

IGLU 2016

Lesekompetenzen von Grundschulkindern
in Deutschland im internationalen Vergleich

Anke Hußmann, Heike Wendt, Wilfried Bos,
Albert Bremerich-Vos, Daniel Kasper, Eva-Maria Lankes,
Nele McElvany, Tobias C. Stubbe, Renate Valtin (Hrsg.)

IGLU 2016

Lesekompetenzen von Grundschulkindern
in Deutschland im internationalen Vergleich



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 978-3-8309-3700-5

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

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Christian Aeverbeck, Münster
Titelfoto: © BillionPhotos.com – 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

Vorwort 11

Wilfried Bos

Kapitel I

IGLU 2016: Wichtige Ergebnisse im Überblick..... 13

Wilfried Bos, Renate Valtin, Anke Hußmann, Heike Wendt und Martin Goy

1	Einleitung.....	13
2	Zentrale Ergebnisse	14
2.1	Lesekompetenzen im internationalen Vergleich.....	14
2.2	Leseselbstkonzept, Lesemotivation und Leseverhalten	17
2.3	Chancengleichheit und soziale Disparitäten	20
2.4	Lehr- und Lernbedingungen	24
3	Resümee.....	26
	Literatur.....	27

Kapitel II

Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2016) 29

Anke Hußmann, Heike Wendt, Daniel Kasper, Wilfried Bos und Martin Goy

1	Die Beteiligung Deutschlands an international vergleichenden Schulleistungsstudien zur Lesekompetenz am Ende der Grundschulzeit.....	30
2	Zentrale Erkenntnisse und Fragestellungen	32
3	IGLU 2016 – ein kooperatives Forschungsprojekt.....	34
4	Anlage und Durchführung von IGLU 2016.....	36
4.1	Die Rahmenkonzeption der Studie.....	36
4.2	Teilnahme und Teilnahmemodalitäten an IGLU 2016.....	39
4.3	Verfahren und Kriterien der Stichprobenziehung	46
5	Entwicklung und Charakteristika der Instrumente.....	52
5.1	Leistungstests.....	52
5.2	Kontextfragebögen.....	56
6	Erhebung.....	59
6.1	Aufbau der Untersuchung	59
6.2	Durchführung der Erhebung.....	60
6.3	Qualitätssicherung	60
7	Stichprobe und Beteiligungsquoten in Deutschland	61
8	Aufbereitung und Analyse der Daten.....	62
9	Skalierung der Leistungstests	63
10	Gewichtung und Schätzung von Stichproben- und Messfehlern.....	65
11	Zur Darstellung und Interpretation der Ergebnisse.....	66
	Literatur.....	73

Kapitel III

Lesekompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse 79

Albert Bremerich-Vos, Heike Wendt und Wilfried Bos

1	Einleitung.....	79
2	Die IGLU-Rahmenkonzeption zur Erfassung von Lesekompetenz.....	80
2.1	Leseintentionen beziehungsweise Textsorten.....	81
2.2	Verstehensprozesse beziehungsweise -leistungen	82
3	Der IGLU-Test zur Erfassung der Lesekompetenz.....	85
3.1	International eingesetzte Testvarianten und der Aufbau von IGLU	85
3.2	Zur curricularen Validität der IGLU-Lesetests	88
4	Das Kompetenzstufenmodell von IGLU 2016.....	91
4.1	Beschreibung der Kompetenzstufen in IGLU 2016.....	92
4.2	Texte und Aufgaben.....	95
5	Zentrale Befunde aus IGLU 2001, 2006 und 2011	110
6	Leseleistungen von Schülerinnen und Schülern im internationalen Vergleich: Ergebnisse 2016.....	111
6.1	Mittlere Leseleistungen im internationalen Vergleich	112
6.2	Mittlere Leseleistungen auf der Gesamtskala: Lesen im Trend	115
6.3	Leistungsheterogenität im internationalen Vergleich 2016.....	121
6.4	Lesekompetenzstufen.....	125
6.5	Kompetenzen beim Lesen von literarischen und informierenden Texten	130
6.6	Ergebnisse zu den Verstehensleistungen	134
7	Zusammenfassung	138
	Literatur.....	140

Kapitel IV

Leseselbstkonzept, Lesemotivation, Leseverhalten und Lesekompetenz 143

Martin Goy, Renate Valtin und Anke Hußmann

1	Einleitung.....	143
1.1	Leseselbstkonzept	144
1.2	Lesemotivation.....	145
1.3	Leseverhalten.....	146
1.4	Familiäre und individuelle Bedingungsfaktoren des Risikos schwacher Lesekompetenz	147
2	Ergebnisse.....	148
2.1	Leseselbstkonzept	149
2.2	Lesemotivation.....	150
2.3	Leseverhalten.....	152
2.4	Familiäre und individuelle Bedingungsfaktoren des Risikos schwacher Lesekompetenz	164
3	Zusammenfassung	169
	Literatur.....	171

Kapitel V**Geschlecht und Lesekompetenz177**

Nele McElvany, Ursula Kessels, Franziska Schwabe und Daniel Kasper

1	Einleitung.....	177
2	Forschungsstand.....	178
2.1	Geschlechterunterschiede in der Lesekompetenz.....	178
2.2	Mögliche Ursachen für Geschlechterunterschiede in der Lesekompetenz: Leseselbstkonzept, Lesemotivation und Leseverhalten	181
3	Ergebnisse	183
3.1	Geschlechterunterschiede in der Lesekompetenz in den Teilnehmerstaaten	183
3.2	Lesekompetenzen von Mädchen und Jungen nach Textsorten	183
3.3	Trends in Geschlechterunterschieden in den Leseleistungen zwischen 2001 und 2016.....	186
3.4	Lesekompetenzen von Mädchen und Jungen in Deutschland nach Kompetenzstufen	187
3.5	Leseselbstkonzept, Lesemotivation und Leseverhalten von Mädchen und Jungen in Deutschland	188
3.6	Bedeutung des Geschlechts für die Lesekompetenz unter Berücksichtigung von Leseselbstkonzept, Lesemotivation und Leseverhalten	189
4	Zusammenfassung	191
	Literatur.....	192

Kapitel VI**Soziale Herkunft und Lesekompetenzen von Schülerinnen
und Schülern195**

Anke Hußmann, Tobias C. Stubbe und Daniel Kasper

1	Einleitung.....	195
2	Soziale Herkunft und Bildungserfolg.....	196
3	Indikatoren der sozialen Herkunft.....	198
4	Zum Zusammenhang von sozialer Herkunft und Lesekompetenzen im internationalen Vergleich.....	201
5	Der Zusammenhang zwischen sozialer Herkunft und Lesekompetenzen in Deutschland	208
6	Zusammenfassung	214
	Literatur.....	214

Kapitel VII**Lesekompetenzen von Schülerinnen und Schülern
mit und ohne Migrationshintergrund219**

Heike Wendt und Knut Schwippert

1	Einleitung.....	219
2	Lesekompetenzen von Schülerinnen und Schülern im internationalen Vergleich nach Sprachgebrauch in der Familie	222
3	Lesekompetenzen von Schülerinnen und Schülern in Deutschland im Trend	224
4	Leseleistungen auf unterschiedlichen Kompetenzstufen im Trend	228
5	Leseselbstkonzept, Lesemotivation und Leseverhalten	230
6	Zusammenfassung	231
	Literatur.....	232

Kapitel VIII**Der Übergang von der Primar- in die Sekundarstufe.....235**

Tobias C. Stubbe, Wilfried Bos und Michael Schurig

1	Einleitung.....	235
2	Schullaufbahnpräferenzen von Lehrkräften und Eltern: Verteilung und Ausmaß der Übereinstimmung.....	239
3	Leistungsrelevante Merkmale der Schülerinnen und Schüler sowie Schullaufbahnpräferenzen der Lehrkräfte.....	241
4	Soziale Merkmale der Schülerinnen und Schüler und Schullaufbahnpräferenzen der Lehrkräfte und der Eltern.....	244
5	Zusammenfassung.....	248
	Literatur.....	248

Kapitel IX**Unterrichtsführung, Sozialklima und kognitive Aktivierung
im Deutschunterricht in vierten Klassen251**

Ruven Stahns, Svenja Rieser und Eva-Maria Lankes

1	Einleitung.....	251
2	Unterrichtsquantität und -qualität als Variablen in einem Angebots-Nutzungs-Modell des Unterrichts.....	252
3	Unterrichtszeit für Deutsch- beziehungsweise Sprachunterricht im internationalen Vergleich.....	253
4	Merkmale lernwirksamen Unterrichts: Befunde zur Unterrichtsqualität.....	256
5	Befunde zur Qualität des Deutschunterrichts in der Grundschule.....	259
6	Schüler- und Lehrkräfteangaben zum Deutschunterricht im Rahmen von IGLU 2016.....	262
7	Ergebnisse.....	264
7.1	Einschätzung des Deutschunterrichts durch die Lernenden und die Lehrkräfte.....	264
7.2	Die Wahrnehmung des Deutschunterrichts und der Zusammenhang mit Leseleistung und Lesemotivation.....	266
8	Zusammenfassung und Ausblick.....	273
	Literatur.....	275

Kapitel X**Förderung von Leseflüssigkeit und Leseverstehen im Leseunterricht.....279**Albert Bremerich-Vos, Ruven Stahns, Anke Hußmann und
Michael Schurig

1	Einleitung.....	279
2	Zum zeitlichen Umfang des Leseunterrichts.....	280
3	Auskünfte zu den Anteilen besonders schwacher und leistungsstarker Schülerinnen und Schüler.....	281
4	Förderung der Leseflüssigkeit.....	283
5	Förderung von Lesestrategien.....	289
6	Zusammenfassung und Bewertung.....	294
	Literatur.....	295

Kapitel XI

Bausteine adaptiven Leseunterrichts angesichts gewachsener Heterogenität297

Albert Bremerich-Vos, Heike Wendt und Anke Hußmann

1	Einleitung.....	297
2	Aspekte von Heterogenität an Grundschulen in Deutschland.....	299
3	Zur Sicht der Schulleitungen auf die Integration neu zugewanderter Schülerinnen und Schüler und Reaktionen in Form von äußerer und innerer Differenzierung.....	302
4	Vorschläge zum adaptiven Lehren beziehungsweise „kompetenten Umgang mit Heterogenität“ im Leseunterricht der Grundschule.....	304
4.1	Förderung des phonologischen Rekodierens.....	305
4.2	Arbeit am Wortschatz.....	306
4.3	Texte in Leichter Sprache.....	308
4.4	Prozessdiagnostik und der „Response-to-Intervention“-Ansatz.....	309
5	Fazit.....	311
	Literatur.....	312

Kapitel XII

Einordnung der IGLU-2016-Befunde in das europäische Rahmenkonzept für gute Leseförderung.....315

Renate Valtin

1	Einleitung.....	315
2	Schaffung einer schriftreichen Umgebung.....	317
2.1	Verbesserung der Lernbedingungen und der Qualität des Unterrichts.....	319
2.2	Steigerung von Teilhabe, Inklusion und Chancengerechtigkeit.....	323
3	Bilanz.....	324
	Literatur.....	325

Anhang A.....329

Anhang B.....336

Abbildungsverzeichnis338

Tabellenverzeichnis.....341

Vorwort

Um die Qualität von Unterricht und Schule nachhaltig zu gestalten, braucht es zuverlässige empirische Erkenntnisse. Im Prozess zunehmender Europäisierung und Globalisierung ist eine auf den internationalen Vergleich gestützte nationale Standortbestimmung zentral. Die Teilnahme Deutschlands an international vergleichenden Schulleistungsstudien wie der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) 2016 ermöglicht dies. Besonders die regelmäßige Teilnahme bietet einzigartige Möglichkeiten des langfristigen Bildungsmonitorings, so etwa hinsichtlich der Frage, inwiefern sich eingeleitete Reformen und Veränderungen im Schulwesen als zielführend oder wirksam erweisen.

Die Ergebnisse aus IGLU liefern wertvolle Informationen zu Schülerinnen und Schülern und ihren Leistungen an einer entscheidenden Schnittstelle des Schulsystems. Daraus gewonnene Erkenntnisse tragen zur empirischen bildungs- und erziehungswissenschaftlichen Grundlagenforschung bei und stellen entscheidende Diskursgrundlagen für Bildungsfragen und Reformen bereit. Ohne die im Rahmen von international vergleichenden Schulleistungsstudien geleistete Forschung wäre beispielsweise die Etablierung von Bildungsstandards, Vergleichsarbeiten oder neuen Kerncurricula wohl nicht möglich gewesen.

Der vorliegende Band stellt Ergebnisse zu IGLU 2016 vor. IGLU ist eine international vergleichende Schulleistungsstudie, die seit 2001 alle fünf Jahre durchgeführt wird. Im Jahr 2016 beteiligte sich Deutschland zum vierten Mal an der Studie. Im Zentrum steht der internationale Vergleich der Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe. Lesen gilt als Kernkompetenz für soziale und kulturelle Teilhabe. Das Lesen von Texten eröffnet Welten. Schülerinnen und Schüler am Ende der Grundschulzeit in Deutschland, so haben es die Befunde vergangener Erhebungszyklen von IGLU wiederholt belegt, weisen unterschiedliche Kompetenzen im Lesen auf; es gibt besonders starke Leserinnen und Leser, aber ebenso einen beachtlichen Anteil an Schülerinnen und Schülern, die nur über ein rudimentäres Leseverständnis verfügen. Die Grundschule hat den Auftrag, Schülerinnen und Schüler mit Werkzeugen und Strategien auszustatten, damit sie gute Leserinnen und Leser werden, damit sie sich Welten aus Texten erschließen, Phantasien entfalten und am gesellschaftlichen Leben aktiv teilnehmen können. Diesen Auftrag hat sie jedoch vor dem Hintergrund sich stetig verändernder Strukturen (Einführung von

Reformen, Curricula, Standards usw.) und einer heterogenen Schülerschaft zu erfüllen. IGLU ist Teil des vom *Bundesministerium für Bildung und Forschung* (BMBF) und von der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) beschlossenen Bildungsmonitorings, um genau dies systematisch und im internationalen Vergleich in den Blick zu nehmen. Dadurch, dass Deutschland sich zum vierten Mal an IGLU beteiligt hat, können erstmalig Veränderungen der Leistungsstände über einen Zeitraum von 15 Jahren berichtet werden. 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, sozialer Herkunft, Migration und Gestaltungsmerkmalen des Unterrichts sowie in Hinblick auf den Übergang von der Primar- in die Sekundarstufe I analysiert. Mit vertiefenden Analysen für Deutschland 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.

Weltweit waren an IGLU 2016 57 Staaten und Regionen mit insgesamt mehr als 312 000 Schülerinnen und Schülern, rund 300 000 Eltern und über 15 000 Lehrerinnen und Lehrern an ungefähr 11 000 Schulen beteiligt. In Deutschland liegen Daten für jeweils rund 4 000 Schülerinnen und Schüler, 3 000 Eltern, 200 Lehrpersonen und 190 Schulleitungen vor. Allen diesen Personen gilt ein besonderes Dankeschön für ihre hohe Bereitschaft, die Leistungstests zu bearbeiten beziehungsweise die Fragebögen auszufüllen und auf diese Weise Informationen zur Verfügung zu stellen. Ohne ihre Beteiligung wären weder die Durchführung noch die vorliegende Berichterstattung möglich gewesen.

Zu danken ist auch den zahlreichen Personen, die zum Gelingen von IGLU 2016 und zur Erstellung dieses Berichts beigetragen haben. Neben den wissenschaftlichen Mitarbeiterinnen und Mitarbeitern und studentischen Hilfskräften im IGLU-Team (siehe Kapitel 2 in diesem Band), den Autorinnen und Autoren und Ansprechpartnerinnen und Ansprechpartnern an der *IEA Hamburg* und am *Boston College* ebenso wie dem Team im Waxmann Verlag dankt das wissenschaftliche Konsortium folgenden Personen für die Unterstützung in den Wochen und Monaten vor Erscheinen dieses Bandes herzlich: Nina Büttner, Jutta Dämmer, Manuela Endberg, Dr. Michael Kandera, Dr. Ramona Lorenz, Melina Marnitz, Nikole Rodrigues, Bettina Röhrich, Alexandra Selent, Dr. Rolf Strietholt, Melek Vural und Lisa Wolf. Ein herzlicher Dank gilt ebenso allen Weggefährtinnen und Weggefährten der letzten 15 Jahre.

Professor Dr. Wilfried Bos
Wissenschaftlicher Leiter von IGLU 2016

Kapitel I

IGLU 2016:

Wichtige Ergebnisse im Überblick

Wilfried Bos, Renate Valtin, Anke Hußmann, Heike Wendt
und Martin Goy

1 Einleitung

Deutschland beteiligt sich seit 2001 regelmäßig an der *Internationalen Grundschul-Lese-Untersuchung* (IGLU), die international den Namen *Progress in International Reading Literacy Study* (PIRLS) trägt. Initiiert wurde diese Studie von der *International Association for the Evaluation of Educational Achievement* (IEA), einem unabhängigen, internationalen Zusammenschluss von Wissenschaftlerinnen und Wissenschaftlern, Forschungseinrichtungen und Regierungsstellen.

IGLU wird im Abstand von fünf Jahren durchgeführt und liefert nicht nur Daten zum Leseverständnis, zur Motivation und zum Leseverhalten von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe, sondern auch Informationen zu Faktoren, die die Lesekompetenz begünstigen, wie familiäre, schulische, curriculare und unterrichtliche Merkmale. Da auch zahlreiche andere Staaten und Regionen an IGLU teilnehmen, ist es möglich, die Ergebnisse im internationalen Vergleich zu verorten. Ferner ist es möglich, belastbare Aussagen zu den Veränderungen zu treffen, die in den vergangenen 15 Jahren in Deutschland stattgefunden haben.

Ein zentrales Anliegen von IGLU/PIRLS ist es, langfristige Entwicklungen in den teilnehmenden Bildungssystemen zu dokumentieren. Die Teilnahme Deutschlands erfolgte auf Beschluss der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) und einer Vereinbarung zwischen der KMK und dem *Bundesministerium für Bildung und Forschung* (BMBF).

An IGLU 2016 beteiligten sich im internationalen Vergleich weltweit 47 Staaten und Regionen als reguläre Teilnehmer und weitere 10 Benchmark-Teilnehmer mit gesonderten Teilnahmebedingungen. In diesem Berichtsband werden primär die Ergebnisse von zentralen Vergleichsstaaten berichtet. Da in IGLU

standardisierte schriftliche Befragungen der teilnehmenden Schülerinnen und Schüler, ihrer Eltern, der unterrichtenden Lehrkräfte sowie der Schulleitungen vorgenommen wurden, liegt ein umfangreicher und gehaltvoller Datensatz vor.

2 Zentrale Ergebnisse

2.1 Lesekompetenzen im internationalen Vergleich

Der Erfassung der Lesekompetenz von Schülerinnen und Schülern in IGLU liegt ein Modell zugrunde, das folgende drei Bereiche umfasst:

- Leseintentionen beziehungsweise das Lesen von literarischen und von Sachtexten,
- Verstehensprozesse, die beim Verständnis und der Verarbeitung der Informationen eines Textes eine Rolle spielen,
- Leseselbstkonzept, Lesemotivation und Leseverhalten.

Der in IGLU 2016 eingesetzte Lesetest erfasst die beiden erstgenannten Bereiche. Das Leseselbstkonzept, die Lesemotivation und das Leseverhalten der Kinder werden mit Hilfe eines Schülerfragebogens erhoben.

Trends in mittleren Leseleistungen im internationalen Vergleich

Mit einem Leistungsmittelwert von 537 Punkten ergibt sich für Deutschland ein Rangplatz im Mittelfeld, der Mittelwert von Deutschland liegt etwa in der Höhe der Mittelwerte aus den Vergleichsgruppen EU (540) und der OECD (540), und Deutschland liegt gleichauf mit Österreich, der Slowakei, Slowenien und Tschechien. Signifikant bessere Leistungen weist aber die große Mehrheit der EU-Staaten auf, nämlich Irland (567), Finnland (566), Polen (565), Nordirland (565), England (559), Lettland (558), Schweden (555), Ungarn (554), Bulgarien (552), Litauen (548), Italien (548), Dänemark (547) und die Niederlande (545). An der Spitze liegen die Russische Föderation (581) und Singapur (576). Ferner sind auch die Schülerinnen und Schüler in den Vereinigten Staaten von Amerika (549) signifikant erfolgreicher. Da eine Punktdifferenz von 40 annähernd dem Leistungszuwachs in einem Lernjahr entspricht (Hornberg, Bos, Buddeberg, Potthoff & Stubbe, 2007), ist der Abstand Deutschlands zu den Spitzenländern beträchtlich.

2001 erzielten Viertklässlerinnen und Viertklässler in Deutschland auf der Gesamtskala Lesen 539 Punkte, im Jahr 2016 537 Punkte. Die Differenz von 2 Punkten ist nicht signifikant. Dieser Befund bleibt, anders als in TIMSS (Wendt et al., 2016) in Bezug auf mathematische und naturwissenschaftliche Kompetenzen beschrieben, auch bestehen, wenn man die Veränderungen in der Zusammensetzung der Schülerschaft berücksichtigt. Im Jahr 2006 gab es zwar ein „Zwischenhoch“ (548), im Jahr 2011 aber bereits nicht mehr (541). In einer Reihe anderer Staaten, auch in der EU, ist es jedoch gelungen, die Leistungen in den vergangenen 15 Jahren zu verbessern. Deshalb hat sich die relative Position Deutschlands erheblich verschlechtert. 2001 lasen Kinder in vier Staaten, die 2016 zur EU zählten (Schweden, Niederlande, England, Bulgarien), signifikant besser, 2016 sind es Schülerinnen und Schüler in 13 EU-Staaten.

Leistungsheterogenität im internationalen Vergleich

Als ein Kriterium für gute Qualität von Bildungssystemen gilt, wenn ein hohes Niveau bei gleichzeitig geringer Streuung der Leistungen erreicht werden kann und damit die Schere zwischen Kindern mit guter und schwacher Leistung vergleichsweise gering ist. In IGLU 2016 ist die Streubreite der Leistungen in Deutschland auffällig hoch. Die Differenz zwischen dem 5. und dem 95. Perzentil, also den Leistungen, die die 5 Prozent der schwächsten Kinder maximal und die 5 Prozent der leistungsstärksten Kinder mindestens erreichen, beträgt 257 Punkte und ist EU-weit nur in Malta bedeutsam größer. Ein ähnliches Bild ergibt sich bei Betrachtung der Standardabweichung. In 17 der 24 teilnehmenden EU-Staaten ist sie geringer.

Während im Verlauf der letzten 15 Jahre in nur wenigen Staaten und Regionen die Streubreite der Leistungen zugenommen hat, ist für Deutschland eine Zunahme zu verzeichnen. Die Standardabweichung vergrößert sich von 66 auf 78 Punkte, im 5. Perzentil wurden 2001 maximal 419, 2016 jedoch nur 395 Punkte erreicht. Auch in Bezug auf das 95. Perzentil sind deutliche Veränderungen zu verzeichnen, diesmal in eine positive Richtung: 2001 waren es minimal 640 Punkte, im Jahr 2016 jedoch 652 Punkte. Die gewachsene Heterogenität resultiert somit sowohl aus geringeren Leistungen der schwächeren als auch aus besseren Leistungen der stärksten Schülerinnen und Schüler.

Anteile von Kindern auf verschiedenen Lesekompetenzstufen

Die Leistungen der Schülerinnen und Schüler auf der Gesamtskala Lesen lassen sich auf fünf unterschiedlichen Kompetenzstufen beschreiben. Ein Leistungsniveau, das der höchsten Kompetenzstufe V entspricht, erreichen in Deutschland 11 Prozent der Schülerinnen und Schüler. Dies entspricht zwar den Werten der Vergleichsgruppen EU und OECD, relativiert sich aber angesichts des Sachverhalts, dass es einigen unserer europäischen Nachbarstaaten gelingt, den Anteil auf der höchsten Kompetenzstufe auf annähernd 20 Prozent oder mehr zu bringen: nämlich Bulgarien, England, Finnland, Irland, Nordirland und Polen.

Hinsichtlich der unteren Kompetenzstufen lässt sich für Deutschland konstatieren: Knapp 6 Prozent der Viertklässlerinnen und Viertklässler in Deutschland weisen allenfalls ein rudimentäres Leseverständnis (Stufe I) auf. In Europa ist dieser Anteil sehr schwacher Leserinnen und Leser nur in Frankreich, der Französischen Gemeinschaft in Belgien, der Slowakei und in Malta nominell größer. Kompetenzstufe III wird in Deutschland von circa 19 Prozent der Viertklässlerinnen und Viertklässler nicht erreicht. Für diese Gruppe ist zu erwarten, dass sie in der Sekundarstufe I mit erheblichen Schwierigkeiten beim Lernen in allen Fächern konfrontiert sein wird, wenn es nicht gelingt, sie maßgeblich zu fördern. Auch im internationalen Vergleich ist diese Quote sehr ungünstig. Von den meisten europäischen Staaten wird sie signifikant unterschritten, signifikant überboten nur von Frankreich, der Französischen Gemeinschaft in Belgien und von Malta.

Ein Vergleich der Ergebnisse der IGLU-Erhebungen zeigt, dass der Anteil sehr guter Leserinnen und Leser signifikant von knapp 9 auf 11 Prozent gestiegen ist. Aber angesichts des Befunds, dass in zahlreichen Ländern dieser Prozentsatz fast doppelt so hoch ist, ist ein derartiger leichter Zuwachs kein Grund zur Zufriedenheit, zumal in den letzten 15 Jahren gleichzeitig die Gruppe der sehr schwachen Viertklässlerinnen und Viertklässler von 3 auf fast 6 Prozent angewachsen ist und der Anteil derjenigen, die Lesekompetenzstufe III nicht erreichen, von 17 auf 19 Prozent gestiegen ist. Auch in anderen europäischen Staaten

sind Veränderungen zu verzeichnen, und zwar sowohl positive als auch negative. So nahm in Slowenien der Anteil der auf Stufe V Lesenden im Vergleich von 2001 und 2016 von 3 auf 11 Prozent zu, zugleich sank die Quote der auf Stufe I und II Lesenden von 33 auf 17 Prozent. In Frankreich hingegen ging der Prozentsatz der auf Stufe V Platzierten von 7 auf 4 zurück und der Anteil der schwachen Leserinnen und Leser auf den Stufen I und II stieg von 23 auf 28 Prozent.

Für die Ergebnisse von IGLU 2011 bescheinigten Bos, Tarelli, Bremerich-Vos und Schwippert (2012, S. 134) Deutschland „ein Leistungsniveau im oberen Drittel der Rangreihe und eine international gesehen relativ geringe Streuung der Leistungen von besonders guten und besonders schwachen Leserinnen und Lesern.“ Davon kann bei IGLU 2016 nicht mehr die Rede sein. Nun sind die Leistungen im zweiten Drittel angesiedelt, im EU-Vergleich schneiden sogar etwa drei Viertel der Teilnehmerstaaten besser ab. Die Leistungsbreite hat zugenommen, an der Spitze der Leistungsskala, aber auch an deren Ende. Dass fast ein Fünftel der Viertklässlerinnen und Viertklässler in Deutschland, mehr als in der weit überwiegenden Zahl der EU-Staaten, Kompetenzstufe III nicht erreicht, ist besorgniserregend.

Lesen von Sachtexten im Vergleich zu literarischen Texten

Die Schülerinnen und Schüler in Deutschland erreichen beim Lesen von Sachtexten 533 Punkte, ein Wert, der etwas unterhalb des Mittelwerts der Vergleichsgruppe EU (539) liegt. Beim Lesen literarischer Texte erreichen die Schülerinnen und Schüler in Deutschland einen mittleren Wert von 542 Punkten, das heißt, sie zeigen vergleichsweise bessere Leistungen im Lesen dieser Textsorte als bei Sachtexten. Dieser Wert entspricht dem der VG EU (542). Die Differenz der Lesekompetenzen bei unterschiedlichen Textsorten beträgt 9 Punkte und ist nicht unerheblich. In einer Reihe von Staaten ist ebenfalls eine derartige Diskrepanz zu beobachten, in Deutschland ist sie jedoch besonders ausgeprägt. Signifikant größer ist sie nur in den USA, in Chile und in der Französischen Gemeinschaft in Belgien. Ausgeglichenere Leistungen in beiden Textsorten erzielen die Kinder in Tschechien, Israel, Finnland, Norwegen (4. Jgst.), Australien und Bulgarien.

Im Trendvergleich ergeben sich bei allen vier IGLU-Erhebungen in Deutschland bessere Ergebnisse beim Lesen literarischer Texte. Für beide Textsorten zeigen sich im Vergleich zwischen 2001 und 2016 Veränderungen; im Lesen von Erzähltexten: 539 (2001), 551 (2006), 545 (2011), 542 (2016), und im Lesen von Sachtexten: 539 (2001), 546 (2006), 538 (2011), 533 (2016). Der Leistungsunterschied ist um 9,5 Punkte signifikant angewachsen, was nur noch in Taiwan und der Französischen Gemeinschaft in Belgien der Fall ist. Auffällig ist, dass die Staaten, in denen es eine Differenz zugunsten der Sachtexte gibt (Russische Föderation, Singapur, Hongkong, Taiwan und Finnland), zu den leistungsstärksten im Lesetest (Gesamtskala) gehören.

Da das Lesen von Sachtexten einen wesentlichen Bestandteil des Lernens in der Sekundarstufe bildet, ist das schwache Abschneiden eine curriculare Herausforderung.

Leistungen bei den textimmanenten und wissensbasierten Verstehensleistungen

Um das Leseverständnis zu erfassen, werden vier Verstehensprozesse unterschieden: (1) explizit angegebene Informationen lokalisieren, (2) einfache Schlussfolgerungen ziehen, (3) komplexe Schlussfolgerungen ziehen, interpretieren

und kombinieren sowie (4) Inhalt und Sprachgebrauch prüfen und bewerten. Die Verstehensprozesse werden zu den *textimmanenten* (1 und 2) und den *wissensbasierten Verstehensleistungen* (3 und 4) zusammengefasst. Es ist unmittelbar einsichtig, dass die letztgenannten Leistungen anspruchsvollere Tätigkeiten darstellen. Handelt es sich bei den textimmanenten Verstehensleistungen vorwiegend um das sorgfältige, gewissenhafte Lesen, geht es bei wissensbasierten Verstehensleistungen um die Aktivierung des eigenen Wissens sowie die reflektierte und kritische Auseinandersetzung mit dem Gelesenen.

Auf der Skala *textimmanente Verstehensleistungen* erreichen Schülerinnen und Schüler in Deutschland einen Wert von 546 Punkten und liegen damit leicht über dem Mittelwert der Vergleichsgruppe EU (542). Für die Skala *wissensbasierte Verstehensleistungen* ergibt sich hingegen für die Schülerinnen und Schüler in Deutschland nur eine mittlere Leistung von 530 Punkten, unterhalb des Mittelwerts der Vergleichsgruppe EU (539). Betrachtet man die Differenz der Testleistungen in diesen beiden Verstehensaspekten, so zeigt sich, dass sich nur in wenigen Staaten ausgeglichene Leistungsprofile ergeben, während sich in der großen Mehrzahl der Staaten die Ergebnisse auf den beiden Skalen signifikant unterscheiden. Mit einer Differenz von 16 Punkten zugunsten der textimmanenten Verstehensleistungen liegt Deutschland fast an der Spitze der Verteilung, gleichauf mit Österreich und Tschechien, nur Frankreich weist eine noch größere Diskrepanz auf. Unter den Staaten, bei denen bedeutend größere Differenzen zugunsten der wissensbasierten Verstehensprozesse zu verzeichnen sind, befinden sich überproportional viele, in denen die Schülerinnen und Schüler besonders leistungsfähig sind.

Zwischen 2001 und 2016 haben sich in Deutschland die Leistungen in diesen beiden Skalen nur wenig verändert. Der Wert für die wissensbasierten Verstehensleistungen betrug 2001 535 Punkte, 2016 sind es 530 Punkte, ein nicht signifikanter Unterschied. Signifikant wird jedoch die Differenz zwischen den beiden Skalen der Verstehensleistungen, ein Befund, der ebenfalls curriculare und unterrichtliche Relevanz hat.

2.2 Leseselbstkonzept, Lesemotivation und Leseverhalten

Leseselbstkonzept. Die Einschätzung der eigenen Fähigkeiten im Lesen ist bei Schülerinnen und Schülern in Deutschland positiv ausgeprägt. Fast drei Viertel der Schülerinnen und Schüler verfügen über ein hohes, nur 5 Prozent über ein niedriges Selbstkonzept. Der *IQB-Bildungstrend 2016* (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017) zeigt ähnliche Befunde für die Schülerinnen und Schüler der vierten Klassenstufe. Über 70 Prozent der Schülerinnen und Schüler schätzen demzufolge ihr fachbezogenes Selbstkonzept im Fach Deutsch als hoch ein (Schipolowski, Wittig, Weinrich & Böhme, 2017). Zwischen 2011 und 2016 zeigen sich keine signifikanten Änderungen im Mittelwert. Jedoch ist auf den drei unteren Lesekompetenzstufen der Anteil an Schülerinnen und Schülern mit hohem Selbstkonzept gesunken. Im Vergleich zu 2011 fallen die Selbsteinschätzungen 2016 somit etwas realistischer aus.

Einen internationalen Vergleich haben wir beim Leseselbstkonzept und auch bei der Lesemotivation nicht vorgenommen. Es kann vermutet werden, dass sich in diesen Skalen, die auf Selbstberichten beruhen, kulturelle Unterschiede in den Antworttendenzen oder unterschiedliche Bezugsnormen in den Teilnehmerstaaten und -regionen spiegeln, die internationale Vergleiche erschweren.

Lesemotivation. Insgesamt gesehen geben die Schülerinnen und Schüler in Deutschland im Mittel an, hoch lesemotiviert zu sein. Beim Vergleich der Ergebnisse der Jahre 2001 und 2016 ist allerdings eine ungünstige Entwicklung festzustellen: Der Anteil der Schülerinnen und Schüler mit hoher Lesemotivation verringert sich signifikant um rund 5 Prozentpunkte, und gleichzeitig erhöht sich der Anteil der Schülerinnen und Schüler mit niedriger Lesemotivation um rund 4 Prozentpunkte. Wenn man die Leseleistungsgruppen betrachtet, so zeigt sich vor allem für die leseschwachen Kinder (Kompetenzstufe I und II) eine missliche Entwicklung: In dieser Leistungsgruppe erhöht sich der Anteil an niedrig lesemotivierten Kindern um rund 11 Prozentpunkte, während der Anteil an hoch lesemotivierten um rund 15 Prozentpunkte sinkt.

Die Erhebungen in den Jahren 2006 und 2011 hatten gegenüber der ersten IGLU-Erhebung im Jahr 2001 leicht positivere Werte in der Lesemotivation gezeigt. Diese Entwicklung hat sich jedoch nicht fortgesetzt, im Gegenteil: 2016 sind die Werte noch ungünstiger als 2001, und fast jedes sechste Kind in der vierten Grundschulklasse hat eine niedrige Lesemotivation.

Leseverhalten. Neben motivationalen Merkmalen wurde in allen IGLU-Erhebungen auch das Leseverhalten der Schülerinnen und Schüler in den Blick genommen. Betrachtet werden hier drei Aspekte: das Lesen zum Vergnügen außerhalb der Schule, die Dauer des täglichen Lesens in der Freizeit sowie das Entleihen von Büchern aus einer Bibliothek.

Lesen zum Vergnügen außerhalb der Schule. In Deutschland geben 17 Prozent der Schülerinnen und Schüler an, nie oder fast nie außerhalb der Schule zum Vergnügen zu lesen. Diese Zahl entspricht in etwa den Ergebnissen der *KIM-Studie* (Kindheit, Internet, Medien) 2016, in der 18 Prozent der befragten Kinder im Alter von 6 bis 13 Jahren angaben, in ihrer Freizeit nicht zu lesen (Feierabend, Plankenhorn & Rathgeb, 2017). Bei IGLU hat sich dieser Anteil von 2001 bis 2016 in der Gesamtstichprobe in Deutschland nicht signifikant verändert. Wie der internationale Vergleich zeigt, ist es jedoch in 12 Teilnehmerstaaten und -regionen gelungen, die Anzahl dieser ‚Nichtleser‘ signifikant zu verringern. Ferner zeigt sich, dass der Anteil an Schülerinnen und Schülern, die jeden Tag oder fast jeden Tag zu ihrem Vergnügen lesen, in Deutschland mit knapp 43 Prozent ebenfalls vergleichsweise hoch ist. Allerdings hat sich dieser Anteil zwischen 2001 und 2016 signifikant verringert. Eine derartige Verringerung trifft auch für eine ganze Reihe der europäischen Nachbarstaaten zu: Während sich in Deutschland der Anteil um 5 Punkte verringert, sind es in anderen Staaten sogar 10 und mehr Prozentpunkte.

Ähnlich wie beim Trend der Lesemotivation ist auch bei dem Merkmal „Lesen zum Vergnügen“ die ungünstigste Entwicklung bei der unteren Leseleistungsgruppe zu verzeichnen: Im Vergleich zur Gruppe der Leistungsstarken findet sich hier die größte Abnahme an Kindern, die jeden Tag oder fast jeden Tag zum Vergnügen lesen, und die größte Zunahme des Anteils derer, die dies nie oder fast nie tun.

Tägliche Lesedauer außerhalb der Schule. Aufgrund der Antworten auf die Frage nach der täglichen Lesedauer („Wie viel Zeit verbringst du außerhalb der Schule an einem normalen Schultag mit Lesen?“) werden zwei Gruppen unterschieden: ‚Wenigleser‘ sind Kinder, die angeben, täglich weniger als 30 Minuten zu lesen, ‚Vielleser‘ berichten, täglich 2 Stunden oder mehr in ihrer Freizeit zu lesen. In Deutschland gehören 40 Prozent zu den ‚Weniglesern‘ und 10 Prozent zu den ‚Viellesern‘. Im europäischen Vergleich sind dies günstige Zahlen: Die entsprechenden Werte sind 46 Prozent bei den ‚Weniglesern‘ und 7 Prozent

bei den ‚Viellesern‘. Wie in den meisten der an IGLU beteiligten Staaten und Regionen hat sich auch in Deutschland der Anteil der ‚Wenigleser‘ von 2011 (als dieses Maß erstmals im internationalen Schülerfragebogen erhoben wurde) auf 2016 signifikant erhöht (von 33 auf 40 %) und der Anteil der ‚Vielleser‘ hat sich im gleichen Zeitraum von 14 Prozent auf 10 Prozent verringert.

Bibliotheksnutzung. Die Antworten auf die Frage „Wie oft leihst du dir Bücher (einschließlich E-Books) aus der Schulbibliothek oder der Bibliothek in deinem Ort aus?“ verweisen auf ungünstige Werte für Deutschland im internationalen Vergleich. Rund 35 Prozent der Schülerinnen und Schülern leihen sich nie oder fast nie Bücher aus. In nur drei Staaten ist dieser Anteil noch höher (Chile, Slowakei und Tschechien). Im Mittel der Vergleichsgruppe EU sind knapp 20 Prozent. Auch der Anteil der Kinder in Deutschland, die mindestens wöchentlich Bücher entleihen (24.2 %), liegt deutlich unter dem EU-Durchschnitt (32.8 %). Problematisch ist vor allem der Befund, dass sich zwischen 2011 und 2016 in Deutschland der Anteil derjenigen, die nie oder fast nie Bücher aus der Bibliothek ausleihen, um 10 Prozentpunkte vergrößert hat, wobei sich im Vergleich der Leistungsgruppen etwas stärkere Zunahmen derjenigen, die nie oder fast nie Bücher entleihen, in der Gruppe der Leseschwachen finden.

Zwar ist in fast allen Staaten und Regionen dieser Anteil der Wenignutzer gestiegen, in Deutschland aber besonders deutlich. In fast allen Staaten und Regionen ist auch der Anteil an Kindern, die häufig beziehungsweise mindestens einmal pro Woche Bücher entleihen, von 2006 zu 2016 signifikant gesunken, in Deutschland um 5 Prozentpunkte. Signifikante Unterschiede zeigen sich allerdings nur bei den leseschwachen Schülerinnen und Schülern auf Kompetenzstufe I und II (um rund 10 %) und bei den Schülerinnen und Schülern auf Kompetenzstufe III (um rund 7 %). Dass sich in den letzten 10 Jahren gerade in der Gruppe der Leseschwachen die größten Veränderungen zeigen, ist durchaus bedenkenswert.

Internationale Trends. Insgesamt deuten die Daten auf einen ungünstigen internationalen Trend im Leseverhalten hin. In fast allen Staaten und Regionen ist der Anteil der Kinder, die fast täglich zum Vergnügen lesen, seit 2001 beziehungsweise 2006 beträchtlich gesunken, in einigen Staaten und Regionen sogar um 10 Prozentpunkte. Was die tägliche Lesedauer betrifft, so hat sich der Anteil derjenigen, die weniger als 30 Minuten täglich außerhalb der Schule lesen, in allen an IGLU beteiligten Staaten erhöht, und zwar durchweg signifikant, und die Quote derjenigen, die dies mehr als 2 Stunden täglich tun, hat etwas abgenommen. In fast allen Staaten und Regionen hat in den letzten 10 Jahren der Anteil der Kinder, die nie oder fast nie Bücher aus der schulischen oder der örtlichen Bibliothek entleihen, signifikant zugenommen, und der Anteil an Kindern, die häufig beziehungsweise mindestens einmal pro Woche Bücher entleihen, ist gesunken.

Die internationalen Daten deuten auf ein verändertes Medienverhalten von Kindern im Grundschulalter hin, wie es auch für Deutschland durch die KIM-Studie (Kindheit, Internet, Medien) zum Freizeit- und Medienverhalten von Kindern im Grundschulalter belegt wird (Feierabend et al., 2017). Im Vergleich zur Nutzung anderer Medien (wie Radio, Handy/Smartphone, Fernsehen und Musikhören) kommt das Lesen von Büchern deutlich seltener vor und wird nur von 5 Prozent der Kinder als Lieblingstätigkeit in der Freizeit genannt.

2.3 Chancengleichheit und soziale Disparitäten

In allen bei IGLU teilnehmenden Staaten und Regionen werden mehr oder wenige große Leistungsunterschiede zwischen verschiedenen Schülergruppen festgestellt. In diesem Buch werden Disparitäten in den Leseleistungen für Mädchen und Jungen, Kindern unterschiedlicher sozialer Herkunft sowie aus Familien mit und ohne Migrationshintergrund betrachtet.

Leistungen von Mädchen und Jungen

In Deutschland erreichen Mädchen (mit 543) leicht bessere Ergebnisse als Jungen (mit 532 Punkten) und haben einen Vorsprung von 11 Punkten auf der Gesamtskala Lesen. Diese Differenz unterscheidet sich nicht von den Unterschieden in den Vergleichsgruppen EU und OECD (je 13 Punkte). Der Leistungsvorsprung der Mädchen gegenüber den Jungen ist in den letzten Jahren fast unverändert geblieben. Ein vertiefender Blick zeigt, dass es insbesondere beim literarischen Lesen im Mittel deutliche Vorteile der Mädchen in den Leseleistungen gibt, während dies beim informierenden Lesen nur minimal der Fall ist. Mädchen haben auch im Leseselbstkonzept, der Lesemotivation und dem Leseverhalten positivere Werte als Jungen. Kontrolliert man diese Unterschiede, findet sich kein bedeutsamer Kompetenzunterschied im Lesen zwischen den Geschlechtern. Jungen und Mädchen mit gleichem Niveau im Leseselbstkonzept, in der Lesemotivation und im Leseverhalten verfügen demnach über gleich gute Lesekompetenzen. Die Frage stellt sich, wie es zu den ungünstigeren motivationalen Merkmalen der Jungen kommt. Ein Erklärungsansatz verweist auf ein geringeres schulisches Engagement von Jungen. Dieser Erklärungsansatz wird auch durch die Daten der IGLU-Erhebung von 2001 gestützt, in der eine Vielzahl motivationaler und emotionaler Merkmale von Schülerinnen und Schülern des vierten Schuljahrs erhoben wurde. Die Befunde von IGLU verweisen darauf, dass eine Leseförderung der Jungen bei ihrer Lesemotivation ansetzen muss. Das differentielle Unterschiedsmuster in Abhängigkeit von den Textsorten verdeutlicht, dass eine stärkere Förderung der Jungen im Bereich des literarischen Lesens den mittleren Unterschied in den Leseleistungen im Vergleich zu den Mädchen reduzieren und damit auch den Gesamtleistungswert der Schülerschaft am Ende der vierten Klassenstufe verbessern könnte.

Soziale Disparitäten in den Leseleistungen

Soziale Disparitäten im Bildungssystem haben in den vergangenen Jahren sowohl in der wissenschaftlichen als auch in der öffentlichen Diskussion viel Aufmerksamkeit erfahren, widersprechen sie doch dem Anspruch auf Chancengleichheit in einer Gesellschaft, wie er in Artikel 3, Absatz 3 des deutschen Grundgesetzes verankert ist. In IGLU 2016 zeigen sich – wie auch in allen früheren IGLU-Erhebungen – für alle Teilnehmerstaaten signifikante Unterschiede in der Lesekompetenz zwischen den Schülerinnen und Schülern aus günstigen und ungünstigen sozialen Lagen. Allerdings muss auch betont werden, dass sich das Ausmaß der sozialen Disparitäten zwischen den Teilnehmerstaaten zum Teil erheblich unterscheidet. Dies kann entweder bedeuten, dass die Unterschiede zwischen den oberen und den unteren sozialen Lagen in einigen Staaten ausgeprägter sind als in anderen oder dass es einigen Staaten besser gelingt als anderen, die primären Herkunftseffekte der Schülerinnen und Schüler im Laufe der ersten vier Schuljahre zu verringern.

Zur Messung der sozialen Disparitäten im internationalen Vergleich wurden drei Indikatoren gewählt: die Anzahl der Bücher im Haushalt, Berufsstatus der Eltern und Bildungsniveau der Eltern. In Bezug auf den Indikator *Bücher im Haushalt* wurde der Leistungsvorsprung im Leseverständnis von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern berechnet. Mit einer Differenz von 54 Punkten liegt Deutschland hinter Ungarn (58) gleichauf mit Bulgarien und der Slowakei an der Spitze der Verteilung; mit einer Differenz von knapp 30 Punkten befinden sich Finnland (29), die Niederlande (29) und Lettland (27) am unteren Ende der Verteilung. Besorgniserregend ist das Ergebnis der Trendanalyse: Deutschland gehört neben der Slowakei, Ungarn und Slowenien zu den vier Staaten, in denen seit 2001 eine signifikante Vergrößerung der sozialen Disparitäten in diesem Indikator zu beobachten ist.

Eine noch größere soziale Disparität gibt es in Deutschland in Bezug auf den Berufsstatus der Eltern. Der Leistungsvorsprung im Leseverständnis von Kindern aus Familien der dritten Berufsgruppe (Akademiker, Techniker und Führungskräfte) vor denen der ersten Berufsgruppe (manuelle Tätigkeiten) beträgt 72 Punkte und damit ungefähr eineinhalb Lernjahre. Auch hier liegt Deutschland signifikant über dem Mittelwert der Vergleichsgruppe EU (55 Punkte). Seit 2006 ist der Leistungsvorsprung in diesem Indikator um 12 Punkte, allerdings nicht signifikant, größer geworden. In Bezug auf das Bildungsniveau zeigt sich ein Unterschied von 48 Punkten im Lesetest zwischen Kindern aus Familien, in denen mindestens ein Elternteil einen tertiären Bildungsabschluss erreicht hat, und Kindern aus Familien, in denen dies nicht der Fall ist. Die Größe dieses Unterschieds hat sich seit 2001 kaum verändert.

Weitere Analysen mit Indikatoren, die nur für den deutschen Datensatz vorliegen, belegen, dass sich die sozialen Disparitäten in Deutschland in den letzten 15 Jahren nicht statistisch signifikant verändert haben. Die Leistungsunterschiede im Lesetest zwischen Kindern aus günstigen und ungünstigen sozialen Lagen liegen bei etwa einem Lernjahr. Allerdings ist die Richtung der Veränderungen fast durchweg negativ, beispielsweise liegt 2016 die Lesekompetenz von Kindern aus nicht armutsgefährdeten Elternhäusern 52 Punkte über der Lesekompetenz von armutsgefährdeten Kindern. Im Jahr 2011 (als erstmalig dieser Indikator in IGLU erfasst wurde) betrug dieser Wert 40 Punkte. Die Differenz von 12 Punkten ist statistisch allerdings nicht signifikant.

Deutschland ist es in den vergangenen 15 Jahren im Bereich der Grundschule nicht gelungen, den Anspruch auf Chancengleichheit im Bildungssystem zu realisieren. Diese Aussage bezieht sich allerdings auf Gesamtdeutschland. Sowohl zwischen als auch innerhalb der Länder bestehen deutliche Unterschiede hinsichtlich der Chancen auf schulische Bildungsteilhabe und Bildungserfolg (u. a. Berkemeyer, Bos, Hermstein, Abendroth & Semper, 2017; Haag, Kocaj, Jansen & Kuhl, 2017).

Migrationsbedingte Disparitäten in den Leseleistungen

Für den schulischen Erfolg ebenso wie für das Leseverständnis ist die Beherrschung der Amts- oder Verkehrssprache unerlässlich. Mit einem Anteil von knapp 17 Prozent der Kinder, die angeben, nie oder nur manchmal zu Hause Deutsch zu sprechen, liegt Deutschland nahe am Mittelwert der Vergleichsgruppe EU (14.5%). Viertklässlerinnen und Viertklässler, die angeben, immer oder fast immer zu Hause Deutsch zu sprechen, erreichen einen Leistungsmittelwert von 549 Punkten. Ihre Mitschülerinnen und Mitschüler, die angeben, nie oder nur manchmal zu Hause Deutsch zu sprechen, erreichen lediglich

509 Punkte. Die Leistungsdifferenz in Deutschland beträgt 40 Punkte und somit fast eine halbe Standardabweichung. Die Befunde aus IGLU 2016 bestätigen, dass Schülerinnen und Schüler, die zu Hause nicht die Testsprache sprechen, in fast allen Teilnehmerstaaten und -regionen niedrigere Leseleistungen aufweisen als ihre Mitschülerinnen und Mitschüler, die die zu Hause die Testsprache sprechen. Allerdings zeigen sich erhebliche Variationen in der Größe des Leistungsunterschieds. Einige Teilnehmerstaaten und -regionen weisen im Vergleich zu Deutschland deutlich niedrigere migrationsbezogene Disparitäten auf. Es bleibt zu prüfen, ob dies durch in den Staaten realisierte Maßnahmen in den Bildungssystemen oder durch die jeweilige Zusammensetzung der Migrantengruppen zu erklären ist.

Wegen des hohen Anteils fehlender Werte (etwa 20% der Eltern haben den Fragebogen nicht ausgefüllt) ist es schwierig, verlässliche Daten über den Anteil von Kindern mit Migrationshintergrund in der IGLU-Stichprobe zu gewinnen. Laut Angaben der Eltern, die den Fragebogen ausgefüllt haben, liegt der Anteil von Schülerinnen und Schülern mit Migrationshintergrund (mindestens ein Elternteil im Ausland geboren) bei rund 32 Prozent und ist vor dem Hintergrund verfügbarer Bildungsstatistiken plausibel (Autorengruppe Bildungsberichterstattung, 2016). Ein vergleichbarer Anteil von 34 Prozent der Viertklässlerinnen und Viertklässler mit Zuwanderungshintergrund wird auch im *IQB-Bildungstrend 2016* (Stanat et al., 2017) berichtet, wobei es auch dort fehlende Werte im Umfang von 20 bis 30 Prozent bei den Elternangaben gab. Laut IQB-Bericht ist der Anteil von Kindern mit Migrationshintergrund gegenüber 2011 um mehr als ein Drittel gestiegen. Auch die IGLU-Erhebungen zeigen 2016 signifikant höhere Anteile als 2001.

Im Vergleich zu ihren Mitschülerinnen und Mitschülern, deren Eltern in Deutschland geboren sind, erzielten Schülerinnen und Schüler mit Migrationshintergrund Leseleistungen, die 24 (ein Elternteil im Ausland geboren) beziehungsweise 49 (beide Elternteile im Ausland geboren) Leistungspunkte geringer ausfallen, ein Unterschied, der einem Lernzuwachs von einem halben beziehungsweise einem ganzen Schuljahr entspricht. Im Vergleich zu 2001 lassen sich 2016 für Kinder, deren Eltern beide im Ausland geboren wurden, signifikant höhere Leseleistungen beobachten. Weniger deutliche, aber dennoch signifikante Zuwächse bestehen auch für ihre Mitschülerinnen und Mitschüler, deren Eltern beide in Deutschland geboren wurden.

Trotz vielfältiger Anstrengungen ist es am Ende der vierjährigen Grundschulzeit in Deutschland nicht gelungen, das bildungspolitische Ziel der Verringerung von zuwanderungsbezogenen Disparitäten zu realisieren. Berücksichtigt werden muss allerdings, dass sich die Zusammensetzung der Schülerinnen und Schüler aus Familien mit Migrationshintergrund in 15 Jahren verändert hat. Die Heterogenität ist größer geworden. Die Förderung von Kindern entsprechend ihren unterschiedlichen individuellen, sprachlichen sowie soziokulturellen Voraussetzungen muss stärker als bisher in den Fokus bildungspolitischer sowie schulischer Bemühungen rücken.

Übergang in eine weiterführende Schule

Eine zentrale Gelenkstelle für das Entstehen von Bildungsbenachteiligung in Deutschland ist der Übergang in die Sekundarstufe. In IGLU wurden die Lehrkräfte danach gefragt, welchen Schulabschluss sie für das jeweilige Kind erwarten, und Eltern, welche Schule ihr Kind nach den Sommerferien besuchen wird. Wir interpretieren die Antworten als Hinweis auf eine Schullaufbahnpräferenz.

Die in diesem Beitrag vorgelegten Trendanalysen zum Übergang von der Grundschule in die Sekundarstufe bei IGLU 2001, 2006, 2011 und 2016 zeigen ein insgesamt relativ konstantes Bild. Veränderungen ergeben sich insbesondere durch die veränderte Schulstruktur. Jedes vierte Kind wechselt den Elternangaben zufolge 2016 auf eine Schule mit mehreren Bildungsgängen – damit ist dieser Anteil mehr als dreimal so hoch wie 15 Jahre zuvor. Knapp die Hälfte der Viertklässlerinnen und Viertklässler wechselt auf ein Gymnasium. Seit 2006 ist dieser Wert relativ konstant. Hinsichtlich der Schullaufbahnpräferenzen der Lehrkräfte zeigen sich eine stetige Abnahme der Nennungen für die Hauptschule und ein entsprechender Anstieg für das Gymnasium.

Wie schon in den vorangegangenen Erhebungen hängen die Schullaufbahnpräferenzen der Lehrkräfte eng mit den Noten der Schülerinnen und Schüler zusammen. Da jedoch auch in IGLU 2016 nur ein relativ geringer Zusammenhang zwischen den Ergebnissen des Lesetests und der Deutschnote besteht, bedeutet dies, dass ebenfalls nur ein relativ geringer Zusammenhang zwischen den Ergebnissen des IGLU-Lesetests und den von den Lehrkräften angegebenen Präferenzen für eine Sekundarschulform besteht.

Bedenklich ist der konstant hohe Zusammenhang zwischen der sozialen Herkunft der Schülerinnen und Schüler und den Schullaufbahnpräferenzen. Auch unter Kontrolle der Lesekompetenz und der kognitiven Fähigkeiten haben Kinder aus bildungsnahen Elternhäusern eine deutlich höhere Chance auf eine Gymnasialpräferenz als Kinder aus bildungsfernen Familien. Im Vergleich der IGLU-Erhebungen von 2001 bis 2016 lässt sich im Zeitverlauf sogar eine Zunahme dieser Chancen feststellen.

Als ein weiterer Kennwert wurde ein „kritischer Wert“ im Lesetest verwendet: Erreicht ein Kind diesen Wert, hat es gute Chancen, als geeignet für ein Gymnasium angesehen zu werden. Insgesamt wird eine Gymnasialpräferenz der Lehrkräfte ab einer Lesekompetenz von 562 Punkten (25 Punkte oberhalb des deutschen Mittelwerts) hinreichend wahrscheinlich. Zwischen den Berufsklassen unterscheiden sich diese Werte jedoch erheblich: Während Kinder mit Eltern der oberen Berufsklassen (Akademiker, Techniker und Führungskräfte) bereits mit einem Wert von 518 (19 Punkte *unterhalb* des deutschen Mittelwertes) gute Chancen auf eine Gymnasialpräferenz ihrer Lehrkräfte haben, liegt der kritische Wert bei Kindern von un- und angelernten Arbeiten bei 620, also 83 Punkte *oberhalb* des deutschen Mittelwertes. Die Differenz zwischen diesen beiden Berufsklassen beträgt etwa 100 Punkte, was gut zwei Lernjahren entsprechen dürfte (Hornberg et al., 2007).

Im Vergleich mit früheren Erhebungen ist der kritische Wert für eine Gymnasialempfehlung gesunken: von 581 beziehungsweise 580 in den Jahren 2001 und 2006 auf 562 bei IGLU 2016. Allerdings hat dies nicht zu einer größeren Chancengerechtigkeit beigetragen. Im Gegenteil: Lag der kritische Wert für die oberste Berufsgruppe 2001 bei 551, sank er 2006 auf 537 und 2016 auf 518. Im gleichen Zeitraum stieg der kritische Wert bei Kindern von un- und angelernten Arbeitern von 601 (2001) auf 614 (2006) und schließlich auf 620 (2016). Diese Werte verweisen darauf, dass beim Übergang in die Sekundarschulen die Chancengerechtigkeit nicht gewahrt ist.

2.4 Lehr- und Lernbedingungen

Mit den Fragebögen für Deutschlehrkräfte und Schulleitungen werden in IGLU zahlreiche schulische, curriculare und unterrichtliche Merkmale erfasst. In diesem Band werden ausgewählte Merkmale zum Deutsch- und Leseunterricht berichtet.

Unterrichtszeit und Unterrichtsqualität im Deutschunterricht

Im Durchschnitt werden in Deutschland laut Auskünften der Lehrkräfte in der Woche 373 Minuten der Unterrichtszeit für Deutsch- beziehungsweise Sprachunterricht verwendet. Dieser Wert liegt signifikant unter dem Mittelwert der Vergleichsgruppe OECD (406 Minuten) und beträchtlich, wenngleich nicht statistisch signifikant, unter dem der Vergleichsgruppe EU (384 Minuten) sowie unter dem internationalen Mittelwert der Teilnehmerstaaten von IGLU 2016 (388 Minuten). Die Quantität der Unterrichtszeit allein besagt jedoch nichts über die Qualität des Unterrichts und die effektiv zum Lernen genutzte Zeit. Aus den Schüler- und Lehrerfragebögen liegen dazu einige Hinweise vor. Gefragt wurde nach verschiedenen Merkmalen des Deutschunterrichts, die in der empirischen Schulforschung als lernwirksam angesehen werden: Klassenführung, Strukturiertheit des Unterrichts und kognitive Aktivierung. Die Einschätzungen von Schülerinnen und Schülern und Lehrkräften liegen jedoch deutlich auseinander. Betrachtet man den Anteil der Schülerinnen und Schüler, die den Unterricht im Hinblick auf das betreffende Merkmal als gering ausgeprägt ansehen, so ergeben sich folgende Zahlen: Klassenführung (31 %), Strukturiertheit des Unterrichts (45 %), kognitive Aktivierung (16 %). Lehrkräfte schätzen ihren Unterricht als sehr viel positiver ein. Bei diesen drei Merkmalen bescheinigen sich mehr als drei Viertel der Befragten eine hohe Unterrichtsqualität. Auch bei der Wahrnehmung des Sozialklimas unterscheiden sich die Befragten: 93,2 Prozent der Lehrkräfte schätzen das Sozialklima als positiv ein, bei den Schülerinnen und Schülern sind es 72,8 Prozent.

Unterrichtszeit und Unterrichtsqualität im Leseunterricht

Nach eigenen Angaben verwenden die befragten Lehrkräfte den Durchschnitt 8 Stunden für Unterricht im Lesen, Schreiben, Sprechen, den Umgang mit Literatur und in weiteren sprachbezogenen Bereichen. Hochgerechnet auf das Schuljahr ist dieser Wert vergleichbar mit dem internationalen Mittelwert (vgl. Mullis, Martin, Foy & Hooper, 2017). Speziell für Leseunterricht und/oder Leseaktivitäten der Schülerinnen und Schüler werden ihren Auskünften zufolge im Mittel drei Schulstunden (136 Minuten) verwendet, wobei die Streuung beträchtlich ist (126 Minuten). Umgerechnet auf das Schuljahr ergeben sich, absolut betrachtet, etwa 90 Stunden für expliziten Leseunterricht, auch über Fächergrenzen hinweg. Deutschland liegt damit weit unter dem internationalen Mittelwert mit knapp 160 Stunden (vgl. ebd.). Berücksichtigt man auch noch das gesamte jährliche Unterrichtsvolumen, das in Deutschland bei etwa 840 Zeitstunden liegt, dann ergibt sich speziell für den Leseunterricht jahresbezogen eine Quote von 11 Prozent. Nur in sechs Teilnehmerstaaten und -regionen ist diese Quote geringer, der internationale Mittelwert beträgt 18 Prozent. Diese Daten zeigen, dass der Anteil des Leseunterrichts am gesamten Unterricht eines vierten Schuljahrs in Deutschland im internationalen Vergleich ausgesprochen gering ist.

In Deutschland zeigt sich eine besonders hohe Differenz zwischen den mit dem IGLU-Lesetest gemessenen Verstehensleistungen. Die mittlere Leistung

bei den wissensbasierten Verstehensleistungen fällt sehr viel niedriger aus als die textimmanente Verstehensleistung. Diese Differenz von 16 Punkten kann als bedeutsam für die Curriculum-Planung angesehen werden und verweist darauf, dass Aufgaben, die dem tieferen Verständnis eines Textes dienen und bei denen ein Bezug zum eigenen Weltwissen hergestellt werden muss, stärker betont werden sollten. Betrachtet man die Einschätzungen der Lehrkräfte dazu, wie oft sie ihren Schülerinnen und Schülern Gelegenheit geben, gezielt an der Entwicklung von Lesestrategien zu arbeiten (z. B. das Gelesene mit ihren eigenen Erfahrungen vergleichen), und wie oft sie gezielt verständnisrelevantes Wissen oder Lesestrategien vermitteln (z. B. den Schülern verschiedene Lesestrategien erklären oder vormachen), zeigen sich im europäischen Vergleich für Deutschland eher ungünstige Werte. Noch stärker unter dem internationalen Mittelwert liegt Deutschland bei Aufgaben, die sich auf Stil und Struktur von Texten beziehen, sowie bei Aufgaben, in denen nach Perspektive oder Absicht des Textes beziehungsweise des Autors gefragt wird.

Lehrkräfte in Deutschland wurden zur Vermittlung von Lesestrategien befragt, wie sie in den länderübergreifenden Bildungsstandards und in den Curricula aller Länder der Bundesrepublik Deutschlands verankert sind. Die Befragung ergab, dass nur eine Minderheit der Schülerinnen und Schüler von Lehrkräften unterrichtet wird, die in nennenswertem Umfang explizit lehren, wie und wann Lesestrategien einzusetzen sind. Obwohl die flexible Anwendung von Lesestrategien ein wichtiges Lernziel ist, werden sie von den Lehrkräften im Unterricht wenig beachtet.

