachsen und den Stadtstaaten Hamburg, Berlin und Bremen erreicht.

Abbildung 6.3: Mittelwerte, Streuungen und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Orthografie*



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

6.1.2 Heterogenität der erreichten Kompetenzen im Jahr 2016

In diesem Abschnitt wird die Streuung der im Fach Deutsch erreichten Kompetenzen in den Ländern untersucht. Günstig ist hierbei ein insgesamt hohes Kompetenzniveau bei gleichzeitig geringer Streuung, sodass auch Schülerinnen und Schüler am unteren Rand der Leistungsverteilung möglichst hohe Kompetenzen zeigen. Als Kennwert der Größe der Kompetenzunterschiede wird in den Abbildungen 6.1 bis 6.3 die Standardabweichung (*SD*) angegeben. Signifikante Abweichungen von der Streuung für Deutschland insgesamt sind fett markiert.

Im Jahr 2016 liegt die Streuung im Kompetenzbereich *Lesen* für Deutschland insgesamt bei 99 Punkten auf der Berichtsmetrik (Abb. 6.1). In Bayern und Thüringen sind die festgestellten Kompetenzwerte signifikant homogener, in Berlin und Bremen signifikant heterogener als in Deutschland insgesamt. Das wünschenswerte Befundmuster eines hohen Kompetenzmittelwerts bei geringer Streuung wird im Vergleich der Länder besonders erfolgreich in Bayern erreicht. In den Ländern Sachsen und Schleswig-Holstein, in denen die Schülerinnen und Schüler ebenfalls signifikant höhere mittlere Kompetenzen im *Lesen* aufweisen, liegt die Streuung hingegen nah an der bundesweiten Standardabweichung.

Im Kompetenzbereich *Zuhören* beträgt die Standardabweichung für Deutschland insgesamt 107 Punkte (Abb. 6.2). Wie im Kompetenzbereich *Lesen* ist die Streuung auch im *Zuhören* in Bayern homogener als in Deutschland insgesamt. Auch in Brandenburg und Sachsen-Anhalt sind die Standardabweichungen im *Zuhören* signifikant geringer als bundesweit. Signifikant heterogenere Kompetenzwerte weisen hingegen die Stadtstaaten Berlin, Hamburg und Bremen auf.

Die Streuung im Kompetenzbereich *Orthografie* beträgt für Deutschland definitionsgemäß 100 Punkte (Abb. 6.3). Signifikante Abweichungen davon finden sich lediglich in Bayern mit einer geringeren und in Bremen mit einer größeren Heterogenität. In den übrigen Ländern bewegen sich die Streuungen nah um den bundesdeutschen Durchschnitt von 100 Punkten.

Als weiterer Indikator für die Heterogenität der erreichten Kompetenzen wird die Breite des Skalenabschnitts zwischen dem 5. und dem 95. Perzentil betrachtet (Abb. 6.1 bis 6.3). Die geringste Breite hat dieser Skalenabschnitt im Kompetenzbereich *Lesen* in Thüringen, gefolgt von Hessen und Sachsen.

Im Kompetenzbereich *Zuhören* weisen die Länder Bayern, Brandenburg und Sachsen-Anhalt, im Bereich *Orthografie* Bayern und Hessen die geringste Breite zwischen dem 5. und dem 95. Perzentil auf.

6.1.3 Zusammenfassung der Befunde für das Jahr 2016

Für das Jahr 2016 im Fach Deutsch zeigt sich zusammenfassend, dass Viertklässlerinnen und Viertklässler in Bayern in allen drei erfassten Kompetenzbereichen sowohl einen signifikant über dem Bundesdurchschnitt liegenden Mittelwert erreichen als auch eine signifikant geringere Heterogenität aufweisen. Entsprechend sind auch die Kompetenzwerte der schwächeren Schülerinnen und Schüler in Bayern vergleichsweise hoch.

In Sachsen und Schleswig-Holstein werden in zwei der drei Kompetenzbereiche – im *Lesen* und im *Zuhören* – signifikant über dem deutschen Gesamtwert liegende Mittelwerte erreicht. Schülerinnen und Schüler im Saarland erzielen in der *Orthografie* signifikant höhere Kompetenzwerte.

Die in Berlin und in Bremen erreichten Kompetenzmittelwerte liegen hingegen in allen drei untersuchten Kompetenzbereichen signifikant unterhalb des deutschen Gesamtmittelwerts. Für Bremen sind zudem in allen drei Domänen vergleichsweise heterogene Kompetenzen zu verzeichnen. Schülerinnen und Schüler in Berlin weisen im *Lesen* und im *Zuhören* eine vergleichsweise große Heterogenität auf.

6.1.4 Vergleich der in den Jahren 2011 und 2016 im Mittel erreichten Kompetenzen

Ein zentrales Anliegen des IQB-Bildungstrends ist die Untersuchung von Veränderungen der in den Ländern erreichten Kompetenzen über die Zeit. Dazu werden zunächst die in den Jahren 2011 und 2016 im Mittel erreichten Kompetenzen verglichen, um festzustellen, ob innerhalb der Länder signifikante Trends vorliegen. Anschließend werden die Veränderungen innerhalb der Länder mit den für Deutschland insgesamt ermittelten Trends verglichen. Die Veränderungen der im Mittel erreichten Kompetenzstände in den im Fach Deutsch getesteten Kompetenzbereichen sind in Abbildung 6.4 dargestellt.

Im Kompetenzbereich *Lesen* zeigt sich für das Jahr 2016 ein im Vergleich zum Jahr 2011 statistisch signifikanter Rückgang um 7 Punkte. Innerhalb der einzelnen Länder haben sich die mittleren Kompetenzwerte im Bereich *Lesen* entweder verringert oder nicht signifikant verändert. Für kein Land lässt sich ein signifikant positiver Trend feststellen. Signifikant geringere Kompetenzwerte als im Jahr 2011 werden im Jahr 2016 in den Ländern Bremen, Sachsen-Anhalt und Thüringen erzielt.

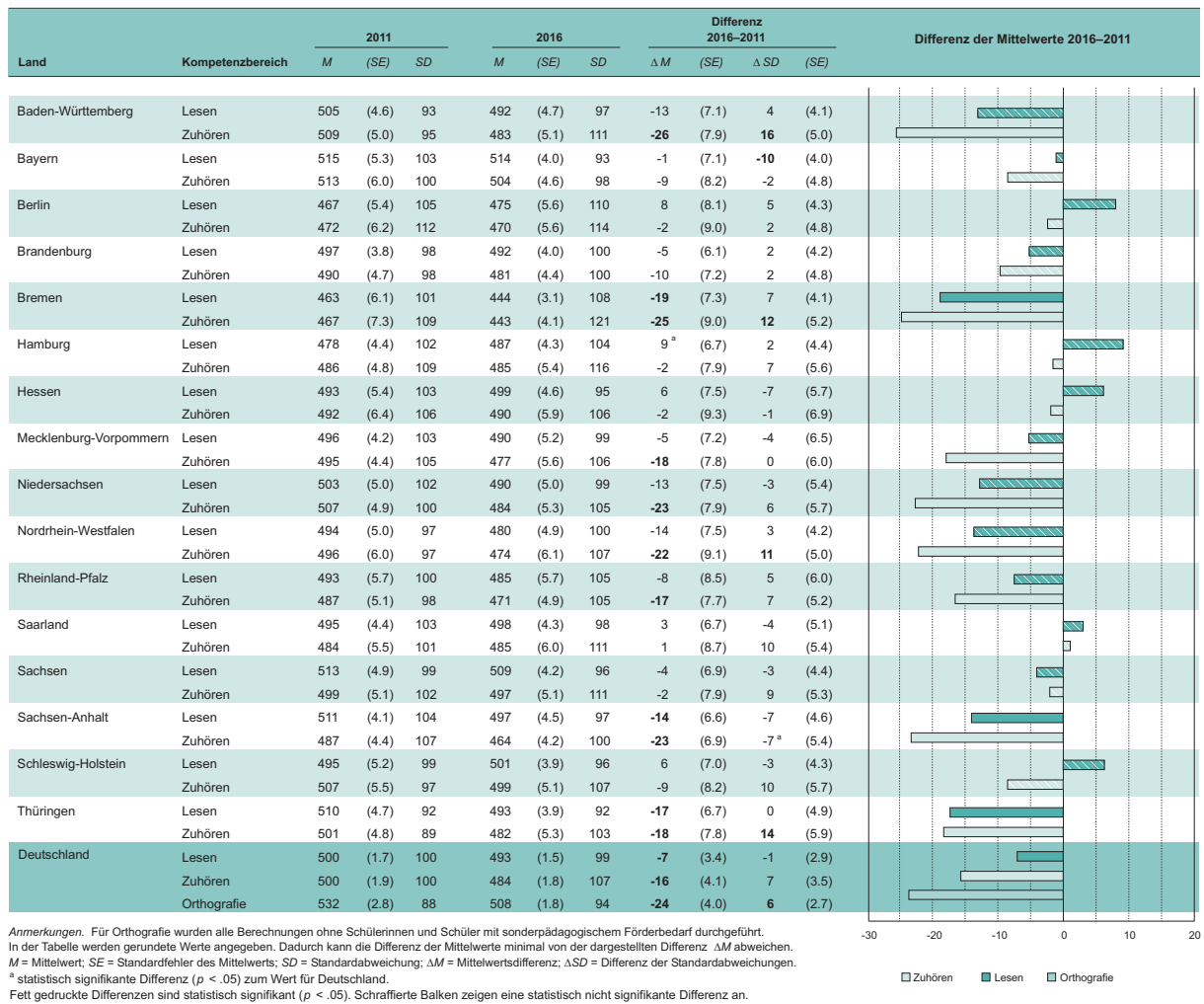
Vergleicht man die länderspezifischen Trends mit dem Trend für Deutschland insgesamt, so ergibt sich im Bereich *Lesen* nur für Hamburg ein signifikanter Unterschied. Die leicht positive Veränderung, die in Hamburg zu verzeichnen ist, setzt sich damit vom negativen Trend für Deutschland insgesamt ab.

Zusätzlich zu Abbildung 6.4 stellt Abbildung 6.5 die jeweils erreichten mittleren Kompetenzen im *Lesen* in den Jahren 2011 und 2016 in Relation zueinander dar. Der auf der x-Achse abgetragene Wert kennzeichnet den Mittelwert für das Jahr 2016, der Wert auf der y-Achse kennzeichnet den Mittelwert für das Jahr 2011. Die x- und y-Achse schneiden sich im Mittelwert des jeweiligen Erhebungsjahres für Deutschland insgesamt. Für den Kompetenzbereich *Lesen* befindet sich der Schnittpunkt bei den Werten $M = 500$ für das Jahr 2011 und $M = 493$ für das Jahr 2016.

Oben rechts und damit im ersten Quadranten des Koordinatensystems befinden sich Länder, die sowohl im Erhebungsjahr 2011 als auch im Erhebungsjahr 2016 Mittelwerte erreichen, die über dem jeweiligen deutschen Gesamtmittelwert liegen. Im dritten Quadranten (unten links) befinden sich Länder, für die in beiden Jahren Mittelwerte unterhalb des jeweiligen Mittelwerts für Deutschland insgesamt festgestellt wurden. Durch unterschiedliche Symbole wird gekennzeichnet, ob die Kompetenzmittelwerte in beiden Erhebungsjahren (Raute), lediglich im Jahr 2011 (Quadrat), nur im Jahr 2016 (Dreieck) oder zu keinem der beiden Messzeitpunkte (Kreis) signifikant vom bundesdeutschen Durchschnitt abweichen.

Die mittleren Kompetenzen im *Lesen* fallen in den Ländern Bayern und Sachsen sowohl im IQB-Ländervergleich 2011 als auch im IQB-Bildungstrend

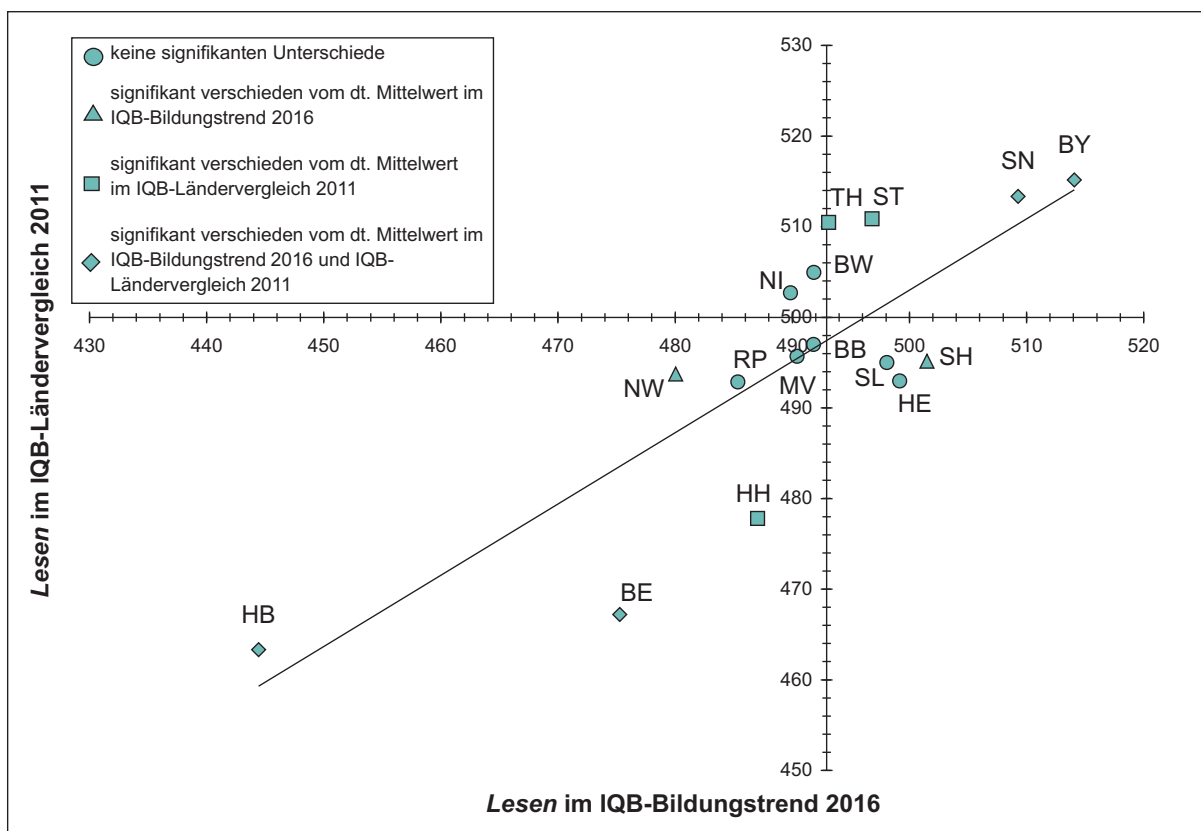
Abbildung 6.4: Vergleich der in den Jahren 2011 und 2016 erreichten Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Fach Deutsch



2016 signifikant höher als in Deutschland insgesamt. In beiden Erhebungsjahren signifikant unterhalb des deutschen Durchschnitts liegen die Kompetenzmittelwerte in Berlin und Bremen.

Eine besonders günstige Entwicklung ist für Schleswig-Holstein zu erkennen. Hier liegt die Lesekompetenz im Jahr 2011 noch im gesamtdeutschen Durchschnitt, im Jahr 2016 werden hingegen überdurchschnittliche Ergebnisse erzielt. Eine ähnliche Entwicklung zeigt sich für Hamburg. Wird hier im Jahr 2011 noch ein signifikant geringerer Kompetenzmittelwert erreicht, weicht dieser im Jahr 2016 nicht mehr vom bundesdeutschen Durchschnitt ab. Weniger günstig entwickelten sich die Lesekompetenzen in Thüringen, Sachsen-Anhalt und in Nordrhein-Westfalen. In Thüringen und in Sachsen-Anhalt werden im Jahr 2011 im Mittel Kompetenzwerte erreicht, die signifikant über dem deutschen Gesamtmittelwert liegen, im Jahr 2016 unterscheiden sich die erzielten Ergebnisse hingegen nicht mehr vom Mittelwert für Deutschland insgesamt. In Nordrhein-Westfalen unterscheidet sich die Lesekompetenz von Viertklässlerinnen und Viertklässlern im Jahr 2011 nicht vom deutschen Gesamtmittelwert, im Jahr 2016 liegt sie jedoch signifikant unterhalb des bundesdeutschen Durchschnitts.

Abbildung 6.5: Ländermittelwerte im Kompetenzbereich *Lesen* im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016



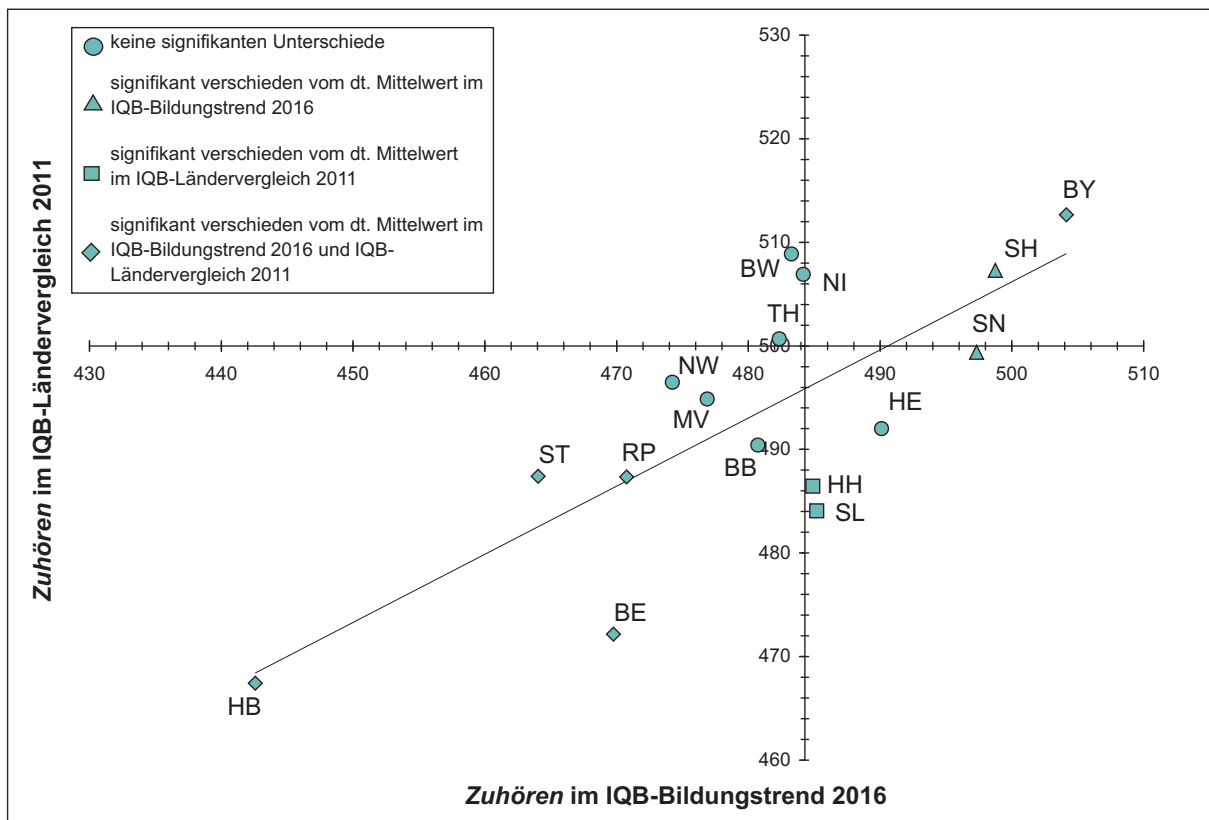
Anmerkungen. BB = Brandenburg; BE = Berlin; BW = Baden-Württemberg; BY = Bayern; HB = Bremen; HE = Hessen; HH = Hamburg; MV = Mecklenburg-Vorpommern; NI = Niedersachsen; NW = Nordrhein-Westfalen; RP = Rheinland-Pfalz; SH = Schleswig-Holstein; SL = Saarland; SN = Sachsen; ST = Sachsen-Anhalt; TH = Thüringen.

Die in Abbildung 6.4 dargestellten Ergebnisse zeigen, dass sich die Kompetenzen deutscher Viertklässlerinnen und Viertklässler im *Zuhören* um 16 Punkte signifikant verringert haben. Wie im Kompetenzbereich *Lesen* werden auch im *Zuhören* im Jahr 2016 in keinem Land signifikant höhere Kompetenzen erreicht als im Jahr 2011. Signifikant negative Trends zeigen sich hingegen für Baden-Württemberg, Bremen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen-Anhalt und Thüringen.

Anders als im Kompetenzbereich *Lesen* weicht im *Zuhören* der Trend keines Landes signifikant vom Trend für Deutschland insgesamt ab.

Abbildung 6.6 zeigt die jeweils in den Erhebungsjahren 2011 und 2016 erreichten mittleren Kompetenzen im *Zuhören* in Relation zueinander. Der Schnittpunkt der Achsen liegt hier bei den Werten $M = 500$ für das Jahr 2011 und $M = 484$ für das Jahr 2016. Wie im Kompetenzbereich *Lesen* erreicht Bayern auch im *Zuhören* zu beiden Messzeitpunkten überdurchschnittliche Ergebnisse. Die Kompetenzwerte der Schülerinnen und Schüler in Berlin, Bremen, Rheinland-Pfalz und Sachsen-Anhalt liegen zu beiden Messzeitpunkten signifikant unterhalb des jeweiligen Mittelwerts für Deutschland insgesamt. Besonders günstig hat sich die Zuhörkompetenz in Sachsen und Schleswig-Holstein entwickelt. Die Mittelwerte beider Länder unterscheiden sich im Jahr 2011 nicht vom gesamtdeutschen Mittelwert, im Jahr 2016 werden dagegen in beiden Ländern überdurchschnittliche Kompetenzwerte erreicht. Auch für Hamburg und das

Abbildung 6.6: Ländermittelwerte im Kompetenzbereich *Zuhören* im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016



Anmerkungen. 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.

Saarland lassen sich positive Entwicklungen verzeichnen. Liegen die Werte beider Länder im Jahr 2011 noch signifikant unterhalb des gesamtdeutschen Mittelwerts, unterscheiden sie sich im Jahr 2016 nicht mehr signifikant vom deutschen Gesamtmittelwert.

Für den Kompetenzbereich *Orthografie* kann an dieser Stelle nur der Trend für Deutschland insgesamt und ohne Berücksichtigung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf berichtet werden (Abb. 6.4). Wie in den Kompetenzbereichen *Lesen* und *Zuhören* verringern sich die im Mittel erreichten Werte der Viertklässlerinnen und Viertklässler auch im Bereich *Orthografie* signifikant. Hier werden im Jahr 2016 im Mittel 24 Punkte weniger erreicht als noch im Jahr 2011.

6.1.5 Vergleich der Heterogenität der in den Jahren 2011 und 2016 erreichten Kompetenzen in den Ländern

Nach der Analyse von Veränderungen in den Kompetenzmittelwerten wird im folgenden Abschnitt dargelegt, wie sich die Streuung dieser Werte in den einzelnen Kompetenzbereichen verändert hat.

Für den Kompetenzbereich *Lesen* geht aus Abbildung 6.4 hervor, dass sich die Standardabweichung in Deutschland insgesamt zwischen den Jahren 2011 und

2016 nicht signifikant verändert hat. Dies gilt auch für die einzelnen Länder mit Ausnahme Bayerns, dort hat sich die Heterogenität im *Lesen* signifikant verringert.

Im Kompetenzbereich *Zuhören* ist die Streuung im Jahr 2016 für Deutschland insgesamt um 7 Punkte größer als im Jahr 2011, dieser Unterschied ist jedoch nicht statistisch signifikant. Innerhalb der Länder hat sich die Streuung der Kompetenzwerte in keinem Land signifikant verringert. Signifikant heterogener verteilt als im Jahr 2011 sind die Zuhörkompetenzen im Jahr 2016 in Baden-Württemberg, Bremen, Nordrhein-Westfalen und Thüringen. Vergleicht man die Veränderungen in den Standardabweichungen der Kompetenzwerte in den Ländern mit der Veränderung in Deutschland insgesamt, so zeigt sich nur für Sachsen-Anhalt eine Abweichung: Die geringfügige Verringerung der Heterogenität in Sachsen-Anhalt unterscheidet sich signifikant von der geringfügigen Vergrößerung der Heterogenität in Deutschland insgesamt.

Die für Deutschland insgesamt festgestellte Streuung im Kompetenzbereich *Orthografie* hat sich zwischen den Messzeitpunkten 2011 und 2016 um 6 Punkte signifikant vergrößert (siehe Abb. 6.4).

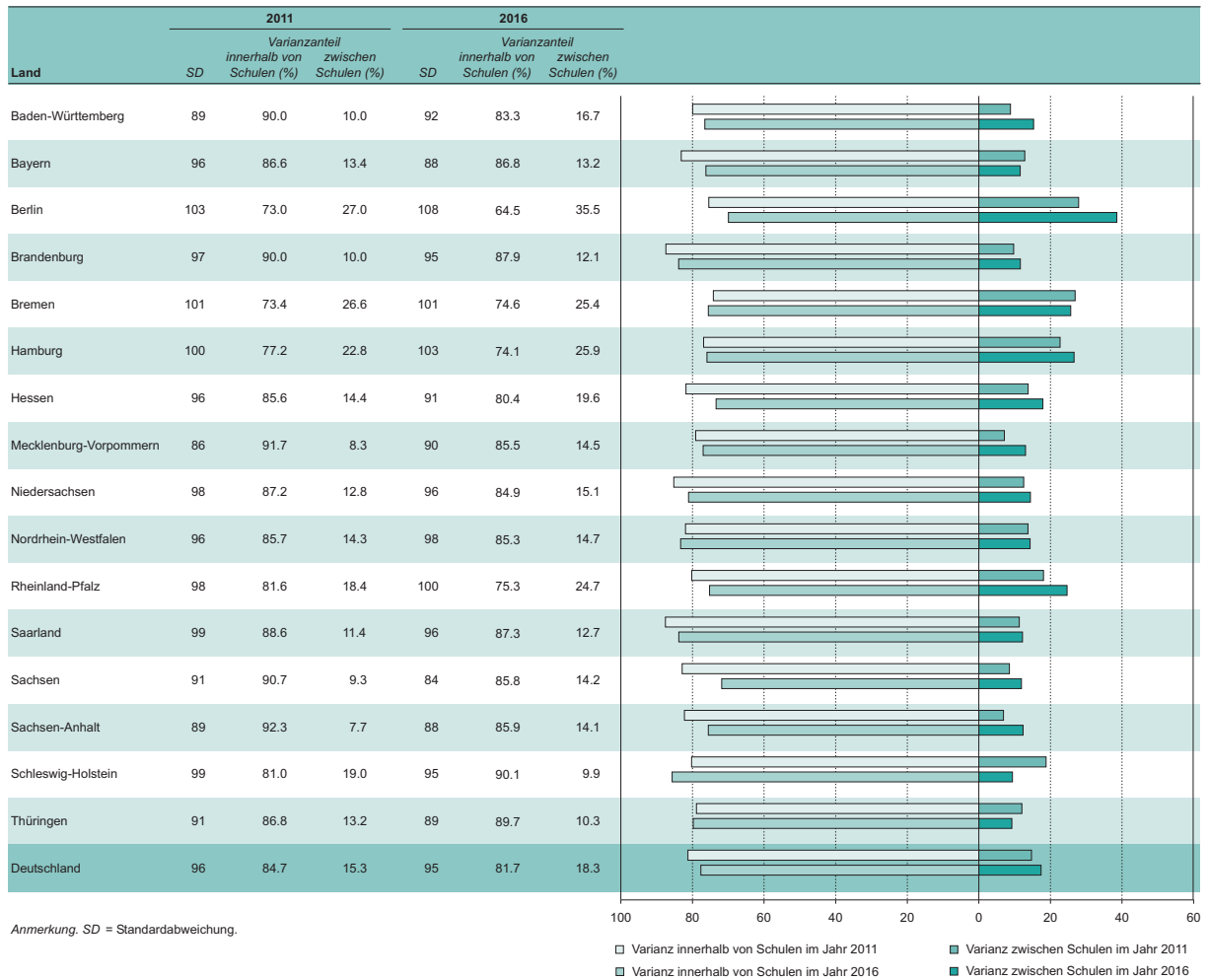
6.1.6 Vergleich der Heterogenität der in den Jahren 2011 und 2016 erreichten Kompetenzen innerhalb und zwischen Schulen

Im letzten Abschnitt wurde beschrieben, wie sich die Heterogenität der von Viertklässlerinnen und Viertklässlern erreichten Kompetenzen im Fach Deutsch zwischen den Jahren 2011 und 2016 verändert hat. Die Gesamtvarianz lässt sich dabei zerlegen in einen Teil, der zwischen Schulen liegt und in einen Teil, der innerhalb von Schulen liegt. Die Varianz zwischen Schulen gibt an, wie stark sich die Kompetenzen der Schülerinnen und Schüler in unterschiedlichen Schulen unterscheiden. Die Varianz innerhalb von Schulen kennzeichnet, mit welcher Heterogenität der Kompetenzen Lehrkräfte in ihren Schulen umzugehen haben. Als lediglich deskriptives Maß für den Varianzanteil, der zwischen Schulen liegt, wird die Intraklassenkorrelation (*ICC*, siehe Kapitel 3.2) verwendet. Veränderungen in den Intraklassenkorrelationen zwischen den Erhebungsjahren 2011 und 2016, die größer als 5 Prozent sind, werden als substantiell betrachtet.

Die Heterogenität, die zwischen Schulen besteht, wird unter anderem davon beeinflusst, ob Kinder mit sonderpädagogischem Förderbedarf an Förder- oder an allgemeinen Schulen unterrichtet werden. Berücksichtigt man Förderschulen bei der Berechnung der Intraklassenkorrelation, so fallen die zwischen Schulen liegenden Unterschiede in der Varianz insbesondere in solchen Ländern hoch aus, in denen ein hoher Anteil der Mädchen und Jungen mit sonderpädagogischem Förderbedarf an Förderschulen unterrichtet wird (Böhme & Weirich, 2012). Für die Berechnung der Varianzunterschiede innerhalb und zwischen Schulen in den Kompetenzbereichen *Lesen* und *Zuhören* wurden daher alle Förderschulen ausgeschlossen. Im Bereich *Orthografie* wurden darüber hinaus alle Schülerinnen und Schüler ausgeschlossen, die einen dokumentierten sonderpädagogischen Förderbedarf haben, da im Jahr 2011 keine Förderschulen einbezogen wurden und die Verteilung der Varianz auch im Trend analysiert werden soll.

In den Tabellenteilen der Abbildungen 6.7 bis 6.9 sind die Standardabweichungen (*SD*) sowie die zwischen und innerhalb von Schulen liegenden Varianzanteile für den jeweiligen Kompetenzbereich aufgeführt. Da bei der

Abbildung 6.7: Varianz innerhalb und zwischen Schulen im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 (ohne Förderschulen) im Kompetenzbereich *Lesen*

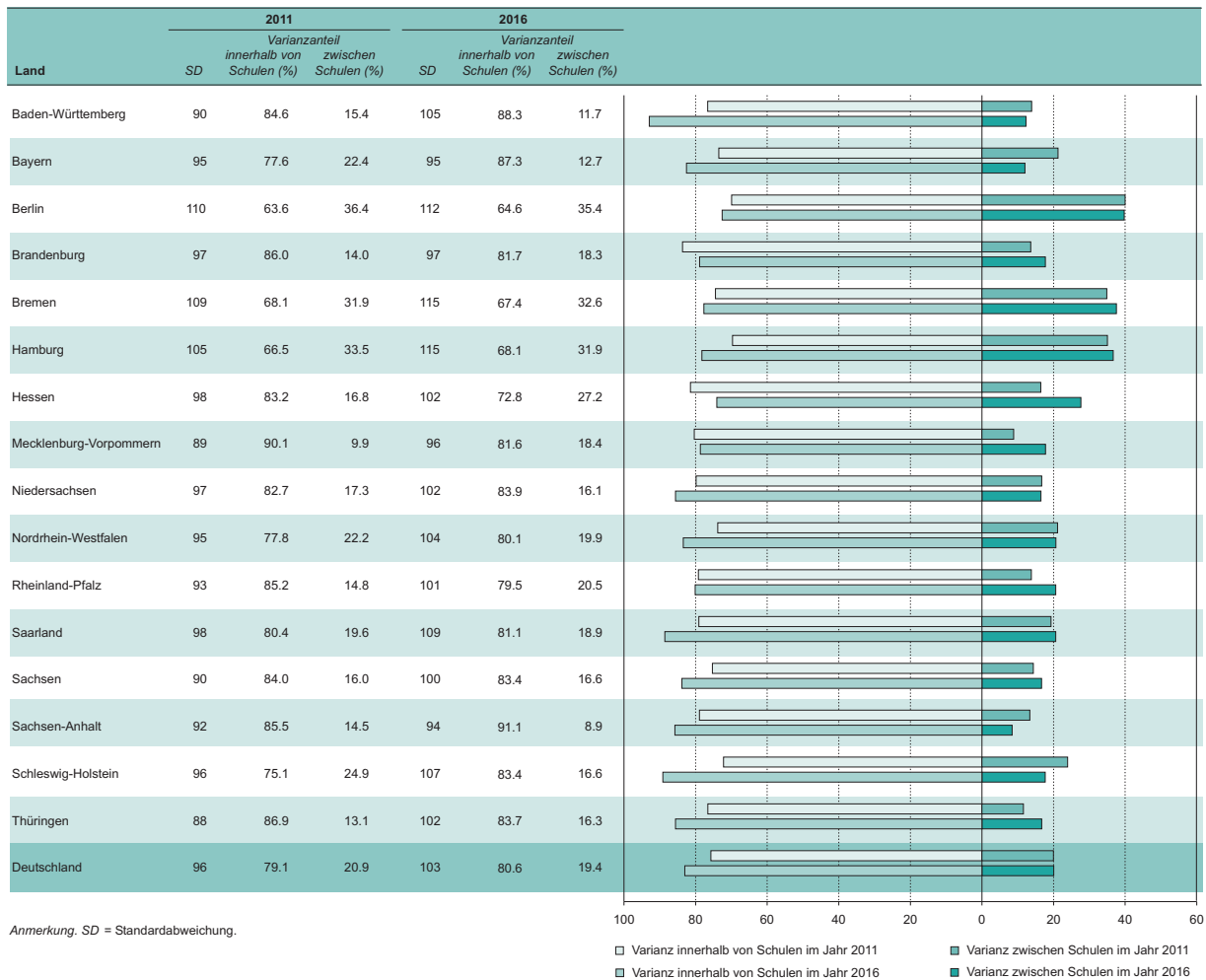


Berechnung alle Förderschulen (*Lesen* und *Zuhören*) beziehungsweise sämtliche Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (*Orthografie*) ausgeschlossen wurden, unterscheiden sich die Standardabweichungen, die in Abbildung 6.4 aufgeführt sind, von den Standardabweichungen in den Abbildungen 6.7 bis 6.9. Im rechten Teil der Abbildungen ist die Varianz in der jeweiligen Stichprobe zusätzlich grafisch dargestellt. Die Gesamtlänge der Balken gibt die Streuung der erreichten Kompetenzen pro Land im Erhebungsjahr 2011 beziehungsweise 2016 an, wobei längere Balken heterogenere Kompetenzen, kürzere Balken homogenere Kompetenzen markieren. Der auf der rechten Seite liegende, dunklere Teil des Balkens verdeutlicht den Anteil der Varianz, der für das jeweilige Land zwischen Schulen liegt. Der hellere Teil des Balkens auf der linken Seite kennzeichnet den Anteil der Varianz, der innerhalb von Schulen liegt.

Im Kompetenzbereich *Lesen* beträgt der Anteil zwischenschulischer Varianz im Erhebungsjahr 2016 für Deutschland insgesamt 18 Prozent (Abb. 6.7). Dies deutet darauf hin, dass sich Viertklässlerinnen und Viertklässler innerhalb von Schulen deutlich in ihren erreichten Kompetenzen unterscheiden.

Betrachtet man die Varianzverteilung im Kompetenzbereich *Lesen* auf Länderebene, zeigen sich zum Teil unterschiedliche Muster. Besonders geringe Kompetenzunterschiede zwischen Schulen finden sich in den Ländern Schleswig-

Abbildung 6.8: Varianz innerhalb und zwischen Schulen im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 (ohne Förderschulen) im Kompetenzbereich *Zuhören*



Holstein und Thüringen, in denen lediglich etwa 10 Prozent der Kompetenzunterschiede zwischen Schulen liegen. Dahingegen liegt in Berlin 36 Prozent der Varianz in den Lesekompetenzen zwischen Schulen. Auch in den Stadtstaaten Bremen und Hamburg sowie in Rheinland-Pfalz sind die Kompetenzunterschiede zwischen Schulen relativ groß. Sie machen in diesen Ländern etwa ein Viertel der Gesamtvarianz aus.

Vergleicht man die zwischenschulischen Kompetenzunterschiede für die Jahre 2011 und 2016, zeigt sich, dass diese in sechs Ländern um mehr als 5 Prozentpunkte zugenommen haben. So hängt das erreichte Kompetenzniveau der Viertklässlerinnen und Viertklässler in Baden-Württemberg, Berlin, Hessen, Mecklenburg-Vorpommern, Rheinland-Pfalz und Sachsen-Anhalt im Jahr 2016 stärker damit zusammen, welche Schule sie besuchen, als im Jahr 2011.

Eine substantielle Verringerung der Varianz zwischen Schulen lässt sich im *Lesen* hingegen für Schleswig-Holstein erkennen. Lassen sich im Jahr 2011 19 Prozent der Varianz auf Unterschiede zwischen den Schulen zurückführen, sind es im Jahr 2016 nur noch knapp 10 Prozent.

Ähnlich wie im *Lesen* kann auch im Kompetenzbereich *Zuhören* im Jahr 2016 etwa ein Fünftel (19 %) der Heterogenität auf Unterschiede zwischen Schulen

Abbildung 6.9: Varianz innerhalb und zwischen Schulen im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 (ohne Förderschüler) im Kompetenzbereich *Orthografie*

und vier Fünftel auf Unterschiede innerhalb von Schulen zurückgeführt werden (Abb. 6.8).

Auf Länderebene finden sich in Sachsen-Anhalt mit einem Varianzanteil von lediglich 9 Prozent besonders geringe Kompetenzunterschiede zwischen Schulen. Auch in Baden-Württemberg und Bayern erreichen Viertklässlerinnen und Viertklässler im *Zuhören* weitgehend unabhängig von der besuchten Schule ähnliche Kompetenzstände. In den drei Stadtstaaten Berlin, Bremen und Hamburg hängen die Unterschiede im erreichten Lernstand hingegen zu etwa einem Drittel von Unterschieden zwischen Schulen ab.

Im Vergleich der Veränderungen der Varianzanteile zwischen den Erhebungsjahren 2011 und 2016 ergeben sich für Deutschland insgesamt keine substanziellen Unterschiede. Auf Länderebene hat sich jedoch in Bayern, Sachsen-Anhalt und Schleswig-Holstein im Jahr 2016 der zwischenschulische Varianzanteil deutlich verkleinert. Hier hängen die Zuhörkompetenzen im Jahr 2016 weniger stark mit der besuchten Schule zusammen als noch im Jahr 2011. In den Ländern Hessen, Mecklenburg-Vorpommern und Rheinland-Pfalz hat sich der Anteil der zwischenschulischen Varianz hingegen deutlich vergrößert.

Im Kompetenzbereich *Orthografie* lassen sich in Deutschland insgesamt im Jahr 2016 etwa 18 Prozent der Heterogenität auf Unterschiede zwischen Schulen zurückführen, rund 82 Prozent der Unterschiede liegen somit innerhalb von Schulen (Abb. 6.9).

Auf Länderebene ist der Anteil zwischenschulischer Varianz in Baden-Württemberg, Bayern und Thüringen vergleichsweise gering. In Hessen hingegen kann etwa ein Viertel der Varianz in der *Orthografie* auf zwischenschulische Unterschiede zurückgeführt werden.

Betrachtet man die Veränderungen der Varianzanteile zwischen den Erhebungsjahren 2011 und 2016, zeigen sich für Deutschland insgesamt kaum Unterschiede. Analysen auf Länderebene konnten für das Jahr 2011 nicht durchgeführt werden.

6.1.7 Zusammenfassung

Für die im Jahr 2016 erreichten Mittelwerte im Fach Deutsch gilt, dass Kinder in Bayern in allen drei getesteten Bereichen Kompetenzen zeigen, die jeweils signifikant über dem deutschen Mittelwert liegen. Gleichzeitig ist auch die Standardabweichung in Bayern in allen drei Domänen vergleichsweise gering. Im Vergleich der Länder wird in Bayern damit das wünschenswerte Befundmuster eines hohen Kompetenzniveaus bei gleichzeitig niedriger Heterogenität im Fach Deutsch durchweg am besten erreicht.

In den Ländern Sachsen und Schleswig-Holstein erreichen die Viertklässlerinnen und Viertklässler im *Lesen* und im *Zuhören* im Mittel signifikant höhere Kompetenzwerte als bundesweit. Dabei unterscheidet sich die Standardabweichung in diesen Ländern nicht von der Streuung in Deutschland insgesamt.

Im Saarland erreichen Schülerinnen und Schüler in der *Orthografie* Kompetenzwerte, die signifikant über dem für Deutschland insgesamt ermittelten Wert liegen. Gleichzeitig unterscheidet sich die Streuung nicht von der Gesamtstreuung für Deutschland.

Weniger günstige Ergebnisse erreichen die Länder Berlin und Bremen, die in allen im Fach Deutsch überprüften Kompetenzbereichen Werte erzielen, die signifikant unterhalb des gesamtdeutschen Mittelwerts liegen. In Bremen sind darüber hinaus in allen drei Domänen die Kompetenzunterschiede zwischen den Schülerinnen und Schülern signifikant größer als in Deutschland insgesamt. In Berlin ist dies im *Lesen* und im *Zuhören* der Fall.

In sechs weiteren Ländern erreichen Schülerinnen und Schüler in jeweils einem der Kompetenzbereiche signifikant geringere Werte als Schülerinnen und Schüler bundesweit. In Nordrhein-Westfalen ist dies im Bereich *Lesen* der Fall, in Rheinland-Pfalz und Sachsen-Anhalt im Bereich *Zuhören*. In Sachsen-Anhalt sind die erreichten Kompetenzen dabei signifikant homogener als in Deutschland insgesamt. Hamburg, Niedersachsen und Thüringen weisen im Bereich *Orthografie* Kompetenzwerte auf, die signifikant unterhalb des deutschen Gesamtwertes liegen. Die Streuungen unterscheiden sich in diesen Ländern aber nicht von der für Deutschland insgesamt festgestellten Streuung.

Signifikant homogenere Kompetenzen als im gesamtdeutschen Vergleich sind in Thüringen im Bereich *Lesen* sowie in Brandenburg im Bereich *Zuhören* zu verzeichnen. Die Kompetenzen der Schülerinnen und Schüler im Bereich *Zuhören* in Hamburg fallen hingegen signifikant heterogener aus als in Deutschland insgesamt.

Betrachtet man die Veränderungen der erzielten Kompetenzen zwischen den Jahren 2011 und 2016, zeigt sich für Deutschland insgesamt in allen drei Domänen eine signifikante Verringerung der erreichten Mittelwerte. In der *Orthografie* hat sich zusätzlich die Heterogenität vergrößert. Innerhalb der Länder variiert die Ausprägung signifikanter Trends zwischen –14 Punkten im *Lesen* in Sachsen-Anhalt und –26 Punkten im *Zuhören* in Baden-Württemberg. Signifikant positive Trends sind für kein Land zu verzeichnen.

In Bremen, Sachsen-Anhalt und Thüringen haben sich die Kompetenzen in den beiden Domänen *Lesen* und *Zuhören* signifikant verringert. In Bremen und Thüringen hat im Kompetenzbereich *Zuhören* zusätzlich die Heterogenität zugenommen. Auch Schülerinnen und Schüler in Baden-Württemberg und Nordrhein-Westfalen erreichen im *Zuhören* im Jahr 2016 signifikant geringere Werte bei steigender Streuung. In Mecklenburg-Vorpommern, Niedersachsen und Rheinland-Pfalz ist lediglich der Trend im *Zuhören* signifikant negativ, die Streuung der Leistung hat sich in diesen Ländern aber kaum verändert.

Literatur

- 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* (S. 357–375). Weinheim: Beltz.
- Böhme, K. & Bremerich-Vos, A. (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 und Basel: Beltz.
- Böhme, K. & Weirich, S. (2012). Der Ländervergleich im Fach Deutsch. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (S. 103–116). Münster: Waxmann.
- Pant, H. A., Böhme, K. & Köller, O. (2012). Das Kompetenzkonzept der Bildungsstandards und die Entwicklung von Kompetenzstufenmodellen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (S. 49–55). Münster: Waxmann.

6.2 Mittelwerte und Streuungen der im Fach Mathematik erreichten Kompetenzen

Nicole Haag und Pauline Kohrt

In diesem Kapitel wird für das Fach Mathematik beschrieben, welche Kompetenzen Schülerinnen und Schüler der 4. Jahrgangsstufe in den einzelnen Ländern im Mittel erreichen. Zudem wird die Heterogenität der erreichten Kompetenzen in den Ländern dargestellt. Die Ergebnisse beziehen sich zum einen auf die Globalskala mathematischer Kompetenzen und zum anderen auf die fünf mathematischen Leitideen *Zahlen und Operationen*, *Raum und Form*, *Muster und Strukturen*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit*. Die in Mathematik eingesetzten Testaufgaben werden in Kapitel 1.3 beschrieben.

Die erreichten Kompetenzen werden auf der Berichtsmetrik dargestellt, die bereits im IQB-Ländervergleich 2011 für die Population der Viertklässlerinnen und Viertklässler in Deutschland zugrunde gelegt wurde. Sowohl für die Globalskala als auch für die einzelnen Kompetenzbereiche wurde hierbei im Jahr 2011 jeweils ein Mittelwert von $M = 500$ und eine Standardabweichung von $SD = 100$ festgelegt (Haag & Roppelt, 2012). Unterschiede in den erreichten Kompetenzen werden als Punktwertdifferenzen auf der Berichtsmetrik dargestellt (vgl. Kapitel 3.2). Um die ländervergleichende Interpretation dieser Differenzen zu veranschaulichen, können diese zu Lernzuwächsen ins Verhältnis gesetzt werden, die im Mittel am Ende der Grundschulzeit zu erwarten sind. Entsprechende Schätzungen für auf den Bildungsstandards basierende Kompetenzmessungen liegen aus Normierungsstudien des IQB aus den Jahren 2006 und 2007 vor. Hierbei handelt es sich um in der 3. und 4. Jahrgangsstufe durchgeführte, bundesweit repräsentative Erhebungen in den Fächern Deutsch und Mathematik. Ein Vergleich der von Schülerinnen und Schülern der 3. und der 4. Jahrgangsstufe im Mittel erreichten Kompetenzen zeigt, dass der Unterschied auf der Globalskala mathematischer Kompetenz etwa 80 Punkte auf der Berichtsmetrik beträgt (Reiss & Winkelmann, 2009).

Nachfolgend stellen wir in Abschnitt 6.2.1 zunächst für das Jahr 2016 dar, welches Kompetenzniveau die Schülerinnen und Schüler im Fach Mathematik in den einzelnen Ländern im Durchschnitt erreicht haben. Zudem wird in Abschnitt 6.2.2 die Heterogenität der innerhalb der Länder erreichten Kompetenzen untersucht. Abschnitt 6.2.3 fasst die Ergebnisse für das Jahr 2016 kurz zusammen. In Abschnitt 6.2.4 wird der Frage nachgegangen, wie sich das durchschnittliche Kompetenzniveau zwischen den Jahren 2011 und 2016 im Trend verändert hat. Trendanalysen für die Heterogenität der erreichten Kompetenzen finden sich in Abschnitt 6.2.5. Anschließend werden in Abschnitt 6.2.6 Veränderungen in der Heterogenität beschrieben, die sowohl innerhalb als auch zwischen Schulen besteht. In Abschnitt 6.2.7 schließlich werden die zentralen Ergebnisse dieses Kapitels zusammengefasst.

6.2.1 Im Mittel erreichte Kompetenzen im Jahr 2016

Neben dem seit der Berichterstattung im Jahr 2015 gesetzten zentralen Fokus auf eine kriteriale Vergleichsperspektive mit den von Schülerinnen und Schülern erreichten Kompetenzstufen (vgl. Kapitel 1.1 und Kapitel 5) sind die Mittelwerte

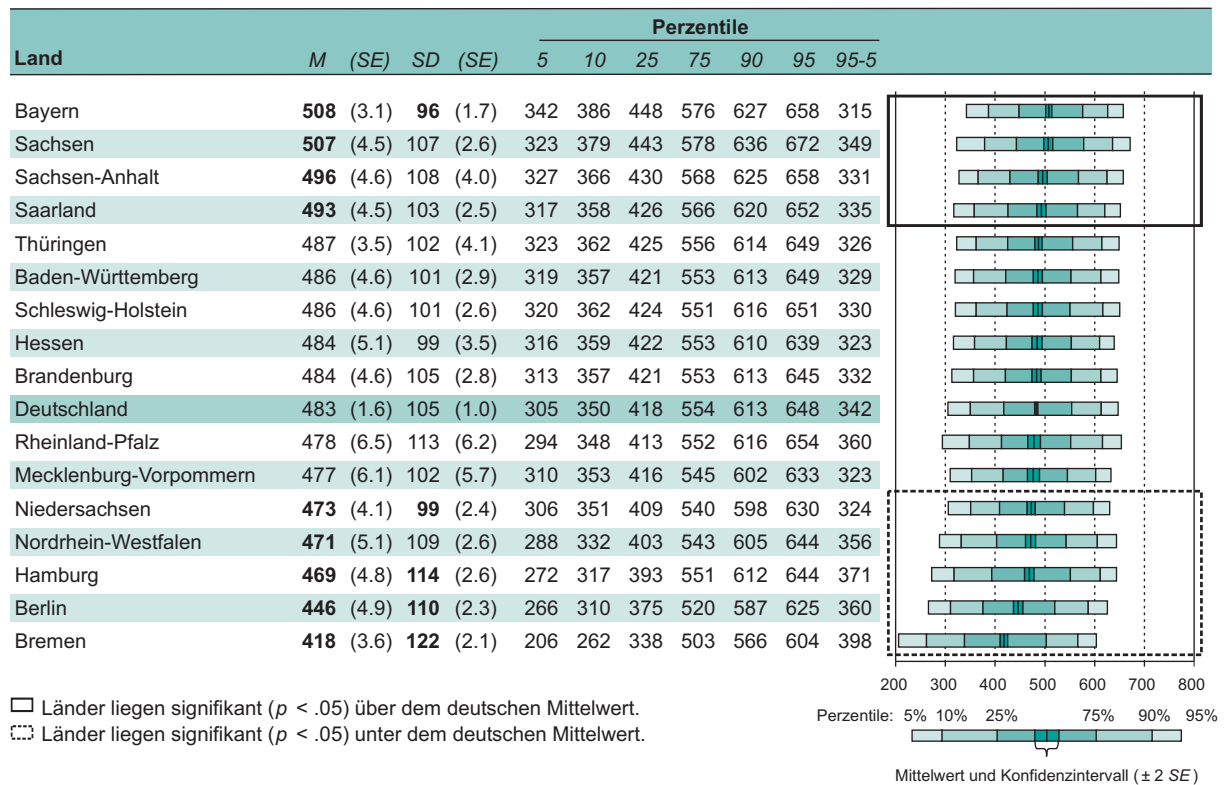
und Streuungen der Kompetenzen auf Länderebene nach wie vor wichtige Indikatoren für das Erreichen bildungspolitischer Ziele. Die für das Fach Mathematik ermittelten Ländermittelwerte werden daher im Folgenden sowohl für die Globalskala mathematischer Kompetenzen (Abb. 6.10) als auch separat für die fünf Leitideen (Abb. 6.11 bis 6.15) dargestellt. Die Länder sind in diesen Abbildungen absteigend nach dem jeweils erreichten Kompetenzmittelwert geordnet.

Im Tabellenteil der Abbildungen werden neben den Mittelwerten (M) auch die Standardfehler (SE), die Standardabweichungen (SD) sowie verschiedene Perzentile aufgeführt. Die Interpretation dieser statistischen Kennwerte wird in Kapitel 3.2 erläutert. Im Abbildungsteil ist durch Einrahmungen gekennzeichnet, welche Ländermittelwerte sich signifikant vom deutschen Gesamtmittelwert im jeweiligen Bereich unterscheiden. Signifikant höhere Landeswerte werden durch einen durchgezogenen Rahmen, signifikant geringere Landeswerte durch einen gestrichelten Rahmen hervorgehoben.

In die Berechnungen wurden jeweils *alle Viertklässlerinnen und Viertklässler* einbezogen – also auch Schülerinnen und Schüler, bei denen ein sonderpädagogischer Förderbedarf festgestellt wurde, auch wenn sie zieldifferent unterrichtet werden (vgl. Kapitel 3.1 für eine Erläuterung dieses Vorgehens).

Für die Globalskala mathematischer Kompetenz (Abb. 6.10) beträgt der Mittelwert im Jahr 2016 für Deutschland insgesamt 483 Punkte. Die in den einzelnen Ländern von den Schülerinnen und Schülern im Mittel erreichten Kompetenzen bewegen sich in einem Bereich von 418 Punkten in Bremen und 508 Punkten in Bayern. Zwischen dem Land mit dem niedrigsten Mittelwert und dem Land mit

Abbildung 6.10: Mittelwerte, Streuungen, Perzentile und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Fach Mathematik (Globalskala)



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

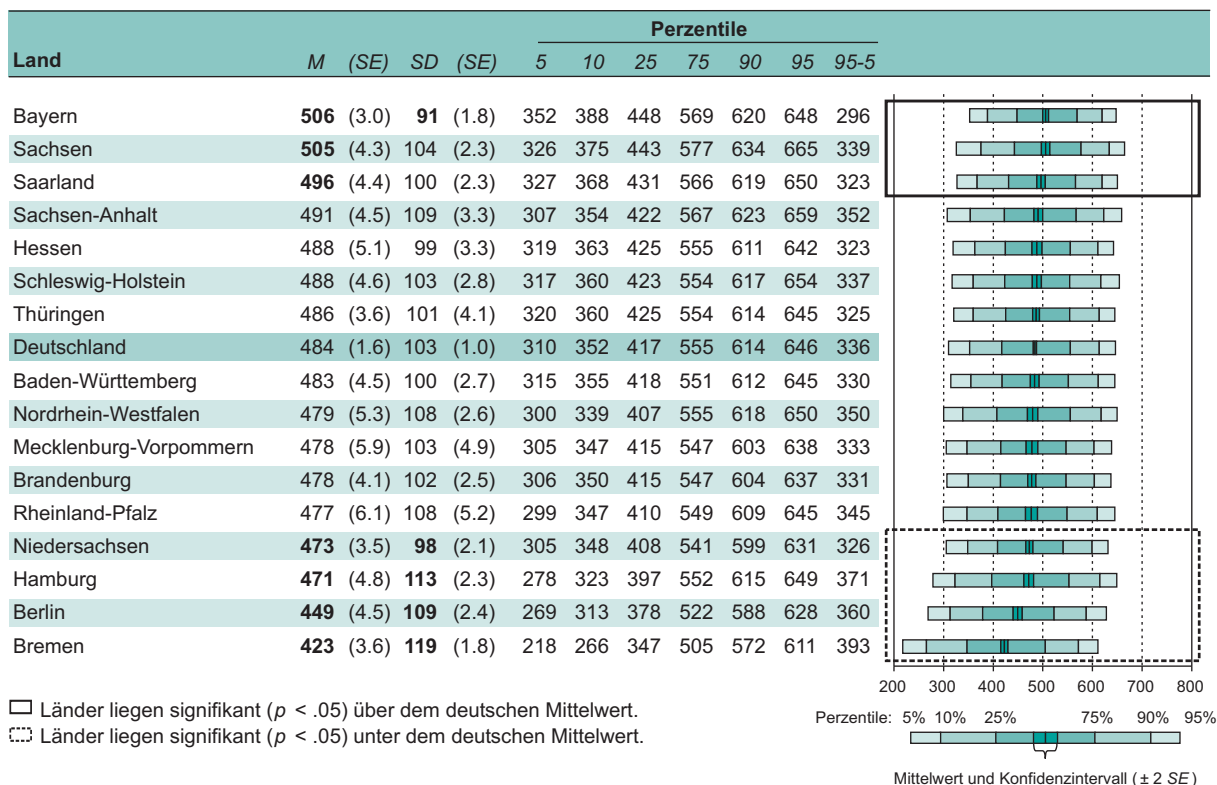
dem höchsten Mittelwert liegen also 90 Punkte auf der Berichtsmetrik, was dem Kompetenzzuwachs entspricht, der zum Ende der Primarstufe in etwas mehr als einem Schuljahr zu erwarten ist.

Die Schülerinnen und Schüler in Bayern, Sachsen, Sachsen-Anhalt und dem Saarland erreichen im Durchschnitt Kompetenzwerte, die signifikant über dem deutschen Mittelwert liegen. In Niedersachsen und Nordrhein-Westfalen sowie in den Stadtstaaten Hamburg, Berlin und Bremen hingegen liegen die mittleren Kompetenzen im Fach Mathematik statistisch signifikant unter dem Wert für Deutschland insgesamt.

Für die einzelnen Leitideen ergibt sich insgesamt ein sehr ähnliches Bild wie für die Globalskala (vgl. Abb. 6.11 bis 6.15). Dies wird auch anhand der hohen Zusammenhänge deutlich, die zwischen den Ländermittelwerten auf der Globalskala und auf den Skalen zu den einzelnen Leitideen bestehen.¹ Spezifische Befundmuster einzelner Länder für die verschiedenen Kompetenzbereiche werden im Folgenden dargestellt.

Im Kompetenzbereich *Zahlen und Operationen* (Abb. 6.11) liegt der deutsche Gesamtmittelwert im Jahr 2016 bei 484 Punkten. Vergleicht man die Zusammensetzung der Gruppe der Länder, in denen signifikant über dem Bundesdurchschnitt liegende Ergebnisse erzielt werden, und die Zusammensetzung der Gruppe der Länder, in denen signifikant unter dem Bundesdurchschnitt liegende Ergebnisse erreicht werden, so zeigt sich eine weitgehende

Abbildung 6.11: Mittelwerte, Streuungen, Perzentile und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Zahlen und Operationen*



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

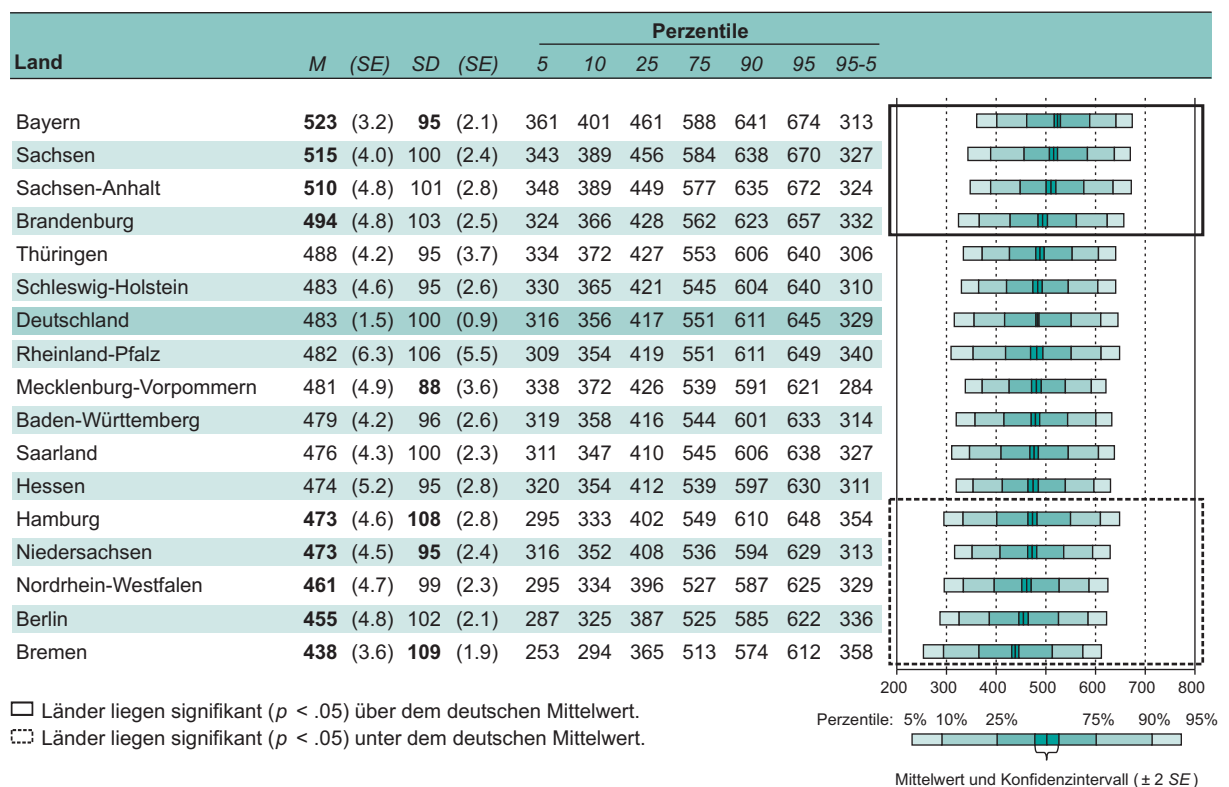
1 Diese Korrelationen reichen von $r = .89$ für den Bereich *Raum und Form* bis $r = .99$ für den Bereich *Zahlen und Operationen*.

Übereinstimmung mit dem Ergebnismuster für die Globalskala, wobei die Anzahl der Länder, deren Ergebnisse sich signifikant vom deutschen Mittelwert unterscheiden, im Bereich *Zahlen und Operationen* kleiner ist. In Bayern, in Sachsen und im Saarland werden in diesem Bereich im Mittel signifikant überdurchschnittliche, in Niedersachsen sowie in den drei Stadtstaaten Hamburg, Berlin und Bremen hingegen signifikant unterdurchschnittliche Kompetenzwerte festgestellt.

Im Kompetenzbereich *Raum und Form* erreichen Schülerinnen und Schüler in Deutschland insgesamt eine mittlere Kompetenz von 483 Punkten (Abb. 6.12). Die Gruppe der Länder, in denen überdurchschnittliche Kompetenzen erzielt werden, umfasst neben Bayern, Sachsen und Sachsen-Anhalt auch Brandenburg. Signifikant unterdurchschnittliche Kompetenzen sind hingegen im Mittel für Hamburg, Niedersachsen, Nordrhein-Westfalen, Berlin und Bremen festzustellen.

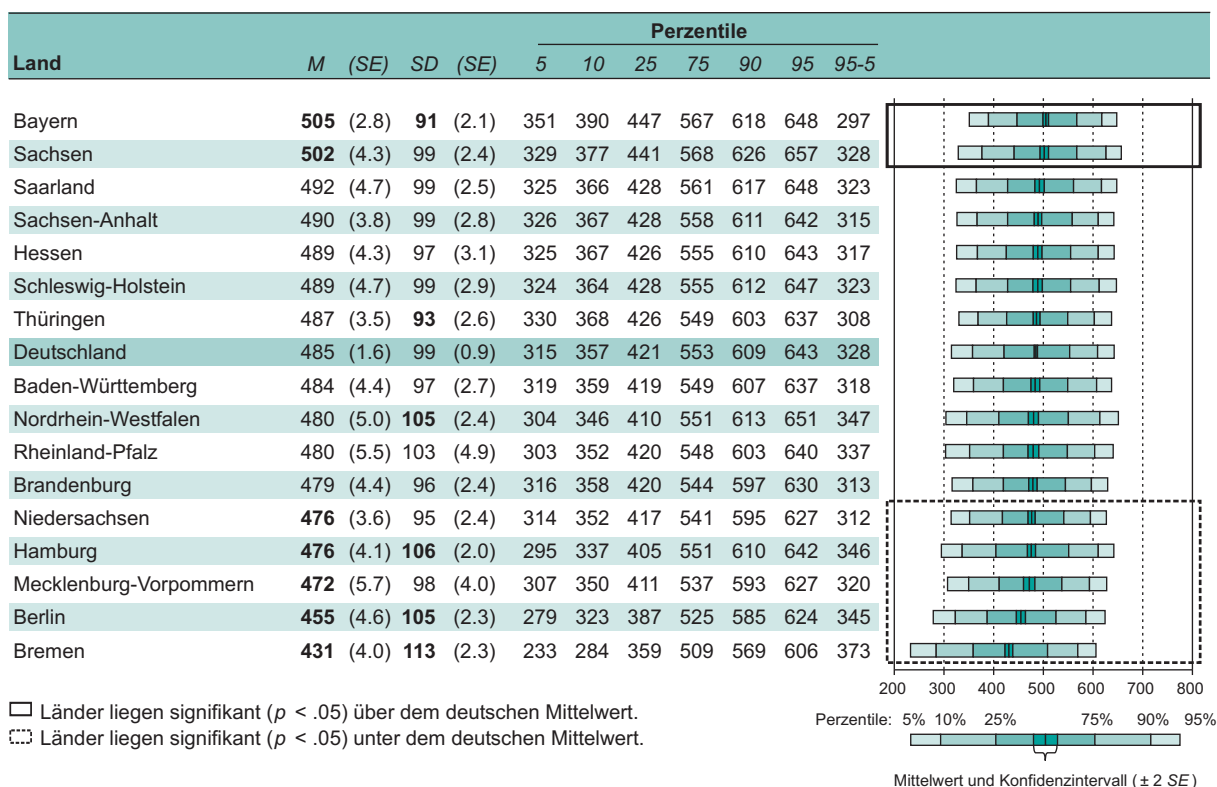
Im Kompetenzbereich *Muster und Strukturen* beträgt der mittlere Kompetenzwert in Deutschland insgesamt 485 Punkte (Abb. 6.13). In diesem Kompetenzbereich erzielen nur die Schülerinnen und Schüler in Bayern und Sachsen signifikant überdurchschnittliche Ergebnisse. Auch im unteren Leistungsbereich zeigen sich einige Verschiebungen im Vergleich zu den Ergebnissen für die Globalskala. Die in Niedersachsen, Hamburg, Berlin und Bremen erreichten Kompetenzmittelwerte liegen im Bereich *Muster und Strukturen* – ebenso wie auf der Globalskala – signifikant unter dem deutschen Mittelwert. Darüber hinaus

Abbildung 6.12: Mittelwerte, Streuungen, Perzentile und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Raum und Form*



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

Abbildung 6.13: Mittelwerte, Streuungen, Perzentile und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Muster und Strukturen*



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

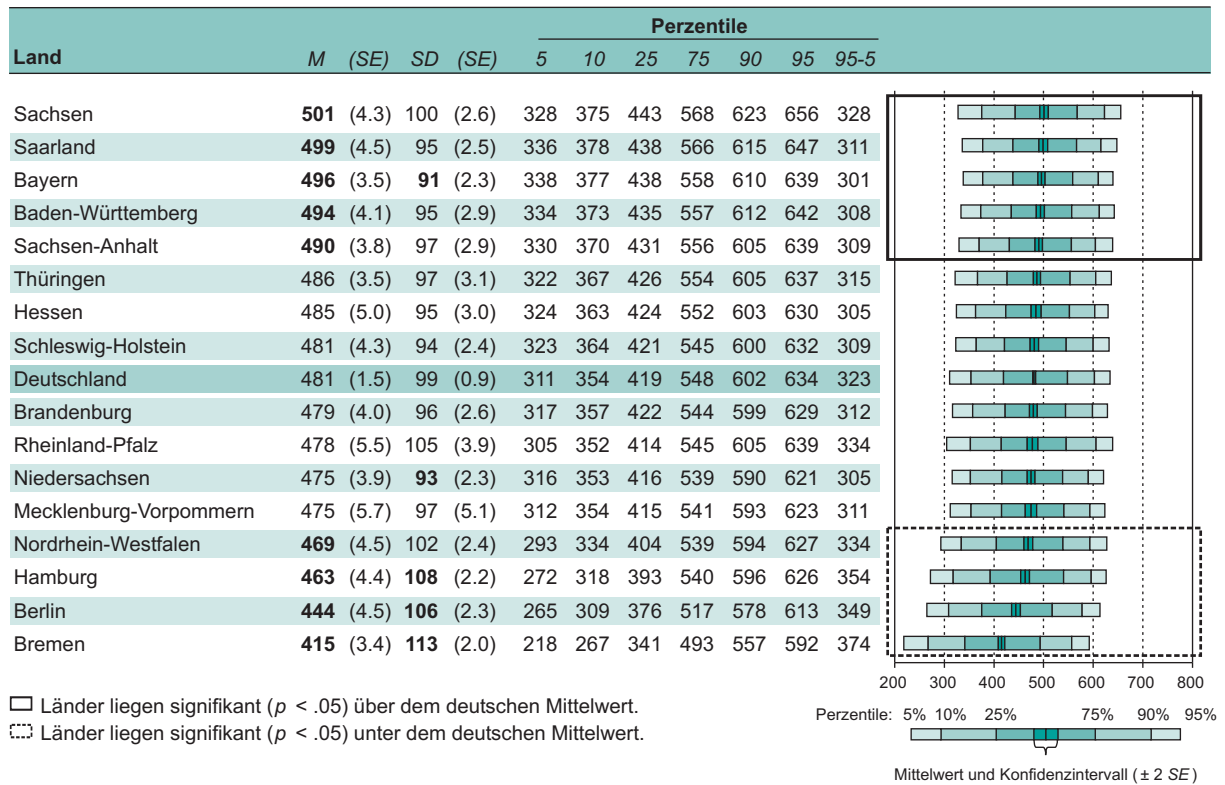
ist in Mecklenburg-Vorpommern ein mittlerer Kompetenzwert festzustellen, der sich signifikant unterhalb des bundesdeutschen Durchschnitts befindet.

Im Kompetenzbereich *Größen und Messen* beträgt der deutsche Gesamtmittelwert 481 Punkte (Abb. 6.14). Neben Sachsen, dem Saarland, Bayern und Sachsen-Anhalt gehört Baden-Württemberg zur Gruppe der Länder, in denen in diesem Bereich überdurchschnittliche Kompetenzwerte erreicht werden. Signifikant unterhalb des Wertes für Deutschland insgesamt liegen hingegen die in Nordrhein-Westfalen, Hamburg, Berlin und Bremen im Durchschnitt erreichten Kompetenzen.

Für den Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit* wurde in Deutschland insgesamt ein Kompetenzmittelwert von 488 Punkten festgestellt (Abb. 6.15). In diesem Kompetenzbereich werden nur in Bayern und Sachsen mittlere Kompetenzwerte erreicht, die signifikant über dem deutschen Mittelwert liegen. Die Landesmittelwerte Niedersachsens, Berlins und Bremens liegen demgegenüber signifikant unterhalb des Bundesdurchschnitts.

Der Abstand zwischen dem jeweils höchsten und niedrigsten Ländermittelwert beträgt 83 Punkte in *Zahlen und Operationen*, 85 Punkte in *Raum und Form*, 74 Punkte in *Muster und Strukturen*, 86 Punkte in *Größen und Messen* sowie 76 Punkte in *Daten, Häufigkeit und Wahrscheinlichkeit*. Ebenso wie für die Globalskala entsprechen diese Differenzen ungefähr dem am Ende der Primarstufe in einem Schuljahr zu erwartenden Lernzuwachs in der mathematischen Kompetenz, der etwa 80 Punkte beträgt (Reiss & Winkelmann, 2009).

Abbildung 6.14: Mittelwerte, Streuungen, Perzentile und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Größen und Messen*

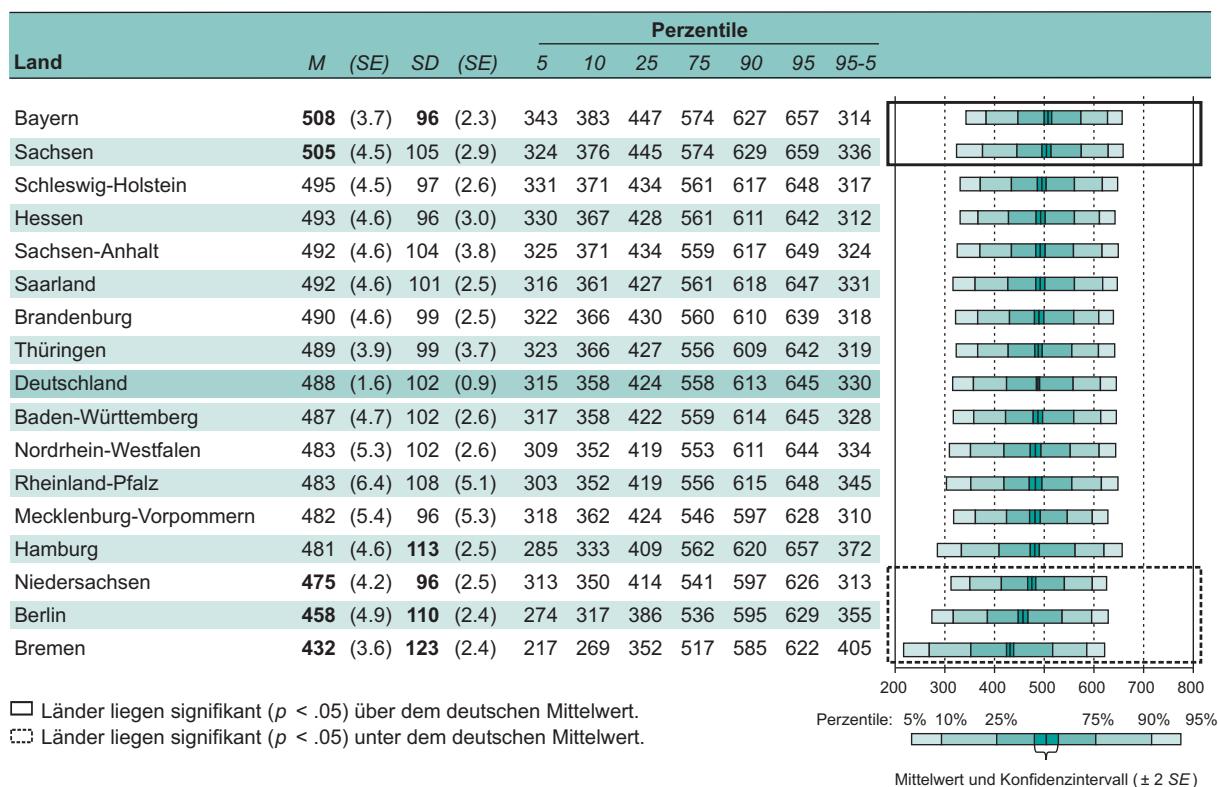


Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

Um relative Stärken und Schwächen zu beschreiben, kann zusätzlich zum Vergleich der in den einzelnen Kompetenzbereichen erzielten Landesmittelwerte mit dem jeweiligen deutschen Gesamtmittelwert auch die Heterogenität der Mittelwerte innerhalb jedes Landes betrachtet werden. Dabei stehen nicht Unterschiede *zwischen* den Ländern, sondern *innerhalb* der Länder im Fokus. Profile der in den Ländern erreichten Kompetenzen können als Differenzen der Landesmittelwerte jeder Leitidee zum Landesmittelwert auf der Globalskala bestimmt werden. Eine grafische Veranschaulichung der Kompetenzprofile kann auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

Die Ergebnisse dieser Analysen zeigen, dass die Kompetenzprofile in den meisten Ländern nur wenig ausgeprägt sind. Signifikante Abweichungen einzelner Bereiche vom Landesmittelwert auf der Globalskala lassen sich nur für vier Länder identifizieren. Signifikant höhere Werte in einzelnen Bereichen finden sich in Bayern für den Bereich *Raum und Form*, in Bremen für die Bereiche *Raum und Form*, *Muster und Strukturen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* und in Sachsen-Anhalt für den Bereich *Raum und Form*. Signifikant geringere Kompetenzen als auf der Globalskala werden hingegen in Bayern im Bereich *Größen und Messen* sowie im Saarland im Bereich *Raum und Form* erreicht. Diese Abweichungen könnten auf länderspezifische Schwerpunktsetzungen im Mathematikunterricht zurückzuführen sein.

Abbildung 6.15: Mittelwerte, Streuungen, Perzentile und Perzentilbänder für die Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Kompetenzbereich *Daten, Häufigkeit und Wahrscheinlichkeit*



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann der Wert in der Spalte 95-5 minimal von der Differenz der entsprechenden Perzentile abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) vom jeweiligen Wert für Deutschland.

6.2.2 Heterogenität der erreichten Kompetenzen im Jahr 2016

In diesem Abschnitt wird die Streuung der im Fach Mathematik erreichten Kompetenzen in den Ländern betrachtet. Hierbei sind Befundmuster wünschenswert, die durch ein insgesamt hohes Niveau mit einer geringen Standardabweichung als Maß der Streuung innerhalb der Schülerschaft gekennzeichnet sind. Auch Schülerinnen und Schüler am unteren Rand der Leistungsverteilung sollten einen möglichst hohen Kompetenzwert erreichen.

Im Jahr 2016 beträgt die Standardabweichung für die Globalskala mathematischer Kompetenz in Deutschland insgesamt 105 Punkte (Abb. 6.10). Wie anhand der durch Fettdruck gekennzeichneten Werte zu erkennen ist, weisen Schülerinnen und Schüler in Bayern und in Niedersachsen im Vergleich mit der für Deutschland festgestellten Streuung signifikant homogenere Kompetenzwerte auf, wohingegen die in den Stadtstaaten Hamburg, Berlin und Bremen erreichten Kompetenzwerte signifikant heterogener sind. Das angestrebte Befundmuster einer hohen mittleren Kompetenz bei gleichzeitiger geringer Streuung wird im Vergleich der Länder in Bayern am besten erreicht. In den anderen drei Ländern, in denen signifikant höhere mittlere Kompetenzwerte erreicht werden als in Deutschland insgesamt, liegen die Standardabweichungen hingegen nahe der bundesweiten Streuung.

Die Standardabweichungen in den einzelnen Kompetenzbereichen (Abb. 6.11 bis Abb. 6.15) liegen für Deutschland insgesamt zwischen 99 Punkten (*Muster und Strukturen* sowie *Größen und Messen*) und 103 Punkten (*Zahlen*

und Operationen). Ebenso wie für die Globalskala sind die erreichten Kompetenzen der Schülerinnen und Schüler in Bremen und in Hamburg für alle Leitideen signifikant heterogener, in Bayern dagegen signifikant homogener als in Deutschland insgesamt. Heterogenere Kompetenzen als in Deutschland insgesamt finden sich darüber hinaus in Berlin für die vier Leitideen *Zahlen und Operationen*, *Muster und Strukturen*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* sowie in Nordrhein-Westfalen für die Leitidee *Muster und Strukturen*. Signifikant homogenere Kompetenzen verglichen mit Deutschland liegen dagegen auch in Niedersachsen für die Leitideen *Zahlen und Operationen*, *Raum und Form*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit*, in Mecklenburg-Vorpommern für die Leitidee *Raum und Form* sowie in Thüringen für die Leitidee *Muster und Strukturen* vor.

Als zusätzlicher Kennwert für die Heterogenität der erreichten Kompetenzen kann die Breite des Skalenabschnitts zwischen dem 5. und dem 95. Perzentil betrachtet werden (Abb. 6.10 bis 6.15). Die Befunde für diesen Indikator decken sich sowohl für die Globalskala als auch für die einzelnen Kompetenzbereiche mit den für die Standardabweichung ermittelten Befunden.

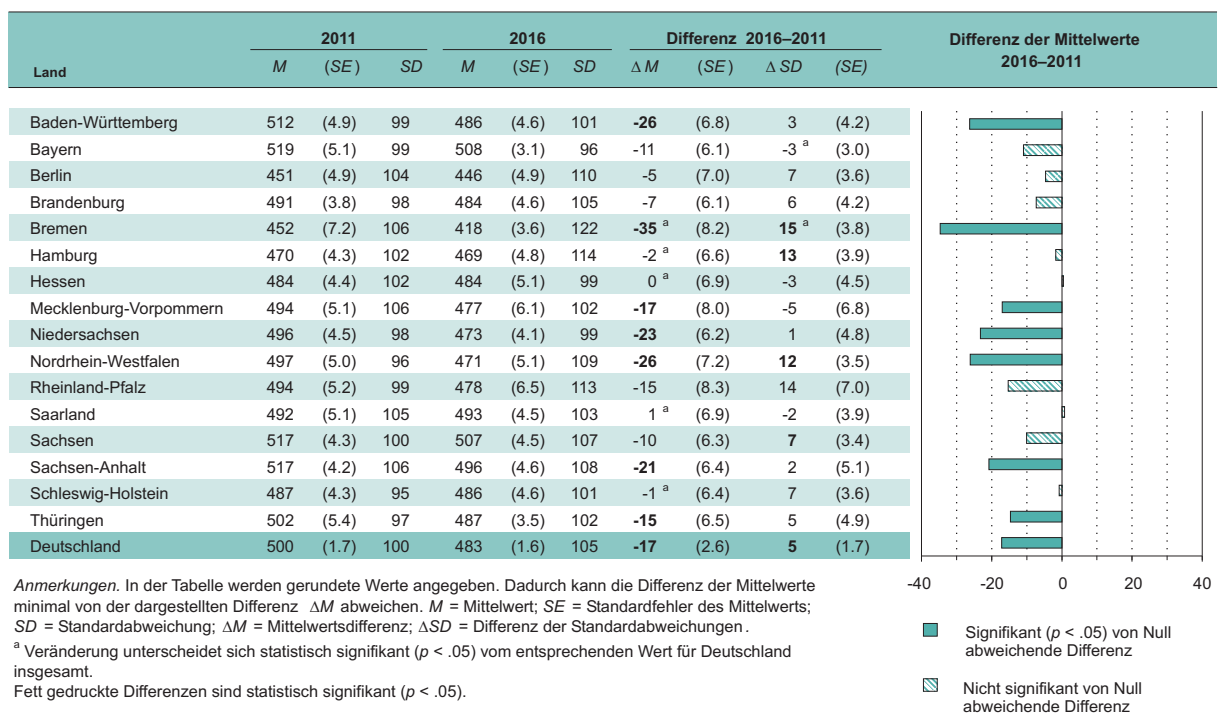
6.2.3 Zusammenfassung der Befunde für das Jahr 2016

Zusammenfassend zeigt sich für das Jahr 2016 im Fach Mathematik, dass die Schülerinnen und Schüler in Bayern und in Sachsen sowohl auf der Globalskala als auch auf den Skalen für die fünf Leitideen Kompetenzmittelwerte erreichen, die signifikant über dem Bundesdurchschnitt liegen. In Bayern ist zudem die Streuung der erreichten Kompetenzen in allen Bereichen signifikant geringer als in Deutschland insgesamt. Das Muster für Bayern zeigt, dass hier auch schwächere Schülerinnen und Schüler vergleichsweise hohe Kompetenzwerte erreichen. Die in Berlin und in Bremen erreichten Kompetenzmittelwerte liegen hingegen sowohl für die Globalskala als auch für alle getesteten Leitideen signifikant unterhalb des deutschen Mittelwerts. Diese beiden Länder verzeichnen zudem vergleichsweise heterogene Kompetenzwerte in allen (Bremen) oder nahezu allen (Berlin) getesteten Kompetenzbereichen. In Niedersachsen werden ebenfalls signifikant geringere Kompetenzwerte auf der Globalskala und in vier von fünf getesteten Leitideen erreicht als im Bundesdurchschnitt; die Kompetenzwerte sind hier jedoch homogener als in Deutschland insgesamt.

6.2.4 Vergleich der in den Jahren 2011 und 2016 im Mittel erreichten Kompetenzen

Ein zentrales Anliegen des IQB-Bildungstrends 2016 besteht in der Untersuchung von Trends über die Zeit. In diesem Abschnitt wird daher geprüft, inwieweit sich die im Durchschnitt erreichten mathematischen Kompetenzen zwischen den Jahren 2011 und 2016 verändert haben. Dazu werden zwei Perspektiven eingenommen (vgl. Kapitel 1.1): Zunächst werden die in den Jahren 2011 und 2016 im Mittel erzielten Kompetenzen verglichen, um festzustellen, ob innerhalb der Länder signifikante Trends vorliegen. Anschließend werden die Veränderungen innerhalb der Länder mit den für Deutschland insgesamt ermittelten Veränderungen verglichen, um zu untersuchen, ob diese signifikant stärker oder schwächer ausfallen.

Abbildung 6.16: Vergleich der in den Jahren 2011 und 2016 erreichten Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe im Fach Mathematik (*Globalskala*)

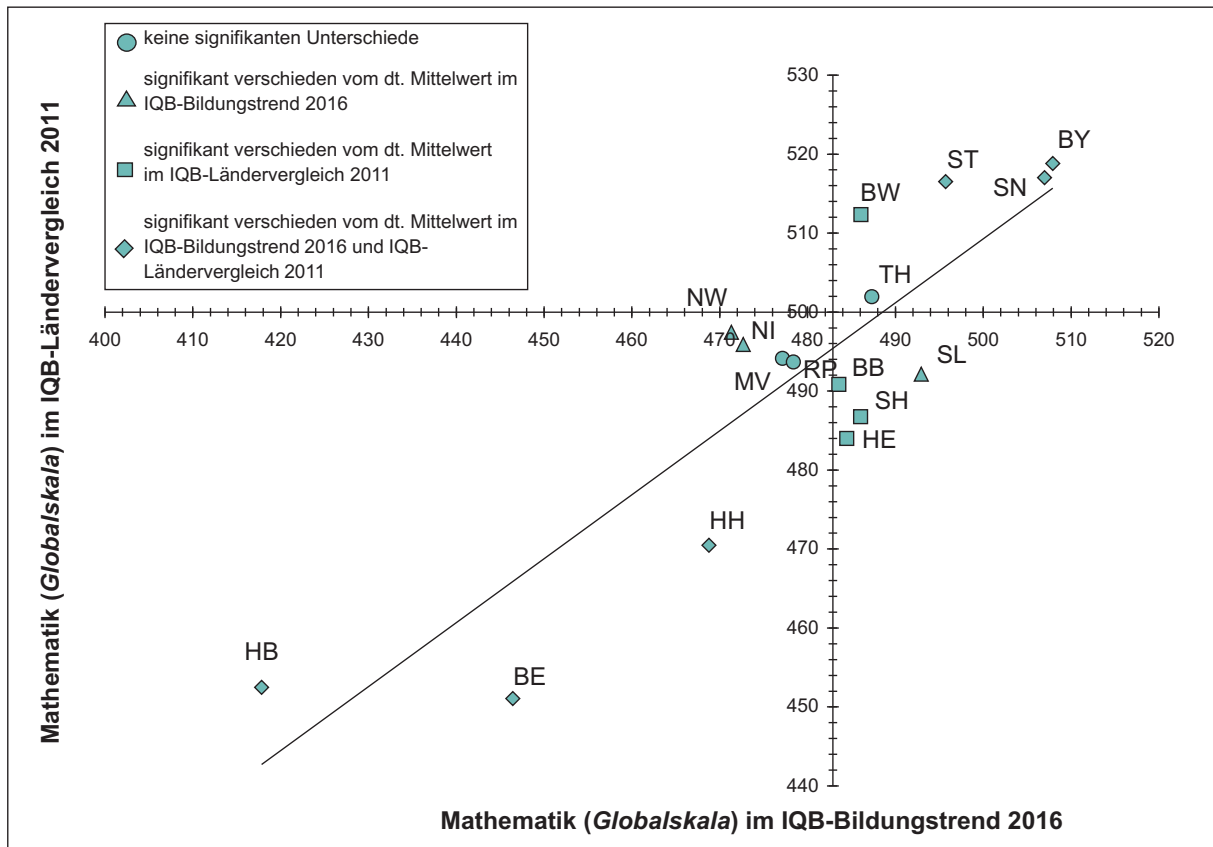


Trendanalysen für die Globalskala

Die Veränderungen in der globalen mathematischen Kompetenz sind in Abbildung 6.16 dargestellt. Für die in Deutschland im Mittel erreichte mathematische Kompetenz ergibt sich vom Jahr 2011 zum Jahr 2016 ein statistisch signifikanter Rückgang um 17 Punkte. Auch in den einzelnen Ländern haben sich die mittleren Kompetenzwerte entweder verringert oder nicht signifikant verändert; ein signifikant positiver Trend findet sich für kein Land. Im Jahr 2016 werden in den Ländern Baden-Württemberg, Bremen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Sachsen-Anhalt und Thüringen signifikant geringere Kompetenzwerte erzielt als im Jahr 2011. Im Vergleich zum bundesweiten Trend fällt diese Differenz für Hamburg, Hessen, das Saarland und Schleswig-Holstein signifikant geringer, für Bremen hingegen signifikant stärker aus.

Abbildung 6.17 stellt die in den Ländern erreichten mittleren Kompetenzen in den beiden Erhebungsjahren 2011 und 2016 in Relation zueinander dar. Der Wert auf der x-Achse entspricht dabei dem Kompetenzmittelwert für das Jahr 2016 und der Wert auf der y-Achse dem Kompetenzmittelwert für das Jahr 2011. Die beiden Koordinatenachsen schneiden sich im deutschen Mittelwert des jeweiligen Erhebungsjahres; der Schnittpunkt befindet sich also bei den Werten $M = 500$ für das Jahr 2011 und $M = 483$ für das Jahr 2016. Im Quadranten I des Koordinatensystems (oben rechts) sind Länder positioniert, in denen die mittleren Kompetenzen in beiden Erhebungsjahren über dem bundesweiten Mittelwert liegen. Im Quadranten III (unten links) befinden sich solche Länder, in denen in beiden Jahren Kompetenzmittelwerte unterhalb des deutschen Mittelwerts erreicht werden. Die statistische Signifikanz der Unterschiede zum Bundesdurchschnitt wird mit unterschiedlichen Symbolen markiert, die kennzeichnen, ob sich der

Abbildung 6.17: Ländermittelwerte im Fach Mathematik (*Globalskala*) im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016



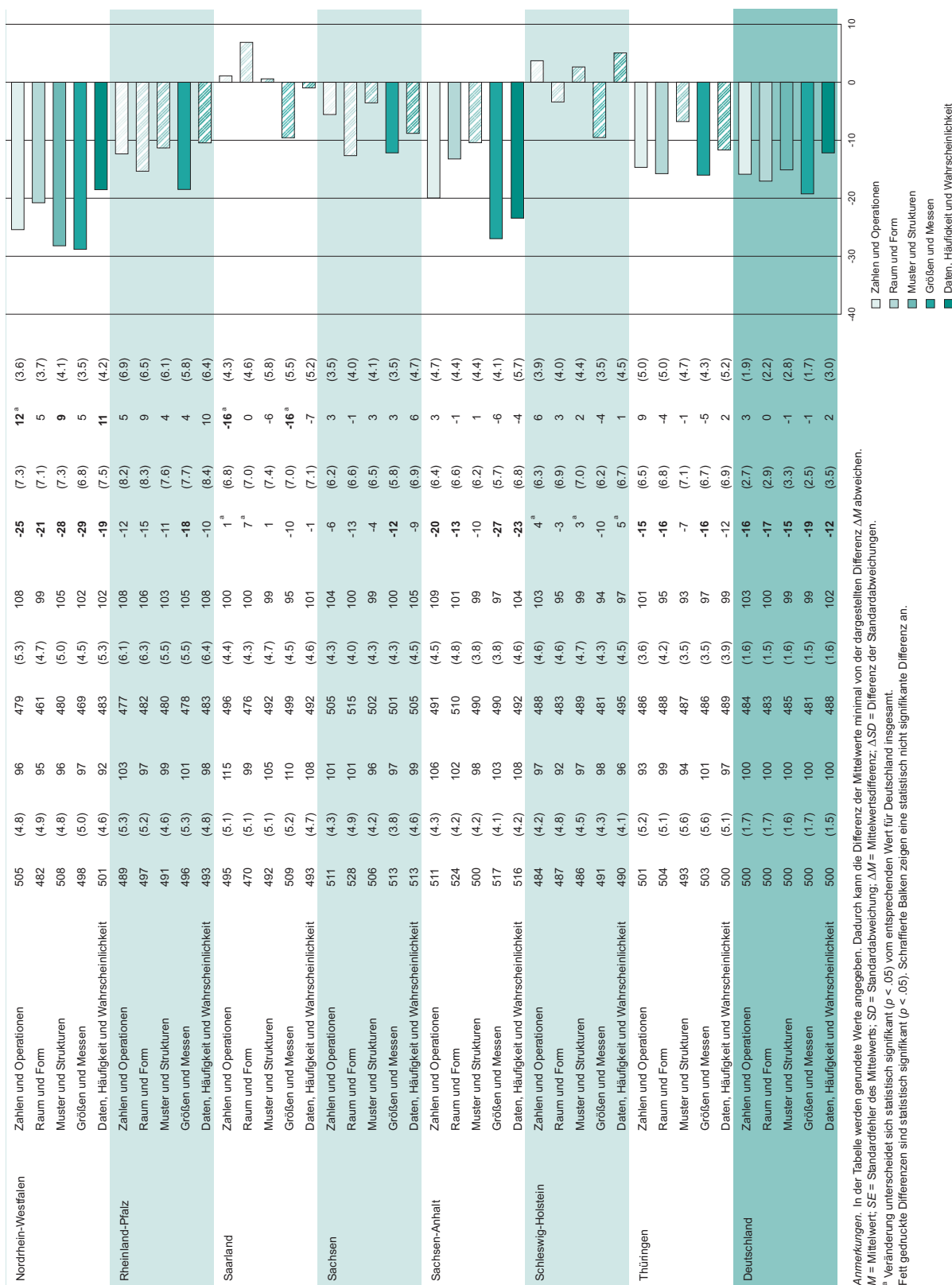
Anmerkungen. BB = Brandenburg; BE = Berlin; BW = Baden-Württemberg; BY = Bayern; HB = Bremen; HE = Hessen; HH = Hamburg; MV = Mecklenburg-Vorpommern; NI = Niedersachsen; NW = Nordrhein-Westfalen; RP = Rheinland-Pfalz; SH = Schleswig-Holstein; SL = Saarland; SN = Sachsen; ST = Sachsen-Anhalt; TH = Thüringen.

Kompetenzmittelwert des jeweiligen Landes in beiden Erhebungen (Raute), nur im Jahr 2011 (Quadrat), nur im Jahr 2016 (Dreieck) oder in keiner der beiden Erhebungen (Kreis) statistisch signifikant vom jeweiligen deutschen Mittelwert unterscheidet.

In den Ländern Bayern, Sachsen und Sachsen-Anhalt fallen die mathematischen Kompetenzen zu beiden Erhebungszeitpunkten signifikant höher, in den drei Stadtstaaten Berlin, Bremen und Hamburg hingegen signifikant geringer aus als der Bundesdurchschnitt. Eine besonders günstige Entwicklung ist im Saarland zu beobachten: Während dort im Jahr 2011 Kompetenzwerte erzielt werden, die sich nicht signifikant vom deutschen Mittelwert unterscheiden, gelingt es im Saarland im Jahr 2016, überdurchschnittliche Ergebnisse zu erreichen. Ähnliche Tendenzen zeigen sich für Brandenburg, Hessen und Schleswig-Holstein: In diesen Ländern werden im Jahr 2011 signifikant geringere mittlere Kompetenzen erreicht als im Bundesdurchschnitt, im Jahr 2016 unterscheiden sich die Kompetenzstände aber nicht mehr signifikant vom deutschen Mittelwert. Weniger günstige Entwicklungen lassen sich für Baden-Württemberg, Niedersachsen und Nordrhein-Westfalen feststellen: Im Jahr 2011 werden in Baden-Württemberg signifikant überdurchschnittliche mittlere Kompetenzwerte erreicht, im Jahr 2016 unterscheiden sich die erreichten Kompetenzwerte jedoch nicht mehr signifikant vom deutschen Gesamtmittelwert. Die in Niedersachsen und Nordrhein-Westfalen

Abbildung 6.18: Vergleich der in den Jahren 2011 und 2016 im Mittel erreichten Kompetenzen von Schülerinnen und Schülern der 4. Jahrgangsstufe in den einzelnen Kompetenzbereichen im Fach Mathematik





Anmerkungen: In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen.
M = Mittelwert, SE = Standardfehler des Mittelwerts, SD = Standardabweichung, ΔM = Mittelwertsdifferenz, ΔSD = Differenz der Standardabweichungen.
^a Veränderung unterscheidet sich statistisch signifikant (p < .05) vom entsprechenden Wert für Deutschland insgesamt.
Fett gedruckte Differenzen sind statistisch signifikant (p < .05). Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

erzielten Kompetenzmittelwerte weichen hingegen im Jahr 2011 nicht signifikant vom deutschen Mittelwert ab, im Jahr 2016 liegen sie jedoch signifikant unter dem Bundesdurchschnitt.

Trendanalysen für die einzelnen Leitideen

Für die einzelnen Leitideen zeigen sich ähnliche Trends wie für die Globalskala (Abb. 6.18). In allen Kompetenzbereichen sind für Deutschland insgesamt signifikante Rückgänge der im Mittel erreichten Kompetenzen zu verzeichnen, die zwischen 12 Punkten (*Daten, Häufigkeit und Wahrscheinlichkeit*) und 19 Punkten (*Größen und Messen*) variieren. Dieses Muster zeigt sich auch für die einzelnen Länder, wobei dort nicht alle Veränderungen signifikant sind. In Baden-Württemberg, Bremen, Niedersachsen und Nordrhein-Westfalen werden im Jahr 2016 für alle untersuchten Kompetenzbereiche signifikant geringere Werte ermittelt als im Jahr 2011. Signifikant negative Trends in vier Bereichen sind für Sachsen-Anhalt, in drei Bereichen für Thüringen festzustellen. In beiden Ländern betrifft dies die Bereiche *Zahlen und Operationen, Raum und Form* sowie *Größen und Messen*, in Sachsen-Anhalt zusätzlich den Bereich *Daten, Häufigkeit und Wahrscheinlichkeit*. In Bayern und Mecklenburg-Vorpommern ergeben sich im Jahr 2016 für die beiden Kompetenzbereiche *Raum und Form* sowie *Größen und Messen* signifikant geringere Werte als im Jahr 2011, in Rheinland-Pfalz und in Sachsen sind nur im Bereich *Größen und Messen* signifikante Rückgänge zu verzeichnen. Für die Länder Berlin, Brandenburg, Hamburg, Hessen, Saarland und Schleswig-Holstein zeigen sich hingegen keine signifikanten Veränderungen der Kompetenzmittelwerte für die einzelnen Leitideen. In diesen Ländern bleiben die mittleren Kompetenzen in den einzelnen Leitideen also stabil, allerdings auf unterschiedlichem Niveau.

Vergleicht man die Veränderungen der für die einzelnen Leitideen ermittelten Kompetenzwerte mit den bundesweiten Trends, so lässt sich feststellen, dass die Kompetenzmittelwerte in einigen Ländern signifikant stabiler sind als dies bundesweit der Fall ist. In Hamburg, Hessen und Schleswig-Holstein sind die Rückgänge für drei der fünf Leitideen geringer als in Deutschland insgesamt. In Hamburg betrifft dies die Leitideen *Raum und Form, Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit*; in Hessen und in Schleswig-Holstein die Leitideen *Zahlen und Operationen, Muster und Strukturen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit*. Stabilere Kompetenzmittelwerte für zwei Leitideen werden im Saarland festgestellt (*Zahlen und Operationen, Raum und Form*). In Berlin zeigen sich nur für den Kompetenzbereich *Größen und Messen* geringere Veränderungen der Kompetenzwerte als im bundesweiten Durchschnitt. Stärkere Rückgänge der Kompetenzmittelwerte und somit ungünstigere Entwicklungen als in Deutschland insgesamt sind hingegen für kein Land festzustellen. Eine Darstellung der Beziehungen zwischen den mittleren Kompetenzen, die bezogen auf die Leitideen in den beiden Erhebungsjahren 2011 und 2016 in den Ländern erreicht werden, kann auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

6.2.5 Vergleich der Heterogenität der in den Jahren 2011 und 2016 erreichten Kompetenzen in den Ländern

Nach den Veränderungen in den mittleren Kompetenzständen, die in den Ländern erreicht werden, wird im Folgenden berichtet, wie sich die Heterogenität der Kompetenzen zwischen den Jahren 2011 und 2016 verändert hat. Dafür werden die Streuungen der Werte auf der Globalskala mathematischer Kompetenz (Abb. 6.16) sowie auf den einzelnen Leitideen (Abb. 6.18) zwischen den beiden Erhebungszeitpunkten verglichen. Die für Deutschland insgesamt festgestellte Streuung hat sich im betrachteten Zeitraum mit einem Zuwachs von 5 Punkten zwar nur leicht, jedoch signifikant vergrößert (Abb. 6.16). Damit sind im Jahr 2016 die Kompetenzen der Schülerinnen und Schüler in Deutschland im Fach Mathematik etwas heterogener als im Jahr 2011. In den einzelnen Ländern zeigen sich im Vergleich der Jahre 2016 und 2011 entweder stabile oder steigende Streuungen; in keinem Land hat die Streuung abgenommen. Eine Vergrößerung der Heterogenität in den mathematischen Kompetenzen ist in Bremen, Hamburg, Nordrhein-Westfalen und Sachsen zu beobachten. Vergleicht man die Veränderungen in den Streuungen der Kompetenzwerte in den Ländern mit den Veränderungen für Deutschland insgesamt, so lässt sich feststellen, dass sich die Heterogenität in Bremen stärker vergrößert hat; in Bayern fällt die Veränderung hingegen signifikant schwächer aus als bundesweit.

Die für Deutschland insgesamt festgestellten Streuungen der Kompetenzwerte in den fünf untersuchten Leitideen zeigen, dass die Heterogenität der erreichten Kompetenzen im betrachteten Zeitraum stabil geblieben ist (Abb. 6.18). In den Ländern ist jedoch kein klares Befundmuster für die Veränderungen der Streuungen erkennbar. Während die Streuungen in den meisten Ländern vergleichsweise stabil blieben, sind – ähnlich wie für die Globalskala – in Bremen, Hamburg und Nordrhein-Westfalen signifikante Anstiege der Streuungen in mehreren Kompetenzbereichen zu verzeichnen. In Berlin und in Brandenburg erhöhte sich die Heterogenität jeweils in einem Kompetenzbereich. Im Gegensatz zur Globalskala zeigen sich in Bayern, Hessen, Mecklenburg-Vorpommern und im Saarland für einzelne Leitideen auch signifikante Rückgänge der Streuungen.

6.2.6 Vergleich der Heterogenität der in den Jahren 2011 und 2016 erreichten Kompetenzen innerhalb und zwischen Schulen

Im vorherigen Abschnitt wurde dargestellt, wie sich die Heterogenität der von Schülerinnen und Schülern im Fach Mathematik erreichten Kompetenzen zwischen den Jahren 2011 und 2016 verändert hat. Unterschiede in den erreichten Kompetenzen können dabei in beiden untersuchten Jahren 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, Stanat & Watermann, 2006). Der Varianzanteil zwischen Schulen verdeutlicht, wie stark sich die Schülerschaft verschiedener Schulen unterscheidet; der Varianzanteil innerhalb von Schulen gibt Hinweise darauf, mit welcher Heterogenität der Schülerkompetenzen Lehrkräfte umgehen müssen. Die Untersuchung der Frage, ob ein hoher Anteil der Heterogenität der erreichten Kompetenzen zwischen oder innerhalb von Schulen liegt, ist unter anderem für die Frage relevant, welche Strategien der Qualitätsentwicklung gewählt werden sollten (Willms, 2006).

In diesem Abschnitt wird daher deskriptiv untersucht, welcher Anteil der Varianz der im Jahr 2016 erreichten Kompetenzen zwischen Schulen liegt, welcher Anteil innerhalb von Schulen liegt und wie sich das Verteilungsmuster seit dem Jahr 2011 verändert hat. Zu diesem Zweck wurden für jedes Land Intra-klassenkorrelationen für die Jahre 2011 und 2016 berechnet (vgl. Kapitel 3.2).

Die Variation der erreichten Kompetenzen zwischen Schulen wird stark davon beeinflusst, ob Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf an Förderschulen oder an allgemeinen Schulen unterrichtet werden. Bezieht man Förderschulen in die Berechnung der Intra-klassenkorrelationen ein, fallen die Intra-klassenkorrelationen insbesondere in denjenigen Ländern deutlich größer aus, in denen ein hoher Anteil der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf an Förderschulen unterrichtet wird (vgl. Haag & Roppelt, 2012). Daher wurden nur allgemeine Schulen in die Berechnung der in diesem Kapitel berichteten Varianzanteile zwischen Schulen und innerhalb von Schulen einbezogen. Die Bemühungen der Länder um eine verstärkte Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf könnten allerdings dazu geführt haben, dass sich die Heterogenität innerhalb der allgemeinen Schulen zwischen den Jahren 2011 und 2016 vergrößert hat (vgl. Kapitel 4 und Kapitel 10.1). Die Veränderungen der Inklusionsquoten in den Ländern sind in Kapitel 4 dargestellt und sollten bei der Interpretation der Veränderungen in den Intra-klassenkorrelationen berücksichtigt werden.

Im Tabellenteil der Abbildung 6.19 sind die Standardabweichungen (*SD*) sowie die zwischen Schulen und innerhalb von Schulen liegenden Varianzanteile der von Schülerinnen und Schülern an allgemeinen Schulen erreichten globalen Mathematikkompetenzen für die Jahre 2011 und 2016 angegeben. Hierbei ist zu beachten, dass die Werte der Standardabweichungen in Abbildung 6.19 geringer ausfallen als die Werte der Standardabweichungen in Abbildung 6.16. Dies liegt daran, dass bei der Berechnung der in Abbildung 6.19 angegebenen Standardabweichungen keine Schülerinnen und Schüler an Förderschulen berücksichtigt wurden.

Der rechte Teil der Abbildung visualisiert die Gesamtvarianz der Kompetenzwerte an allgemeinen Schulen sowie ihre Zerlegung in einen Teil zwischen Schulen und einen Teil innerhalb von Schulen zu beiden Erhebungszeitpunkten. Die Gesamtlänge der Balken gibt dabei die Streuung der erreichten Kompetenzen im jeweiligen Land im Jahr 2011 beziehungsweise im Jahr 2016 an. Längere Balken deuten somit auf heterogenere Kompetenzen, kürzere Balken auf homogenere Kompetenzen innerhalb der Länder hin. Der nach rechts weisende, dunklere Teil des Balkens verdeutlicht den Anteil der Heterogenität, der zwischen den einzelnen Schulen innerhalb eines Landes besteht, wohingegen der nach links weisende, hellere Teil des Balkens den Anteil der Heterogenität kennzeichnet, der innerhalb von Schulen besteht.

Für Deutschland insgesamt liegt sowohl im Jahr 2011 als auch im Jahr 2016 etwas weniger als ein Fünftel der Kompetenzunterschiede zwischen Schulen (2011: 17.4 %, 2016: 17.6 %) und etwas mehr als vier Fünftel innerhalb von Schulen (2011: 82.6 %, 2016: 82.4 %). Dies deutet darauf hin, dass die von Schülerinnen und Schülern einer beliebigen Schule erreichten Kompetenzen eine beträchtliche Heterogenität aufweisen.

Die Varianzanteile zwischen Schulen unterscheiden sich auf Länderebene im Jahr 2016 deutlich. Besonders geringe zwischenschulische Kompetenzunterschiede bestehen mit etwa 8 Prozent in Bayern und Niedersachsen. Dies weist darauf hin, dass sich die Ergebnisse von Lehr-Lern-Prozessen in diesen bei-

Abbildung 6.19: Varianz der mathematischen Kompetenz (*Globalskala*) zwischen Schulen und innerhalb von Schulen im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 (ohne Förderschulen)

den Ländern vergleichsweise wenig zwischen Schulen unterscheiden. Mit knapp 30 Prozent sind die zwischenschulischen Kompetenzunterschiede in Berlin, Bremen, Hamburg und Rheinland-Pfalz hingegen mehr als dreimal so hoch.

Die innerhalb von Schulen bestehende Heterogenität der erreichten Kompetenzen wird im rechten Teil der Abbildung 6.19 durch die Länge der nach links weisenden, hellen Teile der Balken verdeutlicht. Vergleichsweise heterogene Kompetenzwerte innerhalb von allgemeinen Schulen finden sich zum einen in Ländern, in denen insgesamt hohe Standardabweichungen der Kompetenzverteilungen festgestellt wurden (Stadtstaaten Berlin, Bremen und Hamburg), zum anderen aber auch in Ländern, für die vergleichsweise geringe Unterschiede zwischen Schulen beobachtet werden (Brandenburg, Niedersachsen). In den Stadtstaaten Berlin, Bremen und Hamburg liegen sowohl hohe Streuungen als auch hohe Kompetenzunterschiede zwischen Schulen vor. In den Flächenländern zeigt sich hingegen kein besonders starker Zusammenhang zwischen den Standardabweichungen der mathematischen Kompetenz und den zwischenschulischen Kompetenzunterschieden.

Für die vier Länder, in denen überdurchschnittliche globale Mathematikkompetenzen erreicht wurden, zeigen sich eher gering bis durchschnittlich aus-

geprägte Kompetenzunterschiede zwischen Schulen. In Bayern beträgt dieser Unterschied nur 8 Prozent, was darauf hindeutet, dass Schülerinnen und Schüler an verschiedenen Schulen ähnliche Kompetenzwerte erreichen. Für die Gruppe der Länder, in denen unterdurchschnittliche Kompetenzwerte festgestellt wurden, bestehen in einigen Ländern deutlichere Unterschiede zwischen Schulen. So hängen in den Stadtstaaten Berlin, Bremen und Hamburg die erworbenen Kompetenzen stärker als im Bundesdurchschnitt mit der konkreten Schule zusammen, die von den Schülerinnen und Schülern besucht wird. In diesen Ländern sind etwa 30 Prozent der Kompetenzunterschiede zwischen Schülerinnen und Schülern auf Unterschiede zwischen Schulen zurückzuführen. In Niedersachsen sind die zwischenschulischen Unterschiede mit 8 Prozent dagegen deutlich geringer ausgeprägt als im Bundesdurchschnitt.

Vergleicht man die für die Jahre 2011 und 2016 festgestellten zwischenschulischen Kompetenzunterschiede, lässt sich in den meisten Ländern eine weitgehende Stabilität der Unterschiede zwischen Schulen erkennen. Ein substanzieller Anstieg um mehr als 5 Prozentpunkte findet sich in Hamburg und Rheinland-Pfalz, eine substanzielle Verringerung hingegen in Niedersachsen und Thüringen. Eine Beziehung zwischen den Veränderungen der im Mittel erreichten Kompetenzen und den Veränderungen der Unterschiede zwischen Schulen ist nicht zu erkennen.

In Bezug auf die Heterogenität innerhalb von Schulen lassen sich hingegen für einige Länder zwischen den Jahren 2011 und 2016 leicht steigende Werte feststellen. Die in Abschnitt 6.2.5 beschriebenen Veränderungen in der Streuung der erreichten Kompetenzen scheinen also tendenziell eher auf Veränderungen innerhalb von Schulen als zwischen Schulen zurückzuführen sein.

6.2.7 Zusammenfassung

Für das Jahr 2016 lässt sich festhalten, dass die Schülerinnen und Schüler in Bayern, im Saarland, in Sachsen sowie in Sachsen-Anhalt signifikant höhere globale Mathematikkompetenzen erzielen als in Deutschland insgesamt; in Bayern und Sachsen gilt dies auch für alle fünf untersuchten Kompetenzbereiche. Darüber hinaus ist die Streuung der mathematischen Kompetenzen in Bayern vergleichsweise gering. Hier wird das anzustrebende Ergebnismuster eines sowohl hohen als auch homogenen Kompetenzniveaus weitgehend erreicht. Überdurchschnittliche Kompetenzen in einzelnen Bereichen werden außerdem in Baden-Württemberg (*Größen und Messen*), in Brandenburg (*Raum und Form*), im Saarland (*Zahlen und Operationen* sowie *Größen und Messen*) und in Sachsen-Anhalt (*Raum und Form*) erreicht, wobei sich die für Baden-Württemberg und für Brandenburg festgestellten globalen Mathematikkompetenzen nicht signifikant vom deutschen Mittelwert unterscheiden.

Weniger günstige Ergebnisse finden sich in Niedersachsen, Nordrhein-Westfalen sowie in den drei Stadtstaaten Berlin, Bremen und Hamburg. In diesen fünf Ländern erreichen Schülerinnen und Schüler verglichen mit dem bundesdeutschen Durchschnitt im Mittel signifikant geringere Mathematikkompetenzen auf der Globalskala. Die in Bremen und in Berlin ermittelten Werte liegen auch für alle fünf Leitideen signifikant unterhalb des deutschen Mittelwerts. Im Vergleich zum bundesdeutschen Durchschnitt geringere Kompetenzwerte in vier Bereichen werden in Hamburg (*Zahlen und Operationen*, *Raum und Form*, *Muster und Strukturen* sowie *Größen und Messen*) und in Niedersachsen

(*Zahlen und Operationen, Raum und Form, Muster und Strukturen sowie Daten, Häufigkeit und Wahrscheinlichkeit*) erzielt. In Nordrhein-Westfalen weichen die erreichten Kompetenzen in zwei Bereichen nach unten vom deutschen Mittelwert ab (*Raum und Form* sowie *Größen und Messen*). Darüber hinaus werden in Mecklenburg-Vorpommern im Bereich *Muster und Strukturen* signifikant geringere mittlere Kompetenzwerte erreicht als im bundesdeutschen Durchschnitt, wobei sich die in Mecklenburg-Vorpommern erreichte mittlere globale Mathematikkompetenz jedoch nicht signifikant vom deutschen Mittelwert unterscheidet.

Die Streuungen der Kompetenzwerte sind in Bremen und Hamburg sowohl für die globale Mathematikkompetenz als auch für die einzelnen Leitideen höher als in Deutschland insgesamt; in Berlin ist dies für die Globalskala und für vier der fünf Leitideen der Fall. Damit ist die Heterogenität der erreichten Kompetenzen in den Stadtstaaten besonders ausgeprägt und stellt eine besondere Herausforderung dar. Im Gegensatz dazu werden in Niedersachsen sowohl für die Globalskala als auch für vier Leitideen im Vergleich mit Deutschland insgesamt besonders geringe Streuungen der erreichten Kompetenzen festgestellt.

Betrachtet man die Veränderungen der mathematischen Kompetenzen zwischen den Jahren 2011 und 2016, lässt sich in Deutschland insgesamt für die Globalskala mathematischer Kompetenz ein Rückgang von 17 Punkten beobachten. Dies entspricht in der Tendenz den Ergebnissen von TIMSS 2015 für die Primarstufe, wobei der in TIMSS festgestellte Rückgang von 6 Punkten zwischen den Jahren 2011 und 2015 nicht statistisch signifikant ist (Selter, Walter, Walther & Wendt, 2016).

Ein Rückgang der im Durchschnitt erreichten Kompetenzen ist im IQB-Bildungstrend für die meisten Länder zu verzeichnen, wenn auch nur in der Hälfte der Länder mit statistischer Signifikanz. In Baden-Württemberg, Bremen, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Sachsen-Anhalt und Thüringen werden im Jahr 2016 signifikant geringere mittlere Kompetenzen erreicht als im Jahr 2011. Dabei variiert die Höhe der negativen Trends zwischen 15 Punkten in Thüringen und 35 Punkten in Bremen. Positive Trends liegen hingegen für kein Land vor.

Die Streuung der globalen Mathematikkompetenz hat deutschlandweit zwischen den Jahren 2011 und 2016 leicht zugenommen (+5 Punkte). Auch in Bremen, Hamburg, Nordrhein-Westfalen und Sachsen stieg die Heterogenität der erreichten Kompetenzen im betrachteten Zeitraum signifikant an.

Für die einzelnen Leitideen sind ebenfalls statistisch signifikante negative Trends zu beobachten, die in Deutschland insgesamt zwischen 12 Punkten im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* und 19 Punkten im Bereich *Größen und Messen* variieren. Die Veränderungen der Kompetenzwerte erreichen jedoch nicht in allen Ländern statistische Signifikanz. Die für die fünf Leitideen festgestellten Streuungen der erreichten Kompetenzen sind deutschlandweit stabil. Für einzelne Leitideen sind in einigen Ländern Verschiebungen der Streuungen erkennbar, die jedoch kein einheitliches Ergebnismuster bilden.

Literatur

- 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. Vertiefende Analysen im Rahmen von PISA 2000* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Haag, N. & Roppelt, A. (2012). Der Ländervergleich im Fach Mathematik. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 117–127). 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.
- Selter, C., Walter, D., Walther, G. & Wendt, H. (2016). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79–136). 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 7

Geschlechtsbezogene Disparitäten

Stefan Schipolowski, Julia Wittig, Sebastian Weirich und Katrin Böhme

Eine wichtige Rolle in Diskussionen über Bildungsgerechtigkeit spielt die Frage, inwieweit Jungen und Mädchen die gleichen Bildungschancen haben und in der Schule mit vergleichbarem Erfolg gefördert werden. Die bisher vorliegenden Forschungsergebnisse weisen darauf hin, dass bildungsrelevante Unterschiede zwischen Jungen und Mädchen bereits beim Eintritt in das Bildungssystem bestehen (Blossfeld et al., 2009; Stanat & Bergann, 2009). Diese Unterschiede werden im schulischen Kontext wirksam, formen sich weiter aus und führen zu geschlechtsbezogenen Disparitäten im Kompetenzerwerb und in der Bildungsbeteiligung. Gleichzeitig gibt es jedoch auch klare Hinweise darauf, dass sich Unterschiede zwischen Jungen und Mädchen in der Bildungsbeteiligung und im Kompetenzerwerb durch gezielte Förderung reduzieren lassen (vgl. z. B. Blossfeld et al., 2009).

Um auf der Grundlage aktueller Daten für den Primarbereich festzustellen, inwieweit geschlechtsbezogene Disparitäten im Bildungserfolg bestehen, werden im vorliegenden Kapitel Kompetenzunterschiede zwischen Jungen und Mädchen in den Fächern Deutsch und Mathematik in der 4. Jahrgangsstufe untersucht. Zunächst fassen wir in Abschnitt 7.1 den Forschungsstand zu geschlechtsbezogenen Disparitäten in sprachlichen und mathematischen Kompetenzen zusammen. Zudem wird die empirische Befundlage zu Unterschieden zwischen Jungen und Mädchen in ausgewählten motivationalen Merkmalen beschrieben, indem auf das fachbezogene Selbstkonzept und das fachliche Interesse eingegangen wird. Anschließend stellen wir in Abschnitt 7.2 zunächst die Geschlechterdisparitäten für die im Jahr 2016 in den Fächern Deutsch und Mathematik untersuchten Kompetenzbereiche für Deutschland insgesamt dar. Neben Unterschieden in den Kompetenzmittelwerten gehen wir hierbei auch auf die Kompetenzverteilungen ein und beschreiben, welche Anteile der Jungen und Mädchen in den einzelnen Kompetenzbereichen jeweils die Mindeststandards verfehlen, die Regelstandards erreichen oder übertreffen beziehungsweise die Optimalstandards erreichen. Ferner gehen wir auf die länderspezifischen Befunde zu Kompetenzunterschieden zwischen Jungen und Mädchen ein. In Abschnitt 7.3 wird anhand eines Vergleichs der Befunde des IQB-Bildungstrends 2016 mit den Ergebnissen des IQB-Ländervergleichs 2011 überprüft, inwieweit sich die geschlechtsbezogenen Disparitäten in der Primarstufe in den letzten fünf Jahren verändert haben. Im nachfolgenden Abschnitt 7.4 präsentieren wir zudem Befunde zu geschlechtsbezogenen Unterschieden im Selbstkonzept und im fachlichen Interesse und gehen der Frage nach, wie diese mit den erreichten Kompetenzen zusammenhängen. Das Kapitel schließt in Abschnitt 7.5 mit einer Zusammenfassung der zentralen Ergebnisse.

7.1 Geschlechtsbezogene Unterschiede in schulischen Kompetenzen und motivationalen Merkmalen

7.1.1 Geschlechtsbezogene Unterschiede in schulischen Kompetenzen

Sowohl in großen internationalen Schulleistungsstudien wie IGLU/PIRLS¹, TIMSS² und PISA³ als auch in den Studien des IQB zum nationalen Bildungsmonitoring sowie in Metaanalysen und zahlreichen kleineren empirischen Untersuchungen (vgl. etwa Hyde, 1981, 2005; Wieczerkowski & Prado, 1992) wurde vielfach belegt, dass geschlechtsbezogene Unterschiede in den von Jungen und Mädchen erreichten Kompetenzen bestehen. Nachfolgend werden zentrale empirische Befunde der internationalen und nationalen Monitoringstudien zusammenfassend dargestellt, wobei der Schwerpunkt auf den in Tabelle 7.1 dargestellten Ergebnissen der Studien des IQB zum Erreichen der Bildungsstandards liegt. Um die in den verschiedenen Studien ermittelten Gruppenunterschiede trotz unterschiedlicher Metriken und Streuungen der Kompetenzwerte direkt miteinander vergleichen zu können, wird das Effektstärkemaß Cohens d herangezogen (vgl. die Erläuterungen hierzu in Kapitel 3.2).⁴

Geschlechtsbezogene Unterschiede in sprachlichen Kompetenzen

Bereits im Elementarbereich zeigen Mädchen einen leichten Vorsprung im Erwerb sprachlicher Kompetenzen und sind seltener von Störungen der Sprachentwicklung wie Stottern oder besonderen Schwierigkeiten im Lesen und Rechtschreiben betroffen als Jungen (Halpern, 2000). Zu Beginn der Grundschulphase sind geschlechtsbezogene Unterschiede in sprachlichen Kompetenzen zwischen Jungen und Mädchen oft noch gering ausgeprägt (Mengerling, 2005), sie vergrößern sich jedoch bis zum Ende der Grundschulzeit und im Verlauf der Sekundarstufe (Stanat & Bergann, 2009).

Für den Kompetenzbereich *Lesen* im Fach Deutsch hat sich sowohl im Rahmen der IQB-Ländervergleichsstudien und im IQB-Bildungstrend 2015 als auch in den großen internationalen Schulleistungsstudien IGLU/PIRLS und PISA wiederholt gezeigt, dass Mädchen im Mittel über eine höhere Lesekompetenz verfügen als Jungen. Im Primarbereich wurde im Rahmen des IQB-Ländervergleichs 2011 in der 4. Jahrgangsstufe ein signifikanter Kompetenzvorteil der Mädchen gegenüber den Jungen in Höhe von $d = 0.24$ identifiziert (Böhme & Roppelt, 2012). Die internationale PIRLS/IGLU-Studie, die regelmäßig am Ende der Grundschulzeit durchgeführt wird und ebenfalls die Lesekompetenz der Kinder untersucht, ermittelte für die Gesamtskala Lesen allerdings etwas geringere Vorteile der Mädchen. Diese lagen in den verschiedenen Erhebungsjahren bei $d = 0.13$ (IGLU 2001), $d = 0.07$ (IGLU 2006) und $d = 0.08$ (IGLU 2011) (Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Betrachtet man die Leseleistungen

-
- 1 Das Akronym PIRLS steht für *Progress in International Reading Literacy Study*; im deutschen Sprachraum wird diese Studie als *Internationale Grundschul-Lese-Untersuchung* (IGLU) bezeichnet.
 - 2 Das Akronym TIMSS steht für *Trends in International Mathematics and Science Study*.
 - 3 Das Akronym PISA steht für *Programme for International Student Assessment*.
 - 4 Auf mögliche Ursachen geschlechtsbezogener Disparitäten im Bildungserfolg wurde bereits in den Berichtsbänden zum IQB-Ländervergleich 2011 (Böhme & Roppelt, 2012) und zum IQB-Bildungstrend 2015 (Böhme, Sebald, Weirich & Stanat, 2016) eingegangen. Diese Ausführungen sollen an dieser Stelle nicht wiederholt werden.

Tabelle 7.1: Kompetenzunterschiede zwischen Jungen und Mädchen in den IQB-Ländervergleichsstudien 2009, 2011 und 2012 sowie im IQB-Bildungstrend 2015

Deutsch	Primarbereich 2011				Sekundarstufe I 2009				Sekundarstufe I 2015			
	M_J	M_M	$M_J - M_M$	d	M_J	M_M	$M_J - M_M$	d	M_J	M_M	$M_J - M_M$	d
Lesen	488	512	-24	-0.24	488	505	-17	-0.19	487	514	-27	-0.27
Zuhören	499	501	-3	-0.03	491	510	-19	-0.18	488	513	-25	-0.25
Orthografie	484	517	-32	-0.33	475	526	-51	-0.53	478	524	-46	-0.47

Mathematik	Primarbereich 2011				Sekundarstufe I 2012				
	M_J	M_M	$M_J - M_M$	d	M_J	M_M	$M_J - M_M$	d	
Globalskala	508	492	16	0.16	Globalskala	508	492	16	0.16
DHW	503	497	7	0.07	Daten und Zufall	490	510	20	0.20
Größen und Messen	516	483	33	0.34	Messen	491	508	17	0.17
Muster und Strukturen	511	489	22	0.22	Funktionaler Zusammenhang	493	506	13	0.13
Raum und Form	503	497	6	0.06	Raum und Form	497	503	5	0.05
Zahlen und Operationen	507	493	14	0.14	Zahl	490	510	20	0.20

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte von der dargestellten Differenz $M_J - M_M$ abweichen. M_J = Mittelwert in der Gruppe der Jungen; M_M = Mittelwert in der Gruppe der Mädchen; $M_J - M_M$ = Differenz der Mittelwerte von Jungen und Mädchen; d = Effektstärke Cohens d ; DHW = Daten, Häufigkeit und Wahrscheinlichkeit. Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

differenziert für verschiedene Textsorten, so ist der Kompetenzvorsprung der Mädchen für das Lesen und Verstehen literarischer Texte besonders ausgeprägt (Valtin, Bos, Buddeberg, Goy & Potthoff, 2008).

Für Schülerinnen und Schüler der 9. Jahrgangsstufe identifizierten Winkelmann und Groeneveld (2010) im Rahmen des IQB-Ländervergleichs 2009 einen Kompetenzvorsprung der Mädchen im *Lesen* von $d = 0.19$. Für die Neuntklässlerinnen und Neuntklässler, deren Kompetenzen im IQB-Bildungstrend 2015 getestet wurden, fiel die Mittelwertdifferenz mit $d = 0.27$ zwar etwas größer aus, der Unterschied zum Jahr 2009 war jedoch statistisch nicht signifikant (Böhme et al., 2016). In den internationalen PISA-Studien der Jahre 2000 bis 2012 wurde für die Sekundarstufe I ein Geschlechterunterschied in der Lesekompetenz zugunsten der Mädchen von jeweils etwa $d = 0.40$ identifiziert (Hohn, Schiepe-Tiska, Sälzer & Artelt, 2013; Naumann, Artelt, Schneider & Stanat, 2010).⁵

Für den Kompetenzbereich *Zuhören* im Fach Deutsch stellten Böhme und Roppelt (2012) im IQB-Ländervergleich 2011 im Primarbereich keinen statistisch bedeutsamen Kompetenzvorsprung der Mädchen fest. Am Ende der Sekundarstufe I wurde im IQB-Ländervergleich 2009 allerdings ein Kompetenzvorsprung der Mädchen in Höhe von $d = 0.18$ ermittelt (Winkelmann & Groeneveld, 2010). Im IQB-Bildungstrend 2015 lag der Unterschied bei $d = 0.25$, wobei die Differenz zum Jahr 2009 wiederum statistisch nicht signifikant war (Böhme et al., 2016).

5 In der PISA-Studie des Jahres 2015 war der Unterschied zwischen Jungen und Mädchen in der Lesekompetenz nur etwa halb so groß wie in früheren PISA-Studien. Dieser Befund dürfte jedoch auf die erstmals am Computer durchgeführte Testung zurückzuführen sein (Weis et al., 2016).

Für Kompetenzen in der *Orthografie* berichteten Böhme und Roppelt (2012) auf der Grundlage von Daten des IQB-Ländervergleichs 2011 einen Kompetenzvorsprung der Mädchen von $d = 0.33$. Auch in IGLU-E 2006 wurde für den Bereich der Rechtschreibung ein Vorsprung der Mädchen identifiziert ($d = 0.20$; Kowalski, Voss, Valtin & Bos, 2010). Für Jugendliche in der 9. Jahrgangsstufe lag der Kompetenzvorsprung der Mädchen in der *Orthografie* bei $d = 0.53$ im IQB-Ländervergleich 2009 (Winkelmann & Groeneveld, 2010) und bei $d = 0.47$ im IQB-Bildungstrend 2015 (Böhme et al., 2016). Die geschlechtsbezogenen Kompetenzunterschiede sind also in der *Orthografie* sowohl in der Primarstufe als auch in der Sekundarstufe I etwas stärker ausgeprägt als in den rezeptiven Kompetenzbereichen *Lesen* und *Zuhören*.

Geschlechtsbezogene Unterschiede in mathematischen Kompetenzen

Auch für Kompetenzen im Fach Mathematik sind geschlechtsbezogene Unterschiede zu verzeichnen. Metaanalysen zur mathematischen Kompetenz von Jungen und Mädchen sowie nationale und internationale Schulleistungsstudien belegen, dass bereits am Ende der Grundschulzeit Unterschiede zugunsten der Jungen bestehen und 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).

Böhme und Roppelt (2012) berichteten im IQB-Ländervergleich 2011 für den Primarbereich einen Kompetenzvorsprung der Jungen gegenüber den Mädchen auf der Globalskala im Fach Mathematik in Höhe von $d = 0.16$. Allerdings fiel der Vorsprung der Jungen für die einzelnen mathematischen Inhaltsbereiche unterschiedlich groß aus. So lag der größte Kompetenzunterschied zwischen Jungen und Mädchen im Kompetenzbereich *Größen und Messen* vor ($d = 0.34$), während die Unterschiede in den Kompetenzbereichen *Daten, Häufigkeit und Wahrscheinlichkeit* sowie *Raum und Form* deutlich geringer waren ($d = 0.07$ bzw. $d = 0.06$). Im Rahmen der internationalen Schulleistungsstudie TIMSS, die ebenfalls in der 4. Jahrgangsstufe durchgeführt wurde, zeigte sich in den Erhebungen der Jahre 2007, 2011 und 2015 für Deutschland ebenfalls ein Kompetenzvorsprung zugunsten der Jungen, der jeweils statistisch signifikant war, mit rund $d = 0.10$ jedoch vergleichsweise klein ausfiel und über die Zeit signifikant abnahm (Wendt, Steinmayr & Kasper, 2016).

Für die Sekundarstufe I ermittelten Schroeders, Penk, Jansen und Pant (2013) im IQB-Ländervergleich 2012 sowohl für die globale mathematische Kompetenz als auch für alle mathematischen Leitideen, dass Jungen im Durchschnitt signifikant höhere Werte erreichen als Mädchen. Der Unterschied auf der Globalskala war mit $d = 0.16$ genauso groß wie die von Böhme und Roppelt berichtete Geschlechterdifferenz in der Primarstufe. Auch in den internationalen PISA-Studien der Jahre 2009 bis 2015 wurden in Deutschland für die Gesamtskala mathematischer Kompetenz für Jungen höhere Leistungen ermittelt als für Mädchen, wobei die Effektstärke ähnlich ausfiel wie in den Studien des IQB (Frey, Heinze, Mildner, Hochweber & Asseburg, 2010; Hammer et al., 2016; Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013).

7.1.2 Geschlechtsbezogene Unterschiede in motivationalen Merkmalen

Unterschiede zwischen Jungen und Mädchen bestehen auch in motivationalen Merkmalen wie dem fachbezogenen Selbstkonzept und dem fachlichen Interesse, die ihrerseits miteinander sowie mit schulischen Kompetenzen und bildungsbezogenen Wahlentscheidungen zusammenhängen (vgl. etwa Jansen, Schroeders & Stanat, 2013; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005; Murphy & Whitelegg, 2006; Schiefele, Krapp & Schreyer, 1993).

Beim fachbezogenen schulischen Selbstkonzept handelt es sich um die Einschätzung der eigenen Kompetenzen in einem bestimmten Schulfach oder Kompetenzbereich. Solche Fähigkeitseinschätzungen basieren auf den Kompetenzerfahrungen, die eine Schülerin oder ein Schüler in einem bestimmten Schulfach oder Bereich macht. Das Selbstkonzept ist subjektiv geprägt und spiegelt nur bedingt die objektive Kompetenzausprägung wider, da die Einschätzung der eigenen Fähigkeiten in einem bestimmten Bereich unter anderem auch durch Vergleiche mit Mitschülerinnen und Mitschülern sowie den eigenen Leistungen in anderen Fächern beeinflusst wird (Möller & Köller, 2004).

In großen Schulleistungsstudien wurde gezeigt, dass im Fach Deutsch Mädchen ihre eigene Kompetenz im Durchschnitt höher einschätzen als Jungen, während letztere im Fach Mathematik im Mittel eine positivere Einschätzung der eigenen Fähigkeiten abgeben. So berichteten Böhme et al. (2016) auf Basis der Daten zum IQB-Bildungstrend 2015 in der Sekundarstufe I für das Selbstkonzept im Fach Deutsch einen Unterschied zugunsten der Mädchen in Höhe von $d = 0.31$. Die positivere Selbsteinschätzung der Mädchen im Fach Deutsch spiegelte somit in etwa die Höhe des oben beschriebenen Kompetenzvorsprungs der Mädchen in den Kompetenzbereichen *Lesen* und *Zuhören* in dieser Studie wider. Für das Fach Mathematik konnte hingegen gezeigt werden, dass Jungen ihre eigenen Kompetenzen im Mittel positiver einschätzen als Mädchen ($d = 0.36$ im Primarbereich beziehungsweise $d = 0.54$ in der Sekundarstufe I), wobei der Unterschied in der Selbsteinschätzung größer ausfällt als in den erreichten Kompetenzen (Jansen et al., 2013; Wendt et al., 2016).

Das schulfachbezogene Interesse äußert sich durch anhaltende Wertschätzung und positive Emotionen gegenüber den Inhalten eines bestimmten Fachs und durch die Beschäftigung mit diesen Inhalten (vgl. Schiefele, 2009). Für Viertklässlerinnen und Viertklässler in Deutschland konnten Bos et al. (2012) im Rahmen der IGLU-Studie 2011 zeigen, dass Jungen signifikant seltener als Mädchen außerhalb der Schule zum Vergnügen lesen. Dieser Unterschied im Leseverhalten deutet darauf hin, dass Mädchen in der 4. Jahrgangsstufe ein höheres Leseinteresse aufweisen als Jungen.⁶ Auch Ergebnisse des IQB-Bildungstrends 2015 in der 9. Jahrgangsstufe zeigen, dass Unterschiede zwischen Jungen und Mädchen im Interesse für das Fach Deutsch bestehen, wobei Mädchen ein größeres Interesse angaben ($d = 0.36$; Böhme et al., 2016). Für das Fach Mathematik ergab die TIMS-Studie 2015, dass Jungen im Primarbereich im Durchschnitt ein größeres fachliches Interesse⁷ äußern als Mädchen ($d = 0.20$; Wendt et al., 2016). Im IQB-Ländervergleich 2012 wurde in der Sekundarstufe I

6 Für das Leseinteresse werden von Bos et al. (2012) keine Geschlechterunterschiede berichtet.

7 Die Autorinnen und Autoren sprechen von „mathematikbezogenen Einstellungen“ (Wendt et al., 2016, S. 285). Eine genauere Betrachtung der verwendeten Items legt jedoch nahe, dass diese weitgehend das fachliche Interesse erfassen dürften.

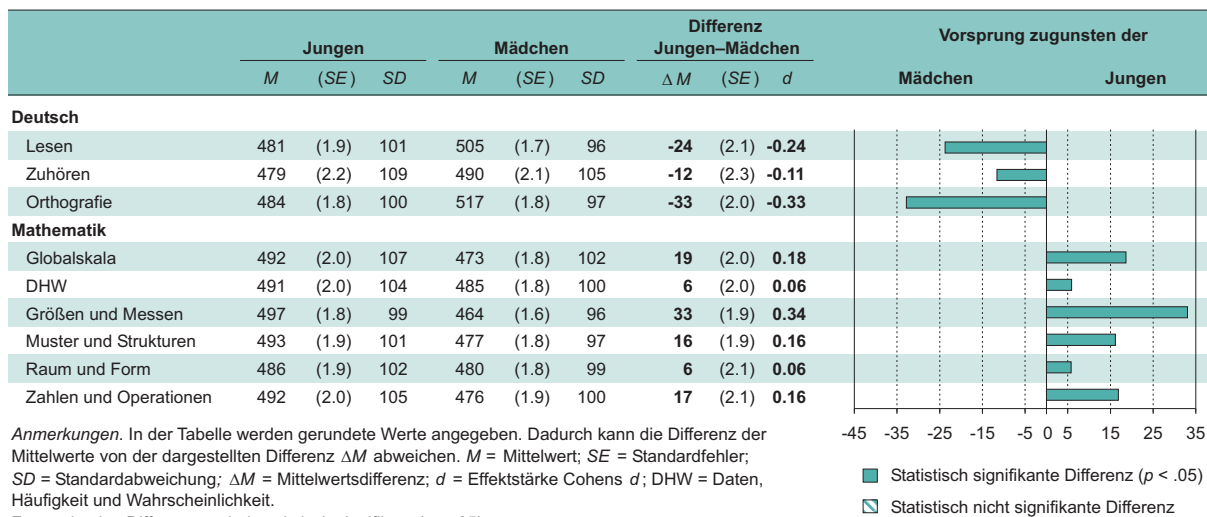
für das Interesse am Fach Mathematik ein noch größerer Effekt zugunsten der Jungen ermittelt ($d = 0.48$; Jansen et al., 2013).

7.2 Geschlechtsbezogene Kompetenzunterschiede im Jahr 2016

7.2.1 Geschlechtsbezogene Unterschiede in den Kompetenzmittelwerten

In diesem Abschnitt wird zunächst für das Jahr 2016 dargestellt, inwieweit sich die von Jungen und Mädchen erreichten Kompetenzmittelwerte in den Fächern Deutsch und Mathematik in Deutschland insgesamt unterscheiden. In Abbildung 7.1 sind die Mittelwerte (M) für Jungen und Mädchen auf der Berichtsmetrik, deren Standardfehler (SE) sowie die Streuung (SD) der von Jungen und Mädchen erreichten Kompetenzen angegeben. Zusätzlich dargestellt ist die Differenz der Kompetenzmittelwerte (ΔM) und der zugehörige Standardfehler, wobei eine positive Differenz einem Kompetenzvorsprung der Jungen entspricht. Um einen Vergleich mit den in Abschnitt 7.1.1 zusammengefassten Ergebnissen anderer Studien zu erleichtern, ist zudem die Effektstärke d angegeben.⁸ Statistisch signifikante Mittelwertdifferenzen sind durch Fettdruck gekennzeichnet. Zudem werden die Unterschiede in den Kompetenzmittelwerten als Balkendiagramm veranschaulicht, wobei nach rechts weisende Balken einen Kompetenzvorsprung der Jungen anzeigen. Statistisch signifikante Unterschiede sind durch ausgefüllte Balken markiert.

Abbildung 7.1: Kompetenzunterschiede zwischen Jungen und Mädchen in den Fächern Deutsch und Mathematik für Deutschland insgesamt



8 Zur Einordnung der im Folgenden beschriebenen Unterschiede in den Kompetenzmittelwerten von Jungen und Mädchen können auch die in Kapitel 6 angegebenen erwarteten Lernzuwächse innerhalb eines Schuljahres herangezogen werden. Diese betragen im Fach Deutsch im *Lesen* sowie im *Zuhören* jeweils etwa 60 Punkte und in der *Orthografie* rund 100 Punkte auf der Berichtsmetrik. Für das Fach Mathematik (*Globalskala*) kann ein Lernzuwachs von 80 Punkten innerhalb eines Schuljahres angenommen werden.

Übereinstimmend mit den oben dargestellten Befunden früherer Schulleistungsstudien zeigt sich für Deutschland insgesamt, dass Mädchen im Fach Deutsch und Jungen im Fach Mathematik signifikant höhere Kompetenzmittelwerte erzielen. Die Größe der geschlechtsbezogenen Disparitäten unterscheidet sich jedoch für die einzelnen im IQB-Bildungstrend 2016 untersuchten Kompetenzbereiche. So liegt im *Lesen* ein signifikanter Kompetenzvorsprung der Mädchen von 24 Punkten auf der Berichtsmetrik vor, während im *Zuhören* mit einer Differenz von 12 Punkten ein zwar ebenfalls signifikanter, jedoch geringerer Vorsprung der Mädchen zu verzeichnen ist. Die größten Kompetenzunterschiede zwischen Jungen und Mädchen im Fach Deutsch ergeben sich für die *Orthografie*; hier beträgt der Kompetenzvorteil der Mädchen im Mittel 33 Punkte.

Im Fach Mathematik ergibt sich ein mittlerer Kompetenzunterschied von 19 Punkten auf der Berichtsmetrik zugunsten der Jungen. Die Differenz zwischen Jungen und Mädchen variiert jedoch zwischen den inhaltsbezogenen Kompetenzbereichen. Den größten Kompetenzvorsprung erzielen Jungen bei Aufgaben zur mathematischen Leitidee *Größen und Messen* ($\Delta M = 33$ Punkte); die geringsten, jedoch ebenfalls signifikanten Kompetenzunterschiede liegen für die Leitideen *Daten, Häufigkeit und Wahrscheinlichkeit* sowie *Raum und Form* vor (jeweils $\Delta M = 6$ Punkte zugunsten der Jungen).

7.2.2 Geschlechtsbezogene Unterschiede in den Kompetenzverteilungen

Neben den durchschnittlichen Kompetenzunterschieden zwischen Jungen und Mädchen ist für eine gezielte Förderung auch relevant, ob sich die Kompetenzverteilungen von Jungen und Mädchen unterscheiden und in welchen Bereichen des Kompetenzspektrums die Unterschiede besonders groß sind.

Die in Abbildung 7.1 dargestellten Ergebnisse zeigen, dass sich die Kompetenzen von Jungen und Mädchen im Jahr 2016 nicht nur im Mittel unterscheiden, sondern auch in der jeweiligen Streuung (vgl. die Spalten *SD*). Sowohl für das Fach Deutsch als auch für das Fach Mathematik gilt, dass in allen untersuchten Bereichen die Kompetenzen der Jungen eine etwas höhere Streuung aufweisen als die der Mädchen. Trotz der relativ geringen absoluten Punktwertunterschiede sind die Differenzen in den Standardabweichungen für den Kompetenzbereich *Lesen* im Fach Deutsch sowie für alle im Fach Mathematik untersuchten Kompetenzbereiche mit Ausnahme von *Größen und Messen* statistisch signifikant.

Wie in Kapitel 1.1 ausgeführt, liegt der Fokus des IQB-Bildungstrends 2016 auf der kriterialen Vergleichsperspektive, also auf der Frage, inwieweit die Bildungsstandards erreicht werden. Der Vergleich der Kompetenzverteilungen von Jungen und Mädchen erfolgt daher im Folgenden mit Bezug auf die Kompetenzstufenmodelle des IQB (vgl. Kapitel 2). Entsprechend sind in Tabelle 7.2 die Anteile der Jungen (J) und Mädchen (M) ausgewiesen, die in den einzelnen Kompetenzbereichen im Jahr 2016 den jeweiligen Mindeststandard verfehlen, den Regelstandard erreichen oder übertreffen beziehungsweise den Optimalstandard erreichen. Zusätzlich wird die Differenz zwischen dem jeweiligen Anteil der Jungen und der Mädchen angegeben (J–M) sowie der zugehörige Standardfehler (*SE*).

Die in Tabelle 7.2 aufgeführten Ergebnisse zeigen, dass – bis auf wenige Ausnahmen bei einzelnen mathematischen Leitideen – auch für die hier berichteten Prozentsätze signifikante geschlechtsbezogene Disparitäten vorliegen. Im Fach

Tabelle 7.2: Anteile der Jungen und der Mädchen, die in den untersuchten Kompetenzbereichen den jeweiligen Mindeststandard verfehlen, mindestens den Regelstandard erreichen bzw. den Optimalstandard erreichen (Angaben in %)

	Mindeststandard nicht erreicht				Regelstandard erreicht oder übertroffen				Optimalstandard erreicht			
	J	M	J-M	(SE)	J	M	J-M	(SE)	J	M	J-M	(SE)
Deutsch												
Lesen	15	10	5	(0.8)	61	70	-9	(1.1)	8	12	-4	(0.7)
Zuhören	12	10	2	(0.8)	67	70	-3	(1.1)	9	11	-2	(0.6)
Orthografie	27	17	10	(0.9)	48	60	-13	(1.1)	6	11	-5	(0.7)
Mathematik												
Globalskala	13	17	-4	(0.7)	66	58	8	(1.0)	16	10	5	(0.8)
Daten, Häufigkeit und Wahrscheinlichkeit	13	15	-1	(0.8)	66	63	3	(1.0)	14	12	2	(0.8)
Größen und Messen	11	19	-8	(0.8)	69	55	15	(1.0)	15	7	8	(0.8)
Muster und Strukturen	13	16	-3	(0.7)	67	60	7	(1.0)	15	10	5	(0.8)
Raum und Form	15	16	-1	(0.9)	63	60	3	(1.0)	14	11	2	(0.8)
Zahlen und Operationen	14	17	-3	(0.8)	66	59	7	(1.0)	16	11	5	(0.8)

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Anteile von Jungen und Mädchen minimal von der dargestellten Differenz J-M abweichen. J = Jungen; M = Mädchen; J-M = Differenz der Anteile von Jungen und Mädchen; SE = Standardfehler.

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

Deutsch ist in allen untersuchten Kompetenzbereichen der Anteil der Jungen, die die Mindeststandards verfehlen, signifikant größer als der entsprechende Anteil der Mädchen. Im Kompetenzbereich *Orthografie* liegt mit 10 Prozentpunkten die größte Differenz vor; hier verfehlen 27 Prozent der Jungen, aber nur 17 Prozent der Mädchen den Mindeststandard. In den Bereichen *Lesen* (5 Prozentpunkte) und *Zuhören* (2 Prozentpunkte) fallen die geschlechtsbezogenen Disparitäten in den Anteilen von Jungen und Mädchen, die den Mindeststandard nicht erreichen, geringer aus. Demgegenüber sind in allen drei Kompetenzbereichen im Fach Deutsch die Anteile der Jungen, die mindestens den Regelstandard oder sogar den Optimalstandard erreichen, signifikant geringer als die entsprechenden Anteile der Mädchen. Auch hier finden sich die größten geschlechtsbezogenen Disparitäten im Kompetenzbereich *Orthografie*: Für das Erreichen der Regelstandards beträgt der Vorteil der Mädchen 13 Prozentpunkte, für das Erreichen der Optimalstandards 5 Prozentpunkte. Im *Lesen* sowie im *Zuhören* fallen die Differenzen der entsprechenden Anteile im mittleren und oberen Leistungsbereich wiederum geringer aus als in der *Orthografie*.

Ein anderes Befundmuster als im Fach Deutsch ergibt sich für das Fach Mathematik. So ist auf der Globalskala der Anteil der Jungen, die den Mindeststandard verfehlen, signifikant geringer als der entsprechende Anteil der Mädchen (Differenz von 4 Prozentpunkten), während die Anteile der Jungen, die mindestens den Regelstandard erreichen beziehungsweise den Optimalstandard erreichen, signifikant höher ausfallen als die entsprechenden Anteile der Mädchen (Differenz von 8 bzw. 5 Prozentpunkten).

Insgesamt verdeutlichen die Ergebnisse zum Erreichen der Bildungsstandards für Jungen und Mädchen, dass geschlechtsbezogene Disparitäten sowohl im unteren als auch im oberen Bereich der hier untersuchten Kompetenzverteilungen vorliegen. Das Muster dieser Disparitäten ähnelt dem Befundmuster, das für die

Kompetenzmittelwerte beschrieben wurde. So sind im Fach Deutsch die Jungen im unteren Bereich der Kompetenzverteilung über- und im oberen Bereich unterrepräsentiert. Im Fach Mathematik zeigt sich das umgekehrte Bild: Hier sind die Anteile der Mädchen im unteren Kompetenzbereich höher und im oberen Kompetenzbereich geringer als die der Jungen.

7.2.3 Geschlechtsbezogene Kompetenzunterschiede in den Ländern

Nachdem die geschlechtsbezogenen Disparitäten in den vorherigen Abschnitten zunächst für Deutschland insgesamt beschrieben wurden, wird im Folgenden auf die Kompetenzunterschiede zwischen Jungen und Mädchen in den Ländern eingegangen. Die Ergebnisse für die Fächer Deutsch und Mathematik im Jahr 2016 sind in Abbildung 7.2 dargestellt. Für alle Kompetenzbereiche im Fach Deutsch sowie für die Globalskala im Fach Mathematik werden für jedes Land die Kompetenzmittelwerte der Jungen (M_J) und der Mädchen (M_M), deren Differenz ($M_J - M_M$) und die zugehörigen Standardfehler (SE) sowie die Effektstärke d angegeben. Die geschlechtsbezogenen Kompetenzunterschiede werden zudem grafisch veranschaulicht. Nach rechts weisende Balken entsprechen einem Kompetenzvorsprung der Jungen. Statistisch signifikante Unterschiede zwischen Jungen und Mädchen sind durch Fettdruck und ausgefüllte Balken gekennzeichnet.⁹

Wie bereits in Abschnitt 7.2.1 beschrieben, liegt in Deutschland insgesamt im Kompetenzbereich *Lesen* ein signifikanter Kompetenzvorsprung der Mädchen von durchschnittlich 24 Punkten vor. Abbildung 7.2 zeigt, dass auch in allen Ländern signifikante geschlechtsbezogene Disparitäten in der Lesekompetenz zugunsten der Mädchen zu verzeichnen sind, wobei jedoch die Größe des Kompetenzunterschieds zwischen den Ländern erheblich variiert. Besonders groß ist der Vorsprung der Mädchen im *Lesen* in den Ländern Mecklenburg-Vorpommern (33 Punkte) und Brandenburg (32 Punkte). Die geringste geschlechtsbezogene Disparität in der Lesekompetenz liegt hingegen im Land Bremen vor (13 Punkte).

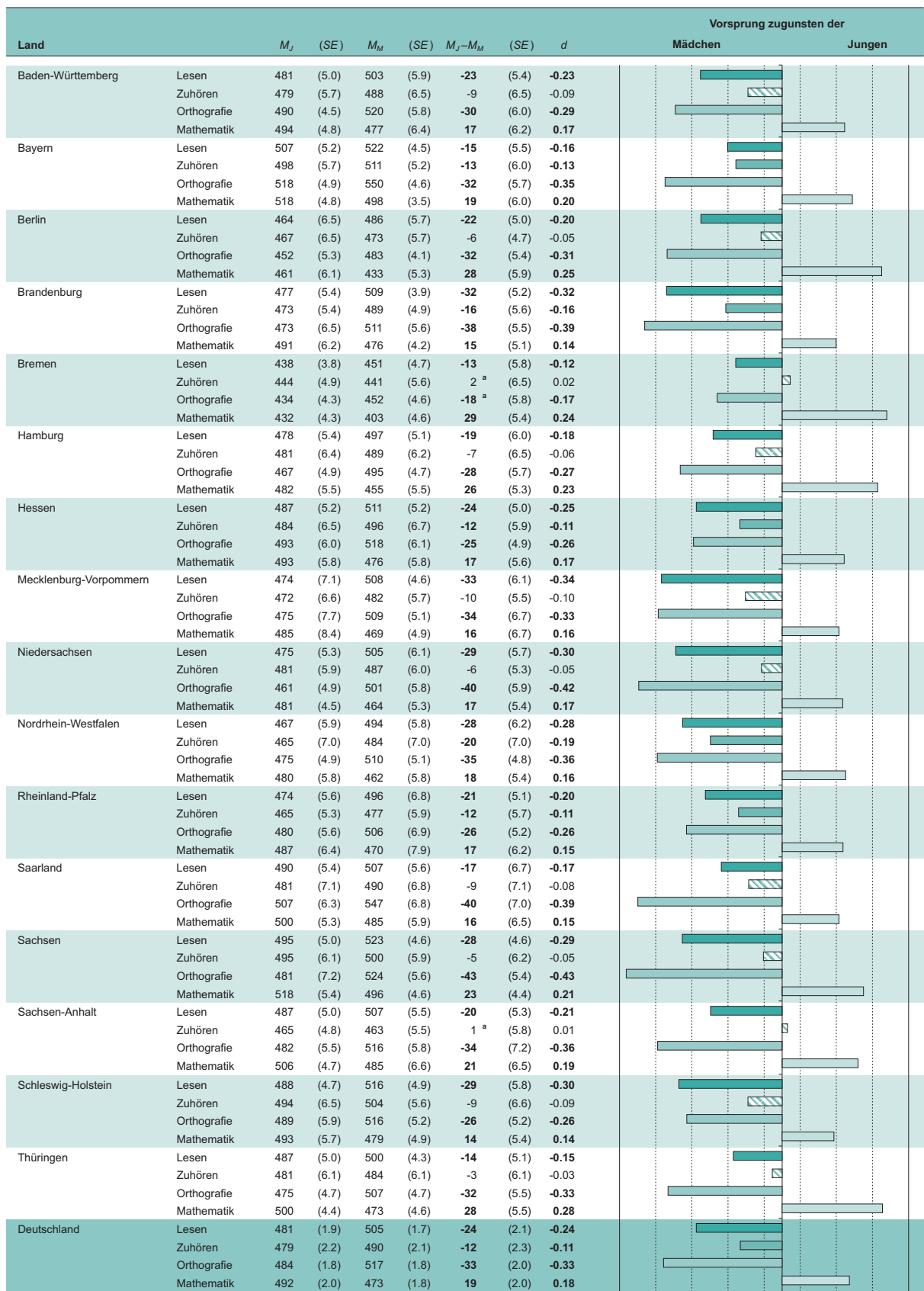
Im Kompetenzbereich *Zuhören* ergibt sich bundesweit ein vergleichsweise kleiner, jedoch statistisch signifikanter Kompetenzvorsprung der Mädchen von 12 Punkten. Signifikant höhere Kompetenzmittelwerte als Jungen erzielen Mädchen in den Ländern Bayern, Brandenburg, Hessen, Nordrhein-Westfalen und Rheinland-Pfalz, wobei der Kompetenzunterschied zwischen Jungen und Mädchen in Nordrhein-Westfalen im Ländervergleich besonders groß ausfällt (20 Punkte). In den übrigen elf Ländern finden sich keine signifikanten geschlechtsbezogenen Disparitäten im *Zuhören*.

In der *Orthografie* besteht mit 33 Punkten für Deutschland insgesamt ein besonders großer Unterschied in den Kompetenzmittelwerten von Jungen und Mädchen. Einen durchweg signifikanten Vorsprung der Mädchen ergibt auch die Analyse der orthografischen Kompetenz in den Ländern. Dabei fällt der Kompetenzvorteil der Mädchen in den Ländern Sachsen (43 Punkte), Niedersachsen (40 Punkte) und dem Saarland (40 Punkte) besonders groß aus, während er im Land Bremen wiederum am niedrigsten ist (18 Punkte).

Wie bereits erläutert, ergeben die Analysen für das Fach Mathematik bundesweit ein anderes Befundmuster als für das Fach Deutsch. Für die Globalskala im Fach Mathematik besteht in Deutschland insgesamt ein Kompetenzvorsprung

9 Ergänzend können Angaben zur Streuung der Kompetenzen von Jungen und Mädchen in den Ländern im Jahr 2016 auf der Webseite des IQB als Zusatzmaterial abgerufen werden (vgl. Tab. 7.1web).

Abbildung 7.2: Kompetenzunterschiede zwischen Jungen und Mädchen in den Fächern Deutsch und Mathematik in den Ländern



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz der Mittelwerte von Jungen und Mädchen abweichen. M_M = Mittelwert in der Gruppe der Mädchen; M_J = Mittelwert in der Gruppe der Jungen; $M_J - M_M$ = Differenz der Mittelwerte von Jungen und Mädchen; SE = Standardfehler; d = Effektstärke Cohens d ; Mathematik = Mathematik (Globalskala).
^a Statistisch signifikante Differenz ($p < .05$) zum Wert für Deutschland.
 Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

-45 -35 -25 -15 -5 0 5 15 25 35
 ■ Lesen ■ Zuhören ■ Orthografie ■ Mathematik
 Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz zwischen Jungen und Mädchen an.

von 19 Punkten zugunsten der Jungen. Auch in allen Ländern ist ein signifikanter Vorsprung der Jungen in der mathematischen Kompetenz zu verzeichnen. Dieser ist in den Ländern Bremen (29 Punkte), Berlin (28 Punkte), Thüringen (28 Punkte) und Hamburg (26 Punkte) besonders ausgeprägt. Die geringste Differenz zwischen den Kompetenzmittelwerten von Jungen und Mädchen im Fach Mathematik ist in Schleswig-Holstein zu verzeichnen (14 Punkte).

7.3 Geschlechtsbezogene Kompetenzunterschiede in den Jahren 2011 und 2016 im Vergleich

In den vorherigen Abschnitten wurden die Unterschiede in den Kompetenzen von Jungen und Mädchen im Jahr 2016 beschrieben. Im Folgenden soll nun mit Hilfe von Trendanalysen der Frage nachgegangen werden, wie sich die von Jungen und Mädchen gegen Ende der 4. Jahrgangsstufe erreichten Kompetenzen in den letzten fünf Jahren entwickelt haben und inwieweit die geschlechtsbezogenen Disparitäten gegenüber dem Jahr 2011 verringert werden konnten.

In einem ersten Schritt wird zunächst separat für die Teilpopulationen der Jungen und der Mädchen dargestellt, inwieweit sich die in den Fächern Deutsch und Mathematik untersuchten Kompetenzen seit dem IQB-Ländervergleich 2011 verändert haben (Abschnitt 7.3.1). Im zweiten Schritt wird geprüft, ob diese Entwicklungen zu einer Veränderung der geschlechtsbezogenen Disparitäten geführt haben (Abschnitt 7.3.2). Die Ergebnisse werden jeweils für Deutschland insgesamt und separat für die 16 Länder dargestellt. Eine Besonderheit ist für den Kompetenzbereich *Orthografie* zu beachten: Die orthografischen Kompetenzen der Schülerinnen und Schüler wurden im Jahr 2011 nur in einer bundesweit repräsentativen Teilstichprobe erhoben, die weder eine Auswertung auf Länderebene erlaubte noch Schülerinnen und Schüler in Förderschulen einschloss. Deshalb können Trendanalysen zur *Orthografie* lediglich für Deutschland insgesamt und nur für Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf durchgeführt werden (vgl. hierzu auch Kapitel 3.2).

7.3.1 Trends in den erreichten Kompetenzen von Jungen und Mädchen

Die Abbildungen 7.3 bis 7.5 zeigen für Deutschland insgesamt sowie für die einzelnen Länder die von Jungen und Mädchen im jeweiligen Kompetenzbereich erreichten Mittelwerte (M) in den Jahren 2011 und 2016, die Differenz der Mittelwerte für die beiden Erhebungszeitpunkte (ΔM) sowie die zugehörigen Standardfehler (SE). Die Veränderung zwischen den Erhebungszeitpunkten wird zudem grafisch veranschaulicht, wobei nach rechts weisende Balken eine Steigerung der erreichten Kompetenzen anzeigen. Statistisch signifikante Veränderungen sind durch Fettdruck und ausgefüllte Balken gekennzeichnet.

Im *Lesen* (vgl. Abb. 7.3) haben sich in der Teilpopulation der Jungen in Deutschland insgesamt die im Mittel erreichten Kompetenzen gegenüber dem Jahr 2011 nicht signifikant verändert. Auf Länderebene sind jedoch signifikante negative Trends in Baden-Württemberg (–16 Punkte), Bremen (–17 Punkte) und Nordrhein-Westfalen (–17 Punkte) zu verzeichnen. Betrachtet man die Teilpopulation der Mädchen, so zeigt sich für Deutschland insgesamt ein signifikanter Rückgang der im Mittel erreichten Kompetenzen um 7 Punkte. Ein im Jahr 2016 signifikant geringerer Kompetenzmittelwert als im Jahr 2011 ergibt

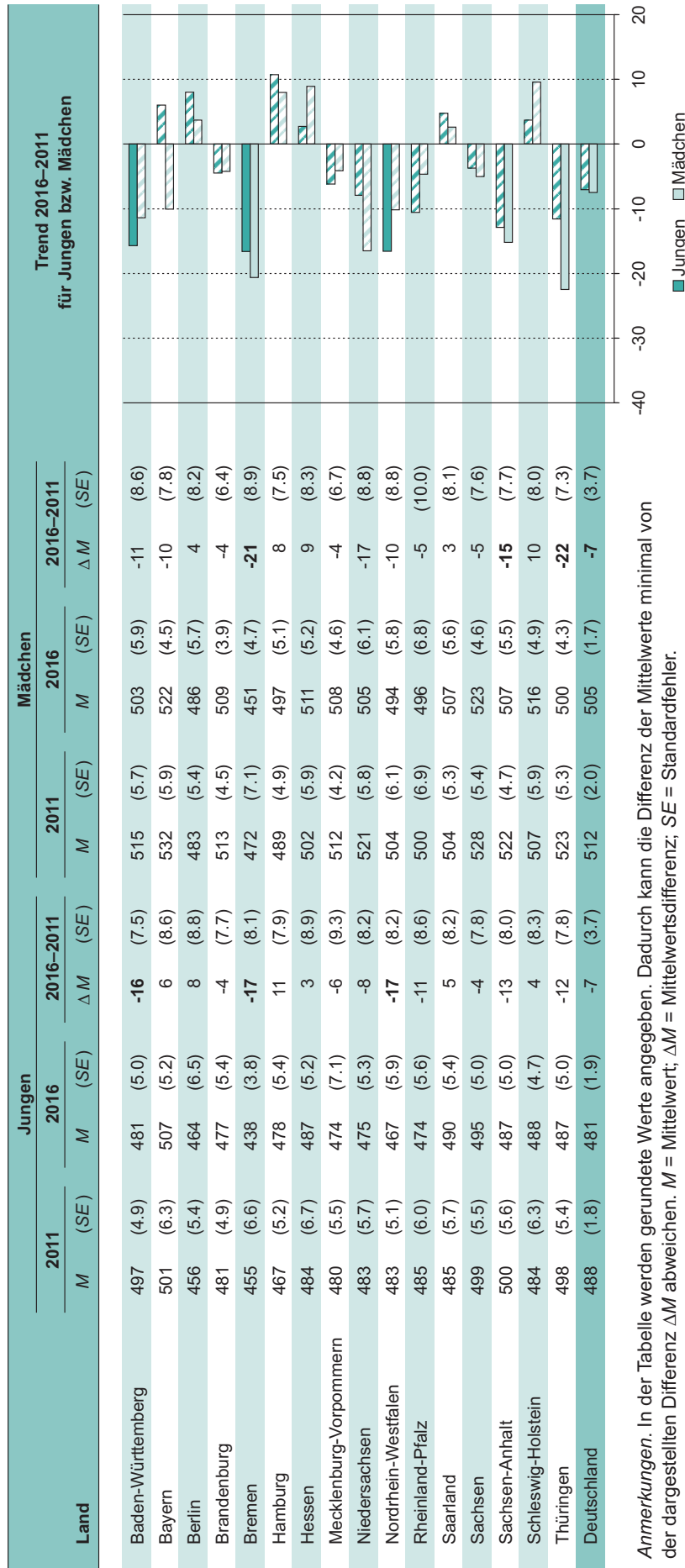
sich für die Mädchen auch in den Ländern Bremen (–21 Punkte), Sachsen-Anhalt (–15 Punkte) und Thüringen (–22 Punkte). Ein signifikanter Kompetenzrückgang in beiden betrachteten Teilpopulationen liegt somit nur im Land Bremen vor. Signifikant positive Entwicklungen der Kompetenzwerte der Jungen und Mädchen sind im *Lesen* für keines der betrachteten Länder zu verzeichnen.

Abbildung 7.4 stellt die Ergebnisse für den Kompetenzbereich *Zuhören* dar. In Deutschland insgesamt hat sich das mittlere Kompetenzniveau in der Teilpopulation der Jungen gegenüber dem Jahr 2011 um 20 Punkte verringert. Auch in den Ländern Baden-Württemberg (–29 Punkte), Bremen (–25 Punkte), Mecklenburg-Vorpommern (–19 Punkte), Niedersachsen (–18 Punkte), Nordrhein-Westfalen (–31 Punkte), Rheinland-Pfalz (–25 Punkte) und Sachsen-Anhalt (–24 Punkte) ist für die Jungen im Durchschnitt ein signifikanter Kompetenzrückgang im *Zuhören* festzustellen. Ein ähnliches Befundmuster ergibt sich für die Teilpopulation der Mädchen: Hier findet sich für Deutschland insgesamt ein signifikanter Rückgang des Kompetenzmittelwerts im *Zuhören* von 11 Punkten. Ein negativer Trend liegt auch für die Länder Baden-Württemberg (–22 Punkte), Bremen (–24 Punkte), Mecklenburg-Vorpommern (–17 Punkte), Niedersachsen (–27 Punkte), Sachsen-Anhalt (–23 Punkte) und Thüringen (–21 Punkte) vor. In fünf Ländern ist somit sowohl für die Jungen als auch für die Mädchen ein signifikanter negativer Trend im Kompetenzbereich *Zuhören* zu beobachten (Baden-Württemberg, Bremen, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt). Auch in diesem Bereich findet sich auf Länderebene für keine der betrachteten Teilpopulationen ein signifikanter Zuwachs der mittleren Kompetenzen.

In der *Orthografie* (ohne Abbildung) zeigt sich in Deutschland insgesamt sowohl in der Teilpopulation der Mädchen als auch in der Teilpopulation der Jungen ein signifikanter Kompetenzrückgang. Dieser beträgt in der Teilpopulation der Jungen 22 Punkte und in der Teilpopulation der Mädchen 26 Punkte.

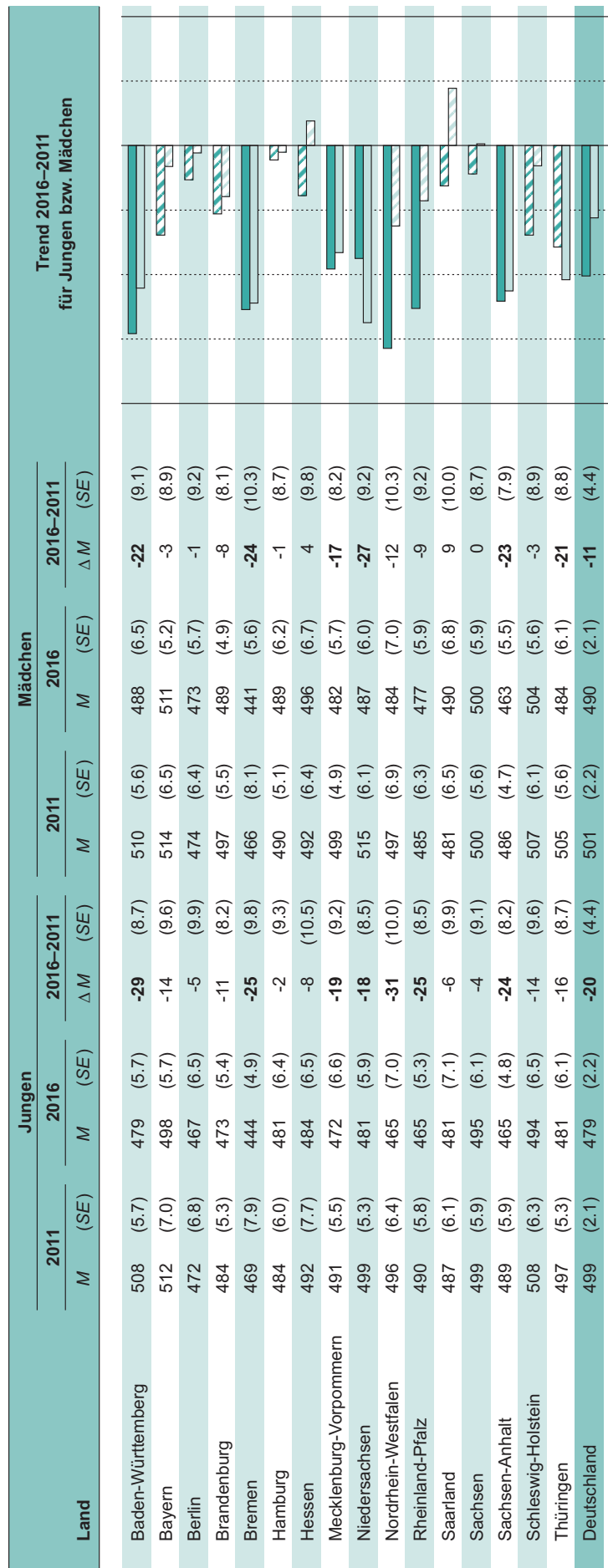
Abbildung 7.5 gibt die Ergebnisse für das Fach Mathematik (*Globalskala*) wieder. In der Teilpopulation der Jungen in Deutschland insgesamt zeigt sich auch für die mathematische Kompetenz im Jahr 2016 ein signifikant geringerer Kompetenzmittelwert als noch im Jahr 2011 (–16 Punkte). Ein signifikant negativer Trend ist für die Jungen auch in den Ländern Baden-Württemberg (–23 Punkte), Bremen (–33 Punkte), Niedersachsen (–21 Punkte), Nordrhein-Westfalen (–26 Punkte) und Sachsen-Anhalt (–17 Punkte) zu verzeichnen. Für die Mädchen fällt das Befundmuster ähnlich aus: Für diese Teilpopulation liegt in Deutschland insgesamt ebenfalls ein signifikanter Rückgang in der mathematischen Kompetenz vor (–19 Punkte). Auch in den Ländern Baden-Württemberg (–30 Punkte), Bayern (–15 Punkte), Bremen (–35 Punkte), Mecklenburg-Vorpommern (–21 Punkte), Niedersachsen (–26 Punkte), Nordrhein-Westfalen (–26 Punkte), Sachsen-Anhalt (–25 Punkte) und Thüringen (–19 Punkte) zeigt sich eine signifikante Verringerung der Kompetenzen und somit unter anderem in allen zuvor genannten Ländern, in denen sich auch das Kompetenzniveau der Jungen signifikant reduziert hat. Wie im *Lesen* und *Zuhören* im Fach Deutsch ist auch im Fach Mathematik auf Länderebene für keine der beiden Teilpopulationen ein signifikanter Kompetenzzuwachs zu beobachten.

Abbildung 7.3: Vergleich der in den Jahren 2011 und 2016 von Jungen beziehungsweise Mädchen der 4. Jahrgangsstufe erreichten Kompetenzen im Kompetenzbereich Lesen im Fach Deutsch



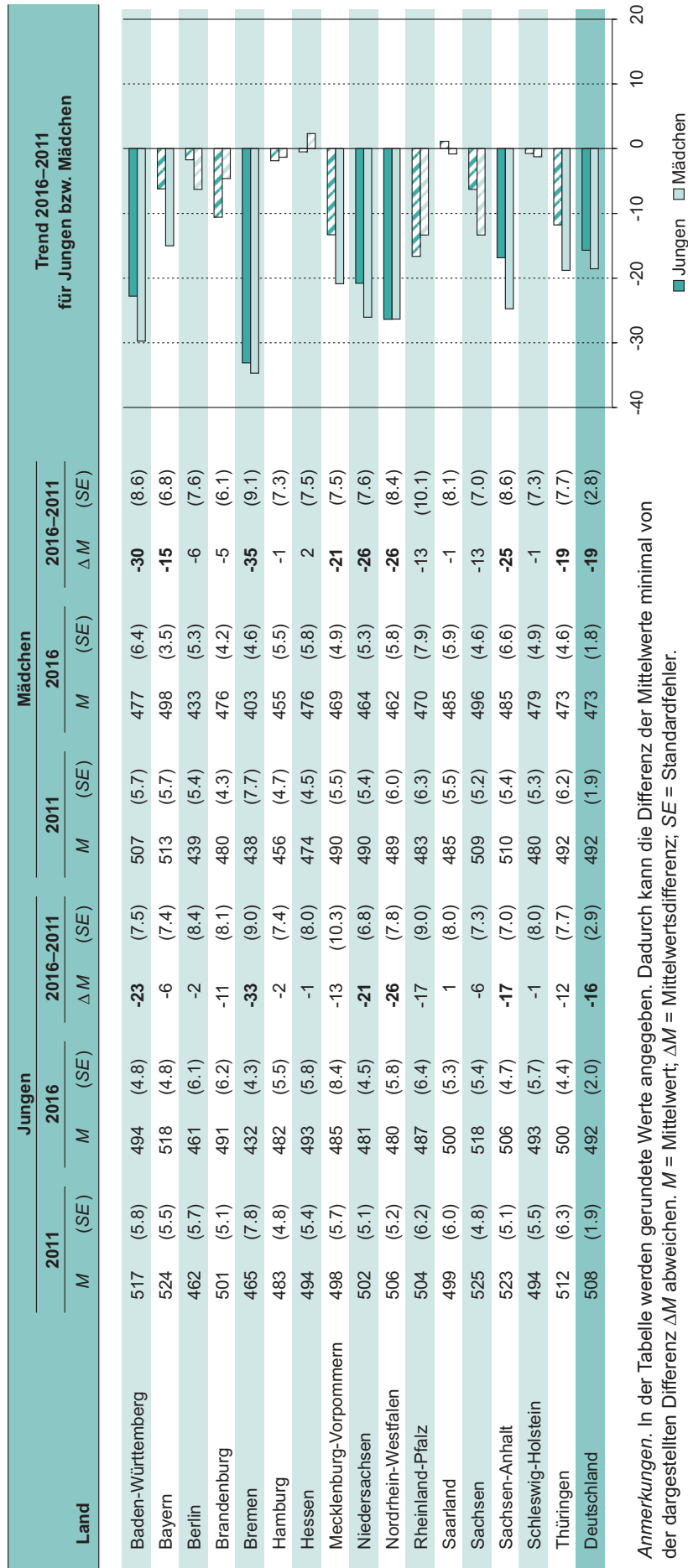
Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M = Mittelwert; ΔM = Mittelwertsdifferenz; SE = Standardfehler. Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$). Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Abbildung 7.4: Vergleich der in den Jahren 2011 und 2016 von Jungen beziehungsweise Mädchen der 4. Jahrgangsstufe erreichten Kompetenzen im Kompetenzbereich Zuhören im Fach Deutsch



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M = Mittelwert; ΔM = Mittelwertsdifferenz; SE = Standardfehler. Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$). Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Abbildung 7.5: Vergleich der in den Jahren 2011 und 2016 von Jungen beziehungsweise Mädchen der 4. Jahrgangsstufe erreichten Kompetenzen im Fach Mathematik (Globalskala)



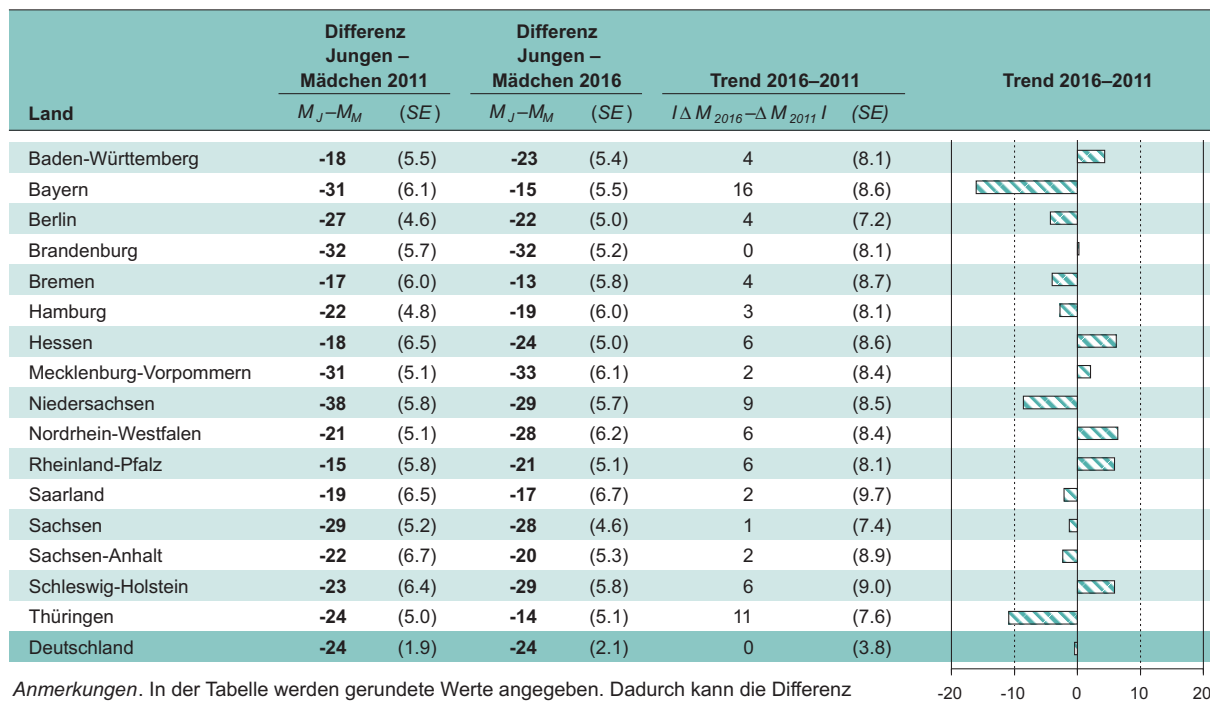
Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M = Mittelwert; ΔM = Mittelwertsdifferenz; SE = Standardfehler. Felt gedruckte Differenzen sind statistisch signifikant ($p < .05$). Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

7.3.2 Trends in geschlechtsbezogenen Disparitäten

Im Folgenden wird der Frage nachgegangen, inwieweit sich die Unterschiede in den von Jungen und Mädchen im Mittel erreichten Kompetenzen seit dem Jahr 2011 in Deutschland insgesamt sowie in den Ländern verändert haben.

In Abbildung 7.6 werden für den Kompetenzbereich *Lesen* im Fach Deutsch die Kompetenzunterschiede zwischen Jungen und Mädchen in den Jahren 2011 und 2016 gegenübergestellt. Angegeben ist zum einen die Differenz zwischen den im Mittel von Jungen und Mädchen erzielten Kompetenzen im Jahr 2011 beziehungsweise 2016 ($M_J - M_M$). Diese Werte stellen die geschlechtsbezogenen Disparitäten in den beiden Erhebungsjahren dar. In der rechten Spalte des Tabellenteils der Abbildung ist zudem angegeben, um wie viele Punkte sich die geschlechtsbezogenen Disparitäten zwischen den Erhebungszeitpunkten verändert haben. Dargestellt ist der Betrag der Differenz zwischen den Kompetenzunterschieden in den Jahren 2016 und 2011 ($|\Delta M_{2016} - \Delta M_{2011}|$). Die angegebene Veränderung der geschlechtsbezogenen Disparitäten in Punktwerten gibt zunächst also keine Auskunft über die Richtung der Veränderung. Diese geht aus der grafischen Darstellung hervor, wobei nach rechts weisende Balken eine Vergrößerung des Kompetenzunterschieds zwischen Jungen und Mädchen seit dem Jahr 2011 anzeigen. Angegeben sind zudem für alle Werte die entsprechenden Standardfehler (*SE*). Signifikante Veränderungen sind durch Fettdruck und ausgefüllte Balken gekennzeichnet.

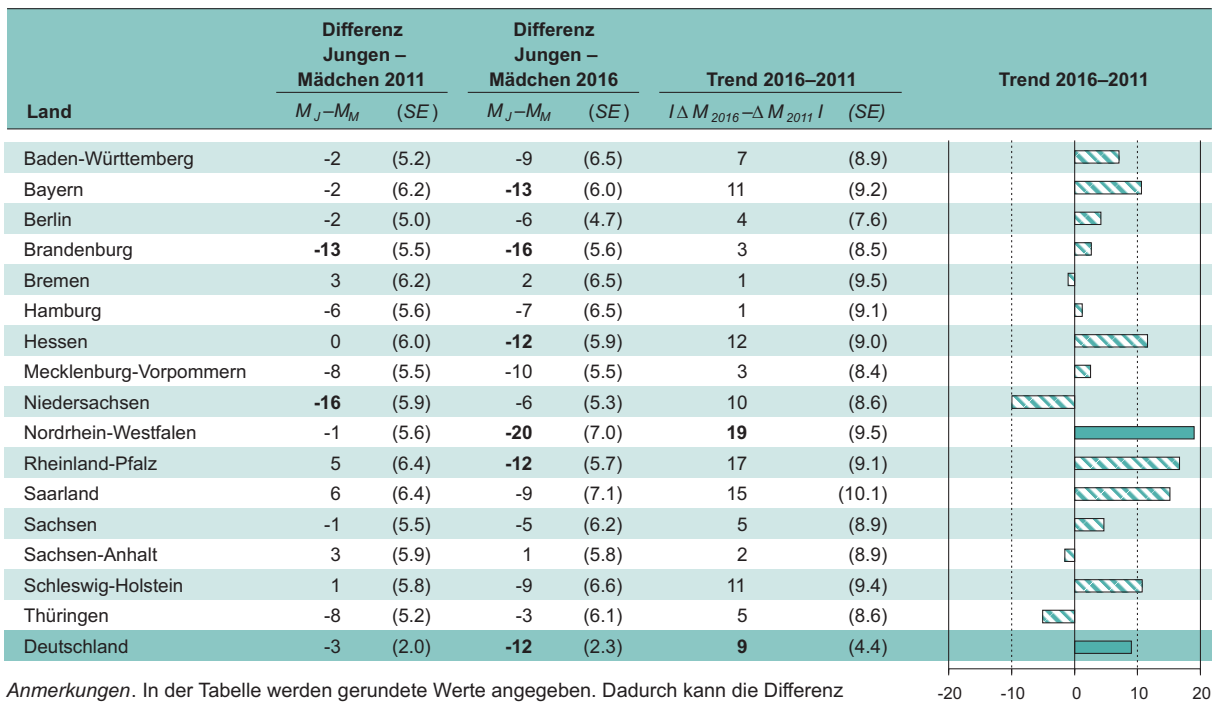
Abbildung 7.6: Kompetenzunterschiede zwischen Jungen und Mädchen im Kompetenzbereich *Lesen* im Fach Deutsch im Trend



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M_M = Mittelwert in der Gruppe der Mädchen; M_J = Mittelwert in der Gruppe der Jungen; *SE* = Standardfehler der Mittelwertsdifferenz; ΔM_{2016} = Mittelwertsdifferenz 2016; ΔM_{2011} = Mittelwertsdifferenz 2011; *d* = Effektstärke Cohens *d*.

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$). Nach links zeigende Balken zeigen eine Verringerung des Geschlechtsunterschieds an, nach rechts zeigende Balken eine Zunahme.

Statistisch signifikante Differenz ($p < .05$)
 Statistisch nicht signifikante Differenz

Abbildung 7.7: Kompetenzunterschiede zwischen Jungen und Mädchen im Kompetenzbereich *Zuhören* im Fach Deutsch im Trend

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M_M = Mittelwert in der Gruppe der Mädchen; M_J = Mittelwert in der Gruppe der Jungen; SE = Standardfehler der Mittelwertsdifferenz; ΔM_{2016} = Mittelwertsdifferenz 2016; ΔM_{2011} = Mittelwertsdifferenz 2011; d = Effektstärke Cohens d .

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$). Nach links zeigende Balken zeigen eine Verringerung des Geschlechtsunterschieds an, nach rechts zeigende Balken eine Zunahme.

■ Statistisch signifikante Differenz ($p < .05$)
 ▨ Statistisch nicht signifikante Differenz

Die Ergebnisse zum Kompetenzbereich *Lesen* zeigen, dass sich zwischen den Erhebungszeitpunkten in den Jahren 2011 und 2016 keine signifikanten Veränderungen in den Geschlechterdisparitäten ergeben haben. Weder in Deutschland insgesamt noch in den einzelnen Ländern haben sich die Disparitäten im *Lesen* demnach statistisch bedeutsam reduziert oder vergrößert.

Ein anderes Befundmuster liegt für den Kompetenzbereich *Zuhören* vor (vgl. Abb. 7.7). In Deutschland insgesamt hat sich der Unterschied in den Kompetenzmittelwerten von Jungen und Mädchen seit dem Jahr 2011 signifikant um 9 Punkte vergrößert. Auf Länderebene findet sich hingegen nur für Nordrhein-Westfalen eine signifikante Zunahme des Kompetenzunterschieds zwischen Jungen und Mädchen (+19 Punkte). In den übrigen Ländern haben die geschlechtsbezogenen Disparitäten seit dem Jahr 2011 nicht signifikant zu- oder abgenommen.¹⁰

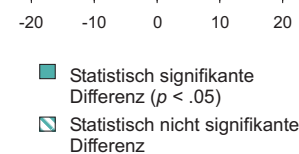
10 Eine Veränderung der Disparitäten liegt vor, wenn sich die Trends der erreichten Kompetenzen zwischen zwei Gruppen signifikant unterscheiden und sich eine Gruppe somit statistisch signifikant stärker verändert als die andere Gruppe. Die statistische Signifikanz der Veränderung wird jedoch nicht nur von der Höhe der Veränderung, sondern auch von der Unsicherheit der Schätzung beeinflusst (siehe Kapitel 3.2 und 13.1). Die Prüfung der statistischen Signifikanz von Gruppenunterschieden in Veränderungen bezieht mehr Unsicherheitsquellen ein als die querschnittliche Signifikanzprüfung von Disparitäten. Daher kann es also vorkommen, dass Disparitäten zwischen zwei Gruppen – hier Jungen und Mädchen – einmal statistisch signifikant sind (etwa im Jahr 2016) und einmal nicht (etwa im Jahr 2011), sich die Veränderungen innerhalb der beiden Gruppen aber nicht signifikant voneinander unterscheiden, da sie mit größerer Unsicherheit behaftet sind.

Abbildung 7.8: Kompetenzunterschiede zwischen Jungen und Mädchen im Fach Mathematik (*Globalskala*) im Trend

Land	Differenz Jungen – Mädchen 2011		Differenz Jungen – Mädchen 2016		Trend 2016–2011		Trend 2016–2011
	$M_J - M_M$	(SE)	$M_J - M_M$	(SE)	$ \Delta M_{2016} - \Delta M_{2011} $	(SE)	
Baden-Württemberg	10	(5.9)	17	(6.2)	7	(8.7)	
Bayern	11	(4.5)	19	(6.0)	9	(7.5)	
Berlin	23	(4.4)	28	(5.9)	5	(7.4)	
Brandenburg	21	(5.4)	15	(5.1)	6	(7.6)	
Bremen	27	(5.7)	29	(5.4)	2	(7.9)	
Hamburg	27	(4.2)	26	(5.3)	1	(6.8)	
Hessen	20	(4.7)	17	(5.6)	3	(7.4)	
Mecklenburg-Vorpommern	8	(4.6)	16	(6.7)	8	(8.2)	
Niedersachsen	11	(5.3)	17	(5.4)	5	(7.6)	
Nordrhein-Westfalen	18	(5.0)	18	(5.4)	0	(7.4)	
Rheinland-Pfalz	20	(6.8)	17	(6.2)	3	(9.3)	
Saarland	14	(5.5)	16	(6.5)	2	(8.6)	
Sachsen	15	(5.0)	23	(4.4)	7	(6.8)	
Sachsen-Anhalt	13	(6.3)	21	(6.5)	8	(9.1)	
Schleswig-Holstein	14	(6.6)	14	(5.4)	0	(8.6)	
Thüringen	21	(6.3)	28	(5.5)	7	(8.4)	
Deutschland	16	(1.8)	19	(2.0)	3	(2.8)	

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M_M = Mittelwert in der Gruppe der Mädchen; M_J = Mittelwert in der Gruppe der Jungen; SE = Standardfehler der Mittelwertsdifferenz; ΔM_{2016} = Mittelwertsdifferenz 2016; ΔM_{2011} = Mittelwertsdifferenz 2011; d = Effektstärke Cohens d .

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$). Nach links zeigende Balken zeigen eine Verringerung des Geschlechtsunterschieds an, nach rechts zeigende Balken eine Zunahme.



Im Kompetenzbereich *Orthografie* (ohne Abbildung) hat sich der für Deutschland insgesamt ermittelte Kompetenzunterschied zwischen Jungen und Mädchen zwischen den Erhebungszeitpunkten nicht signifikant verändert.

Auch für das Fach Mathematik (*Globalskala*; vgl. Abb. 7.8) zeigt sich – wie bereits im *Lesen* und in der *Orthografie* im Fach Deutsch – für Deutschland insgesamt keine signifikante Veränderung des Kompetenzunterschieds zwischen Jungen und Mädchen im untersuchten Fünfjahreszeitraum. Dies gilt auch für die einzelnen Länder.

7.4 Geschlechtsbezogene Unterschiede in motivationalen Merkmalen

7.4.1 Geschlechtsbezogene Unterschiede im fachbezogenen Selbstkonzept und im fachlichen Interesse

Nachdem in den vorherigen Abschnitten geschlechtsbezogene Disparitäten in den untersuchten Kompetenzen beschrieben wurden, werden in diesem Abschnitt Unterschiede zwischen Jungen und Mädchen im fachbezogenen Selbstkonzept und im fachlichen Interesse analysiert.

Das Selbstkonzept und das Interesse in den Fächern Deutsch und Mathematik wurden im IQB-Bildungstrend 2016 bei ungefähr zwei Dritteln der Stichprobe erfasst, wobei die Auswahl der befragten Schülerinnen und Schüler auf Klassenebene zufällig getroffen wurde. Für die im Folgenden darge-

Abbildung 7.9: Fachbezogenes Selbstkonzept und fachliches Interesse von Jungen und Mädchen in den Fächern Deutsch und Mathematik

Merkmal		niedrig	mittel	hoch	<i>M</i>	<i>SD</i>	$M_J - M_M$	(<i>SE</i>)	<i>d</i>
Deutsch									
Selbstkonzept	Jungen	9%	22%	69%	3.1	0.7	-0.18	(0.01)	-0.28
	Mädchen	5%	15%	80%	3.3	0.6			
Interesse	Jungen	21%	29%	50%	2.8	0.8	-0.27	(0.01)	-0.36
	Mädchen	11%	25%	64%	3.1	0.7			
Mathematik									
Selbstkonzept	Jungen	8%	14%	78%	3.3	0.7	0.30	(0.02)	0.39
	Mädchen	15%	21%	64%	3.0	0.8			
Interesse	Jungen	11%	16%	73%	3.3	0.8	0.25	(0.02)	0.30
	Mädchen	16%	22%	61%	3.0	0.8			

Anmerkungen. Im Balkendiagramm werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. *M* = Mittelwert; *SD* = Standardabweichung; $M_M - M_J$ = Differenz der Mittelwerte von Jungen und Mädchen; *SE* = Standardfehler der Mittelwertsdifferenz; *d* = Effektstärke Cohens *d*. Die Balken geben die Anteile von Schülerinnen und Schülern mit niedrigem (Skalenwert ≤ 2), mittlerem (Skalenwert > 2 und < 3) und hohem (Skalenwert ≥ 3) Selbstkonzept bzw. Interesse an.

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

stellten Analysen wurden die fehlenden Werte multipel imputiert (Graham, 2009; Schafer & Graham, 2002). Die Skala zur Erfassung des fachbezogenen Selbstkonzepts im Fach Deutsch enthielt vier Aussagen, wie beispielsweise „Ich bin gut in Deutsch“. Die Kinder wurden gebeten, auf einer vierstufigen Skala von 1 = „stimmt gar nicht“ bis 4 = „stimmt genau“ zu beurteilen, inwieweit diese Aussagen für sie selbst zutreffen. Die Erhebung des fachlichen Interesses erfolgte ebenfalls anhand von vier Aussagen, wie etwa „Deutsch ist ein Lieblingsfach von mir“. Hierbei kam dieselbe vierstufige Antwortskala zum Einsatz wie bei der Erfassung des Selbstkonzepts. Für das Fach Mathematik wurden gleichlautende Fragen vorgelegt. Anhand der Angaben zu den jeweiligen Aussagen konnte pro Fach für jede Schülerin und jeden Schüler ein Skalenwert für das fachbezogene Selbstkonzept und ein Skalenwert für das fachliche Interesse gebildet werden. In Anlehnung an das Vorgehen von Jansen und Kollegen (2013) wurde auf Basis dieser Skalenwerte der Anteil der Schülerinnen und Schüler ermittelt, die über ein „niedriges“ (Skalenwert ≤ 2), ein „mittleres“ (Skalenwert > 2 und < 3) oder ein „hohes“ (Skalenwert ≥ 3) Selbstkonzept beziehungsweise Interesse verfügen.

Abbildung 7.9 stellt die Anteile der Jungen und der Mädchen mit niedrigem, mittlerem und hohem fachspezifischen Selbstkonzept beziehungsweise fachlichem Interesse in den Fächern Deutsch und Mathematik dar. Rechts neben den horizontalen Stapelbalken sind die jeweiligen Skalenmittelwerte (*M*) sowie die Standardabweichungen der Skalenwerte (*SD*) angegeben. Die Differenz der Mittelwerte zwischen den Teilpopulationen der Jungen und der Mädchen kann der Spalte $M_J - M_M$ entnommen werden, wobei auch der zugehörige Standardfehler (*SE*) angegeben ist. Zusätzlich aufgeführt ist das Effektstärkemaß Cohens *d*.

In beiden untersuchten Fächern schätzen sowohl die Jungen als auch die Mädchen ihre eigenen Fähigkeiten und ihr Interesse überwiegend als hoch ein. So liegt der Anteil derjenigen, die nach eigenen Angaben über ein hohes Interesse beziehungsweise ein hohes Selbstkonzept verfügen, je nach Fach, Skala und Geschlecht zwischen 50 Prozent und 80 Prozent. Trotz dieser generell ho-

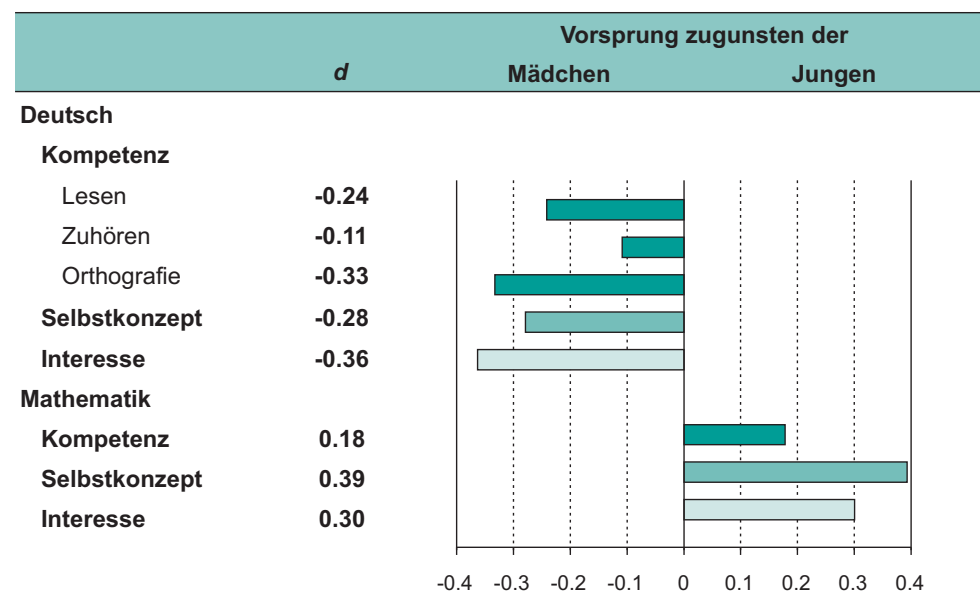
hen Zustimmung zu den im Fragebogen präsentierten Aussagen sind sowohl für das fachbezogene Selbstkonzept als auch für das fachliche Interesse signifikante Geschlechterunterschiede zu verzeichnen. Während Mädchen im Fach Deutsch ihre eigene Kompetenz im Mittel höher einschätzen als Jungen ($d = 0.28$) und auch ein größeres Interesse am Fach Deutsch zeigen ($d = 0.36$), schätzen sich Jungen im Fach Mathematik im Durchschnitt als fähiger ein ($d = 0.39$) und geben ein im Mittel größeres Interesse an mathematischen Inhalten an ($d = 0.30$) als Mädchen.

7.4.2 Geschlechtsbezogene Unterschiede in Selbstkonzept und Interesse im Vergleich zum erreichten Kompetenzniveau

Wie bereits in Abschnitt 7.1.2 erwähnt, kann die Ausprägung des fachbezogenen Selbstkonzepts und das fachliche Interesse die Kompetenzentwicklung beeinflussen (vgl. auch Marsh & Martin, 2011). Bei den entsprechenden Selbsteinschätzungen handelt es sich jedoch um subjektive Angaben, die nicht perfekt mit den von den Schülerinnen und Schülern erreichten Kompetenzen übereinstimmen. Daher sollen im Folgenden die Unterschiede zwischen Jungen und Mädchen in den motivationalen Merkmalen den geschlechtsbezogenen Disparitäten in den Kompetenzen in den Fächern Deutsch und Mathematik gegenübergestellt werden.

Die entsprechenden Befunde sind in Abbildung 7.10 zusammengefasst. Im Fach Deutsch wurden Kompetenzvorteile der Mädchen gefunden, die zwischen $d = 0.11$ (*Zuhören*) und $d = 0.33$ (*Orthografie*) betragen. Die Geschlechter-

Abbildung 7.10: Geschlechterunterschiede in den erreichten Kompetenzen, im fachbezogenen Selbstkonzept und im fachlichen Interesse für die Fächer Deutsch und Mathematik



Anmerkungen. d = Effektstärke Cohens d .

Fett gedruckte d -Werte sind statistisch signifikant ($p < .05$).

Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

unterschiede im Selbstkonzept und im Interesse entsprechen damit in etwa dem Kompetenzunterschied in der *Orthografie*. Im Fach Mathematik wurde hingegen ein Kompetenzvorsprung der Jungen von $d = 0.18$ festgestellt. Hierzu passt, dass Jungen sich in Mathematik im Mittel auch als kompetenter einschätzen und ein ausgeprägteres Interesse am Fach bekunden. Die Effektstärken für die motivationalen Merkmale fallen jedoch etwa doppelt so groß aus wie für die tatsächlichen Kompetenzunterschiede. Der Geschlechterunterschied in den Selbsteinschätzungen zur mathematischen Kompetenz ist also deutlich größer als der Unterschied in den Testergebnissen.

Abschließend soll untersucht werden, inwieweit sich das Selbstkonzept und das Interesse der Schülerinnen und Schüler in Abhängigkeit davon unterscheiden, welches Kompetenzniveau sie erreicht haben. Hierzu sind in Abbildung 7.11 die durchschnittlichen Skalenwerte für das Selbstkonzept beziehungsweise das Interesse für diejenigen Jungen und Mädchen abgetragen, die im jeweiligen Kompetenzbereich Kompetenzstufe I (unter Mindeststandard) beziehungsweise Kompetenzstufe V (Optimalstandard) erreicht haben und somit über niedrige beziehungsweise besonders hohe Kompetenzen in diesem Kompetenzbereich verfügen. Zusätzlich werden auch die Werte für das mittlere Kompetenzniveau (Stufe III; Regelstandard) angegeben. Die folgende Beschreibung der Ergebnisse ist rein deskriptiv; es wurden keine Prüfungen auf statistische Signifikanz vorgenommen.

Für das fachbezogene Selbstkonzept ergibt sich in allen untersuchten Kompetenzbereichen der Befund, dass das Selbstkonzept von Kindern auf höheren Kompetenzstufen tendenziell positiver ausfällt als das Selbstkonzept von Kindern auf niedrigeren Kompetenzstufen. Dies gilt sowohl für die Teilpopulation der Jungen als auch für die Teilpopulation der Mädchen. Die Unterschiede zwischen den Kindern auf der höchsten Kompetenzstufe V und der niedrigsten Kompetenzstufe I fallen im Fach Deutsch jedoch relativ gering aus; sie betragen für alle untersuchten Kompetenzbereiche nur rund eine halbe Stufe auf der vierstufigen Antwortskala (vgl. Abschnitt 7.4.1). Im Fach Mathematik hingegen ist der Unterschied im Selbstkonzept zwischen den Kindern auf Kompetenzstufe V und I mit rund einer Stufe auf der Antwortskala deutlich stärker ausgeprägt. Über das gesamte Kompetenzspektrum und alle untersuchten Kompetenzbereiche verfügen Mädchen im Fach Deutsch und Jungen im Fach Mathematik über ein positiveres Selbstkonzept als die jeweils andere Teilpopulation.

Ein anderes Befundmuster ergibt sich teilweise für das fachliche Interesse. So ist in Abbildung 7.11 für die Kompetenzbereiche *Lesen* und *Zuhören* im Fach Deutsch weder für die Jungen noch für die Mädchen ein klarer Zusammenhang zwischen der erreichten Kompetenzstufe und der durchschnittlichen Ausprägung des fachlichen Interesses erkennbar. Auch im Kompetenzbereich *Orthografie* unterscheidet sich die Ausprägung des Interesses kaum in Abhängigkeit von der erreichten Kompetenzstufe. Anders hingegen verhält es sich im Fach Mathematik: Hier ist – ähnlich wie beim mathematischen Selbstkonzept – sowohl für Jungen als auch in der Teilpopulation der Mädchen ein positiver Zusammenhang zwischen fachlichem Interesse und erreichter Kompetenzstufe festzustellen. In allen untersuchten Kompetenzbereichen im Fach Deutsch zeigen Mädchen sowohl auf Kompetenzstufe I als auch auf den Kompetenzstufen III und V ein größeres Interesse am Fach Deutsch als Jungen, wohingegen im Fach Mathematik Jungen auf allen Kompetenzniveaus ein höheres Interesse angeben als Mädchen.

Abbildung 7.11: Selbstkonzept und Interesse von Jungen und Mädchen in den Fächern Deutsch und Mathematik nach erreichter Kompetenzstufe in den untersuchten Kompetenzbereichen



7.5 Zusammenfassung

Die in diesem Kapitel dargestellten Ergebnisse zeigen, dass im Jahr 2016 in allen untersuchten Kompetenzbereichen in den Fächern Deutsch und Mathematik gegen Ende der 4. Jahrgangsstufe signifikante Unterschiede zwischen den von Jungen und den von Mädchen im Durchschnitt erreichten Kompetenzen bestehen. Die Unterschiede fallen dabei stereotyp aus: Mädchen verfügen im Fach Deutsch im Mittel über höhere Kompetenzen als Jungen, wobei der Unterschied in der *Orthografie* am größten ist. Im *Zuhören* zeigt sich hingegen nur ein relativ geringer Kompetenzvorsprung der Mädchen. Demgegenüber erzielen Jungen im Fach Mathematik sowohl auf der Globalskala als auch in allen inhaltlichen Kompetenzbereichen (Leitideen) höhere Kompetenzen als Mädchen. Diese Befunde entsprechen im Wesentlichen den Ergebnissen anderer Schulleistungsstudien der letzten Jahre für den Primarbereich.

In Analysen zum Erreichen der Bildungsstandards zeigt sich, dass die geschlechtsbezogenen Disparitäten sowohl im unteren Bereich als auch im oberen Bereich der Kompetenzverteilungen zu verzeichnen sind. So ist im Fach Deutsch in allen untersuchten Kompetenzbereichen der Anteil der Jungen, die den Mindeststandard verfehlen, größer als der entsprechende Anteil der Mädchen, während der Anteil der Jungen, die den Optimalstandard erreichen, jeweils niedriger ausfällt. Im Fach Mathematik ergibt sich das umgekehrte Bild: Hier sind die Anteile der Mädchen im unteren Kompetenzbereich höher und im oberen Kompetenzbereich geringer als die der Jungen.

Die Ergebnisse auf Länderebene ähneln den Ergebnissen für Deutschland insgesamt. So findet sich in allen Ländern ein signifikanter Kompetenzvorsprung der Mädchen im *Lesen* und in der *Orthografie* im Fach Deutsch sowie ein signifikanter Vorsprung der Jungen im Fach Mathematik (*Globalskala*). Eine Ausnahme bildet lediglich der Kompetenzbereich *Zuhören* im Fach Deutsch, für den in der Mehrzahl der Länder keine signifikanten geschlechtsbezogenen Disparitäten vorliegen. Besonders geringe geschlechtsbezogene Disparitäten finden sich im Fach Deutsch in allen untersuchten Kompetenzbereichen im Land Bremen. Demgegenüber ist in den Ländern Brandenburg und Nordrhein-Westfalen in allen drei Bereichen ein signifikanter Vorsprung der Mädchen zu verzeichnen, der in der Tendenz größer ausfällt als in Deutschland insgesamt. Auf der Globalskala im Fach Mathematik sind die geringsten geschlechtsbezogenen Disparitäten in Schleswig-Holstein zu finden, während sie in Berlin, Bremen, Hamburg und Thüringen besonders groß ausfallen.

Vergleicht man die Ergebnisse für das Jahr 2016 mit den entsprechenden Befunden aus dem Jahr 2011, so zeigt sich sowohl für Deutschland insgesamt als auch auf Länderebene, dass sich die geschlechtsbezogenen Disparitäten in den untersuchten Kompetenzen im genannten Fünfjahreszeitraum nicht signifikant verändert haben. Die einzige Ausnahme hiervon ist der Kompetenzbereich *Zuhören* im Fach Deutsch: In diesem Bereich hat sich zwischen den Jahren 2011 und 2016 der Unterschied in den Kompetenzmittelwerten für Jungen und Mädchen in Deutschland insgesamt um 9 Punkte signifikant erhöht. Auf Länderebene ist dieser Trend jedoch nur in einem Land (Nordrhein-Westfalen) statistisch signifikant (+19 Punkte).

Neben der Entwicklung kognitiver Kompetenzen besteht ein weiteres Ziel des schulischen Unterrichts darin, bestimmte motivationale Aspekte wie das fachbezogene Selbstkonzept und das fachliche Interesse der Kinder zu fördern. Sowohl Jungen als auch Mädchen geben im Durchschnitt ein hohes Interesse für die Fächer Deutsch und Mathematik an. Zudem schätzt eine große Mehrheit in beiden Teilpopulationen die eigenen Kompetenzen in diesen Fächern als hoch ein. Dennoch zeigen sich auch in diesen Einschätzungen im Mittel Unterschiede zwischen Jungen und Mädchen, die in dieselbe Richtung weisen wie die geschlechtsbezogenen Disparitäten in den untersuchten Kompetenzen. Die Ergebnisse im Fach Mathematik können als Hinweis darauf gewertet werden, dass Mädchen ihre Kompetenz in diesem Fach geringer einschätzen als gleich kompetente Jungen.

Insgesamt weisen die in diesem Kapitel dargestellten Befunde darauf hin, dass in der Primarstufe nach wie vor gezielte Förderung erforderlich ist, um geschlechtsbezogene Disparitäten in kognitiven Kompetenzen und motivationalen Merkmalen zu reduzieren.

Literatur

- Blossfeld, H.-P., Bos, W., Hannover, B., Lenzen, D., Müller-Böling, D., Prenzel, M. & Wößmann, L. (2009). *Geschlechterdifferenzen im Bildungssystem. Jahresgutachten 2009 des Aktionsrats Bildung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Böhme, K. & Roppelt, A. (2012). Geschlechtsbezogene Disparitäten. In P. Stanat, G. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 173–190). Münster: Waxmann.
- Böhme, K., Sebald, S., Weirich, S. & Stanat, P. (2016). Geschlechtsbezogene Disparitäten. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 377–408). Münster: Waxmann.
- Bos, W., Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Lesekompetenzen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 91–135). Münster: Waxmann.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576.
- Frey, A., Heinze, A., Mildner, D., Hochweber, J. & Asseburg, R. (2010). Mathematische Kompetenz von PISA 2003 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel & S. Wolfgang (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 153–176). Münster: Waxmann.
- Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3. Aufl.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hammer, S., Reiss, K., Lehner, M. C., Heine, J.-H., Sälzer, C. & Heinze, A. (2016). Mathematische Kompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 219–248). Münster: Waxmann.
- Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In M. Prenzel, C. Sälzer, E. Klieme, & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 217–244). Münster: Waxmann.
- 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.
- Jansen, M., Schroeders, U. & Stanat, P. (2013). Motivationale Schülermerkmale in Mathematik und den Naturwissenschaften. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 348–365). 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.
- Marsh, H. W. & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81, 59–77.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O. & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76, 397–416.
- Mengering, F. (2005). Bärenstark – Empirische Ergebnisse der Berliner Sprachstands-erhebung an Kindern im Vorschulalter. *Zeitschrift für Erziehungswissenschaft*, 8, 256–262.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55, 19–27.
- Murphy, P. & Whitelegg, E. (2006). Girls and physics: continuing barriers to ‘belonging’. *Curriculum Journal. Special Issue: Science learning and citizenship*, 17, 281–305.

- 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.
- 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.
- Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenz im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 47–98). Münster: Waxmann.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schiefele, U. (2009). Motivation. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 151–177). Heidelberg: Springer.
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 120–148.
- Schroeders, U., Penk, C., Jansen, M. & Pant, H. A. (2013). Geschlechtsbezogene Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 249–274). 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.
- 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.
- Weis, M., Zehner, F., Sälzer, C., Strohmaier, A., Artelt, C. & Pfof, M. (2016). Lesekompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. (S. 249–284). Münster: Waxmann.
- Wendt, H., Steinmayr, R. & Kasper, D. (2016). Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 317–332). 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.

Kapitel 8

Soziale Disparitäten

Nicole Haag, Aleksander Kocaj, Malte Jansen und Poldi Kuhl

8.1 Einleitung

Verschiedene Schulleistungsstudien konnten zeigen, dass die schulische Kompetenzentwicklung von Kindern und Jugendlichen in erheblichem Maße mit dem sozialen Status ihrer Eltern zusammenhängt (z. B. Kuhl, Haag, Federlein, Weirich & Schipolowski, 2016; Kuhl, Siegle & Lenski, 2013; Müller & Ehmke, 2016; Richter, Kuhl & Pant, 2012; Stubbe, Schwippert & Wendt, 2016; Wendt, Stubbe & Schwippert, 2012). Schülerinnen und Schüler aus Familien mit einem höheren sozioökonomischen Status erreichen konsistent deutlich höhere Kompetenzwerte als Schülerinnen und Schüler aus Familien mit einem geringeren sozioökonomischen Status.

Der enge Zusammenhang zwischen sozialer Herkunft und schulischen Kompetenzen ist unter anderem darauf zurückzuführen, dass Kinder und Jugendliche in unterschiedlichen sozialen, kulturellen und ökonomischen Verhältnissen aufwachsen. Damit verbunden unterscheiden sich auch die bildungsbezogenen Ressourcen und häuslichen Lerngelegenheiten, die ihnen zur Verfügung stehen. Diese Unterschiede in den Ausgangsvoraussetzungen können zu sozialen Ungleichheiten im weiteren Kompetenzerwerb und Bildungsverlauf führen, die als soziale Disparitäten bezeichnet werden (z. B. Baumert, Stanat & Watermann, 2006; Maaz, Baumert, Gresch & McElvany, 2010).

Soziale Disparitäten im Kompetenzerwerb treten bereits in der Primarstufe auf. In Deutschland kommt ihnen eine besonders große Bedeutung zu, da die in diesem Bildungsabschnitt erreichten Kompetenzen ausschlaggebend für den Übergang in die weiterführende Schule und somit für den weiteren Bildungsweg der Schülerinnen und Schüler sind. In Deutschland zeigen sich deutliche Kompetenzvorsprünge von Viertklässlerinnen und Viertklässlern aus sozial privilegierten Elternhäusern in den Bereichen Mathematik, Naturwissenschaften und Lesekompetenz, die in einer ähnlichen Höhe auch in anderen europäischen Ländern festgestellt werden können (Stubbe et al., 2016; Stubbe, Tarelli & Wendt, 2012; Wendt et al., 2012).

Auch im IQB-Ländervergleich 2011 wurden für Schülerinnen und Schüler aus Familien mit hohem sozioökonomischen Status deutlich höhere Kompetenzwerte in den Bereichen *Lesen, Zuhören* und im Fach Mathematik in der Primarstufe ermittelt (Richter et al., 2012). Die Ausprägung der sozialen Disparitäten unterschied sich zwischen den Ländern, wobei die Disparitäten in den Stadtstaaten Berlin, Bremen und Hamburg sowie in anderen deutschen Großstädten vergleichsweise groß waren (Richter et al., 2012; siehe auch Stubbe, Bos & Hornberg, 2008).

Dieses Kapitel führt die Analyse von Zusammenhängen zwischen der sozialen Herkunft von Viertklässlerinnen und Viertklässlern und den von ihnen erreichten Kompetenzen im Ländervergleich fort. Dazu wird im Folgenden zunächst die Erfassung der sozialen Herkunft im IQB-Bildungstrend 2016 und im IQB-Ländervergleich 2011 beschrieben und es werden die in den Analysen verwendeten Indikatoren sozialer Disparitäten dargestellt (Abschnitt 8.2). Nach Hinweisen zur Stichprobe und der Ergebnisdarstellung (Abschnitt 8.3) werden die Verteilungen des sozioökonomischen Status in den Ländern berichtet (Abschnitt 8.4). Dabei wird zum einen bestimmt, inwieweit sich die Schülerinnen und Schüler zwischen den Ländern in ihrem sozialen Hintergrund unterscheiden. Zum anderen wird untersucht, inwieweit sich der durchschnittliche sozioökonomische Status in den Ländern seit dem IQB-Ländervergleich 2011 verändert hat. Die Abschnitte 8.5 und 8.6 stellen die sozialen Disparitäten in den 16 Ländern für die Fächer Deutsch und Mathematik dar. Dabei werden die Zusammenhänge zwischen den Indikatoren der sozialen Herkunft und den Kompetenzen der Schülerinnen und Schüler ermittelt, um die Ausprägung der sozialen Disparitäten in den einzelnen Ländern für das Jahr 2016 und im Trend zu bestimmen. Das Kapitel schließt in Abschnitt 8.7 mit einer Zusammenfassung und Diskussion der Ergebnisse.

8.2 Indikatoren sozialer Disparitäten

Zur Beschreibung der sozialen Herkunft von Schülerinnen und Schülern werden in der empirischen Bildungsforschung in der Regel die berufliche Qualifikation und die berufliche Tätigkeit der Eltern herangezogen. Die Informationen zur Ausbildung und zum Beruf von Mutter und Vater wurden im IQB-Bildungstrend 2016 ebenso wie im IQB-Ländervergleich 2011 von den Eltern erfragt und anhand der *International Standard Classification of Occupation (ISCO)* kategorisiert (International Labour Office, 2012). Die ISCO-Klassifikation bildet die Basis für die Bildung von Indikatoren für den sozioökonomischen Status, wie etwa den sozioökonomischen Index (ISEI, Ganzeboom, 2010; Ganzeboom, de Graaf, Treiman & de Leeuw, 1992) oder die EGP¹-Klassifikation (Erikson, Goldthorpe & Portocarero, 1979). Da diese beiden Indikatoren zentral für die Analysen in diesem Kapitel sind, werden sie im Folgenden kurz beschrieben.

8.2.1 Sozioökonomischer Index (HISEI) und sozialer Gradient

Der sozioökonomische Index ISEI (Ganzeboom, 2010; Ganzeboom et al., 1992) ist ein international vergleichbarer Index für den Status der beruflichen Tätigkeit Erwachsener unter Berücksichtigung des Einkommens und des Bildungsniveaus. Auf Grundlage des ISEI werden die Berufe anhand einer Skala in eine Rangreihe gebracht, die Werte zwischen 10 Punkten und 89 Punkten annehmen kann. Niedrige Werte charakterisieren Berufe, die mit einem niedrigen sozioökonomischen Status assoziiert sind (z. B. Hilfsarbeitskräfte), während hohe Werte Berufe kennzeichnen, die einem hohen sozioökonomischen Status entsprechen (z. B. leitende Verwaltungsbedienstete und Führungskräfte). Um den sozioökonomischen Status einer Familie abzubilden, wird jeweils der höchste ISEI-Wert der Eltern

1 Das Akronym EGP leitet sich aus den Nachnamen der Autoren Erikson, Goldthorpe und Portocarero ab.

(Mutter oder Vater) herangezogen, der entsprechend als *Highest ISEI* (HISEI) der Familie bezeichnet wird.²

Zur Bestimmung der Ausprägung sozialer Disparitäten im Kompetenzerwerb anhand des HISEI wird in Schulleistungsstudien der sogenannte soziale Gradient geschätzt. Dieser beschreibt den Zusammenhang zwischen dem sozioökonomischen Status der Familie und den erreichten Kompetenzausprägungen der Schülerinnen und Schüler als Steigungskoeffizient einer linearen Regression der Kompetenzwerte auf den HISEI. Dabei wird der HISEI so transformiert, dass er für die Gesamtstichprobe der Schülerinnen und Schüler in Deutschland einen Mittelwert (M) von 0 und eine Standardabweichung (SD) von 1 aufweist. So ist es möglich, direkt an der Steigung des sozialen Gradienten (b) abzulesen, wie viele Punkte auf der Berichtsmetrik Schülerinnen und Schüler durchschnittlich mehr erzielen würden, wenn der HISEI ihrer Familie um eine Standardabweichung höher wäre.³

8.2.2 EGP-Klassifikation

Als weiterer Indikator der sozialen Herkunft wird die EGP-Klassifikation herangezogen, die neben dem Beruf der Eltern (ISCO-Klassifikation) und der dafür nötigen Ausbildung weitere Merkmale der beruflichen Tätigkeit berücksichtigt. Auf dieser Grundlage können qualitative Unterschiede zwischen verschiedenen Berufsgruppen anschaulicher dargestellt werden. Die EGP-Klassifikation bestimmt den sozioökonomischen Status verschiedener sozialer Gruppen nach Art der Tätigkeit (manuell, nichtmanuell, landwirtschaftlich), der Stellung im Beruf (selbstständig, abhängig beschäftigt), der Weisungsbefugnis (keine, niedrige, hohe) sowie der erforderlichen Qualifikationen (keine, niedrige, hohe). Anhand dieser Merkmale lassen sich sieben Gruppen von der „oberen Dienstklasse“ (EGP-Klasse I) bis zu „un- und angelernten Arbeitern“ (EGP-Klasse VII) unterscheiden. In Anlehnung an das Vorgehen im Rahmen von PISA 2006 (Ehmke & Baumert, 2008) und der früheren IQB-Ländervergleichsstudien (Knigge & Leucht, 2010; Kuhl et al., 2016; Kuhl et al., 2013; Richter et al., 2012) wird im IQB-Bildungstrend 2016 jedem Kind die EGP-Klasse zugeordnet, die dem höchsten EGP-Wert der Eltern (Mutter oder Vater) entspricht.

Zur effizienten und anschaulichen Darstellung werden jedoch nicht die Werte aller EGP-Klassen berichtet, sondern Extremgruppen gebildet, um so die erreichten Kompetenzen von Heranwachsenden vergleichen zu können, deren Eltern hoch bzw. niedrig qualifizierten Berufsgruppen zuzuordnen sind. Die erste Extremgruppe fasst Viertklässlerinnen und Viertklässler zu einer Gruppe zusammen, deren Eltern beispielsweise in freien akademischen Berufen beziehungsweise als führende Angestellte oder Beamte im höheren Dienst arbeiten und deren Tätigkeiten durch hohe Verantwortung, Entscheidungsbefugnis sowie Autonomie gekennzeichnet sind (EGP-Klassen I–II). Die zweite EGP-Extremgruppe (EGP-

2 Die Vorschrift zur Bildung des ISEI wurde seit dem IQB-Ländervergleich 2011 überarbeitet. Damit die Ergebnisse für das Jahr 2011 direkt mit den Ergebnissen für das Jahr 2016 vergleichbar sind, wurde der ISEI für das Jahr 2011 anhand der aktualisierten Vorschrift neu gebildet. Die in diesem Kapitel berichteten Werte für die Verteilung des HISEI in den Ländern und für die sozialen Gradienten im Jahr 2011 können daher von den im Berichtsband zum IQB-Ländervergleich 2011 (Richter et al., 2012) dargestellten Werten abweichen.

3 Im Folgenden wird die Bezeichnung „sozialer Gradient“ synonym mit der Bezeichnung „Steigung des sozialen Gradienten“ verwendet.

Klassen V–VII) umfasst Schülerinnen und Schüler, deren Eltern beispielsweise un- und angelernte Arbeiter, Vorarbeiter, Meister oder Techniker in manuellen Arbeitsprozessen sind oder Dienstleistungstätigkeiten mit weitgehend manuellem Charakter und geringem Anforderungsniveau ausüben.

8.3 Anmerkungen zur Ergebnisdarstellung

Wie in großen Schulleistungsstudien üblich, wurden fehlende Werte im HISEI und in den EGP-Klassen durch multiple Imputationen ersetzt (Kuhl et al., 2016; Kuhl et al., 2013; Müller & Ehmke, 2013; Richter et al., 2012; vgl. Kapitel 13.1). Die Informationen zum Beruf und zur beruflichen Stellung der Eltern wurden im IQB-Bildungstrend 2016 ausschließlich von den Eltern erfragt, da nicht davon ausgegangen werden kann, dass alle Schülerinnen und Schüler in der Primarstufe diese Informationen zuverlässig angeben können. Aufgrund der unterschiedlichen Rücklaufquoten des Elternfragebogens in den Ländern fehlen diese Informationen teilweise für einen erheblichen Teil der Schülerinnen und Schüler (vgl. Kapitel 3.1).

Die Schätzung sozialer Disparitäten ist allgemein mit einem gewissen Grad an Unsicherheit behaftet, der mit steigendem Anteil fehlender Werte größer wird. Daher muss für Länder, in denen ein großer Anteil der Daten fehlt, davon ausgegangen werden, dass die vorliegenden Daten verzerrte Schätzungen liefern. Gemäß einem von der Amtschefskommission „Qualitätssicherung in Schulen“ in ihrer 74. Sitzung beschlossenen Vorgehen sind Ergebnisse zu sozialen und zuwanderungsbezogenen Disparitäten im IQB-Bildungstrend 2016 unter Vorbehalt zu berichten, wenn die erforderlichen Angaben für weniger als 80 Prozent, jedoch für mindestens 70 Prozent der Schülerinnen und Schüler eines Landes vorliegen. Liegen für weniger als 70 Prozent der Schülerinnen und Schüler Angaben vor, so werden für die betreffenden Länder keine Ergebnisse berichtet (vgl. Kapitel 3.1).

Da in Berlin, Bremen, Hamburg und im Saarland für über 30 Prozent der Schülerinnen und Schüler keine Angaben zum sozialen Hintergrund ihrer Familien vorliegen, können für diese Länder folglich keine sozialen Disparitäten berichtet werden. In Baden-Württemberg, Bayern, Brandenburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen, Sachsen-Anhalt und Schleswig-Holstein fehlen die entsprechenden Angaben für 20 bis 30 Prozent der Schülerinnen und Schüler. Die Ergebnisse dieser Länder stehen daher unter Vorbehalt und werden in den Tabellen entsprechend gekennzeichnet. Da die sozialen Disparitäten in der Mehrzahl der Länder unter Vorbehalt berichtet werden, wird bei der Ergebnisdarstellung in den Tabellen und Abbildungen auf eine Sortierung der Länder nach den Werten der sozialen Indikatoren beziehungsweise Disparitäten verzichtet. Die Länder sind in den Tabellen und Abbildungen stattdessen immer alphabetisch geordnet.

8.4 Verteilung des sozioökonomischen Status der Schülerinnen und Schüler

Abbildung 8.1 zeigt die Verteilung des sozioökonomischen Status der Viertklässlerinnen und Viertklässler in Deutschland insgesamt sowie in den einzelnen Ländern. Im linken Teil der Abbildung sind der Mittelwert (*M*), der Standardfehler des Mittelwerts (*SE*) sowie die Standardabweichung (*SD*) und der Standardfehler der Standardabweichung (*SE*) der Standardabweichung des HISEI angegeben. Im rechten Teil werden die Verteilungen des HISEI anhand von Perzentilbändern dargestellt (vgl. Kapitel 3.2).

Der mittlere HISEI beträgt im Jahr 2016 in Deutschland 50.6 Punkte. Die Ländermittelwerte bewegen sich zwischen 52.2 Punkten in Hessen und 46.5 Punkten in Sachsen-Anhalt, wobei der HISEI in Mecklenburg-Vorpommern, Sachsen-Anhalt und Thüringen signifikant unter dem Bundesdurchschnitt liegt. Die Standardabweichung des HISEI liegt in Deutschland insgesamt bei 20.7 Punkten. Die Streuungen des HISEI sind in Brandenburg, Mecklenburg-Vorpommern, Sachsen-Anhalt und Thüringen signifikant geringer ausgeprägt als bundesweit. Familien innerhalb dieser Länder unterscheiden sich also signifikant weniger stark im sozioökonomischen Status als Familien in Deutschland insgesamt.

Abbildung 8.1: Verteilung des sozioökonomischen Status (HISEI) der Familien von Viertklässlerinnen und Viertklässlern in den Ländern

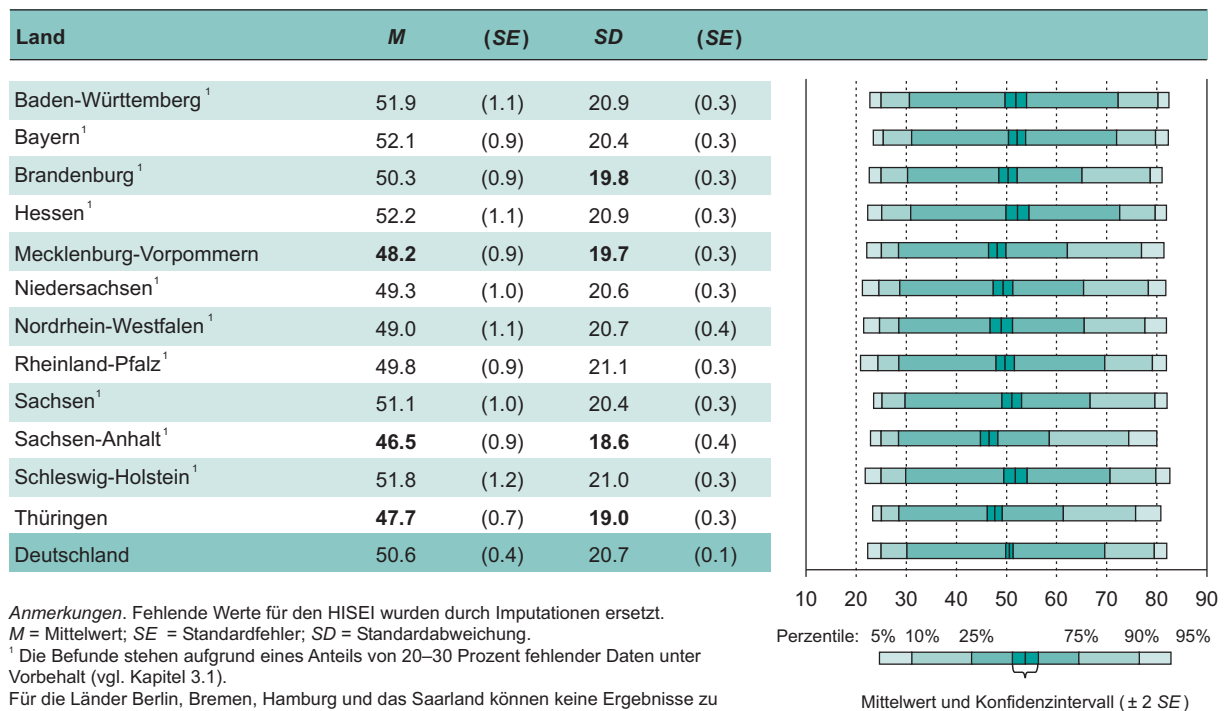


Abbildung 8.2: Veränderungen des sozioökonomischen Status (HISEI) in den Ländern zwischen den Jahren 2011 und 2016

Land	2011			2016			Differenz				Differenz der Mittelwerte 2016–2011	
	M	(SE)	SD	M	(SE)	SD	ΔM	(SE)	ΔSD	(SE)		d
Baden-Württemberg ¹	54.3	(1.0)	20.6	51.9	(1.1)	20.9	-2.4	(1.5)	0.2	(0.4)	-0.08	
Bayern ¹	51.7	(1.2)	20.4	52.1	(0.9)	20.4	0.5	(1.4)	0.0	(0.5)	0.02	
Brandenburg ¹	50.5	(1.0)	19.6	50.3	(0.9)	19.8	-0.1	(1.4)	0.1	(0.5)	-0.01	
Hessen ¹	51.1	(1.0)	20.9	52.2	(1.1)	20.9	1.1	(1.5)	0.0	(0.4)	0.04	
Mecklenburg-Vorpommern	49.9	(1.0)	19.9	48.2	(0.9)	19.7	-1.8	(1.3)	-0.2	(0.5)	-0.06	
Niedersachsen ¹	49.6	(0.7)	19.6	49.3	(1.0)	20.6	-0.2	(1.2)	1.0	(0.4)	-0.01	
Nordrhein-Westfalen ¹	49.6	(1.1)	20.2	49.0	(1.1)	20.7	-0.6	(1.5)	0.5	(0.5)	-0.02	
Rheinland-Pfalz ¹	49.8	(0.9)	19.8	49.8	(0.9)	21.1	-0.1	(1.3)	1.3	(0.4)	0.00	
Sachsen ¹	49.2	(0.9)	18.7	51.1	(1.0)	20.4	1.9	(1.3)	1.7	(0.5)	0.07	
Sachsen-Anhalt ¹	47.1	(0.7)	19.2	46.5	(0.9)	18.6	-0.6	(1.1)	-0.6	(0.5)	-0.02	
Schleswig-Holstein ¹	51.1	(1.0)	20.1	51.8	(1.2)	21.0	0.6	(1.5)	0.9	(0.4)	0.02	
Thüringen	47.3	(1.1)	19.2	47.7	(0.7)	19.0	0.3	(1.3)	-0.2	(0.4)	0.01	
Deutschland	50.8	(0.4)	20.3	50.6	(0.4)	20.7	-0.2	(0.5)	0.4	(0.2)	-0.01	

Anmerkungen. M = Mittelwert; SE = Standardfehler des Mittelwerts; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.

¹ Die Befunde stehen aufgrund eines Anteils von 20–30 Prozent fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1). Für die Länder Berlin, Bremen, Hamburg und das Saarland können keine Ergebnisse zu sozialen Disparitäten berichtet werden, da die erforderlichen Informationen für weniger als 70 Prozent der Schülerinnen und Schüler vorliegen (vgl. Kapitel 3.1). Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

■ Signifikant ($p < .05$) von Null abweichende Differenz
 □ Nicht signifikant von Null abweichende Differenz

Abbildung 8.2 zeigt die Veränderungen des mittleren HISEI sowie der Streuung des HISEI zwischen den Jahren 2011 und 2016. Sowohl in Deutschland insgesamt als auch in den Ländern sind die mittleren HISEI-Werte stabil geblieben. Im betrachteten Zeitraum zeigen sich lediglich geringe Veränderungen des mittleren HISEI von bis zu etwa 2 Punkten, die jedoch in keinem Land statistisch signifikant sind. Die Streuung des HISEI ist in Deutschland insgesamt signifikant angestiegen, mit weniger als einem Punkt ist der Zuwachs in der Heterogenität jedoch nicht substantiell. Auch in den Ländern Niedersachsen, Rheinland-Pfalz, Sachsen und Schleswig-Holstein haben sich die Streuungen seit dem Jahr 2011 signifikant vergrößert, wobei der Anstieg hier ebenfalls jeweils gering ist. Diese nur geringen Zunahmen der Standardabweichungen des HISEI weisen darauf hin, dass sich die Heterogenität des sozioökonomischen Status weder in Deutschland noch in den Ländern substantiell verändert hat.

Neben dem mittleren Niveau sowie der Streuung des sozioökonomischen Status in Deutschland und in den einzelnen Ländern wird im IQB-Bildungstrend 2016 erstmals auch untersucht, wie sich Schülerinnen und Schüler mit unterschiedlichem sozioökonomischen Status über die Schulen verteilen. Diese Informationen zur sozioökonomischen Heterogenität der Schülerschaft in den Schulen bilden einen Hintergrund zur Interpretation der in Kapitel 6 berichteten Analysen zur Heterogenität in den erreichten Kompetenzen. Analysen zum Zusammenhang zwischen der sozioökonomischen Heterogenität und der Heterogenität der erreichten Kompetenzen in den Ländern werden im IQB-Bildungstrend 2016 jedoch nicht vorgelegt.

Um die Unterschiede in der sozialen Zusammensetzung zwischen Schulen zu bestimmen, wird die Varianz der HISEI-Werte zerlegt in einen Anteil, der zwischen einzelnen Schulen besteht und in einen Anteil, der innerhalb von Schulen besteht. Hierzu werden analog zu Kapitel 6 Intraklassenkorrelationen berechnet, die angeben, welcher Anteil der Gesamtvarianz des HISEI auf Unterschiede zwischen Schulen zurückzuführen ist (vgl. Kapitel 3.2). Multipliziert man die Intraklassenkorrelation mit 100, ergibt dies den prozentualen Anteil der Varianz im sozioökonomischen Status zwischen Schulen an der Gesamtvarianz, die im jeweiligen Land vorliegt.