Da in den vorangegangenen IGLU-Erhebungen unterschiedliche Indikatoren für die Qualität des Deutschunterrichts gebildet wurden, sind Trendvergleiche nicht möglich. Aber auch die früheren Erhebungen deuteten darauf hin, dass im Deutschunterricht wenig kognitive Aktivierung der Schülerinnen und Schüler stattfindet. In einer Analyse der Lehrerangaben in IGLU 2006 in Bezug auf kognitives Anregungspotential, Methodenvielfalt und individuelle Förderung haben Lankes und Carstensen (2007) fünf verschiedene Typen von Leseunterricht herausgestellt. In Deutschland vorherrschend waren „enggeführter Unterricht im Klassenverband ohne individuelle Unterstützung“ und „wenig anregender, wenig abwechslungsreicher Klassenunterricht ohne individuelle Förderung“. Für IGLU 2011 zeigten Tarelli, Lankes, Drossel und Gegenfurtner (2012), dass im Vergleich mit anderen EU-Staaten in Deutschland die Indikatoren zum kognitiven Anregungspotential des Leseunterrichts (gezielte Vermittlung von verständnisrelevantem Wissen oder Lesestrategien; Anwendung von Lesestrategien) eher im unterdurchschnittlichen Bereich liegen. Auch die Werte für Sozialformen des Unterrichts als Voraussetzung für Individualisierung waren relativ niedrig. Laut Bos et al. (2012) sind die Veränderungen im Leseunterricht der Teilnehmerstaaten – im Vergleich mit 2006 – nur geringfügig, was auf eine Stabilität der Unterrichtskulturen schließen lässt, die sich nur langfristig ändern lassen.

Förderung für leseschwache Schülerinnen und Schüler

Über das Ausmaß der schulischen Förderung von leseschwachen Schülerinnen und Schülern geben zwei Indikatoren von IGLU Auskunft. Im internationalen Lehrerfragebogen wurde erfasst, bei wie vielen Kindern Förderbedarf besteht und wie viele tatsächlich eine schulische Förderung erhalten. Nach Ansicht der Lehrkräfte besteht bei rund 17 Prozent der Schülerinnen und Schüler Förderbedarf; diese Zahl entspricht in etwa dem Anteil der Kinder unterhalb

von Kompetenzstufe III. Förderunterricht erhalten jedoch nur knapp 11 Prozent. Bemerkenswerterweise haben die Lehrkräfte 2011 bei 23 Prozent der Viertklässlerinnen und Viertklässler einen Förderbedarf gesehen (Tarelli et al., 2012) und offensichtlich einen strengeren Maßstab angelegt, denn im Jahr 2011 erreichten 15 Prozent der Schülerinnen und Schüler nicht die Kompetenzstufe III, während diese Zahl 2016 bei rund 19 Prozent liegt. Hätten die Lehrkräfte vergleichbare Kriterien angelegt, hätten sie 2016 mehr – und nicht weniger – Kindern einen Förderbedarf bescheinigen müssen. In Deutschland besteht somit eine große Diskrepanz zwischen Förderbedarf und tatsächlicher Förderung, die in anderen EU-Staaten deutlich geringer ausfällt.

Die Angaben der Lehrkräfte beziehen sich summarisch auf den Anteil der Kinder in der Klasse, nicht auf individuelle Schülerinnen und Schüler. Im nationalen Ergänzungsfragebogen wurden die Lehrkräfte gebeten, für jedes einzelne Kind anzugeben, ob es eine schulische Förderung benötigt und ob es sie erhält. Diese Angaben können auf die Leistungen des Kindes im IGLU-Lesetest bezogen werden. Bereits in IGLU 2006 zeigte sich, dass von den leseschwachen Schülerinnen und Schülern (Kompetenzstufe I und II) nur jedes dritte Kind eine besondere schulische Förderung erhielt (Valtin, Hornberg, Buddeberg, Kowoll & Potthoff, 2010). Dies ist in 2016 nicht anders. Nach wie vor erhält die große Mehrheit der Schülerinnen und Schüler in Deutschland (66 Prozent), bei denen laut IGLU-Lesetest ein Förderbedarf besteht, da sie unterhalb der Kompetenzstufe III liegen, keinerlei spezielle schulische Förderung.

3 Resümee

Seit der ersten IGLU-Erhebung im Jahr 2001 haben die Projektteams in Deutschland Handlungsfelder und Maßnahmen aufgezeigt, die auch im Jahr 2017 aktuell und von Relevanz sind (Bos et al., 2003; Valtin, Bos, Hornberg & Schwippert, 2007; Bos, Tarelli, Bremerich-Vos & Schwippert, 2012). Diese Maßnahmen betreffen institutionelle, schulstrukturelle und unterrichtliche Faktoren:

- Verstärkung des kognitiven Anregungspotentials im Leseunterricht
- gezielte Förderung für lesestarke Schülerinnen und Schüler
- gezielte Förderung für leseschwächere Schülerinnen und Schüler
- Verbesserung der Leseleistung bei ‚wissensbasierten‘ Leseaufgaben
- verstärkter Einsatz von Sachtexten im Leseunterricht
- Verbesserung der Chancengerechtigkeit für Schülerinnen und Schüler unabhängig von ihrem Geschlecht, ihrer sozialen Herkunft, ihrem Migrationshintergrund oder anderen Merkmalen
- Förderung gezielter Elternarbeit
- Verbesserung der Lehreraus- und -fortbildung
- Ausweitung und Verbesserung von Aus- und Fortbildungsangeboten gezielt im Bereich des adaptiven Unterrichtens
- Verstärkung des Angebots an Ganztagsgrundschulen
- Erhöhung der Durchlässigkeit im Schulsystem
- verstärkte Implementation von Leseförderung in der Sekundarstufe I
- Sicherung einer qualitativ hochwertigen Bildung und Erziehung im Elementarbereich.

Eine Einordnung dieser Handlungsempfehlungen in ein europäisches Rahmenkonzept der Leseförderung findet sich in Kapitel 12 in diesem Band.

Literatur

- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann. Zugriff am 03.10.2016 unter <http://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/bildungsbericht-2016>
- Berkemeyer, N., Bos, W., Hermstein, B., Abendroth, S. & Semper, I. (2017). *Chancenspiegel – eine Zwischenbilanz. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002*. Gütersloh: Bertelsmann Stiftung.
- 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.
- Bos, W., Lankes, E.-M., Schwippert, K., Valtin, R., Voss, A., Badel, I. & Plafmeier, N. (2003). Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin. (Hrsg.), *Erste Ergebnisse aus IGLU* (S. 69–142). Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.). (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Feierabend, S., Plankenhorn, T. & Rathgeb, T. (Hrsg.). (2017). *KIM-Studie 2016. Kindheit, Internet, Medien*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest.
- Haag, N., Kocaj, A., Jansen, M. & Kuhl, P. (2017). Soziale Disparitäten. In P. Stanat, S. Schipolowski, C. Rojsk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 213–235). Münster: Waxmann.
- Hornberg, S., Bos, W., Buddeberg, I., Potthoff, B. & Stubbe, T. C. (2007). Anlage und Durchführung von IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 21–46). Münster: Waxmann.
- Lankes, E.-M. & Carstensen, C. H. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 161–193). Münster: Waxmann.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Zugriff am 14.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/international-results/>
- Schipolowski, S., Wittig, J., Weirich, S. & Böhme, K. (2017). Geschlechtsbezogene Disparitäten. In P. Stanat, S. Schipolowski, C. Rojsk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 187–212). Münster: Waxmann.
- Stanat, P., Schipolowski, S., Rojsk, C., Weirich, S. & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Tarelli, I., Lankes, E.-M., Drossel, K. & Gegenfurtner, A. (2012). Lehr- und Lernbedingungen an Grundschulen 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. 137–169). Münster: Waxmann.
- Valtin, R., Bos, W., Hornberg, S. & Schwippert, K. (2007). Zusammenschau und Schlussfolgerungen. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 329–348). Münster: Waxmann.

- Valtin, R., Hornberg, S., Buddeberg, M., Kowoll, M. E. & Potthoff, B. (2010). Schülerinnen und Schüler mit Leseproblemen – eine ökosystemische Betrachtungsweise. 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. 43–90). Münster: Waxmann.
- Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K. & Kasper, D. (Hrsg.). (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.

Kapitel II

Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2016)

Anke Hußmann, Heike Wendt, Daniel Kasper, Wilfried Bos und Martin Goy

Im Jahr 2001 wurde als Folgestudie der *International Association for the Evaluation of Educational Achievement (IEA) Reading Literacy Study (RLS)* aus den Jahren 1990/1991 die *Progress in International Reading Literacy Study (PIRLS)* ins Leben gerufen, die weiterhin in der Verantwortung der IEA steht. Die IEA ist ein unabhängiger, internationaler Zusammenschluss von Wissenschaftlerinnen und Wissenschaftlern, Forschungseinrichtungen und Regierungsstellen und führt seit 1959 international vergleichende Schulleistungsstudien durch. PIRLS gehört seit 2001 zu den Kernstudien der IEA, wird alle fünf Jahre durchgeführt und trägt in Deutschland den Namen *Internationale Grundschul-Lese-Untersuchung (IGLU)*. Zentrales Anliegen der IEA ist es, mit dieser Studie langfristige Entwicklungen in den teilnehmenden Bildungssystemen unter Berücksichtigung curricularer Vorgaben und weiterer zentraler Rahmenbedingungen zu schulischen Lehr- und Lernumgebungen zu dokumentieren.

Deutschland nimmt seit 2001 regelmäßig auch in den Jahren 2006, 2011 und 2016 mit der vierten Jahrgangsstufe an der Studie teil (Bos et al., 2003; Bos et al., 2007; Bos, Tarelli, Bremerich-Vos & Schwippert, 2012). Neben der Berichterstattung aktueller Befunde aus IGLU 2016 in nationaler und internationaler Perspektive ermöglicht diese Beteiligung auch, den Blick auf Entwicklungen im Grundschulwesen seit 2001 zu richten. Die Teilnahme an IGLU erfolgt auf Beschluss der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK)* und in Vereinbarung mit der KMK und dem *Bundesministerium für Bildung und Forschung (BMBF)*. Sie ist Teil der systematischen Beobachtung des deutschen Bildungssystems im Rahmen der Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (KMK, 2015).

Als internationale Schulleistungsstudie stellt IGLU einen Vergleichsmaßstab für Schülerleistungen bereit, mit dem gezielt jene Vergleiche in den Blick ge-

nommen werden können, die aus deutscher Perspektive von Erkenntnisinteresse sind, beispielsweise zu weiteren Mitgliedstaaten der Europäischen Union. Damit unterscheidet sich IGLU von anderen Studien, die ebenfalls als Teil der KMK-Gesamtstrategie durchgeführt werden, zum Beispiel der *IQB-Bildungstrend* in der Primarstufe, der zur Überprüfung der in den Bildungsstandards formulierten Kompetenzanforderungen Vergleiche zwischen den Ländern der Bundesrepublik Deutschland erlaubt (zuletzt: Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017). IGLU stellt als standardisierte und wissenschaftlich verantwortete Studie eine wichtige Datenbasis für die erziehungswissenschaftliche Forschung bereit. Die gewonnenen Daten dienen als Grundlagen, bildungspolitisch relevante Fragestellungen untersuchen und damit Wissen über Faktoren erlangen zu können, die die Qualität und Entwicklung von Bildungssystemen beeinflussen.

1 Die Beteiligung Deutschlands an international vergleichenden Schulleistungsstudien zur Lesekompetenz am Ende der Grundschulzeit

IGLU ist als Trendstudie konzipiert. In IGLU 2016 wurden zum vierten Mal die Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe untersucht. Vorläuferstudie von IGLU ist die *Reading Literacy Study* (RLS), die zuvor – ebenfalls von der IEA – in den Jahren 1990/1991 durchgeführt wurde. Beteiligt hatten sich damals 31 Länder (Elley, 1992, 1994). Deutschland nahm mit repräsentativen Stichproben für die westlichen und für die östlichen Länder der Bundesrepublik teil (Lehmann, Peek, Pieper & van Stritzky, 1995). Das Leseverständnis wurde in dritten und achten Klassen getestet, anders als in IGLU stand zum damaligen Zeitpunkt noch nicht das Ende der Grundschulzeit im Fokus der Untersuchung. Erhoben wurden Kompetenzen beim Lesen erzählender, darlegender und dokumentierender Texte anhand verschiedener Lesetexte und überwiegend mit Aufgaben im *Multiple-Choice*-Format.

Im Jahr 2001 wurde zum ersten Mal PIRLS – beziehungsweise in Deutschland IGLU – als international vergleichende Schulleistungsstudie durchgeführt. Die IEA stützte sich beim Aufbau des Erhebungsdesigns im Wesentlichen auf die Erfahrungen und Erkenntnisse aus der RL-Studie, setzte jedoch auch deutliche Modifikationen um. Das Leseverstehen wurde anhand zweier Leseintentionen und vier Leseverstehensprozessen untersucht (siehe Kapitel 3 in diesem Band). Um einen direkten Vergleich zwischen Ergebnissen aus der RL-Studie von 1990/91 und PIRLS/IGLU zu erreichen, wurden in neun Teilnehmerstaaten im Rahmen der Erhebung von PIRLS 2001 wiederholt Lesetests der RL-Studie eingesetzt (Martin, Mullis, Gonzalez & Kennedy, 2003). Deutschland war an dieser vergleichenden Untersuchung nicht beteiligt, die Ergebnisse aus IGLU 2001 und IGLU 2006 konnten jedoch mit jenen der RL-Studie kontrastiert werden (Bos et al., 2007).

An IGLU 2001 beteiligten sich international 35 Staaten und Regionen. In Deutschland wurden die eingesetzten Lesetests in 12 der 16 Länder um weitere Inhaltsbereiche ergänzt: An einem zweiten Testtag wurden mathematische, naturwissenschaftliche, orthographische Kompetenzen und die Schreibkompetenz der Schülerinnen und Schüler erhoben. Die Aufgaben zur Testung mathematischer und naturwissenschaftlicher Kompetenzen wurden zum Teil aus dem Aufgabenpool der *Trends in International Mathematics and Science Study*

(TIMSS 1995) für die Primarstufe entnommen, so dass eine internationale Verankerung gelang. Mit dieser nationalen Erweiterung wurden verlässliche Befunde zu Kompetenzen gewonnen, die am Ende der Grundschulzeit von Bedeutung sind.

Darüber hinaus entschieden sich in IGLU 2001 einige der Länder der Bundesrepublik Deutschland für eine Erweiterung der Stichprobe (*oversampling*), um repräsentative Ergebnisse für einzelne Länder zu erhalten und damit Vergleiche zwischen diesen Ländern (Baden-Württemberg, Bayern, Brandenburg, Bremen und Hessen) durchführen zu können. Für Nordrhein-Westfalen waren Analysen auf Länderebene möglich, weil im Rahmen der regulären Stichprobenziehung für Deutschland aus diesem Land ausreichend Schulen gezogen wurden. In Thüringen ergänzte das Ministerium des Landes die Stichprobe um weitere Schulen, mit dem Ziel, Schulentwicklungsprozesse unterstützen zu können. Bei der internationalen Auswertung wurden alle zusätzlich einbezogenen Schulen nicht berücksichtigt, sie dienten ausschließlich nationalen Analysen, wie detailliert dargelegt ist (Bos et al., 2003; Bos et al., 2004; Bos, Lankes, Prenzel, Valtin & Walther, 2005).

Die zweite Erhebung von IGLU wurde im Jahr 2006 durchgeführt. Beteiligt waren 45 Staaten und Regionen. Die Stichprobe in Deutschland wurde erweitert, so dass erstmalig Vergleiche zwischen allen Ländern möglich wurden. Pro Land wurden 25 Schulen gezogen, mit Ausnahme von Nordrhein-Westfalen, wo die Daten von 35 Schulen einbezogen wurden. Wie in IGLU 2001 so wurden auch in IGLU 2006 die internationalen um nationale Testkomponenten ergänzt, in 2006 um solche zur Erfassung orthographischer Kompetenzen. Da für Deutschland bereits eine Teilnahme an TIMSS 2007 (Bos et al., 2008) festgelegt war, wurde in 2006 auf die zusätzliche Testung mathematischer und naturwissenschaftlicher Kompetenzen verzichtet. Ausführliche Berichte zu den Ergebnissen der internationalen und nationalen Analysen liegen vor (Bos et al., 2007; Mullis, Martin, Kennedy & Foy, 2007).

Die dritte Erhebung von IGLU fand im Jahr 2011 statt. In Deutschland wurde sie zeitgleich mit TIMSS 2011 durchgeführt. Beteiligt mit der vierten Jahrgangsstufe (oder der äquivalenten Klassenstufe) waren 45 Staaten und Regionen als reguläre Teilnehmer. Erstmals wurden in IGLU 2011 auch neun zusätzliche Regionen als sogenannte Benchmark-Teilnehmer aufgenommen, die ihre Ergebnisse zwar auf der internationalen Leistungsskala verorten möchten, deren Leistungsdaten aber nicht in die Berechnung des internationalen Mittelwerts eingehen. Aufgrund der zeitgleichen Durchführung von IGLU und TIMSS, entschied man in Deutschland (sowie in 36 anderen Teilnehmerstaaten und -regionen) in einer gemeinsamen Stichprobe zu erheben. Die Umsetzung erfolgte an zwei Testtagen, an denen in der einen Hälfte der Schulen die IGLU-Testkomponente am ersten Tag und die TIMSS-Testkomponente am zweiten Tag durchgeführt wurde, während in der zweiten Hälfte der Schulen die Testungen in umgekehrter Reihenfolge stattfanden. Getestet wurden sowohl Lesekompetenzen (IGLU) als auch mathematische und naturwissenschaftliche Kompetenzen (TIMSS). Die Stichprobe umfasste rund 200 Schulen. Die 16 Länder der Bundesrepublik Deutschland wurden gemäß ihrer Größe und Schulanzahl adäquat durch die Stichprobe abgebildet. Die nationalen und internationalen Befunde sind umfassend dokumentiert (Bos, Tarelli et al., 2012; Bos, Wendt, Köller & Selter, 2012; Mullis, Martin, Foy & Drucker, 2012; Wendt, Stubbe, Schwippert & Bos, 2015).

Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe im Trend zu untersuchen, setzt eine einheitliche Linie von Zyklus zu Zyklus voraus. Jeder Erhebungszyklus zeichnet sich jedoch durch Besonderheiten beziehungsweise Veränderungen aus. Für IGLU 2016 sind im Vergleich zu den vorangegangenen Erhebungszyklen folgende Spezifika zu benennen:

- Für die Einordnung der Leistungsergebnisse von Grundschulkindern steht nicht nur der internationale Bezugsrahmen zur Verfügung, sondern auch der Vergleich im Trend der Erhebungen aus 2001, 2006 und 2011. Damit sind erstmalig auf solider Datengrundlage Trendaussagen zu Kompetenzveränderungen am Ende der Grundschulzeit in den letzten 15 Jahren möglich.
- Im Gegensatz zu IGLU 2001 und IGLU 2006 lässt die Datengrundlage von IGLU 2016 keine Analyse auf der Ebene der Länder der Bundesrepublik Deutschland zu. Dies ist bereits seit IGLU 2011 nicht mehr möglich. Für Analysen zu Lesekompetenzen am Ende der Grundschulzeit im Vergleich der Länder sei daher auf den aktuell vorliegenden *IQB-Bildungstrend 2016* verwiesen (Stanat et al., 2017).

2 Zentrale Erkenntnisse und Fragestellungen

Die einleitend beschriebene Zielsetzung von IGLU, den Ertrag von Bildungssystemen in international vergleichender Perspektive zu dokumentieren, wird hinsichtlich ihrer Möglichkeiten und Grenzen vielfältig diskutiert und reflektiert (z.B. Baumert, 2016; BMBF, 2001; Bos, Postlethwaite & Gebauer, 2010; Klieme, 2013). Herausgestellt werden in dieser Auseinandersetzung beispielsweise die verschiedenen Funktionen und Ansprüche, die mit Schulleistungsstudien verbunden sind (z.B. Howie & Plomp, 2005). Nach Klieme und Vieluf (2013) sind Schulleistungsstudien hinsichtlich ihrer Funktionen zu unterscheiden, die sie für unterschiedliche gesellschaftliche Gruppen erfüllen, zum Beispiel für die Politik, die Öffentlichkeit oder die Forschung. Aus dieser Perspektive betrachtet liefern Schulleistungsstudien wie IGLU als Systemmonitoringstudien *Indikatoren*, anhand derer sich die Strukturen, Funktionen und Erträge ebenso wie Disparitäten in Leistungen und die Bildungsteilhabe in Bildungssystemen beobachten, beschreiben und in einen internationalen Referenzrahmen einordnen lassen. Die Repräsentativität und die damit verbundene Aussagekraft der *Daten* erlauben zugleich, erziehungswissenschaftliche Grundlagenforschung zu betreiben, ebenso wie bildungspolitisch relevante Fragestellungen zu untersuchen. Es wird *Wissen* über Faktoren generiert, die die Qualität und Entwicklung von Bildungssystemen – sowohl in international vergleichender als auch in nationaler Perspektive – beeinflussen.

Mit der Durchführung international vergleichender Schulleistungsstudien wie IGLU verfolgt die IEA somit nicht nur die Feststellung dessen, was Bildungssysteme leisten. Zugleich werden potentielle Einflussfaktoren des Kompetenzerwerbs berücksichtigt, um weitere Erkenntnisse über die Bedingungen und Möglichkeiten der Verbesserung schulischer Förderung zu erhalten (IEA, 2012). Als potentielle Einflussfaktoren werden zentrale Rahmenbedingungen des Lesekompetenzerwerbs und des Unterrichts in der Grundschule untersucht (siehe Abbildung 2.1, Abschnitt 4.1.2). Die Lesekompetenzen der Schülerinnen und Schüler werden dazu anhand von Leistungstests ermittelt. Die Erfassung der Rahmenbedingungen erfolgt auf der Grundlage einer schriftlichen Befragung der Schülerinnen und Schüler, ihrer Eltern, der Lehrkräfte und der Schulleitungen so-

wie anhand der Einschätzung von Experten (siehe Abschnitt 5.2). Insofern bietet IGLU vielfältige Erkenntnismöglichkeiten, sowohl für Staaten, die regelmäßig an der Studie teilnehmen (z. B. um Leistungsveränderungen im Zeitverlauf zu beobachten), als auch für Staaten, die zum ersten Mal an IGLU teilnehmen (z. B. für den Leistungsvergleich mit anderen Staaten). IGLU gibt einen Einblick in die Bedingungen des Lehrens und Lernens am Ende der Grundschulzeit und trägt so zum Verständnis innerhalb und zwischen Bildungssystemen und ihrer Erträge bei. Aus diesen Erkenntnissen lassen sich Hinweise für Verbesserungen sowie weitergehende Forschungsbedarfe ableiten.

Die Teilnahme an IGLU 2016 und die damit verbundene Erhebung von umfassenden empirischen Daten erlauben repräsentative Rückschlüsse auf die Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe mit folgendem Erkenntnisgewinn:

- Die standardisierte Durchführung und die statistisch angemessene Auswertung elaborierter und erprobter Kompetenztestungen für den Bereich Leseverständnis lassen eine zuverlässige Einschätzung des Leistungsniveaus von Viertklässlerinnen und Viertklässlern in Deutschland im internationalen Vergleich zu.
- Die vierte Teilnahme an IGLU ermöglicht, Trends zu den Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe im internationalen Vergleich einzuschätzen. Die Auswertung erlaubt für die vergangenen 15 Jahre eine Bilanz der Entwicklungen der Leseleistungen am Ende der Grundschulzeit und der Rahmenbedingungen, in die der Lesekompetenzerwerb eingebettet ist.
- Die Befragung von Schülerinnen und Schülern zu ihrem Leseselbstkonzept, ihrer Lesemotivation und zu ihrem Lernverhalten ermöglicht eine zuverlässige Einschätzung der Bedeutsamkeit dieser Aspekte für die Leistungen von Viertklässlerinnen und Viertklässlern in Deutschland im internationalen Vergleich.
- Die Befragung von weiteren schulischen Akteurinnen und Akteuren (z. B. Fachlehrkräften und Schulleitungen) zu zentralen Bedingungsfaktoren schulischen Lernens ermöglicht eine zuverlässige Einschätzung der Bedeutsamkeit dieser Aspekte für die Testleistungen von Viertklässlerinnen und Viertklässlern in Deutschland im internationalen Vergleich.
- Die Befragung von Eltern, deren Kinder an IGLU teilgenommen haben, zu individuellen Voraussetzungen ermöglicht eine zuverlässige Einschätzung der Bedeutsamkeit dieser Aspekte für die Lesekompetenzentwicklung von Viertklässlerinnen und Viertklässlern in Deutschland im internationalen Vergleich.
- Die Stichprobenziehung ermöglicht, belastbare Aussagen zum Leistungsniveau und zur Bedeutsamkeit von individuellen Lernvoraussetzungen und Kontextfaktoren für Subpopulationen zu treffen (z. B. hinsichtlich der Aspekte Geschlecht, soziale Herkunft und Migrationshintergrund).

Der vorliegende Berichtsband stellt die Ergebnisse des internationalen Vergleichs aus deutscher Perspektive zu IGLU 2016 dar. Konzeptionelle und technische Besonderheiten sind für die Zielgruppe der Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe dokumentiert. Vertiefende Fragestellungen werden in Nachfolgepublikationen aufgegriffen und behandelt. Im Fokus dieses Berichts steht die Beantwortung folgender Fragestellungen:

1. Welches Lesekompetenzniveau erreichen Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe in Deutschland im Jahr 2016? Wie lassen sich die Ergebnisse im internationalen Vergleich bewerten? Haben sich die Ergebnisse seit 2001 verändert?
2. Wie lassen sich die IGLU-Leistungskennwerte auf Kompetenzstufen einordnen? Wie groß sind die Gruppen der auffällig leistungsschwachen und leistungsstarken Schülerinnen und Schüler? Wie unterscheiden sich die Ergebnisse von IGLU 2001, 2006, 2011 und 2016 diesbezüglich?
3. Welche Ergebnisse erzielen die Viertklässlerinnen und Viertklässler in Deutschland differenziert nach Textsorten und Verstehensprozessen? Wie unterscheiden sich die Ergebnisse von IGLU 2001, 2006, 2011 und 2016? Sind beobachtbare Veränderungen für bestimmte Subgruppen konstant?
4. Welche Bedeutung haben die individuellen Lernvoraussetzungen und Kontextfaktoren für die Lesekompetenzen? Welche Veränderungen von Lehr- und Lernbedingungen lassen sich seit 2001 beobachten?

3 IGLU 2016 – ein kooperatives Forschungsprojekt

Die Realisierung internationaler Vergleichsstudien wie IGLU ist aufwendig und erfordert die Zusammenarbeit verschiedener Organisationen, Institutionen und Personen auf nationaler und internationaler Ebene. Die IEA hat das *TIMSS & PIRLS International Study Center (ISC)* am *Boston College* in Chestnut Hill, Massachusetts, USA mit dem internationalen Management der Studie beauftragt. Es steht unter der Leitung von Ina V. S. Mullis, Professor an der *Lynch School of Education, Boston College*, sowie Michael O. Martin, Research Professor an der *Lynch School of Education, Boston College*. Das ISC verantwortet das Design und die Implementation der Studie, die internationale Koordination der Entwicklung der Instrumente und der Erhebungsprozeduren sowie die Qualität der Datenerhebung. Die internationale Skalierung (siehe Abschnitt 9) wird am ISC durchgeführt, auch der internationale Ergebnisbericht wird dort verfasst. Für Stichprobenziehung, Dokumentation der nationalen Stichproben und Berechnung der internationalen Stichprobengewichte kooperiert die internationale Studienleitung am ISC mit *Statistics Canada* in Ottawa, Ontario und der Abteilung der Stichprobenziehung an der *IEA Hamburg*.

Verantwortung für die Vorbereitung und Durchführung der Studie in jedem der an IGLU 2016 beteiligten Staaten trägt ein nationaler Projektkoordinator (der sogenannte *National Research Coordinator, NRC*). International vorgegebene Richtlinien sind dabei verbindlich. In Deutschland obliegt Prof. Dr. Wilfried Bos am *Institut für Schulentwicklungsforschung (IFS)* an der *Technischen Universität Dortmund* diese Verantwortung. Diese Aufgabe hat er gemeinsam mit Frau Dr. Heike Wendt, ebenfalls am IFS, wahrgenommen. Die Durchführung von IGLU in Deutschland wird zu gleichen Teilen vom BMBF und von der KMK finanziert.

Für die Analyse der Studienergebnisse und die Berichtslegung in Deutschland ist ein nationales Konsortium unter der Federführung des IFS verantwortlich, dem folgende Wissenschaftlerinnen und Wissenschaftler angehören:

Prof. Dr. Wilfried Bos	(NRC, Wissenschaftlicher Leiter von IGLU 2016 in Deutschland und Sprecher des Konsortiums) – Professur für Bildungsforschung und Qualitätssicherung am Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund
Dr. Heike Wendt	(NRC und Co-Projektleitung) – Akademische Rätin am Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund
Dr. Anke Hußmann	(Projektleitung) – Wissenschaftliche Mitarbeiterin am Institut für Schulentwicklungsforschung (IFS) an der Technischen Universität Dortmund
Prof. Dr. Albert Bremerich-Vos (i.R.)	(Mitglied des Konsortiums) – bis 2017 Professur für Linguistik und Sprachdidaktik an der Universität Duisburg-Essen
Prof. Dr. Nele McElvany	(Mitglied des Konsortiums) – Professur für Empirische Bildungsforschung mit dem Schwerpunkt Lehren und Lernen im schulischen Kontext an der Technischen Universität Dortmund, Geschäftsführende Direktorin des Instituts für Schulentwicklungsforschung (IFS)
Prof. Dr. Eva-Maria Lankes	(Kooptiertes Mitglied des Konsortiums) – Professur für Schulpädagogik an der Technischen Universität München, School of Education sowie Leitung der Qualitätsagentur des Staatsinstituts für Schulqualität und Bildungsforschung München
Prof. Dr. Tobias C. Stubbe	(Kooptiertes Mitglied des Konsortiums) – Professur für Schulpädagogik und Empirische Schulforschung am Institut für Erziehungswissenschaft an der Georg-August-Universität Göttingen
Prof. Dr. Renate Valtin (i.R.)	(Kooptiertes Mitglied des Konsortiums) – bis 2008 Professur für Grundschulpädagogik an der Humboldt-Universität zu Berlin, Präsidentin der Europäischen Lesegesellschaften

Ein großer Teil der Projektarbeit erfolgte hauptverantwortlich am IFS unter der Leitung von Dr. Anke Hußmann und Dr. Heike Wendt. Die methodische Betreuung der Studie am IFS oblag Dr. Daniel Kasper. Zu den beteiligten wissenschaftlichen Mitarbeiterinnen und Mitarbeitern zählen: Martin Goy, Maike Hoeft (bis Januar 2016), Dr. Svenja Rieser (bis Oktober 2017), Daniel Scott Smith (bis Juni 2016), Dr. Ruven Stahns und Anna Vaskova (bis Mai 2016). Das Projektteam wurde von vielen studentischen Hilfskräften unterstützt, darunter Cihan Günes, Ruth Engel (bis Juni 2017), Frederike Joosten, Donieta Jusufi, Adriana Kilisch, Katharina Roth, Lisa Schmitt (bis Oktober 2017), Maximilian Schulz, Felix Senger (bis August 2016), Charlotte Siepman und Katja Weber. Mit der Organisation der nationalen Datenerhebung und -verarbeitung sowie der Aufgabenkodierung hat das IFS die *IEA Hamburg* beauftragt.

4 Anlage und Durchführung von IGLU 2016

Im vorliegenden Bericht werden zentrale Ergebnisse aus IGLU 2016 für Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe vorgestellt. Im Vordergrund stehen die Ergebnisse des internationalen Vergleichs. Diese werden ergänzt um auf Deutschland fokussierte Analysen zu weiterführenden Fragestellungen. In diesem Abschnitt wird ein Überblick über das der Studie zugrundeliegende Rahmenmodell und die untersuchten Kompetenzbereiche gegeben. Vorgestellt werden darüber hinaus die Bildungssysteme, die an IGLU 2016 teilgenommen haben, sowie jene, die bei Vergleichen im Trend berichtet werden. Weitere Informationen, die dem besseren Verständnis beim Lesen des Berichtes dienen, sind ebenfalls dokumentiert, darunter die Teilnahmemodalitäten zum aktuellen Erhebungszyklus und zu den vergangenen Erhebungszyklen in den Jahren 2001, 2006 und 2011.

4.1 Die Rahmenkonzeption der Studie

Unter Berücksichtigung von Rahmenbedingungen, die für das Lernen in der Schule von Bedeutung sind, untersucht IGLU im internationalen Vergleich Leseleistungen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe. Von besonderem Interesse ist dabei die Frage, wie erfolgreich Schülerinnen und Schüler in den Bildungssystemen der teilnehmenden Staaten und Regionen am Ende der vierten Jahrgangsstufe grundlegende Kompetenzen erworben haben und ob sich Teilgruppen von Schülerinnen und Schülern identifizieren lassen, die sich in ihren Kompetenzen systematisch voneinander unterscheiden. Die Studie ist folglich so angelegt, dass gesicherte Aussagen über die Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich getroffen werden können. Die theoretischen und konzeptionellen Grundlagen der Studie sind umfassend in einem Rahmenkonzept formuliert, das Expertinnen und Experten aus unterschiedlichen Teilnehmerstaaten erarbeitet haben (Mullis, Martin & Sainsbury, 2015). Es wurde erstmals zu IGLU 2001 vorgestellt und im Verlauf der folgenden Zyklen kontinuierlich angepasst (Campbell, Kelly, Mullis, Martin & Sainsbury, 2001; Mullis, Kennedy, Martin & Sainsbury, 2006; Mullis, Martin, Kennedy, Trong & Sainsbury, 2009).

Das Verständnis von Lesekompetenz ist durch die aus dem angloamerikanischen Raum stammende *Literacy*-Konzeption definiert (siehe Abschnitt 4.1.1). Der nationalen Berichterstattung liegt darüber hinaus ein theoretisches Rahmenmodell zugrunde, das der Beschreibung des Zusammenhangs von Schülerleistungen und deren Bedingungen dient (siehe Abschnitt 4.1.2).

4.1.1 *Literacy* als Rahmenkonzept für die Testentwicklung

Theoretisch und konzeptionell liegt dem Verständnis von Lesekompetenz in IGLU die angloamerikanische *Literacy*-Konzeption zugrunde. Es handelt sich um ein pragmatisches Konzept von Grundbildung, in dem grundlegende Kompetenzen definiert sind, die in der Wissensgesellschaft von Bedeutung sind. Die Definition von *reading literacy*, auf die man sich in IGLU stützt, basiert auf einer Vielzahl an Theorien, in denen dieses Konzept als ein konstruktiver und

interaktiver Prozess verstanden wird (Anderson & Pearson, 1984; Chall, 1983; Kintch, 1998, 2012, 2013; Ruddell & Unrau, 2004; Rumelhart, 1985):

„Reading literacy is the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment.“ (Mullis et al., 2015, S. 12)

Die in IGLU erfasste Lesekompetenz wird im Sinne von *reading literacy* als die Fähigkeit verstanden, solche schriftsprachlichen Formen verstehen und nutzen zu können, die gesellschaftlich erforderlich und/oder für den Einzelnen von Bedeutung sind. Sie beschreibt demnach die Fähigkeit, in unterschiedlichen, für die Lebensbewältigung praktisch bedeutsamen Verwendungssituationen lesen zu können. Betont werden bei diesem Verständnis von Lesen die kulturelle Bedeutung von Bildungsinhalten, Partizipationsmöglichkeiten und die Entwicklung von Kompetenzen für ein selbstgesteuertes und lebenslanges Lernen. Es wird davon ausgegangen, dass junge Leserinnen und Leser auf unterschiedliche Weise die Bedeutung eines Textes konstruieren und das Leseverständnis im Zusammenhang mit Textsorten und Leseanlässen steht. Kinder lesen also, um zu lernen, um sich mit anderen Leserinnen und Lesern in der Schule und im Alltag auszutauschen und zu ihrem Vergnügen (ebd.). In IGLU werden drei Bereiche der Lesekompetenz betrachtet (siehe vertiefend die Ausführungen in den Kapiteln 3 und 4 in diesem Band):

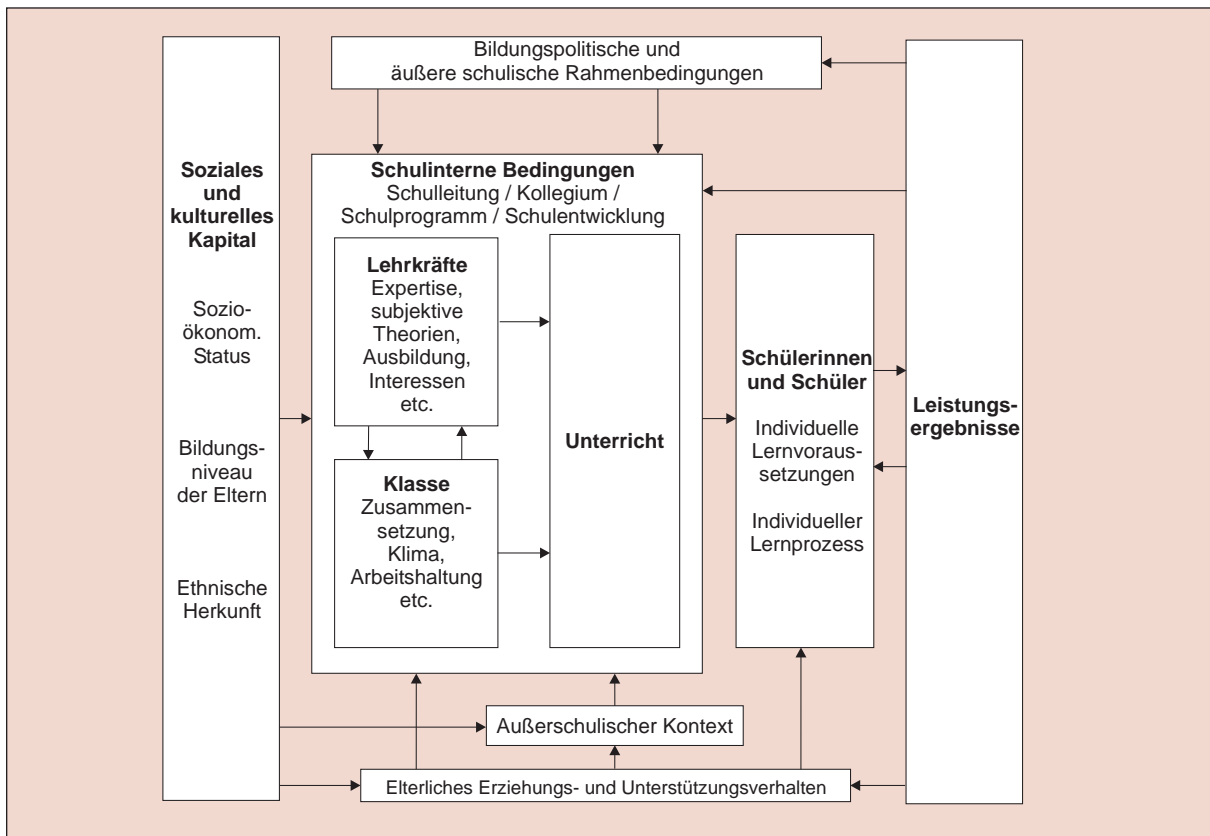
- Leseintentionen,
- Leseverstehensprozesse,
- Leseselbstkonzept, Lesemotivation und Leseverhalten.

Intentionen und Verstehensprozesse des Lesens von Schülerinnen und Schülern werden mit den eingesetzten Kompetenztests erhoben, Angaben zum Leseselbstkonzept, zur Lesemotivation und zum Leseverhalten und zu Einstellungen der Schülerinnen und Schüler mittels Fragebögen.

4.1.2 Rahmenmodell der Bedingungen schulischer Leistungen

Bereits zur nationalen Berichtslegung zu IGLU 2001 wurde ein theoretisches Rahmenmodell vorgestellt, das der Beschreibung des Zusammenhangs von Schülerleistungen und deren Bedingungen dienen sollte. In diesem Rahmenmodell werden Lehren und Lernen im Kontext gesellschaftlicher Ausgangsbedingungen verortet (vgl. Baumert & Weiß, 2002; Lankes et al., 2003; Prenzel & Doll, 2002; Wang, Haeterl & Walberg, 1993; Weinert & Helmke, 1997). Über die Erhebungszyklen hinweg hat es sich als Rahmenmodell zur Beschreibung schulischer Leistungen bewährt. Berücksichtigt werden darin die vielfältigen gesellschaftlichen, institutionellen und individuellen Bedingungen, in die das Lehren und Lernen in Grundschulen eingebettet ist. Bis heute hat das Modell nichts von seiner Bedeutung und Aktualität eingebüßt. Auch im vorliegenden Bericht dient es als Orientierungs- und Beschreibungsrahmen, um die Ebenen, die potentiell Einfluss auf die Qualität der Leistungen ausüben, in ihrem Beziehungsgeflecht betrachten zu können. Das Modell veranschaulicht, dass die erzielten Leistungen der Schülerinnen und Schüler als Ergebnis komplexer

Abbildung 2.1: Rahmenmodell zur Betrachtung des Zusammenhangs von Schülerleistungen und deren Bedingungen



IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Lernprozesse in schulischen und außerschulischen Kontexten zu verstehen sind. Berücksichtigt sind gesellschaftliche und institutionelle Rahmenbedingungen sowie individuelle und familiäre Voraussetzungen (siehe Abbildung 2.1).

Dargestellt sind im Zentrum des Modells zum einen die Schülerinnen und Schüler, deren individuellen Lernvoraussetzungen und Lernprozesse in direktem Zusammenhang mit Leistungen beziehungsweise Leistungsergebnissen zu betrachten sind. Zum anderen sind die schulinternen Rahmenbedingungen, die Lehrkräfte, die Klassen und der Unterricht von zentraler Bedeutung. Als weitere im Modell aufgeführte Bedingungsfaktoren werden außerschulische und familiäre Merkmale, der soziokulturelle Hintergrund und das elterliche Erziehungs- und Unterstützungsverhalten einbezogen. Berücksichtigt werden darüber hinaus institutionelle Merkmale wie die bildungspolitischen und äußeren schulischen Rahmenbedingungen. Markiert durch die Pfeile, sind im Modell auch Hinweise auf Wechselwirkungen zwischen den Faktoren gegeben: Ergebnisse dieser Studie, so ist beispielsweise zu entnehmen, wirken auf bildungspolitische und äußere schulische Rahmenbedingungen sowie individuelle Lernvoraussetzungen und -prozesse.

Für die Analyse und die statistische Betrachtung von Unterschieden in den Schülerleistungen bieten die im Modell aufgezeigten Bedingungsfaktoren und deren komplexe wechselseitige Verflechtungen wertvolle Anhaltspunkte. Deshalb werden in IGLU die institutionellen und gesellschaftlichen Rahmenbedingungen, von denen angenommen wird, dass sie die Bereitstellung und Nutzung schulischer Lerngelegenheiten beeinflussen, mit einer standardisierten schriftlichen Befragung der teilnehmenden Schülerinnen und Schüler, ihrer Eltern, der

unterrichtenden Lehrkräfte und der Schulleitungen erfasst. Die so erhobenen Rahmendaten ermöglichen in der vertiefenden Analyse, Zusammenhänge zwischen den Leistungsergebnissen und den genannten Hintergrundmerkmalen zu beleuchten. Aussichtsreich ist dies insbesondere im Hinblick auf die Beschreibung und Erklärung von geschlechtsbezogenen, sozialen oder kulturellen Disparitäten in den Schülerleistungen (siehe auch Kapitel 5, 6 und 7 in diesem Band).

Die Rahmenbedingungen des Lehrens und Lernens werden in IGLU nicht nur mit Hilfe der administrierten Fragebögen an die Schulleitungen, die Lehrerinnen und Lehrer, die Schülerinnen und Schüler und Eltern erfasst. Zur Interpretation und Einordnung der Befunde dienen weitere Instrumente als Informationsquellen (Hooper, Mullis & Martin, 2015):

Ergänzt wird die internationale Berichterstattung der Leistungsergebnisse in IGLU durch die sogenannte *Enzyklopädie* (Mullis, Martin, Goh & Prendergast, 2017), in der die Schulsysteme der einzelnen Teilnehmerstaaten und -regionen beschrieben sind. Den Kapiteln der Enzyklopädie liegen einheitliche Kriterien zugrunde und sie werden von den nationalen Studienkoordinatoren verfasst. Schwerpunkte der Kapitel sind die Beschreibung der Organisation und der Strukturen des jeweiligen Bildungssystems und der fachspezifischen Curricula. Tabellarische Darstellungen zu demographischen Charakteristika der beteiligten Bildungssysteme, Inhalte und der Aufbau von Curricula sowie Spezifika der Lehrerinnen- und Lehrerbildung sind weitere Bestandteile der Kapitel. Erfasst werden die in der Enzyklopädie präsentierten Informationen von der internationalen Studienleitung mit Hilfe von standardisierten Fragebögen (*Curriculum Questionnaire*), die von ausgewählten Fachexpertinnen und Fachexperten der Teilnehmerstaaten und -regionen bearbeitet werden.

4.2 Teilnahme und Teilnahmemodalitäten an IGLU 2016

4.2.1 Teilnehmende 2016

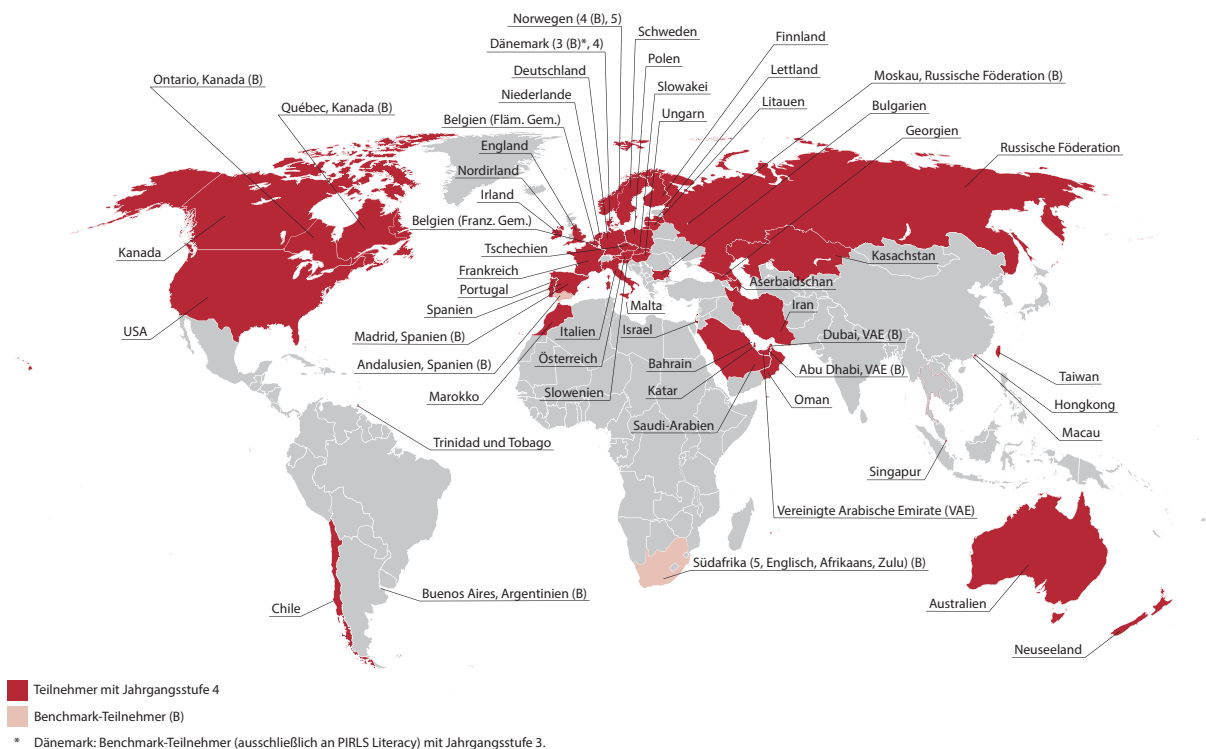
An IGLU 2016 haben sich weltweit 57 Bildungssysteme beteiligt. Es handelt sich überwiegend um souveräne Staaten. Teilgenommen haben auch Regionen, die als solche in Bildungsfragen weitgehend autonom handeln und deswegen in der internationalen Berichterstattung als Teilnehmer aufgeführt werden (z.B. die Flämische Gemeinschaft in Belgien). Insgesamt haben 47 Staaten und Regionen als reguläre Teilnehmer an der Studie teilgenommen (siehe Abbildung 2.2). Zehn sind Benchmark-Teilnehmer. Bei diesen handelt es sich um weitere Bildungssysteme, die ihre Ergebnisse auf der internationalen Ergebnisskala verorten möchten, deren Leistungsdaten aber nicht in die Berechnung des internationalen Mittelwerts eingehen.

Eine Besonderheit in IGLU 2016 stellt die Erweiterungsstudie PIRLS *Literacy* dar. In der Vergangenheit zeigte sich, dass in den Staaten, in denen die Lesekompetenzen der Viertklässlerinnen und Viertklässler deutlich unterhalb des internationalen Niveaus lagen, die Kompetenzen der leistungsschwächeren Kinder mit den regulären IGLU-Instrumenten nicht angemessen erfasst werden konnten. Für den Erhebungszyklus in 2016 entschied die IEA daher, IGLU um die Studienkomponente PIRLS *Literacy* zu erweitern (Mullis & Martin, 2015). PIRLS *Literacy* richtet sich auf Basiskompetenzen des Lesens, orientiert sich an der IGLU-Rahmenkonzeption und ist daher ähnlich wie IGLU aufgebaut. In PIRLS *Literacy* wurden zehn Lesetexte eingesetzt, von denen zwei (ein literari-

scher und ein informierender Text) auch in den regulären Testheften von IGLU rotieren (siehe Tabelle 2.3, Abschnitt 5.1.2). Drei Teilnehmer (Ägypten, Kuwait und Südafrika mit Jgst. 4) haben ausschließlich an PIRLS *Literacy* teilgenommen. Sie sind nicht Teil dieser Berichterstattung. Es gibt aber auch Teilnehmer, die als reguläre Teilnehmer an IGLU und zusätzlich an PIRLS *Literacy* teilgenommen haben (Iran und Marokko). Deutschland hat nicht an PIRLS *Literacy* teilgenommen, sondern lediglich als regulärer Teilnehmer an IGLU.

Kein Bestandteil dieses Berichts sind Ergebnisse der Teilnehmer aus PIRLS *Literacy*. Eine Ausnahme stellt Dänemark dar. Dänemark hat sich, neben der regulären Teilnahme mit der vierten Jahrgangsstufe, zusätzlich mit der dritten Jahrgangsstufe als Benchmark-Teilnehmer und hierbei ausschließlich an PIRLS *Literacy* beteiligt. Dieser Kohortenvergleich kann für Deutschland im europäischen Vergleich von Interesse sein. Daher wird Dänemark auf der Gesamtskala Lesen (siehe Kapitel 3, Abbildung 3.6 in diesem Band) als regulärer Teilnehmer mit der vierten Jahrgangsstufe und zusätzlich als Benchmark-Teilnehmer mit der dritten Jahrgangsstufe aufgeführt.

Abbildung 2.2: Staaten und Regionen, die an IGLU 2016 teilgenommen haben



In diesem Berichtsband sind die Ergebnisse der Schülerinnen und Schüler auf der Gesamtskala Lesen somit für alle beteiligten Staaten, Regionen und Benchmark-Teilnehmer, einschließlich Dänemark mit Jahrgangsstufe 3 als Benchmark-Teilnehmer an PIRLS *Literacy*, dokumentiert.

Wie in den vorangegangenen Erhebungszyklen werden bei der Darstellung der Ergebnisse Vergleichsgruppen (VG) gebildet. In diese Vergleichsgruppen werden jene Staaten und Staatengruppen aufgenommen, die für einen Vergleich mit Deutschland geeignet sind (z.B. aufgrund eines ähnlichen kulturellen Hintergrunds oder einer ähnlichen wirtschaftlichen Situation). Dabei handelt es sich

um die an IGLU 2016 teilnehmenden Mitglieder der *Europäischen Union* (EU) und der *Organisation for Economic Co-operation and Development* (OECD). In den Ergebnisdarstellungen sind in den nachfolgenden Kapiteln nicht nur Mittelwerte für einzelne Teilnehmer, sondern auch Mittelwerte für die beiden Vergleichsgruppen der EU (VG EU) und OECD (VG OECD) dargestellt. Die Teilnehmer der Vergleichsgruppen sind in Tabelle 2.1 aufgeführt.

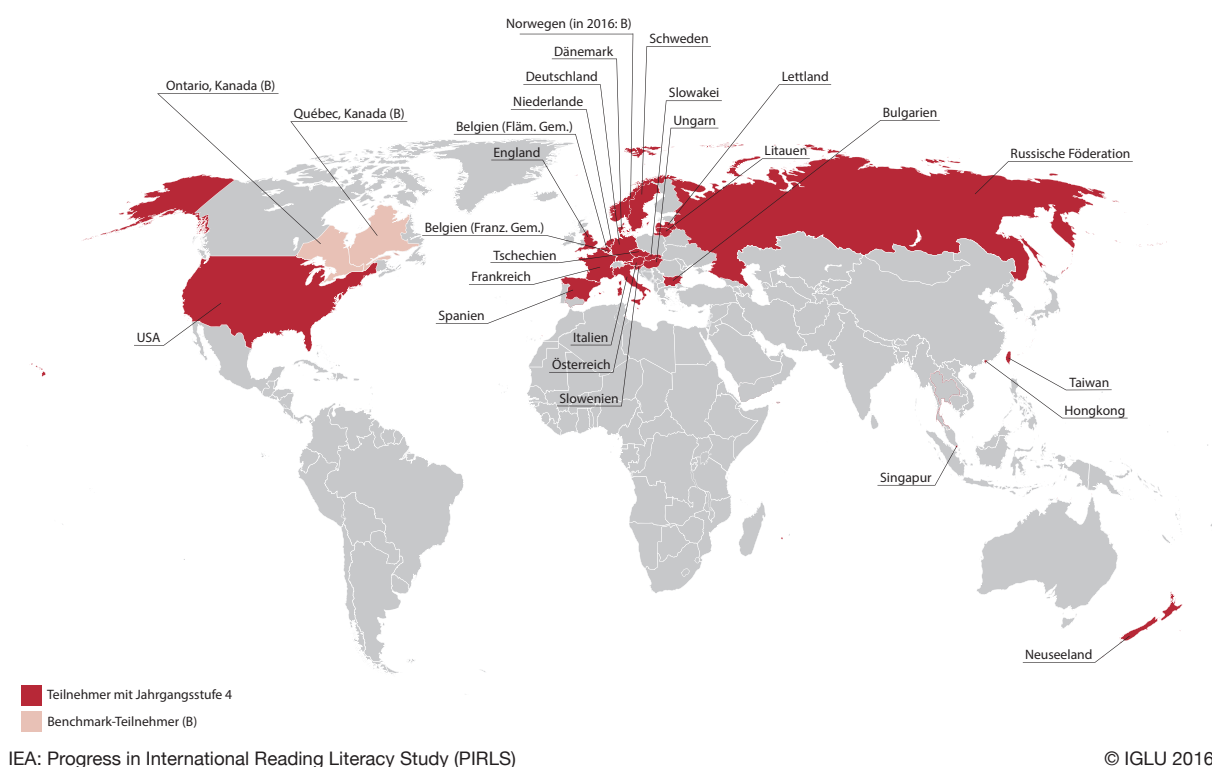
In der Berichterstattung zum Erhebungszyklus von IGLU 2016 werden lediglich bei der Darstellung der Gesamtskala Lesen sämtliche der oben genannten Teilnehmer berücksichtigt. Da der Fokus auf den Leistungen von Schülerinnen und Schülern in Deutschland liegt, werden in den übrigen Ergebnisdarstellungen zu IGLU 2016 ausschließlich Ergebnisse derjenigen Staaten, Regionen und Benchmark-Teilnehmer dokumentiert, auf die eines der folgenden Kriterien zutrifft: (1) Teilnehmer ist Mitglied der EU, (2) Teilnehmer gehört der OECD an, (3) Teilnehmer, hat auf der Gesamtskala Lesen besser oder nicht signifikant von Deutschland verschieden abgeschnitten.

Tabelle 2.1: Vergleichsgruppen in IGLU 2016

VG EU		VG OECD	
Belgien (Fläm. Gem.)	Nordirland	Australien	Neuseeland
Belgien (Franz. Gem.)	Österreich	Belgien (Fläm. Gem.)	Niederlande
Bulgarien	Polen	Belgien (Franz. Gem.)	Nordirland
Dänemark	Portugal	Chile	Norwegen
Deutschland	Schweden	Dänemark	Österreich
England	Slowakei	Deutschland	Polen
Finnland	Slowenien	England	Portugal
Frankreich	Spanien	Finnland	Schweden
Irland	Tschechien	Frankreich	Slowakei
Italien	Ungarn	Irland	Slowenien
Lettland		Israel	Spanien
Litauen		Italien	Tschechien
Malta		Kanada	Ungarn
Niederlande		Lettland	USA
	24		28

4.2.2 Teilnehmende IGLU 2001, 2006, 2011 und 2016

Einige Staaten, Regionen und Benchmark-Teilnehmer haben, wie Deutschland auch, an vergangenen Erhebungszyklen von IGLU teilgenommen. Für die Darstellung von Trends werden in diesem Bericht insgesamt 24 reguläre Teilnehmer und drei Benchmark-Teilnehmer betrachtet (siehe Abbildung 2.3).

Abbildung 2.3: Teilnehmer der Trendvergleiche zu IGLU 2001, 2006, 2011 und 2016

Berücksichtigt werden diejenigen Bildungssysteme, (1) die zusätzlich zu IGLU 2016 an mindestens zwei weiteren IGLU-Zyklen teilgenommen haben und (2) zugleich Mitglied der EU und/oder der OECD sind und/oder auf der Gesamtskala Lesen in IGLU 2016 signifikant besser oder nicht signifikant verschieden von Deutschland abgeschnitten haben. Im Trendvergleich nicht berücksichtigt werden daher die Teilnehmer Georgien, Iran, Katar, Marokko, Trinidad und Tobago sowie Benchmark-Teilnehmer Südafrika (5. Jgst.). Ebenfalls im Trendvergleich nicht berücksichtigt werden Israel und Polen aufgrund von Änderungen in den Übersetzungen der Testinstrumente oder in den Erhebungsbedingungen. Eine Ausnahme ist bei der hier vorgenommenen Trendauswahl zu berücksichtigen: Die Flämische Gemeinschaft in Belgien wird in die Trendgruppe einbezogen, obwohl sie neben IGLU 2016 an lediglich einem weiteren Erhebungszyklus (IGLU 2006) teilgenommen hat. Grund hierfür ist der für den vorliegenden Berichtsband aus deutscher Perspektive relevante Vergleich.

4.2.3 Teilnahmebedingungen der Studie

Bildungssysteme miteinander systematisch zu vergleichen, setzt die Einhaltung strenger methodischer Standards voraus. In diesen Standards ist definiert, mit welchen Schülerinnen und Schülern und unter welchen Bedingungen sich einzelne Bildungssysteme an der Studie beteiligen dürfen (Martin, Mullis & Hooper, 2017). Die Teilnahmebedingungen, die international vorgegeben sind, werden im Folgenden für IGLU 2016 resümiert.

4.2.3.1 Definition der Untersuchungspopulation

Ziel von IGLU ist es, den Ertrag von Bildungssystemen in international vergleichender Perspektive festzustellen. Dazu werden systematisch Leistungen von Schülerinnen und Schülern der vierten Jahrgangsstufe in den Blick genommen. Angestrebt wird, komplette Schülerjahrgänge miteinander zu vergleichen. Für die Definition der zu untersuchenden Schülerpopulation (Zielpopulation) steht dabei – wie in den vorangegangenen Erhebungszyklen – zunächst das Kriterium der formalen Beschulungszeit im Vordergrund:

Für alle Teilnehmerstaaten sind als Zielpopulation alle Schülerinnen und Schüler definiert, die sich nach der *International Standard Classification of Education* im vierten Jahr formaler Beschulung befinden (ISCED Level 1, vgl. UIS, 2015). Dies trifft in den meisten Teilnehmerstaaten auf Schülerinnen und Schüler der vierten Jahrgangsstufe zu. Um ergänzend zu dieser Referenzkategorie eine entwicklungsgerechte Passung der Leistungstests und der Durchführungsbedingungen der Tests sicherzustellen, ist als zusätzliches Kriterium ein Durchschnittsalter von mindestens 9 Jahren und 6 Monaten der definierten Zielpopulation zum Testzeitpunkt bestimmt worden (Martin, Mullis & Foy, 2015).

Ausschöpfung der Zielpopulation

Internationale Vergleichbarkeit setzt unter anderem voraus, dass die definierte Zielpopulation in allen Teilnehmerstaaten möglichst vollständig ausgeschöpft wird. Als erfüllt gilt diese Vorgabe, wenn alle Schülerinnen und Schüler, die unter die Definition der internationalen Zielpopulation fallen, die Möglichkeit erhalten, an der Studie teilnehmen zu können. Ausgehend von der international definierten Zielpopulation muss jeder Teilnehmer die international vorgegebenen Richtlinien zur Stichprobenziehung an die Spezifika des eigenen Systems anpassen. Ein Handbuch zur Stichprobenziehung (Martin et al., 2017) dient dabei als grundlegender Orientierungsrahmen. Unterstützung bei der Umsetzung dieses Vorgehens bieten Expertinnen und Experten von *Statistics Canada* und der *IEA Hamburg*.

Manche nationalen Besonderheiten hindern einzelne Teilnehmer, den internationalen Vorgaben zu entsprechen. Beispielsweise gibt es in fast jedem Schulsystem einzelne Schulen oder Gruppen von Schülerinnen und Schülern, die nicht an der Erhebung teilnehmen können oder aus nationalen Gründen nicht teilnehmen sollen. Sogenannte *Ausschlüsse von der Zielpopulation* können nach eindeutig von der internationalen Studienleitung definierten Kriterien von jedem Teilnehmer vorgenommen werden:

Ausgeschlossen werden können auf *Schülerebene* solche Schülerinnen und Schüler, die aus körperlichen, emotionalen oder geistigen Gründen nicht in der Lage sind, den Test selbstständig zu bearbeiten, oder Schülerinnen und Schüler, die weniger als ein Jahr in der Testsprache unterrichtet wurden und deren Muttersprache nicht die Testsprache ist.

Auf *Schulebene* können jene Schulen ausgeschlossen werden, die in besonders schwer erreichbaren Regionen liegen, die weniger als vier Schülerinnen und Schüler im vierten Jahrgang beschulen (d.h. eine besonders geringe Schülerzahl aufweisen), die in Lehrplan oder Schulstruktur vom nationalen Schulsystem abweichen oder ausschließlich Schülerinnen und Schüler unterrichten, auf die die Schülerebene betreffenden Ausschlusskriterien zutreffen. Ausschlüsse dürfen auf

Schul- und auf Schülerebene insgesamt nicht mehr als 5 Prozent der international definierten Zielpopulation betragen (Joncas & Foy, 2011).

Unter Berücksichtigung dieser Ausschlusskriterien wird für jeden Teilnehmer die *nationale, effektiv erreichte Zielpopulation (national effective target population)* bestimmt. Diese beschreibt, welche Schülerinnen und Schüler einer Population tatsächlich durch die Stichprobe der teilnehmenden Schülerinnen und Schüler repräsentiert werden. Auftretende Abweichungen von der international definierten Zielpopulation in den internationalen Berichten sind auf der Grundlage dieser Beschreibung nachvollziehbar dokumentiert, bedeutsame Abweichungen sind hervorgehoben. Informationen zu den Ausschlüssen der Teilnehmer in IGLU 2016 sowie zu den Teilnehmern der vorangegangenen Erhebungszyklen in 2001, 2006 und 2011 finden sich unter anderem im Anhang A des vorliegenden Bandes.

4.2.3.2 Veränderte Teilnahmemodalitäten in Deutschland seit IGLU 2001

Veränderungen in der Zusammensetzung der Stichproben können unter anderem durch Ausschlüsse von der Zielpopulation begründet (siehe Abschnitt 4.2.3.1), oder auch Folge von Veränderungen der Teilnahmemodalitäten in den einzelnen Staaten und Regionen sein.