Abbildung 8.3: Varianz des sozioökonomischen Status (HISEI) innerhalb und zwischen Schulen in den Jahren 2011 und 2016

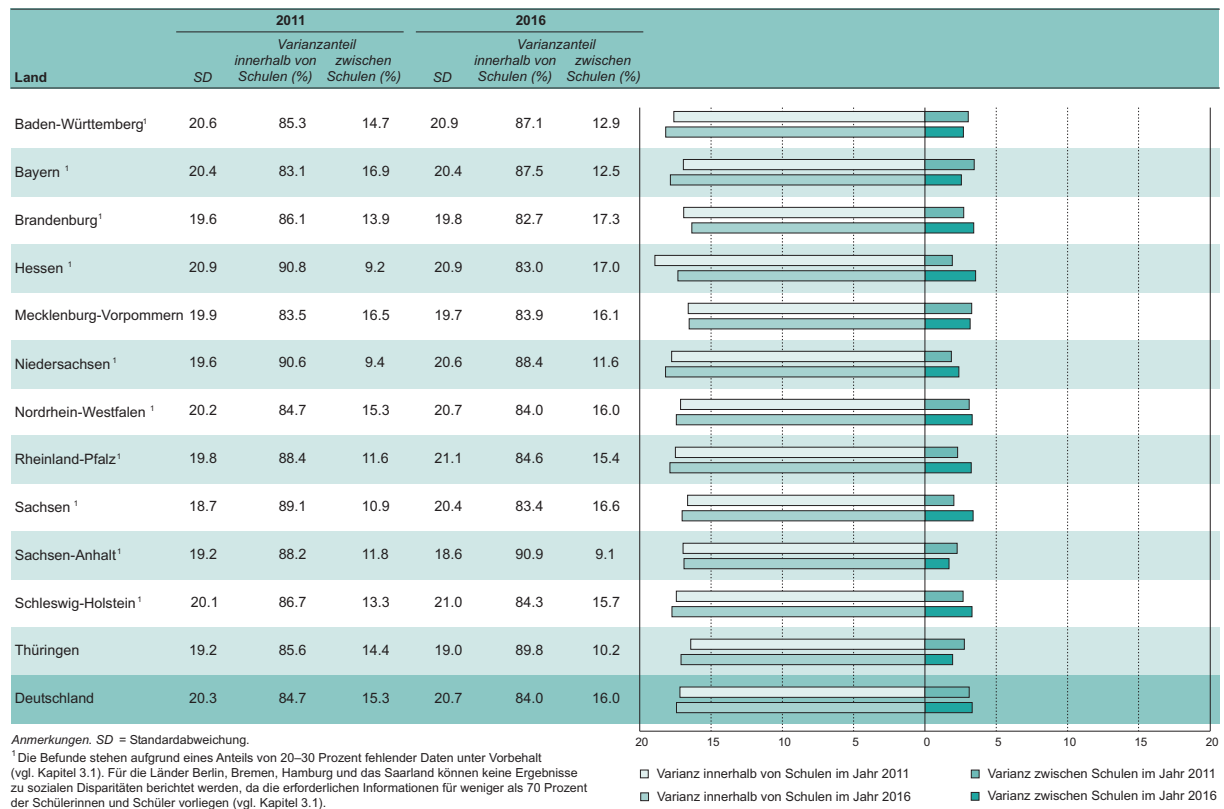


Abbildung 8.3 zeigt die Heterogenität des HISEI innerhalb von Schulen und zwischen Schulen im Trend. Im Tabellenteil der Abbildung sind die Standardabweichungen (*SD*) sowie die zwischen Schulen und innerhalb von Schulen liegenden Varianzanteile der HISEI-Werte für die Jahre 2011 und 2016 angegeben. Der rechte Teil der Abbildung zeigt die Gesamtvarianz des sozioökonomischen Status sowie ihre Zerlegung in einen Teil zwischen Schulen und einen Teil innerhalb von Schulen zu beiden Erhebungszeitpunkten. Die Gesamtlänge der Balken gibt dabei die Streuung des HISEI im Jahr 2011 beziehungsweise 2016 im jeweiligen Land an. Je länger der Balken eines Landes insgesamt ist, desto größere soziale Unterschiede bestehen insgesamt zwischen Familien. Der nach rechts weisende, dunklere Teil des Balkens verdeutlicht den Varianzanteil, der zwischen den einzelnen Schulen innerhalb eines Landes besteht, wohingegen der nach links weisende, hellere Teil des Balkens den Anteil der Varianz kennzeichnet, der innerhalb von Schulen besteht. Diese Darstellung ermöglicht es, die Heterogenität des sozialen Status innerhalb von Schulen und zwischen Schulen über die Länder hinweg zu vergleichen. Größere Unterschiede zwischen Schulen können als eine stärkere soziale Trennung der Schülerinnen und Schüler interpretiert werden. Die in diesem Kapitel präsentierten Analysen zur Heterogenität des HISEI innerhalb von Schulen und zwischen Schulen sind jedoch rein deskriptiv, die Unterschiede zwischen Ländern und zwischen den Jahren 2011 und 2016 wurden nicht auf Signifikanz geprüft.

Betrachtet man den HISEI für das Jahr 2016 in Deutschland insgesamt, liegt etwa ein Sechstel der Unterschiede (16 %) im sozialen Status zwischen Schulen. Die Unterschiede im sozialen Status zwischen Schulen sind somit etwas kleiner als die in Kapitel 6 berichteten Unterschiede zwischen Schulen in den erreichten Kompetenzen (*Lesen*: 18.3 %, *Zuhören*: 19.4 %, *Orthografie*: 17.9 %, ...)

Mathematik (*Globalskala*): 17.6 %). In den Ländern bewegen sich die Anteile der zwischenschulischen Unterschiede im HISEI zwischen 9.1 Prozent in Sachsen-Anhalt und 17.3 Prozent in Brandenburg. Diese eher geringen Unterschiede zwischen den Ländern in den Intraklassenkorrelationen und die zwischen den Ländern ebenfalls nur wenig variierenden Standardabweichungen im HISEI weisen insgesamt darauf hin, dass die soziale Heterogenität innerhalb von Schulen über die dargestellten Länder hinweg vergleichsweise ähnlich ist.

Ebenso wie die Streuung des sozialen Status insgesamt sind auch die Anteile der Unterschiede im sozioökonomischen Status, die zwischen den Schulen bestehen, zwischen den Jahren 2011 und 2016 weitgehend konstant geblieben. Lediglich in den Ländern Hessen und Sachsen lässt sich im betrachteten Zeitraum eine etwas stärkere Zunahme des prozentualen Varianzanteils zwischen Schulen um mindestens 5 Prozentpunkte feststellen, die auf steigende Unterschiede im sozioökonomischen Status zwischen Schulen hindeutet.

8.5 Soziale Gradienten

Die Abbildungen 8.4 bis 8.7 zeigen die für die Kompetenzbereiche *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch sowie für die *Globalskala* im Fach Mathematik festgestellten sozialen Gradienten für die Jahre 2011 und 2016. Die in den Abbildungen 8.4 bis 8.7 sowie in Tabelle 8.1 berichteten Ergebnisse umfassen den Achsenabschnitt (a), die Steigung des sozialen Gradienten (b) und den Anteil der durch den sozialen Gradienten aufgeklärten Varianz in den Kompetenzwerten (R^2). Der Achsenabschnitt entspricht dem Kompetenzwert, den diejenigen Schülerinnen und Schüler im Mittel erreicht haben, deren HISEI dem durchschnittlichen HISEI in Deutschland entspricht. Die Steigung des sozialen Gradienten gibt an, um wie viele Punkte die mittlere Leistung von Schülerinnen und Schülern auseinanderliegt, deren HISEI sich um eine Standardabweichung unterscheidet. Hohe Steigungskoeffizienten weisen dabei auf einen engen Zusammenhang zwischen der sozialen Herkunft und den Kompetenzen von Schülerinnen und Schülern hin. Der Unterschied zwischen den für die Länder festgestellten sozialen Gradienten und dem für Deutschland insgesamt gemessenen sozialen Gradienten wird zudem auf Signifikanz geprüft. Die Varianzaufklärung ist ein Maß für den Anteil der Kompetenzunterschiede, der durch die Unterschiede im sozioökonomischen Status der Familien erklärt werden kann.

Neben den für das Jahr 2016 festgestellten Werten umfassen die Abbildungen 8.4 bis 8.7 auch die entsprechenden Werte für das Jahr 2011. Darüber hinaus wird die Veränderung der sozialen Gradienten zwischen den Jahren 2011 und 2016 berichtet (Δb) und auf statistische Signifikanz geprüft. Die Trendanalysen für den Bereich *Orthografie* lassen sich nur auf Bundesebene durchführen (vgl. Kapitel 3.2), da für das Jahr 2011 keine länderspezifische Auswertung möglich ist und somit für die Länder keine Veränderungswerte gegenüber 2016 berichtet werden können. Für diesen Bereich werden daher die sozialen Gradienten in den Ländern für das Jahr 2016 in Tabelle 8.1 und die bundesweite Veränderung der sozialen Gradienten seit dem Jahr 2011 in Abbildung 8.6 dargestellt.

Im Jahr 2016 bestehen in allen Ländern und in Deutschland insgesamt statistisch signifikante Zusammenhänge zwischen der von den Schülerinnen und Schülern erreichten Lesekompetenz und dem HISEI ihrer Familien (Abb. 8.4).

Abbildung 8.4: Soziale Gradienten im Kompetenzbereich *Lesen* im Fach Deutsch in den Jahren 2011 und 2016 im Vergleich

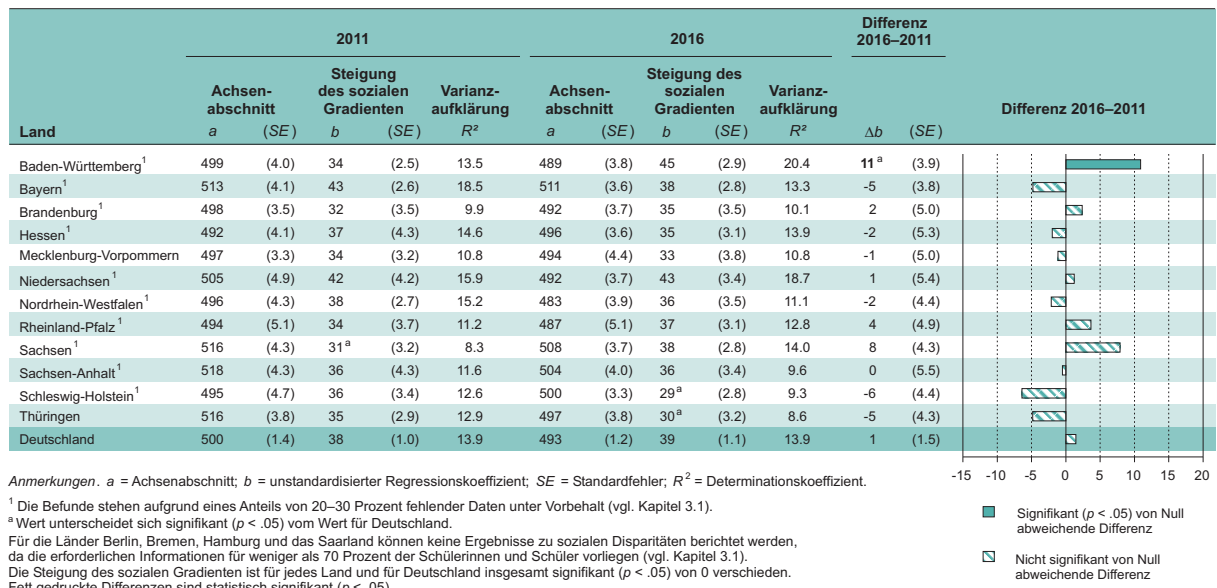
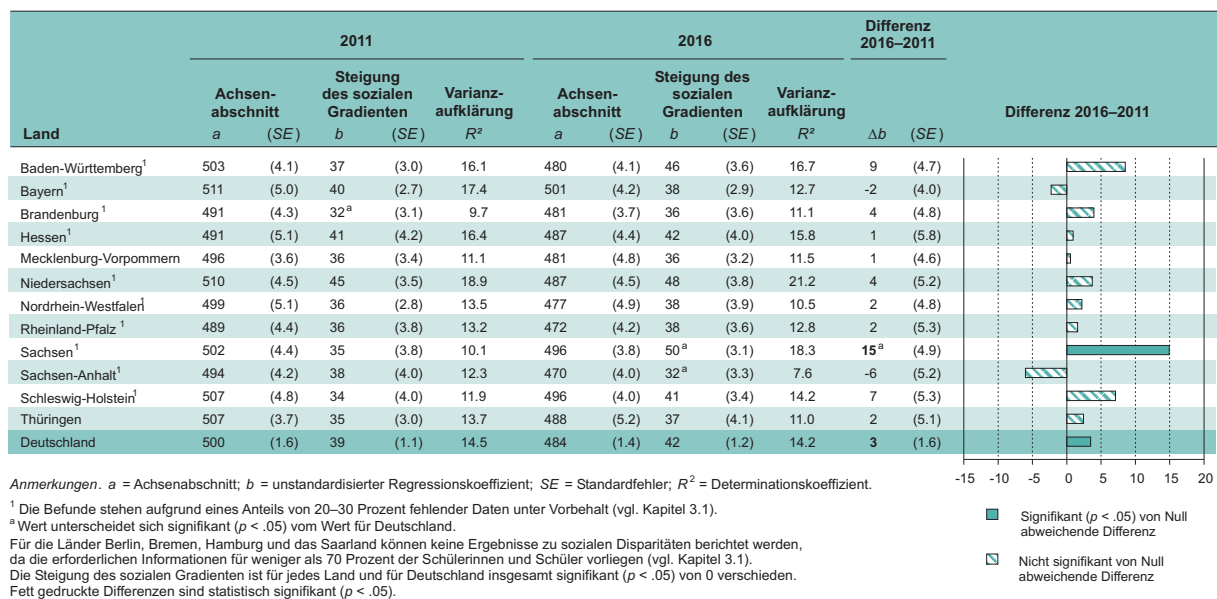


Abbildung 8.5: Soziale Gradienten im Kompetenzbereich *Zuhören* im Fach Deutsch in den Jahren 2011 und 2016 im Vergleich



Die im Bereich *Lesen* erreichten Kompetenzen unterscheiden sich also systematisch in Abhängigkeit vom sozioökonomischen Status der Familie, wobei Schülerinnen und Schüler mit einem höheren sozioökonomischen Status höhere Kompetenzwerte erreichen. Für Deutschland insgesamt beträgt der soziale Gradient im Bereich *Lesen* 39 Punkte und der HISEI erklärt etwa 14 Prozent der Varianz in der Lesekompetenz der Schülerinnen und Schüler. Die sozialen Gradienten variieren zwischen 29 Punkten in Schleswig-Holstein und 45 Punkten in Baden-Württemberg. Dabei fällt der soziale Gradient in Schleswig-Holstein und in Thüringen signifikant niedriger aus als in Deutschland insgesamt.

Vergleicht man die sozialen Gradienten zwischen den Jahren 2011 und 2016, zeigt sich sowohl für Deutschland insgesamt als auch für die einzelnen Länder eine weitgehende Stabilität. Lediglich in Baden-Württemberg ist der so-

ziale Gradient um 11 Punkte gestiegen, was eine signifikante Zunahme sozialer Disparitäten im *Lesen* zeigt.

Auch im Kompetenzbereich *Zuhören* ist der soziale Gradient in allen Ländern und in Deutschland insgesamt positiv und statistisch signifikant (Abb. 8.5). Die erreichten Kompetenzen im *Zuhören* hängen also ebenfalls stark mit dem sozioökonomischen Status der Familie zusammen. Der für das Jahr 2016 in Deutschland insgesamt festgestellte soziale Gradient beträgt in diesem Bereich 42 Punkte. Wie auch im *Lesen* werden im Bereich *Zuhören* deutschlandweit etwa 14 Prozent der Varianz in den erreichten Kompetenzen durch den HISEI erklärt. Die sozialen Gradienten der Länder liegen zwischen 32 Punkten in Sachsen-Anhalt und 50 Punkten in Sachsen. Beide Länder unterscheiden sich damit signifikant vom Wert für Deutschland insgesamt: Während in Sachsen-Anhalt im Vergleich mit dem bundesdeutschen Gradienten besonders geringe Zusammenhänge zwischen der Zuhörkompetenz von Schülerinnen und Schülern und ihrem sozioökonomischen Status bestehen, ist in Sachsen eine besonders enge Kopplung zwischen den in diesem Bereich erzielten Kompetenzwerten und dem sozialen Status zu erkennen.

Tabelle 8.1: Soziale Gradienten im Kompetenzbereich *Orthografie* im Fach Deutsch in den Ländern für das Jahr 2016

Land	Achsenabschnitt		Steigung des sozialen Gradienten		Varianzaufklärung
	<i>a</i>	(<i>SE</i>)	<i>b</i>	(<i>SE</i>)	<i>R</i> ²
Baden-Württemberg ¹	503	(4.0)	31	(3.4)	9.2
Bayern ¹	532	(3.5)	31	(2.9)	9.3
Brandenburg ¹	491	(5.1)	31	(3.7)	8.1
Hessen ¹	503	(4.9)	32	(3.1)	11.8
Mecklenburg-Vorpommern	495	(5.2)	30	(4.2)	7.8
Niedersachsen ¹	483	(4.0)	36	(3.2)	13.6
Nordrhein-Westfalen ¹	494	(3.8)	30	(3.4)	8.2
Rheinland-Pfalz ¹	495	(5.4)	25	(3.2)	6.2
Sachsen ¹	502	(6.0)	27	(4.0)	6.6
Sachsen-Anhalt ¹	504	(4.3)	27	(3.7)	5.4
Schleswig-Holstein ¹	501	(4.5)	27	(3.4)	7.0
Thüringen	495	(3.6)	34	(3.2)	9.9
Deutschland	500	(1.3)	31	(1.2)	8.9

Anmerkungen. *a* = Achsenabschnitt; *b* = unstandardisierter Regressionskoeffizient; *SE* = Standardfehler; *R*² = Determinationskoeffizient.

¹ Die Befunde stehen aufgrund eines Anteils von 20–30 Prozent fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

Für die Länder Berlin, Bremen, Hamburg und das Saarland können keine Ergebnisse zu sozialen Disparitäten berichtet werden, da die erforderlichen Informationen für weniger als 70 Prozent der Schülerinnen und Schüler vorliegen (vgl. Kapitel 3.1).

Die Steigung des sozialen Gradienten ist für jedes Land und für Deutschland insgesamt signifikant ($p < .05$) von 0 verschieden.

Abbildung 8.6: Soziale Gradienten im Kompetenzbereich *Orthografie* im Fach Deutsch in den Jahren 2011 und 2016 im Vergleich

Land	2011			2016			Differenz 2016–2011		Differenz 2016–2011
	Achsenabschnitt <i>a</i>	Steigung des sozialen Gradienten <i>b</i>	Varianzaufklärung <i>R</i> ²	Achsenabschnitt <i>a</i>	Steigung des sozialen Gradienten <i>b</i>	Varianzaufklärung <i>R</i> ²	Δb	(SE)	
Deutschland	530 (2.5)	26 (1.7)	9.1	507 (1.6)	28 (1.2)	8.5	1	(2.1)	

Anmerkungen. Mittelwerte wurden ohne Einbeziehung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf berechnet.
a = Achsenabschnitt; *b* = unstandardisierter Regressionskoeffizient; *SE* = Standardfehler; *R*² = Determinationskoeffizient.
 Die Steigung des sozialen Gradienten ist für Deutschland insgesamt signifikant (*p* < .05) von 0 verschieden.

■ Signifikant (*p* < .05) von Null abweichende Differenz
 □ Nicht signifikant von Null abweichende Differenz

Im Bereich *Zuhören* stieg der soziale Gradient zwischen den Jahren 2011 und 2016 leicht, jedoch statistisch signifikant an (+3 Punkte). In den Ländern zeigt sich nur in Sachsen ein signifikanter Anstieg sozialer Disparitäten um 15 Punkte.

Tabelle 8.1 zeigt die für das Jahr 2016 festgestellten sozialen Gradienten für den Bereich *Orthografie* im Fach Deutsch, die ebenfalls in allen untersuchten Ländern statistisch signifikant sind. Der soziale Gradient für Deutschland fällt mit 31 Punkten hier etwas geringer aus als in den Bereichen *Lesen* und *Zuhören*, was darauf hindeutet, dass die im Bereich *Orthografie* erworbenen Kompetenzen etwas weniger stark vom sozioökonomischen Status der Familien abhängen. Deutschlandweit werden etwa 9 Prozent der Varianz durch den HISEI erklärt. Auch die Länderunterschiede in der Höhe der sozialen Gradienten sind etwas geringer ausgeprägt; die entsprechenden Werte liegen zwischen 25 Punkten in Rheinland-Pfalz und 36 Punkten in Niedersachsen. Die Unterschiede zum deutschen Mittelwert erreichen aber für kein Land statistische Signifikanz, was auf vergleichbare Zusammenhänge zwischen dem sozialen Status und den erreichten Kompetenzen in allen Ländern hinweist.

Für den Kompetenzbereich *Orthografie* sind die bundesweiten Veränderungen der sozialen Disparitäten zwischen den Jahren 2011 und 2016 in Abbildung 8.6 dargestellt. Ähnlich wie im *Lesen* zeigt sich auch in der *Orthografie* nur eine geringe, statistisch nicht signifikante Veränderung des sozialen Gradienten. Wie in Kapitel 3.2 erläutert, können die länderspezifischen Veränderungen nicht bestimmt werden.

Für die Globalskala im Fach Mathematik liegt der soziale Gradient im Jahr 2016 für Deutschland insgesamt bei 41 Punkten (Abb. 8.7), wobei durch den HISEI 13 Prozent der Varianz in den Kompetenzen aufgeklärt wird. Auch im Fach Mathematik bestehen in allen Ländern statistisch signifikante Zusammenhänge zwischen dem sozialen Status und den erreichten Kompetenzen. Die sozialen Gradienten der Länder variieren zwischen 35 Punkten in Mecklenburg-Vorpommern und 44 Punkten in Baden-Württemberg und in Sachsen, weichen jedoch nicht signifikant vom Wert für Deutschland insgesamt ab.

Im Fach Mathematik zeigt sich zwischen den Jahren 2011 und 2016 für Deutschland insgesamt ein statistisch signifikanter, jedoch geringer Anstieg des sozialen Gradienten um 4 Punkte, der auf eine leichte Zunahme der sozialen Disparitäten hinweist. Die Veränderungen der sozialen Gradienten in den einzelnen Ländern sind im Fach Mathematik jedoch alle nicht statistisch signifikant.

Abbildung 8.7: Soziale Gradienten im Fach Mathematik (*Globalskala*) in den Jahren 2011 und 2016 im Vergleich

Land	2011			2016			Differenz 2016–2011		Differenz 2016–2011
	Achsenabschnitt <i>a</i>	Steigung des sozialen Gradienten <i>b</i>	Varianz-aufklärung <i>R</i> ²	Achsenabschnitt <i>a</i>	Steigung des sozialen Gradienten <i>b</i>	Varianz-aufklärung <i>R</i> ²	Δb	(SE)	
Baden-Württemberg ¹	506 (4.0)	39 (2.7)	16.6	483 (3.7)	44 (3.5)	18.8	5 (4.4)		
Bayern ¹	517 (4.2)	36 (3.0)	14.8	505 (2.8)	36 (2.6)	11.0	0 (3.9)		
Brandenburg ¹	491 (3.2)	29 ^a (3.5)	8.1	484 (4.3)	38 (3.5)	10.9	9 (5.0)		
Hessen ¹	483 (3.8)	32 (6.2)	10.6	482 (4.4)	37 (3.4)	14.4	5 (7.1)		
Mecklenburg-Vorpommern	496 (4.5)	34 (3.3)	9.8	481 (5.1)	35 (4.7)	11.6	1 (5.7)		
Niedersachsen ¹	498 (4.1)	38 (3.5)	14.2	475 (3.1)	43 (3.4)	18.7	6 (4.9)		
Nordrhein-Westfalen ¹	499 (4.0)	36 (2.4)	14.0	474 (3.9)	39 (3.2)	10.4	3 (4.0)		
Rheinland-Pfalz ¹	496 (4.4)	37 (3.4)	13.7	480 (5.9)	38 (3.0)	11.6	1 (4.6)		
Sachsen ¹	520 (3.7)	36 (3.3)	11.0	506 (4.0)	44 (2.8)	14.5	8 (4.4)		
Sachsen-Anhalt ¹	523 (4.2)	36 (3.5)	10.9	504 (3.9)	43 (4.4)	11.5	7 (5.6)		
Schleswig-Holstein ¹	486 (3.7)	34 (3.1)	12.6	484 (3.7)	37 (3.2)	13.0	3 (4.5)		
Thüringen	507 (5.2)	31 (3.5)	9.4	493 (3.2)	38 (3.7)	11.6	7 (5.1)		
Deutschland	500 (1.4)	36 (1.2)	12.8	483 (1.3)	41 (1.1)	13.0	4 (1.6)		

Anmerkungen. *a* = Achsenabschnitt; *b* = unstandardisierter Regressionskoeffizient; SE = Standardfehler; *R*² = Determinationskoeffizient.

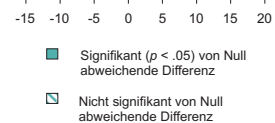
¹ Die Befunde stehen aufgrund eines Anteils von 20–30 Prozent fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

^a Wert unterscheidet sich signifikant ($p < .05$) vom Wert für Deutschland.

Für die Länder Berlin, Bremen, Hamburg und das Saarland können keine Ergebnisse zu sozialen Disparitäten berichtet werden, da die erforderlichen Informationen für weniger als 70 Prozent der Schülerinnen und Schüler vorliegen (vgl. Kapitel 3.1).

Die Steigung des sozialen Gradienten ist für jedes Land und für Deutschland insgesamt signifikant ($p < .05$) von 0 verschieden.

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

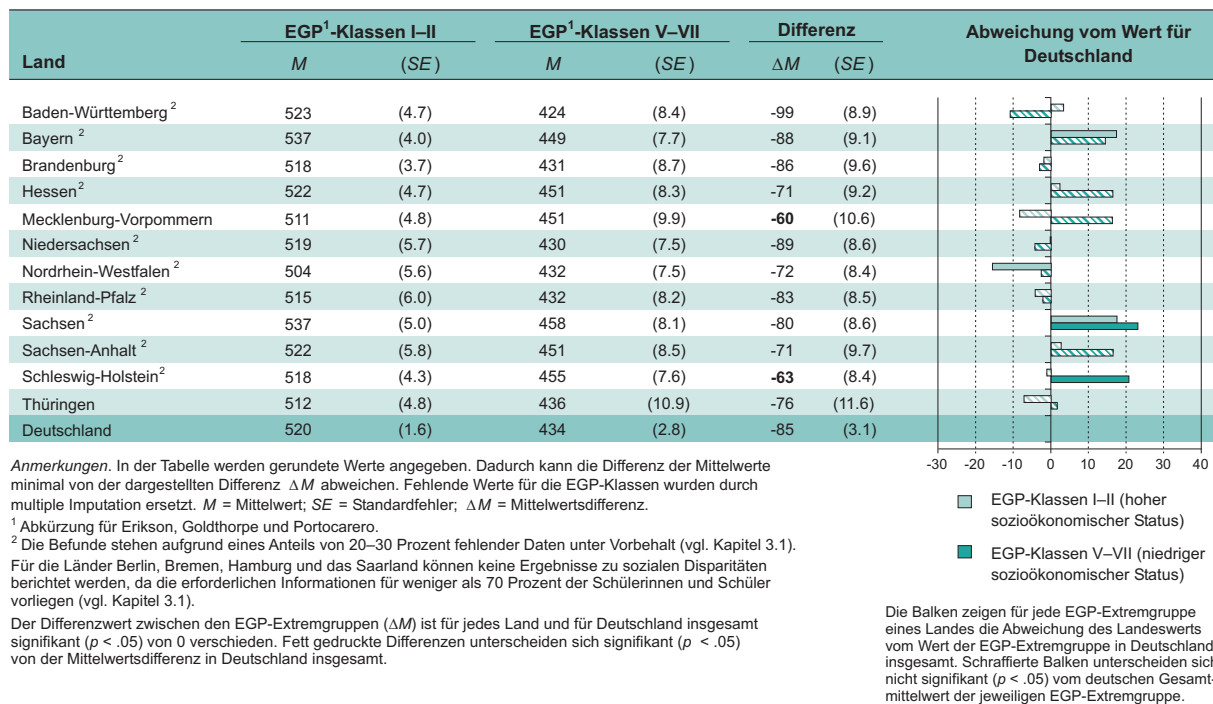


8.6 Kompetenzniveau nach EGP-Extremgruppen

Zusätzlich zu den sozialen Gradienten wird die EGP-Klassifikation als Grundlage für die Beschreibung sozialer Disparitäten herangezogen. Dabei werden aus den sieben EGP-Klassen zwei Extremgruppen gebildet und miteinander verglichen. Wie bereits in Abschnitt 8.2.2 erläutert, umfasst die erste Extremgruppe Schülerinnen und Schüler aus Familien mit einem hohen sozioökonomischen Status (EGP-Klassen I–II) und die zweite Extremgruppe Schülerinnen und Schüler aus Familien mit einem niedrigen sozioökonomischen Status (EGP-Klassen V–VII). Die Abbildungen 8.8 bis 8.11 enthalten die von Schülerinnen und Schülern dieser Gruppen im Jahr 2016 jeweils erreichten mittleren Kompetenzen und die Differenzen zwischen den Gruppen für die Bereiche *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch sowie für die *Globalskala* im Fach Mathematik.

Um ein genaueres Bild davon zu erhalten, inwieweit es in den verschiedenen Ländern gelingt, günstige Bedingungen für den Kompetenzerwerb sowohl für Kinder aus sozial schwachen als auch für Kinder aus sozial besser gestellten Familien zu schaffen, werden im rechten Teil der Abbildung die Ländermittelwerte der beiden EGP-Extremgruppen mit dem deutschen Gesamtmittelwert der jeweiligen EGP-Extremgruppe verglichen. Nach rechts weisende ausgefüllte Balken zeigen an, dass in diesem Land Schülerinnen und Schüler der entsprechenden EGP-Extremgruppe im Vergleich zum Gesamtwert dieser EGP-Extremgruppe in Deutschland im Mittel signifikant höhere Kompetenzwerte erzielen. Nach links weisende ausgefüllte Balken zeigen hingegen signifikant geringere Kompetenzwerte in dieser EGP-Extremgruppe an. Schraffierte Balken unterscheiden sich nicht signifikant vom Mittelwert für Deutschland insgesamt.

Abbildung 8.8: Vergleich der durchschnittlichen Kompetenzwerte der EGP-Extremgruppen im Kompetenzbereich *Lesen* im Fach Deutsch für das im Jahr 2016



8.6.1 Kompetenzniveau nach EGP-Extremgruppen im Jahr 2016

Im Kompetenzbereich *Lesen* unterscheiden sich die in Deutschland im Mittel erreichten Kompetenzen zwischen den EGP-Extremgruppen um 85 Punkte zugunsten von Schülerinnen und Schülern aus sozial privilegierten Familien (Abb. 8.8). Vergleicht man diesen Unterschied mit dem Kompetenzzuwachs im *Lesen*, der zwischen der 3. und 4. Jahrgangsstufe etwa 60 Punkte umfasst (Bremerich-Vos & Böhme, 2009; vgl. auch Kapitel 6.1), entspricht die Differenz der EGP-Extremgruppen ungefähr dem im Zeitraum von anderthalb Schuljahren zu erwartenden Lernzuwachs. Die Differenz fällt in Mecklenburg-Vorpommern und in Schleswig-Holstein signifikant geringer aus als im bundesdeutschen Durchschnitt. Signifikant höhere Unterschiede der EGP-Gruppen als in Deutschland insgesamt ergeben sich hingegen für kein Land.

Vergleicht man die im Mittel erreichten Kompetenzen der EGP-Extremgruppen mit dem jeweiligen Wert dieser Gruppen für Deutschland insgesamt, zeigen sich einige Unterschiede zwischen den Ländern. In Sachsen gelingt es Schülerinnen und Schülern aus beiden EGP-Extremgruppen, überdurchschnittliche Kompetenzwerte zu erzielen. In Bayern, Nordrhein-Westfalen und Schleswig-Holstein hingegen weichen die Kompetenzmittelwerte jeweils nur für eine der beiden EGP-Extremgruppen vom deutschen Durchschnittswert ab. Schülerinnen und Schüler aus sozial privilegierten Familien erreichen in Bayern überdurchschnittliche Kompetenzwerte, in Nordrhein-Westfalen hingegen unterdurchschnittliche Werte. In Schleswig-Holstein schneiden Schülerinnen und Schüler mit niedrigem sozioökonomischen Status besser ab als der bundesdeutsche Durchschnitt, was in diesem Land zur besonders geringen Differenz der beiden EGP-Extremgruppen beiträgt.

Abbildung 8.9: Vergleich der durchschnittlichen Kompetenzwerte der EGP-Extremgruppen im Kompetenzbereich *Zuhören* im Fach Deutsch für das Jahr 2016

Land	EGP ¹ -Klassen I–II		EGP ¹ -Klassen V–VII		Differenz		Abweichung vom Wert für Deutschland
	M	(SE)	M	(SE)	ΔM	(SE)	
Baden-Württemberg ²	517	(5.0)	409	(9.1)	-108	(10.0)	
Bayern ²	526	(4.2)	442	(8.2)	-84	(8.7)	
Brandenburg ²	507	(5.1)	409	(8.9)	-98	(10.7)	
Hessen ²	518	(6.1)	432	(9.4)	-86	(11.1)	
Mecklenburg-Vorpommern	500	(5.8)	416	(9.8)	-84	(10.7)	
Niedersachsen ²	519	(5.6)	412	(9.5)	-107	(10.1)	
Nordrhein-Westfalen ²	502	(6.4)	420	(9.3)	-82	(9.5)	
Rheinland-Pfalz ²	500	(5.4)	411	(7.8)	-88	(8.9)	
Sachsen ²	534	(5.5)	433	(8.9)	-101	(10.8)	
Sachsen-Anhalt ²	487	(5.1)	425	(9.0)	-63	(9.8)	
Schleswig-Holstein ²	526	(5.0)	434	(9.1)	-92	(9.8)	
Thüringen	508	(6.1)	424	(13.1)	-84	(13.6)	
Deutschland	514	(1.8)	419	(3.2)	-95	(3.4)	

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. Fehlende Werte für die EGP-Klassen wurden durch multiple Imputation ersetzt. M = Mittelwert; SE = Standardfehler; ΔM = Mittelwertsdifferenz.

¹ Abkürzung für Erikson, Goldthorpe und Portocarero.

² Die Befunde stehen aufgrund eines Anteils von 20–30 Prozent fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

Für die Länder Berlin, Bremen, Hamburg und das Saarland können keine Ergebnisse zu sozialen Disparitäten berichtet werden, da die erforderlichen Informationen für weniger als 70 Prozent der Schülerinnen und Schüler vorliegen (vgl. Kapitel 3.1).

Der Differenzwert zwischen den EGP-Extremgruppen (ΔM) ist für jedes Land und für Deutschland insgesamt signifikant ($p < .05$) von 0 verschieden. Fett gedruckte Differenzen unterscheiden sich signifikant ($p < .05$) von der Mittelwertsdifferenz in Deutschland insgesamt.

-30 -20 -10 0 10 20 30 40

EGP-Klassen I–II (hoher sozioökonomischer Status)
EGP-Klassen V–VII (niedriger sozioökonomischer Status)

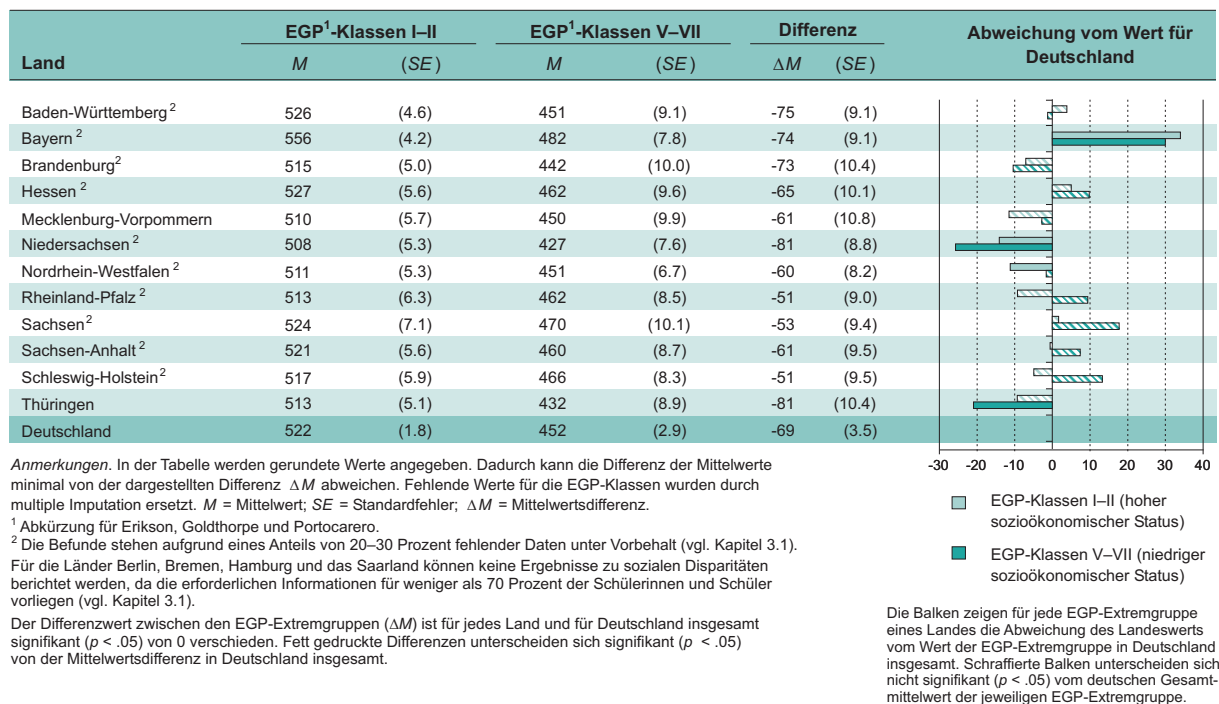
Die Balken zeigen für jede EGP-Extremgruppe eines Landes die Abweichung des Landeswerts vom Wert der EGP-Extremgruppe in Deutschland insgesamt. Schraffierte Balken unterscheiden sich nicht signifikant ($p < .05$) vom deutschen Gesamtmittelwert der jeweiligen EGP-Extremgruppe.

Die für den Kompetenzbereich *Zuhören* festgestellte Differenz zwischen den EGP-Extremgruppen beträgt für Deutschland insgesamt 95 Punkte (Abb. 8.9), wobei auch hier Schülerinnen und Schüler aus Familien mit hohem sozioökonomischen Status höhere Kompetenzwerte erreichen. Da der Kompetenzunterschied im *Zuhören* zwischen der 3. und 4. Jahrgangsstufe (ebenso wie im *Lesen*) etwa 60 Punkte beträgt, entspricht der Unterschied zwischen den EGP-Extremgruppen dem in etwas mehr als anderthalb Schuljahren erwarteten Lernzuwachs (Behrens, Böhme & Krelle, 2009; vgl. auch Kapitel 6.1). In Sachsen-Anhalt fällt die Differenz signifikant geringer aus als im bundesdeutschen Durchschnitt, wohingegen kein Land signifikant größere Disparitäten aufweist.

Im Kompetenzbereich *Zuhören* ergeben sich vor allem für Schülerinnen und Schüler aus sozial privilegierten Familien signifikante Abweichungen der Ländermittelwerte in beide Richtungen von den in Deutschland im Mittel erreichten Kompetenzen dieser Gruppe. Während Schülerinnen und Schüler mit einem hohen sozioökonomischen Status in Mecklenburg-Vorpommern, Rheinland-Pfalz und Sachsen-Anhalt im Vergleich zum deutschen Mittelwert dieser EGP-Extremgruppe signifikant geringere Kompetenzen aufweisen, erreichen diese Schülerinnen und Schüler in Bayern, Sachsen und Schleswig-Holstein überdurchschnittliche Kompetenzen im *Zuhören*. Für Schülerinnen und Schüler mit einem niedrigen sozioökonomischen Status besteht hingegen nur in Bayern eine signifikante positive Abweichung vom deutschen Mittelwert dieser Gruppe.

Abbildung 8.10 zeigt die Ergebnisse der EGP-Extremgruppen für den Kompetenzbereich *Orthografie*. In diesem Bereich fällt die Differenz der EGP-Extremgruppen mit 69 Punkten deutlich geringer aus als in den Kompetenzbereichen *Lesen* und *Zuhören*. Diese Differenz unterscheidet sich zudem in keinem Land signifikant von der für Deutschland insgesamt ermittelten Differenz. Auch im Vergleich mit dem Kompetenzzuwachs zwischen der 3. und der 4. Jahrgangsstufe, der im Bereich *Orthografie* etwa 100 Punkte beträgt (vgl.

Abbildung 8.10: Vergleich der durchschnittlichen Kompetenzwerte der EGP-Extremgruppen im Kompetenzbereich *Orthografie* im Fach Deutsch für das Jahr 2016

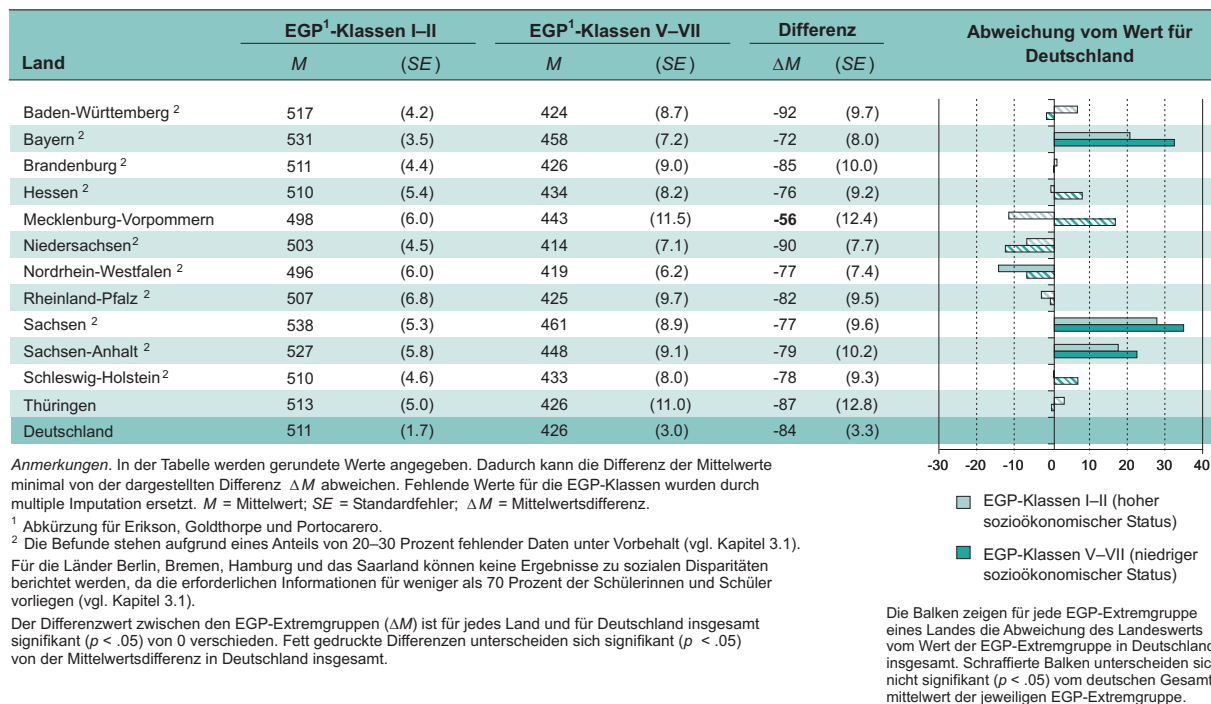


Kapitel 6.1), fällt die Differenz der EGP-Extremgruppen im Bereich *Orthografie* kleiner aus; sie entspricht dem in etwas mehr als zwei Dritteln eines Schuljahres zu erwartenden Lernzuwachs.

Im Vergleich zum jeweiligen Gruppenmittelwert für Deutschland können jedoch auch für den Kompetenzbereich *Orthografie* einige Länderunterschiede in den Ergebnissen der EGP-Extremgruppen festgestellt werden. Signifikante Abweichungen für beide EGP-Extremgruppen zeigen sich in Bayern und in Niedersachsen: Während Schülerinnen und Schüler beider Gruppen in Bayern im Mittel überdurchschnittliche Kompetenzen erreichen, fallen die Ergebnisse für beide Gruppen in Niedersachsen signifikant ungünstiger aus als im bundesweiten Durchschnitt. Signifikant geringere Kompetenzwerte für eine der beiden EGP-Extremgruppen finden sich in zwei Ländern: In Nordrhein-Westfalen erreichen Schülerinnen und Schüler mit hohem sozioökonomischen Status unterdurchschnittliche Kompetenzwerte, in Thüringen ist dies bei Schülerinnen und Schülern mit niedrigem sozioökonomischen Status der Fall.

Bezogen auf die Globalskala mathematischer Kompetenzen beträgt die Differenz zwischen den EGP-Extremgruppen in Deutschland insgesamt 84 Punkte zugunsten von Kindern aus Familien mit einem hohen sozioökonomischen Status (Abb. 8.11), was ungefähr dem Kompetenzzuwachs zwischen der 3. und der 4. Jahrgangsstufe entspricht, der bei etwa 80 Punkten liegt (Reiss & Winkelmann, 2009, vgl. auch Kapitel 6.2). In Mecklenburg-Vorpommern ist diese Differenz signifikant geringer als im bundesdeutschen Durchschnitt.

Betrachtet man die im Mittel erreichten mathematischen Kompetenzen von Schülerinnen und Schülern der EGP-Extremgruppen im Ländervergleich, so lässt sich feststellen, dass die Mittelwerte beider Gruppen zumeist in dieselbe Richtung vom bundesdeutschen Durchschnitt der jeweiligen Gruppe abweichen. So gelingt es in Bayern, Sachsen und Sachsen-Anhalt für beide EGP-Extremgruppen, besonders hohe Kompetenzwerte zu erzielen. Abweichungen

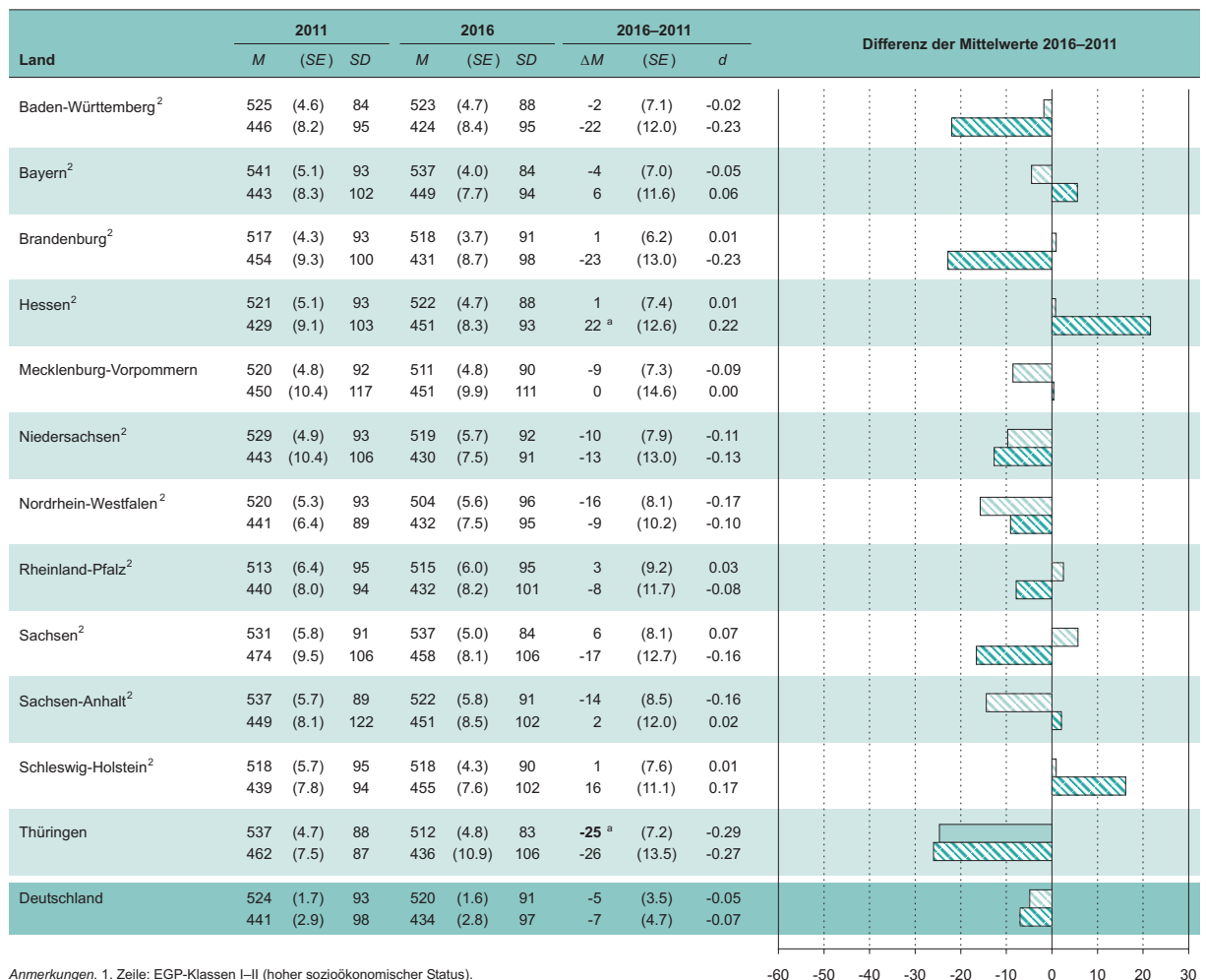
Abbildung 8.11: Vergleich der durchschnittlichen Kompetenzwerte der EGP-Extremgruppen im Fach Mathematik (*Globalskala*) für das Jahr 2016

für nur eine der beiden EGP-Extremgruppen zeigen sich lediglich in Nordrhein-Westfalen. Dort erreichen Schülerinnen und Schüler mit hohem sozioökonomischen Status signifikant geringere Kompetenzwerte als Schülerinnen und Schüler dieser Gruppe bundesweit.

8.6.2 Kompetenzniveau nach EGP-Extremgruppen in den Jahren 2011 und 2016 im Vergleich

Im folgenden Abschnitt wird dargestellt, wie sich die von Schülerinnen und Schülern der EGP-Extremgruppen im Mittel erreichten Kompetenzen in den Ländern zwischen den Jahren 2011 und 2016 verändert haben. Die Abbildungen 8.12 bis 8.15 enthalten die Ergebnisse der Trendanalysen. Im Tabellenteil dieser Abbildungen sind für jedes Land und für Deutschland insgesamt jeweils für die Jahre 2011 und 2016 die von Schülerinnen und Schülern der EGP-Extremgruppen im Mittel erreichten Kompetenzwerte (*M*), die Standardfehler der Mittelwerte (*SE*) und die Standardabweichungen der Kompetenzwerte (*SD*) angegeben. Die letzten drei Spalten des Tabellenteils der Abbildungen beziehen sich auf die Veränderungen der im Durchschnitt erreichten Kompetenzen innerhalb jeder Gruppe. Die Veränderungen (ΔM) sowie ihre Standardfehler (*SE*) werden in Punkten auf der Berichtsmetrik angegeben. Da sich innerhalb der Gruppen zwischen den Jahren 2011 und 2016 teilweise auch die Streuungen der erreichten Kompetenzen verändert haben, wird zusätzlich eine standardisierte Mittelwertsdifferenz für die Kompetenzunterschiede zwischen den Jahren 2011 und 2016 angegeben (*d*) (vgl. Kapitel 3.2). Die Differenzen zwischen den Kompetenzen, die innerhalb der Gruppen in den Jahren 2011 und 2016 erreicht wurden, sind zusätzlich in einem Balkendiagramm visualisiert. Statistisch signifikant

Abbildung 8.12: Mittelwerte und Streuungen der erreichten Kompetenzen im Kompetenzbereich *Lesen* im Fach Deutsch nach EGP-Extremgruppe und Land in den Jahren 2011 und 2016 im Vergleich



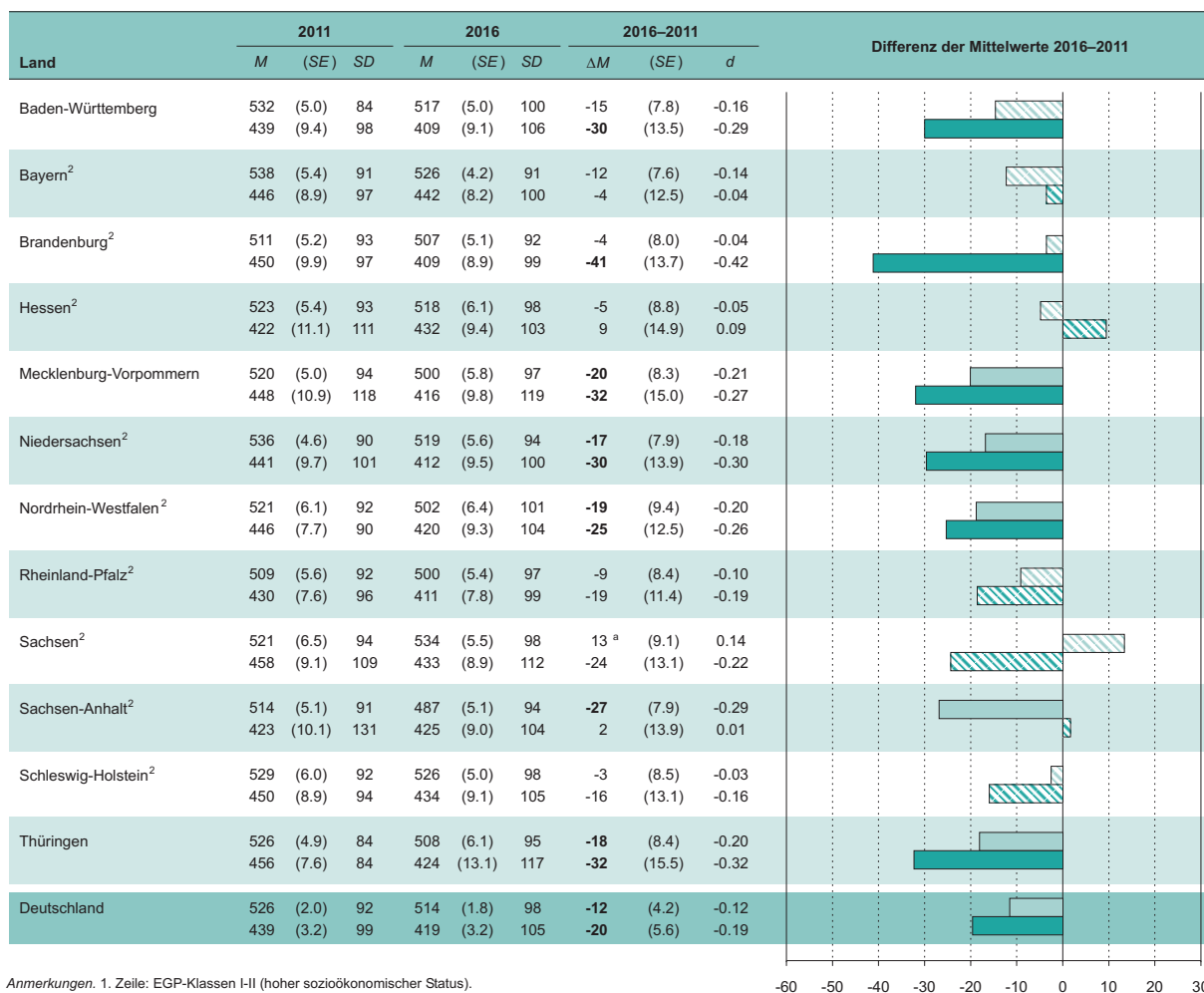
Anmerkungen. 1. Zeile: EGP-Klassen I–II (hoher sozioökonomischer Status).
 2. Zeile: EGP-Klassen V–VII (niedriger sozioökonomischer Status).
 M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.
¹ Abkürzung für Erikson, Goldthorpe und Portocarero. ² Die Befunde stehen aufgrund eines Anteils von 20–30 Prozent fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).
^a Wert unterscheidet sich signifikant ($p < .05$) vom Wert für Deutschland.
 Für die Länder Berlin, Bremen, Hamburg und das Saarland können keine Ergebnisse zu sozialen Disparitäten berichtet werden, da die erforderlichen Informationen für weniger als 70 Prozent der Schülerinnen und Schüler vorliegen (vgl. Kapitel 3.1).
 Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).
 Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

EGP¹-Klassen I–II (hoher sozioökonomischer Status)
 EGP¹-Klassen V–VII (niedriger sozioökonomischer Status)

te Unterschiede sind durch vollständig ausgefüllte Balken, statistisch nicht signifikante Unterschiede durch schraffierte Balken gekennzeichnet.

Im Kompetenzbereich *Lesen* zeigen sich für Deutschland insgesamt für keine der beiden EGP-Extremgruppen signifikante Veränderungen der im Mittel erreichten Kompetenzen (Abb. 8.12). Auf der Ebene der Länder ist ausschließlich in Thüringen ein signifikant negativer Trend der erreichten Kompetenzwerte für Schülerinnen und Schüler mit hohem sozialen Status zu verzeichnen, der darüber hinaus signifikant stärker ausgeprägt ist als die für Deutschland insgesamt festgestellte Verringerung in dieser Gruppe. In Hessen fällt die Veränderung der von Schülerinnen und Schülern mit niedrigem sozioökonomischen Status erreichten Kompetenzwerte signifikant positiver aus als bundesweit.

Im Kompetenzbereich *Zuhören* sind sowohl für Schülerinnen und Schüler mit hohem sozioökonomischen Status als auch für Schülerinnen und Schüler mit niedrigem sozioökonomischen Status signifikante Kompetenzrückgänge zu erkennen (Abb. 8.13), die für Schülerinnen und Schüler mit niedrigem sozioökonomischen

Abbildung 8.13: Mittelwerte und Streuungen der erreichten Kompetenzen im Kompetenzbereich *Zuhören* im Fach Deutsch nach EGP-Extremgruppe und Land in den Jahren 2011 und 2016 im Vergleich

Anmerkungen. 1. Zeile: EGP-Klassen I-II (hoher sozioökonomischer Status).

2. Zeile: EGP-Klassen V-VII (niedriger sozioökonomischer Status).

M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.

¹ Abkürzung für Erikson, Goldthorpe und Portocarero. ² Die Befunde stehen aufgrund eines Anteils von 20–30 Prozent fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

^a Wert unterscheidet sich signifikant ($p < .05$) vom Wert für Deutschland.

Für die Länder Berlin, Bremen, Hamburg und das Saarland können keine Ergebnisse zu sozialen Disparitäten berichtet werden, da die erforderlichen Informationen für weniger als 70 Prozent der Schülerinnen und Schüler vorliegen (vgl. Kapitel 3.1). Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Status tendenziell etwas größer ausfallen. Auch in Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen und Thüringen sind im Jahr 2016 für beide Gruppen geringere Kompetenzen festzustellen als im Jahr 2011. In drei Ländern zeigen sich hingegen differenzielle Kompetenzentwicklungen für einzelne EGP-Extremgruppen: In Baden-Württemberg und in Brandenburg ergibt sich nur für die Gruppe der Schülerinnen und Schüler mit niedrigem sozioökonomischen Status, in Sachsen-Anhalt hingegen für die Gruppe der Schülerinnen und Schüler mit hohem sozioökonomischen Status eine signifikante Verringerung der erreichten Kompetenzen; die Kompetenzen der jeweils anderen Gruppe bleiben stabil. In Sachsen heben sich die leicht, jedoch nicht signifikant steigenden Kompetenzwerte für Schülerinnen und Schüler mit hohem sozioökonomischen Status zudem signifikant vom bundesweiten negativen Trend für diese Gruppe ab.

Veränderungen der orthografischen Kompetenzen der einzelnen EGP-Extremgruppen können nur bundesweit bestimmt werden. Auch in diesem Bereich zei-

Abbildung 8.14: Mittelwerte und Streuungen der erreichten Kompetenzen im Kompetenzbereich *Orthografie* im Fach Deutsch nach EGP-Extremgruppe in den Jahren 2011 und 2016 im Vergleich

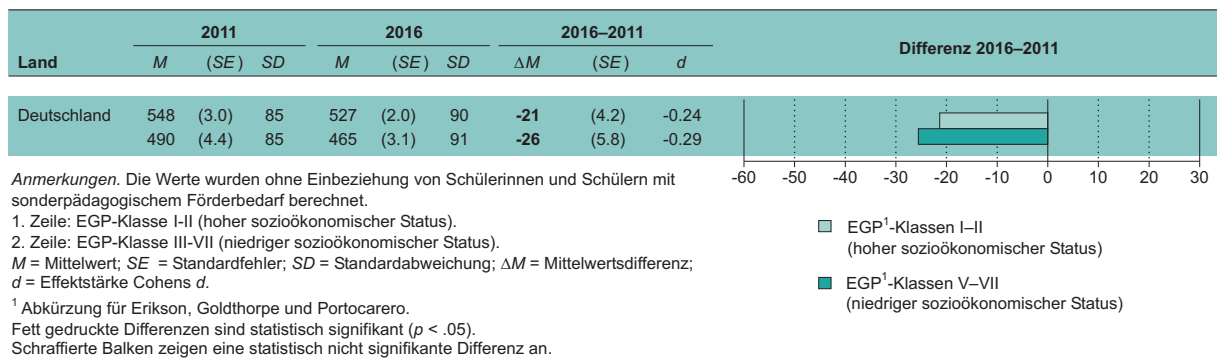
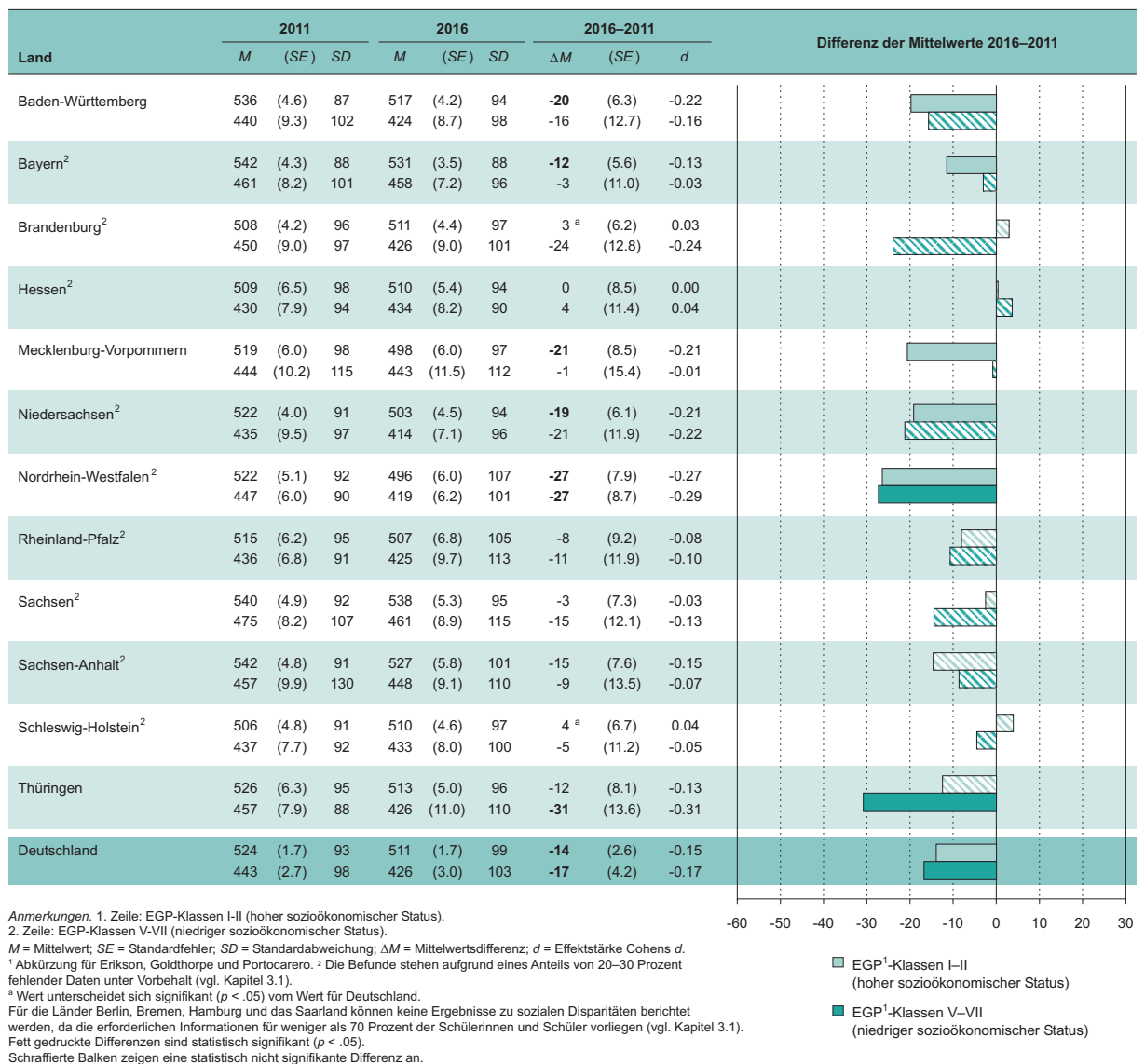


Abbildung 8.15: Mittelwerte und Streuungen der erreichten Kompetenzen im Fach Mathematik (*Globalskala*) nach EGP-Extremgruppe und Land in den Jahren 2011 und 2016 im Vergleich



gen sich für beide betrachteten EGP-Extremgruppen im Jahr 2016 signifikant geringere Kompetenzwerte gegenüber dem Jahr 2011 (Abb. 8.14).

Deutschlandweit betrachtet sind auch im Fach Mathematik die erreichten Kompetenzwerte seit dem Jahr 2011 in beiden EGP-Extremgruppen statistisch signifikant zurückgegangen. Auf der Ebene der Länder findet sich dieses Ergebnismuster aber nur in Nordrhein-Westfalen wieder. In einigen Ländern fallen die Kompetenzmittelwerte im Jahr 2016 vor allem bei Schülerinnen und Schülern mit hohem sozioökonomischen Status geringer aus als im Jahr 2011. Dies gilt für Baden-Württemberg, Bayern, Mecklenburg-Vorpommern und Niedersachsen. Demgegenüber verringern sich die im Mittel erreichten Kompetenzwerte für Schülerinnen und Schüler mit niedrigem sozioökonomischen Status nur in Thüringen statistisch signifikant. Im Fach Mathematik gelingt es in Brandenburg und in Schleswig-Holstein, das Kompetenzniveau der Schülerinnen und Schüler mit hohem sozioökonomischen Status signifikant stabiler zu halten als im bundesweiten Durchschnitt.

8.7 Zusammenfassung und Diskussion

In diesem Kapitel wurde untersucht, inwieweit die in den Fächern Deutsch und Mathematik erreichten Kompetenzen der Viertklässlerinnen und Viertklässler in Deutschland mit dem sozialen Status ihrer Familien zusammenhängen. Die in diesem Kapitel berichteten Analysen stehen aufgrund des hohen Anteils fehlender Angaben zum sozioökonomischen Status der Familien für die Mehrzahl der Länder unter Vorbehalt. Da in Berlin, Bremen, Hamburg und im Saarland für über 30 Prozent der Schülerinnen und Schüler keine Angaben zum sozialen Hintergrund ihrer Familien vorliegen, können für diese Länder keine sozialen Disparitäten berichtet werden (vgl. Kapitel 3.1).

Für das Jahr 2016 ist festzuhalten, dass sich die Länder im durchschnittlichen HISEI und auch in der Streuung der HISEI-Werte nur geringfügig unterscheiden. Ein ähnlicher Befund zeigte sich bereits in früheren IQB-Studien sowohl in der Primarstufe (Richter et al., 2012) als auch in der Sekundarstufe (Kuhl et al., 2016; Kuhl et al., 2013). Familien in den Ländern Mecklenburg-Vorpommern, Sachsen-Anhalt und Thüringen verfügen jedoch über einen im Mittel etwas geringeren und etwas homogeneren sozioökonomischen Status. In Brandenburg entsprechen die HISEI-Werte zwar dem bundesdeutschen Durchschnitt, die Streuung des HISEI ist jedoch geringer als in Deutschland insgesamt. In diesen vier Ländern fallen die sozialen Unterschiede zwischen Familien also geringer aus als in Deutschland insgesamt.

Im Vergleich der Jahre 2011 und 2016 zeigt sich eine weitgehende Stabilität des mittleren sozioökonomischen Status. In Deutschland insgesamt sowie in einigen Ländern kann jedoch eine geringfügige Zunahme der Heterogenität des sozialen Status festgestellt werden. Im IQB-Bildungstrend 2016 wird erstmals auch die Verteilung der sozialen Heterogenität auf die Schulen analysiert und die Veränderung zwischen den Jahren 2011 und 2016 beschrieben. Die sozialen Unterschiede zwischen Schülerinnen und Schülern in Deutschland sind zu etwa fünf Sechsteln (84 %) auf Unterschiede innerhalb von Schulen und zu einem Sechstel (16 %) auf Unterschiede zwischen Schulen zurückzuführen. Diese Werte sind in allen Ländern sowohl im Jahr 2011 als auch im Jahr 2016 ähnlich hoch ausgeprägt; ein etwas größerer Anstieg der Unterschiede zwischen Schulen findet sich nur in Hessen und in Sachsen.

Die Kopplung zwischen den erreichten Kompetenzen der Viertklässlerinnen und Viertklässler und dem sozialen Status ihrer Familien wird anhand der Steigung der sozialen Gradienten bestimmt. Für alle untersuchten Bereiche liegen sowohl deutschlandweit als auch in allen Ländern signifikante soziale Gradienten vor, die auf ausgeprägte soziale Disparitäten hinweisen. Die sozialen Gradienten fallen im Bereich *Orthografie* mit etwa 30 Punkten etwas geringer aus als in den Bereichen *Lesen* und *Zuhören* sowie im Fach Mathematik. Dies deutet darauf hin, dass die orthografischen Kompetenzen etwas weniger stark vom sozialen Status der Familie abhängen als die Kompetenzen in den anderen untersuchten Bereichen. Zwischen den Ländern zeigen sich insgesamt nur geringe Unterschiede im Grad der Kopplung sozialer Herkunftsmerkmale mit den erreichten Kompetenzen, die sich auf die Bereiche *Lesen* und *Zuhören* beschränken. Geringer ausgeprägte Werte des sozialen Gradienten als für Deutschland insgesamt zeigen sich im *Lesen* für Schleswig-Holstein und Thüringen sowie im *Zuhören* für Sachsen-Anhalt. Signifikant höher ausgeprägte soziale Gradienten bestehen hingegen lediglich im Bereich *Zuhören* in Sachsen.

Ein Vergleich der sozialen Gradienten zwischen den Jahren 2011 und 2016 zeigt auf Bundesebene weitgehende Stabilität. Im Kompetenzbereich *Zuhören* sowie im Fach Mathematik nehmen die sozialen Disparitäten geringfügig zu; in den Bereichen *Lesen* und *Orthografie* werden hingegen bundesweit keine Veränderungen identifiziert. Auf Länderebene sind lediglich in zwei Ländern in jeweils einem Bereich steigende Disparitäten zu verzeichnen: In Baden-Württemberg sind die Kompetenzen im *Lesen*, in Sachsen im Bereich *Zuhören* im Jahr 2016 stärker an den sozialen Status gekoppelt als im Jahr 2011.

Im Vergleich der Ländermittelwerte von zwei EGP-Extremgruppen zeigt sich, dass die Unterschiede zwischen den Schülerinnen und Schülern aus Familien mit einem hohen sozioökonomischen Status (EGP-Klassen I–II) und den Schülerinnen und Schülern mit einem niedrigen sozioökonomischen Status (EGP-Klassen V–VII) in allen Ländern substantiell sind und zwischen den Ländern nur wenig variieren. Besonders geringe Disparitäten können teilweise auf die Ergebnisse einzelner EGP-Extremgruppen zurückgeführt werden: Während sich die vergleichsweise geringen sozialen Disparitäten im *Lesen* in Schleswig-Holstein durch das relativ hohe Kompetenzniveau der Schülerinnen und Schüler aus sozial schwachen Familien erklären lassen, sind die geringen sozialen Disparitäten im *Zuhören* in Sachsen-Anhalt auf das vergleichsweise geringe Kompetenzniveau von Schülerinnen und Schülern aus sozial privilegierten Familien zurückzuführen.

Bei der Analyse von Trends können für die von den Schülerinnen und Schülern der beiden EGP-Extremgruppen erreichten Kompetenzen weitgehend vergleichbare Veränderungen über die Zeit festgestellt werden. Im Bereich *Zuhören* und im Fach Mathematik bestehen jedoch tendenziell stärkere Kompetenzrückgänge für Schülerinnen und Schüler mit niedrigem sozioökonomischen Status.

Zusammenfassend zeigt sich anhand der Daten des IQB-Bildungstrends 2016 erneut, dass soziale Herkunftsmerkmale der Familie für den Kompetenzerwerb in der Primarstufe nach wie vor sehr relevant sind, wobei die Ausprägung dieses Zusammenhangs zwischen den untersuchten Kompetenzbereichen variiert. Die Kopplung zwischen dem sozialen Status der Familien und dem Kompetenzerwerb von Viertklässlerinnen und Viertklässlern im Fach Deutsch ist zwischen den Jahren nahezu unverändert geblieben und hat im Fach Mathematik sogar leicht zugenommen. Ziel bildungspolitischer und schulpraktischer Bemühungen sollte

es daher auch künftig sein, Bedingungen zu schaffen, die zur Verringerung sozialer Disparitäten beitragen und Schülerinnen und Schülern unabhängig von ihrer sozialen Ausgangslage möglichst gute Entwicklungschancen bieten.

Literatur

- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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.
- 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. 198–218). 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.
- 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. (2010). *A new international socio-economic index [ISEI] of occupational status for the International Standard Classification of Occupation 2008 [ISCO-08] constructed with data from the ISSP 2002–2007; with an analysis of quality of educational measurement in ISSP*. Vortrag auf der Annual Conference of International Social Survey Programme, Lissabon.
- 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. (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.
- Kuhl, P., Haag, N., Federlein, F., Weirich, S. & Schipolowski, S. (2016). Soziale Disparitäten. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 409–430). Münster: Waxmann.
- Kuhl, P., Siegle, T. & Lenski, A. E. (2013). Soziale Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 275–296). Münster: Waxmann.
- 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.
- Müller, K. & Ehmke, T. (2013). Soziale Herkunft als Bedingung der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 245–275). Münster: Waxmann.
- Müller, K. & Ehmke, T. (2016). Soziale Herkunft und Kompetenzerwerb. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 285–316). 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.
- Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011* (S. 191–208). Münster: Waxmann.
- 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.
- Stubbe, T. C., Schwippert, K. & Wendt, H. (2016). Soziale Disparitäten der Schülerleistungen in Mathematik und Naturwissenschaften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 299–316). Münster: Waxmann.
- Stubbe, T. C., Tarelli, I. & Wendt, H. (2012). Soziale Disparitäten der Schülerleistungen in Mathematik und Naturwissenschaften. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 231–246). Münster: Waxmann.
- Wendt, H., Stubbe, T. C. & Schwippert, K. (2012). Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 175–190). Münster: Waxmann.

Kapitel 9

Zuwanderungsbezogene Disparitäten

Camilla Rjosk, Nicole Haag, Birgit Heppt und Petra Stanat

9.1 Analysen zuwanderungsbezogener Disparitäten als Gegenstand des Bildungsmonitorings

Angesichts steigender Zahlen von Kindern und Jugendlichen mit Zuwanderungshintergrund im Bildungssystem der Bundesrepublik Deutschland gewinnt auch die Frage an Bedeutung, inwieweit die schulbezogene Integration dieser Kinder und Jugendlichen gelingt (vgl. z. B. Autorengruppe Bildungsberichterstattung, 2016). Ein zentraler Aspekt ist dabei der schulische Kompetenzerwerb als wichtiger Bestandteil kultureller Integration (vgl. Esser, 2006) und als entscheidende Voraussetzung für die weitere Bildungslaufbahn sowie für die gesellschaftliche Teilhabe. Daher ist es ein allgemein anerkanntes bildungspolitisches Ziel, zuwanderungsbezogene Disparitäten in den erreichten Kompetenzen soweit wie möglich zu reduzieren. Entsprechend wird im Bildungsmonitoring fortlaufend untersucht, inwieweit dieses Ziel erreicht wird, das heißt, inwieweit sich die von Schülerinnen und Schülern aus zugewanderten Familien und die von Schülerinnen und Schülern ohne Zuwanderungshintergrund erreichten Kompetenzen unterscheiden und bestehende Disparitäten reduziert werden konnten.¹

Schulische Integration bezieht sich jedoch nicht nur auf den Kompetenzerwerb, sondern auch darauf, ob sich Schülerinnen und Schüler in der Schule angenommen und wohl fühlen und ob sie Freude am Lernen haben. Soziale Integration und schulische Motivation gelten im schulischen Kontext zusätzlich zum Kompetenzerwerb als wichtige Ziele, da sie unter anderem das soziale und individuelle Lernverhalten von Schülerinnen und Schülern und somit deren weitere Kompetenzentwicklung positiv beeinflussen (Roeser, Eccles & Sameroff, 2000). Um die Lernsituation von Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund umfassender zu beschreiben, werden daher Merkmale der sozialen Integration und schulischen Motivation zunehmend auch im Bildungsmonitoring in den Blick genommen (z. B. Haag et al., 2016; Stanat, Christen & Segeritz, 2010).

1 Wie im Kapitel zu zuwanderungsbezogenen Disparitäten des IQB-Bildungstrends 2015 ausführlicher diskutiert (Haag, Böhme, Rjosk & Stanat, 2016), birgt die Beschreibung zuwanderungsbezogener Disparitäten, vor allem die differenzierte Darstellung von Ergebnissen für einzelne Herkunftsgruppen, einige Gefahren: Eventuelle Kompetenznachteile können auf eine den Gruppen inhärente „Andersartigkeit“ zurückgeführt werden, die zuwanderungsbezogenen Kategorisierungen könnten einen ausgrenzenden Effekt haben und somit der Entwicklung einer Gesellschaft zuwider laufen, in der Pluralität Normalität ist (vgl. z. B. Foroutan, 2010). Empirische Befunde zu zuwanderungsbezogenen Disparitäten ermöglichen es jedoch, weiterhin bestehende Benachteiligungen sichtbar zu machen und eine Grundlage für Bemühungen um deren Verringerung zu schaffen. Daher ist es wichtig, die Kompetenzentwicklung von Schülerinnen und Schülern mit Zuwanderungshintergrund im Bildungssystem fortlaufend zu beobachten.

Im vorliegenden Kapitel werden die im IQB-Bildungstrend 2016 erfassten Kompetenzen in den Fächern Deutsch und Mathematik bei Schülerinnen und Schülern mit Zuwanderungshintergrund untersucht und mit den Kompetenzen verglichen, die von Schülerinnen und Schülern ohne Zuwanderungshintergrund erreicht werden. Dabei liegt ein zusätzlicher Fokus auf Trendanalysen, die Anhaltspunkte dafür geben, inwieweit es in den Schulen in der Primarstufe zwischen den Jahren 2011 und 2016 gelungen ist, zuwanderungsbezogene Disparitäten zu reduzieren. In weiteren Analysen werden wichtige familiäre Hintergrundmerkmale der Schülerinnen und Schüler – der sozioökonomische Status sowie der Bildungshintergrund der Eltern und die zu Hause gesprochene Sprache – berücksichtigt, um Hinweise darauf zu erhalten, inwieweit die zuwanderungsbezogenen Disparitäten mit diesen Faktoren in Zusammenhang stehen.

Zusätzlich wird untersucht, inwieweit sich Kinder aus zugewanderten Familien und Kinder ohne Zuwanderungshintergrund im Jahr 2016 in verschiedenen Merkmalen der sozialen Integration und schulischen Motivation unterscheiden. Dabei geht es um das Zugehörigkeitsgefühl zu den Mitschülerinnen und Mitschülern, die allgemeine Schulzufriedenheit sowie die Lernfreude als Merkmal der schulischen Motivation.

Bevor die Ergebnisse zur kompetenzbezogenen und sozialen Integration sowie zur Lernfreude dargestellt werden, wird zunächst erläutert, anhand welcher Merkmale der Zuwanderungshintergrund erfasst wird. Anschließend werden die Anteile von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund in den Ländern im Jahr 2016 sowie deren Veränderungen gegenüber dem Jahr 2011 beschrieben. Als Beitrag zum Schwerpunktthema des IQB-Bildungstrends 2016, „Umgang mit Heterogenität“ (vgl. Kapitel 1.1), wird in diesem Kapitel zudem genauer untersucht, wie sich die Kinder mit Zuwanderungshintergrund auf die Schulen in den einzelnen Ländern verteilen und wie sich die Verteilung zwischen den Jahren 2011 und 2016 verändert hat.

9.2 Methodisches Vorgehen

9.2.1 Erfassung des Zuwanderungshintergrunds

Eine fortlaufende Beobachtung zuwanderungsbezogener Disparitäten erfordert eine konsistente und kontinuierliche Verwendung von Gruppierungsmerkmalen. Die Klassifikation des Zuwanderungshintergrunds im IQB-Bildungstrend 2016 orientiert sich deshalb am Vorgehen in internationalen Studien wie IGLU², TIMSS³ und PISA⁴ sowie den bereits durchgeführten IQB-Ländervergleichs- und IQB-Bildungstrendstudien (z. B. Haag et al., 2016; Pöhlmann, Haag & Stanat, 2013; Rauch, Mang, Härtig & Haag, 2016; Schwippert, Wendt & Tarelli, 2012; Wendt, Schwippert & Stubbe, 2016).

Um den *Zuwanderungsstatus* bestimmen zu können, wurden die Schülerinnen und Schüler in den Erhebungen zum IQB-Bildungstrend 2016 gebeten, in einem Fragebogen ihr eigenes Geburtsland sowie das ihrer Eltern anzugeben. Für die Auswertung wurden diese Informationen mit den Angaben der Eltern über

2 Das Akronym IGLU steht für *Internationale Grundschul-Lese-Untersuchung*.

3 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.

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

ihre eigenen Geburtsländer, das Geburtsland ihres Kindes sowie das Geburtsland der Großeltern⁵ des Kindes zusammengeführt. Lagen von den Eltern keine Informationen vor, wurden die Angaben der Schülerinnen und Schüler herangezogen.

Gemäß dem Vorgehen in den bisherigen Ländervergleichs- und Bildungstrendstudien des IQB werden im vorliegenden Kapitel folgende vier Gruppen unterschieden:

- *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 Zuwanderungsgeneration*: beide Elternteile sind im Ausland geboren, die Schülerin oder der Schüler selbst ist in Deutschland geboren;
- *Schülerinnen und Schüler der ersten Zuwanderungsgeneration*: sowohl beide Elternteile als auch die Schülerin oder der Schüler selbst sind im Ausland geboren.

Zusätzlich zum Zuwanderungsstatus wurden die Schülerinnen und Schüler im IQB-Bildungstrend 2016 anhand des Geburtslands ihrer Eltern einer *Herkunftsgruppe* zugeordnet (vgl. Olczyk et al., 2016), um potenzielle Benachteiligungen⁶ bestimmter Teilgruppen zu identifizieren. In den Analysen werden die größten Herkunftsgruppen in Deutschland betrachtet und wie folgt definiert:

- *Türkei*: mindestens ein Elternteil ist in der Türkei geboren;
- *Gebiet der ehemaligen Sowjetunion* (im Folgenden kurz: *ehemalige Sowjetunion*): mindestens ein Elternteil ist in Russland oder in einer anderen ehemaligen Sowjetrepublik geboren;
- *Polen*: mindestens ein Elternteil ist in Polen geboren;
- *Gebiet des ehemaligen Jugoslawien* (im Folgenden kurz: *ehemaliges Jugoslawien*): mindestens ein Elternteil ist in Bosnien und Herzegowina, dem Kosovo, Kroatien, Mazedonien, Montenegro, Serbien oder Slowenien geboren;
- *arabische Länder*: mindestens ein Elternteil ist im Irak, in Syrien oder einem anderen Mitgliedsstaat der arabischen Liga mit arabischer Bevölkerungsmehrheit geboren (vgl. Flores, 2013);
- *anderes Land*: mindestens ein Elternteil ist in einem hier nicht genannten Land geboren oder die Eltern sind in zwei unterschiedlichen Ländern im Ausland geboren.

5 Das Geburtsland der Großeltern des Kindes wurde ebenfalls erfragt, um die dritte Zuwanderergeneration bestimmen zu können. Diese Gruppe umfasst in Deutschland geborene Kinder mit zwei in Deutschland geborenen Elternteilen und mindestens zwei im Ausland geborenen Großelternanteilen (vgl. Olczyk, Seuring, Will & Zinn, 2016). Der Anteil von Schülerinnen und Schülern der dritten Zuwanderergeneration kann der auf der Webseite des IQB als Zusatzmaterial zur Verfügung gestellten Tabelle 9.1web entnommen werden. Da dieser Anteil jedoch sehr gering ist (deutschlandweit 3 %), wird die dritte Zuwanderergeneration in weitergehenden Analysen nicht separat aufgeführt, sondern – in Übereinstimmung mit internationalen Studien sowie bisherigen IQB-Ländervergleichs- und IQB-Bildungstrendstudien – mit der Gruppe der Schülerinnen und Schüler ohne Zuwanderungshintergrund zusammengefasst.

6 Der Begriff „Benachteiligung“ wird in diesem Kapitel im beschreibenden Sinne verwendet. Er kennzeichnet, dass sich die Lage einer Gruppe ungünstiger darstellt als die einer anderen Gruppe, ohne dabei eine Annahme über mögliche Ursachen zu treffen.

Zudem wurde sowohl für den Zuwanderungsstatus als auch für die Herkunftsgruppen jeweils eine zusätzliche Gruppe gebildet, die als *nicht zuzuordnen* bezeichnet wird. In ihr wurden alle Schülerinnen und Schüler zusammengefasst, für die nicht alle notwendigen Angaben zur Bestimmung des Zuwanderungshintergrundes vorlagen. Diese Gruppe mit fehlenden Informationen ist für Hamburg und das Saarland so groß, dass die Ergebnisse der zuwanderungsbezogenen Analysen mit erheblicher Unsicherheit behaftet sind. Die Ergebnisse dieser beiden Länder stehen daher unter Vorbehalt und werden in allen Ergebnisdarstellungen entsprechend gekennzeichnet (vgl. Kapitel 3.1).⁷

9.2.2 Ergebnisdarstellung und -interpretation

In der Ergebnisdarstellung werden verschiedene Vergleichsperspektiven eingenommen, um die erreichten Kompetenzen der oben beschriebenen Schülergruppen getrennt nach Zuwanderungsstatus und Herkunftsland zu beschreiben. Bei der Interpretation der verschiedenen Vergleichsperspektiven ist zunächst zu berücksichtigen, dass die Fallzahlen für einige Gruppen in den Ländern teilweise sehr gering sind. Da verlässliche Aussagen daher nicht durchgehend möglich sind, werden einzelne Zuwanderergruppen für ländervergleichende Analysen zu größeren Gruppen zusammengefasst (vgl. Abschnitt 9.3.1).

Wurden in früheren IQB-Ländervergleichsstudien Ergebnisse für Schülerinnen und Schüler aus zugewanderten Familien nur für diejenigen Länder in der Bundesrepublik Deutschland berichtet, in denen der Anteil mindestens einer der Zuwanderungsgruppen (ein Elternteil im Ausland geboren, beide Elternteile im Ausland geboren) mehr als 10 Prozent betrug (vgl. Böhme, Tiffin-Richards, Schipolowski & Leucht, 2010; Haag, Böhme & Stanat, 2012; Pöhlmann et al., 2013), werden diese Ergebnisse seit dem IQB-Bildungstrend 2015 für alle Länder berichtet. Da aufgrund der aktuellen Entwicklungen in der Zuwanderung nach Deutschland davon auszugehen ist, dass die Anteile von Schülerinnen und Schülern aus zugewanderten Familien auch in den nächsten Jahren weiter steigen werden, können auf diese Weise zukünftige Entwicklungen in den zuwanderungsbezogenen Disparitäten besser abgebildet werden (Haag et al., 2016).

Bei der Interpretation der Befunde ist weiter zu beachten, dass die Ergebnisse in einigen Ländern aufgrund des geringen Anteils von Kindern aus zugewanderten Familien auf sehr kleinen Stichproben beruhen (z. B. in den ostdeutschen Flächenländern). In anderen Ländern sind die Ergebnisse der Stichproben nicht hinreichend aussagekräftig, weil der Zuwanderungsstatus vieler Kinder nicht bestimmt werden konnte (insbesondere in Hamburg und im Saarland). Zur Einordnung der Befunde werden in den Abbildungen daher auch die Stichprobengrößen der einzelnen Gruppen aufgeführt.

⁷ Die Ergebnisse länderspezifischer Analysen werden unter Vorbehalt berichtet, wenn für 20–30 Prozent der Schülerschaft unvollständige Informationen vorliegen. Dieser Prozentwert bezieht sich auf die in Kapitel 3.1 dargestellten Anteile der Angaben von Schülerinnen und Schülern und/oder ihren Eltern an der Gesamtheit der für die Teilnahme an der Studie durch Zufall ausgewählten Schülerschaft eines Landes. Im Gegensatz dazu beziehen sich die in einigen Tabellen des Kapitels 9 angegebenen Anteile von Kindern mit fehlenden Daten zum Zuwanderungshintergrund auf den Anteil an der Gesamtzahl von Kindern, die an den Kompetenztests des IQB-Bildungstrends teilgenommen haben (siehe z. B. Tab. 9.1). Dadurch kommt es zu geringfügigen Unterschieden zwischen den Angaben zum Anteil fehlender Daten in Kapitel 3.1 und Kapitel 9.

9.3 Kinder aus zugewanderten Familien und Kinder ohne Zuwanderungshintergrund in den Ländern

In diesem Abschnitt wird zunächst berichtet, welche Anteile von Schülerinnen und Schülern der 4. Jahrgangsstufe in den Ländern im Jahr 2016 aus zugewanderten Familien stammen und welche Veränderungen sich seit dem IQB-Ländervergleich 2011 ergeben haben. Zudem wird die Verteilung von Kindern aus zugewanderten Familien in den Schulen der Primarstufe, also die zugewanderungsbezogene Zusammensetzung der Schülerschaft an Schulen, länderweise dargestellt und es wird beschrieben, wie sich diese seit dem Jahr 2011 verändert hat.

9.3.1 Anteile von Viertklässlerinnen und Viertklässlern nach Zuwanderungsstatus

In Tabelle 9.1 sind die Anteile der Schülerinnen und Schüler nach Zuwanderungsstatus sowohl für Deutschland insgesamt als auch für die einzelnen Länder dargestellt. Außerdem ist in Tabelle 9.1 angegeben, welcher Anteil der Schülerinnen und Schüler aufgrund von fehlenden Angaben zum eigenen Geburtsland oder zum Geburtsland der Eltern keiner der Gruppen zugeordnet werden konnte.

Im IQB-Bildungstrend 2016 ließ sich bundesweit der Zuwanderungsstatus von ungefähr 8 Prozent der Schülerinnen und Schüler nicht bestimmen, wobei dieser Anteil zwischen den Ländern erheblich variiert. Dies dürfte vor allem auf die unterschiedlichen Rücklaufquoten für die Fragebogen zurückzuführen sein, die durch länderspezifische Regelungen zum Datenschutz bedingt sind (vgl. Abschnitt 3.1.2). Besonders hoch ist die Quote fehlender Werte in Hamburg und im Saarland.

Bundesweit wie auch in der Mehrheit der Länder ist der Anteil fehlender Werte relativ zum IQB-Ländervergleich 2011 weitgehend konstant geblieben. In Bayern, Bremen, Hessen und Niedersachsen hat sich der Anteil fehlender Werte signifikant verringert; in Bremen ist dieser Rückgang mit mehr als 10 Prozentpunkten besonders deutlich. Im Saarland stieg der Anteil fehlender Werte hingegen seit dem Jahr 2011 um etwa 5 Prozentpunkte an.

Betrachtet man die Schülerinnen und Schüler, deren Zuwanderungsstatus bestimmt werden konnte (vgl. Tab. 9.1; linker Tabellenbereich), liegt der Anteil der Viertklässlerinnen und Viertklässler mit Zuwanderungshintergrund in Deutschland bei insgesamt etwa 34 Prozent und hat sich seit dem Jahr 2011 signifikant um fast 9 Prozentpunkte erhöht. Die höchsten Anteile finden sich mit über 40 Prozent in den Stadtstaaten Berlin, Bremen und Hamburg sowie in den westdeutschen Flächenländern Baden-Württemberg, Hessen und Nordrhein-Westfalen, die geringsten Anteile mit etwa 10 Prozent in den ostdeutschen Flächenländern. In fast allen Ländern ist ein statistisch bedeutsamer Anstieg des Anteils von Kindern mit Zuwanderungshintergrund seit 2011 zu verzeichnen, der mit über 10 Prozentpunkten in Baden-Württemberg, Bremen, Niedersachsen, Nordrhein-Westfalen und Rheinland-Pfalz besonders hoch ist.

Viele der Schülerinnen und Schüler aus zugewanderten Familien sind in Deutschland geboren (zweite Generation bzw. ein Elternteil im Ausland geboren) und haben damit die gesamte Schulzeit im deutschen Bildungssystem verbracht. Besonders stark vertreten ist die Gruppe der zweiten Generation in den Stadtstaaten Bremen (28 %) und Hamburg (25 %) sowie in Baden-Württemberg, Hessen und Nordrhein-Westfalen (jeweils über 20 %). In diesen Ländern – mit

Tabelle 9.1: Prozentuale Anteile der Viertklässlerinnen und Viertklässler nach Zuwanderungsstatus in den Ländern in den Jahren 2011 und 2016 im Vergleich

Land	ohne Zuwanderungshintergrund		mit Zuwanderungshintergrund				nicht zuzuordnen									
	gültige % ¹	(SE)	mit Zuwanderungshintergrund		mit Zuwanderungshintergrund		nicht zuzuordnen									
			gültige % ¹	+/- (SE)	gültige % ¹	+/- (SE)	gültige % ¹	+/- (SE)								
Baden-Württemberg	55.7	(2.4)	44.3	(2.4)	15.1	(1.3)	5.2	(2.1)	8.5	(2.1)	3.6	(0.6)	1.4	10.8	(1.6)	1.9
Bayern	68.6	(2.1)	31.4	(2.1)	6.7	(0.9)	2.2	(2.0)	3.0	(2.0)	3.7	(0.5)	1.6	6.3	(0.8)	-3.1
Berlin	58.6	(1.6)	41.4	(1.6)	5.8	(1.1)	3.1	(1.2)	0.4	(1.2)	5.7	(0.7)	2.3	15.2	(1.3)	-0.2
Brandenburg	89.3	(0.9)	10.7	(0.9)	5.3	(0.7)	2.0	(0.6)	2.5	(0.6)	1.6	(0.3)	0.8	1.8	(0.3)	-0.6
Bremen	47.5	(1.3)	52.5	(1.3)	11.0	(1.2)	3.3	(1.3)	3.3	(1.3)	8.2	(0.7)	4.4	3.7	(0.5)	-12.9
Hessen	57.8	(1.9)	42.2	(1.9)	9.9	(1.0)	4.0	(1.5)	4.2	(1.5)	4.8	(0.7)	1.7	1.3	(0.3)	-3.0
Mecklenburg-Vorpommern	90.4	(1.0)	9.6	(1.0)	5.1	(0.5)	0.9	(0.6)	2.5	(0.6)	2.4	(0.5)	1.8	1.9	(0.4)	0.2
Niedersachsen	68.0	(1.9)	32.0	(1.9)	11.0	(1.0)	4.1	(1.4)	4.6	(1.4)	4.4	(0.8)	2.3	1.2	(0.3)	-1.9
Nordrhein-Westfalen	59.6	(2.4)	40.4	(2.4)	10.2	(1.3)	1.9	(1.7)	6.1	(1.7)	4.1	(0.7)	2.2	13.1	(1.2)	2.9
Rheinland-Pfalz	62.8	(1.9)	37.2	(1.9)	10.1	(0.9)	2.5	(1.4)	5.2	(1.4)	4.3	(0.6)	2.4	11.6	(1.6)	-3.4
Sachsen	89.6	(1.1)	10.4	(1.1)	3.9	(0.7)	2.7	(0.4)	0.9	(0.4)	1.5	(0.4)	0.3	7.3	(1.0)	0.9
Sachsen-Anhalt	90.5	(0.9)	9.5	(0.9)	3.1	(0.7)	1.4	(0.6)	1.6	(0.6)	1.5	(0.3)	0.1	2.0	(0.4)	-1.1
Schleswig-Holstein	76.1	(1.5)	23.9	(1.5)	9.0	(0.8)	3.6	(1.3)	4.6	(1.3)	2.7	(0.5)	0.7	11.0	(1.2)	1.3
Thüringen	91.6	(1.1)	8.4	(1.1)	2.6	(0.5)	1.0	(0.6)	1.0	(0.6)	1.6	(0.5)	0.6	2.0	(0.5)	0.3
Hamburg ²	51.5	(2.2)	48.5	(2.2)	4.2	(1.2)	1.2	(1.9)	2.1	(1.9)	4.9	(0.9)	0.9	17.4	(1.6)	-1.8
Saarland ²	71.9	(1.8)	28.1	(1.8)	4.6	(1.2)	3.0	(1.3)	-0.2	(1.3)	3.5	(0.6)	1.8	17.1	(1.5)	4.7
Deutschland	66.4	(0.7)	33.6	(0.7)	8.9	(0.4)	2.8	(0.6)	4.3	(0.6)	3.8	(0.2)	1.7	8.3	(0.5)	-0.2

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. SE = Standardfehler; +/- = Veränderung gegenüber IQB-Ländervergleich 2011. Erste Generation: Sowohl beide Elternteile als auch das Kind sind im Ausland geboren. Zweite Generation: Beide Elternteile sind im Ausland geboren, das Kind selbst ist in Deutschland geboren.

¹ gültige % = Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind. ² Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

Ausnahme von Nordrhein-Westfalen – sowie in Berlin haben außerdem besonders viele Schülerinnen und Schüler einen im Ausland geborenen Elternteil (16 % und mehr).

Der in Deutschland insgesamt zu beobachtende Anstieg des Anteils von Kindern mit Zuwanderungshintergrund ist vor allem auf Erhöhungen des Anteils von Kindern der zweiten Generation und von Kindern mit einem im Ausland geborenen Elternteil zurückzuführen. Besonders stark stiegen die Anteile der Schülerinnen und Schüler der zweiten Generation in den Ländern Baden-Württemberg (+9 %), Nordrhein-Westfalen (+6 %) und Rheinland-Pfalz (+5 %) sowie um 4 bis 5 Prozentpunkte in Hessen, Niedersachsen und Schleswig-Holstein. In den meisten dieser Länder sowie zusätzlich in Berlin, Brandenburg, Bremen, Sachsen und dem Saarland erhöhten sich auch die Anteile der Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil signifikant.

In Tabelle 9.1 wird zudem deutlich, dass der Anteil zugewanderter Schülerinnen und Schüler der ersten Generation sowohl bundesweit (knapp 4 %) als auch innerhalb der Länder (maximal 8 % in Bremen) recht klein ist und in den Ländern seit dem Jahr 2011 weitgehend konstant blieb. Ein signifikanter Anstieg des Anteils von Schülerinnen und Schülern der ersten Generation, der mit rund 2 bis 4 Prozentpunkten allerdings nicht sehr groß ausfällt, ist in den Ländern Berlin, Bremen, Niedersachsen, Nordrhein-Westfalen und Rheinland-Pfalz zu verzeichnen.⁸ Bei der Interpretation dieser Zahlen ist zu beachten, dass für neu zugewanderte Schülerinnen und Schüler mit einer nichtdeutschen Herkunftssprache keine Pflicht zur Teilnahme am IQB-Bildungstrend bestand, wenn sie weniger als ein Jahr in deutscher Sprache unterrichtet wurden und nicht in der Lage waren, Deutsch zu lesen oder zu sprechen (vgl. Kapitel 3.1). Somit mussten zugewanderte Schülerinnen und Schüler mit geringen Deutschkenntnissen bereits vor dem Frühjahr 2015 eine Schule in Deutschland besuchen, um an der Erhebung zum IQB-Bildungstrend 2016 teilzunehmen. Da der deutliche Anstieg der Neuzuwanderung jedoch erst im Laufe des Jahres 2015 einsetzte (Bundesamt für Migration und Flüchtlinge, 2016), dürfte ein großer Anteil dieser Gruppe nicht von der Stichprobe erfasst worden sein.

Die Befunde bundesweiter Analysen werden getrennt für Kinder ohne Zuwanderungshintergrund, Kinder mit einem im Ausland geborenen Elternteil, Kinder der zweiten Zuwanderungsgeneration und Kinder der ersten Zuwanderungsgeneration aufgeführt. Da die Fallzahlen der zugewanderten Schülerinnen und Schüler der ersten Generation innerhalb der meisten Länder für tragfähige Auswertungen zu klein sind, werden in den länderspezifischen Analysen die erste und zweite Generation zur Gruppe *beide Elternteile im Ausland geboren* zusammengefasst. Bevor die erreichten Kompetenzen von Schülerinnen und Schülern nach Zuwanderungsstatus in Abschnitt 9.4.1 beschrieben werden, wird dargestellt, wie sich die Kinder mit Zuwanderungshintergrund auf die Schulen der Länder verteilen.

⁸ 13 Prozent der Schülerinnen und Schüler der ersten Generation sind nach Angabe der Lehrkräfte im Jahr 2014 oder später als Geflüchtete nach Deutschland gekommen. Für rund 3 Prozent der Kinder liegen keine Angaben vor.

9.3.2 Verteilung von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund auf die Schulen in der Primarstufe

Die steigende Anzahl von Kindern und Jugendlichen aus zugewanderten Familien hat zur Folge, dass die Schülerschaft an Schulen in der Primarstufe zunehmend heterogener wird. Der Grad der zugewanderungsbezogenen Heterogenität kann jedoch zwischen Schulen erheblich variieren, etwa aufgrund von Unterschieden in der Bevölkerungsstruktur des Einzugsgebiets von Schulen oder auch aufgrund des unterschiedlichen Schulwahlverhaltens von Eltern mit und ohne Zuwanderungshintergrund (vgl. Karsten et al., 2006; Kristen, 2008). Zusätzlich zum Anteil von Kindern mit Zuwanderungshintergrund in den Ländern wird deshalb im IQB-Bildungstrend 2016 erstmals untersucht, wie sich diese Schülerinnen und Schüler in den Ländern auf die Schulen verteilen. Diese Informationen bilden einen Hintergrund zur Interpretation der Analysen zur Heterogenität erreichter Kompetenzen (Kapitel 6) und zur sozialen Heterogenität (Kapitel 8).

Abbildung 9.1 zeigt, wie sich die Schülerinnen und Schüler mit mindestens einem im Ausland geborenen Elternteil auf die Schulen der Länder in der Primarstufe in den Jahren 2011 und 2016 verteilen.⁹ Veränderungen in diesen Verteilungen werden dabei nur deskriptiv dargestellt und nicht auf statistische Signifikanz geprüft. Es werden fünf Gruppen von Schulen unterschieden: Schulen, in denen der Anteil von Kindern mit Zuwanderungshintergrund

- (1) bei weniger als 5 Prozent,
- (2) zwischen 5 und 19 Prozent,
- (3) zwischen 20 und 39 Prozent,
- (4) zwischen 40 und 59 Prozent oder
- (5) bei mindestens 60 Prozent liegt.

Die Prozentangaben innerhalb der Balken in Abbildung 9.1 geben an, welcher prozentuale Anteil der Schulen eines Landes der jeweiligen Schulgruppe zuzuordnen ist. Grau gefärbte Balken kennzeichnen die Verteilung im Jahr 2016, türkis gefärbte Balken die Verteilung im Jahr 2011.

Bundesweit zeigt sich, dass im Jahr 2016 an rund 13 Prozent der Schulen der Anteil von Viertklässlerinnen und Viertklässlern mit Zuwanderungshintergrund bei unter 5 Prozent liegt. In etwa je einem Drittel aller Schulen liegt der Anteil von Kindern mit Zuwanderungshintergrund „zwischen 5 und 19 Prozent“ beziehungsweise „zwischen 20 und 39 Prozent“.

Vergleicht man die Anteile von Kindern mit Zuwanderungshintergrund an Schulen zwischen den Jahren 2011 und 2016, so zeigt sich bundesweit, dass sich im Jahr 2016 die Anzahl der Schulen, in denen bis zu 39 Prozent („weniger als 5 %“, „5–19 %“ und „20–39 %“) der Kinder einen Zuwanderungshintergrund haben, reduziert hat. Demgegenüber vergrößerten sich die Gruppen der Schulen mit höheren Anteilen von Kindern mit Zuwanderungshintergrund („40–59 %“ sowie „mindestens 60 %“). Dies weist darauf hin, dass sich die Schulen bei einem insgesamt gestiegenen Anteil von Schülerinnen und Schülern aus zugewanderten Familien im Jahr 2016 und einer etwas gleichmäßigeren Verteilung auf die fünf Schulgruppen etwas stärker in ihrer zugewanderungsbezogenen Zusammensetzung gleichen als im Jahr 2011.

⁹ Die Darstellung umfasst allgemeine Schulen und Förderschulen und beruht auf gewichteten Werten auf Schulebene. Für weitere Informationen zu Kindern mit Zuwanderungshintergrund an allgemeinen Schulen und Förderschulen siehe Kapitel 10.3.

Abbildung 9.1: Verteilung von Schülerinnen und Schülern mit Zuwanderungshintergrund auf die Schulen in den Ländern in den Jahren 2011 und 2016



In der zuwanderungsbezogenen Zusammensetzung der Schulen im Jahr 2016 unterscheiden sich die Länder zudem deutlich, wobei die Differenzen zwischen ostdeutschen Flächenländern einerseits und Stadtstaaten sowie einigen westdeutschen Flächenländern andererseits besonders auffällig sind und nachfolgend beschrieben werden. Wie aufgrund der vergleichsweise geringen Anteile von Kindern mit Zuwanderungshintergrund in den ostdeutschen Ländern zu erwarten ist (vgl. Tab. 9.1), finden sich dort im Jahr 2016 besonders viele Kinder in Schulen mit einem niedrigen Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund. Bei einem Drittel bis rund der Hälfte der Schulen dieser Länder liegt der Anteil von Kindern aus zugewanderten Familien bei unter 5 Prozent. Analog ist der Anteil von Schulen mit besonders vielen Kindern mit Zuwanderungshintergrund („mindestens 60 %“) sehr klein, in manchen Ländern gibt es sogar überhaupt keine Schulen dieser Gruppe. Im Vergleich zur Verteilung im Jahr 2011 hat sich der Anteil von Schulen mit wenigen Kindern mit Zuwanderungshintergrund jedoch auch in den ostdeutschen Ländern – mit Ausnahme von Thüringen mit einem gleichbleibenden Anteil – reduziert. Der Anteil von Schulen mit besonders vielen Kindern mit Zuwanderungshintergrund ist im Vergleich der Jahre 2011 und 2016 in den ostdeutschen Ländern dagegen stabil.

In den Stadtstaaten und den westdeutschen Flächenländern Baden-Württemberg, Hessen und Nordrhein-Westfalen verhält es sich hingegen umgekehrt. Hier liegt der Anteil von Kindern mit Zuwanderungshintergrund an rund einem Drittel bis der Hälfte der Schulen bei mindestens 40 Prozent, mit Ausnahme von Nordrhein-Westfalen, wo dieser Anteil bei 29 Prozent liegt. Es gibt damit besonders wenige Schulen mit einem sehr geringen Anteil und relativ viele Schulen mit einem besonders hohen Anteil von Kindern mit Zuwanderungshintergrund. Im Vergleich der Verteilungen der Jahre 2011 und 2016 zeigt sich, dass der Prozentsatz von Schulen mit einem geringen Zuwandereranteil in diesen Ländern weitgehend stabil geblieben ist und nur in Berlin abgenommen hat. Der Anteil von Schulen mit besonders vielen Kindern mit Zuwanderungshintergrund ist in Hamburg, Hessen und Nordrhein-Westfalen stabil geblieben, erhöht hat sich der Anteil in Baden-Württemberg und Bremen, verringert dagegen in Berlin. In Berlin sind im Jahr 2016 also mehr Schulen in den Schulgruppen mit einem mittleren Zuwandereranteil zu finden.

Zusammenfassend lässt sich festhalten, dass sich die Lernbedingungen von Viertklässlerinnen und Viertklässlern an Schulen der Primarstufe in Deutschland mit Blick auf zuwanderungsbezogene Zusammensetzung zwischen den Ländern und auch zwischen den einzelnen Schulen innerhalb der Länder deutlich unterscheiden. Bei einer Zunahme des Zuwandereranteils an den Schulen insgesamt wird tendenziell im Jahr 2016 gegenüber dem Jahr 2011 deutlich, dass sich die Unterschiede zwischen den Schulen in der zuwanderungsbezogenen Zusammensetzung ihrer Schülerschaft etwas reduziert haben.

9.4 Erreichte Kompetenzen von Viertklässlerinnen und Viertklässlern

9.4.1 Erreichte Kompetenzen von Viertklässlerinnen und Viertklässlern nach Zuwanderungsstatus

In diesem Abschnitt wird für die Fächer Deutsch und Mathematik berichtet, welches Kompetenzniveau Kinder ohne Zuwanderungshintergrund und Kinder aus zugewanderten Familien im Jahr 2016 durchschnittlich bis zur 4. Jahrgangsstufe erreicht haben und wie sich diese Kompetenzen zwischen den Jahren

2011 und 2016 verändert haben. Die Ergebnisse werden zunächst für Deutschland insgesamt nach Zuwanderungsstatus berichtet. Hierbei werden die folgenden Gruppen unterschieden: Kinder mit einem im Ausland geborenen Elternteil, Kinder der zweiten Zuwanderungsgeneration und Kinder der ersten Zuwanderungsgeneration.

Anschließend werden die in den Jahren 2011 und 2016 erreichten Kompetenzen in jedem Kompetenzbereich auf Länderebene beschrieben. Aufgrund der geringen Anteile von Schülerinnen und Schülern der ersten Generation kann hier nicht zwischen der ersten und zweiten Zuwanderungsgeneration unterschieden werden (vgl. Abschnitt 9.3.1), stattdessen werden beide Gruppen zur Kategorie *beide Elternteile im Ausland geboren* zusammengefasst. Zusätzlich wird sowohl für die Analysen auf Bundesebene als auch für die länderspezifischen Analysen jeweils grafisch dargestellt, inwieweit die von Kindern ohne Zuwanderungshintergrund und von Kindern mit Zuwanderungshintergrund durchschnittlich erreichten Kompetenzen im Jahr 2016 vom bundesweiten Gesamtmittelwert abweichen.¹⁰

Beim Vergleich der zuwanderungsbezogenen Disparitäten zwischen den Ländern sowie bei der Interpretation der Veränderungen ist durchgängig zu beachten, dass die Ergebnisse für Hamburg und das Saarland aufgrund des erheblichen Anteils an fehlenden Werten unter Vorbehalt stehen (vgl. Kapitel 3.1).

Informationen zur Darstellung der Ergebnisse in den Jahren 2011 und 2016 im Vergleich

In diesem Kapitel werden Trendabbildungen mit Informationen über die erreichten Kompetenzen von Viertklässlerinnen und Viertklässlern in den Jahren 2011 und 2016 sowie über Veränderungen in den erreichten Kompetenzen präsentiert. Weitere Tabellen und Abbildungen zu Trendanalysen und zu Ergebnissen querschnittlicher Analysen für das Jahr 2016, die im Text durch den Zusatz „web“ gekennzeichnet sind, können auf der Webseite des IQB als Zusatzmaterial abgerufen werden.

Zu den Trendanalysen, deren Befunde nachfolgend dargestellt werden (siehe Abb. 9.2 bis 9.4 und 9.6 bis 9.11), sind für Deutschland insgesamt und für jedes Land zunächst die prozentualen Anteile der untersuchten Gruppen im Jahr 2016 sowie Veränderungen in diesen Anteilen seit 2011 dargestellt. Weiterhin werden für die Jahre 2011 und 2016 die Stichprobengrößen (N), die im Mittel erreichten Kompetenzwerte (M), die Standardfehler der Mittelwerte (SE) und die Standardabweichungen der Kompetenzwerte (SD) angegeben. Für jeden der beiden Erhebungszeitpunkte wird außerdem berichtet, ob sich die Mittelwerte der Schülerinnen und Schüler mit Zuwanderungshintergrund signifikant vom jeweiligen Mittelwert der Schülerinnen und Schüler ohne Zuwanderungshintergrund unterscheiden. Dies ist in den Tabellenteilen der Abbildungen mit dem hochgestellten Buchstaben „a“ hinter dem jeweiligen Mittelwert gekennzeichnet. Die letzten drei Spalten des Tabellenteils beziehen sich auf die Veränderungen der im Durchschnitt erreichten Kompetenzen innerhalb jeder Gruppe. Die Veränderungen (ΔM) sowie ihre Standardfehler (SE) werden in Punkten auf der Berichtsmetrik angegeben. Zudem wird geprüft, ob sich die Disparitäten in Deutschland und gegebenenfalls innerhalb der Länder

¹⁰ Die Abweichungen vom bundesweiten Gesamtmittelwert werden jedoch im Text nicht beschrieben; aus Platzgründen geht die Darstellung nur auf die Unterschiede zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund ein.

signifikant verändert haben. Dies wäre der Fall, wenn sich die *Trends* zwischen Schülerinnen und Schülern ohne Zuwanderungshintergrund und Schülerinnen und Schülern aus zugewanderten Familien signifikant unterscheiden.

Signifikante Gruppenunterschiede in Veränderungen werden ebenfalls durch den hochgestellten Buchstaben „a“ in der Spalte ΔM markiert. Zusätzlich wird die standardisierte Mittelwertsdifferenz (d) für die Kompetenzunterschiede zwischen 2011 und 2016 angegeben (vgl. Kapitel 3.2). Die Differenzen kennzeichnen, wie stark sich die im Mittel erreichten Kompetenzen *innerhalb* der jeweiligen Gruppe über die Zeit verändert haben.

Im rechten Teil der Abbildungen wird jeweils grafisch dargestellt, inwieweit die Kompetenzen, die die einzelnen Zuwanderergruppen in Deutschland und in den Ländern im Jahr 2016 erreichen, vom bundesweiten Gesamtmittelwert abweichen. Damit bilden die grafischen Darstellungen zu den Trendanalysen etwas anderes ab als die übrigen Trendabbildungen im Berichtsband, die üblicherweise zeitliche Veränderungen illustrieren. Die Visualisierung der Kompetenzveränderungen ist im Zusatzmaterial auf der Website des IQB enthalten, in denen für jede Gruppe die Differenzen zwischen den im Jahr 2011 und 2016 im Durchschnitt erreichten Kompetenzen in einem Balkendiagramm dargestellt sind.

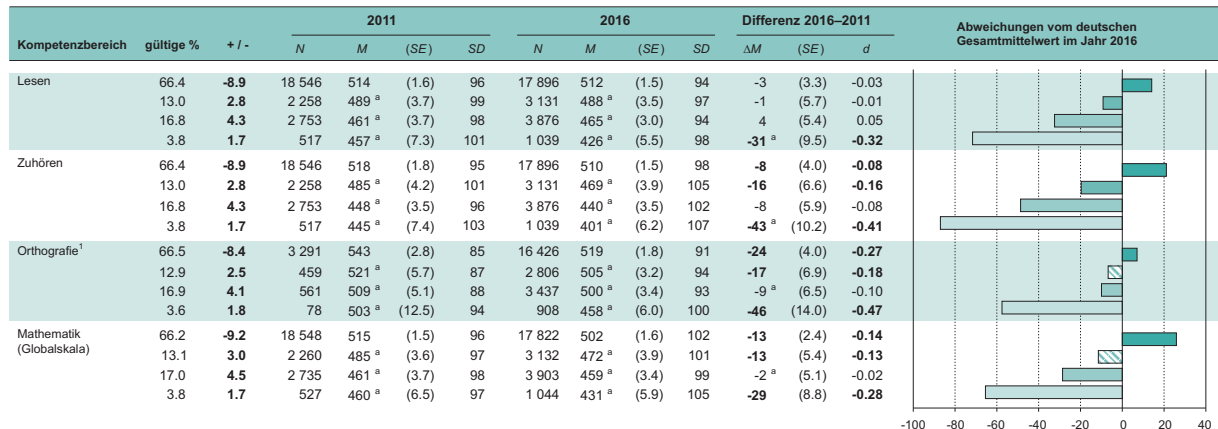
Erreichte Kompetenzen in den Fächern Deutsch und Mathematik nach Zuwanderungsstatus in Deutschland

In Abbildung 9.2 ist differenziert nach Zuwanderungsstatus dargestellt, welches Kompetenzniveau Viertklässlerinnen und Viertklässler in Deutschland insgesamt in den Kompetenzbereichen *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch sowie im Fach Mathematik (*Globalskala*) in den Jahren 2011 und 2016 erreicht haben. Signifikante Unterschiede zwischen Kindern mit und ohne Zuwanderungshintergrund sind durch den hochgestellten Buchstaben „a“ gekennzeichnet.

Im Jahr 2016 bestehen für alle betrachteten Gruppen ausgeprägte zuwanderungsbezogene Disparitäten. Schülerinnen und Schüler aus zugewanderten Familien erreichen in allen vier Kompetenzbereichen beziehungsweise Teilbereichen im Durchschnitt signifikant geringere Kompetenzen als Schülerinnen und Schüler ohne Zuwanderungshintergrund. Diese Unterschiede sind für Kinder der ersten Generation, die selbst mit ihren Eltern nach Deutschland zugewandert sind, durchgängig am stärksten ausgeprägt. Dieses Ergebnis könnte damit zusammenhängen, dass ein Teil dieser Schülerinnen und Schüler nicht die gesamte Schullaufbahn in Deutschland verbracht hat. Die Lernzeit dieser Schülerinnen und Schüler innerhalb des deutschen Bildungssystems ist somit geringer als für Schülerinnen und Schüler der zweiten Generation, die in Deutschland geboren sind. Entsprechend sind die Kompetenznachteile in der zweiten Zuwanderergeneration deutlich kleiner als in der ersten Zuwanderergeneration, was ein Hinweis auf gelingende Integrationsprozesse sein kann. Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil weisen durchgängig die geringsten Disparitäten auf. Für alle Gruppen sind die Disparitäten im *Zuhören* besonders ausgeprägt, während sie in *Orthografie* vergleichsweise gering ausfallen (vgl. Abb. 9.1web).¹¹

¹¹ Erläuterungen zu Abbildungen mit Ergebnissen querschnittlicher Analysen für das Jahr 2016, wie beispielsweise in Abbildung 9.1web dargestellt, werden weiter unten gegeben (siehe Abschnitt zu den erreichten Kompetenzen im Bereich *Orthografie*).