Im Vergleich der Erhebungszyklen von IGLU 2001, 2006, 2011 und 2016 zeigt sich, dass je nach Erhebungsjahr und je nach Land unterschiedliche Verpflichtungsgrade der Teilnahme bestehen. So war zum Erhebungszyklus im Jahr 2001 die Teilnahme an IGLU noch in allen Ländern der Bundesrepublik Deutschland freiwillig. Ein Kind konnte jederzeit die Bearbeitung der Tests ablehnen oder abbrechen, selbst wenn von den Eltern eine Einverständniserklärung vorlag (ein solcher Fall wurde jedoch nicht bekannt) (vgl. Lankes et al., 2003, S. 26).

Erst seit IGLU 2006 ist einheitlich definiert, dass die Teilnahme am Leistungstest für alle Schülerinnen und Schüler verpflichtend ist. Die Bearbeitung des Schülerfragebogens unterliegt jedoch bundesweit unterschiedlichen Regelungen (siehe Tabelle 2.2). Während die eine Hälfte der Länder die Befragung der Schülerinnen und Schüler (teil-)verpflichtend durchführt (Berlin, Brandenburg, Bremen, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt und Thüringen), nehmen in der anderen Hälfte der Länder Schülerinnen und Schüler nach Vorlage der schriftlichen Einverständniserklärung ihrer Eltern freiwillig an der Befragung teil (Baden-Württemberg, Bayern, Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Sachsen und Schleswig-Holstein).

Unterschiedliche Verpflichtungsgrade gelten auch in den einzelnen Ländern im Hinblick auf die Teilnahme an der schriftlichen Befragung der Lehrkräfte sowie der Schulleitungen. Die Rechtslage eines Landes kann Lehrkräfte und Schulleitungen zur Teilnahme an der Befragung in IGLU verpflichten, teilverpflichten oder zur freiwilligen Teilnahme lediglich ermuntern. Teilverpflichtungen schließen ein, dass Angaben zur Schule verpflichtend sind, Angaben zu personenbezogenen Daten hingegen freiwillig. Unverändert für alle Länder und über die Erhebungszyklen hinweg ist die freiwillige Teilnahme an der Elternbefragung.

Tabelle 2.2: Länderspezifische Modalitäten der Teilnahme an IGLU 2001, 2006, 2011 und 2016

Länder	Verpflichtungsgrade der Teilnahme an den Erhebungsinstrumenten an öffentlichen Schulen				
	Leistungstest ^A	Schülerfragebogen ^A	Lehrkräftefragebogen ^A	Schulfragebogen ^A	Elternfragebogen
Baden-Württemberg	verpflichtend ¹	freiwillig	freiwillig	freiwillig	freiwillig
Bayern	verpflichtend ¹	freiwillig	freiwillig	freiwillig	freiwillig
Berlin	verpflichtend	teilverpflichtend ²	teilverpflichtend ²	teilverpflichtend ²	freiwillig
Brandenburg	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Bremen	verpflichtend	verpflichtend ³	teilverpflichtend ⁴	teilverpflichtend ⁴	freiwillig
Hamburg	verpflichtend	freiwillig	freiwillig ⁵	teilverpflichtend	freiwillig
Hessen	verpflichtend	verpflichtend	teilverpflichtend	teilverpflichtend	freiwillig
Mecklenburg-Vorpommern	verpflichtend	verpflichtend ⁶	verpflichtend	verpflichtend	freiwillig
Niedersachsen	verpflichtend ¹	verpflichtend	verpflichtend	verpflichtend	freiwillig
Nordrhein-Westfalen	verpflichtend	freiwillig	verpflichtend	verpflichtend	freiwillig
Rheinland-Pfalz	verpflichtend	freiwillig	teilverpflichtend ⁷	teilverpflichtend ⁷	freiwillig
Saarland	verpflichtend	freiwillig	teilverpflichtend ⁸	teilverpflichtend ⁸	freiwillig
Sachsen	verpflichtend	freiwillig	freiwillig	freiwillig ⁹	freiwillig
Sachsen-Anhalt	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Schleswig-Holstein	verpflichtend	freiwillig	freiwillig	freiwillig	freiwillig
Thüringen	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig

A = Im Jahr 2001 war die Teilnahme an IGLU in allen Ländern der Bundesrepublik Deutschland freiwillig.

1 = In Baden-Württemberg, Bayern und Niedersachsen war im Jahr 2006 die Teilnahme am Test freiwillig.

2 = In Berlin war in den Jahren 2006 und 2011 die Teilnahme am Schüler-, Lehrkräfte- und Schulfragebogen verpflichtend.

3 = In Bremen war in den Jahren 2006 und 2011 die Teilnahme am Schülerfragebogen freiwillig.

4 = In Bremen war im Jahr 2006 die Teilnahme am Lehrer- und Schulfragebogen verpflichtend.

5 = In Hamburg war in den Jahren 2006 und 2011 die Teilnahme am Lehrkräftefragebogen teilverpflichtend.

6 = In Mecklenburg-Vorpommern war im Jahr 2006 die Teilnahme am Schülerfragebogen freiwillig.

7 = In Rheinland-Pfalz war im Jahr 2006 die Teilnahme am Lehrer- und Schulfragebogen verpflichtend.

8 = Im Saarland war in den Jahren 2006 und 2011 die Teilnahme am Lehrkräfte- und Schulfragebogen verpflichtend.

9 = In Sachsen war im Jahr 2006 die Teilnahme am Schulfragebogen teilverpflichtend.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Für Schulen in privater Trägerschaft gilt mehrheitlich, dass die Teilnahme an der gesamten Studie freiwillig ist. In manchen Ländern wie Baden-Württemberg, Bayern oder Hamburg ist für diese Schulen lediglich die Teilnahme an den Leistungstests verpflichtend.

In Tabelle 2.2 sind die Verpflichtungsgrade der länderspezifischen Teilnahme an öffentlichen Schulen für die Erhebungszyklen von IGLU in den Jahren 2001, 2006, 2011 und 2016 zusammenfassend dargestellt. Die Tabelle dokumentiert die zum Zeitpunkt für 2016 geltenden Teilnahmemodalitäten pro Land und pro Erhebungsinstrument. Abweichungen in den Teilnahmebedingungen im Vergleich der Erhebungszyklen 2001, 2006 und 2011 sind durch Fußnoten mit entsprechenden Erläuterungen ausgewiesen. Insbesondere für Trendanalysen sind diese Informationen relevant, weil sie Hintergründe für mögliche veränderte Stichprobenzusammensetzungen und Teilnahmequoten der verschiedenen befragten Akteursgruppen liefern.

4.3 Verfahren und Kriterien der Stichprobenziehung

In IGLU erfolgt die Stichprobenziehung nach einem zweistufigen stratifizierten Clusterdesign (Martin et al., 2017; Joncas & Foy, 2011). Es wird eine zufällige Auswahl aus allen Schulen der national effektiv erreichten Zielpopulation gezogen und anschließend eine oder mehrere Klassen in den ausgewählten Schulen ermittelt. Die Auswahl der Schulen erfolgt in zwei Stufen:

In der *ersten Stufe* wird aus einer Liste (*Sampling Frame*), auf der alle Schulen verzeichnet sind, die Teil der national effektiv erreichten Zielpopulation sind, eine Zufallsstichprobe von Schulen gezogen. Die Mindestanzahl beträgt nach den Vorgaben der internationalen Studienleitung 150 Schulen. Da Schulen mit einer größeren Anzahl an Schülerinnen und Schülern in der vierten Jahrgangsstufe auch einen höheren Anteil an der Population haben, wird die Information zur Größe der Jahrgangsstufe 4 bei der Stichprobenziehung berücksichtigt. Zur Steigerung der Effizienz der Stichprobe und der Präzision der Ergebnisse wird die Methode der Stratifizierung angewendet: Schulen, die als Teil der national effektiv erreichten Zielpopulation gelten, werden nach bestimmten Merkmalen (*Strata*) kategorisiert (*geschichtet*). Mögliche *Strata* sind die geographische Lage (z. B. Länder der Bundesrepublik Deutschland) oder der Schultyp (z. B. Förder- vs. Regelschule, privat vs. öffentlich). Diese zur Verfügung stehenden Informationen über die Schulen werden bei der Stichprobenziehung berücksichtigt, um sicherzustellen, dass sich die tatsächliche Verteilung der Schulen innerhalb der *Strata* auch in der Stichprobe widerspiegelt und spezifische Gruppen der Population adäquat repräsentiert sind.

Innerhalb der in der ersten Stufe ausgewählten Schulen erfolgt in der *zweiten Stufe* die Ziehung der Klassenstichproben. Jede Klasse hat dieselbe Wahrscheinlichkeit, in die Stichprobe zu gelangen. Mit der Auswahl der Klassen sind automatisch auch die Schülerinnen und Schüler bestimmt, da – abgesehen von den oben beschriebenen Ausnahmen – alle Schülerinnen und Schüler einer Klasse am Test teilnehmen sollen. Nach internationalen Vorgaben sind mindestens 4000 Schülerinnen und Schüler auszuwählen (ebd.).

4.3.1 Schul- und Schülerteilnahmequoten

In jeder Studie ist mit Stichprobenausfällen zu rechnen. Um Datenverzerrungen aufgrund nicht teilnehmender Schulen oder Personengruppen in IGLU gering zu halten, sind seitens der internationalen Studienleitung Vorgaben für minimale Beteiligungsquoten definiert (Martin et al., 2017; Joncas & Foy, 2011). Teilnehmende Staaten und Regionen, die die vorgeschriebenen Rücklaufquoten nicht erreichen, werden in der internationalen Berichterstattung bei weniger gravierenden Abweichungen mit einer Fußnote und entsprechender Erklärung in den Ländervergleich einbezogen. Erweisen sich die Rücklaufquoten als zu gering, werden die Ergebnisse gesondert berichtet.

Auf *Schulebene* liegt die minimal zu erreichende Beteiligungsquote bei 85 Prozent der ursprünglich ausgewählten Schulen. Fehlen weniger als 15 Prozent der ursprünglich ausgewählten Schulen, können Schulen, die nicht teilgenommen haben, durch Ersatzschulen im Nachrückverfahren angesprochen werden. Dabei handelt es sich um Schulen, die bereits im Rahmen der Stichprobenziehung ausgewählt wurden. Dieses Vorgehen ermöglicht, die Anforderungen an die Stichprobengröße insgesamt erfüllen zu können. Ersatzschulen müssen den ur-

sprünglich gezogenen Schulen in den für die Stratifizierung bestimmten Merkmalen ähneln.

Auf *Klassenebene* liegt die minimal zu erreichende Beteiligungsquote bei 95 Prozent der ursprünglich ausgewählten Klassen. Für nicht teilnehmende Klassen müssen Klassen in Ersatzschulen gefunden werden. Es besteht in diesen Fällen nicht die Möglichkeit, diese durch Parallelklassen zu ersetzen. Auch für die Befragung der Lehrerinnen und Lehrer ist für die spätere Berichtslegung eine Mindestbeteiligung von 85 Prozent vorgegeben. Wie im Falle der Beteiligungsquote von Klassen, so dürfen auch Lehrkräfte den internationalen Vorgaben zufolge nicht ersetzt werden (ebd.).

Für die *Schülerebene* ist eine Beteiligungsquote von mindestens 85 Prozent der Schülerinnen und Schüler festgelegt. Zusätzlich gilt, dass die Beteiligung in den Klassen nicht unter 50 Prozent sinken darf. Um Rücklaufquoten im geforderten Maße zu erreichen, wird an Schulen ein Nachttest durchgeführt, an denen am Testtag mehr als 10 Prozent der Schülerinnen und Schüler fehlen. Wird trotz Nachttests die Mindestrücklaufquote nicht erreicht, wird die Klasse und damit auch die ganze Schule von der Studie ausgeschlossen.

Alternativ zu den genannten Beteiligungsquoten ist zu gewährleisten, dass die kombinierte Schul- und Schülerteilnahmequote bei mindestens 75 Prozent liegt, das heißt, der durch Ausschluss auf Schul-, Klassen- und Schüler Ebene bedingte Anteil an nicht teilnehmenden Schülerinnen und Schülern darf 25 Prozent nicht überschreiten.

Die Bewertung der Qualität der Stichprobe auf Schüler Ebene erfolgt auch danach, ob die Teilnahme an der Leistungsmessung zu verwertbaren Ergebnissen geführt hat. Als Interpretationsgrundlage dienen dabei die Kompetenzmittelwerte, anhand derer aufzuzeigen ist, dass für mindestens 85 Prozent der Schülerinnen und Schüler innerhalb der Stichprobe zuverlässige Leistungswerte geschätzt werden können. Die Schätzung des Leistungswertes gilt als zuverlässig, wenn eine Schülerin oder ein Schüler mehr Aufgaben richtig bearbeitet hat als durch Raten richtig zu lösen wären.

4.3.2 Besonderheiten der Stichproben im internationalen Vergleich

Die Tabelle A.4 in Anhang A dieses Bandes gibt einen Überblick über zentrale Kennwerte der Stichproben der an IGLU 2016 teilnehmenden Bildungssysteme. Zusätzlich sind hier auch die Angaben für die Teilnehmer an PIRLS *Literacy* aufgenommen. Die Tabelle dokumentiert Teilnahmen an den Erhebungszyklen von IGLU sowie Besonderheiten bezüglich der nationalen Zielpopulationen, der Schul- und Schülerteilnahmequoten sowie der Leistungsmessungen. Norwegen als Teilnehmer mit Jahrgangsstufe 5, die Benchmark-Teilnehmer sowie die Teilnehmer an PIRLS *Literacy* werden gesondert aufgeführt (siehe Abschnitt 4.2.1). Im Anhang A sind zudem in Tabelle A.1 diese Besonderheiten entsprechend für IGLU 2001, in Tabelle A.2 für IGLU 2006 und in Tabelle A.3 für IGLU 2011 dargestellt. In der ersten Spalte dieser Tabellen wird durch Ziffernfußnoten auf Besonderheiten der Stichproben verwiesen, zudem wird auf Besonderheiten in einzelnen Ländern durch an die Ländernamen angefügte Buchstabenfußnoten (Spalte 2) hingewiesen. Erläuterungen geben Aufschluss über mögliche Einschränkungen, die sich für eine international vergleichende Interpretation der Ergebnisse ergeben (siehe Abschnitt 4.3.4).

In den Spalten 3, 4 und 5 von Tabelle A.4 ist gekennzeichnet, welche Teilnehmer und Benchmark-Teilnehmer von IGLU 2016 auch in 2011, in 2006 beziehungsweise in 2001 beteiligt waren. Die Tabellenspalten 6 bis 9 verweisen auf Besonderheiten bezüglich der Definition und Ausschöpfung der nationalen Zielpopulation, die Tabellenspalten 10 bis 14 auf Besonderheiten bezüglich der *Schul- und Schülerteilnahmequoten*. In Tabellenspalte 15 sind *Besonderheiten bezüglich der Leistungsmessung* dokumentiert.

Das Durchschnittsalter der getesteten Schülerinnen und Schüler ist der sechsten Tabellenspalte zu entnehmen. England, Malta, Neuseeland sowie Trinidad und Tobago wählten als äquivalente nationale Zielpopulation gemäß den Kriterien ‚formale Beschulungszeit‘ und ‚Durchschnittsalter des Schülerjahrgangs‘ eine höhere Jahrgangsstufe.

In Spalte 8 wird der *Ausschöpfungsgrad der nationalen Zielpopulation in Prozent bezogen auf die internationale Vorgabe* (100%) illustriert. Hier zeigt sich für die Teilnehmer Georgien und Kanada, dass die internationale Zielpopulation nicht vollständig durch die nationale Zielpopulation abgedeckt wird. Für Georgien ist der geringere Ausschöpfungsgrad dadurch bedingt, dass hier nur Schülerinnen und Schüler getestet wurden, die in Georgisch und Aserbaidschanisch unterrichtet werden. In Kanada wurden nur Schülerinnen und Schüler aus den Provinzen Alberta, Britisch Kolumbien, Manitoba, Neubraunschweig, Neufundland, Ontario, Québec und Saskatchewan getestet.

In Spalte 9 sind die *Ausschlüsse von der Zielpopulation* als Gesamtausschlussquote in Prozent aufgeführt. Hier zeigt sich, dass die internationale Vorgabe von maximal 5 Prozent von 13 Teilnehmern und drei Benchmark-Teilnehmern sowie von Dänemark (3. Jahrgangsstufe, PIRLS *Literacy*) überschritten wird. Deutschland gehört zu den Teilnehmern, deren Gesamtausschlussquote auf der Ebene der nationalen Zielpopulation unter 5 Prozent liegt (für Details zu Stichprobe und Beteiligungsquoten in Deutschland siehe Abschnitt 7).

Tabellenspalten 10 bis 14 sind die *Schulteilnahmequoten* (ohne und mit Ersatzschulen), die *Schülerteilnahmequoten* sowie die *Gesamtteilnahmequoten* (ohne und mit Ersatzschulen) zu entnehmen. Während die meisten Teilnehmer und Benchmark-Teilnehmer die internationalen Vorgaben erfüllen, zeigt sich für Belgien (Flämische Gemeinschaft), Hongkong, Kanada, die Niederlande, Nordirland, die USA und Südafrika (5. Jahrgangsstufe), dass sie eine *Schulteilnahmequote* von 85 Prozent nur unter Hinzunahme von Ersatzschulen erreichen. Québec gelingt dies auch unter Berücksichtigung der Ersatzschulen nicht. Eine *Schülerteilnahmequote* von 85 Prozent erreichen alle an IGLU 2016 teilnehmenden Bildungssysteme. Eine *kombinierte Schüler- und Schulgesamtteilnahmequote* von 75 Prozent erreichen Hongkong, die Niederlande und die USA hingegen nur unter Hinzunahme von Ersatzschulen. Québec liegt auch mit Ersatzschulen noch unterhalb dieser Vorgabe.

Tabellenspalte 15 dokumentiert den Anteil der Schülerinnen und Schüler ohne skalierbare Leistungswerte im Lesetest bei IGLU 2016. Da diese Information erst seit der Erhebung im Jahr 2011 für IGLU berichtet wird, finden sich entsprechende Angaben nicht in den Tabellen A.1 und A.2. Wie Tabelle A.4 zu entnehmen ist, gibt es bei IGLU 2016 keinen Teilnehmer oder Benchmark-Teilnehmer, der die internationale Vorgabe von maximal 15 Prozent an Schülerinnen und Schülern ohne skalierbare Leistungswerte überschreitet.

4.3.3 Bedeutsamkeit der Ausschlussquote für den internationalen Vergleich

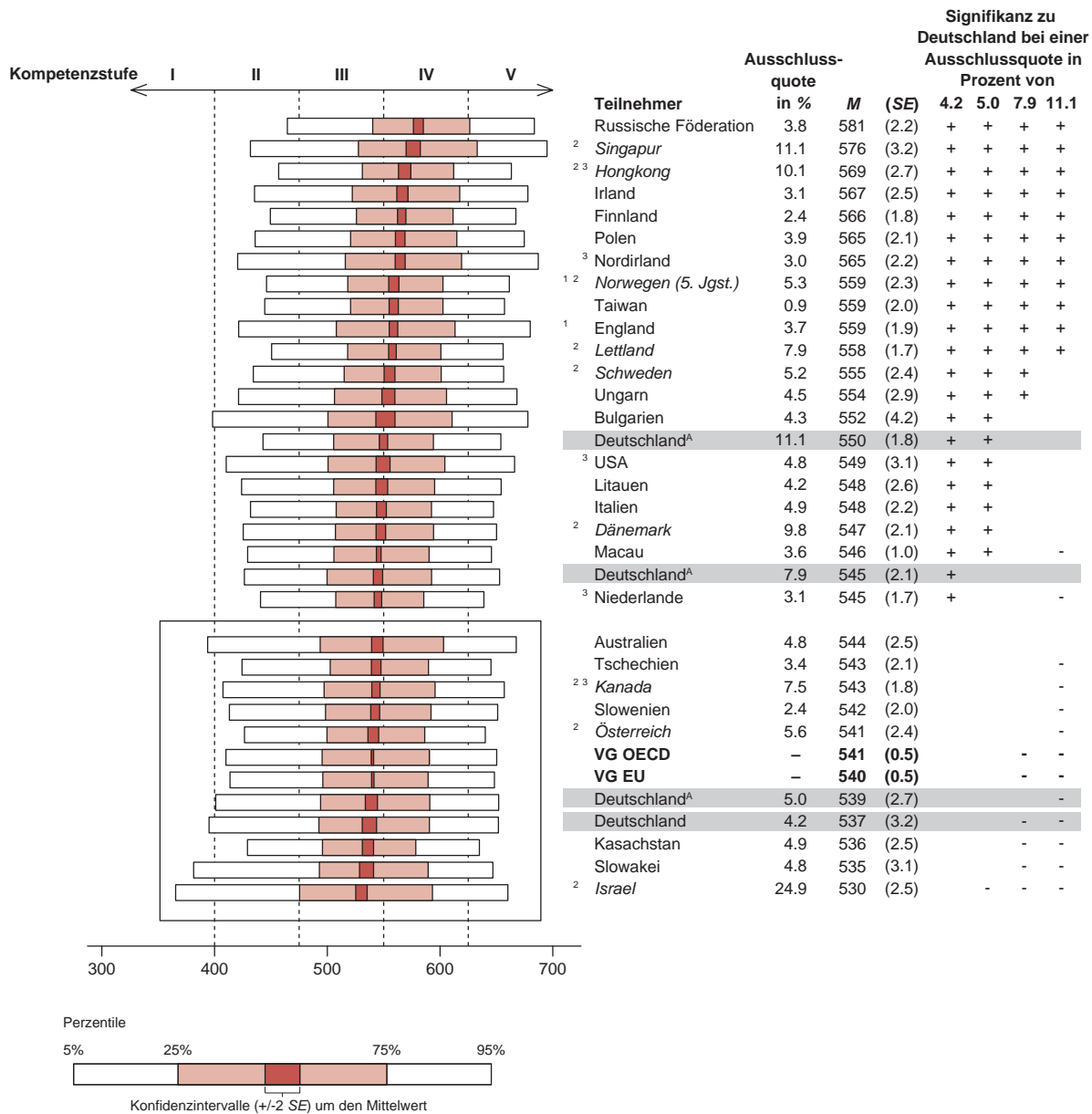
Die *Ausschlüsse von der Zielpopulation* liegen für Deutschland mit einer Gesamtausschlussquote von 4.2 Prozent im Toleranzbereich, das heißt unterhalb von 5 Prozent (siehe Abschnitt 4.3.2). Bei manchen Teilnehmern ist sie deutlich geringer (z. B. in Taiwan mit 0.9% oder in Finnland mit 2.4%), andere Teilnehmer wie Singapur (11.1%), Hongkong (10.1%), Dänemark (9.8%), Lettland (7.9%) oder Österreich (5.6%) erreichen Ausschlussquoten, die deutlich über der internationalen Vorgabe liegen.

In Deutschland ist die Gesamtausschlussquote im Vergleich zu IGLU 2011 (1.9%) oder auch IGLU 2006 (0.7%) deutlich gestiegen. Von den 4.2 Prozent sind 1.4 Prozentpunkte der Ausschlüsse auf Schulebene zu begründen, 2.8 Prozentpunkte auf Schülerebene (2% Schülerinnen und Schüler, die weniger als ein Jahr in der Testsprache unterrichtet wurden und deren Muttersprache nicht die Testsprache ist, und 0.8 Prozent Kinder, die aus Sicht der Lehrkräfte körperlich, emotional oder aus geistigen Gründen nicht in der Lage waren, den Test selbstständig zu bearbeiten). Die erhöhte Gesamtausschlussquote in 2016 ist damit mit dem erhöhten Anteil von Ausschlussgründen auf Schülerebene zu begründen und hängt mit den, im Zuge von Inklusion und der seit 2015 zu beobachtenden, ansteigenden Zuwanderung nach Deutschland, gestellten Herausforderungen an das Bildungswesen zusammen.

Unterstellte man, dass Ausschlüsse von der Zielpopulation in erster Linie Schülerinnen und Schüler mit eher geringen Kompetenzen betreffen, so wäre zu fragen, welche Mittelwerte die Viertklässlerinnen und Viertklässler in Deutschland bei vergleichbaren Ausschlussquoten erzielt hätten. Die Abbildung 2.4 zeigt Antworten auf diese hypothetische Frage. Dargestellt sind verschiedene Verortungen Deutschlands auf der Gesamtskala Lesen (siehe Kapitel 3, Abbildung 3.6 in diesem Band), die sich ergäben, wenn für Deutschland Ausschlussraten zugrunde lägen, die vergleichbar sind mit denen ausgewählter Teilnehmer, die die Vorgabe von 5 Prozent überschreiten. Als Referenzpunkte sind in der Abbildung die realen Kompetenzmittelwerte für Deutschland und für diejenigen Teilnehmerstaaten angegeben, die auf der Gesamtskala Lesen signifikant besser als oder nicht signifikant verschieden von Deutschland abschneiden. Der rechten Seite der Abbildung ist zu entnehmen, ob die Leistungsmittelwerte der dargestellten Teilnehmer unter Berücksichtigung der Ausschlussquoten signifikant besser oder schlechter abschneiden als die Leistungsmittelwerte für Deutschland. Symbolisiert sind signifikant höhere Werte durch ein Plus (+), signifikant niedrigere Werte durch ein Minus (-).

Abbildung 2.4 veranschaulicht, dass eine Erhöhung der Ausschlussquoten für Deutschland im Rahmen der internationalen Vorgaben (max. 5%) keinen bedeutsamen Einfluss auf das Abschneiden der Viertklässlerinnen und Viertklässler in Deutschland im internationalen Vergleich gehabt hätte. Es ist im Umkehrschluss auch zu vermuten, dass die Leistungsmittelwerte der Teilnehmer Singapur, Hongkong, Dänemark oder Lettland bei geringeren Ausschlussquoten unterhalb der für diese Teilnehmer gefundenen Leistungsmittelwerte lägen. Ausschlussquoten in Höhe von 7.9 Prozent (wie in Lettland) oder 11.1 Prozent (wie in Singapur) würden zwar den Leistungsmittelwert der Schülerinnen und Schüler in Deutschland und damit auch die Platzierung Deutschlands im Ländervergleich signifikant positiv beeinflussen. Insgesamt aber bleibt festzuhalten, dass die Ausschlussquoten anderer Teilnehmer nur begrenzt als Begründung für das Abschneiden Deutschlands im internationalen Vergleich herangezogen

Abbildung 2.4: Veränderung des Mittelwerts von Viertklässlerinnen und Viertklässlern in Deutschland beim Vergleich der Leistung auf der Gesamtskala Lesen in Abhängigkeit von der Ausschlussquote



□ Nicht statistisch signifikant vom deutschen Mittelwert abweichende Staaten ($p > .05$).
 Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse gegangen werden muss.
 + = Statistisch signifikant positive Abweichungen ($p < .05$) zu dem für Deutschland ermittelten Leistungsmittelwert unter Berücksichtigung der in der Spaltenüberschrift angegebenen Ausschlussquote.
 - = Statistisch signifikant negative Abweichungen ($p < .05$) zu dem für Deutschland ermittelten Leistungsmittelwert unter Berücksichtigung der in der Spaltenüberschrift angegebenen Ausschlussquote.
 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 A = Berechnung für Deutschland nach einem höheren Ausschluss von Schülerinnen und Schülern mit geringen Kompetenzwerten.

werden können. Denn sogar unter Ausschluss von 11.1 Prozent der leistungsschwächsten Viertklässlerinnen und Viertklässler aus der deutschen Stichprobe würden etwa Schülerinnen und Schüler in der Russischen Föderation, Singapur, Hongkong, Irland, Finnland oder auch in Polen und Nordirland weiterhin signifikant bessere Leistungsergebnisse im Lesen erzielen als Viertklässlerinnen und Viertklässler in Deutschland.

4.3.4 IGLU 2016 Fußnotensystem zur Klassifikation von Teilnahmebedingungen

Die Besonderheiten der Teilnahmebedingungen sind in den Tabellen A.1, A.2, A.3 und A.4 dokumentiert (siehe Anhang A in diesem Band). Im vorliegenden Berichtsband wurde in Anlehnung an die internationale Berichterstattung für die Darstellung der Ergebnisse ein differenziertes Fußnotensystem entwickelt, das auf die vier Erhebungszyklen von IGLU angewendet werden kann. Die Fußnoten benennen Besonderheiten der Stichproben der einzelnen Teilnehmer. Mit dieser Übersicht ist für eine international-vergleichende Interpretation von Ergebnissen eine Grundlage gegeben, Einschränkungen der Vergleichbarkeit systematisch zu betrachten, die aus Besonderheiten der nationalen Stichproben resultieren. Die nachfolgend aufgeführten Fußnoten werden analog zur Berichtslegung zu IGLU 2011 (Tarelli, Wendt, Bos & Zylowski, 2012) zur Kennzeichnung von Besonderheiten nationaler Stichproben verwendet:

- | | | |
|---|---|---|
| 1 | = | Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe. |
| 2 | = | Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben. |
| 3 | = | Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben. |

Die Fußnoten 4, 5 und 6, die im Rahmen der Berichterstattung zu IGLU 2011 (vgl. ebd.), TIMSS 2011 (Wendt, Tarelli, Bos, Frey & Vennemann, 2012) beziehungsweise TIMSS 2015 (Wendt, Bos, Kasper, Walzebug, Goy & Jusufi, 2016) definiert wurden, kommen in der Berichtslegung zu IGLU 2016 nicht vor. Um eine Konsistenz der Fußnotennummerierung in der IGLU- und TIMSS-Berichterstattung zu gewährleisten, werden diese Nummern bei IGLU 2016 nicht anderweitig vergeben.

Für IGLU 2016 wurden im Vergleich zu IGLU 2011 und analog zur Berichterstattung bei TIMSS 2015 zwei weitere Fußnoten (7 und 8) definiert, die zum einen die Teilnahme an PIRLS *Literacy* (Fußnote 7), zum anderen Besonderheiten im Hinblick auf Trendvergleiche mit den Zyklen IGLU 2001, IGLU 2006 und IGLU 2011 (Fußnote 8) betreffen (siehe Anhang A.4 in diesem Band).

Kursivschreibung der Ländernamen und Kennzahlen

Ebenfalls analog zur Berichtslegung von IGLU 2011 (Tarelli et al., 2012) wird neben Fußnoten durch Kursivschreibung der Staatennamen auf eine eingeschränkte Vergleichbarkeit hingewiesen. Bei IGLU 2016 erfolgte dies, wenn für Teilnehmer mindestens eins der beiden Kriterien zutrifft:

- die Gesamtausschlussquote liegt über der internationalen Vorgabe von 5 Prozent (siehe Tabelle A.4 in diesem Band; Spalte 9; *Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) in %*),
- die Schüler- und Schulgesamtteilnahmequote (mit Ersatzschulen) liegt unter 75 Prozent (siehe Tabelle A.4 in diesem Band; Spalte 14; *Gesamtteilnahmequote in %*),

Kursive Formatierungen der Namen der Teilnehmer und Benchmark-Teilnehmer in den Abbildungen und Tabellen in diesem Band werden durch folgende Fußnote erläutert: „Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss“. In Trenddarstellungen weisen entsprechende Kursivsetzungen darauf hin, dass in den betreffenden Bildungssystemen Besonderheiten in den Erhebungsbedingungen vorliegen und Interpretationen von Ergebnissen im Vergleich der Studienzyklen nur unter deren Berücksichtigung erfolgen sollten (siehe hierzu auch die Erläuterungen zu Fußnote 8 sowie die Tabellen A.1 bis A.4 im Anhang A in diesem Band).

5 Entwicklung und Charakteristika der Instrumente

In IGLU werden zur Erfassung der Lesekompetenzen Leistungstests eingesetzt. Die für das schulische Lernen relevanten Hintergrundinformationen werden mit Hilfe von Fragebögen auf System-, Schul-, Lehrer-, Schüler- und Elternebene umfassend erhoben.

5.1 Leistungstests

Die Leistungstests, die in IGLU eingesetzt werden, dienen der Ermittlung von Leistungsständen der Schülerinnen und Schüler im Bereich des Leseverstehens. Die so ermittelten Schülerleistungen erlauben, Aussagen im internationalen Vergleich über Leistungsunterschiede zwischen gesamten Schülerpopulationen und zwischen spezifischen Schülergruppen für die übergreifenden Kompetenzskalen und die Subdomänen zu treffen (siehe Abschnitt 4.1). Die Leistungen sollen für die übergreifende Gesamtskala zur Messung der Lesekompetenz und für die dem Rahmenkonzept entsprechenden Subskalen auswertbar sein (siehe Kapitel 3 in diesem Band).

Die in IGLU eingesetzten Leistungstests bestehen aus verschiedenen Erzähl- und Sachtexten und den dazu entwickelten Verständnisaufgaben (Martin et al., 2015; Mullis et al., 2015). In der Regel werden ungefähr 60 Prozent der Aufgabenblöcke zum Zwecke der Verlinkung mit den zurückliegenden Erhebungszyklen aus den vorangegangenen Erhebungszyklen übernommen. Dabei stammen etwa 75 Prozent aus den letzten und knapp 25 Prozent aus dem jeweils vorletzten Zyklus. Etwa 40 Prozent der Blöcke werden für einen Erhebungszyklus neu entwickelt. Die Auswahl der Texte und die Entwicklung der Testaufgaben erfolgen in einem kooperativen Prozess, an dem Vertreterinnen und Vertreter der Teilnehmerstaaten, die internationale Studienleitung und eine Gruppe aus ausgewählten internationalen Expertinnen und Experten (der sogenannten *Reading Development Group*) beteiligt sind. Aufgaben können aus jedem Teilnehmerstaat vorgeschlagen und eingereicht werden. Nach der Einreichung

erfolgt eine umfassende Begutachtung. Die *Reading Development Group* wählt aus den Vorschlägen jenes Textmaterial aus, das authentischen Leseerfahrungen von Viertklässlerinnen und Viertklässlern in unterschiedlichen Kulturen entspricht und den Anforderungen der Messung von Lesekompetenzen im Hinblick auf den Umfang und den Inhalt gerecht wird. Geprüft wird das Material unter anderem hinsichtlich der Angemessenheit der Themen, der Informationsdichte und des Niveaus darin auftretender linguistischer Charakteristika. Besonderes Augenmerk wird zudem auf Aspekte geschlechtsspezifischer, kultureller und religiöser Fairness der Texte gelegt. Texte, die den Prozess der Begutachtung erfolgreich durchlaufen haben, werden der Gruppe der *National Research Coordinators* vorgestellt. Diese Gruppe prüft erneut die Inhalte für die teilnehmenden Staaten und Regionen. Nach Bestimmung der Texte werden – ebenfalls in Kooperation mit den Teilnehmerstaaten – Testaufgaben zu den Texten konzipiert. Verbindlich definierte Kriterien, die in einem Leitfaden formuliert sind, gewährleisten, dass die Aufgaben die in der Rahmenkonzeption festgelegten Inhalte und Anforderungen angemessen repräsentieren (Martin et al., 2017; Mullis & Prendergast, 2017).

Diesem Prozess folgen im Rahmen des Feldtests ein Jahr vor der Haupterhebung die Erprobung der Textauswahl und die Überprüfung der empirischen Qualität der Aufgaben. Von insgesamt 12 im Feldtest neu eingesetzten Texten wurden sechs Texte in der Hauptuntersuchung eingesetzt (Mullis & Prendergast, 2017). Ausgewählt wurden diejenigen Texte, deren Testaufgaben für die Schülerinnen und Schüler aller Teilnehmerstaaten vergleichbare Messeigenschaften aufweisen konnten.

5.1.1 Charakteristika der Testaufgaben

In IGLU 2016 wurden 12 Lesetexte und entsprechende Aufgaben zu den Texten eingesetzt, davon sechs als informierende und sechs als erzählende Lesetexte. Vier von diesen 12 Texten wurden neu ergänzt, die anderen acht stammen aus den Erhebungszyklen 2001, 2006 und 2011 (Mullis & Prendergast, 2017). Zwei der vier neu in IGLU 2016 eingesetzten Texte wurden aus der ergänzenden Studienkomponente PIRLS *Literacy* (siehe Abschnitt 4.2.1 in diesem Kapitel) in den Aufgabenpool übernommen, um eine Vergleichbarkeit der Testleistungen in IGLU und PIRLS *Literacy* herzustellen (ebd.). Dabei handelt es sich um einen Sach- und einen Erzähltext in leichterer Sprache (siehe auch Kapitel 3, Tabelle 3.1 in diesem Band). Durch den Einsatz von in den Erhebungszyklen 2001, 2006 oder 2011 verwendeter und neuer Texte findet eine Verknüpfung der vergangenen Erhebungszyklen mit dem aktuellen statt, so dass eine Einschätzung der Veränderung in den Leistungsergebnissen im Trend ermöglicht ist.

Alle Lesetexte sind kindgerecht aufbereitet, ansprechend illustriert und sprachlich klar formuliert (siehe die beiden Beispieltex te in Kapitel 3, Abschnitte 4.2.1 und 4.2.2 in diesem Band). An jeden Text schließen sich 12 bis 17 Aufgaben an, die in einem geschlossenen (*Multiple Choice*) oder offenen Antwortformat erstellt sind und auf unterschiedlichen Schwierigkeitsniveaus verschiedene Aspekte des Leseverständnisses erfragen. Bei geschlossenen Formaten wählen Schülerinnen und Schüler aus vier vorgegebenen Antworten die richtige Lösung aus. Bei offenen Formaten ist von den Schülerinnen und Schülern verlangt, ihre Antworten in einem Textfeld frei zu formulieren. Die Auswertung der Aufgaben erfolgt nach international einheitlichen Auswertungskriterien.

5.1.2 Rotation der Testaufgaben

Jede Schülerin beziehungsweise jeder Schüler bearbeitet während der Testsitzung ein Testheft. Darin enthalten sind zwei Lesetexte mit dazugehörigen Aufgaben.

Wie in den meisten internationalen Leistungsstudien, liegt auch der Testkonzeption in IGLU ein multiples Matrixdesign zugrunde, das eine Rotation der 12 Texte (6 Erzähl- und 6 Sachtexte) über 16 unterschiedliche Testhefte vorsieht (siehe Tabelle 2.3). Das Grundprinzip dieses Designs sieht vor, dass nicht allen Schülerinnen und Schülern sämtliche Aufgaben vorgelegt werden, sondern jede Schülerin und jeder Schüler lediglich einen Teil der Aufgaben bearbeitet. Eine Schülerin beziehungsweise ein Schüler bearbeitet jeweils zwei ausgewählte Lesetexte (einen Erzähl- und einen Sachtext), die in einem Testheft arrangiert sind.

Die in IGLU gezogene Stichprobe ist ausreichend groß, so dass mit Hilfe statistischer Methoden Leistungswerte dank des multiplen Matrixdesigns für die Population präzise geschätzt werden können. Dieses Testdesign bietet den Vorteil, auch die Leistung von Schülerinnen oder Schülern vergleichen zu können, die unterschiedliche Testteile bearbeitet haben.

Tabelle 2.3: Testheftdesign in IGLU 2016

		Testheft															Geschichtenheft
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Teil 1	Inf 1	Lit 2	Inf 2	Lit 2	Inf 2	Lit 4	Inf 4	Lit 5	Inf 5	Lit 5	Lit 2	Inf 2	Inf 4	Lit 5	Inf 5	Lit 6	
Teil 2	Lit 1	Inf 1	Lit 2	Inf 3	Lit 3	Inf 3	Lit 4	Inf 4	Lit 1	Inf 5	Inf 3	Lit 4	Lit 1	Inf 1	Lit 3	Inf 6	

Lesetexte aus PIRLS *Literacy*.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

5.1.3 Freigegebene Aufgaben

Informationen zu der Art und den Inhalten der in IGLU eingesetzten Lesetexte und Aufgaben werden nach der Berichtslegung der Öffentlichkeit zugänglich gemacht. Die internationale Studienleitung wird im Frühjahr 2018 Lesetexte und Aufgabenblöcke der IGLU 2016-Testhefte veröffentlichen. Die in diesem Bericht verwendeten Aufgabenbeispiele entstammen dieser Auswahl (siehe Kapitel 3 in diesem Band). In den letzten Jahren hat die IEA ihr Verfahren der Nutzung von Testaufgaben verändert. Eine Nutzung von Aufgaben muss bei der IEA beantragt werden.¹

¹ Weiterführende Informationen sind unter folgendem Link verfügbar: <http://www.iea.nl/permissions.html>

5.1.4 Ergänzende Aufgaben zum Leseverständnis auf Wort- und Satzebene

In Deutschland wurden im Rahmen von IGLU 2016 ergänzend zu den international vorgegebenen IGLU-Tests zwei Testkomponenten des *Leseverständnistests für Erst- bis Siebtklässler, Version II* (ELFE II) eingesetzt (Lenhard, Lenhard & Schneider, 2017). Der ELFE II dient der Erfassung der Leseflüssigkeit, der Lesegenauigkeit und des Leseverständnisses auf Wort-, Satz- und Textebene und differenziert nach dem Kompetenzniveau von Schülerinnen und Schülern vom Ende der ersten bis zum Beginn der siebten Klassenstufe. Es handelt sich um einen standardisierten Test, der sich bewährt hat, um zum Beispiel Defizite im Schriftspracherwerb frühzeitig zu erkennen und gegebenenfalls unterstützende Maßnahmen einleiten zu können.

In IGLU 2016 wurden die ELFE II-Untertests zum Wort- und Satzverständnis eingesetzt, nicht aber zum Textverständnis. Beide Tests wurden im Anschluss an den IGLU-Test als *Papier-und-Bleistift-Tests* von den Schülerinnen und Schülern bearbeitet (siehe Abschnitt 6.1). Bei dem Wortverständnistest (75 Items) geht es darum, passend zu einem dargestellten Bild eines von vier möglichen Wörtern zu erkennen und anzukreuzen. Ist beispielsweise eine Ente abgebildet, so müssen Schülerinnen und Schüler aus den vier Antwortalternativen „Ente“ – „Stein“ – „Blume“ – „Fisch“ das zum Bild passende Wort erkennen und ankreuzen. Beim Satzverständnistest (36 Items) hingegen ist aus fünf Antwortalternativen dasjenige Wort zu unterstreichen, das den vorgegebenen Satz sinnvoll ergänzt. Ein Beispiel hierzu wäre der Satz „Mit einem Füller – Bein – Kuchen – Kopf – Hals kann man schreiben.“

Der Test ist sowohl in der Einzeldiagnostik als auch in der Testung großer Gruppen und Stichproben verwendbar. Die Durchführung erfolgte nach Manual standardisiert. Für den Wortverständnistest und für den Satzverständnistest hatten die Schülerinnen und Schüler jeweils 3 Minuten Zeit. Für jede richtig gelöste Aufgabe wurde ein Punkt gegeben. Nicht erreichte oder ausgelassene Aufgaben sind als falsch gewertet worden.

Die *odds-even Split-Half-Reliabilität* erreicht mit .99 für den Wortverständnistest und .98 für den Satzverständnistest ähnliche hohe Werte wie sie in der Normstichprobe festgestellt worden sind (.98 für den Wortverständnistest und .94 für den Satzverständnistest). Da für den ELFE II schulmonatsgenaue Normwerttabellen vorliegen, wurde tagesgenau geprüft, in welchem Schulmonat der vierten Klassenstufe sich die Schülerinnen und Schüler befunden haben. 61.8 Prozent der Schülerinnen und Schüler befanden sich im Bereich von sieben bis acht Schulmonaten und 38.2 Prozent im Bereich von neun bis zehn Schulmonaten. Beide Normwerttabellen wurden angewendet und verglichen. Dabei wurden keine statistisch bedeutsamen Differenzen in logistischen Regressionen und Varianzvergleichen festgestellt, die auf eine Notwendigkeit der Anwendung unterschiedlicher Normwerttabellen hindeuten (Wortverständnistest: *beta stand* = -.48; *t* = -0.48; Satzverständnistest: *beta stand* = -.88; *t* = -0.90). Da eine größere Zahl von Schülerinnen und Schülern sieben bis acht Schulmonate in der vierten Klassenstufe aufwies, wurden in der Folge die entsprechenden Normwerttabellen für beide Testteile verwendet.

Abschließend wurden die Ergebnisse der Testteile kombiniert, um diagnostische Schwellenwerte abzuleiten. Dafür wurden, wenn gültige Werte für beide Testteile vorlagen, die normierten Werte für das Wortverständnis und das Satz-

verständnis summiert und anhand einer Normtabelle für die ELFE-II-Kurzform erneut in normierte Werte umgewandelt.

Die beiden Untertests aus ELFE II wurden in IGLU 2016 mit dem Ziel eingesetzt, insbesondere die Leistungen der leseschwachen Schülerinnen und Schüler (Kompetenzstufen I und II) differenzierter erfassen zu können als in den Erhebungszyklen zuvor, da über die Leistungen dieser Kinder bislang kaum Aussagen getroffen werden konnten. Mit dem Lesen auf Wortebene sind hierarchieniedrige Prozesse des Lesens angesprochen, die eine wichtige Grundlage für das Beherrschen hierarchiehöherer Prozesse (= Lesen auf Satz- und Textebene) darstellen (z.B. Christmann & Groeben, 2001). Die Befunde der in IGLU 2016 eingesetzten Untertests aus ELFE II liefern damit differenzierte Informationen zu dem Leseverstehen von Schülerinnen und Schülern auf der Schwelle von hierarchieniedrigen zu hierarchiehöheren Leseprozessen. Befunde hierzu werden in den Kapiteln 3 und 10 im vorliegenden Band vorgestellt.

5.2 Kontextfragebögen

Mit Hilfe von schriftlichen Befragungen der an IGLU teilnehmenden Schülerinnen und Schüler und der an ihrem Lernprozess beteiligten Akteurinnen und Akteure können Hintergrundinformationen erhoben werden, die mit schulischem Lernen – insbesondere in Bezug auf den Erwerb von Lesekompetenzen bezogen – im Zusammenhang stehen können (Hooper & Fishbein, 2017). Wie schon für die Leistungstests beschrieben (siehe Abschnitt 5.1), erfolgt auch die Entwicklung der Fragebogenitems in einem kooperativen Prozess der IGLU-Teilnehmerstaaten. Vorschläge für Fragen zur Erfassung von Hintergrundmerkmalen können aus allen Teilnehmerstaaten eingereicht werden. Nach der Einreichung erfolgt eine Begutachtung durch Expertinnen und Experten. Zu jedem Studienzyklus werden neue Fragebogenitems entwickelt und eingereicht. Dies geschieht etwa dann, wenn mit aus den Fragebögen gewonnenen Informationen Unterschiede in schulischen Leistungen von Schülerinnen und Schülern verschiedener Teilnehmerstaaten und Regionen umfassender als zuvor erfasst und damit interpretiert werden können. Maßgeblich ist die inhaltliche Nähe zu und der Erkenntnisgewinn mit den durch die Rahmenkonzeption fokussierten Inhalten und Anforderungen (siehe Abschnitt 4.1). Sichergestellt wird auch, dass die ausgewählten Fragen die Testziele der IGLU-Rahmenkonzeption repräsentieren und sich in allen Teilnehmerstaaten einsetzen lassen. Die Überprüfung der Qualität erfolgt im Feldtest. Die von der internationalen Studienleitung bestimmten Inhalte der Kontextfragebögen sind für alle Teilnehmerstaaten verbindlich (Mullis & Martin, 2015). Da sich diese Inhalte nicht zwingend mit Fragen decken, die auch Wissenschaft und Politik in einzelnen Staaten vorrangig beschäftigen, besteht die Möglichkeit der nationalen Erweiterung von Fragebögen. So können Fragebogenitems national ergänzend zu den international verbindlich vorgegebenen Fragen in einem Staat eingesetzt werden. In Deutschland wurden der Schüler-, Eltern-, Lehrkräfte- und der Schulfragebogen um nationale Ergänzungen erweitert.

5.2.1 Datenschutzrechtliche Begutachtung der Kontextfragebögen

Die in IGLU 2016 implementierten Verfahren der Datenerhebung und die Instrumente wurden durch die Kultusministerien der Länder datenschutzrechtlich geprüft und anschließend – teilweise mit kleineren länderspezifischen Anpassungen – für die Erhebung zugelassen.

5.2.2 Inhalte der Kontextfragebögen

Die in IGLU 2016 eingesetzten Fragebögen thematisieren unter anderem die nachfolgend aufgeführten Aspekte (Hußmann, Wendt, Bos & Rieser, 2018).

Der *Schülerfragebogen* umfasst Fragen zum soziodemografischen und soziokulturellen Hintergrund des Kindes (z.B. Alter, Geschlecht, Sprachgebrauch zu Hause, Muttersprache) und Fragen zum Leseverhalten, zur Lesemotivation und zum Leseselbstkonzept. Erfragt werden die Teilnahme an außerunterrichtlichen Angeboten (Ganztags- und Förderangeboten) und Freizeitaktivitäten.

Informationen, die aus dem *Elternfragebogen* gewonnen werden, ergänzen die Angaben der Schülerinnen und Schüler zur familiären Lernumwelt. Darin erfragte Themen sind unter anderem die einer Familie zur Verfügung stehenden Ressourcen, soziodemographische Daten der Eltern (z.B. höchster Bildungsabschluss, Fragen zur Berufstätigkeit, Zusammensetzung der Familie), der Bildungshintergrund der Eltern, Bildungsaspirationen der Eltern, ebenso wie frühe Lernerfahrungen des Kindes sowie lernunterstützende Aktivitäten und Fragen zum kognitiven Anregungsgehalt der familiären Lernumwelt.

Der *Fragebogen für die Deutschlehrkräfte* dient der systematischen Erfassung des Unterrichts und spezifischer Klassenmerkmale. Neben Personenmerkmalen (Alter, Geschlecht, Lehrerfahrung) sind Fragen zur Ausbildung und Qualifizierung, zu Merkmalen und zur Ausstattung der Klasse, zu Einstellungen unterrichtsrelevanter Aspekte, Unterrichtspraxis und Lehrmethoden, lernrelevanten Voraussetzungen der Schülerinnen und Schüler sowie persönlichen Einstellungen zum Lesen zu beantworten. Zusätzliche Informationen wie Schullaufbahneempfehlungen, Förderbedarfe oder Schulnoten werden mit einer *Schülerteilnahmeliste* erhoben, die ebenfalls durch eine Lehrkraft ausgefüllt wird.

Mit dem von der Schulleitung auszufüllenden *Schulfragebogen* werden Informationen zu Schulcharakteristika (Schulgröße, Lehrkräfte, Unterrichtszeiten, soziale und geographische Lage), zur Schulorganisation, zum Kontext und zur Ausstattung einer Schule, pädagogischen Zielsetzungen, zu Curricula, Ressourcen in Form von technischer und materieller Ausstattung, zum sozialen Klima, zu Kooperationsstrukturen, Einbindung der Eltern, außercurricularen Aktivitäten, zur Rolle der Schulleitung, zum Umgang mit Heterogenität, zur Ganztagsbetreuung sowie zu Angeboten zum Lesen gewonnen.

Ergänzend zu diesen Fragebögen enthält der *Curriculumfragebogen* Fragen dazu, inwieweit die Curricula der teilnehmenden Länder und Regionen durch die Inhalte der IGLU-Tests abgebildet werden. Ausgefüllt wird der Fragebogen, wie in Abschnitt 4.1.2 erläutert, von Fachexpertinnen und Fachexperten der Staaten und Regionen. Fragen darin umfassen Rahmenbedingungen der Umsetzung des Curriculums, zum Beispiel wer über das jeweilige Curriculum entscheidet, ob und in welcher Form die Implementation des Curriculums evaluiert wird oder ob und, wenn ja, inwiefern Lehrkräfte bei der Umsetzung des Curriculums

unterstützt werden. Die Ergebnisse dieser Befragung gehen in die *Enzyklopädie* (Mullis et al., 2017; Wendt, Walzebug, Bos, Smith & Bremerich-Vos, 2017) ein.

5.2.3 Übersetzung und Gestaltung der Testinstrumente

Die internationale Studienleitung stellt zur Durchführung der Studie jedem Teilnehmerstaat die Test- und Befragungsinstrumente in englischer Sprache bereit. An Schulen eingesetzt werden die Tests und Fragebögen in der Sprache des Teilnehmerlandes beziehungsweise der Teilnehmerregion. Jedes nationale Forschungsteam trägt Verantwortung dafür, die Testinstrumente und Fragebögen in den für den nationalen Kontext relevanten Unterrichtssprachen zu erstellen (Martin et al., 2015; Mullis & Prendergast, 2017). Um ein hohes Maß an Qualität und die Gestaltung nach einheitlichen Standards zu gewährleisten, ist international präzise festgelegt, nach welchen Richtlinien und Abläufen die Übersetzung und die Gestaltung der Instrumente erfolgen sollen. Bei der Übersetzung der Testaufgaben oder Fragen in die jeweilige(n) Unterrichtssprache(n) muss die internationale Vergleichbarkeit gewahrt bleiben. Im Speziellen gilt für die Leistungstests, dass sich durch die Übersetzung weder eine Aufgabenstellung noch eine Antwortoption in ihrem Schwierigkeitsgrad von dem international vorgegebenen Original unterscheiden darf. Hingegen ist bei der Übersetzung der international vorgegebenen Fragen in den Kontextfragebögen zu prüfen, ob sie auch im nationalen Kontext einen Sinn ergeben. Gegebenenfalls ist auf Besonderheiten des nationalen Schulsystems Rücksicht zu nehmen, so dass es in Absprache mit der internationalen Studienleitung zu nationalen Anpassungen kommen kann.

Die Einhaltung der Prozeduren sowie die Qualität der Übersetzung werden von der internationalen Studienleitung sorgfältig kontrolliert. Nationale Anpassungen der international vorgegebenen Instrumente sind nur zulässig, sofern sie durch kulturelle Unterschiede begründet werden können (Yu & Ebbs, 2011). Die Qualität der Übersetzung in andere Sprachen und mögliche nationale Adaptionen werden von Übersetzerinnen und Übersetzern der IEA durch Rückübersetzung ins Englische kritisch überprüft. Dabei identifizierte Abweichungen von Inhalt und Layout des englischen Originals werden in Abstimmung mit der nationalen Studienleitung korrigiert und zulässige nationale Abweichungen dokumentiert.

In Deutschland erfolgte die Übersetzung der Testaufgaben und Fragebögen wie auch im vergangenen Studienzyklus zu IGLU 2011 in Kooperation mit dem IGLU-Forschungsteam aus Österreich. Umgesetzt wurde sie von einer professionellen Übersetzerin, die auf eine breite Erfahrung im sensiblen Umgang mit Begriffen und Formulierungen im Rahmen von Schulleistungstudien zurückgreifen kann und bereits bei IGLU 2011 diese Prozesse begleitet hat. Im Anschluss an die Übersetzung wurden die Testinstrumente und Fragebögen erneut von Expertinnen und Experten kritisch begutachtet und – sofern sinnvoll – mit den Übersetzungen aus den vorangegangenen Studienzyklen abgeglichen, gegebenenfalls hatte dies kleinere sprachliche Anpassungen zur Folge.

Verzichtet wurde, anders als noch in IGLU 2011, auf die Übersetzung des Elternanschreibens in verschiedene Sprachen. Es lag also nur in deutscher Sprache vor. Das Elternanschreiben informiert die Eltern der teilnehmenden Schülerinnen und Schüler über die Studie, die Organisation der Studien, Komponenten der Leistungsmessung, Inhalte der Befragung und datenschutz-

rechtliche Grundlagen. Teil dieses Elternschreibens ist auch die schriftliche Bestätigung des Einverständnisses zur Teilnahme des Kindes.

6 Erhebung

Die Haupterhebung von IGLU 2016 fand in Deutschland in der Zeit vom 2. Mai bis zum 3. Juni 2016 statt.

6.1 Aufbau der Untersuchung

Der international vorgegebene Testablauf stellt sicher, dass alle teilnehmenden Schülerinnen und Schüler genügend Zeit zur Bearbeitung der Testaufgaben und zur Beantwortung der Fragen des Schülerfragebogens haben. In Deutschland fanden die Testung und die Befragung der Schülerinnen und Schüler an einem Vormittag statt. Die Schülerinnen und Schüler lösten am Testtag zuerst die international vorgegebenen IGLU-Testaufgaben (siehe Abschnitt 5.1). Zwischen dem ersten und dem zweiten Teil des Testhefts fand eine 10-minütige Pause statt. Für die Bearbeitung beider Testteile waren je 40 Minuten vorgesehen. Dies ergab eine Gesamttestzeit von 80 Minuten (siehe Tabelle 2.4).

Ergänzend wurden in Deutschland im Anschluss an den IGLU-Test kleinere Tests in allen IGLU-Klassen durchgeführt. Diese umfassen mit dem Kurztest zu figuralem Denken einen Teilbereich kognitiver Fähigkeiten (KFT, N2, Heller &

Tabelle 2.4: Untersuchungsablauf von IGLU 2016

Beginn der Testsitzung: Verteilung des Materials, Begrüßung, allgemeine Einweisung	ca. 10 min.
Bearbeitung Lesetest Teil I	40 min.
Pause	10 min.
Einweisung in den Lesetest Teil II	ca. 5 min.
Bearbeitung Lesetest Teil II	40 min.
Pause [Einsammeln der Testhefte und Austeilen der Schüler- und Elternfragebogen]	20 min.
Einweisung in die Bearbeitung des kognitiven Fähigkeitstests (KFT, N2)	ca. 5 min.
Bearbeitung des kognitiven Fähigkeitstests (KFT, N2)	7 min.
Einweisung in die Bearbeitung des Wortverständnistests (ELFE II)	ca. 4 min.
Bearbeitung des Wortverständnistests (ELFE II)	3 min.
Einweisung in die Bearbeitung des Satzverständnistests (ELFE II)	ca. 2.5 min.
Bearbeitung des Satzverständnistests (ELFE II)	3 min.
Einweisung in den Schülerfragebogen	ca. 5 min.
Bearbeitung des Schülerfragebogens	ca. 45 min.
Beenden der Testsitzung: Einsammeln der Materialien	ca. 5 min.
Reine Bearbeitungszeit	ca. 138 min.
Gesamtzeit	ca. 204.5 min.

Perleth, 2000) sowie Leseverstehen auf Wort- und Satzebene (ELFE II, Lenhard et al., 2017), wie in Abschnitt 5.1.4 erläutert. Die Bearbeitungszeit für den KFT-Subtest betrug 7 Minuten, für die ELFE-Untertests je 3 Minuten.

Der Testtag endete für die teilnehmenden Schülerinnen und Schüler mit der Bearbeitung des Schülerfragebogens. Für die Bearbeitung des Schülerfragebogens hatten die Schülerinnen und Schüler 45 Minuten Zeit.

6.2 Durchführung der Erhebung

Die Datenerhebung wurde, wie auch in allen vorangegangenen Erhebungsrounden, in Deutschland durch die *IEA Hamburg* durchgeführt. Die Erhebung folgte hoch standardisiert und kontrolliert den internationalen Vorgaben entsprechend. Die Arbeit der *IEA Hamburg* umfasste den Kontakt mit den Schulen und die Auswahl der Testleiterinnen und Testleiter, die zumeist Lehramts-, Erziehungswissenschafts- oder Psychologiestudierende höherer Semester waren, und über Erfahrungen im Bereich der Testdurchführung verfügten. Sämtliche Testleiterinnen und Testleiter nahmen vor der Erhebungsphase an einer verbindlichen Schulung zur Testdurchführung teil und erhielten ein schriftliches Manual mit genauen Vorgaben zur Durchführung des Testtages. Sie waren dazu angehalten, den international vorgegebenen Testablauf akkurat einzuhalten. Nach standardisierter Vorgabe wurde den Schülerinnen und Schülern am Testtag jeder Testteil ausführlich erklärt und die Durchführung anhand von Beispielen erläutert. Die Kinder hatten jederzeit die Möglichkeit, Verständnisfragen zur Testbearbeitung zu stellen, die nicht auf den Inhalt bezogen waren.

6.3 Qualitätssicherung

Um verlässliche Aussagen aus international-vergleichenden Schulleistungsstudien wie IGLU ableiten zu können, ist es notwendig, dass der Leistungstest in allen teilnehmenden Staaten unter vergleichbaren Voraussetzungen durchgeführt wird. Mit der Teilnahme an IGLU verpflichten sich alle Staaten zur Einhaltung aller Vorgaben und zur Implementation umfassender qualitätssichernder Maßnahmen.

Unabhängige Expertinnen und Experten wurden von der IEA beauftragt, in allen Teilnehmerstaaten stichprobenartig die Einhaltung der internationalen Vorgaben zu überprüfen. In Deutschland führte das sogenannte *International Quality Control Monitoring* (IQCM) Dr. Kristina Frey vom Institut für Erziehungswissenschaft der Universität Münster durch. Sie begutachtete in einer international vorgegebenen Stichprobe von 10 Prozent sämtlicher Testklassen den Ablauf der Testdurchführung. Beobachtungen wurden in von der internationalen Studienleitung vorgegebene Bögen dokumentiert und mündliche Befragungen mit Schulkoordinatorinnen, Schulkoordinatoren und Testleitungen im Anschluss an die besuchten Testungen durchgeführt. Bei der Qualitätskontrolle in Deutschland wurden keine Mängel in der Einhaltung der Erhebungsstandards festgestellt.

Neben dem IQCM wurde ein von der nationalen Studienleitung verantwortetes Qualitätsmonitoring, das sogenannte *National Quality Control Monitoring* (NQCM), an weiteren 20 zufällig ausgewählten Testschulen durch Mitarbeiterinnen des IFS durchgeführt. Es fanden Qualitätsbeobachtungen am Testvormittag statt, die um Interviews mit der Schulkoordination zur Testqualität, Organisation und zum Belastungsempfinden der Schulen ergänzt und anschlie-

ßend umfassend dokumentiert wurden. Weder die Durchführung des NQCM noch die des IQCM konnten Mängel in der Einhaltung der Erhebungsstandards in Deutschland feststellen.

7 Stichprobe und Beteiligungsquoten in Deutschland

Die für Deutschland zufällig gezogene Schulstichprobe umfasst 208 Schulen, aus allen Ländern der Bundesrepublik. Das *Sampling* ist umfassend im technischen Report dokumentiert (Martin et al., 2017). Die Stichprobenziehung erfolgte gemeinsam und analog zur Stichprobenziehung für TIMSS 2015 (vgl. Wendt et al., 2016). Für die Testung wurde an jeder Schule eine vierte Klasse gezogen. An allen 208 getesteten Schulen konnte die Leistungsmessung planmäßig durchgeführt werden.

Von den 4277 Schülerinnen und Schülern der Stichprobe nahmen 3959 Schülerinnen und Schüler tatsächlich am IGLU-Test teil. Die Nicht-Teilnahme von 318 Schülerinnen und Schülern ist auf verschiedene Gründe zurückzuführen (o.g. Ausschlussgründe, Verlassen der Schule vor dem Testtag oder Abwesenheit am Testtag). Deutschland erreicht damit, ähnlich wie in IGLU 2011, eine Gesamtteilnahmequote (für Schüler *und* Schulen) von 93 Prozent exklusive, beziehungsweise 95 Prozent inklusive Ersatzschulen (siehe Tabelle A.4 im Anhang A in diesem Band). Auf erfreulich hohe Akzeptanz stieß IGLU auch bei Lehrkräften und Schulleitungen. Aus jeder IGLU-Klasse liegt mindestens ein Lehrerfragebogen vor, und die Rücklaufquote der IGLU-Schulfragebögen beträgt für Deutschland 93 Prozent. Die Rücklaufquote der Elternfragebögen fiel in IGLU 2016 allerdings mit 72 Prozent erkennbar kleiner aus als etwa in IGLU 2011 (80%).

In Tabelle 2.5 sind die Rücklaufquoten sowie zentrale Hintergrundmerkmale der Schülerinnen und Schüler im Vergleich von 2001, 2006 und 2011 zu 2016 dargestellt. In Bezug auf die Komposition der Schülerschaft zeigt sich, dass sich der Anteil an Mädchen in der IGLU-Stichprobe und das durchschnittliche Alter praktisch nicht geändert hat. Im Vergleich zu den vorangegangenen Erhebungsrunden ist der Anteil von Kindern mit Migrationshintergrund (beide Eltern im Ausland geboren) in IGLU 2016 mit 19 Prozent statistisch signifikant höher, spiegelt aber auch reale Veränderungen in der Komposition der Schülerschaft an Grundschulen in Deutschland wider (Autorengruppe Bildungsberichterstattung, 2016; Statistisches Bundesamt, 2017). Gleiches gilt für den ebenfalls im Vergleich zu 2011 statistisch signifikanten höheren Anteil an Schülerinnen und Schüler mit besonderen Unterstützungsbedarfen, dieser liegt in IGLU 2016 bei rund 7 Prozent.

Tabelle 2.5: Beteiligungszahlen, Rücklaufquoten und zentrale Charakteristika zu IGLU 2001, 2006, 2011 und 2016

	IGLU 2001	IGLU 2006	IGLU 2011	IGLU 2016
Schulen				
Anzahl	211	405	198	208
Beteiligung am Schulfragebogen (%)	95	96	96	93
Klassen				
Anzahl	393	405	198	208
Lehrkräfte				
Anzahl	393	418	222	227
Beteiligung am Lehrerfragebogen (%)	92	96	96	93
Schülerinnen und Schüler				
Anzahl	8997	8302	4241	4277
Durchschnittsalter	10.5	10.5	10.4	10.3
Mädchen (%)	50	49	49	49
Familie mit hohem sozioökonomischen Status (%) (mindestens ein Elternteil ist Akademiker oder Führungskraft)	29	29	32	29
Kinder mit Migrationshintergrund (%) (beide Elternteile im Ausland geboren)	14*	15*	16	19
Kinder mit besonderen Unterstützungsbedarfen (%)	-	-	5*	7
Testteilnahme				
Anzahl	7633	7899	4000	3959
Teilnahmequote (%)	88	94	96	95
Beteiligung am Schülerfragebogen (%)	98	96	91	88
Eltern				
Beteiligung am Elternfragebogen (%)	89	87	80	72

* = Unterschied zu IGLU 2016 statistisch signifikant ($p < .05$).

8 Aufbereitung und Analyse der Daten

In Deutschland war, wie auch in den vergangenen Erhebungsrunden, die *IEA Hamburg* mit der Organisation der nationalen Datenverarbeitung, -aufbereitung und der Aufgabenkodierung beauftragt. Die Dateneingabe und -verarbeitung erfolgte technologiegestützt, so dass die Testhefte elektronisch eingelesen werden konnten. Neben einer niedrigen Fehlerquote im Prozess der Datenverarbeitung hat dies den Vorteil, auch zu einem späteren Zeitpunkt schnell und problemlos auf die Antworten einzelner Schülerinnen und Schüler zugreifen zu können, beispielsweise wenn individuelle Antworten überprüft werden müssen. Antworten, die von der Scan-Software nicht sicher erkannt werden konnten, wurden von ausgebildeten Mitarbeiterinnen und Mitarbeitern nachträglich verifiziert. Die Datensätze wurden abschließend mit Hilfe einer speziell entwickelten Software kritisch auf ungültige Daten und Inkonsistenzen geprüft.

Kodierung der Leistungstests

Die Schülerlösungen der bearbeiteten offenen Testaufgaben wurden nach eng umrissenen internationalen Vorgaben bewertet und für die Analysen kodiert. Während bei Mehrfachwahlaufgaben die richtige Lösung eindeutig festgelegt ist, weisen Testfragen mit offenem Antwortformat (siehe Abschnitt 5.1.1 sowie Kapitel 3) einen kleinen Bewertungsspielraum auf.

In Deutschland wurde die Kodierung der Fragen mit offenem Antwortformat der *IEA Hamburg* übertragen und dort von erfahrenen Kodierern und Kodierern übernommen, bevorzugt von solchen, die Deutsch studieren. Zur Vorbereitung wurden die Kodierern und Kodierer mit Beispiellösungen intensiv für ihre Aufgabe geschult. Sie erhielten eine Kodieranweisung mit genauen Beschreibungen für richtige und falsche Lösungen zu jeder einzelnen Aufgabe. Die Güte der Kodierungen wurde anhand von Doppelkodierungen geprüft. Dazu wurden 200 zufällig ausgewählte Antworten zu jeder Aufgabe von einer zweiten Person beurteilt und es wurde die Übereinstimmung zwischen der Erst- und der Zweitkodierung bestimmt. In IGLU 2016 ist, wie auch in den Erhebungszyklen zuvor, die durchschnittliche Übereinstimmung über alle Aufgaben hinweg in allen Staaten sehr hoch (Martin et al., 2017).

9 Skalierung der Leistungstests

Mit der in IGLU eingesetzten Leistungstestung soll von den in den Testheften gezeigten Schülerantworten auf die zu messende Kompetenz von Viertklässlerinnen und Viertklässlern geschlossen werden. Die Verortung von Schülerinnen und Schülern auf einer Kompetenzskala wird Skalierung genannt. Zur Bestimmung individueller Leistungswerte auf den Kompetenzdomänen wird in IGLU auf statistische Modelle zurückgegriffen, die auf der *Item Response Theory* basieren (IRT, Boomsma, van Duijn & Snijders, 2000; Linden v. d. & Hambleton, 1997). Die IRT geht von der Annahme aus, dass die beobachtbaren Antworten einer Person in einem Test durch eine nicht beobachtbare, testbezogene Fähigkeit erklärbar sind. Die beobachtbaren Antworten werden verwendet, um die Höhe der nicht beobachtbaren (*latenten*) Fähigkeiten zu schätzen. Mit der IRT können die Schwierigkeiten der Aufgaben und die Kompetenzen der Personen (auch als *Fähigkeiten* bezeichnet) auf derselben Skala abgebildet werden, das heißt die Schwierigkeiten der Aufgaben und die Fähigkeiten der Personen sind auf einer Metrik direkt nebeneinander vergleichbar. Damit ist zudem möglich, Personen die unterschiedliche Aufgaben derselben Metrik bearbeitet haben, hinsichtlich ihrer Fähigkeiten zu vergleichen.

Die beschriebene Eigenschaft von IRT-Modellen macht dieses Verfahren für IGLU (und vergleichbare Studien wie TIMSS) besonders interessant. Bedingt durch das in IGLU implementierte Testheftdesign (Multi-Matrix-Design, siehe Tabelle 2.3) bearbeiten einzelne teilnehmende Schülerinnen und Schüler immer nur eine Auswahl der Testaufgaben des IGLU-Aufgabenpools (Foy, Brossman & Galia, 2011). Durch die Etablierung einer IRT-Skala für den IGLU-Aufgabenpool wird gewährleistet, dass dennoch ein direkter Vergleich dieser Schülerleistungen möglich ist – eine ausreichend große Stichprobe und Überlappungen der Testheftinhalte ist dabei vorausgesetzt (z. B. Kolen, 1981; Kolen & Brennan, 2004; Lord, 1980). IGLU zielt nicht auf Individualdiagnostik, sondern dient ausschließlich der Beschreibung von Kompetenzverteilungen in den untersuchten Populationen. Da diese Verteilungen möglichst genau geschätzt werden sollen, wird in IGLU

zur Ermittlung der Schülerleistung (*Personenparameter*) der *Plausible-Value*-Ansatz gewählt (Mislevy, 1991; Mislevy, Beaton, Kaplan & Sheehan, 1992); ein Verfahren, dem die Theorie der Multiplen Imputation zugrunde liegt (Rubin, 1987).

Die Grundidee des *Plausible-Value*-Ansatzes ist, die nicht beobachtbare wahre Fähigkeit einer Person als prinzipiell unbekannt zu betrachten. Diese nicht bekannte Fähigkeit wird dann durch ‚plausible Werte‘ abgebildet. Zur Bestimmung der plausiblen Werte wird neben der Information über die Testleistung der Person eine Vielzahl von Informationen aus den Hintergrundfragebögen mitberücksichtigt. Auf Basis dieser Information wird eine (bedingte) Wahrscheinlichkeitsverteilung für die Fähigkeit einer Person erstellt. *Plausible Values* sind jene Werte, die aus dieser bedingten Verteilung per Zufall gezogen werden. Durch die Ziehung mehrerer *Plausible Values* wird berücksichtigt, dass die Bestimmung eines fehlenden Wertes immer auch mit Unsicherheit behaftet ist. Als Konvention hat sich die Ziehung von fünf Werten etabliert. Zur angemessenen Schätzung von Populationskennwerten sind alle *Plausible Values* zu berücksichtigen, und bei der Bestimmung der Stichprobenunsicherheiten auch deren Varianz. Analysen mit den Leistungsdaten müssen entsprechend immer fünfmal erfolgen und die Ergebnisse dann zusammengefasst werden.

Skalierungsmodelle

Im vorliegenden Bericht werden die Kompetenzen der Schülerinnen und Schüler auf Grundlage der Skalierung berichtet, die die internationale Studienleitung durchgeführt hat (Martin et al., 2017). Die Skalierung der Leistungsdaten aller Teilnehmerstaaten ist komplex und zeitintensiv, ihr gehen umfangreiche Datenprüfungs- und Aufbereitungsschritte voraus. Für IGLU 2016 erfolgte erst nach Abschluss aller Prüfungen und einem Review der Ergebnisse durch alle beteiligten nationalen Studienleitungen die Weitergabe der Datensätze aller Teilnehmerstaaten und Regionen an die nationalen Studienleitungen im August 2017. Zentrale Modellparameter werden in der technischen Dokumentation zur Studie veröffentlicht (ebd.).

Die internationale Skalierung erfolgte im Zeitraum von Herbst 2016 bis Sommer 2017. Die Qualitätssicherung der Berechnungen wurde durch Expertinnen und Experten des US-amerikanischen *Educational Testing Service* (ETS) durchgeführt. Ausgangspunkt der Skalierung ist die umfassende Prüfung der Messeigenschaften jeder einzelnen Testaufgabe nach festgelegten Kriterien (ebd.). Im Ergebnis dieser Prüfungen standen für die Ermittlung der Kompetenzmittelwerte im Lesen 169 administrierte Testaufgaben zur Verfügung.

Die Bestimmung der *Plausible Values* erfolgte zunächst unter Nutzung eines mehrdimensionalen dreiparametrischen logistischen Modells (3-PL-Modell), in dem neben der Schwierigkeit der Aufgabe auch deren Trennschärfe und (bei Aufgaben im *Multiple-Choice*-Format) ein Parameter für die Ratewahrscheinlichkeit berücksichtigt wird. Geschätzt wurden die Modellparameter mit dem Computerprogramm *Parscale* (Muraki & Bock, 1999) unter Verwendung des *Marginal-Maximum-Likelihood* (MML) Ansatzes. Anschließend wurde für die Schätzung der Personenparameter (*Plausible Values*) die Software *MGROUP* (Sheehan, 1985) verwendet. Um genauere Schätzer der Personenparameter zu erhalten, kommt in IGLU 2016 für die Skalierung der Subskalen (wie auch schon in IGLU 2011) ein mehrdimensionales Antwortmodell zur Anwendung, mit dem die in den Daten enthaltene Zusammenhangsstruktur besser abgebildet werden kann. Zuvor, das heißt in den vorangegangenen Erhebungszyklen 2001 und 2006, wur-

den die Kompetenzwerte für die Subskalen (je zwei Leseintentionen und zwei Verstehensprozesse, siehe Kapitel 3 in diesem Band) separat ermittelt. Um die Vergleichbarkeit zu den vorangegangenen Studienzyklen sicherzustellen wurden im vorliegenden Berichtsband die im Rahmen der nationalen Berichtslegung von IGLU 2011 neu skalierten Daten zu IGLU 2001 und 2006 genutzt (vgl. auch Tarelli et al., 2012). Abweichungen zur Berichtslegung zu IGLU 2001 und 2006 (Bos et al., 2003; Bos et al., 2007; Mullis, Martin, Gonzalez & Kennedy, 2003; Mullis et al., 2007) sind demnach durch die Veränderung des Skalierungsmodells bedingt.

10 Gewichtung und Schätzung von Stichproben- und Messfehlern

Das wesentliche Merkmal einer repräsentativen Stichprobe besteht darin, von den Verhältnissen in einer Stichprobe auf die Grundgesamtheit schließen zu können. Dies setzt voraus, dass die Stichprobe nicht verzerrt und die Grundgesamtheit angemessen repräsentiert ist. In IGLU finden sich zwei Ursachen, die zu Stichprobenverzerrungen führen können: Erstens hat aufgrund des Designs der Stichprobenauswahl (siehe Abschnitt 7) nicht jede Schülerin beziehungsweise jeder Schüler dieselbe Wahrscheinlichkeit, in die Stichprobe zu gelangen. So hängt die Wahrscheinlichkeit unter anderem von der Zügigkeit der Schulen ab: In einer Schule mit zwei Jahrgangsklassen hätte ein Schulkind eine Wahrscheinlichkeit von 50 Prozent in die Stichprobe zu gelangen. In einer Schule mit vier Jahrgangsklassen wäre es dagegen nur eine Wahrscheinlichkeit von 25 Prozent. Zweitens kann auch der Stichprobenausfall eine Verzerrung bedeuten. Nur selten sind Ausfälle zufällig; vielmehr weisen sie oft einen Zusammenhang mit studienrelevanten Merkmalen auf. So ist es denkbar, dass leistungsschwache Schülerinnen und Schüler mit höherer Wahrscheinlichkeit am Testtag fehlen als durchschnittliche oder leistungsstarke Schülerinnen und Schüler. Werden Stichprobenverzerrungen nicht angemessen korrigiert, können inferenzstatistische Methoden zu falschen Schlussfolgerungen führen.

In IGLU wird diesem Problem dadurch begegnet, dass für jedes getestete Schulkind ein statistisches Gewicht berechnet wird. Die Gewichte werden dann bei der Berechnung aller Statistiken verwendet, zum Beispiel bei einem Mittelwert oder bei Prozentangaben. Die Daten der teilnehmenden Schülerinnen und Schüler tragen dadurch in kontrollierter Art und Weise unterschiedlich stark zu der Berechnung der Statistiken bei.

Da in IGLU keine reinen Zufallsstichproben vorliegen, sondern sogenannte Clusterstichproben gezogen werden, kann die Bestimmung des Standardfehlers nicht mit sonst üblichen Analyseverfahren vorgenommen werden, da sonst der Standardfehler systematisch unterschätzt werden würde. Eine präzisere Bestimmung des Standardfehlers erlauben sogenannte *Jackknife*-Verfahren. Diese Verfahren bestimmen die Variabilität der Schätzung von Populationskennwerten (wie z. B. die Leseleistung der in IGLU 2016 getesteten Grundschülerinnen und Grundschüler) durch ein wiederholtes Schätzen dieser Werte aus Substichproben, was die Möglichkeit bietet, Stichprobenfehler zu schätzen, ohne zugleich die Annahme einfacher Zufallsstichproben voraussetzen zu müssen. Durch die Reduzierung der Freiheitsgrade führt diese Methode zu akkurateren Schätzungen der Standardfehler.

Die korrekte Bestimmung des Standardfehlers ist sehr wichtig, weil er genutzt wird, um zu ermitteln, ob sich zwei Gruppen signifikant voneinander unterscheiden. Im vorliegenden Bericht werden entsprechend alle Standardfehler mit solch einem Verfahren (*Jackknife Repeated Replication Technique*, JRR) geschätzt. In diesem Bericht wird für die Entscheidung über die Signifikanz eine Irrtumswahrscheinlichkeit von $\alpha = .05$ zugrunde gelegt.

11 Zur Darstellung und Interpretation der Ergebnisse

Zur Darstellung der im vorliegenden Band berichteten Ergebnisse in Tabellen und Abbildungen werden verschiedene statistische Kennwerte verwendet. Die wichtigsten Kennwerte sind nachfolgend erläutert. Darüber hinausgehende Begriffserklärungen und technische Grundlagen sind in dem technischen Bericht der internationalen Studienleitung (Martin et al., 2017) sowie in einschlägiger Fachliteratur dokumentiert.

Mittelwerte und Standardabweichungen von Leistungsdaten

Die in diesem Bericht dokumentierten Ergebnisse und Vergleiche basieren auf der internationalen Stichprobe und den internationalen Kompetenz- und Fragebogenskalen von IGLU 2016 (Martin et al., 2017; Mullis & Martin, 2015). Zur Darstellung der Leistungswerte wurde für die erste Erhebung von IGLU 2001 und für alle nachfolgenden Zyklen von IGLU ein Mittelwert (M für arithmetisches Mittel) von 500 Punkten und eine Standardabweichung (SD für *Standard Deviation*) von 100 Punkten festgelegt (Mullis et al., 2003; Martin et al., 2003). Die Wahl der Einheiten für diese Skala basiert auf Konventionen. Werte, die nahe beim Mittelwert liegen, kommen häufiger vor als Extremwerte. Oft ergibt sich eine Normalverteilung, wie sie in Abbildung 2.5 (siehe Seite 68) dargestellt ist.

Die durchschnittliche Streuung der Werte um den Mittelwert wird durch die Standardabweichung statistisch gekennzeichnet. Im Bereich einer Standardabweichung über und unter dem Mittelwert (d.h. in Abbildung 2.5 im Bereich von 400 bis 600 Punkten) liegen rund zwei Drittel (68.3%) aller Testwerte der internationalen Population. Bei zwei Standardabweichungen erhöht sich dies auf 95.5 Prozent und bei drei Standardabweichungen auf 99.7 Prozent.

Differenzen und damit Veränderungen von Leistungsmittelwerten zwischen den Erhebungszyklen, wie sie im Trend berichtet werden, werden im vorliegenden Band mitunter durch das Symbol Delta (Δ) gekennzeichnet.

Internationaler Mittelwert versus Skalenmittelwert

IGLU ist als Trendstudie konzipiert, um Veränderungen in den mittleren Leistungsniveaus von Schülerinnen und Schülern der teilnehmenden Staaten und Regionen über die Zeit hinweg darstellen zu können. Der Darstellung der Ergebnisse liegt dieselbe Skala zugrunde, damit die Schülerleistungen, die zu den verschiedenen Zyklen erzielt wurden, miteinander verglichen werden können. Die Vergleichbarkeit der Ergebnisse wird sichergestellt, indem in jedem Zyklus Aufgaben aus den vorangegangenen Studienzyklen erneut eingesetzt werden (siehe Abschnitt 5.1). Die in IGLU 2016 ermittelten Ergebnisse und Daten zu Testaufgaben, die zu mehreren Erhebungszyklen eingesetzt wurden, können in einer gemeinsamen Skalierung mit den Daten von 2011 verankert werden (Martin et al., 2017; Foy et al., 2011). In gleicher Weise lassen sich die Daten

aus IGLU 2011 wiederum zu denen aus IGLU 2006 beziehungsweise IGLU 2001 in Beziehung setzen. Letztlich bilden die Daten der ersten IGLU-Erhebung im Jahr 2001 den Referenzpunkt für alle Folgezyklen. Veränderungen in den Leistungen zwischen den einzelnen Zyklen können auch 2016 auf der Basis des in IGLU 2001 bestimmten Mittelwerts von $M = 500$ verglichen und beurteilt werden. Dieser Referenzwert wird in IGLU als *Skalenmittelwert* bezeichnet.

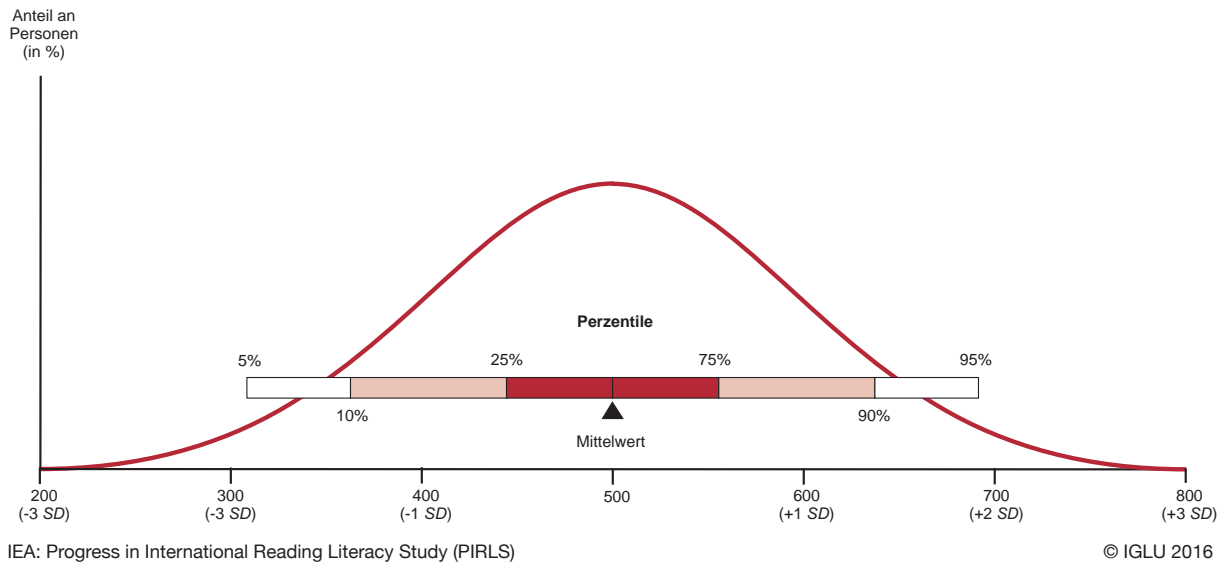
Der *internationale Mittelwert* bezeichnet den Wert, der mit jeder IGLU-Erhebung neu berechnet wird: Es handelt sich dabei um den Mittelwert, der aus allen Mittelwerten der jeweiligen Teilnehmerstaaten eines Zyklus gebildet wird (ohne Berücksichtigung der Benchmark-Teilnehmer, siehe Abschnitt 4.2.1). Im Gegensatz zum Skalenmittelwert variiert der internationale Mittelwert zwischen den Erhebungszyklen. Setzen sich zum Beispiel die Teilnehmer einer Erhebung aus Staaten zusammen, die im Vergleich zu jenen Staaten, die an IGLU 2001 teilgenommen haben, insgesamt leistungsstärker sind, ergibt sich ein internationaler Mittelwert der größer als 500 ist. In IGLU 2016 beträgt der internationale Mittelwert für Lesen 521 Punkte ($SD = 78$). Die an IGLU 2016 teilnehmenden Staaten sind im Durchschnitt leistungsstärker als die Teilnehmer an IGLU im Jahr 2001. Bei der Berichterstattung zu den Hintergrunddaten wird nicht zwischen einem internationalen Mittelwert und einem Skalenmittelwert unterschieden, weil die Fragebogeninhalte im Gegensatz zu den Leistungstests in sich heterogene Inhalte abdecken, die von Erhebung zu Erhebung teilweise variieren.

Median, Perzentile und Perzentilbänder

Neben dem Mittelwert und der Standardabweichung werden auch Perzentilwerte berichtet. Perzentile informieren, ebenso wie die Standardabweichung, über die Variation der Werte. Ein Perzentilwert gibt an, wie viel Prozent der untersuchten Personen einen Wert erreichen oder darunter bleiben. Liegt beispielsweise der Leistungswert zum 5. Perzentil bei 318, bedeutet dies, dass 5 Prozent der untersuchten Schülerinnen und Schüler einen Punktwert von 318 oder geringer erreichen. Zugleich bedeutet dies aber auch, dass 95 Prozent der Schülerinnen und Schüler einen Wert erreichen, der höher als 318 ist. Entsprechend trennt der Punktwert des 25. Perzentils das untere Leistungsviertel ab, der des 75. Perzentils das obere Leistungsviertel. Das 50. Perzentil, auch Median genannt, trennt die Verteilung in zwei Hälften mit je gleicher Personenzahl.

Im vorliegenden Bericht werden die Perzentilwerte tabellarisch oder graphisch in Form von Perzentilbändern (siehe Abbildung 2.5) dargestellt. Passen sich die Werte einer Normalverteilung an (siehe ebd.), ergibt sich ein symmetrisches Perzentilband. In diesem Fall fällt auch der Median mit dem Mittelwert zusammen.

Abbildung 2.5: Normalverteilung mit Perzentilen



Kompetenzstufen

Die Leistungsskala in IGLU deckt ein breites Kompetenzspektrum ab. Die in IGLU eingesetzten Lesetexte und Testaufgaben thematisieren verschiedene Inhalte und aktivieren verschiedene Prozesse kognitiver Anforderungen (siehe Abschnitt 4.1.1). Um die erreichten Kompetenzwerte der getesteten Schülerinnen und Schüler inhaltlich interpretieren zu können, werden in IGLU Kompetenzstufen gebildet. Ausführlich beschrieben und anhand von Beispielaufgaben illustriert sind die Kompetenzstufen in Kapitel 3 in diesem Band.

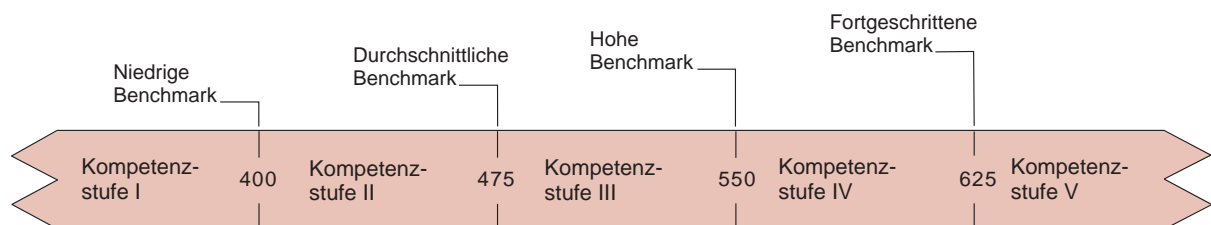
Die Entwicklung und Beschreibung der Kompetenzstufen wird von einem international zusammengesetzten Expertengremium unter der Leitung der internationalen Studienleitung vorgenommen. Bei der Entwicklung von Kompetenzstufen wurden als zentrale Bezugspunkte auf den Leistungsskalen sogenannte *Benchmarks* festgelegt (Mullis & Martin, 2015; Martin, Mullis, Foy & Stanco, 2012). Die vier *Benchmarks* (400, 475, 550, 625) teilen die Leistungsskala in fünf Abschnitte, die in Deutschland als *Kompetenzstufen* bezeichnet werden. Die niedrige *Benchmark* liegt bei 400 Punkten. Sie beschreibt den Grenzbereich zwischen Kompetenzstufe I und Kompetenzstufe II. Die durchschnittliche *Benchmark* liegt bei 475 Punkten und markiert den Beginn von Kompetenzstufe III. Die hohe und die fortgeschrittene *Benchmark* liegen bei 550 beziehungsweise 625 Punkten. Mit ihrer Überschreitung beginnen Kompetenzstufe IV beziehungsweise Kompetenzstufe V (siehe Abbildung 2.6).

Die Anordnung der internationalen *Benchmarks* und die sich daraus ergebenden Intervalle für die Kompetenzstufen sind in Abbildung 2.6 grafisch dargestellt. Das unter der niedrigen *Benchmark* liegende Intervall umfasst in Deutschland Leistungen der Kompetenzstufe I. Schülerinnen und Schüler, die Kompetenzstufe I erreichen, sind nicht in der Lage, die für Kompetenzstufe II charakteristischen, dieser Schülerkohorte entsprechend relativ einfachen Aufgaben zu lösen. Das Kompetenzniveau von Schülerinnen und Schülern am unteren Ende der Leistungsskala lässt sich auf der Basis der eingesetzten IGLU-Leistungstests und aufgrund der geringen Anzahl richtig gelöster Aufgaben nicht differenziert beschreiben. Interpretationen zur Schülergruppe auf diesem Kompetenzniveau sind daher zunächst in dem Sinne möglich, dass lediglich aus-

gesagt werden kann, dass Schülerinnen und Schüler dieser Kompetenzstufe nicht die mit dem internationalen IGLU-Test festgelegten Mindestanforderungen erreichen. Die Erweiterung um ELFE II (Untertests auf Wort- und Satzebene) in IGLU 2016 ermöglicht erstmals, die Leseleistungen dieser Schülerinnen und Schüler auf Wort- und Satzebene konkreter zu beschreiben (siehe Abschnitt 5.1.4).

Der Beschreibung und Entwicklung der Kompetenzstufen liegt eine spezifische Auswahl von Aufgaben zugrunde, die für das auf den vier Benchmarks gezeigte Leistungsniveau charakteristisch sind. Um zu dieser Aufgabenauswahl zu gelangen, werden Schülerinnen und Schüler ausgewählt, die folgende mittlere Leistungswerte erreichten: 395–405, 470–480, 545–555, 620–630. Die Leistungen dieser Schülerinnen und Schüler lassen sich damit im Bereich von fünf Punkten unter bis fünf Punkte über einem *Benchmark*-Wert verorten. Diese Zuordnung ist aufgrund eines Verfahrens zulässig, das die Darstellung von Schülerfähigkeiten und Aufgabenschwierigkeiten auf einer gemeinsamen Skala erlaubt (siehe Abschnitt 9). Zur Beschreibung mittlerer *Benchmarks* werden Aufgaben herangezogen, die von mindestens 65 Prozent der zugeordneten Schülerinnen und Schüler gelöst werden, zugleich aber von weniger als 50 Prozent der Schülerinnen und Schüler der darunter liegenden *Benchmark*. Für die Kompetenzstufen am oberen und unteren Ende der Leistungsskala gelten diesem Vorgehen folgend leicht angepasste Kriterien (Mullis & Martin, 2015). In Tabellen und Abbildungen dieses Bandes wird für Kompetenzstufen die Abkürzung ‚KS‘ genutzt, wo dies der besseren Darstellbarkeit dient.

Abbildung 2.6: Beziehung von Benchmarks und Kompetenzstufen



IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Standardfehler und Signifikanz

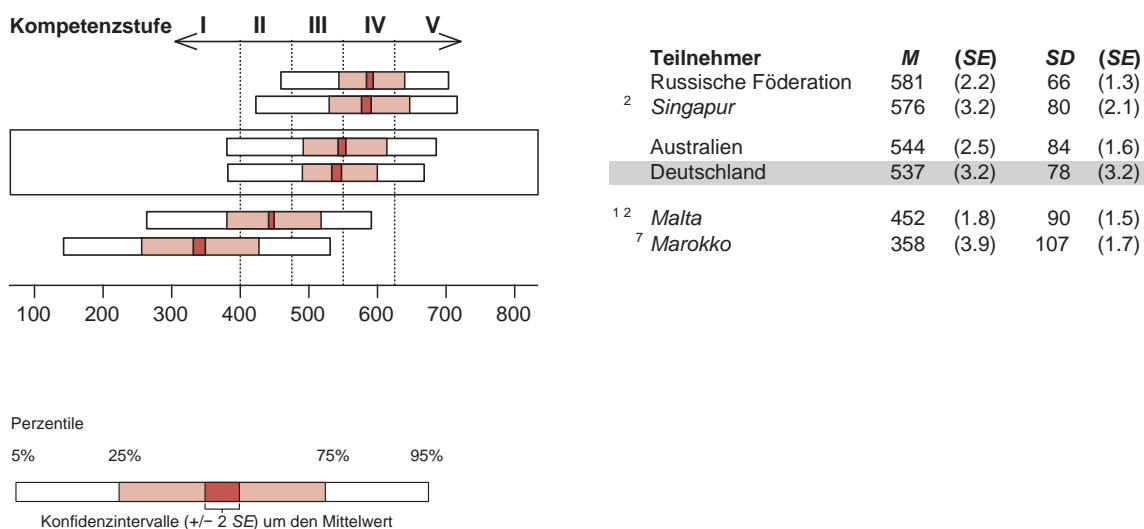
Die Repräsentativität einer Stichprobe, wie sie in IGLU gegeben ist, stellt sicher, dass die anhand dieser Stichprobe erlangten Erkenntnisse auf die Grundgesamtheit übertragen werden können. Es gibt jedoch in jeder Stichprobe kleinere oder größere Abweichungen von der Grundgesamtheit. In der Regel ist die Abweichung umso geringer, je größer die gewonnenen Stichproben ausfallen. Berechnet man beispielsweise den Mittelwert für jede dieser Stichproben, lässt sich für diese Mittelwerte ähnlich wie für die Messwerte innerhalb einer Stichprobe eine Streuung berechnen. Diese Streuung beziehungsweise das berechnete Streuungsmaß bezeichnet den Standardfehler (*SE* für *Standard Error*) oder auch Standardschätzfehler.

Der Standardfehler quantifiziert die Unsicherheit der aus Stichprobendaten geschätzten Populationswerte. Je geringer er ausfällt, desto genauer ist der geschätzte Populationskennwert. Für den Standardfehler gilt, dass im Bereich von rund zwei Standardfehlern (1.96) unter beziehungsweise über einem errechneten

Mittelwert mit 95-prozentiger Wahrscheinlichkeit der ‚wahren‘ Populationswert liegt.

Für die Interpretation der Ergebnisse stellt der Standardfehler ein wichtiges Maß dar. Er dient der Einschätzung, ob sich zwei oder mehr Gruppen signifikant voneinander unterscheiden. Im vorliegenden Bericht werden (wie auch in den vergangenen Berichten zu IGLU 2001, 2006 und 2011) Perzentilbänder zur Ergebnisdarstellung angegeben, denen Konfidenzintervalle der Mittelwerte und dabei berechnete Standardfehler zugrunde liegen. Die Konfidenzintervalle (Vertrauensintervalle) der Mittelwerte (siehe Abbildung 2.7) geben an, in welchem Wertebereich der wahre Populationswert mit 95-prozentiger Wahrscheinlichkeit liegt. Zugrunde gelegt wird das Intervall von über beziehungsweise unter 1.96 Standardfehlern um die jeweilige Statistik. Überlappen sich beispielsweise zwei Konfidenzintervalle nicht, wie in Abbildung 2.7 die Konfidenzintervalle um die mittlere Leseleistung der Russischen Föderation oder Singapur und Deutschland, entspricht dies einem signifikanten Unterschied mit einer Irrtumswahrscheinlichkeit von $\alpha = .05$.

Abbildung 2.7: Darstellung von Perzentilbändern mit Konfidenzintervallen am Beispiel der Leseleistung in IGLU 2016



□ Nicht statistisch signifikant vom deutschen Mittelwert abweichende Staaten ($p > .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

7 = Teilnahme an IGLU 2016 und PIRLS Literacy (Iran, Marokko) bzw. ausschließlich an PIRLS Literacy (Dänemark, 3. Jgst.).

Die Kennwerte für Iran und Marokko werden in Anlehnung an die internationale Berichterstattung als Mittelwert der beiden Studien dargestellt.

Mehrebenenanalysen

Die Betrachtung der Leseleistung im Trend zeigt für die einzelnen teilnehmenden Staaten und Regionen unterschiedliche Verläufe; bei manchen Staaten und Regionen lassen sich positive Veränderungen nachzeichnen, bei anderen keine oder gar negative (siehe Kapitel 3 in diesem Band). Über die Zeit haben sich neben der Testleistung aber auch die Rahmenbedingungen (z.B. Anteil der Schülerinnen und Schülern mit Migrationshintergrund in der Population) geändert. Eine Möglichkeit, die Bedeutung dieser Veränderungen in den Rahmenbedingungen für die Leistungsentwicklungen von Viertklässlerinnen und Viert-

klässlern im Lesen abzuschätzen und damit quasi einen Nettoeffekt der Trendentwicklung zu ermitteln, stellt die Anwendung von Trendmodellen dar. Ermittelt wird die Leistungsdifferenz zwischen den Studienzyklen, die sich ergeben hätte, wenn die Verteilung der soziodemografischen Merkmale in der Schülerschaft gleichgeblieben wäre. Dieses Verfahren wurde beispielsweise von der OECD in ihrem internationalen Bericht über Trends in der Lesekompetenz zwischen 2000 und 2009, die im Rahmen des *Programme for International Student Assessment* (PISA) ermittelt wurden, angewendet (OECD, 2010, S. 49 ff.; Ehmke, Klieme & Stanat, 2013). Um die Leistungsveränderungen über die Zeit nach Kontrolle von Merkmalen zu berücksichtigen, wurde pro teilnehmendes Land eine lineare Regression für die Schülerleistungen bestimmt. Als erklärende Faktoren der Leistung wurden der Messzeitpunkt, die soziodemografischen Merkmale und die Interaktionen zwischen den soziodemografischen Merkmalen und dem Messzeitpunkt berücksichtigt. In einem ähnlichen Vorgehen wurden auch im Rahmen der nationalen Berichtslegung zu TIMSS 2015 Trendanalysen für Deutschland vorgestellt (vgl. Wendt et al., 2016; darin insbesondere Kasper, Wendt, Bos & Köller, 2016).

In diesem Berichtsband wird zur Betrachtung der Leistungsentwicklung im Trend in Deutschland ein methodisches Vorgehen gewählt, das demjenigen der OECD-Studie und demjenigen in TIMSS 2015 in wesentlichen Teilen entspricht. Wie schon für die Trendanalysen in TIMSS 2015 für Deutschland wurde in Abweichung zu dem Vorgehen in der OECD-Studie kein lineares Regressionsmodell verwendet, sondern ein lineares Modell mit Zufallseffekten (ein sogenanntes Mehrebenenmodell; McCulloch, Searle & Neuhaus, 2008). Die Verwendung dieses Modells anstelle der linearen Regression scheint in dem vorliegenden Kontext, wie auch in TIMSS 2015, aus methodischer Sicht geboten. Eine Anwendung der linearen Regression würde voraussetzen, dass die Leistungen der Schülerinnen und Schüler nach Kontrolle aller erklärenden Variablen unabhängig voneinander sind (Werner, 1997). Allerdings ist aus der empirischen Bildungsforschung bekannt, dass die individuelle Schülerleistung auch vom Klassen- beziehungsweise Schulkontext abhängt. Die Leistungen von Schülerinnen und Schülern innerhalb einer Klasse und Schule sind also im Allgemeinen nicht unabhängig voneinander. Da die Schülerinnen und Schüler in der IGLU-Stichprobe im Klassenverbund getestet werden (siehe Abschnitt 4.3), ist davon auszugehen, dass auch ihre Leistungen in der IGLU-Stichprobe (ohne Kontrolle der Klassen- und Schulzugehörigkeit) voneinander abhängig sind. Das verwendete Mehrebenenmodell berücksichtigt diese Abhängigkeiten in den Schülerleistungen: Indem die klassenbedingten Gemeinsamkeiten in der Leistung durch einen Interceptwert, der allen Schülerinnen und Schülern einer Klasse gemeinsam ist, modelliert und repräsentiert werden, können die Abhängigkeiten zwischen den Schülerleistungen kontrolliert werden. Da die untersuchten Klassen außerdem eine Zufallsauswahl aus allen zur definierten Zielpopulation gehörenden Klassen in Deutschland darstellen, konnte die konkrete Ausprägung des Interceptwerts als zufällig betrachtet werden, weswegen ein sogenanntes *random intercept model* berechnet wurde (Raudenbush & Bryk, 2002).

Anders als in der OECD-Studie, aber vergleichbar zum Vorgehen in TIMSS 2015, basieren die Mehrebenenmodelle auf nicht imputierten Datensätzen. Dieses Vorgehen ermöglicht es, dass die Ergebnisse der hier berichteten Analysen mit den Trendergebnissen im internationalen Vergleich aus den Kapiteln in diesem Band abgeglichen werden können: Abweichungen zwischen den Trendschätzern in diesem und den anderen Kapiteln aufgrund eines unterschiedlichen Umgangs

mit fehlenden Werten sind demnach auszuschließen. In Bezug auf den Studienzyklus 2011 wurde der kombinierte Datensatz aus TIMSS und IGLU verwendet, um eine Vergleichbarkeit zu den Analysen der im Rahmen der nationalen Berichterstattungen publizierten Ergebnisse zu den Hintergrundmerkmalen der Schülerinnen und Schüler gewährleisten zu können (vgl. Bos, Tarelli et al., 2012; Bos, Wendt et al., 2012).

Das hier beschriebene Verfahren entspricht weitestgehend dem methodischen Vorgehen von Van Damme und Bellens (2016), die die Trendentwicklung in TIMSS für ausgewählte Länder analysiert haben. Die Mehrebenenanalysen wurden mit dem Softwareprogramm SAS/STAT Software, Version 9.4 (TS1M1) von SAS System für Windows durchgeführt. Durch die Implementierung einer eigens für diese Analysen entwickelten Funktion in dem Programm konnten die *Sampling*-Varianz und die Varianz zu Lasten der *Plausible Values* bei den Analysen berücksichtigt werden.

Staatenvergleiche

Durch die weltweite Beteiligung von 57 Bildungssystemen an IGLU 2016 (siehe Abschnitt 4.2.1), eröffnen sich zahlreiche Möglichkeiten, die Leistungen der Schülerinnen und Schüler in Deutschland mit denen von Kindern in anderen Staaten und Regionen zu vergleichen. Zu klären ist dabei immer die Relevanz solcher Vergleiche sowie die Frage, in welcher Weise Vergleiche sinnvoll sind oder ob sie aufschlussreiche Informationen liefern. So ist etwa zu fragen, in welchem Ausmaß ein Vergleich von Bildungssystemen überhaupt tragfähig ist. Einen Überblick zu dieser Thematik bietet die von der internationalen Studienleitung veröffentlichte *Enzyklopädie* (siehe Abschnitt 4.1.2), in der nachzulesen ist, wie die Bildungssysteme der Teilnehmer strukturiert und inhaltlich ausgestaltet sind (Mullis et al., 2017). Die Beschreibung der Bildungssysteme folgt einer einheitlich vorgegebenen Gliederung, so dass relevante Informationen systematisch strukturiert für alle Teilnehmer vorliegen (für Deutschland siehe Wendt, Walzebug, Bos, Smith und Bremerich-Vos, 2017). Neben inhaltlichen Fragen ist des Weiteren zu klären, ob unter Berücksichtigung nationaler Besonderheiten der Stichproben (siehe Abschnitt 4.3) der Vergleich einzelner Staaten zulässig ist. So ist ein Vergleich mit Regionen wie den Benchmark-Teilnehmern nur unter Einschränkungen möglich (siehe Abschnitt 4.2.1), da bei Benchmark-Teilnehmern Besonderheiten eines Schulwesens im Vordergrund stehen könnten, die möglicherweise nicht für einen ganzen Staat zutreffen. Die internationale Studienleitung weist deshalb in Absprache mit den nationalen Studienleitungen Benchmark-Teilnehmer stets getrennt in den Ergebnisdarstellungen aus.

Um sinnvolle Vergleiche zu Deutschland in der vorliegenden Berichterstattung dokumentieren zu können, werden in IGLU Vergleichsgruppen gebildet (siehe Abschnitt 4.2.1, Tabelle 2.1). Ihnen gehören Staaten oder Staatengruppen an, die sich hinsichtlich des kulturellen und ökonomischen Hintergrunds oder der wirtschaftlichen Situation ähneln und sich daher für einen Staatenvergleich mit Deutschland eignen. Aus bewährten Gründen (vgl. Bos et al., 2003; Bos et al., 2007; Bos, Tarelli et al., 2012) stehen auch in diesem Bericht zu IGLU 2016 die Vergleichsgruppen der Teilnehmerstaaten der EU (VG EU) und die der OECD (VG OECD) zur Verfügung.

Die Darstellung der Ergebnisse in den nachfolgenden Kapiteln orientiert sich an den vorgestellten Entscheidungen: In Kapitel 3 werden die Ergebnisse für alle an IGLU 2016 teilnehmenden Staaten und Regionen auf der Gesamtskala Lesen dargestellt. Die Benchmark-Teilnehmer werden graphisch am unteren

Ende der aufgelisteten Teilnehmer durch einen kleinen Absatz getrennt aufgelistet. Darunter findet sich auch Dänemark mit Jahrgangsstufe 3 als Benchmark-Teilnehmer der Studienkomponente *PIRLS Literacy* (siehe Abschnitt 4.2.1). Die Leistungswerte der Benchmark-Teilnehmer fließen aus den bereits genannten Gründen nicht in die Berechnung des internationalen Mittelwerts ein. Bei allen anderen Ergebnissen, die in Tabellen und Abbildungen zu IGLU 2016 dokumentiert werden, werden diejenigen Teilnehmer und Benchmark-Teilnehmer berichtet, auf die mindestens eins der folgenden drei Kriterien zutrifft: (1) Mitglied der EU, (2) Mitglied der OECD und/oder (3) im Vergleich zu Deutschland signifikant bessere oder nicht signifikant unterschiedliche Leistung auf der Gesamtskala Lesen. Bei der Dokumentation von Trends im internationalen Vergleich werden diejenigen Bildungssysteme berücksichtigt, die (1) zusätzlich zu IGLU 2016 an mindestens zwei weiteren IGLU-Zyklen teilgenommen haben und (2) zugleich zum Erhebungszeitpunkt von IGLU 2016 Mitglied der EU und/oder der OECD sind und/oder auf der Gesamtskala Lesen in IGLU 2016 signifikant bessere oder nicht signifikant unterschiedliche Leistungen im Vergleich zu Deutschland erzielt haben. Eine Ausnahme ist die Flämische Gemeinschaft in Belgien, die in Trenddarstellungen berücksichtigt wird, obwohl sie neben IGLU 2016 nur an einem weiteren Zyklus (IGLU 2006) teilgenommen hat (siehe Abschnitt 4.2.2). In keiner Trenddarstellung werden Mittelwerte der Vergleichsgruppen VG EU und VG OECD berichtet, da sich über die vier Zyklen die Zusammensetzungen der Vergleichsgruppen durch andere Studienteilnehmer und neue Mitgliedschaften geändert haben.

Literatur

- Anderson, R. & Pearson, P. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. Pearson (Hrsg.), *Handbook of reading research* (S. 255–291). White Plains, NY: Longman.
- Autorengruppe Bildungsberichterstattung (2016). *Bildung und Migration*. Zugriff am 18.11.2017 unter https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/h_web2016.pdf
- Baumert, J. (2016). Leistungen, Leistungsfähigkeit und Leistungsgrenzen der empirischen Bildungsforschung. *Zeitschrift für Erziehungswissenschaft*, 19 (1), 215–253.
- Baumert, J. & Weiß, M. (2002). Föderalismus und Gleichwertigkeit der Lebensverhältnisse. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000 – Die Länder der Bundesrepublik im Vergleich* (S. 39–53). Opladen: Leske + Budrich.
- BMBF – Bundesministerium für Bildung und Forschung (Hrsg.). (2001). *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente*. Bonn: BMBF Publik.
- Boomsma, A., van Duijn, M. A. J. & Snijders, T. A. B. (Hrsg.). (2000). *Essay on item response theory*. New York: Springer.
- Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C. & Walther, G. (Hrsg.). (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2007). *IGLU 2006: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R. & Walther, G. (Hrsg.). (2004). *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.

- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schulleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Valtin, R. & Walther, G. (Hrsg.). (2005). *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien*. Münster: Waxmann.
- Bos, W., Postlethwaite, T. N. & Gebauer, M. M. (2010). Potentiale, Grenzen und Perspektiven internationaler Schulleistungsforschung. In R. Tippelt & B. Schmidt (Hrsg.), *Handbuch Bildungsforschung* (S. 275–295). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.). (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.). (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Campbell, J. R., Kelly, D. L., Mullis, I. V. S., Martin, M. O. & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001* (2. Aufl.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Chall, J. (1983). *Stages of reading development*. New York: McGraw-Hill.
- Christmann, U. & Groeben, N. (2001). Psychologie des Lesens. In B. Franzmann, K. Hasemann, D. Löffler & E. Schön (Hrsg.), *Handbuch Lesen* (S. 145–207). Baltmannsweiler: Schneider Hohengehren.
- Ehmke, T., Klieme, E. & Stanat, P. (2013). Veränderungen der Lesekompetenz von PISA 2000 und PISA 2009. Die Rolle von Unterschieden in den Bildungswegen und in der Zusammensetzung der Schülerschaft. *Zeitschrift für Pädagogik*, 59, 132–150.
- Elley, W. B. (1992). *How in the world do students read? IEA study of reading literacy*. Den Haag: IEA.
- Elley, W. B. (1994). *The IEA study of reading literacy: Achievement and instruction in thirty-two school systems*. Oxford: Pergamon.
- Foy, P., Brossman, B. & Galia, J. (2011). Scaling TIMSS and PIRLS 2011 achievement data. In M. O. Martin & I. V. S. Mullis (Hrsg.), *TIMSS and PIRLS methods and procedures*. Zugriff am 17.10.2016 unter <http://timssandpirls.bc.edu/methods/index.html>
- Heller, K. A. & Perleth, C. (2000). *KFT 4–12+R. Kognitiver Fähigkeitstest für 4. bis 12. Klasse, Revision*. Göttingen: Beltz Test.
- Hooper, M., & Fishbein, B. (2017). Developing the PIRLS 2016 context questionnaires. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Hrsg.), *Methods and procedures in PIRLS 2016*. Zugriff am 18.11.2017 unter <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-2.html>
- Hooper, M., Mullis, I. V. S. & Martin, M. O. (2015). PIRLS 2016 context questionnaire framework. In I. V. S. Mullis & M. O. Martin (Hrsg.), *PIRLS 2016 assessment framework* (2. Auflage, S. 31–54). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Howie, S. & Plomp, T. (2005). International comparative studies of education and large-scale change. In N. Bascia, A. Cumming, A. Datnow, K. Leithwood & D. Livingstone (Hrsg.), *International handbook of educational policy*. (Springer international handbooks of education, Bd. 13, S. 75–99). Berlin: Springer.
- Hußmann, A., Wendt, H., Bos, W. & Rieser, S. (Hrsg.). (2018, i. Vorb.). *IGLU 2016. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster: Waxmann.
- Joncas, M. & Foy, P. (2011). Sample design in TIMSS and PIRLS. In M. O. Martin & I. V. S. Mullis (Hrsg.), *TIMSS and PIRLS methods and procedures*. Zugriff am 16.11.2012 unter http://timssandpirls.bc.edu/methods/pdf/TP_Sampling_Design.pdf
- IEA – International Association for the Evaluation of Educational Achievement. (2012). *Studies*. Zugriff am 16.11.2012 unter <http://www.iea.nl/studies.html>
- Kasper, D., Wendt, H., Bos, W. & Köller, O. (2016). Trends in mathematischen und naturwissenschaftlichen Kompetenzen am Ende der Grundschulzeit in Deutschland. 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. 367–382). Münster: Waxmann.

- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kintsch, W. (2012). Psychological models of reading comprehension and their implications for assessments. In J. Sabatini, E. Albro & T. O'Reilly (Hrsg.), *Measuring up: Advances in how to assess reading ability* (S. 21–37). Plymouth: Rowman & Littlefield Publishers.
- Kintsch, W. (2013). Revisiting the construction-integration model of text comprehension and its implications for instruction. In D. Alvermann, N. Unrau & R. Ruddell (Hrsg.), *Theoretical models and processes of reading* (S. 807–841). Newark, DE: International Reading Association.
- Klieme, E. (2013). Bildung unter undemokratischem Druck? Anmerkungen zur Kritik der PISA-Studie. In S. Lin-Klitzing, D. Di Fuccia & G. Müller-Frerich (Hrsg.), *Zur Vermessung von Schule* (S. 37–51). Bad Heilbrunn: Klinkhardt.
- Klieme, E. & Vieluf, S. (2013). Schulische Bildung im internationalen Vergleich. Ein Rahmenmodell für Kontextanalysen in PISA. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (Zeitschrift für Pädagogik. Beiheft 59, S. 229–246). Weinheim: Beltz.
- KMK – 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 17.10.2016 unter <https://www.kmk.org/themen/qualitaetsicherung-in-schulen/bildungsmonitoring.htm>
- Kolen, M. J. (1981). Comparison of traditional and item response theory methods for equating tests. *Journal of Educational Measurement*, 18 (1), 1–11.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking. Methods and practices* (2. Aufl.). New York: Springer.
- Lankes, E.-M., Bos, W., Mohr, I., Plaßmeier, N., Schwippert, K., Sibberns, H. & Voss, A. (2003). Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU) und ihrer Erweiterung um Mathematik und Naturwissenschaften (IGLU-E). In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 7–28). Münster: Waxmann.
- Lehmann, R. H., Peek, R., Pieper, I. & Stritzky, R. v. (1995). *Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen*. Weinheim: Beltz.
- Lenhard, W., Lenhard, A. & Schneider, W. (2017). *ELFE II. Ein Leseverständnistest für Erst- bis Siebtklässler – Version II*. Göttingen: Hogrefe.
- Linden, W. v. d. & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York: Springer.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- Martin, M. O., Mullis, I. V. S. & Foy, P. (2015). Assessment design for PIRLS, PIRLS Literacy, and ePIRLS in 2016. In I. V. S. Mullis & M. O. Martin (Hrsg.), *PIRLS 2016 assessment framework* (2. Auflage, S. 55–70). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S., Foy, P. & Stanco, G. M. (2012). *TIMSS 2011 international results in science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S., Gonzalez, E. J. & Kennedy, A. M. (2003). *Trends in children's reading literacy achievement 1991–2001: IEA's study of trends in reading literacy achievement in primary school in nine countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S. & Hooper, M. (Hrsg.). (2017). *Methods and procedures in PIRLS 2016*. Zugriff am 18.11.2017 unter <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>
- McCulloch, C. E., Searle, S. R. & Neuhaus, J. M. (2008). *Generalized, linear, and mixed models*. Hoboken, NJ: Wiley.
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177–196.

- Mislevy, R. J., Beaton, A. E., Kaplan, B. & Sheehan, K. M. (1992). Estimation population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29 (2), 133–161.
- Mullis, I. V. S., Kennedy, A. M., Martin, M. O. & Sainsbury, M. (Hrsg.). (2006). *PIRLS 2006. Assessment framework and specifications* (2. Aufl.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O. & Sainsbury, M. (2015). PIRLS 2016 reading framework. In I. V. S. Mullis & M. O. Martin (Hrsg.), *PIRLS 2016 assessment framework* (2. Auflage, S. 11–30). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S. & Martin, M. O. (Hrsg.). (2015). *PIRLS 2016 assessment framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Drucker, K. T. (2012). *PIRLS 2011 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Goh, S. & Prendergast, C. (Hrsg.). (2017). *PIRLS 2016 Encyclopedia: Education policy and curriculum in reading*. Zugriff am 18.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/encyclopedia/>
- Mullis, I. V. S., Martin, M. O., Gonzalez, E. J. & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary schools*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M. & Foy, P. (2007). *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary school in 40 countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L. & Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., & Prendergast, C. O. (2017). Developing the PIRLS 2016 achievement items. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Hrsg.), *Methods and Procedures in PIRLS 2016* (S. 1.1–1.29). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Zugriff am 20.11.2017: <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-1.html>
- Muraki, E. & Bock, D. (1999). *PARSCALE 3.5: IRT item analysis and test scoring for rating-scale data* [Computer software]. Chicago, IL: Scientific Software,
- OECD – Organisation for Economic Co-operation and Development. (2010). *PISA 2009 assessment framework. Key competencies in reading, mathematics and science*. Paris: OECD.
- Prenzel, M. & Doll, J. (2002). Bildungsqualität von Schulen: Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen [Beiheft]. *Zeitschrift für Pädagogik*, 45.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods*. London: Sage Publication.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley + Sons.
- Ruddell, R. & Unrau, N. (Hrsg.). (2004). Read as a meaning-construction process: The reader, the text, and the teacher. In R. Ruddell & N. Unrau (Hrsg.), *Theoretical models and processes of reading* (5. Auflage, S. 1462–1521). Newark, DE: International Reading Association.
- Rumelhart, D. (1985). Toward an interactive model of reading. In H. Singer & R. Ruddell (Hrsg.), *Theoretical models and the processes of reading* (3. Auflage, S. 722–750). Newark, DE: International Reading Association.
- Sheehan, K. M. (1985). *M-GROUP. Estimation of group effects in multivariate models* [Computer software]. Princeton, NJ: Educational Testing Service.
- Statistisches Bundesamt (2017). *Bildung und Kultur. Allgemeinbildende Schulen*. Fachserie 11, Reihe 1, 2016/2017. Zugriff am 18.11.2017 unter https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100177004.pdf;jsessionid=29D3FD1BA09C7F36C1A0202842B06C01.InternetLive1?__blob=publicationFile

- Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S. & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- UIS – UNESCO Institute for Statistics. (2015). *ISCED 2011 operational manual. Guidelines for classifying national education programmes and related qualifications*. Zugriff am 03.08.2016 unter <http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx>
- Tarelli, I., Wendt, H., Bos, W. & Zyłowski, 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 Damme, J. & Bellens, K. (2016). Countries strive towards more quality and equity in education: Do they show success or failure? Evidence from TIMSS 2003 and 2011, for Grade 4. In M. Rosén, K. Yang Hansen & U. Wolff (Hrsg.), *Cognitive abilities and educational outcomes. A Festschrift in honour of Jan-Eric Gustafsson* (S. 127–148). Cham: Springer.
- Wang, M., Haeterl, G. & Walberg, H. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63 (3), 249–294.
- Weinert, F. E. & Helmke, A. (Hrsg.). (1997). *Entwicklung im Grundschulalter*. Weinheim: Beltz.
- 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* (S. 31–78). Münster: Waxmann.
- Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K. & Kasper, D. (Hrsg.). (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Wendt, H., Bos, W., Tarelli, I., Vaskova, A. & Walzebug, A. (Hrsg.). (2016). *IGLU & TIMSS 2011. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster: Waxmann.
- Wendt, H., Stubbe, T. C., Schwippert, K. & Bos, W. (Hrsg.). (2015). *10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011*. Münster: Waxmann
- Wendt, H., Tarelli, I., Bos, W., Frey, K. & Vennemann, M. (2012). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2011). In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–68). Münster: Waxmann.
- Wendt, H., Walzebug, A., Bos, W., Smith, D. S. & Bremerich-Vos, A. (2017). Germany. In I. V. S. Mullis, M. O. Martin, S. Goh & C. Prendergast (Hrsg.), *PIRLS 2016 encyclopedia: Education policy and curriculum in reading*. Zugriff am 18.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/encyclopedia/countries/germany/>
- Werner, J. (1997). *Lineare Statistik*. Weinheim: Beltz.
- Yu, A. & Ebbs, D. (2011). Translation and translation verification. In M. O. Martin & I. V. S. Mullis (Hrsg.), *TIMSS and PIRLS methods and procedures*. Zugriff am 16.11.2012 unter http://timssandpirls.bc.edu/methods/pdf/TP_Translation_Verif.pdf

Kapitel III

Lesekompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse

Albert Bremerich-Vos, Heike Wendt und Wilfried Bos

1 Einleitung

Der Hauptauftrag der Grundschule besteht darin, für alle Kinder Bedingungen zu schaffen, die für den Erwerb grundlegender Bildung möglichst günstig sind. Dabei spielt die Sprache als Gegenstand des Nachdenkens und als Mittel, die Welt und sich selbst zu verstehen und sich mit anderen zu verständigen, eine zentrale Rolle. Mit „Sprache“ ist nicht nur Deutsch gemeint, sondern die Vielfalt der Sprachen, die von den Kindern in den Unterricht mitgebracht werden. In einer schriftgeprägten Umgebung eignen sich Kinder Schrift zunächst in Form von Logogrammen an. Sie identifizieren Geschriebenes als Bild, orientieren sich also an hervorstechenden visuellen Merkmalen, „lesen“ also zum Beispiel „ARAL“ als Symbol für „Tankstelle“. Im Rahmen des schulischen Schriftspracherwerbs wird diese Art des „Lesens“ abgelöst durch alphabetisches und orthographisches Lesen. Auf der Basis der Kenntnis von Buchstaben-Laut-Zuordnungen werden Wörter und größere Einheiten nun sukzessive von links nach rechts gelesen, und zunehmend sind die Kinder nicht mehr darauf angewiesen, lautierend oder silbenorientiert zu lesen, sondern die Zielwörter werden mehr und mehr direkt erfasst. Orthographische Merkmale wie die Verdopplung von Konsonantenbuchstaben und die Großschreibung werden als das Lesen erleichternd erkannt und genutzt. Gelingt der Leseprozess, können die Schülerinnen und Schüler am Ende der Grundschulzeit längere Texte selbstständig lesen und verstehen und an Gesprächen über Texte kundig teilnehmen. So lassen sich neue Aspekte der „Welt“ erschließen, im Umgang mit literarischen Texten zum Beispiel bislang fremde Verhaltensweisen und Gefühle. Beim Lesen von Sachtexten und Text-Bild-Kombinationen kann Wissen vertieft und es können neue Wissensbereiche erschlossen werden. Auch die ästhetische Dimension des Geschriebenen kann in den Blick geraten, zum Beispiel Gespür für ungewöhnlichen Sprachgebrauch und symbolische Bedeutungen entwickelt werden (Spinner, 2016). Die Erfahrung,

dass Lesen als solches nicht zuletzt in emotionaler Hinsicht befriedigend sein kann, trägt zu einer positiven Einstellung zum Lesen bei, die wiederum die Leseleistung befördert.

Mit der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) wird das Leseverständnis von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe alle fünf Jahre im internationalen Vergleich untersucht. Deutschland beteiligt sich seit 2001 regelmäßig an der Studie. In diesem Kapitel werden anschließend an eine Erläuterung der Testkonzeption die Ergebnisse der Erhebung aus dem Jahr 2016 berichtet und mit Ergebnissen aus früheren Erhebungsrounden verglichen.

Das Kapitel ist wie folgt gegliedert: Zunächst wird die der IGLU zugrunde liegende Rahmenkonzeption zur Erfassung der Lesekompetenz skizziert, wobei nach Leseintentionen beziehungsweise Textsorten und Verstehensleistungen beziehungsweise -prozessen differenziert wird (siehe Abschnitt 2). In Abschnitt 3 wird auf den IGLU-Test und seine curriculare Validität eingegangen. Hier wird auch über Ergebnisse eines in Deutschland zusätzlich eingesetzten Tests berichtet, mit dem Leistungen besonders schwacher Leserinnen und Leser erfasst wurden. Abschnitt 4 ist dem Kompetenzstufenmodell von IGLU 2016 gewidmet. Es wird anhand von Aufgaben zu einem literarischen Text und einem Sachtext illustriert. Im fünften Abschnitt werden die Ergebnisse von IGLU 2016 im Einzelnen dargelegt. Die hier leitenden Fragestellungen sind zu Beginn dieses Abschnitts aufgeführt. Besonderer Wert wird auf die Darstellung von Trends gelegt, ist die Erhebung 2016 doch die vierte nach IGLU 2001, 2006 und 2011.

2 Die IGLU-Rahmenkonzeption zur Erfassung von Lesekompetenz

Lesekompetenz (*reading literacy*) wird verstanden als Fähigkeit, gesellschaftlich und/oder individuell wertgeschätzte Texte verschiedener Art (bzw. Textsorten oder -muster) zu verstehen und zu nutzen. Verstehen wird als Konstruktion von Bedeutung konzipiert, wobei drei zentrale Absichten unterschieden werden: um zu lernen, Vergnügen zu haben und an der Gemeinschaft der Lesenden innerhalb und außerhalb der Schule teilzuhaben (Mullis, Martin & Sainsbury, 2015). Diese Auffassung von Lesekompetenz steht im Einklang mit verschiedenen aktuell diskutierten Theorien, wonach das Verstehen von Texten als konstruktiv-interaktiver Prozess zu begreifen ist (z.B. Kintsch, 2012). Im Umgang mit Texten aktivieren Leserinnen und Leser sprachliches Wissen, mehr oder weniger automatisierte kognitive und metakognitive Strategien und (bereichsspezifisches) Vorwissen beziehungsweise „Weltwissen“. Sie haben es mit einer Vielzahl an Textsorten zu tun, in tradierten Papierformaten und in wachsendem Maß auch mit digitalen Versionen. In der Definition von Lesekompetenz wird nicht nur die individuelle, sondern auch die soziale Dimension des Lesens berücksichtigt. Im Rahmen von Anschlusskommunikation innerhalb und außerhalb der Schule können Schülerinnen und Schüler ihr Verständnis des jeweiligen Textes zur Diskussion stellen, andere Perspektiven kennenlernen und ihr primäres Verständnis in Reaktion auf die Beiträge anderer vertiefen oder auch korrigieren.