Abbildung 9.2: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Unterschiede nach Zuwanderungsstatus in den Jahren 2011 und 2016 im Vergleich und Abweichungen vom deutschen Gesamtmittelwert pro Kompetenzbereich in den Fächern Deutsch und Mathematik im Jahr 2016 in Deutschland



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen.
 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren)
 2. Zeile: Kinder mit einem im Ausland geborenen Elternteil
 3. Zeile: Kinder der 2. Generation (beide Elternteile im Ausland geboren, das Kind selbst in Deutschland geboren).
 4. Zeile: Kinder der 1. Generation (sowohl beide Elternteile als auch das Kind im Ausland geboren).
 +/- = Veränderung gegenüber IQB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung;
 ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.
¹ Die Berechnungen für den Kompetenzbereich Orthografie des Tabellenteils wurden ohne Schülerinnen und Schüler mit sonderpädagogischen Förderbedarf durchgeführt, die im Abbildungsteil angegebenen Werte beziehen sich hingegen auf querschnittliche Analysen aller Schülerinnen und Schüler.
^a signifikante Differenz (p < .05) zu Kindern ohne Zuwanderungshintergrund.
 Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz zum deutschen Gesamtmittelwert im jeweiligen Kompetenzbereich an.
 Fett gedruckte Differenzen sind statistisch signifikant (p < .05).

In Abbildung 9.2 sind zudem Veränderungen in den durchschnittlich erzielten Kompetenzen zwischen den Jahren 2011 und 2016 numerisch aufgeführt, die in Abbildung 9.2web grafisch veranschaulicht werden. Signifikante Trends der erreichten Kompetenzen innerhalb der Gruppen sind in Abbildung 9.2 durch Fettdruck von Differenzwerten hervorgehoben. Besonders große Veränderungen in den durchschnittlich erreichten Kompetenzen zeigen sich für die Gruppe von Schülerinnen und Schülern, die mit ihren Eltern nach Deutschland zugewandert sind (erste Generation). Kinder der ersten Generation erreichen im Jahr 2016 insbesondere im *Zuhören* (–43 Punkte) und in der *Orthografie* (–46 Punkte), aber auch im *Lesen* (–31 Punkte) und in *Mathematik* (–29 Punkte) deutlich geringere Kompetenzen als im Jahr 2011. Bei den Kindern der zweiten Generation haben sich dagegen die im Durchschnitt erreichten Kompetenzen im Verlauf von fünf Jahren in keinem Kompetenzbereich signifikant verändert. Für Kinder mit einem im Ausland geborenen Elternteil und für Kinder ohne Zuwanderungshintergrund ist wiederum eine signifikante Abnahme der im Mittel erreichten Kompetenzen in den Bereichen *Zuhören*, *Orthografie* und in *Mathematik* zu verzeichnen.

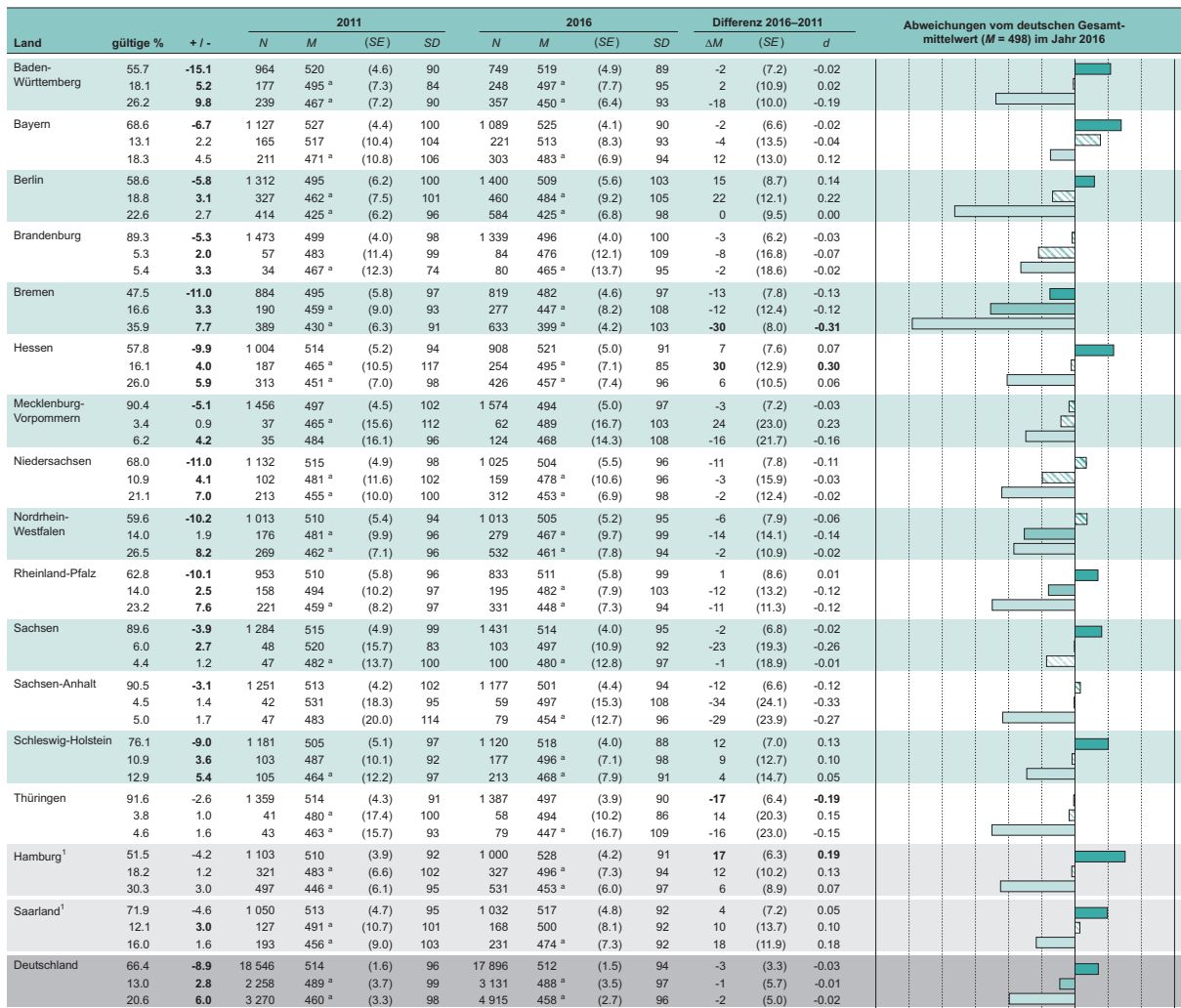
Die zuwanderungsbezogenen Disparitäten haben sich für Kinder der ersten Generation in den Bereichen *Lesen* und *Zuhören* zwischen den Jahren 2011 und 2016 signifikant vergrößert. So haben die im Mittel erreichten Kompetenzen dieser Gruppe stärker abgenommen als die im Mittel erreichten Kompetenzen der Kinder ohne Zuwanderungshintergrund (siehe hochgestelltes „a“ in der Spalte ΔM). Für Kinder der zweiten Generation haben sich die Disparitäten in den Bereichen *Orthografie* und *Mathematik* hingegen signifikant verringert, was auf die stärkere Abnahme erreichter Kompetenzen in der Gruppe von Kindern ohne Zuwanderungshintergrund zurückzuführen ist (vgl. auch Abb. 9.2web).

Erreichte Kompetenzen im Bereich Lesen nach Zuwanderungsstatus in den Ländern

Abbildung 9.3 zeigt, dass im Jahr 2016 im Lesen erhebliche zuwanderungsbezogene Disparitäten im Kompetenzerwerb bestehen. Der Abstand zwischen Kindern ohne Zuwanderungshintergrund und Kindern mit zwei im Ausland geborenen Elternteilen liegt in Deutschland insgesamt bei 54 Punkten, was dem zu erwartenden Kompetenzzuwachs von etwas weniger als einem Schuljahr entspricht (60 Punkte, Böhme & Bremerich-Vos, 2009; Kapitel 6.1). Der Abstand von Schülerinnen und Schülern mit einem im Ausland geborenen Elternteil im Vergleich zu Kindern ohne Zuwanderungshintergrund ist mit 24 Punkten auf der Kompetenzskala deutlich kleiner, aber ebenfalls signifikant.

Die länderspezifischen Analysen zeigen, dass die Mittelwertsunterschiede im Lesen zwischen Kindern mit zwei im Ausland geborenen Elternteilen und

Abbildung 9.3: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich Lesen im Fach Deutsch nach Zuwanderungsstatus und Land in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016



Anmerkungen: In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen.
 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren).

2. Zeile: Kinder mit einem im Ausland geborenen Elternteil.

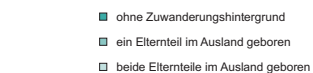
3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.

ΔM = Veränderung gegenüber IQB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.

¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

^a signifikante Differenz (p < .05) zu Kindern ohne Zuwanderungshintergrund.

Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Fett gedruckte Differenzen sind statistisch signifikant (p < .05).



Kindern ohne Zuwanderungshintergrund im Jahr 2016 in nahezu allen Ländern statistisch signifikant sind. Eine Ausnahme bildet Mecklenburg-Vorpommern mit der im Vergleich der Länder kleinsten Differenz, die nicht statistisch signifikant ist. Insgesamt variiert die Größe der Differenzen zwischen den Ländern erheblich. Während in den Stadtstaaten die größten Disparitäten zwischen den beiden Gruppen zu finden sind (Berlin: 84 Punkte, Bremen: 83 Punkte, Hamburg: 75 Punkte), fällt der Unterschied in einigen der ostdeutschen Flächenländer mit einem relativ geringen Anteil von Schülerinnen und Schülern aus zugewanderten Familien am niedrigsten aus.

Die Kompetenzen, die Schülerinnen und Schüler mit einem im Ausland und einem in Deutschland geborenen Elternteil im *Lesen* im Durchschnitt erreichen, unterscheiden sich in etwa der Hälfte der Länder nicht signifikant von den durchschnittlichen Kompetenzen der Kinder ohne Zuwanderungshintergrund. Signifikant geringere Kompetenzen von Kindern mit einem im Ausland geborenen Elternteil finden sich in Baden-Württemberg, Berlin, Bremen, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, Schleswig-Holstein und Hamburg. Besonders groß sind die Nachteile dieser Gruppe in Bremen (35 Punkte), Nordrhein-Westfalen (38 Punkte) und Hamburg (32 Punkte), siehe auch Abbildung 9.3web.

Bundesweit ergeben sich für die drei Gruppen keine signifikanten Unterschiede zwischen der im Jahr 2011 und der im Jahr 2016 im Durchschnitt erreichten Lesekompetenz.¹² Innerhalb der Länder sind die Kompetenzen von Kindern mit Zuwanderungshintergrund in nahezu allen Ländern stabil geblieben. Lediglich in Bremen ist für die Gruppe von Kindern mit zwei im Ausland geborenen Elternteilen eine signifikante Abnahme zu verzeichnen (–30 Punkte); in Hessen haben hingegen Kinder mit einem im Ausland geborenen Elternteil im Jahr 2016 statistisch signifikant höhere Kompetenzen erreicht als im Jahr 2011 (+30 Punkte). Die Kompetenzen von Kindern ohne Zuwanderungshintergrund fallen im Jahr 2016 in Thüringen (–17 Punkte) signifikant niedriger aus als im Jahr 2011. In Hamburg ist entgegen dem bundesweiten Trend eine signifikante Steigerung der im Durchschnitt erreichten Kompetenzen bei Schülerinnen und Schülern ohne Zuwanderungshintergrund um 17 Punkte zu verzeichnen (vgl. auch Abb. 9.4web).

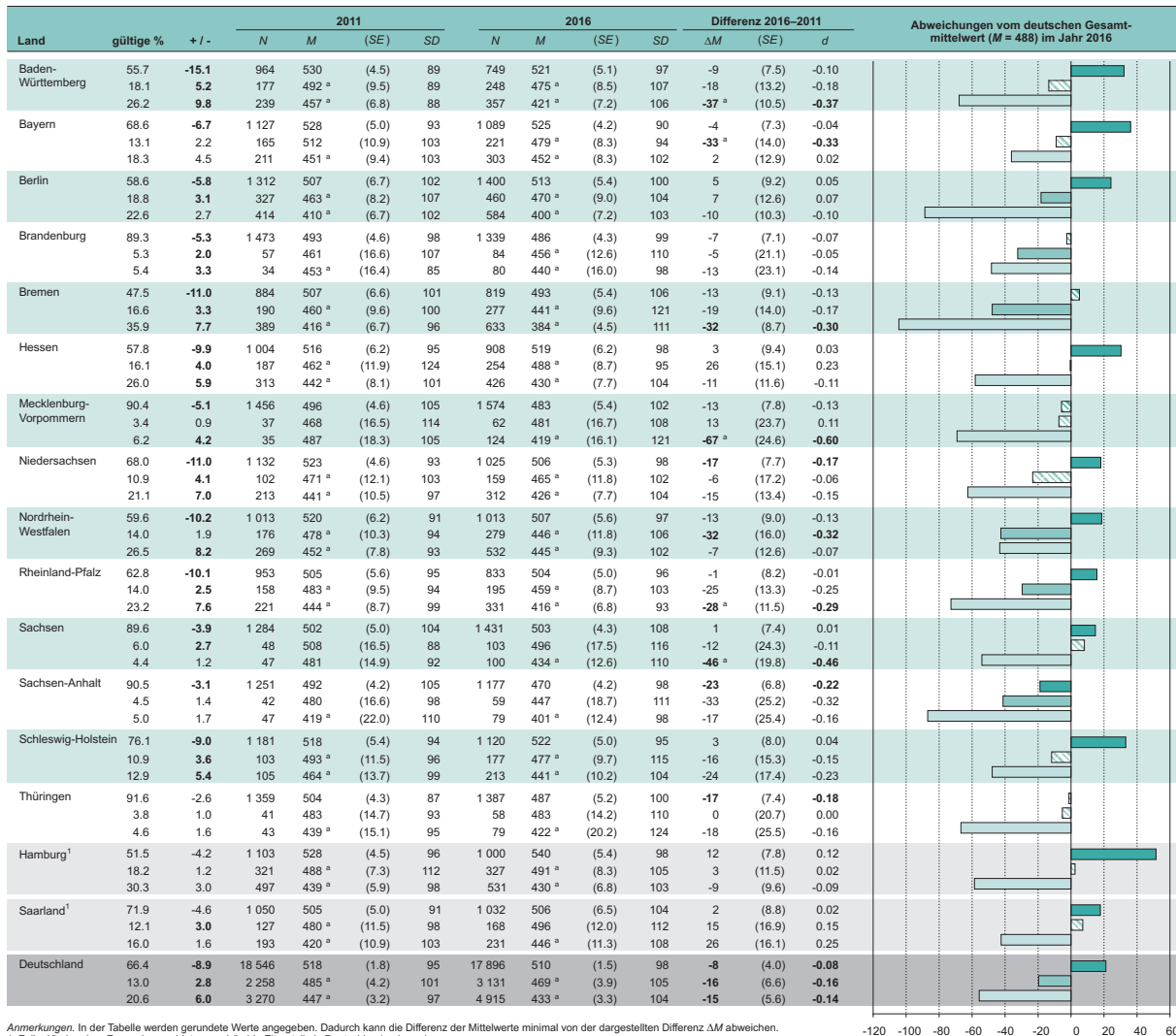
Weder bundesweit noch in den einzelnen Ländern unterscheiden sich die Trends signifikant zwischen Kindern mit Zuwanderungshintergrund und Kindern ohne Zuwanderungshintergrund; damit haben sich auch die zuwanderungsbezogenen Disparitäten im *Lesen* seit dem Jahr 2011 nicht signifikant verändert.

Erreichte Kompetenzen im Bereich *Zuhören* nach Zuwanderungsstatus in den Ländern

Für den Kompetenzbereich *Zuhören* sind im Jahr 2016 ähnliche Befundmuster festzustellen wie für den Kompetenzbereich *Lesen*; die Differenzen zwischen den Gruppen fallen jedoch im Allgemeinen etwas größer aus (siehe Abb. 9.4). Bundesweit erreichen Schülerinnen und Schüler aus zugewanderten Familien im Durchschnitt signifikant niedrigere Kompetenzen als Schülerinnen und Schüler

12 Die in Kapitel 6.1 berichteten bundesweiten Veränderungen im Kompetenzbereich *Lesen* weichen von den hier dargestellten Werten ab, da sich die in Abbildung 9.3 berichteten Ergebnisse ausschließlich auf Kinder beziehen, die eindeutig einem Zuwanderungsstatus zugeordnet werden können.

Abbildung 9.4: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich *Zuhören* im Fach Deutsch nach Zuwanderungsstatus und Land in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen.
 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren).
 2. Zeile: Kinder mit einem im Ausland geborenen Elternteil.
 3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.
 +/- = Veränderung gegenüber IQB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung;
 ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.
¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).
^a signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund.
 Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

ohne Zuwanderungshintergrund. Die Disparitäten betragen 77 Punkte für Kinder mit zwei im Ausland geborenen Elternteilen, was bei einer Differenz von 60 Punkten zwischen der 3. und 4. Jahrgangsstufe dem zu erwartenden Lernzuwachs von etwas mehr als einem Schuljahr entspricht (Behrens, Böhme & Krelle, 2009; vgl. Kapitel 6.1), und 41 Punkte für Kinder mit einem im Ausland geborenen Elternteil. Für Schülerinnen und Schüler mit zwei im Ausland geborenen Elternteilen sind in allen Ländern signifikante Kompetenznachteile zu verzeichnen, die in Berlin (113 Punkte), Bremen (109 Punkte) und Hamburg (110 Punkte) besonders ausgeprägt sind. Die Kompetenznachteile von Schülerinnen und Schülern mit einem im Ausland geborenen Elternteil sind demgegenüber deutlich geringer und nicht in allen Ländern statistisch signifikant (vgl. auch Abb. 9.5web).

Bundesweit zeigen sich signifikante Unterschiede zwischen den im Jahr 2011 und den im Jahr 2016 im Durchschnitt erreichten Kompetenzen im *Zuhören* für alle drei Schülergruppen. In etwa der Hälfte der Länder haben die Kinder mit

Zuwanderungshintergrund im Jahr 2016 durchschnittlich geringere Kompetenzen erreicht als im Jahr 2011. Besonders ausgeprägt ist diese Abnahme für Kinder mit zwei im Ausland geborenen Elternteilen in Mecklenburg-Vorpommern (–67 Punkte) und Sachsen (–46 Punkte). Kinder mit einem im Ausland geborenen Elternteil erzielten im Jahr 2016 in Bayern (–33 Punkte) und Nordrhein-Westfalen (–32 Punkte) signifikant geringere Kompetenzen als im Jahr 2011. In Niedersachsen (–17 Punkte), Sachsen-Anhalt (–23 Punkte) und Thüringen (–17 Punkte) erreichen Kinder *ohne* Zuwanderungshintergrund im Jahr 2016 signifikant geringere Kompetenzen als im Jahr 2011.

Zudem haben sich die zuwanderungsbezogenen Disparitäten im *Zuhören* in einigen Ländern signifikant verändert. Der Unterschied zwischen den Kompetenzen, die Schülerinnen und Schüler mit zwei im Ausland geborenen Elternteilen erreichen, und den Kompetenzen von Kindern ohne Zuwanderungshintergrund hat sich in Baden-Württemberg, Mecklenburg-Vorpommern, Rheinland-Pfalz und Sachsen signifikant vergrößert. In Bayern haben sich die zuwanderungsbezogenen Disparitäten für Kinder mit einem im Ausland geborenen Elternteil im Jahr 2016 gegenüber dem Jahr 2011 vergrößert (vgl. auch Abb. 9.6web).

Erreichte Kompetenzen im Bereich *Orthografie* nach Zuwanderungsstatus in den Ländern

Die Kompetenzen im Bereich *Orthografie* wurden im Jahr 2011 nur in einer bundesweit repräsentativen Teilstichprobe erhoben, die weder eine Auswertung auf Länderebene erlaubte, noch Schülerinnen und Schüler in Förderschulen einbezog. Deshalb können Trendaussagen im Kompetenzbereich *Orthografie* lediglich für Deutschland insgesamt getroffen werden, nicht jedoch für die einzelnen Länder. Zudem beziehen sich diese Aussagen nur auf Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf (vgl. Kapitel 3.2). Aus diesem Grund werden für den Bereich *Orthografie*, abweichend von den Darstellungen zu den übrigen Kompetenzbereichen, die Ergebnisse im Jahr 2016 sowie die Veränderungen zwischen den Jahren 2011 und 2016 in getrennten Abbildungen präsentiert.

Informationen zur Darstellung der Ergebnisse im Jahr 2016

In den Abschnitten zu Kompetenzen der Schülerinnen und Schüler im Bereich *Orthografie*, zu Merkmalen der sozialen Integration und zur Lernfreude (Abschnitt 9.5) sowie in zusätzlich zur Verfügung stehenden Onlinematerialien werden die Ergebnisse gesondert für das Jahr 2016 berichtet. Die Mittelwerte von Schülerinnen und Schülern mit Zuwanderungshintergrund werden darin den Werten von Schülerinnen und Schülern ohne Zuwanderungshintergrund für das Jahr 2016 gegenübergestellt. Dazu sind in der Abbildung 9.5, den Tabellen 9.4 und 9.5 sowie im Zusatzmaterial auf der Website des IQB neben den Mittelwerten (M), Standardfehlern (SE) und Standardabweichungen (SD) auch standardisierte Mittelwertsdifferenzen (d) angegeben. Für die querschnittlichen Analysen im Jahr 2016 beziehen sich die standardisierten Mittelwertsdifferenzen auf Unterschiede zwischen der jeweiligen Zuwanderergruppe und der Gruppe der Kinder ohne Zuwanderungshintergrund. Positive d -Werte kennzeichnen einen Vorteil von Schülerinnen und Schülern aus zugewanderten Familien im Vergleich zur Gruppe von Kindern ohne Zuwanderungshintergrund. Im Gegensatz zu den

Tabellenteilen in Abbildungen zu Trendanalysen gibt die standardisierte Mittelwertsdifferenz (d) also *nicht* Veränderungen von Kompetenzen an, sondern Unterschiede zwischen Kindern mit Zuwanderungshintergrund und Kindern ohne Zuwanderungshintergrund.

Abbildung 9.5 zeigt, dass auch im Bereich *Orthografie* im Jahr 2016 statistisch signifikante Kompetenznachteile für Kinder aus zugewanderten Familien bestehen. Diese sind allerdings geringer ausgeprägt als im *Lesen* und *Zuhören*. In der *Orthografie* erreichen Kinder mit zwei im Ausland geborenen Elternteilen deutschlandweit im Mittel ein um 26 Punkte und Kinder mit einem im Ausland geborenen Elternteil ein um 14 Punkte geringeres Kompetenzniveau als Kinder ohne Zuwanderungshintergrund. Bei einer Punktwertdifferenz von 100 Punkten zwischen der 3. und 4. Jahrgangsstufe (vgl. Kapitel 6.1) entsprechen die ermittelten Disparitäten in der *Orthografie* somit weniger als dem Lernzuwachs, der in einem Drittel Schuljahr zu erwarten ist.

Betrachtet man die zugewanderungsbezogenen Disparitäten im Jahr 2016 im Ländervergleich, so wird deutlich, dass die Mittelwertsdifferenzen zum Nachteil von Kindern mit zwei im Ausland geborenen Elternteilen gegenüber Kindern ohne Zuwanderungshintergrund in einem Großteil der Länder statistisch signifikant sind (durch Fettdruck hervorgehoben). Auch im Bereich *Orthografie* variiert die Größe der Disparitäten zwischen den Ländern erheblich. Die größten Gruppenunterschiede bestehen in Bremen (52 Punkte), Mecklenburg-Vorpommern (40 Punkte) und Hamburg (59 Punkte), wohingegen sie in Bayern (18 Punkte), Nordrhein-Westfalen (21 Punkte) und Rheinland-Pfalz (16 Punkte) kleiner ausfallen. In Brandenburg, Sachsen, Sachsen-Anhalt, Schleswig-Holstein und im Saarland unterscheiden sich die Gruppen nicht signifikant. Auch die im Durchschnitt von Schülerinnen und Schülern mit einem im Ausland und einem in Deutschland geborenen Elternteil erreichten Kompetenzen im Bereich *Orthografie* unterscheiden sich in der Hälfte der Länder nicht signifikant vom Mittelwert der Kinder ohne Zuwanderungshintergrund. Besonders große Kompetenznachteile dieser Gruppe finden sich in Sachsen (29 Punkte), Sachsen-Anhalt (34 Punkte) und Hamburg (24 Punkte), wohingegen sie in Bayern und im Saarland mit 6 bis 7 Punkten relativ gering sind. Besonders auffällig ist zudem, dass in Bayern und dem Saarland sowohl Kinder ohne Zuwanderungshintergrund als auch beide Zuwanderergruppen im Mittel Kompetenzen erreichen, die signifikant über dem deutschen Gesamtmittelwert liegen.

Tabelle 9.2 zeigt, dass bundesweit die Kompetenzen im Bereich *Orthografie* zwischen den Jahren 2011 und 2016 in allen Gruppen signifikant abgenommen haben. Gleichzeitig haben sich die zugewanderungsbezogenen Disparitäten in diesem Zeitraum nicht signifikant verändert (vgl. auch Abb. 9.7web).

Abbildung 9.5: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede und Abweichungen vom deutschen Gesamtmittelwert im Kompetenzbereich *Orthografie* im Fach Deutsch nach Zuwanderungsstatus und Land im Jahr 2016



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente pro Land minimal von 100 abweichen.
 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren)
 2. Zeile: Kinder mit einem im Ausland geborenen Elternteil
 3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.
 N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; d = Effektstärke Cohens d.
¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).
 Fett gedruckte Mittelwertsangaben unterscheiden sich statistisch signifikant ($p < .05$) von den Mittelwerten von Kindern ohne Zuwanderungshintergrund. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz zum deutschen Gesamtmittelwert ($M = 504$) an.

■ ohne Zuwanderungshintergrund
 ■ ein Elternteil im Ausland geboren
 ■ beide Elternteile im Ausland geboren

Tabelle 9.2: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich *Orthografie* im Fach Deutsch nach Zuwanderungsstatus in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016

Land	gültige %	+ / -	2011				2016				Differenz 2016–2011		
			N	M	(SE)	SD	N	M	(SE)	SD	ΔM	(SE)	<i>d</i>
Deutschland	66.5	-8.4	3 291	543	(2.8)	85	16 426	519	(1.8)	91	-24	(4.0)	-0.27
	12.9	2.5	459	521	(5.7)	87	2 806	505	(3.2)	94	-17	(6.9)	-0.18
	20.5	6.0	639	508	(4.8)	89	4 345	492	(3.2)	95	-16	(6.2)	-0.17

Anmerkungen. Die Berechnungen wurden ohne Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf durchgeführt. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen und die Summe der Prozente kann minimal von 100 abweichen.

1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren)

2. Zeile: Kinder mit einem im Ausland geborenen Elternteil

3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.

+/- = Veränderung gegenüber IQB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; *d* = Effektstärke Cohens *d*.

^a signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund.

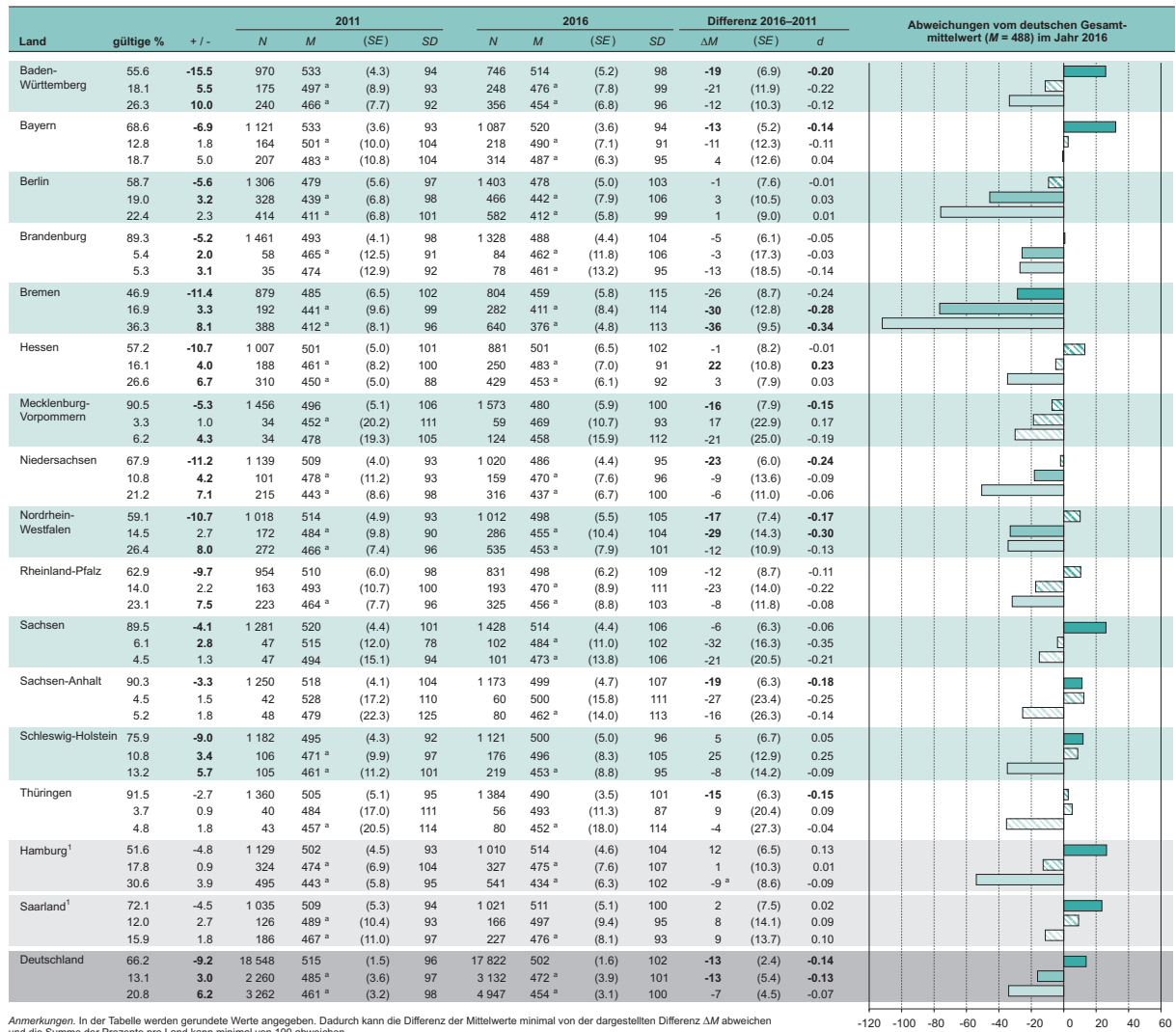
Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

Erreichte Kompetenzen im Fach Mathematik nach Zuwanderungsstatus in den Ländern

Im Fach Mathematik entsprechen die zuwanderungsbezogenen Disparitäten im Jahr 2016 in etwa denen im *Lesen* im Fach Deutsch. Bezogen auf die Global-skala mathematischer Kompetenzen beträgt der Abstand zwischen Kindern ohne Zuwanderungshintergrund und Kindern mit zwei im Ausland geborenen Elternteilen in Deutschland insgesamt 48 Punkte (vgl. Abb. 9.6) und ist damit ähnlich groß wie der entsprechende Unterschied im *Lesen* im Fach Deutsch von 54 Punkten. Für das Fach Mathematik entsprechen diese Disparitäten einem zu erwartenden Lernzuwachs von mehr als einem halben Schuljahr (80 Punkte entsprechen einem Schuljahr, Reiss & Winkelmann, 2009; vgl. Kapitel 6.2). Für Kinder mit einem im Ausland geborenen Elternteil liegt der Kompetenznachteil gegenüber Kindern ohne Zuwanderungshintergrund bei 30 Punkten. Damit sind auch die Disparitäten dieser Gruppen für das Fach Mathematik und den Bereich *Lesen* in etwa vergleichbar (23 Punkte).

In nahezu allen Ländern erreichen Kinder mit zwei im Ausland geborenen Elternteilen signifikant geringere Kompetenzen in Mathematik als Schülerinnen und Schüler ohne Zuwanderungshintergrund. Besonders große Kompetenznachteile für Schülerinnen und Schüler mit zwei im Ausland geborenen Elternteilen zeigen sich erneut in den Stadtstaaten (Berlin: 66 Punkte, Bremen: 83 Punkte, Hamburg: 80 Punkte). Die geringsten Disparitäten liegen in Brandenburg (27 Punkte) und Mecklenburg-Vorpommern (22 Punkte) vor, wobei Mecklenburg-Vorpommern das einzige Land ist, in dem der Unterschied nicht statistisch signifikant ist. Auch die Schülerinnen und Schüler mit einem im Ausland und einem in Deutschland geborenen Elternteil erzielen im Durchschnitt Kompetenzen, die in der Mehrzahl der Länder signifikant unter dem Mittelwert von Kindern ohne Zuwanderungshintergrund liegen. Besonders ausgeprägt sind die Disparitäten in Bremen (48 Punkte), Hamburg (39 Punkte) und Nordrhein-Westfalen (43 Punkte); gering und statistisch nicht signifikant sind sie hingegen in Mecklenburg-Vorpommern (11 Punkte) und Schleswig-Holstein (4 Punkte) (vgl. auch Abb. 9.8web).

Abbildung 9.6: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede in Mathematik (*Globalskala*) nach Zuwanderungsstatus und Land in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016



Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen und die Summe der Prozente pro Land kann minimal von 100 abweichen.
 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren).
 2. Zeile: Kinder mit einem im Ausland geborenen Elternteil.
 3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.
 +/- = Veränderung gegenüber IGB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung;
 ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.
¹ Die Befunde stützen sich auf einen erhöhten Anteil fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).
^a signifikante Differenz (p < .05) zu Kindern ohne Zuwanderungshintergrund.
 Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Fett gedruckte Differenzen sind statistisch signifikant (p < .05).

Für die Trendanalysen zeigt sich in Abbildung 9.6, dass die von Kindern mit zwei im Ausland geborenen Elternteilen erreichten Kompetenzen bundesweit im Verlauf von fünf Jahren stabil geblieben sind. Die Kompetenzwerte von Kindern mit einem im Ausland geborenen Elternteil (-13 Punkte) und von Kindern ohne Zuwanderungshintergrund (-13 Punkte) sind hingegen im Jahr 2016 signifikant geringer ausgeprägt als im Jahr 2011.

Innerhalb der Länder sind die Veränderungen der Kompetenzen von Kindern mit Zuwanderungshintergrund nur vereinzelt statistisch signifikant. Lediglich in Bremen erzielen Kinder mit zwei im Ausland geborenen Elternteilen im Jahr 2016 signifikant geringere Kompetenzen als im Jahr 2011 (-36 Punkte). Auch für Kinder mit einem im Ausland geborenen Elternteil sind in Bremen (-30 Punkte) und Nordrhein-Westfalen (-29 Punkte) signifikant geringere Kompetenzmittelwerte zu verzeichnen. Entgegen dem bundesweiten Trend ist in Hessen (+22 Punkte) eine signifikante Steigerung der im Durchschnitt er-

reichten Kompetenzen für Schülerinnen und Schüler mit einem im Ausland geborenen Elternteil zu beobachten. Die Kompetenzen von Kindern ohne Zuwanderungshintergrund fallen im Jahr 2016 in sieben Ländern (Baden-Württemberg, Bayern, Mecklenburg-Vorpommern, Niedersachsen, Nordrhein-Westfalen, Sachsen-Anhalt und Thüringen) signifikant niedriger aus als im Jahr 2011. Am stärksten ausgeprägt (–19 bis –23 Punkte) sind diese Unterschiede in Baden-Württemberg, Niedersachsen und Sachsen-Anhalt (vgl. auch Abb. 9.9web).

Die zuwanderungsbezogenen Disparitäten im Fach Mathematik sind zwischen den Jahren 2011 und 2016 konstant geblieben. Lediglich in Hamburg haben sich die Unterschiede zwischen Kindern mit zwei im Ausland geborenen Elternteilen und Kindern ohne Zuwanderungshintergrund signifikant vergrößert.

9.4.2 Erreichte Kompetenzen von Viertklässlerinnen und Viertklässlern verschiedener Herkunftsgruppen

In diesem Abschnitt werden die zuwanderungsbezogenen Disparitäten nach verschiedenen Herkunftsgruppen differenziert. Die Analysen dienen dazu, besondere Benachteiligungen zu identifizieren, die für einzelne Teilgruppen bestehen könnten, und die Entwicklung solcher Disparitäten im Bildungsmonitoring kontinuierlich zu beschreiben (vgl. Abschnitt 9.2.1). In den Analysen werden analog zum IQB-Ländervergleich 2011 insgesamt fünf Herkunftsregionen unterschieden, aus denen in den Jahren 2011 und 2016 die größten Zuwanderergruppen in Deutschland stammten: *Türkei*, *ehemalige Sowjetunion*, *Polen*, *ehemaliges Jugoslawien* und für die übrigen Regionen die Gruppe *anderes Land*. Für das Jahr 2016 können aufgrund der ausreichend großen Fallzahlen außerdem die im Mittel erreichten Kompetenzwerte für die Herkunftsgruppe *arabische Länder* bestimmt werden. Die Ergebnisse von Analysen zuwanderungsbezogener Disparitäten der sechs Herkunftsgruppen für das Jahr 2016 können auf der Webseite des IQB als Zusatzmaterial abgerufen werden.¹³ Informationen zur Darstellung und Interpretation der im Folgenden präsentierten Abbildungen sind Abschnitt 9.4.1 zu entnehmen.

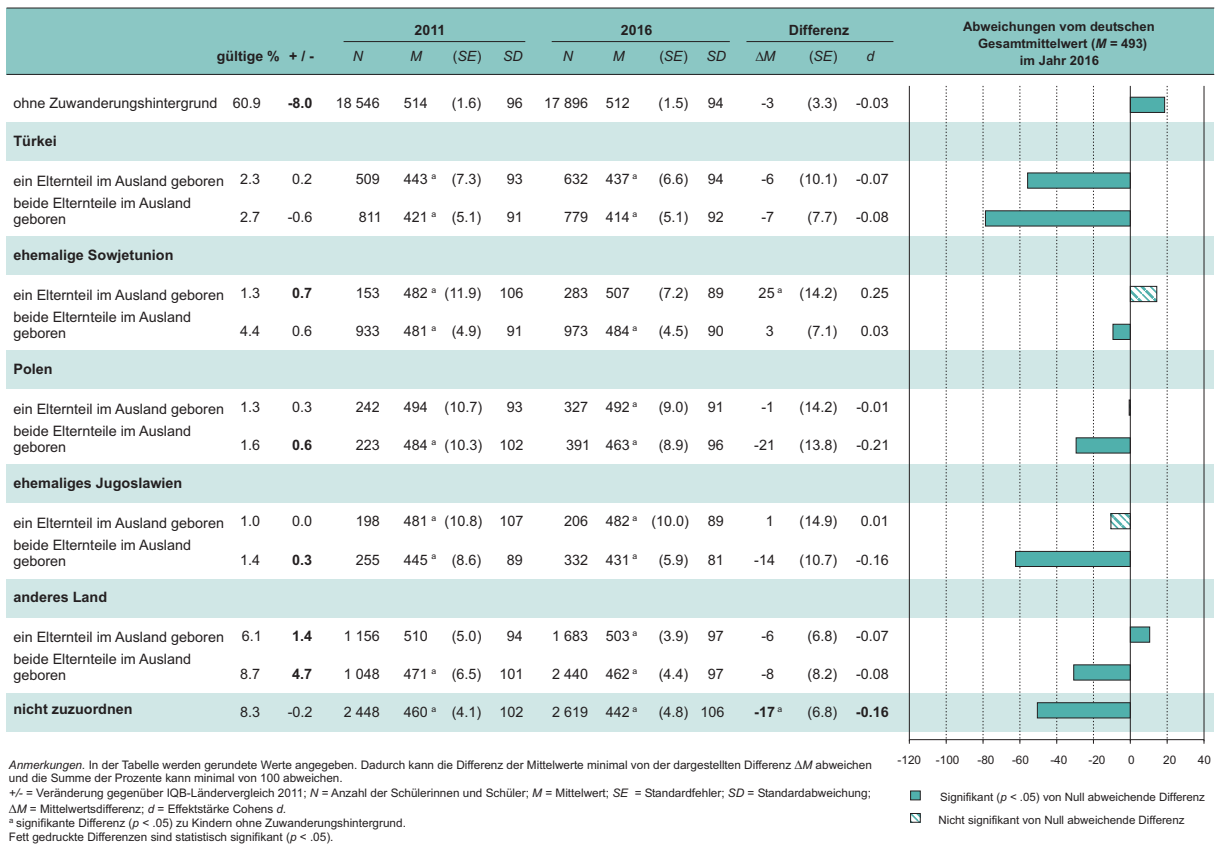
Erreichte Kompetenzen im Bereich *Lesen* nach Herkunftsgruppen

Abbildung 9.7 zeigt, dass bundesweit im Jahr 2016 für nahezu alle Herkunftsgruppen statistisch signifikante Unterschiede zu Kindern ohne Zuwanderungshintergrund bestehen (vgl. auch Abb. 9.10web). Eine Ausnahme bilden Kinder mit einem in der ehemaligen Sowjetunion geborenen Elternteil, deren Kompetenzen sich im Mittel nicht signifikant von denen der Kinder ohne Zuwanderungshintergrund unterscheiden. Zudem haben sich die im Durchschnitt erreichten Lesekompetenzen in keiner der betrachteten Herkunftsgruppen zwischen 2011 und 2016 signifikant verändert (vgl. auch Abb. 9.11web).

Die Höhe der zuwanderungsbezogenen Disparitäten ist im Bereich *Lesen* für nahezu alle Herkunftsgruppen weitgehend stabil geblieben, da sich die

¹³ Die Kinder der Herkunftsgruppe *arabische Länder* wurden für die Trendanalysen der Herkunftsgruppe *anderes Land* zugeordnet. Entsprechend unterscheidet sich die Zusammensetzung der Gruppe *anderes Land* zwischen den im vorliegenden Kapitel berichteten Trendanalysen und den im Zusatzmaterial verfügbaren Analysen für das Jahr 2016.

Abbildung 9.7: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich *Lesen* im Fach Deutsch nach Herkunftsgruppen in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016



Veränderungen in den erreichten Kompetenzen zwischen den Jahren 2011 und 2016 nicht bedeutsam zwischen den Gruppen unterscheiden. Nur für die Gruppe von Kindern mit einem in der ehemaligen Sowjetunion geborenen Elternteil ist eine signifikante Verringerung der Disparitäten festzustellen.¹⁴

Erreichte Kompetenzen im Bereich *Zuhören* nach Herkunftsgruppen

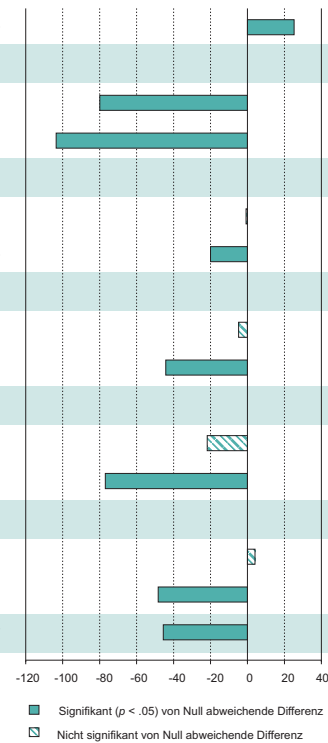
In Abbildung 9.8 sind die Kompetenzmittelwerte im Jahr 2016 und deren Veränderungen in den jeweiligen Herkunftsgruppen für den Kompetenzbereich *Zuhören* dargestellt. Im Jahr 2016 zeigen sich für alle Herkunftsgruppen signifikante zugewanderungsbezogene Disparitäten, das heißt, Schülerinnen und Schüler mit Zuwanderungshintergrund erreichen im Durchschnitt durchgängig gerin-

¹⁴ Eine Veränderung der Disparitäten liegt vor, wenn sich die Trends der erreichten Kompetenzen zwischen zwei Gruppen signifikant unterscheiden und sich eine Gruppe somit statistisch signifikant stärker verändert als die andere Gruppe. Die statistische Signifikanz der Veränderung wird jedoch nicht nur von der Höhe der Veränderung, sondern auch von der Unsicherheit der Schätzung beeinflusst (siehe Kapitel 3.2 und 13.1). Die Prüfung der statistischen Signifikanz von Gruppenunterschieden in Veränderungen bezieht mehr Unsicherheitsquellen ein als die querschnittliche Signifikanzprüfung von Disparitäten. Daher kann es also vorkommen, dass Disparitäten zwischen zwei Gruppen, wie zum Beispiel in Abbildung 9.7 bei Kindern mit einem in einem *anderen Land* geborenen Elternteil und Kindern ohne Zuwanderungshintergrund in den Jahren 2011 und 2016 einmal statistisch signifikant sind (im Beispiel im Jahr 2016) und einmal nicht (im Beispiel im Jahr 2011), sich die Veränderungen innerhalb der beiden Gruppen aber nicht signifikant voneinander unterscheiden, da sie mit größerer Unsicherheit behaftet sind.

Abbildung 9.8: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich *Zuhören* im Fach Deutsch nach Herkunftsgruppen in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016

	gültige %	+ / -	2011				2016				Differenz			Abweichungen vom deutschen Gesamtmittelwert (M = 484) im Jahr 2016
			N	M	(SE)	SD	N	M	(SE)	SD	ΔM	(SE)	d	
ohne Zuwanderungshintergrund	60.9	-8.0	18 546	518	(1.8)	95	17 896	510	(1.5)	98	-8	(4.0)	-0.08	
Türkei														
ein Elternteil im Ausland geboren	2.3	0.2	509	433 ^a	(8.6)	96	632	404 ^a	(6.8)	99	-28	(11.4)	-0.29	
beide Elternteile im Ausland geboren	2.7	-0.6	811	408 ^a	(5.5)	91	779	381 ^a	(5.9)	97	-27 ^a	(8.7)	-0.29	
ehemalige Sowjetunion														
ein Elternteil im Ausland geboren	1.3	0.7	153	485 ^a	(13.7)	115	283	484 ^a	(8.2)	95	-1	(16.2)	-0.01	
beide Elternteile im Ausland geboren	4.4	0.6	933	472 ^a	(5.3)	90	973	464 ^a	(4.8)	100	-8	(7.8)	-0.08	
Polen														
ein Elternteil im Ausland geboren	1.3	0.3	242	493 ^a	(11.1)	94	327	479 ^a	(9.1)	96	-14	(14.7)	-0.14	
beide Elternteile im Ausland geboren	1.6	0.6	223	468 ^a	(9.5)	96	391	440 ^a	(9.8)	102	-28	(14.0)	-0.29	
ehemaliges Jugoslawien														
ein Elternteil im Ausland geboren	1.0	0.0	198	479 ^a	(11.3)	104	206	463 ^a	(11.6)	100	-16	(16.5)	-0.16	
beide Elternteile im Ausland geboren	1.4	0.3	255	435 ^a	(8.3)	91	332	407 ^a	(7.6)	87	-28	(11.7)	-0.31	
anderes Land														
ein Elternteil im Ausland geboren	6.1	1.4	1 156	507 ^a	(5.3)	94	1 683	488 ^a	(3.7)	102	-19	(7.2)	-0.19	
beide Elternteile im Ausland geboren	8.7	4.7	1 048	455 ^a	(5.9)	98	2 440	436 ^a	(4.8)	104	-19	(8.2)	-0.19	
nicht zuzuordnen	8.3	-0.2	2 448	457 ^a	(4.1)	102	2 619	439 ^a	(5.8)	115	-18	(7.8)	-0.17	

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen und die Summe der Prozente kann minimal von 100 abweichen.
 +/- = Veränderung gegenüber IGB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung;
 ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d.
^a = signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund.
 Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).



gere Kompetenzen als Kinder ohne Zuwanderungshintergrund (vgl. auch Abb. 9.12web).

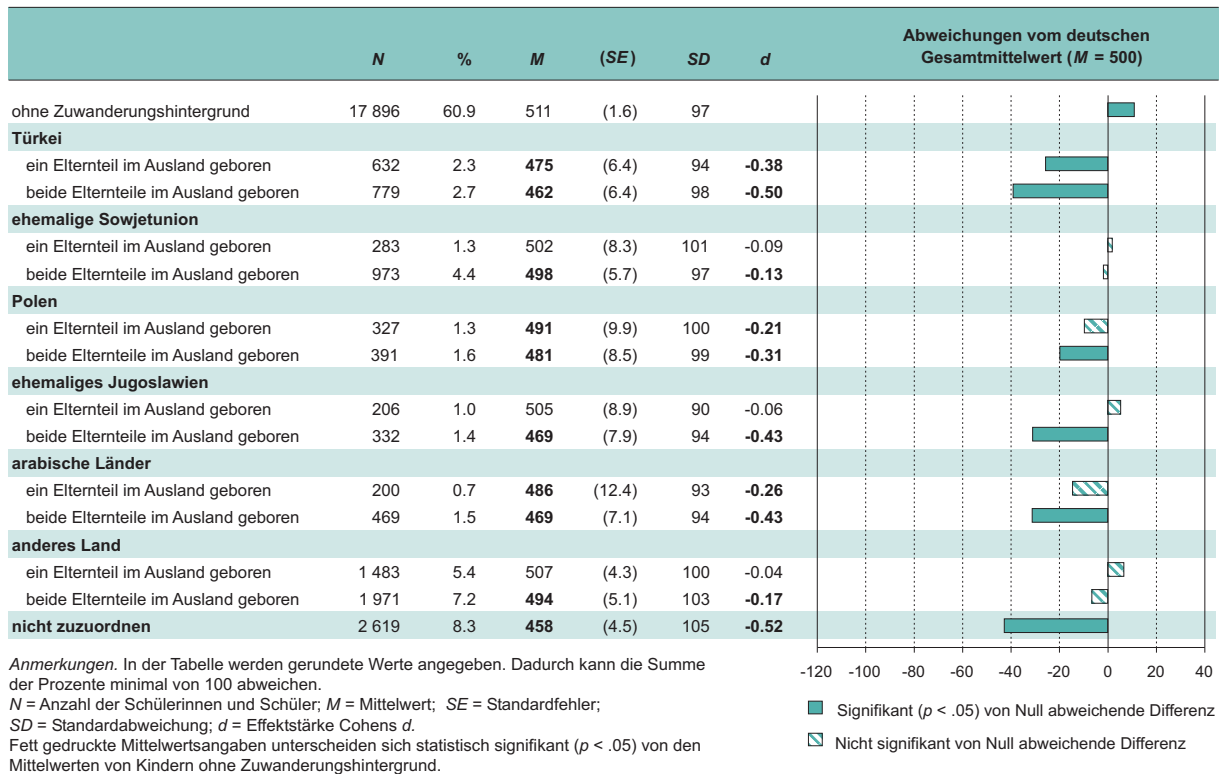
Im Gegensatz zum Bereich *Lesen* ist in der Mehrzahl der Gruppen im Bereich *Zuhören* das mittlere Kompetenzniveau im Jahr 2016 geringer als im Jahr 2011. Dieser Rückgang ist besonders ausgeprägt für Kinder der Herkunftsgruppe Türkei (-28 bzw. -27 Punkte) sowie für Kinder, deren Eltern beide aus Polen (-28 Punkte) oder aus dem ehemaligen Jugoslawien zugewandert sind (-28 Punkte) (vgl. auch Abb. 9.13web).

Auch im Kompetenzbereich *Zuhören* ist die Höhe der zugewanderungsbezogenen Disparitäten weitgehend stabil geblieben. Nur für die Gruppe von Kindern mit zwei in der Türkei geborenen Elternteilen haben sich die Disparitäten gegenüber dem Jahr 2011 vergrößert.

Erreichte Kompetenzen im Bereich *Orthografie* nach Herkunftsgruppen

Auch für den Kompetenzbereich *Orthografie* liegen im Jahr 2016 für die Mehrzahl der Gruppen zugewanderungsbezogene Disparitäten vor (siehe Abb. 9.9). Wie in Abschnitt 9.4.1 sind auch hier für den Bereich *Orthografie* die Ergebnisse separat für das Jahr 2016 und für den Vergleich der Jahre 2011 und 2016 abgebil-

Abbildung 9.9: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich *Orthografie* im Fach Deutsch nach Herkunftsgruppen sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016



det.¹⁵ Die dargestellten Veränderungen in Tabelle 9.3 zeigen zudem, dass sich bei Kindern mit zwei in Polen geborenen Elternteilen (–50 Punkte), bei Kindern der Herkunftsgruppe anderes Land (–27 bzw. –24 Punkte) sowie bei Kindern ohne Zuwanderungshintergrund (–24 Punkte) die erreichten Kompetenzen zwischen den Jahren 2011 und 2016 signifikant reduziert haben. In allen anderen Gruppen sind keine signifikanten Trends zu verzeichnen (siehe Tab. 9.3; vgl. auch Abb. 9.14web).

Wie im *Lesen* und *Zuhören* ist das Ausmaß der zuwanderungsbezogenen Disparitäten auch in der *Orthografie* weitgehend konstant geblieben. Nur für die Gruppe von Kindern mit einem im ehemaligen Jugoslawien geborenen Elternteil ist eine signifikante Verringerung der Disparitäten im Vergleich zu den Disparitäten im Jahr 2011 festzustellen.

15 Da im IQB-Ländervergleich 2011 die Orthografieaufgaben nicht an Förderschulen eingesetzt wurden, beziehen sich Trendanalysen nur auf Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf. Analysen erreichter Kompetenzen im Bereich *Orthografie* im Jahr 2016 beziehen sich hingegen auf alle Schülerinnen und Schüler.

Tabelle 9.3: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Kompetenzbereich *Orthografie* im Fach Deutsch nach Herkunftsgruppen in den Jahren 2011 und 2016 im Vergleich

	gültige %	+ / -	2011				2016				Differenz		
			N	M	(SE)	SD	N	M	(SE)	SD	ΔM	(SE)	<i>d</i>
ohne Zuwanderungshintergrund	61.4	-7.9	3 291	543	(2.8)	85	16 426	519	(1.8)	91	-24	(4.0)	-0.27
Türkei													
ein Elternteil im Ausland geboren	2.3	0.1	104	492 ^a	(8.7)	74	551	481 ^a	(6.9)	90	-11	(11.3)	-0.13
beide Elternteile im Ausland geboren	2.7	-0.9	168	479 ^a	(7.1)	80	648	470 ^a	(6.7)	93	-9	(10.0)	-0.11
ehemalige Sowjetunion													
ein Elternteil im Ausland geboren	1.3	0.6	32	503 ^a	(17.8)	99	266	506	(9.2)	98	3	(20.2)	0.03
beide Elternteile im Ausland geboren	4.6	0.6	188	512 ^a	(8.2)	88	915	502 ^a	(5.9)	94	-10	(10.4)	-0.11
Polen													
ein Elternteil im Ausland geboren	1.3	0.3	44	529	(15.2)	84	287	499 ^a	(9.7)	96	-30	(18.2)	-0.34
beide Elternteile im Ausland geboren	1.6	0.6	48	537	(14.0)	85	355	487 ^a	(9.0)	95	-50	(16.8)	-0.56
ehemaliges Jugoslawien													
ein Elternteil im Ausland geboren	1.0	0.1	42	492 ^a	(17.0)	85	181	513	(9.6)	88	21 ^a	(19.7)	0.25
beide Elternteile im Ausland geboren	1.4	0.6	39	509	(17.2)	91	290	478 ^a	(8.4)	87	-32	(19.2)	-0.35
anderes Land													
ein Elternteil im Ausland geboren	6.1	1.1	237	540	(6.1)	87	1 521	513	(4.3)	93	-27	(7.8)	-0.30
beide Elternteile im Ausland geboren	8.6	4.5	196	522 ^a	(7.6)	91	2 137	498 ^a	(4.5)	97	-24	(9.1)	-0.25
nicht zuzuordnen	7.8	0.1	362	493 ^a	(6.9)	89	2 076	472 ^a	(4.6)	97	-21	(8.6)	-0.23

Anmerkungen. Diese Berechnungen wurden ohne Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf durchgeführt. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen und die Summe der Prozente kann minimal von 100 abweichen.

+/- = Veränderung gegenüber IQB-Ländervergleich 2011; N = Anzahl der Schülerinnen und Schüler; M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; *d* = Effektstärke Cohens *d*.

^a signifikante Differenz ($p < .05$) zu Kindern ohne Zuwanderungshintergrund.

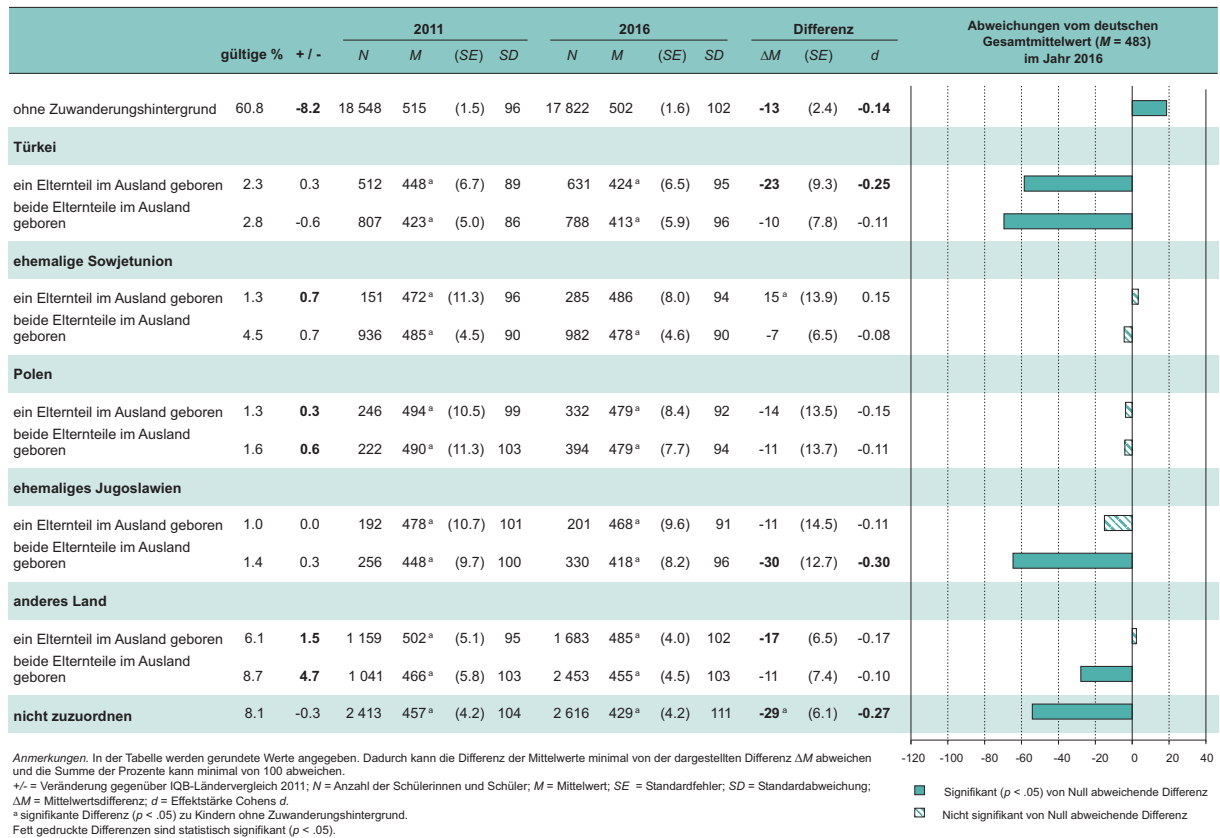
Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

Erreichte Kompetenzen im Fach Mathematik nach Herkunftsgruppen

Abbildung 9.10 schließlich stellt die Disparitäten im Fach Mathematik getrennt nach Herkunftsgruppen für die Jahre 2011 und 2016 dar. Fast alle Herkunftsgruppen erreichen im Jahr 2016 im Durchschnitt signifikant geringere mathematische Kompetenzen als Kinder ohne Zuwanderungshintergrund (vgl. auch Abb. 9.15web). Eine Ausnahme bilden Kinder mit einem in der ehemaligen Sowjetunion geborenen Elternteil, für die im Jahr 2016 keine signifikanten zuwanderungsbezogenen Disparitäten bestehen.

Die Trendanalysen zeigen eine signifikante Abnahme der durchschnittlichen mathematischen Kompetenzen von Kindern mit einem in der Türkei geborenen Elternteil (-23 Punkte), von Kindern mit zwei im ehemaligen Jugoslawien geborenen Elternteilen (-30 Punkte), von Kindern mit einem in einem anderen Land geborenen Elternteil (-17 Punkte) sowie von Kindern ohne

Abbildung 9.10: Mittelwerte und Streuungen der erreichten Kompetenzen sowie Gruppenunterschiede im Fach Mathematik (*Globalskala*) nach Herkunftsgruppen in den Jahren 2011 und 2016 im Vergleich sowie Abweichungen vom deutschen Gesamtmittelwert im Jahr 2016



Zuwanderungshintergrund (-13 Punkte) im Vergleich zu den im Jahr 2011 erreichten Kompetenzen (vgl. auch Abb. 9.16web). Die zuwanderungsbezogenen Disparitäten sind dabei weitgehend konstant geblieben. Nur für die Gruppe der Kinder mit einem in der ehemaligen Sowjetunion geborenen Elternteil haben sich die Disparitäten signifikant verringert.

9.4.3 Merkmale des familiären Hintergrunds von Viertklässlerinnen und Viertklässlern

Bei der Interpretation zuwanderungsbezogener Disparitäten ist zu berücksichtigen, dass sich Kinder mit und ohne Zuwanderungshintergrund oft auch in weiteren familiären Merkmalen unterscheiden, die zur Entstehung der beobachteten Kompetenzunterschiede zwischen Kindern mit und ohne Zuwanderungshintergrund beitragen können. Entsprechende Analysen weisen darauf hin, dass die soziale Lage zugewanderter Familien, das Bildungsniveau der Eltern sowie die in der Familie gesprochene Sprache hierbei zentral sind (vgl. Haag et al., 2012; Haag et al., 2016; Pöhlmann et al., 2013; Siegert & Olszenka, 2016; Stanat et al., 2010). Je günstiger die soziale Lage von Eltern ist und je mehr bildungsrelevante Ressourcen ihnen zur Verfügung stehen, umso eher ist es ihnen möglich, ihren Kindern Wissen und Fertigkeiten zu vermitteln und sie in ihrer Schullaufbahn zu unterstützen. Die in der Familie gesprochene Sprache gilt zudem als Indikator für außerschulische Lerngelegenheiten zum Erwerb der Instruktionssprache.

In weiterführenden Analysen anhand der Daten des IQB-Bildungstrends 2016 (vgl. Tab. 9.2web) werden diese Hintergrundmerkmale für Kinder ohne Zuwanderungshintergrund und für Kinder mit einem beziehungsweise mit zwei im Ausland geborenen Elternteilen aufgeführt. Es zeigen sich in nahezu allen Ländern bedeutsame zuwanderungsbezogene Unterschiede im sozioökonomischen Status, im Bildungsniveau der Eltern sowie im Sprachgebrauch in der Familie. Sowohl Familien, in denen beide Elternteile im Ausland geboren sind, als auch Familien, in denen ein Elternteil im Ausland geboren ist, weisen im Durchschnitt einen niedrigeren sozioökonomischen Status und ein geringeres Bildungsniveau (Anzahl der Bildungsjahre) auf als Familien ohne Zuwanderungshintergrund. Erwartungsgemäß geben Familien mit Zuwanderungshintergrund zudem häufiger als Familien ohne Zuwanderungshintergrund an, nur manchmal oder nie zu Hause Deutsch zu sprechen.

Zusätzlich durchgeführte Regressionsanalysen ermöglichen es, zu untersuchen, inwieweit die zuwanderungsbezogenen Disparitäten im Kompetenzerwerb mit Unterschieden in diesen Hintergrundmerkmalen in Zusammenhang stehen. Die Tabellen 9.3web bis 9.6web geben die Ergebnisse für jeden der vier Kompetenzbereiche im Jahr 2016 wieder. Sie zeigen, dass der sozioökonomische Status und das Bildungsniveau der Eltern erwartungsgemäß positiv mit den erreichten Kompetenzen in den vier Bereichen zusammenhängen (vgl. auch Kapitel 8). Außerdem erzielen Kinder, die nur manchmal oder nie zu Hause Deutsch sprechen, in allen Bereichen ein signifikant geringeres Kompetenzniveau als Kinder, die zu Hause immer Deutsch sprechen. Unter Berücksichtigung der familiären Hintergrundmerkmale reduzieren sich die zuwanderungsbezogenen Disparitäten. Für mehrere Zuwanderergruppen, insbesondere für Kinder, deren Eltern aus der Türkei oder aus arabischen Ländern zugewandert sind, bleiben jedoch in allen Kompetenzbereichen signifikante Disparitäten bestehen. Eine Ausnahme bildet lediglich der Kompetenzbereich *Orthografie*. Unter Berücksichtigung des Sprachgebrauchs zeigen sich für keine der Herkunftsgruppen zuwanderungsbezogene Kompetenznachteile; für einzelne Teilgruppen von Schülerinnen und Schülern mit Zuwanderungshintergrund ist nach Kontrolle des Sprachgebrauchs in *Orthografie* sogar ein signifikanter Kompetenzvorteil gegenüber Kindern ohne Zuwanderungshintergrund zu verzeichnen.

Inwieweit sich die Hintergrundmerkmale der Viertklässlerinnen und Viertklässler zwischen den Jahren 2011 und 2016 verändert haben, ist in den Tabellen 9.7web bis 9.9web in Abhängigkeit vom Zuwanderungsstatus dargestellt. Insgesamt wird eine weitgehende Stabilität der durchschnittlichen Merkmalsausprägungen in den Jahren 2011 und 2016 deutlich, auch wenn sich in einzelnen Ländern für manche Gruppen der sozioökonomische Status, das Bildungsniveau der Eltern sowie die Häufigkeit, mit der zu Hause Deutsch gesprochen wird, leicht verändert haben. Der Vergleich der Regressionsanalysen zwischen den Jahren 2011 und 2016 (siehe Tab. 9.10web bis 9.13web) zeigt, dass die Veränderungen erreichter Kompetenzen in den einzelnen Herkunftsgruppen weitgehend unabhängig von Veränderungen in familiären Hintergrundmerkmalen sind.

9.5 Merkmale der sozialen Integration und schulischen Motivation von Viertklässlerinnen und Viertklässlern nach Zuwanderungsstatus im Jahr 2016

Schulische Integration bedeutet, dass sich Schülerinnen und Schüler in der Schule angenommen und wohl fühlen und sie Freude am Lernen haben. Um festzustellen, inwieweit dies gelingt, werden im IQB-Bildungstrend 2016 die wahrgenommene soziale Eingebundenheit in der Klasse, die Schulzufriedenheit sowie die Lernfreude in den Fächern Deutsch und Mathematik untersucht. Zur Erfassung dieser Merkmale wurden Skalen mit jeweils drei bis vier Aussagen (Items) im Fragebogen für Schülerinnen und Schüler eingesetzt, die sich an Skalen aus anderen *Large-Scale-Assessments* orientieren (Bos et al., 2010; Furthmüller, 2014; Pekrun, Goetz & Frenzel, 2005). Ein Beispiel für ein Item zur Erfassung der *sozialen Eingebundenheit* ist „Ich habe wirklich gute Freunde in der Klasse.“ Zur Erfassung der *Schulzufriedenheit* wurde beispielsweise die Aussage „Ich fühle mich wohl in meiner Schule.“ eingesetzt. Aussagen zur Ermittlung der *Lernfreude* pro Fach sind beispielweise „Ich freue mich auf die Deutschstunde/Mathematikstunde.“ Die Schülerinnen und Schüler gaben auf einer vierstufigen Antwortskala an, wie sehr sie den Items jeweils zustimmen (1 = „stimmt gar nicht“ bis 4 = „stimmt genau“). Die Werte wurden anschließend pro Skala gemittelt, sodass die resultierenden Skalenwerte zwischen 1 und 4 variieren können und der theoretische Mittelwert je Skala bei $M = 2.5$ Punkten liegt. Fehlende Werte auf diesen Variablen wurden multipel imputiert (Graham, 2009; Schafer & Graham, 2002).

Die Ergebnisse der Analysen sind den Tabellen 9.4 und 9.5 zu entnehmen. Darin sind für jedes Land und für Deutschland insgesamt die Mittelwerte (M), die Standardfehler der Mittelwerte (SE) und die Standardabweichungen (SD) angegeben. Signifikante Unterschiede zwischen Kindern ohne Zuwanderungshintergrund und den beiden Gruppen von Schülerinnen und Schülern mit Zuwanderungshintergrund (Kinder mit einem im Ausland geborenen Elternteil und Kinder mit zwei im Ausland geborenen Elternteilen) sind durch Fettdruck der Mittelwerte und der standardisierten Mittelwertsdifferenzen (d) gekennzeichnet (vgl. Kapitel 3.2).

Soziale Eingebundenheit nach Zuwanderungsstatus im Jahr 2016

Der durchschnittliche Wert für die wahrgenommene soziale Eingebundenheit liegt in allen Ländern und in allen Gruppen zwischen etwa 2.7 und 2.9 Punkten und damit oberhalb des theoretischen Mittelwerts der Skala von 2.5 Punkten (siehe Tab. 9.4, vgl. auch die grafische Veranschaulichung in Abb. 9.17web). Zudem zeigen sich weder bundesweit noch innerhalb der Länder signifikante Unterschiede zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund. Unabhängig vom Zuwanderungshintergrund fühlen sich Viertklässlerinnen und Viertklässler somit insgesamt – sowohl bundesweit als auch innerhalb der einzelnen Länder – gut in den sozialen Kontext der Schule eingebunden.

Tabelle 9.4: Mittelwerte und Streuungen der sozialen Eingebundenheit in der Schulklasse und der Schulzufriedenheit sowie Gruppenunterschiede nach Zuwanderungsstatus und Land im Jahr 2016

Land	soziale Eingebundenheit				Schulzufriedenheit			
	M	(SE)	SD	d	M	(SE)	SD	d
Baden-Württemberg	2.91	(0.0)	0.61		2.95	(0.0)	0.75	
	2.89	(0.0)	0.59	-0.05	2.93	(0.1)	0.79	-0.02
	2.88	(0.0)	0.61	-0.06	3.06	(0.1)	0.75	0.14
Bayern	2.88	(0.0)	0.59		2.94	(0.0)	0.76	
	2.88	(0.0)	0.59	0.00	2.94	(0.1)	0.76	0.00
	2.90	(0.0)	0.58	0.03	3.03	(0.1)	0.73	0.12
Berlin	2.80	(0.0)	0.61		2.94	(0.0)	0.73	
	2.80	(0.0)	0.61	-0.01	3.00	(0.1)	0.76	0.08
	2.80	(0.0)	0.63	0.00	2.97	(0.0)	0.78	0.04
Brandenburg	2.74	(0.0)	0.63		2.89	(0.0)	0.75	
	2.74	(0.1)	0.67	0.01	2.93	(0.1)	0.78	0.05
	2.66	(0.1)	0.61	-0.13	2.72	(0.1)	0.81	-0.22
Bremen	2.85	(0.0)	0.62		2.99	(0.0)	0.76	
	2.91	(0.0)	0.62	0.11	3.11	(0.1)	0.79	0.15
	2.89	(0.0)	0.63	0.07	3.16	(0.0)	0.76	0.23
Hessen	2.87	(0.0)	0.57		2.92	(0.0)	0.72	
	2.86	(0.0)	0.57	-0.02	2.92	(0.1)	0.72	0.00
	2.80	(0.0)	0.57	-0.12	3.04	(0.0)	0.74	0.17
Mecklenburg-Vorpommern	2.75	(0.0)	0.58		2.87	(0.0)	0.77	
	2.69	(0.1)	0.62	-0.11	2.78	(0.1)	0.86	-0.11
	2.66	(0.1)	0.63	-0.15	2.98	(0.1)	0.78	0.14
Niedersachsen	2.79	(0.0)	0.61		2.92	(0.0)	0.77	
	2.75	(0.0)	0.67	-0.06	2.93	(0.1)	0.75	0.02
	2.80	(0.0)	0.64	0.02	3.02	(0.1)	0.80	0.14
Nordrhein-Westfalen	2.83	(0.0)	0.61		2.99	(0.0)	0.76	
	2.83	(0.1)	0.62	0.01	2.99	(0.1)	0.79	0.00
	2.85	(0.0)	0.58	0.05	3.03	(0.1)	0.80	0.05
Rheinland-Pfalz	2.86	(0.0)	0.61		3.02	(0.0)	0.75	
	2.82	(0.1)	0.60	-0.05	3.05	(0.1)	0.73	0.04
	2.87	(0.0)	0.60	0.02	3.13	(0.1)	0.70	0.16
Sachsen	2.79	(0.0)	0.57		2.96	(0.0)	0.71	
	2.73	(0.1)	0.62	-0.09	2.92	(0.1)	0.68	-0.06
	2.75	(0.1)	0.54	-0.07	2.94	(0.1)	0.74	-0.03
Sachsen-Anhalt	2.77	(0.0)	0.61		2.99	(0.1)	0.76	
	2.80	(0.1)	0.61	0.05	3.08	(0.1)	0.75	0.12
	2.84	(0.1)	0.55	0.11	3.25	(0.1)	0.69	0.35
Schleswig-Holstein	2.85	(0.0)	0.61		2.88	(0.0)	0.78	
	2.80	(0.0)	0.59	-0.09	2.96	(0.1)	0.79	0.10
	2.79	(0.1)	0.61	-0.10	3.07	(0.1)	0.67	0.27
Thüringen	2.79	(0.0)	0.61		3.02	(0.0)	0.72	
	2.73	(0.1)	0.64	-0.10	3.13	(0.2)	0.69	0.15
	2.82	(0.1)	0.56	0.05	3.07	(0.1)	0.73	0.07
Hamburg ¹	2.92	(0.0)	0.56		3.01	(0.0)	0.70	
	2.87	(0.0)	0.57	-0.10	3.03	(0.1)	0.72	0.03
	2.91	(0.0)	0.60	-0.02	3.12	(0.0)	0.72	0.16
Saarland ¹	2.90	(0.0)	0.62		2.97	(0.0)	0.78	
	2.92	(0.0)	0.57	0.04	2.96	(0.1)	0.83	-0.02
	2.88	(0.0)	0.59	-0.03	3.06	(0.1)	0.74	0.11
Deutschland	2.84	(0.0)	0.60		2.95	(0.0)	0.75	
	2.84	(0.0)	0.61	0.00	2.96	(0.0)	0.76	0.01
	2.85	(0.0)	0.60	0.02	3.04	(0.0)	0.76	0.12

Anmerkungen. 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren)

2. Zeile: Kinder mit einem im Ausland geborenen Elternteil

3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.

M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; d = Effektstärke Cohens d.

¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

Fett gedruckte Mittelwertangaben unterscheiden sich statistisch signifikant ($p < .05$) von den Mittelwerten von Kindern ohne Zuwanderungshintergrund.

Schulzufriedenheit nach Zuwanderungsstatus im Jahr 2016

Für die Schulzufriedenheit zeigt Tabelle 9.4, dass die Mittelwerte in allen Ländern und in allen Gruppen ebenfalls über dem theoretischen Mittelwert liegen und zwischen 2.7 und 3.3 Punkten variieren. Bundesweit ist die Schulzufriedenheit bei Schülerinnen und Schülern mit zwei im Ausland geborenen Elternteilen statistisch signifikant höher als bei Schülerinnen und Schülern ohne Zuwanderungshintergrund, der Unterschied ist jedoch sehr klein ($d = 0.12$). Zwischen Kindern mit einem im Ausland geborenen Elternteil und Kindern ohne Zuwanderungshintergrund bestehen in Deutschland insgesamt keine signifikanten Unterschiede.

Auch in nahezu allen Ländern sind die Mittelwerte der Schulzufriedenheit für Kinder mit Zuwanderungshintergrund ähnlich hoch wie für Kinder ohne Zuwanderungshintergrund. In Bremen, Hessen, Sachsen-Anhalt, Schleswig-Holstein und Hamburg sind kleine bis mittelgroße Unterschiede in der durchschnittlichen Schulzufriedenheit zu verzeichnen, die darauf hinweisen, dass Kinder mit zwei im Ausland geborenen Elternteilen signifikant zufriedener sind als Kinder ohne Zuwanderungshintergrund (vgl. auch Abb. 9.18web).

Die Befunde zur Schulzufriedenheit deuten also – ebenso wie die Ergebnisse zur sozialen Eingebundenheit – darauf hin, dass sich die Kinder in Deutschland unabhängig von ihrem Zuwanderungsstatus in ihren Schulen gut integriert fühlen und dass Kinder mit Zuwanderungshintergrund zum Teil sogar etwas zufriedener mit ihrer Schule sind als Kinder ohne Zuwanderungshintergrund.

Lernfreude in den Fächern Deutsch und Mathematik nach Zuwanderungsstatus im Jahr 2016

Für die Lernfreude im Fach Deutsch zeigen sich insgesamt etwas niedrigere Mittelwerte als für die Merkmale der sozialen Integration (siehe Tab. 9.5; vgl. auch Abb. 9.19web). Sie liegen zumeist geringfügig unter dem theoretischen Mittelwert von 2.5 Punkten und variieren zwischen 2.2 und 2.8 Punkten. In Deutschland insgesamt ist die Lernfreude im Fach Deutsch bei Schülerinnen und Schülern mit zwei im Ausland geborenen Elternteilen ($M = 2.5$) statistisch signifikant höher als bei Schülerinnen und Schülern ohne Zuwanderungshintergrund ($M = 2.3$), der Unterschied ist jedoch klein ($d = 0.18$). Die Lernfreude von Kindern mit einem im Ausland geborenen Elternteil ($M = 2.4$) im Fach Deutsch weicht hingegen nicht signifikant von der Lernfreude von Kindern ohne Zuwanderungshintergrund ab.

In der Hälfte der Länder ist die Lernfreude im Fach Deutsch bei Kindern mit zwei im Ausland geborenen Eltern signifikant höher als bei Kindern ohne Zuwanderungshintergrund, bei Differenzen im kleinen bis mittleren Bereich. Dies gilt für Baden-Württemberg, Bayern, Berlin, Bremen, Hessen, Niedersachsen, Schleswig-Holstein und Hamburg. In fast keinem Land bestehen statistisch signifikante Unterschiede in der durchschnittlichen Lernfreude im Fach Deutsch zwischen Kindern mit einem im Ausland geborenen Elternteil und Kindern ohne Zuwanderungshintergrund. Lediglich in Bremen ist der Mittelwert der Kinder mit einem im Ausland geborenen Elternteil ($M = 2.6$) signifikant höher als der Mittelwert der Kinder ohne Zuwanderungshintergrund ($M = 2.5$).

Die auf das Fach Mathematik bezogene Lernfreude (siehe Tab. 9.5; vgl. auch Abb. 9.20web) liegt wiederum in fast allen Ländern und in allen Gruppen

Tabelle 9.5: Mittelwerte und Streuungen der Lernfreude im Fach Deutsch und im Fach Mathematik sowie Gruppenunterschiede nach Zuwanderungsstatus und Land im Jahr 2016

Land	Lernfreude im Fach Deutsch				Lernfreude im Fach Mathematik			
	<i>M</i>	(<i>SE</i>)	<i>SD</i>	<i>d</i>	<i>M</i>	(<i>SE</i>)	<i>SD</i>	<i>d</i>
Baden-Württemberg	2.27	(0.1)	0.84		2.50	(0.1)	0.97	
	2.26	(0.1)	0.89	-0.01	2.54	(0.1)	0.99	0.03
	2.48	(0.1)	0.90	0.25	2.78	(0.1)	1.01	0.28
Bayern	2.32	(0.0)	0.83		2.60	(0.0)	0.97	
	2.41	(0.1)	0.89	0.10	2.62	(0.1)	0.95	0.02
	2.46	(0.1)	0.82	0.18	2.86	(0.1)	0.93	0.27
Berlin	2.33	(0.0)	0.81		2.50	(0.0)	0.95	
	2.39	(0.1)	0.79	0.08	2.52	(0.1)	0.97	0.02
	2.50	(0.0)	0.84	0.21	2.78	(0.1)	0.98	0.29
Brandenburg	2.41	(0.0)	0.81		2.63	(0.0)	0.96	
	2.43	(0.1)	0.80	0.03	2.62	(0.1)	0.93	-0.02
	2.37	(0.1)	0.75	-0.05	2.73	(0.1)	0.93	0.11
Bremen	2.46	(0.0)	0.87		2.55	(0.1)	0.97	
	2.62	(0.1)	0.84	0.19	2.75	(0.1)	1.00	0.21
	2.76	(0.1)	0.89	0.34	2.88	(0.1)	0.97	0.34
Hessen	2.26	(0.0)	0.81		2.65	(0.1)	0.96	
	2.26	(0.1)	0.79	0.01	2.64	(0.1)	0.92	-0.01
	2.42	(0.0)	0.81	0.20	2.85	(0.1)	0.96	0.21
Mecklenburg-Vorpommern	2.30	(0.0)	0.80		2.65	(0.0)	0.95	
	2.24	(0.1)	0.81	-0.07	2.55	(0.2)	1.06	-0.11
	2.42	(0.1)	0.78	0.16	2.79	(0.1)	0.93	0.14
Niedersachsen	2.34	(0.1)	0.83		2.55	(0.1)	0.98	
	2.34	(0.1)	0.84	0.01	2.53	(0.1)	1.00	-0.01
	2.48	(0.1)	0.89	0.17	2.69	(0.1)	1.02	0.14
Nordrhein-Westfalen	2.40	(0.0)	0.82		2.68	(0.0)	0.98	
	2.47	(0.1)	0.85	0.08	2.68	(0.1)	1.00	0.00
	2.53	(0.1)	0.86	0.15	2.74	(0.1)	1.02	0.06
Rheinland-Pfalz	2.31	(0.1)	0.86		2.61	(0.0)	0.97	
	2.39	(0.1)	0.83	0.10	2.68	(0.1)	0.94	0.07
	2.39	(0.1)	0.87	0.09	2.85	(0.1)	0.96	0.25
Sachsen	2.32	(0.0)	0.81		2.59	(0.0)	0.90	
	2.40	(0.1)	0.83	0.10	2.38	(0.1)	1.02	-0.23
	2.41	(0.1)	0.76	0.11	2.63	(0.1)	0.96	0.04
Sachsen-Anhalt	2.46	(0.1)	0.82		2.58	(0.1)	0.95	
	2.32	(0.1)	0.82	-0.17	2.76	(0.1)	1.00	0.19
	2.70	(0.1)	0.86	0.29	2.93	(0.1)	0.87	0.38
Schleswig-Holstein	2.28	(0.0)	0.80		2.51	(0.0)	0.96	
	2.44	(0.1)	0.85	0.19	2.66	(0.1)	0.98	0.16
	2.58	(0.1)	0.79	0.38	2.75	(0.1)	0.87	0.26
Thüringen	2.47	(0.0)	0.81		2.67	(0.0)	0.94	
	2.57	(0.1)	0.80	0.13	2.80	(0.2)	1.06	0.12
	2.47	(0.1)	0.83	0.00	3.03	(0.1)	0.91	0.38
Hamburg ¹	2.38	(0.0)	0.76		2.51	(0.1)	0.89	
	2.46	(0.1)	0.80	0.10	2.54	(0.1)	0.94	0.03
	2.70	(0.1)	0.80	0.42	2.78	(0.1)	0.92	0.30
Saarland ¹	2.32	(0.0)	0.85		2.61	(0.1)	0.99	
	2.34	(0.1)	0.86	0.02	2.73	(0.1)	1.02	0.12
	2.37	(0.1)	0.80	0.06	2.76	(0.1)	0.97	0.15
Deutschland	2.34	(0.0)	0.82		2.60	(0.0)	0.97	
	2.38	(0.0)	0.85	0.04	2.61	(0.0)	0.98	0.01
	2.49	(0.0)	0.86	0.18	2.78	(0.0)	0.99	0.19

Anmerkungen. 1. Zeile: Kinder ohne Zuwanderungshintergrund (beide Elternteile in Deutschland geboren)

2. Zeile: Kinder mit einem im Ausland geborenen Elternteil

3. Zeile: Kinder mit zwei im Ausland geborenen Elternteilen.

M = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung; *d* = Effektstärke Cohens *d*.

¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

Fett gedruckte Mittelwertsangaben unterscheiden sich statistisch signifikant ($p < .05$) von den Mittelwerten von Kindern ohne Zuwanderungshintergrund.

oberhalb des theoretischen Mittelwerts. Die Mittelwerte der Lernfreude variieren insgesamt zwischen 2.4 und 3.0 Punkten. In Deutschland insgesamt ist auch im Fach Mathematik die Lernfreude bei Schülerinnen und Schülern mit zwei im Ausland geborenen Elternteilen ($M = 2.8$) statistisch signifikant höher als bei Schülerinnen und Schülern ohne Zuwanderungshintergrund ($M = 2.6$), der Unterschied ist jedoch wiederum klein ($d = 0.19$). Für die Lernfreude von Kindern mit einem im Ausland geborenen Elternteil ($M = 2.6$) bestehen deutschlandweit im Fach Mathematik keine signifikanten Unterschiede zu Kindern ohne Zuwanderungshintergrund.

Auch in mehr als der Hälfte der Länder gaben Kinder mit zwei im Ausland geborenen Elternteilen durchschnittlich eine signifikant höhere Lernfreude an als Kinder ohne Zuwanderungshintergrund. Besonders ausgeprägt sind diese Unterschiede in Bremen, Thüringen, Sachsen-Anhalt und Hamburg. Erneut findet sich nur in Bremen ein signifikanter Unterschied für Kinder mit einem im Ausland geborenen Elternteil, die eine etwas höhere mathematische Lernfreude angeben ($M = 2.8$) als Kinder ohne Zuwanderungshintergrund ($M = 2.6$). Die Ergebnisse weisen also insgesamt darauf hin, dass Kinder mit Zuwanderungshintergrund über eine mindestens genauso hohe Lernfreude verfügen wie Kinder ohne Zuwanderungshintergrund; teilweise ist ihre Lernfreude sogar signifikant höher ausgeprägt. Die in den vorherigen Abschnitten dargestellten Disparitäten im Kompetenzerwerb spiegeln sich demnach nicht in der Lernfreude von Kindern mit Zuwanderungshintergrund am Ende der 4. Jahrgangsstufe wider.

9.6 Zusammenfassung und Diskussion

In diesem Kapitel wurde zunächst untersucht, inwieweit sich bei den von Viertklässlerinnen und Viertklässlern in den Fächern Deutsch und Mathematik erreichten Kompetenzen zuwanderungsbezogene Disparitäten zeigen. Diese Analysen geben Anhaltspunkte dafür, wie gut die auf den Kompetenzerwerb bezogene schulische Integration von Kindern mit Zuwanderungshintergrund in den Ländern gelingt. Zusätzlich zu den schulischen Kompetenzen wurden auch Merkmale der sozialen Integration und der schulischen Motivation in den Blick genommen. Mit Trendanalysen wurde außerdem untersucht, wie sich die Verteilung von Kindern mit Zuwanderungshintergrund in den Ländern, die zuwanderungsbezogene Zusammensetzung der Schülerschaft in den Schulen und die Disparitäten im Kompetenzerwerb zwischen den Jahren 2011 und 2016 verändert haben.

Im betrachteten Zeitraum von fünf Jahren hat der Anteil von Kindern aus zugewanderten Familien in Deutschland insgesamt um 9 Prozentpunkte zugenommen (vgl. Autorengruppe Bildungsberichterstattung, 2016). Auch in nahezu allen Ländern ist der Anteil der Schülerinnen und Schüler mit Zuwanderungshintergrund statistisch signifikant gestiegen. Diese Zunahme ist vor allem auf höhere Zahlen von Kindern mit einem im Ausland geborenen Elternteil sowie von Kindern der zweiten Generation zurückzuführen. Der Anteil von Kindern der ersten Generation, die selbst im Ausland geboren und mit ihren Familien nach Deutschland zugewandert sind, hat sich dagegen zwischen den Jahren 2011 und 2016 deutlich weniger stark erhöht; bundesweit stieg er um nur rund 2 Prozentpunkte an.

Die Analysen zur Verteilung der Kinder mit Zuwanderungshintergrund auf die Schulen in Deutschland verdeutlichen, dass sich die Schulen in ihrer zuwanderungsbezogenen Zusammensetzung stark unterscheiden. Während bun-

desweit in etwa 13 Prozent der Schulen weniger als 5 Prozent der Kinder einen Zuwanderungshintergrund haben, liegt der Anteil in rund einem Viertel der Schulen bei mindestens 40 Prozent. In je einem Drittel der Schulen kommen zwischen 5 und 19 Prozent beziehungsweise zwischen 20 und 39 Prozent der Schülerinnen und Schüler aus zugewanderten Familien. Diese Verteilungen haben sich im Vergleich zur Verteilung im Jahr 2011 verändert: Der Anteil von Schulen mit besonders wenigen Kindern mit Zuwanderungshintergrund ist gesunken und der Anteil von Schulen mit einer höheren Anzahl von Kindern mit Zuwanderungshintergrund ist gestiegen. Im Jahr 2016 sind sich somit bei insgesamt gestiegenem Anteil von Kindern mit Zuwanderungshintergrund die einzelnen Schulen in ihrer zuwanderungsbezogenen Schulzusammensetzung etwas ähnlicher als im Jahr 2011.

In den Analysen zuwanderungsbezogener Disparitäten im Kompetenzerwerb wurde geprüft, inwieweit zwischen Kindern mit Zuwanderungshintergrund und Kindern ohne Zuwanderungshintergrund Unterschiede in den erreichten Kompetenzen bestehen und inwieweit sich diese Disparitäten über den Untersuchungszeitraum von fünf Jahren (2011–2016) verändert haben.

Im Fach Deutsch bestehen im Kompetenzbereich *Lesen* sowohl bundesweit als auch in nahezu allen Ländern signifikante Nachteile für Kinder aus zugewanderten Familien. Die Kompetenznachteile von Kindern mit zwei im Ausland geborenen Elternteilen sind in fast allen Ländern statistisch signifikant, wohingegen die Kompetenznachteile von Kindern mit einem im Ausland geborenen Elternteil deutlich geringer ausgeprägt sind und nicht in allen Ländern statistische Signifikanz erreichen. Im Trend sind die zuwanderungsbezogenen Disparitäten im Bereich *Lesen* weitgehend stabil geblieben. Lediglich für Kinder der ersten Generation haben sich die Disparitäten auf Bundesebene verstärkt.

Im Kompetenzbereich *Zuhören* sind die Unterschiede zwischen Kindern ohne Zuwanderungshintergrund und Kindern aus zugewanderten Familien insgesamt größer als im *Lesen*. Erneut sind die Disparitäten für Kinder mit zwei im Ausland geborenen Elternteilen in allen Ländern signifikant und stärker ausgeprägt als für Kinder mit einem im Ausland geborenen Elternteil, für die zudem die Unterschiede zu Kindern ohne Zuwanderungshintergrund nicht in allen Ländern statistisch signifikant sind. Seit dem Jahr 2011 haben sich die Kompetenzen im *Zuhören* sowohl von Kindern mit Zuwanderungshintergrund als auch von Kindern ohne Zuwanderungshintergrund signifikant verringert. Auch in diesem Bereich sind die zuwanderungsbezogenen Disparitäten in den Ländern weitgehend stabil geblieben. Bundesweit haben sich die Disparitäten für Kinder der ersten Generation sowie für Kinder, deren Eltern in der Türkei geboren sind, jedoch signifikant vergrößert.

Die im Bereich *Orthografie* identifizierten Kompetenzunterschiede zwischen Kindern mit und ohne Zuwanderungshintergrund fallen im Vergleich zu den anderen Kompetenzbereichen im Fach Deutsch deutlich kleiner und auch nicht in allen Ländern statistisch signifikant aus. In Deutschland insgesamt haben sich die Kompetenzen im Bereich *Orthografie* in allen Gruppen von Kindern aus zugewanderten Familien und bei Kindern ohne Zuwanderungshintergrund zwischen den Jahren 2011 und 2016 signifikant reduziert, bei weitgehend gleichbleibenden Kompetenzunterschieden zwischen Kindern mit und ohne Zuwanderungshintergrund. Für die Gruppe der Kinder der zweiten Generation sowie für Kinder mit einem im ehemaligen Jugoslawien geborenen Elternteil haben sich die Disparitäten in Deutschland insgesamt signifikant verringert.

Auch im Fach Mathematik bestehen im Jahr 2016 Kompetenzunterschiede zwischen Kindern mit und ohne Zuwanderungshintergrund, die in ihrer Größenordnung in etwa den Disparitäten im *Lesen* entsprechen. Die Kompetenznachteile von Kindern mit zwei im Ausland geborenen Elternteilen sind in fast allen Ländern statistisch signifikant, die Kompetenznachteile von Kindern mit einem im Ausland geborenen Elternteil fallen hingegen deutlich geringer aus und erreichen nicht in allen Ländern statistische Signifikanz. Im Fach Mathematik haben sich bundesweit die von Kindern der ersten Generation, von Kindern mit einem im Ausland geborenen Elternteil sowie von Kindern ohne Zuwanderungshintergrund erreichten Kompetenzen zwischen den Jahren 2011 und 2016 signifikant reduziert bei gleichzeitig weitgehender Stabilität der zuwanderungsbezogenen Disparitäten. Nur für Kinder der zweiten Generation und Kinder mit einem in der ehemaligen Sowjetunion geborenen Elternteil haben sich die Disparitäten signifikant verringert.

Insgesamt zeigen sich im Jahr 2016 also in allen betrachteten Kompetenzbereichen zuwanderungsbezogene Disparitäten, die am stärksten im *Zuhören* und am geringsten in der *Orthografie* ausgeprägt sind und die seit dem Jahr 2011 relativ stabil geblieben sind. Analysen auf Bundesebene weisen darauf hin, dass die größten Disparitäten für Kinder der ersten Generation sowie für Kinder, deren Eltern aus der Türkei, dem ehemaligen Jugoslawien oder einem arabischen Land zugewandert sind, bestehen. Diese Schülerinnen und Schüler werden also besondere Unterstützung benötigen, um in der Sekundarstufe I erfolgreich zu sein.

Dieses allgemeine Befundmuster stimmt mit bisherigen Studien zu zuwanderungsbezogenen Disparitäten im deutschen Bildungssystem überein (vgl. z. B. Haag et al., 2012; Schwippert et al., 2012; Wendt et al., 2016). So wurden auch im IQB-Ländervergleich 2011 in den Ländern in der Bundesrepublik Deutschland zum Teil große Unterschiede in den erreichten Kompetenzen im Fach Deutsch und Mathematik zwischen Viertklässlerinnen und Viertklässlern mit und ohne Zuwanderungshintergrund deutlich. Besonders ausgeprägt waren auch hier die Disparitäten für Kinder mit zwei im Ausland geborenen Elternteilen sowie für Kinder, deren Eltern aus der Türkei zugewandert sind.

Vergleicht man die zuwanderungsbezogenen Disparitäten zwischen verschiedenen Kompetenzbereichen, so waren die Disparitäten im Bereich *Zuhören* bereits im IQB-Ländervergleich 2011 am größten (Haag et al., 2012). Auch im IQB-Bildungstrend 2015, in dem die Kompetenzen von Neuntklässlerinnen und Neuntklässlern untersucht wurden, werden für den Bereich *Zuhören* besonders große zuwanderungsbezogene Disparitäten berichtet. Die vergleichsweise geringeren Disparitäten zwischen Viertklässlerinnen und Viertklässlern mit und ohne Zuwanderungshintergrund im Kompetenzbereich *Orthografie* spiegeln sich ebenfalls in den Ergebnissen des IQB-Bildungstrends 2015 für Neuntklässlerinnen und Neuntklässler wider. Für die Bereiche *Zuhören* und *Orthografie* scheint somit über verschiedene Bildungsetappen hinweg ein vergleichsweise stabiles Befundmuster in den zuwanderungsbezogenen Disparitäten zu bestehen.

Etwas weniger einheitlich sind die Befunde für den Kompetenzbereich Mathematik: Während sich die zuwanderungsbezogenen Disparitäten für 15-jährige Schülerinnen und Schüler in PISA 2012 seit dem Jahr 2003 verringert haben (Gebhardt, Rauch, Mang, Sälzer & Stanat, 2013), sind sie im Grundschulbereich weitgehend stabil geblieben: Vergleichbar mit den Befunden des IQB-Bildungstrends 2016 zeigten auch die TIMS-Studien, dass sich Disparitäten in mathematischen Kompetenzen bundesweit an deutschen Schulen in der Primarstufe in

den Jahren 2007, 2011 und 2015 nicht signifikant verändert haben (Wendt et al., 2016).

Die signifikante Abnahme der in den Fächern Deutsch und Mathematik erreichten Kompetenzen in fast allen Gruppen – sowohl mit als auch ohne Zuwanderungshintergrund – weisen darauf hin, dass diese Entwicklung nicht auf einzelne Gruppen mit besonders starken Verringerungen der Kompetenzen zurückzuführen ist. Die weitgehende Stabilität der Disparitäten lässt darauf schließen, dass es den Schulen der Primarstufe in den Ländern im Vergleich zum Jahr 2011 nicht besser gelingt, zugewanderungsbezogene Disparitäten zu reduzieren.

Die Ergebnisse der zusätzlichen Analysen unter Berücksichtigung familiärer Hintergrundmerkmale weisen darauf hin, dass die schulischen Kompetenzen nach wie vor stark vom sozialen Hintergrund der Familie abhängen (vgl. auch Kapitel 8). So stehen auch zugewanderungsbezogene Disparitäten teilweise mit dem sozioökonomischen Status der Familie, dem Bildungshintergrund der Eltern und der Häufigkeit, mit der in der Familie Deutsch gesprochen wird, in Zusammenhang. Nach statistischer Kontrolle dieser Merkmale reduzieren sich die Kompetenznachteile der Schülerinnen und Schüler mit Zuwanderungshintergrund zwar deutlich, für einige Teilgruppen sind aber weiterhin substantielle Disparitäten zu beobachten. Insbesondere für Kinder, deren Familien aus der Türkei, dem ehemaligen Jugoslawien oder einem arabischen Land zugewandert sind, bleiben Kompetenznachteile bestehen. Zudem bestätigt sich die Bedeutung der in der Familie gesprochenen Sprache: Bei vergleichbarem sozioökonomischen Status der Familie und Bildungshintergrund sowie Geburtsland der Eltern sind die in den Fächern Deutsch und Mathematik erzielten Kompetenzen bei Schülerinnen und Schülern, die in der Familie „manchmal Deutsch“ oder „nie Deutsch“ sprechen, geringer ausgeprägt als bei Schülerinnen und Schülern mit ausschließlich deutscher Familiensprache. Diese Befunde unterstreichen nochmals, wie wichtig es ist, Kindern nichtdeutscher Herkunftssprache geeignete Lernmöglichkeiten für den Erwerb und die kontinuierliche Weiterentwicklung bildungssprachlicher Kompetenzen in der Instruktionssprache Deutsch zur Verfügung zu stellen (Kempert et al., 2016; Schneider et al., 2012; Stanat & Felbrich, 2013).

Neben zugewanderungsbezogenen Disparitäten im Kompetenzerwerb wurden im IQB-Bildungstrend 2016 erstmals für die Primarstufe auch Analysen zu Indikatoren der sozialen Integration und der schulischen Motivation berichtet. Die Befunde zur sozialen Eingebundenheit und Schulzufriedenheit zeigen, dass sich die Kinder unabhängig von ihrem Zuwanderungsstatus insgesamt gut in ihren Schulen integriert fühlen und Kinder mit Zuwanderungshintergrund zum Teil sogar etwas zufriedener in der Schule sind als Kinder ohne Zuwanderungshintergrund. Dies entspricht den Befunden entsprechender Analysen für Neuntklässlerinnen und Neuntklässler (Haag et al., 2016), die ebenfalls unabhängig vom Zuwanderungshintergrund in fast allen Ländern ein hohes Maß an Zugehörigkeit zur Schule berichten. Im IQB-Bildungstrend 2016 wurde zudem die Lernfreude der Schülerinnen und Schüler in den Fächern Deutsch und Mathematik untersucht. Diese ist bei Kindern aus zugewanderten Familien ebenso hoch, teilweise sogar höher ausgeprägt als bei Kindern ohne Zuwanderungshintergrund. Die zugewanderungsbezogenen Nachteile im Kompetenzerwerb spiegeln sich also nicht in der Lernfreude von Kindern mit Zuwanderungshintergrund wider. Die Diskrepanz zwischen den erreichten Kompetenzen und den motivationalen Merkmalen könnte möglicherweise darauf hinweisen, dass die motivationalen Ressourcen von Schülerinnen und Schülern mit Zuwanderungshintergrund nicht vollständig ausgeschöpft werden. Das entspricht unter

anderem den Ergebnissen von Untersuchungen in der Sekundarstufe, die ebenfalls eine höhere Lernmotivation beziehungsweise Lernfreude bei Schülerinnen und Schülern mit Zuwanderungshintergrund feststellen (vgl. Stanat & Christensen, 2006; Stanat et al., 2010).

Die in diesem Kapitel berichteten Befunde zu den zuwanderungsbezogenen Disparitäten in den Fächern Deutsch und Mathematik geben Hinweise darauf, in welchen Bereichen weiterhin verstärkter Förderbedarf für Kinder aus zugewanderten Familien besteht. Dieser ist im Bereich *Zuhören*, aber auch in den Bereichen *Lesen* und Mathematik besonders groß, wohingegen die zuwanderungsbezogenen Disparitäten in der *Orthografie* geringer ausgeprägt sind. Anhand der differenzierten Analysen nach Herkunftsgruppen konnten zudem genauere Muster der Disparitäten im Kompetenzerwerb innerhalb der heterogenen Gruppe der Kinder mit Zuwanderungshintergrund in den Jahren 2011 und 2016 identifiziert werden. Dabei weisen die Ergebnisse dieses Kapitels erneut darauf hin, dass die zur Verfügung stehenden sprachlichen Lerngelegenheiten für den Kompetenzerwerb zentral sind. Zu den wichtigsten Aufgaben der Schulen der Primarstufe wie des gesamten Bildungssystems gehört es daher, alle Schülerinnen und Schüler sprachlich in die Lage zu versetzen, die verfügbaren Bildungsangebote zu nutzen. Insgesamt zeigen die Ergebnisse, dass in Deutschland noch weitere Anstrengungen unternommen werden müssen, um dieses Ziel zu erreichen.

Literatur

- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: W. Bertelsmann Verlag.
- 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* (S. 357–375). Weinheim: Beltz.
- Böhme, K. & Bremerich-Vos, A. (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.
- 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.
- Bos, W., Strieholt, R., Goy, M., Stubbe, T. C., Tarelli, I. & Hornberg, S. (2010). *IGLU 2006. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Bundesamt für Migration und Flüchtlinge. (2016). *Wanderungsmonitoring: Erwerbsmigration nach Deutschland. Jahresbericht 2015*. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Esser, H. (2006). *Sprache und Integration. Die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten*. Frankfurt am Main: Campus.
- Flores, A. (2013). Die arabische Welt – ökonomische und soziale Gegebenheiten. In G. T. Schneiders (Hrsg.), *Die Araber im 21. Jahrhundert: Politik, Gesellschaft, Kultur* (S. 17–34). Wiesbaden: Springer.
- Foroutan, N. (2010). Neue Deutsche, Postmigranten und Bindungs-Identitäten. Wer gehört zum neuen Deutschland? *Aus Politik und Zeitgeschichte*, 46–47, 9–15.
- Furthmüller, P. (2014). *Skalenverzeichnis. Skalen und Indizes der Scientific-Use-Files 2005 bis 2009. Studie zur Entwicklung von Ganztagschulen (StEG)*. München: Deutsches Jugendinstitut e.V.

- Gebhardt, M., Rauch, D., Mang, J., Sälzer, C. & Stanat, P. (2013). Mathematische Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen für Deutschland* (S. 275–308). Münster: Waxmann.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576.
- Haag, N., Böhme, K., Rjosk, C. & Stanat, P. (2016). Zuwanderungsbezogene Disparitäten. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 431–480). Münster: Waxmann.
- Haag, N., Böhme, K. & Stanat, P. (2012). Zuwanderungsbezogene Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 209–235). Münster: Waxmann.
- Karsten, S., Felix, C., Ledoux, G., Meijnen, W., Roeleveld, J. & van Schooten, E. (2006). Choosing segregation or integration? The extent and effects of ethnic segregation in Dutch cities. *Education and Urban Society*, 38, 228–247.
- Kempert, S., Edele, A., Rauch, D., Wolf, K. M., Paetsch, J., Darsow, A., Maluch, J. & Stanat, P. (2016). Die Rolle der Sprache für zuwanderungsbezogene Ungleichheiten im Bildungserfolg. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 157–241). Wiesbaden: Springer.
- Kristen, C. (2008). Primary school choice and ethnic school segregation in German elementary schools. *European Sociological Review*, 24, 495–510.
- Olczyk, M., Seuring, J., Will, G. & Zinn, S. (2016). Migranten und ihre Nachkommen im deutschen Bildungssystem: Ein aktueller Überblick. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 33–70). Wiesbaden: Springer.
- Pekrun, R., Goetz, T. & Frenzel, A. C. (2005). *Achievement Emotions Questionnaire – Mathematics (AEQ-M). User's manual*. Department of Psychology, University of Munich.
- Pöhlmann, C., Haag, N. & Stanat, P. (2013). Zuwanderungsbezogene Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 297–330). Münster: Waxmann.
- Rauch, D., Mang, J., Härtig, H. & Haag, N. (2016). Naturwissenschaftliche Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 317–348). 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.
- Roeser, R. W., Eccles, J. S. & Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *The Elementary School Journal*, 100(5), 443–471.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, T., Roßbach, H.G., Rothweiler, M. & Stanat, P. (2012). *Expertise „Bildung durch Sprache und Schrift (BISS)“ – Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung*. Berlin: Bundesministerium für Bildung und Forschung.
- Schwippert, K., Wendt, H. & Tarelli, I. (2012). Lesekompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 191–173). Münster: Waxmann.
- Siegert, M. & Olszenka, N. (2016). Ethnische Ungleichheit in der Sekundarstufe I. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 543–595). Wiesbaden: Springer.

- Stanat, P. & Christensen, G. S. (2006). *Where immigrant students succeed: A comparative review of performance and engagement in PISA 2003*. Paris: OECD.
- Stanat, P. & Felbrich, A. (2013). Sprachförderung als Voraussetzung für die Sicherung von Mindeststandards im Bildungssystem: Ansatzpunkte und Herausforderungen. In D. Deißner (Hrsg.), *Chancen bilden: Wege zu einer gerechteren Bildung – ein internationaler Erfahrungsaustausch* (S. 79–100). Wiesbaden: Springer.
- Stanat, P., Segeritz, M. & Christensen, G. (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. 31–58). Münster: Waxmann.
- Wendt, H., Schwippert, K. & Stubbe, T. C. (2016). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 317–331). Münster: Waxmann.

Kapitel 10

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in Deutschland

10.1 Einleitung

Cornelia Gresch, Camilla Rjosk, Aleksander Kocaj und
Petra Stanat

Die schulische Entwicklung von Kindern und Jugendlichen mit einem sonderpädagogischen Förderbedarf (SPF) steht seit etwa einem Jahrzehnt verstärkt im Blickpunkt aller Akteure im Bildungsbereich. Nach der Definition der Kultusministerkonferenz (KMK) haben Schülerinnen und Schüler einen SPF, wenn sie „in ihren Bildungs-, Entwicklungs- und Lernmöglichkeiten so beeinträchtigt sind, dass sie im Unterricht der allgemeinen Schule ohne sonderpädagogische Unterstützung nicht ausreichend gefördert werden können“ (KMK, 1994, S. 5)¹. Es werden in Deutschland insgesamt acht Förderschwerpunkte unterschieden: (1) „Lernen“, (2) „Sprache“, (3) „emotionale und soziale Entwicklung“, (4) „geistige Entwicklung“, (5) „körperlich-motorische Entwicklung“, (6) „Hören“, (7) „Sehen“ und (8) „Unterricht kranker Schülerinnen und Schüler“ (KMK, 1994, 1998). Dabei kann eine Schülerin oder ein Schüler auch eine Beeinträchtigung in mehreren Förderschwerpunkten gleichzeitig aufweisen. Kinder und Jugendliche mit SPF in den Förderschwerpunkten „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“ (LSE) bilden laut amtlicher Statistik neben Kindern und Jugendlichen mit dem Förderschwerpunkt „geistige Entwicklung“ die größte Gruppe der Schülerinnen und Schüler mit SPF (KMK, 2017; vgl. auch Kapitel 4). In einigen Ländern werden die drei Förderschwerpunkte „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“ aufgrund ihrer teilweise auftretenden Überschneidungen zusammengefasst zu einer gemeinsamen Kategorie (LSE).

Nachdem Schülerinnen und Schüler mit SPF in Deutschland lange Zeit gesonderte Förderschulen besuchten, steigt der Anteil von Kindern und Jugendlichen mit SPF, die an allgemeinen Schulen unterrichtet werden, seit 15 Jahren kontinuierlich an. Ein wichtiger Katalysator dieser Entwicklung ist die UN-Behindertenrechtskonvention (UN-BRK), in der das Recht von Menschen mit Behinderung auf eine inklusive² Schulbildung verankert ist (Wrase, 2015, 2017) und die in Deutschland im Jahr 2009 in Kraft trat. Die Umsetzung der UN-BRK

-
- 1 Bei Kindern, die ausschließlich eine Lese-Rechtschreib-Schwäche (Dyslexie) und/oder eine Rechenschwäche (Dyskalkulie) aufweisen, liegt nach dieser Definition kein SPF vor.
 - 2 Über die Definition von Inklusion, die Zielgruppe inklusiver Beschulung und Formen ihrer Umsetzung besteht keine Einigkeit. Für eine ausführliche Darstellung der Debatte siehe z. B. Grosche (2015), Hinz (2002) oder Wocken (2009, 2014).

und die damit verbundene Anpassung der Bildungssysteme stellt für die Länder in Deutschland eine zentrale Herausforderung dar (Baumert, 2016). Dies schlägt sich auch in den Anforderungen nieder, die an Analysen zum Bildungsmonitoring gestellt werden. Betroffen sind unter anderem methodische Fragen, die mit der Einbeziehung von Schülerinnen und Schülern in Schulleistungsstudien verbunden sind. Ferner wird in der überarbeiteten Gesamtstrategie der KMK zum Bildungsmonitoring das Ziel definiert, Schulleistungsstudien verstärkt zur Untersuchung praktischer Schlüsselfragen der Schul- und Unterrichtsentwicklung zu nutzen, und der Umgang mit Heterogenität als besonders relevantes Thema identifiziert. Im IQB-Bildungstrend 2016 wird dieses Schwerpunktthema mit besonderem Fokus auf Schülerinnen und Schüler mit SPF aufgegriffen (vgl. Kapitel 1.1).

Die Einbeziehung von Kindern und Jugendlichen mit SPF in Schulleistungsstudien wie dem IQB-Bildungstrend 2016 ist aus verschiedenen Gründen relevant. Zum einen werden Schülerinnen und Schüler mit SPF teilweise nach einem allgemeinen Lehrplan, und damit auch nach den Bildungsstandards, unterrichtet („zielgleich“), sodass sie zur Zielpopulation von Untersuchungen zum Erreichen der Bildungsstandards in Deutschland gehören und in die Erhebungen einzubeziehen sind. Zum anderen bezieht das Anliegen des Monitorings, die Bildungssituation von Schülerinnen und Schülern anhand von empirischen Daten zu beschreiben, auch Kinder und Jugendliche mit SPF ein, um Stärken und Schwächen von Bildungssystemen zu identifizieren und diese Informationen für die Schul- und Unterrichtsentwicklung zu nutzen. Auch vor dem Hintergrund des zentralen Inklusionsgedankens, Kindern und Jugendlichen mit SPF die vollständige Teilhabe an schulischer Bildung zu ermöglichen, ist ihre Einbeziehung in Studien des Bildungsmonitorings anzustreben (Davies & Elliott, 2012). In der UN-BRK (2008) wird daher ausdrücklich darauf hingewiesen, dass die Entwicklung eines inklusiveren Bildungssystems durch empirische Daten überprüft werden sollte (Artikel 31, UN-BRK, 2008), auch um Steuerungswissen für die Umsetzung der Konvention zur Verfügung zu stellen.

Große Schulleistungsstudien können einen wichtigen Beitrag zur Beschreibung der Lernsituation von Kindern und Jugendlichen mit SPF leisten. Der IQB-Bildungstrend 2016 bietet die Möglichkeit, ausgewählte Aspekte der Lernsituation von Viertklässlerinnen und Viertklässlern mit SPF in Deutschland genauer zu untersuchen, wobei der Fokus auf deskriptiven Analysen liegt. Die Datengrundlage weist besondere Potenziale für weitergehende Analysen auf, hat jedoch auch einige methodische Grenzen.

Eine Stärke des IQB-Bildungstrends 2016 liegt in der großen Stichprobe von Schülerinnen und Schülern der 4. Jahrgangsstufe, die bundesweit repräsentativ für Schülerinnen und Schüler an allgemeinen Schulen ist. Bezogen auf Förderschulen besteht die Einschränkung, dass nur Schulen mit den Förderschwerpunkten „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“ an der Studie teilgenommen haben, da bei diesen Schulen davon auszugehen ist, dass ein substantieller Anteil der Schülerinnen und Schüler zur selbstständigen Bearbeitung der Kompetenztests grundsätzlich in der Lage ist (vgl. Kapitel 3.1). Der IQB-Bildungstrend 2016 bietet damit eine Datengrundlage, die für Kinder der 4. Jahrgangsstufe mit SPF in den drei genannten Förderschwerpunkten an allgemeinen Schulen und Förderschulen bundesweit repräsentativ ist. Sie ermöglicht auch Analysen, in denen zwischen Kindern des Förderschwerpunkts „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ differenziert wird.

Eine weitere Stärke des IQB-Bildungstrends 2016 ist, dass neben standardisierten Kompetenztests umfangreiche Informationen zu individuellen Lernvoraussetzungen und Merkmalen der schulischen Lernumgebung der Schülerinnen und Schüler erhoben wurden. Dies erfolgte mit Fragebogen für Schülerinnen und Schüler, Eltern, Lehrkräfte und Schulleitungen. Zudem wurde das Schwierigkeitsniveau der Testmaterialien für die Schülergruppe mit SPF angepasst (vgl. Kapitel 3.1 und 13.2).³ Mit der vorliegenden Datengrundlage sind Beschreibungen der Lernsituation von Schülerinnen und Schülern mit SPF möglich, die über Analysen anhand von amtlichen Schulstatistiken deutlich hinausgehen (Gresch & Piezunka, 2015; KMK, 2016).

Grenzen der Datengrundlage, die im IQB-Bildungstrend 2016 zur Verfügung steht, betreffen die querschnittliche Anlage der Studie, die Art der erhobenen Informationen und die länderspezifischen Stichprobengrößen. Anhand von Querschnittsdaten, die in der Studie erhoben wurden, lassen sich keine individuellen Entwicklungsverläufe für die Kompetenzen von Schülerinnen und Schülern mit SPF über die Zeit verfolgen; es sind lediglich Kohortenvergleiche in sogenannten Trendanalysen zwischen den Jahren 2011 und 2016 möglich. Auch Analysen, die sich auf die Qualität von Lehr-Lern-Prozessen in Schulen beziehen, können anhand der erhobenen Daten nicht durchgeführt werden. Ebenso muss auf ländervergleichende Analysen für Kinder mit SPF verzichtet werden, da die Stichprobe dieser Gruppe von Schülerinnen und Schülern in den einzelnen Ländern zu klein ist, um belastbare Aussagen treffen zu können.

Trotz der Einschränkungen ermöglichen es die Daten des IQB-Bildungstrends 2016, zur empirischen Beschreibung der Lernsituation von Schülerinnen und Schülern mit SPF beizutragen. Die Forschung dazu steht aktuell noch am Anfang, sodass es sowohl auf inhaltlicher als auch auf methodischer Ebene eine Reihe ungeklärter Fragen gibt, die sich anhand von Daten aus großen Schulleistungsstudien untersuchen lassen. Eine wesentliche Frage betrifft die zur Erfassung eines SPF in Deutschland herangezogenen Kriterien. Ein zentrales Kriterium bildet die amtliche Feststellung eines SPF. Darüber hinaus können Schülerinnen und Schüler jedoch auch ohne amtliche Feststellung sonderpädagogisch gefördert werden. Vor diesem Hintergrund geht Kapitel 10.2 der Frage nach, inwieweit sich die nach den Kriterien der amtlichen Diagnose und sonderpädagogischen Förderung erfassten Schülergruppen mit SPF überschneiden und welche Förderung die Kinder erhalten.

Weiterhin stellt sich die Frage, inwieweit die Diagnose eines SPF und die von Schülerinnen und Schülern mit SPF besuchte Schulart (Förderschule oder allgemeine Schule) mit bestimmten Hintergrundmerkmalen zusammenhängen. Die entsprechenden Analysen in Kapitel 10.3 untersuchen, ob sich die Gruppe der Kinder mit SPF und die Gruppe der Kinder ohne SPF im Hinblick auf die Geschlechterzusammensetzung, die soziale Herkunft und den Zuwanderungshintergrund unterscheiden. Auch Unterschiede zwischen Kindern mit SPF an Förderschulen und Kindern mit SPF an allgemeinen Schulen werden hinsichtlich dieser Hintergrundmerkmale analysiert.

3 Die Anpassung von Testmaterialien für Schülerinnen und Schüler mit SPF wird als besondere Herausforderung der Einbeziehung dieser Schülergruppe in Studien des *Large-Scale-Assessments* diskutiert (Kocaj et al., 2016; Lane & Leventhal, 2015). Kapitel 13.2 dieses Berichtsbands beschreibt den Entwicklungsprozess angepasster Testaufgaben für Kinder mit SPF im Rahmen des IQB-Bildungstrends 2016 sowie die empirische Prüfung der Aufgabengüte.