2.1 Leseintentionen beziehungsweise Textsorten

Für Viertklässlerinnen und Viertklässler sind in erster Linie zwei der genannten Ziele beziehungsweise Zwecke mit dem Lesen verbunden: Lesen aus Interesse beziehungsweise zum Vergnügen und Lesen, um zu lernen. In der IGLU-Rahmenkonzeption wird angenommen, dass insbesondere fiktional-narrative Texte mit dem Ziel gelesen werden, sich unterhalten zu lassen. Das schließt nicht aus, dass zumindest einige der Textsorten, die hier in Frage kommen, durchaus auch explizit belehrenden Charakter haben, zum Beispiel bestimmte Fabeln oder auch Kalendergeschichten. In fiktional-narrativen Texten werden Sachverhalte als wirklich dargestellt, ohne dass ein Bezug dieser Darstellung auf unsere Wirklichkeit beansprucht wird. Leserinnen und Leser solcher Texte werden involviert in Ereignisse und Schauplätze, setzen sich mit Figuren und deren Motiven, Gedanken und Gefühlen auseinander, ihren Handlungen, deren Konsequenzen und so weiter, aber auch mit dem Wie des Erzählens, rhetorischen und stilistischen Mitteln, der Gestaltung der Erzählstimme, das heißt des „Standorts“ des Erzählers innerhalb oder außerhalb seiner Geschichte, und der Perspektive der Darstellung. Die textseitigen Anforderungen an die Lesenden können beträchtlich variieren. So wird womöglich nicht chronologisch erzählt, sondern es werden Ereignisse vorweggenommen oder der Erzähler geht hinter das bereits Erzählte zurück, die Perspektiven der Darstellung mögen wechseln und so weiter. In die Interaktion mit den Texten bringen Leserinnen und Leser vielfältige Erfahrungen ein, zum Beispiel solche, die mit den Beziehungen zwischen Emotionen und dem Spektrum ihrer Ausdrucksformen zu tun haben, sie verfügen über Kenntnisse von Textsorten beziehungsweise Genres, die sich zum Beispiel in der Erwartung manifestieren, dass sich in einem narrativen Text etwas ereignen wird, was vom normalen Lauf der Dinge abweicht.

In IGLU werden ausschließlich narrative Texte eingesetzt, andere Gattungen bleiben außer Betracht. Die Übertragung lyrischer Texte in viele andere Sprachen ist ausgesprochen schwierig, und auf Spielbarkeit hin angelegte szenische Texte sind zwar in den Curricula der Länder der Bundesrepublik Deutschland genannt, aber in den Lehrplänen anderer Teilnehmerstaaten und -regionen nicht vorgesehen (Mullis, Martin, Goh & Prendergast, 2017).

Was das Lesen mit dem vorrangigen Ziel angeht, sich Informationen anzueignen und sie zu nutzen, so kann man im Hinblick auf Viertklässlerinnen und Viertklässler zugespitzt formulieren: Bislang haben sie primär gelernt, wie man liest, nun sollen sie primär lesen, um zu lernen. Die Textthemen variieren beträchtlich, wie das breite Spektrum der Gegenstände des Sachunterrichts zeigt. Auch die Textsorten sind vielfältig. So wird über Fakten berichtet, Objekte werden beschrieben, es wird instruiert, wie ein Spiel zu gestalten oder ein Gegenstand zu gebrauchen ist, oder es wird argumentiert. Auch wenn es sich um Exemplare derselben Textsorte handelt, können die Anforderungen erheblich differieren, und zwar sowohl in inhaltlicher, struktureller als auch in sprachlicher Hinsicht. So ist die Instruktion, wie man „17 und 4“ zu spielen hat, im Vergleich mit einer Anleitung zum Schachspiel inhaltlich einfach und eine Darstellung der Abfolge von Ereignissen in „mimetischer“, chronologischer Folge ist leichter zu verstehen als etwa in folgender Form: Bevor Y stattfand, ereignete sich X. Informationen werden nicht nur in Form von Fließtexten präsentiert, sondern zum Beispiel auch listenförmig; handelt es sich um Fließtexte, sind sie oft kombiniert mit Tabellen, Graphiken oder Diagrammen. Das erleichtert das Verstehen zum Beispiel dann, wenn im Fließtext dargelegte begriffliche Relationen zu-

sätzlich in Form von Schemata dargeboten werden. Es kann sich aber auch erschwerend auswirken. Das mag zum Beispiel der Fall sein, wenn einer Graphik Informationen abzulesen sind, die im Fließtext selbst überhaupt keine Rolle spielen.

2.2 Verstehensprozesse beziehungsweise -leistungen

Lesen Viertklässlerinnen und Viertklässler fiktionale Erzähltexte und Sachtexte, so können der Rahmenkonzeption von IGLU zufolge vier jeweils breit gefächerte Verstehensprozesse zum Einsatz kommen: (1) Sie konzentrieren sich auf explizit angegebene Informationen und rufen sie ab; (2) sie ziehen einfache Schlussfolgerungen; (3) sie ziehen komplexe Schlussfolgerungen, das heißt, sie interpretieren und verknüpfen textinterne Informationen mit textunabhängig verfügbarem Vorwissen; (4) sie prüfen und bewerten den Inhalt des Textes und Aspekte seiner sprachlichen Gestaltung (Mullis et al., 2015, S. 18 ff.).

Zu (1): Fragt man Kinder, was sie an einem Text besonders wichtig, interessant, auffällig und so weiter finden, dann beziehen sie sich oft auf eine einzelne Information, zum Beispiel ein Merkmal einer Figur, das für das Verständnis des Textes nicht unbedingt relevant ist. Bei den IGLU-Testaufgaben zum Lokalisieren explizit angegebener Informationen geht es hingegen darum, zum Beispiel anzugeben, welche Figuren vorkommen, Fragen nach Ort und Zeit des Geschehens zu beantworten, das zentrale Thema zu nennen, sofern es im Text explizit formuliert ist oder Bedeutungen von Wörtern und Phrasen zu erläutern, falls die Erläuterungen im Text selbst zu finden sind. In erster Linie geht es hier um Einheiten wie Wort, Phrase und Satz, zuweilen müssen auch verstreute Informationen miteinander verknüpft werden. Zentrale Voraussetzung dafür, dass die Kinder mit Aufgaben dieses Zuschnitts kompetent umgehen können, ist, dass sie weitgehend flüssig lesen können.

Zu (2): Wer Texte liest, setzt voraus, dass sie kohärent sind. Wie die inhaltlichen Beziehungen zwischen Sätzen und größeren Einheiten beschaffen sind, kann durch Kohäsionsmittel wie Konjunktionen (*weil*), Adverbien (*deshalb*), Präpositionen (*wegen*), aber auch durch Nomen (*Grund, Folge*) und Verben (*dazu führen, dass ...*) angezeigt werden. Oft fehlen diese Mittel aber, und die Art der inhaltlichen Beziehung muss von der Leserin, dem Leser erschlossen werden. Das ist leicht in einem Fall wie *Emma war krank. Sie ging nicht zur Schule.*, schwieriger bei *Emma war krank. Sie ging zur Schule.* Beim ersten Beispiel ist auf eine kausale Beziehung zu schließen, beim zweiten auf eine konzessive (*Obwohl Emma krank war, ging sie zur Schule.* Oder: *Emma war krank, trotzdem ging sie zur Schule.*). Im ersten Fall vollziehen geübte Leserinnen und Leser den Schluss automatisch, im zweiten dürften auch sie einige Aufmerksamkeit aufwenden müssen. Autorinnen und Autoren beziehungsweise Erzählerinnen und Erzähler muten den Lesenden bewusst einfache Schlüsse zu; deshalb spricht man hier auch von intendierten Inferenzen. So mag sich aus der Folge der Handlungen einer Figur leicht ein Aspekt ihres Charakters erschließen lassen. Im Rahmen einschlägiger Aufgaben haben die Kinder zum Beispiel zu schließen,

- dass ein Ereignis ein anderes verursacht;
- wie eine im Text nicht explizit benannte Beziehung zwischen Figuren beschaffen ist.

Zu (3): Wer liest, bringt seine Erfahrungen in den Prozess ein. So können Lesende auf der Basis von im Text explizit zu findenden Informationen eine literarische Figur mit weiteren Merkmalen ausstatten, die ihnen bekannte Personen haben. Eine solche Interpretation lässt sich als komplexe Schlussfolgerung rekonstruieren, die nicht mehr als vom Autor beziehungsweise Erzähler intendiert, sondern als elaborativ anzusehen ist. Interpretationen können sich zwar auch auf einzelne Textstellen oder -passagen beziehen, sie betreffen aber vor allem das Verständnis des Textganzen. Es ist nicht immer leicht, die elaborativen Schlüsse, die beim Interpretieren im Spiel sind, von einfachen, textbasierten Schlüssen zu unterscheiden. Vorwissen und Erfahrungen spielen beim Interpretieren eine größere Rolle als beim einfachen Schließen, die Unterschiede sind demnach gradueller Natur (Mullis, Martin, Kennedy, Trong & Sainsbury, 2009). Insofern in Interpretationen die je individuellen Perspektiven der Lesenden stärker zum Tragen kommen, ist davon auszugehen, dass ihr Spektrum recht breit ist. Im Rahmen von IGLU-Testaufgaben zum Interpretieren sollen die Viertklässlerinnen und Viertklässler zum Beispiel

- in Kenntnis des Textganzen schließen, wie das Textthema lauten könnte;
- angeben, wie sie die „Stimmung“ oder den „Ton“ einer Geschichte verstehen;
- Alternativen zu Handlungen von Figuren in Erwägung ziehen.

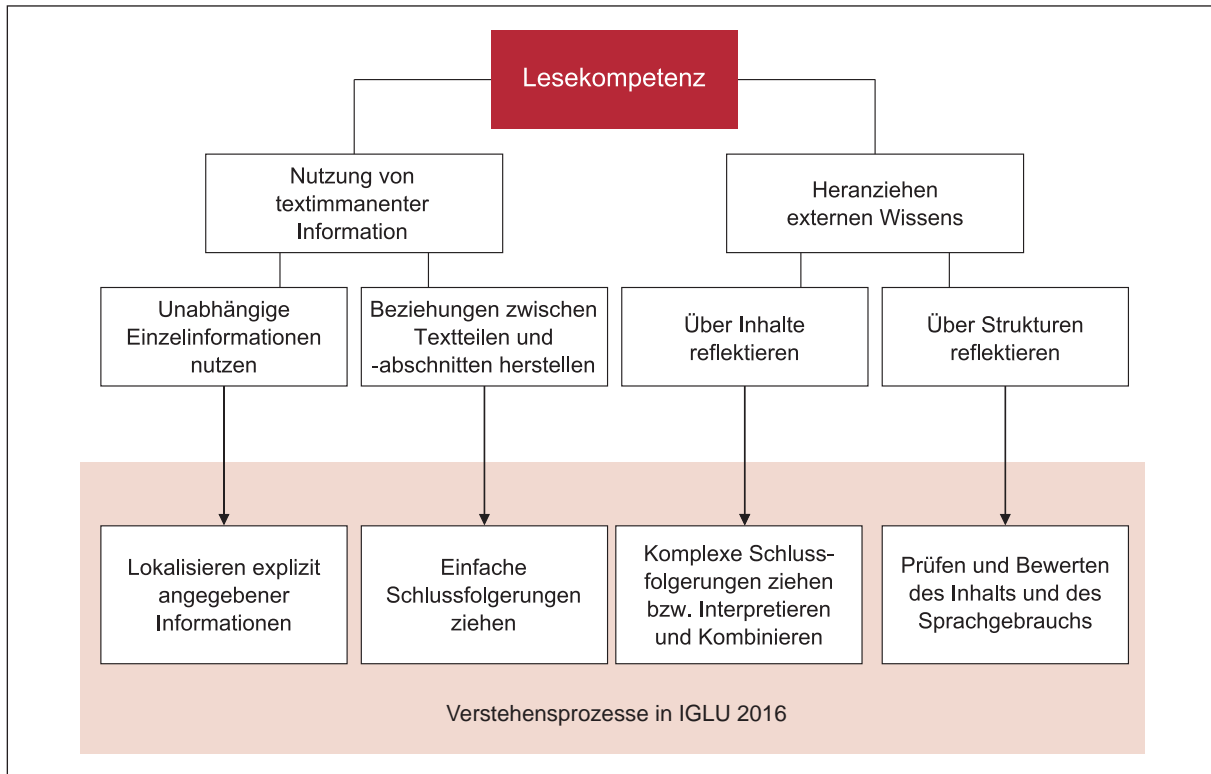
Zu (4): Auch Viertklässlerinnen und Viertklässler sind in der Lage, Aspekte von Texten, soweit sie sie verstanden beziehungsweise interpretiert haben, zu beurteilen und zu bewerten. Handlungsweisen von Figuren in literarischen Texten erscheinen – vor dem Hintergrund des eigenen Weltwissens – als mehr oder weniger angemessen, Beschreibungen werden als zu weitschweifig oder auch als zu detailliert bewertet, Instruktionen als nicht klar genug beurteilt, Berichte als zu stark ausgeschmückt. Die eine oder andere Stützung einer These wird als fragwürdig angesehen, insofern sie Informationen aus anderen, für zuverlässig gehaltenen Quellen widerspricht, und so weiter. Nicht nur Aspekte des Inhalts und der Struktur, sondern auch Merkmale des Sprachgebrauchs können thematisch werden. So mögen Metaphern und Vergleiche auffallen und in einer Reihe von Hinsichten bewertet werden, die Verständlichkeit von einzelnen Sätzen, Abschnitten oder auch des ganzen Textes mag als positiv oder negativ eingestuft werden. Unabhängig davon, ob die Stellungnahmen sich auf Inhaltliches, Strukturelles oder Sprachliches beziehen: Nicht mehr die Konstruktion von Bedeutung steht im Zentrum, sondern die kritische Betrachtung des Textes, bei der auf vielfältige textexterne Ressourcen zurückzugreifen ist. Im Rahmen von IGLU-Testaufgaben, die sich auf diesen Verstehensaspekt beziehen, wird etwa verlangt

- zu beurteilen, ob eine Information ausführlich genug und ob sie klar formuliert ist;
- anzugeben, für wie wahrscheinlich man es hält, dass die im Text dargestellten Ereignisse tatsächlich stattfinden könnten;
- einzuschätzen, ob beziehungsweise inwiefern der Titel eines Textes zu seinem zentralen Thema passt.

Obwohl es von Fall zu Fall nicht einfach ist zu entscheiden, ob für die Lösung einer Aufgabe eine einfache (intendierte) Schlussfolgerung oder eine komplexe (elaborative) Inferenz nötig ist (Ballstaedt, Mandl, Schnotz & Tergan, 1981), lässt es sich rechtfertigen, die vier Verstehensprozesse zwei Klassen zuzuweisen: Lokalisieren und einfache Inferenzen auf der einen Seite können als textbasierte

Verstehensprozesse angesehen werden, elaborative Inferenzen und Bewertungen von Textinhalt, -struktur und sprachlichen Aspekten auf der anderen als vorwissensbasierte Prozesse. Abbildung 3.1 veranschaulicht die im Rahmen von IGLU angenommene Struktur der Lesekompetenz.

Abbildung 3.1: Die theoretische Struktur der Lesekompetenz in IGLU



Es liegt nahe anzunehmen, dass insbesondere das Lokalisieren einer explizit formulierten Information durchgängig einfacher ist als etwa das erfahrungsgestützte Interpretieren größerer Abschnitte oder des ganzen Textes. Diese Annahme ist aber problematisch. Eine Lokalisierungsaufgabe kann zum Beispiel dann schwierig sein, wenn die Aufgabenformulierung nur teilweise oder gar nicht wörtlich mit der gesuchten Information übereinstimmt. Ob die Information leicht zu lokalisieren ist, hängt darüber hinaus von Merkmalen des jeweiligen Textes ab, ist also textseitig bedingt. Relevant sind etwa die Textlänge, Aspekte des Wortschatzes, inhaltliche Dichte, syntaktische Komplexität und der Grad der Transparenz der Struktur beziehungsweise Ordnung des Textes. Ein Beispiel: Ist in einer Aufgabe nach einem Ort gefragt und in einem kurzen Text kommt nur eine Ortsbezeichnung vor, in einem längeren dagegen eine Vielzahl von Ortsnamen, dann ist die Lokalisierungsaufgabe im zweiten Fall voraussichtlich schwieriger. So mag es dazu kommen, dass eine Aufgabe dieses Typs auch schwieriger ist als eine, deren Lösung vorwissensbasierte Verstehensleistungen voraussetzt. „Thus, the nature of the text can impact the difficulty of the question asked, across and within the four types of comprehension processes“ (Mullis et al., 2015).

3 Der IGLU-Test zur Erfassung der Lesekompetenz

Im Folgenden werden zunächst wesentliche Aspekte der Testung im internationalen Rahmen beschrieben. Anschließend wird dargelegt, wie im Rahmen einer nationalen Zusatzerhebung Facetten der Leseflüssigkeit erhoben wurden. Dann wird darauf eingegangen, inwiefern die Testaufgaben als *curricular valide* angesehen werden können. In einem nächsten Schritt werden exemplarisch zwei Lesetexte, ein fiktional-narrativer und ein Sachtext, und die Aufgaben dazu vorgestellt. Schließlich wird das im Rahmen von IGLU 2016 verwendete Modell zur Beschreibung der Stufen von Lesekompetenz präsentiert.

3.1 International eingesetzte Testvarianten und der Aufbau von IGLU

Für 2016 wurde neben den seit 2001 etablierten IGLU-/PIRLS-Lesetests unter dem Akronym PIRLS *Literacy* eine weitere Version entwickelt. Mit PIRLS *Literacy* soll in den unteren Leistungsbereichen besser differenziert werden. Dieser Test bietet sich für die teilnehmenden Staaten und Regionen an, in denen noch viele Viertklässlerinnen und Viertklässler insbesondere dann Probleme haben, wenn vorwissensbasierte Verstehensleistungen gefragt sind. PIRLS *Literacy* liegt dieselbe Testkonzeption wie IGLU zugrunde, die Lesetexte sind aber kürzer und die Aufgaben leichter, insofern stärker auf Lokalisieren und einfache Schlussfolgerungen gesetzt wird. Staaten und Regionen, die diese Version wählten, konnten ausschließlich sie nutzen oder auch eine Kombination von PIRLS und PIRLS *Literacy*, etwa indem sie PIRLS *Literacy* in vierten und PIRLS in fünften Klassen einsetzten.

Deutschland hat sich mit Schülerinnen und Schülern der vierten Jahrgangsstufe an IGLU beteiligt. Bei den Lesetests handelt sich um einen Papier- und Bleistift-Test. Der Zielvorgabe entsprechend lagen den Kindern im Rahmen eines rotierten Testdesigns (siehe Kapitel 2, Abschnitt 5.1.2 in diesem Band) insgesamt sechs Erzähl- und sechs Sachtexte vor, die in 16 Testheften präsentiert wurden. Jedes Kind bearbeitete ein Heft, das je einen Erzähl- und einen Sachtext enthielt.

Die Länge der Erzähltexte beträgt im Mittel etwa 800 Wörter. Neben zwei Ich-Erzählungen kommen vier mit einem personalen Erzähler beziehungsweise Erzählerin vor. Die Handlung wird jeweils von wenigen menschlichen Hauptfiguren getragen, in einem Fall, einem fabelähnlichen Text, in dem allerdings auch menschliche Akteure eine Rolle spielen, sind zwei Wölfe die (sprechenden) Protagonisten. Die Charaktere sind nicht komplex, einige ihrer Merkmale sind öfter explizit genannt, andere müssen aus ihren Handlungen sukzessive erschlossen werden. In zwei Texten erreichen die Hauptfiguren aufgrund von Ehrlichkeit beziehungsweise eines kreativen Experiments ihre Handlungsziele, in einem Text ändert der Ich-Erzähler aufgrund einer einschneidenden Erfahrung seine Einstellung. In zwei Texten verhelfen Figuren anderen dazu, dass ein früherer, glücklicherer Zustand wiederhergestellt wird. In dem fabelartigen Text schließlich, dem einzigen mit unglücklichem Ausgang, endet der Helfende in Gefangenschaft. Durchgängig wird linear erzählt, häufig kommt direkte, das Geschehen vergegenwärtigende Rede vor. Nur in einem Text finden sich kurze innere Monologe. Der Wortschatz dürfte, soweit das ohne Kenntnis des passiven Wortschatzes der einzelnen Kinder vor Ort, zu beurteilen ist, vertraut sein, die syntaktischen Strukturen sind in der Regel einfach. Es dominieren einfache Hauptsätze und Satzreihen, Satzgefüge bestehen in der großen Mehrheit

Tabelle 3.1: Sach- und Erzähltexte in IGLU 2016

Erzähltexte	Sachtexte
IGLU-Lesetexte	
<p>Goldfell² – Diese Tiergeschichte erzählt von Heldentum und den Konsequenzen rücksichtslosen Verhaltens.</p> <p>Marie und das rote Huhn – Diese Geschichte porträtiert ein Mädchen, das sich der Herausforderung stellt, sich um ein rotes Huhn zu kümmern.</p> <p>Der leere Topf³ – Diese in China spielende traditionelle Erzählung beinhaltet eine moralische Botschaft über die Bedeutung von Ehrlichkeit.</p> <p>Oliver und der Greif – In dieser Fantasiegeschichte trifft ein Junge namens Oliver einen alten Greif in einem Garten und entscheidet sich ihm zu helfen.</p>	<p>Leonardo Da Vinci¹ – Dieser biografische Text beschreibt die Erfindungen Leonardo Da Vincis und warum er seiner Zeit voraus war.</p> <p>Die grüne Meeresschildkröte und die Reise ihres Lebens – Dieser Text schildert den Lebenslauf einer weiblichen grünen Meeresschildkröte vom Zeitpunkt des Schlüpfens aus einem Ei bis zu der Zeit, in der sie selbst Eier legt.</p> <p>Wo ist der Honig?³ – Dieser Text beschreibt die Beziehung zwischen dem Honiganzeiger (einem Spechtvogel) und den afrikanischen Einwohnern der Borana mittels Erklärungen, Fotografien und grafischen Darstellungen.</p> <p>Inlandpferde – Dieser Artikel beschreibt die Geschichte und Eigenschaften der Inlandpferde in ihrer Entwicklung und der Menschen, die in ihrer Umgebung leben.</p>
Gemeinsam in IGLU und PIRLS <i>Literacy</i> eingesetzte Lesetexte	
<p>Blumen auf dem Dach¹ – Diese zeitgemäße Geschichte erzählt von der Freundschaft zwischen Generationen.</p> <p>Pemba Sherpa – Diese im Himalaya-Gebirge spielende moderne Erzählung gibt die Geschichte eines jungen Mädchens wieder, das sich dazu entschließt ein Sherpa zu sein.</p>	<p>Haie² – Dieser Artikel informiert über Haie mittels vielfältiger Textformate, der Verwendung von Zwischenüberschriften, einem beschrifteten Diagramms sowie Fotografien.</p> <p>Wie haben wir fliegen gelernt? – Dieser geschichtliche Text erläutert die Entwicklung des modernen Flugzeuges.</p>

¹ Lesetext aus IGLU 2001.² Lesetext aus IGLU 2006.³ Lesetext aus IGLU 2011.

der Fälle aus einem Haupt- und einem Nebensatz, komplexere Konstruktionen sind selten. Das Vorwissen, das für das Lösen der Aufgaben erforderlich ist, sollte in lebensweltlich vertrauten Kontexten erworben worden sein. Immer finden sich zusätzlich bildliche Darstellungen. Sie haben ausschließlich ästhetische und illustrierende Funktionen, für das Verständnis des jeweiligen Textes müssen sie nicht herangezogen werden.

Zwei der sechs Sachtexte sind thematisch auf Meerestiere bezogen, in dem einen geht es um den Lebenszyklus von Schildkröten, in dem anderen nach dem Muster eines Eintrags in einem Kinderlexikon um Merkmale von Haien im Allgemeinen und von drei Haiarten im Besonderen. Auch zwei weitere Texte handeln von Tieren. In einem dieser Texte stehen, unter anderem in historischer Perspektive, Inlandpferde im Zentrum, in dem anderen wird dargelegt, wie in Ostafrika Mitglieder eines Stammes und eine Vogelart auf der Suche nach Nahrung „kooperieren“. Hinzu kommen ein Text zu Aspekten der Biographie

eines Universalgelehrten zur Zeit der Renaissance und einer zur Historie der Entwicklung von Fluggeräten. Die Sachtexte sind im Mittel etwas länger als die literarischen Texte. Zuweilen kommen nicht erläuterte Fach- und Fremdwörter (*Magnetfeld, Technologie, Frequenz, Elektrorezeptoren, potentiell*) vor. Sie mögen die Lektüre erschweren, in keinem Fall ist die Kenntnis ihrer Bedeutung aber für die Lösung der Aufgaben erforderlich. Es dominieren einfache Sätze und Satzgefüge. Mit der Dominanz der einfachen Sätze geht in wenigen Fällen allerdings eine Verdichtung der Information einher, die das Verständnis erschweren kann, zum Beispiel für Kinder mit Migrationshintergrund. Ein Beispiel für eine solche Verdichtung ist die Ersetzung eines Nebensatzes (*Ihre Fähigkeit zu riechen scheint zuzunehmen, wenn ihr Hunger wächst.*) durch eine Präpositionalphrase (*Ihre Fähigkeit zu riechen scheint mit wachsendem Hunger zuzunehmen.*). Die Gliederung der Sachtexte wird durch Zwischenüberschriften erleichtert, zusätzlich durch Fotos, geographische Karten und schematische Darstellungen, die zum Teil mit separaten Kurztexten kombiniert sind. Teilweise dienen die nicht verbalen Elemente nur der Illustration, teilweise erleichtern sie aber auch das Verständnis des im Fließtext Dargestellten. Die in IGLU/PIRLS verwendeten Sachtexte wurden ausschließlich von Autorinnen und Autoren verfasst, die erhebliche Expertise haben, was das Schreiben für Kinder angeht. Thematisch decken die Texte ein Spektrum ab, das für die beteiligten Bildungssysteme als repräsentativ angesehen werden kann.

Zu den 12 Texten gab es insgesamt 175 Aufgaben, für deren Lösung jeweils 1 Punkt bis maximal 3 Punkte vergeben wurden. Es wurden 86 *Multiple-Choice*-Aufgaben und 89 offene Aufgaben gestellt. 90 Aufgaben bezogen sich auf literarische und 85 Aufgaben auf Sachtexte (Mullis, Martin, Foy & Hooper, 2017).

Die Verteilung der Aufgaben beziehungsweise Items im Hinblick auf Verstehensprozesse ist aus der Tabelle 3.2 ersichtlich und im Vergleich zu IGLU 2001, 2006 und 2011 dargestellt (vgl. Bos et al., 2003, S. 96; Bos, Valtin, Voss, Hornberg & Lankes, 2007, S. 86; Bremerich-Vos, Tarelli & Valtin, 2012, S. 75). Es ist eine deutliche Erweiterung des Aufgabenpools bei vergleichbaren Gewichtungen der Verstehensprozesse festzustellen.

Tabelle 3.2: Prozentuale Anteile der Verstehensprozesse bei IGLU 2016 in Prozent

Verstehensprozesse	2001		2006		2011		2016	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Lokalisieren explizit angegebener Informationen	25	26	31	22	33	22	50	29
Einfache Schlussfolgerungen ziehen	27	28	43	28	46	28	53	30
Komplexe Schlussfolgerungen ziehen bzw. Interpretieren und Kombinieren	31	32	34	37	38	37	47	27
Prüfen und Bewerten des Inhalts und des Sprachgebrauchs	15	15	18	14	18	13	25	14
Gesamt	98	100	126	100	135	100	175	100

Differenzen zu 100 Prozent sind im Rundungsverfahren begründet.

Es kann angenommen werden, dass eine Reihe von Kindern, die viele Aufgaben zum Textverstehen nicht richtig gelöst haben, noch nicht flüssig genug lesen können. Üblicherweise werden vier Facetten von Leseflüssigkeit unterschieden: Dekodiergenauigkeit, Automatisierung, Lesegeschwindigkeit und prosodisch phrasiertes Lesen (NICHD, 2000, Kapitel 3; Daane et al., 2005; Nix, 2011, S. 55 ff.). Wer zum Beispiel häufig vorkommende Wörter noch lautierend oder syllabierend liest und diese Wörter nur mit Mühe und langsam erfasst, hat aufgrund der begrenzten Kapazität des Arbeitsgedächtnisses kaum Ressourcen für die Verarbeitung größerer Einheiten auf Satzebene und darüber hinaus zur Verfügung (siehe Kapitel 10 in diesem Band). Insofern ist die Fähigkeit, flüssig zu lesen, als hierarchieniedrige Teilfähigkeit des Lesens anzusehen. Für die Testung aller Facetten dieser Fähigkeit wäre es nötig, die Kinder laut vorlesen zu lassen. Das ist im Rahmen eines *Large-Scale-Assessments* wie IGLU nicht leistbar. Um aber auf ökonomische Weise nähere Auskünfte über die Ausprägungen einiger dieser Dimensionen zu bekommen, wurden in Deutschland im Jahr 2016 zusätzlich auf die Wort- und die Satzebene bezogene Teile des Lesetests ELFE II (Lenhard, Lenhard & Schneider, 2017) eingesetzt (siehe auch Kapitel 2, Abschnitt 5.1.4 in diesem Band). Für die Testung des Wort- und des Satzverständnisses ist hier jeweils ein Zeitlimit vorgesehen. Die Kinder haben zunächst anzugeben, welches von vier Wörtern zu einem Bild passt. So ist zum Beispiel eine Maus abgebildet und zur Auswahl stehen die Wörter *Maus*, *Mann*, *Maul* und *Haus*. Hier handelt es sich jeweils um phonologisch ähnliche Einsilber. Es gibt insgesamt 75 Aufgaben zur Bild-Wort-Zuordnung und es kommen ein- bis viersilbige Wörter vor. Auf der Satzebene geht es darum, in insgesamt 36 Fällen zu entscheiden, welches von fünf Wörtern in einen gegebenen Satzrahmen passt. Dabei handelt es sich um Substantive, Verben, Adjektive, Präpositionen oder Konjunktionen. Ein Beispiel: *Florian kauft sich ein Computerspiel, (nachdem, obwohl, dafür, damit, außer) es teuer ist*. Auf diese Weise werden also Dekodiergenauigkeit und -geschwindigkeit ermittelt.

Auf dieser Basis lassen sich insbesondere die Leistungen der sehr leseschwachen Kinder genauer charakterisieren, das heißt derjenigen Viertklässlerinnen und Viertklässler, die über Kompetenzstufe I nicht hinauskommen.

3.2 Zur curricularen Validität der IGLU-Lesetests

Was Kinder am Ende der Grundschulzeit können sollen, ist in den „Bildungsstandards im Fach Deutsch für den Primarbereich“ fixiert, die von der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) 2004 verabschiedet wurden. Es sind vier zentrale Kompetenzbereiche ausgewiesen (KMK, 2005, S. 7), die in Tabelle 3.3 abgedruckt sind.

Unter der Überschrift „über Lesefähigkeiten verfügen“ wird zunächst der „Globalstandard“ „altersgemäße Texte sinnverstehend lesen“ genannt und dann ein auf fiktionale Texte bezogener Aspekt: „lebendige Vorstellungen beim Lesen und Hören literarischer Texte entwickeln“. Unter der Rubrik „über Leseerfahrungen verfügen“ ist unter anderem genannt, dass zwischen Sach- und Gebrauchstexten und literarischen Texten (erzählenden, lyrischen und szenischen) unterschieden werden kann. Unter „Texte präsentieren“ wird verstanden, dass man sinngestaltend vorlesen und auch auswendig vortragen kann. Am stärksten ausdifferenziert ist der Bereich „Texte erschließen“. Hier finden sich die folgenden Standards (KMK, 2005, S. 12):

Tabelle 3.3: Kompetenzbereiche des Faches Deutsch der KMK-Bildungsstandards

Sprechen und Zuhören	Schreiben	Lesen – mit Texten und Medien umgehen
– Gespräche führen	– über Schreibfertigkeiten verfügen	– über Lesefähigkeiten verfügen
– zu anderen sprechen	– richtig schreiben	– über Leseerfahrungen verfügen
– verstehend zuhören	– Texte planen	– Texte erschließen
– szenisch spielen	– Texte schreiben	– Texte präsentieren
– über Lernen sprechen	– Texte überarbeiten	
Methoden und Arbeitstechniken		
Methoden und Arbeitstechniken werden jeweils im Zusammenhang mit den Inhalten jedes einzelnen Kompetenzbereichs erworben.		
Sprache und Sprachgebrauch untersuchen		
– grundlegende sprachliche Strukturen und Begriffe kennen		
– sprachliche Verständigung untersuchen		
– an Wörtern, Sätzen, Texten arbeiten		
– Gemeinsamkeiten und Unterschiede von Sprachen entdecken		

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

- Verfahren zur Orientierung nutzen;
- gezielt einzelne Informationen suchen;
- Texte genau lesen;
- Verstehenshilfen anwenden;
- Texte mit eigenen Worten wiedergeben;
- Textaussagen erfassen und wiedergeben;
- Aussagen mit Textstellen belegen;
- eigene Gedanken entwickeln, Stellung nehmen, mit anderen darüber sprechen;
- Sensibilität für Gedanken und Gefühle anderer zeigen (bei literarischen Texten);
- Unterschiede und Gemeinsamkeiten von Texten finden;
- handelnd mit Texten umgehen.

Die Länder der Bundesrepublik Deutschland verpflichteten sich, die Standards ab dem Schuljahr 2005/2006 zu implementieren und neue Lehrpläne an ihnen auszurichten. In einigen Ländern geschah das schon früher, nämlich für das Schuljahr 2004/2005 (auf Einzelnachweise wird im Folgenden verzichtet, vgl. Deutscher Bildungsserver, 2011).

Die Datierung der zum Zeitpunkt der Testung, das heißt im Mai 2016, in den einzelnen Ländern gültigen Lehrpläne für das Fach Deutsch in der Grundschule ist aus Tabelle 3.4 ersichtlich.

In den Curricula der Länder wird die von der KMK gewählte Bezeichnung des Kompetenzbereichs („Lesen – mit Texten und Medien umgehen“) zuweilen modifiziert, so in Bayern (2014) und Hamburg (2011). Hier ist von „Lesen – mit Texten und weiteren (bzw. anderen) Medien umgehen“ die Rede, wo-

Tabelle 3.4: Lehrpläne für das Fach Deutsch in der Grundschule nach Ländern der Bundesrepublik Deutschland und Veröffentlichungszeiträumen vor und nach IGLU 2001, 2006, 2011 und 2016

Land	Jahr	Im Mai 2016 für die 4. Jahrgangsstufe gültige Lehrpläne				neue Lehrpläne gültig ab dem Schuljahr
		vor Oktober 2004	zwischen 2004 und 2006	zwischen 2006 und 2011	zwischen 2011 und 2016	
Bayern	2014	-	-	-	✓	-
Hessen	2011	-	-	-	✓	-
Hamburg	2011	-	-	-	✓	-
Thüringen	2010	-	-	✓	-	-
Sachsen	2009	-	-	✓	-	-
Saarland	2009	-	-	✓	-	-
Nordrhein-Westfalen	2008	-	-	✓	-	-
Sachsen-Anhalt	2007	-	-	✓	-	-
Niedersachsen	2006	-	✓	-	-	2016/2017
Rheinland-Pfalz	2005	-	✓	-	-	-
Mecklenburg-Vorpommern	2004	✓	-	-	-	-
Bremen	2004	✓	-	-	-	-
Brandenburg	2004	✓	-	-	-	2017/2018
Berlin	2004	✓	-	-	-	2017/2018
Baden-Württemberg	2004	✓	-	-	-	2016/2017
Schleswig-Holstein	1997	✓	-	-	-	-

Rosa markierte Zeilen indizieren Länder, in denen die Lehrpläne vor der Verabschiedung der KMK-Bildungsstandards erlassen wurden.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

mit dem Sachverhalt Rechnung getragen wird, dass auch Texte (Print-)Medien sind. Der Globalstandard „altersgemäße Texte sinnverstehend lesen“ ist in wörtlicher Übernahme fast durchgängig zu finden. Häufig findet sich die in den Standards vorgegebene Gliederung in vier Bereiche, es werden aber auch neue Teilüberschriften vergeben. Vielfach ist das Bemühen erkennbar, die Standards in eine systematischere Form zu „übersetzen“. So heißt es im Lehrplan für Bayern: Flüssiges, „angemessen schnelles und genaues Lesen bilden die Grundlage für die erfolgreiche Anwendung von Lesestrategien“. Hier wird auf die Differenz von hierarchieniedrigen und -hohen Teilfähigkeiten abgehoben. Auf Lesestrategien, einen Begriff, der in den Standards nicht vorkommt, wird nicht nur hier, sondern in einer Reihe von Lehrplänen Bezug genommen. Dabei kommen unter anderem Strategien zur Orientierung in einem Text zur Sprache (detailliertes, selektives, überfliegendes Lesen, so in Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz, Thüringen), aber auch prozessbezogene Strategien, wobei die Stadien vor dem Lesen, während des Lesens und nach dem Lesen unterschieden werden (Baden-Württemberg, Bayern, Berlin/Brandenburg, Bremen/Mecklenburg-Vorpommern, Hamburg).

Festzuhalten ist, dass Kinder in vierten Klassen nach Auskunft aller Lehrpläne Kenntnisse von „altersangemessenen“ Sach- und literarischen, insbesondere narrativen Texten haben (sollten), dass sie diese Texte flüssig lesen können (sollten) und dass sie über eine Reihe von Lesestrategien verfügen (sollten), die es ihnen unter anderem gestatten, nicht nur wesentliche Einzelinformationen zu lokalisieren, sondern auch Textteile und das Textganze zusammenzufassen. Sie sollten auch dazu in der Lage sein, Texte elaborativ zu verarbeiten, sie also „anzureichern“, zum Beispiel indem sie text- und vorwissensbasiert Schlüsse ziehen. Insofern kann davon ausgegangen werden, dass die IGLU-Testung curricular valide ist. Es handelt sich allerdings nur um die intendierten Curricula, nicht um die im Unterricht vor Ort realisierten beziehungsweise implementierten Curricula (siehe Kapitel 10 in diesem Band).

4 Das Kompetenzstufenmodell von IGLU 2016

Um eine inhaltliche Interpretation der von den Viertklässlerinnen und Viertklässlern erreichten Testwerte zu ermöglichen, werden in IGLU wie in anderen großen Schulleistungsstudien Kompetenzstufen gebildet. Das Testverfahren gestattet es, Schülerfähigkeiten und Aufgabenschwierigkeiten auf einer gemeinsamen Skala abzubilden. Auf internationaler Ebene wurde vereinbart, diese kontinuierliche Skala in gleich große Abschnitte zu unterteilen. Deren Grenzen liegen bei 400, 475, 550 und 625 Skaleneinheiten. Die Intervalle, die zwischen diesen Referenzwerten (*benchmarks*) liegen, werden hier als Kompetenzstufen bezeichnet (siehe Kapitel 2, Abbildung 2.7 in diesem Band). Um bestimmen zu können, welche Leseleistungen von Kindern auf den verschiedenen Stufen erwartet werden können, wurde wie folgt verfahren: Zunächst wurden die Kinder ausgewählt, die zwischen 390 und 410, 465 und 485, 540 und 560 und 615 bis 635 Punkte erreichten, also die Referenzwerte plus beziehungsweise minus 10 Punkte. In einem zweiten Schritt wurde in Anlehnung an das internationale Vorgehen für diese Gruppen für alle Aufgaben jeweils die mittlere Lösungshäufigkeit berechnet. Im dritten Schritt wurden die Aufgaben identifiziert, die auf den vier Referenzpunkten verankert sind. Dabei waren für *Multiple-Choice*-Aufgaben zwei Kriterien maßgeblich: Eine Aufgabe wurde dann auf einer Benchmark verankert, wenn sie von den Schülerinnen und Schülern mit Leistungen im Bereich des jeweiligen Referenzwerts mit einer Wahrscheinlichkeit von mindestens 65 Prozent und von den Kindern auf der Stufe darunter mit einer Wahrscheinlichkeit von weniger als 50 Prozent gelöst wurden. Im Bereich der niedrigen Benchmark (390 bis 410) greift nur das erste Kriterium.

Bei Aufgaben mit offenen Formaten, bei denen die Kinder nicht raten können, war das Kriterium weniger streng. Eine offene Aufgabe wurde dann auf einer bestimmten Benchmark verortet, wenn sie von mindestens 50 Prozent der Viertklässlerinnen und Viertklässler mit Leistungen in diesem Bereich korrekt gelöst wurde. Wurde eine offene Aufgabe von weniger als 50 Prozent der Kinder auf fortgeschrittenem (*advanced*) Niveau gelöst, wurde sie als zu schwer eingestuft.

Die ausgewählten Aufgaben werden im Hinblick auf die damit verbundenen Anforderungen analysiert und interpretiert. Leitend sind hier in erster Linie die im IGLU-Modell angenommenen vier Verstehensprozesse. Berücksichtigt werden auch die Aufgabenformate im Allgemeinen (geschlossen versus offen) und einige ihrer Merkmale im Besonderen. Relevant ist zum Beispiel, ob und,

falls ja, in welchem Umfang die Aufgabenformulierung mit der im Text zu lokalisierenden Information identisch ist. Bei *Multiple-Choice*-Aufgaben fällt ins Gewicht, wie plausibel die nicht zutreffenden Optionen, die Distraktoren, sind. Die Aufgabenschwierigkeit wird darüber hinaus von Fall zu Fall durch Distraktoren im Stimulustext selbst beeinflusst (Lumley, Routitsky, Mendelovits & Ramalingam, 2012). Sollen für Aufgaben auf den jeweiligen Stufen charakteristische Anforderungen bestimmt werden, muss von Spezifika der einzelnen Aufgaben abgesehen und deren gemeinsamer Nenner bestimmt werden. Mit diesem Abstraktionsschritt geht unvermeidlich ein Informationsverlust einher. Die für die Klasse der Aufgaben auf einer Stufe typischen Anforderungen sind darüber hinaus von den Anforderungen zu unterscheiden, die für die Aufgaben auf den anderen Kompetenzstufen charakteristisch sind. Einerseits wird also angestrebt, die Differenzen möglichst klar herauszustellen. Andererseits ist zu bedenken, dass die zugrunde liegende Skala kontinuierlich ist und dass sich die Anforderungen auf den verschiedenen Stufen nicht durchgängig qualitativ (kategorial) unterscheiden, sondern nur graduell (Olsen & Nilsen, 2017).

4.1 Beschreibung der Kompetenzstufen in IGLU 2016

Die Deskriptionen der einzelnen Kompetenzstufen beziehen sich auf die von den Kindern in Deutschland erbrachten Leseleistungen. Für Charakterisierungen der Benchmarks auf internationaler Ebene sei auf Mullis et al. (2017) verwiesen.

Tabelle 3.5: Kompetenzstufen und Skalenwerte IGLU 2016

Kompetenzstufe		Skalenbereich der Fähigkeit
V	Unter Bezug auf Textpassagen bzw. den Gesamttext Informationen ordnen und Aussagen selbstständig interpretierend und kombinierend begründen	> 625
IV	Für die Herstellung von Kohärenz auf der Ebene des Textes relevante Aspekte des Inhalts und der Darstellung erfassen und komplexe Schlüsse ziehen	551–625
III	„Verstreute“ Informationen verständlich miteinander verknüpfen	476–550
II	Explizit angegebene Informationen identifizieren und auf lokaler Ebene Kohärenz herstellen	400–475
I	Rudimentäres Leseverständnis	< 400

Kompetenzstufe V: Unter Bezug auf Textpassagen beziehungsweise den Gesamttext Informationen ordnen und Aussagen selbstständig interpretierend und kombinierend begründen

Schülerinnen und Schüler auf dieser Stufe lösen über die Aufgaben auf den unteren Kompetenzstufen hinaus mit größerer Wahrscheinlichkeit weitere Aufgaben, bei denen vornehmlich Interpretationsleistungen zu erbringen sind. Am schwierigsten sind offene Aufgaben, bei denen nicht einzelne Wörter, sondern mehrere Sätze zu schreiben sind. Sind – ausschließlich im Umgang mit informierenden Texten – noch einfache Schlussfolgerungen gefragt, so handelt es

sich in erster Linie um *Multiple-Choice*-Aufgaben mit mindestens einem sehr plausiblen Distraktor. Bei den anderen Aufgaben zu Sachtexten, die auf dieser Stufe deutlich dominieren, geht es unter anderem darum, in Tabellen, in denen einige Zellen bereits beschriftet sind, weitere Zellen auszufüllen. Dazu müssen über den jeweiligen Text verstreute Informationen geordnet werden. Größere Textsegmente sind auch heranzuziehen, wenn anzugeben ist, wie Geographie, Vegetation und Klima Aussehen und Verbreitung einer Tierart beeinflusst haben. In weiteren Aufgaben werden Szenarien geboten, die in den Texten selbst nicht vorkommen. Um darauf bezogene Fragen zu beantworten, sind jeweils mehrere Informationen, etwa zu Sinnesleistungen und zu Verhaltensweisen von Tieren, miteinander zu verknüpfen. Aufgaben zu narrativen Texten verlangen zum Beispiel die begründete Bewertung eines alternativen Texttitels, die Beschreibung des Charakters von Protagonisten und deren Rechtfertigung durch Nennung einiger ihrer Handlungen. Darüber hinaus sind Wandlungen von Gefühlen und Meinungen von Figuren im Textverlauf darzustellen und zu erklären.

Kompetenzstufe IV: Für die Herstellung von Kohärenz auf der Ebene des Textes relevante Aspekte des Inhalts und der Darstellung erfassen und komplexe Schlüsse ziehen

Kinder auf dieser Stufe lösen nicht nur Aufgaben auf den Stufen I bis III mit hoher Wahrscheinlichkeit. Sie bearbeiten mit höherer Wahrscheinlichkeit auch Aufgaben erfolgreich, bei denen zum einen textbasierte Verstehensleistungen, überwiegend einfache Schlussfolgerungen, gefragt sind, zum anderen vorwissensbasiertes Verstehen. Aufgaben mit offenem Format werden seltener gelöst als solche im *Multiple-Choice*-Format. Bei einigen der Lokalisierungsaufgaben sind zweiteilige Antworten gefragt. Sind einfache Schlüsse zu ziehen, müssen in der Regel nicht Beziehungen zwischen benachbarten Sätzen, sondern größere Segmente bedacht werden. So ist zum Beispiel anzukreuzen, welche Zusammenfassung eines Abschnitts beziehungsweise welche Hauptaussage zutreffend ist. Es sind unter anderem Handlungen mit Motiven, Zielen und Folgen und Emotionen mit Ausdrucksverhalten zu verknüpfen. Aufgaben, bei denen es um Interpretieren geht, beziehen sich fast ausschließlich auf literarische Texte. Hier ist die Angabe von Motiven für Handlungen gefragt, zuweilen sind im Textverlauf sich ändernde Beziehungen zwischen Figurenmerkmalen, Handlungen und Emotionen zu erklären. Eine Reihe von Aufgaben zum Bewerten bezieht sich auf Aspekte der Darstellungsweise. So soll angegeben werden, inwiefern bildliche Darstellungen das Textverstehen, unter anderem den Nachvollzug der Textgliederung oder das Verständnis einer Metapher, erleichtern.

Kompetenzstufe III: „Verstreute“ Informationen verständlich miteinander verknüpfen

Viertklässlerinnen und Viertklässler, die Stufe III erreichen, lösen über die für Stufe II charakteristischen Aufgaben hinaus mit größerer Wahrscheinlichkeit offene und geschlossene Aufgaben, die in etwa zwei Drittel der Fälle textbasierte Verstehensleistungen verlangen. Offene Aufgaben, bei denen nur wenig zu schreiben ist, sind im Mittel schwieriger als Aufgaben im *Multiple-Choice*-Format. Einige Lokalisierungsaufgaben sind so formuliert, dass der Bezug zu den gesuchten Informationen weniger deutlich ist als auf Stufe I. Sie beziehen sich auf explizit benannte Handlungen, Ereignisse und Gefühle, öfter sind mehrere benachbarte Informationen anzugeben. Sind einfache Schlüsse zu ziehen, so ist zwar auch noch das Verständnis lokaler Kohärenz gefragt. Bei einer grö-

ßeren Zahl von Aufgaben ist es aber nötig, mehrere über den Text „verstreute“ Informationen zu verknüpfen, die sich unter anderem auf Attribute und Motive zentraler Figuren beziehen. Das gilt auch für Aufgaben zum Interpretieren und Bewerten. Hier ist unter anderem das Ordnen von Ereignissen gefragt. In wenigen Fällen ist das Verständnis des Textganzen vorausgesetzt, etwa dann, wenn im *Multiple-Choice*-Format nach der geeignetsten alternativen Textüberschrift gefragt wird beziehungsweise das Textthema anzugeben ist.

Kompetenzstufe II: Explizit angegebene Informationen identifizieren und auf lokaler Ebene Kohärenz herstellen

Kinder auf dieser Stufe lösen vor allem *Multiple-Choice*- und offene Aufgaben, bei denen das Lokalisieren explizit angegebener Informationen gefragt ist. Sie finden sich in der Regel in einem einzigen (auch komplexen) Satz beziehungsweise in benachbarten Sätzen. Zuweilen sind Teile der Aufgabe und der gesuchten Information identisch formuliert. In einigen Fällen sind einfache Schlüsse zu ziehen. Auf lokaler Ebene, das heißt unter Berücksichtigung benachbarter Sätze, sind die Kinder in der Lage, nicht explizit angegebene Beziehungen zu erkennen, zum Beispiel zwischen Gründen und Handlungen und Ursachen und Vorgängen. Auf dieser Stufe werden nur wenige Aufgaben gemeistert, für die Interpretieren nötig ist.

Kompetenzstufe I: Rudimentäres Leseverständnis

Die Leistungen der Schülerinnen und Schüler, die unterhalb der Schwelle von 390 bis 410 Punkten liegen, können im Rahmen des IGLU-Modells nicht näher charakterisiert werden. Deshalb wurden in Deutschland zusätzlich die auf das Wort- und Satzverstehen bezogenen Teile des Leseverständnistests ELFE II (Lenhard et al., 2017) eingesetzt. Die Kinder auf Kompetenzstufe I erreichen hier im Mittel auf Wortebene 41.8 von 75 Punkten, auf Satzebene 13.8 von 36 Punkten. Im Lauf von jeweils drei Minuten lösen sie also etwa die Hälfte der Aufgaben, bei denen eines von vier Wörtern zu unterstreichen ist, das zu einem Bild passt, und etwas mehr als ein Drittel der Aufgaben, bei denen eines von fünf Wörtern markiert werden muss, das in einen vorgegebenen Satzrahmen passt.

Tabelle 3.6: Mittlere Lösungshäufigkeiten in den Teilen des Leseverständnistests ELFE II nach IGLU-Kompetenzstufen

	Kompetenzstufen											
	insg.		I		II		III		IV		V	
	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)	M	(SE)
Wortverständnis (max. 75 Punkte)	55.5	(0.4)	41.8	(3.3)	50.5	(1.7)	55.5	(0.8)	57.9	(0.9)	60.6	(1.4)
Satzverständnis (max. 36 Punkte)	24.2	(0.2)	13.8	(2.0)	19.6	(0.7)	23.7	(0.4)	26.4	(0.4)	29.1	(0.6)

4.2 Texte und Aufgaben

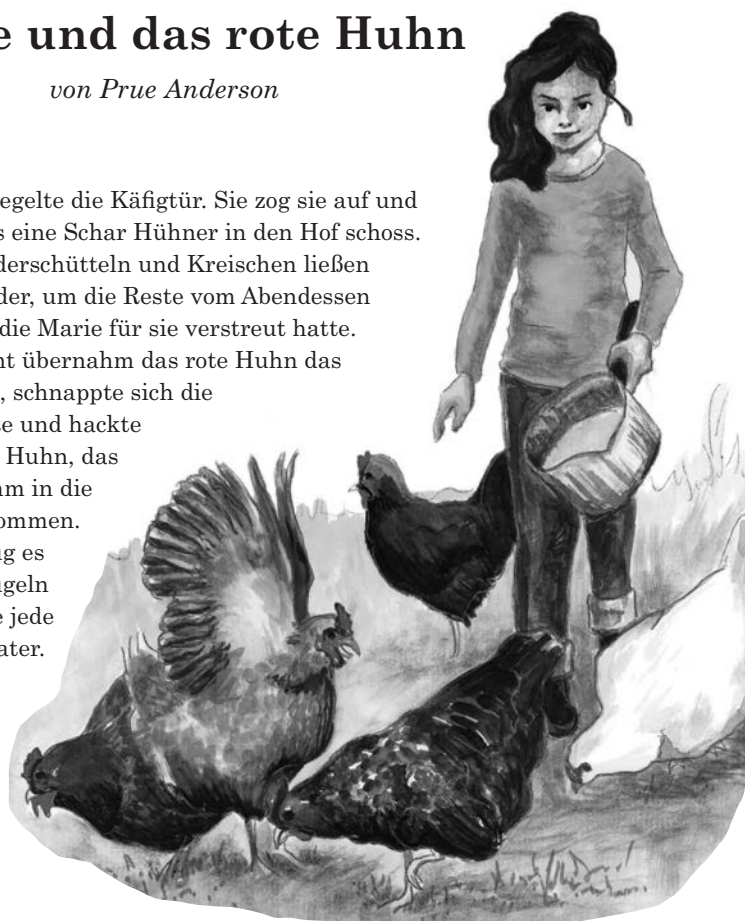
Auf den folgenden Seiten sind zwei Texte abgedruckt, die den Kindern im Rahmen der Testung im Jahr 2016 vorlagen. Es handelt sich um einen der sechs narrativen Texte und einen der sechs Sachtexte. Darüber hinaus sind alle von den Schülerinnen und Schülern zu bearbeitenden Aufgaben aufgeführt, wobei auch angegeben ist, welche Verstehensprozesse jeweils angenommen werden. Einige Aufgaben wurden ausgewählt, um die Anforderungen beziehungsweise Leistungen auf verschiedenen Kompetenzstufen zu illustrieren. Eine Aufgabe wird dann auf einer bestimmten Kompetenzstufe verortet, wenn sie von mindestens 65 Prozent (im Fall einer *Multiple-Choice*-Aufgabe) beziehungsweise von mindestens 50 Prozent (bei einer offenen Aufgabe) der Kinder auf dieser Stufe gelöst wird und wenn für die Kinder auf einer Stufe darunter eine erfolgreiche Bearbeitung weniger wahrscheinlich ist.

4.2.1 Ein fiktionaler Erzähltext, die Aufgaben und Beispiele für Anforderungen auf verschiedenen Kompetenzstufen

Marie und das rote Huhn

von Prue Anderson

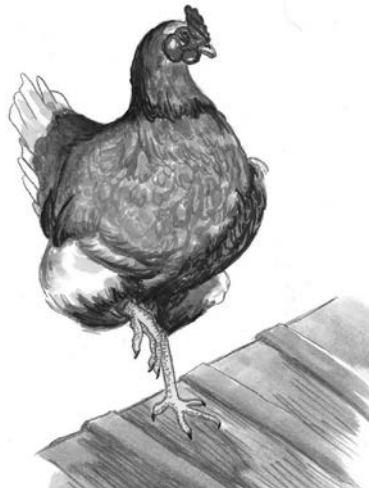
Marie entriegelte die Käfigtür. Sie zog sie auf und lächelte, als eine Schar Hühner in den Hof schoss. Mit viel Federschütteln und Kreischen ließen sie sich nieder, um die Reste vom Abendessen zu fressen, die Marie für sie verstreut hatte. Wie gewohnt übernahm das rote Huhn das Kommando, schnappte sich die besten Reste und hackte nach jedem Huhn, das es wagte, ihm in die Quere zu kommen. Dabei schlug es mit den Flügeln und machte jede Menge Theater.



„Warum lassen die anderen Hühner sich von dem roten Huhn so herumkommandieren?“, hatte Marie ihre Mutter gefragt.

„Hühner haben eine Hackordnung“, erklärte ihre Mutter. „Das mutigste und stärkste Huhn hat das Kommando. Es darf nach allen anderen Hühnern hacken, aber keins darf nach ihm hacken. Das nächste Huhn in der Hackordnung kann nach allen hacken außer nach dem obersten Huhn, und so geht das bis ganz runter, also tut einem das arme Huhn ganz unten wirklich leid. Hühner haben gerne eine Anführerin, die sie herumscheucht.“

Aber Marie war anderer Ansicht. Jeden Abend musste sie die Hühner wieder in ihrem Käfig einsperren, damit Füchse und Eulen sie sich nicht holten. Das war ihre Aufgabe. Jeder in ihrer großen Familie hatte Aufgaben. Wenn es dunkel wurde, gingen alle Hühner gerne wieder zurück in ihren Käfig. Das heißt, alle außer dem roten Huhn. Es tat so, als würde es in Richtung Käfigtür gehen, und rannte dann im letzten Moment zur Seite und wartete darauf, dass Marie ihm hinterherjagte.



Ein weiterer Trick war, sich in die Mitte des Hofes zu setzen. Sobald Marie nahe genug war, um sich zu bücken und es hochzuheben, schlug das Huhn heftig mit den Flügeln, sodass Marie es nicht greifen konnte, dann rannte es wieder weg. Schließlich, wenn das rote Huhn entschieden hatte, dass Marie ihm genug hinterhergejagt war, ging es ganz ruhig von allein in den Käfig. Seine kleinen roten Augen glänzten triumphierend, als Marie die Käfigtür hinter ihm zuschlug.

Marie hatte versucht, das rote Huhn zu locken, indem sie ihm abends sein Lieblingsfutter in den Käfig stellte, aber das Huhn ließ sich nicht bestechen. Marie hatte ausprobiert, Topfdeckel laut gegeneinanderzuschlagen, um das rote Huhn zu erschrecken. Aber damit machte sie den anderen Hühnern so viel Angst, dass sie zwei Tage lang keine Eier legten.



Marie ging ihren Vater suchen. „Ich muss diesem roten Huhn eine Lehre erteilen“, sagte sie. „Ich werde es die ganze Nacht draußen lassen, dann muss es sich gegen die Füchse und die Eulen wehren. Das wird es lehren, in seinen Käfig zu gehen, wenn ich es sage.“

„Marie“, sagte Papa und wandte sich ihr zu. „Ein Huhn kann sich nicht gegen eine Eule oder einen Fuchs wehren und wir brauchen unsere Hühner. Wir brauchen jedes Ei, das wir bekommen können.“

Er lächelte. „Außerdem würdest du damit dein Problem nicht lösen, denn dann würde einfach das nächste Huhn in der Hackordnung den Platz des roten Huhns einnehmen.“ Er wandte sich wieder seiner Arbeit zu.

Marie stampfte in die Küche. „Mama, mir gefällt meine Aufgabe nicht, ich brauche eine andere Aufgabe.“

„Also, das ist einfach“, sagte Mama. „Du kannst das Abendessen kochen.“ Marie betrachtete den großen Haufen Essen auf dem Tisch, das gewaschen und kleingeschnitten und gerührt werden musste, damit ihre ganze Familie satt wurde.

„Du kannst das Geschirr abwaschen.“ Marie betrachtete den Berg schmutziger Töpfe und Pfannen vom letzten Abend, die sich noch im Spülbecken stapelten.

„Du kannst auf das Baby aufpassen.“ Marie betrachtete ihre kleine Schwester, die sich zufrieden eine Banane ins Gesicht, in die Haare und auf die Kleidung schmierte.

Marie zog sich aus der Küche zurück. „Ich hätte gerne deine Aufgabe“, rief ihre Mutter ihr nach.



Als Marie an diesem Abend im Hof dem roten Huhn hinterherjagte, sah sie, wie sich auf der anderen Seite des Hofzauns eine Eule niederließ.



Die Eule schwebte mit ihren großen weißen Flügeln über das Gras, schnappte sich mit den Krallen eine Maus und flog wieder zurück in den Schatten. Das brachte Marie auf eine Idee. Am nächsten Tag besorgte sich Marie Draht und weißen Stoff und bastelte zwei große Flügel, die sie am Ende einer langen Stange befestigte. Sie erklärte ihrem Bruder Simon ihren Plan.

Als Marie an diesem Abend wie gewohnt hinter dem roten Huhn herjagte, kam Simon mit der Stange, die Marie gebastelt hatte, in den Hof. Er ließ die Flügel in Richtung des roten Huhns nach unten schweben. Das Huhn hörte auf zu rennen und plusterte seine Federn auf, kreischte wild und schlug mit den Flügeln, bereit, sich dem Angreifer zu stellen. Aber die weißen Flügel kamen immer näher. Das Kreischen des roten Huhns verstummte. Es kauerte sich flach auf den Boden, den Schnabel nach oben gestreckt, um zu hacken, falls es die Gelegenheit dazu bekommen sollte. Plötzlich griff Marie ein.



Marie schrie die weißen Flügel an. Sie schlug mit den Händen nach den Flügeln und stieß die Flügel weg. Ihr Bruder hob die Flügel nach oben. Noch einmal bewegten sie sich nach unten und Marie wehrte sie ab. Mit seinen kleinen Knopfaugen beobachtete das rote Huhn alles, während es neben Maries Füßen auf dem Boden kauerte. Schließlich gaben die furchterregenden Flügel auf und flogen weg.

Marie bückte sich und streckte die Hände aus. Schnell schmiegte sich das rote Huhn hinein und legte sanft den Schnabel auf Maries Arm. Marie fühlte das kleine Herz des Huhns in seiner gefiederten Brust schlagen, als sie es zurück zum Käfig trug. Sie streichelte das Huhn, bis sein Herz langsamer schlug, dann setzte sie es in den Käfig und lächelte Simon an.

„Jetzt bist du in der Hackordnung ganz oben“, lachte Simon.

Betrachtet man diesen Text im Kontext der gegenwärtigen Debatten über schwierighkeitsrelevante Merkmale von literarischen Erzähltexten (z. B. Rosebrock, 2016), dann ergibt sich Folgendes: Die Handlung ist einsträngig, das Ziel der Hauptfigur beziehungsweise das von ihr erstrebte „Gut“ konkret und damit ist auch die Komplikation (die Konfrontation des Huhns mit der „Eule“) leicht zu erfassen. Die Auflösung ist eindeutig; Marie löst ihr Problem im Umgang mit dem Huhn, es wird zukünftig gehorsam sein. Für das Verständnis von Komplikation und Auflösung ist kein spezielles (historisches, psychologisches usw.) Wissen vorausgesetzt. Die Zahl der Figuren ist klein, ihre inneren Zustände bleiben außer Betracht, nur Ausdrucksverhalten wird namhaft gemacht (Lachen, Lächeln, Stampfen). Der Erzähler ist nicht Teil der erzählten Welt, er kommentiert nicht und sagt nicht mehr, als die Hauptfigur weiß, erlebte Rede und innerer Monolog spielen keine Rolle, wohl aber direkte Rede, das heißt szenische Präsentation des Geschehens. Nur die Darstellung der zeitlichen Abfolge des Geschehens fällt etwas komplexer aus: Auf den episodischen Einstieg folgt eine Rückwendung, angezeigt durch das Verb im Plusquamperfekt (... *hatte Marie ihre Mutter gefragt*). In den folgenden Abschnitten werden sich wiederholende Geschehnisse zusammenfassend („iterativ“) dargestellt (bis *dass sie zwei Tage lang keine Eier legten*). Ab dann wird chronologisch erzählt. Zur sprachlichen Ebene: In lexikalischer Hinsicht dürfte der Text, soweit das „aus der Ferne“ zu beurteilen ist, für Viertklässlerinnen und Viertklässler kaum Probleme bereiten. Die meisten Substantive sind Konkreta, die Bedeutungen seltener Verben (*entriegeln, kauern*) sind kontextuell leicht erschließbar, was mit dem für das Verständnis zentralen Wort *Hackordnung* gemeint ist, wird explizit erläutert. „Veraltete“ Lexeme, Neologismen, innovative Metaphern und Symbole kommen nicht vor. Die syntaktischen Strukturen hingegen sind teilweise schwerer zu erschließen. Neben einfachen Hauptsätzen kommen zahlreiche aus mehreren Teilsätzen bestehende Sätze vor, die einander teils neben-, teils untergeordnet sind. Dabei sind auch Nebensätze zweiten Grades zu verzeichnen, das heißt solche, die von einem anderen Nebensatz abhängen wie in folgendem Beispiel: *Sobald Marie nahe genug war, um sich zu bücken und es hochzuheben, schlug das Huhn heftig mit den Flügeln, sodass Marie es nicht greifen konnte, dann rannte es wieder weg.*

Summarisch betrachtet, handelt es sich um einen inhaltlich und strukturell einfachen Erzähltext, dessen Verständnis allerdings aufgrund syntaktisch bedingter „Hürden“ erschwert sein könnte.

Wie aus Abbildung 3.2 ersichtlich wird, erfordern 4 der 16 Aufgaben das Lokalisieren von explizit angegebenen Informationen und bei weiteren 4 Aufgaben sind einfache Schlussfolgerungen zu ziehen, indem Beziehungen zwischen Textteilen und -abschnitten hergestellt werden. 5 Aufgaben sind dem Verstehensprozess *Komplexe Schlussfolgerungen ziehen bzw. Interpretieren und Kombinieren* zugeordnet und 3 dem *Prüfen und Bewerten des Inhalts und des Sprachgebrauchs*.

Die Zuordnungen von Aufgaben und Verstehensprozessen beruhen auf interpretativen Leistungen der Testkonstrukteure. Insofern ist es naheliegend, dass im Einzelfall auch eine alternative Deutung als plausibel erscheinen mag. So könnte man bei Aufgabe 8 (*Was meint der Vater mit seiner Äußerung?*) den Standpunkt vertreten, dass es hier weniger um Bewerten als um Interpretieren geht.

Abbildung 3.2: Aufgaben zum Erzähltext *Marie und das rote Huhn*

<p>1 Was tut Marie am Anfang der Geschichte?</p> <p><input type="radio"/> A ein Huhn fangen</p> <p><input checked="" type="radio"/> B die Hühner füttern</p> <p><input type="radio"/> C Eier suchen</p> <p><input type="radio"/> D Federn sammeln</p> <p>Kompetenzstufe I (.88/.91) L</p>	<p>2 Wie zeigt die Autorin dir, welchen Charakter das rote Huhn hat?</p> <p><input type="radio"/> A Sie beschreibt, wie das rote Huhn aussieht.</p> <p><input type="radio"/> B Sie beschreibt das Lieblingsfutter des roten Huhns.</p> <p><input type="radio"/> C Sie beschreibt, wo das rote Huhn wohnt.</p> <p><input checked="" type="radio"/> D Sie beschreibt, wie sich das rote Huhn benimmt.</p> <p>Kompetenzstufe II (.80/.85) B</p>	<p>3 Warum tut Maries Mutter das Huhn leid, das ganz unten in der Hackordnung ist?</p> <p><input type="radio"/> A <u>Alle anderen Hühner können nach ihm hacken.</u></p> <p>Kompetenzstufe III (.49/.63) S</p>	<p>4 Warum trickst das rote Huhn Marie aus?</p> <p><input type="radio"/> A <u>Das Huhn ist an der Spitze der Hackordnung.</u></p> <p>Kompetenzstufe V (.17/.17) I</p>
<p>5 Warum schlägt Marie die Käfigtür zu?</p> <p><input checked="" type="radio"/> A Marie ist wütend.</p> <p><input type="radio"/> B Die Tür geht schwer zu.</p> <p><input type="radio"/> C Ein Fuchs kommt.</p> <p><input type="radio"/> D Das rote Huhn entwischt.</p> <p>Kompetenzstufe III (.38/.49) S</p>	<p>6 Marie will, dass das rote Huhn in den Käfig geht. Welche zwei Dinge tut Marie, die nicht funktionieren?</p> <p><input type="radio"/> A <u>1. Das Huhn hochheben</u></p> <p><input type="radio"/> B <u>2. Das Huhn mit Futter locken</u></p> <p>Kompetenzstufe III (.76/.84) L</p>	<p>7 Warum ist Papa gegen Maries Idee?</p> <p><input type="radio"/> A Das rote Huhn könnte sich wehtun.</p> <p><input type="radio"/> B Das rote Huhn würde keine Eier mehr legen.</p> <p><input checked="" type="radio"/> C Papa meint, dass das rote Huhn getötet werden würde.</p> <p><input type="radio"/> D Papa will dem Huhn eine Lehre erteilen.</p> <p>Kompetenzstufe IV (.68/.63) L</p>	<p>8 Papa sagt, dann würde einfach das nächste Huhn in der Hackordnung den Platz des roten Huhns einnehmen. Was meint er damit?</p> <p><input type="radio"/> A <u>Er meint, dass das rote Huhn sterben würde und dass das nächste Huhn in der Hackordnung dann so sein würde wie das rote Huhn.</u></p> <p>Kompetenzstufe V (.22/.24) B</p>
<p>9 Warum sagt Mama: „Ich hätte gerne deine Aufgabe“?</p> <p><input type="radio"/> A Marie tut ihrer Mama leid.</p> <p><input type="radio"/> B Marie sollte im Haushalt mehr Aufgaben übernehmen.</p> <p><input type="radio"/> C Mama kümmert sich sehr gerne um Hühner.</p> <p><input checked="" type="radio"/> D Marie soll verstehen, dass Mama schwerere Aufgaben hat.</p> <p>Kompetenzstufe III (.56/.56) S</p>	<p>10 Wie kommt Marie auf ihre Idee?</p> <p><input type="radio"/> A Maries Bruder Simon erzählt ihr von dem Plan.</p> <p><input checked="" type="radio"/> B Marie sieht, wie eine Eule eine Maus fängt.</p> <p><input type="radio"/> C Maries Papa erzählt ihr von Eulen.</p> <p><input type="radio"/> D Marie sieht Draht und weißen Stoff.</p> <p>Kompetenzstufe III (.72/.83) L</p>	<p>11 Warum befestigt Marie weiße Flügel an einer Stange?</p> <p><input type="radio"/> A damit es wie Hühnerfedern aussieht</p> <p><input type="radio"/> B um eine Entscheidung zu treffen</p> <p><input checked="" type="radio"/> C damit es wie eine Eule aussieht</p> <p><input type="radio"/> D um Simon zu beeindrucken</p> <p>Kompetenzstufe II (.82/.87) S</p>	<p>12 Marie „schlug mit den Händen nach den Flügeln und stieß die Flügel weg“. Marie möchte, dass das Huhn etwas denkt. Was soll es denken?</p> <p><input checked="" type="radio"/> A dass Marie das Huhn rettet</p> <p><input type="radio"/> B dass Marie wütend auf das Huhn ist</p> <p><input type="radio"/> C dass Marie sich vor der Eule fürchtet</p> <p><input type="radio"/> D dass Marie mit der Eule spielt</p> <p>Kompetenzstufe II (.73/.80) I</p>
<p>13 Wie Marie ist, merkst du an den Dingen, die sie tut. Beschreibe, wie Marie ist, und nenne zwei Beispiele aus der Geschichte, die das zeigen.</p> <p><input type="radio"/> A <u>Sie ist klug, weil sie das Huhn in den Käfig bekam. Sie gibt nicht auf, denn sie versucht es 4- oder 5-mal.</u></p> <p>Kompetenzstufe V (.17/.16) I</p>	<p>14 Warum ist Marie am Ende der Geschichte ganz oben in der Hackordnung? Erkläre deine Antwort mit Informationen aus der Geschichte.</p> <p><input type="radio"/> A <u>Marie gab vor, das Huhn zu retten.</u></p> <p>Kompetenzstufe III (.50/.72) I</p>	<p>15 Was wird das rote Huhn wohl tun, wenn Marie die Hühner das nächste Mal in ihren Käfig bringt?</p> <p><input type="radio"/> A <u>Es wird folgsam sein.</u></p> <p>Kompetenzstufe III (.71/.82) I</p>	<p>16 Warum wäre „Marie findet einen Weg“ eine andere gute Überschrift für diese Geschichte? Nenne einen Grund.</p> <p><input type="radio"/> A <u>Sie findet einen Weg, das Huhn in den Käfig zu bekommen, ohne dass sie darum kämpfen muss.</u></p> <p>Kompetenzstufe IV (.41/.53) B</p>

L = Lokalisieren explizit angegebener Informationen.

S = Einfache Schlussfolgerungen ziehen.

I = Komplexe Schlussfolgerungen ziehen bzw. Interpretieren und Kombinieren.

B = Prüfen und Bewerten des Inhalts und des Sprachgebrauchs.

Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: <http://www.iea.nl/permissions.html>.

Für die Kodierung und die Vergabe von Punkten wurden Richtlinien erstellt, die allgemeine Hinweise und Beispiele für akzeptable und nicht akzeptable Antworten enthalten. Dabei kann die Entscheidung, welche Antwort als gerade noch akzeptabel beziehungsweise als schon nicht mehr akzeptabel einzustufen ist, in dem einen oder anderen Fall als problematisch angesehen werden. Die Hinweise zur Kodierung der offenen Aufgabe 3 zum Beispiel lauten:

Akzeptabel: In der Antwort wird deutlich, dass alle anderen Hühner nach dem Huhn hacken oder dass es selbst nach keinem hacken kann oder dass es die schlechtesten Bedingungen hat.

Beispiele: Alle anderen Hühner können nach ihm hacken. – Es kann nach keinem anderen Huhn hacken. – Es wird herumkommandiert. – Das oberste Huhn kommandiert es weiter herum. – Es bekommt die schlechtesten Futterstücke.

Nicht akzeptabel: Aus der Antwort ergibt sich, dass die Hackordnung missverstanden wurde, oder sie ist vage, nicht auf den Text bezogen oder eine Wiederholung von Teilen der Aufgabe.

Beispiele: Es könnte krank sein. – Alle anderen Hühner sitzen über ihm. – Weil es ein langer Weg ist, Chef zu werden. – Weil Hühner gern einen herrischen Chef haben.

Hier ließe sich einwenden, dass auch die Antwort *Alle anderen Hühner sitzen über ihm.* noch akzeptabel sein kann.

Ein weiteres Beispiel dafür, wie schwierig es sein kann, eine Grenze zwischen einer akzeptablen und einer nicht akzeptablen Antwort zu ziehen, sind die Kodierrichtlinien für die offene Aufgabe 8:

Akzeptabel: Es ist verstanden, dass das Muster Bestand hat, auch wenn das rote Huhn getötet wird; Maries Problem ist damit nicht gelöst; das nächste Huhn wird denselben Ärger/dasselbe Problem bereiten; es wird sich auf dieselbe Weise verhalten und genauso ärgern wie das rote Huhn.

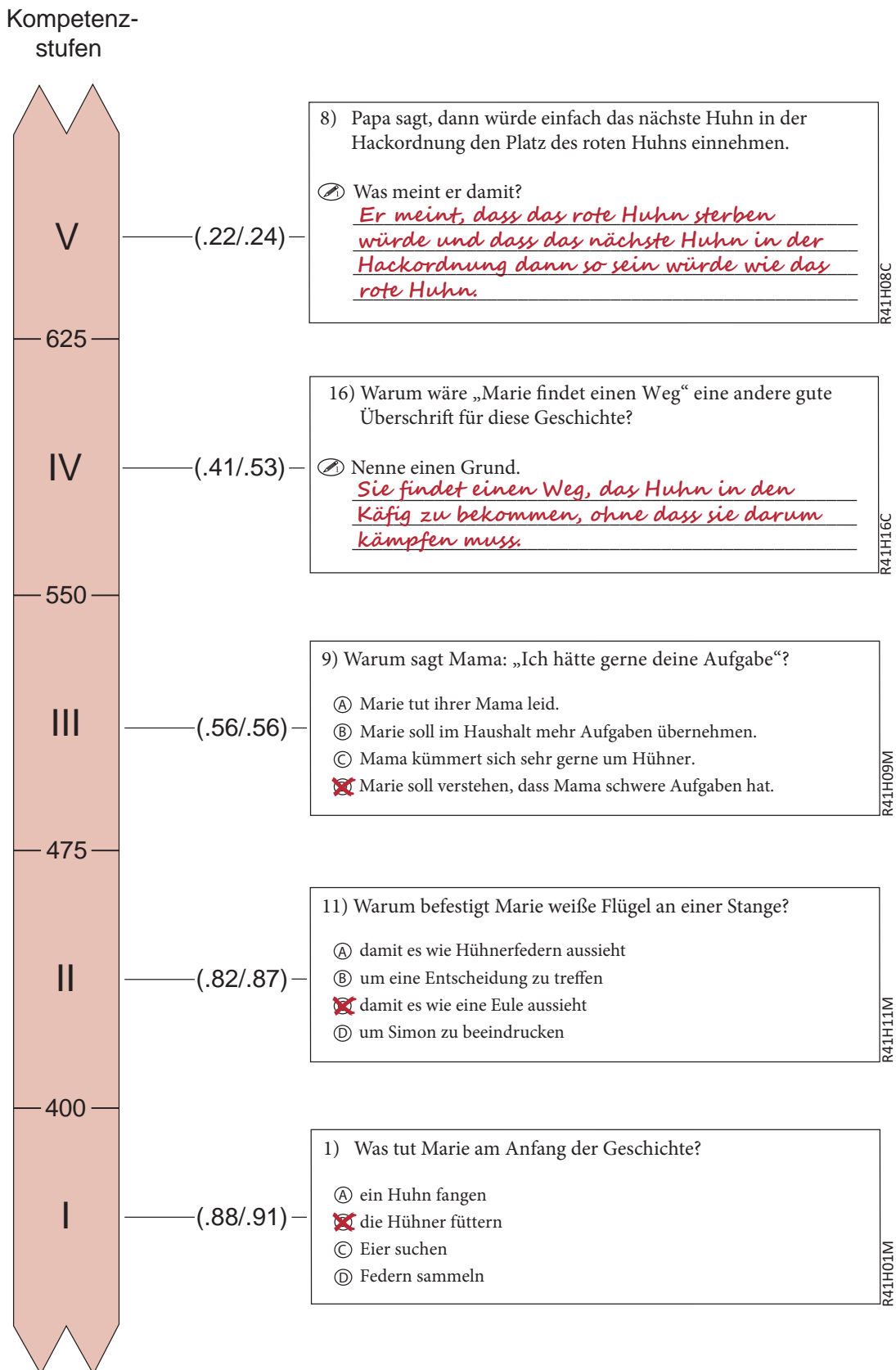
Beispiele: Das nächste Huhn wäre dann oben und es wäre keine Lösung zuzulassen, dass der Fuchs das rote Huhn tötet. – Wenn das rote Huhn getötet wird, wird das nächste Huhn in der Hackordnung dasselbe machen. – Dann wird ein anderes Huhn Marie ärgern. – Er meint, dass das rote Huhn sterben würde und dass das nächste Huhn in der Hackordnung dann so sein würde wie das rote Huhn. – Maries Probleme würden nicht gelöst.

Nicht akzeptabel: Die Antwort bezieht sich auf eine andere Henne, die das rote Huhn als Führerin ersetzt, oder auch nicht, aber sie zeugt nicht davon, dass das Muster beziehungsweise das Problem damit fortbestehen würde.

Beispiele: Das zweite Huhn wird das oberste sein. – Ein anderes Huhn wird als Chefin übernehmen. – Wenn das rote Huhn weg ist, wird es durch eine neue Führerin ersetzt. – Das Huhn, das am zweitstärksten ist, würde seinen Platz einnehmen. – Es bedeutet, dass das rote Huhn getötet werden wird. – Dass das nichts ändern wird. – Der Vater will ein anderes Huhn an die erste Stelle setzen.

Hier kann gefragt werden, worin der qualitative Unterschied zwischen der akzeptablen Antwort *Dann wird ein anderes Huhn Marie ärgern.* und den ersten vier als nicht akzeptabel gewerteten Antworten besteht.

Abbildung 3.3: Kompetenzstufen und Beispielaufgaben (literarischer Text)



Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.
Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: <http://www.iea.nl/permissions.html>.

4.2.2 Ein Sachtext, die Aufgaben und Beispiele für Anforderungen auf verschiedenen Kompetenzstufen

Die Grüne Meeresschildkröte und die Reise ihres Lebens

Aus *Die Reisen der Schildkröte*
von Gary Miller



Heraus aus dem Sand

Es ist eine sternklare Nacht im August. Ein Nest mit Eiern ist über sechzig Zentimeter tief unter der Oberfläche eines Strands in Costa Rica im Sand vergraben. Im Nest liegen über 100 Eier der Grünen Meeresschildkröte, jedes davon ist ungefähr so groß wie ein Golfball.

Eines der Schildkrötenbabys beginnt sich zu bewegen, um aus seinem Ei zu schlüpfen. Der Schlüpfling zerbricht mit der scharfen Spitze seines Schnabels die Schale seines Eis. Noch immer im Sand vergraben, befreit sich das Schildkrötenbaby. Bald ist im ganzen Nest Leben und Bewegung.

Das Schildkrötenbaby benutzt seine Flossen, um immer weiter nach oben zu klettern. Es kann über einen Tag dauern, bis es die Oberfläche des Sands erreicht.

Hinein ins Wasser

Wenn der Schlüpfling die Oberfläche des Sandes erreicht, zieht ihn das Mondlicht an, das sich auf dem Ozean spiegelt. Zum Glück scheinen keine Lichter an einer Straße oder einem Haus in der Nähe. Diese Lichter können ein Schildkrötenbaby durcheinanderbringen. Sie können dafür sorgen, dass es in die falsche Richtung kriecht, weg vom Meer.

Die Reise des Schlüpfings zum Wasser ist ein Rennen ums Überleben. Er ist nicht größer als eine Walnuss. Krebse und Vögel, wie zum Beispiel Nachtreiher, schnappen sich einige der anderen Schildkrötenbabys am Strand. Aber dieses Schildkrötenbaby schafft es bis zum Wasser.

Die schäumende Brandung schiebt das Schildkrötenbaby zurück. Es kämpft, um gegen die sich brechenden Wellen anzuschwimmen. Der Schlüpfing schwimmt den ersten Tag und die erste Nacht durch und er wird zwei Tage lang nicht langsamer.

Hinaus aufs offene Meer

Die Reise des Schildkrötenbabys durch das offene Meer nennt man oft die „verlorenen Jahre“. Wissenschaftler wissen wenig über diese Phase des Lebens einer Grünen Meeresschildkröte. Wahrscheinlich bewegt sie sich mit den Strömungen und lässt sich mit Matten aus Seetang treiben.



Der Schlüpfing frisst wahrscheinlich Krabben, kleine Quallen und Schnecken, die im und um den Seetang treiben. Leider gibt es im Meer auch Plastik und Müll, den die Menschen wegwerfen. Das zu fressen könnte für die Schildkröte tödlich sein.

Im Meer gibt es noch viele andere Gefahren. Raubtiere wie zum Beispiel Haie schwimmen unter der kleinen Schildkröte und über ihr fliegen große Vögel. Zum Glück bietet ihr die Färbung ihres Panzers etwas Schutz. Die Unterseite ist fast weiß, damit Haie, die unter ihr schwimmen, sie im Sonnenlicht nicht entdecken. Die Oberseite des Panzers ist dunkel, sodass sich die Schildkröte von oben betrachtet vom dunklen Wasser nicht abhebt.

Größer und grün werden

Nach mehreren Jahren ist sie ein Jungtier geworden. Sie ist kein Schlüpfing mehr, aber auch noch nicht erwachsen. Ihr Panzer ist nun so groß geworden wie ein Essteller. Jetzt ist es an der Zeit, das offene Meer zu verlassen und in die warmen Küstengewässer Floridas in den USA zu ziehen.

Mit dem größeren Panzer ist sie jetzt sicherer, als sie es als Schlüpfing war. Manchmal schlürft sie eine Qualle, aber meistens frisst sie jetzt Algen und Seegras.

Während sie langsam größer wird, vergehen die Jahre. Sie zieht zu Nahrungsgründen weiter weg von der Küste, wo sie erwachsen wird.

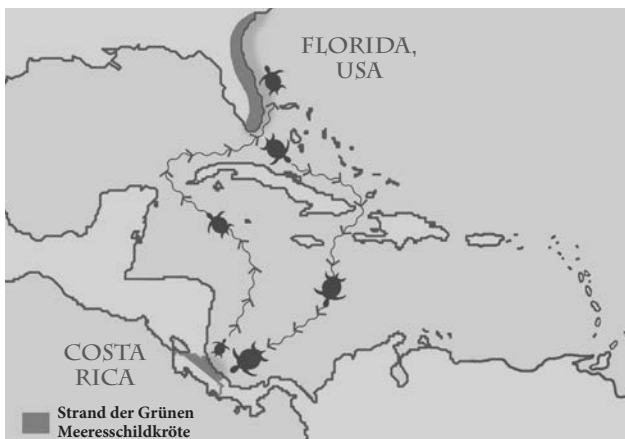
Nachts ruht sie sich im Wasser unter Felsen und Felsvorsprüngen aus, wobei sie bis zu fünf Stunden den Atem anhält. Jeden Tag kehrt sie zum selben Platz mit Seegras zurück,

das „Schildkrötengras“ genannt wird. Wie ein Rasenmäher hält die Schildkröte die Seegraswiese kurz. Durch das Fressen von Seegras und Algen färbt sich ihr Körperfett grün. So ist die Grüne Meeresschildkröte zu ihrem Namen gekommen!



Zurück zum Sand

Wenn die Schildkröte ungefähr 26 Jahre alt ist, ist ihr ausgewachsener Panzer über 90 cm lang, und sie wiegt ungefähr 135 Kilo. Jetzt bricht sie zu einem neuen Abenteuer auf. Sie beginnt ihre lange Reise zurück an den Strand, an dem sie geboren wurde. Sie wird selbst Eier legen.



Die Meeresschildkröte muss vielleicht über 950 Kilometer schwimmen, aber sie ist für die Reise gut gerüstet. Ihre Flossen sind wie Flügel. Sie fliegt durch das Wasser.

Wissenschaftler sind noch dabei herauszufinden, wie Meeresschildkröten ihren Weg durch den Ozean finden. Sie vermuten, dass die Schildkröten Veränderungen des Magnetfelds der Erde spüren können. Das könnte den Schildkröten helfen, eine Art geistige Landkarte anzulegen. Auch die Erinnerung an Chemikalien oder Gerüche im Wasser könnte ihnen helfen, ihren Weg zu finden.



Sobald die Grüne Meeresschildkröte wieder an ihrem Geburtsort ist, sucht sie sich einen Partner. Ein paar Wochen später wartet sie, bis es dunkel ist, und kriecht an den Strand.

Die nächste Generation

Einmal aus dem Wasser, fällt es der Schildkröte schwer, sich an Land zu bewegen. Sie kriecht zu einer Stelle, von der die Flut ihre Eier nicht wegspülen kann. Mit ihren Vorderflossen gräbt sie eine breite Grube. Das wird ihr Nest. Mit den Hinterflossen hebt sie in der Grube ein kleineres Loch aus.

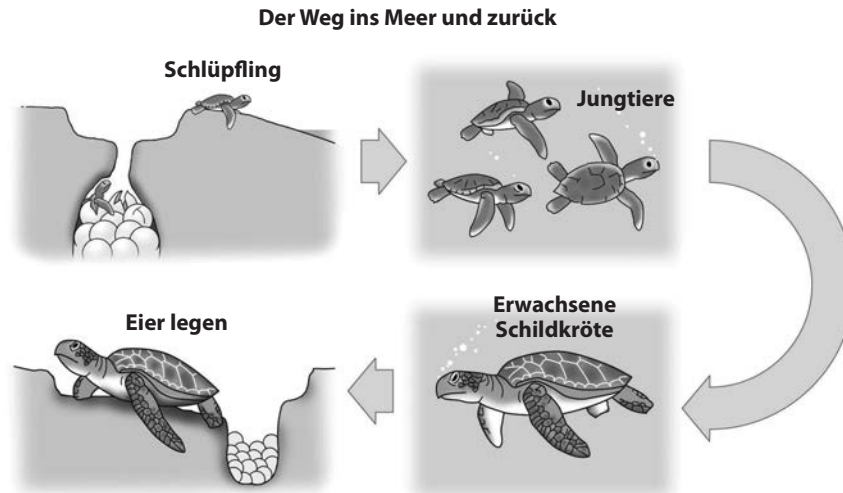
Nach zwei Stunden harter Arbeit ist sie bereit, über 100 lederartige weiße Eier in das kleinere, tiefere Loch zu legen. Sie häuft Sand darauf. Dann wirft sie Sand auf das ganze Nest.

In den folgenden zwei Monaten wird sie noch drei Nester graben und Eier hineinlegen. Nach zwei Monaten befreien sich die neuen Schlüpflinge aus ihrer Schale und beginnen ihre eigene Reise.

Das Fortbestehen der Schildkröten

Wenn die erwachsene Schildkröte all ihre Eier gelegt hat, bricht sie wieder zu ihren Nahrungsgründen vor der Küste Floridas auf. Alle paar Jahre kehren sie und andere erwachsene Schildkröten an diesen Strand zurück, um weitere Eier zu legen.

Jede Grüne Meeresschildkröte tut das ihr ganzes Leben lang, das bis zu 80 Jahre dauern kann. In dieser Zeit werden Tausende Babys der Grünen Meeresschildkröte geboren und machen sich auf den Weg ins offene Meer.



Dieser für Viertklässlerinnen und Viertklässler recht lange Text beginnt wie ein Erzähltext. Das Geschehen wird raumzeitlich situiert und es ist von einem bestimmten Schildkrötenbaby die Rede. Bereits im ersten Abschnitt deutet sich aber an, dass dieses Baby als Prototyp anzusehen ist, als Vertreter seiner Art, von dessen individuellen Merkmalen abgesehen wird (*Es kann über einen Tag dauern, bis es die Oberfläche des Sands erreicht.*). Die Lebensabschnitte kommen in „natürlicher“ Abfolge zur Sprache, die letzte Etappe zwischen Erwachsensein und Tod wird als Wiederkehr des Gleichen dargestellt. Rückverweise kommen nicht vor, ein Vorverweis (*Sie wird selbst Eier legen.*) dürfte das Verständnis nicht beeinträchtigen. Der Autor verfährt nicht ausschließlich deskriptiv, sondern streut – sozusagen Partei für die Schildkröte ergreifend – Wertungen ein (*zum Glück, leider*). Die Gliederung wird durch Zwischenüberschriften erleichtert, die resultierenden Abschnitte sind vergleichsweise kurz. Vier photographische Abbildungen haben im Wesentlichen dekorative Funktion, allenfalls die dritte könnte funktional sein, insofern hier eine Matte aus Seetang dargestellt ist. Die Landkarte illustriert den „Reiseweg“ von Puerto Rico nach Florida und umgekehrt. Für die Lösung der Aufgaben (s.u.) ist ihr Verständnis nicht nötig. Anders verhält es sich mit der schematischen Darstellung am Textende. Deren Teile korrespondieren mit den Abschnittüberschriften. Sie erleichtern insofern das Verständnis der Textgliederung unter der Voraussetzung, dass die durch die Pfeile signalisierte, partiell ungewöhnliche Leserichtung eingehalten wird.

Auf lexikalischer Ebene dürfte der Text, soweit das im „Lehnstuhl“ zu beurteilen ist, in der Regel keine größeren Schwierigkeiten bereiten. Das schließt ein, dass in der konkreten Unterrichtssituation bei einzelnen Kindern oder Subgruppen lexikalische Lücken deutlich werden können. Dass das Fachwort *Schlüpfling* und *Schildkrötenbaby* synonym verwendet werden, ist kontextuell leicht zugänglich.

Sollte die Bedeutung eines Wortes wie *Algen* nicht bekannt sein, ist das hier un-
schädlich, reicht doch das semantische Merkmal [*essbar*]. Nicht erläutert wird
die Phrase *Veränderungen des Magnetfelds der Erde*. Wie sie zu verstehen ist,
dürfte auch den meisten Erwachsenen nicht klar sein. Folgerichtig spielt sie in
keiner Aufgabe eine Rolle. In syntaktischer Hinsicht bietet der Text ebenfalls
kaum Hürden. Es dominieren einfache Hauptsätze, Satzreihen und Satzgefüge
mit einem Nebensatz oder zwei koordinierten Nebensätzen.

Aus Abbildung 3.4 wird ersichtlich, wie die 16 Aufgaben den vier Ver-
stehensprozessen zugeordnet wurden. 4 der 16 Aufgaben erfordern das Lokali-
sieren von explizit angegebenen Informationen, bei 7 Aufgaben sind einfache
Schlussfolgerungen zu ziehen. Drei Aufgaben sind dem Verstehensprozess *Kom-
plexe Schlussfolgerungen ziehen bzw. Interpretieren und Kombinieren* zugeordnet
und 2 dem *Prüfen und Bewerten des Inhalts und des Sprachgebrauchs*.


Was die Kodierung der offenen Aufgaben angeht, so ist es hier wie beim lite-
rarischen Text nicht immer einfach, die Grenzen zwischen noch akzeptablen und
nicht mehr akzeptablen Antworten zu bestimmen. So lauten die Richtlinien für
Aufgabe 10:

*Eine akzeptable Antwort enthält den Hinweis auf grünes Futter: Seegras. –
Sie fressen Algen und Seetang, das macht sie grün. – Sie frisst etwas, was grün
ist. – Gras. – Sie isst Pflanzen.*

*Nicht akzeptabel sind Antworten ohne oder eine falsche Angabe von
Ursachen. Sie sind auch vage, haben keinen Bezug zum Text oder sind
Wiederholungen von Teilen der Aufgabe: Sie hat grünes Körperfett. – So bekam
sie ihren Namen. – Sie wird grün von dem, was sie isst. – Sie wird erwachsen. –
Ihr Fett ist grün, weil sie eine grüne Schildkröte ist. – Weil sie alt ist.*

Für *Sie frisst etwas, was grün ist* wird also ein Punkt vergeben, nicht aber für
Sie wird grün von dem, was sie isst. Hier ist zwar ein Bedeutungsunterschied ge-
geben, denn in der als nicht akzeptabel eingestuften Antwort ist nicht explizit ge-
sagt, dass das Futter grün ist. Insofern das aber gemeint sein dürfte, kann gefragt
werden, ob hier nicht zu streng verfahren wird.

Abbildung 3.4: Aufgaben zum Sachtext *Die Grüne Meeresschildkröte und die Reise ihres Lebens*

<p>R41T01M</p> <p>1 Worum geht es im ersten Abschnitt „Heraus aus dem Sand“?</p> <p><input type="radio"/> A darum, wie verschiedene Meeresschildkröten aussehen</p> <p><input type="radio"/> B darum, wie Meeresschildkröten schwimmen lernen</p> <p><input type="radio"/> C darum, was Meeresschildkröten gerne fressen</p> <p><input checked="" type="radio"/> D darum, wie Meeresschildkröten aus dem Ei schlüpfen</p> <p>Kompetenzstufe II (.85/.93) S</p>	<p>R41T02C</p> <p>2 „Eines der Schildkrötenbabys beginnt sich zu bewegen, um aus seinem Ei zu schlüpfen.“</p> <p>Schreibe die ersten zwei Dinge auf, die der Schlüpfing als Nächstes tut.</p> <p><input type="radio"/> <u>1. Es zerreißt die Schale.</u></p> <p><input type="radio"/> <u>2. Es geht zum Wasser.</u></p> <p>Kompetenzstufe III (.53/.62) L</p>	<p>R41T03C</p> <p>3 Wenn der Schlüpfing die Oberfläche des Sands erreicht, was hilft ihm, in die richtige Richtung zu kriechen?</p> <p><input type="radio"/> <u>Mondlicht</u></p> <p>Was kann die Schlüpfinge durcheinanderbringen?</p> <p><input type="radio"/> <u>Lichter von Autos</u></p> <p>Kompetenzstufe III (.42/.52) S</p>	<p>R41T04C</p> <p>4 Warum ist die Reise des Schlüpfings zum Wasser ein „Rennen ums Überleben“? Benutze den Text, um deine Antwort zu erklären.</p> <p><input type="radio"/> <u>Wegen der Fressfeinde/ Feinde</u></p> <p>Kompetenzstufe III (.57/.80) S</p>															
<p>R41T05M</p> <p>5 Was macht der Schlüpfing als Erstes, wenn er es hinter die sich brechenden Wellen geschafft hat?</p> <p><input type="radio"/> A er sucht nach den anderen Schlüpfingen</p> <p><input checked="" type="radio"/> B er schwimmt weiter, weit aufs Meer hinaus</p> <p><input type="radio"/> C er ruht sich im Seetang aus</p> <p><input type="radio"/> D er sucht sich Futter</p> <p>Kompetenzstufe III (.58/.64) L</p>	<p>R41T06C</p> <p>6 Im Text werden verschiedene Arten genannt, wie die Menschen die Umwelt für Schildkröten gefährlicher machen. Nenne ein Beispiel.</p> <p><input type="radio"/> <u>Schadstoffe</u></p> <p>Kompetenzstufe II (.78/.91) S</p>	<p>R41T07C</p> <p>7 Die Farbe des Panzers schützt den Schlüpfing vor Raubtieren. Wie schützt ihn die Farbe vor Vögeln?</p> <p><input type="radio"/> <u>Die obere Seite der Schildkröte ist dunkel, so dass sie sich nicht vom dunklen Wasser abhebt.</u></p> <p>Wie schützt ihn die Farbe vor Haien?</p> <p><input type="radio"/> <u>Die untere Seite der Schildkröte ist weiß, so dass es für Haie von unten wie Sonnenlicht aussieht.</u></p> <p>Kompetenzstufe V (.26/.35) I</p>	<p>R41T08C</p> <p>8 Wann hält eine Meeresschildkröte bis zu 5 Stunden lang den Atem an?</p> <p><input type="radio"/> <u>Beim Schlafen</u></p> <p>Kompetenzstufe III (.60/.69) L</p>															
<p>R41T09M</p> <p>9 Was steht im Sachtext über die Fressgewohnheiten einer erwachsenen Grünen Meeresschildkröte?</p> <p><input type="radio"/> A Sie sucht unter Felsen und Felsvorsprüngen nach Futter.</p> <p><input type="radio"/> B Sie schwimmt weit, um Futter zu finden.</p> <p><input checked="" type="radio"/> C Sie schwimmt jeden Tag zum Fressen an denselben Ort.</p> <p><input type="radio"/> D Sie nutzt Gerüche im Wasser, um Futter zu finden.</p> <p>Kompetenzstufe IV (.59/.64) S</p>	<p>R41T10C</p> <p>10 Warum wird das Körperfett der Meeresschildkröte grün?</p> <p><input type="radio"/> <u>Sie fressen Algen, das macht sie grün.</u></p> <p>Kompetenzstufe III (.59/.76) L</p>	<p>R41T11C</p> <p>11 Welche Informationen liefert der Text zur Größe und zum Futter der Meeresschildkröte in jeder Phase ihres Lebens? Fülle die folgende Tabelle aus. Drei Felder sind schon ausgefüllt.</p> <table border="1"> <thead> <tr> <th>Lebensphase</th> <th>Größe</th> <th>Futter</th> </tr> </thead> <tbody> <tr> <td>Ei</td> <td><u>Golfball</u></td> <td>Im Ei ist Nahrung vorhanden.</td> </tr> <tr> <td>Schlüpfing</td> <td><u>Walnuss</u></td> <td><u>Krabben, Quallen, Schnecken</u></td> </tr> <tr> <td>Jungtier</td> <td>Essteller</td> <td><u>Algen, Seegras, Quallen</u></td> </tr> <tr> <td>Erwachsene Schildkröte</td> <td><u>90 cm</u> <u>135 kg</u></td> <td>Algen und Seegras</td> </tr> </tbody> </table> <p>Kompetenzstufe IV (.42/.44)¹ I</p>	Lebensphase	Größe	Futter	Ei	<u>Golfball</u>	Im Ei ist Nahrung vorhanden.	Schlüpfing	<u>Walnuss</u>	<u>Krabben, Quallen, Schnecken</u>	Jungtier	Essteller	<u>Algen, Seegras, Quallen</u>	Erwachsene Schildkröte	<u>90 cm</u> <u>135 kg</u>	Algen und Seegras	<p>R41T12M</p> <p>12 Wie alt ist eine weibliche Grüne Meeresschildkröte, wenn sie zum ersten Mal aufbricht, um Eier zu legen?</p> <p><input type="radio"/> A ungefähr 3 Jahre</p> <p><input type="radio"/> B ungefähr 10 Jahre</p> <p><input checked="" type="radio"/> C ungefähr 26 Jahre</p> <p><input type="radio"/> D ungefähr 80 Jahre</p> <p>Kompetenzstufe III (.74/.76) S</p>
Lebensphase	Größe	Futter																
Ei	<u>Golfball</u>	Im Ei ist Nahrung vorhanden.																
Schlüpfing	<u>Walnuss</u>	<u>Krabben, Quallen, Schnecken</u>																
Jungtier	Essteller	<u>Algen, Seegras, Quallen</u>																
Erwachsene Schildkröte	<u>90 cm</u> <u>135 kg</u>	Algen und Seegras																
<p>R41T13M</p> <p>13 Welches Verhalten einer erwachsenen weiblichen Grünen Meeresschildkröte verstehen die Wissenschaftler noch nicht vollständig?</p> <p><input type="radio"/> A wie sie über 950 Kilometer schwimmen kann</p> <p><input type="radio"/> B wie sie ein Nest für ihre Eier baut</p> <p><input type="radio"/> C wie sie es verhindert, von Raubtieren gefressen zu werden</p> <p><input checked="" type="radio"/> D wie sie den richtigen Strand für das Eierlegen findet</p> <p>Kompetenzstufe IV (.49/.56) S</p>	<p>R41T14C</p> <p>14 Hier siehst du eine Abbildung aus dem Sachtext.</p>  <p>Was kannst du durch die Abbildung besser verstehen?</p> <p><input type="radio"/> <u>Was die einzelnen Lebensabschnitte sind.</u></p> <p>Kompetenzstufe III (.54/.51) B</p>	<p>R41T15M</p> <p>15 Wie zeigt der Autor dir, dass die Grüne Meeresschildkröte etwas besonderes ist?</p> <p><input type="radio"/> A Er bittet dich mitzuhelfen, sie zu schützen.</p> <p><input checked="" type="radio"/> B Er erzählt dir, was für erstaunliche Dinge sie tut.</p> <p><input type="radio"/> C Er beschreibt, wie schön sie aussieht.</p> <p><input type="radio"/> D Er warnt, dass nur noch wenige Schildkröten leben.</p> <p>Kompetenzstufe IV (.61/.65) B</p>	<p>R41T16M</p> <p>16 Der Sachtext ist in Abschnitte mit Überschriften unterteilt. Wovon handeln die einzelnen Abschnitte?</p> <p><input type="radio"/> A von verschiedenen Gefahren für Meeresschildkröten</p> <p><input checked="" type="radio"/> B von verschiedenen Lebensphasen einer Meeresschildkröte</p> <p><input type="radio"/> C von verschiedenen Meeresschildkröten-Arten</p> <p><input type="radio"/> D von verschiedenen Ansichten über Meeresschildkröten</p> <p>Kompetenzstufe III (.71/.68) B</p>															

L = Lokalisieren explizit angegebener Informationen.

S = Einfache Schlussfolgerungen ziehen.

I = Komplexe Schlussfolgerungen ziehen bzw. Interpretieren und Kombinieren.

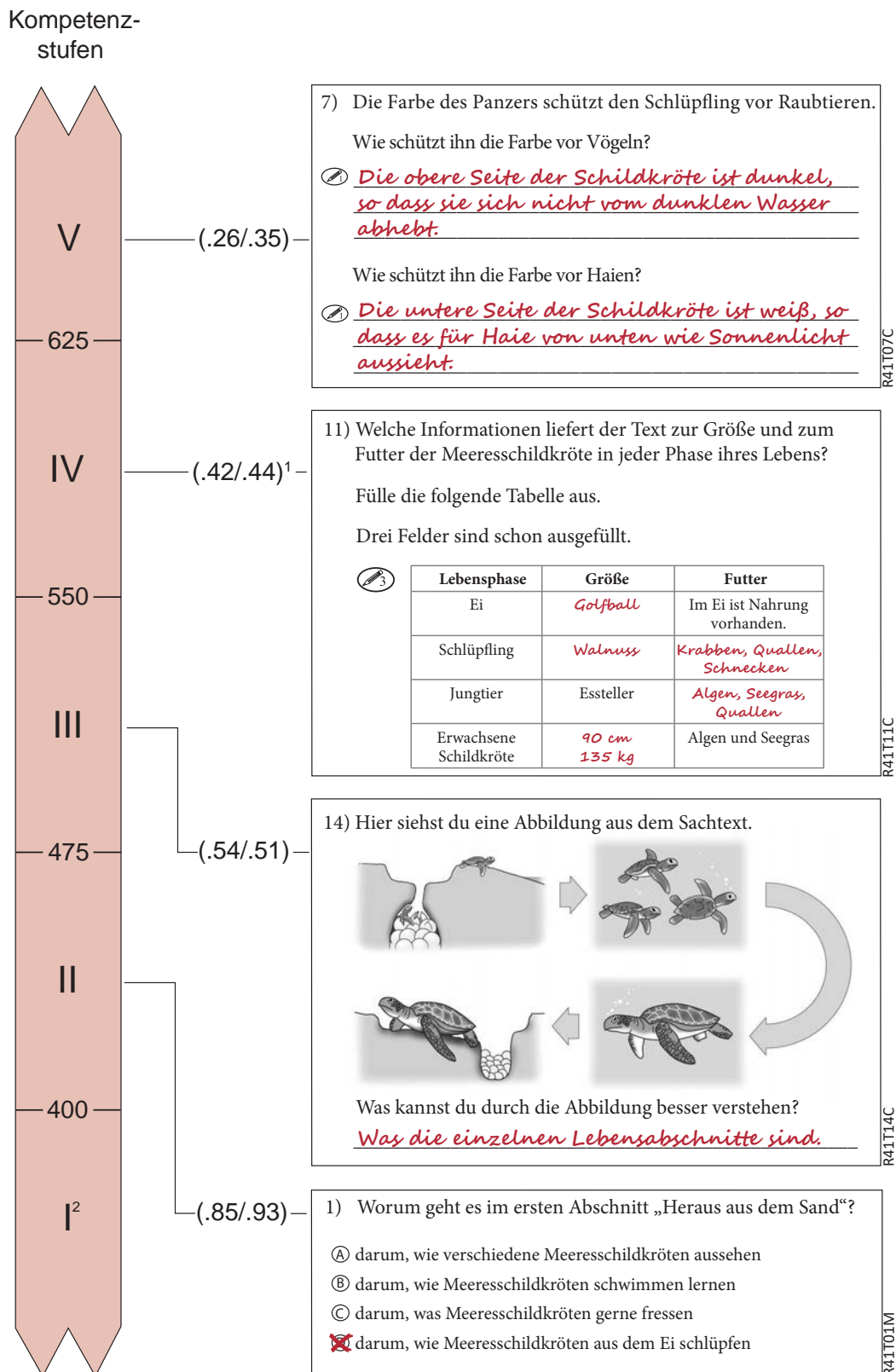
B = Prüfen und Bewerten des Inhalts und des Sprachgebrauchs.

Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.

1 = Lösungshäufigkeit für eine fast bzw. komplett gelöste Aufgabe (2 oder 3 Punkte).

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: <http://www.iea.nl/permissions.html>.

Abbildung 3.5: Kompetenzstufen und Beispielaufgaben (Sachtext)



Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.

1 = Lösungshäufigkeit für eine fast bzw. komplett korrekt gelöste Aufgabe (2 oder 3 Punkte).

2 = Für die Kompetenzstufe I liegt für diesen Text keine Beispielaufgabe vor.

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: <http://www.iea.nl/permissions.html>.

5 Zentrale Befunde aus IGLU 2001, 2006 und 2011

Deutschland beteiligt sich seit 2001 regelmäßig an IGLU (siehe Kapitel 2 in diesem Band). Damit liegen repräsentative und belastbare Daten vor, die es ermöglichen, die Lesekompetenzen von Viertklässlerinnen und Viertklässlern über einen Zeitraum von 15 Jahren zu beschreiben und Ergebnisse in einen internationalen Vergleichsrahmen einzuordnen. Im Folgenden werden zunächst wesentliche Befunde aus IGLU 2001, 2006 und 2011 zusammengefasst (vgl. Bos et al., 2003; Bos, Valtin, Hornberg et al., 2007; Bos, 2012). Abweichungen bei den Punktwerten und Signifikanzen ergeben sich aus der Reskalierung der Daten im Jahr 2011 und optimierten Signifikanztests (siehe Kapitel 2):

- In IGLU lag die auf der Gesamtskala Lesen beschriebene Lesekompetenz von Viertklässlerinnen und Viertklässlern in Deutschland bei 539 (2001), 548 (2006) und 541 (2011) Punkten. Damit zeigte sich eine signifikante Verbesserung von 2001 zu 2006, jedoch dann eine signifikante Verschlechterung von 2006 zu 2011.
- Im internationalen Vergleich lag Deutschland in IGLU 2001 und 2006 im oberen Drittel der Rangreihe, wobei 2001 nur für fünf Teilnehmerstaaten und -regionen (Schweden, Niederlande, England, Bulgarien, Ontario) und 2006, trotz gesteigerter Teilnehmerzahl, nur für sieben Teilnehmer (Russische Föderation, Hongkong, Singapur, Luxemburg sowie die kanadischen Provinzen Alberta und Britisch Kolumbien und Ontario) signifikant bessere mittlere Schülerleistungen festzustellen waren. In IGLU 2011 ließ sich Deutschland, bei zunehmender Zahl an Teilnehmerländern sowie einigen Ländern mit deutlichen Leistungszuwächsen, nur noch im Mittelfeld verorten. Für 15 Teilnehmerstaaten und -regionen konnten 2011 signifikant bessere mittlere Schülerleistungen festgestellt werden.
- Hingegen zeigte sich konstant in allen Erhebungsrunden in Deutschland eine im internationalen Vergleich relativ geringe Streuung.
- Die Anteile an Schülerinnen und Schülern mit vergleichsweise schwachen beziehungsweise starken Leseleistungen blieben weitestgehend unverändert. 2001 lag der Anteil leistungsschwacher Schülerinnen und Schüler bei 17 und 2011 bei 15,4 Prozent. Die bessere Quote 2006 (12,3 %) wurde 2011 nicht mehr erreicht (15,4 %) (Bos et al., 2012, S. 105). Vor dem Hintergrund der Curricula für die Grundschule und der schulischen Anforderungen in der Sekundarstufe I wurden diese Befunde als kritisch interpretiert. Hinzu kam, dass die Anteile zwischen 8,6 und 10,8 Prozent der Schülerinnen und Schülern mit Leistungen auf dem Niveau der höchsten Kompetenzstufe im internationalen Vergleich gering ausfielen (ebd.).
- Die Kinder in Deutschland zeigten 2001 keine unterschiedlichen Leistungen beim Lesen verschiedener Textsorten. In den beiden folgenden Erhebungen ergaben sich zunehmend signifikant bessere Leistungen beim Lesen literarischer Texte.
- Die Viertklässlerinnen und -klässler in Deutschland sind seit 2001 im internationalen Vergleich vergleichsweise schwach, wenn wissensbasierte Verstehensleistungen verlangt sind.
- Im Rahmen von IGLU 2001 konnte für einige und in 2006 für alle Länder der Bundesrepublik Deutschland durch die Erweiterung der Stichprobe ein Ländervergleich realisiert werden (Bos et al., 2004; Bos, Hornberg et al., 2008). Ab 2011 wird er – nicht nur für die Domäne Lesen – vom *Institut zur Qualitätsentwicklung im Bildungswesen (IQB)* in Berlin durchgeführt.

6 Leseleistungen von Schülerinnen und Schülern im internationalen Vergleich: Ergebnisse 2016

Im Folgenden werden zentrale Ergebnisse zu den mit IGLU 2016 erhobenen Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe im internationalen Vergleich berichtet. Einen zentralen Aspekt des Kapitels bildet der Vergleich der Befunde im Rahmen von IGLU 2001, 2006, 2011 und 2016. Leitend sind die folgenden Fragestellungen:

- Welche mittleren Leseleistungen zeigen Schülerinnen und Schüler in Deutschland am Ende der vierten Jahrgangsstufe im internationalen Vergleich? Wie unterscheiden sich die Ergebnisse von IGLU 2001, 2006, 2011 und 2016 (siehe Abschnitte 6.1 und 6.2)?
- Wie groß ist die Streuung der Leistungen der Schülerinnen und Schüler in Deutschland am Ende der vierten Jahrgangsstufe im internationalen Vergleich? Hat sich die Heterogenität seit 2001 verändert (siehe Abschnitt 6.3)?
- Wie lassen sich die Leistungskennwerte auf Kompetenzstufen einordnen? Wie groß sind die Gruppen der auffällig leistungsschwachen und leistungsstarken Kinder? Wie unterscheiden sich diesbezüglich die Ergebnisse von IGLU 2001, 2006, 2011 und 2016 (siehe Abschnitt 6.4)?
- Zeigen Schülerinnen und Schüler in Deutschland relative Stärken oder Schwächen, wenn nach literarischen und informierenden Texten unterschieden wird? Wie unterscheiden sich die Ergebnisse von IGLU 2001, 2006, 2011 und 2016 (siehe Abschnitt 6.5)?
- Welche Ergebnisse erzielen die Viertklässlerinnen und Viertklässler in Deutschland im internationalen Vergleich, differenziert nach Verstehensleistungen? Wie unterscheiden sich die Ergebnisse von IGLU 2001, 2006, 2011 und 2016 (siehe Abschnitt 6.6)?

Für Ergebnisse zu lesebezogenen Einstellungen sowie Unterschieden in den Lesekompetenzen in Abhängigkeit von Geschlecht, sozialer Herkunft und Migrationshintergrund sei auf die Kapitel 5, 6 und 7 in diesem Band verwiesen.

An IGLU 2016 nahmen 47 Staaten teil, darunter Norwegen mit der fünften Jahrgangsstufe. Belgien war zweimal vertreten, einmal für den flämischen und einmal für den französischen Sprachraum. Darüber hinaus gab es elf Benchmark-Teilnehmer, nämlich die Länder Abu Dhabi, Dubai, Dänemark (3. Jahrgangsstufe), Norwegen (4. Jahrgangsstufe) und Südafrika (5. Jahrgangsstufe) und die Städte beziehungsweise Regionen Andalusien, Buenos Aires, Madrid, Moskau, Ontario und Québec. Die Ergebnisse der Benchmark-Teilnehmer wurden zwar auf der Gesamtskala Lesen verortet; gehen aber nicht in die Berechnung des internationalen Mittelwerts ein (siehe Kapitel 2 in diesem Band).

In diesem Kapitel werden die Ergebnisse für alle Teilnehmer auf der Gesamtskala Lesen berichtet (siehe Abbildung 3.6); für alle weiteren Abbildungen und Tabellen wird eine Auswahl an Staaten anhand folgender Kriterien getroffen:

- (1) Teilnehmer, die Mitglieder der *Europäischen Union* (EU) sind: Belgien, Bulgarien, Dänemark, Deutschland, England, Finnland, Frankreich, Irland, Italien, Lettland, Litauen, Malta, Niederlande, Nordirland, Österreich, Polen, Portugal, Schweden, Slowakei, Slowenien, Spanien, Tschechien und Ungarn;
- (2) Teilnehmer, die der *Organisation for Economic Co-operation and Development* (OECD) angehören: Australien, Belgien, Chile, Dänemark, Deutschland, England, Finnland, Frankreich, Irland, Israel, Italien, Kanada, Lettland, Neuseeland, die Niederlande, Nordirland, Norwegen, Österreich, Polen,

- Portugal, Schweden, die Slowakei, Slowenien, Spanien, Tschechien, Ungarn und die USA;
- (3) Teilnehmer, die auf der Gesamtskala Lesen nicht signifikant schlechtere Ergebnisse erzielt haben als Deutschland.

Eine vollständige Darstellung der Ergebnisse für alle Teilnehmerstaaten und -regionen ist im Bericht der internationalen Studienleitung nachzulesen (Mullis et al., 2017)

Zur Einschätzung der Leseleistungen der Schülerinnen und Schüler in Deutschland werden in Abbildungen und Tabellen, in denen die Ergebnisse aus 2016 dargestellt werden, neben dem internationalen Mittelwert als Vergleichswerte die Mittelwerte der Staaten angegeben, die Mitglieder der EU beziehungsweise der OECD sind (siehe auch Kapitel 2, Abschnitt 4.2.1 in diesem Band). Diese werden als Vergleichsgruppe EU (VG EU) beziehungsweise Vergleichsgruppe OECD (VG OECD) in den Abbildungen und Tabellen aufgeführt. Aus Gründen der Vergleichbarkeit wird bei allen Darstellungen, in denen die Ergebnisse mehrerer Studienzyklen verglichen werden, auf die Angabe von Vergleichsgruppenwerten verzichtet, da sich die Anzahl der Teilnehmerstaaten und Zugehörigkeiten in den letzten 15 Jahren verändert haben und Veränderungen in Testergebnissen nur schwer zu interpretieren sind.

Das Format der Bezeichnungen der Teilnehmer in den Abbildungen sowie die beigeestellten Fußnoten verweisen auf Besonderheiten der nationalen Untersuchungspopulationen und -stichproben. Die Bedeutung der einzelnen Fußnoten wird ausführlich in Kapitel 2 erläutert. Um die internationale Vergleichbarkeit der Ergebnisse zu gewährleisten, verlangt die internationale Studienleitung die Einhaltung strenger methodischer Standards. Teilnehmerstaaten, die aus unterschiedlichen Gründen diese Vorgaben nicht eingehalten haben, sind durch eine Kursivsetzung ihres Namens in den Abbildungen und Tabellen gekennzeichnet. Dies trifft auf 15 der 47 Teilnehmer sowie auf 5 der 11 Benchmark-Teilnehmer zu. Während für die anderen 32 Teilnehmer (ohne Benchmark-Teilnehmer) – zu denen auch Deutschland zählt – die Vergleichbarkeit der Ergebnisse gegeben ist, gilt für die kursiv gesetzten Teilnehmer, dass auf eine Repräsentativität der Stichprobe für alle Viertklässlerinnen und Viertklässler des jeweiligen Teilnehmerstaats beziehungsweise der Region nur mit Einschränkung geschlossen werden kann. Für die vergleichende Interpretation der Ergebnisse Deutschlands mit kursiv gesetzten Teilnehmerstaaten und -regionen ist unter Berücksichtigung der Tabellen im Anhang dieses Bandes für den Einzelfall zu prüfen, inwiefern ein differentes Abschneiden durch Unterschiede in den Teilnahmebedingungen bedingt sein könnte.

6.1 Mittlere Leseleistungen im internationalen Vergleich

In diesem Abschnitt werden zunächst für die Gesamtskala Lesen die mittleren Testleistungen im internationalen Vergleich betrachtet. In Abbildung 3.6 sind für alle Teilnehmer an IGLU 2016 die jeweiligen Mittelwerte (M), ihre Standardabweichungen (SD) sowie die Standardfehler (SE) dieser beiden Werte angegeben. Darüber hinaus vermitteln die dargestellten Perzentilbänder (siehe Kapitel 2, Abschnitt 12 in diesem Band) einen visuellen Eindruck von der Streuung von Leistungsergebnissen innerhalb der teilnehmenden Staaten und Regionen. Des Weiteren sind der Mittelwert für alle Teilnehmerstaaten (interna-

tionaler Mittelwert) sowie die Mittelwerte für die VG OECD und die VG EU aufgeführt. Teilnehmer mit einem besonderen Teilnahmestatus (Benchmark-Teilnehmer) sind in der Abbildung 3.6 separat dargestellt. Die Mittelwerte der Teilnehmerstaaten sind innerhalb der einzelnen Gruppen jeweils in einer Rangreihe angeordnet. Bei der Interpretation der Ergebnisse ist zu beachten, dass die Anordnung der Staaten nach Rangplätzen nicht notwendigerweise Schlüsse auf Unterschiede in der Lesekompetenz zwischen den Staaten zulässt. Die deskriptiven, numerischen Differenzen können zufällige Unterschiede darstellen, die statistisch nicht bedeutsam sind. Deshalb sind in der Abbildung 3.6 die Perzentilbänder jener Teilnehmer durch eine schwarze Umrandung gekennzeichnet, deren Leistungsmittelwerte sich trotz deskriptiver Unterschiede nicht signifikant vom Kompetenzmittelwert für Deutschland (537) unterscheiden.

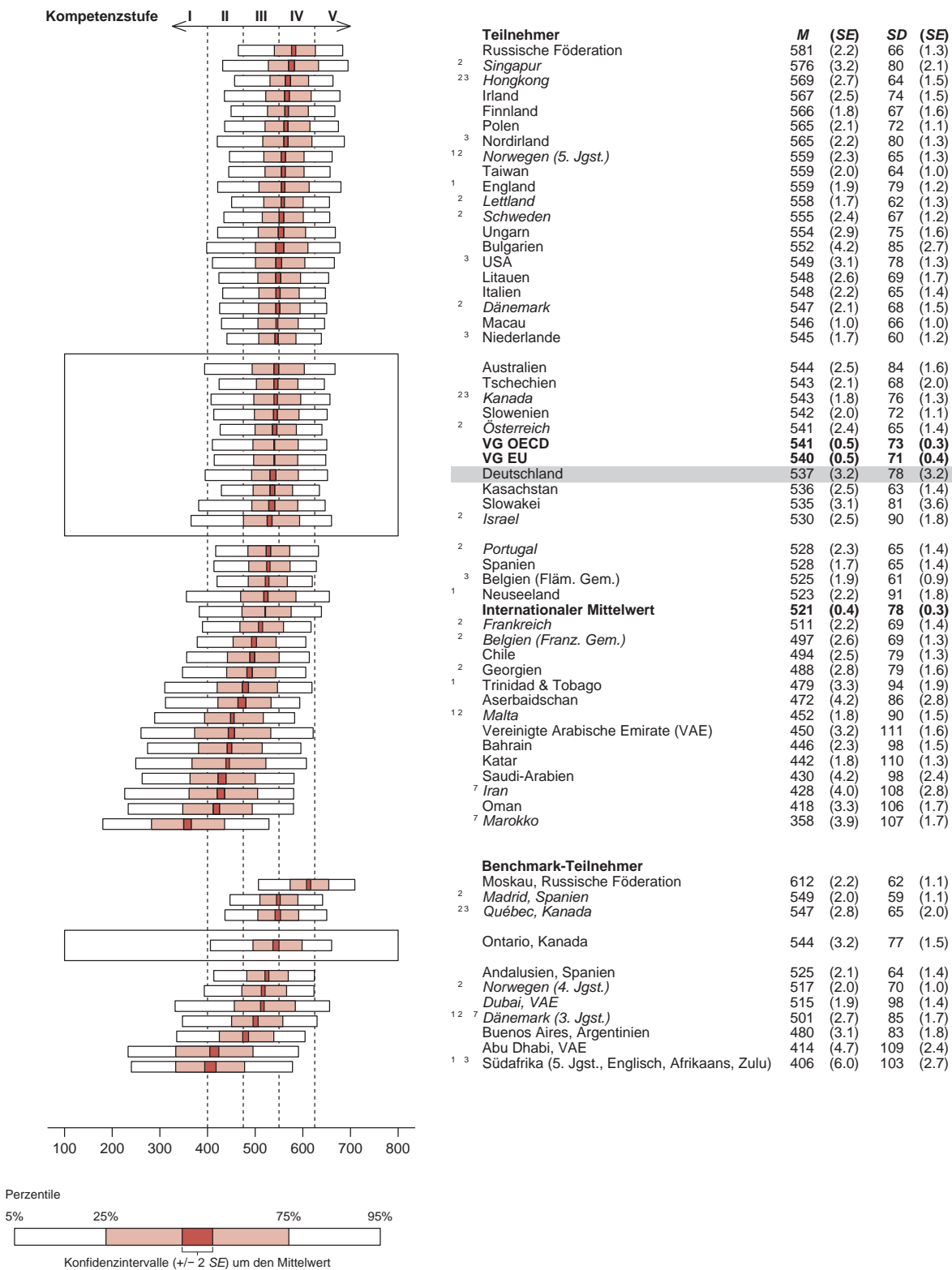
Leistungsmittelwerte in IGLU 2016

Wie Abbildung 3.6 zu entnehmen ist, reicht das Spektrum der mittleren Leistungswerte von 581 Punkten für die Russische Föderation bis zu 358 Punkten in Marokko. Der internationale Mittelwert beträgt 521 Punkte. Signifikant über dem internationalen Skalenmittelwert von 500, der in IGLU 2001 festgelegt wurde, liegen 35 Teilnehmerstaaten, darunter auch Deutschland. Mit einem Leistungsmittelwert von 537 Punkten ergibt sich für Deutschland ein Rangplatz im Mittelfeld und damit auf dem Niveau des Mittelwerts aller Teilnehmerstaaten, die Mitglieder der EU (VG EU) (540) oder der OECD sind (VG OECD) (541) sind, sowie ferner der folgenden Teilnehmerstaaten und -regionen: Australien (544), Ontario (544), Tschechien (543), Kanada (543), Slowenien (542), Österreich (541), Kasachstan (536), die Slowakei (535) und Israel (530).

Am oberen Ende der Leistungsskala lassen sich die Russische Föderation (581) und Singapur (576) verorten. Der Unterschied zu Deutschland beträgt mit 43 beziehungsweise 39 Punkten etwa ein Lernjahr (Hornberg et al., 2007). Insgesamt fallen für 20 teilnehmende Staaten und Regionen und drei Benchmark-Teilnehmer die mittleren Schülerleistungen signifikant besser aus als in Deutschland. Dies sind, neben der Russischen Föderation (581) und Singapur (576), Hongkong (569), Irland (567), Finnland (566), Polen (565), Nordirland (565), Norwegen (5. Jgst.) (559), Taiwan (559), England (559), Lettland (558), Schweden (555), Ungarn (554), Bulgarien (552), die USA (549), Litauen (548), Italien (548), Dänemark (547), Macau (546) und die Niederlande (545). Für die Benchmark-Teilnehmer lassen sich in den Städten Moskau (612) und Madrid (549) sowie in der kanadischen Provinz Québec (547) signifikant höhere mittlere Testleistungen feststellen.

Für 18 Teilnehmerstaaten und sieben Benchmark-Teilnehmer sind die mittleren Schülerleistungen signifikant geringer als in Deutschland. Hierzu zählen die EU-Mitglieder Portugal (528), Spanien (528), Belgien mit seinen flämischen (525) und französischen (497) Gemeinschaften, Frankreich (511) und Malta (452) sowie die OECD-Mitglieder Neuseeland (523) und Chile (494).

Abbildung 3.6: Testleistungen der Schülerinnen und Schüler im internationalen Vergleich – Gesamtskala Lesen



□ Nicht statistisch signifikant vom deutschen Mittelwert abweichende Staaten ($p > .05$).
 Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 7 = Teilnahme an IGLU 2016 und PIRLS Literacy (Iran, Marokko) bzw. ausschließlich an PIRLS Literacy (Dänemark, 3. Jgst.).
 Die Kennwerte für Iran und Marokko werden in Anlehnung an die internationale Berichterstattung als Mittelwert der beiden Studien dargestellt.

6.2 Mittlere Leseleistungen auf der Gesamtskala: Lesen im Trend

2001 bis 2016: Leseleistungen auf der Gesamtskala Lesen im Trend im internationalen Vergleich

Deutschland beteiligt sich wie 19 andere Staaten beziehungsweise Benchmark-Teilnehmer im Jahr 2016 bereits zum vierten Mal an IGLU. Durch den Vergleich der Ergebnisse der vier Messzeitpunkte lassen sich Trends der Leseleistungen von Schülerinnen und Schülern am Ende der Grundschulzeit betrachten.