Förderschulen und allgemeine Schulen bieten unterschiedliche Rahmenbedingungen für die Kompetenz- und Motivationsentwicklung von Kindern und Jugendlichen mit SPF. Wichtig ist daher die Frage, inwieweit die besuchte Schulart (Förderschule oder allgemeine Schule) mit den schulischen Kompetenzen und der schulischen Motivation von Schülerinnen und Schülern mit SPF zusammenhängt. Dies wird in Kapitel 10.4 untersucht.

Damit werden im Kapitel 10 drei zentrale Fragestellungen aufgegriffen, die für eine länderübergreifende Beschreibung der Lernsituation von Kindern mit SPF in Förderschulen und allgemeinen Schulen relevant sind und eine Grundlage für weiterführende Analysen zu Bedingungen und Prozessen des schulischen Lernens von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf bilden.

Literatur

- Baumert, J. (2016). Bildungsgerechtigkeit in Deutschland – ein Überblick. In D. Döring (Hrsg.), *Auf der Suche nach der richtigen Ordnung: Königsteiner Forum 2014* (S. 63–98). Frankfurt am Main: Societäts-Verlag.
- Davies, M. & Elliott, S. N. (2012) Inclusive assessment and accountability: Policy to evidence-based practices. *International Journal of Disability, Development and Education*, 59(1), 1–6.
- Gresch, C. & Piezunka, A. (2015). Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich „Lernen“) an Regelschulen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 181–220). Wiesbaden: Springer VS.
- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer VS.
- Hinz, A. (2002). Von der Integration zur Inklusion – terminologisches Spiel oder konzeptionelle Weiterentwicklung? *Zeitschrift für Heilpädagogik*, 53, 354–361.
- KMK (1994) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1994). *Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland. Beschluß vom 06.05.1994*. Zugriff am 03.08.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopae94.pdf>
- KMK (1998) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1998). *Empfehlungen zum Förderschwerpunkt Unterricht kranker Schülerinnen und Schüler. Beschluß vom 20.03.1998*. Zugriff am 03.08.2017 unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/1998/1998_03_20-Empfehlung-Foerderschwerpunkt-krank-Schueler.pdf
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Sonderpädagogische Förderung in Schulen 2005 bis 2014*. Zugriff am 03.08.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_210_SoPae_2014.pdf
- KMK (2017) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017). *Übersicht der Kultusministerkonferenz. Inklusives Schulsystem – Entwicklungen in den Ländern*. Berlin, Stand: 10.01.2017.
- Kocaj, A., Haag, N., Weirich, S., Kuhl, P., Pant, H. A. & Stanat, P. (2016). Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (S. 212–234). Weinheim: Beltz Juventa.
- Lane, S. & Leventhal, B. (2015). Psychometric challenges in assessing English language learners and students with disabilities. *Review of Research in Education*, 39, 165–214.

- UN-BRK. (2008). Gesetz zu dem Übereinkommen der Vereinten Nationen vom 13. Dezember 2006 über die Rechte von Menschen mit Behinderungen (UN-Behindertenrechtskonvention) vom 21.12.2008, § 2.
- Wocken, H. (2009). Inklusion & Integration. Ein Versuch, die Integration vor der Abwertung und die Inklusion vor Träumereien zu bewahren. In A.-D. Stein, I. Niedek & S. Krach (Hrsg.), *Integration und Inklusion auf dem Wege ins Gemeinwesen* (S. 204–234). Bad Heilbrunn: Klinkhardt.
- Wocken, H. (2014). Frei herumlaufende Irrtümer. Eine Warnung vor pseudoinklusiven Betörungen. *Gemeinsam Leben, 1*, 52–62.
- Wrase, M. (2015). Die Implementation des Rechts auf inklusive Schulbildung nach der UN-Behindertenrechtskonvention und ihre Evaluation aus rechtlicher Perspektive. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 41–74). Wiesbaden: Springer VS.
- Wrase, M. (2017). Das Recht auf inklusive Beschulung nach der UN-Behindertenrechtskonvention – rechtliche Anforderungen und Umsetzung in Deutschland. In B. Lütje-Klose, M. A. Boger, B. Hopmann & P. Neumann (Hrsg.), *Inklusion in die Leistungsgesellschaft. Band 1: Menschenrechtliche, theoretische und professionsbezogene Perspektiven* (S. 17–25). Bad Heilbrunn: Klinkhardt.

10.2 Amtlich festgestellter sonderpädagogischer Förderbedarf und sonderpädagogische Förderung

Cornelia Gresch, Jenny Kölm und Aleksander Kocaj

10.2.1 Einleitung

Spätestens seit Inkrafttreten der UN-Behindertenrechtskonvention (UN-BRK) im Jahr 2009 werden in allen Ländern in der Bundesrepublik Deutschland Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) verstärkt an allgemeinen Schulen unterrichtet (KMK, 2011). Für Schulleistungsstudien wie dem IQB-Bildungstrend und auch für die amtlichen Statistiken ist es sinnvoll, diese Schülerinnen und Schüler möglichst präzise erfassen zu können. Nur so ist es möglich, ihre Lernsituation zu beschreiben (vgl. Kapitel 10.3 und 10.4) und, wie in Artikel 31 der UN-BRK (2008) gefordert, die Entwicklung eines inklusiven Bildungssystems¹ zu analysieren. Es gibt jedoch unterschiedliche Herangehensweisen bei der Erfassung dieser Schülergruppe. So kann ein amtlich festgestellter SPF als Grundlage herangezogen werden. Allerdings werden Schülerinnen und Schüler mit SPF vermehrt auch unabhängig von amtlichen Verfahren sonderpädagogisch gefördert (z. B. KMK, 2016). Deshalb empfehlen Piezunka, Gresch, Sälzer und Kroth (2016), in Schulleistungsstudien sowohl Informationen zum amtlich festgestellten SPF als auch zur sonderpädagogischen Förderung zu erheben. Auf diese Weise kann sichergestellt werden, dass möglichst alle Schülerinnen und Schüler mit SPF erfasst werden.

Auch für die amtliche Statistik spielt diese Unterscheidung zwischen amtlichem Feststellungsverfahren und sonderpädagogischer Förderung eine Rolle. Die Kultusministerkonferenz (KMK) sammelt gezielt Informationen zur sonderpädagogischen Förderung (vgl. KMK, 2016, 2017a), wobei in den Ländern häufig auf Angaben zum amtlich festgestellten SPF zurückgegriffen wird (vgl. Malecki, 2013). Im Falle einer sonderpädagogischen Förderung unabhängig von einem Feststellungsverfahren bestehen jedoch Unsicherheiten, die in einzelnen Ländern dazu führen können, dass nicht alle Schülerinnen und Schüler erfasst werden (vgl. KMK, 2016).

Im IQB-Bildungstrend 2016 werden beide Kriterien parallel erhoben: (1) die amtliche Feststellung eines SPF und (2) die sonderpädagogische Förderung. Anhand dieser Daten wird im vorliegenden Beitrag untersucht, wie stark sich die nach diesen Kriterien erfassten Schülergruppen mit SPF überschneiden beziehungsweise unterscheiden.

Die Analysen zu dieser Fragestellung, die im vorliegenden Kapitel dargestellt werden, beziehen sich auf Schülerinnen und Schüler mit den Förderschwerpunkten „Lernen (L)“, „Sprache (S)“ und „emotionale und soziale Entwicklung (E)“, die nachfolgend zusammenfassend als „LSE“ bezeichnet werden. Die Ergebnisse werden dabei soweit wie möglich auch nach Schulart (Förderschulen und allgemeine Schulen) differenziert dargestellt. Abschnitt 10.2.2 beschreibt zu-

1 In dem vorliegenden Beitrag bezeichnet der Begriff „Inklusion“ die Beschulung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an allgemeinen Schulen. Für eine kritische Auseinandersetzung mit der Bezeichnung und dem zugrunde liegenden Verständnis vgl. beispielsweise Grosche (2015), Hinz (2002) und Wocken (2009; 2014).

nächst die beiden Kriterien zur Erfassung eines SPF genauer, um im Anschluss die anhand dieser Kriterien erfassten Schülergruppen empirisch gegenüberzustellen (Abschnitt 10.2.3). Das Kapitel schließt mit einer Diskussion der Implikationen, die sich aus den Unterschieden für eine länderübergreifende Erfassung von Schülerinnen und Schülern mit SPF und somit auch für die IQB-Bildungstrends und weitere *Large-Scale-Assessments* im Bildungsbereich ergeben (Abschnitt 10.2.4).

10.2.2 Kriterien zur Erfassung eines sonderpädagogischen Förderbedarfs

Die Kriterien zur Bestimmung der Schülergruppe mit SPF und die anschließende Zuweisung einer Schülerin oder eines Schülers zu einer Förderschule oder allgemeinen Schule haben sich in den letzten 20 Jahren verändert. In Einklang mit internationalen Konventionen zur inklusiven beziehungsweise integrativen Beschulung (z. B. Salamanca-Erklärung, *The United Nations Educational, Scientific and Cultural Organization*, UNESCO, 1994) wurde der bis dahin verwendete Begriff der „Sonderschulbedürftigkeit“ im Jahr 1994 durch den Begriff des „sonderpädagogischen Förderbedarfs“ abgelöst (KMK, 1994). Damit sollte unter anderem betont werden, dass nicht nur Förderschulen, sondern auch allgemeine Schulen für den Unterricht von Kindern und Jugendlichen mit SPF zuständig sein können, sofern die einzelnen Schülerinnen und Schüler dort mit sonderpädagogischer Unterstützung hinreichend gefördert werden (KMK, 1994, 2011).

Bundesweit spielen amtliche Feststellungsverfahren eine wesentliche Rolle für die Zuweisung eines SPF. Dabei bestehen jedoch Unterschiede zwischen den Ländern, unter anderem hinsichtlich der Förderschwerpunkte, für die ein amtliches Feststellungsverfahren notwendig ist, des Zeitpunkts der Feststellung und der am Verfahren beteiligten Akteure (vgl. Piezunka et al., 2016; Sälzer, Gebhard, Müller & Pauly, 2015). In den Feststellungsverfahren werden wissenschaftlich fundierte Diagnoseverfahren eingesetzt, allerdings kommen unterschiedliche Methoden zum Einsatz (Kany & Schöler, 2009; Sälzer et al., 2015), die auf verschiedene Aspekte der schulischen Entwicklung abzielen (Petermann & Petermann, 2006). In den meisten Ländern trifft anschließend die Schulaufsichtsbehörde die endgültige Entscheidung darüber, an welcher Schulart die Schülerin oder der Schüler unterrichtet werden soll, wobei die elterlichen Wünsche in unterschiedlichem Maße berücksichtigt werden (z. B. Sälzer et al., 2015). Die Überweisung an eine Förderschule setzt dabei in der Regel ein amtliches Feststellungsverfahren voraus. Im Gegensatz dazu wird an allgemeinen Schulen teilweise von einem amtlichen Feststellungsverfahren abgesehen (vgl. KMK, 2017a, 2017b). Diese Praxis, die im Zuge der Inklusion zunehmend Anwendung findet, betrifft insbesondere die Förderschwerpunkte LSE, da argumentiert wird, dass hier eine klare Abgrenzung zu einer typischen kindlichen Entwicklung oftmals schwierig ist (Klauer & Lauth, 1997; Linderkamp & Grünke, 2007; Ricking, 2005). Die Förderschwerpunkte „Sehen“, „Hören“ und „körperliche und motorische Entwicklung“ sind dagegen in der Regel eindeutiger zu bestimmen, da es sich um medizinisch feststellbare Behinderungen handelt (z. B. Bos, Müller & Stubbe, 2010).

Eine sonderpädagogische Förderung ohne amtliches Feststellungsverfahren erfolgt in einigen Ländern auf der Basis pauschaler Ressourcenzuweisungen an die allgemeinen Schulen (KMK, 2016). Eine Schule erhält dabei zusätzliche Ressourcen für sonderpädagogische Unterstützung, die nicht von der Anzahl

der Schülerinnen und Schüler mit einer amtlichen Diagnose eines SPF abhängt, sondern von der sozialen Zusammensetzung der Schülerschaft, die über einen Sozialindex ermittelt wird (z. B. Rauer & Schuck, 2007). Wie die zusätzlichen Ressourcen eingesetzt werden, bleibt der Schule überlassen. Teilweise liegt der Entscheidung, welche Schülerinnen und Schüler durch eine sonderpädagogische Lehrkraft gefördert werden, ein schulinternes Verfahren zugrunde, bei dem bestimmt wird, ob ein SPF besteht. Teilweise wird aber auch unabhängig von einem solchen Verfahren individuell gefördert. Sonderpädagogische Lehrkräfte können dabei unterschiedliche Funktionen übernehmen, wie etwa Kleingruppenförderung durchführen, vollständige Klassen (Schülerinnen und Schüler mit und ohne SPF) im Team-Teaching beziehungsweise allein unterrichten oder das Kollegium der allgemeinen Schule beraten (Lütje-Klose & Neumann, 2015; Prengel, 2013).

Bei der teilweisen Entkoppelung sonderpädagogischer Förderung von einem amtlichen Feststellungsverfahren wird insbesondere im Primarbereich auf die damit verbundene Flexibilität der Förderung verwiesen (Prengel, 2013), die es erlaubt, kurzfristig auf Unterstützungsbedarfe einzelner Schülerinnen und Schüler zu reagieren. Ferner wird argumentiert, dass Feststellungsverfahren eine stigmatisierende Wirkung haben könnten (Pfahl, 2010; Schumann, 2007). Darüber hinaus verweisen Kritiker auf das Etikettierungs-Ressourcen-Dilemma (vgl. Füssel & Kretschmann, 1993), wonach Schulen vermehrt sonderpädagogische Diagnosen anstreben, um zusätzliche Ressourcen zu erhalten. Auf der anderen Seite wird argumentiert, dass Feststellungsverfahren wichtig sind, um Ressourcen gezielter bereitstellen zu können und die Förderung auf individuelle Bedürfnisse der Kinder und Jugendlichen abzustimmen (Ahrbeck, 2014; Wrase, 2015, 2017). Erfolgt eine sonderpädagogische Förderung ohne Feststellungsverfahren, bleibt letztlich offen, unter welchen Bedingungen Schülerinnen und Schüler Förderung aus pauschal zugewiesenen Ressourcen erhalten und wie sonderpädagogische Förderung individuell umgesetzt wird.

Aufgrund der beschriebenen Unterschiede in der Bestimmung eines sonderpädagogischen Förderbedarfs ist unklar, inwieweit sich die beiden Kriterien – amtlich festgestellter Förderbedarf einerseits und sonderpädagogische Förderung andererseits – überlappen. Die Klärung dieser Frage ist für die Beschreibung der betroffenen Schülergruppen und ihrer Lernsituation jedoch zentral. Anhand von Daten des IQB-Bildungstrends 2016, in dem sowohl der amtlich festgestellte SPF als auch sonderpädagogische Förderung parallel erhoben wurde, können die anhand dieser beiden Kriterien erfassten Schülergruppen einander gegenübergestellt werden. „Sonderpädagogische Förderung“ wird dabei in Anlehnung an Boban und Hinz (2008) als Förderung eines Kindes durch eine sonderpädagogische Lehrkraft definiert. Angesichts unterschiedlicher Konzeptualisierungen von sonderpädagogischer Förderung (z. B. Glück, 2015; Miller & Kottmann, 2016) werden in einer vertiefenden Darstellung auch weitere Formen der besonderen individuellen Unterstützung in Schulen berücksichtigt.

10.2.3 Schülerinnen und Schüler mit amtlich festgestelltem sonderpädagogischem Förderbedarf und sonderpädagogischer Förderung im IQB-Bildungstrend 2016

Informationen über amtlich festgestellten SPF und sonderpädagogische Förderung wurden im IQB-Bildungstrend 2016 jeweils von den Lehrkräften erfragt. Für jede einzelne Schülerin und jeden einzelnen Schüler wurde angegeben, in welchen Förderschwerpunkten gegebenenfalls ein amtlich festgestellter SPF vorliegt. Die im Folgenden dargestellten Analysen beziehen alle Kinder ein, für die mindestens einer der Förderschwerpunkte „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ angegeben wurde. Hatte die Lehrkraft zusätzlich weitere Förderschwerpunkte aufgeführt, werden diese Schülerinnen und Schüler ebenfalls einbezogen. Schülerinnen und Schüler, für die ein SPF ausschließlich in anderen Förderschwerpunkten („geistige Entwicklung“, „körperliche und motorische Entwicklung“, „Hören“, „Sehen“ oder „Unterricht kranker Schülerinnen und Schüler“) besteht, wurden aus den Analysen ausgeschlossen.

Als Schülerinnen und Schüler mit sonderpädagogischer Förderung werden hier diejenigen Kinder betrachtet, die längerfristig (mindestens sechs Monate)² zum Beispiel in Einzel- oder Kleingruppensettings durch eine Sonderpädagogin beziehungsweise einen Sonderpädagogen gefördert werden. Dabei spielt es keine Rolle, ob diese Kinder ein amtliches Feststellungsverfahren durchlaufen haben oder nicht. Es handelt sich somit um eine enge Definition der sonderpädagogischen Förderung, die an sonderpädagogische Lehrkräfte gebunden ist. Um weitere Arten der Förderung zu erfassen, die gegebenenfalls auch als „sonderpädagogische Förderung“ wahrgenommen wird (z. B. Miller & Kottmann, 2016), wurde geprüft, ob die einzelnen Schülerinnen und Schüler längerfristig durch anderes pädagogisches Personal unterstützt werden und ob sie sonstige Unterstützung, beispielsweise durch eine Assistenz, therapeutische Maßnahmen oder auch Sprach- oder Leseförderung, erhalten.

Eine Eingrenzung der Schülerinnen und Schüler mit sonderpädagogischer Förderung auf die Förderschwerpunkte LSE kann lediglich annäherungsweise erfolgen, da nur für Schülerinnen und Schüler mit einem amtlich festgestellten sonderpädagogischen Förderbedarf auch Informationen zum konkreten Förderschwerpunkt vorliegen. Allerdings ist anzunehmen, dass die sonderpädagogisch geförderten Schülerinnen und Schüler überwiegend in den Förderschwerpunkten LSE gefördert werden, da es sich mit über 80 Prozent um die anteilig größte Gruppe aller sonderpädagogisch geförderten Kinder handelt (vgl. KMK, 2016; siehe auch Kapitel 4). Zudem wurden Schülerinnen und Schüler mit anderen Förderschwerpunkten (z. B. „geistige Entwicklung“ oder „Hören“) aus den Analysen ausgeschlossen. Schließlich wird insbesondere in den Bereichen LSE teilweise auf eine amtliche Feststellung verzichtet (vgl. KMK, 2017b).

Amtlich festgestellter sonderpädagogischer Förderbedarf an allgemeinen Schulen und an Förderschulen

Den Daten des IQB-Bildungstrends 2016 zufolge weisen in Deutschland insgesamt 5.6 Prozent aller Schülerinnen und Schüler einen amtlich festgestellten SPF in den Förderschwerpunkten LSE (inklusive mehrerer Förderschwerpunkte)

2 Die zeitliche Festlegung auf mindestens sechs Monate erfolgt in Anlehnung an die Behinderungsdefinition nach §2 SGB IX (vgl. Autorengruppe Bildungsberichterstattung, 2014).

auf. Fast 54 Prozent dieser Kinder besuchen eine allgemeine Schule und etwa 46 Prozent eine Förderschule. An allgemeinen Schulen beträgt der Anteil der Schülerinnen und Schüler mit einem amtlich festgestellten SPF in den Schwerpunkten LSE damit rund 3 Prozent. Für insgesamt 0.1 Prozent der am IQB-Bildungstrend 2016 teilnehmenden Schülerinnen und Schüler an allgemeinen Schulen liegen keine Angaben darüber vor, ob ein sonderpädagogischer Förderbedarf besteht oder nicht.

Sonderpädagogische Förderung

Die Auswertungen zur sonderpädagogischen Förderung beschränken sich auf allgemeine Schulen, da davon auszugehen ist, dass an Förderschulen die Unterrichtsgestaltung grundsätzlich durch sonderpädagogische Lehrkräfte³ erfolgt. Für die folgenden Analysen wird – unabhängig von der amtlichen Feststellung eines SPF – danach unterschieden, ob eine Schülerin beziehungsweise ein Schüler an einer allgemeinen Schule sonderpädagogische Förderung (ggf. gekoppelt mit weiteren Formen der besonderen Unterstützung) erhält oder nicht.

Der Anteil der Schülerinnen und Schüler an allgemeinen Schulen, die sonderpädagogisch gefördert werden, beträgt 3 Prozent. Dieser Anteil ist somit genauso hoch wie der Anteil der Schülerinnen und Schüler an allgemeinen Schulen, die einen festgestellten sonderpädagogischen Förderbedarf in den Förderschwerpunkten LSE aufweisen (vgl. oben). Etwa 96 Prozent der Kinder an allgemeinen Schulen erhalten keine besondere Unterstützung durch eine sonderpädagogische Lehrkraft. Für knapp 1 Prozent der Schülerinnen und Schüler liegen keine Angaben zur sonderpädagogischen Förderung vor.

Amtliches Feststellungsverfahren und sonderpädagogische Förderung

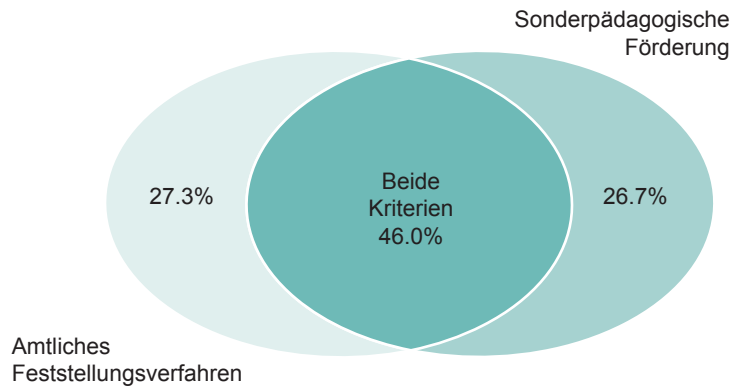
Wie oben ausgeführt, kann eine Förderung durch sonderpädagogische Lehrkräfte unabhängig von einem amtlichen Feststellungsverfahren erfolgen. In einem nächsten Schritt wird nun untersucht, inwieweit anhand der beiden Kriterien dieselben Schülerinnen und Schüler erfasst werden. Dies ist in Abbildung 10.1 dargestellt. Kinder mit fehlenden Angaben werden nicht berücksichtigt.

Die Grundgesamtheit umfasst alle Schülerinnen und Schüler an allgemeinen Schulen, bei denen ein amtlich festgestellter SPF besteht oder die sonderpädagogische Förderung erhalten (= 100 %). Von dieser Gesamtgruppe weisen 46 Prozent einen amtlich festgestellten SPF auf und erhalten zudem Förderung durch eine sonderpädagogische Lehrkraft. Für jeweils rund 27 Prozent der Schülerinnen und Schüler wurde entweder nur ein amtlich festgestellter SPF (27.3 %) oder nur eine sonderpädagogische Förderung (26.7 %) berichtet.

Dass auch Schülerinnen und Schüler ohne amtliches Feststellungsverfahren sonderpädagogisch gefördert werden, ist nicht überraschend, da in den meisten Ländern Formen der sonderpädagogischen Förderung auf Basis pauschaler Ressourcenzuweisungen möglich sind. Es fällt jedoch auf, dass Schülerinnen und Schüler mit amtlich festgestelltem SPF in den Schwerpunkten LSE teil-

3 Allerdings sind nicht nur an allgemeinen Schulen, sondern auch an Förderschulen in nicht unerheblichem Umfang Lehrkräfte mit anderen Lehramtsabschlüssen tätig, die mit Ausnahme der Durchführung von sonderpädagogischen Feststellungsverfahren alle anderen sonderpädagogisch relevanten Tätigkeiten (Erziehen, Unterrichten, Diagnostizieren, Fördern) übernehmen können.

Abbildung 10.1: Überlappung der Anteile von Schülerinnen und Schülern mit amtlich festgestelltem sonderpädagogischem Förderbedarf (SPF) und sonderpädagogischer Förderung an allgemeinen Schulen⁴



weise keine Förderung durch eine Sonderpädagogin beziehungsweise einen Sonderpädagogen erhalten. Eine mögliche Erklärung besteht darin, dass sich die „sonderpädagogische Förderung“ im Rahmen der vorliegenden Analysen auf eine längerfristige, mindestens sechsmonatige Förderung des Kindes durch eine sonderpädagogische Lehrkraft bezieht. Wie oben ausgeführt, muss die Förderung des Kindes jedoch nicht zwingend durch sonderpädagogische Lehrkräfte erfolgen. In der Praxis kann dies zum Beispiel bedeuten, dass sonderpädagogische Lehrkräfte die allgemeinpädagogischen Lehrkräfte zwar beraten, selbst aber nicht unterrichten und die Förderung durch die allgemeinen Lehrkräfte erfolgt.

Im Folgenden wird dargestellt, inwieweit Schülerinnen und Schüler mit amtlich festgestelltem SPF weitere Formen der besonderen Unterstützung im Schulalltag erhalten (vgl. Tab. 10.1). Betrachtet man ausschließlich die Schülerinnen und Schüler mit einem amtlich festgestellten SPF in den Schwerpunkten LSE, werden etwa 63 Prozent von ihnen durch sonderpädagogische Lehrkräfte gefördert. Dies schließt Schülerinnen und Schüler ein, die neben der sonderpädagogischen Förderung weitere Unterstützung erhalten. Darüber hinaus bekommen rund 12 Prozent der Kinder mit SPF an allgemeinen Schulen längerfristig besondere Unterstützung durch weiteres pädagogisches Personal und etwa 12 Prozent andere Formen der Unterstützung in Form von Assistenzen, therapeutischen Maßnahmen oder Sprach- und Leseförderung. Diese Anteile sind deutlich höher als die entsprechenden Anteile in der Gruppe von Schülerinnen und Schülern ohne amtlich festgestellten SPF⁵ (vgl. Tab. 10.1, rechte Spalte). Es verbleibt jedoch eine Gruppe von rund 14 Prozent der Kinder, die an allgemeinen Schulen trotz festgestelltem SPF in den Schwerpunkten LSE keine besondere Unterstützung im Unterricht zu erhalten scheint.

Eine mögliche Erklärung dafür, dass ein Teil der Schülerinnen und Schüler mit diagnostiziertem SPF nach Angabe der Schule keine besondere Förderung erhält, könnte in der zeitlichen Abfolge zwischen amtlichem Feststellungsverfahren und sonderpädagogischer Förderung liegen. Die Zuweisung sonderpädagogischer Unterstützungsressourcen erfolgt häufig zu Beginn eines Schuljahres, sodass bei Feststellungsverfahren, die erst später im Schuljahr stattfinden und positiv aus-

4 Es handelt sich um eine schematische Darstellung, in der die Flächenverhältnisse nicht exakt proportional zu den Prozentangaben sind.

5 Schülerinnen und Schüler ohne SPF erhalten teilweise auch Förderung für besonders Leistungsstarke. Diese sind ebenfalls unter „andere Formen der Unterstützung“ zusammengefasst.

Tabelle 10.1: Sonderpädagogische Förderung und weitere Unterstützung von Schülerinnen und Schülern an allgemeinen Schulen in Abhängigkeit von einem amtlich festgestellten SPF

Sonderpädagogische Förderung und weitere Unterstützung im Schulalltag	Amtlich festgestellter SPF in den Bereichen LSE (inkl. Doppeldiagnosen)	
	ja	nein
Längerfristige Förderung durch eine sonderpädagogische Lehrkraft	62.5	1.2
Längerfristig besondere Unterstützung durch weiteres pädagogisches Personal	11.9	6.4
Andere Formen der Unterstützung	11.6	4.3
Keine besondere Unterstützung	14.1	88.2
Gesamt	100	100

Anmerkungen. Anteile in Prozent. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. Schülerinnen und Schüler mit fehlenden Angaben sind ausgeschlossen. SPF = sonderpädagogischer Förderbedarf; LSE = Förderschwerpunkte „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“. Andere Formen der Unterstützung = Unterstützung durch Assistenz, therapeutische Maßnahmen, Förderung für besonders leistungsstarke Schülerinnen und Schüler (nur in der Gruppe ohne SPF) oder Sprach- und Leseförderung.

fallen, gegebenenfalls noch keine Ressourcen für sonderpädagogische Förderung freigegeben werden konnten.

Eine weitere mögliche Erklärung liegt darin, dass sonderpädagogische Lehrkräfte teilweise auch durch die Erstellung von Förderplänen oder die Beratung allgemeinpädagogischer Lehrkräfte und Eltern indirekt fördernd tätig sind, ohne selbst mit dem Kind zu arbeiten. Inwieweit diese Form der indirekten Förderung durch Beratungsarbeit der sonderpädagogischen Lehrkräfte hier vorliegt, ob tatsächlich trotz Feststellung keine sonderpädagogische Förderung erfolgt oder ob weitere Ursachen eine Rolle spielen, kann an dieser Stelle nicht geklärt werden.

10.2.4 Implikationen für die länderübergreifende Erfassung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf

Das Konzept des SPF ist als Grundlage für aussagekräftige statistische Darstellungen mit einfachen und klaren Kriterien eine besondere Herausforderung, da bislang weder in der Literatur noch in der Praxis Einigkeit darüber besteht, unter welchen Bedingungen eine Schülerin beziehungsweise ein Schüler als Person mit SPF oder sonderpädagogischer Förderung erfasst wird. Sowohl der amtlich festgestellte SPF als auch die sonderpädagogische Förderung können hierbei eine Rolle spielen. Hinzu kommt, dass auch diese Kriterien in sehr unterschiedlicher Weise definiert werden: Den amtlichen Feststellungsverfahren liegen in den Ländern unterschiedliche Diagnoseverfahren zugrunde, und die sonderpädagogische Förderung erfolgt mit unterschiedlichen Maßnahmen. Diese Heterogenität sollte nach Möglichkeit bei der statistischen Erfassung der betroffenen Schülergruppe in den IQB-Bildungstrendstudien und weiteren bundesweiten Studien berücksichtigt werden. Im ersten Schritt empfiehlt es sich, grundsätzlich – bei bundesweiten Analysen ebenso wie bei Ländervergleichen – sowohl die amtliche Feststellung als auch die sonderpädagogische Förderung als Kriterien zu berücksichtigen. Dabei sollte das jeweilige Verständnis dieser Kriterien in den Ländern benannt und beschrieben werden. Nur auf dieser Grundlage ist es möglich, die Situation der Schülergruppe, um die es in den Bemühungen um Inklusion unter anderem geht, genauer zu beschreiben und fortlaufend zu beobachten.

Literatur

- Ahrbeck, B. (2014). Schulische Inklusion – Möglichkeiten, Dilemmata und Widersprüche. *Soziale Passagen*, 6, 5–19.
- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Boban, I. & Hinz, A. (2008). Sonderpädagogische Förderung in der Allgemeinen Schule (Integration) und in Sonderschulen. In K.-H. Arnold, O. Graumann & A. Rachkockine (Hrsg.), *Handbuch Förderung. Grundlagen, Bereiche und Methoden der individuellen Förderung von Schülern* (S. 410–419). Weinheim: Beltz.
- Bos, W., Müller, S. & Stubbe, T. C. (2010). Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 375–397). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Füssel, H.-P. & Kretschmann, R. (1993). *Gemeinsamer Unterricht für behinderte und nicht-behinderte Kinder*. Witterschlick/Bonn: Wehle.
- Glück, C. (2015). Forschung zur Inklusion von Kindern mit Sprachstörungen. In M. Grohnfeldt (Hrsg.), *Inklusion im Förderschwerpunkt Sprache* (S. 129–144). Stuttgart: Kohlhammer.
- Gresch, C. & Piezunka, A. (2015). Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich „Lernen“) an Regelschulen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 181–220). Wiesbaden: Springer VS.
- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer VS.
- Hinz, A. (2002). Von der Integration zur Inklusion – terminologisches Spiel oder konzeptionelle Weiterentwicklung? *Zeitschrift für Heilpädagogik*, 53(9), 354–361.
- Kany, W. & Schöler, H. (2009). *Diagnostik schulischer Lern- und Leistungsschwierigkeiten. Ein Leitfaden*. Stuttgart: Kohlhammer.
- Klauer, K. J. & Lauth, G. W. (1997). Lernbehinderungen und Leistungsschwierigkeiten bei Schülern. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie: Pädagogische Psychologie* (Bd. 3: Psychologie des Unterrichts und der Schule, S. 701–738). Göttingen: Hogrefe.
- KMK (1994) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1994). *Empfehlungen zur sonderpädagogischen Förderung in den Schulen in der Bundesrepublik Deutschland. Beschluß vom 06.05.1994*. Zugriff am 08.06.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopae94.pdf>
- KMK (2011) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2011). *Inklusive Bildung von Kindern und Jugendlichen mit Behinderungen in Schulen. Beschluss der Kultusministerkonferenz (KMK) vom 20.10.2011*. Zugriff am 13.07.2012 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2011/2011_10_20-Inklusive-Bildung.pdf
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Sonderpädagogische Förderung in Schulen 2005 bis 2014*. Zugriff am 08.06.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_210_SoPae_2014.pdf
- KMK (2017a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017). *Definitionenkatalog zur Schulstatistik 2017*. Zugriff am 10.09.2017 unter <https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Defkat2017.pdf>
- KMK (2017b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017). *Übersicht der Kultusministerkonferenz. Inklusives Schulsystem – Entwicklungen in den Ländern*. Berlin, Stand: 10.01.2017.
- Linderkamp, F. & Grünke, M. (2007). Lern- und Verhaltensstörungen: Klassifikation, Prävalenz & Prognostik. *Lern- und Verhaltensstörungen. Genese-Diagnostik-Intervention* (S. 14–28). Weinheim und Basel: Beltz.

- Lütje-Klose, B. & Neumann, P. (2015): Die Rolle der Sonderpädagogik im Rahmen der Lehrerinnen- und Lehrerprofessionalisierung für eine inklusive schulische Bildung. In T. Häcker & M. Walm (Hrsg.), *Inklusion – Herausforderung für Schulentwicklung und Lehrerbildung* (S. 101–116). Bad Heilbrunn: Klinkhardt.
- Malecki, A. (2013). Sonderpädagogische Förderung in Deutschland – eine Analyse der Datenlage in der Schulstatistik. *Wirtschaft und Statistik*, 5, 356–365.
- Miller, S. & Kottmann, B. (2016). Grundschullehrkräfte im Entscheidungsdilemma zwischen Fördern und Selektieren. In B. Amrhein (Hrsg.), *Diagnostik im Kontext inklusiver Bildung* (S. 154–167). Bad Heilbrunn: Klinkhardt.
- Petermann, U. & Petermann, F. (Hrsg.). (2006). *Diagnostik sonderpädagogischen Förderbedarfs*. Göttingen: Hogrefe.
- Pfahl, L. (2010). *Disability Studies. Bd. 7: Techniken der Behinderung. Der deutsche Lernbehinderungsdiskurs, die Sonderschule und ihre Auswirkungen auf Bildungsbiografien*. Bielefeld: transcript.
- Piezunka, A., Gresch, C., Sälzer, C. & Kroth, A. (2016). Identifizierung von Schülerinnen und Schülern nach Vorgaben der UN-BRK in bundesweiten Erhebungen: Sonderpädagogischer Förderbedarf, sonderpädagogische Förderung oder besondere Unterstützung? In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion. 62. Beiheft der Zeitschrift für Pädagogik* (S. 190–211). Weinheim: Beltz.
- Prenzel, A. (2013). *Inklusive Bildung in der Primarstufe. Eine wissenschaftliche Expertise für den Grundschulverband*. Frankfurt: Grundschulverband.
- Rauer, W. & Schuck, K. D. (2007). Hamburger Grundschulen und Grundschulklassen mit einer formellen Integrationsorganisation. In W. Bos, C. Gröhlich & M. Pietsch (Hrsg.), *KESS 4 – Lehr- und Lernbedingungen in Hamburger Grundschulen* (S. 219–254). Münster: Waxmann.
- Ricking, H. (2005). Der „Overlap“ von Lern- und Verhaltensstörungen. *Sonderpädagogik*, 35, 235–248.
- Sälzer, C., Gebhard, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 129–152). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schumann, B. (2007). „Ich schäme mich ja so!“. *Die Sonderschule für Lernbehinderte als „Schonraumfalle“*. Bad Heilbrunn: Klinkhardt.
- UNESCO (1994) = United Nations Educational, Scientific and Cultural Organization. (1994). The Salamanca Statement and Framework for Action on Special Needs Education. Adopted by the World Conference on Special Needs Education: Access and Quality. Salamanca, Spain, 7–10 June 1994. UNESCO.
- Wocken, H. (2009). Inklusion & Integration. Ein Versuch, die Integration vor der Abwertung und die Inklusion vor Träumereien zu bewahren. In A.-D. Stein, I. Niedek & S. Krach (Hrsg.), *Integration und Inklusion auf dem Wege ins Gemeinwesen* (S. 204–234). Bad Heilbrunn: Klinkhardt.
- Wocken, H. (2014). Frei herumlaufende Irrtümer. Eine Warnung vor pseudoinklusiven Betörungen. *Gemeinsam Leben*, 1, 52–62.
- Wrase, M. (2015). Die Implementation des Rechts auf inklusive Schulbildung nach der UN-Behindertenrechtskonvention und ihre Evaluation aus rechtlicher Perspektive. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 41–74). Wiesbaden: Springer VS.
- Wrase, M. (2017). Das Recht auf inklusive Beschulung nach der UN-Behindertenrechtskonvention – rechtliche Anforderungen und Umsetzung in Deutschland. In B. Lütje-Klose, M. A. Boger, B. Hopmann & P. Neumann (Hrsg.), *Inklusion in die Leistungsgesellschaft. Band 1: Menschenrechtliche, theoretische und professionsbezogene Perspektiven* (S. 17–25). Bad Heilbrunn: Klinkhardt.

10.3 Hintergrundmerkmale von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an Förderschulen und an allgemeinen Schulen

Jenny Kölm, Cornelia Gresch und Nicole Haag

10.3.1 Einleitung

Wie erfolgreich Kinder und Jugendliche in der Schule sind, hängt mit einer Reihe von Faktoren zusammen. Hierzu gehören neben Merkmalen der Unterrichtsqualität auch verschiedene Hintergrundmerkmale der Schülerinnen und Schüler, insbesondere das Geschlecht, der sozioökonomische Status (SES) der Familie sowie der Zuwanderungshintergrund. In den Kapiteln 7 bis 9 wird untersucht, inwieweit diese Hintergrundmerkmale mit den Kompetenzen assoziiert sind, die Schülerinnen und Schüler in den Fächern Deutsch und Mathematik bis zur 4. Jahrgangsstufe erreicht haben. Dabei bestätigt sich der Befund, dass in den erreichten Kompetenzen teilweise erhebliche geschlechtsbezogene, soziale und zugewanderungsbezogene Disparitäten bestehen. Im vorliegenden Kapitel wird geprüft, inwieweit diese Hintergrundmerkmale auch mit dem amtlich festgestellten sonderpädagogischen Förderbedarf (SPF) in den Förderschwerpunkten „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ zusammenhängen. Ferner wird der Frage nachgegangen, ob Schülerinnen und Schüler mit einem SPF in Abhängigkeit von den drei genannten Hintergrundmerkmalen eher eine Förderschule oder eher eine allgemeine Schule besuchen.

Verschiedene Gründe sprechen dafür, dass die amtliche Feststellung eines SPF mit dem Geschlecht des Kindes, mit dem sozioökonomischen Status und mit dem Zuwanderungshintergrund der Familie assoziiert sein könnte. So hat sich wiederholt gezeigt, dass Kinder aus sozial schwachen Familien und Kinder aus zugewanderten Familien geringere Kompetenzwerte in bestimmten Bereichen erzielen als Kinder aus sozial besser gestellten Familien und Kinder ohne Zuwanderungshintergrund; zudem liegen bedeutsame geschlechtsbezogene Disparitäten in verschiedenen Kompetenzen vor (vgl. Kapitel 7 bis 9). Daher sind für die Feststellung eines SPF, die mit dem allgemeinen Leistungs- und Entwicklungsstand zusammenhängen sollte,¹ ebenfalls entsprechende Disparitäten zu erwarten.

Zusätzlich könnten diese Hintergrundmerkmale aber auch unabhängig von den erreichten Kompetenzen beeinflussen, ob bei einem Kind ein SPF festgestellt wird oder nicht. Das amtliche Feststellungsverfahren kann bei Bedarf (je nach Schulgesetz) durch Lehrkräfte, andere institutionelle Akteure (z. B. Ärztinnen und Ärzte des Gesundheitsamtes bei der Schuleingangsuntersuchung) oder durch die Eltern eingeleitet werden (vgl. Sälzer, Gebhardt, Müller & Pauly, 2015). Dabei wäre es möglich, dass stereotype Wahrnehmungsmuster, Urteilsverzerrungen oder Vorurteile von Lehrkräften oder anderen institutionellen Akteuren dazu führen, dass bei bestimmten Schülergruppen häufiger ein Verfahren eingeleitet und häufiger die Diagnose eines SPF vergeben wird (z. B.

¹ Dies gilt vor allem für den Förderschwerpunkt „Lernen“, aber auch bei der Feststellung eines SPF in den Förderschwerpunkten „Sprache“ sowie „emotionale und soziale Entwicklung“ sollte der allgemeine Leistungs- und Entwicklungsstand berücksichtigt werden.

Gomolla, 2013; Lanfranchi, 2007). Auf Seiten der Eltern könnten Unterschiede im Wissen über Möglichkeiten der Förderung, Sprachbarrieren, bürokratische Hürden oder unterschiedliche Bewertungen der sonderpädagogischen Förderung zu ungleichen Verteilungen der Diagnose eines SPF beitragen. Insbesondere bei der Feststellung der Förderschwerpunkte „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ gilt die Abgrenzung zur normalen kindlichen Entwicklung als schwierig (vgl. Bos, Müller & Stubbe, 2010), weshalb systematische Verzerrungen nicht auszuschließen sind. Dementsprechend werden in der Literatur die Verlässlichkeit und die Objektivität der Diagnose verschiedentlich in Frage gestellt (z. B. Grosche, 2015; Kottmann, 2007; Mand, 2003).

Auch bei der Wahl der Schulart (Förderschule oder allgemeine Schule) könnten Hintergrundmerkmale von Schülerinnen und Schülern mit SPF über die erreichten Kompetenzen hinaus eine Rolle spielen. In den meisten Ländern in der Bundesrepublik Deutschland trifft die Schulaufsichtsbehörde die Entscheidung über die Schulart von Kindern und Jugendlichen mit SPF, wobei die Eltern je nach Land in unterschiedlichem Maße Mitsprachemöglichkeiten haben (vgl. Sälzer et al., 2015). Somit sind hier ebenfalls verschiedene Akteure beteiligt, die die Entscheidung für eine Förderschule oder eine allgemeine Schule systematisch beeinflussen könnten.

Ungleiche Verteilungen der Diagnose eines SPF und des Besuchs einer Förderschule, die mit den betrachteten Hintergrundmerkmalen zusammenhängen, können als weiteres Indiz für die Herkunftsabhängigkeit von Schulerfolg gesehen werden. Allerdings ist die Interpretation aufgrund der verschiedenen, teilweise gegenläufigen Wirkungszusammenhänge komplex und sollte mit Vorsicht erfolgen. Beispielsweise muss die Diagnose eines SPF für Schülerinnen und Schüler nicht zwangsläufig als Nachteil gesehen werden, sondern diese kann gezielt von Eltern oder Lehrkräften angestrebt werden, um einem Kind zusätzliche Förderung zukommen zu lassen. Dieses Kapitel beschreibt die Verteilungen rein deskriptiv, das heißt, die berichteten Ergebnisse lassen keine Schlussfolgerungen über die Ursachen beobachteter Zusammenhänge zu.

Im vorliegenden Kapitel wird anhand von Daten des IQB-Bildungstrends 2016 bestimmt, inwieweit die Diagnose eines SPF und die besuchte Schulart von Kindern mit SPF mit dem Geschlecht, der sozialen Herkunft und dem Zuwanderungshintergrund assoziiert sind. Dabei werden zwei Perspektiven eingenommen: Zum einen wird jeweils für die betrachteten Gruppen (Mädchen und Jungen, Kinder mit hohem und niedrigem SES, Kinder mit und ohne Zuwanderungshintergrund) bestimmt, welcher Anteil von ihnen die Diagnose eines SPF erhalten hat. Die Gruppenunterschiede werden dabei sowohl für den SPF insgesamt als auch für die verschiedenen Förderschwerpunkte untersucht. Zum anderen wird geprüft, inwieweit sich Schülerinnen und Schüler mit und ohne SPF in den untersuchten Hintergrundmerkmalen unterscheiden. Innerhalb der Gruppe mit SPF werden außerdem Schülerinnen und Schüler mit SPF an Förderschulen und an allgemeinen Schulen getrennt betrachtet.

Untersucht werden folgende Fragen:

- 1) Inwieweit unterscheidet sich die Häufigkeit, mit der bei Kindern ein SPF in den Förderschwerpunkten „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ festgestellt wird nach (a) Geschlecht, (b) sozialer Herkunft und (c) Zuwanderungshintergrund?
- 2) Inwieweit unterscheiden sich Kinder mit und ohne SPF einerseits und Kinder mit SPF an allgemeinen Schulen und Förderschulen andererseits in

der Verteilung (a) des Geschlechts, (b) der sozialen Herkunft und (c) des Zuwanderungshintergrunds?

In Abschnitt 10.3.2 folgen zunächst einige methodische Vorbemerkungen, bevor die beiden Fragestellungen für die Hintergrundmerkmale Geschlecht (Abschnitt 10.3.3), soziale Herkunft (Abschnitt 10.3.4) und Zuwanderungshintergrund (Abschnitt 10.3.5) untersucht werden. Abschließend werden die Ergebnisse zusammengefasst und in den bisherigen Forschungsstand eingeordnet (Abschnitt 10.3.6).

10.3.2 Methodische Vorbemerkungen

Stichprobe und Erfassung der Hintergrundmerkmale

In die Analysen dieses Kapitels werden Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf in den drei Förderschwerpunkten „Lernen“ (L), „Sprache“ (S) sowie „emotionale und soziale Entwicklung“ (E) einbezogen. Neben den genannten drei Förderschwerpunkten wird außerdem die Gruppe der Kinder mit L, S oder E und Förderbedarf in weiteren Förderschwerpunkten in den Analysen separat betrachtet.² Schülerinnen und Schüler, für die keine Angaben zum Förderschwerpunkt vorliegen oder die ausschließlich Förderschwerpunkte außerhalb von L, S, oder E aufweisen, werden aus den Analysen ausgeschlossen.

Die Auswahl und die Erfassung der Hintergrundmerkmale orientieren sich an den Kapiteln 7 bis 9 dieses Berichtsbandes. Das Geschlecht der Kinder wurde sowohl von den Lehrkräften als auch von den Schülerinnen und Schülern selbst erfragt. Als Indikator für die soziale Herkunft wird die EGP-Klassifikation (Erikson & Goldthorpe, 2002; Erikson, Goldthorpe & Portocarero, 1979) genutzt, die auf den Angaben der Eltern zu ihrem ausgeübten Beruf, ihrer beruflichen Ausbildung sowie weiterer Merkmale der beruflichen Tätigkeit basiert und in Kapitel 8 beschrieben wird. Zur Kontrastierung werden anhand der sieben EGP-Klassen zwei Extremgruppen gebildet: Schülerinnen und Schüler aus Familien mit einem hohen Sozialstatus bilden die erste Extremgruppe (EGP-Klassen I–II) und Schülerinnen und Schüler aus Familien mit einem niedrigen Sozialstatus die zweite Extremgruppe (EGP-Klassen V–VII). Die zwei übrigen EGP-Klassen werden zu einer Gruppe mit mittlerem Sozialstatus zusammengefasst. Die Erfassung des Zuwanderungshintergrunds erfolgt in Anlehnung an Kapitel 9 anhand des Geburtslands des Kindes und der Eltern. Kinder aus zugewanderten Familien werden danach differenziert, ob ein Elternteil oder beide Elternteile im Ausland geboren sind. Schülerinnen und Schüler, deren Zuwanderungshintergrund nicht bestimmt werden kann, werden aus den zugewanderungsbezogenen Analysen ausgeschlossen.

Bei der Interpretation der Ergebnisse ist zu beachten, dass die in diesem Kapitel durchgeführten Analysen trotz der insgesamt großen Stichprobe des IQB-Bildungstrends 2016 (vgl. Kapitel 3.1) zum Teil auf geringen Fallzahlen beruhen und dadurch mit einer höheren Unsicherheit behaftet sind (vgl. Kapitel 3.2). Dies ist darauf zurückzuführen, dass einige Kombinationen aus Hintergrundmerkmalen und Förderschwerpunkten in der Population selten auftreten.

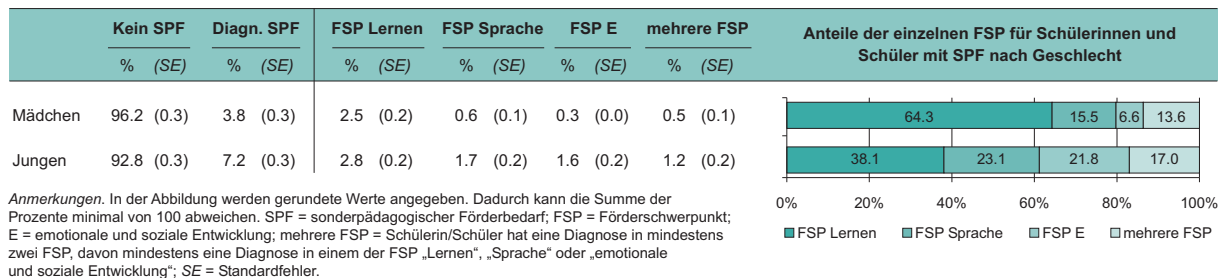
² Sie umfasst Kinder mit mindestens einem der Schwerpunkte L, S oder E sowie gegebenenfalls einem oder mehreren weiteren sonderpädagogischen Förderschwerpunkten.

10.3.3 Geschlecht

In Abbildung 10.2 sind auf der linken Seite die prozentualen Anteile der Mädchen und der Jungen mit und ohne SPF dargestellt. In der Mitte werden die Anteile der Kinder mit SPF nach Förderschwerpunkten differenziert. Das Balkendiagramm auf der rechten Seite weist ergänzend die Verteilung der verschiedenen Förderschwerpunkte bei allen Mädchen beziehungsweise Jungen mit SPF aus. Insgesamt liegt bei 3.8 Prozent aller Mädchen und bei 7.2 Prozent aller Jungen ein diagnostizierter SPF in mindestens einem der Schwerpunkte L, S oder E vor.³ Jungen haben somit fast doppelt so häufig einen festgestellten SPF wie Mädchen; der Unterschied ist statistisch signifikant. In beiden Geschlechtergruppen kommt der Förderschwerpunkt „Lernen“ am häufigsten vor (2.5 % aller Mädchen und 2.8 % aller Jungen).

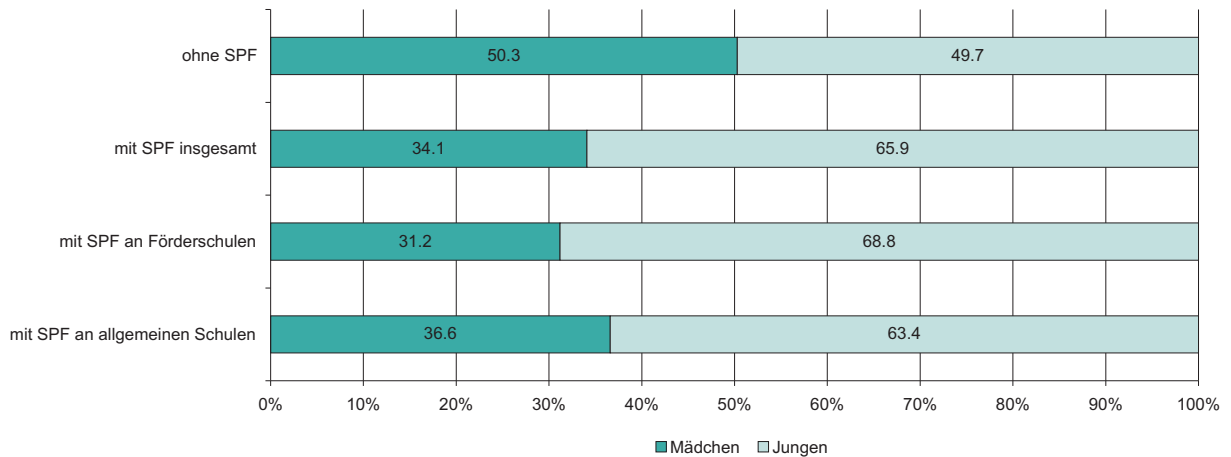
Bei der Verteilung der Förderschwerpunkte innerhalb der beiden Gruppen (rechte Seite der Abbildung 10.2) finden sich weitere Unterschiede: In der Gruppe der Jungen mit SPF tritt vergleichsweise häufiger der Förderschwerpunkt „emotionale und soziale Entwicklung“ auf als in der Gruppe der Mädchen mit SPF (Jungen: 21.8 %, Mädchen: 6.6 %), der Schwerpunkt „Lernen“ kommt dagegen bei Mädchen mit SPF anteilig deutlich häufiger vor (64.3 %) als bei Jungen (38.1 %). Die Verteilung der Förderschwerpunkte bei Kindern mit SPF unterscheidet sich signifikant zwischen Mädchen und Jungen.

Abbildung 10.2: Prozentuale Anteile von Kindern ohne und mit SPF in den Förderschwerpunkten „Lernen“, „Sprache“, „emotionale und soziale Entwicklung“ und mit mehreren Förderschwerpunkten nach Geschlecht



In einem zweiten Schritt (Abb. 10.3) werden die Geschlechteranteile in der Schülerschaft ohne SPF und in der Schülerschaft mit SPF betrachtet. Die Ergebnisse zeigen, dass bei den Kindern mit SPF der Anteil der Jungen mit 65.9 Prozent signifikant höher ist als bei Kindern ohne SPF (49.7 %). Werden darüber hinaus die Kinder mit SPF an allgemeinen Schulen und Förderschulen verglichen, finden sich Unterschiede in der Geschlechterverteilung nach der besuchten Schulart: Der Jungenanteil an Förderschulen fällt mit 68.8 Prozent signifikant höher aus als der Jungenanteil in der Gruppe der Kinder mit SPF an allgemeinen Schulen (63.4 %).

3 Der Anteil fällt etwas niedriger aus als in der amtlichen Statistik, da Schülerinnen und Schüler, die ausschließlich einen oder mehrere Förderschwerpunkte außerhalb von L, S oder E aufweisen, aus den hier berichteten Analysen ausgeschlossen wurden.

Abbildung 10.3: Prozentuale Anteile der Mädchen und Jungen an Kindern ohne SPF, mit SPF insgesamt, mit SPF an Förderschulen und mit SPF an allgemeinen Schulen

Anmerkungen. In der Abbildung werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. SPF = sonderpädagogischer Förderbedarf.

Insgesamt weisen Jungen also nicht nur häufiger die Diagnose eines SPF auf als Mädchen, sondern sie besuchen zudem auch häufiger eine Förderschule. Dass der Jungenanteil an Förderschulen vergleichsweise hoch ausfällt, ist dabei nicht ausschließlich darauf zurückzuführen, dass bei Jungen häufiger ein SPF diagnostiziert wird: Jungen mit SPF werden auch seltener an einer allgemeinen Schule unterrichtet als Mädchen mit SPF. Dies zeigt sich auch daran, dass der Jungenanteil an Förderschulen höher ist als der Anteil der Jungen an Kindern mit SPF an allgemeinen Schulen.

10.3.4 Soziale Herkunft

Analog zu den Analysen zu Geschlechterunterschieden ist in Abbildung 10.4 dargestellt, wie hoch der Anteil der Kinder mit und ohne SPF in Familien mit hohem, mittlerem und niedrigem sozioökonomischen Status ist. Zusätzlich zeigt die Abbildung die Verteilung der einzelnen Förderschwerpunkte bei allen Kindern mit einer SPF-Diagnose nach sozioökonomischem Status.

Insgesamt weisen Kinder aus Familien mit einem niedrigen SES (EGP-Klassen V–VII) mit 10.6 Prozent mehr als dreimal so häufig die Diagnose eines SPF in mindestens einem der Schwerpunkte L, S oder E auf wie Kinder aus Familien mit einem hohen SES (EGP-Klassen I–II; 3.2 %). Dieser Unterschied ist signifikant. Zudem unterscheiden sich die beiden Extremgruppen signifikant in der Verteilung der Förderschwerpunkte: Während ungefähr die Hälfte der Kinder mit SPF aus Familien mit niedrigem SES (51.8 %) den Förderschwerpunkt „Lernen“ aufweist, liegt der entsprechende Anteil bei Kindern aus Familien mit hohem SES nur bei 39.9 Prozent. Die anderen Förderschwerpunkte (S, E und mehrere Förderschwerpunkte) kommen folglich bei Kindern aus Familien mit hohem SES anteilig etwas häufiger vor als bei Kindern aus Familien mit niedrigem SES.

Abbildung 10.4: Prozentuale Anteile von Kindern ohne und mit SPF in den Förderschwerpunkten „Lernen“, „Sprache“, „emotionale und soziale Entwicklung“ und mit mehreren Förderschwerpunkten nach sozioökonomischem Status

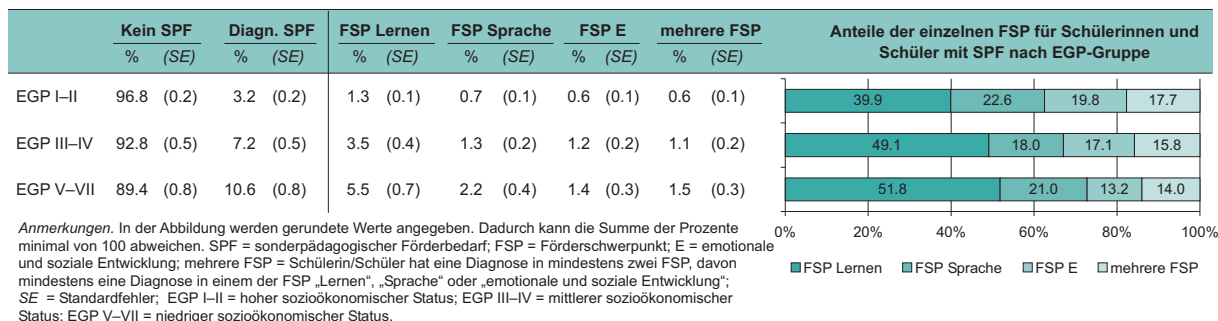
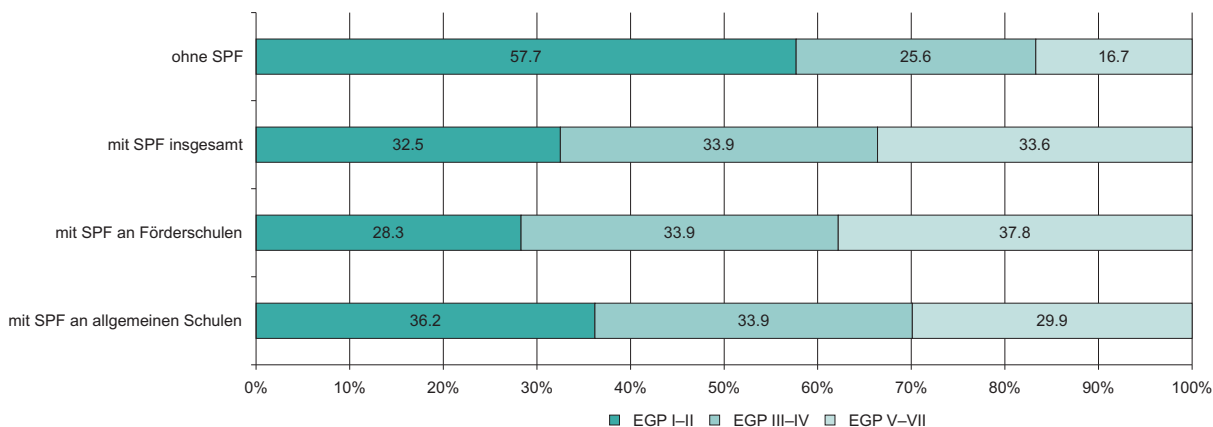


Abbildung 10.5 zeigt die Verteilung von Kindern mit niedrigem, mittlerem und hohem SES innerhalb verschiedener Gruppen von Schülerinnen und Schülern. Dabei ist ersichtlich, dass sich die Verteilung der sozialen Herkunft zwischen Kindern mit und ohne SPF signifikant unterscheidet: Während mit 57.7 Prozent über die Hälfte aller Schülerinnen und Schüler ohne SPF aus Familien mit einem hohen SES stammen, sind bei Schülerinnen und Schülern mit SPF alle drei EGP-Gruppen nahezu gleich stark vertreten. Der Vergleich zwischen Kindern mit SPF an Förderschulen und allgemeinen Schulen zeigt zudem, dass der Anteil von Schülerinnen und Schülern aus Familien mit einem niedrigen SES an Förderschulen mit 37.8 Prozent signifikant höher ausfällt als an allgemeinen Schulen mit 29.9 Prozent. Umgekehrt verhält es sich bei Schülerinnen und Schülern mit SPF aus Familien mit hohem SES, die an Förderschulen einen Anteil von 28.3 Prozent und an allgemeinen Schulen einen Anteil von 36.2 Prozent ausmachen.

Abbildung 10.5: Prozentuale Anteile der Kinder aus Familien mit niedrigem, mittlerem und hohem sozioökonomischem Status an Kindern ohne SPF, mit SPF insgesamt, mit SPF an Förderschulen und mit SPF an allgemeinen Schulen



Anmerkungen. In der Abbildung werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. SPF = sonderpädagogischer Förderbedarf; EGP I-II = hoher sozioökonomischer Status; EGP III-IV = mittlerer sozioökonomischer Status; EGP V-VII = niedriger sozioökonomischer Status.

Insgesamt weisen Schülerinnen und Schüler aus Familien mit niedrigem SES also häufiger die Diagnose eines SPF auf und besuchen häufiger eine Förderschule als Kinder aus Familien mit hohem SES. Auch hier erklärt die häufigere Diagnose bei Kindern aus Familien mit niedrigem SES nicht allein ihren höheren Anteil an Förderschulen, sondern Kinder mit SPF werden auch seltener an einer allgemei-

nen Schule unterrichtet, wenn sie aus Familien mit einem niedrigen SES stammen. Dies ist wiederum daran erkennbar, dass der Anteil der Kinder mit niedrigem SES an Förderschulen höher ist als ihr Anteil an Kindern mit SPF an allgemeinen Schulen.

10.3.5 Zuwanderungshintergrund

Abbildung 10.6 stellt jeweils die Unterschiede in den Anteilen der Kinder mit und ohne SPF nach Zuwanderungshintergrund dar. Auch hier werden die jeweiligen Anteile zudem nach Förderschwerpunkten differenziert ausgewiesen.

Bei 4.9 Prozent aller Schülerinnen und Schüler ohne Zuwanderungshintergrund liegt in mindestens einem der Förderschwerpunkte L, S oder E die Diagnose eines SPF vor. Bei Kindern mit einem im Ausland geborenen Elternteil liegt der entsprechende Anteil bei 5.6 Prozent und bei Kindern mit zwei im Ausland geborenen Eltern bei 5.3 Prozent. Diese Unterschiede sind allerdings statistisch nicht signifikant. Die Daten des IQB-Bildungstrends 2016 geben in dieser bundesweiten Analyse also keine Hinweise darauf, dass zwischen dem Zuwanderungshintergrund und der Diagnose eines SPF ein systematischer Zusammenhang besteht.

Betrachtet man die Verteilung auf die verschiedenen Förderschwerpunkte, so zeigt sich jedoch, dass Kinder, deren Eltern beide im Ausland geboren sind, häufiger eine Diagnose im Förderschwerpunkt „Sprache“ (29.6 %) aufweisen als Kinder mit einem im Ausland geborenen Elternteil (21.1 %) und als Kinder ohne Zuwanderungshintergrund (19.3 %). Eine Diagnose in den Förderschwerpunkten „Lernen“ (40.8 %) und „emotionale und soziale Entwicklung“ (13.7 %) kommt bei Kindern mit zwei im Ausland geborenen Elternteilen dagegen etwas seltener vor als bei Kindern mit einem im Ausland geborenen Elternteil (L: 48.8 % bzw. E: 19.5 %) und bei Kindern ohne Zuwanderungshintergrund (L: 47.8 % bzw. E: 17.9 %). Die Verteilung der Förderschwerpunkte bei Kindern mit zwei im Ausland geborenen Elternteilen unterscheidet sich signifikant von den Verteilungen bei Kindern ohne Zuwanderungshintergrund und bei Kindern mit einem im Ausland geborenen Elternteil.

Abbildung 10.6: Prozentuale Anteile von Kindern ohne und mit SPF in den Förderschwerpunkten „Lernen“, „Sprache“, „emotionale und soziale Entwicklung“ und mit mehreren Förderschwerpunkten nach Zuwanderungsstatus

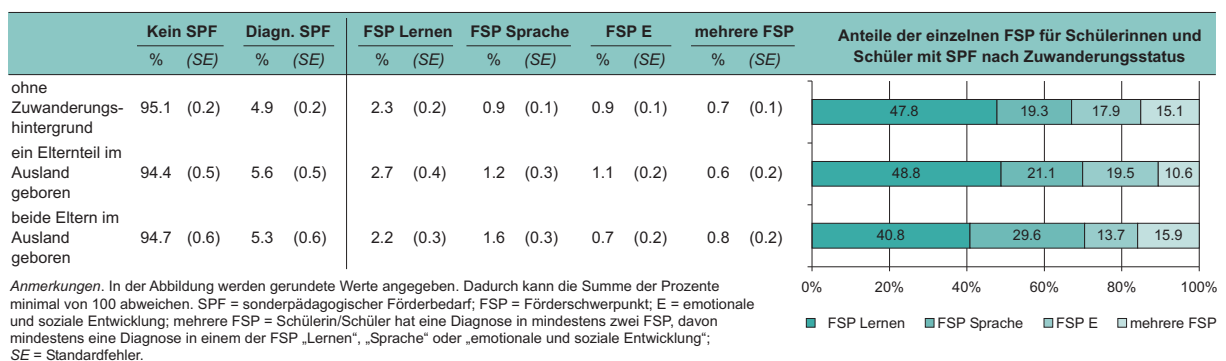
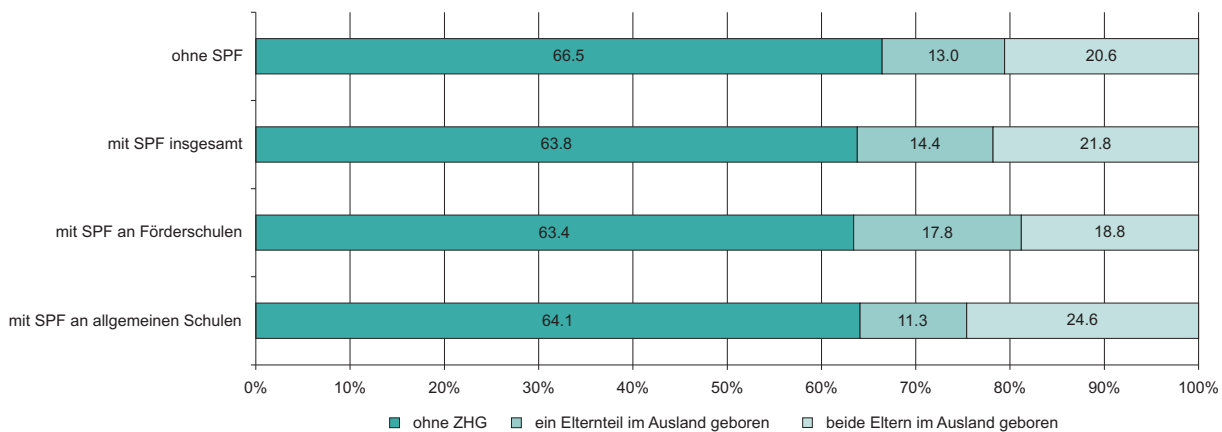


Abbildung 10.7: Prozentuale Anteile der Kinder aus Familien ohne Zuwanderungshintergrund, mit einem im Ausland geborenen Elternteil und mit zwei im Ausland geborenen Elternteilen an Kindern ohne SPF, mit SPF insgesamt, mit SPF an Förderschulen und mit SPF an allgemeinen Schulen



Anmerkungen. In der Abbildung werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. SPF = sonderpädagogischer Förderbedarf; ZHG = Zuwanderungshintergrund.

Abbildung 10.7 stellt die Anteile von Kindern mit unterschiedlichem Zuwanderungsstatus innerhalb verschiedener Gruppen dar. Ein Vergleich von Kindern mit und ohne Diagnose eines SPF ergibt, dass keine signifikanten Unterschiede in den Anteilen von Kindern ohne Zuwanderungshintergrund, Kindern mit einem im Ausland geborenen Elternteil und Kindern mit zwei im Ausland geborenen Elternteilen bestehen. Werden Kinder mit SPF an allgemeinen Schulen und Förderschulen verglichen, findet sich dagegen ein Unterschied in der Verteilung des Zuwanderungshintergrunds nach der besuchten Schulart: Der Anteil der Kinder mit einem im Ausland geborenen Elternteil an Förderschulen fällt mit 17.8 Prozent signifikant höher aus als der Anteil dieser Gruppe an Kindern mit SPF an allgemeinen Schulen (11.3 %).

Insgesamt lassen sich keine signifikanten Unterschiede zwischen Kindern mit und ohne Zuwanderungshintergrund in der Häufigkeit der Diagnose eines SPF nachweisen. Kinder aus Familien mit einem im Ausland geborenen Elternteil besuchen allerdings häufiger eine Förderschule als Kinder mit zwei im Ausland geborenen Elternteilen und als Kinder ohne Zuwanderungshintergrund.

10.3.6 Zusammenfassung und Diskussion

In diesem Kapitel wurde anhand der bundesweit repräsentativen Datengrundlage des IQB-Bildungstrends 2016 untersucht, inwiefern die Diagnose eines SPF und die besuchte Schulart von Kindern mit SPF mit den Hintergrundmerkmalen Geschlecht, soziale Herkunft und Zuwanderungshintergrund zusammenhängen. Es wurde zwei Fragen nachgegangen: Erstens wurde geprüft, inwieweit sich die Kinder abhängig von diesen drei Hintergrundmerkmalen in der Häufigkeit der SPF-Diagnose unterscheiden. Dabei wurde auch auf die Verteilung der Förderschwerpunkte eingegangen. Zweitens wurde untersucht, wie sich die jeweiligen Hintergrundmerkmale bei Kindern mit und ohne SPF sowie bei Kindern mit SPF an Förderschulen und allgemeinen Schulen verteilen.

Insbesondere bezogen auf das Geschlecht und die soziale Herkunft zeigen sich sowohl für das Vorliegen einer Diagnose als auch für den Besuch einer

Förderschule deutliche Unterschiede. Jungen haben häufiger einen festgestellten SPF als Mädchen und besuchen häufiger eine Förderschule. Diese Befunde stehen in Einklang mit den Ergebnissen regionaler Untersuchungen und Analysen von Daten der amtlichen Statistik (vgl. KMK, 2016; Lehmann & Hoffmann, 2009; Wild et al., 2015; Wocken & Gröhlich, 2009). Ferner kommt die Diagnose im Förderschwerpunkt emotionale und soziale Entwicklung bei Jungen häufiger vor als bei Mädchen.

Vielfältige Ursachen können für diese Befunde verantwortlich sein. Eine mögliche Erklärung ist, dass unterschiedliche Kriterien für die Diagnosestellung und Förderschulüberweisung angewendet werden. Da der Förderschwerpunkt „emotionale und soziale Entwicklung“ eher verhaltensbasiert ist und auf Beeinträchtigungen im Erleben und sozialen Handeln abzielt, könnte die unterschiedliche Verteilung ein Hinweis darauf sein, dass Mädchen im Durchschnitt häufiger als Jungen Verhaltensweisen zeigen, die im schulischen Kontext als angemessen und wünschenswert wahrgenommen werden. Dies wird auch als mögliche Ursache für Unterschiede in der Bildungsbeteiligung und im Bildungserfolg von Jungen und Mädchen diskutiert (vgl. z. B. Hannover & Kessels, 2011).

Auch für die soziale Herkunft finden sich deutliche Unterschiede. In Einklang mit den Ergebnissen früherer Studien haben Kinder aus Familien mit einem niedrigen SES deutlich häufiger einen festgestellten SPF in mindestens einem der Förderschwerpunkte L, S oder E und besuchen häufiger eine Förderschule als Kinder aus Familien mit einem hohen SES (z. B. Wocken, 2005). Ähnlich wie soziale Disparitäten in den erreichten Kompetenzen könnten diese Unterschiede darauf zurückgeführt werden, dass Schülerinnen und Schüler aus Familien mit hohem SES über einen besseren Zugang zu bildungsbezogenen Ressourcen und außerschulischen Lerngelegenheiten verfügen (z. B. Baumert, Watermann & Schümer, 2003). Insbesondere der Befund, dass Kinder mit SPF aus Familien mit hohem SES häufiger an allgemeinen Schulen beschult werden als Kinder mit SPF aus Familien mit niedrigem SES, könnte darauf hinweisen, dass sozial besser gestellte Eltern häufiger davon ausgehen, dass gemeinsamer Unterricht mit Vorteilen für ihre Kinder verbunden ist und sich stärker dafür einsetzen, dass ihr Kind an einer allgemeinen Schule beschult wird.

Hervorzuheben ist ein Befund, der auf Grundlage der vorliegenden Daten erstmals auf Bundesebene sichtbar ist: Dass Jungen und Kinder aus sozial schwachen Familien häufiger eine Förderschule besuchen, ist nicht alleine darauf zurückzuführen, dass sie häufiger die Diagnose eines SPF bekommen. Darüber hinaus werden diese Kinder, wenn sie eine Diagnose erhalten haben, auch seltener an einer allgemeinen Schule unterrichtet. Dies steht im Einklang mit den Ergebnissen regionaler Untersuchungen (vgl. Stranghöner et al., 2017).

Betrachtet man die Befunde zu Schülerinnen und Schülern mit Zuwanderungshintergrund, sind anders als in bisherigen Veröffentlichungen auf Grundlage von Daten der amtlichen Statistik (z. B. Diefenbach, 2007; Kemper & Weishaupt, 2011; Kornmann, 2013; Powell & Wagner, 2014) die Unterschiede zwischen Kindern mit und ohne Zuwanderungshintergrund weniger ausgeprägt. Dass im Gegensatz zu den Befunden der amtlichen Statistik keine bedeutsamen Unterschiede in der Häufigkeit der SPF-Diagnose bestehen, könnte in der jeweiligen Erfassung des Zuwanderungshintergrunds liegen: Während in der amtlichen Statistik die Staatsbürgerschaft der Kinder erfasst wird (deutsche Staatsbürgerschaft oder die eines anderen Landes), wurde in den vorliegenden Analysen das Geburtsland der Kinder und ihrer Eltern zugrunde gelegt. Nach der im IQB-Bildungstrend 2016 gewählten Art der Erfassung weisen deut-

lich mehr Schülerinnen und Schüler einen Zuwanderungshintergrund auf, die darüber hinaus im Mittel einen höheren Bildungserfolg erzielen, als die Gruppe der Kinder mit ausländischer Staatsbürgerschaft (vgl. Gresch & Kristen, 2011). Mit dem Zuwanderungshintergrund assoziierte Unterschiede bestehen zum einen darin, dass Kinder mit einem im Ausland geborenen Elternteil häufiger eine Förderschule besuchen als Kinder, deren Eltern beide im Ausland geboren sind, und als Kinder ohne Zuwanderungshintergrund. Zum anderen haben Kinder mit zwei im Ausland geborenen Elternteilen häufiger den Förderschwerpunkt „Sprache“ als Kinder ohne Zuwanderungshintergrund und als Kinder mit einem im Ausland geborenen Elternteil. Dieser Befund ist besonders hervorzuheben, da er anzeigt, dass bei Kindern, die Deutsch nicht als Muttersprache gelernt haben, anteilig häufiger ein sonderpädagogischer Förderbedarf in diesem Förderschwerpunkt diagnostiziert wird. Dies könnte als Hinweis für die in der Literatur geäußerte Vermutung interpretiert werden, dass die Abgrenzung zwischen einem sonderpädagogischen Förderbedarf und geringen Deutschkenntnissen bei der Diagnosestellung schwierig ist (vgl. Gomolla, 2013; Kronig, 2003).

In diesem Kapitel konnte ein grundlegender deskriptiver Überblick der Häufigkeitsverteilungen verschiedener Hintergrundmerkmale in Verbindung mit SPF und der Schulart dargestellt werden. Es liegt eine Reihe von Forschungsarbeiten vor, die zeigen, dass die hier einzeln betrachteten Hintergrundmerkmale zum Teil stark miteinander zusammenhängen (z. B. Diehl, Hunkler & Kristen, 2016). Vor diesem Hintergrund sollten die Verschränkungen dieser Faktoren Gegenstand weiterer Forschung sein, um Aussagen über gemeinsame und spezifische Effekte der betrachteten Merkmale auf die Diagnose von SPF und die Wahl der Schulart treffen zu können.

Literatur

- Baumert, J., Watermann, R. & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und psychologisches Mediationsmodell. *Zeitschrift für Erziehungswissenschaft*, 6(1), 46–72.
- Bos, W., Müller, S. & Stubbe T. C. (2010). Abgehängte Bildungsinstitutionen: Hauptschulen und Förderschulen. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 375–397). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Diefenbach, H. (2007). *Kinder und Jugendliche aus Migrantenfamilien im deutschen Bildungssystem. Erklärungen und empirische Befunde*. Wiesbaden: Springer VS.
- Diehl, C., Hunkler, C. & Kristen, C. (2016). Ethnische Ungleichheiten im Bildungsverlauf: Eine Einführung. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 3–33). Wiesbaden: Springer VS.
- 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.
- Gresch, C. & Kristen, C. (2011). Staatsbürgerschaft oder Migrationshintergrund? Ein Vergleich unterschiedlicher Operationalisierungsweisen am Beispiel der Bildungsbeteiligung. *Zeitschrift für Soziologie*, 40(3), 208–227.
- Gomolla, M. (2013). Fördern und Fordern allein genügt nicht! Mechanismen institutioneller Diskriminierung von Migrantenkindern und -jugendlichen im deutschen Schulsystem. In G. Auernheimer (Hrsg.), *Schieflagen im Bildungssystem. Die Benachteiligung der Migrantenkinder* (5. Aufl., S. 87–102). Wiesbaden: Springer VS.

- Grosche, M. (2015). Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen*, (S. 17–39). Wiesbaden: Springer VS.
- 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.
- Kemper, T. & Weishaupt, H. (2011). Zur Bildungsbeteiligung ausländischer Schüler an Förderschulen – unter besonderer Berücksichtigung der spezifischen Staatsangehörigkeit. *Zeitschrift für Heilpädagogik*, 62(10), 419–431.
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Sonderpädagogische Förderung in Schulen 2005 bis 2014*. Zugriff am 08.06.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_210_SoPae_2014.pdf
- Kottmann, B. (2007). Die Feststellung von sonderpädagogischem Förderbedarf: Benachteiligung der Benachteiligten. In I. Demmer-Dieckmann & A. Textor (Hrsg.), *Integrationsforschung und Bildungspolitik im Dialog* (S. 99–108). Bad Heilbrunn: Klinkhardt.
- Kormann, R. (2013). Die Überrepräsentation ausländischer Kinder und Jugendlicher in Sonderschulen mit dem Schwerpunkt Lernen. In G. Auernheimer (Hrsg.), *Schief lagen im Bildungssystem. Die Benachteiligung der Migrantenkinder* (5. Aufl., S. 71–85). Wiesbaden: Springer VS.
- Kronig, W. (2003). Das Konstrukt des leistungsschwachen Immigrantenkindes. *Zeitschrift für Erziehungswissenschaft*, 6(1), 126–141.
- Lanfranchi, A. (2007). Sonderklassenversetzung und integrative Förderung: Denken und handeln Lehrpersonen kulturell neutral? *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 2, 128–141.
- Lehmann, R. & Hoffmann, E. (2009). *BELLA: Berliner Erhebung arbeitsrelevanter Basiskompetenzen von Schülerinnen und Schülern mit Förderbedarf „Lernen“*. Münster: Waxmann.
- Mand, J. (2003). *Lern- und Verhaltensprobleme in der Schule*. Stuttgart: Kohlhammer.
- Powell, J. J. W. & Wagner, S. J. (2014). An der Schnittstelle Ethnie und Behinderung benachteiligt. Jugendliche mit Migrationshintergrund an deutschen Sonderschulen weiterhin überrepräsentiert. In G. Wansing & M. Westphal (Hrsg.), *Behinderung und Migration. Inklusion, Diversität, Intersektionalität* (S. 177–199). Wiesbaden: Springer VS.
- Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen: Forschungsstand, Herausforderungen und Perspektiven*, (S. 129–152). Wiesbaden: Springer VS.
- Stranghöner, D., Hollmann, J., Otterpohl, N., Wild, E., Lütje-Klose, B. & Schwinger, M. (2017). Inklusion vs. Exklusion: Schulsetting und Lese-Rechtschreibentwicklung von Kindern mit Förderschwerpunkt Lernen. *Zeitschrift für pädagogische Psychologie*, 31(2), 125–136.
- Wild, E., Schwinger, M., Lütje-Klose, B., Yotyodying, S., Gorges, J., Stranghöner, D., Neumann, P., Serke, B. & Kurnitzki, S. (2015). Schülerinnen und Schüler mit dem Förderschwerpunkt Lernen in inklusiven und exklusiven Förderarrangements: Erste Befunde des BiLieF-Projektes zu Leistung, sozialer Integration, Motivation und Wohlbefinden. *Unterrichtswissenschaft*, 43(1), 7–21.
- Wocken, H. (2005). *Andere Länder, andere Schüler? Vergleichende Untersuchungen von Förderschülern in den Bundesländern Brandenburg, Hamburg und Niedersachsen*. Forschungsbericht. Zugriff am 26.06.2017 unter <http://bidok.uibk.ac.at/download/wocken-forschungsbericht.pdf>
- Wocken, H. & Gröhlich, C. (2009). Kompetenzen von Schülerinnen und Schülern an Hamburger Förderschulen. In W. Bos, M. Bonsen & C. Gröhlich (Hrsg.), *KESS 7 – Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7* (S. 133–142). Münster: Waxmann.

10.4 Schulische Kompetenzen und schulische Motivation von Kindern mit sonderpädagogischem Förderbedarf an Förderschulen und an allgemeinen Schulen

Aleksander Kocaj, Poldi Kuhl, Nicole Haag, Pauline Kohrt und Petra Stanat

10.4.1 Einleitung

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) werden zunehmend in allgemeinen Schulen statt in Förderschulen unterrichtet (vgl. Kapitel 4; KMK, 2016). Dieser Ausbau des gemeinsamen Lernens ist eine der umfassendsten Schulreformen im deutschen Bildungssystem in den letzten 20 Jahren (Baumert, 2016). Dies wirft die Frage auf, inwieweit die Beschulung in Förderschulen oder allgemeinen Schulen mit der schulischen Entwicklung von Kindern und Jugendlichen mit SPF zusammenhängt. Kinder und Jugendliche können in Abhängigkeit von der besuchten Schulart unterschiedliche Lern- und Entwicklungschancen erhalten, die zu Unterschieden in ihrer schulischen Entwicklung beitragen (Baumert, Stanat & Watermann, 2006; Wocken, 2005).

In diesem Kapitel werden schulische Kompetenzen und Merkmale der schulischen Motivation von Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen miteinander verglichen, um Hinweise zu gewinnen, ob Schülerinnen und Schüler mit SPF von einer Beschulung in Förderschulen oder in allgemeinen Schulen profitieren. Dabei werden die Schulartunterschiede für die Förderschwerpunkte „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ separat bestimmt. Insbesondere für Kinder und Jugendliche mit diesen Förderschwerpunkten wird zunehmend gefordert, auf eine separate Beschulung in Förderschulen zu verzichten und sie stattdessen in allgemeinen Schulen zu unterrichten (Autorengruppe Bildungsberichterstattung, 2014; Gresch & Piezunka, 2015).

Das Kapitel ist wie folgt gegliedert: In Abschnitt 10.4.2 wird zunächst kurz zusammengefasst, welche Argumente für eine Beschulung in Förderschulen oder eine Beschulung in allgemeinen Schulen sprechen und der Forschungsstand dazu skizziert (für eine ausführliche Darstellung der theoretischen Grundlagen und der Befundlage siehe Kocaj, Kuhl, Jansen, Pant & Stanat, 2017 und Kocaj, Kuhl, Kroth, Pant & Stanat, 2014). Daran schließt sich eine Beschreibung des methodischen Vorgehens an, in der insbesondere auf methodische Herausforderungen bei der Analyse von Schulartunterschieden eingegangen wird (Abschnitt 10.4.3). Im Anschluss werden die Ergebnisse zu Unterschieden von Kindern mit SPF in Förderschulen und allgemeinen Schulen in den schulischen Kompetenzen (Abschnitt 10.4.4) und in Merkmalen der schulischen Motivation (Abschnitt 10.4.5) berichtet. Abschließend werden die Ergebnisse zusammengefasst und diskutiert (Abschnitt 10.4.6).

10.4.2 Zusammenhänge der besuchten Schulart mit den schulischen Kompetenzen und der schulischen Motivation von Kindern mit sonderpädagogischem Förderbedarf

Die Frage, in welcher Schulart Kinder und Jugendliche mit SPF stärker in ihrer schulischen Entwicklung profitieren, wird seit längerer Zeit kontrovers diskutiert. Dabei geht es vor allem um Fragen zum Zusammenhang zwischen der Schulart und den erreichten schulischen Kompetenzen sowie um die schulische Motivation von Schülerinnen und Schülern mit SPF.

Als Argument für eine separate Beschulung wird angeführt, dass Förderschulen besser als allgemeine Schulen auf die speziellen Bedürfnisse von Kindern und Jugendlichen mit SPF eingehen könnten. Lehrkräfte in Förderschulen erhielten eine sonderpädagogische Ausbildung, durch die sie die Leistungsentwicklung von Schülerinnen und Schülern mit SPF besser fördern könnten als Lehrkräfte in allgemeinen Schulen (Ahrbeck, 2014). Außerdem werde die Unterrichtsgestaltung in Förderschulen stärker auf die Lernvoraussetzungen von Kindern und Jugendlichen mit SPF ausgerichtet (Vaughn, Elbaum & Schumm, 1996; Wocken, 2005). Dadurch könnten Lehrkräfte in Förderschulen stärker auf individuelle Lernvoraussetzungen der Schülerinnen und Schüler eingehen (Wocken, 2005). Es wird jedoch kontrovers diskutiert, ob der Unterricht in Förderschulen besser geeignet ist, die Leistungsentwicklung von Kindern und Jugendlichen mit SPF zu fördern, als der Unterricht in allgemeinen Schulen (vgl. zusammenfassend Werning & Lütje-Klose, 2016).

Im Hinblick auf die Entwicklung der schulischen Motivation wird argumentiert, dass Förderschulen einen Schonraum für Kinder und Jugendliche mit SPF bereitstellen können, in dem sie vor schulischem Versagen, Leistungsdruck und sozialer Ausgrenzung geschützt sind (Tent, Witt, Zschoche-Lieberum & Bürger, 1991; Wocken, 2005). Außerdem verglichen sich Kinder und Jugendliche mit SPF in Förderschulen mit Mitschülerinnen und Mitschülern, die ähnliche Lernschwierigkeiten und Lernvoraussetzungen haben (Renick & Harter, 1989). Diese sozialen Vergleiche mit leistungsschwächeren Mitschülerinnen und Mitschülern könnten zu einer höheren schulischen Motivation führen (sogenannte Fischteicheffekte, Marsh, 1987; Tracey & Marsh, 2000). In allgemeinen Schulen hingegen würden sich Kinder und Jugendliche mit SPF mit leistungsstärkeren Mitschülerinnen und Mitschülern vergleichen und könnten dadurch eine geringere schulische Motivation aufweisen (Bear, Minke & Manning, 2002).

Fürsprecher einer gemeinsamen Beschulung verweisen darauf, dass Kinder und Jugendliche mit SPF in allgemeinen Schulen eine anregende Lernumgebung vorfinden, die sich positiv auf ihre schulischen Kompetenzen auswirkt (Wocken, 2005). Die Lernumgebung in allgemeinen Schulen ließe sich im Vergleich zu Förderschulen durch eine anspruchsvollere Unterrichtsgestaltung und höhere Erwartungshaltungen der Lehrkräfte beschreiben (Peetsma, Vergeer, Roeleveld & Karsten, 2001). Der gemeinsame Unterricht biete zudem die Möglichkeit, dass Kinder und Jugendliche mit SPF in allgemeinen Schulen die Lösungsstrategien leistungsstärkerer Mitschülerinnen und Mitschüler übernehmen und durch soziale Lernprozesse in ihrer Leistungsentwicklung profitieren könnten (Gersten et al., 2009; Justice, Logan, Lin & Kaderavek, 2014).

Außerdem weisen Befürworter einer gemeinsamen Beschulung darauf hin, dass Schülerinnen und Schüler mit SPF hinsichtlich ihrer schulischen Motivation in allgemeinen Schulen stärker als kompetente Lernende wahrgenommen und positiver bewertet würden als in Förderschulen (Elbaum, 2002;

Pfahl, 2012). Die höheren Leistungserwartungen in allgemeinen Schulen könnten von Kindern und Jugendlichen mit SPF übernommen werden und zur höheren schulischen Motivation beitragen (Elbaum, 2002). Im Gegensatz dazu könnten Kinder und Jugendliche in Förderschulen eher als leistungsschwach stigmatisiert werden (Tracey & Marsh, 2000). Die Zuschreibung geringer Fähigkeiten und Lernpotenziale könnte von Schülerinnen und Schülern wahrgenommen und internalisiert werden und zu geringeren Fähigkeitseinschätzungen und Erfolgserwartungen führen (Knigge, 2017; Pfahl, 2012).

Zusammenfassend lassen sich sowohl Argumente für eine Beschulung in Förderschulen als auch Argumente zugunsten von allgemeinen Schulen anführen. Der Forschungsstand zu schulischen Kompetenzen verweist mehrheitlich auf neutrale bis positive Effekte zugunsten einer Beschulung von Kindern und Jugendlichen mit SPF in allgemeinen Schulen (Ruijs & Peetsma, 2009; Stranghöner et al., 2017). Auf Grundlage des IQB-Ländervergleichs 2011 (Stanat, Pant, Böhme & Richter, 2012) zeigte sich, dass Kinder mit SPF in allgemeinen Schulen höhere Kompetenzen im *Lesen*, *Zuhören* und im Fach Mathematik aufwiesen als vergleichbare Kinder mit SPF in Förderschulen (Kocaj et al., 2014). Im Gegensatz dazu verweist die Mehrzahl der Studien auf eine höhere schulische Motivation zugunsten von Schülerinnen und Schülern in Förderschulen beziehungsweise in speziellen Förderklassen (Bear et al., 2002). Auch im IQB-Ländervergleich 2011 berichteten Kinder mit SPF in Förderschulen über eine höhere schulische Motivation als Kinder mit SPF in allgemeinen Schulen (Kocaj et al., 2017). Es gibt aber auch Studien, die keine eindeutigen Zusammenhänge zwischen der Schulart und der schulischen Motivation von Kindern mit SPF feststellten (Elbaum, 2002; Peetsma et al., 2001; Wild et al., 2015). Die vorliegenden Analysen knüpfen an Analysen zum IQB-Ländervergleich 2011 an und erweitern diese um Ergebnisse zum Kompetenzbereich *Orthografie* und um Ergebnisse für Schülerinnen und Schüler mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ (Kocaj et al., 2017; Kocaj et al., 2014).

10.4.3 Methodisches Vorgehen

Stichprobe

Die in diesem Kapitel berichteten Analysen basieren auf den Daten der Teilstichprobe von Schülerinnen und Schülern mit einem festgestellten SPF in den Bereichen „Lernen“, „Sprache“ oder „emotionale und soziale Entwicklung“ (vgl. Kapitel 10.2), die an der Kompetenztestung teilnahmen. Kinder mit mehreren Förderschwerpunkten wurden nicht in die Analysen einbezogen, um differenzierte Zusammenhänge für die einzelnen Förderschwerpunkte betrachten zu können. Insgesamt gingen Daten von 2456 Schülerinnen und Schülern mit SPF in die Analysen ein, von denen 1196 Kinder eine Förderschule und 1260 Kinder eine allgemeine Schule besuchten.¹ Tabelle 10.2 zeigt die Verteilung der Schülerinnen und Schüler auf Förderschulen und allgemeine Schulen differenziert nach Förderschwerpunkt und nach Teilnahme an den Kompetenztests in den Fächern Deutsch und Mathematik. Der Fokus des vorliegenden Kapitels liegt auf einem bundesweiten Vergleich von Kindern mit SPF in Förderschulen und all-

¹ Die Stichprobengrößen beziehen sich auf Kinder mit SPF, die Kompetenztests im Fach Deutsch bearbeiteten. Die leicht abweichenden Stichprobengrößen für das Fach Mathematik können Tabelle 10.2 entnommen werden.

Tabelle 10.2: Stichprobenumfänge nach Förderschwerpunkt und bearbeiteten Kompetenztests in den Fächern Deutsch und Mathematik

Schülergruppe	Fach	N insgesamt	N Förderschule	N allgemeine Schule
Schülerinnen und Schüler mit SPF (insgesamt)	Deutsch	2 456	1 196	1 260
	Mathematik	2 457	1 190	1 267
Schülerinnen und Schüler mit FSP „Lernen“	Deutsch	1 272	488	784
	Mathematik	1 279	486	793
Schülerinnen und Schüler mit FSP „Sprache“	Deutsch	773	584	189
	Mathematik	764	581	183
Schülerinnen und Schüler mit FSP „emotionale und soziale Entwicklung“	Deutsch	411	124	287
	Mathematik	414	123	291

Anmerkungen. N = Anzahl der Schülerinnen und Schüler; SPF = sonderpädagogischer Förderbedarf; FSP = Förderschwerpunkt.

gemeinen Schulen; länderspezifische Analysen wurden aufgrund der begrenzten Stichprobengrößen nicht vorgenommen (vgl. Kapitel 10.1).

Messinstrumente

Zur Erfassung der im IQB-Bildungstrend 2016 untersuchten Kompetenzen in den Fächern Deutsch und Mathematik kamen bei Schülerinnen und Schülern mit SPF in Förderschulen und in allgemeinen Schulen spezielle Testhefte zum Einsatz, die angepasste Aufgaben beinhalteten (vgl. Kapitel 13.2). Als Merkmale der schulischen Motivation wurden das akademische Selbstkonzept sowie das Interesse in den Fächern Deutsch und Mathematik erfasst. Beim fachspezifischen akademischen Selbstkonzept handelt es sich um das von den Kindern selbsteingeschätzte Fähigkeitsniveau, das anhand der Zustimmung zu jeweils vier Aussagen auf einer vierstufigen Skala erhoben wurde (z. B. „Ich bin gut in Deutsch/Mathematik.“; 1 = „stimmt gar nicht“ bis 4 = „stimmt genau“), wobei höhere Werte für höhere Ausprägungen im akademischen Selbstkonzept stehen. Die Erfassung des fachspezifischen Interesses erfolgte ebenfalls über die Zustimmung der Kinder zu jeweils vier Aussagen auf einer vierstufigen Skala, bei der höhere Werte ein höheres Interesse indizieren (z. B. „Deutsch/Mathematik macht mir Spaß.“; 1 = „stimmt gar nicht“ bis 4 = „stimmt genau“).

Methodische Herausforderungen und Vorgehen bei der Schätzung von Schularteffekten

Eine grundlegende methodische Herausforderung bei der Analyse von Unterschieden zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen sind Selektionseffekte (Duncan, Magnuson & Ludwig, 2004). Darunter wird verstanden, dass der Besuch einer Schulart nicht zufällig erfolgt, sondern dass Schülerinnen und Schüler mit SPF unter anderem auf Grundlage ihrer schulischen Kompetenzen beziehungsweise ihrer Leistungsfähigkeit einer Förderschule oder allgemeinen Schule zugewiesen werden. Liegt nur ein gerin-

ger Förderbedarf vor, steigt die Wahrscheinlichkeit, dass ein Kind in einer allgemeinen Schule und nicht in einer Förderschule unterrichtet wird (Gresch & Piezunka, 2015; Sälzer, Gebhardt, Müller & Pauly, 2015). Darüber hinaus unterscheiden sich Schülerinnen und Schüler mit SPF in Förderschulen und allgemeinen Schulen auch in weiteren Hintergrundmerkmalen wie zum Beispiel im Zuwanderungshintergrund und im sozialen Status (vgl. Kapitel 10.3).

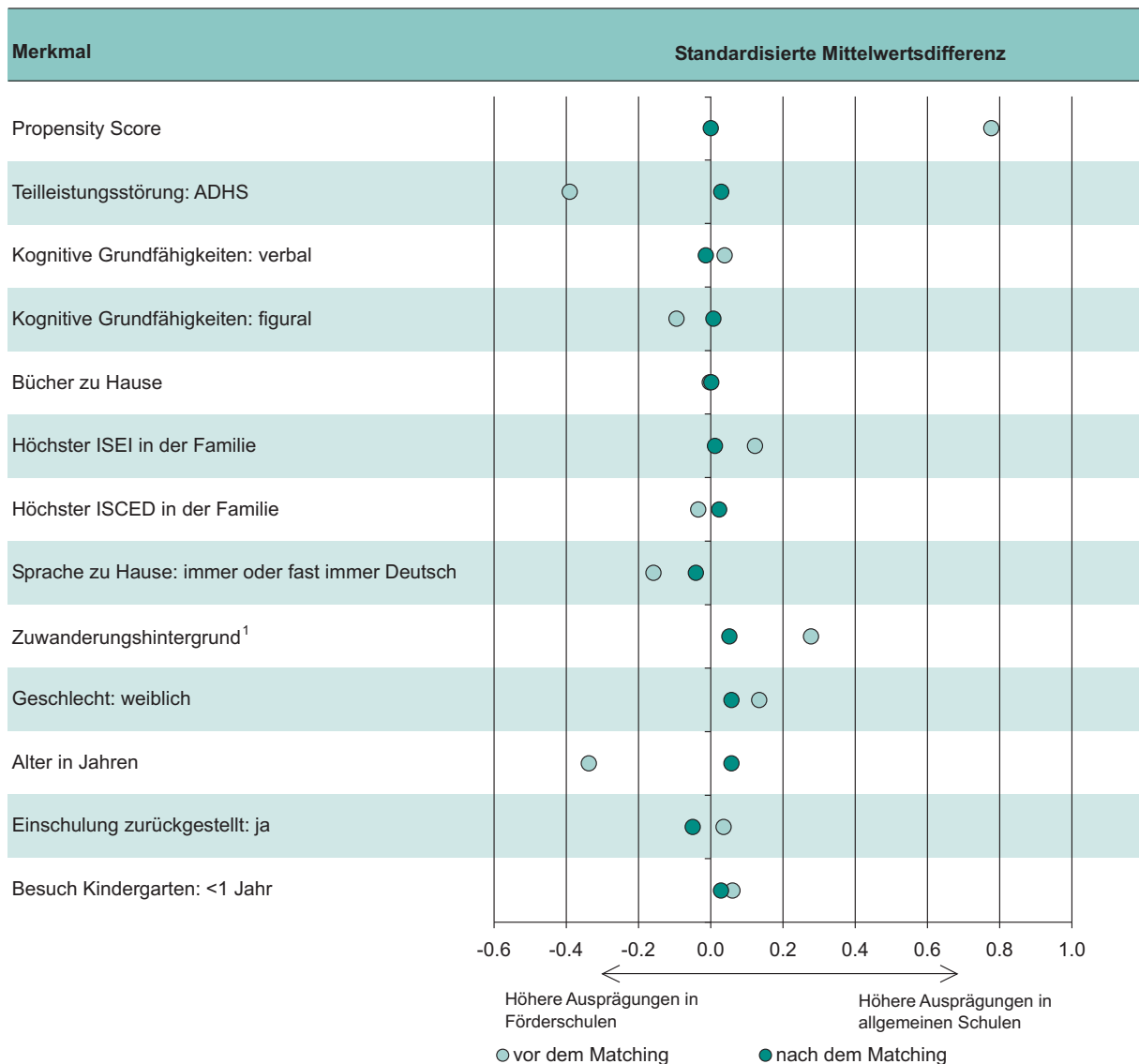
Vielfältige Faktoren können somit dazu führen, dass sich Schülerinnen und Schüler mit SPF in Förderschulen und in allgemeinen Schulen in Merkmalen unterscheiden, die mit den schulischen Kompetenzen und der schulischen Motivation zusammenhängen. Bleiben diese Unterschiede in den Lernvoraussetzungen und Hintergrundmerkmalen unberücksichtigt, kann dies zu einer Überschätzung von Schularteffekten führen (Baumert et al., 2006). In diesem Fall ließen sich potenzielle Unterschiede zwischen Schülerinnen und Schülern mit SPF in Förderschulen und in allgemeinen Schulen nicht eindeutig auf die Schulart zurückführen. Unterschiede könnten vielmehr auch durch unterschiedliche Lernvoraussetzungen verursacht worden sein, die bereits vor der Zuweisung zu Förderschulen und allgemeinen Schulen bestanden.

In den vorliegenden Analysen werden daher potenzielle Unterschiede in den Lernvoraussetzungen von Kindern mit SPF in Förderschulen und allgemeinen Schulen durch *Propensity-Score-Matching-Verfahren* (Rosenbaum & Rubin, 1983) statistisch kontrolliert. Bei diesem methodischen Verfahren werden anhand von lern- und leistungsrelevanten Merkmalen der Schülerinnen und Schüler (z. B. kognitive Grundfähigkeiten, sozialer Status der Familien) „statistische Zwillinge“ gebildet. Jedem Kind mit SPF aus einer allgemeinen Schule wird ein in diesen Merkmalen sehr ähnliches Kind aus einer Förderschule zugeordnet. Anschließend werden nur diese „statistischen Zwillinge“ in ihren schulischen Kompetenzen und ihrer schulischen Motivation miteinander verglichen. In Einklang mit vorangegangenen Analysen zum IQB-Ländervergleich 2011 wird ein *Full-Matching-Verfahren* (Stuart, 2010) eingesetzt, wobei soweit wie möglich dieselben Hintergrundmerkmale wie im IQB-Ländervergleich 2011 berücksichtigt werden (Kocaj et al., 2014).

Die Ergebnisse des Propensity-Score-Matchings zeigen, dass sich die Kinder mit SPF in Förderschulen und allgemeinen Schulen vor dem Matching in vielen Merkmalen unterscheiden (Abb. 10.8).² Die hellen Punkte in Abbildung 10.8 kennzeichnen die mittleren Unterschiede zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen auf den Hintergrundmerkmalen und dem *Propensity Score* (Zusammenfassung aller berücksichtigten Hintergrundmerkmale zu einem Kennwert) vor dem Matching. Die dunklen Punkte zeigen die Unterschiede zwischen den Schülergruppen nach dem Matching und weisen darauf hin, dass diese zu vernachlässigen sind. Somit konnten durch das Matching unterschiedliche Lernvoraussetzungen und Hintergrundmerkmale der Kinder mit SPF in Förderschulen und in allgemeinen Schulen statistisch kontrolliert werden, wodurch sich Unterschiede in den schulischen Kompetenzen und der schulischen Motivation zwischen diesen Gruppen mit höherer Sicherheit auf die Schulart zurückführen lassen. Im Folgenden werden zunächst Unterschiede zwischen Schülerinnen und Schülern mit SPF *vor* dem Matching berichtet. Im Anschluss werden die Schulartvergleiche *nach* dem Matching berichtet, bei dem unter-

2 Zusätzlich zu den in Abbildung 10.8 dargestellten Merkmalen wurde aufgrund der Unterschiede zwischen den Ländern im Anteil an Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen (vgl. Kapitel 4) auch das Land beim Propensity-Score-Matching berücksichtigt.

Abbildung 10.8: Unterschiede zwischen Kindern mit SPF in Förderschulen und in allgemeinen Schulen vor und nach dem Propensity-Score-Matching



Anmerkungen. Die Abbildung bezieht sich auf die gesamte Stichprobe der Kinder mit sonderpädagogischem Förderbedarf, für die gültige Kompetenzwerte im Fach Deutsch vorlagen. ADHS = Aufmerksamkeitsdefizit-/Hyperaktivitätsstörung; ISEI = *International Socio-economic Index of Occupational Status* (sozioökonomischer Index); ISCED = *International Standard Classification of Education* (Bildungsabschluss der Eltern).
¹Zuwanderungshintergrund (dichotom): 1 = mindestens ein Elternteil im Ausland geboren.

schiedliche Lernvoraussetzungen und Hintergrundmerkmale der Kinder mit SPF statistisch kontrolliert werden.

10.4.4 Vergleich der schulischen Kompetenzen zwischen Kindern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen

Tabelle 10.3 zeigt die mittleren Kompetenzwerte der Kinder mit SPF in Förderschulen und allgemeinen Schulen differenziert nach Förderschwerpunkt und Kompetenzbereich vor dem Propensity-Score-Matching. Kinder mit dem Förderschwerpunkt „Lernen“ weisen in allgemeinen Schulen in allen getesteten Kompetenzbereichen – mit Ausnahme des Kompetenzbereichs *Orthografie*

Tabelle 10.3: Unterschiede in den schulischen Kompetenzen zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen vor dem Propensity-Score-Matching

Schülergruppe	Kompetenzbereich	Förderschule			allgemeine Schule			Differenz		
		M	(SE)	SD	M	(SE)	SD	ΔM	(SE)	d
Schülerinnen und Schüler mit FSP „Lernen“	Lesen	282	(7.9)	93	347	(6.3)	90	65	(11.1)	0.71
	Zuhören	273	(8.7)	107	353	(9.0)	108	80	(14.0)	0.74
	Orthografie	323	(10.6)	106	347	(8.7)	100	23	(14.7)	0.23
	Mathematik (Globalskala)	237	(10.2)	101	303	(7.3)	94	66	(12.5)	0.68
Schülerinnen und Schüler mit FSP „Sprache“	Lesen	375	(6.0)	79	400	(13.8)	91	26	(15.4)	0.31
	Zuhören	359	(8.3)	92	380	(17.8)	101	21	(20.0)	0.22
	Orthografie	395	(7.8)	83	403	(15.8)	88	8	(17.5)	0.09
	Mathematik (Globalskala)	386	(7.9)	87	425	(10.4)	87	39	(12.7)	0.45
Schülerinnen und Schüler mit FSP „emotionale und soziale Entwicklung“	Lesen	388	(9.0)	91	410	(11.1)	90	22	(14.3)	0.25
	Zuhören	385	(12.7)	102	427	(11.6)	90	42	(17.9)	0.45
	Orthografie	394	(9.6)	92	411	(12.4)	96	16	(15.4)	0.17
	Mathematik (Globalskala)	401	(7.1)	101	423	(14.4)	101	22	(15.7)	0.22

Anmerkungen. Positive Werte der Differenz stehen für höhere Ausprägungen in allgemeinen Schulen, negative Werte für höhere Ausprägungen in Förderschulen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. *M* = Mittelwert; *SE* = Standardfehler; *SD* = Standardabweichung; ΔM = Mittelwertsdifferenz; *d* = Effektstärke Cohens *d*; FSP = Förderschwerpunkt.

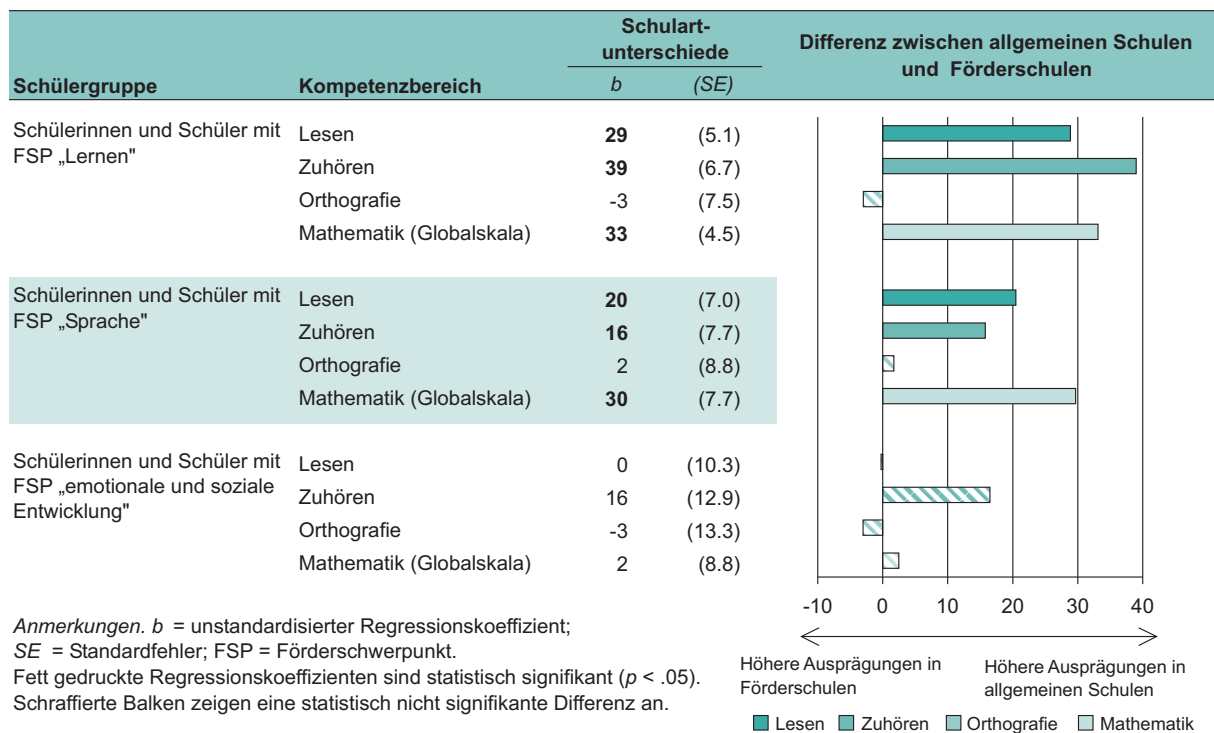
Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

– deutlich höhere schulische Kompetenzen auf als Kinder in Förderschulen. In den Förderschwerpunkten „Sprache“ und „emotionale und soziale Entwicklung“ sind die Schulartunterschiede in den erreichten Kompetenzen der Schülerinnen und Schüler deutlich geringer. Die Kompetenzvorteile zugunsten von allgemeinen Schulen sind für Kinder mit dem Förderschwerpunkt „Sprache“ nur im Fach Mathematik und für Kinder mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ nur im Kompetenzbereich *Zuhören* signifikant.

In Abbildung 10.9 sind die Mittelwertsunterschiede in den schulischen Kompetenzen zwischen Förderschulen und allgemeinen Schulen nach dem Propensity-Score-Matching abgebildet. Nach rechts weisende Balken zeigen höhere schulische Kompetenzen von Kindern mit SPF in allgemeinen Schulen an; nach links weisende Balken zeigen höhere schulische Kompetenzen von Kindern mit SPF in Förderschulen an. Unter Berücksichtigung individueller Unterschiede in den Lernvoraussetzungen und weiterer bedeutsamer Hintergrundmerkmale durch das Propensity-Score-Matching reduzieren sich die Kompetenzunterschiede zum Teil zwar deutlich, sind aber je nach Förderschwerpunkt und Kompetenzbereich weiterhin statistisch signifikant.

Kinder mit dem Förderschwerpunkt „Lernen“ in allgemeinen Schulen erreichen in den Kompetenzbereichen *Lesen* und *Zuhören* sowie im Fach Mathematik signifikant höhere Kompetenzwerte als vergleichbare Kinder in Förderschulen; im Kompetenzbereich *Orthografie* ist der Schulartunterschied hingegen nicht signifikant. Für Kinder mit dem Förderschwerpunkt „Sprache“ fallen die Kompetenzunterschiede zugunsten von allgemeinen Schulen etwas geringer aus als für Kinder mit dem Förderschwerpunkt „Lernen“. In den Kompetenzbereichen *Lesen* und *Zuhören* sowie im Fach Mathematik sind die Kompetenzunterschiede aber

Abbildung 10.9: Unterschiede in den schulischen Kompetenzen zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen *nach* dem Propensity-Score-Matching



ebenfalls statistisch signifikant. Bei Kindern mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ lassen sich nach dem Propensity-Score-Matching für keinen der getesteten Kompetenzbereiche statistisch signifikante Unterschiede in Abhängigkeit von der Schulart nachweisen.

10.4.5 Vergleich der schulischen Motivation zwischen Kindern mit sonderpädagogischem Förderbedarf in Förderschulen und allgemeinen Schulen

Tabelle 10.4 zeigt die Mittelwertsunterschiede der Kinder mit SPF in Förderschulen und in allgemeinen Schulen für die untersuchten Merkmale der schulischen Motivation vor dem Propensity-Score-Matching. Kinder mit dem Förderschwerpunkt „Lernen“ weisen demnach in Förderschulen ein signifikant höheres akademisches Selbstkonzept und ein stärkeres Interesse am Fach Mathematik auf als in allgemeinen Schulen. Im Gegensatz dazu sind für das Fach Deutsch keine signifikanten motivationalen Unterschiede zwischen Kindern mit dem Förderschwerpunkt „Lernen“ in Förderschulen und allgemeinen Schulen festzustellen. Für Kinder mit dem Förderschwerpunkt „Sprache“ ergeben sich insgesamt keine signifikanten Schulartunterschiede in den betrachteten motivationalen Merkmalen. Für Kinder mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ schließlich zeigen sich ähnliche Ergebnisse wie für Kinder mit dem Förderschwerpunkt „Lernen“: Kinder in Förderschulen weisen ein höheres akademisches Selbstkonzept und ein stärkeres Interesse am Fach Mathematik auf, im Fach Deutsch sind die Schulartunterschiede in den motivationalen Merkmalen hingegen nicht signifikant.

Tabelle 10.4: Unterschiede in Merkmalen der schulischen Motivation zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen vor dem Propensity-Score-Matching

Schülergruppe	Merkmale der schulischen Motivation	Förderschule			allgemeine Schule			Differenz		
		M	(SE)	SD	M	(SE)	SD	ΔM	(SE)	d
Schülerinnen und Schüler mit FSP „Lernen“	Selbstkonzept Deutsch	3.22	(0.07)	0.78	3.06	(0.06)	0.74	-0.17	(0.10)	-0.22
	Selbstkonzept Mathematik	3.19	(0.04)	0.87	2.75	(0.09)	0.93	-0.44	(0.10)	-0.48
	Interesse Deutsch	3.09	(0.09)	0.89	3.20	(0.05)	0.73	0.10	(0.11)	0.13
	Interesse Mathematik	3.29	(0.05)	0.87	3.06	(0.08)	0.92	-0.23	(0.10)	-0.25
Schülerinnen und Schüler mit FSP „Sprache“	Selbstkonzept Deutsch	3.01	(0.06)	0.68	2.88	(0.08)	0.64	-0.13	(0.10)	-0.19
	Selbstkonzept Mathematik	3.14	(0.05)	0.86	3.01	(0.15)	0.84	-0.13	(0.16)	-0.15
	Interesse Deutsch	2.97	(0.07)	0.76	2.86	(0.10)	0.81	-0.11	(0.13)	-0.14
	Interesse Mathematik	3.32	(0.04)	0.77	3.24	(0.15)	0.79	-0.08	(0.16)	-0.10
Schülerinnen und Schüler mit FSP „emotionale und soziale Entwicklung“	Selbstkonzept Deutsch	3.08	(0.08)	0.71	3.06	(0.11)	0.68	-0.03	(0.13)	-0.04
	Selbstkonzept Mathematik	3.29	(0.07)	0.85	2.93	(0.09)	0.92	-0.36	(0.11)	-0.40
	Interesse Deutsch	2.97	(0.10)	0.80	2.85	(0.10)	0.85	-0.12	(0.14)	-0.14
	Interesse Mathematik	3.35	(0.06)	0.92	2.99	(0.12)	1.04	-0.36	(0.13)	-0.36

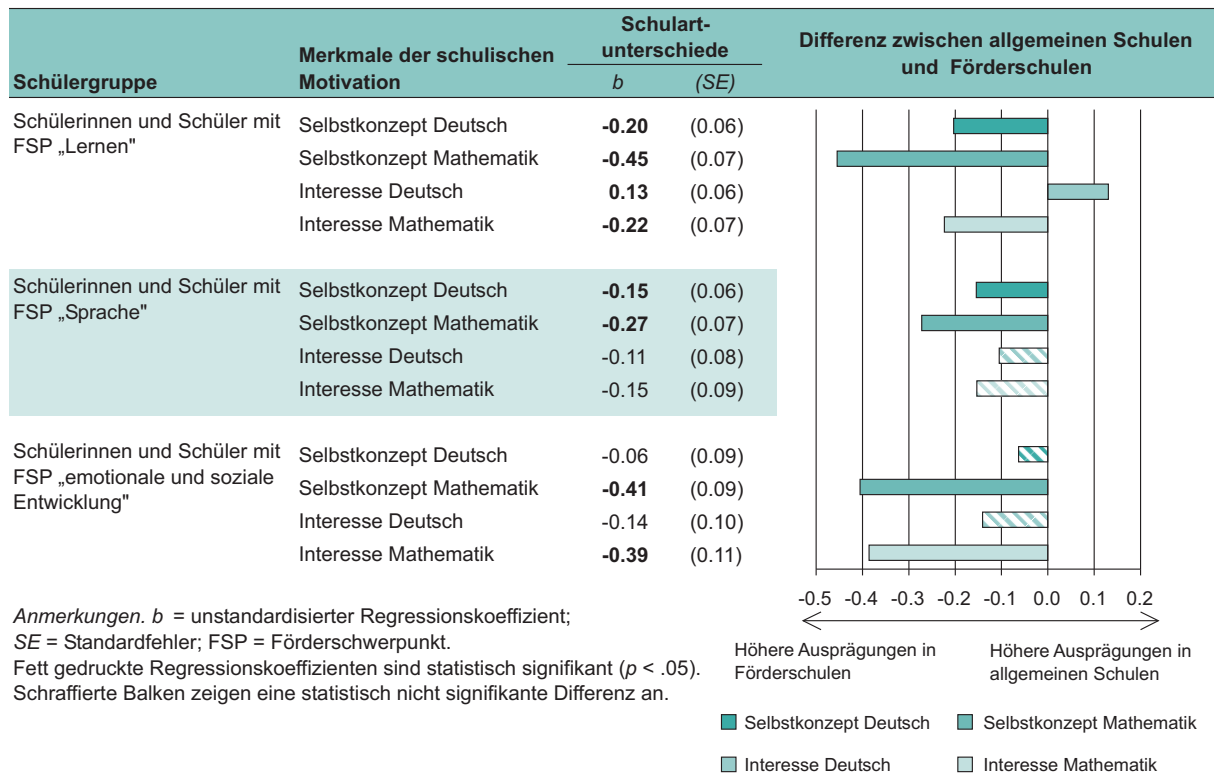
Anmerkungen. Positive Werte der Differenz stehen für höhere Ausprägungen in allgemeinen Schulen, negative Werte für höhere Ausprägungen in Förderschulen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Differenz der Mittelwerte minimal von der dargestellten Differenz ΔM abweichen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; ΔM = Mittelwertsdifferenz; d = Effektstärke Cohens d ; FSP = Förderschwerpunkt.

Fett gedruckte Differenzen sind statistisch signifikant ($p < .05$).

In Abbildung 10.10 sind die Mittelwertsunterschiede und Standardfehler für die motivationalen Merkmale der Kinder mit SPF in Förderschulen und in allgemeinen Schulen nach dem Propensity-Score-Matching abgebildet. Nach rechts weisende Balken zeigen eine höhere schulische Motivation von Kindern mit SPF in allgemeinen Schulen an; nach links weisende Balken zeigen eine höhere schulische Motivation von Kindern mit SPF in Förderschulen an. Durch das Propensity-Score-Matching vergrößern sich die motivationalen Unterschiede im Vergleich zu den Analysen vor dem Matching (Tab. 10.4). Dies könnte daran liegen, dass in den Analysen vor dem Matching Unterschiede zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen in den schulischen Kompetenzen und in weiteren Hintergrundmerkmalen nicht berücksichtigt werden. Diese unterschiedlichen Lernvoraussetzungen könnten zu schwächeren Zusammenhängen zwischen der Schulart und der schulischen Motivation führen. Im Rahmen der Analysen nach dem Propensity-Score-Matching werden unter anderem auch die individuellen schulischen Kompetenzen zur Vorhersage von Unterschieden in der schulischen Motivation einbezogen, um motivationale Unterschiede stärker mit der besuchten Schulart in Verbindung bringen zu können.

Kinder mit dem Förderschwerpunkt „Lernen“ in Förderschulen weisen ein signifikant höheres akademisches Selbstkonzept in den Fächern Deutsch und Mathematik sowie ein stärkeres Interesse am Fach Mathematik auf als Kinder mit vergleichbaren Lernvoraussetzungen in allgemeinen Schulen. Das Interesse

Abbildung 10.10: Unterschiede in Merkmalen der schulischen Motivation zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen *nach* dem Propensity-Score-Matching



am Fach Deutsch bei Kindern in allgemeinen Schulen ist dagegen signifikant höher als bei vergleichbaren Kindern in Förderschulen. Kinder mit dem Förderschwerpunkt „Sprache“ haben in Förderschulen ein signifikant höheres akademisches Selbstkonzept in den Fächern Deutsch und Mathematik als vergleichbare Kinder in allgemeinen Schulen. Für das fachspezifische Interesse ergeben sich keine signifikanten Unterschiede in Abhängigkeit von der Schulart. Bei Kindern mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ zeigen sich vor allem für das Fach Mathematik motivationale Schulartunterschiede: Kinder in Förderschulen weisen im Fach Mathematik ein deutlich höheres akademisches Selbstkonzept und Interesse auf als vergleichbare Kinder in allgemeinen Schulen; für die motivationalen Merkmale im Fach Deutsch ergeben sich nach dem Propensity-Score-Matching hingegen keine signifikanten Unterschiede in Abhängigkeit von der Schulart.

10.4.6 Zusammenfassung und Diskussion

Beim Vergleich der schulischen Kompetenzen und der schulischen Motivation zwischen Schülerinnen und Schülern mit SPF in Förderschulen und allgemeinen Schulen lassen sich nur für die Förderschwerpunkte „Lernen“ und „Sprache“ signifikante Zusammenhänge zwischen der Schulart und den schulischen Kompetenzen feststellen. Demnach profitieren Kinder mit dem Förderschwerpunkt „Lernen“ in besonderem Maße von einer gemeinsamen Beschulung in allgemeinen Schulen. Die Kompetenzunterschiede zugunsten von Kindern in allgemeinen Schulen entsprechen in den Kompetenzbereichen *Lesen* und *Zuhören*

ungefähr dem Lernzuwachs, der in einem halben Schuljahr zu erwarten ist (vgl. Kapitel 6.1). Im Fach Mathematik entsprechen die Kompetenzunterschiede zugunsten von Kindern in allgemeinen Schulen einem Lernzuwachs von etwas weniger als einem halben Schuljahr (vgl. Kapitel 6.2).

Für Kinder mit dem Förderschwerpunkt „Sprache“ zeigen sich ebenfalls insgesamt höhere schulische Kompetenzen in allgemeinen Schulen, allerdings sind die Unterschiede im Vergleich zum Förderschwerpunkt „Lernen“ geringer ausgeprägt. In den Kompetenzbereichen *Lesen* und *Zuhören* entsprechen die Unterschiede etwa einem Drittel Schuljahr und im Fach Mathematik weniger als einem halben Schuljahr zugunsten der Kinder in allgemeinen Schulen. Im Gegensatz dazu unterscheiden sich Kinder mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ in Förderschulen und allgemeinen Schulen in keinem der getesteten Kompetenzbereiche. Für den Kompetenzbereich *Orthografie* sind in keinem Förderschwerpunkt bedeutsame Unterschiede zwischen Kindern mit SPF in Förderschulen und allgemeinen Schulen zu verzeichnen.

Diese förderschwerpunktspezifischen Ergebnisse könnten unter anderem auf unterschiedliche Schülerzusammensetzungen und auf unterschiedliche Lehrpläne zurückzuführen sein. Die Gruppenvergleiche vor dem Matching (Tab. 10.3) zeigen, dass die Schülergruppen mit den Förderschwerpunkten „Sprache“ und „emotionale und soziale Entwicklung“ unabhängig von ihrer Schulart höhere schulische Kompetenzen erzielen als Kinder mit dem Förderschwerpunkt „Lernen“. Somit kann angenommen werden, dass Schülerinnen und Schüler mit diesen beiden Förderschwerpunkten in Förderschulen in leistungsstärkeren Lerngruppen unterrichtet werden als Kinder in Förderschulen mit dem Förderschwerpunkt „Lernen“. In leistungsstärkeren Lerngruppen könnte unabhängig von der besuchten Schulart ein anspruchsvoller Unterricht mit einem anregenden Lernklima durchgeführt werden (Dreeben & Barr, 1988), der zu höheren schulischen Kompetenzen beiträgt. Des Weiteren wird in Förderschulen mit den Förderschwerpunkten „Sprache“ und „emotionale und soziale Entwicklung“ nach den Lehrplänen der allgemeinen Schule gearbeitet (KMK, 1998; KMK, 2000), anders als beim Förderschwerpunkt „Lernen“, für den reduzierte Lehrpläne gelten (KMK, 1999). Gemeinsam könnten diese Unterschiede in der Schülerzusammensetzung und den Lehrplänen dazu beitragen, dass die Schulartunterschiede für Schülerinnen und Schüler mit dem Förderschwerpunkt „Lernen“ stärker ausgeprägt sind als für Kinder mit den Förderschwerpunkten „Sprache“ und „emotionale und soziale Entwicklung“.

Vergleicht man die Ergebnisse der Analysen zum IQB-Bildungstrend 2016 und zum IQB-Ländervergleich 2011, so zeigen sich übereinstimmend Kompetenzvorteile in den Bereichen *Lesen* und *Zuhören* sowie im Fach Mathematik zugunsten von Kindern mit SPF in allgemeinen Schulen (Kocaj et al., 2014). Allerdings sind die Kompetenzunterschiede im IQB-Bildungstrend 2016 deutlich geringer ausgeprägt als im IQB-Ländervergleich 2011. Eine mögliche Erklärung dafür könnte sein, dass durch die fortschreitende Inklusion zunehmend auch leistungsschwächere Schülerinnen und Schüler mit SPF in allgemeinen Schulen unterrichtet werden. Eventuell profitieren leistungsstärkere Schülerinnen und Schüler mit SPF aber mehr vom gemeinsamen Unterricht in allgemeinen Schulen (Gersten et al., 2009). Im IQB-Ländervergleich 2011 war der Anteil leistungsstärkerer Kinder mit SPF in allgemeinen Schulen möglicherweise höher als in den vorliegenden

Analysen. Das könnte dazu beitragen, dass die Kompetenzunterschiede im IQB-Ländervergleich 2011 stärker ausgeprägt sind als im IQB-Bildungstrend 2016.³

Im Gegensatz zu den schulischen Kompetenzen weisen Kinder mit SPF in Förderschulen insgesamt eine höhere schulische Motivation auf (fachspezifisches akademisches Selbstkonzept und Interesse) als vergleichbare Kinder mit SPF in allgemeinen Schulen. Diese motivationalen Unterschiede zugunsten von Förderschulen zeigen sich für alle untersuchten Förderschwerpunkte, sind jedoch für Kinder mit dem Förderschwerpunkt „Lernen“ und „emotionale und soziale Entwicklung“ tendenziell stärker ausgeprägt als für Kinder mit dem Förderschwerpunkt „Sprache“. Außerdem sind die Schulartunterschiede für das akademische Selbstkonzept stärker ausgeprägt als für das Interesse. Das höhere akademische Selbstkonzept von Kindern mit SPF in Förderschulen lässt sich durch soziale Vergleichsprozesse erklären (vgl. Abschnitt 10.4.2). In Förderschulen vergleichen sich Kinder mit SPF mit durchschnittlich leistungsschwächeren Mitschülerinnen und Mitschülern. Dies hat zur Folge, dass Kinder mit SPF in Förderschulen positivere Fähigkeitselbsteinschätzungen entwickeln (Renick & Harter, 1989). Im Gegensatz dazu vergleichen sich Kinder mit SPF in allgemeinen Schulen im Mittel mit leistungsstärkeren Mitschülerinnen und Mitschülern ohne SPF. Diese Vergleiche könnten dazu führen, dass sich Kinder mit SPF in allgemeinen Schulen als weniger kompetent einschätzen und eine geringere schulische Motivation aufweisen als vergleichbare Kinder in Förderschulen (Tracey & Marsh, 2000).

Zusammenfassend lassen sich gegenläufige Zusammenhänge zwischen der Schulart mit schulischen Kompetenzen und schulischer Motivation feststellen. Während Kinder mit SPF in allgemeinen Schulen im Durchschnitt höhere schulische Kompetenzen erzielen, weisen Kinder mit SPF in Förderschulen eine höhere schulische Motivation auf. Dabei scheinen je nach Förderschwerpunkt und Fach differenzierte Zusammenhänge zwischen Schulart und Merkmalen der schulischen Entwicklung zu bestehen. Die vorliegenden Ergebnisse weisen auf die Herausforderung hin, den Unterricht so zu gestalten, dass er den individuellen Bedürfnissen von Kindern mit SPF sowohl im Hinblick auf ihre Kompetenzentwicklung als auch auf ihre motivationale Entwicklung gerecht wird (Prenzel, 2013).

Literatur

- Ahrbeck, B. (2014). Schulische Inklusion – Möglichkeiten, Dilemmata und Widersprüche. *Soziale Passagen*, 6, 5–19.
- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag.
- Baumert, J. (2016). Bildungsgerechtigkeit in Deutschland – ein Überblick. In D. Döring (Hrsg.), *Auf der Suche nach der richtigen Ordnung: Königsteiner Forum 2014* (S. 63–98). Frankfurt am Main: Societäts-Verlag.
- 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.

3 Dafür spricht auch, dass sich Kinder mit SPF in Förderschulen und allgemeinen Schulen im IQB-Ländervergleich 2011 stärker in ihren Lernvoraussetzungen und weiteren Hintergrundmerkmalen unterscheiden als in den vorliegenden Analysen (Abb. 10.8).

- Bear, G. G., Minke, K. M. & Manning, M. A. (2002). Self-concept of students with learning disabilities: A meta-analysis. *School Psychology Review*, 31, 405–427.
- Dreeben, R. & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, 61, 129–142.
- Duncan, G. J., Magnuson, K. A. & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1, 59–80.
- Elbaum, B. (2002). The self-concept of students with learning disabilities: A meta-analysis of comparisons across different placements. *Learning Disabilities Research & Practice*, 17, 216–226.
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P. & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79, 1202–1242.
- Gresch, C. & Piezunka, A. (2015). Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich „Lernen“) an Regelschulen. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 181–220). Wiesbaden: Springer VS.
- Justice, L. M., Logan, J. A. R., Lin, T.-J. & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science*, 25, 1722–1729.
- KMK (1998) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1998). *Empfehlungen zum Förderschwerpunkt Sprache. Beschluss vom 26.06.1998*. Zugriff am 03.08.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sprache.pdf>
- KMK (1999) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1999). *Empfehlungen zum Förderschwerpunkt Lernen. Beschluss vom 01.10.1999*. Zugriff am 03.08.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/sopale.pdf>
- KMK (2000) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2000). *Empfehlungen zum Förderschwerpunkt emotionale und soziale Entwicklung. Beschluss vom 10.03.2000*. Zugriff am 03.08.2017 unter <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2000/emotsozentw.pdf>
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Sonderpädagogische Förderung in Schulen 2005 bis 2014*. Zugriff am 03.08.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Dok_210_SoPae_2014.pdf
- Knigge, M. (2017). Wie ein negatives gesellschaftliches Ansehen der besuchten Schulform das Lernen erschweren kann. *Sonderpädagogische Förderung in Brandenburg*, 1, 26–33.
- Kocaj, A., Kuhl, P., Jansen, M., Pant, H. A. & Stanat, P. (2017). *Educational placement and achievement motivation of students with special educational needs*. Manuskript in Vorbereitung.
- Kocaj, A., Kuhl, P., Kroth, A. J., Pant, H. A. & Stanat, P. (2014). Wo lernen Kinder mit sonderpädagogischem Förderbedarf besser? Ein Vergleich schulischer Kompetenzen zwischen Regel- und Förderschulen in der Primarstufe. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 165–191.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280–295.
- Peetsma, T., Vergeer, M., Roeleveld, J. & Karsten, S. (2001). Inclusion in education: Comparing pupils' development in special and regular education. *Educational Review*, 53, 125–135.
- Pfahl, L. (2012). Bildung, Behinderung und Agency. Eine wissenssoziologische Untersuchung der Folgen schulischer Segregation und Inklusion. In R. Becker & H. Solga (Hrsg.), *Soziologische Bildungsforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 52* (S. 415–436). Wiesbaden: Springer VS.
- Prenzel, A. (2013). *Inklusive Bildung in der Primarstufe. Eine wissenschaftliche Expertise für den Grundschulverband*. Frankfurt am Main: Grundschulverband.
- Renick, M. J. & Harter, S. (1989). Impact of social comparisons on the developing self-perceptions of learning disabled students. *Journal of Educational Psychology*, 81, 631–638.

- Rosenbaum, P. R. & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.
- Ruijs, N. M. & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4, 67–79.
- Sälzer, C., Gebhardt, M., Müller, K. & Pauly, E. (2015). Der Prozess der Feststellung sonderpädagogischen Förderbedarfs in Deutschland. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 129–152). Wiesbaden: Springer VS.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Stranghöner, D., Hollmann, J., Otterpohl, N., Wild, E., Lütje-Klose, B. & Schwinger, M. (2017). Inklusion versus Exklusion: Schulsetting und Lese-Rechtschreibentwicklung von Kindern mit Förderschwerpunkt Lernen. *Zeitschrift für Pädagogische Psychologie*, 31, 125–136.
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25, 1–21.
- Tent, L., Witt, M., Zschoche-Lieberum, C. & Bürger, W. (1991). Über die pädagogische Wirksamkeit der Schule für Lernbehinderte. *Zeitschrift für Heilpädagogik*, 5, 289–320.
- Tracey, D. K. & Marsh, H. W. (2000). Self-concepts of primary students with mild intellectual disabilities: Issues of measurement and educational placement. In R. G. Craven & H. W. Marsh (Hrsg.), *Self-concept theory, research and practice: Advances for the new millennium* (S. 419–425). Sydney: SELF Research Centre, University of Western Sydney.
- Vaughn, S., Elbaum, B. E. & Schumm, J. S. (1996). The effects of inclusion on the social functioning of students with learning disabilities. *Journal of Learning Disabilities*, 29, 598–608.
- Werning, R. & Lütje-Klose, B. (2016). *Einführung in die Pädagogik bei Lernbeeinträchtigungen*. München: Ernst Reinhardt.
- Wild, E., Schwinger, M., Lütje-Klose, B., Yotyodying, S., Gorges, J., Stranghöner, D., Neumann, P., Serke, B. & Kurnitzki, S. (2015). Schülerinnen und Schüler mit dem Förderschwerpunkt Lernen in inklusiven und exklusiven Förderarrangements: Erste Befunde des BiLieF-Projektes zu Leistung, sozialer Integration, Motivation und Wohlbefinden. *Unterrichtswissenschaft*, 43, 7–21.
- Wocken, H. (2005). *Andere Länder, andere Schüler? Vergleichende Untersuchung von Förderschülern in den Bundesländern Brandenburg, Hamburg und Niedersachsen*. Zugriff am 03.08.2017 unter <http://bidok.uibk.ac.at/download/wocken-forschungsbericht.pdf>

Kapitel 11

Leistungsstarke Schülerinnen und Schüler in Deutschland

Claudia Neuendorf, Poldi Kuhl und Malte Jansen

11.1 Einleitung

Der Umgang mit zunehmender Heterogenität in Schule und Unterricht gehört zu den zentralen Themen der empirischen Bildungsforschung. Nachdem die Aufmerksamkeit in den letzten Jahren vor allem auf Schülerinnen und Schülern mit vergleichsweise schwachen Leistungen und ungünstigen Lernvoraussetzungen lag (vgl. Kapitel 10 in diesem Bericht), geraten nun zunehmend auch leistungsstarke Schülerinnen und Schüler in den Fokus. Die Forderung nach Bildungsgerechtigkeit soll sich auf alle Schülerinnen und Schüler beziehen und neben der Förderung von Kindern mit ungünstigeren Lernvoraussetzungen auch eine angemessene Förderung leistungsstarker Kinder umfassen (z. B. Hammer et al., 2016).

Mit Verweis auf nationale und internationale Schulleistungsstudien, bei denen der Anteil leistungsstarker Schülerinnen und Schüler in Deutschland nach Einschätzung der Kultusministerkonferenz (KMK) vergleichsweise gering ist, hat die KMK vor zwei Jahren eine Strategie zur Förderung leistungsstarker Schülerinnen und Schüler verabschiedet (KMK, 2015) sowie gemeinsam mit dem Bund eine Initiative „zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler“ beschlossen (KMK, 2016). Auch das neue Rahmenprogramm des Bundesministeriums für Bildung und Forschung (BMBF) für die empirische Bildungsforschung konstatiert Handlungsbedarf bezüglich der Förderung „leistungsstarker und leistungsfähiger“ Schülerinnen und Schüler (BMBF, 2017, S. 11). Diese Schülerinnen und Schüler sollen in den kommenden Jahren intensiver untersucht, Ansatzpunkte für ihre Förderung identifiziert und Fortschritte in der Förderung überprüft werden.

Dieses Kapitel nimmt daher die Gruppe der leistungsstarken Schülerinnen und Schüler in den Blick. Zunächst wird eine operationale Definition von Leistungsstärke im schulischen Bereich vorgestellt. Auf dieser Basis werden die Anteile leistungsstarker Schülerinnen und Schüler in den Fächern Deutsch und Mathematik bestimmt (Abschnitt 11.2). Dabei wird auch der Frage nachgegangen, wie stark sich diese Anteile in den untersuchten Fächern und Kompetenzbereichen im IQB-Bildungstrend 2016 unterscheiden und wie groß der Überschneidungsbereich ist. In die Analysen dieses Kapitels gehen daher nur diejenigen Schülerinnen und Schüler ein, die sowohl am Kompetenztest im Fach Deutsch als auch an der Erhebung im Fach Mathematik teilgenommen haben.¹

¹ Da die Erhebung an den Schulen auf zwei Testtage verteilt war und pro Testtag nur Aufgaben zu einem der beiden Fächer vorgelegt wurden, konnte es vorkommen, dass einzelne Schülerinnen und Schüler bei der Erhebung zu einem Fach anwesend waren, bei der Erhebung zum anderen Fach aber nicht (vgl. Kapitel 3.1). Insgesamt wurden aus diesem Grund 4,8 Prozent der Schülerinnen und Schüler der Gesamtstichprobe von den Analysen des vorliegenden Kapitels ausgeschlossen.

Im Anschluss werden die leistungsstarken Viertklässlerinnen und Viertklässler anhand zentraler individueller, familiärer und motivationaler Merkmale charakterisiert (Abschnitt 11.3). Schließlich wird dargestellt, welche Fördermaßnahmen aktuell an Schulen im Primarbereich zur besonderen Förderung leistungsstarker Schülerinnen und Schüler angeboten werden und in welchem Umfang leistungsstarke Schülerinnen und Schüler diese auch nutzen (Abschnitt 11.4).

11.2 Leistungsstarke Viertklässlerinnen und Viertklässler in den Fächern Deutsch und Mathematik

Eine grundlegende Herausforderung bei der Untersuchung von leistungsstarken Schülerinnen und Schülern besteht darin, diese Gruppe zu definieren. Bislang fehlt eine klare Definition von Leistungsstärke im schulischen Bereich (Köller & Baumert, 2017). Für die im Rahmen des IQB-Bildungstrends 2016 vorgenommene Beschreibung leistungsstarker Schülerinnen und Schüler bieten die KMK-Bildungsstandards einen kriterialen Referenzrahmen (vgl. Kapitel 2), auf den sich auch die Förderstrategie der KMK bezieht (KMK, 2015) und der somit zur Definition dieser Gruppe genutzt werden kann. Gemäß den auf Basis der Bildungsstandards entwickelten Kompetenzstufenmodellen im Primarbereich erbringen Kinder, die in einem Kompetenzbereich die Kompetenzstufe V und somit den Optimalstandard erreichen, Leistungen, die „bei sehr guten oder ausgezeichneten individuellen Lernvoraussetzungen und der Bereitstellung geeigneter Lerngelegenheiten innerhalb und außerhalb der Schule erreicht werden können und die bei Weitem die Erwartungen der KMK-Bildungsstandards übertreffen“ (vgl. Kapitel 2.1, S. 50).

Im folgenden Abschnitt wird der Begriff Leistungsstärke im IQB-Bildungstrend 2016 anhand dieser Definition operationalisiert, wobei die Anteile leistungsstarker Schülerinnen und Schüler dargestellt werden, die sich bei ihrer Anwendung ergeben. Im ersten Schritt wird hierbei auf die einzelnen Fächer Mathematik und Deutsch eingegangen, bevor in einem zweiten Schritt eine fächerübergreifende Perspektive als Grundlage für die folgenden Abschnitte eingenommen wird.

11.2.1 Leistungsstarke Schülerinnen und Schüler im Fach Mathematik

Auf der Grundlage der Bildungsstandards der KMK und der darauf basierenden Kompetenzstufenmodelle werden Schülerinnen und Schüler im Fach Mathematik als leistungsstark definiert, wenn sie auf der Globalskala den Optimalstandard (Kompetenzstufe V) erreichen. Diese Schülerinnen und Schüler „bearbeiten mathematische Problemstellungen in allen Inhaltsbereichen auch in einem unbekanntem Kontext angemessen, sicher und flexibel“ (vgl. Kapitel 2.3, S. 73). Im IQB-Bildungstrend 2016 liegt dieser Anteil in Deutschland insgesamt bei 13.1 Prozent und damit 3.0 Prozentpunkte niedriger als im IQB-Ländervergleich 2011, in dem 16.1 Prozent der Viertklässlerinnen und Viertklässler im Fach Mathematik den Optimalstandard erreichten (vgl. Kapitel 5.2).

11.2.2 Leistungsstarke Schülerinnen und Schüler im Fach Deutsch

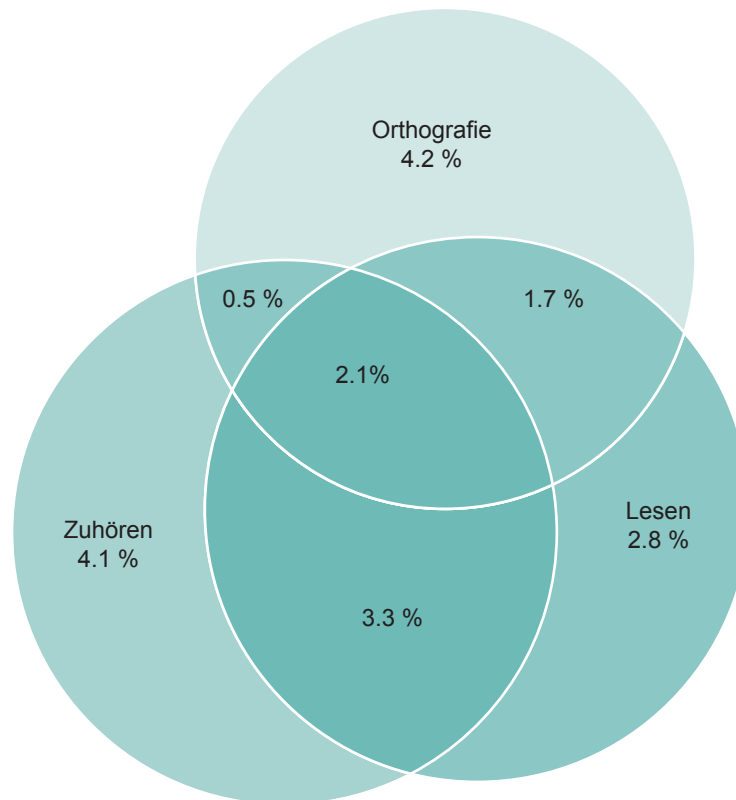
Für das Fach Deutsch gestaltet sich die Festlegung, welche Schülerinnen und Schüler als leistungsstark einzuordnen sind, komplexer. Anders als für das Fach Mathematik wird in den Kompetenzstufenmodellen und Studien des IQB für das Fach Deutsch keine Globalskala gebildet, da die Annahme einer globalen Deutschkompetenz aus empirischen und fachdidaktischen Gründen nicht sinnvoll ist (vgl. Kapitel 1.2). Vielmehr werden die Ergebnisse im Fach Deutsch im IQB-Bildungstrend 2016 ausschließlich separat für die drei untersuchten Kompetenzbereiche berichtet. In den Bereichen *Lesen*, *Zuhören* und *Orthografie* erreichen demnach bundesweit 10.2 Prozent, 10.3 Prozent beziehungsweise 8.6 Prozent der Schülerinnen und Schüler den Optimalstandard (vgl. Kapitel 5.1). Im Vergleich zum Jahr 2011 haben sich diese Anteile im *Lesen* und in der *Orthografie* signifikant verringert. Dabei ist die Veränderung im *Lesen* mit einem Rückgang von 1.9 Prozentpunkten geringer als in der *Orthografie* mit 3.9 Prozentpunkten. Im *Zuhören* hat sich der Anteil leistungsstarker Schülerinnen und Schüler gegenüber 2011 nicht signifikant verändert.

Werden die Kompetenzbereiche im Fach Deutsch gemeinsam betrachtet, zeigt sich, dass insgesamt 18.7 Prozent der Viertklässlerinnen und Viertklässler in mindestens einem der drei Kompetenzbereiche den Optimalstandard erreichen (vgl. Tab. 11.1). Jeweils gut 4 Prozent der Viertklässlerinnen und Viertklässler und somit ein großer Teil derjenigen, die in mindestens einem Kompetenzbereich besonders leistungsstark sind, erreicht den Optimalstandard entweder nur im *Zuhören* oder nur in der *Orthografie*. Etwas geringer ist der Anteil derjenigen, die nur im *Lesen*, nicht aber in einem der anderen Kompetenzbereiche, den Optimalstandard erreicht. Zwischen *Lesen* und *Zuhören* sind die Überschneidungen am größten: 3.3 Prozent der Schülerinnen und Schüler erreichen in beiden Bereichen den Optimalstandard. Sowohl im *Lesen* als auch in der *Orthografie* erreichen 1.7 Prozent den Optimalstandard, während die Überschneidung zwischen *Zuhören* und *Orthografie* mit lediglich 0.5 Prozent am geringsten ist. Etwa 2 Prozent der Schülerinnen und Schüler erreichen in allen drei Kompetenzbereichen den Optimalstandard (vgl. Abbildung 11.1).

Um Leistungsstärke für das Fach Deutsch bereichsübergreifend zu operationalisieren, sind verschiedene Ansätze denkbar. So könnte man sich einerseits auf die erreichten Kompetenzen im *Lesen* fokussieren, das als grundlegende Kulturtechnik für den weiteren Wissensaufbau zentral ist (Kirsch, de Jong, McQueen, Mendelovits & Monseur, 2002). Andererseits umfassen die Bildungsstandards der KMK im Fach Deutsch neben dem *Lesen* weitere Kompetenzbereiche, denen ebenfalls ein wichtiger Stellenwert in der Entwicklung von Sprachhandlungskompetenz als Basis schulischen Lernens beigemessen wird (KMK, 2005). Möchte man alle Kompetenzbereiche des Faches Deutsch gleich gewichten, stellt sich die Frage, ob leistungsstarke Schülerinnen und Schüler sich in allen Bereichen durch besonders hohe Leistungen auszeichnen müssen (und-Verknüpfung; enge Definition) oder ob herausragende Leistungen in einem Kompetenzbereich ausreichen (oder-Verknüpfung; weite Definition).

Im Rahmen des vorliegenden Berichts wird eine vergleichsweise weite Definition zugrunde gelegt. Als leistungsstark im Fach Deutsch werden Kinder bezeichnet, wenn sie den Optimalstandard in mindestens einem Kompetenzbereich erreicht und gleichzeitig den Regelstandard (Kompetenzstufe III) in keinem der drei getesteten Kompetenzbereiche verfehlt haben. Tabelle 11.1 stellt dar, wie viele Kinder nach dieser Definition im Fach Deutsch zu den

Abbildung 11.1: Anteil der Schülerinnen und Schüler, die in den Kompetenzbereichen *Lesen*, *Zuhören* und *Orthografie* den Optimalstandard erreichen



Leistungsstarken gehören. Darüber hinaus wird dargestellt, welcher Anteil der Schülerinnen und Schüler aufgrund des nicht erreichten Regelstandards in einem oder mehreren der anderen Kompetenzbereiche nicht zu den Leistungsstarken gezählt wird. Nach der hier zugrunde gelegten Definition gelten 16.4 Prozent der Kinder als leistungsstark im Fach Deutsch. Nur 2.3 Prozent der Kinder erreichen zwar in mindestens einem der getesteten Kompetenzbereiche im Fach Deutsch den Optimalstandard, verfehlen aber den Regelstandard in mindestens einem weiteren Kompetenzbereich und werden daher hier nicht zur Gruppe der leistungsstarken Schülerinnen und Schüler gezählt. Der Großteil dieser Kinder, die in mindestens einem Bereich die Kompetenzstufe V erlangt, aber in mindestens einem anderen Kompetenzbereich die Kompetenzstufe III verfehlt haben, erreicht den Optimalstandard im Kompetenzbereich *Zuhören*, verfehlt aber den Regelstandard in *Lesen* oder in der *Orthografie* (1.3 %, vgl. Tab. 11.1).

11.2.3 Fächerübergreifend leistungsstarke Schülerinnen und Schüler

Die Frage, inwieweit besondere fächerübergreifende Leistungsstärken bestehen, wurde in Schulleistungsstudien bislang kaum analysiert. Eine Ausnahme bilden die Analysen, die im Rahmen der Studien IGLU² 2011 und TIMSS³ 2011 für die Leistungen von Viertklässlerinnen und Viertklässlern im *Lesen*, in Mathematik und in den Naturwissenschaften vorgenommen wurden (Bos et al., 2012;

2 Das Akronym IGLU steht für *Internationale Grundschul-Lese-Untersuchung*.

3 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.

Tabelle 11.1: Anteil der Schülerinnen und Schüler, die den Optimalstandard in einem oder mehreren Kompetenzbereichen des Faches Deutsch erreichen

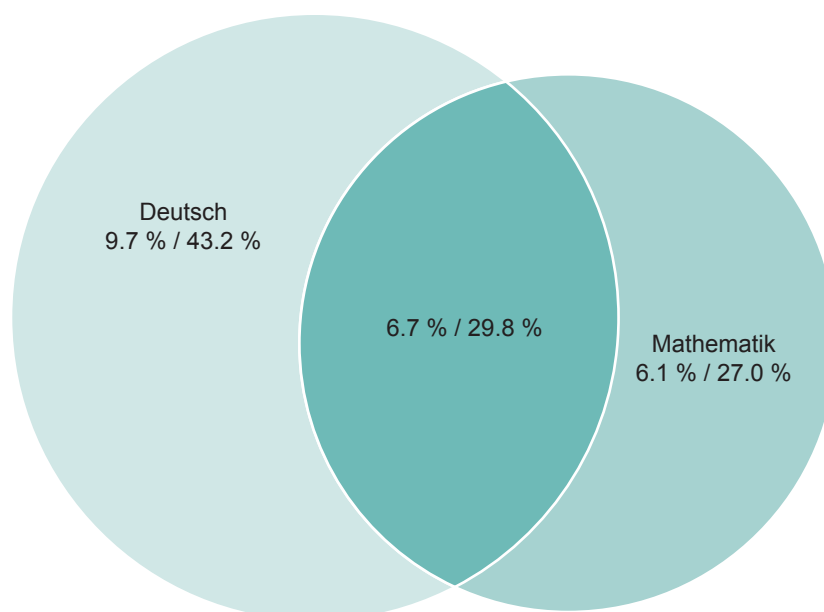
Optimalstandard erreicht	Gesamt (in %)	davon leistungsstark ¹	
		ja	nein
nur Lesen	2.8	2.6	0.2
nur Zuhören	4.1	2.8	1.3
nur Orthografie	4.2	3.6	0.5
Lesen und Zuhören	3.3	3.1	0.3
Zuhören und Orthografie	0.5	0.5	0.0
Lesen und Orthografie	1.7	1.7	0.0
Lesen, Zuhören und Orthografie	2.1	2.1	-
Gesamt	18.7	16.4	2.3

Anmerkung. ¹ Als leistungsstark werden Schülerinnen und Schüler bezeichnet, die in mindestens einem Kompetenzbereich den Optimalstandard erreichen und in keinem Kompetenzbereich den Regelstandard verfehlen.

Wendt, Willems, Tarelli, Euen & Bos, 2013). Dabei wurde festgestellt, dass die Kompetenzen von Schülerinnen und Schülern über die untersuchten Fächer hinweg relativ homogen waren.

Im Folgenden werden für den IQB-Bildungstrend 2016 die Leistungsprofile leistungsstarker Schülerinnen und Schüler in den Fächern Deutsch und Mathematik beschrieben. Dabei wird für das Fach Mathematik das Erreichen des Optimalstandards auf der Globalskala herangezogen; für das Fach Deutsch wird, wie oben beschrieben, von Leistungsstärke gesprochen, wenn in mindestens einem der drei getesteten Kompetenzbereiche der Optimalstandard erreicht und gleichzeitig in keinem Kompetenzbereich der Regelstandard verfehlt wird.

Nach dieser Definition sind insgesamt 22.5 Prozent aller Schülerinnen und Schüler in mindestens einem der beiden Fächer leistungsstark. Wie Abbildung 11.2 darstellt, erreichen 27.0 Prozent der so definierten Leistungsstarken

Abbildung 11.2: Anteile der Schülerinnen und Schüler, die im Fach Deutsch und/oder im Fach Mathematik leistungsstark sind (Gesamtanteil / Anteil an Leistungsstarken)

nur im Fach Mathematik, nicht aber im Fach Deutsch den Optimalstandard, während 43.2 Prozent der Leistungsstarken demgegenüber im Fach Deutsch, nicht aber im Fach Mathematik zu den Leistungsstarken zu zählen sind. Ein knappes Drittel der leistungsstarken Kinder ist sowohl im Fach Mathematik als auch im Fach Deutsch und damit fächerübergreifend leistungsstark (vgl. Abbildung 11.2).

Dass die Mehrzahl der leistungsstarken Schülerinnen und Schüler nur in einem der beiden Fächer leistungsstark ist, spricht für eine gewisse Domänenspezifität von Leistungsstärke. Um beantworten zu können, ob sich die leistungsstarken Kinder untereinander systematisch unterscheiden sowie von Schülerinnen und Schülern zu unterscheiden sind, die im Sinne der hier verwendeten Definition als nicht leistungsstark gelten, werden die weiteren Ergebnisse im Folgenden für vier Gruppen separat dargestellt. Verglichen werden (1) Schülerinnen und Schüler, die nur im Fach Deutsch leistungsstark sind, (2) Schülerinnen und Schüler, die nur im Fach Mathematik leistungsstark sind, (3) Schülerinnen und Schüler, die in beiden Fächern leistungsstark sind, und (4) Schülerinnen und Schüler, die weder im Fach Deutsch noch im Fach Mathematik als leistungsstark identifiziert wurden (im Folgenden auch *Vergleichsgruppe* genannt; diese umfasst 77.5 Prozent aller Viertklässlerinnen und Viertklässler).

11.3 Merkmale leistungsstarker Schülerinnen und Schüler

In diesem Abschnitt sollen die leistungsstarken Schülerinnen und Schüler im IQB-Bildungstrend 2016 anhand verschiedener Hintergrundmerkmale charakterisiert werden. Dabei werden zum einen solche Merkmale einbezogen, von denen bekannt ist, dass sie in einem fachspezifischen Zusammenhang mit den von Schülerinnen und Schülern erzielten Leistungen stehen, wie beispielsweise Geschlecht und motivationale Merkmale. Bei diesen Merkmalen werden daher Unterschiede zwischen den drei Gruppen leistungsstarker Schülerinnen und Schüler erwartet. Zum anderen werden Merkmale wie der sozioökonomische Hintergrund oder das kulturelle Kapital betrachtet, für die bekannt ist, dass sie über unterschiedliche Fächer hinweg einen Einfluss auf die schulische Kompetenzentwicklung haben (Helmke & Weinert, 1997). Hier ist aufgrund der fächerübergreifenden Befundmuster davon auszugehen, dass die drei oben beschriebenen Gruppen Leistungsstarker sich in ähnlicher Weise von der Vergleichsgruppe unterscheiden.

11.3.1 Geschlecht

Geschlechterunterschiede in schulischen Kompetenzen sind ein Thema, das anhaltend erhebliche Aufmerksamkeit erfährt, da vom Bildungssystem erwartet wird, dass es Jungen und Mädchen die gleichen Entwicklungschancen bietet. Tatsächlich haben in den vergangenen Jahren zahlreiche Studien bedeutsame Unterschiede in den von Jungen und Mädchen im Durchschnitt erreichten Kompetenzen belegt. Studienübergreifend konnten dabei Kompetenzvorteile für Mädchen im sprachlichen Bereich gezeigt werden, während sich für Jungen ein Kompetenzvorsprung im mathematischen Bereich ergibt. Dieses Befundmuster bestätigt sich auch im IQB-Bildungstrend 2016 (vgl. Kapitel 7).

Die Existenz von Differenzen in Kompetenzmittelwerten lässt jedoch keine direkten Schlüsse darauf zu, inwieweit auch an den Rändern der Kompetenz-

verteilung Geschlechterunterschiede bestehen. Es lässt sich also nicht unmittelbar ableiten, wie viele Jungen beziehungsweise Mädchen besonders leistungsschwach oder besonders leistungsstark sind. Arbeiten zur *greater male variability hypothesis* zeigen, dass bei Jungen und Männern in vielen kognitiven Fähigkeiten eine größere Streuung vorliegt, die dazu führt, dass sie gegenüber Mädchen und Frauen in den Extremgruppen – also sowohl unter den besonders Leistungsstarken als auch bei den besonders Leistungsschwachen – überrepräsentiert sind (Baye & Monseur, 2016; Penner, 2008). Damit übereinstimmend zeigte die oben bereits beschriebene kompetenzbereichsübergreifende Analyse der Daten von IGLU 2011 und TIMSS 2011 (Wendt et al., 2013), dass in der Gruppe der Schülerinnen und Schüler mit herausragenden Leistungen in den Bereichen *Lesen*, *Mathematik* und *Naturwissenschaften*, die etwa vier Prozent aller Schülerinnen und Schüler umfasste, 58,3 Prozent Jungen waren. In allen anderen Leistungsgruppen hingegen waren keine signifikanten Geschlechterunterschiede zu verzeichnen.

Die Ergebnisse der Analyse im IQB-Bildungstrend 2016 sind in Tabelle 11.2 dargestellt. Erwartungsgemäß zeigt sich, dass bereits in der 4. Jahrgangsstufe die Gruppe der Kinder, die im Sinne der hier verwendeten Definition nur im Fach *Deutsch* leistungsstark sind, zu 67 Prozent aus Mädchen besteht, während die Gruppe der nur im Fach *Mathematik* leistungsstarken Kinder zu 70 Prozent Jungen umfasst. Neben den prozentualen Anteilen von Jungen und Mädchen in den verschiedenen Gruppen sind in Tabelle 11.2 auch die Chancenverhältnisse (*Odds Ratios, OR*) dargestellt. Diese sind Maße der Effektstärke und auch bei unterschiedlichen Gruppengrößen vergleichbar. Als Referenzgruppe dient jeweils die Vergleichsgruppe der Kinder, die in keinem der getesteten Fächer leistungsstark sind. Das ermittelte Chancenverhältnis von 2.1 für Mädchen in der Gruppe der nur in *Deutsch* Leistungsstarken bedeutet, dass die Chance, dieser Gruppe anzugehören und nicht der Vergleichsgruppe, für Mädchen mehr als doppelt so hoch ist wie für Jungen. Hingegen ist die Chance von Mädchen, der Gruppe der nur in *Mathematik* Leistungsstarken anzugehören, 0,4-mal so hoch und damit deutlich geringer als für Jungen.

Die Geschlechterverteilung in der Gruppe der Kinder, die in beiden Fächern leistungsstark sind, unterscheidet sich dagegen nicht signifikant von derjenigen in der Gruppe der Kinder, die in keinem der beiden Fächer zu den leistungsstarken gehören; sie ist in diesen beiden Gruppen nahezu ausgeglichen (vgl. Tab. 11.2). Diese Ergebnisse stellen eine Erweiterung bisheriger Analysen dar. Zum einen wird deutlich, dass Jungen eher eine Stärke im Fach *Mathematik* und Mädchen eher eine Stärke im Fach *Deutsch* aufweisen. Zum anderen zeigt sich jedoch, dass die Anteile von Jungen und Mädchen in der Gruppe der Kinder, die in beiden Fächern zugleich leistungsstark sind, ähnlich hoch sind. Dieser Befund ist umso erfreulicher, da die Gruppe dieser vielseitig leistungsstarken Kinder keineswegs deutlich kleiner ausfällt als die anderen beiden Teilgruppen.

11.3.2 Familiärer Hintergrund

Internationale und nationale Schulleistungsstudien haben wiederholt gezeigt, dass es einen bedeutsamen Zusammenhang zwischen den Kompetenzen der Schülerinnen und Schüler und deren sozialem Hintergrund in Deutschland gibt (vgl. Kapitel 8; Richter, Kuhl & Pant, 2012; Stubbe, Schwippert & Wendt, 2016; Wendt, Stubbe & Schwippert, 2012). Vor diesem Hintergrund wäre zu erwarten,

Tabelle 11.2: Vergleich der Verteilungen der Hintergrundmerkmale in den vier Leistungsgruppen

Merkmale	Vergleichsgruppe ¹		Leistungsstark in Deutsch			Leistungsstark in Mathematik			Leistungsstark in Deutsch und Mathematik		
	%	(SE)	%	(SE)	OR	%	(SE)	OR	%	(SE)	OR
Geschlecht											
Jungen ¹	51.0	(0.5)	32.8	(1.8)	–	70.1	(2.5)	–	52.1	(2.4)	–
Mädchen	48.9	(0.5)	67.1	(1.8)	2.1	29.7	(2.5)	0.4	47.8	(2.4)	1.0
familiäre Merkmale											
höchste elterliche EGP-Klasse²											
EGP I–II	50.3	(0.8)	75.0	(1.7)	2.3	75.0	(2.1)	2.5	85.0	(1.6)	4.1
EGP III–IV ¹	28.7	(0.6)	19.0	(1.6)	–	17.1	(1.9)	–	11.8	(1.4)	–
EGP V–VII	20.9	(0.6)	6.0	(1.1)	0.4	7.9	(1.3)	0.6	3.2	(0.7)	0.4
elterliche Bildung											
mind. ein Elternteil mit Abitur	22.8	(0.6)	48.2	(2.0)	2.9	51.6	(2.7)	3.3	64.7	(2.8)	5.9
kein Elternteil mit Abitur ¹	50.7	(0.8)	36.5	(2.2)	–	34.3	(2.6)	–	24.2	(2.4)	–

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente minimal von 100 abweichen. SE = Standardfehler; OR = Odds Ratios (Chancenverhältnisse); EGP I–II = hoher sozioökonomischer Status; EGP III–IV = mittlerer sozioökonomischer Status; EGP V–VII = niedriger sozioökonomischer Status.

¹ Referenzgruppe für Chancenverhältnisse. ² EGP-Klassen nach Erikson & Goldthorpe, 2002; Erikson, Goldthorpe & Portocarero, 1979. Fett gedruckte Werte unterscheiden sich statistisch signifikant ($p < .05$) von der Gruppe der nicht leistungsstarken Schülerinnen und Schüler.

dass auch in der Gruppe der leistungsstarken Schülerinnen und Schüler ein größerer Anteil der Kinder aus Familien mit hohem sozialem Status kommt als in der Gesamtpopulation. Diese Annahme soll im Folgenden geprüft werden.

Als Indikator des sozialen Hintergrunds der Schülerinnen und Schüler wird in Schulleistungsstudien unter anderem die EGP-Klassifikation beruflicher Tätigkeiten herangezogen (Erikson & Goldthorpe, 2002; Erikson, Goldthorpe & Portocarero, 1979; vgl. Kapitel 8), die auch die Grundlage der folgenden Analysen bildet. Dabei werden Schülerinnen und Schüler aus Familien mit einem hohen sozialen Status (EGP-Klassen I–II) und einem niedrigen sozialen Status (EGP-Klassen V–VII) den Kindern aus Familien mit einem mittleren sozialen Status (EGP-Klassen III–IV) gegenübergestellt. Zusätzlich wird ein Indikator des Bildungshintergrunds gebildet, der angibt, ob mindestens ein Elternteil über den Abschluss Abitur verfügt.

Im unteren Teil der Tabelle 11.2 sind die Ergebnisse zum sozialen Hintergrund der drei Gruppen leistungsstarker Schülerinnen und Schüler und der Vergleichsgruppe dargestellt. Es sind die prozentualen Anteile der sozialen Merkmalskategorien für die jeweilige Leistungsgruppe aufgeführt. Zusätzlich wird das Chancenverhältnis (OR) berichtet, das jeweils angibt, um wie viel höher die relative Chance für ein Kind mit dem betrachteten familiären Hintergrundmerkmal gegenüber einem Kind mit dem Referenzmerkmal ist, in der jeweiligen Gruppe Leistungsstarker zu sein statt in der Vergleichsgruppe. Es zeigt sich durchgängig, dass Kinder aus Familien mit einem hohen sozialen Status häufiger unter den Leistungsstarken anzutreffen sind als in der Vergleichsgruppe. So haben Kinder von Eltern der höheren Dienstklassen (EGP-Klassen I–II) eine etwa zweieinhalbmal so hohe Chance, im Fach Deutsch oder Mathematik leistungsstark zu sein, als Kinder aus Familien mit einem mittleren sozialen Status. Bei Kindern, die in

beiden Fächern leistungsstark sind, ist das Chancenverhältnis noch unausgeglichener. Für Kinder aus einer Familie mit hohem sozialem Status ist die Chance, in dieser Gruppe statt in der Vergleichsgruppe zu sein, 4.1-mal so hoch wie bei Kindern aus Familien mit einem mittleren sozialen Status. Das Chancenverhältnis für Kinder mit niedrigem sozialen Status liegt bei 0.4 bis 0.6; das heißt, die Chance, leistungsstark im Fach Deutsch und/oder im Fach Mathematik zu sein, ist für sie also etwa halb so groß wie für Kinder mit mittlerem sozialem Status.

Auch für den Bildungshintergrund der Eltern zeigen sich signifikante Unterschiede zwischen den vier untersuchten Gruppen. So ist der Anteil derjenigen Kinder, deren Eltern Abitur haben, in der Gruppe der Leistungsstarken in einem Fach signifikant höher (48.2 % im Fach Deutsch bzw. 51.6 % im Fach Mathematik gegenüber 22.8 % in der Vergleichsgruppe) und bei der Gruppe, die in beiden Fächern leistungsstark ist, mit 64.7 Prozent am höchsten. Auch die Analysen der Chancenverhältnisse zeigen ein vergleichbares Muster. Kinder von Eltern mit Abitur haben eine um das Dreifache erhöhte Chance, in Deutsch oder Mathematik leistungsstark zu sein verglichen mit Kindern von Eltern, die kein Abitur haben. Für Kinder, die in beiden Fächern leistungsstark sind, erhöht sich die Chance auf das Sechsfache (Tab. 11.2).

Trotz dieser Befunde, die dafür sprechen, dass Kinder aus bildungsnäheren Elternhäusern höhere Chancen haben, in der Schule Spitzenleistungen zu erbringen, stammt ein Viertel der leistungsstarken Kinder im Fach Deutsch oder Mathematik aus Familien mit einer mittleren oder einfachen sozialen Lage und jeweils mehr als ein Drittel hat Eltern ohne Allgemeine Hochschulreife. Es scheint an vielen Schulen im Primarbereich also bereits zu gelingen, das Potenzial einzelner Kinder in Deutsch oder Mathematik trotz weniger günstiger Voraussetzungen zur Entfaltung zu bringen. Dennoch ist der Unterschied zur Vergleichsgruppe, in der etwa die Hälfte der Kinder aus einer mittleren oder einfachen sozialen Lage stammt beziehungsweise kein Elternteil mit Abitur hat, sehr deutlich.

Bei der Gruppe der Schülerinnen und Schüler, die in beiden Fächern zu den leistungsstarken Kindern gehören, liegt der entsprechende Anteil von Kindern aus Familien mit weniger günstigen Voraussetzungen bei 15.0 Prozent (EGP III–IV und EGP V–VII) beziehungsweise 24.2 Prozent (kein Elternteil mit Abitur) und damit noch niedriger als bei den zuvor beschriebenen Gruppen (Tab. 11.2). Für Spitzenleistungen in mehreren Bereichen scheinen familiäre Merkmale also in noch größerem Maße relevant zu sein als für Spitzenleistungen in einzelnen Fächern.

11.3.3 Motivationale Merkmale

Neben den zuvor betrachteten Merkmalen sind auch motivationale Schülermerkmale für den Erwerb von Kompetenzen bei allen Schülerinnen und Schülern bedeutsam (Dai, Moon & Feldhusen, 1998; Marsh & O'Mara, 2008). Zudem wird die Entwicklung von Interesse sowie Lern- und Leistungsbereitschaft an sich auch als ein genuines Bildungsziel von Schule betrachtet (Clausen, 2002; Jansen, Schroeders & Stanat, 2013).

Befunde aus der Begabtenforschung sprechen dafür, dass die intrinsische Motivation, die aus einer herausfordernden Aufgabe entsteht, eine wesentliche Rolle bei der Entwicklung von Hochbegabung spielt (Dai et al., 1998; Phillips & Lindsay, 2006). Entsprechend wird befürchtet, dass das Anspruchsniveau eines

Unterrichts, der sich häufig an den durchschnittlichen Schülerinnen und Schülern einer Klasse orientiert, zu gering sein könnte (Gronostaj & Vock, 2014; Kiuru et al., 2015). Dies könne wiederum zu einer Verringerung der Lernmotivation der potenziell leistungsstarken Schülerinnen und Schüler führen und deren Bildungskarrieren damit langfristig negativ beeinflussen (Baumert, Roeder, Sang & Schmitz, 1986; Vock & Gronostaj, 2017). Vor diesem Hintergrund sollen auch Ausprägungen motivationaler Merkmale in diesem Kapitel analysiert werden. Hierzu werden die drei Gruppen leistungsstarker Schülerinnen und Schüler den Kindern der Vergleichsgruppe in ihren fächerübergreifenden und fachspezifischen motivationalen Merkmalen gegenübergestellt.

Fächerübergreifend liegen im IQB-Bildungstrend 2016 Einschätzungen der Eltern zur Lern- und Leistungsbereitschaft ihrer Kinder vor (vgl. Bos et al., 2010). Eltern gaben zu fünf Aussagen (z. B. „Mein Kind lernt leicht und braucht wenig Hilfe.“) ihre Einschätzung auf einer vierstufigen Skala an (von „stimme nicht zu“ bis „stimme völlig zu“). Die Angaben zu fachspezifischen motivationalen Schülermerkmalen stammen aus dem Schülerfragebogen. Die Kinder gaben jeweils auf einer vierstufigen Skala von „stimmt gar nicht“ bis „stimmt genau“ ihre Zustimmung zu Aussagen über das Fach Deutsch beziehungsweise Mathematik an. Im Folgenden werden drei motivationale Aspekte mit Hilfe von Fragebogenskalen untersucht, die sich an Befragungen aus anderen *Large-Scale-Assessments* orientieren: das fachbezogene akademische Selbstkonzept (vgl. Richter, Böhme, Bastian-Wurzel, Pant & Stanat, 2014), das fachbezogene Interesse und die fachspezifische Langeweile (vgl. Lenski et al., 2016; Pekrun, Goetz & Frenzel, 2005). Für jedes Konstrukt wurden drei bis vier Aussagen zur Bewertung vorgelegt, zum Beispiel „Ich bin gut in Deutsch.“ (akademisches Selbstkonzept), „Ich interessiere mich für Deutsch.“ (fachbezogenes Interesse) oder „Im Deutschunterricht langweile ich mich sehr.“ (fachspezifische Langeweile).

Die Ergebnisse zu den motivationalen Merkmalen sind für die vier untersuchten Gruppen in Tabelle 11.3 dargestellt. Es zeigt sich, dass Eltern leistungsstarker Schülerinnen und Schüler ihre Kinder in allen Gruppen durchschnittlich als lern- und leistungsbereiter beschreiben als Eltern von Kindern der Vergleichsgruppe. Die Effektstärken⁴ von $d = 0.55$ bis $d = 0.89$ deuten in der Höhe auf einen mittleren bis großen Effekt und somit eine deutlich ausgeprägtere Lern- und Leistungsbereitschaft bei den leistungsstarken Kindern hin.

Die fachspezifischen Ergebnisse zeigen ein differenziertes Bild. So korrespondiert das fachbezogene Selbstkonzept deutlich mit fachlicher Leistungsstärke. Kinder, die in Deutsch leistungsstark sind, berichten ein höheres fachbezogenes Selbstkonzept im Fach Deutsch ($d = 0.57$) als Kinder der Vergleichsgruppe. Kinder, die im Fach Mathematik leistungsstark sind, weisen ein höheres Selbstkonzept im Fach Mathematik auf als Kinder der Vergleichsgruppe ($d = 0.76$). Hingegen ist das Selbstkonzept im Fach Deutsch bei Kindern, die ausschließlich im Fach Mathematik leistungsstark sind, sowie das Selbstkonzept im Fach Mathematik bei Kindern, die ausschließlich im Fach Deutsch leistungsstark sind, nur geringfügig positiver ausgeprägt als in der Vergleichsgruppe ($d = 0.20$ bzw. $d = 0.21$). Kinder, die in beiden Fächern als leistungsstark einzustufen sind, berichten auch in beiden Fächern ein höheres Selbstkonzept als Kinder der Vergleichsgruppe ($d = 0.60$ bzw. $d = 0.74$). Die Selbsteinschätzungen der leistungsstarken Kinder scheinen also insgesamt recht realistisch zu sein.

4 Zur Erklärung des Kennwerts Cohens d wird auf das Kapitel 3.2 verwiesen.

Tabelle 11.3: Motivationale Merkmale von Schülerinnen und Schülern der vier Leistungsgruppen

Motivationale Merkmale	Vergleichsgruppe			Leistungsstark in Deutsch				Leistungsstark in Mathematik				Leistungsstark in Deutsch und Mathematik			
	M	(SE)	SD	M	(SE)	SD	d	M	(SE)	SD	d	M	(SE)	SD	d
überfachlich															
Lern- und Leistungsbereitschaft (Elterneinschätzung)	2.80	(0.01)	0.68	3.21	(0.03)	0.60	0.61	3.17	(0.03)	0.59	0.55	3.39	(0.02)	0.53	0.89
Deutsch															
Selbstkonzept	3.11	(0.01)	0.65	3.47	(0.01)	0.49	0.57	3.24	(0.02)	0.58	0.20	3.49	(0.02)	0.49	0.60
Interesse	2.94	(0.01)	0.76	3.07	(0.02)	0.70	0.17	2.72	(0.03)	0.76	-0.29	2.91	(0.04)	0.73	-0.05
Langeweile	1.71	(0.01)	0.78	1.52	(0.02)	0.63	-0.25	1.73	(0.03)	0.74	0.03	1.60	(0.03)	0.66	-0.13
Mathematik															
Selbstkonzept	3.09	(0.01)	0.80	3.25	(0.02)	0.69	0.21	3.69	(0.01)	0.48	0.76	3.67	(0.03)	0.48	0.74
Interesse	3.13	(0.01)	0.84	3.09	(0.03)	0.83	-0.05	3.55	(0.02)	0.63	0.50	3.46	(0.03)	0.67	0.40
Langeweile	1.80	(0.01)	0.81	1.66	(0.02)	0.68	-0.19	1.63	(0.03)	0.70	-0.21	1.63	(0.03)	0.67	-0.22

Anmerkungen. M = Mittelwert; SE = Standardfehler; SD = Standardabweichung; d = Effektstärke Cohens d. Fett gedruckte Werte zeigen eine statistisch signifikante Abweichung ($p < .05$) von der Vergleichsgruppe an.

Auch im fachbezogenen Interesse spiegeln sich die fachlichen Stärken der Kinder in Deutsch beziehungsweise Mathematik wider. Die Unterschiede sind aber weniger ausgeprägt als beim Selbstkonzept. So zeigen Schülerinnen und Schüler, die nur im Fach Deutsch leistungsstark sind, ein etwas höheres Interesse am Fach Deutsch ($d = 0.17$), aber ein ähnliches Interesse am Fach Mathematik wie Kinder der Vergleichsgruppe. Kinder, die nur im Fach Mathematik leistungsstark sind, weisen ein geringeres Interesse am Fach Deutsch ($d = -0.29$), aber ein höheres Interesse am Fach Mathematik ($d = 0.50$) auf als die Vergleichsgruppe. Schülerinnen und Schüler, die in beiden Fächern leistungsstark sind, interessieren sich im Mittel für das Fach Deutsch genauso stark wie die Vergleichsgruppe, haben aber ein höheres Interesse am Fach Mathematik ($d = 0.40$). Die Unterschiede im Interesse zwischen den betrachteten Gruppen sind im Fach Mathematik damit größer als im Fach Deutsch.

Die Ergebnisse zur fachbezogenen Langeweile verdeutlichen, dass leistungsstarke Schülerinnen und Schüler durchgängig eine etwas geringere Ausprägung von Langeweile berichten als Kinder der Vergleichsgruppe ($-0.13 \leq d \leq -0.25$). Eine Ausnahme bilden hier die Leistungsstarken im Fach Mathematik, die sich bei der empfundenen Langeweile im Fach Deutsch nicht von den Kindern der Vergleichsgruppe unterscheiden. Entgegen mancher Befürchtungen (z. B. Phillips & Lindsay, 2006) scheinen leistungsstarke Viertklässlerinnen und Viertklässler nicht überdurchschnittlich gelangweilt zu sein. Im Gegenteil zeigen sie sogar geringere Ausprägungen von Langeweile als die Schülerinnen und Schüler der Vergleichsgruppe.

11.4 Schulische Umwelt

Während der vorangegangene Abschnitt der Frage nachging, welche Merkmale leistungsstarke Schülerinnen und Schüler kennzeichnen, soll im Folgenden das Augenmerk auf die schulische Umwelt leistungsstarker Kinder gerichtet werden.

Die Förderung leistungsstarker Schülerinnen und Schüler kann einerseits innerhalb des Unterrichts stattfinden, indem Kinder unterschiedlicher Leistungsfähigkeit beispielsweise differenzierte Angebote erhalten. Ein solcher differenzierter, adaptiver Unterricht stellt die Basis einer angemessenen Förderung aller Schülerinnen und Schüler dar; auch der leistungsstarken. Neben Maßnahmen der Unterrichtsgestaltung gibt es zusätzliche schulische und außerschulische Förderangebote, die auf den Prinzipien der Akzeleration, des Enrichments und der Gruppierung basieren (Vock & Gronostaj, 2017). *Akzelerationsmaßnahmen* ermöglichen es leistungsstarken Schülerinnen und Schülern, die Schule schneller zu durchlaufen. Dazu gehören eine vorzeitige Einschulung, eine kürzere Verweildauer in der Eingangsphase oder das Überspringen von Jahrgangsstufen. Eine weitere Möglichkeit besteht darin, Teile des Unterrichts in höheren Jahrgangsstufen zu besuchen, wenn beispielsweise eine besondere Stärke in einem bestimmten Fach vorliegt. Mit *Enrichment* wird die Erweiterung des Lernangebots durch eine Anreicherung oder Vertiefung der Lerninhalte bezeichnet, was innerhalb und außerhalb des Unterrichts, beispielsweise in Form von schulischen, aber auch außerschulischen Zusatzangeboten erfolgen kann (BMBF, 2015; KMK, 2015). Unter *Gruppierung* werden Maßnahmen zusammengefasst, bei denen leistungsstarke Schülerinnen und Schüler in leistungshomogenen Gruppen gefördert werden. Bei diesen Prinzipien handelt es sich nicht um trennscharfe Kategorien, zumal Fördermaßnahmen meist auf einer Kombination dieser Prinzipien beruhen (Vock & Gronostaj, 2017).

Im Folgenden soll auf Grundlage der Daten des IQB-Bildungstrends 2016 dargestellt werden, welche außerunterrichtlichen Förderangebote für leistungsstarke Schülerinnen und Schüler an Schulen im Primarbereich in Deutschland bereitgestellt werden und inwieweit diese durch die Schülerinnen und Schüler auch wahrgenommen werden. Mit diesem Fokus auf außerunterrichtliche schulische Angebote werden Gestaltungsmerkmale des Unterrichts somit im Rahmen dieses Kapitels explizit nicht in den Blick genommen.

11.4.1 Außerunterrichtliche schulische Förderangebote für leistungsstarke Schülerinnen und Schüler

Zur Erfassung schulischer Fördermaßnahmen für leistungsstarke Kinder wurden Schulleiterinnen und Schulleiter an allgemeinen Schulen⁵ im Fragebogen um Auskunft gebeten, welche außerunterrichtlichen Zusatzangebote zur fachbezogenen Förderung leistungsstarker Schülerinnen und Schüler an ihrer Schule existieren. Der Fragebogen deckte dabei verschiedene Maßnahmen ab, die hier unter den oben erläuterten Begriffen Akzeleration und Enrichment zusammengefasst werden (vgl. Tab. 11.4). Bei den dargestellten Kategorien handelt es sich um eine Auswahl an Zusatzangeboten, die speziell auf leistungsstarke Kinder abzielen.

5 An Förderschulen wurden den Schulleitungen andere Fragen gestellt, sodass sich dieser Abschnitt ausschließlich auf allgemeine Schulen bezieht.

Tabelle 11.4: Außerunterrichtliche Fördermaßnahmen für Leistungsstarke – Angebote der Schulen

Maßnahmen zur Förderung Leistungsstarker	Anteil an Schulen, an denen das Angebot vorliegt		
	%	(SE)	% k.A.
Akzeleration			
jahrgangsübergreifende Schuleingangsphase	16.8	(1.1)	0.7
Überspringen einer Jahrgangsstufe ¹	78.0	(1.2)	8.8
Teile des Unterrichts in höheren Klassenstufen ¹	53.7	(1.4)	17.2
Enrichment (innerschulisch)			
außerunterrichtliche Angebote für Leistungsstarke in Deutsch ²	24.6	(1.3)	6.1
außerunterrichtliche Angebote für Leistungsstarke in Mathematik ²	29.2	(1.4)	6.3
Enrichment (außerschulisch)			
Teilnahme an Wettbewerben und Olympiaden ¹	75.6	(1.0)	9.5
Anregung zur Teilnahme an außerschulischen Bildungsangeboten ¹	80.3	(1.0)	6.5

Anmerkungen. Die Berechnungen beziehen sich ausschließlich auf allgemeine Schulen. SE = Standardfehler; k.A. = keine Angabe.

¹ Instruktion: „Bitte geben Sie an, welche der folgenden Maßnahmen Ihre Schule zur Förderung von Schülerinnen und Schülern ergreift, die besonders leistungsstark sind oder ein hohes Potenzial erkennen lassen.“ Antwortmöglichkeiten: „ja“/„nein“.

² Instruktion: „Bitte geben Sie an, ob an Ihrer Schule für die Jahrgangsstufe 4 die folgenden außerunterrichtlichen Angebote bestehen und falls ja, wie häufig sie stattfinden.“ „gar nicht“ wurde hier als „nein“, alle übrigen Antwortmöglichkeiten („unregelmäßig bis einmal im Monat“/„mehrmals im Monat bis einmal pro Woche“/„2-3 mal pro Woche“/„4-5 mal pro Woche“) als „ja“ kodiert.

In Tabelle 11.4 ist dargestellt, wie hoch der Anteil der allgemeinen Schulen im Primarbereich ist, an denen das jeweilige Förderangebot vorliegt. Für die Akzelerationsmaßnahmen zeigt sich, dass ein Großteil der Schulen leistungsstarken Schülerinnen und Schülern die Möglichkeit bietet, Jahrgangsstufen zu überspringen. Etwa die Hälfte der Schulleitungen gab zudem an, dass es an ihrer Schule die Möglichkeit gibt, Teile des Unterrichts in höheren Jahrgangsstufen zu besuchen. Eine jahrgangsübergreifende Schuleingangsphase, die eine höhere Flexibilität bezüglich der Verweildauer in der Primarstufe bietet, gibt es an 16.8 Prozent der Schulen.

Auf die Frage nach einer ganzen Reihe außerunterrichtlicher schulischer Zusatzangebote (Enrichment) gaben 24.6 beziehungsweise 29.2 Prozent der Schulleitungen an, an ihren Schulen außerunterrichtliche fachliche Angebote explizit für leistungsstarke Kinder im Fach Deutsch beziehungsweise Mathematik anzubieten. Deutlich häufiger gibt es an den Schulen fachliche Angebote für Kinder mit Schwierigkeiten in den Fächern Mathematik (65.5 %) oder Deutsch (76.3 %) (ohne Tabelle). Der Fokus von außerunterrichtlichen Zusatzangeboten liegt demnach stärker auf der Förderung vergleichsweise schwacher Schülerinnen und Schüler.

Zur Förderung leistungsstarker und potenziell leistungsfähiger Schülerinnen und Schüler gab ein großer Teil der Schulleitungen an, die Teilnahme an außerschulischen Förderangeboten (wie beispielsweise Wettbewerbe, Olympiaden oder Kinder-Unis) anzuregen und zu ermöglichen. Dies ist in etwa 80 Prozent der teilnehmenden Schulen der Fall.

Die hier dargestellten Analysen zeigen, dass nur relativ wenige Schulen im Primarbereich fachspezifische außerunterrichtliche Zusatzangebote zur Förderung Leistungsstarker machen. Zentrale Maßnahmen der Schulen sind hingegen das Überspringen von Jahrgangsstufen sowie der Hinweis auf außerschulische Fördermöglichkeiten. Insgesamt ist festzuhalten, dass nur knapp 15 Prozent der

Schulen die hier beispielhaft aufgeführten Enrichmentangebote zur Förderung leistungsstarker Kinder noch nicht aufgreifen.

11.4.2 Nutzung außerunterrichtlicher Förderangebote durch Leistungsstarke

Um zu bestimmen, inwieweit die Förderangebote der Schulen für leistungsstarke Schülerinnen und Schüler von den Kindern in Anspruch genommen werden, wurden im Rahmen des IQB-Bildungstrends 2016 die Eltern dazu befragt, ob ihre Kinder in ihrer bisherigen Schullaufbahn Akzelerationsmaßnahmen genutzt haben und ob sie aktuell in der 4. Jahrgangsstufe außerunterrichtliche oder außerschulische Maßnahmen zur Leistungsförderung in Anspruch nehmen. Hierbei wurden alle Angaben der Eltern berücksichtigt, auch wenn die Teilnahme nur unregelmäßig bis einmal pro Monat erfolgte, denn gerade auch Veranstaltungen wie Schülerakademien oder Ferienkurse, die nicht wöchentlich stattfinden, scheinen einen positiven Effekt zur Förderung leistungsstarker Kinder und Jugendlicher zu haben (Kim, 2016). Da die elterlichen Angaben viele fehlende Werte aufweisen, werden diese in Tabelle 11.5 gesondert ausgewiesen.

Anhand von Tabelle 11.5 wird deutlich, dass Kinder, die in der 4. Jahrgangsstufe leistungsstark sind, seltener frühzeitig eingeschult wurden und auch die Primarstufe im Mittel nicht schneller durchlaufen haben als die Kinder der Vergleichsgruppe. Dieses Ergebnis scheint zunächst zu überraschen. Allerdings sind die frühzeitige Einschulung oder das Überspringen von Jahrgangsstufen Maßnahmen, die in der Regel Unterforderung und damit verbundene negative Entwicklungen verhindern sollen. Diese Maßnahmen führen jedoch nicht bei allen Kindern dazu, dass sie in ihrer neuen Vergleichsgruppe zu den leistungs-

Tabelle 11.5: Außerunterrichtliche Fördermaßnahmen für Leistungsstarke – Nutzung durch Schülerinnen und Schüler

Maßnahmen für Leistungsstarke	Vergleichsgruppe			Leistungsstarke in Deutsch			Leistungsstarke in Mathematik			Leistungsstarke in Deutsch und Mathematik		
	%	(SE)	% k.A.	%	(SE)	% k.A.	%	(SE)	% k.A.	%	(SE)	% k.A.
Akzeleration												
vorzeitige Einschulung	14.6	(0.6)	29.7	8.9	(1.1)	17.1	8.5	(1.3)	16.3	6.6	(0.9)	12.7
akzeleriert ¹	2.6	(0.2)	26.9	1.8	(0.5)	15.9	2.2	(0.7)	14.2	1.7	(0.5)	11.3
Enrichment (schulische Förderung)												
erhält außerunterrichtliche Förderung für Leistungsstarke in Deutsch	7.4	(0.4)	31.3	6.0	(1.2)	19.0	5.2	(1.0)	17.8	6.7	(1.3)	13.9
erhält außerunterrichtliche Förderung für Leistungsstarke in Mathematik	7.1	(0.4)	31.8	7.5	(1.2)	19.4	12.8	(1.8)	17.4	14.2	(1.7)	13.5
Enrichment (außerschulische Förderung)												
erhält außerschulische Förderung in Deutsch	6.3	(0.3)	31.8	4.6	(0.9)	20.6	6.3	(1.2)	18.0	8.4	(1.4)	16.1
erhält außerschulische Förderung in Mathematik	6.0	(0.3)	32.0	5.1	(0.8)	20.6	6.1	(1.1)	18.1	6.6	(1.3)	16.0

Anmerkungen. SE = Standardfehler. k.A. = keine Angabe. ¹ Einschulung nach 2012.

Fett gedruckte Prozentwerte weichen statistisch signifikant ($p < .05$) von der Vergleichsgruppe ab.

stärksten gehören; Klassenwiederholungen sind in der Folge recht häufig (vgl. Gronostaj & Vock, 2014; Vock, Penk & Köller, 2013). Insgesamt werden Akzelerationsmaßnahmen, wie das Überspringen von Klassenstufen, relativ zurückhaltend eingesetzt (Rambo & McCoach, 2012).

Betrachtet man die Nutzung von Fördermaßnahmen für die hier gebildeten Gruppen der Leistungsstarken sowie für die Vergleichsgruppe, so zeigt sich zunächst, dass Viertklässlerinnen und Viertklässler die schulischen und außerschulischen Fördermaßnahmen insgesamt eher selten wahrnehmen. In allen Gruppen geben jeweils weniger als 10 Prozent der Eltern an, dass ihre Kinder schulische beziehungsweise außerschulische Förderung für Leistungsstarke erhalten. Eine Ausnahme bildet das Fach Mathematik. Hier nehmen insbesondere die in diesem Fach leistungsstarken Kinder an entsprechenden schulischen Angeboten teil. Das gilt sowohl für Schülerinnen und Schüler, die nur im Fach Mathematik den Optimalstandard erreichen (12.8 % Teilnahme an schulischen Fördermaßnahmen für Leistungsstarke im Fach Mathematik), als auch für Schülerinnen und Schüler, die in den Fächern Mathematik und Deutsch leistungsstark sind (14.2 % Teilnahme an schulischen Fördermaßnahmen für Leistungsstarke im Fach Mathematik). Während die Förderangebote im Fach Mathematik vor allem die in Mathematik leistungsstarken Kinder erreichen, nehmen an den Förderangeboten im Fach Deutsch fast alle Gruppen gleichermaßen teil. Eine Ausnahme hiervon bilden die Kinder, die nur in Mathematik leistungsstark sind. Sie besuchen Förderangebote im Fach Deutsch signifikant seltener als Kinder, die in keinem der beiden Fächer leistungsstark sind. Besuche außerschulischer Fördermaßnahmen sind insgesamt vergleichsweise selten. In der Häufigkeit des Besuchs außerschulischer Förderangebote für Leistungsstarke unterscheiden sich die leistungsstarken Schülerinnen und Schüler im IQB-Bildungstrend 2016 nicht signifikant von den Kindern der Vergleichsgruppe.

Insgesamt nehmen nach Aussagen der Eltern nur wenige leistungsstarke Schülerinnen und Schüler die hier untersuchten außerunterrichtlichen Fördermaßnahmen wahr. Diese Ergebnisse decken sich auch mit den Resultaten der Studie StEG-P, die außerunterrichtliche Angebote von Ganztagsgrundschulen untersucht hat (StEG-Konsortium, 2016). Es sei allerdings nochmals darauf hingewiesen, dass die untersuchten Fördermaßnahmen nur einen Teil möglicher Angebote abbilden und insbesondere die Förderung leistungsstarker Schülerinnen und Schüler direkt im schulischen Unterricht ein wichtiger Bereich ist, der an dieser Stelle nicht untersucht werden konnte.

11.5 Fazit

Mit dem vorliegenden Kapitel wurde erstmals im Rahmen der IQB-Ländervergleiche beziehungsweise IQB-Bildungstrends ein vertiefender Blick auf leistungsstarke Schülerinnen und Schüler gerichtet. Dabei wurde zunächst die Gruppe der leistungsstarken Schülerinnen und Schüler fächerübergreifend definiert. Anschließend wurden die leistungsstarken Viertklässlerinnen und Viertklässler im IQB-Bildungstrend 2016 mit Bezug auf verschiedene Hintergrundmerkmale charakterisiert. Schließlich wurden das Angebot und die Nutzung von Maßnahmen der außerunterrichtlichen Förderung leistungsstarker Schülerinnen und Schüler untersucht.

Dabei wird deutlich, dass 22.5 Prozent der Viertklässlerinnen und Viertklässler leistungsstark im Sinne des Erreichens des Optimalstandards in den Fächern

Deutsch und/oder Mathematik sind. Der Anteil der leistungsstarken Kinder ist im Fach Deutsch höher als im Fach Mathematik. Etwa ein Drittel der leistungsstarken Kinder zeigt besonders gute Leistungen in beiden Fächern.

In Übereinstimmung mit früheren Studien bestätigen die Analysen, dass schulische Leistungsstärke in einem engen Zusammenhang mit persönlichen und familiären Merkmalen steht. Bereits in der 4. Jahrgangsstufe sind es häufiger Mädchen, die im sprachlichen Bereich leistungsstark sind, und häufiger Jungen, die im Fach Mathematik besonders hohe Leistungen erzielen. In der Gruppe der Kinder, die in beiden Fächern zugleich leistungsstark sind, sind Mädchen und Jungen in etwa gleich stark vertreten. Dies ist ein erfreulicher Befund, der etwa ein Drittel der leistungsstarken Schülerinnen und Schüler betrifft. Leistungsstarke Kinder kommen überwiegend aus Familien mit einem hohen sozialen Status. Aber auch in den niedrigen sozialen Statusgruppen ist ein Teil der Kinder der Gruppe der Leistungsstarken im Sinne der hier verwendeten Definition zuzuordnen. Jeweils ein Drittel der leistungsstarken Kinder im Fach Deutsch oder Mathematik haben Eltern, die selbst kein Abitur erworben haben; bei den Kindern, die in beiden Fächern leistungsstark sind, ist es etwa ein Viertel. Diese Kinder auf ihrer Bildungslaufbahn weiterhin optimal zu fördern, wird eine wichtige Aufgabe sein, um einerseits jedem Kind angemessene Entwicklungschancen zu bieten und andererseits die Kopplung zwischen Herkunft und schulischem Erfolg zu verringern.

Für die motivationalen Merkmale zeichnet sich insgesamt ein positives Bild ab. Es zeigt sich, dass leistungsstarke Kinder von ihren Eltern als deutlich leistungsbereiter eingeschätzt werden als Kinder, die in keinem der untersuchten Fächer leistungsstark sind. Die leistungsstarken Viertklässlerinnen und Viertklässler weisen ein positives schulisches Selbstkonzept auf. Dieses ist in demjenigen Fach besonders ausgeprägt, in dem ihre Stärke liegt, aber in geringerem Maße auch in dem jeweils anderen Fach. Sie zeichnen sich durchgängig durch ein etwas geringeres Gefühl der Langeweile im Fachunterricht aus und zeigen ein fachspezifisch höheres Interesse an dem Fach, in dem ihre Stärke liegt.

Die Förderung leistungsstarker Schülerinnen und Schüler steht neben der Sicherung von Mindeststandards und der Inklusion von Kindern mit sonderpädagogischem Förderbedarf im aktuellen Fokus der Bildungspolitik. Derzeit fällt jedoch die außerunterrichtliche schulische Förderung Leistungsstarker noch weit hinter die außerunterrichtliche Förderung von leistungsschwächeren Schülerinnen und Schülern zurück. Außerunterrichtliche Maßnahmen, die fachbezogen sind und sich an Leistungsstärke richten, werden zwar relativ häufig von den Schulen im Primarbereich angeboten, aber nur von einem kleinen Teil der leistungsstarken Schülerinnen und Schüler genutzt. Eine häufigere Nutzung außerunterrichtlicher Förderangebote durch leistungsstarke Kinder im Vergleich zu Schülerinnen und Schülern ohne besondere Leistungsstärken in den hier betrachteten Fächern zeigt sich lediglich im Fach Mathematik.

Der große und sicherlich äußerst relevante Bereich der Förderung leistungsstarker Schülerinnen und Schüler innerhalb des Unterrichts wurde in den vorliegenden Analysen struktureller Maßnahmen an Primarschulen in Deutschland ausgeklammert. Wie auch in den aktuellen Strategien und Rahmenprogrammen gefordert (BMBF, 2017; KMK, 2015; 2016), sollte es daher Ziel weiterer Forschungsvorhaben sein, die Untersuchung der Gelingensbedingungen guten Unterrichts für heterogene Lerngruppen zu vertiefen.

Literatur

- Baumert, J., Roeder, P. M., Sang, F. & Schmitz, B. (1986). Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen. *Zeitschrift für Pädagogik*, 32, 639–660.
- Baye, A. & Monseur, C. (2016). Gender differences in variability and extreme scores in an international context. *Large-scale Assessments in Education*, 4, 1.
- BMBF (2015) = Bundesministerium für Bildung und Forschung. (2015). Begabte Kinder finden und fördern. *Ein Wegweiser für Eltern, Erzieherinnen und Erzieher, Lehrerinnen und Lehrer*. Berlin: BMBF. Zugriff am 27.07.2017 unter: https://www.bmbf.de/pub/Begabte_Kinder_finden_und_foerdern.pdf
- BMBF (2017) = Bundesministerium für Bildung und Forschung. (2017). *Rahmenprogramm empirische Bildungsforschung des Bundesministeriums für Bildung und Forschung*. Zugriff am 27.07.2017 unter <http://www.empirische-bildungsforschung-bmbf.de/media/content/Rahmenprogramm%20empirische%20Bildungsforschung%20%28BITV%29.pdf>
- Bos, W., Wendt, H., Ünlü, A., Valtin, R., Euen, B., Kasper, D. & Tarelli, I. (2012). Leistungsprofile von Viertklässlerinnen und Viertklässlern in Deutschland. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 227–259). Münster: Waxmann.
- Bos, W., Strieholt, R., Goy, M., Stubbe, T. C., Tarelli, I. & Hornberg, S. (2010). *IGLU 2006. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Clausen, M. (2002). *Unterrichtsqualität: eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 29). Münster: Waxmann.
- Dai, D. Y., Moon, S. M. & Feldhusen, J. F. (1998). Achievement motivation and gifted students: A social cognitive perspective. *Educational Psychologist*, 33, 45–63.
- Erikson, R., & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *The 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. *The British Journal of Sociology*, 30, 415–441.
- Gronostaj, A. & Vock, M. (2014). Effekte schulischer Begabtenförderung. In M. Stamm (Hrsg.), *Handbuch Talententwicklung* (S. 427–440). Bern: Verlag Hans Huber.
- Hammer, S., Reiss, K., Lehner, M. C., Heine, J.-H., Sälzer, C. & Heinze, A. (2016). Mathematische Kompetenz in PISA 2015. Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 219–248). Münster: Waxmann.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Bd. 3, S. 71–176). Göttingen: Hogrefe.
- Jansen, M., Schroeders, U. & Stanat, P. (2013). Motivationale Schülermerkmale in Mathematik und den Naturwissenschaften. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 347–365). Münster: Waxmann.
- Kim, M. (2016). A meta-analysis of the effects of enrichment programs on gifted students. *Gifted Child Quarterly*, 60, 102–116.
- Kirsch, I., Jong, J. de, McQueen, J., Mendelovits, J. & Monseur, C. (2002). *Lesen kann die Welt verändern. Leistung und Engagement im Ländervergleich. Ergebnisse von PISA 2000*. Paris: OECD.
- Kiuru, N., Nurmi, J.-E., Leskinen, E., Torppa, M., Poikkeus, A.-M., Lerkkanen, M.-K. & Niemi, P. (2015). Elementary school teachers adapt their instructional support according to students' academic skills: A variable and person-oriented approach. *International Journal of Behavioral Development*, 39, 391–401.
- KMK (2005) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich (Jahrgangsstufe 4). Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Förderstrategie für leistungsstarke Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 11.06.2015*. Berlin: KMK.

- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 10.11.2016*. Berlin: KMK.
- Köller, O., & Baumert, J. (2017). Hochleistende Schülerinnen und Schüler im mehr- und zweigliedrigen System. In M. Neumann, M. Becker, J. Baumert, K. Maaz & O. Köller (Hrsg.), *Zweigliedrigkeit im deutschen Schulsystem – Potenziale und Herausforderungen in Berlin* (S. 227–254). Münster: Waxmann.
- Lenski, A. E., Hecht, M., Penk, C., Milles, F., Mezger, M., Heitmann, P., Stanat, P. & Pant, H. A. (2016). *IQB-Ländervergleich 2012. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Berlin: Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen.
- Marsh, H. W. & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin*, 34(4), 542–552.
- Pekrun, R., Goetz, T. & Frenzel, A. C. (2005). *Achievement Emotions Questionnaire – Mathematics (AEQ-M). User's manual*. Department of Psychology, University of Munich.
- Penner, A. M. (2008). Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *American Journal of Sociology*, 114, 138–170.
- Phillips, N. & Lindsay, G. (2006). Motivation in gifted students. *High Ability Studies*, 17(1), 57–73.
- Rambo, K. E. & McCoach, D. B. (2012). Teacher attitudes toward subject-specific acceleration. *Journal for the Education of the Gifted*, 35, 129–152.
- Richter, D., Böhme, K., Bastian-Wurzel, J., Pant, H. A. & Stanat, P. (2014). *IQB-Ländervergleich 2011. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Berlin: Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen.
- Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten des Kompetenzerwerbs. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011* (S. 191–208). Münster: Waxmann.
- Stubbe, T. C., Schwippert, K. & Wendt, H. (2016). Soziale Disparitäten der Schülerleistungen in Mathematik und Naturwissenschaften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 299–316). Münster: Waxmann.
- StEG-Konsortium (2016). *Ganztagsschule: Bildungsqualität und Wirkungen außerunterrichtlicher Angebote. Ergebnisse der Studie zur Entwicklung von Ganztagsschulen 2012–2015*. Frankfurt am Main: Projekt StEG.
- Vock, M. & Gronostaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht* (Schriftenreihe des Netzwerk Bildung, Bd. 40). Berlin: Friedrich-Ebert-Stiftung, Abt. Studienförderung.
- Vock, M., Penk, C. & Köller, O. (2013). Wer überspringt eine Schulklasse? Befunde zum Klassenüberspringen in Deutschland. *Psychologie in Erziehung und Unterricht*, 61, 153–164. doi: 10.2378/peu2013.art22d
- Wendt, H., Willems, A. S., Tarelli, I., Euen, B. & Bos, W. (2013). Ausreichend geförderte Talente? Zu den deutschen Ergebnissen von leistungsstarken Viertklässlerinnen und Viertklässlern in IGLU 2011 und TIMSS 2011. In C. Fischer (Hrsg.), *Schule und Unterricht adaptiv gestalten. Fördermöglichkeiten für benachteiligte Kinder und Jugendliche* (Münstersche Gespräche zur Pädagogik, Bd. 29, S. 23–34). Münster: Waxmann.
- Wendt, H., Stubbe, T. C. & Schwippert, K. (2012). Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 175–190). Münster: Waxmann.

Kapitel 12

Qualifikation von Lehrkräften und Einschätzungen zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf

Camilla Rjosk, Lars Hoffmann, Dirk Richter, Alexandra Marx und Cornelia Gresch

12.1 Einleitung

Lehrkräfte spielen für die Kompetenzentwicklung von Schülerinnen und Schülern eine zentrale Rolle (vgl. z. B. Baumert et al., 2010; Hattie, 2009, 2012). Daher werden in den Studien des IQB zum Bildungsmonitoring immer auch Befragungen von Lehrerinnen und Lehrern durchgeführt und für beschreibende Analysen genutzt. Seit dem IQB-Ländervergleich 2011 beschäftigt sich ein Kapitel der Berichterlegung mit den Lehrkräften der jeweils teilnehmenden Schülerinnen und Schüler und charakterisiert diese anhand verschiedener Hintergrundmerkmale. Neben demografischen Merkmalen, wie etwa dem Geschlecht und dem Alter, wird insbesondere die berufliche Qualifikation der Lehrkräfte untersucht. Es wird der Frage nachgegangen, inwieweit Schülerinnen und Schüler von Lehrkräften, die das von ihnen unterrichtete Fach studiert haben beziehungsweise eine Lehrbefähigung für das Fach erworben haben, höhere Kompetenzen erreichen. Diese Analysen werden im IQB-Bildungstrend 2016 fortgeführt. Bezogen auf das Geschlecht und das Alter wird dabei auch untersucht, inwieweit sich die Zusammensetzung der Lehrerschaft zwischen den Jahren 2011 und 2016 verändert hat.^{1,2}

Zusätzlich zur Fortführung der Analysen aus früheren Studien des IQB wird im vorliegenden Kapitel das Thema „Umgang mit Heterogenität“ aufgegriffen, das einen Schwerpunkt des IQB-Bildungstrends 2016 bildet, wobei sich die Analysen vor allem auf Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) beziehen (vgl. Kapitel 10). Der gemeinsame Unterricht von Kindern und Jugendlichen mit SPF und Kindern und Jugendlichen ohne SPF (im Folgenden auch kurz als „gemeinsamer Unterricht“ bezeichnet) ist mit besonderen Herausforderungen für Lehrerinnen und Lehrer verbunden. Daher werden ausgewählte Ergebnisse der Lehrkräftebefragung dargestellt, die dieses Thema

-
- 1 Für das studierte Fach können keine Trendanalysen durchgeführt werden, da Merkmale der beruflichen Qualifikation im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 unterschiedlich erfasst wurden (s. u.).
 - 2 In die Lehrkräftebefragungen wurden alle Lehrerinnen und Lehrer einbezogen, die in den teilnehmenden Klassen zum Erhebungszeitpunkt Deutsch und/oder Mathematik unterrichteten. Dabei kam es vor, dass Schülerinnen und Schüler von mehreren Lehrkräften im Fach Deutsch oder Mathematik unterrichtet wurden oder Teilungsunterricht stattfand. Dies hat zur Folge, dass die Zielpopulation der einzubeziehenden Lehrkräfte nicht eindeutig definiert ist und über die Repräsentativität der Lehrkräftestichprobe im IQB-Bildungstrend 2016 und im IQB-Ländervergleich 2011 streng genommen keine Aussage gemacht werden kann. Daher werden die in diesem Kapitel untersuchten Merkmale der Lehrkräfte rein deskriptiv analysiert und es wird auf inferenzstatistische Prüfungen verzichtet. Die Daten bieten jedoch die Möglichkeit, Analysen anhand einer bundesweiten Lehrkräftestichprobe von per Zufall ausgewählten Schulen durchzuführen.

betreffen. Dies umfasst Qualifikationen im Bereich der Sonderpädagogik sowie Erfahrungen, wahrgenommenen Fortbildungsbedarf und Einschätzungen zum gemeinsamen Unterricht von Kindern mit SPF und Kindern ohne SPF.

Im Folgenden werden zunächst die demografischen Merkmale der befragten Lehrkräfte für die Fächer Deutsch und Mathematik auf Länderebene dargestellt (Abschnitt 12.2). Es folgt eine kurze Zusammenfassung des Forschungsstands zur beruflichen Qualifikation von Lehrkräften und deren Bedeutung für den Lernerfolg von Schülerinnen und Schülern sowie eine Beschreibung der Befunde des IQB-Bildungstrends 2016 zu dieser Fragestellung (Abschnitt 12.3). Anschließend geht das Kapitel auf Fragen zum gemeinsamen Unterricht ein. Es wird eine Reihe von Analysen präsentiert, die Anhaltspunkte darauf geben, welche Qualifikationen, Erfahrungen und Einschätzungen Lehrkräfte zu gemeinsamem Unterricht in der Primarstufe haben und wie diese miteinander zusammenhängen (Abschnitt 12.4). Das Kapitel schließt mit einer Zusammenfassung und Diskussion der Ergebnisse (Abschnitt 12.5).

12.2 Demografische Hintergrundmerkmale der Lehrkräfte

Die in diesem Kapitel dargestellten Analysen basieren auf einer schriftlichen Befragung von Lehrkräften, deren Klassen im IQB-Bildungstrend 2016 teilnahmen. Einbezogen werden Angaben von Lehrkräften an allgemeinen Schulen, die die Fächer Deutsch oder Mathematik in den getesteten Klassen unterrichteten. Insgesamt umfasst die Stichprobe 1784 Lehrkräfte, von denen 552 nur das Fach Deutsch, 517 nur das Fach Mathematik und 715 beide Fächer in den teilnehmenden Klassen unterrichteten. Angaben über die Teilnahmequoten der Lehrkräftebefragung in den Ländern sind in Kapitel 3.1 zu finden. Auch hier gilt, dass bei einem hohen Anteil fehlender Werte die Zuverlässigkeit der Analysen eingeschränkt sein kann. Da in Hamburg und im Saarland für einen hohen Anteil der Deutsch- beziehungsweise Mathematikklassen keine Angaben der Lehrkräfte vorlagen, sind die Ergebnisse dieser Länder mit besonders großer Unsicherheit behaftet und werden unter Vorbehalt berichtet. Im Saarland bezieht sich dies insbesondere auf die Angaben zu demografischen Hintergrundmerkmalen und zur Qualifikation der Lehrkräfte.³

12.2.1 Lehrkräfte im Fach Deutsch

Angaben zur Anzahl, zum Geschlecht und zum Alter der befragten Deutschlehrkräfte in den Jahren 2011⁴ und 2016 sind in Tabelle 12.1 dargestellt. Im Jahr 2016 liegt der Anteil weiblicher Lehrkräfte bei den am IQB-Bildungstrend teilneh-

3 In Hamburg lagen für 67 Prozent der Klassen im Fach Deutsch und für 65 Prozent der Klassen im Fach Mathematik Angaben der Lehrkräfte vor. Im Saarland lagen die entsprechenden Quoten bei 56 Prozent im Fach Deutsch und bei 52 Prozent im Fach Mathematik. Für weitere 39 Prozent der Klassen (Deutsch) beziehungsweise 40 Prozent (Mathematik) füllten die Lehrkräfte im Saarland zwar einen Fragebogen aus, gaben jedoch nicht explizit an, mit der Verwendung der personenbezogenen Angaben in der Auswertung einverstanden zu sein. Die zuletzt genannte Personengruppe wird deshalb nicht in die Darstellung demografischer Hintergrundmerkmale und Qualifikationsmerkmale einbezogen.

4 Die in diesem Kapitel berichteten Werte für Lehrkräfte im Jahr 2011 weichen von den im Bericht zum IQB-Ländervergleich 2011 angegebenen Werten leicht ab, da sich die Angaben im vorliegenden Kapitel ausschließlich auf Lehrkräfte an allgemeinen Schulen beziehen. Zudem werden die Verteilungen hier fachspezifisch dargestellt.

Tabelle 12.1: Demografische Angaben zu den Deutschlehrkräften der am IQB-Ländervergleich 2011 und am IQB-Bildungstrend 2016 teilnehmenden Klassen allgemeiner Schulen

Land	2011						2016							
	Geschlecht		Alter in Jahren (in %)				Geschlecht		Alter in Jahren (in %)					
	N	weiblich (in %)	bis 29	30–39	40–49	50–59	ab 60	N	weiblich (in %)	bis 29	30–39	40–49	50–59	ab 60
Baden-Württemberg	62	85.5	9.7	17.7	9.7	48.4	14.5	72	88.9	9.7	33.3	27.8	18.1	11.1
Bayern	72	87.5	7.0	19.7	31.0	32.4	9.9	72	88.9	6.9	22.2	38.9	22.2	9.7
Berlin	95	93.7	3.2	10.5	36.8	35.8	13.7	113	83.2	5.5	24.5	21.8	40.0	8.2
Brandenburg	68	94.1	3.0	14.9	38.8	41.8	1.5	65	95.4	3.1	18.5	26.2	47.7	4.6
Bremen	72	93.1	4.2	29.2	16.7	37.5	12.5	74	94.6	2.8	26.8	29.6	26.8	14.1
Hessen	73	95.9	13.7	20.5	20.5	32.9	12.3	75	90.7	9.5	20.3	29.7	24.3	16.2
Mecklenburg-Vorpommern	69	92.8	1.4	1.4	46.4	44.9	5.8	85	96.5	3.6	7.2	25.3	56.6	7.2
Niedersachsen	68	92.6	4.5	19.4	34.3	26.9	14.9	77	93.5	10.4	20.8	28.6	27.3	13.0
Nordrhein-Westfalen	73	93.2	11.1	25.0	16.7	34.7	12.5	88	95.5	8.0	18.2	30.7	29.5	13.6
Rheinland-Pfalz	71	87.3	12.9	27.1	22.9	30.0	7.1	84	90.5	19.0	26.2	34.5	13.1	7.1
Sachsen	63	93.7	6.5	8.1	24.2	50.0	11.3	74	97.3	4.1	12.2	32.4	32.4	18.9
Sachsen-Anhalt	72	95.8	0.0	2.9	37.7	52.2	7.2	66	98.5	4.6	9.2	21.5	53.8	10.8
Schleswig-Holstein	64	93.8	6.3	19.0	30.2	30.2	14.3	70	94.3	4.3	18.8	37.7	26.1	13.0
Thüringen	74	90.5	7.2	6.9	12.5	38.9	33.3	80	92.5	3.8	18.8	21.3	37.5	18.8
Hamburg ¹	79	92.4	78	3.8	21.8	37.2	15.4	79	92.4	2.5	31.6	30.4	21.5	13.9
Saarland ¹	67	88.1	65	20.0	33.8	9.2	18.5	47	78.7	17.0	46.8	25.5	4.3	6.4
Deutschland	1 142	91.9	1 127	7.0	17.7	27.5	36.6	1 221	92.0	7.0	21.7	28.8	30.7	11.7

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente der Altersangaben pro Land minimal von 100 abweichen. N = Anzahl der Lehrkräfte mit gültigen Angaben.
¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

menden Deutschlehrkräften in der Primarstufe bundesweit bei etwa 92 Prozent. Der Frauenanteil ist in allen Ländern hoch und variiert zwischen 79 Prozent im Saarland und 99 Prozent in Sachsen-Anhalt. Die Geschlechterzusammensetzung der befragten Deutschlehrkräfte unterscheidet sich zwischen den Jahren 2011 und 2016 kaum.

Für die Altersverteilung der Deutschlehrkräfte im Jahr 2016 ergeben sich deutliche Länderunterschiede. Liegt der Anteil jüngerer Deutschlehrkräfte (unter 40 Jahren) im bundesweiten Durchschnitt bei 29 Prozent, so ist dieser Wert in nahezu allen westdeutschen Ländern höher, wobei die höchsten Anteile in Baden-Württemberg (43 %) und Rheinland-Pfalz (45 %) sowie im Saarland (64 %) zu finden sind. Am niedrigsten sind die Anteile jüngerer Lehrkräfte in den ostdeutschen Flächenländern, insbesondere in Mecklenburg-Vorpommern (11 %) und Sachsen-Anhalt (14 %). Dementsprechend hoch fällt in den ostdeutschen Ländern der Anteil älterer Deutschlehrkräfte (ab 50 Jahren) aus, wobei die höchsten Anteile mit jeweils rund 64 Prozent wiederum in Mecklenburg-Vorpommern und Sachsen-Anhalt zu finden sind. Am geringsten ist der Anteil älterer Lehrkräfte in Rheinland-Pfalz (20 %) sowie im Saarland (11 %).

Bundesweit betrachtet sind jüngere Deutschlehrkräfte in den Lehrkräftestichproben im Jahr 2016 etwas stärker vertreten als im Jahr 2011. Dies ist insbesondere in den Ländern Baden-Württemberg, Berlin, Sachsen-Anhalt sowie im Saarland der Fall.

12.2.2 Lehrkräfte im Fach Mathematik

Die Angaben zu Anzahl, Geschlecht und Alter der befragten Mathematiklehrkräfte, die der Tabelle 12.2 entnommen werden können, zeigen aufgrund der sich überschneidenden Stichproben erwartungsgemäß ein ähnliches Bild wie für die Deutschlehrkräfte. Der Anteil weiblicher Lehrkräfte ist im Fach Mathematik bundesweit mit 87 Prozent etwas geringer als im Fach Deutsch und variiert leicht zwischen den Ländern, wobei der geringste Anteil im Saarland (79 %) und der höchste Anteil in Mecklenburg-Vorpommern (97 %) vorliegt. Auch im Fach Mathematik ist die Geschlechterzusammensetzung teilnehmender Lehrkräfte in den Jahren 2011 und 2016 ähnlich.

Wie im Fach Deutsch ist auch im Fach Mathematik der Anteil jüngerer Lehrkräfte (unter 40 Jahren) in den westdeutschen Ländern deutlich höher als in den ostdeutschen Ländern. Besonders hohe Anteile liegen im Jahr 2016 mit jeweils rund 41 Prozent in Baden-Württemberg und Rheinland-Pfalz sowie im Saarland (58 %) vor. Die geringsten Anteile jüngerer Mathematiklehrkräfte sind in Brandenburg (9 %) und in Sachsen-Anhalt (8 %) zu finden. Entsprechend fällt der Anteil älterer Mathematiklehrkräfte (ab 50 Jahren) in den ostdeutschen Flächenländern deutlich höher aus als in den westdeutschen Ländern, wobei die höchsten Anteile für Brandenburg (71 %) und Sachsen-Anhalt (76 %), die niedrigsten Anteile für Rheinland-Pfalz (22 %) und für das Saarland (14 %) zu verzeichnen sind. Jüngere Lehrkräfte im Fach Mathematik sind bundesweit sowie vor allem in Baden-Württemberg, Berlin und Mecklenburg-Vorpommern im Jahr 2016 etwas stärker vertreten als im Jahr 2011.

Tabelle 12.2: Demografische Angaben zu den Mathematiklehrkräften der am IQB-Ländervergleich 2011 und am IQB-Bildungstrend 2016 teilnehmenden Klassen allgemeiner Schulen

Land	2011						2016								
	Geschlecht			Alter in Jahren (in %)			Geschlecht			Alter in Jahren (in %)					
	N	weiblich (in %)		N	bis 29	30–39	40–49	50–59	ab 60	N	bis 29	30–39	40–49	50–59	ab 60
Baden-Württemberg	66	78.8	6.1	21.2	7.6	47.0	18.2	63	82.5	63	7.9	33.3	27.0	17.5	14.3
Bayern	72	87.5	7.0	19.7	31.0	33.8	8.5	71	87.3	71	7.0	19.7	36.6	22.5	14.1
Berlin	96	82.3	2.1	8.5	26.6	43.6	19.1	107	84.1	105	7.6	24.8	22.9	27.6	17.1
Brandenburg	65	89.2	4.7	10.9	39.1	42.2	3.1	68	82.4	66	3.0	6.1	19.7	57.6	13.6
Bremen	75	88.0	5.4	21.6	25.7	33.8	13.5	78	83.3	74	6.8	23.0	27.0	20.3	23.0
Hessen	63	87.3	9.5	17.5	20.6	38.1	14.3	76	88.2	75	10.7	25.3	29.3	21.3	13.3
Mecklenburg-Vorpommern	66	97.0	0.0	1.5	47.0	47.0	4.5	85	96.5	83	6.0	10.8	22.9	49.4	10.8
Niedersachsen	68	79.4	7.4	20.6	26.5	27.9	17.6	76	84.2	75	6.7	25.3	26.7	26.7	14.7
Nordrhein-Westfalen	72	88.9	15.7	22.9	17.1	32.9	11.4	86	94.2	86	5.8	24.4	29.1	27.9	12.8
Rheinland-Pfalz	71	83.1	14.3	22.9	22.9	34.3	5.7	82	86.6	82	17.1	24.4	36.6	12.2	9.8
Sachsen	64	85.9	1.6	7.9	38.1	41.3	11.1	69	91.3	68	5.9	13.2	25.0	32.4	23.5
Sachsen-Anhalt	71	90.1	1.4	1.4	41.4	47.1	8.6	62	91.9	62	4.8	3.2	16.1	58.1	17.7
Schleswig-Holstein	64	87.5	6.6	26.2	16.4	41.0	9.8	66	87.9	65	4.6	9.2	36.9	23.1	26.2
Thüringen	73	86.3	5.6	9.7	30.6	44.4	9.7	74	87.8	74	1.4	16.2	23.0	43.2	16.2
Hamburg ¹	74	75.7	5.5	23.3	17.8	37.0	16.4	77	84.4	77	2.6	29.9	35.1	23.4	9.1
Saarland ¹	66	86.4	20.3	35.9	7.8	14.1	21.9	43	79.1	43	16.3	41.9	27.9	4.7	9.3
Deutschland	1 126	85.7	1 109	6.9	16.8	26.1	12.3	1 183	87.2	1 169	7.0	20.5	27.6	29.5	15.3

Anmerkungen. In der Tabelle werden gerundete Werte angegeben. Dadurch kann die Summe der Prozente der Altersangaben pro Land minimal von 100 abweichen. N = Anzahl der Lehrkräfte mit gültigen Angaben.
¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

12.3 Berufliche Qualifikation von Lehrkräften und erreichte Kompetenzen von Schülerinnen und Schülern

12.3.1 Forschungsstand

Als Indikator für die berufliche Qualifikation von Lehrkräften kann im deutschen Bildungssystem die sogenannte Lehrbefähigung herangezogen werden, die beschreibt, für welche Schulstufen und welche Fächer Lehrkräfte ausgebildet worden sind. Die Erteilung der Lehrbefähigung erfolgt in der Regel nach dem *Master of Education* beziehungsweise dem Staatsexamen und dem anschließenden Vorbereitungsdienst (KMK, 2012, 2013, 2017). Den aktuellen Rahmenvereinbarungen der KMK aus dem Jahr 2013 zufolge umfasst das Lehramtsstudium für den Primarbereich bildungswissenschaftliche, grundschulpädagogische und grundschuldidaktische Anteile sowie fachdidaktische und fachwissenschaftliche Anteile in den Fächern Deutsch, Mathematik und einem weiteren Lernbereich, bei dem es sich auch um eine sonderpädagogische Schwerpunktsetzung handeln kann.⁵ Eines dieser Fächer beziehungsweise einer dieser Lernbereiche wird als Schwerpunkt mit umfangreicherem Stundenumfang gewählt. Für beide Phasen der Lehramtsausbildung – Hochschulstudium und Vorbereitungsdienst – gelten die von der KMK verabschiedeten Standards für die Lehrerbildung in den Bildungswissenschaften (KMK, 2004) sowie die ländergemeinsamen Anforderungen für die Fachwissenschaften und Fachdidaktiken (KMK, 2008). Diese Vorgaben sollen dazu beitragen, dass Personen, die eine Lehrbefähigung erwerben, über die für die Gestaltung von Unterricht erforderlichen professionellen Kompetenzen verfügen. Die aktuellen Ausbildungsordnungen der Länder unterscheiden sich unter anderem darin, wie viele und welche Fächer für das Studium und den Vorbereitungsdienst vorgegeben sind und ob Lehrkräfte als sogenannte Generalisten mit Grundlagenwissen in allen oder mehreren Fächern oder eher als Spezialisten für einzelne Fächer ausgebildet werden (Porsch, 2017).

Aufgrund schulspezifischer Erfordernisse werden Lehrkräfte häufig auch in Fächern eingesetzt, in denen sie kein vertieftes Studium absolviert haben. In der Primarstufe ist darüber hinaus das Klassenlehrerprinzip verbreitet, bei dem die Klassenlehrkraft mehrere Fächer in ihrer Klasse unterrichtet (Schorch, 2009; vgl. Porsch, 2016). In der Praxis bedeutet dies, dass Lehrkräfte nicht nur das Fach unterrichten, das schwerpunktmäßig Gegenstand ihres Studiums und Vorbereitungsdienstes war, sondern auch Unterricht in anderen Fächern erteilen (fachfremdes Unterrichten). Der Begriff „fachfremdes Unterrichten“ wird hier analog zu bisherigen Ländervergleichs- und Bildungstrendstudien des IQB verwendet, die sich überwiegend auf die Sekundarstufe I bezogen. Als „fachfremd“ werden in diesem Kapitel alle Lehrerinnen und Lehrer bezeichnet, die Deutsch beziehungsweise Mathematik unterrichten, ohne dieses Fach als Studienfach oder als Schwerpunkt mit umfangreicherem Stundenumfang studiert zu haben.⁶

5 Diese Angaben zum Lehramtsstudium gelten für Studierende, die derzeit eine Lehrerausbildung durchlaufen. Lehrkräfte, die bereits im Schuldienst tätig sind, haben zum Teil nicht alle aufgeführten Ausbildungsbestandteile absolviert.

6 Die hier verwendete Definition von „fachfremd“ ist für die Primarstufe also sehr eng, da im Lehramtsstudium für die Primarstufe fachdidaktische und fachwissenschaftliche Anteile in den Fächern Deutsch und Mathematik in den meisten Ländern mittlerweile zumindest in geringem Umfang obligatorische Bestandteile des Studiums sind (vgl. KMK, 2013; Porsch, 2016, 2017).

Es kann angenommen werden, dass Lehrkräfte, die fachfremd eingesetzt werden, über ein geringeres fachliches und fachdidaktisches Wissen in diesem Fach verfügen als Lehrkräfte, die das Fach studiert haben. Diese Wissensbereiche gelten neben Überzeugungen, Werthaltungen sowie motivationalen Orientierungen und selbstregulativen Fähigkeiten als zentrale Aspekte professioneller Kompetenz von Lehrkräften (siehe dazu die COACTIV-Studie⁷; z. B. Kunter et al., 2011) und sollten sich auf die Qualität der fachlichen Lerngelegenheiten und die Lernergebnisse der Schülerinnen und Schüler auswirken (vgl. auch Baumert et al., 2010). Die empirische Befundlage, insbesondere in der Primarstufe, ist hierzu jedoch nicht eindeutig.

Sowohl in der IQB-Ländervergleichsstudie 2011 als auch in TIMSS⁸ 2011 zeigte sich, dass Kinder in fachfremd unterrichteten Grundschulklassen im Durchschnitt tendenziell geringere Kompetenzen im Lesen, im Zuhören, in Mathematik und in den Naturwissenschaften erreichten (Porsch & Wendt, 2015; Richter, Kuhl, Reimers & Pant, 2012). Weiterführende Analysen mit Daten des IQB-Ländervergleichs 2011 fanden jedoch unter Kontrolle von Hintergrundmerkmalen der Schülerinnen und Schüler sowie der Zusammensetzung der Klassen keine Zusammenhänge zwischen fachfremdem Unterrichten und den Kompetenzen der Kinder (Ziegler & Richter, 2017). Auch in den Analysen der TIMS-Studie aus dem Jahr 2015 ergaben sich keine Kompetenzunterschiede in Abhängigkeit davon, ob die jeweiligen Lehrkräfte fachfremd unterrichteten oder nicht (Porsch & Wendt, 2016). Ebenso konnten Tiedemann und Billmann-Mahecha (2007) keine Unterschiede im Kompetenzzuwachs bei fachfremd unterrichteten und nicht fachfremd unterrichteten Klassen identifizieren, den sie für Kinder von der 3. zur 4. Jahrgangsstufe im Fach Mathematik und im Bereich Orthografie untersuchten.

Für die Sekundarstufe I zeigten die IQB-Ländervergleichsstudie 2012 und der IQB-Bildungstrend 2015 hingegen, dass Schülerinnen und Schüler geringere Kompetenzen erreichen, wenn sie von fachfremd eingesetzten Lehrkräften unterrichtet werden (Hoffmann & Richter, 2016; Richter, Kuhl, Haag & Pant, 2013). Diese Effekte ließen sich insbesondere für nicht-gymnasiale Schularten nachweisen und waren in den Fächern Deutsch, Englisch und Mathematik statistisch signifikant.

Insgesamt lässt sich festhalten, dass die Befundlage zum fachfremden Unterrichten in der Primarstufe heterogen ausfällt, während in der Sekundarstufe I Schülerinnen und Schüler bei fachfremd unterrichtenden Lehrkräften in der Tendenz geringere Kompetenzen erreichen. Vor diesem Hintergrund soll im Rahmen des IQB-Bildungstrends 2016 erneut geprüft werden, ob fachfremdes Unterrichten mit den Kompetenzen von Viertklässlerinnen und Viertklässlern unter Berücksichtigung weiterer Merkmale der Kinder und der Lehrkräfte zusammenhängt.

7 Das Akronym COACTIV steht für *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung mathematischer Kompetenz*.

8 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.

12.3.2 Befunde

Im IQB-Bildungstrend 2016 wurden die Lehrkräfte der teilnehmenden Schulklassen unter anderem nach ihrem Lehramtsstudium und nach ihren Studienfächern gefragt. Basierend auf diesen Angaben kann bestimmt werden, ob Lehrkräfte fachfremd unterrichten. Als fachfremd unterrichtend werden, wie oben erläutert, alle Lehrkräfte kategorisiert, die das von ihnen unterrichtete Fach nicht als Studienfach belegt hatten.⁹

Die in Tabelle 12.3 aufgeführten Anteile fachfremd unterrichtender Lehrkräfte verdeutlichen, dass im Jahr 2016 bundesweit rund 19 Prozent der Lehrkräfte im Fach Deutsch und 31 Prozent der Lehrkräfte im Fach Mathematik in den am IQB-Bildungstrend teilnehmenden Klassen fachfremd unterrichten, bei deutlich variierenden Anteilen zwischen den Ländern. Relativ hohe Anteile fachfremd erteilten Unterrichts in den teilnehmenden Klassen finden sich im Fach Deutsch in Baden-Württemberg (32 %), Rheinland-Pfalz (46 %) und im Saarland (41 %). Besonders gering ist dieser Anteil in Mecklenburg-Vorpommern (4 %), Sachsen (3 %) und Thüringen (1 %). Im Fach Mathematik zeichnet sich ein ähnliches Bild ab, wobei die Anteile fachfremd unterrichtender Lehrkräfte in fast allen

Tabelle 12.3: Prozentuale Anteile fachfremd unterrichtender Lehrkräfte der am IQB-Bildungstrend 2016 teilnehmenden Klassen in den Fächern Deutsch und Mathematik nach Land

Land	Deutsch		Mathematik	
	N	Anteil fachfremd unterrichtender Lehrkräfte in %	N	Anteil fachfremd unterrichtender Lehrkräfte in %
Baden-Württemberg	72	31.9	63	57.1
Bayern	72	11.1	71	32.4
Berlin	106	27.4	101	41.6
Brandenburg	63	15.9	62	21.0
Bremen	72	19.4	77	35.1
Hessen	75	16.0	75	28.0
Mecklenburg-Vorpommern	76	3.9	77	5.2
Niedersachsen	76	21.1	75	45.3
Nordrhein-Westfalen	88	15.9	86	16.3
Rheinland-Pfalz	85	45.9	83	60.2
Sachsen	64	3.1	64	9.4
Sachsen-Anhalt	60	6.7	56	5.4
Schleswig-Holstein	69	8.7	66	39.4
Thüringen	77	1.3	71	1.4
Hamburg ¹	78	25.6	77	36.4
Saarland ¹	46	41.3	43	69.8
Deutschland	1 179	18.7	1 147	31.2

Anmerkungen. N = Anzahl der Lehrkräfte mit gültigen Angaben.

¹ Die Befunde stehen aufgrund eines erhöhten Anteils fehlender Daten unter Vorbehalt (vgl. Kapitel 3.1).

9 Im IQB-Bildungstrend 2016 wird also als Indikator für fachfremdes Unterrichten das Studienfach der Lehrkräfte herangezogen (vgl. auch die TIMS-Studie; Porsch & Wendt, 2015). Im IQB-Ländervergleich 2011 wurde hingegen die Lehrbefähigung der Lehrkräfte für die einzelnen Fächer erhoben, was zwar ebenfalls als ein Indikator für die studierten Fächer gelten kann, mit der direkten Erfassung des Studienfachs jedoch nicht identisch ist. Daher werden die Ergebnisse der beiden Studien nicht direkt gegenübergestellt. Die Befunde für das Jahr 2011 sind im Berichtsband zum IQB-Ländervergleich 2011 zu finden (Richter et al., 2012).

Ländern höher sind als im Fach Deutsch. Die höchsten Anteile fachfremd erteilten Unterrichts im Fach Mathematik sind erneut für Baden-Württemberg (57 %), Rheinland-Pfalz (60 %) und im Saarland (70 %) festzustellen, die geringsten Anteile liegen auch hier in Mecklenburg-Vorpommern (5 %), Sachsen-Anhalt (5 %) und Thüringen (1 %) vor.

Einige Lehrkräfte, die in Schulen der Primarstufe in Deutschland unterrichten, haben kein reguläres grundständiges Lehramtsstudium absolviert, sondern sind als sogenannte Quereinsteiger¹⁰ in den Schuldienst eingetreten. Der Anteil der am IQB-Bildungstrend teilnehmenden Lehrkräfte, die im Jahr 2016 in der 4. Jahrgangsstufe als Quereinsteiger unterrichten, ist bundesweit mit jeweils rund 1 Prozent im Fach Deutsch und Mathematik sehr gering. Aus diesem Grund wird auf eine länderspezifische Beschreibung dieser Gruppe verzichtet.