In Abbildung 3.7 sind für 27 Teilnehmerstaaten und -regionen und Benchmark-Teilnehmer die Ergebnisse aufgeführt, die an mindestens drei Erhebungen teilgenommen haben und zur Vergleichsgruppe EU und/oder zur Vergleichsgruppe OECD gehören und/oder die auf der Gesamtskala Lesen nicht signifikant geringere Leistungen erzielt haben als Deutschland. Darüber hinaus werden für die Flämische Gemeinschaft in Belgien Ergebnisse berichtet, obwohl hier nur Daten aus zwei Zyklen vorliegen. Beim Vergleich mit Publikationen zu IGLU 2001 und 2006 (vgl. Bos, Valtin, Hornberg et al., 2007; Bos et al., 2003; Mullis et al., 2003; Mullis, Martin, Kennedy & Foy, 2007) ist zu beachten, dass sich Abweichungen von den hier dargestellten Ergebnissen ergeben können, da für eine optimierte Trendberichterstattung 2012 eine Neuskalierung der Werte erfolgt ist und Methoden der Berechnung von Standardfehlern verändert wurden (siehe Kapitel 2, Abschnitt 9 in diesem Band).

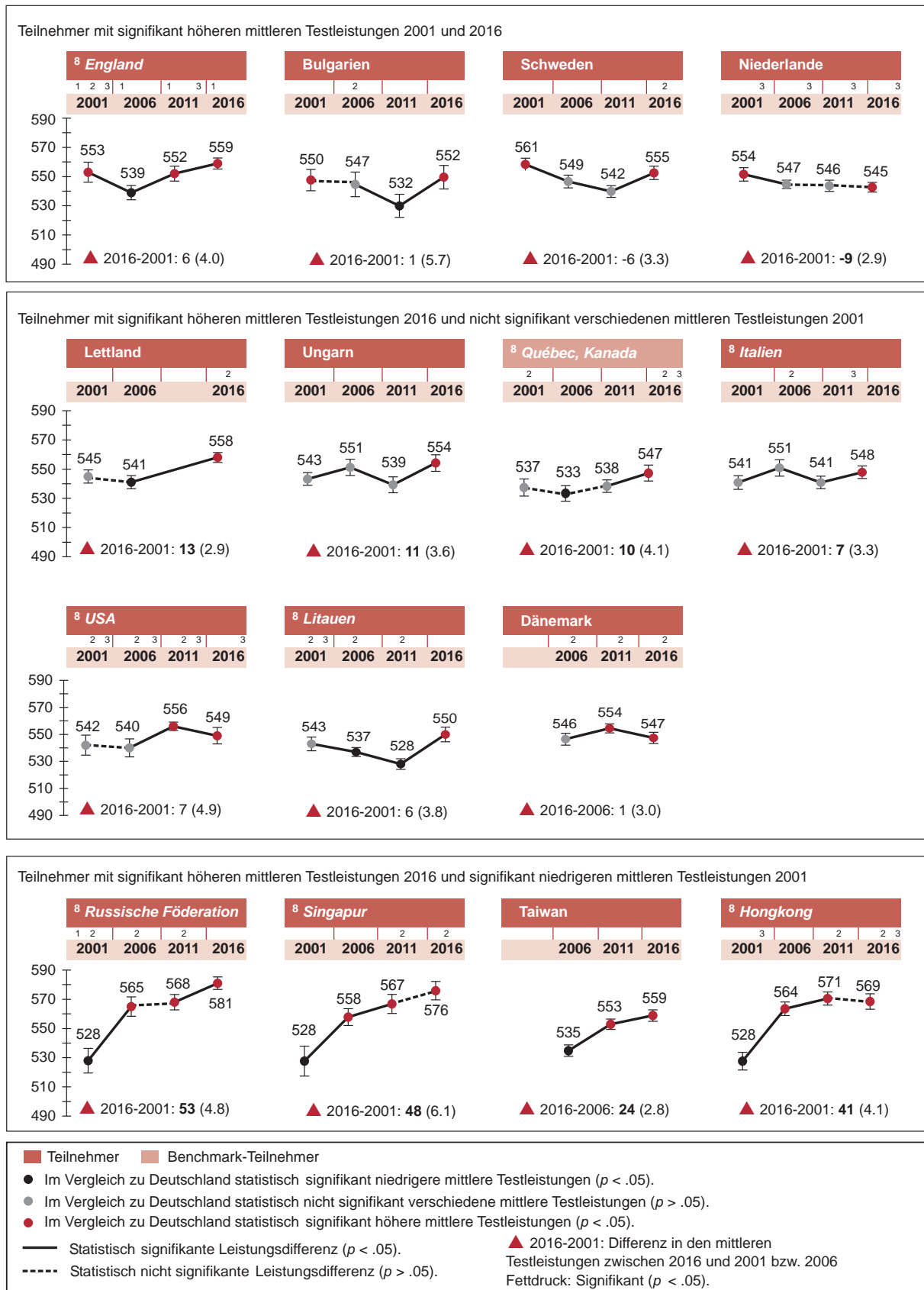
In Abbildung 3.7 sind für alle Vergleichsländer die jeweiligen Mittelwerte (M) für die einzelnen Erhebungszyklen angegeben (vgl. Mullis et al., 2017, Exhibit 1.4). Länder auf der linken Seite der Darstellung (siehe Seite 116) erzielen 2016 signifikant bessere Schülerleistungen; auf der rechten Seite sind Teilnehmerstaaten und -regionen und Benchmark-Teilnehmer angeordnet, für die sich 2016 vergleichbare oder signifikant schlechtere Testleistungen beobachten lassen. Darüber hinaus sind die Teilnehmer hinsichtlich ihrer Leistungsentwicklungen und der daraus resultierenden Position im Vergleich mit Deutschland gruppiert. Sie zeigt sich in der jeweiligen Differenz zu den Leistungsmittelwerten in Deutschland und ist durch unterschiedliche Färbungen der Punkte gekennzeichnet (für die Kennzahlen vgl. Mullis et al., 2017, Exhibit 1.2). So signalisiert beispielsweise eine rote Färbung des Punktes, dass sich zum jeweiligen Messzeitpunkt eine statistisch signifikant höhere mittlere Testleistung als in Deutschland beobachten ließ.

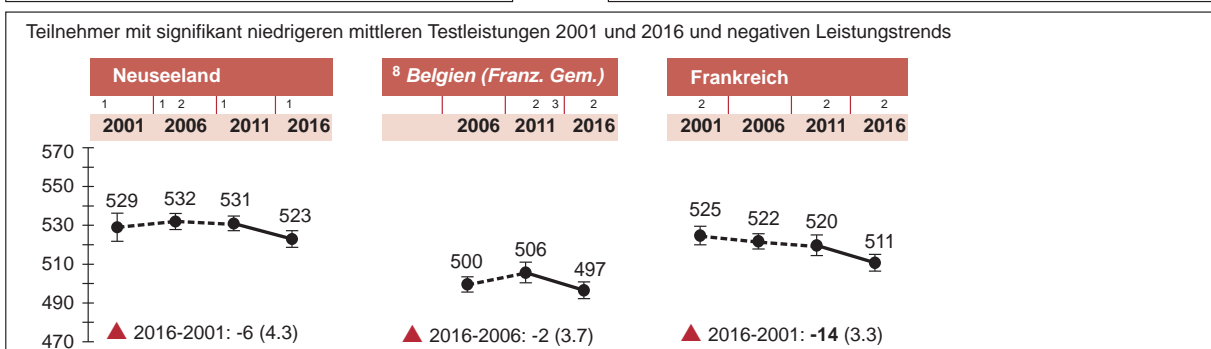
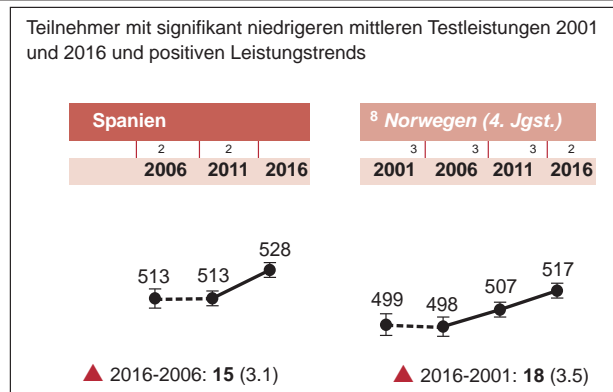
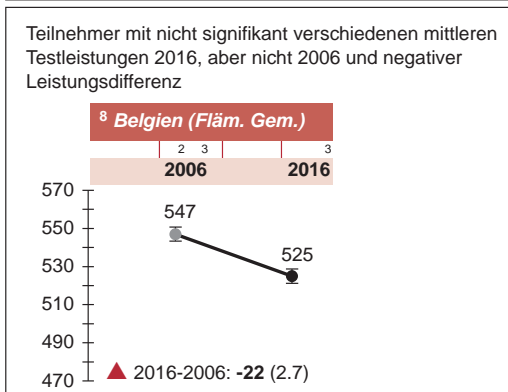
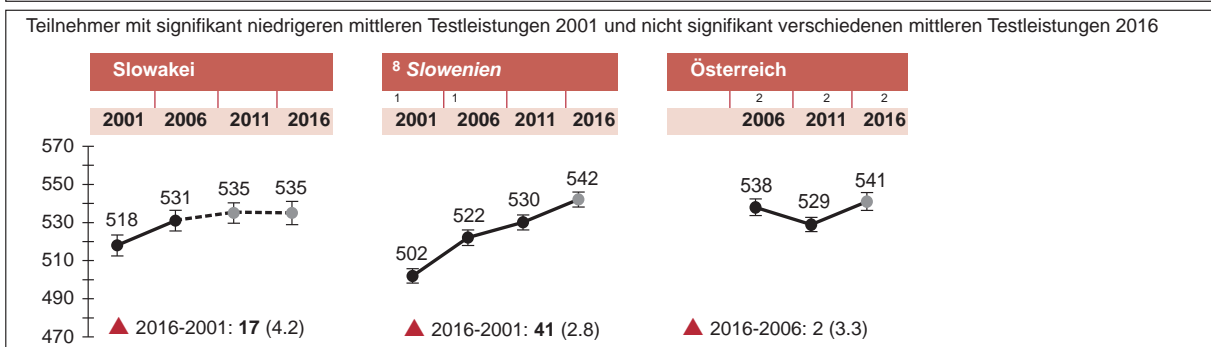
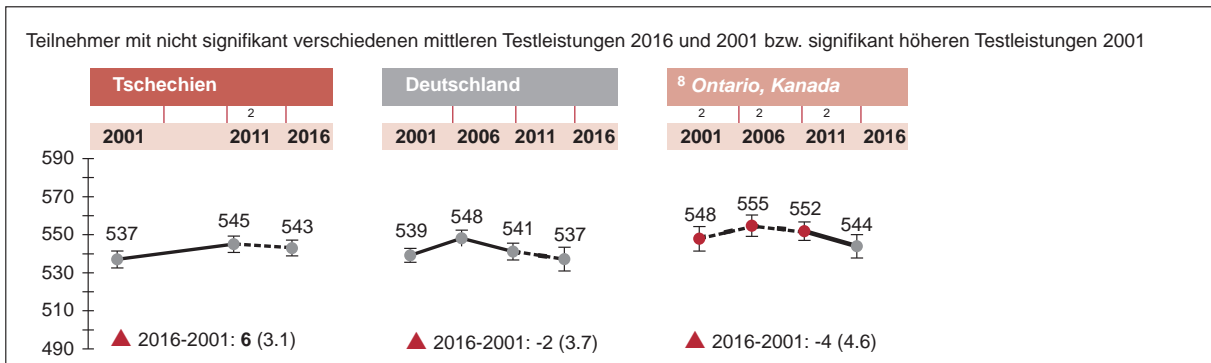
Die Färbung der Trendlinien, mit denen die Punkte verbunden sind, verdeutlicht für den jeweiligen Teilnehmerstaat, inwieweit sich zwischen den Messzeitpunkten eine signifikante Veränderung der Leistungsmittelwerte feststellen lässt (für die Kennzahlen vgl. Mullis et al., 2017, Exhibit 1.4). Eine durchgezogene Linie verweist auf eine statistisch signifikante Leistungsdifferenz. Eine gepunktete Linie steht für statistisch nicht signifikante Unterschiede zwischen Erhebungszyklen.

Als Kennwerte für den Leistungstrend zwischen 2016 und 2001 (bzw. 2006 für Staaten mit Erstteilnahme in 2006) sind der Differenzwert in den mittleren Schülerleistungen ($\blacktriangle_{2016-2001}$) und der dazugehörige Standardfehler angegeben. Signifikante Leistungstrends sind durch Fettdruck des Differenzwerts gekennzeichnet.

Für Deutschland wird aus der Darstellung ersichtlich, dass sich im Vergleich zu 2001 keine signifikanten Unterschiede in der mittleren Leseleistung von Viertklässlerinnen und Viertklässlern feststellen lassen. Der Differenzwert von -2 Punkten zwischen 2001 und 2016 ist statistisch nicht signifikant. Die signifi-

Abbildung 3.7: Relative Positionierung zu Deutschland und Vergleich der Testleistungen zwischen IGLU 2001, 2006, 2011 und 2016 – Gesamtskala Lesen





Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den hier betrachteten Studienzyklen.
 Die Ergebnisse von Israel und Polen werden aufgrund der nicht gegebenen Vergleichbarkeit zwischen den hier betrachteten Studienzyklen nicht berichtet.

kante Verbesserung der Leistung (+9), die sich zwischen 2001 und 2006 gezeigt hatte, wurde 2011 durch eine signifikante Leistungsverschlechterung (-7) aufgehoben. Im Vergleich mit 2006 fallen die Leseleistungen von Grundschulkindern in Deutschland 2016 signifikant geringer aus, nämlich um 10 Leistungspunkte (vgl. Mullis et al., 2017, Exhibit 1.4), im Vergleich mit 2011 (541) um 4 Punkte, wobei diese Differenz nicht signifikant ist. Dieser Befund entspricht dem, den das IQB kürzlich ermittelte. Unter dem Titel „IQB-Bildungstrend 2016“ legte es einen Bericht vor, in dem die Leistungen von Viertklässlerinnen und Viertklässlern in allen Ländern der Bundesrepublik Deutschland in den Bereichen Lesen, Zuhören und Orthographie im Jahr 2016 mit denen im Jahr 2011 verglichen wurden. Für das Lesen zeigte sich, dass die mittlere Testleistung im Jahr 2011 (500 Punkte) im Vergleich zum Jahr 2016 um 7 Punkte unterschritten wurde. Diese Differenz ist, anders als die bei IGLU, signifikant (vgl. Wittig & Weirich, 2017, S. 158f.). Die Modelle, die der IGLU- und der IQB-Testung der Lesekompetenz zugrunde liegen, haben vieles gemeinsam. Es gibt aber auch gewichtige Differenzen (vgl. zu Details Pietsch, Böhme, Robitzsch & Stubbe, 2009; Böhme et al., 2014). Insofern sind die geringen Unterschiede, die sich beim Vergleich der Ergebnisse zu den Leseleistungen in den Jahren 2016 und 2011 bei IGLU (-4) und beim IQB (-7) ergeben, bemerkenswert.

Deutschland ist einer von 11 der hier dargestellten 27 Trend-Teilnehmer (neben England, Bulgarien, Schweden, USA, Litauen, Dänemark, Ontario, Österreich, Neuseeland und der Französischen Gemeinschaft in Belgien), für die sich im Vergleich von 2001 und 2016 keine signifikanten *Veränderungen in den Testleistungen* beobachten lassen.

Hingegen können 11 Staaten und Regionen positive Leistungstrends von 2001 zu 2016 verzeichnen. Die größten Leistungszuwächse sind für die Russische Föderation (+53), Singapur (+48), Hongkong (+41), Slowenien (+41), Norwegen (4. Jgst., +18) und die Slowakei (+17) zu konstatieren.

Frankreich (-14) und die Niederlande (-9) zählen zu den Staaten, für die sich von 2001 zu 2016 negative Leistungstrends feststellen lassen. Auch für die Flämische Gemeinschaft in Belgien, die sich nur 2006 und 2016 beteiligte, ist eine solche Entwicklung beobachtbar (-22). Im Vergleich Deutschlands mit anderen europäischen Ländern ist damit festzuhalten, dass sich in drei Staaten beziehungsweise Regionen die Leistungen verschlechtert haben. Keine Veränderungen zeigten sich in Deutschland und der Französischen Gemeinschaft in Belgien, Dänemark, Österreich (jeweils seit 2006), sowie in Bulgarien, England und Schweden.

Betrachtet man nur die Länder und Regionen, die an allen vier Erhebungszyklen teilgenommen haben, dann zeigt sich im Hinblick auf die Position von Deutschland Folgendes: 2001 ließen sich für lediglich vier Teilnehmerstaaten (England, Schweden, Bulgarien und die Niederlande) sowie Ontario im Vergleich zu Deutschland signifikant höhere Leistungsmittelwerte beobachten. Im Jahr 2016 sind es 14 Teilnehmer und ein Benchmark-Teilnehmer. Lettland, Ungarn, Litauen, die USA, Italien, Dänemark und Québec konnten sich seit 2001, als die dort erreichten Leistungen nicht von denen in Deutschland verschieden waren, positiv vom Leistungsmittelwert für Deutschland absetzen. Während sich in der Russischen Föderation, in Singapur, Taiwan und Hongkong im Jahr 2001 noch signifikant schlechtere Schülerleistungen als in Deutschland beobachten ließen, werden dort seit 2006 durchgängig signifikant bessere Leistungen erzielt.

Für Tschechien und Ontario ist sowohl 2001 als auch 2016 kein signifikanter Unterschied zu den Leistungsmittelwerten für Deutschland festzustellen.

Die Slowakei (+17) und Slowenien (+41) zeigen deutlich positive Leistungstrends. Sie schnitten im Jahr 2001 schlechter ab als Deutschland, im Jahr 2016 unterscheiden sich die Mittelwerte nicht mehr signifikant.

Bei den Teilnehmern, für die sich im Vergleich zu Deutschland 2016 signifikant geringere Leistungsmittelwerte feststellen lassen, sind verschiedene Entwicklungen zu beobachten: Die Flämische Gemeinschaft in Belgien schnitt im Jahr 2006 wie Deutschland, im Jahr 2016 aber signifikant schlechter ab. Für Frankreich führt der negative Leistungstrend (-14) zwischen 2001 und 2016 zu einer Vergrößerung des Abstandes zu Deutschland. Für Neuseeland und Belgien (Franz. Gemeinschaft) ergeben sich aufgrund konstanter Leistungsstände keine Veränderungen der Positionen, bezogen auf Deutschland. Das gilt auch für Spanien und Norwegen (4. Jgst.), obwohl dort positive Leistungstrends zu verzeichnen sind.

2001 bis 2016: Leseleistungen auf der Gesamtskala Lesen im Trend in Deutschland

Für eine Interpretation der Ergebnisse von IGLU 2016 ist es wichtig, die im Vergleich mit den drei vorangegangenen Studienzyklen veränderte Schülerpopulation (siehe Kapitel 2, Abschnitt 7 in diesem Band) zu berücksichtigen. Komplexe Trendanalysen bieten die Möglichkeit, Veränderungen der Leseleistungen zwischen 2001, 2006, 2011 und 2016 zu modellieren, die sich ergeben hätten, wenn die Zusammensetzung der Schülerschaft hinsichtlich verschiedener soziodemographischer Merkmale gleich geblieben wäre.

Entsprechende Analysen können Hinweise darauf liefern, inwieweit der bei IGLU 2016 im Vergleich zu 2001 beobachtbare gestiegene Anteil von Kindern mit Migrationshintergrund sowie niedrigem sozioökonomischen Status, und damit veränderte Rahmenbedingungen schulischen Lernens, als Erklärungen für die veränderten Leistungsmittelwerte herangezogen werden sollten. Vor dem Hintergrund der Interpretationen der Ergebnisse des IQB-Bildungstrends (vgl. Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017), aber auch der Ergebnisse solcher Analysen für TIMSS 2015 (vgl. Kasper et al., 2016) erweist sich eine solche Annahme als plausibel: So konnte beispielsweise für die mittleren Schülerleistungen von Viertklässlerinnen und Viertklässlern in Deutschland in Mathematik und den Naturwissenschaften bei TIMSS 2015 ebenfalls keine signifikante Veränderung zwischen 2007, 2011 und 2015 festgestellt werden. Bei Modellierung der Leistungstrends unter Berücksichtigung von Veränderungen der Struktur der Schülerschaft zeigten sich jedoch positive Leistungstrends in Form eines Zuwachses von 12 Punkten (*Mathematik*) beziehungsweise 22 Punkten (*Naturwissenschaften*), die sich vorsichtig als Hinweis auf die Verbesserung der Bildungsqualität unter herausfordernden Bedingungen interpretieren lassen (ebd.).

Tabelle 3.7 zeigt, wie sich die Leistungstrends von Viertklässlerinnen und Viertklässlern in Deutschland in IGLU unter Kontrolle von Schülercharakteristika darstellen. Die Modelle 1 und 2 dienen als Referenzmodelle für die Modelle 3 und 4. Die für Modell 1 dargestellten Koeffizienten entsprechen den bereits in Abbildung 3.7 für Deutschland illustrierten Differenzen der mittleren Leseleistungen im Vergleich zu 2016: Die mittlere Schülerleistung beträgt im Jahr 2016 537 Punkte. In allen drei vorangegangenen Erhebungszyklen fielen die mittleren Schülerleistungen besser aus, allerdings ist nur der Unterschied zu 2006 (10,3 Punkte) signifikant. Für Modell 2 wurden die Leistungstrends unter Berücksichtigung der Mehrebenenstruktur der Daten ge-

Tabelle 3.7: Veränderung der Schülerleistungen zwischen IGLU 2001 und 2016 unter Kontrolle von Schülermerkmalen – Gesamtskala Lesen

	Lineare Regression		Mehrebenenregressionen					
	b	(SE)	Modell 1		Modell 2		Modell 3	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Konstante	537.3	(3.2)**	533.5	(4.4)**	551.8	(4.9)**	550.8	(5.7)**
Testzeitpunkt								
IGLU 2001	1.8	(3.4) ^{ns}	5.7	(5.0) ^{ns}	-0.6	(4.7) ^{ns}	1.8	(5.9) ^{ns}
IGLU 2006	10.3	(3.6)**	9.4	(5.6)*	4.9	(4.4) ^{ns}	7.5	(5.7) ^{ns}
IGLU 2011	3.9	(4.0) ^{ns}	5.4	(5.3) ^{ns}	-4.1	(4.9) ^{ns}	-5.9	(5.9) ^{ns}
Merkmale								
Geschlecht ^A					-5.7	(2.4)**	-5.7	(2.4)**
Familie mit hohem sozioökonomischen Status ^B					22.4	(2.5)**	26.9	(6.5)**
Migrationshintergrund (ein Elternteil) ^C					-12.0	(3.7)**	-13.7	(9.2) ^{ns}
Migrationshintergrund (beide Elternteile) ^D					-27.2	(3.8)**	-27.6	(9.1)**
Interaktionen								
Familie mit hohem sozioökonomischen Status ^B * IGLU 2001							-6.0	(7.8) ^{ns}
Familie mit hohem sozioökonomischen Status ^B * IGLU 2006							-9.8	(8.3) ^{ns}
Familie mit hohem sozioökonomischen Status ^B * IGLU 2011							0.4	(7.4) ^{ns}
Migrationshintergrund (ein Elternteil) ^C * IGLU 2001							-0.1	(12.2) ^{ns}
Migrationshintergrund (beide Elternteile) ^D * IGLU 2001							3.1	(11.4) ^{ns}
Migrationshintergrund (ein Elternteil) ^C * IGLU 2006							3.2	(12.8) ^{ns}
Migrationshintergrund (beide Elternteile) ^D * IGLU 2006							-6.7	(11.3) ^{ns}
Migrationshintergrund (ein Elternteil) ^C * IGLU 2011							0.2	(11.8) ^{ns}
Migrationshintergrund (beide Elternteile) ^D * IGLU 2011							8.9	(11.8) ^{ns}
Kovariaten								
Kognitive Fähigkeiten ^E					17.7	(1.0)**	17.7	(0.9)**
Erklärter Anteil der Gesamtvarianz	0.3							
Erklärter Anteil auf Individualebene (70.3%)			25.0		44.0		44.1	

b = Regressionsgewichte (unstandardisiert).

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

A = Geschlecht (0 = Mädchen; 1 = Jungen), siehe Kapitel 5 in diesem Band.

B = Sozioökonomischer Status nach Angabe der Eltern (0 = Manuell Tätige, Angestellte und kleinere Unternehmer;

1 = Akademiker, Techniker und Führungskräfte), siehe Kapitel 6 in diesem Band.

C = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren;

1 = Ein Elternteil im Ausland geboren), siehe Kapitel 7 in diesem Band.

D = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren;

1 = Beide Elternteile im Ausland geboren), siehe Kapitel 7 in diesem Band.

E = Skalierter Index der kognitiven Fähigkeiten (Messinvarianz über die Zyklen angenommen), siehe Kapitel 2 in diesem Band.

Das Gesamtmodell zu dem hier dargestellten Ausschnitt ist im Anhang B dargestellt, siehe Modelle 1, 12 und 13. Die lineare Regression ist nicht Teil des Gesamtmodells.

schätzt. Damit wird berücksichtigt, dass die in IGLU getesteten Schülerinnen und Schüler im Klassenverband unterrichtet werden und insofern mehr gemeinsame Charakteristika (z.B. dieselbe Lehrkraft, weitgehend identische Lerngelegenheiten) haben als Kinder aus verschiedenen Klassen. Als ein Ergebnis dieser Analyse ist zunächst festzuhalten, dass etwa 30 Prozent aller Unterschiede der beobachteten Schülerleistungen auf Unterschiede zwischen den Schulen beziehungsweise Klassen und etwa 70 Prozent auf individuelle Unterschiede zwischen den einzelnen Schülerinnen und Schülern innerhalb der Klassen zurückzuführen sind. Diese Varianzanteile entsprechen Werten aus vergleichbaren Studien (vgl. Kasper, Wendt, Bos & Köller, 2016). Ein durchaus beachtlicher Anteil der Differenzen bei den Leseleistungen ist also auf Unterschiede zwischen den Schulen beziehungsweise Klassen zurückzuführen.

Unter Berücksichtigung der Mehrebenenstruktur zeigt sich in Modell 2 in Bezug auf die Leistungstrends ein mit Modell 1 vergleichbares Ergebnis, wobei die signifikanten Unterschiede in den mittleren Schülerleistungen zwischen 2006 und 2016 nominell mit 9.4 Punkten etwas geringer ausfallen. In Modell 3 werden die Leistungstrends unter Berücksichtigung zentraler Schülercharakteristika modelliert. Die Ergebnisse für die einzelnen Schülergruppen werden ausführlich in den Kapiteln 5, 6 und 7 in diesem Band diskutiert und daher an dieser Stelle nicht ausgeführt. Als Ergebnis ist zunächst festzuhalten, dass mithilfe der vier hier genutzten Indikatoren etwa 44 Prozent aller Leistungsunterschiede zwischen den Schülerinnen und Schülern erklärt werden können. In Bezug auf die Unterschiede zwischen den Studienzyklen ist festzustellen, dass sich der höhere Leistungsmittelwert für 2006 nicht mehr als signifikant nachweisen lässt. In Modell 4 werden durch Aufnahme von sogenannten Interaktionen Veränderungen der Anteile von Kindern mit Migrationshintergrund und der verschiedenen sozio-ökonomischen Statusgruppen ergänzend berücksichtigt. Als Ergebnis ist festzuhalten: Auch unter Berücksichtigung der Veränderung relevanter Merkmale der Schülerschaft lässt sich über die IGLU-Erhebungszyklen hinweg – anders als bei TIMSS – kein signifikant positiver Leistungstrend beobachten. Die Leistungen bleiben stabil. Die Trendanalysen liefern insofern keine Hinweise auf eine verbesserte Qualität schulischer Leseförderung.

6.3 Leistungsheterogenität im internationalen Vergleich 2016

Bei der Einschätzung der Qualität der Lesekompetenz von Schülerinnen und Schülern ist nicht nur der Mittelwert der Leistungen von Interesse, sondern auch die Streuung der Testwerte. Diese gibt darüber Auskunft, wie groß die Differenz zwischen Personen mit guter und mit schwacher Leistung ist. Als ein Kriterium für gute Qualität oder Erfolg von Bildungssystemen kann ein hohes Leistungsniveau bei gleichzeitig geringer Streuung der Leistungen angesehen werden.

In der Tabelle 3.8 ist die Leistungsstreuung durch Perzentile illustriert, angegeben ist das 5., 25., 75. und 95. Perzentil (siehe Kapitel 2, Abschnitt 11 in diesem Band). Die Perzentilbänder vom 5. bis zum 95. Perzentil illustrieren, wie in Abbildung 3.6 dargestellt, die Streuung der Lesewerte für die mittleren 90 Prozent der Schülerinnen und Schüler, das heißt diejenigen, die nicht zu den jeweils fünf Prozent mit den niedrigsten und höchsten Leistungswerten gehören. Der Leistungswert etwa, der dem 75. Perzentil entspricht, sagt aus, dass 75 Prozent einen niedrigeren oder gleich großen und 25 Prozent einen höheren

Tabelle 3.8: Lesekompetenzen nach Perzentilen im Vergleich: IGLU 2001, 2006, 2011 und 2016 – Gesamtskala Lesen

					IGLU 2016						IGLU 2011														
					Perzentile						Perzentile														
2001	2006	2011	2016	Teilnehmer ^A	SD	(SE)	5	(SE)	25	(SE)	75	(SE)	95	(SE)	SD	(SE)	5	(SE)	25	(SE)	75	(SE)	95	(SE)	
1 2	2	2		⁸ Russische Föderation	66	(1.3)	465	(5.2)	540	(2.8)	626	(2.7)	684	(3.3)	66	(1.7)	455	(5.2)	526	(2.9)	+ 614	(2.4)	+ 672	(2.9)	+
		2	2	⁸ Singapur	80	(2.1)	432	(7.8)	528	(4.4)	633	(3.5)	695	(3.9)	80	(1.8)	421	(7.0)	519	(4.6)	+ 623	(3.9)	687	(4.4)	+
	3		2 3	⁸ Hongkong	64	(1.5)	457	(9.3)	531	(2.6)	612	(3.0)	663	(4.2)	61	(1.3)	460	(5.0)	534	(3.8)	612	(1.6)	662	(2.5)	
				Taiwan	64	(1.0)	445	(5.1)	521	(2.9)	603	(2.3)	657	(2.7)	67	(1.2)	434	(4.3)	512	(3.4)	599	(1.6)	653	(6.0)	+
1 2 3	1	1	3 1	⁸ England	79	(1.2)	421	(6.4)	508	(3.1)	613	(2.2)	680	(3.3)	82	(1.4)	404	(8.7)	+ 500	(4.0)	609	(2.4)	678	(4.7)	
				Lettland	62	(1.3)	451	(5.0)	518	(2.4)	601	(2.3)	656	(2.3)	-		-		-		-		-		
				Schweden	67	(1.2)	434	(5.7)	515	(3.1)	601	(3.2)	656	(3.5)	65	(1.0)	426	(3.5)	502	(3.2)	+ 585	(2.4)	+ 643	(3.3)	
				Ungarn	75	(1.6)	421	(5.5)	506	(4.6)	606	(3.6)	668	(4.0)	78	(2.1)	397	(10.2)	+ 493	(3.1)	+ 594	(2.8)	+ 656	(3.2)	
	2			Bulgarien	85	(2.7)	398	(11.0)	501	(6.5)	611	(3.5)	678	(4.1)	82	(2.6)	382	(9.0)	482	(4.8)	+ 589	(3.1)	+ 652	(3.7)	
2 3	2 3	2 3	3	⁸ USA	78	(1.3)	410	(5.9)	501	(4.1)	604	(3.3)	666	(4.4)	73	(1.0)	+ 428	(3.5)	- 510	(2.1)	607	(1.2)	671	(3.0)	
2 3	2	2		⁸ Litauen	69	(1.9)	425	(8.5)	507	(3.3)	596	(3.2)	656	(4.5)	66	(1.2)	412	(3.4)	487	(3.3)	574	(1.5)	630	(2.5)	
	2	3		⁸ Italien	65	(1.4)	432	(6.2)	508	(3.2)	592	(2.4)	647	(3.1)	66	(1.3)	427	(6.8)	500	(3.2)	+ 586	(2.1)	645	(2.1)	-
	2	2	2	Dänemark	68	(1.5)	425	(6.3)	507	(3.0)	594	(2.5)	650	(3.9)	64	(0.9)	+ 438	(3.8)	514	(2.0)	599	(1.6)	652	(3.7)	
3	3	3	3	Niederlande	60	(1.2)	441	(5.4)	508	(2.5)	586	(2.1)	639	(2.9)	54	(0.9)	+ 454	(3.1)	- 510	(3.5)	583	(1.8)	+ 631	(2.4)	+
			2	Tschechien	68	(2.0)	424	(7.0)	503	(2.8)	590	(2.3)	645	(3.6)	61	(1.4)	434	(5.4)	509	(2.2)	587	(2.5)	639	(2.7)	
1	1			⁸ Slowenien	72	(1.1)	413	(6.3)	498	(3.8)	592	(2.2)	651	(3.6)	70	(0.9)	405	(7.9)	487	(2.7)	+ 579	(2.2)	+ 637	(2.5)	+
	2	2	2	Österreich	65	(1.4)	427	(4.4)	500	(2.9)	586	(2.2)	640	(3.6)	63	(1.0)	418	(3.5)	487	(2.0)	+ 573	(1.6)	+ 626	(3.7)	
				Deutschland	78	(3.2)	395	(11.5)	493	(4.2)	591	(2.8)	652	(4.3)	66	(1.3)	+ 425	(7.3)	- 499	(3.0)	586	(1.8)	646	(4.6)	
				Slowakei	81	(3.6)	381	(15.5)	493	(3.8)	589	(2.7)	647	(2.2)	69	(1.9)	+ 408	(11.1)	495	(2.8)	+ 582	(2.5)	+ 638	(3.3)	+
	2	2		Spanien	65	(1.4)	413	(5.1)	486	(2.2)	573	(1.4)	628	(2.2)	68	(1.2)	393	(5.2)	+ 469	(3.4)	+ 561	(2.5)	+ 618	(3.3)	
	2 3		3	⁸ Belgien (Fläm. Gem.)	61	(0.9)	420	(3.8)	486	(2.6)	567	(2.3)	620	(2.6)	-		-		-		-		-		
1	1 2	1	1	Neuseeland	91	(1.8)	356	(6.8)	469	(3.4)	586	(3.1)	656	(3.0)	88	(1.2)	373	(3.4)	- 474	(3.0)	592	(4.5)	666	(4.6)	
	2		2	Frankreich	69	(1.4)	389	(5.2)	468	(2.8)	559	(2.2)	617	(4.0)	68	(1.3)	401	(5.5)	475	(3.9)	- 568	(1.9)	626	(3.9)	
		2 3	2	⁸ Belgien (Franz. Gem.)	69	(1.3)	378	(4.9)	454	(3.1)	544	(2.5)	606	(2.8)	65	(1.6)	391	(7.9)	- 466	(3.9)	- 551	(2.0)	606	(3.5)	
Benchmark-Teilnehmer																									
2			2 3	⁸ Québec, Kanada	65	(2.0)	437	(7.5)	506	(4.3)	591	(3.2)	650	(5.2)	62	(1.2)	434	(6.7)	498	(2.7)	+ 579	(1.5)	+ 634	(2.8)	+
2	2	2		⁸ Ontario, Kanada	77	(1.5)	406	(7.0)	495	(4.8)	598	(4.0)	660	(5.1)	73	(1.5)	+ 423	(2.3)	- 506	(2.4)	603	(3.0)	663	(3.8)	
3	3	3	2	⁸ Norwegen (4. Jgst.)	70	(1.0)	393	(6.0)	472	(2.8)	566	(2.5)	623	(3.8)	61	(0.9)	+ 398	(4.4)	467	(2.9)	+ 550	(2.5)	+ 601	(3.4)	+

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.
 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.
 A = Zur Auswahl der Trendteilnehmer siehe Kapitel 2, Abschnitt 4.2 in diesem Band.
 + = Prozentwert in 2016 signifikant höher als 2011 bzw. 2006 bzw. 2001 (p < .05).
 - = Prozentwert in 2016 signifikant niedriger als 2011 bzw. 2006 bzw. 2001 (p < .05).

IGLU 2006					IGLU 2001				
Perzentile					Perzentile				
SD (SE)	5 (SE)	25 (SE)	75 (SE)	95 (SE)	SD (SE)	5 (SE)	25 (SE)	75 (SE)	95 (SE)
69 (2.0)	443 (9.5) +	523 (4.4) +	612 (2.9) +	671 (2.7) +	66 (4.2)	412 (12.9) +	488 (5.1) +	574 (4.6) +	627 (4.0) +
77 (1.6)	420 (5.8)	512 (4.9) +	612 (2.8) +	672 (3.2) +	92 (3.7) -	348 (10.6) +	479 (7.2) +	592 (4.6) +	658 (5.4) +
59 (1.2) +	460 (4.7)	527 (2.9)	605 (1.8) +	655 (2.9)	63 (1.7)	415 (6.4) +	491 (5.0) +	571 (4.0) +	622 (3.2) +
64 (1.0)	420 (3.9) +	497 (3.0) +	579 (2.0) +	633 (4.7) +	-	-	-	-	-
87 (1.6) -	383 (8.0) +	486 (4.6) +	598 (2.3) +	673 (5.1)	87 (1.7) -	395 (6.3) +	501 (4.4)	612 (4.5)	685 (5.3)
63 (1.2)	433 (9.2)	501 (4.9) +	585 (3.1) +	639 (3.3) +	62 (1.3)	440 (4.9)	505 (3.3) +	586 (2.4) +	640 (3.4) +
64 (1.3)	437 (3.6)	512 (3.4)	592 (2.3) +	647 (5.5)	66 (1.2)	445 (4.5)	521 (4.7)	605 (1.7)	663 (2.1)
70 (1.8)	427 (6.4)	507 (5.0)	599 (3.5)	658 (2.3)	66 (1.2) +	428 (4.4)	502 (2.4)	589 (2.9) +	643 (3.8) +
83 (2.4)	397 (10.0)	498 (6.5)	604 (3.4)	673 (6.0)	83 (2.8)	400 (11.9)	502 (4.5)	607 (2.1)	671 (3.8)
74 (1.3) +	409 (7.6)	494 (3.5)	592 (3.8) +	653 (7.3)	83 (2.0) -	389 (8.9)	492 (4.7)	601 (4.2)	663 (2.8)
57 (1.1) +	440 (5.1)	500 (2.0)	577 (2.2)	627 (4.5)	64 (1.3) +	433 (4.4)	502 (4.0)	589 (2.3)	642 (3.6)
68 (1.4)	435 (5.3)	507 (3.0)	599 (4.3)	658 (3.3) -	71 (1.4) -	415 (6.5)	496 (3.2) +	590 (3.1)	649 (2.7)
70 (1.2)	418 (4.5)	505 (3.6)	594 (1.8)	649 (2.9)	-	-	-	-	-
53 (0.9) +	457 (3.3) -	513 (1.8)	584 (1.8)	631 (2.1) +	57 (1.2)	458 (4.1) -	517 (3.8) -	593 (2.9) -	645 (3.6)
-	-	-	-	-	65 (1.4)	421 (5.2)	496 (1.9)	582 (3.0) +	634 (4.7)
71 (0.9)	395 (3.7) +	476 (2.2) +	571 (1.7) +	629 (2.6) +	72 (1.5)	373 (6.4) +	456 (2.8) +	551 (2.7) +	611 (3.0) +
64 (1.4)	427 (3.5)	498 (4.1)	582 (2.2)	636 (5.3)	-	-	-	-	-
67 (1.2) +	430 (4.9) -	508 (3.0) -	593 (2.3)	647 (2.4)	67 (1.0) +	419 (3.9) -	497 (3.1)	586 (1.9)	640 (1.9) +
74 (2.1)	394 (6.3)	488 (4.0)	582 (2.6) +	639 (2.9) +	70 (1.7) +	389 (9.7)	477 (2.7) +	566 (1.8) +	623 (3.9) +
71 (1.6) -	390 (4.1) +	468 (3.4) +	561 (1.7) +	622 (3.8)	-	-	-	-	-
56 (1.0) +	451 (3.7) -	512 (1.8) -	585 (1.9) -	636 (4.6) -	-	-	-	-	-
87 (1.3)	374 (3.0) -	478 (2.5) -	592 (2.1)	664 (4.0)	93 (1.9)	360 (4.7)	472 (5.9)	593 (4.5)	668 (5.1) -
67 (1.0)	406 (2.5) -	478 (2.4) -	568 (2.1) -	626 (4.7)	70 (1.6)	403 (5.2)	481 (2.8) -	573 (1.8) -	636 (4.5) -
69 (1.3)	381 (2.7)	455 (3.2)	547 (2.9)	608 (4.1)	-	-	-	-	-
63 (1.1)	422 (6.9)	493 (3.4) +	577 (3.8) +	632 (4.0) +	63 (1.2)	429 (8.0)	496 (3.1)	581 (3.6) +	640 (3.5)
71 (1.3) +	433 (4.7) -	510 (4.0) -	603 (4.3)	666 (4.7)	77 (1.5)	411 (8.0)	500 (4.2)	602 (4.6)	667 (3.9)
67 (1.4) +	378 (3.8) +	457 (5.5) +	544 (2.4) +	598 (3.7) +	81 (1.6) -	351 (5.0) +	450 (4.1) +	556 (2.8) +	620 (6.0)

Wert erzielt haben. Die Leistungswerte, die die Schülerinnen und Schüler der einzelnen Teilnehmerstaaten auf den jeweiligen Perzentilen erzielen, sind für die Teilnehmer, die an IGLU 2001, 2006, 2011 und 2016 teilgenommen haben, in Tabelle 3.8 dargestellt (vgl. auch Mullis et al., 2017, Appendix F.1).

Betrachtet man die in Abbildung 3.6 und Tabelle 3.8 dargestellten Perzentile und Perzentilbänder, fällt auf, dass es erhebliche Unterschiede in der Streuung der Lesekompetenzen zwischen den Teilnehmerstaaten und -regionen gibt.

In Deutschland beträgt der Unterschied zwischen dem 5. und dem 95. Perzentil 257 Punkte. In 20 Staaten und Regionen fallen die Unterschiede signifikant geringer aus, in 13 Staaten signifikant höher. Für die signifikant besseren Teilnehmer reichen die Streuungen von 280 Punkten in Bulgarien bis 198 Punkte in den Niederlanden. In 12 der 20 Staaten, darunter auch die Russische Föderation (219) und Hongkong (206), fällt die Differenz signifikant geringer aus als in Deutschland.

Im europäischen Vergleich erweist sich die Leistungsstreuung in Deutschland als verhältnismäßig groß (VG EU: 234 Punkte). Lediglich in Malta (294) fällt die Streuung signifikant höher aus. 15 der 24 europäischen Teilnehmer erzielen eine deutlich geringere Streuung und damit homogenere Leistungen: die Niederlande mit einem Leistungsunterschied von 198 Punkten, Belgien (Fläm. Gem.) (200), Lettland (205), Österreich (213), Spanien (215), Italien (215), Portugal (216), Finnland (218), Tschechien (221), Schweden (222), Dänemark (225), Frankreich (228), Belgien (Franz. Gem.) (228), Litauen (231) und Polen (239).

Ein ähnliches Bild zeigt sich bei Betrachtung der Standardabweichung (*SD*), die in Deutschland mit 78 Punkten relativ hoch ausfällt. Dieser Wert ist dahingehend zu interpretieren, dass gut zwei Drittel der Viertklässlerinnen und Viertklässler in Deutschland Leistungswerte im Skalenbereich von 459 bis 615 Punkten erzielen.

Im internationalen Vergleich fällt die Leistungsstreuung in 23 Staaten und Regionen signifikant geringer aus. Die Standardabweichung für die Vergleichsgruppe EU liegt bei 71 und die der Vergleichsgruppe OECD bei 73. Eine signifikant geringere Standardabweichung ergibt sich für 17 der 23 Teilnehmerstaaten, die gemeinsam mit Deutschland Mitglieder in der Europäischen Union sind.

2001 bis 2016: Streuungen der Testwerte im Trend

In Tabelle 3.8 sind für die Teilnehmerstaaten und -regionen, die für den Trendvergleich zur Verfügung stehen, neben den Angaben zur Streuung der Testwerte im Jahr 2016, auf die bereits eingegangen wurde, auch die entsprechenden Werte für 2011, 2006 und 2001 dargestellt. So können Unterschiede in der Streuung der Werte zwischen 2001 und 2016 beschrieben werden. Dargestellt sind die Standardabweichungen und das 5., 25., 75. und das 95. Perzentil sowie die dazugehörigen Standardfehler. Sind Unterschiede im Vergleich zu IGLU 2016 signifikant, so sind diese mit einem „+“ beziehungsweise einem „-“ gekennzeichnet. Die Teilnehmer sind in dieser Tabelle nach der Rangfolge auf der Gesamtskala Lesen gereiht (siehe Abbildung 3.6).

Für Deutschland ergeben sich zwischen 2016 und den vorangegangenen Studienzyklen bedeutsame Veränderungen bei der Streuung der Testwerte. Die Standardabweichung (*SD*) lag zwischen 2001 und 2011 bei 66 beziehungsweise 67 Punkten und ist im Jahr 2016 signifikant auf 78 Punkte gestiegen. Auch bei der Betrachtung der Streuung der Mittelwerte für die mittleren 90 Prozent der Viertklässlerinnen und Viertklässler in Deutschland wird die zunehmende Heterogenität in den Schülerleistungen deutlich: 2016 beträgt der Abstand

zwischen dem 5. und dem 95. Perzentil insgesamt 257 Punkte, 2011 liegt er ebenso wie 2001 bei 221 Punkten, 2006 sogar noch geringfügig niedriger bei 217 Punkten. Im Vergleich mit den vorangegangenen Erhebungszyklen sind die Schülerleistungen in Deutschland im Jahr 2016 damit deutlich heterogener.

Betrachtet man das 5. Perzentil, also den Wert, den die 5 Prozent leistungsschwächsten Schülerinnen und Schüler maximal erreichen, so liegt er im Jahr 2016 bei 395, im Jahr 2011 bei 425, im Jahr 2006 bei 430 und im Jahr 2001 bei 419 Punkten. Bezogen auf 2016 sind diese Differenzen jeweils signifikant. Das 95. Perzentil, also der Wert, den die 5 Prozent leistungsstärksten Schülerinnen und Schüler mindestens erreichen, liegt 2016 bei 652, 2011 bei 646, 2006 bei 647 und 2001 bei 640 Punkten. Nur der Unterschied zwischen 2001 und 2016 ist statistisch signifikant. Im Vergleich zu 2001 erzielen in 2016 auch die leistungsstärksten Kinder (95. Perzentil) höhere Testwerte.

Die zunehmende Heterogenität in den Schülerleistungen in Deutschland ist damit sowohl auf geringere Leistungen der leistungsschwächeren Viertklässlerinnen und Viertklässler als auch auf bessere Leistungen der leistungsstärksten Schülerinnen und Schüler zurückzuführen.

Eine von 2001 bis 2016 signifikant zunehmende Heterogenität der Leistungen lässt sich, gemessen an der Standardabweichung, ebenfalls in Ungarn, der Slowakei und Litauen beobachten. Für die Flämische Gemeinschaft in Belgien fällt sie im Jahr 2016 ebenfalls signifikant größer aus als im Jahr 2006. In Belgien erzielen allerdings die Schülerinnen und Schüler 2016 im Vergleich zu 2006 auf allen Leistungsniveaus signifikant schlechtere Leistungsergebnisse. In Ungarn ist die gewachsene Heterogenität durch bessere Leistungen der stärkeren Schülerinnen und Schüler zu erklären.

In Singapur, England, Spanien und Norwegen (4. Jgst.) hingegen fallen die Schülerleistungen im Jahr 2016 homogener aus als in 2001 (bzw. 2006 für Spanien).

In England ist dies auf bessere Leistungen der leistungsschwächeren Schülerinnen und Schüler zurückzuführen. In Singapur, Spanien und Norwegen (4. Jgst.) ist die Reduktion von Leistungsheterogenität mit gestiegenen Leistungen auf allen Perzentilen verbunden und damit besonders erfreulich. Unter den 14 Teilnehmern mit positiven Trends bei den Leistungsmittelwerten sind diese drei Länder damit besonders hervorzuheben.

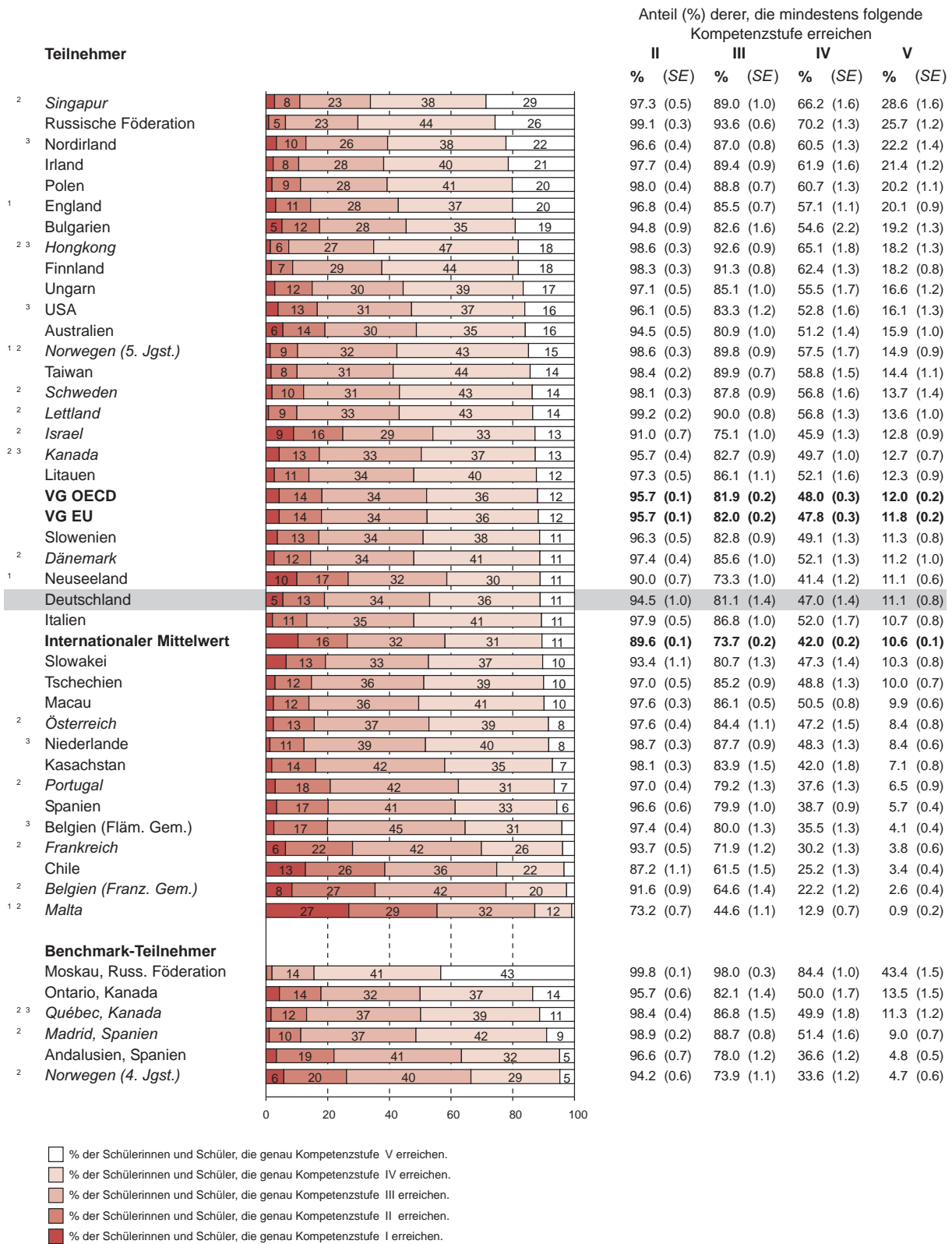
6.4 Lesekompetenzstufen

In Abschnitt 4.1 wurden fünf Stufen der Lesekompetenz vorgestellt. Für die einzelnen Kompetenzstufen können die Anteile der Schülerinnen und Schüler bestimmt werden, die ein Leistungsniveau erreichen, das den jeweiligen Stufen entspricht. Im Folgenden wird zunächst die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen im internationalen Vergleich betrachtet. Anschließend werden die Veränderungen von 2001 bis 2016 dargestellt.

Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen

Im linken Teil der Abbildung 3.8 sind die Anteile derjenigen Schülerinnen und Schüler illustriert, deren Leistungen einer bestimmten Kompetenzstufe zuzuordnen sind. In der Tabelle im rechten Teil der Abbildung sind zusätzlich die kumulierten Anteile angegeben, das heißt die Anteile der Kinder, die jeweils min-

Abbildung 3.8: Prozentuale Verteilung der Schülerinnen und Schüler auf die fünf Kompetenzstufen (Lesen) im internationalen Vergleich



% der Schülerinnen und Schüler, die genau Kompetenzstufe V erreichen.
 % der Schülerinnen und Schüler, die genau Kompetenzstufe IV erreichen.
 % der Schülerinnen und Schüler, die genau Kompetenzstufe III erreichen.
 % der Schülerinnen und Schüler, die genau Kompetenzstufe II erreichen.
 % der Schülerinnen und Schüler, die genau Kompetenzstufe I erreichen.

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

- 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
- 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
- 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

destens eine bestimmte Kompetenzstufe erreicht haben. Die Staaten sind entsprechend den Anteilen der Schülerinnen und Schüler, die der Kompetenzstufe V zugeordnet wurden, in absteigender Reihenfolge angeordnet.

Unter den Teilnehmern ist in Singapur mit 28.6 Prozent der mit Abstand größte Anteil auf der Kompetenzstufe V verortet. Es folgt eine Gruppe von Teilnehmern, bei denen ebenfalls mehr als 20 Prozent dieser Stufe zuzuordnen sind: die Russische Föderation (25.7%), Nordirland (22.2%), Irland (21.4%), Polen (20.2%) und England (20.1%). Für die unteren Kompetenzstufen I und II ergibt sich, dass lediglich in der Russischen Föderation (6.4%), Hongkong (7.4%) und Finnland (8.7%) die Anteile unter 10 Prozent liegen.

Für Deutschland ergibt sich folgendes Bild: Die Leistungen von 11.1 Prozent der Schülerinnen und Schüler sind auf Kompetenzstufe V verortet. Dies entspricht in etwa den Werten der Vergleichsgruppe EU (VG EU: 11.8%) und der Vergleichsgruppe OECD (VG OECD: 12.0%). Auf Kompetenzstufe IV sind 35.9 Prozent und auf Kompetenzstufe III 34.1 Prozent zu verzeichnen. Immerhin 18.9 Prozent der Kinder in Deutschland erreichen nicht die Kompetenzstufe III. Es ist zu erwarten, dass diese Schülergruppe in der Sekundarstufe I in vielen Fächern mit erheblichen Schwierigkeiten konfrontiert sein wird.

2001 bis 2011: Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen im Trend

Tabelle 3.9 zeigt die prozentualen Anteile von Schülerinnen und Schülern auf den fünf Kompetenzstufen für IGLU 2001, 2006, 2011 und 2016. Die Staaten und Regionen sind nach dem Mittelwert der Lesekompetenz in einer Rangreihe angeordnet. Signifikante Veränderungen im Vergleich zu IGLU 2001 sind mit einem „+“ oder einem „-“ gekennzeichnet.

Der Vergleich von IGLU 2001 und 2016 zeigt, dass der Anteil der Schülerinnen und Schüler in Deutschland, die lediglich Leistungen auf dem Niveau der Kompetenzstufe I erreichen, statistisch signifikant von 3.0 auf 5.5 Prozent gestiegen ist. Für die Kompetenzstufe V zeigt sich ebenfalls ein gestiegener Anteil: Lag er im Jahr 2001 bei 8.6 Prozent, so erreichen im Jahr 2016 11.1 Prozent der Kinder die Kompetenzstufe V.

Positive Veränderungen, also größere Anteile von Schülerinnen und Schülern auf den höheren Kompetenzstufen und geringere Anteile auf den niedrigen Kompetenzstufen, zeigen sich im Jahr 2016 im Vergleich zum Jahr 2001 für die Russische Föderation, Singapur, Hongkong, Slowenien und Norwegen (4. Jgst.).

Tabelle 3.9: Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen in IGLU 2001, 2006, 2011 und 2016 (Angaben in Prozent)

					IGLU 2016					IGLU 2011				
					Kompetenzstufen					Kompetenzstufen				
					I	II	III	IV	V	I	II	III	IV	V
2001	2006	2011	2016	Teilnehmer ^A	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)
1 2	2	2		⁸ Russische Föderation	0.9 (0.3)	5.5 (0.5)	23.4 (1.0)	44.5 (1.0)	25.7 (1.2)	1.0 (0.2)	7.5 (1.0)	28.7 (1.1)	43.5 (1.3) +	19.3 (1.2) +
		2	2	⁸ Singapur	2.7 (0.5)	8.3 (0.7)	22.8 (1.1)	37.5 (1.5)	28.6 (1.6)	3.2 (0.4)	9.6 (0.8)	25.4 (1.3)	37.4 (1.2)	24.4 (1.6)
	3	2	2 3	⁸ Hongkong	1.4 (0.3)	6.0 (0.7)	27.5 (1.4)	46.9 (1.5)	18.2 (1.3)	0.8 (0.2)	6.1 (0.6)	26.6 (1.3)	48.9 (1.3)	17.6 (1.2)
				Taiwan	1.6 (0.2)	8.5 (0.6)	31.1 (1.1)	44.4 (1.2)	14.4 (1.1)	2.1 (0.3)	10.7 (0.6)	31.9 (1.0)	42.5 (1.2)	12.8 (0.9)
1 2 3	1	1	3 1	⁸ England	3.2 (0.4)	11.3 (0.6)	28.4 (0.9)	37.0 (1.1)	20.1 (0.9)	4.6 (0.5) -	12.6 (1.0)	28.8 (1.1)	35.6 (1.0)	18.3 (1.1)
			2	Lettland	0.8 (0.2)	9.2 (0.8)	33.2 (1.3)	43.3 (1.4)	13.6 (1.0)	-	-	-	-	-
			2	Schweden	1.9 (0.3)	10.2 (0.8)	31.1 (1.1)	43.1 (1.7)	13.7 (1.4)	2.4 (0.3)	12.3 (0.9)	38.4 (1.4) +	37.9 (1.4) +	9.0 (0.8) +
				Ungarn	2.9 (0.5)	12.1 (0.8)	29.5 (1.2)	38.9 (1.1)	16.6 (1.2)	5.4 (0.7) -	13.6 (0.9)	33.4 (1.1)	35.4 (1.0) +	12.2 (0.9) +
	2			Bulgarien	5.2 (0.9)	12.2 (1.2)	28.1 (1.3)	35.4 (1.3)	19.2 (1.3)	7.0 (1.0)	15.9 (1.1) -	31.9 (1.2)	34.0 (1.5) +	11.2 (0.8) +
2 3	2 3	2 3	3	⁸ USA	3.9 (0.5)	12.7 (0.9)	30.5 (1.1)	36.7 (1.4)	16.1 (1.3)	2.5 (0.3) +	11.4 (0.5)	30.3 (0.6)	38.5 (0.7)	17.3 (0.7)
2 3	2	2		⁸ Litauen	2.7 (0.5)	10.7 (0.9)	33.5 (1.3)	40.3 (1.3)	12.8 (0.9)	3.5 (0.4)	16.6 (1.1) -	41.0 (1.3) +	32.9 (1.2) +	6.0 (0.5) +
	2	3		⁸ Italien	2.1 (0.5)	11.1 (0.8)	34.8 (1.3)	41.4 (1.7)	10.7 (0.8)	2.3 (0.4)	12.7 (1.0)	38.9 (1.0) +	36.4 (1.3)	9.7 (0.7)
	2	2	2	Dänemark	2.6 (0.4)	11.8 (0.8)	33.5 (1.0)	40.8 (1.1)	11.2 (1.0)	1.4 (0.2) +	10.3 (0.7)	33.0 (0.9)	42.9 (1.2)	12.5 (0.8)
3	3	3	3	Niederlande	1.3 (0.3)	11.0 (0.8)	39.4 (1.3)	39.9 (1.1)	8.4 (0.6)	0.4 (0.2) +	9.5 (0.8)	41.9 (1.1)	41.6 (1.4) +	6.5 (0.5) +
		2		Tschechien	3.0 (0.5)	11.8 (0.7)	36.4 (1.0)	38.8 (1.0)	10.0 (0.7)	1.7 (0.5)	11.0 (0.8)	37.5 (1.2)	41.6 (1.2)	8.2 (0.9)
1	1			⁸ Slowenien	3.7 (0.5)	13.5 (0.8)	33.7 (0.9)	37.8 (1.1)	11.3 (0.8)	4.5 (0.6)	16.0 (0.8)	37.9 (1.0) +	33.9 (1.1) +	7.7 (0.7) +
	2	2	2	Österreich	2.4 (0.4)	13.2 (0.9)	37.2 (0.9)	38.8 (1.3)	8.4 (0.8)	2.9 (0.3)	16.7 (0.8) -	41.3 (1.1) +	33.8 (1.3) +	5.2 (0.5) +
				Deutschland	5.5 (1.0)	13.4 (0.9)	34.1 (1.0)	35.9 (1.1)	11.1 (0.8)	2.4 (0.3) +	13.0 (0.8)	38.5 (1.2)	36.5 (1.2)	9.5 (0.8)
				Slowakei	6.6 (1.1)	12.7 (0.8)	33.4 (1.1)	37.0 (1.3)	10.3 (0.8)	4.2 (0.8)	13.6 (0.9)	37.7 (1.0)	36.6 (1.2) +	7.8 (0.6) +
	2	2		Spanien	3.4 (0.6)	16.7 (0.8)	41.2 (0.8)	33.0 (0.9)	5.7 (0.4)	6.1 (0.7) -	21.5 (1.0) -	41.7 (1.2) +	26.8 (1.3) +	3.9 (0.5) +
	2 3		3	⁸ Belgien (Fläm. Gem.)	2.6 (0.4)	17.4 (1.2)	44.6 (1.1)	31.4 (1.1)	4.1 (0.4)	-	-	-	-	-
1	1 2	1	1	Neuseeland	10.0 (0.7)	16.6 (0.7)	31.9 (1.0)	30.3 (1.0)	11.1 (0.6)	8.3 (0.5) +	16.9 (0.7)	30.2 (1.2)	30.9 (1.0) -	13.7 (0.7) -
2		2	2	Frankreich	6.3 (0.5)	21.8 (1.0)	41.7 (1.2)	26.4 (1.1)	3.8 (0.6)	4.7 (0.8)	20.1 (1.1)	40.6 (0.8)	29.3 (1.4)	5.2 (0.5)
		2 3	2	⁸ Belgien (Franz. Gem.)	8.4 (0.9)	27.0 (1.0)	42.4 (1.1)	19.7 (1.1)	2.6 (0.4)	6.2 (1.1)	23.4 (1.2) +	44.9 (1.3)	23.2 (1.3)	2.3 (0.5)
				Benchmark-Teilnehmer										
2			2 3	⁸ Québec, Kanada	1.6 (0.4)	11.5 (1.3)	37.0 (1.5)	38.6 (1.6)	11.3 (1.2)	1.8 (0.3)	13.1 (0.9)	42.0 (1.3)	36.1 (1.6) +	7.0 (0.7) +
2	2	2		⁸ Ontario, Kanada	4.3 (0.6)	13.5 (1.1)	32.1 (1.1)	36.5 (1.4)	13.5 (1.5)	2.7 (0.4) +	12.3 (1.0)	31.3 (1.4)	38.5 (1.1)	15.2 (1.3)
3	3	3	2	⁸ Norwegen (4. Jgst.)	5.8 (0.6)	20.4 (0.9)	40.3 (1.1)	28.9 (1.0)	4.7 (0.6)	5.2 (0.7) -	23.9 (1.5) -	46.0 (1.3) +	23.1 (1.5) +	1.9 (0.4) +

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

A = Zur Auswahl der Trendteilnehmer siehe Kapitel 2, Abschnitt 4.2 in diesem Band.