Nach der deskriptiven Darstellung der Anteile fachfremd unterrichtender Lehrkräfte in den einzelnen Fächern werden im Folgenden Zusammenhänge zwischen der fachlichen Qualifikation der Lehrkraft und den von ihren Schülerinnen und Schülern erreichten Kompetenzen untersucht. Dies erfolgt mit Hilfe von Mehrebenenanalysen, um die hierarchische Datenstruktur angemessen zu berücksichtigen und gleichzeitig Kontrollvariablen auf der Ebene der Lehrkräfte und auf der Ebene der Schülerinnen und Schüler einzubeziehen (Raudenbush & Bryk, 2002).¹¹ Auf der Ebene der Lehrkräfte (Klassenebene) werden demografische Hintergrundmerkmale (Alter, Geschlecht) und Angaben zu ihrer Qualifikation (fachfremdes Unterrichten)¹² berücksichtigt. Als Kontrollvariablen auf der Ebene der Schülerinnen und Schüler (Individualebene) werden Angaben zum Geschlecht und zum dokumentierten sonderpädagogischen Förderbedarf der Schülerinnen und Schüler (vgl. Kapitel 3.1), zu ihrer Familiensprache und zum sozioökonomischen Status ihrer Eltern (HISEI: *Highest International Socio-Economic Index*, vgl. Ganzeboom, 2010; Ganzeboom, De Graaf & Treiman, 1992) sowie die Ergebnisse eines Tests zur Erfassung kognitiver Grundfähigkeiten (Schroeders, Schipolowski & Wilhelm, 2017; vgl. auch Wilhelm, Schroeders & Schipolowski, 2014) in die Analysen aufgenommen.¹³

Die in Tabelle 12.4 dargestellten Ergebnisse zeigen, dass weder in den untersuchten Kompetenzbereichen *Lesen*, *Zuhören*, *Orthografie* noch im Fach Mathematik systematische Nachteile für fachfremd unterrichtete Schülerinnen und Schüler bestehen. Auch in Modellen ohne statistische Kontrolle der in Tabelle 12.4 aufgeführten Hintergrundmerkmale ergeben sich keine statistisch signifikanten Kompetenzunterschiede der Kinder in Klassen, deren Lehrkräfte das entsprechende Fach studiert haben, im Vergleich zu Kindern in fachfremd unterrichteten Klassen.

10 In der Praxis wird außerdem zwischen Quer- und Seiteneinsteigern unterschieden. Während Quereinsteiger ein berufsbegleitendes Referendariat absolvieren müssen, sind Seiteneinsteiger dazu nicht verpflichtet. Da im Rahmen des IQB-Bildungstrends 2016 keine Angaben zum Referendariat erhoben wurden, wird in diesem Kapitel nicht zwischen Quer- und Seiteneinsteigern differenziert, sondern beide Gruppen werden unter der Bezeichnung „Quereinsteiger“ zusammengefasst.

11 Alle Analysen wurden mit Mplus 7.3 (Muthén & Muthén, 1998–2012) unter Verwendung des *Full Information Maximum Likelihood*-Ansatzes durchgeführt.

12 Auf die Unterscheidung zwischen Lehrkräften, die das entsprechende Fach studiert haben, und Quereinsteigern wurde aufgrund geringer Fallzahlen für Quereinsteiger verzichtet.

13 Alle nominalen Prädiktorvariablen wurden wie folgt zu dichotomen Variablen mit den Ausprägungen 0 und 1 transformiert (Dummy-Kodierung): Geschlecht: 0 (männlich) und 1 (weiblich); dokumentierter sonderpädagogischer Förderbedarf: 0 (kein SPF) und 1 (SPF); Familiensprache: 0 (immer Deutsch) und 1 (manchmal oder immer eine andere Sprache); Fachfremd: 0 (Deutsch bzw. Mathematik im Rahmen des Lehramts studiert) und 1 (fachfremd unterrichtend). Alle übrigen Prädiktorvariablen wurden z-standardisiert.

Tabelle 12.4: Ergebnisse von Mehrebenenanalysen zur Vorhersage erreichter Kompetenzen von Schülerinnen und Schülern in den Fächern Deutsch und Mathematik durch Merkmale der Lehrkraft

Modellelemente	Deutsch Lesen ¹		Deutsch Zuhören ¹		Deutsch Orthografie ¹		Mathematik Globalskala ²	
	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)
Ordinatenabschnitt	489	(5.6)	490	(6.1)	497	(6.6)	500	(4.9)
Ebene der Schülerinnen und Schüler								
weibliches Geschlecht	19	(1.6)	8	(1.9)	29	(1.8)	-24	(1.5)
nicht nur deutsche Familiensprache	-21	(2.1)	-38	(2.5)	-6	(2.3)	-14	(1.9)
Sozioökonomischer Status (HISEI)	20	(1.0)	22	(1.2)	17	(1.1)	18	(1.0)
Kognitive Grundfähigkeit	41	(1.1)	36	(1.0)	30	(1.1)	55	(1.0)
Sonderpädagogischer Förderbedarf	-70	(5.9)	-64	(7.3)	-80	(6.0)	-63	(5.4)
Ebene der Lehrkräfte								
Fachfremd	-2	(3.5)	-1	(3.8)	-3	(4.4)	-2	(2.5)
Alter	1	(1.1)	0	(1.3)	0	(1.3)	1	(1.0)
weibliches Geschlecht	6	(4.4)	7	(4.6)	-2	(5.3)	5	(3.6)
Varianzanteile								
<i>R</i> ² Ebene der Schülerinnen und Schüler	0.33	(0.0)	0.27	(0.0)	0.22	(0.0)	0.45	(0.0)
<i>R</i> ² Ebene der Lehrkräfte	0.01	(0.0)	0.01	(0.0)	0.00	(0.0)	0.01	(0.0)

Anmerkungen. Die Kompetenzen auf der Ebene der Schülerinnen und Schüler sowie auf der Ebene der Klassen beziehungsweise Lehrkräfte werden mit den Gewichten der jeweiligen Ebene geschätzt. *b* = Regressionskoeffizient; *SE* = Standardfehler; *R*² = Bestimmtheitsmaß.

¹ *N*_{Kinder} = 23 196, *N*_{Lehrkräfte} = 1 208; ² *N*_{Kinder} = 22 462, *N*_{Lehrkräfte} = 1 167.
Fett gedruckte Regressionskoeffizienten sind statistisch signifikant (*p* < .05).

12.4 Berufliche Qualifikation und Einschätzungen von Lehrkräften zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf

12.4.1 Forschungsstand

Die unterrichtlichen Rahmenbedingungen an allgemeinen Schulen haben sich in den letzten Jahrzehnten erheblich gewandelt. Nachdem in Deutschland Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf lange Zeit überwiegend an Förderschulen unterrichtet wurden, steigt seit etwa 15 Jahren der Anteil von Kindern und Jugendlichen mit SPF, die eine allgemeine Schule besuchen, kontinuierlich an. Verbunden mit diesem Anstieg sind neue Anforderungen an das Professionswissen von Lehrkräften und an ihr unterrichtliches Handeln. Entsprechend sehen die im Jahr 2013 verabschiedeten Rahmenvereinbarungen der KMK über die Ausbildung für ein Lehramt in der Primarstufe explizit die Vermittlung von Basisqualifikationen in den Themenbereichen Umgang mit Heterogenität und Inklusion sowie die Möglichkeit einer sonderpädagogischen Schwerpunktsetzung im Studium vor (KMK, 2013). Neben Professionswissen sind auch Überzeugungen, Werthaltungen sowie motivationale Orientierungen und selbstregulative Fähigkeiten für professionelles Handeln von Lehrkräften (vgl. Abschnitt 12.3.1; Kunter et al., 2011) und für gelingende Anpassungsprozesse bei Veränderungen im Bildungssystem zentral. Dies gilt gerade auch für den gemeinsamen Unterricht von Schülerinnen und Schülern mit SPF und

Kindern und Jugendlichen ohne SPF, der von einer zunehmenden Anzahl von Lehrkräften umzusetzen ist.

Eine allgemein wichtige Überzeugung im Lehrerberuf ist das Erleben von Selbstwirksamkeit, die bei der Planung und Durchführung von Unterricht eine Rolle spielt. Sie bezieht sich auf die Selbsteinschätzung von Lehrkräften, den Unterricht für die Schülerinnen und Schüler ihrer Klasse angemessen gestalten und bewältigen zu können (vgl. Bandura, 1997; Schwarzer & Jerusalem; 2002). Ebenso können Einstellungen¹⁴ von Lehrkräften, also „Vorstellungen und Annahmen über schul- und unterrichtsbezogene Prozesse mit einer bewertenden Komponente“ (Kunter & Pohlmann, 2009, S. 267), das unterrichtliche Handeln leiten.

Aktuelle Analysen, die in Deutschland anhand von Daten regionaler Stichproben durchgeführt wurden, weisen darauf hin, dass Lehrkräfte gegenüber dem gemeinsamen Unterricht tendenziell hohe Selbstwirksamkeitserwartungen und positive Einstellungen aufweisen. Diese Einschätzungen gehen in der Tendenz mit einer erhöhten Bereitschaft einher, gemeinsamen Unterricht zu erteilen, und scheinen bei Lehrkräften, die über Erfahrungen mit gemeinsamem Unterricht verfügen, höher ausgeprägt zu sein (vgl. Bosse et al., 2016; Urton, Wilbert & Hennemann; 2015; für einen Überblick des Forschungsstands zu Überzeugungen und Einstellungen von Lehrkräften zum gemeinsamen Unterricht im deutschsprachigen Raum siehe Gasterstädt & Urban, 2016; Ruberg & Porsch, 2017). Gleichzeitig scheint bei vielen Lehrkräften ein ausgeprägter Fortbildungsbedarf zum gemeinsamen Unterricht zu bestehen. So gaben 68 Prozent aller im IQB-Bildungstrend 2011 befragten Grundschullehrkräfte an, einen hohen bis sehr hohen Fortbildungsbedarf zu „Integration/Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf“ zu haben. Zu diesem Themenfeld und zu den Themenfeldern „Förderung lernschwacher Schülerinnen und Schüler“ (69 %) sowie „Binnendifferenzierung/individuelle Förderung“ (62 %) waren die Anteile der Lehrkräfte, die einen hohen bis sehr hohen Fortbildungsbedarf berichteten, am größten.

Um die Befundlage anhand von aktuellen Daten zu erweitern, werden auf Basis von Informationen aus der Lehrkräftebefragung im Rahmen des IQB-Bildungstrends 2016 Qualifikationen und Einschätzungen von Lehrkräften zum gemeinsamen Unterricht beschrieben. Dabei wird zunächst auf Aspekte der Qualifikation, des Fortbildungsbedarfs und der Erfahrungen zum gemeinsamen Unterricht eingegangen. Weiterhin werden Überzeugungen zum gemeinsamen Unterricht und Einstellungen zur gemeinsamen Beschulung von Kindern mit und Kindern ohne SPF sowie zum Unterrichten der am IQB-Bildungstrend teilnehmenden Klasse analysiert. Im Folgenden werden die Ergebnisse ausschließlich für Deutschland insgesamt berichtet. Differenzierte Ergebnisdarstellungen auf Länderebene können auf der Webseite des IQB als Zusatzmaterial abgerufen werden. Diese Abbildungen und Tabellen sind durch den Zusatz „web“ gekennzeichnet.

14 Die Begriffe „Überzeugung“ und „Einstellung“ werden in Studien teilweise synonym verwendet (siehe Ruberg & Porsch, 2017).

12.4.2 Befunde

In den im Folgenden dargestellten Analysen liegt der Fokus auf denjenigen Lehrkräften, die mindestens ein Kind mit SPF in der am IQB-Bildungstrend teilnehmenden Klasse unterrichten.¹⁵ Dies ist bei 929 der insgesamt 1 784 Lehrkräfte (52 %) der Fall. Im Text wird ausschließlich auf die für diese Gruppe ermittelten Ergebnisse eingegangen. Der Vollständigkeit halber werden in den Tabellen und Abbildungen jedoch auch die Kennwerte für die Gruppe der Lehrkräfte angegeben, die kein Kind mit SPF in der am IQB-Bildungstrend teilnehmenden Klasse unterrichten.¹⁶

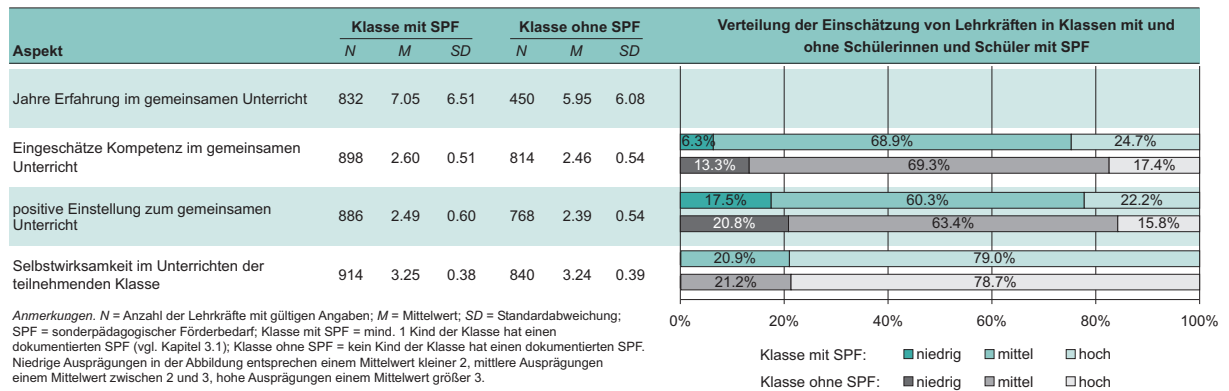
Zunächst erfolgt eine Beschreibung der Qualifikation, des Fortbildungsbedarfs und der Erfahrung der Lehrkräfte im Bereich Sonder- und Inklusionspädagogik und gemeinsamer Unterricht von Kindern mit und Kindern ohne SPF als Indikatoren professionsbezogener Hintergrundmerkmale in diesem Bereich. Zur Erfassung von Merkmalen der Qualifikation wurden Angaben der Lehrkräfte zu Ausbildung und Fächerwahl herangezogen und es wurde untersucht, wie viele Personen entweder ein sonderpädagogisches Lehramt oder das Fach Inklusionspädagogik beziehungsweise Sonderpädagogik im Rahmen ihres Lehramtsstudiums studiert haben. Insgesamt zeigen die Ergebnisse, dass in der Primarstufe derzeit nur wenige Lehrkräfte über eine spezifische sonder- oder inklusionspädagogische Qualifikation durch ihr Studium verfügen. In der Gruppe der Lehrkräfte, die in der am IQB-Bildungstrend 2016 teilnehmenden Klasse mindestens ein Kind mit SPF unterrichten, beträgt der Anteil derjenigen, die entweder ein sonderpädagogisches Lehramt oder das Fach Inklusionspädagogik beziehungsweise Sonderpädagogik im Rahmen ihres Lehramtsstudiums studiert haben, lediglich 3 Prozent.

Weitere Informationen über Merkmale der Qualifikation bietet der von Lehrkräften berichtete Fortbildungsbedarf. Die Lehrkräfte gaben ihren Fortbildungsbedarf im Bereich „Integration/Inklusion von Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf“ auf einer vierstufigen Antwortskala im Fragebogen an (1 = „überhaupt kein Bedarf“ bis 4 = „sehr hoher Bedarf“). Mit rund 72 Prozent ist der Anteil der Lehrkräfte, die nach eigener Einschätzung einen hohen oder sehr hohen Fortbildungsbedarf in diesem Bereich haben, insgesamt als sehr hoch einzustufen (vgl. auch Tab. 12.1web).

Um auch die praktische Erfahrung der Lehrkräfte im Unterrichten von Kindern mit SPF beschreiben zu können, wurden sie gebeten, in einem offenen Antwortformat anzugeben, wie viele Jahre sie bereits Erfahrung mit gemeinsamem Unterricht haben. Die Ergebnisse zeigen, dass die von den Lehrkräften berichtete Erfahrung im Durchschnitt 7 Jahre umfasst (Abb. 12.1), wobei nur sehr wenige Lehrkräfte (rund 1 %) angaben, weniger als ein Jahr Erfahrung mit dem Unterrichten von Kindern mit SPF zu haben. Am häufigsten (18 %) wurden 4 Jahre Erfahrung genannt (vgl. Tab. 12.2web für Angaben auf Länderebene).

15 Diese Kategorisierung erfolgte unabhängig davon, ob das betreffende Kind an den Kompetenztests und Befragungen teilgenommen hat oder nicht.

16 Dabei ist zu berücksichtigen, dass diese Lehrkräfte, deren teilnehmende Klasse kein Kind mit SPF umfasst, bezogen auf ihre unterrichtlichen Erfahrungen zu Kindern mit SPF eine heterogene Gruppe bilden (38 % von ihnen haben noch nie ein Kind mit SPF unterrichtet, 55% haben bereits entsprechende Erfahrungen).

Abbildung 12.1: Erfahrungen und Einschätzungen von Lehrkräften an allgemeinen Schulen zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf


Die beschriebenen Qualifikationen und Erfahrungen sagen jedoch nichts darüber aus, wie Lehrkräfte aktuell den gemeinsamen Unterricht von Kindern mit und Kindern ohne SPF einschätzen. Daher wurden im IQB-Bildungstrend 2016 auch selbsteingeschätzte Kompetenzen und Einstellungen zum gemeinsamen Unterricht sowie Selbstwirksamkeitsüberzeugungen bezogen auf das Unterrichten der am IQB-Bildungstrend teilnehmenden Klasse erhoben. Zur Erfassung der selbsteingeschätzten Kompetenz sollten die Lehrkräfte angeben, wie gut sie sich durch ihre Ausbildung, Fortbildung und Berufserfahrung auf insgesamt sieben verschiedene Aspekte des gemeinsamen Unterrichts (beispielsweise Einsatz von individuellen Fördermaßnahmen und differenzierte Leistungsbewertung) vorbereitet fühlen (vierstufige Antwortskala: 1 = „sehr schlecht“ bis 4 = „sehr gut“). Im Mittel liegen die von den Lehrkräften selbst eingeschätzten Kompetenzen zum gemeinsamen Unterricht bei $M = 2.6$ (Abb. 12.1; vgl. Abb. 12.1web für Angaben auf Länderebene), was in etwa dem theoretischen Mittelpunkt der Skala von $M = 2.5$ entspricht. Ein Blick auf Abbildung 12.1 und die dort dargestellten Balkendiagramme zeigt ferner, dass sich nur sehr wenige der Lehrkräfte (6%), die in der am IQB-Bildungstrend 2016 teilnehmenden Klasse mindestens ein Kind mit SPF unterrichten, schlecht vorbereitet fühlen, gemeinsamen Unterricht zu erteilen. Rund ein Viertel dieser Lehrkräfte fühlt sich hingegen sehr gut darauf vorbereitet.

Die Einstellungen zum gemeinsamen Unterricht wurden ebenfalls auf einer Skala abgebildet. Diese fasst die Stellungnahmen der Lehrkräfte zu insgesamt acht Aussagen zusammen (z. B. „In Klassen mit gemeinsamem Unterricht entwickeln sich die Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf besser als an einer Förderschule.“), die jeweils auf einer vierstufigen Antwortskala von 1 („trifft nicht zu“) bis 4 („trifft zu“) eingeschätzt wurden. Im Durchschnitt entspricht der empirisch ermittelte Wert ($M = 2.5$) für die Einstellungen der Lehrkräfte zum gemeinsamen Unterricht dem theoretischen Mittelpunkt der Skala. Allerdings unterscheiden sich die Lehrkräfte in ihren Einstellungen erheblich: Während rund 18 Prozent der Lehrkräfte kaum Vorteile im gemeinsamen Unterricht von Kindern mit und Kindern ohne SPF sehen, sind rund 22 Prozent der Lehrkräfte dem gemeinsamen Unterricht gegenüber sehr positiv eingestellt (Abb. 12.1; vgl. Abb. 12.2web für Angaben auf Länderebene). Die Mehrheit der Lehrerinnen und Lehrer (60%) äußert moderate Einstellungen zum gemeinsamen Unterricht.

Zusätzlich wurden die Selbstwirksamkeitsüberzeugungen der Lehrkräfte im Hinblick auf das Unterrichten der am IQB-Bildungstrend teilnehmenden Klasse

erfragt, um beschreiben zu können, wie sie ihre eigenen Fähigkeiten in der Gestaltung von Lernprozessen für diese Lerngruppe einschätzen. Die Erhebung dieses Merkmals erfolgte mit insgesamt zehn Aussagen (z. B. „Ich kann den Unterricht so organisieren, dass alle Schülerinnen und Schüler in ihrem eigenen Lerntempo zum Ziel kommen können.“), die von den Lehrkräften anhand einer vierstufigen Antwortskala (1 = „trifft nicht zu“ bis 4 = „trifft zu“) bewertet wurden. Insgesamt zeigt die Analyse, dass Lehrkräfte in Klassen, in denen sie mindestens ein Kind mit SPF gemeinsam mit Kindern ohne SPF unterrichten, ihre Selbstwirksamkeit sehr positiv einschätzen ($M = 3.3$). Der Anteil von Lehrkräften mit einer sehr niedrig ausgeprägten Selbstwirksamkeit liegt lediglich bei rund 2 Prozent. Die Angaben der großen Mehrzahl von Lehrkräften (79 %) weisen auf sehr hohe Ausprägungen der Selbstwirksamkeit im Unterrichten der teilnehmenden Klasse hin (Abb. 12.1; vgl. Abb. 12.3web für Angaben auf Länderebene). Konkret scheinen Lehrkräfte also auch dann überzeugt zu sein, die Lernprozesse für die teilnehmende Klasse angemessen gestalten zu können, wenn diese sowohl Kinder mit SPF als auch Kinder ohne SPF umfasst.

Um abschließend zu prüfen, inwieweit die von den Lehrkräften berichteten Erfahrungen und Einschätzungen miteinander zusammenhängen, wurden in einem nächsten Schritt Korrelationen berechnet.¹⁷ Dabei wurden wiederum nur die Angaben der Lehrkräfte einbezogen, die in der am IQB-Bildungstrend 2016 teilnehmenden Klasse mindestens ein Kind mit SPF unterrichten. In Übereinstimmung mit der oben zusammengefassten Befundlage zeigt sich hierbei, dass Lehrkräfte mit längerer Erfahrung im gemeinsamen Unterricht auch ihre darauf bezogene Kompetenz als höher einschätzen ($r = .30, p < .05$). Hingegen hängt der Umfang an Erfahrung kaum damit zusammen, wie positiv die Einstellungen der Lehrkräfte zum gemeinsamen Unterricht von Kindern mit SPF und Kindern ohne SPF an allgemeinen Schulen sind ($r = .08, p < .05$). Lehrkräfte, die ihre Kompetenzen zum gemeinsamen Unterricht als hoch beurteilen, weisen jedoch tendenziell positivere Einstellungen zum gemeinsamen Unterricht auf ($r = .34, p < .05$). Die Selbstwirksamkeit im Unterrichten der gesamten Klasse schließlich steht kaum mit der Erfahrung im gemeinsamen Unterricht in Zusammenhang ($r = .08, p < .05$), sie ist jedoch positiv mit den Einstellungen zum gemeinsamen Unterricht ($r = .35, p < .05$) und mit der eingeschätzten Kompetenz zum gemeinsamen Unterricht ($r = .21, p < .05$) assoziiert.

12.5 Zusammenfassung und Diskussion

In diesem Kapitel wurden die Deutsch- und Mathematiklehrkräfte allgemeiner Schulen, deren Klassen am IQB-Bildungstrend 2016 teilgenommen haben, anhand demografischer Merkmale und Merkmale ihrer beruflichen Qualifikation beschrieben. In einer vertiefenden Analyse wurde untersucht, ob Schülerinnen und Schüler von fachfremd unterrichtenden Lehrkräften ein geringeres Kompetenzniveau erreichen als Schülerinnen und Schüler von Lehrkräften, die das jeweilige Unterrichtsfach studiert haben. Weiterhin wurden zum Themenfeld „Umgang mit Heterogenität“, das im IQB-Bildungstrend 2016 einen Schwerpunkt bildet, Analysen durchgeführt, die sich auf das Unterrichten von Kindern mit SPF an

¹⁷ Auf die Untersuchung von Zusammenhängen mit der im Lehramtsstudium erworbenen Qualifikation der Lehrkräfte wurde aufgrund der geringen Fallzahlen von Lehrkräften, die im Rahmen des Studiums Qualifikationen im sonder- oder inklusionspädagogischen Bereich erworben haben, verzichtet.

allgemeinen Schulen beziehen. Im Folgenden werden die zentralen Befunde des Kapitels zusammengefasst und zu den Ergebnissen bisheriger Studien in Beziehung gesetzt.

In den Analysen zu den demografischen Merkmalen von Deutsch- und Mathematiklehrkräften werden vor allem Unterschiede zwischen den ostdeutschen und den westdeutschen Ländern deutlich. Die Ergebnisse zur Altersverteilung zeigen, dass die Lehrkräfte in den ostdeutschen Flächenländern im Durchschnitt älter sind als in den westdeutschen Ländern. Ein Vergleich der Altersverteilungen im IQB-Ländervergleich 2011 und im IQB-Bildungstrend 2016 macht deutlich, dass der Anteil der jüngeren Lehrkräfte bundesweit gestiegen ist, wobei diese Verschiebung in Baden-Württemberg und in Berlin sowie bezogen auf ein Fach in Mecklenburg-Vorpommern, in Sachsen-Anhalt und im Saarland¹⁸ besonders ausgeprägt ist. Im Einklang mit früheren Ländervergleichs- und Bildungstrendstudien des IQB (Hoffmann & Richter, 2016; Richter et al., 2012; Richter et al., 2013) und Modellrechnungen der KMK zum Lehrereinstellungsbedarf und Lehrkräfteangebot (KMK, 2015) ist also erneut zu resümieren, dass sich insbesondere die ostdeutschen Flächenländer den mit einer Lehrerpensionierungswelle verbundenen Herausforderungen werden stellen müssen.

Die deskriptiven Analysen zur Qualifikation der Lehrkräfte in den Fächern Deutsch und Mathematik haben gezeigt, dass der Anteil fachfremd unterrichtender Lehrkräfte sowohl für Deutschland insgesamt als auch innerhalb der einzelnen Länder im Fach Mathematik höher ausfällt als im Fach Deutsch, wobei sich der Anteil zwischen den Ländern teilweise deutlich unterscheidet. Dieses Muster stimmt mit Befunden für die Sekundarstufe überein (IQB-Ländervergleich 2012; Richter et al., 2013; IQB-Bildungstrends 2015; Hoffmann & Richter, 2016), die ebenfalls höhere Anteile fachfremden Unterrichts im Fach Mathematik als in den Fächern Deutsch und Englisch bei ausgeprägten Länderunterschieden zeigten. Außerdem korrespondieren die Ergebnisse mit Befunden aus Befragungen von Schulleiterinnen und Schulleitern der Sekundarstufe I in den PISA-Studien¹⁹ 2003 und 2006, die insbesondere für die mathematisch-naturwissenschaftlichen Fächer einen Mangel an qualifizierten Fachlehrkräften feststellten (vgl. Autorengruppe Bildungsberichterstattung, 2008). Inwieweit die verbindlichen Regelungen der KMK zur Fachlehrerausbildung in der Primarstufe (vgl. Rahmenvereinbarungen der KMK, 2013) zur Reduzierung der Unterschiede im Anteil fachfremden Unterrichts zwischen Mathematik und Deutsch sowie zwischen Ländern beitragen, wird sich in zukünftigen Untersuchungen zeigen.

Als weiteres Merkmal der Qualifikation von Lehrkräften wurde in diesem Kapitel untersucht, wie viele Personen die am IQB-Bildungstrend 2016 teilnehmenden Klassen unterrichten, ohne ein Lehramtsstudium absolviert zu haben (Quereinsteiger). Bundesweit ist diese Quote sowohl im Fach Deutsch als auch im Fach Mathematik sehr gering (je rund 1 %) und liegt unterhalb der für die Sekundarstufe ermittelten Anteile für die Fächer Deutsch und Englisch (6.2 % bzw. 5.9 %; vgl. Hoffmann & Richter, 2016). Aufgrund der geringen Fallzahl wurde die Gruppe der Quereinsteiger nicht vertiefend untersucht.

Weiterführende Analysen zeigten, dass die Qualifikation der Lehrkräfte in den Fächern Deutsch und Mathematik (fachfremdes Unterrichten) nicht signifikant mit den von ihren Schülerinnen und Schülern erreichten Kompetenzen zusammenhängt. Dieser Befund entspricht den Ergebnissen der Studien von Porsch

18 Im Saarland sind die Ergebnisse aufgrund eines erhöhten Anteils fehlender Angaben mit einer besonders großen Unsicherheit behaftet und stehen unter Vorbehalt.

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

und Wendt (2016), Tiedemann und Billmann-Mahecha (2007) sowie Ziegler und Richter (2017). Das Ergebnis deutet darauf hin, dass das Fachstudium an sich nicht prädiktiv ist für die von Schülerinnen und Schülern in den Fächern Deutsch und Mathematik in der Primarstufe erworbenen Kompetenzen. Damit unterscheidet sich der Befund von den Ergebnissen entsprechender Analysen in der Sekundarstufe I, die Kompetenznachteile bei fachfremd unterrichteten Schülerinnen und Schülern insbesondere an nicht-gymnasialen Schularten nachgewiesen haben (Hoffmann & Richter, 2016, Richter et al., 2013). Die mit einem Fachstudium verbundene Expertise von Lehrkräften scheint also in weiterführenden Schulen von größerer Bedeutung für den Lernerfolg zu sein als im Primarbereich. Möglicherweise wirkt sich diese erst bei der Vermittlung komplexerer Unterrichtsinhalte in der Sekundarstufe auf die Schülerleistungen aus. Ab welcher Klassenstufe die fachliche Qualifikation der Lehrkraft von Bedeutung für den Lernerfolg ihrer Schülerinnen und Schüler ist, bedarf weiterer Untersuchungen.

Bei der Interpretation der Kompetenzunterschiede in der vorliegenden Untersuchung muss berücksichtigt werden, dass fachliche Weiterbildungen, an denen die Lehrkräfte im Laufe ihres Schuldienstes teilgenommen haben, in den vorliegenden Analysen nicht einbezogen wurden. Es kann also nicht ausgeschlossen werden, dass die Lehrkräfte zusätzliche Fortbildungen in den Fächern absolviert haben, die sie fachfremd unterrichten. Darüber hinaus ist zu berücksichtigen, dass die Ausbildungsordnungen für das Lehramt einem zeitlichen Wandel unterliegen und sich somit auch die fachlichen und fachdidaktischen Qualifikationen von Lehrkräften, die zu verschiedenen Zeitpunkten ausgebildet wurden, unterscheiden. Es bedarf daher genauerer Untersuchungen zur Frage, welche Kompetenzen durch ein Studium des Lehramtes für die Primarstufe erworben werden und wie sich diese Kompetenzen in unterrichtlichem Handeln niederschlagen. Da durch die Anwendung des Klassenlehrerprinzips in der Primarstufe häufig die gleiche Lehrkraft Deutsch und Mathematik in einer Klasse unterrichtet, sind Fragen der Ausbildung in diesen zentralen Fächern relevant (vgl. Porsch & Wendt, 2016).

In weiteren Analysen wurden ausgewählte Merkmale der Qualifikation, Erfahrung und Einschätzungen von Lehrkräften zum gemeinsamen Unterricht von Kindern mit SPF und Kindern ohne SPF untersucht. Dabei lag der Fokus auf denjenigen Lehrkräften an allgemeinen Schulen, die in der am IQB-Bildungstrend 2016 teilnehmenden Klasse mindestens ein Kind mit SPF unterrichten, was bei gut der Hälfte der befragten Lehrkräfte der Fall war. Nur wenige dieser Lehrerinnen und Lehrer verfügen über eine spezifische sonder- oder inklusionspädagogische Qualifizierung durch ihr Lehramtsstudium und ein Großteil von ihnen gibt an, einen hohen oder sehr hohen Fortbildungsbedarf im Bereich gemeinsamer Unterricht zu haben (72 %). Dies entspricht weitgehend den Ergebnissen des IQB-Ländervergleichs 2011, in dem knapp 68 Prozent der Lehrkräfte einen hohen oder sehr hohen Fortbildungsbedarf zu diesem Thema angaben. Eine Abnahme dieses Bedarfs ist über den Zeitraum von fünf Jahren also nicht zu verzeichnen. Dabei muss allerdings offen bleiben, inwieweit dieser Befund darauf zurückzuführen ist, dass das Fortbildungsangebot unzureichend ist oder aber die Lehrkräfte das vorhandene Angebot nicht nutzen. Für die Zukunft sollte dies genauer bestimmt werden, um dem offenbar immer noch hohen Bedarf an Fortbildung zu Inklusion und Integration durch die Bereitstellung qualitativ hochwertiger Angebote, die potenziell alle Schulen erreichen können, Rechnung zu tragen. Die meisten Lehrkräfte berichteten allerdings, bereits seit einigen Jahren Erfahrung im Umgang mit gemeinsamem Unterricht zu haben, und schät-

zen ihre auf gemeinsamen Unterricht bezogene Kompetenz moderat ein: Sie fühlen sich mehrheitlich weder besonders gut noch besonders schlecht durch ihre Ausbildung, Fortbildung und Berufserfahrung auf den gemeinsamen Unterricht vorbereitet. Auch die Einstellungen der Lehrerinnen und Lehrer zum gemeinsamen Unterricht von Kindern mit und Kindern ohne SPF an allgemeinen Schulen im Vergleich zur Beschulung an Förderschulen liegen überwiegend im mittleren Bereich. Insgesamt verfügen die Lehrkräfte über eine hohe Selbstwirksamkeit im Unterrichten aller Kinder der teilnehmenden Klasse. Ob die hier untersuchten Einschätzungen von Lehrkräften mit schulischen Rahmenbedingungen zusammenhängen, unter denen sie unterrichten, und mit den Kompetenzen der von ihnen unterrichteten Kinder korrelieren, sind Fragen, die in weiterführenden Analysen untersucht werden sollten.

Literatur

- Autorengruppe Bildungsberichterstattung. (2008). *Bildung in Deutschland 2008. Ein indikatorergestützter Bericht mit einer Analyse zu Übergängen im Anschluss an den Sekundarbereich II*. Bielefeld: W. Bertelsmann Verlag.
- Bandura, A. (1997): *Self-efficacy: The exercise of control*. New York: Freeman.
- 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, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47, 133–180.
- Bosse, S., Henke, T., Jäntsch, C., Lambrecht, J. Vock, M. & Spörer, N. (2016). Die Entwicklung der Einstellung zum inklusiven Lernen und der Selbstwirksamkeit von Grundschullehrkräften. *Empirische Sonderpädagogik*, 1, 103–116.
- Ganzeboom, H. B. G. (2010). *A new international socio-economic index [ISEI] of occupational status for the International Standard Classification of Occupation 2008 [ISCO-08] constructed with data from the ISSP 2002–2007; with an analysis of quality of educational measurement in ISSP*. Vortrag auf der Annual Conference of International Social Survey Programme, Lisbon.
- Ganzeboom, H. B. G., De Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1–56.
- Gasterstädt, J. & Urban, M. (2016). Einstellung zu Inklusion? Implikationen aus Sicht qualitativer Forschung im Kontext der Entwicklung inklusiver Schule. *Empirische Sonderpädagogik*, 8(1), 54–66.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hattie, J. (2012). *Visible learning for teachers. Maximizing impact on learning*. London: Routledge.
- Hoffmann, L. & Richter, D. (2016). Aspekte der Aus- und Fortbildung von Deutsch- und Englischlehrkräften im Ländervergleich. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 481–507). Münster: Waxmann.
- KMK (2004) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). *Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004*. Berlin: KMK.
- KMK (2008) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2008). *Ländergemeinsame inhaltliche Anforderungen. Aspekte der Aus- und Fortbildung von Mathematik- und Naturwissenschaftslehrkräften im Ländervergleich für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung. Beschluss der Kultusministerkonferenz vom 16.10.2008 i.d.F. vom 08.12.2008*. Berlin: KMK.
- KMK (2012) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2012). *Ländergemeinsame Anforderungen für die Ausgestaltung des Vorbereitungsdienstes und die abschließende Staatsprüfung. Beschluss der Kultusministerkonferenz vom 06.12.2012*. Berlin: KMK.

- KMK (2013) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2013). *Rahmenvereinbarung über die Ausbildung und Prüfung für ein Lehramt der Grundschule bzw. Primarstufe (Lehramtstyp 1). Beschluss der Kultusministerkonferenz vom 28.02.1997 i. d. F. vom 10.10.2013*. Berlin: KMK.
- KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Lehrereinstellungsbedarf und -angebot in der Bundesrepublik Deutschland 2014–2025. Zusammengefasste Modellrechnungen der Länder*. Berlin: KMK.
- KMK (2017) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017). *Sachstand in der Lehrerbildung (Sachstand 07.03.2017)*. Zugriff am 20.08.2016 unter https://www.kmk.org/fileadmin/Dateien/pdf/Bildung/AllgBildung/2017-03-07__Sachstand_LB_o_EW.pdf
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften: Ergebnisse aus dem Forschungsprogramm COACTIV*. Münster: Waxmann.
- Kunter, M. & Pohlmann, B. (2009). Lehrer. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 261–282). Heidelberg: Springer Medizin.
- Muthén, L. K. & Muthén, B. O. (1998–2012). *Mplus: Statistical Analysis with Latent Variables: User's Guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Porsch, R. (2016). Fachfremd unterrichten in Deutschland. Definition – Verbreitung – Auswirkungen. *Die deutsche Schule*, 108, 9–32.
- Porsch, R. (2017). Spezialisten oder Generalisten? Eine Betrachtung der Fachausbildung von Grundschullehrerinnen und -lehrern in Deutschland. In M. Radhoff & S. Wieckert (Hrsg.), *Grundschule im Wandel* (S. 151–162). Hamburg: Verlag Dr. Kovac.
- Porsch, R. & Wendt, H. (2015). Welche Rolle spielt der Studienschwerpunkt von Sachunterrichtslehrkräften für ihre Selbstwirksamkeit und die Leistungen ihrer Schülerinnen und Schüler? In H. Wendt, T. C. Stubbe, K. Schwippert & W. Bos (Hrsg.), *10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011* (S. 161–183). Münster: Waxmann.
- Porsch, R. & Wendt, H. (2016). Aus- und Fortbildung von Mathematik- und Sachunterrichtslehrkräften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 189–204). Münster: Waxmann.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.
- Richter, D., Kuhl, P., Haag, N. & Pant, H. A. (2013). Aspekte der Aus- und Fortbildung von Mathematik- und Naturwissenschaftslehrkräften im Ländervergleich. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 367–390). Münster: Waxmann.
- Richter, D., Kuhl, P., Reimers, H. & Pant, H. A. (2012). Aspekte der Aus- und Fortbildung von Lehrkräften in der Primarstufe. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 237–250). Münster: Waxmann.
- Ruberg, C. & Porsch, R. (2017): Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion: Ein systematisches Review deutschsprachiger Forschungsarbeiten. *Zeitschrift für Pädagogik*, 63, 393–415.
- 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.
- Schroeders, U., Schipolowski, S. & Wilhelm, O. (2017). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 3. und 4. Jahrgangsstufe*. Manuskript in Vorbereitung.
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen. Zeitschrift für Pädagogik, Beiheft 44* (S. 28–53). Weinheim: Beltz.
- Tiedemann, J. & Billmann-Mahecha, E. (2007). Macht das Fachstudium einen Unterschied?. Zur Rolle der Lehrerexpertise für Lernerfolg und Motivation in der Grundschule. *Zeitschrift für Pädagogik*, 53, 58–73.

- Urton, K., Wilbert, J. & Hennemann, T. (2015). Die Einstellung zur Integration und die Selbstwirksamkeit von Lehrkräften. *Psychologie in Erziehung und Unterricht*, 62(2), 147–157.
- Wilhelm, O., Schroeders, U. & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8-10)*. Göttingen: Hogrefe.
- Ziegler, C. & Richter, D. (2017). Der Einfluss fachfremden Unterrichts auf die Schülerleistung: Können Unterschiede in der Klassenzusammensetzung zur Erklärung beitragen? *Unterrichtswissenschaft*, 45, 136–155.

Kapitel 13

Grundlagen der Kompetenzmessung im IQB-Bildungstrend 2016

Anlage, Durchführung und Auswertung des IQB-Bildungstrends 2016 wurden bereits in Kapitel 3 im Überblick dargestellt. Diese Beschreibung wird in Kapitel 13.1 vertieft, um die zentralen methodischen Aspekte der Durchführung und der statistischen Auswertung des IQB-Bildungstrends 2016 transparent zu machen. Das Vorgehen orientierte sich dabei an Methoden, die sich in den IQB-Ländervergleichsstudien 2009, 2011 und 2012 sowie im IQB-Bildungstrend 2015 bewährt haben (Hecht, Roppelt & Siegle, 2013; Sachse, Haag & Weirich, 2016; Weirich, Haag & Roppelt, 2012). Zudem erfolgte eine Orientierung an den internationalen Schulleistungstudien, insbesondere an PISA (OECD, 2009, 2012) und PIRLS/IGLU (Tarelli, Wendt, Bos & Zylowski, 2012). Kapitel 13.1 zeigt auf, wie diese Verfahren und Methoden im Rahmen des IQB-Bildungstrends 2016 konkret angewendet wurden. Anschließend wird in Kapitel 13.2 auf die angepassten Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf eingegangen. Dabei werden einerseits Prinzipien der Aufgabenkonstruktion und andererseits psychometrische Eigenschaften dieser Aufgaben beschrieben.

13.1 Testdesign und Auswertung des IQB-Bildungstrends 2016

Sebastian Weirich, Nicole Haag und Karoline A. Sachse

13.1.1 Testdesign

Im IQB-Bildungstrend 2016 wurde überprüft, inwieweit Schülerinnen und Schüler der 4. Jahrgangsstufe zentrale Kompetenzziele erreichen, die in den Bildungsstandards der KMK für die Fächer Deutsch und Mathematik in der Primarstufe (Granzer et al., 2009) definiert sind. Weiterhin wurde ermittelt, inwieweit sich die Kompetenzstände seit dem ersten IQB-Ländervergleich in der Primarstufe im Jahr 2011 (Stanat, Pant, Böhme, & Richter, 2012) verändert haben.

Die Erfassung der Kompetenzen der Schülerinnen und Schüler erfolgte mit Testaufgaben, die zentrale Bildungsstandards in den zwei Fächern operationalisieren. Wie in Kapitel 3.1 erwähnt, wurde ein Teil dieser Aufgaben bereits 2011 eingesetzt, ein anderer Teil wurde für den IQB-Bildungstrend 2016 neu entwickelt. Im Fach Deutsch erfassten die Aufgaben Kompetenzen von Schülerinnen und Schülern in den Bereichen *Lesen*, *Zuhören* und *Orthografie* (vgl. Kapitel 1.2) und im Fach Mathematik Kompetenzen bezogen auf die Leitideen *Zahlen und*

Operationen, Raum und Form, Muster und Strukturen, Größen und Messen sowie Daten, Häufigkeit und Wahrscheinlichkeit. Die Ergebnisse zu den fünf Leitideen lassen sich zudem zu einem Wert für die globale Mathematikkompetenz zusammenfassen (vgl. Kapitel 1.3).

Alle an der Erhebung teilnehmenden Schülerinnen und Schüler erhielten jeweils ein Testheft, das nur eine Teilmenge aller Aufgaben umfasste, wobei auch nicht alle Testhefte sämtliche Kompetenzbereiche beziehungsweise Leitideen abdeckten. Ein solches Vorgehen – das auch als *Item Sampling* bezeichnet wird (Mislevy, Beaton, Kaplan & Sheehan, 1992; Mislevy, Johnson & Muraki, 1992) – ist in groß angelegten Schulleistungsstudien üblich, um die untersuchten Kompetenzen möglichst breit zu erfassen. Dies erfordert den Einsatz einer großen Anzahl von Aufgaben, die in ihrer Gesamtheit nicht von jeder einzelnen Schülerin und jedem einzelnen Schüler innerhalb einer zumutbaren Testzeit bearbeitet werden können. Damit die unterschiedlichen Teilmengen von Aufgaben bei der Auswertung dennoch auf einer gemeinsamen Skala abgebildet und die Leistungen verglichen werden können, wurden die Testhefte so gestaltet, dass verschiedene Hefte zum Teil dieselben Aufgaben enthielten. Auf diese Weise wurden die Aufgaben direkt innerhalb eines Testhefts oder indirekt über mehrere Testhefte hinweg miteinander verbunden (*Verlinkung*). Ein solches komplexes Testdesign, das auch aus mehreren Teildesigns bestehen kann, wird oft als *Multiple-Matrix-Design* bezeichnet (Gonzalez & Rutkowski, 2010).

Das Testdesign für den IQB-Bildungstrend 2016 setzt sich aus fünf verschiedenen Teildesigns zusammen, die für verschiedene Zielgruppen entwickelt wurden. Die Teildesigns unterscheiden sich teilweise hinsichtlich der Auswahl und der Zusammenstellung von Testaufgaben. Für das Fach Deutsch wurden drei Teildesigns erstellt: für Schülerinnen und Schüler ohne sonderpädagogischen Förderbedarf (SPF), für Schülerinnen und Schüler mit SPF an allgemeinen Schulen und für Schülerinnen und Schüler mit SPF an Förderschulen. Für das Fach Mathematik wurden zwei Teildesigns erstellt: für Schülerinnen und Schüler ohne SPF und für Schülerinnen und Schüler mit SPF.

Die Erstellung des Testdesigns für den IQB-Bildungstrend 2016 erfolgte in mehreren, aufeinander aufbauenden Schritten. Zunächst wurden Aufgaben zu Aufgabenblöcken gruppiert und diese dann wiederum zu Testheften zusammengestellt. Dieser Prozess wird im Folgenden näher erläutert.

Gruppierung der Aufgaben zu Aufgabenblöcken

Im ersten Schritt der Erstellung des Testdesigns wurden die einzelnen Aufgaben Aufgabenblöcken zugeordnet. Eine Aufgabe bestand aus einer oder mehreren Teilaufgaben, den sogenannten Items. Beispielsweise wurde den Schülerinnen und Schülern im Fach Deutsch für die Kompetenzbereiche *Lesen* und *Zuhören* eine Reihe von Items zur Bearbeitung vorgelegt, die sich auf jeweils einen Stimulustext (Lese- oder Hörtext) beziehen. Innerhalb eines Aufgabenblocks wechselten sich leichte und schwierige Items ab, wobei jeder Aufgabenblock mit eher leichten Items begann.

Ein Aufgabenblock umfasste im Fach Deutsch jeweils 20 Minuten und im Fach Mathematik jeweils 10 Minuten. Im Fach Deutsch konnte ein Block aus nur einer Aufgabe oder aus mehreren Aufgaben bestehen; im Fach Mathematik bestand ein Block immer aus mehreren Aufgaben. In beiden Fächern enthielt ein Aufgabenblock ausschließlich Aufgaben aus einem Kompetenzbereich.

Ein Teil der Blöcke wurde unverändert aus dem IQB-Ländervergleich 2011 übernommen. Dies betraf 4 Blöcke aus dem Kompetenzbereich *Zuhören*, 9 Blöcke aus dem Bereich *Lesen*, 12 Blöcke aus dem Bereich *Orthografie* und 31 Blöcke aus dem Fach Mathematik. Neben den unverändert übernommenen Blöcken kamen neu zusammengestellte Blöcke zum Einsatz. Diese Blöcke enthielten zum Teil neu entwickelte Aufgaben und zum Teil bereits im IQB-Ländervergleich 2011 eingesetzte Aufgaben.

Um die Messung der Kompetenzen von Schülerinnen und Schülern mit SPF zu optimieren, wurden in den Fächern Deutsch und Mathematik für den IQB-Bildungstrend 2016 zusätzliche Aufgaben entwickelt und in einer vorangehenden Pilotierungsstudie erprobt (vgl. Kapitel 13.2). Aufgaben und Items, die sich in dieser Pilotierungsstudie als empirisch geeignet erwiesen hatten, wurden für den IQB-Bildungstrend 2016 zu neuen Aufgabenblöcken gruppiert. Diese Blöcke kamen größtenteils auch bei Schülerinnen und Schülern ohne SPF zum Einsatz, um die Teildesigns miteinander zu verlinken. Eine Ausnahme bildet der Kompetenzbereich *Orthografie*, für den keine gesonderten Aufgaben für Schülerinnen und Schüler mit SPF entwickelt wurden. Deshalb wurden zur optimierten Erfassung der orthografischen Kompetenzen von Schülerinnen und Schülern mit SPF besonders leichte Aufgaben genutzt, die im Rahmen der Vergleichsarbeiten (VERA) für die 3. Jahrgangsstufe entwickelt und pilotiert, dann aber nicht im VERA-Haupttest eingesetzt worden sind.

Während in den Kompetenzbereichen *Lesen* und *Zuhören* eine Aufgabe immer genau einem Block zugeordnet war, konnte eine Aufgabe im Bereich *Orthografie* in mehreren Blöcken vorkommen. Im Fach Mathematik unterschied sich die Zuordnung der Aufgaben zu Blöcken zwischen den beiden Teildesigns, da für Schülerinnen und Schüler mit SPF teilweise neue Blöcke aus bestehenden Aufgaben zusammengestellt wurden. Innerhalb eines Teildesigns war jede Aufgabe immer genau einem Block zugeordnet.

Für den Kompetenzbereich *Zuhören* sowie für einige Aufgaben im Bereich *Orthografie* war zu berücksichtigen, dass aus organisatorischen Gründen allen Schülerinnen und Schülern innerhalb einer Klasse dieselben Stimuli auditiv präsentiert werden mussten. Dadurch waren die Möglichkeiten zum Einsatz angepasster Testaufgaben für inklusiv beschulte Schülerinnen und Schüler mit SPF in diesen Kompetenzbereichen eingeschränkt.

In Tabelle 13.1 ist dargestellt, welche Anzahl von Aufgabenblöcken, Aufgaben und Items in jedem Fach, jedem Kompetenzbereich und jedem Teildesign zum Einsatz kam. Da zum Zwecke der Verlinkung ein Großteil der Aufgaben und Items sowohl bei Schülerinnen und Schülern ohne SPF als auch bei Schülerinnen und Schülern mit SPF eingesetzt wurden, addieren sich die Zahlen in den Spalten „Anzahl Aufgabenblöcke“, „Anzahl Aufgaben“ und „Anzahl Items“ nicht zur Gesamtanzahl der in der Studie eingesetzten Blöcke beziehungsweise Aufgaben und Items.

Im Fach Mathematik bearbeiteten Schülerinnen und Schüler mit SPF eine Teilmenge der für Schülerinnen und Schüler ohne SPF ausgewählten Aufgaben. Die Gesamtzahl der eingesetzten Aufgaben und Items entspricht somit den in Tabelle 13.1 für Schülerinnen und Schüler ohne SPF angegebenen Zahlen. Die Globalskala mathematischer Kompetenz umfasst sämtliche Items zu allen mathematischen Leitideen. Die Itemzahlen der fünf Leitideen summieren sich jedoch nicht zur Anzahl der Items für die Globalskala auf. Dies ist darauf zurückzuführen, dass einige Items zwei Leitideen zugeordnet sind (vgl. Kapitel 1.3): Solche Items werden in der Tabelle für jede Leitidee separat gezählt.

Tabelle 13.1: Anzahl der im IQB-Bildungstrend 2016 eingesetzten Aufgabenblöcke, Aufgaben und Items

Eingesetzt bei	Fach	Kompetenzbereich	Anzahl Aufgabenblöcke	Anzahl Aufgaben	Anzahl Items
SuS ohne SPF	Deutsch	Lesen	15	17	134
		Zuhören	11	14	110
		Orthografie	15	53	146
	Mathematik	Globalskala	48	237	400
		Zahlen und Operationen	9	76	120
		Raum und Form	10	50	75
		Muster und Strukturen	8	39	66
		Größen und Messen	11	60	117
	Daten, Häufigkeit und Wahrscheinlichkeit	10	46	78	
SuS mit SPF	Mathematik	Globalskala	19	97	167
		Zahlen und Operationen	4	31	42
		Raum und Form	4	21	33
		Muster und Strukturen	3	17	31
		Größen und Messen	4	22	46
		Daten, Häufigkeit und Wahrscheinlichkeit	4	15	30
SuS mit SPF an allgemeinen Schulen	Deutsch	Lesen	11	17	106
		Zuhören	11	14	106
		Orthografie	8	53	139
SuS mit SPF an Förderschulen	Deutsch	Lesen	8	10	66
		Zuhören	6	9	45
		Orthografie	3	41	41

Anmerkungen. SuS = Schülerinnen und Schüler; SPF = sonderpädagogischer Förderbedarf. Im Fach Mathematik wurden für alle Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf unabhängig von der besuchten Schulart dieselben Aufgaben eingesetzt.

Zusammenstellung von Testheften

Im zweiten Schritt wurden die Aufgabenblöcke zu Testheften zusammengestellt. Bei der Erstellung des Testdesigns für den IQB-Bildungstrend 2016 war zu berücksichtigen, dass (a) für die Trendschätzung eine größtmögliche Konsistenz hinsichtlich der Aufgaben- und Blockpositionen mit dem Testdesign des IQB-Ländervergleichs 2011 gewährleistet werden sollte, (b) Schülerinnen und Schüler mit SPF unabhängig von ihrer besuchten Schulart in die Testung einbezogen wurden und (c) die Administration von Zuhöraufgaben sowie Lückensatz-Diktataufgaben im Kompetenzbereich *Orthografie* über CD-Spieler zur Folge hatte, dass die entsprechenden Stimuli in allen Testheftversionen innerhalb einer Klasse gleich sein mussten.

Tabelle 13.2: Anzahl der Testhefte je Teildesign im IQB-Bildungstrend 2016

Fach	Teildesign	Zielgruppe	Anzahl Testhefte
Deutsch	1. Teildesign	SuS ohne SPF	39 ¹
	2. Teildesign	SuS mit SPF an allgemeinen Schulen	39 ^{1, 2}
	3. Teildesign	SuS mit SPF an Förderschulen	17 ²
Mathematik	1. Teildesign	SuS ohne SPF	44
	2. Teildesign	SuS mit SPF (an Förderschulen und an allgemeinen Schulen)	19

Anmerkungen. SuS = Schülerinnen und Schüler; SPF = sonderpädagogischer Förderbedarf.

¹15 der 39 Testhefte in Teildesign 1 und 2 waren identisch. ²4 Testhefte in Teildesign 2 und 3 waren identisch.

Für das Fach Deutsch wurden drei Teildesigns und für das Fach Mathematik zwei Teildesigns erstellt (vgl. Tab. 13.2). Jedes Testheft in jedem Teildesign enthielt acht verschiedene Aufgabenblöcke à 10 Minuten Bearbeitungszeit (Mathematik) beziehungsweise vier verschiedene Aufgabenblöcke à 20 Minuten Bearbeitungszeit (Deutsch), sodass jedes Testheft innerhalb von 80 Minuten zu bearbeiten war. In jedem Teildesign wurde ein teilweise balanciertes, unvollständiges Blockdesign umgesetzt (*Incomplete Block Design*; Frey, Hartig & Rupp, 2009). „Unvollständig“ bedeutet hier, dass nicht jeder Aufgabenblock einmal mit jedem anderen Aufgabenblock in einem Testheft gemeinsam auftritt, und somit manche Aufgabenblöcke nur indirekt miteinander verbunden sind.

Ein Design gilt dann als balanciert, wenn jeder Aufgabenblock an jeder Blockposition im Testdesign gleich häufig auftritt (Frey et al., 2009). In einem solchen balancierten Testdesign sind alle Aufgaben in gleicher Weise von im Testverlauf eventuell nachlassender Motivation und Konzentration der Testteilnehmerinnen und Testteilnehmer betroffen, sodass die Itemparameter trotz möglicher Ermüdungseffekte unverzerrt geschätzt werden können. Daher wurden im IQB-Bildungstrend 2016 alle Aufgabenblöcke bezüglich der Blockpositionen weitestgehend ausbalanciert.

Das erste Teildesign im Fach Deutsch bestand aus 39 Testheften und enthielt sowohl Aufgaben des IQB-Ländervergleichs 2011 als auch neue Aufgaben aus dem Aufgabenpool. Obwohl dieses Teildesign für Schülerinnen und Schüler ohne SPF entwickelt wurde, enthielt es auch SPF-Aufgabenblöcke, um eine vertikale Verlinkung zwischen den heterogenen Schülerpopulationen herstellen zu können (A. von Davier, Carstensen & M. von Davier, 2008).

Das zweite Teildesign im Fach Deutsch wurde für Schülerinnen und Schüler mit SPF an allgemeinen Schulen konzipiert und bestand ebenfalls aus 39 Testheften. Anders als im ersten Teildesign kamen hier mehrheitlich SPF-Aufgabenblöcke zum Einsatz, die besonders leichte Aufgaben (*Orthografie*) oder eigens für Schülerinnen und Schüler mit SPF entwickelte Testaufgaben enthielten (*Lesen und Zuhören*).

Das dritte Teildesign im Fach Deutsch war auf Schülerinnen und Schüler mit SPF an Förderschulen ausgerichtet und bestand aus 17 Testheften.

Auch im Fach Mathematik wurde das erste Teildesign für Schülerinnen und Schüler ohne SPF entwickelt; es bestand aus 44 Testheften. In diesem Teildesign kamen sowohl Aufgaben des IQB-Ländervergleichs 2011 als auch neue Aufgaben aus dem IQB-Aufgabenpool sowie zur vertikalen Verlinkung mit der SPF-Population neu entwickelte Aufgaben für Schülerinnen und Schüler mit SPF zum Einsatz.

Das zweite Teildesign im Fach Mathematik wurde für Schülerinnen und Schüler mit SPF an allgemeinen Schulen und an Förderschulen konzipiert und bestand aus 19 Testheften.

An die Kompetenztests schlossen sich in den Erhebungen ein kurzer Test zur Erfassung kognitiver Grundfähigkeiten sowie ein Fragebogen an (vgl. Kapitel 3.1 zum Ablauf der Testung).

Verteilung der Testhefte auf Schülerinnen und Schüler

Die Testhefte aller Testdesigns wurden zufällig auf Schulklassen der jeweiligen Zielgruppen verteilt. Testhefte im Fach Deutsch konnten aufgrund der enthaltenen Zuhör- und Orthografieaufgaben innerhalb einer Klasse nicht variiert werden; alle Schülerinnen und Schüler einer Klasse erhielten folglich dieselbe Testheftversion. Im Fach Mathematik wurden die Testheftversionen auch innerhalb einer Klasse zufällig auf die Schülerinnen und Schüler verteilt. Die Zuweisung der Testhefte erfolgte jeweils so, dass jedes einzelne Testheft innerhalb eines Teildesigns von ungefähr gleich vielen Schülerinnen und Schülern bearbeitet wurde.

13.1.2 Skalierung der Kompetenztests

Die Aufgaben der Kompetenztests im IQB-Bildungstrend 2016 wurden fächerweise und kompetenzbereichsspezifisch entwickelt, wobei für jeden Kompetenzbereich (Deutsch) und jede Leitidee (Mathematik) jeweils von einem eindimensionalen Konstrukt ausgegangen wurde. Konstrukte wie die im IQB-Bildungstrend 2016 erfassten Kompetenzen können als latente, also nicht direkt beobachtbare Variablen verstanden werden. Daher wurden zur Kompetenzschätzung der Schülerinnen und Schüler in den einzelnen Domänen statistische Verfahren zur Modellierung latenter Variablen genutzt. Diese Modelle werden im Folgenden näher beschrieben.

Das statistische Modell

Das im IQB-Bildungstrend 2016 zur Skalierung der Kompetenzdaten verwendete statistische Modell mit latenten Variablen basiert auf der probabilistischen Testtheorie beziehungsweise der *Item Response Theory* (IRT; Embretson & Reise, 2000; Hambleton, Swaminathan & Rogers, 1991). Im Rahmen der IRT wurden Modelle entwickelt, mit denen eine funktionale Beziehung zwischen mindestens einer latenten Variablen (z. B. der Lesekompetenz im Fach Deutsch) und einer manifesten kategorialen Variablen (z. B. Item gelöst / Item nicht gelöst) formalisiert wird. Es wird angenommen, dass die Lesekompetenz nicht direkt beobachtbar ist und nur indirekt über die Schülerantworten auf die Aufgaben im Kompetenztest erschlossen werden kann. Mit Hilfe des Modells werden genaue Annahmen über den Zusammenhang zwischen der Kompetenz und der Wahrscheinlichkeit einer richtigen Itemantwort getroffen. Demnach ist die Wahrscheinlichkeit, ein bestimmtes Item zu lösen, umso größer, je höher die vorhandene Kompetenz ist.

Anders als in den vorangegangenen IQB-Ländervergleichsstudien (Hecht et al., 2013; Sachse et al., 2016; Weirich et al., 2012) wurde im IQB-Bildungstrend

2016 aus der Familie der IRT-Modelle nicht mehr das Rasch-Modell (Adams & Wu, 2007) zur Auswertung verwendet, sondern ein restringiertes zweiparametrisches IRT-Modell. Dieses wurde jedoch so parametrisiert, dass es dem Rasch-Modell möglichst ähnlich ist (siehe unten). Das 2pl- oder Birnbaum-Modell (Birnbaum, 1968) erlaubt als eine Erweiterung des Rasch-Modells, unterschiedliche (also von 1 abweichende) Trennschärfen für einzelne Items oder Itemgruppen zu modellieren.

Die Grundgleichung des 2pl-Modells für dichotome Daten modelliert die Wahrscheinlichkeit $P(X_{ik}=1)$, mit der eine Schülerin oder ein Schüler k mit Kompetenz θ_k ein bestimmtes Item i mit der Schwierigkeit β_i und der Trennschärfe α_i korrekt lösen kann:

$$P(X_{ik} = 1) = \frac{e^{\alpha_i(\theta_k - \beta_i)}}{1 + e^{\alpha_i(\theta_k - \beta_i)}}$$

Setzt man in der oben stehenden Gleichung α_i für jedes Item auf den festen Wert 1, entspricht die Gleichung der des Rasch-Modells. X_{ik} ist eine Zufallsvariable, deren Realisation die konkret beobachtete Itemantwort bezeichnet (0 im Falle einer Falschantwort und 1 im Falle einer Richtigantwort). Die Itemschwierigkeiten β_i und Personenfähigkeiten θ_k liegen auf einer kontinuierlichen Skala, die einen nach oben und nach unten unbegrenzten Wertebereich aufweist und deren Nullpunkt zunächst unbestimmt ist. Sie wird als Logit-Skala bezeichnet. Dasselbe gilt (theoretisch) auch für die Trennschärfen α_i , wobei Werte nahe Null oder kleiner als Null unplausibel und zumeist auf Fehler bei der Kodierung der Itemantworten zurückzuführen sind. Im IQB-Bildungstrend 2016 wurden die Personenparameter aus dieser Metrik auf eine leichter verständliche und interpretierbare Metrik transformiert.

Die Entscheidung, das 2pl-Modell anstelle des Rasch-Modells zu verwenden, resultiert aus dem Befund, dass die neu entwickelten Items für Schülerinnen und Schüler mit SPF empirisch eine im Mittel von 1 abweichende Trennschärfe besitzen, also nicht raschhomogen sind. So liegt die Trennschärfe von SPF-Items für den Kompetenzbereich *Lesen* im Mittel bei 1.31 statt bei 1.0. Praktisch bedeutet dies: Würde das Rasch-Modell mit einer für alle Items auf 1.0 fixierten Trennschärfe genutzt werden, um etwa für zwei Personengruppen (beispielsweise Jungen und Mädchen) zu schätzen, wie stark sie sich in ihrer mittleren Kompetenz unterscheiden, so würde auf Grundlage der SPF-Items eine größere Differenz ermittelt werden als auf Grundlage der Items, die für Schülerinnen und Schüler ohne SPF entwickelt wurden. Um verzerrte Varianz- und Differenzschätzungen zu vermeiden, wurde im IQB-Bildungstrend 2016 daher nicht das Rasch-Modell, sondern das 2pl-Modell angewendet und eine von 1 abweichende Trennschärfe für diejenigen Items zugelassen, die für Kinder mit SPF entwickelt wurden.

Zwischen den im IQB-Bildungstrend 2016 und im IQB-Ländervergleich 2011 verwendeten Skalierungsmodellen bestehen folglich Unterschiede. Um jedoch valide Tendaussagen treffen zu können, müssen die Testmodelle für die beiden Erhebungszeitpunkte äquivalent sein. Aus diesem Grund wurde im IQB-Bildungstrend 2016 kein freies (unrestringiertes) 2pl-Modell verwendet, bei dem für jedes einzelne Item sowohl ein Schwierigkeits- als auch ein Trennschärfeparameter geschätzt wird. Stattdessen wurde ein restringiertes 2pl-Modell spezifiziert, bei dem die Trennschärfen sämtlicher Items, die bereits im IQB-Ländervergleich 2011 zum Einsatz kamen, auf den Wert 1 fixiert

waren. Für die neu entwickelten SPF-Items wurde hingegen ein abweichender Trennschärfeparameter zugelassen, der jedoch gemäß der Modellspezifikation identisch für alle Items desselben Kompetenzbereichs war.

Parameterschätzung

Zur Schätzung der Modellparameter wurde das Paket TAM (Robitzsch, Kiefer & Wu, 2017) des Statistikprogramms R (R Core Team, 2015) verwendet, in das die Schätzmethode *Marginal Maximum Likelihood* (MML) implementiert ist. Bei der MML-Schätzung wird in der Regel eine Verteilungsannahme für die Personenfähigkeiten getroffen, üblicherweise die Normalverteilungsannahme. Im IQB-Bildungstrend 2016 erfolgte die Parameterschätzung, wie in internationalen Schulleistungsstudien üblich, in einem zweischrittigen Verfahren: Zunächst wurden die Itemschwierigkeiten geschätzt (Kalibrierung der Items) und anschließend wurden die Schülerinnen und Schüler auf der gemeinsamen Schwierigkeits- und Kompetenzskala verortet (Schätzung der Personenparameter).

Kalibrierung der Items

Unter der Kalibrierung von Items versteht man die Schätzung der Itemschwierigkeiten. Zu diesem Zweck wurde für jeden Kompetenzbereich und jede Leitidee separat ein eindimensionales restringiertes 2pl-Modell spezifiziert. Da sich die Zielgruppen der verschiedenen Teildesigns systematisch in ihrer mittleren Kompetenz unterscheiden und zugleich die Verteilung der Aufgaben und Testitems nicht zufällig oder gleichverteilt auf die Testhefte der jeweiligen Teildesigns erfolgte, musste bei Verwendung der MML-Schätzmethode bereits für die Kalibrierung ein Hintergrundmodell spezifiziert werden, um Verzerrungen zu vermeiden (DeMars, 2002). Ein Hintergrundmodell oder auch Populationsmodell beinhaltet in der Regel Informationen über Fähigkeitsunterschiede zwischen Gruppen in der Population. Bei der Schätzung der Itemparameter wird somit nicht von einer normalverteilten Fähigkeit in der Population ausgegangen, was bei den verschiedenen Zielgruppen im IQB-Bildungstrend 2016 nicht realistisch wäre, sondern es wird eine in Abhängigkeit der Hintergrundvariablen differenzierte Population zugrunde gelegt. Mögliche Verzerrungen in der Itemparameterschätzung können dadurch vermieden werden, da bei der Schätzung berücksichtigt wird, zu welchem Teildesign (vgl. Tab. 13.2) das entsprechende Testheft gehört (DeMars, 2002; M. von Davier, Gonzalez & Mislevy, 2009). Die so gewonnenen Itemparameterschätzungen bilden die Basis für die weiteren Auswertungsschritte, insbesondere für die Verlinkung mit dem IQB-Ländervergleich 2011.

Obwohl ein Großteil der Items in vorhergehenden Studien bereits erprobt werden konnte, wurden sämtliche Items nach der Kalibrierung nochmals auf ihre psychometrische Güte geprüft. Zur Einschätzung der psychometrischen Eignung eines jeden Items wurden die Kriterien Itemschwierigkeit, Trennschärfe und Modellpassung herangezogen. Die Prüfung erfolgte separat für jedes Teildesign. So wurde beispielsweise ein für die Gruppe der Schülerinnen und Schüler mit SPF entwickeltes Item, das in der Gruppe der Schülerinnen und Schüler ohne SPF eine schlechte Passung an das Modell aufwies, für diese Gruppe aus der Analyse ausgeschlossen.

Schätzung der Personenparameter

Zur Schätzung der Personenparameter stehen in der IRT mehrere Verfahren zur Verfügung. Ziel ist es, die Kompetenzwerte der Personen ohne systematische Verzerrungen mit der größtmöglichen Genauigkeit zu schätzen. Dabei ist nicht die Genauigkeit individueller Kompetenzwerte, sondern die Genauigkeit aggregierter Statistiken auf Gruppenebene entscheidend, wie zum Beispiel Mittelwerte und Standardabweichungen der Gesamtpopulation oder von Teilpopulationen. Schätzungen der Personenparameter, die auf Individualebene große Genauigkeit erreichen (z. B. *Weighted Likelihood Estimates* = WLE; Warm, 1989), führen mitunter zu verzerrten Schätzungen auf Gruppenebene. In großen Schulleistungsstudien hat sich deshalb der auf dem Prinzip der Multiplen Imputation (Rubin, 1987; Schafer & Graham, 2002) beruhende Plausible-Values-Ansatz bewährt, der zu erwartungstreuen Lage- und Dispersionsparameterschätzungen auf Gruppenebene führt (Mislevy, Beaton et al., 1992; M. von Davier et al., 2009). Anstelle eines einzelnen Kompetenzwertes wird hierbei für jede Person eine individuelle Wahrscheinlichkeitsverteilung ihrer Kompetenz modelliert, und es werden deren Lage- sowie Streuungsparameter geschätzt. Für jede Person werden dann mehrere „plausible Werte“ (*Plausible Values*) zufällig aus dieser individuellen Verteilung gezogen. Damit die Verteilung der Plausible Values möglichst genau die tatsächliche Verteilung der Kompetenzen der Schülerinnen und Schüler repräsentiert, muss in der Regel auch bei diesem Verfahren ein Hintergrundmodell spezifiziert werden, das Annahmen über die Populationsverteilung beinhaltet. Dabei gilt der Grundsatz, dass alle Variablen, über deren Zusammenhang mit den Kompetenzwerten in späteren Analysen Aussagen getroffen werden sollen, bereits bei der Plausible-Value-Ziehung im Hintergrundmodell zu berücksichtigen sind, um Verzerrungen zu vermeiden (z. B. Frey, Carstensen, Walter, Rönnebeck & Gomolka, 2008).

Die Daten, die in das Hintergrundmodell im IQB-Bildungstrend 2016 Eingang fanden, beruhen einerseits auf Angaben der Schulen und andererseits auf Angaben in den Schüler- und Elternfragebogen. Bei der Erfassung von Hintergrunddaten über Fragebogen kommt es unvermeidlich zu fehlenden Werten. In einigen Fällen fehlen die Angaben nicht vollständig zufällig, sondern bestimmte Personengruppen tendieren eher dazu, bestimmte Angaben auszulassen als andere. In solchen Fällen sollten zur Behandlung der fehlenden Werte statistische Verfahren zum Einsatz kommen, mit denen der Mechanismus modelliert werden kann, der für das Zustandekommen der fehlenden Werte verantwortlich ist (Lüdtke, Robitzsch, Trautwein & Köller, 2007). Hierzu eignen sich insbesondere imputationsbasierte Verfahren, in denen fehlende Werte ersetzt (imputiert) werden. Wie im IQB-Ländervergleich 2012 (Hecht et al., 2013) und im IQB-Bildungstrend 2015 (Sachse et al., 2016) wurde auch im IQB-Bildungstrend 2016 zur Imputation der Hintergrunddaten das Verfahren *Multivariate Imputation by Chained Equations* verwendet, das im Paket mice (van Buuren, 2007; van Buuren & Groothuis-Oudshoorn, 2011) für die Statistiksoftware R (R Core Team, 2015) implementiert ist. Bei diesem Verfahren werden die fehlenden Werte einer Variablen anhand der vorhandenen Informationen aller anderen Variablen geschätzt. Für jedes Land erfolgte die Imputation separat in einem Modell, in das unter anderem folgende Hintergrundmerkmale einbezogen wurden: Geschlecht, Schulart, Noten, Vorliegen eines sonderpädagogischen Förderbedarfs, familiäre Herkunftsmerkmale (u. a. Zuwanderungshintergrund, soziale Herkunft und Familiensprache), motivationale Merkmale (u. a. fachspezifisches Selbstkonzept und Interesse), ein Maß für kognitive Grundfähigkeiten sowie die Klassenmittel-

werte ausgewählter Variablen. Neben diesen Hintergrundmerkmalen wurden das Teildesign, die Kompetenzwerte der Schülerinnen und Schüler in Form von WLEs sowie die Klassenmittelwerte dieser Kompetenzwerte in die Imputation einbezogen, um ihre Zusammenhänge mit den Hintergrundvariablen zu berücksichtigen.

Einige der Hintergrundvariablen weisen hohe Abhängigkeiten untereinander auf (Multikollinearität). Werden multikollineare Variablen bei der Ziehung der Plausible Values in das Hintergrundmodell aufgenommen, führt dies häufig zu Schätzproblemen. Daher wurden die Einzelvariablen zunächst in einer Hauptkomponentenanalyse in zueinander unkorrelierte (orthogonale) Faktoren überführt. Konkret wurden nach der Imputation so viele Hauptkomponenten extrahiert, dass 95 Prozent der Gesamtvarianz aller Hintergrundvariablen erklärt werden konnte. Die Anzahl der extrahierten Hauptkomponenten liegt im IQB-Bildungstrend 2016 je nach Land zwischen 66 und 72. Bei der länderweisen Ziehung der Plausible Values wurden diese Hauptkomponenten als Hintergrundmodell in der Software TAM (Robitzsch et al., 2017) spezifiziert. Um für eine valide Trendschätzung sicherzustellen, dass die länderweise gezogenen Plausible Values länderübergreifend auf der Metrik des IQB-Ländervergleichs 2011 liegen, erfolgte eine Fixierung der Itemschwierigkeiten. Dazu wurden die mit dem Gesamtdatensatz geschätzten Itemparameter über ein lineares Transformationsverfahren auf die Logitmetrik des IQB-Ländervergleichs 2011 transformiert.

Zur Berücksichtigung der Ladungs- und Korrelationsstruktur zwischen den Kompetenzbereichen innerhalb eines Faches wurde in analoger Weise verfahren wie im IQB-Ländervergleich 2011. Im Fach Deutsch erfolgte die Bestimmung der Plausible Values separat für jedes Land in einem mehrdimensionalen Modell für die drei Kompetenzbereiche *Lesen*, *Zuhören* und *Orthografie* unter Einbeziehung der für das jeweilige Land extrahierten Hauptkomponenten. Dabei wurde jede Testaufgabe ausschließlich dem Kompetenzbereich zugeordnet, für dessen Messung sie konstruiert worden war. Diesem Vorgehen liegt die Annahme zugrunde, dass es sich bei den Kompetenzbereichen um klar trennbare, jedoch miteinander korrelierte Konstrukte handelt.

Da die Kompetenzbereiche (Leitideen) im Fach Mathematik nicht immer klar voneinander abzugrenzen sind (vgl. Kapitel 1.3), wurden die Plausible Values im Fach Mathematik für jedes Land und jede Leitidee separat in jeweils einem eindimensionalen Modell bestimmt, ohne die Mehrfachladungsstruktur explizit zu modellieren. Die mit diesem Vorgehen verbundene geringfügige Unterschätzung der latenten Korrelationen zwischen den einzelnen Kompetenzbereichen wurde in Kauf genommen, um andernfalls durch die Mehrfachladungsstruktur entstehende Probleme zu vermeiden (zur genaueren Begründung dieses Vorgehens siehe Kapitel 12, Abschnitt 12.2.2 in Weirich et al., 2012).

Bestimmung der Transformationsvorschrift für die Berichtsmetrik

Durch die Verwendung von IRT-Modellen liegen die berechneten Kompetenzwerte zunächst auf der Logit-Skala vor. Da der Wertebereich der Logit-Skala die gesamten reellen Zahlen umfasst und negative Werte im Falle von Testergebnissen nicht intuitiv interpretierbar sind, wird in den großen Bildungsstudien in der Regel eine Transformation dieser Skala vorgenommen, wobei die inhaltliche Bedeutung der Kompetenzskalen unverändert bleibt. Für Mittelwert

und Standardabweichung dieser Berichtsmetrik werden vorab festgelegte Werte für eine bestimmte Referenzpopulation definiert. Im IQB-Bildungstrend 2016 wurde dabei dieselbe Transformation wie im IQB-Ländervergleich 2011 gewählt: In jedem Fach und Kompetenzbereich lag in der Referenzpopulation des IQB-Ländervergleichs 2011 der Mittelwert bei 500 Punkten und die Standardabweichung bei 100 Punkten. Die Ergebnisse des IQB-Bildungstrends 2016 wurden auf dieser Metrik abgebildet.

Eine Ausnahme ist der Kompetenzbereich *Orthografie* im Fach Deutsch: Hier galt die Population der Viertklässlerinnen und Viertklässler im Jahr 2016 als Referenzpopulation (vgl. Kapitel 3.2). Der Grund hierfür ist, dass *Orthografie* im IQB-Ländervergleich 2011 in einer Stichprobe erhoben wurde, die keine Förderschulen beinhaltete und damit nicht vollständig repräsentativ für die Schülerinnen und Schüler in Deutschland war (vgl. Kapitel 3.1).

13.1.3 Trendschätzung

Die Datenstruktur des IQB-Bildungstrends 2016 in Verbindung mit dem IQB-Ländervergleich 2011 entspricht einem *Non-Equivalent Groups with Anchor Test Design* (Holland, 2007; Kolen & Brennan, 2004; A. von Davier et al., 2008). Hierbei werden Personengruppen aus verschiedenen Populationen über ein Set gemeinsamer Items (Ankeritems) verbunden. Im IQB-Bildungstrend 2016 geht es dabei um den Vergleich der Population der Viertklässlerinnen und Viertklässler im Jahr 2011 mit der Population der Viertklässlerinnen und Viertklässler im Jahr 2016. Die Ankeritems wurden in beiden Erhebungen in unveränderter Form vorgelegt und gewährleisteten eine gemeinsame Metrik.

Die Trendschätzung selbst erfolgte in mehreren Schritten und in praktisch äquivalenter Weise zum IQB-Bildungstrend 2015 (für eine ausführliche Beschreibung siehe Sachse et al., 2016).

Differenzielles Itemfunktionieren

Obwohl die Ankeritems in den Jahren 2011 und 2016 in derselben Form zum Einsatz kamen, ist nicht garantiert, dass ihre empirische Schwierigkeit zu beiden Testzeitpunkten konstant ist. Haben einzelne Ankeritems zum zweiten Testzeitpunkt eine (relativ zu den anderen Ankeritems) andere Schwierigkeit als zum ersten Testzeitpunkt, spricht man von Differenziellem Itemfunktionieren (*Differential Item Functioning* = DIF; Camilli, 1992; Penfield & Camilli, 2007; Sireci & Rios, 2013) beziehungsweise *Item Parameter Drift* (IPD). IPD kann zu systematisch verzerrten Trendschätzungen führen und entsteht beispielsweise, wenn im Jahr 2016 eine Kohorte von Schülerinnen und Schülern mit bestimmten Inhalten nicht mehr vertraut ist, die fünf Jahre zuvor noch allgemein bekannt waren und in der Aufgabenentwicklung verwendet wurden. Im Rahmen der Aufgabenentwicklung wird versucht, solche Probleme so weit wie möglich zu vermeiden. Im Einklang mit dem Vorgehen im IQB-Bildungstrend 2015 (Sachse et al., 2016) wurden Items, deren IPD signifikant über 0.64 oder signifikant unter -0.64 Logits lag (was hohem DIF gemäß der ETS-Klassifikation entspricht; siehe Penfield & Algina, 2006), von der Verlinkung ausgeschlossen, um potenzielle Verzerrungen der Trendschätzung durch IPD zu vermeiden.

Tabelle 13.3 weist für jeden Kompetenzbereich und jede Leitidee die Anzahl der zur Verlinkung genutzten Items (Ankeritems) aus. Ebenfalls ist angegeben, wie viele Items je Kompetenzbereich aufgrund von IPD von der Verlinkung ausgeschlossen wurden.

Tabelle 13.3: Anzahl der Items, Ankeritems und Größe des Linkfehlers je Fach und Kompetenzbereich auf der Berichtsmetrik

Fach	Kompetenzbereich	Anzahl Items gesamt	davon Ankeritems	davon mit IPD	Linkfehler
Deutsch	Lesen	155	67	4	2.54
	Zuhören	115	35	1	3.09
	Orthografie	148	104	1	2.47
Mathematik	Globalskala	400	250	6	1.06
	Zahlen und Operationen	120	81	1	1.38
	Raum und Form	75	52	2	1.82
	Muster und Strukturen	66	39	0	2.48
	Größen und Messen	117	66	0	1.13
	Daten, Häufigkeit und Wahrscheinlichkeit	78	60	4	2.73

Anmerkungen. IPD = Item Parameter Drift. Items des Fachs Mathematik, die zwei Leitideen zuzuordnen sind, werden in der Tabelle entsprechend mehrfach gezählt. Die Anzahl der Items aller fünf Leitideen summiert sich daher nicht zur Anzahl der Items für die Globalskala auf. Die Anzahl der tatsächlich zur Verlinkung genutzten Items entspricht der Anzahl der Ankeritems minus der Anzahl der Items mit IPD.

Quantifizierung der Unsicherheit der Trendschätzungen

Jede statistische Schätzung ist mit Unsicherheit verbunden, die verschiedene Ursachen haben kann. Für große Schulleistungsstudien spielen vor allem der Messfehler und der Stichprobenfehler eine zentrale Rolle. Für die Schätzung, wie sich die erreichten Kompetenzen zwischen den Jahren 2011 und 2016 verändert haben, muss darüber hinaus der Linkingfehler als eine dritte mögliche Unsicherheitsquelle berücksichtigt werden.

Im IQB-Bildungstrend 2016 setzt sich der Standardfehler $\hat{\sigma}_\mu$ für Kennwerte, die sich nur auf einen Zeitpunkt beziehen, aus Mess- und Stichprobenfehler zusammen:

$$\hat{\sigma}_\mu = \sqrt{\hat{\sigma}_{\mu_{meas}}^2 + \hat{\sigma}_{\mu_{samp}}^2}.$$

Für zwei Messzeitpunkte existieren demzufolge zwei Standardfehler $\hat{\sigma}_{\mu_{2011}}$ und $\hat{\sigma}_{\mu_{2016}}$. Für den Standardfehler von Trendschätzungen $\hat{\sigma}_{\mu_{2016}-\mu_{2011}}$ wurden diese beiden Fehlerkomponenten und zusätzlich der Linkingfehler $\hat{\sigma}_{\mu_{2016}-\mu_{2011}link}$ berücksichtigt:

$$\hat{\sigma}_{\mu_{2016}-\mu_{2011}} = \sqrt{\hat{\sigma}_{\mu_{2011}}^2 + \hat{\sigma}_{\mu_{2016}}^2 + \hat{\sigma}_{\mu_{2016}-\mu_{2011}link}^2}$$

Aus Platzgründen wird an dieser Stelle auf eine nähere Beschreibung der unterschiedlichen Unsicherheitsquellen und deren Berücksichtigung in den Berechnungen verzichtet. Eine ausführliche Beschreibung hierzu kann jedoch dem technischen Kapitel im Berichtsband zum IQB-Bildungstrend 2015 (Sachse et al., 2016) entnommen werden.

Literatur

- Adams, R. J. & Wu, M. L. (2007). The mixed-coefficients multinomial logit model: A generalized form of the Rasch model. In M. von Davier & C. H. Carstensen (Hrsg.), *Multivariate and mixture distribution Rasch models* (S. 57–75). New York: Springer.
- Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Hrsg.), *Statistical theories of mental test scores* (S. 395–479). Reading, MA: Addison-Wesley.
- Camilli, G. (1992). A conceptual analysis of differential item functioning in terms of a multidimensional item response model. *Applied Psychological Measurement*, 16(2), 129–147.
- DeMars, C. E. (2002). Incomplete data and item parameter estimates under JMLE and MML estimation. *Applied Measurement in Education*, 15, 15–31.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. 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. 375–397). Münster: Waxmann.
- Frey, A., Hartig, J. & Rupp, A. A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice*, 28(3), 39–53.
- Gonzalez, E. & Rutkowski, L. (2010). Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. *IEA-ETS Research Institute Monograph*, 3, 125–156.
- Granzer, D., Köller, O., Bremerich-Vos, A., van den Heuvel-Panhuizen, M., Reiss, K. & Walther, G. (Hrsg.). (2009). *Bildungsstandards Deutsch und Mathematik*. Weinheim: Beltz.
- Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park: Sage.
- Hecht, M., Roppelt, A. & Siegle, T. (2013). Testdesign und Auswertung des Ländervergleichs. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 391–402). Münster: Waxmann.
- Holland, P. W. (2007). A framework and history for score linking. In N. J. Dorans, M. Pommerich & P. W. Holland (Hrsg.), *Linking and aligning scores and scales* (S. 5–30). New York: Springer.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practice*. New York: Springer.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung: Probleme und Lösungen. *Psychologische Rundschau*, 58, 103–117.
- 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.
- Mislevy, R. J., Johnson, E. G. & Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Educational and Behavioral Statistics*, 17(2), 131–154.
- OECD. (2009). *PISA 2006 technical report*. Paris: OECD.
- OECD. (2012). *PISA 2009 technical report*. Retrieved from <http://dx.doi.org/10.1787/9789264167872-en>
- Penfield, R. D. & Algina, J. (2006). A generalized DIF effect variance estimator for measuring unsigned differential test functioning in mixed format tests. *Journal of Educational Measurement*, 43, 295–312.
- Penfield, R. D. & Camilli, G. (2007). Differential item functioning and item bias. In C. R. Rao & S. Sinharay (Hrsg.), *Handbook of statistics* (S. 125–168). New York: Elsevier.
- R Core Team. (2015). R: A language and environment for statistical computing (Version 3.2.2). Wien: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

- Robitzsch, A., Kiefer, T. & Wu, M. (2017). TAM: Test analysis modules (Version R package version 2.0-37). Retrieved from <https://CRAN.R-project.org/package=TAM>
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Sachse, K., Haag, N. & Weirich, S. (2016). Testdesign und Auswertung des IQB-Bildungstrends 2015: Technische Grundlagen. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015* (S. 509–526). Münster: Waxmann.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Sireci, S. G. & Rios, J. A. (2013). Decisions that make a difference in detecting differential item functioning. *Educational Research and Evaluation*, 19(2-3), 170–187.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik*. Münster: Waxmann.
- Tarelli, I., Wendt, H., Bos, W. & Zylowski, A. (2012). Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2011). In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–67). Münster: Waxmann.
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16, 219–242.
- van Buuren, S. & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1–67.
- von Davier, A., Carstensen, C. H. & von Davier, M. (2008). Linking competencies in horizontal, vertical, and longitudinal settings and measuring growth. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 121–149). New York: Hogrefe & Huber.
- von Davier, M., Gonzalez, E. & Mislevy, R. J. (2009). What are plausible values and why are they useful? *IERI Monograph Series: Issues and Methodologies in Large Scale Assessments*, 2, 9–36.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Weirich, S., Haag, N. & Roppelt, A. (2012). Testdesign und Auswertung des Ländervergleichs: Technische Grundlagen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik* (S. 277–290). Münster: Waxmann.

13.2 Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

Katrin Böhme, Pauline Kohrt, Nicole Haag und Sebastian Weirich

Im IQB-Bildungstrend 2016 wurde überprüft, inwieweit Schülerinnen und Schüler der 4. Jahrgangsstufe zentrale Kompetenzziele in den Fächern Deutsch und Mathematik erreichen. Hierbei sollten auch die Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) differenziert erfasst werden, um für das Bildungsmonitoring Informationen zur Verfügung zu stellen, inwieweit die Förderung dieser Schülerinnen und Schüler gelingt (vgl. Kapitel 10.1). Zur Erfassung der Kompetenzen von Schülerinnen und Schülern mit SPF sollten nach Möglichkeit angepasste Testaufgaben eingesetzt werden, die den besonderen kognitiven und motivationalen Voraussetzungen dieser Kinder besser gerecht werden als die bislang zur Kompetenzmessung verwendeten Tests (Elliott, Davies & Kettler, 2012; Weigert, 2012). Solche angepassten Testaufgaben wurden für den IQB-Bildungstrend 2016 für die Kompetenzbereiche *Lesen* und *Zuhören* im Fach Deutsch sowie für das Fach Mathematik entwickelt. Das vorliegende Kapitel erläutert die theoretische Fundierung der Konstruktion dieser angepassten Testaufgaben, beschreibt den Entwicklungsprozess und illustriert die resultierenden Testaufgaben anhand von Beispielen. Ferner werden die Messeigenschaften der angepassten Testaufgaben im IQB-Bildungstrend 2016 dargestellt und es wird diskutiert, inwieweit sie die in den Bildungsstandards verankerten Konstrukte in ähnlicher Weise erfassen wie die regulären Testaufgaben der IQB-Bildungstrendstudien im Primarbereich.

13.2.1 Empirische und konzeptionelle Grundlagen der Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

Um die Berücksichtigung von Schülerinnen und Schülern mit SPF im nationalen Bildungsmonitoring besser gewährleisten zu können, muss sichergestellt werden, dass die in den Erhebungen eingesetzten Testverfahren auch für diese Zielgruppe genaue und zuverlässige Aussagen über die Ausprägung der interessierenden Kompetenzen gestatten. Dies ist nach aktuellen empirischen Erkenntnissen nur eingeschränkt der Fall (vgl. Müller, 2014).

So berichten beispielsweise Südkamp, Pohl, Hardt, Jordan und Duchhardt (2015) eine geringe Passung zwischen den Aufgabenschwierigkeiten im Nationalen Bildungspanel (*National Educational Panel Study*, NEPS) und den Kompetenzen von Schülerinnen und Schülern mit SPF: Die regulären NEPS-Aufgaben erwiesen sich für diese Schülerinnen und Schüler als deutlich zu schwer. Dies führt zu einer geringen Präzision der Kompetenzschätzungen und zu einer unzureichenden Differenzierung der Kompetenzen, über die Schülerinnen und Schüler mit SPF verfügen.

Für die auf den Bildungsstandards basierenden Aufgaben, die im IQB-Ländervergleich 2011 in der Primarstufe in den Fächern Deutsch und Mathematik eingesetzt wurden, konnte zwar nachgewiesen werden, dass sie prinzipiell geeignet sind, die Kompetenzen von Kindern mit SPF valide und reliabel zu erfassen.

sen, aber auch hier zeigten sich bestimmte Einschränkungen (Kocaj et al., 2016). Kritisch ist insbesondere, dass der Anteil fehlender Werte bei Schülerinnen und Schülern mit SPF in dieser Studie vergleichsweise groß war, was vermutlich darauf hinweist, dass diese Kinder die Aufgaben aufgrund von Zeitmangel nicht bis zum Schluss bearbeiten konnten. Ferner stellten Kocaj et al. (2016) eine geringe Differenzierungsfähigkeit im relevanten Kompetenzspektrum fest. Auch im IQB-Ländervergleich 2011 erwiesen sich somit reguläre Testaufgaben, die nicht gezielt für Kinder mit SPF entwickelt und erprobt wurden, als zu schwer und waren nur eingeschränkt geeignet, um die von diesen Schülerinnen und Schülern erreichten Kompetenzen akkurat und differenziert zu erfassen.

Es ist jedoch davon auszugehen, dass Schülerinnen und Schüler mit SPF im Hinblick auf die von ihnen erworbenen Kompetenzen eine sehr heterogene Gruppe darstellen und sie in unterschiedlichem Maße zumindest teilweise über die in den Bildungsstandards beschriebenen Kompetenzen verfügen und somit zum Beispiel in der Lage sind, aus Lesetexten Informationen zu entnehmen und Zusammenhänge zu erschließen. Aufgrund ihres Förderbedarfs ist aber für einige Kinder der kognitive Zugang zu den Stimuli in den Testaufgaben, also das Verständnis der Lese- und Hörtexte sowie der Zugang zu den Aufgabenstellungen erschwert. Daher können Schülerinnen und Schüler mit SPF ihre vorhandenen Kompetenzen in diesen Tests möglicherweise nicht vollständig zeigen. Die üblicherweise in Schulleistungsstudien eingesetzten Testverfahren sind zudem aufgrund des geringen Umfangs von Items im unteren Leistungsbereich in der Regel nur begrenzt geeignet, zwischen unterschiedlich schwachen Lernerinnen und Lernern zu differenzieren (Müller 2014).

Ziel der Entwicklung angepasster Testaufgaben für den IQB-Bildungstrend 2016 war es daher, Aufgaben zu entwickeln, die für Schülerinnen und Schüler mit SPF besser zugänglich und insgesamt eher leicht sind. Diese Testaufgaben sollen ihnen besser ermöglichen, vorhandene Kompetenzen zu zeigen, ohne während der Testbearbeitung auf Barrieren zu stoßen, die in der Formulierung der Instruktion oder der Gestaltung der Aufgaben liegen. Zu diesem Zweck wurden die Rezeption von Stimuli und Aufgaben erleichtert und die kognitiven Anforderungen durch sprachliche und visuelle Adaptionen verringert. Hierdurch soll die Validität der Kompetenzmessung für die Zielgruppe der Schülerinnen und Schüler mit SPF optimiert und die Differenzierung im unteren Leistungsbereich erhöht werden.

Bei der Entwicklung angepasster Testaufgaben für den IQB-Bildungstrend 2016 wurde auf die Theorie der kognitiven Belastung (*Cognitive Load Theory*, Chandler & Sweller, 1991) und die Theorie der Zugänglichkeit (*Accessibility Theory*, Beddow, 2012) zurückgegriffen. Die Theorie der kognitiven Belastung wurde ursprünglich entwickelt, um das schulische Lernen durch eine optimierte Unterrichtsgestaltung zu unterstützen und Empfehlungen zur lernförderlichen Gestaltung von Lernmaterialien zu unterbreiten. Sie geht unter anderem davon aus, dass in Lernsituationen die begrenzte Kapazität des Arbeitsgedächtnisses berücksichtigt werden muss. Eine Lernsituation sollte demnach so gestaltet werden, dass unnötige Belastungen des Arbeitsgedächtnisses vermieden werden.

Die Theorie der Zugänglichkeit beschreibt das Zusammenspiel von Merkmalen eines Tests mit Merkmalen von Personen, die diesen Test bearbeiten. Es geht dabei um die Frage, welche Merkmale eines Tests dazu führen können, dass Testteilnehmerinnen und Testteilnehmer ihre Kompetenzen nicht vollständig demonstrieren können, sodass die Validität des Testergebnisses eingeschränkt ist.

Anhand der Theorie der kognitiven Belastung lassen sich die Annahmen der Accessibility Theory weiter konkretisieren und Strategien zur Erleichterung des Zugangs zu Testaufgaben entwickeln.

Um die Zugänglichkeit von Testaufgaben zu erhöhen, können verschiedene Strategien eingesetzt werden. Diese beziehen sich unter anderem auf die sprachliche Gestaltung der eingesetzten Stimuli (bspw. Lese- und Hörtexte) und Iteminstruktionen, auf Visualisierungen und auf das Layout der Testinstrumente, also die Anordnung von Aufgaben, Items und Visualisierungen auf den Testheftseiten (Beddow, 2012). Für den englischen Sprachraum gibt es zahlreiche Empfehlungen, wie sprachliche Barrieren in Testaufgaben vermieden werden können (vgl. Kettler, Elliott & Beddow, 2009). Im Hinblick auf die sprachliche Gestaltung zugänglicher Testaufgaben wird unter anderem empfohlen, sowohl den Stimulus als auch die Items möglichst klar zu formulieren und Texte in kurze Abschnitte zu gliedern (Kettler et al., 2009). Ferner sollten Pronomen, Passivformen und Verneinungen vermieden werden (vgl. Clark, Nguyen, & Sweller, 2006; Haladyna, Downing, & Rodriguez, 2002; Kettler et al., 2011; Mayer & Moreno, 2003).

Viele der regulären Aufgaben in Schulleistungstests umfassen Inhalte, die nicht unmittelbar für die Erfassung des Zielkonstrukts relevant sind. Beispielsweise können Abbildungen neben Lesetexten platziert sein, die den Test ansprechender gestalten, ohne für das Verständnis des Textes erforderlich zu sein. Für Schülerinnen und Schüler ohne SPF sind solche Abbildungen in der Regel motivierend und stellen keine zusätzliche kognitive Belastung dar. Für Schülerinnen und Schüler, die Probleme mit der Aufmerksamkeitssteuerung haben, können sie hingegen ablenkend wirken und dadurch zu einer Barriere werden, die das Textverstehen und die Bearbeitung der Aufgaben erschwert (Elliott et al., 2012).

Wie die genannten Prinzipien bei der Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit SPF für den IQB-Bildungstrend 2016 umgesetzt wurden, ist im folgenden Abschnitt dargestellt.

13.2.2 Entwicklung und Pilotierung angepasster Testaufgaben für den IQB-Bildungstrend 2016 in den Fächern Deutsch und Mathematik

Die Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit SPF erfolgte im Zeitraum von Dezember 2014 bis Juli 2015 und orientierte sich eng am Vorgehen und an den Erfahrungen bei der Entwicklung regulärer Testaufgaben für die Bildungstrendstudien des IQB. Damit sollte sichergestellt werden, dass die Struktur der angepassten Testaufgaben der Struktur der regulären Testaufgaben so weit wie möglich ähnelt, sofern dies für die Zielgruppe der Schülerinnen und Schüler mit SPF sinnvoll erschien. Detaillierte Informationen zu den regulären Testaufgaben des IQB-Bildungstrends 2016 finden sich in den Kapiteln 1.2 und 1.3 dieses Berichts.

An der Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit SPF wirkten sowohl fachdidaktische Expertinnen und Experten für die Fächer Deutsch und Mathematik im Primarbereich als auch Expertinnen und Experten für Sonderpädagogik und Inklusionspädagogik mit (vgl. Kapitel 3.1). Alle beteiligten Personen wurden in umfangreichen Schulungen mit den wesentlichen Grundsätzen der Konstruktion bildungsstandardbasierter Testaufgaben vertraut gemacht.

Anpassungen im Fach Deutsch

Für die Kompetenzbereiche *Lesen* und *Zuhören* stand die Entwicklung angepasster Stimuli, also angepasster Lese- und Hörtexte, im Mittelpunkt.¹ Hierfür wurden gezielt Texte verfasst und authentische Texte systematisch modifiziert. Die bei der Konstruktion der angepassten Testaufgaben angewendeten Strategien orientierten sich unter anderem an Erkenntnissen zur Entwicklung zugänglicher Leistungstests (Elliott, Kettler, Beddow & Kurz, 2011) sowie an empirischen Befunden zu (bildungs-)sprachlichen Merkmalen, die das Verständnis von Texten beeinflussen (vgl. z. B. Heppt, 2016; Schleppegrell, 2002). Die schriftsprachlichen Anpassungen in den Testaufgaben betrafen sowohl semantisch-lexikalische Aspekte, die auf die Wortbedeutung bezogen waren, als auch morphologisch-syntaktische Aspekte der Wort- und Satzbildung sowie phonetisch-phonologische Aspekte der Sprachlaute. Die sprachlichen Anpassungen lehnten sich auch an Erkenntnisse zu verständniserleichternden beziehungsweise -erschwerenden sprachlichen Merkmalen aus den Forschungsbereichen der Bildungssprache (vgl. Heppt, 2016) sowie der Leichten Sprache an. Informationen zum Konzept der Leichten Sprache sind in einem Exkurs in Abbildung 13.1 dargestellt.

Abbildung 13.1: Exkurs – Leichte Sprache

Leichte Sprache ist eine Form der barrierefreien Kommunikation und soll Menschen mit kognitiven oder sensorischen Einschränkungen den eigenständigen und selbst gesteuerten Zugang zu aller Art von Information ermöglichen (Maaß, 2015). Leichte Sprache ist ein auch gesetzlich verankertes Mittel, um bestimmten Zielgruppen Informationen und Kommunikation zugänglich zu machen (Bock, 2015). Um die informationelle Selbstbestimmung gewährleisten zu können, wird in Texten, die in Leichte Sprache übersetzt wurden, die Informationsdichte reduziert, außerdem greifen entsprechende Texte auf Beispiele und Erklärungen zurück. Texte in Leichter Sprache verzichten auf Passivkonstruktionen sowie komplexe und abstrakte Nominalphrasen und benennen Handlungsträger in aktiv formulierten Sätzen (Maaß, 2015). Komposita können durch Bindestriche oder Mediopunkte in leichter erfassbare Wortbausteine zerlegt werden (Oomen-Welke, 2015).

In der öffentlichen Diskussion wird Leichte Sprache häufig kritisch hinterfragt und mit dem Vorwurf konfrontiert, sie würde zu einem Verfall oder zumindest einer Verflachung der Sprache beitragen. Befürworter Leichter Sprache argumentieren allerdings, dass sie dazu beiträgt, Menschen mit Einschränkungen eine Teilhabe in allen Bereichen des Alltags und der Kultur zu ermöglichen. Texte in Leichter Sprache „erschließen den Leserinnen und Lesern Gegenstände, zu denen sie andernfalls keinen Zugang hätten“ (Maaß, 2015, S. 7). Ferner hat Leichte Sprache eine Brückenfunktion. Im Sinne von

¹ Bei der Entwicklung angepasster Testaufgaben für das Fach Deutsch lag der Fokus zunächst auf diesen zwei Kompetenzbereichen. Für den Kompetenzbereich *Orthografie* wurden auf Basis empirischer Informationen aus früheren Studien des IQB reguläre Aufgaben und Items ausgewählt, die sich bereits für Schülerinnen und Schüler der 3. Jahrgangsstufe als sehr einfach erwiesen hatten und die über gute Messeigenschaften verfügten. Diese Aufgaben wurden für den IQB-Bildungstrend 2016 in separaten Blöcken zusammengestellt.

Feilke (2015) ist sie eine transitorische Norm, die nicht als Ziel, sondern als unterstützendes Mittel im Sinne eines *Scaffolding* der Kompetenzentwicklung zu verstehen ist.

Wie bei der sprachlichen Anpassung konkret vorgegangen wurde, soll nachfolgend anhand der Abbildungen 13.2 bis 13.4 mit Hilfe eines Beispiels für den Kompetenzbereich *Lesen* erläutert werden (Gerlach et al., 2016). Bei dem gewählten Textauszug handelt es sich nicht um eine im IQB-Bildungstrend 2016 eingesetzte Testaufgabe, sondern um einen als Beispiel konstruierten Text, der einige der angewendeten Anpassungsprinzipien veranschaulichen soll.²

In Abbildung 13.2 sind Anpassungen auf Ebene des gesamten Textes sowie auf semantisch-lexikalischer Ebene illustriert. Anpassungen auf phonetisch-phonologischer Ebene sowie morphologisch-syntaktischer Ebene sind in Abbildung 13.3 markiert.³ Abbildung 13.4 zeigt den nach den Anpassungen resultierenden Lesetext. Dieser ist wesentlich kürzer und sprachlich stark vereinfacht. Die wesentlichen Inhalte des Textes sind jedoch erhalten geblieben.

Abbildung 13.2: Markierung von Anpassungen auf Textebene sowie auf semantisch-lexikalischer Ebene

<p>Heute war ein wunderbarer Ferientag. Nicht nur, dass den ganzen Tag die Sonne lachte. Nein, wir waren auch fröhlich und <u>ausgelassen</u>. Schon bevor ich heute aufstand, saß ein Vögelchen an meinem Fenster und piepste mich ganz sanft wach. Ich zog meine Schlappen an und rannte gut gelaunt in die Stube, wo schon meine Eltern beim Frühstück saßen. Und dann wurde ich überrascht: Sie hatten einen Ausflug zum Bauernhof geplant. Da wollte ich schon immer mal hin! Bevor ich mich anzog, tanzte ich vor Freude durch das ganze Wohnzimmer ...</p>	<ul style="list-style-type: none"> • irrelevante Informationen • weniger gebräuchliche Worte • Rückverweise • Synonyme
---	--

Auf der Textebene sowie der semantisch-lexikalischen Ebene (vgl. Abb. 13.2) wurden zunächst Informationen entfernt, die für das Textverständnis und die Aufgabenbearbeitung irrelevant sind, da sie weder direkt erfragt werden noch für das Ziehen von Schlüssen herangezogen werden müssen (grau hinterlegt). Außerdem wurde auf die Verwendung von Vor- und Rückverweisen verzichtet (durch Kästen und Pfeile markiert). Ferner achteten die Aufgabenentwicklerinnen und -entwickler auf die Konsistenz von Begriffen und verwendeten daher keine Synonyme (durch gepunktete Ovale gekennzeichnet). Zudem kamen nur hochfrequente Begriffe zum Einsatz, deren zu erwartender Bekanntheitsgrad für Grundschulkinder über die *ChildLex*-Datenbank (Schroeder, Würzner, Heister,

2 Für die Abschätzung von Trends müssen die Testaufgaben in mehreren aufeinanderfolgenden Bildungstrendstudien unter vergleichbaren Bedingungen eingesetzt werden. Da eine valide Kompetenzschätzung voraussetzt, dass die Aufgaben den Schülerinnen und Schülern unbekannt sind, können diese nicht veröffentlicht werden. Die angepassten Testaufgaben, die im IQB-Bildungstrend 2016 eingesetzt wurden, sind aber äquivalent zu der hier dargestellten Aufgabe gestaltet.

3 In dem Beispiel wurde auf den verschiedenen sprachlichen Ebenen jeweils nur eine Auswahl der angepassten Elemente markiert, um die Konstruktionsprinzipien nachvollziehbar zu machen. Es handelt sich nicht um eine vollständige Darstellung aller in den neu entwickelten Testaufgaben umgesetzten Anpassungen. Beispielsweise wurden in einigen Lesetexten Abschnitte durch ein geeignetes Layout besonders klar abgesetzt, was im hier gewählten Beispiel nicht illustriert werden konnte.

Geyken & Kliegl, 2015) geprüft wurde (Gegenbeispiele mit gestrichelter Linie unterstrichen).

Abbildung 13.3: Markierung von Anpassungen auf phonetisch-phonologischer Ebene und morphologisch-syntaktischer Ebene

<p>Heute war ein wunderbarer <u>Ferientag</u>. Nicht nur, dass den ganzen Tag die Sonne lachte. Nein, wir waren auch fröhlich und ausgelassen. Schon bevor ich heute aufstand, saß ein Vögelchen an meinem Fenster und <u>piepste</u> mich ganz <u>sansf</u> wach. Ich zog meine Schlappen an und rannte gut gelaunt in die Stube, wo schon meine Eltern beim Frühstück saßen. Und dann wurde ich überrascht: Sie <u>hatten</u> einen Ausflug zum <u>Bauernhof</u> <u>geplant</u>. Da wollte ich schon immer mal hin! <u>Bevor ich mich</u> <u>anzög</u>, tanzte ich vor Freude durch das ganze <u>Wohnzimmer</u> ...</p>	<ul style="list-style-type: none"> • <u>Komposita</u> • komplexe Satzstruktur • <u>Konsonantencluster</u> • <u>weite Verbklammer</u> • <u>Temporalsatz</u>
--	--

Auf phonetisch-phonologischer Ebene (vgl. Abb. 13.3) vermieden die Aufgabenentwicklerinnen und -entwickler Wörter mit Konsonantenclustern (durch gepunktete Einrahmung markiert) sowie nicht lautgetreue Wörter, um schwachen Leserinnen und Lesern das Dekodieren zu erleichtern. Auf morphologisch-syntaktischer Ebene wurde auf Passivformen und Negativformulierungen sowie Temporalsätze (bspw. mit „bevor“, durch gestrichelte Einrahmung markiert) verzichtet. Ferner wurden vorrangig Nominativ- und Dativkonstruktionen verwendet und möglichst oft einfache Satzkonstruktionen eingesetzt (Konstruktion mit der Reihenfolge Subjekt-Verb-Objekt; Gegenbeispiel fett gepunktet unterstrichen). Zudem wurde auf enge Verbklammern (Gegenbeispiel eingerahmt) und die Verwendung von Präsensformen geachtet.

In Anlehnung an die Prinzipien der Leichten Sprache (vgl. Abb. 13.1) wurde die Bindestrich-Trennung von Komposita eingesetzt, allerdings nur dann, wenn die Einzelbestandteile einen höheren Bekanntheitsgrad hatten als das Kompositum selbst (Komposita in Abb. 13.3 durch Ovale gekennzeichnet). Da diese Anpassung einen Eingriff in die orthografische Korrektheit der Kompositaschreibung bedeutet, erfolgte eine kritische Abwägung, welche Auswirkungen solche Schreibungen auf die Kompetenzentwicklung der Schülerinnen und Schüler haben könnten. Da Kompositaschreibungen mit Mediopunkt oder Bindestrich, die mitunter auch in Büchern für Leseanfänger eingesetzt werden (vgl. Bertram & Schulmeyer, 2014), hier aber eine Brückenfunktion haben, wurde darauf als unterstützendes Mittel im Sinne eines Scaffolding zurückgegriffen. Um die Kinder während der Testung nicht mit abweichenden Schreibungen zu verunsichern, wurden alle Kinder zu Testbeginn darüber informiert, dass diese in einzelnen Aufgaben vorkommen können. Angepasste Testaufgaben, die abweichende Schreibungen enthielten, waren zudem in den Testheften mit einem Symbol markiert.

Abbildung 13.4: Resultierender Text

Heute ist ein toller Ferien-Tag. Die Sonne lacht und wir sind fröhlich. Ich wache auf, ziehe meine Haus-Schuhe an und renne gut gelaunt ins Wohnzimmer. Im Wohnzimmer sitzen schon meine Eltern und essen. Sie sagen: „Wir machen einen Ausflug zum Bauern-Hof.“ Ich wollte schon immer mal zum Bauern-Hof! Ich tanze vor Freude durch das ganze Wohnzimmer. Dann ziehe ich mich an. ...

Anpassungen im Fach Mathematik

Bei der Entwicklung angepasster Testaufgaben zur Erfassung mathematischer Kompetenzen wurde insbesondere darauf geachtet, sehr einfache mathematische Inhalte zu operationalisieren. Hierbei kamen verschiedene Strategien zum Einsatz, die sich zum einen auf das Zahlenmaterial und zum anderen auf die Strukturierung der Inhalte und Anforderungen bezogen. Beispielsweise wurde der Zahlenraum, in dem die Schülerinnen und Schüler verschiedene Rechenoperationen durchführen sollten, so gewählt, dass beim Lösen der Aufgabe keine Zehnerübergänge ($6 + 3$ statt $7 + 5$) notwendig wurden. Auch die Struktur verbaler oder als Gleichung dargestellter Arbeitsaufträge wurde vereinfacht, indem der mathematische Anfangszustand und nicht der Zielzustand ($A + B = ?$ statt $? + B = C$) dargestellt wurde. Nachfolgend finden sich in den Abbildungen 13.5 und 13.6 zwei Beispiele für angepasste Testaufgaben in den Kompetenzbereichen *Zahlen und Operationen* sowie *Größen und Messen*. Das Beispiel für den Kompetenzbereich *Zahlen und Operationen* in Abbildung 13.5 illustriert die beiden genannten mathematischen Aspekte: die Wahl des Zahlenmaterials und die bewusst einfachere Strukturierung der Gleichungen im Stimulus.⁴


Darüber hinaus wurde das Prinzip umgesetzt, sprachliche Anforderungen der Mathematikaufgaben so einfach wie möglich zu halten. Ähnlich wie bei den im Fach Deutsch vorgenommenen Anpassungen wurden für die Aufgabenstellungen kurze Sätze formuliert, vorrangig bekannte und einfache Wörter genutzt und Aufzählungen vermieden (vgl. Abb. 13.5 und Abb. 13.6). Auch bei den Antworten der Kinder wurde der Umfang der erforderlichen schriftlichen Sprachproduktion möglichst gering gehalten. So wurden beispielsweise weder freie Begründungen noch Antwortsätze verlangt und die Kinder konnten ihre Antworten an markierten Stellen in zum Teil vorformulierten Sätzen eintragen (vgl. Abb. 13.6). Ferner wurde das Verständnis der Aufgabenstellung durch Visualisierungen unterstützt. Auf Bilder und Visualisierungen ohne verständnisfördernde Funktion wurde verzichtet.


⁴ Die hier dargestellten Beispielaufgaben wurden in der Pilotierung der angepassten Aufgaben im Herbst 2015 empirisch erprobt, aber nicht im IQB-Bildungstrend 2016 eingesetzt.


Abbildung 13.5: Beispiel für eine angepasste Testaufgabe für den Kompetenzbereich *Zahlen und Operationen* im Fach Mathematik


Welche Aufgabe passt? Kreuze an.

17 Äpfel liegen auf dem Tisch. Die Kinder essen 4 Äpfel.

 - 17 = 4

17 + 4 = 


 - 4 = 17

17 - 4 = 




Das Beispiel für den Kompetenzbereich *Zahlen und Operationen* (Abb. 13.5) zeigt ein geschlossenes Item im *Multiple-Choice*-Format, bei dem die Schülerinnen und Schüler aus vier vorgegebenen Optionen die passende – hier die vierte – Option auswählen sollen. Für das Lesen des Stimulus („Welche Aufgabe passt? Kreuze an.“) und die Bearbeitung dieses Items wurde eine Bearbeitungszeit von 90 Sekunden kalkuliert.




Abbildung 13.6: Beispiel für eine angepasste Testaufgabe für den Kompetenzbereich *Größen und Messen* im Fach Mathematik

Was kosten die Spielsachen? Schreibe auf.



Preis:

Die Puppe kostet _____ Euro.

Das Beispiel für den Kompetenzbereich *Größen und Messen* (Abb. 13.6) zeigt ein halboffenes Item, bei dem die Schülerinnen und Schüler eine Zahl angeben sollen. Die korrekte Antwort kann – wie in allen äquivalenten Items im Fach Mathematik – entweder als Ziffer („6“) oder als Zahlwort („sechs“) notiert werden. Im letzteren Fall spielt die orthografische Richtigkeit der Antwort keine Rolle. Für das Lesen des Stimulus („Was kosten die Spielsachen? Schreibe auf.“) und die Bearbeitung dieses Items wurde eine Bearbeitungszeit von zwei Minuten veranschlagt.

Pilotierung der angepassten Testaufgaben

Im Fach Deutsch wurden sieben angepasste Aufgaben mit insgesamt 59 Items für den Kompetenzbereich *Lesen* und ebenfalls sieben Aufgaben mit insgesamt 39 Items für den Kompetenzbereich *Zuhören* entwickelt. Für die Erfassung der mathematischen Kompetenzen konnten 64 angepasste Aufgaben mit insgesamt 96 Items entwickelt werden, die sich auf alle fünf inhaltlichen Kompetenzbereiche verteilten.

Die neu entwickelten Testaufgaben wurden im Herbst 2015 in den Ländern Berlin, Bremen und Nordrhein-Westfalen zu Beginn der 5. Jahrgangsstufe pilotiert.⁵ An der Studie nahmen 1461 Schülerinnen und Schüler in allgemeinen Schulen und Förderschulen teil. An den allgemeinen Schulen wiesen 171 Schülerinnen und Schüler einen SPF in den Bereichen „Lernen“, „Sprache“ oder „emotionale und soziale Entwicklung“ auf. An Förderschulen nahmen 645 Schülerinnen und Schüler an der Studie teil. Die Schülerinnen und Schüler waren im Durchschnitt 10,8 Jahre alt. Der Jungenanteil in der Stichprobe lag bei 59 Prozent, und 26 Prozent der teilnehmenden Kinder hatten einen Zuwanderungshintergrund (mindestens ein Elternteil im Ausland geboren).

In dieser Pilotierungsstudie waren die psychometrischen Eigenschaften der angepassten Testaufgaben konsistent mit dem angenommenen Messmodell (Rasch-Modell), sodass die Aufgaben im IQB-Bildungstrend 2016 zum Einsatz kommen konnten. Die anhand der Daten des IQB-Bildungstrends 2016 durchgeführten psychometrischen Analysen werden in den nachfolgenden Abschnitten dargestellt.

13.2.3 Empirische Befunde zu psychometrischen Eigenschaften, Missinganteilen und zur Konstruktäquivalenz regulärer und angepasster Testaufgaben im IQB-Bildungstrend 2016

Methodisches Vorgehen

Reguläre Testaufgaben sind für die Messung der Kompetenzen von Schülerinnen und Schülern mit SPF nur bedingt geeignet (Kocaj et al., 2016; Südkamp et al., 2015; vgl. auch Abschnitt 13.2.1). Vor diesem Hintergrund stellt sich die Frage, ob die für den IQB-Bildungstrend 2016 entwickelten angepassten Testaufgaben die Kompetenzen von Grundschulkindern mit SPF besser erfassen als die regulären Testaufgaben. Zur Bestimmung der psychometrischen Eigenschaften der Aufgaben wurden zunächst die beiden Aufgabensets (angepasste und reguläre Aufgaben) in drei Schülergruppen jeweils separat analysiert. Diese Gruppen sind (1) Schülerinnen und Schüler ohne SPF, (2) Schülerinnen und Schüler mit SPF an allgemeinen Schulen und (3) Schülerinnen und Schüler mit SPF an Förderschulen. Als Messmodell wurde dabei jeweils das eindimensionale Rasch-Modell zugrunde gelegt. Folgende psychometrische Eigenschaften wurden untersucht: die mittlere Schwierigkeit der Items pro Gruppe (p -Werte), ihre Trennschärpen (punkt-biseriale Korrelationen) und ihre Passung zum Rasch-Modell. Die Modellpassung wurde anhand des sogenannten *Infit* oder *Weighted Mean Square* (WMNSQ, Wright & Masters, 1982) für jedes einzelne Item berechnet. Dieser Kennwert beschreibt die Abweichung der empiri-

5 Dabei handelt es sich um die drei Länder, in denen im IQB-Bildungstrend 2016 eine erweiterte Gruppe von Schülerinnen und Schülern mit SPF getestet werden sollte.

schen Item-Charakteristik-Kurve (ICC) von der durch das Modell vorhergesagten ICC. Verläuft die empirische Kurve flacher als die durch das Modell implizierte, ist das ein Indiz für eine zu geringe Trennschärfe des entsprechenden Items; der WMNSQ nimmt dann Werte größer als 1 an. Der WMNSQ sollte einen Grenzwert von 1.20 oder, bei konservativer Betrachtung, 1.15 nicht überschreiten (Pohl & Carstensen, 2012). Im umgekehrten Fall, also bei einer empirisch steiler verlaufenden Kurve, ist der WMNSQ kleiner als 1 und die Trennschärfe größer als durch das Modell angenommen. Als Grenzwert wurde in diesem Fall ein WMNSQ kleiner als 0.80 festgelegt.

Neben den psychometrischen Eigenschaften der Items wurde auch der Anteil fehlender Werte zwischen den beiden Aufgabensets verglichen. Fehlende Werte oder *Missings* kommen zustande, wenn Schülerinnen und Schüler die ihnen vorgelegten Items nicht vollständig bearbeiten: Beispielsweise überspringen Schülerinnen und Schüler einzelne Items im Testverlauf und geben für diese Items keine Antworten (*missing by intention*), andere Antworten sind uneindeutig und können nicht kodiert werden (z. B. Ankreuzen mehrerer Optionen in Multiple-Choice-Items, *missing invalid response*). Eine unvollständige Bearbeitung aller Items eines Testheftes aufgrund von Zeitmangel schlägt sich in fehlenden Werten am Ende von Aufgabenblöcken (*missing not reached*) nieder. Eine besondere Herausforderung bei der Entwicklung angepasster Testaufgaben besteht darin, fehlende Werte möglichst zu vermeiden. Daher wurde auch geprüft, ob sich die angepassten Testaufgaben in dieser Hinsicht von den regulären Testaufgaben unterscheiden und ob die Aufgabenblöcke von Schülerinnen und Schülern mit SPF in der verfügbaren Zeit vollständig bearbeitet werden können.

Ferner wurde untersucht, ob beide Aufgabensets dasselbe Konstrukt erfassen. Die empirische Ähnlichkeit zweier Konstrukte kann anhand latenter Korrelationen (Kolen & Brennan, 2004) bestimmt werden. Dies ist auch möglich, wenn zwei theoretisch identische Konstrukte – wie im vorliegenden Fall – mit verschiedenen Sets an Aufgaben innerhalb desselben Tests erfasst werden.

In der Vergangenheit wurden die Zusammenhänge zwischen Tests der Ländervergleichs- bzw. Bildungstrendstudien des IQB und anderen großen Schulleistungsstudien, die jeweils inhaltlich äquivalente Konstrukte erfassten, untersucht. Diese Befunde können als Anhaltspunkt für die zu erwartenden Korrelationen zwischen den beiden Aufgabensets im IQB-Bildungstrend 2016 herangezogen werden. Für die Lesekompetenz in der Primarstufe ermittelten Pietsch, Böhme, Robitzsch und Stubbe (2009) in einer gemeinsamen Skalierung von Daten aus der Normierung der Bildungsstandards für die Primarstufe und IGLU⁶ 2006 eine latente Korrelation von $r = .84$. Ähnlich hoch war der Zusammenhang zwischen dem IQB-Ländervergleich 2011 in der Primarstufe und IGLU 2011: $r = .85$ für das einparametrische Modell und $r = .86$ für das zweiparametrische Modell (Böhme et al., 2014). Für den mathematischen Bereich in der Primarstufe wurde zwischen dem IQB-Ländervergleich 2011 und TIMSS⁷ 2011 ein Zusammenhang von $r = .93$ ermittelt (Böhme et al., 2014). Für die Lesekompetenz in der Sekundarstufe I berichteten Jude, Hartig, Schipolowski, Böhme und Stanat (2013) eine latente Korrelation von $r = .85$ zwischen PISA⁸ 2009 und dem IQB-Ländervergleich 2009. Für das Fach Mathematik fanden

6 Das Akronym IGLU steht für *Internationale Grundschul-Lese-Untersuchung*.

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 PISA steht für *Programme for International Student Assessment*.

Hartig und Frey (2012) zwischen der Normierung der Bildungsstandards in der Sekundarstufe I im Jahr 2006 und PISA 2006 eine Korrelation von $r = .94$. Für die Primar- und die Sekundarstufe ergeben sich also Zusammenhänge in sehr ähnlicher Größenordnung.

Psychometrische Eigenschaften regulärer und angepasster Testaufgaben im IQB-Bildungstrend 2016

Die Ergebnisse der psychometrischen Analysen für die angepassten Testaufgaben im IQB-Bildungstrend 2016 sind in Tabelle 13.4 zusammengefasst. Die Tabelle enthält drei Spalten, die die Befunde für Schülerinnen und Schüler ohne SPF, Schülerinnen und Schüler mit SPF an allgemeinen Schulen sowie Schülerinnen und Schüler mit SPF an Förderschulen zeigen. Innerhalb dieser Spalten wird noch einmal zwischen den regulären Testaufgaben (jeweils links) und den angepassten Testaufgaben (jeweils rechts) unterschieden. Es wird deutlich, dass die angepassten Testaufgaben für alle Schülerinnen und Schüler deutlich leichter sind als die regulären Aufgaben. Wichtig ist hierbei, dass die angepassten Aufgaben auch für Schülerinnen und Schüler mit SPF das gesamte Schwierigkeitsspektrum abdecken. So ist beispielsweise im Kompetenzbereich *Lesen* keine der regulären Testaufgaben sehr leicht: Der größte p -Wert der regulären Aufgaben liegt für Kinder mit SPF an allgemeinen Schulen und an Förderschulen bei $p = .78$. Dies bedeutet, dass selbst das leichteste Item von nur 78 Prozent der Kinder mit SPF gelöst wird. Für die angepassten Testaufgaben liegt der maximale p -Wert dagegen bei $p = .93$ für Kinder an Förderschulen und bei $p = .98$ für Kinder mit SPF an allgemeinen Schulen. Die Differenzierung im unteren Bereich des Kompetenzspektrums gelingt mit den angepassten Testaufgaben also deutlich besser als mit den regulären Testaufgaben.

Ob die eingesetzten Testitems gut zwischen Schülerinnen und Schülern mit unterschiedlich stark ausgeprägten Kompetenzen unterscheiden können, wird anhand der Trennschärfen der Items ermittelt. Für die Gruppe der Kinder mit SPF zeigt sich im Fach Deutsch, dass die Trennschärfen der angepassten Aufgaben in beiden Kompetenzbereichen durchgängig höher sind als die Trennschärfen der regulären Testaufgaben. Für das Fach Mathematik liegen die Trennschärfen der angepassten Testaufgaben im Mittel für alle drei Schülergruppen in etwa auf dem gleichen Niveau wie die Trennschärfen der regulären Testaufgaben. Für keine der angepassten Testaufgaben sind bei Kindern mit SPF extrem geringe Trennschärfen zu verzeichnen: Die angepassten Aufgaben differenzieren im unteren Kompetenzspektrum im Fach Deutsch also besser als die regulären Testaufgaben und im Fach Mathematik ähnlich gut.

Im Hinblick auf die Passung zum Rasch-Modell liegen im Bereich *Zuhören* in beiden Itemsets kaum Modellverletzungen vor; alle eingesetzten Items zeigen in allen drei Schülergruppen rasch-homogene Eigenschaften. Für den Bereich *Lesen* ergeben sich jedoch für etwa ein Fünftel der neu entwickelten Testaufgaben in der Gruppe der Schülerinnen und Schüler mit SPF an Förderschulen WMNSQ-Werte, die auf eine zu geringe Itemdiskrimination hinweisen. Dies muss jedoch nicht zwangsläufig bedeuten, dass diese Items tatsächlich schlecht messen. In einer anschließenden Analyse mit einem zweiparametrischen IRT-Modell (Birnbaum, 1968) zeigen die neu entwickelten Testaufgaben im Kompetenzbereich *Lesen* in der Gruppe der Schülerinnen und Schüler mit SPF an Förderschulen Diskriminationsparameter zwischen 0.64 und 3.01. Der

Tabelle 13.4: Psychometrische Eigenschaften der Items in regulären und angepassten Testaufgaben im IQB-Bildungstrend 2016

		Kinder ohne SPF		Kinder mit SPF an allgemeinen Schulen		Kinder mit SPF an Förderschulen	
		reguläre Items aus dem Jahr 2011	ange- passte Items	reguläre Items aus dem Jahr 2011	ange- passte Items	reguläre Items aus dem Jahr 2011	ange- passte Items
Deutsch (Lesen)							
Lösungshäufigkeit (p -Wert)	<i>M</i>	0.59	0.89	0.33	0.71	0.35	0.68
	Min	0.01	0.70	0.04	0.32	0.05	0.27
	Max	0.93	0.97	0.78	0.98	0.78	0.93
Trennschärfe (punkt-biseriale Korrelation)	<i>M</i>	0.30	0.33	0.33	0.47	0.36	0.51
	Min	0.02	0.16	-0.08	0.27	0.14	0.26
	Max	0.51	0.49	0.62	0.64	0.54	0.70
Modellpassung (<i>N</i> Items)	WMNSQ < 0.80	0	0	2	3	0	5
	WMNSQ > 1.15	1	0	6	7	1	10
	WMNSQ > 1.20	0	0	3	2	0	8
	insgesamt	78	15	77	49	32	49
Deutsch (Zuhören)							
Lösungshäufigkeit (p -Wert)	<i>M</i>	0.55	0.80	0.37	0.70	0.43	0.67
	Min	0.01	0.25	0.03	0.15	0.03	0.11
	Max	0.98	0.95	0.87	0.94	0.92	0.96
Trennschärfe (punkt-biseriale Korrelation)	<i>M</i>	0.24	0.21	0.30	0.34	0.29	0.33
	Min	0.03	0.07	-0.05	0.09	0.10	0.10
	Max	0.44	0.31	0.51	0.54	0.47	0.50
Modellpassung (<i>N</i> Items)	WMNSQ < 0.80	0	0	0	1	0	0
	WMNSQ > 1.15	0	0	0	2	0	0
	WMNSQ > 1.20	0	0	0	1	0	0
	insgesamt	40	20	39	27	23	39
Mathematik (Globalskala)							
Lösungshäufigkeit (p -Wert)	<i>M</i>	0.53	0.90	0.39	0.72	0.37	0.69
	Min	0.01	0.42	0.06	0.22	0.03	0.21
	Max	0.96	0.98	0.82	0.98	0.79	0.94
Trennschärfe (punkt-biseriale Korrelation)	<i>M</i>	0.37	0.32	0.48	0.44	0.45	0.48
	Min	0.10	0.10	0.20	0.21	0.03	0.21
	Max	0.59	0.72	0.70	0.61	0.68	0.61
Modellpassung (<i>N</i> Items)	WMNSQ < 0.80	0	3	5	0	4	0
	WMNSQ > 1.15	6	2	12	4	10	4
	WMNSQ > 1.20	1	1	6	1	7	2
	insgesamt	256	60	84	60	84	60

Anmerkungen. SPF = sonderpädagogischer Förderbedarf; WMNSQ = standardisierter Weighted Mean Square; *M* = Mittelwert. Kriterien zur Beurteilung der Modellpassung: Ein leichter Misfit entspricht einem WMNSQ < 0.80 oder WMNSQ > 1.15, starker Misfit entspricht einem WMNSQ > 1.20 (vgl. OECD, 2012; Pohl & Carstensen, 2012).

große Anteil auffälliger WMNSQ-Werte kann also auf die Uneinheitlichkeit der Trennschärfen zurückgeführt werden – die Items messen nicht per se schlecht, sie zeigen jedoch keine rasch-homogenen Eigenschaften. Aus diesem Grund wurde für die Auswertung des IQB-Bildungstrends 2016 ein restringiertes zweiparametrisches Modell anstelle des Rasch-Modells verwendet (für eine genauere Beschreibung des Vorgehens siehe Kapitel 13.1).

Für die Messung der globalen mathematischen Kompetenz zeigen die angepassten Testaufgaben in allen drei Zielgruppen sogar eine bessere Modellpassung als die regulären Testaufgaben.

Insgesamt lässt sich festhalten, dass die psychometrischen Eigenschaften der angepassten Testaufgaben zufriedenstellend sind: Sie decken im Kompetenzspektrum der Schülerinnen und Schüler mit SPF auch das untere Ende der Verteilung ab und differenzieren sehr gut zwischen Kindern mit höheren und niedrigeren Kompetenzen. Aufgrund der uneinheitlichen Trennschärfen der angepassten Items im Kompetenzbereich *Lesen* ist die Passung des Rasch-Modells allerdings eingeschränkt. Daraus resultieren jedoch keine systematischen Über- oder Unterschätzungen der erfassten Kompetenzen.

Anteile fehlender Werte in den regulären und angepassten Testaufgaben im IQB-Bildungstrend 2016

Empirische Befunde zu verschiedenen Arten fehlender Werte für die regulären und die angepassten Testaufgaben im IQB-Bildungstrend 2016 sind in Tabelle 13.5 zusammengestellt. Es wird deutlich, dass die Anteile der aufgrund von Zeitmangel nicht bearbeiteten Items (*missing not reached*) in den Kompetenzbereichen *Lesen* und *Zuhören* für die angepassten Testaufgaben mit jeweils unter einem Prozent durchgängig geringer sind als für die regulären Testaufgaben, bei denen diese Anteile zwischen 1 und 3 Prozent liegen. Im Fach Mathematik liegen die Anteile fehlender Werte bei nicht erreichten Items für Schülerinnen und Schüler mit SPF an allgemeinen Schulen und Förderschulen bei 2.5 bis 3.5 Prozent für die angepassten Aufgaben. Auch diese Anteile sind wiederum deutlich geringer als die entsprechenden Anteile bei den regulären Testaufgaben, die zwischen 6 und 7 Prozent liegen.

Bei den angepassten Testaufgaben sind aber nicht nur die Anteile fehlender Werte gering, die dadurch zustande kommen, dass die Schülerinnen und Schüler das Ende eines Aufgabenblocks nicht erreichen. Auch die Anteile der ausgelassenen, also absichtlich nicht bearbeiteten Testitems (*missing by intention*) sind für alle Schülergruppen und alle Kompetenzbereiche für die angepassten Aufgaben deutlich geringer als für die regulären Testaufgaben. Das Gleiche gilt für nicht gültige Antworten (*missing invalid response*), etwa wenn in einer Multiple-Choice-Aufgabe nicht nur eine Option, sondern mehrere Optionen markiert wurden. Die geringeren Anteile an *missing invalid response* können durch die geringere Schwierigkeit dieser Items bedingt sein oder als Hinweis gedeutet werden, dass die Instruktionen bei den angepassten Testaufgaben besser verstanden wurden.

Tabelle 13.5: Prozentuale Anteile richtig und falsch gelöster Items und fehlender Werte in regulären und angepassten Testaufgaben im IQB-Bildungstrend 2016

	Kinder ohne SPF		Kinder mit SPF an allgemeinen Schulen		Kinder mit SPF an Förderschulen	
	reguläre Items aus dem Jahr 2011	angepasste Items	reguläre Items aus dem Jahr 2011	angepasste Items	reguläre Items aus dem Jahr 2011	angepasste Items
Deutsch (Lesen)						
falsch gelöst	28.7	6.3	45.1	17.5	45.4	18.8
richtig gelöst	66.5	92.0	41.0	75.5	39.1	72.1
missing by intention	3.6	1.2	9.6	5.0	10.8	7.1
missing invalid response	0.8	0.3	1.6	1.1	1.9	1.3
missing not reached	0.5	0.2	2.7	0.9	2.9	0.7
missings insgesamt	4.8	1.7	13.9	7.0	15.5	9.1
Deutsch (Zuhören)						
falsch gelöst	31.3	12.4	40.4	21.6	43.2	24.6
richtig gelöst	61.8	86.3	43.6	73.0	41.8	69.4
missing by intention	5.5	1.0	12.5	4.6	12.0	4.7
missing invalid response	1.2	0.2	2.3	0.7	2.3	1.1
missing not reached	0.2	0.0	1.3	0.1	0.8	0.3
missings insgesamt	6.9	1.3	16.1	5.3	15.1	6.0
Mathematik (Globalskala)						
falsch gelöst	29.9	6.5	37.0	16.8	37.2	17.8
richtig gelöst	58.0	91.3	40.6	74.4	39.1	71.9
missing by intention	9.1	1.5	14.0	5.5	14.7	6.1
missing invalid response	1.3	0.2	2.1	0.7	2.2	0.9
missing not reached	1.7	0.5	6.3	2.6	6.8	3.3
missings insgesamt	12.1	2.2	22.4	8.8	23.7	10.3

Anmerkungen. SPF = sonderpädagogischer Förderbedarf; missing by intention = ausgelassene Antworten; missing invalid response = ungültige Antworten; missing not reached = fehlende Antworten am Ende des Testhefts.

Konstruktäquivalenz regulärer und angepasster Testaufgaben im IQB-Bildungstrend 2016

Die latenten Korrelationen zwischen den regulären und den angepassten Testaufgaben im IQB-Bildungstrend liegen für die gesamte Schülerschaft für den Bereich *Lesen* bei $r = .80$, für den Bereich *Zuhören* bei $r = .76$ und für die mathematische Kompetenz bei $r = .83$. Im IQB-Bildungstrend 2016 fallen die latenten Korrelationen somit sowohl für die beiden Kompetenzbereiche des Faches Deutsch (*Lesen* und *Zuhören*) als auch für die globale mathematische Kompetenz etwas geringer aus, als vor dem Hintergrund der oben berichteten empirischen Befunde zu Zusammenhängen zwischen Kompetenzen in Tests verschiedener Studien erwartbar gewesen wäre.

Trotz der hohen inhaltlichen Übereinstimmung zwischen den regulären und den angepassten Testaufgaben des IQB-Bildungstrends 2016 scheinen die vergleichsweise großen Schwierigkeitsunterschiede zwischen den Aufgabensets eine vollständige Homogenität des abgebildeten Konstrukts nicht gewährleisten zu können. Zudem ist die Operationalisierung der erfassten Kon-

strukturen weder aus konzeptioneller noch aus formaler Sicht vollständig identisch. Eine Untergliederung von Lesetexten in Abschnitte oder das Vermeiden von Informationen, die für die Aufgabebearbeitung nicht unmittelbar relevant sind, erleichtern zwar beispielsweise das Textverständnis, verändern aber auch die Art und Weise, in der ein Text gelesen und verstanden wird. Unterschiede zwischen den regulären und angepassten Testaufgaben liegen damit teilweise in der Zielstellung selbst begründet, Aufgaben zu entwickeln, die im unteren Kompetenzspektrum ausreichend differenzieren. Dass die Sets der regulären und angepassten Testaufgaben in unterschiedlichen Bereichen des Kompetenzspektrums differenzieren, hat allerdings auch messmethodische Konsequenzen: Durch die starken Diskrepanzen in den Schwierigkeiten dieser beiden Aufgabensets können die Messäquivalenz eingeschränkt und die Schätzung des Zusammenhangs beider Itemsets verzerrt sein (Kolen & Brennan, 2004; Lane & Leventhal, 2015).

13.2.4 Fazit

Die Entwicklung angepasster Testaufgaben für Schülerinnen und Schüler mit SPF für den Einsatz im IQB-Bildungstrend 2016 erfolgte theoriegeleitet und erwies sich im Hinblick auf die psychometrischen Eigenschaften und die angestrebte Reduktion der Anteile fehlender Werte als erfolgreich. Durch sprachliche und grafische Anpassungen der Aufgabeninstruktionen, der Stimulustexte und der Items wurden die Aufgaben vereinfacht und die kognitive Belastung während der Bearbeitung verringert. Dies zeigt sich unter anderem darin, dass die angepassten Testaufgaben im unteren Bereich des Kompetenzspektrums besser differenzieren als die regulären Testaufgaben.

Unklar ist allerdings, inwieweit die Anpassungen auch zu einer gewissen Veränderung der erfassten Konstrukte geführt haben. Die latenten Korrelationen zwischen angepassten und regulären Testaufgaben fallen insgesamt etwas geringer aus als erwartet. Dies könnte zum Teil auf den beträchtlichen Unterschied in der mittleren Schwierigkeit der beiden Itemgruppen zurückzuführen sein (vgl. z. B. auch Kolen & Brennan, 2004). Ob oder inwieweit mit den vorgenommenen Anpassungen an Stimuli und Instruktionen auch eine inhaltliche Veränderung der intendierten Konstrukte einhergeht, bedarf vertiefter Analysen, die auch qualitative Zugänge zu dieser Frage berücksichtigen.

Somit lassen sich aus den in diesem Kapitel dargestellten Befunden zu angepassten Testaufgaben für Schülerinnen und Schüler mit SPF zwei zentrale Herausforderungen für die weitere Entwicklung entsprechender Aufgaben ableiten: Zum einen sollte sowohl theoretisch als auch empirisch geprüft werden, wie und in welchem Ausmaß Testaufgaben angepasst werden können, ohne die Äquivalenz der erfassten Konstrukte zu gefährden. Dies betrifft allerdings nicht nur die Frage, ob beide Aufgabensets dasselbe Konstrukt abbilden, sondern auch, ob sie das Konstrukt ähnlich breit abdecken können. Beispielsweise lässt sich empirisch zeigen, dass einige der in den Bildungsstandards formulierten Anforderungen erst auf den höheren Kompetenzstufen relevant werden (bspw. Schlüsse durch Integration von in mehreren Textabschnitten verteilten Informationen ziehen). Die Lesetexte der angepassten Testaufgaben sind aber gezielt so gestaltet, dass alle zugehörigen Items möglichst einfach sind. Schülerinnen und Schüler mit SPF, die höheren Kompetenzanforderungen prin-

ziell gewachsen wären, erhalten in den angepassten Testaufgaben daher kaum Gelegenheit, dieses Können auch unter Beweis zu stellen.

Zum anderen stellt sich die Frage, ob angepasste Testaufgaben für Schülerinnen und Schüler mit unterschiedlichen Förderschwerpunkten ähnlich zugänglich und damit ähnlich gut geeignet sind. Die in diesem Kapitel beschriebenen angepassten Testaufgaben dienen der Erfassung von Kompetenzen bei Schülerinnen und Schülern mit SPF in den Bereichen „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“. Hierbei bleibt jedoch offen, ob alle Schülerinnen und Schüler mit SPF in den verschiedenen Bereichen in gleicher Weise von den Anpassungen der Testaufgaben profitieren können. Gerade die Heterogenität der Kompetenzen und Bedürfnisse von Schülerinnen und Schülern mit sonderpädagogischen Förderbedarfen legt nahe, dass eine einheitliche und für alle Schülerinnen und Schüler dieser heterogenen Zielgruppe in gleicher Weise vorgenommene Anpassung keine optimale Lösung darstellt. Auch dies ist ein wichtiges Desiderat zukünftiger Forschungs- und Entwicklungsvorhaben bei der Einbeziehung von Schülerinnen und Schülern mit SPF in ein umfassendes Bildungsmonitoring.

Literatur

- Beddow, P. A. (2012). Accessibility theory for enhancing the validity of test results for students with special needs. *International Journal of Disability, Development and Education*, 59(1), 97–111.
- Bertram, R. & Schulmeyer, H. (2014). *Leserabe – Die allergefährlichsten Piraten der Welt*. Ravensburg: Ravensburger Buchverlag.
- Bock, B. M. (2015). Anschluss ermöglichen und die Vermittlungsaufgabe ernst nehmen – 5 Thesen zur Leichten Sprache. *Didaktik Deutsch*, 38, 9–17.
- Böhme, K., Richter, D., Weirich, S., Haag, N., Wendt, H., Bos, W., Pant, H. A. & Stanat, P. (2014). Messen wir dasselbe? Zur Vergleichbarkeit des IQB-Ländervergleichs 2011 mit den internationalen Studien IGLU und TIMSS 2011. *Unterrichtswissenschaft*, 42(4), 342–365.
- Clark, R. C., Nguyen, F. & Sweller, J. (2006). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. San Francisco, CA: Jossey-Bass.
- Chandler, P. & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332.
- Elliott, S. N., Davies, M. & Kettler, R. J. (2012). Australian students with disabilities Accessing NAPLAN: Lessons from a decade of inclusive assessment in the United States. *International Journal of Disability, Development and Education*, 59(1), 7–19.
- Elliott, S. N., Kettler, R. J., Beddow, P. A. & Kurz, A. (Hrsg.). (2011). *Handbook of accessible achievement tests for all students: Bridging the gaps between research, practice, and policy*. New York: Springer.
- Feilke, H. (2015). Transitorische Normen. Argumente zu einem didaktischen Normbegriff. *Didaktik Deutsch*, 38, 115–136.
- Gerlach, T., Böhme, K., Haag, N., Krelle, M., Börnert, M. & Grosche, M. (2016). *Adaption von Leseaufgaben für Kinder mit sonderpädagogischem Förderbedarf*. Vortrag auf der 4. GEBF-Tagung, Berlin, 09.–11. März 2016.
- Haladyna, T. M., Downing, S. M. & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309–333.
- Hartig, J., & Frey, A. (2012). Validität des Tests zur Überprüfung des Erreichens der Bildungsstandards in Mathematik: Zusammenhänge mit den bei PISA gemessenen Kompetenzen und Varianz zwischen Schulen und Schulformen. *Diagnostica*, 58, 3–14.
- Heppt, B. (2016). *Verständnis von Bildungssprache bei Kindern mit deutscher und nicht-deutscher Familiensprache* (Doktorarbeit). Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät. Zugriff am 08.08.2017 unter <https://edoc.hu-berlin.de/handle/18452/18186>.

- Jude, N., Hartig, J., Schipolowski, S., Böhme, K. & Stanat, P. (2013). Definition und Messung von Lesekompetenz – PISA und die Bildungsstandards. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung*. 59. Beiheft der Zeitschrift für Pädagogik (S. 200–228). Weinheim: Beltz Juventa.
- Kettler, R. J., Elliott, S. N. & Beddow, P. A. (2009). Modifying achievement test items: A theory-guided and data-based approach for better measurement of what students with disabilities know. *Peabody Journal of Education*, 84(4), 529–551.
- Kettler, R. J., Rodriguez, M. C., Bolt, M., Elliott, S. N., Beddow, P. A. & Kurz, A. (2011). Modified multiple-choice items for alternate assessments: Reliability, difficulty, and differential boost. *Applied Measurement in Education*, 24, 210–234.
- Kocaj, A., Haag, N., Weirich, S., Kuhl, P., Pant, H. A. & Stanat, P. (2016). Aspekte der Testgüte bei der Erfassung schulischer Kompetenzen von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In V. Moser & B. Lütje-Klose (Hrsg.), *Schulische Inklusion*. 62. Beiheft der Zeitschrift für Pädagogik (S. 212–234). Weinheim: Beltz Juventa.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices*. New York: Springer.
- Lane, S. & Leventhal, B. (2015). Psychometric challenges in assessing English language learners and students with disabilities. *Review of Research in Education*, 39, 165–214.
- Maaß, C. (2015). Leichte Sprache – Zugang zu fachlichen Kontexten ermöglichen. *Didaktik Deutsch*, 38, 3–8.
- Mayer, R. E. & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist*, 38, 43–52.
- Müller, S. (2014). *Data Rich, Information Poor? Chancen und Grenzen vergleichender Leistungsmessung an der Förderschule mit dem Förderschwerpunkt Lernen*. Münster: Waxmann.
- OECD (Hrsg.). (2012). *PISA 2009 technical report*. Paris: OECD.
- Oomen-Welke, I. (2015). Leichte Sprache, Einfache Sprache und Deutsch als Zweitsprache. *Didaktik Deutsch*, 38, 24–32.
- Pohl, S. & Carstensen, C. H. (2012). *NEPS Technical Report – Scaling the data of the competence tests*. Bamberg: Otto-Friedrich-Universität, Nationales Bildungspanel.
- 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.
- Schleppegrell, M. J. (2002). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431–459.
- Schroeder, S., Würzner, K. M., Heister, J., Geyken, A. & Kliegl, R. (2015). childLex – Eine lexikalische Datenbank zur Schriftsprache für Kinder im Deutschen. *Psychologische Rundschau*, 66, 155–165.
- Südkamp, A., Pohl, S., Hardt, K., Jordan, A. K. & Duchhardt, C. (2015). Kompetenzmessung in den Bereichen Lesen und Mathematik bei Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf. In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 243–272). Wiesbaden: Springer.
- Weigert, S. C. (2012). Aligning and inventing practices to achieve inclusive assessment policies: A decade of work toward optimal access for US students with disabilities 2001–2011. *International Journal of Disability, Development and Education*, 59(1), 21–36.
- Wright, B. D. & Masters, G. N. (1982). *Rating scale analysis. Rasch measurement*. Chicago, IL: MESA Press.

Kapitel 14

Zusammenfassung und Einordnung der Befunde

Petra Stanat, Stefan Schipolowski, Camilla Rjosk, Sebastian Weirich und Nicole Haag

Im IQB-Bildungstrend 2016 wurde für den Primarbereich zum zweiten Mal überprüft, inwieweit schulische Erträge in den Fächern Deutsch und Mathematik den Kompetenzerwartungen der Bildungsstandards entsprechen, die von der Kultusministerkonferenz (KMK) vereinbart worden sind. Die erste Überprüfung dieser länderübergreifend geltenden Kompetenzziele im Primarbereich erfolgte im IQB-Ländervergleich 2011. Damit war es im Rahmen des vorliegenden Berichts für die 4. Jahrgangsstufe erstmals möglich, nicht nur die von Schülerinnen und Schülern im Jahr 2016 erreichten Kompetenzen zu beschreiben, sondern auch anhand von Trendanalysen zu prüfen, inwieweit sich die Ergebnismuster seit dem Jahr 2011 verändert haben. Durch die wiederholte Messung im zeitlichen Abstand von fünf Jahren hat sich das Spektrum der möglichen Analysen und damit der Informationsgehalt der Ergebnisse deutlich erweitert.

Am IQB-Bildungstrend 2016 nahm eine für alle 16 Länder in der Bundesrepublik Deutschland repräsentative Stichprobe von insgesamt mehr als 29 000 Schülerinnen und Schülern der 4. Jahrgangsstufe teil. Die in den Erhebungen eingesetzten Testaufgaben wurden auf Basis der Bildungsstandards der KMK unter Federführung des IQB und in enger Zusammenarbeit mit Fachdidaktikerinnen und Fachdidaktikern von Lehrkräften entwickelt. Im Fach Deutsch untersuchte der IQB-Bildungstrend 2016 die Kompetenzbereiche beziehungsweise Teilbereiche *Lesen*, *Zuhören* und *Orthografie*. Im Fach Mathematik wurden die fünf in den Bildungsstandards beschriebenen inhaltsbezogenen Kompetenzbereiche (Leitideen) *Zahlen und Operationen*, *Raum und Form*, *Muster und Strukturen*, *Größen und Messen* sowie *Daten, Häufigkeit und Wahrscheinlichkeit* erfasst. Zusätzlich wurde eine Globalskala mathematischer Kompetenz gebildet, die alle Leitideen zusammenfasst. Zur inhaltlichen Interpretation der von den Schülerinnen und Schülern erreichten Testleistungen dienen Kompetenzstufenmodelle, die das IQB auf Grundlage der Bildungsstandards der KMK entwickelt hat. Anhand dieser Modelle lässt sich beschreiben, welche Anforderungen Schülerinnen und Schüler, die ein bestimmtes Testergebnis erzielt haben, bewältigen können. Ferner kann festgestellt werden, inwieweit die Kompetenzen der Kinder im jeweiligen Fach und Kompetenzbereich den Erwartungen entsprechen, die mit den Bildungsstandards festgelegt wurden.

Im vorliegenden Bericht wurden die Ergebnisse zu den von Schülerinnen und Schülern erreichten Kompetenzen unter drei Vergleichsperspektiven ausgewertet (vgl. auch Kapitel 1.1):

Unter einer *kriterialen Vergleichsperspektive* gingen die Analysen der Frage nach, wie sich die Viertklässlerinnen und Viertklässler im Jahr 2016 in den einzelnen Ländern auf die Stufen der in Kapitel 2 beschriebenen Kompetenz-

stufenmodelle verteilen. Dabei wurde jeweils untersucht, welcher Anteil der Schülerinnen und Schüler mindestens die Regelstandards erreicht beziehungsweise die Mindeststandards nicht erreicht hat. Ferner wurde das Erreichen der Optimalstandards analysiert.

Unter einer *ipsativen Vergleichsperspektive* konnten im IQB-Bildungstrend 2016 – wie oben bereits erwähnt – auch Veränderungen über die Zeit beschrieben werden (Trends). Die Verknüpfung der ipsativen mit der kriterialen Perspektive erlaubt eine Aussage darüber, inwieweit sich die Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen in den einzelnen Ländern zwischen den Jahren 2011 und 2016 verändert haben – ob also zum Beispiel der Anteil der Kinder, deren Kompetenzen den Anforderungen der Mindeststandards nicht entsprechen, reduziert und der Anteil der Kinder, die mindestens die Regelstandards erreichen, erhöht werden konnte.

Zusätzlich spielt die *soziale Vergleichsperspektive* in den Analysen des IQB-Bildungstrends 2016 eine Rolle, die zum Beispiel Aussagen darüber ermöglicht, in welchen Ländern der Anteil der Schülerinnen und Schüler, die die Regelstandards erreichen, im Vergleich zu anderen Ländern besonders hoch oder besonders niedrig ist.

Ein weiterer Analyseschwerpunkt der Studien des IQB zum Bildungsmonitoring betrifft die Frage, inwieweit Unterschiede in den von Schülerinnen und Schülern erreichten Kompetenzen mit bestimmten Hintergrundmerkmalen in Zusammenhang stehen. Untersucht wurden Kompetenzunterschiede zwischen Mädchen und Jungen (Geschlechterdisparitäten), Zusammenhänge zwischen Merkmalen der sozialen Herkunft und den erreichten Kompetenzen (soziale Disparitäten) sowie Leistungsdifferenzen zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund (zuwanderungsbezogene Disparitäten). Auch wenn die Erwartung, dass ein Bildungssystem ungleiche Eingangsvoraussetzungen vollständig ausgleichen kann, unrealistisch ist, so gilt es doch als allgemein akzeptiertes bildungspolitisches Ziel, die mit den genannten Hintergrundmerkmalen der Schülerinnen und Schüler verbundenen Disparitäten so weit wie möglich zu reduzieren. Daher wurde sowohl für das Jahr 2016 als auch im Vergleich der Jahre 2011 und 2016 überprüft, inwieweit dies in den Ländern erreicht werden konnte.

Alle beschriebenen Analyseschwerpunkte zusammen – die Überprüfung des Erreichens der Bildungsstandards, die Prüfung von Veränderungen über die Zeit, die ländervergleichende Perspektive und der differenzierte Blick auf Teilgruppen der Schülerschaft – geben Bildungspolitik und Bildungsverwaltung einen Überblick über Stärken und Schwächen der 16 Bildungssysteme im Primarbereich. Insbesondere liefern die Trendschätzungen Anhaltspunkte dafür, inwieweit die Maßnahmen der Qualitätsentwicklung, die in den letzten fünf Jahren in den Bildungssystemen der einzelnen Länder umgesetzt wurden, erfolgreich waren und wo weiterer Handlungsbedarf besteht.

Darüber hinaus nimmt der IQB-Bildungstrend 2016 einzelne Fragestellungen in den Blick, die neben *Ergebnissen* von Lehr-Lern-Prozessen auch zentrale *Bedingungen* dieser Prozesse betreffen. Den Ausgangspunkt der Analysen bildet die Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (KMK, 2015a), wonach Schulleistungsstudien auch zur Untersuchung praktischer Schlüsselfragen der Schul- und Unterrichtsentwicklung genutzt werden sollen. Das für den Bericht über den IQB-Bildungstrend 2016 gewählte Schwerpunktthema bezieht sich primär auf das von der KMK als besonders wichtig herausgestellte Themenfeld „Umgang mit Heterogenität“, wobei der

Fokus der Zusatzanalysen auf Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf (SPF) liegt. Aber auch auf leistungsstarke Schülerinnen und Schüler, zu deren Lernsituation innerhalb des Themenfeldes „Umgang mit Heterogenität“ ebenfalls Forschungsbedarf gesehen wird, geht der Berichtsband ein. In einem weiteren Zusatzkapitel wurden die bereits seit dem Ländervergleich 2011 vom IQB durchgeführten Analysen zur Qualifikation von Lehrkräften fortgeführt. Darüber hinaus wurden auch in diesem Zusatzkapitel Aspekte des Schwerpunktthemas „Umgang mit Heterogenität“ aufgegriffen und Ergebnisse der Lehrkräftebefragung ausgewertet, die sich auf gemeinsamen Unterricht von Kindern mit SPF und Kindern ohne SPF beziehen.

Im Folgenden werden die zentralen Ergebnisse der im Rahmen des IQB-Bildungstrends 2016 durchgeführten Analysen für die Fächer Deutsch und Mathematik knapp zusammengefasst. Die Befunde werden jeweils für die Kompetenzbereiche *Lesen*, *Zuhören* und *Orthografie*¹ im Fach Deutsch und für die Globalskala im Fach Mathematik dargestellt. In der Zusammenfassung und Einordnung der Befunde werden zentrale Tendenzen skizziert sowie besondere Erfolge und Herausforderungen benannt.

14.1 Erreichen der Bildungsstandards in den Ländern im Jahr 2016

Die Bildungsstandards der KMK beziehen sich auf Bildungsabschlüsse beziehungsweise Bildungsabschnitte. Es handelt sich um Regelstandards, die festlegen, über welche Kompetenzen die Schülerinnen und Schüler im Durchschnitt verfügen sollten, wenn sie das Ende der jeweiligen Bildungsetappe erreicht haben. Auf der Grundlage dieser Zielvorgaben für die Fächer Deutsch und Mathematik im Primarbereich wurden Kompetenzstufenmodelle entwickelt, die nicht nur Leistungen auf dem Niveau des Regelstandards beschreiben, sondern das gesamte Kompetenzspektrum abdecken und es in sinnvoll interpretierbare Abschnitte unterteilen (vgl. Kapitel 2). Die Kompetenzstufenmodelle legen zudem fest, auf welcher Stufe die Schülerinnen und Schüler den *Mindeststandard*, den *Regelstandard*, den *Regelstandard plus* beziehungsweise den *Optimalstandard* erreicht haben.

In der nachfolgenden Zusammenfassung der Ergebnisse für die Kompetenzstufenverteilungen im Jahr 2016 (vgl. Kapitel 5) wird zu jedem Kompetenzbereich jeweils auf drei Fragen eingegangen: Welcher Anteil der Schülerinnen und Schüler erreicht mindestens die Regelstandards, verfehlt die Mindeststandards und erreicht die Optimalstandards?

Im Fach Deutsch erreichen oder übertreffen im Jahr 2016 bundesweit im Bereich *Lesen* knapp 66 Prozent, im Bereich *Zuhören* gut 68 Prozent und im Bereich *Orthografie* rund 54 Prozent der Viertklässlerinnen und Viertklässler den Regelstandard. Den Mindeststandard verfehlen in diesen Kompetenzbereichen jeweils knapp 13 Prozent, fast 11 Prozent und rund 22 Prozent der Schülerinnen und Schüler. Kompetenzen, die dem jeweiligen Optimalstandard entsprechen, erreichen in Deutschland insgesamt sowohl im *Lesen* als auch im *Zuhören* etwa 10 Prozent und in der *Orthografie* fast 9 Prozent der Schülerinnen und Schüler.

1 Da die Erhebung zum Kompetenzbereich *Orthografie* im IQB-Ländervergleich 2011 so angelegt war, dass keine Analysen auf Länderebene möglich sind und keine Schülerinnen und Schüler an Förderschulen einbezogen wurden, konnten die Trendanalysen für diesen Kompetenzbereich nur für Deutschland insgesamt und für Kinder ohne SPF durchgeführt werden.

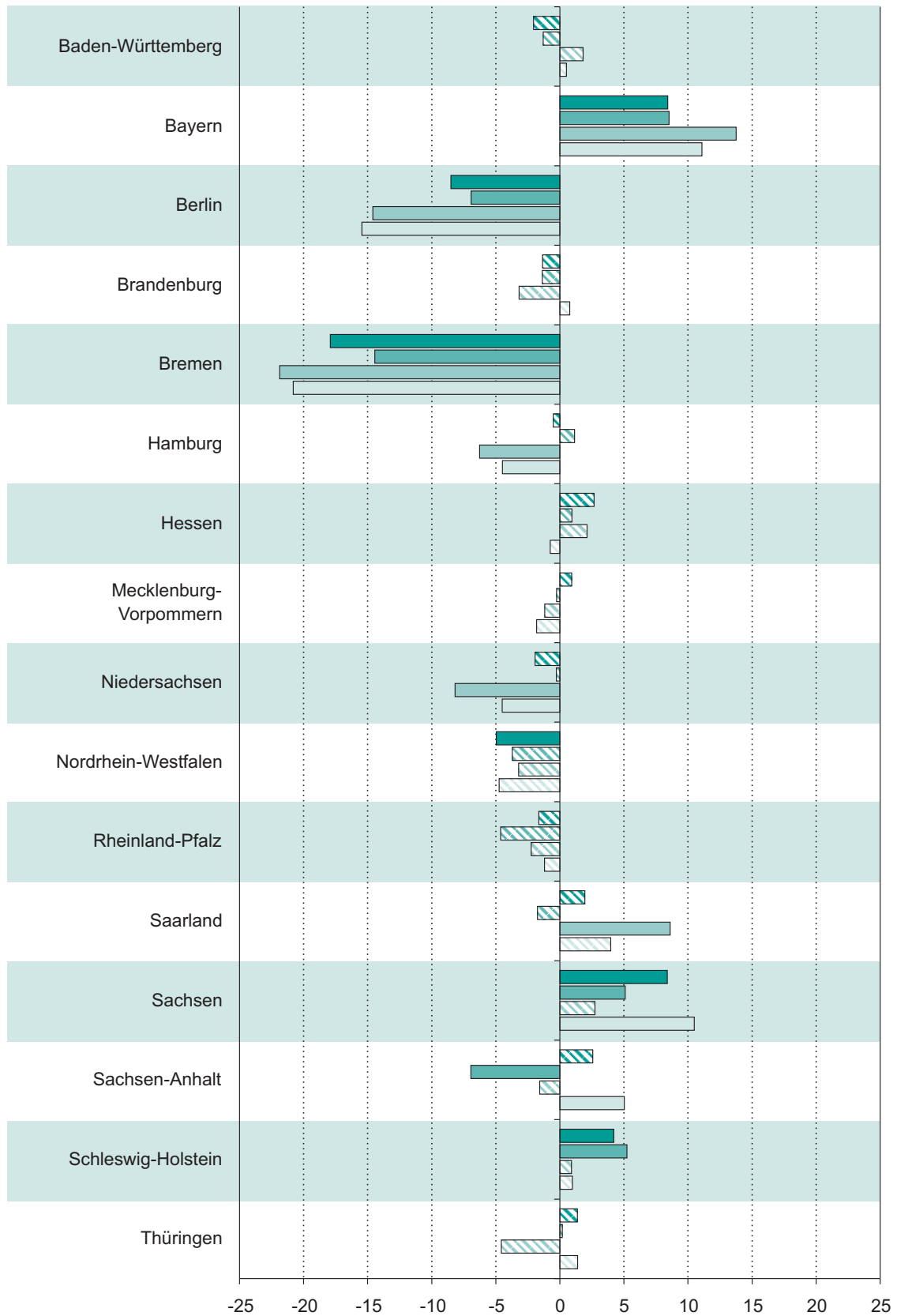
Im Fach Mathematik (*Globalskala*) erreichen oder übertreffen den KMK-Regelstandard gut 62 Prozent der Schülerinnen und Schüler in Deutschland insgesamt, wohingegen rund 15 Prozent den Mindeststandard verfehlen. Etwa 13 Prozent der Viertklässlerinnen und Viertklässler können im Fach Mathematik Kompetenzen nachweisen, die dem Optimalstandard entsprechen.

In den Abbildungen 14.1 und 14.2 ist dargestellt, inwieweit der Anteil der Schülerinnen und Schüler, die mindestens den Regelstandard erreichen (Abb. 14.1) beziehungsweise den Mindeststandard verfehlen (Abb. 14.2), in den einzelnen Ländern und Kompetenzbereichen vom jeweiligen Wert abweicht, der für Deutschland insgesamt identifiziert wurde. Abbildung 14.3 stellt die entsprechenden Ergebnisse für das Erreichen der Optimalstandards dar. Nach rechts weisende Balken zeigen, dass der Anteil im jeweiligen Land größer ist als in Deutschland insgesamt, bei nach links weisenden Balken ist der Anteil im jeweiligen Land kleiner als bundesweit. Statistisch nicht signifikante Abweichungen sind mit schraffierten Balken gekennzeichnet. Interpretiert werden sollten nur statistisch signifikante Unterschiede, wobei zusätzlich zur statistischen Signifikanz auch die Größe der Abweichungen zu beachten ist. Aufgrund der großen Fallzahlen, auf denen die Analysen basieren, können auch sehr kleine Unterschiede, die praktisch kaum ins Gewicht fallen, das statistische Signifikanzniveau erreichen.

Die Ergebnismuster für die einzelnen Länder ergeben ein differenziertes Bild. Im Land Bayern gelingt es in den Fächern Deutsch und Mathematik durchgängig besonders gut, die Regel- und Mindeststandards zu sichern, und auch der Optimalstandard wird in drei der vier untersuchten Bereiche in Bayern von einem signifikant höheren Anteil der Schülerinnen und Schüler erreicht als bundesweit. Bezogen auf die Regelstandards, teilweise aber auch bezogen auf die Mindest- und Optimalstandards, fällt das Ergebnismuster zudem in Sachsen für beide Fächer recht günstig aus. Auf der anderen Seite werden in Bremen und Berlin im Jahr 2016 in allen Kompetenzbereichen die Regelstandards seltener erreicht oder übertroffen und die Mindeststandards häufiger verfehlt als dies deutschlandweit der Fall ist. In Bremen ist zudem der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, in allen untersuchten Kompetenzbereichen geringer als bundesweit; in Berlin ist dies nur im Bereich *Orthografie* und im Fach Mathematik der Fall.

Für die anderen Länder, in denen signifikante Abweichungen von den bundesweiten Ergebnissen zu verzeichnen sind, unterscheiden sich die Befundmuster nach Fächern und Kompetenzbereichen. Besondere Stärken im Fach Deutsch sind sowohl bezogen auf den Regelstandard als auch bezogen auf den Mindeststandard im Saarland für den Bereich *Orthografie* und in Schleswig-Holstein für die Bereiche *Lesen* und *Zuhören* zu verzeichnen. Zudem ist der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, in Hamburg im Bereich *Zuhören* und im Saarland im Bereich *Orthografie* signifikant höher als bundesweit. Relative Schwächen im Fach Deutsch bestehen im Bereich *Lesen* in Nordrhein-Westfalen (Regelstandard) und Thüringen (Optimalstandard), im Bereich *Zuhören* in Sachsen-Anhalt (Regel- und Optimalstandard) und Rheinland-Pfalz (Optimalstandard) und im Bereich *Orthografie* in Hamburg und Niedersachsen (Regel-, Mindest- und Optimalstandard) sowie in Brandenburg und Thüringen (Optimalstandard). Im Fach Mathematik liegen im Vergleich zu den Ergebnissen für Deutschland insgesamt relative Stärken in Sachsen-Anhalt (Regel- und Mindeststandard) und relative Schwächen in Hamburg (Regel- und Mindeststandard), Niedersachsen (Regel- und Optimalstandard) sowie Nordrhein-

Abbildung 14.1: Abweichung der für die 4. Jahrgangsstufe im Jahr 2016 ermittelten Anteile von Schülerinnen und Schülern, die den Regelstandard erreichen oder übertreffen, vom jeweiligen Anteil für Deutschland insgesamt (in Prozentpunkten)



■ Lesen im Fach Deutsch ■ Zuhören im Fach Deutsch ■ Orthografie im Fach Deutsch ■ Mathematik (Globalskala)
 Anmerkung. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Abbildung 14.2: Abweichung der für die 4. Jahrgangsstufe im Jahr 2016 ermittelten Anteile von Schülerinnen und Schülern, die den Mindeststandard nicht erreichen, vom jeweiligen Anteil für Deutschland insgesamt (in Prozentpunkten)



Abbildung 14.3: Abweichung der für die 4. Jahrgangsstufe im Jahr 2016 ermittelten Anteile von Schülerinnen und Schülern, die den Optimalstandard erreichen, vom jeweiligen Anteil für Deutschland insgesamt (in Prozentpunkten)



■ Lesen im Fach Deutsch ■ Zuhören im Fach Deutsch ■ Orthografie im Fach Deutsch ■ Mathematik (Globalskala)
 Anmerkung. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Westfalen (Mindeststandard) vor. Die genannten Abweichungen sind zwar alle statistisch signifikant, betragen jedoch teilweise nur wenige Prozentpunkte und sollten in diesem Fall nicht überinterpretiert werden.

14.2 Erreichen der Bildungsstandards in den Ländern im Trend

Auch für die Veränderungen in den Kompetenzstufenverteilungen über die Zeit, die in den Abbildungen 14.4 bis 14.6 dargestellt sind, ergibt sich ein differenziertes Bild (vgl. Kapitel 5). Nach rechts weisende Balken in den Abbildungen zeigen, dass der Anteil von Kindern im jeweiligen Land, die mindestens den Regelstandard erreichen, den Mindeststandard verfehlen beziehungsweise den Optimalstandard erreichen, im Jahr 2016 größer ist als im Jahr 2011, bei nach links weisenden Balken ist der entsprechende Anteil im Jahr 2016 kleiner als im Jahr 2011.

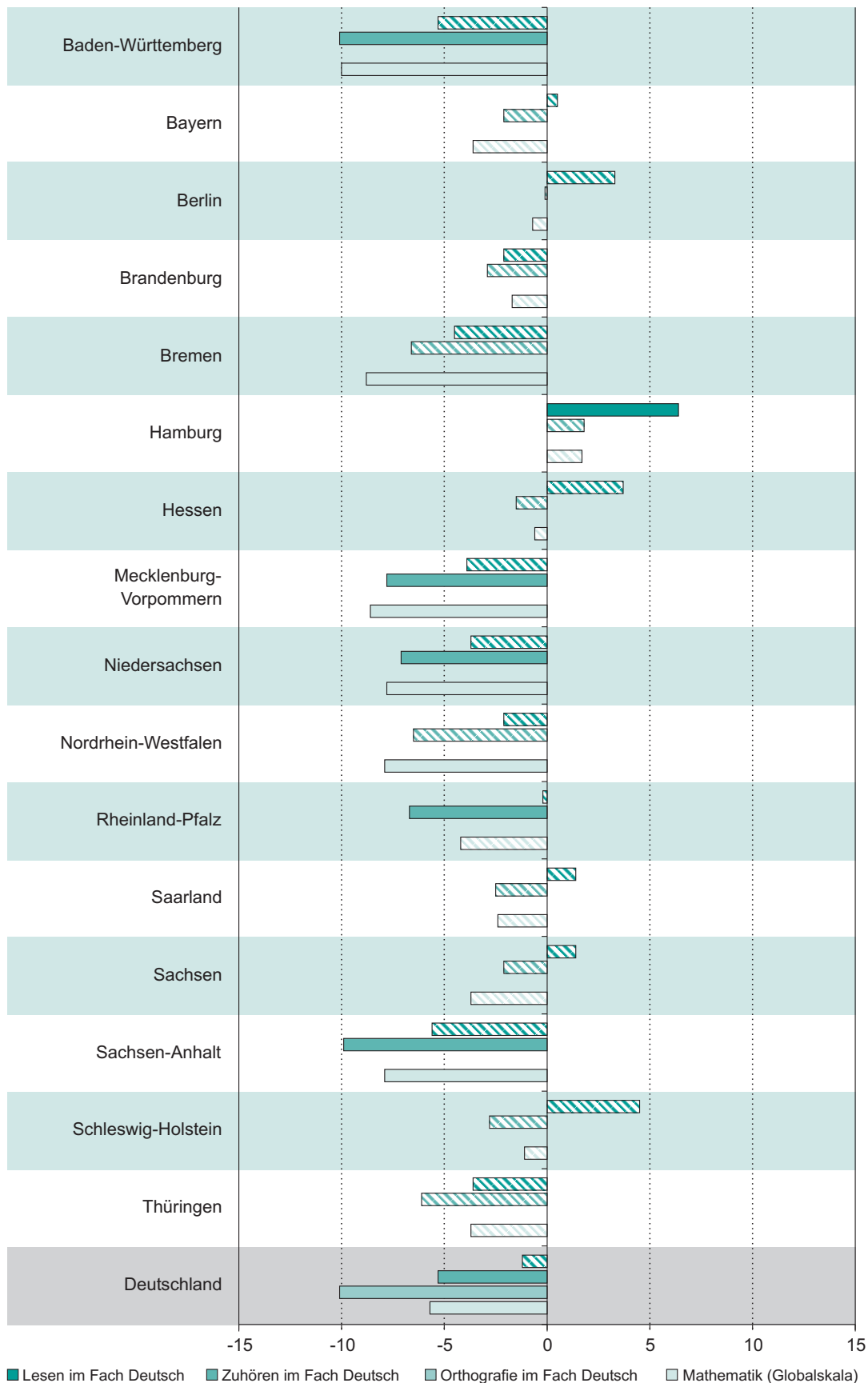
Im Fach Deutsch sind die Ergebnisse für Deutschland insgesamt im Bereich *Lesen* zwischen den Jahren 2011 und 2016 weitgehend stabil geblieben. Lediglich der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, hat sich im *Lesen* signifikant reduziert, wobei die Veränderung mit nur knapp 2 Prozentpunkten klein ist. Für die Bereiche *Zuhören* und *Orthografie* sind hingegen signifikant negative Trends zu verzeichnen, die größer ausfallen. Der Anteil der Viertklässlerinnen und Viertklässler, die mindestens den Regelstandard erreichen, ist im *Zuhören* um etwa 5 Prozentpunkte, in der *Orthografie* sogar um etwa 10 Prozentpunkte gesunken. Gleichzeitig ist der Anteil derjenigen, die den Mindeststandard verfehlen, um gut 3 Prozentpunkte beziehungsweise um fast 8 Prozentpunkte gestiegen. In der *Orthografie* hat sich zudem der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, um knapp 4 Prozentpunkte signifikant reduziert.

Im Fach Mathematik ist für Deutschland insgesamt ebenfalls ein signifikant negativer Trend zu verzeichnen. Hier ist der Anteil der Viertklässlerinnen und Viertklässler, die den Regelstandard erreichen oder übertreffen, um knapp 6 Prozentpunkte gesunken und der Anteil der Kinder, die den Mindeststandard verfehlen, um fast 4 Prozentpunkte gestiegen. Den Optimalstandard erreicht im Jahr 2016 ein um 3 Prozentpunkte signifikant geringerer Anteil der Schülerinnen und Schüler als im Jahr 2011.

Auch innerhalb der Länder sind vor allem für den Kompetenzbereich *Zuhören* und für das Fach Mathematik einige ungünstige Entwicklungen festzustellen. Da für die *Orthografie* auf Länderebene keine Trends bestimmt werden konnten (vgl. Kapitel 3.2), fehlen in den Abbildungen 14.4 bis 14.6 für diesen Kompetenzbereich länderbezogene Ergebnisdarstellungen.

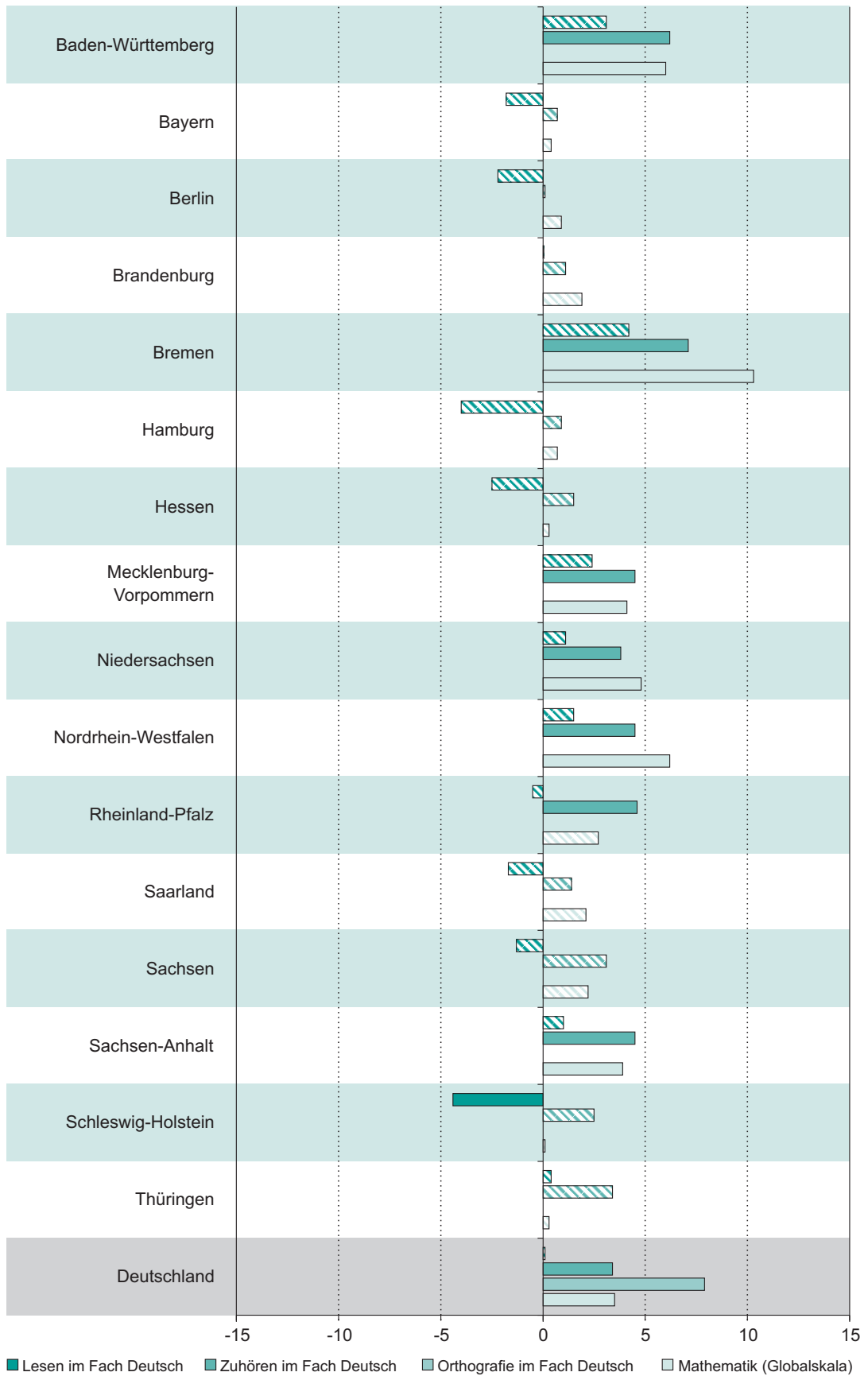
Für das Erreichen der Regelstandards sind signifikant negative Trends sowohl im *Zuhören* als auch im Fach Mathematik für die Länder Baden-Württemberg, Mecklenburg-Vorpommern, Niedersachsen und Sachsen-Anhalt zu verzeichnen. Im Fach Mathematik hat sich zudem in Bremen und in Nordrhein-Westfalen der Anteil der Schülerinnen und Schüler, die den Regelstandard erreichen, zwischen den Jahren 2011 und 2016 signifikant reduziert. Auch bezogen auf den Mindeststandard fallen die Entwicklungen in diesen Ländern (einschl. Bremen und Nordrhein-Westfalen) für den Bereich *Zuhören* und das Fach Mathematik ungünstig aus. In Rheinland-Pfalz liegt im Bereich *Zuhören* ebenfalls ein signifikant negativer Trend vor, sowohl mit Bezug auf den Regelstandard als auch mit Bezug auf den Mindeststandard. Signifikant positive Entwicklungen las-

Abbildung 14.4: Veränderungen in den Anteilen der Viertklässlerinnen und Viertklässler, die den Regelstandard erreichen oder übertreffen, zwischen den Jahren 2011 und 2016 (in Prozentpunkten)



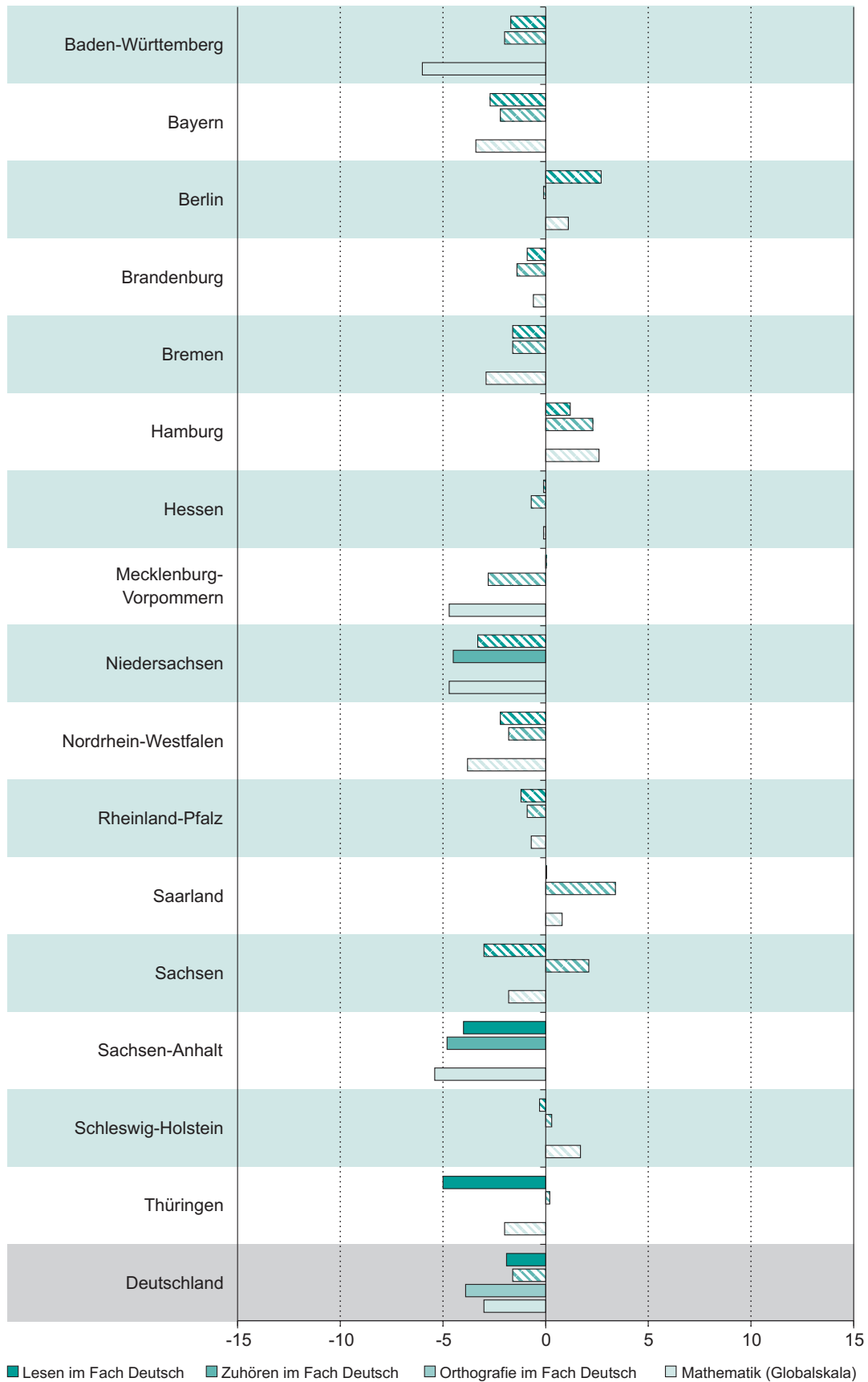
Anmerkungen. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Für den Kompetenzbereich Orthografie konnten keine Trendanalysen auf Länderebene durchgeführt werden.

Abbildung 14.5: Veränderungen in den Anteilen der Viertklässlerinnen und Viertklässler, die den Mindeststandard nicht erreichen, zwischen den Jahren 2011 und 2016 (in Prozentpunkten)



Anmerkungen. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Für den Kompetenzbereich Orthografie konnten keine Trendanalysen auf Länderebene durchgeführt werden.

Abbildung 14.6: Veränderungen in den Anteilen der Viertklässlerinnen und Viertklässler, die den Optimalstandard erreichen, zwischen den Jahren 2011 und 2016 (in Prozentpunkten)



Anmerkungen. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Für den Kompetenzbereich Orthografie konnten keine Trendanalysen auf Länderebene durchgeführt werden.

sen sich nur in zwei Fällen nachweisen: In Hamburg konnte der Anteil der Schülerinnen und Schüler, die im Kompetenzbereich *Lesen* den Regelstandard erreichen oder übertreffen, gesteigert und in Schleswig-Holstein konnte der Anteil der Viertklässlerinnen und Viertklässler, die im Kompetenzbereich *Lesen* den Mindeststandard verfehlen, verringert werden.

Bezogen auf die Optimalstandards sind innerhalb der Länder ebenfalls vereinzelt statistisch signifikante Trends zu beobachten, die ausschließlich negativ sind. So hat sich der Anteil der Schülerinnen und Schüler, die den Optimalstandard erreichen, in Sachsen-Anhalt im *Lesen*, im *Zuhören* und im Fach Mathematik, in Niedersachsen im Bereich *Zuhören* und im Fach Mathematik, in Thüringen im Bereich *Lesen* sowie in Baden-Württemberg und in Mecklenburg-Vorpommern im Fach Mathematik signifikant reduziert.

14.3 Durchschnittliches Niveau der Kompetenzen im Jahr 2016 und im Trend

Zusätzlich zu den Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen wurden im IQB-Bildungstrend 2016 auch die im Durchschnitt erreichten Kompetenzen im Ländervergleich untersucht (vgl. Kapitel 6). Für Deutschland insgesamt sind die Skalen in den untersuchten Kompetenzbereichen so definiert, dass sie im Jahr 2011 einen Mittelwert von $M = 500$ und eine Streuung von $SD = 100$ aufweisen.² In Abbildung 14.7 ist für das Jahr 2016 im Überblick dargestellt, inwieweit die Kompetenzmittelwerte der Schülerinnen und Schüler in den einzelnen Ländern vom bundesweiten Durchschnitt abweichen. Die Ergebnismuster sind dabei erwartungsgemäß den Ergebnismustern für das Erreichen der Bildungsstandards recht ähnlich.

Durchgängig signifikant positive Abweichungen von den deutschen Durchschnittswerten sind wiederum für Bayern zu verzeichnen. Auch in Sachsen fallen die von Schülerinnen und Schülern im Durchschnitt erreichten Kompetenzen in fast allen Kompetenzbereichen (*Lesen*, *Zuhören*, Mathematik) signifikant höher aus als in Deutschland insgesamt. In einzelnen Kompetenzbereichen sind zudem positive Abweichungen für das Saarland (*Orthografie*, Mathematik), für Sachsen-Anhalt (Mathematik) und für Schleswig-Holstein (*Lesen*, *Zuhören*) zu beobachten. Signifikant unter dem Bundesdurchschnitt liegen die Mittelwerte hingegen durchgängig in Berlin und Bremen. Vereinzelt sind zudem signifikant negative Abweichungen für Hamburg und Niedersachsen (*Orthografie*, Mathematik), Nordrhein-Westfalen (*Lesen*, Mathematik), Rheinland-Pfalz und Sachsen-Anhalt (*Zuhören*) sowie Thüringen (*Orthografie*) zu verzeichnen.

Zwischen dem höchsten und dem niedrigsten Kompetenzmittelwert, der im Fach Deutsch in den Ländern erreicht wird, beträgt der Abstand 70 Punkte im Bereich *Lesen*, 61 Punkte im Bereich *Zuhören* und 91 Punkte im Bereich *Orthografie*. Setzt man diese Differenzen mit den groben Schätzungen für den Kompetenzzuwachs in Beziehung, der am Ende der Primarstufe innerhalb eines Schuljahres zu erwarten ist (vgl. Kapitel 6.1), so entsprechen diese Mittelwertsunterschiede in allen drei Bereichen ungefähr einem Schuljahr Lernzeit. Im Fach Mathematik liegt der Abstand zwischen dem höchsten und dem

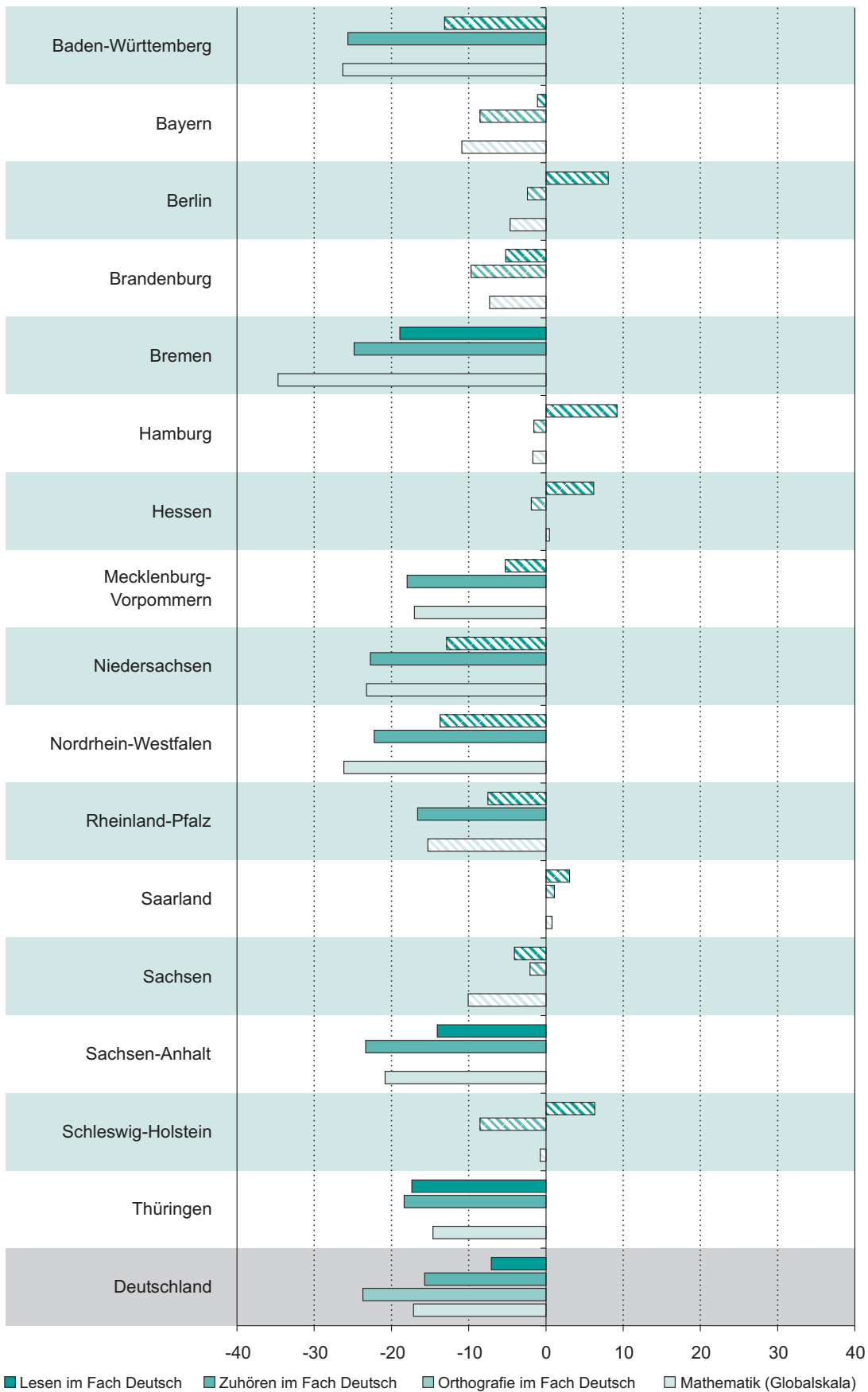
² Eine Ausnahme bildet der Kompetenzbereich *Orthografie*, für den sich der Mittelwert von 500 und die Standardabweichung von 100 auf die im Jahr 2016 getestete Population beziehen (vgl. Kapitel 3.2).

Abbildung 14.7: Abweichung der von den Viertklässlerinnen und Viertklässlern im Jahr 2016 in den Ländern erreichten mittleren Kompetenzwerte vom deutschen Gesamtmittelwert



Anmerkung. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an.

Abbildung 14.8: Veränderungen in den mittleren Kompetenzwerten der Viertklässlerinnen und Viertklässler zwischen den Jahren 2011 und 2016



Anmerkungen. Schraffierte Balken zeigen eine statistisch nicht signifikante Differenz an. Für den Kompetenzbereich Orthografie konnten keine Trendanalysen auf Länderebene durchgeführt werden.

niedrigsten Landesmittelwert bei 90 Punkten, was ebenfalls einem Unterschied von etwa einem Schuljahr entspricht (vgl. Kapitel 6.2).

Auch die Trendschätzungen für die Mittelwerte replizieren weitgehend die Ergebnismuster der Trendschätzungen für das Erreichen der Bildungsstandards, wobei die Veränderungen der Mittelwerte häufiger signifikant sind (vgl. Abb. 14.8). In Deutschland insgesamt sind die Mittelwerte, die im Jahr 2016 erreicht werden, in beiden Fächern signifikant geringer als im Jahr 2011. Im *Lesen* umfasst die Verringerung lediglich 7 Punkte, in den anderen untersuchten Bereichen ist sie jedoch deutlich größer. Die im Jahr 2016 im Durchschnitt erreichten Kompetenzen liegen im *Zuhören* um 16 Punkte, in der *Orthografie* um 24 Punkte und in Mathematik um 17 Punkte unter dem bundesdeutschen Mittelwert für das Jahr 2011. Auch innerhalb der Länder sind weder für das Fach Deutsch noch für das Fach Mathematik signifikant positive Trends zu verzeichnen. Signifikant negative Trends ergeben sich dagegen sowohl für die beiden Kompetenzbereiche im Fach Deutsch, die auf Länderebene in die Trendanalysen einbezogen werden können (*Lesen* und *Zuhören*), als auch für das Fach Mathematik in Bremen, Sachsen-Anhalt und Thüringen, für den Bereich *Zuhören* und für das Fach Mathematik in den Ländern Baden-Württemberg, Mecklenburg-Vorpommern, Niedersachsen und Nordrhein-Westfalen sowie für den Bereich *Zuhören* in Rheinland-Pfalz. In allen anderen Ländern (Bayern, Berlin, Brandenburg, Hamburg, Hessen, Saarland, Sachsen, Schleswig-Holstein) sind die im Durchschnitt erreichten Kompetenzen zwischen den Jahren 2011 und 2016 in den drei im Trend untersuchten Bereichen stabil geblieben.

14.4 Geschlechtsbezogene, soziale und zugewanderungsbezogene Disparitäten

In den intensiv geführten Diskussionen über Bildungsgerechtigkeit kommt der Frage, inwieweit die Förderung der Kompetenzentwicklung in verschiedenen Gruppen von Kindern und Jugendlichen gelingt, eine zentrale Bedeutung zu. Die Bildungssysteme aller Länder zielen darauf ab, bestehende Ungleichheiten zu reduzieren. Daher wurden im IQB-Bildungstrend 2016 Analysen zu geschlechtsbezogenen, sozialen und zugewanderungsbezogenen Disparitäten durchgeführt.

Geschlechtsbezogene Disparitäten

Die Ergebnisse der Analysen geschlechtsbezogener Disparitäten (vgl. Kapitel 7) zeigen, dass im Jahr 2016 in allen untersuchten Kompetenzbereichen in den Fächern Deutsch und Mathematik signifikante Unterschiede zwischen den von Jungen und den von Mädchen im Durchschnitt erreichten Kompetenzen bestehen. Die Unterschiede fallen stereotyp aus: Im Fach Deutsch erreichen Mädchen im Mittel höhere Kompetenzen als Jungen, wobei die Differenz in der *Orthografie* am größten ist (33 Punkte), gefolgt von der Differenz im Bereich *Lesen* (24 Punkte). Im *Zuhören* hingegen ist der Kompetenznachteil der Jungen geringer ausgeprägt (12 Punkte). Demgegenüber erzielten Jungen im Fach Mathematik sowohl auf der Globalskala (19 Punkte) als auch in allen inhaltlichen Kompetenzbereichen (von 6 Punkten im Bereich *Daten, Häufigkeit und Wahrscheinlichkeit* bis zu 33 Punkten im Bereich *Größen und Messen*) im Durchschnitt höhere Kompetenzwerte als Mädchen. Geschlechtsbezogene

Disparitäten sind dabei sowohl im unteren als auch im mittleren und oberen Bereich der Kompetenzverteilungen zu verzeichnen.

Innerhalb der Länder fällt das Ergebnismuster sehr ähnlich aus wie in Deutschland insgesamt. In allen Ländern besteht ein signifikanter Kompetenzvorteil der Mädchen im *Lesen* und in der *Orthografie* im Fach Deutsch sowie ein signifikanter Kompetenzvorteil der Jungen im Fach Mathematik. Eine Ausnahme bildet lediglich der Kompetenzbereich *Zuhören* im Fach Deutsch, für den in der Mehrzahl der Länder keine signifikanten Geschlechterunterschiede zu verzeichnen sind. Besonders klein sind die geschlechtsbezogenen Disparitäten im Fach Deutsch in allen untersuchten Kompetenzbereichen im Land Bremen. Demgegenüber ist in den Ländern Brandenburg und Nordrhein-Westfalen in allen drei Bereichen ein signifikanter Vorsprung der Mädchen zu verzeichnen, der in der Tendenz jeweils größer ausfällt als in Deutschland insgesamt. Auf der Globalskala im Fach Mathematik sind die geringsten geschlechtsbezogenen Disparitäten in Schleswig-Holstein zu finden, während sie in Berlin, Bremen, Hamburg und Thüringen besonders groß ausfallen.

Zwischen den Jahren 2011 und 2016 haben sich die Geschlechterunterschiede in den untersuchten Kompetenzbereichen weder in Deutschland insgesamt noch innerhalb der Länder signifikant verändert. Die einzige Ausnahme bildet der Bereich *Zuhören* im Fach Deutsch, in dem sich der Unterschied zwischen den Kompetenzmittelwerten der Jungen und Mädchen in Deutschland insgesamt um 9 Punkte signifikant erhöht hat. Auf Länderebene ist dieser Trend jedoch nur in einem Land (Nordrhein-Westfalen) statistisch signifikant.

Sowohl Jungen als auch Mädchen geben im Durchschnitt an, ein hohes Interesse an den Fächern Deutsch und Mathematik zu haben. Auch die eigenen Kompetenzen schätzt eine große Mehrheit in beiden Teilpopulationen als hoch ein. Insgesamt scheint also die fachbezogene Motivation der Schülerinnen und Schüler gegen Ende der Primarschulzeit recht günstig zu sein. Gleichzeitig bestehen jedoch Unterschiede zwischen Jungen und Mädchen in den untersuchten motivationalen Merkmalen, die in dieselbe Richtung weisen wie die geschlechtsbezogenen Disparitäten in den untersuchten Kompetenzen. Das Ergebnismuster für das Fach Mathematik kann dabei als Hinweis darauf gewertet werden, dass Mädchen ihre mathematische Kompetenz niedriger einschätzen als gleich kompetente Jungen. Da sich das Selbstkonzept der eigenen Fähigkeiten auf die weitere Interessens- und Kompetenzentwicklung und auf späteres Kurs- und Berufswahlverhalten auswirken kann, trägt dies möglicherweise zu einer Verfestigung oder Verstärkung geschlechtsbezogener Disparitäten im Bereich Mathematik bei.

Soziale Disparitäten

Die Frage, inwieweit der Kompetenzerwerb von Schülerinnen und Schülern mit Merkmalen ihres sozialen Hintergrunds zusammenhängt und in welchem Ausmaß es gelingt, diese Kopplung zu reduzieren, ist eine der zentralen Fragestellungen des Bildungsmonitorings. Daher wurden auch im IQB-Bildungstrend 2016 soziale Disparitäten untersucht und geprüft, inwieweit sich diese seit dem Jahr 2011 verändert haben (vgl. Kapitel 8). Aufgrund des hohen Anteils fehlender Daten zum sozioökonomischen Hintergrund der Schülerinnen und Schüler konnten für die Stadtstaaten Berlin, Bremen und Hamburg sowie für das Saarland keine so-

zialen Disparitäten berichtet werden. Damit liefern die Befunde dieser Analysen ein unvollständiges Bild der sozialen Disparitäten in Deutschland.

In Übereinstimmung mit Befunden früherer Studien zeigen die Ergebnisse des IQB-Bildungstrends für das Jahr 2016, dass sich die soziale Zusammensetzung der Schülerinnen und Schüler zwischen den Ländern, die in die Analysen einbezogen werden konnten, nur geringfügig unterscheidet.³ Im Vergleich der Jahre 2011 und 2016 ist der mittlere sozioökonomische Hintergrund weitgehend stabil geblieben. Die soziale Heterogenität, die anhand der Streuung des sozioökonomischen Status erfasst wurde, ist jedoch sowohl in Deutschland insgesamt als auch in einzelnen Ländern (Niedersachsen, Rheinland-Pfalz, Sachsen und Schleswig-Holstein) leicht gestiegen. Analysen der Verteilung dieser Heterogenität zeigen, dass die Unterschiede im sozioökonomischen Hintergrund der Schülerinnen und Schüler bundesweit zu etwa 85 Prozent innerhalb von Schulen und zu 15 Prozent zwischen Schulen liegen. Diese Verteilung unterscheidet sich weder zwischen den Ländern noch zwischen den Jahren 2011 und 2016 substantiell. Im Primarbereich ist also die soziale Zusammensetzung der Schülerschaft, die Lehrkräfte in ihren Klassen unterrichten, recht heterogen.

Die Kopplung zwischen sozialer Herkunft und den von Viertklässlerinnen und Viertklässlern erreichten Kompetenzen ist im Jahr 2016 für alle untersuchten Kompetenzbereiche sowohl deutschlandweit als auch in allen Ländern substantiell: Ein höherer Sozialstatus geht mit höheren Werten in den Kompetenztests einher. Allerdings fallen die sozialen Gradienten⁴ im Bereich *Orthografie* mit knapp 30 Punkten etwas geringer aus als in den Bereichen *Lesen* und *Zuhören* sowie im Fach Mathematik (jeweils etwa 40 Punkte). Demnach scheinen die orthografischen Kompetenzen etwas weniger stark vom sozialen Status der Familie abzuhängen als die anderen Kompetenzen. Zwischen den Ländern, die in die Analysen einbezogen werden konnten, bestehen im Jahr 2016 keine substantiellen Unterschiede im Grad der Kopplung des sozialen Hintergrunds mit den erreichten Kompetenzen.

Auch im Vergleich der sozialen Gradienten für die Jahre 2011 und 2016 ergeben sich kaum Unterschiede. Im Kompetenzbereich *Zuhören* und im Fach Mathematik haben die Disparitäten zwar in Deutschland insgesamt etwas zugenommen, die Veränderungen sind jedoch gering und in den beiden anderen Kompetenzbereichen, *Lesen* und *Orthografie*, sind bundesweit keine signifikanten Veränderungen zu verzeichnen. Nur in zwei Ländern fällt in jeweils einem Bereich die Kopplung zwischen sozialer Herkunft und erreichtem Kompetenzniveau im Jahr 2016 höher aus als im Jahr 2011: in Baden-Württemberg im Bereich *Lesen* und in Sachsen im Bereich *Zuhören*.

3 In früheren Studien, in denen alle Länder in die Analysen einbezogen werden konnten, unterschieden sich Familien in den Stadtstaaten tendenziell stärker im sozioökonomischen Status als Familien in den Flächenländern.

4 Soziale Gradienten beschreiben den linearen Zusammenhang zwischen dem sozioökonomischen Status der Familie und den erreichten Kompetenzen der Schülerinnen und Schüler. Sie geben an, wie viele Punkte auf der Berichtsmetrik Schülerinnen und Schüler durchschnittlich mehr erzielen würden, wenn der Wert für den sozioökonomischen Status ihrer Familie um eine Standardabweichung höher wäre.

Zuwanderungsbezogene Disparitäten

Inwieweit in der schulischen Kompetenzentwicklung zuwanderungsbezogene Disparitäten bestehen, gehört ebenfalls zu den zentralen Fragestellungen des Bildungsmonitorings und wurde daher auch im IQB-Bildungstrend 2016 erneut untersucht (vgl. Kapitel 9). Die Ergebnisse zeigen, dass sich der Anteil der Viertklässlerinnen und Viertklässler mit Zuwanderungshintergrund (mindestens ein Elternteil im Ausland geboren) seit dem Jahr 2011 in Deutschland insgesamt um 9 Prozentpunkte erhöht hat und im Jahr 2016 bei etwa 34 Prozent liegt. Ein Anstieg ist, wenn auch in unterschiedlichem Umfang, in fast allen Ländern zu beobachten und vor allem auf gestiegene Anteile von Kindern mit einem im Ausland geborenen Elternteil sowie von Kindern der zweiten Generation (in Deutschland geborene Kinder zugewanderter Eltern) zurückzuführen. Der Anteil von Kindern, die selbst im Ausland geboren und mit ihren Familien nach Deutschland zugewandert sind (erste Generation), hat sich dagegen zwischen den Jahren 2011 und 2016 deutlich weniger stark erhöht; bundesweit stieg er nur um rund 2 Prozentpunkte an.⁵

Analysen zur Verteilung der Kinder mit Zuwanderungshintergrund auf die Schulen in Deutschland verdeutlichen, dass sich die Schulen im Primarbereich in ihrer zuwanderungsbezogenen Zusammensetzung stark unterscheiden. Während bundesweit in etwa 13 Prozent der Schulen weniger als 5 Prozent der Kinder einen Zuwanderungshintergrund haben, liegt der Anteil von Kindern aus zugewanderten Familien in rund einem Viertel der Schulen bei mindestens 40 Prozent. Diese Verteilung hat sich im Vergleich zum Jahr 2011 verändert: Der Anteil an Schulen mit besonders wenigen Kindern mit Zuwanderungshintergrund ist gesunken und der Anteil an Schulen mit einer höheren Anzahl von Kindern mit Zuwanderungshintergrund ist gestiegen.

Im Fach Deutsch bestehen im Jahr 2016 in allen untersuchten Kompetenzbereichen signifikante Nachteile für Kinder aus zugewanderten Familien, die im Bereich *Zuhören* am größten und im Bereich *Orthografie* am kleinsten sind. Bundesweit fallen die Disparitäten für Kinder der ersten Zuwanderergeneration erwartungsgemäß am stärksten aus, während sie für Kinder der zweiten Zuwanderergeneration deutlich geringer sind. Dies könnte darauf hinweisen, dass die schulische Integration im Generationenverlauf zunehmend besser gelingt. Zudem sind die Kompetenznachteile von Kindern mit zwei im Ausland geborenen Elternteilen in fast allen Ländern, die Kompetenznachteile von Kindern mit einem im Ausland geborenen Elternteil dagegen nur in einigen Ländern statistisch signifikant. Allerdings variiert die Ausprägung der Disparitäten erheblich zwischen den Ländern. Besonders große Disparitäten sind in den Stadtstaaten Berlin, Bremen und Hamburg zu finden, wobei die Ergebnisse für Hamburg aufgrund eines erheblichen Anteils fehlender Werte unter Vorbehalt stehen (vgl. Kapitel 3.1). In einigen der ostdeutschen Flächenländer fallen die Kompetenznachteile besonders gering aus, allerdings ist der Anteil an Schülerinnen und Schülern aus zugewanderten Familien in diesen Ländern auch relativ klein.

5 Der deutliche Anstieg der Neuzuwanderung, der im Laufe des Jahres 2015 einsetzte, spiegelt sich in den Daten des IQB-Bildungstrends 2016 noch nicht wieder. Neu zugewanderte Schülerinnen und Schüler mit einer nichtdeutschen Herkunftssprache, die weniger als ein Jahr in deutscher Sprache unterrichtet wurden und nicht in der Lage waren, Deutsch zu lesen oder zu sprechen, nahmen nicht an der Testung teil (vgl. Kapitel 3.1). Also wurden Kinder, die nach dem Frühjahr 2015 ins deutsche Schulsystem gekommen sind, in der Regel nicht in die Erhebung einbezogen.

Im Fach Mathematik bestehen im Jahr 2016 ebenfalls Kompetenzunterschiede zuungunsten der Kinder mit Zuwanderungshintergrund, die in ihrer Größenordnung in etwa den Disparitäten im *Lesen* entsprechen.

Den Trendanalysen zufolge sind die mit dem Zuwanderungshintergrund verbundenen Unterschiede in den erreichten Kompetenzen in beiden untersuchten Fächern zwischen den Jahren 2011 und 2016 weitgehend stabil geblieben. Die Verringerung zuwanderungsbezogener Disparitäten in der Kompetenzentwicklung stellt somit sowohl im Fach Deutsch als auch im Fach Mathematik weiterhin eine Herausforderung dar.

Die zuwanderungsbezogenen Disparitäten lassen sich teilweise auf die sozialen Herkunftsmerkmale zurückführen. Zudem bestätigt sich die Bedeutung der in der Familie gesprochenen Sprache: Unter Berücksichtigung des sozioökonomischen Status der Familie, des Bildungsniveaus und des Geburtslands der Eltern sind die in den Fächern Deutsch und Mathematik erzielten Kompetenzen bei Schülerinnen und Schülern, die in der Familie „manchmal Deutsch“ oder „nie Deutsch“ sprechen, geringer ausgeprägt als bei Schülerinnen und Schülern mit ausschließlich deutscher Familiensprache. Diese Befunde unterstreichen nochmals, wie wichtig es ist, gerade auch Kindern nichtdeutscher Herkunftssprache geeignete Lerngelegenheiten für den Erwerb und die kontinuierliche Weiterentwicklung bildungssprachlicher Kompetenzen in der Instruktionssprache Deutsch zur Verfügung zu stellen.

Zusätzlich zu Disparitäten im Kompetenzerwerb wurden im Bericht über den IQB-Bildungstrend 2016 erstmals für den Primarbereich auch Analysen zu Indikatoren der sozialen Integration und der schulischen Motivation (fachspezifische Lernfreude) der Schülerinnen und Schüler präsentiert. Die Befunde zeigen, dass sich die Kinder unabhängig von ihrem Zuwanderungsstatus in ihren Schulen insgesamt gut integriert fühlen und dass Kinder mit Zuwanderungshintergrund zum Teil sogar noch etwas zufriedener mit ihrer Schule sind als Kinder ohne Zuwanderungshintergrund. Die schulische Motivation ist bei Schülerinnen und Schülern aus zugewanderten Familien mindestens ebenso hoch ausgeprägt wie bei Schülerinnen und Schülern ohne Zuwanderungshintergrund.

14.5 Umgang mit Heterogenität

Zusätzlich zur detaillierten Beschreibung der von Viertklässlerinnen und Viertklässlern erreichten Kompetenzen und der Disparitäten im Kompetenzerwerb wurden im Rahmen des IQB-Bildungstrends 2016 Zusatzanalysen zum Themenfeld „Umgang mit Heterogenität“ durchgeführt, das laut Gesamtstrategie der KMK zum Bildungsmonitoring bildungspolitisch von zentraler Bedeutung ist (KMK, 2015a). Der Fokus der Analysen liegt dabei zum einen auf Kindern mit sonderpädagogischem Förderbedarf, die in den aktuellen Bemühungen um die Umsetzung von Inklusion eine zentrale Rolle spielen (vgl. Kapitel 10). Zum anderen werden deskriptive Analysen zu besonders leistungsstarken Schülerinnen und Schülern dargestellt (vgl. Kapitel 11), die in den letzten Jahren ebenfalls verstärkt in das Blickfeld der Bildungspolitik gerückt sind. So hat die KMK eine Strategie zur Förderung leistungsstarker Schülerinnen und Schüler verabschiedet (KMK, 2015b) und gemeinsam mit dem Bund eine Initiative „zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler“ beschlossen (KMK, 2016).

Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf

Am IQB-Bildungstrend 2016 nahmen auch Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) an allgemeinen Schulen und an Förderschulen teil. Damit bietet der IQB-Bildungstrend 2016 eine bundesweit repräsentative Datengrundlage für Kinder der 4. Jahrgangsstufe mit SPF in den Förderschwerpunkten „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“. Anhand dieser Daten wurden deskriptive Analysen zu drei Fragestellungen durchgeführt: (1) Übereinstimmung zwischen amtlich festgestelltem sonderpädagogischem Förderbedarf und sonderpädagogischer Förderung, (2) Zusammenhänge verschiedener Hintergrundmerkmale von Schülerinnen und Schülern mit der Feststellung eines sonderpädagogischen Förderbedarfs und der besuchten Schulart sowie (3) Zusammenhänge der von Schülerinnen und Schülern mit SPF besuchten Schulart mit den von ihnen erreichten Kompetenzen und motivationalen Merkmalen.

In der Praxis werden Schülerinnen und Schüler mit SPF in den Ländern anhand von unterschiedlichen Kriterien erfasst, wobei einerseits der amtlich festgestellte Förderbedarf und andererseits die sonderpädagogische Förderung eine Rolle spielen. Da unklar ist, inwieweit die anhand dieser beiden Kriterien identifizierten Schülergruppen übereinstimmen, wurde dieser Frage auf der Grundlage von Daten des IQB-Bildungstrends 2016 nachgegangen (vgl. Kapitel 10.2). Die Ergebnisse zeigen, dass zwischen den Kriterien deutliche Überschneidungen bestehen, die damit erfassten Gruppen aber nicht deckungsgleich sind. Die Heterogenität des Vorgehens bei der Zuweisung eines SPF-Status sollte daher bei der Erfassung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in bundesweiten Erhebungen berücksichtigt werden.

In den Analysen von Zusammenhängen verschiedener Hintergrundmerkmale von Viertklässlerinnen und Viertklässlern (Geschlecht, soziale Herkunft und Zuwanderungshintergrund) mit der Feststellung eines SPF und der besuchten Schulart (Förderschule oder allgemeine Schule) zeigt sich, dass vor allem das Geschlecht und die soziale Herkunft bedeutsam sind (vgl. Kapitel 10.3): Jungen und Kinder aus Familien mit einem niedrigen sozialen Status haben häufiger einen festgestellten SPF und besuchen häufiger eine Förderschule als Mädchen und Kinder aus Familien mit einem hohen sozioökonomischen Status. Zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund sind im IQB-Bildungstrend 2016 hingegen weder für die Häufigkeit der Diagnose eines SPF insgesamt noch für die Häufigkeit des Besuchs einer Förderschule signifikante Unterschiede zu verzeichnen.

Um Hinweise zur Frage zu gewinnen, in welcher Schulart Schülerinnen und Schüler mit SPF in den Förderschwerpunkten „Lernen“, „Sprache“ sowie „emotionale und soziale Entwicklung“ günstigere Ergebnisse erzielen, wurden – unter Berücksichtigung einer Reihe von Hintergrundmerkmalen der Kinder – Vergleiche zwischen Förderschulen und allgemeinen Schulen durchgeführt. Dabei wurden in den Fächern Deutsch und Mathematik sowohl die erreichten Kompetenzen als auch motivationale Merkmale (fachspezifisches akademisches Selbstkonzept und Interesse) in den Blick genommen. Die Ergebnisse weisen auf gegenläufige Zusammenhänge der besuchten Schulart mit den schulischen Kompetenzen einerseits und mit der schulischen Motivation andererseits hin (vgl. Kapitel 10.4). Während Kinder mit SPF in allgemeinen Schulen im Durchschnitt höhere Kompetenzen erzielen, weisen Kinder mit SPF in Förderschulen insgesamt eine höhere schulische Motivation auf, wobei die Stärke

dieser Zusammenhänge je nach Kompetenzbereich und Förderschwerpunkt variiert. Die Kompetenzunterschiede zugunsten von allgemeinen Schulen fallen für Kinder mit dem Förderschwerpunkt „Lernen“ stärker aus als für Kinder mit dem Förderschwerpunkt „Sprache“; für Kinder mit dem Förderschwerpunkt „emotionale und soziale Entwicklung“ zeigen sich hingegen keine bedeutsamen Kompetenzunterschiede in Abhängigkeit von der besuchten Schulart. Hinsichtlich der schulischen Motivation bestehen Unterschiede zugunsten von Förderschulen, die bei Kindern mit dem Förderschwerpunkt „Lernen“ und „emotionale und soziale Entwicklung“ tendenziell stärker ausgeprägt sind als bei Kindern mit dem Förderschwerpunkt „Sprache“. Gerade auch in allgemeinen Schulen besteht damit eine zentrale Herausforderung darin, Unterricht so zu gestalten, dass er den individuellen Bedürfnissen von Kindern mit SPF nicht nur im Hinblick auf ihre Kompetenzentwicklung, sondern auch in Bezug auf ihre motivationale Entwicklung gerecht wird.

Leistungsstarke Schülerinnen und Schüler

Im IQB-Bildungstrend 2016 wurden leistungsstarke Schülerinnen und Schüler anhand der in der Studie erfassten Kompetenzen identifiziert. Im Fach Deutsch wurden Kinder als leistungsstark eingeordnet, wenn sie in mindestens einem der drei getesteten Kompetenzbereiche (*Lesen, Zuhören, Orthografie*) den Optimalstandard erreicht und gleichzeitig in keinem der Kompetenzbereiche den Regelstandard verfehlt haben. Im Fach Mathematik erfolgte die Zuordnung zur Gruppe leistungsstarker Schülerinnen und Schüler, wenn sie auf der Globalskala den Optimalstandard erreicht haben. Die Ergebnisse deskriptiver Analysen zeigen, dass insgesamt knapp 23 Prozent der Viertklässlerinnen und Viertklässler in mindestens einem der beiden untersuchten Fächer leistungsstark sind (vgl. Kapitel 11). Dabei liegt der Anteil der Kinder, die nur im Fach Deutsch leistungsstark sind, bei etwa 10 Prozent und im Fach Mathematik bei ungefähr 6 Prozent. Knapp 7 Prozent der Schülerinnen und Schüler erweisen sich in beiden Fächern als leistungsstark. Dabei bestätigen sich die bekannten Geschlechterunterschiede, wonach bereits in der 4. Jahrgangsstufe Mädchen häufiger als Jungen im Fach Deutsch und Jungen häufiger als Mädchen im Fach Mathematik besonders hohe Leistungen erzielen. In der Gruppe der Kinder, die in beiden Fächern zugleich leistungsstark sind, sind jedoch Mädchen und Jungen in etwa gleich stark vertreten.

Leistungsstarke Kinder wachsen zwar überwiegend in Familien mit einem hohen sozialen Status und Bildungsniveau auf, im Fach Deutsch und im Fach Mathematik hat jedoch jeweils ein Drittel von ihnen Eltern, die selbst kein Abitur erworben haben; in der Gruppe der Kinder, die in beiden Fächern leistungsstark sind, ist es etwa ein Viertel. Schülerinnen und Schüler unabhängig vom sozialen Status und Bildungsniveau des Elternhauses in ihrer Bildungslaufbahn optimal zu fördern, bleibt mit Blick auf das Ziel, die Kopplung zwischen sozialer Herkunft und schulischem Erfolg zu verringern, eine wichtige bildungspolitische Aufgabe.

Die analysierten motivationalen Merkmale (Elterneinschätzung der Leistungsbereitschaft, selbst berichtetes fachliches Interesse und Selbstkonzept) sind bei leistungsstarken Schülerinnen und Schülern insgesamt hoch ausgeprägt. Erfreulich ist ferner, dass bei ihnen das Gefühl der Langeweile im Fachunterricht sogar noch etwas weniger ausgeprägt ist als bei Schülerinnen und Schülern, die in keinem der untersuchten Fächer besondere Leistungsstärken aufweisen.

Insgesamt scheint es Schulen im Primarbereich also zu gelingen, den Unterricht so zu gestalten, dass er auch für Kinder, die bereits das Niveau der Optimalstandards erreicht haben, motivational anregend ist.

14.6 Qualifikation von Lehrkräften und Einschätzungen zu gemeinsamem Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf

Da Lehrkräfte für die Kompetenzentwicklung von Schülerinnen und Schülern eine zentrale Rolle spielen, werden in den Studien des IQB zum Bildungsmonitoring immer auch Befragungen von Lehrerinnen und Lehrern durchgeführt und für deskriptive Analysen genutzt. Neben einer Beschreibung der Lehrerschaft anhand demografischer Merkmale wird in den Berichten insbesondere die berufliche Qualifikation der Lehrkräfte untersucht und der Frage nachgegangen, inwieweit Schülerinnen und Schüler von Lehrkräften, die im Rahmen ihres Lehramtsstudiums das von ihnen unterrichtete Fach studiert beziehungsweise eine Lehrbefähigung für das Fach erworben haben, höhere Kompetenzen erreichen. Diese Analysen wurden im IQB-Bildungstrend 2016 fortgeführt (vgl. Kapitel 12).

Darüber hinaus wurden in der Lehrkräftebefragung im IQB-Bildungstrend 2016 erstmals Fragen gestellt, die sich auf den gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf beziehen und somit ebenfalls das Schwerpunktthema „Umgang mit Heterogenität“ aufgreifen. Dies umfasst insbesondere die Qualifikation im Bereich der Sonderpädagogik, den wahrgenommenen Fortbildungsbedarf sowie Erfahrungen mit gemeinsamem Unterricht und verschiedene Einschätzungen dazu.

An der Lehrkräftebefragung im IQB-Bildungstrend 2016 nahmen bundesweit insgesamt 1784 Lehrkräfte an allgemeinen Schulen teil.⁶ Die Ergebnisse zur Qualifikation zeigen, dass der Anteil der Lehrkräfte, die das von ihnen unterrichtete Fach weder als Studienfach noch als Schwerpunkt mit umfangreicherem Stundenumfang studiert haben, sowohl in Deutschland insgesamt als auch innerhalb der Länder im Fach Mathematik höher ist als im Fach Deutsch. Bundesweit liegt der Anteil der fachfremd unterrichtenden Lehrkräfte in der Stichprobe des IQB-Bildungstrends 2016 bei knapp 19 Prozent im Fach Deutsch und bei gut 31 Prozent im Fach Mathematik. Besonders hoch ist der Anteil fachfremd erteilten Mathematikunterrichts in den Ländern Baden-Württemberg, Rheinland-Pfalz und dem Saarland, besonders niedrig hingegen in den Ländern Mecklenburg-Vorpommern, Sachsen-Anhalt und Thüringen. Der Anteil von Quereinsteigern, die ohne ein Lehramtsstudium die teilnehmenden Klassen unterrichten, ist mit je einem Prozent der Deutsch- und Mathematiklehrkräfte dagegen sehr gering.

Analysen zur Frage, ob die Qualifikation von Lehrkräften in den Fächern Deutsch und Mathematik mit den von ihren Schülerinnen und Schülern erreichten Kompetenzen zusammenhängen, weisen darauf hin, dass dies nicht der Fall ist. Weder im Fach Deutsch noch im Fach Mathematik sind die Kompetenzunterschiede zwischen Schülerinnen und Schülern, deren Lehrkräfte fachfremd unterrichten, und Schülerinnen und Schülern, deren Lehrkräfte das jeweilige

6 Da in Hamburg und im Saarland für einen hohen Anteil der teilnehmenden Klassen keine Angaben der Lehrkräfte vorliegen, stehen die Ergebnisse dieser Länder unter Vorbehalt. Im Saarland bezieht sich dies insbesondere auf die Angaben zu demografischen Hintergrundmerkmalen und zur Qualifikation der Lehrkräfte.

Fach im Lehramtsstudium studiert haben, statistisch signifikant. Dieser Befund weicht von den Ergebnissen entsprechender Analysen in der Sekundarstufe I ab, die Kompetenznachteile bei fachfremd unterrichteten Schülerinnen und Schülern insbesondere an nicht-gymnasialen Schularten nachgewiesen haben. Demnach scheint die mit einem Fachstudium verbundene Expertise von Lehrkräften in weiterführenden Schulen von größerer Bedeutung für den Lernerfolg der Schülerinnen und Schüler zu sein als im Primarbereich.

In den Analysen der Lehrkräftebefragung zu verschiedenen Aspekten des gemeinsamen Unterrichts von Kindern mit SPF und Kindern ohne SPF lag der Fokus auf denjenigen Lehrerinnen und Lehrern an allgemeinen Schulen, die in der am IQB-Bildungstrend 2016 teilnehmenden Klasse mindestens ein Kind mit sonderpädagogischem Förderbedarf unterrichten. Dies ist bei gut der Hälfte der befragten Lehrkräfte der Fall.⁷ Nur wenige dieser Lehrerinnen und Lehrer haben eine sonder- oder inklusionspädagogische Qualifikation im Rahmen ihres Lehramtsstudiums erworben und viele von ihnen geben einen hohen Fortbildungsbedarf im Bereich des gemeinsamen Unterrichts von Kindern mit und Kindern ohne SPF an. Gleichzeitig verfügen die meisten Lehrkräfte jedoch über mehrjährige Erfahrung im gemeinsamen Unterrichten von Kindern mit und Kindern ohne SPF. Im Durchschnitt umfasst die von den Lehrkräften berichtete Erfahrung 7 Jahre, und nur sehr wenige Lehrkräfte (rund 1 Prozent) geben an, weniger als ein Jahr Erfahrung mit gemeinsamem Unterricht zu haben.

Ihre auf gemeinsames Unterrichten von Kindern mit und Kindern ohne SPF bezogenen Kompetenzen schätzen die Lehrkräfte moderat ein: Die meisten von ihnen fühlen sich weder besonders gut noch besonders schlecht durch ihre Ausbildung, Fortbildung und Berufserfahrung darauf vorbereitet. Auch die Einschätzungen der Lehrerinnen und Lehrer zur Frage, ob sich Kinder mit SPF an allgemeinen Schulen besser entwickeln als an Förderschulen, sind weder eindeutig positiv noch eindeutig negativ. Gleichzeitig sind die Lehrkräfte überzeugt, den Unterricht in der am IQB-Bildungstrend 2016 teilnehmenden Klasse, in der mindestens eine Schülerin oder ein Schüler einen SPF hat, gut gestalten zu können.

14.7 Fazit

Insgesamt ergibt der IQB-Bildungstrend 2016 für die von Viertklässlerinnen und Viertklässlern in den Fächern Deutsch und Mathematik erreichten Kompetenzen ein Bild, das teilweise auf Stabilität, teilweise aber auch auf eher ungünstige Veränderungen über die Zeit hinweist. Dies umfasst negative Trends, die bundesweit vor allem in den Kompetenzbereichen *Zuhören* und *Orthografie* im Fach Deutsch sowie im Fach Mathematik recht ausgeprägt sind. Auch innerhalb der Länder sind kaum signifikant positive Veränderungen zu verzeichnen, wobei erhebliche Unterschiede darin bestehen, welches Kompetenzniveau die Schülerinnen und Schüler im Jahr 2016 erreichen. Bei der Interpretation der Trendergebnisse ist ferner zu beachten, dass sich im Zeitraum zwischen den Jahren 2011 und 2016 in den Schulsystemen der Länder einige Veränderungen vollzogen haben. So hat sich die Heterogenität der Schülerinnen und Schüler unter anderem dadurch erhöht, dass der Anteil der Kinder aus zugewanderten

⁷ Der Anteil der Schülerinnen und Schüler mit SPF, die in der 4. Jahrgangsstufe an allgemeinen Schulen unterrichtet werden, ist zwischen den Jahren 2011 und 2016 um etwa 17 Prozent gestiegen (vgl. Kapitel 4).

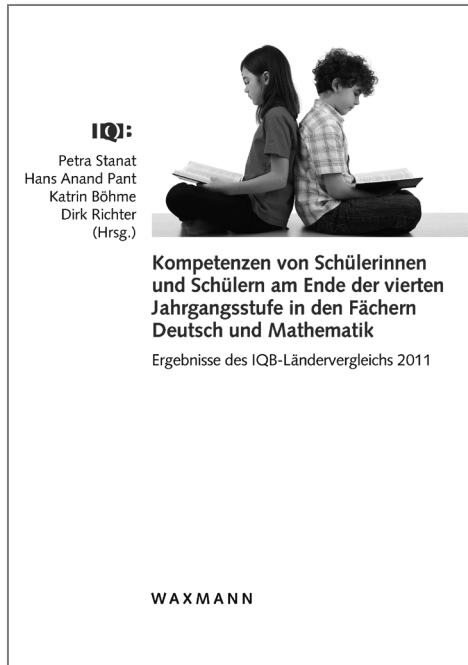
Familien an deutschen Schulen im Primarbereich gestiegen ist (vgl. Kapitel 9). Durch die Umsetzung der Inklusion besuchen zudem mehr Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf eine allgemeine Schule, wodurch sich die Zusammensetzung der Schülerschaft in dieser Schulart zusätzlich verändert hat (vgl. Kapitel 4). Je nach Bundesland waren im Untersuchungszeitraum weitere Veränderungen im Schulsystem zu bewältigen, wie beispielsweise der beginnende Generationswechsel im Kollegium der Lehrkräfte (vgl. Kapitel 12). Angesichts solcher Veränderungen kann es unter Umständen auch als Erfolg zu bewerten sein, wenn schulische Erträge stabil bleiben.

Wie die Ergebnismuster einzuschätzen und welche Schlussfolgerungen daraus zu ziehen sind, wird innerhalb der einzelnen Länder zu diskutieren sein. Zentral ist dabei die Frage, wie Schülerinnen und Schüler in den Fächern und Kompetenzbereichen, in denen Optimierungsbedarf besteht, durch eine gezielte Weiterentwicklung des Unterrichts besser gefördert werden können und wie die Rahmenbedingungen und Prozesse im Bildungssystem gestaltet werden müssen, damit dies gelingen kann. Eine der größten Herausforderungen besteht im Umgang mit der Heterogenität der Schülerschaft, die voraussichtlich auch in Zukunft weiter zunehmen wird. Diese Herausforderung kann nur durch gemeinsames Handeln aller Akteure im Schulsystem bewältigt werden.

Literatur

- KMK (2015a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Beschluss der 350. Kultusministerkonferenz vom 11.06.2015*. Zugriff am 15.08.2016 unter https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf
- KMK (2015b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Förderstrategie für leistungsstarke Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 11.06.2015*. Zugriff am 14.09.2017 unter http://www.kmk.org/fileadmin/Dateien/pdf/350-KMK-TOP-011-Fu-Leistungsstarke_-_neu.pdf
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 10.11.2016*. Zugriff am 14.09.2017 unter https://www.bmbf.de/files/Initiative_Leistungsstarke_Beschluss.pdf

UNSERE BUCHEMPFEHLUNG



Petra Stanat, Hans Anand Pant,
Katrin 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

2012, 300 Seiten, br., 32,90 €,
ISBN 978-3-8309-2777-8

E-Book: 29,99 €,
ISBN 978-3-8309-7777-3

In diesem Band wird über den ersten Ländervergleich des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) im Primarbereich berichtet. Im Fach Mathematik werden in dieser Studie erstmals die Kompetenzstände von Schülerinnen und Schülern der vierten Jahrgangsstufe, die in den Ländern der Bundesrepublik erreicht werden, verglichen.

Neben der Untersuchung der Kompetenzstände werden in diesem Bericht auch Zusammenhänge zwischen den erreichten Kompetenzen und verschiedenen Hintergrundmerkmalen der Schülerinnen und Schüler analysiert. Ergänzend werden Informationen zu den im Primarbereich tätigen Lehrkräften sowie zu Maßnahmen der Sprach- und Leseförderung in der Grundschule berichtet. Im Fokus der Studie stehen jedoch die Kompetenzen Lesen und Zuhören im Fach Deutsch sowie die mathematische Kompetenz, insgesamt und differenziert nach den fünf inhaltlichen Leitideen. Die repräsentativen Erhebungen für den Ländervergleich fanden 2011 an insgesamt 1.349 Schulen statt.



UNSERE BUCHEMPFEHLUNG



Heike Wendt, Wilfried Bos,
Christoph Selter, Olaf Köller,
Knut Schwippert, Daniel Kasper
(Hrsg.)

TIMSS 2015

Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich

2016, 408 Seiten, br., 39,90 €,
ISBN 978-3-8309-3566-7

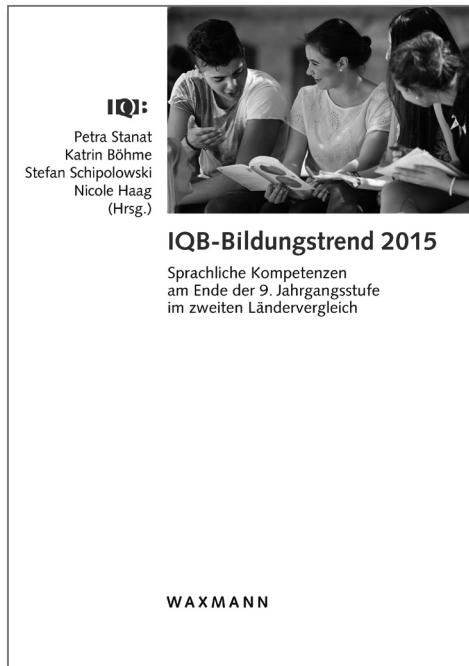
E-Book: Open Access

Im Jahr 2015 beteiligte sich Deutschland zum dritten Mal an der Grundschuluntersuchung *Trends in International Mathematics and Science Study* (TIMSS 2015). Mit TIMSS werden alle vier Jahre die Fachleistungen von Schülerinnen und Schülern der vierten Jahrgangsstufe in den Bereichen Mathematik und Naturwissenschaften im internationalen Vergleich untersucht.

In diesem Band werden die Ergebnisse von TIMSS 2015 vor dem Hintergrund von Themen dargestellt, die das Lehren und Lernen an Grundschulen in Deutschland verändert und den Bildungsdiskurs der letzten Jahre besonders geprägt haben. Im Fokus stehen Leistungsdisparitäten zwischen Jungen und Mädchen sowie Kindern unterschiedlicher sozialer und kultureller Herkunft. Darüber hinaus werden zentrale Lehr- und Lernbedingungen in den Blick genommen, wobei Gestaltungsmerkmale des Unterrichts und die Aus- und Fortbildung von Lehrkräften, Lernbedingungen an Ganztagschulen sowie die Inanspruchnahme von Nachhilfe differenziert betrachtet werden. Darüber hinaus werden der Übergang von der Grundschule in die Sekundarstufe I und soziale Kompetenzen von Grundschulkindern untersucht.



UNSERE BUCHEMPFEHLUNG



Petra Stanat, Katrin Böhme,
Stefan Schipolowski, Nicole Haag
(Hrsg.)

IQB-Bildungstrend 2015 Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich

2016, 544 Seiten, br., 44,90 €,
ISBN 978-3-8309-3535-3

E-Book: Open Access

Im IQB-Bildungstrend 2015 wird über die Ergebnisse des zweiten Ländervergleichs des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) in den sprachlichen Fächern berichtet. Untersucht werden Kompetenzen von Schülerinnen und Schülern der 9. Jahrgangsstufe in den Fächern Deutsch, Englisch und Französisch. Ein Fokus liegt dabei auf Trendanalysen, die zeigen, inwieweit sich das erreichte Kompetenzniveau seit dem IQB-Ländervergleich 2009 verändert hat. Die Referenzgröße bilden die länderübergreifenden Bildungsstandards der Kultusministerkonferenz, die fächerspezifisch festlegen, welche Kompetenzen Schülerinnen und Schüler bis zu einem bestimmten Punkt in ihrer Schullaufbahn entwickelt haben sollen.

Neben der Untersuchung der sprachlichen Kompetenzen in den Bereichen *Lesen, Zuhören* und *Orthografie* im Fach Deutsch sowie im *Lese- und Hörverstehen* in den fremdsprachlichen Fächern werden in diesem Bericht auch geschlechtsbezogene, soziale und zugewanderungsbezogene Disparitäten analysiert und ebenfalls überprüft, inwieweit hierfür seit dem Jahr 2009 Veränderungen festzustellen sind. Ergänzend werden Befunde zur Aus- und Fortbildung von Deutsch- und Englischlehrkräften berichtet.



UNSERE BUCHEMPFEHLUNG



Anke Hußmann, Wilfried Bos,
Albert Bremerich-Vos, Daniel Kasper,
Eva-Maria Lankes, Nele McElvany,
Tobias C. Stubbe, Renate Valtin,
Heike Wendt (Hrsg.)

IGLU 2016

Lesekompetenzen von
Grundschulkindern
in Deutschland im
internationalen Vergleich

2017, br., ISBN 978-3-8309-3700-5

Im Jahr 2016 beteiligte sich Deutschland zum vierten Mal an der Schulleistungsstudie *Internationale Grundschul-Lese-Untersuchung* (IGLU). Mit IGLU wird seit 2001 alle fünf Jahre das Leseverständnis von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe untersucht.

In diesem Band werden die Ergebnisse von IGLU 2016 in Deutschland im internationalen Vergleich berichtet. Die Ergebnisse werden vor dem Hintergrund von Themen dargestellt, die das Lehren und Lernen an Grundschulen in Deutschland verändert und den Bildungsdiskurs der letzten Jahre besonders geprägt haben. Die Lesekompetenzen der Grundschul Kinder werden anhand von zentralen Merkmalen wie Geschlecht, soziale Herkunft, Migrationsstatus und Gestaltungsmerkmalen des Unterrichts sowie im Hinblick auf den Übergang von der Primar- in die Sekundarstufe I analysiert. Mit vertiefenden Analysen und der Einordnung der Ergebnisse in den aktuellen Forschungsstand stellt der Band eine differenzierte und anschlussfähige Bestandsaufnahme zur Leistungsfähigkeit der Grundschule dar und beschreibt Entwicklungen seit 2001.

Der Bericht wendet sich an eine Leserschaft, die an bildungspolitischen, pädagogischen und fachdidaktischen Fragen interessiert ist.