+ = Prozentwert in 2016 signifikant höher als 2011 bzw. 2006 bzw. 2001 ($p < .05$).

- = Prozentwert in 2016 signifikant niedriger als 2011 bzw. 2006 bzw. 2001 ($p < .05$).

IGLU 2006					IGLU 2001				
Kompetenzstufen					Kompetenzstufen				
I	II	III	IV	V	I	II	III	IV	V
% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)
1.6 (0.5)	8.7 (0.8) -	28.8 (1.7) +	42.1 (1.5)	18.9 (1.5) +	3.8 (1.2) -	15.8 (1.1) -	41.6 (1.7) -	33.4 (1.9) +	5.4 (0.9) +
3.3 (0.4)	10.5 (0.8) -	28.3 (1.2) +	38.6 (1.3)	19.4 (1.4) +	9.8 (1.4) -	14.2 (1.0) -	31.2 (1.4) -	32.5 (1.5) +	12.3 (1.4) +
0.9 (0.2)	6.7 (0.8)	30.4 (1.2)	47.4 (1.2)	14.6 (1.0) +	3.4 (0.6) -	15.2 (1.2) -	42.7 (1.2) -	34.1 (1.6) +	4.5 (0.6) +
3.0 (0.4) -	13.4 (0.9) -	40.6 (1.1) +	36.1 (1.1) +	6.9 (0.7) +	-	-	-	-	-
6.8 (0.7) -	14.8 (0.9) -	30.7 (1.2)	32.3 (1.2) +	15.5 (0.9) +	5.5 (0.7) -	12.1 (0.8)	28.0 (1.1)	34.3 (1.0)	20.1 (1.4)
1.8 (0.4) -	12.4 (1.2) -	40.3 (1.1) +	37.2 (1.1) +	8.3 (0.8) +	1.5 (0.4)	11.9 (0.9) -	38.1 (1.8) -	40.1 (2.0)	8.5 (0.9) +
1.8 (0.4)	10.5 (1.0)	35.2 (1.3) +	42.0 (1.2)	10.5 (0.9)	1.5 (0.3)	8.1 (0.7)	30.9 (1.2)	43.9 (1.0)	15.5 (1.0)
2.6 (0.5)	11.2 (1.2)	33.2 (1.2) +	39.0 (1.4)	14.0 (0.9)	2.5 (0.3)	12.5 (0.9)	36.4 (1.9) -	39.1 (1.6)	9.6 (0.9) +
5.2 (1.0)	12.6 (1.4)	30.6 (1.5)	35.1 (1.6)	16.5 (1.4)	5.0 (0.9)	12.0 (1.0)	28.7 (1.3)	36.9 (1.6)	17.5 (1.2)
4.0 (0.6)	14.3 (1.1)	34.7 (1.1) +	35.2 (1.3)	11.8 (1.2) +	6.0 (0.7) -	13.6 (1.3)	30.0 (1.4)	35.7 (1.2)	14.6 (1.1)
1.3 (0.3) +	12.6 (0.9)	43.4 (1.2) +	37.4 (1.4)	5.3 (0.8) +	2.0 (0.4)	12.6 (1.2)	37.4 (1.5)	39.1 (1.6)	9.0 (1.0) +
1.8 (0.4)	11.3 (1.2)	34.5 (1.3)	38.7 (1.3)	13.6 (1.4)	3.4 (0.6)	14.1 (1.2) -	35.0 (1.2)	37.0 (1.4)	10.5 (0.9)
3.5 (0.4)	11.4 (0.8)	33.3 (1.0)	40.6 (1.3)	11.2 (0.8)	-	-	-	-	-
0.6 (0.2)	8.4 (0.7) +	41.7 (1.0)	43.0 (1.1)	6.4 (0.5) +	0.8 (0.3)	7.5 (0.9) +	37.8 (1.4)	43.6 (1.4)	10.3 (0.9)
-	-	-	-	-	2.7 (0.4)	14.1 (1.0)	38.6 (1.4)	37.5 (1.4)	7.1 (0.8) +
5.6 (0.5) -	18.9 (0.9) -	39.0 (0.9) +	30.9 (1.0) +	5.7 (0.6) +	8.9 (0.6) -	24.2 (1.0) -	41.5 (1.1) -	22.5 (1.1) +	3.0 (0.4) +
2.4 (0.4)	13.7 (1.0)	38.8 (1.0)	37.4 (1.6)	7.6 (0.9)	-	-	-	-	-
2.6 (0.3) +	10.7 (0.8) +	34.9 (1.2)	41.1 (1.4) -	10.8 (0.9)	3.0 (0.4) +	14.0 (0.8)	36.5 (0.9)	38.0 (1.1)	8.6 (0.6) +
5.6 (0.9)	14.5 (1.0)	36.5 (1.0) +	35.3 (1.4)	8.0 (0.6) +	6.1 (0.8)	18.0 (1.0) -	42.0 (1.1) -	29.3 (1.4) +	4.6 (0.8) +
6.5 (0.8) -	21.6 (1.0) -	40.9 (1.2)	26.4 (1.0) +	4.6 (0.5)	-	-	-	-	-
0.8 (0.2) +	9.4 (0.8) +	40.5 (1.0) -	42.0 (1.4) -	7.4 (0.6) -	-	-	-	-	-
7.9 (0.6) +	16.1 (0.8)	31.2 (0.9)	31.6 (0.9)	13.2 (0.7) -	10.0 (1.0)	16.0 (1.2)	29.4 (1.1)	30.5 (1.2)	14.0 (1.2) -
4.1 (0.4) +	19.6 (1.0)	41.2 (1.0)	29.8 (1.0) -	5.2 (0.6)	4.6 (0.6)	18.0 (0.9) +	40.4 (1.1)	29.8 (1.2)	7.2 (0.8) -
7.9 (0.7)	26.5 (1.1)	42.2 (1.1)	20.7 (1.1)	2.7 (0.4)	-	-	-	-	-
2.6 (0.4)	14.9 (1.1) -	41.8 (1.4) +	34.4 (1.7)	6.3 (0.8) +	2.0 (0.4)	13.9 (1.3)	41.5 (1.5) -	34.6 (1.7)	8.0 (0.7) +
2.0 (0.5) +	10.6 (0.9) +	33.1 (1.6)	38.2 (1.4)	16.0 (1.1)	3.8 (0.6)	12.6 (1.1)	33.4 (1.3)	34.8 (1.3)	15.4 (1.2)
8.2 (0.8) -	25.3 (1.4) -	44.5 (1.2) +	20.4 (1.1) +	1.6 (0.3) +	11.8 (0.9) -	22.9 (1.2) -	37.6 (1.5) -	23.4 (1.3) +	4.2 (0.8)

6.5 Kompetenzen beim Lesen von literarischen und informierenden Texten

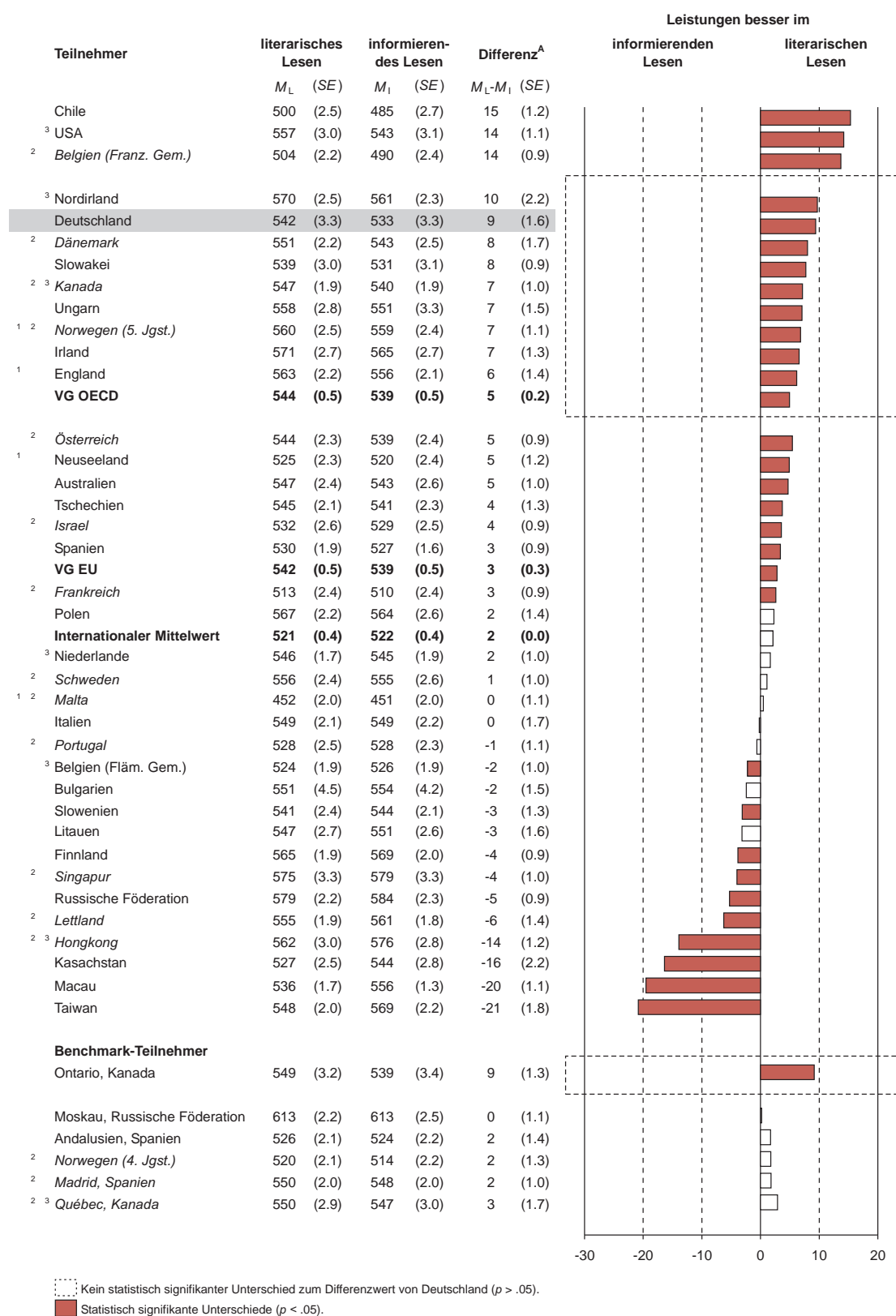
Im Rahmen von IGLU haben die Schülerinnen und Schüler Aufgaben zu narrativ-literarischen, das heißt fiktionalen, und zu faktualen beziehungsweise Sachtexten zu lösen. Das Testdesign ermöglicht es, die Ergebnisse getrennt nach Sachtexten und literarischen Texten darzustellen (siehe Abschnitt 2.1). Es kann also ermittelt werden, ob beide Textsorten auf demselben Niveau verstanden werden oder ob bei der einen oder anderen Textsorte relative Stärken beziehungsweise Schwächen zu konstatieren sind. Aus den Befunden können sich Hinweise zum Beispiel auf curriculare Schwerpunktsetzungen und unterrichtliche Lerngelegenheiten ergeben.

In Abbildung 3.9 ist der Unterschied zwischen den Leistungen beim Lesen von literarischen Texten und informierenden Texten für IGLU 2016 dargestellt. Wie der Abbildung zu entnehmen ist, zeigen sich für Deutschland (9 Punkte), wie für 19 weitere Teilnehmer, signifikant höhere Leistungen für das Lesen literarischer Texte. Die Differenz unterscheidet sich zwar nicht signifikant von der für die Vergleichsgruppe OECD, jedoch weisen signifikant höhere Unterschiede zugunsten des literarischen Lesens nur Chile, die USA und die Französische Gemeinschaft in Belgien auf. Die Differenz für die Vergleichsgruppe EU fällt hingegen signifikant niedriger aus als die für Deutschland. Für 13 Teilnehmerstaaten und -regionen zeigen sich keine signifikanten Unterschiede zwischen dem Lesen von literarischen Texten und von Sachtexten. Für zehn Teilnehmerstaaten, darunter alle asiatischen Teilnehmer, ergeben sich signifikant höhere Werte für das Lesen von Sachtexten.

In Abbildung 3.10 sind die Ergebnisse von IGLU 2001, 2006, 2011 und 2016 in Bezug auf die Leistungen beim Lesen von literarischen und informierenden Texten dargestellt (vgl. Mullis et al., 2017, Exhibit 3.4). Für auf der linken Seite dargestellte Staaten und Regionen sowie Benchmark-Teilnehmer (siehe Seite 132) lassen sich auf der Gesamtskala Lesen signifikant positive Veränderungen der Schülerleistungen von 2001 beziehungsweise 2006 zu 2016 beobachten (siehe Abbildung 3.7). Auf der rechten Seite sind Staaten und Regionen sowie Benchmark-Teilnehmer angeordnet, bei denen sich keine signifikanten Veränderungen oder negativen Leistungstrends beobachten lassen. Darüber hinaus sind die Teilnehmer hinsichtlich der Trends im Umgang mit den beiden Textsorten gruppiert. Dabei interessiert, ob Differenzen zwischen dem Verstehen von literarischen und von Sachtexten auszumachen sind beziehungsweise in welchem Ausmaß sie sich verändert haben. Wie die Befunde im Einzelnen zu interpretieren sind, ist abhängig vom präferierten Bezugsrahmen. Wer es zum Beispiel als vordringlich ansieht, in der Grundschule das Verständnis fiktionaler Texte zu fördern, mag höhere Testwerte im Bereich literarischer Texte und auch ein Anwachsen der Differenz über die vier Testzeitpunkte für unproblematisch oder auch wünschenswert halten. Wer dagegen darauf setzt, dass die Disparitäten möglichst gering sind beziehungsweise sich verringern sollten, kommt zum entgegengesetzten Schluss.

Aus Abbildung 3.10 geht hervor, dass für nur fünf Teilnehmer, Deutschland, die Französische Gemeinschaft in Belgien, die Slowakei, Singapur und Taiwan, eine signifikante Zunahme der Leistungsunterschiede im Umgang mit literarischen und Sachtexten zu beobachten ist. In Taiwan und Singapur beruht sie auf einem weit größeren Zuwachs im Bereich des Lesens von Sachtexten, in Deutschland und Belgien auf im Vergleich mit 2001 niedrigeren Leistungen beim

Abbildung 3.9: Differenz zwischen den Testleistungen nach Textsorten im internationalen Vergleich



Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

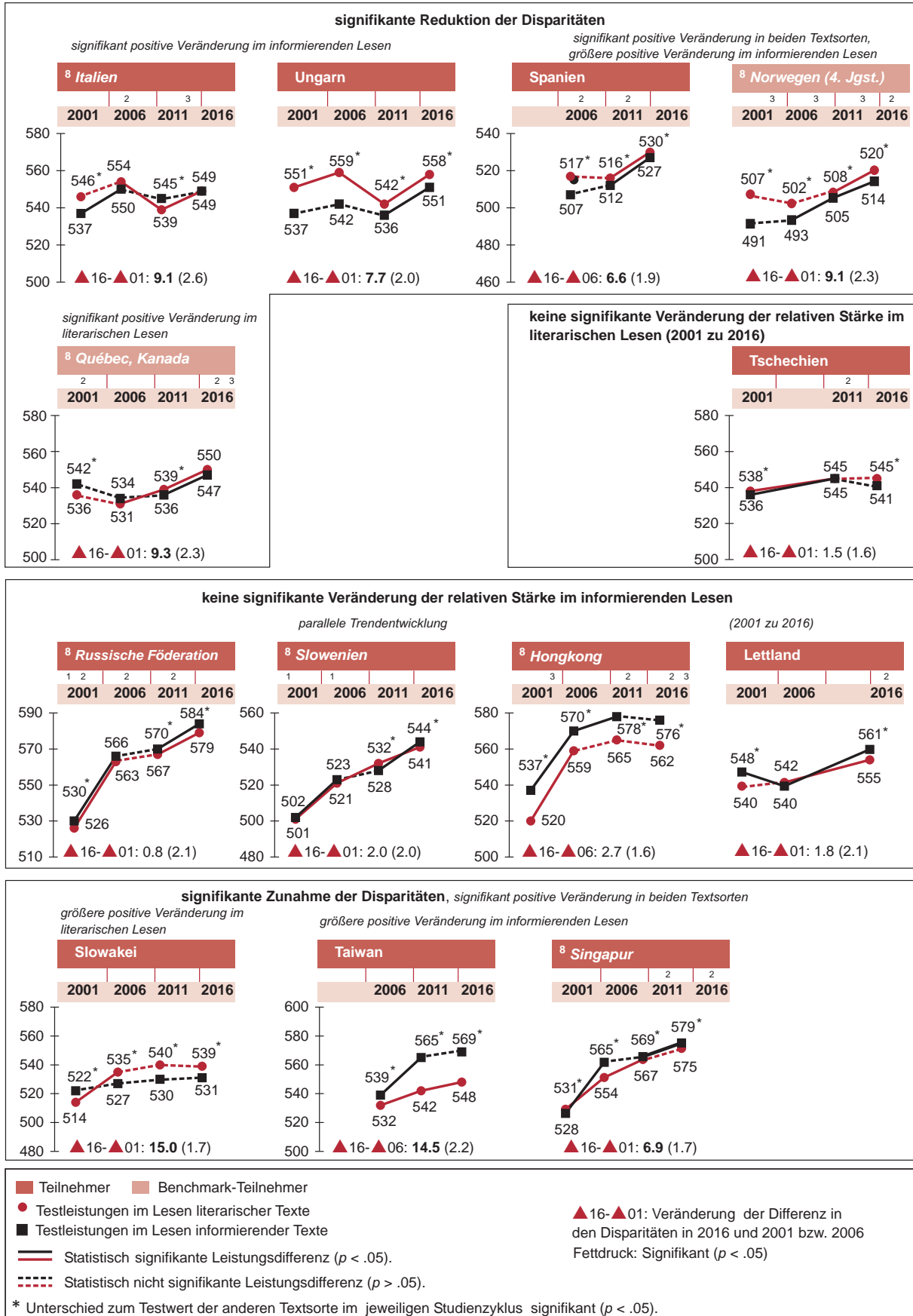
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

Abbildung 3.10: Veränderung der Differenz zwischen den Testleistungen nach Textsorten in IGLU 2001, 2006, 2011 und 2016 im internationalen Vergleich

Teilnehmer mit positiven Veränderungen auf der Gesamtskala Lesen



■ Teilnehmer ■ Benchmark-Teilnehmer

● Testleistungen im Lesen literarischer Texte

■ Testleistungen im Lesen informierender Texte

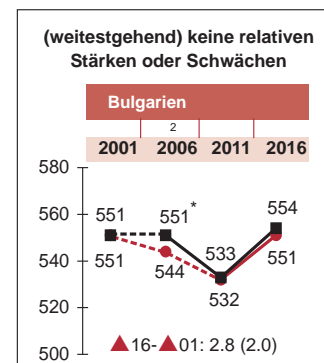
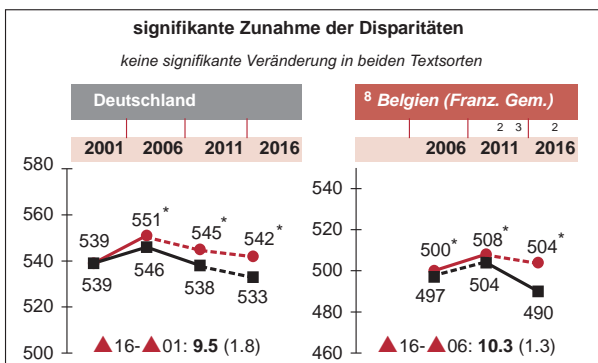
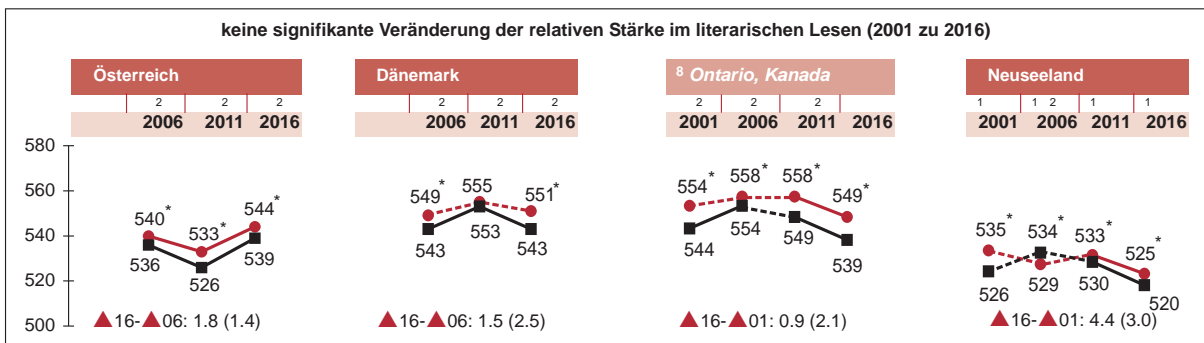
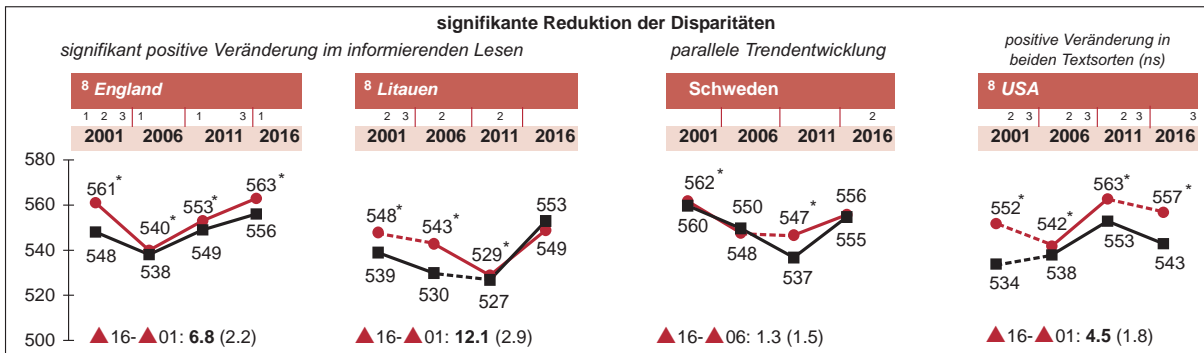
— Statistisch signifikante Leistungsdifferenz ($p < .05$).

- - - Statistisch nicht signifikante Leistungsdifferenz ($p > .05$).

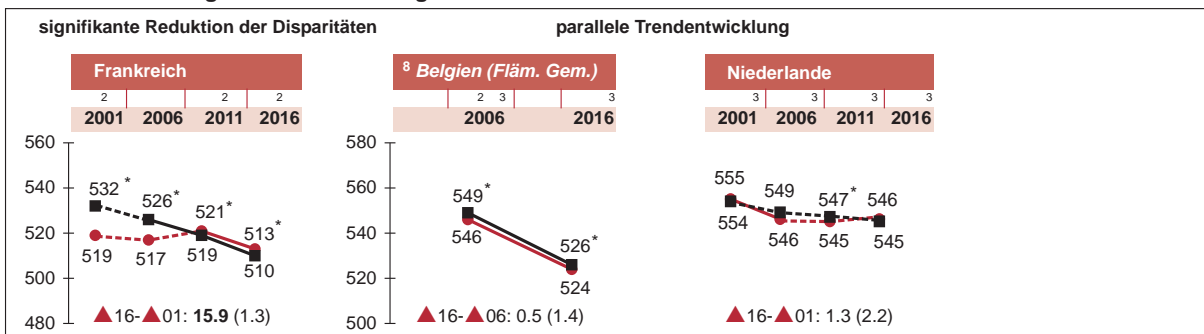
* Unterschied zum Testwert der anderen Textsorte im jeweiligen Studienzyklus signifikant ($p < .05$).

▲16-▲01: Veränderung der Differenz in den Disparitäten in 2016 und 2001 bzw. 2006
Fettdruck: Signifikant ($p < .05$)

Teilnehmer mit nicht signifikanten Veränderungen auf der Gesamtskala Lesen



Teilnehmer mit negativen Veränderungen auf der Gesamtskala Lesen



Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den hier betrachteten Studienzyklen.
 Die Ergebnisse von Israel und Polen werden aufgrund der nicht gegebenen Vergleichbarkeit zwischen den hier betrachteten Studienzyklen nicht berichtet.

Lesen von Sachtexten und leicht gestiegenen Werten für das Lesen von literarischen Texten. Beide Trends sind – je für sich betrachtet – allerdings nicht signifikant.

Für zehn Teilnehmerstaaten und -regionen lässt sich zwischen 2001 beziehungsweise 2006 und 2016 eine signifikante Verringerung der Unterschiede feststellen. Bei 12 Staaten und Regionen haben sich keine Veränderungen ergeben.

Für 5 von 13 Teilnehmerstaaten und Regionen mit positiven Leistungstrends auf der Gesamtskala Lesen (siehe Seite 132) ergeben sich parallele Entwicklungen in beiden Textsorten. Für weitere sechs Teilnehmer, Italien, Ungarn, Spanien, Norwegen (4. Jgst.), Taiwan und Singapur, sind größere Zuwächse im informierenden Lesen zu beobachten. Für Québec und die Slowakei fallen insbesondere gestiegene Leistungen beim literarischen Lesen ins Gewicht.

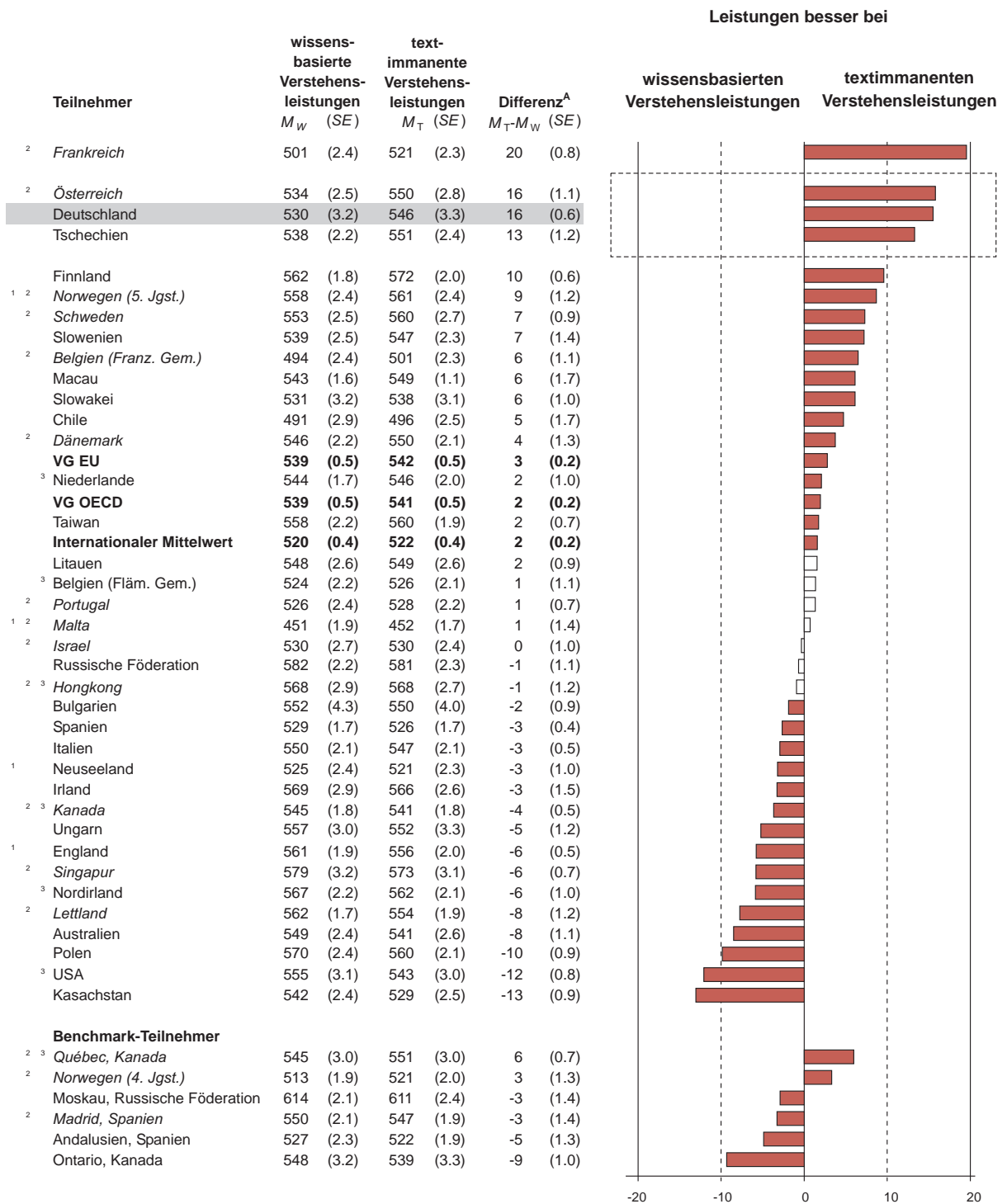
Für die meisten Teilnehmerstaaten und -regionen mit nicht signifikanten oder negativen Veränderungen auf der Gesamtskala Lesen (siehe Seite 133) lassen sich keine bedeutsamen Veränderungen der Leistungen, differenziert nach Textsorten, erkennen.

6.6 Ergebnisse zu den Verstehensleistungen

Das Lesekompetenzmodell von IGLU sieht die Unterscheidung von vier Verstehensleistungen vor, die sich in zwei Subskalen zusammenfassen lassen (siehe Abschnitt 2.2). In der Subskala ‚textimmanente Verstehensleistungen‘ wurden die Leseaufgaben zusammengefasst, die den Prozessen ‚Lokalisieren explizit angegebener Informationen‘ und ‚Einfache Schlussfolgerungen ziehen‘ zugeordnet worden sind. Aufgaben zu den Prozessen ‚Komplexe Schlussfolgerungen ziehen: Interpretieren und Kombinieren‘ sowie ‚Prüfen und Bewerten des Inhalts und des Sprachgebrauchs‘ zählen zur Subskala ‚wissensbasierte Verstehensleistungen‘. Die Zuordnungen von Aufgaben und Verstehensleistungen beruhen auf Expertenurteilen. Darauf, dass es zuweilen schwierig sein mag, zu einem Konsens zu gelangen, insbesondere dann, wenn zu entscheiden ist, ob ein einfacher (intendierter) oder ein komplexer (elaborativer) Schluss zu ziehen ist, wurde bereits hingewiesen (ebd.).

Abbildung 3.11 zeigt die Unterschiede zwischen den Mittelwerten bei den textimmanenten und wissensbasierten Verstehensleistungen für die Teilnehmerstaaten und -regionen. In Deutschland sowie in 16 weiteren Staaten und Regionen schneiden die Schülerinnen und Schüler bei textimmanenten Verstehensleistungen signifikant besser ab. Für Deutschland ergibt sich eine Differenz zwischen den Verstehensleistungen von 16 Punkten. Lediglich in Frankreich fällt diese Differenz signifikant höher aus, und nur in Österreich und Tschechien hat sie eine Größenordnung wie in Deutschland. Damit liegt der Unterschied zwischen den Mittelwerten für die textimmanenten und die wissensbasierten Verstehensleistungen in Deutschland signifikant über den Differenzen für die Vergleichsgruppe EU (3), die Vergleichsgruppe OECD (2) und für den internationalen Mittelwert (2), in den die Ergebnisse aller an IGLU 2016 beteiligten Staaten und Regionen einfließen. Für Litauen, Belgien, Portugal, Malta, Israel, die Russische Föderation und Hongkong ergeben sich keine signifikanten Unterschiede zwischen den Mittelwerten für die unterschiedlichen Verstehensleistungen. Die Kinder in 19 Teilnehmerstaaten und -regionen erreichen bei den wissensbasierten Verstehensleistungen signifikant bessere Ergebnisse als bei den textimmanenten, darunter überproportional viele derer, bei

Abbildung 3.11: Differenzen zwischen den Testleistungen nach Verstehensprozessen im internationalen Vergleich



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland ($p > .05$).

Statistisch signifikante Unterschiede ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

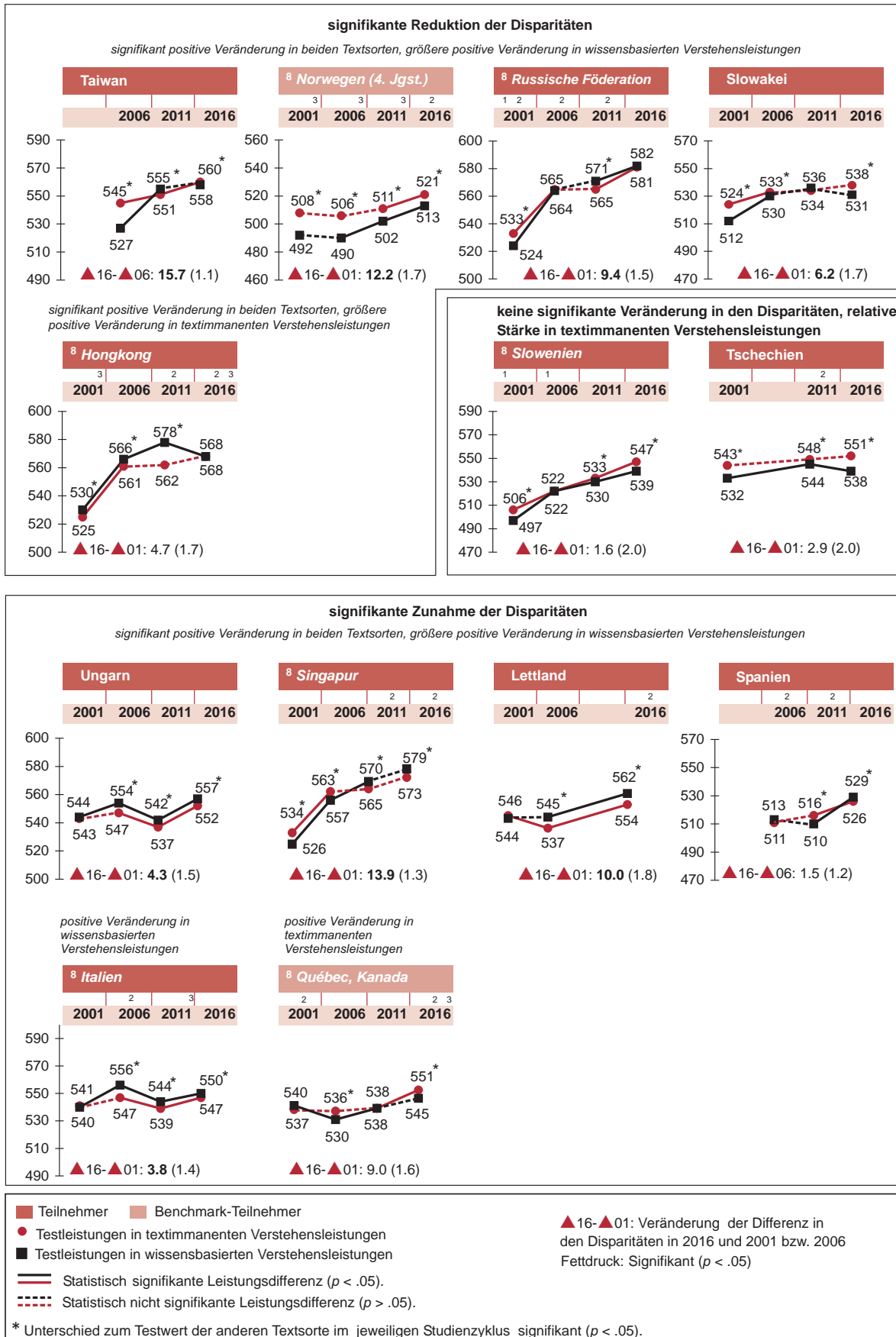
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

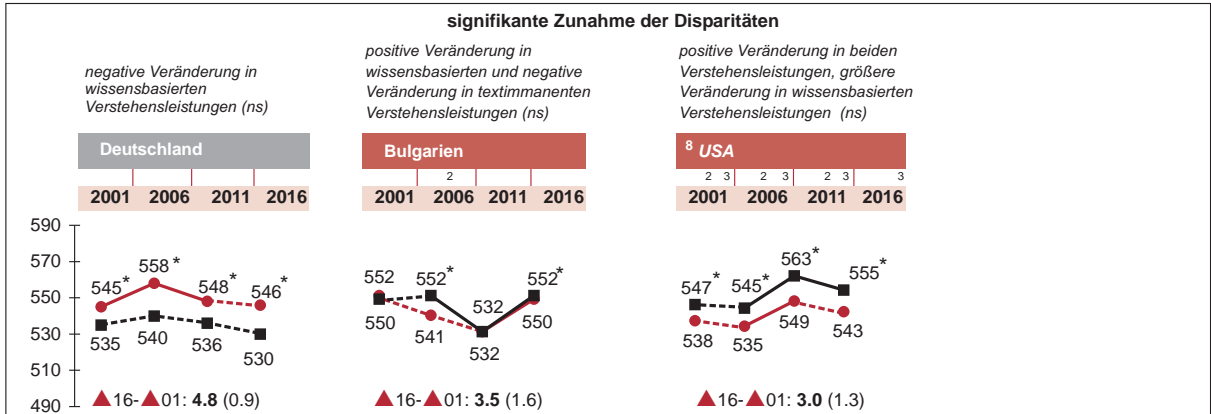
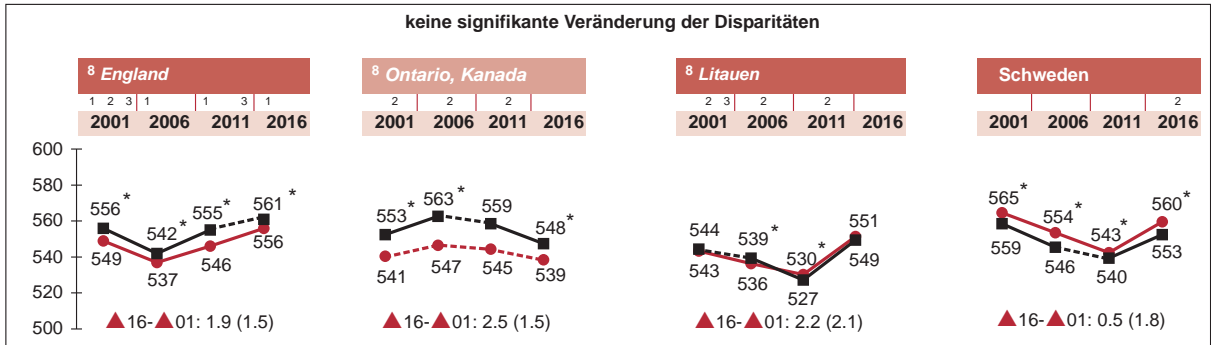
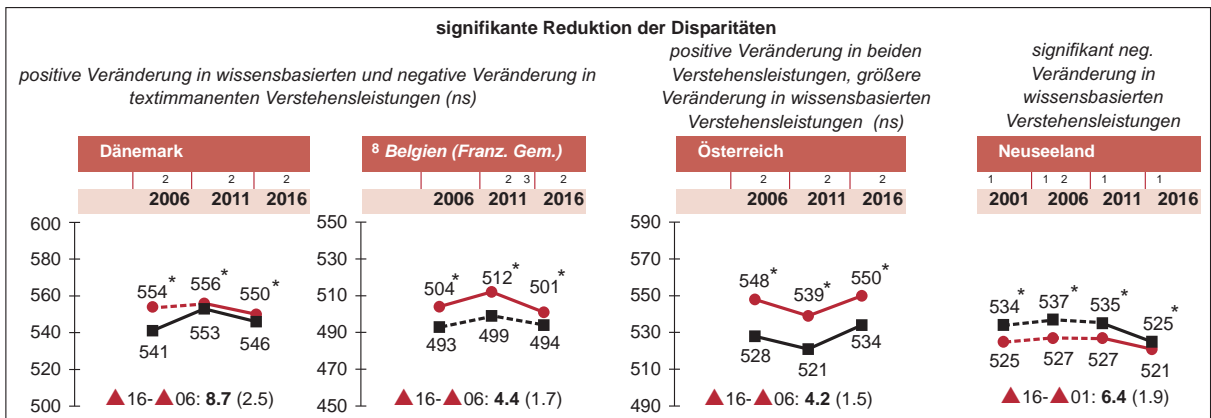
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

Abbildung 3.12: Veränderung der Differenz zwischen den Testleistungen nach Textsorten in IGLU 2001, 2006, 2011 und 2016 im internationalen Vergleich

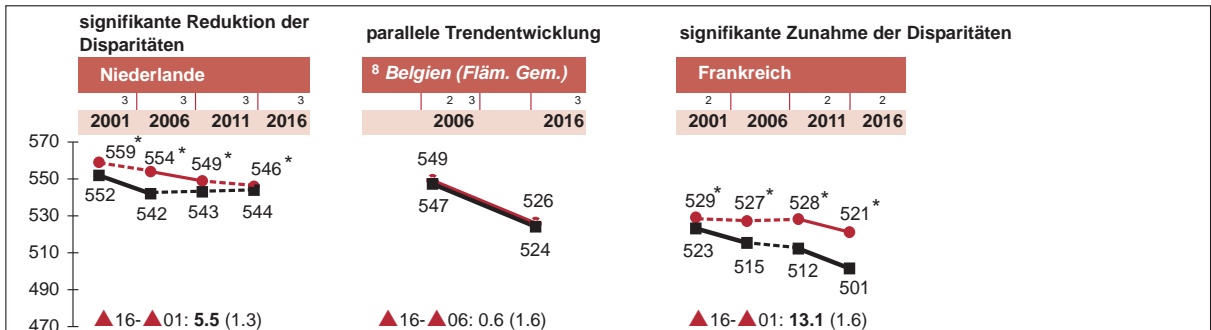
Teilnehmer mit positiven Veränderungen auf der Gesamtskala Lesen



Teilnehmer mit nicht signifikanten Veränderungen auf der Gesamtskala Lesen



Teilnehmer mit negativen Veränderungen auf der Gesamtskala Lesen



Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den hier betrachteten Studienzyklen.

Die Ergebnisse von Israel und Polen werden aufgrund der nicht gegebenen Vergleichbarkeit zwischen den hier betrachteten Studienzyklen nicht berichtet.

denen auch auf der Gesamtskala Lesen bessere Leistungen zu verzeichnen sind. Das ist nicht verwunderlich, sind die schwierigeren Leseaufgaben doch die, deren Lösung vornehmlich auf wissensbasierten kognitiven Prozessen beruht (siehe Abschnitte 2.2 und 4.1).

Für zehn Teilnehmer vergrößerten sich die Differenzen bei den Verstehensleistungen im Verlauf der Testzyklen, und zwar in Singapur, Frankreich, Lettland, Québec, Deutschland, Ungarn, Italien, Bulgarien, den USA und Spanien. In zehn Staaten nahmen die Differenzen ab, besonders eindrücklich in Taiwan, Norwegen (4. Jgst.) und der Russischen Föderation.

7 Zusammenfassung

Im Jahr 2016 erreichen die Viertklässlerinnen und Viertklässler in Deutschland einen Mittelwert von 537 Punkten und damit einen Platz im Mittelfeld. Dieser Wert unterscheidet sich nicht bedeutsam vom EU- und OECD-Durchschnitt. In der EU sind die Ergebnisse für Österreich, die Slowakei, Slowenien und Tschechien mit denen in Deutschland vergleichbar, signifikant schlechter schneiden nur Malta, Belgien (Flämische und Französische Gemeinschaft), Frankreich, Portugal und Spanien ab. Deutlich erfolgreicher ist aber die große Mehrheit der zur EU gehörenden Staaten, nämlich Irland, Finnland, Polen, Nordirland, England, Lettland, Schweden, Ungarn, Bulgarien, Litauen, Italien, Dänemark und die Niederlande. Was die OECD betrifft, so sind unter anderem die Schülerinnen und Schüler in den USA signifikant erfolgreicher.

Betrachtet man die Ergebnisse im Trend, das heißt im Vergleich mit Resultaten im Jahr 2001, so ergibt sich für Deutschland keine statistisch relevante Veränderung. Dieser Befund bleibt auch bestehen, wenn man die Veränderungen in der Zusammensetzung der Schülerschaft berücksichtigt. Im Jahr 2006 gab es zwar ein „Zwischenhoch“, es verschwand aber bereits im Jahr 2011. Vergleichbare Leistungen wie im Jahr 2001 zeigten zum Beispiel auch England, Schweden, Dänemark, die USA, Österreich, die Französische Gemeinschaft in Belgien, Bulgarien und Litauen. Für die Niederlande und Frankreich sind signifikante Verschlechterungen zu verzeichnen, für 13 Teilnehmer dagegen positive Veränderungen. Insofern verwundert es nicht, dass sich die relative Position Deutschlands verschlechtert hat. 2001 lasen Kinder in vier Staaten signifikant besser, 2016 waren es Schülerinnen und Schüler in 20 Staaten.

2016 fällt die Streuung der Leistungen in Deutschland besonders hoch aus. Die Differenz zwischen dem 5. und dem 95. Perzentil, also den Leistungen, die die 5 Prozent der schwächsten Kinder maximal und die 5 Prozent der leistungsstärksten Kinder mindestens erreichen, beträgt 257 Punkte und ist EU-weit nur in Malta bedeutsam größer. Ein ähnliches Bild ergibt sich bei Betrachtung der Standardabweichung. In 17 der EU-Staaten ist sie signifikant geringer.

Im Trend zeigt sich, dass die Leistungen in Deutschland 2016 deutlich heterogener ausfallen als im Jahr 2001. Die Standardabweichung vergrößert sich von 67 auf 78 Punkte, im 5. Perzentil wurden im Jahr 2001 maximal 419, im Jahr 2016 jedoch nur 395 Punkte erreicht. Auch in Bezug auf das 95. Perzentil sind deutliche Veränderungen zu verzeichnen: 2001 wurden hier minimal 640, 2016 aber 652 Punkte erzielt. Die gewachsene Heterogenität resultiert somit sowohl aus geringeren Leistungen der schwächeren als auch aus besseren Leistungen der stärksten Schülerinnen und Schüler. Nur in wenigen anderen Ländern ist über die

Erhebungszyklen hinweg eine signifikante Zunahme der Heterogenität zu vermerken.

Leseleistungen auf dem Niveau der Kompetenzstufe V zeigen im Jahr 2016 11 Prozent der Kinder in Deutschland, etwa halb so viele wie beispielsweise in Nordirland, Irland, Polen, England und Bulgarien. Gleichwohl entspricht dieser Anteil etwa dem in den anderen EU- und OECD-Ländern. Allenfalls rudimentäres Leseverständnis auf Kompetenzstufe I ist knapp 6 Prozent der Viertklässlerinnen und Viertklässler in Deutschland zu attestieren. In Europa ist der Anteil sehr schwacher Leserinnen und Leser nur in Frankreich, der Französischen Gemeinschaft in Belgien, der Slowakei und in Malta nominell größer. Kompetenzstufe III wird in Deutschland von ca. 19 Prozent der Viertklässlerinnen und Viertklässler nicht erreicht. Auch diese Quote ist, international betrachtet, recht hoch. Von den meisten europäischen Staaten wird sie signifikant unterschritten, signifikant überboten nur von Frankreich, der Französischen Gemeinschaft in Belgien und Malta.

Betrachtet man die Befunde für Deutschland im Trend, das heißt, vergleicht man die Ergebnisse für 2001 und 2016, dann zeigt sich, dass der Anteil sehr guter Leserinnen und Leser von knapp 9 auf 11 Prozent gestiegen ist. Zugleich wuchs die Gruppe der sehr schwachen Viertklässlerinnen und Viertklässler von 3 auf fast 6 Prozent und auch der Anteil derjenigen, die Stufe III nicht erreichen, stieg von 17 auf 19 Prozent. Zum Vergleich die – gegenläufigen – Trends in zwei anderen europäischen Ländern, die seit 2001 teilnehmen: In Slowenien nahm der Anteil der auf Stufe V Lesenden im Vergleich von 2001 und 2016 von 3.0 auf 11.3 Prozent zu, zugleich sank die Quote der auf Stufe I oder II Lesenden von 33.1 auf 17.2 Prozent. In Frankreich dagegen ging der Prozentsatz der auf Stufe V Platzierten von 7.2 auf 3.8 zurück und der Anteil der schwachen Leserinnen und Leser auf den Stufen I und II stieg von 22.6 auf 28.1 Prozent.

Kinder in Deutschland haben signifikant höhere Kompetenzen im Umgang mit literarischen Texten als beim Lesen von Sachtexten. Das ist bei knapp der Hälfte der anderen Teilnehmer nicht anders, die Differenz ist in Deutschland aber besonders ausgeprägt. Markanter fällt sie nur in den USA, in Chile und in der Französischen Gemeinschaft in Belgien aus. Der Mittelwert der Differenz für die VG EU liegt bedeutend unter dem für Deutschland. Auffällig ist, dass unter den Ländern, in denen es eine Differenz zugunsten der Sachtexte gibt, mit Russische Föderation, Singapur, Hongkong, Taiwan und Finnland die leistungsstärksten zu finden sind.

Taiwan, die Slowakei, Singapur, die Französische Gemeinschaft in Belgien und Deutschland sind die einzigen Länder, in denen die Differenz der Leistungspunkte für das Lesen der beiden Textsorten seit 2001 signifikant zugenommen hat. In Taiwan und Singapur ist dies auf das überproportionale Anwachsen der Punkte für das Lesen von Sachtexten zurückzuführen, in Belgien und Deutschland dagegen sind die Leistungen beim Lesen von Sachtexten geringer als im Jahr 2001 und zugleich beim Umgang mit literarischen Texten leicht erhöht.

Markante Disparitäten ergeben sich auch, wenn die Verstehensprozesse, die dem IGLU-Modell zufolge beim Lesen im Spiel sind, genauer betrachtet werden. Mit 16 Punkten zugunsten der als ‚textimmanent‘ bezeichneten Prozesse fällt die Differenz in Deutschland (wie in Österreich und Tschechien) besonders deutlich aus. Nur in Frankreich ist sie signifikant größer, in den EU- und OECD-Staaten im Mittel signifikant kleiner. Unter den Ländern, bei denen bedeutend größere Differenzen zugunsten der wissensbasierten Verstehensprozesse zu verzeichnen

sind, befinden sich überproportional viele, in denen die Kinder besonders leistungsfähig sind.

In Deutschland wie in neun weiteren Teilnehmerstaaten und -regionen vergrößerten sich die Differenzen im Zeitraum von 2001 bis 2016. In weiteren zehn Staaten nahmen sie ab, in sieben kam es nicht zu relevanten Veränderungen.

Im Rahmen einer Bewertung der Ergebnisse von IGLU 2011 notierten Bos et al. (2012, S. 134) unter anderem „für Deutschland grundsätzlich ein Leistungsniveau im oberen Drittel der Rangreihe und eine international gesehen relativ geringe Streuung der Leistungen von besonders guten und besonders schwachen Leserinnen und Lesern.“ Davon kann bei IGLU 2016 nicht mehr die Rede sein. Nun sind die Leistungen im zweiten Drittel angesiedelt, im EU-Vergleich schneiden sogar etwas mehr als die Hälfte der Teilnehmer besser ab. Die Leistungsbreite hat zugenommen, erfreulicherweise an der Spitze der Leistungsskala, aber auch an deren Ende. Dass fast ein Fünftel der Viertklässlerinnen und Viertklässler in Deutschland, mehr als in der weit überwiegenderen Zahl der EU-Länder, Kompetenzstufe III nicht erreicht, ist besorgniserregend. Wie bei den früheren IGLU-Zyklen zeigt sich, dass die Kinder in Deutschland beim Lesen von Sachtexten schlechter abschneiden als im Umgang mit literarischen Texten. Angesichts der Bedeutung, die das Verstehen von Sachtexten in den Schulformen der Sekundarstufe I hat, ist auch dieses Ergebnis bedenklich. Das gilt ebenso für die – über die verschiedenen Testungen hinweg zu beobachtende – Differenz von textimmanenten und wissensbasierten Verstehensleistungen. In Deutschland fällt sie, zuungunsten letzterer, besonders groß aus. Dabei ist in der Regel gerade wissensbasiertes Verstehen für die Lösung schwierigerer Aufgaben nötig.

Insofern sollten die Ergebnisse von IGLU 2016 Anlass sein, sich intensiver einer Herausforderung zu stellen, die man als Frage formulieren kann:

Kann man den Leseunterricht in der Grundschule so gestalten, dass die Aufgaben zum einen anspruchsvoller werden und zum anderen den heterogener gewordenen Lernvoraussetzungen der Schülerinnen und Schüler Rechnung getragen wird? Dass die Antwort auf diese Frage nicht auf eine Quadratur des Kreises hinauslaufen muss, zeigen die IGLU-Ergebnisse in einer Reihe anderer Länder.

Literatur

- Ballstaedt, S. P., Mandl, H., Schnotz, W. & Tergan, S.-O. (1981). *Texte verstehen, Texte gestalten*. München: Urban & Schwarzenberg.
- 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.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (Hrsg.). (2008). *IGLU-E 2006. – Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R. & Walther, G. (Hrsg.). (2004). *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Schwippert, K., Valtin, R., Voss, A., Badel, I. & Plafmeier, N. (2003). Lesekompetenzen deutscher Grundschulherinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin. (Hrsg.), *Erste Ergebnisse aus IGLU* (S. 69–142). Münster: Waxmann.

- Bos, W., Valtin, R., Hornberg, S., Buddeberg, I., Goy, M. & Voss, A. (2007). Internationaler Vergleich 2006: Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006 – Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 109–160). Münster: Waxmann.
- Bos, W., Valtin, R., Voss, A., Hornberg, S. & Lankes, E.-M. (2007). Konzepte der Lesekompetenz in IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006 – Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 81–107). Münster: Waxmann.
- Böhme, K., Richter, D., Weirich, S., Haag, N., Wendt, H., Bos, W., Pant, H. P. & Stanat, P. (2014). Messen wir dasselbe? Zur Vergleichbarkeit des IQB-Ländervergleichs 2011 mit den internationalen Studien IGLU und TIMSS 2011. *Unterrichtswissenschaft*, 42 (4), 1–24.
- Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Das Konzept von Lesekompetenz in IGLU 2011. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011 – Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 69–89). Münster: Waxmann.
- Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., Oranje, A. & Goldstein, A. (2005). *The nation's report card. Fourth-grade students' reading aloud: NAEP 2002 Special study of oral reading*, Washington, D.C.: U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics.
- Deutscher Bildungsserver (Hrsg.). (2011, 23. Februar). *Rahmenrichtlinien/Lehrpläne für die Grundschule*. Zugriff am 23.11.2017 unter <https://www.bildungsserver.de/Rahmenrichtlinien-Lehrplaene-fuer-die-Grundschule-1660-de.html>
- Kasper, D., Wendt, H., Bos, W. & Köller, O. (2016). Trends in mathematischen und naturwissenschaftlichen Kompetenzen am Ende der Grundschulzeit in Deutschland. 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. 367–382). Münster: Waxmann.
- Kintsch, W. (2012). Psychological models of reading comprehension and their implications for assessments. In J. Sabatini, E. Albro & T. O'Reilly (Hrsg.), *Measuring up: Advances in how to assess reading ability* (S. 21–37). Plymouth: Rowman & Littlefield Publishers.
- KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Wolters-Kluwer.
- Lenhard, W., Lenhard, A. & Schneider, W. (2017). *ELFE II. Ein Leseverständnistest für Erst- bis Siebtklässler – Version II*. Göttingen: Hogrefe.
- Lumley, T., Routitsky, A., Mendelovits, J. & Ramalingam, D. (2012). *A framework for predicting item difficulty in reading tests*. Paper presented at the American Educational Research Association Meeting, Vancouver. Zugriff am 23.11.2017 unter <http://research.acer.edu.au/pisa/5>
- Mullis, I. V. S. (2012). Using scale anchoring to interpret the TIMSS and PIRLS 2011 achievement scales. In M. O. Martin & I. V. S. Mullis (Hrsg.), *Methods and procedures in TIMSS and PIRLS* (S. 1–6). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (Hrsg.). (2017). *PIRLS 2016 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Goh, S. & Prendergast, C. (Hrsg.). (2017). *PIRLS 2016 Encyclopedia: Education policy and curriculum in reading*. Chestnut Hill: Boston College.
- Mullis, I. V. S., Martin, M. O., Gonzalez, E. J. & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary schools*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M. & Foy, P. (2007). *PIRLS 2006 international Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

- Mullis, I. V. S., Martin, M. O., Kennedy, A., Trong, K. & Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O. & Sainsbury, M. (2015). PIRLS 2016 reading framework. In I. V. S. Mullis & M. O. Martin (Hrsg.), *PIRLS 2016 assessment framework* (2. Auflage, S. 11–29). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- NICHD – National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel: Teaching children to read – An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: NICHD.
- Nix, D. (2011). *Förderung der Leseflüssigkeit. Theoretische Fundierung und empirische Überprüfung eines kooperativen Lautlese-Verfahrens im Deutschunterricht*. Weinheim: Juventa.
- Olsen, R. V. & Nilsen, T. (2017). Standard setting in PISA and TIMSS and how these procedures can be used nationally. In S. Blömeke & J.-E. Gustafsson (Hrsg.), *Standard setting in education* (S. 69–84). Wien: Springer.
- 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.
- Rosebrock, C. (2016). Anforderungen von Sach- und Informationstexten, Anforderungen literarischer Texte. In A. Bertschi-Kaufmann (Hrsg.), *Lesekompetenz – Leseleistung – Leseförderung* (S. 58–75). Seelze: Klett-Kallmeyer.
- Stanat, P., Schipolowski, S., Rojsk, C., Weirich, S. & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Spinner, K. (2016). Lese- und Literaturdidaktik. In A. Pompe, K. H. Spinner & J. Ossner (Hrsg.), *Deutschdidaktik Grundschule* (S. 140–198). Berlin: Erich Schmidt.
- Wittig, J. & Weirich, S. (2017). Mittelwerte und Streuungen der im Fach Deutsch erreichten Kompetenzen. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016 – Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 153–167). Münster: Waxmann.

Kapitel IV

Leseselbstkonzept, Lesemotivation, Leseverhalten und Lesekompetenz

Martin Goy, Renate Valtin und Anke Hußmann

1 Einleitung

In der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) wird Lesekompetenz als Fähigkeit verstanden, verschiedene Arten von Texten, die gesellschaftlich relevant sind und/oder individuell wertgeschätzt werden, zu verstehen und zu nutzen. Dabei werden verschiedene Anlässe des Lesens berücksichtigt, die sich für Kinder am Ende der Grundschulzeit typischerweise ergeben: Kinder lesen unter anderem, um etwas zu lernen, weil sie Freude am Lesen haben, und um sich mit anderen in verschiedenen Kontexten wie Schule, Familie oder Freundeskreis über das auszutauschen, was sie gelesen haben (Mullis & Martin, 2015). Diese Auffassung von Lesekompetenz als *reading literacy* nimmt somit explizit Bezug auf motivationale und verhaltensbezogene Aspekte. Damit wird berücksichtigt, dass die Fähigkeit, Texte zu lesen und zu verstehen, nicht allein von kognitiven Faktoren wie dem Vor- und Weltwissen der Lesenden oder dem Einsatz von Lesestrategien abhängt, sondern auch von weiteren Merkmalen der Leserinnen und Leser, die dazu beitragen, ob und wie die kognitiven Faktoren auch genutzt werden (siehe auch Kapitel 2 und 3 in diesem Band).

Modellen zu Determinanten der Lesekompetenz zufolge ist anzunehmen, dass Kompetenzüberzeugungen (wie das Leseselbstkonzept) und weitere motivationale Faktoren (wie die intrinsische Lesemotivation) das Verhalten (wie die Lesehäufigkeit und die Lesemenge) beeinflussen und dies wiederum auf das Leseverständnis wirkt. Das Leseselbstkonzept und die Lesemotivation hängen daher wie das Leseverhalten mit den Leseleistungen zusammen und sind pädagogisch auch deshalb bedeutsam, weil sie vergleichsweise gut zu beeinflussen sind und somit in Unterricht und Schule Ansatzpunkte für die Entwicklung und Förderung der Lesekompetenz von Kindern und Heranwachsenden bieten (Artelt et al., 2007; Möller & Schiefele, 2004).

In diesem Kapitel werden die Befunde aus IGLU 2016 zum Leseselbstkonzept, zur Lesemotivation und zum Leseverhalten der Schülerinnen und Schüler berichtet und die Zusammenhänge dieser Variablen mit ihren Leseleistungen betrachtet. Die Anlage von IGLU als Trenderhebung erlaubt zudem Aussagen über Veränderungen seit 2001, 2006 beziehungsweise 2011. Vor dem Hintergrund des Befunds, dass fast ein Fünftel der Viertklässlerinnen und Viertklässler in Deutschland Kompetenzstufe III nicht erreicht und davon auszugehen ist, dass sie mit erheblichen Schwierigkeiten beim Lernen in allen Fächern in der Sekundarstufe I konfrontiert sein werden (siehe Kapitel 3 in diesem Band), wird zudem in den Blick genommen, in welchem Zusammenhang das Leseselbstkonzept, die Lesemotivation und das Leseverhalten mit einer Verringerung des Risikos stehen, der Gruppe der leistungsschwachen Leserinnen und Leser anzugehören. Gleichzeitig wird betrachtet, welche Bedeutung dabei Merkmale wie elterlicher Berufsstatus, Migrationshintergrund oder die Lesesozialisation in der Familie haben.

1.1 Leseselbstkonzept

Der Mensch entwickelt nicht nur von seiner Umwelt, sondern auch von sich selbst ein Bild. Dieses ‚Selbst-Bild‘ wird als eine interne, kognitive Repräsentation des Wissens über sich selbst und der darin eingeschlossenen Überzeugungen aufgefasst. Es stellt keine einheitliche Größe dar, sondern ein differenziertes, vielschichtiges und teilweise hierarchisch geordnetes System mit unterschiedlichen Facetten, wie Studien zur Struktur des Selbstkonzepts zeigen (Harter, 1999; Wagner & Valtin, 2004). Akademische oder schulbezogene Selbstkonzepte gehören zu diesen Facetten und umfassen generalisierte fachspezifische Selbsteinschätzungen der eigenen bildungs- beziehungsweise schulbezogenen Fähigkeiten (Möller & Köller, 2004). Sie gelten als zentrale Komponenten der Schülerpersönlichkeit (Pekrun & Helmke, 1991; König, Wagner & Valtin, 2011) und entwickeln sich in erster Linie im Schulkontext, wo Kinder und Jugendliche regelmäßig Bewertungen ihrer Fähigkeiten erhalten, die sie als Grundlage für die Ausbildung generalisierter Selbstbewertungen verwenden können (Köller, 2004). Schulbezogene Selbstkonzepte können leistungsbezogenes Verhalten erklären und vorhersagen und, vermittelt über motivationale Variablen, Lernprozesse im jeweiligen Lernbereich fördern, wobei zwischen einem mathematischen und einem verbalen Selbstkonzept differenziert wird (Möller & Trautwein, 2015).

Das Leseselbstkonzept ist dem verbalen Selbstkonzept zuzuordnen und bereits von den ersten Schuljahren an von Bedeutung im Rahmen der Leseentwicklung. So zeigt eine Studie zum Lesen mit Grundschulkindern beispielsweise, dass das allgemeine schulische Selbstkonzept, das Leseselbstkonzept und die Leistungen im Lesen bereits kurz nach Schuleintritt miteinander zusammenhängen und sich die Zusammenhänge im Laufe der Grundschulzeit verstärken (Chapman & Tunmer, 1997; Chapman, Tunmer & Prochnow, 2000). In der Berliner Längsschnittstudie NOVARA wiesen vor allem Kinder mit schwachen Leistungen niedrige Selbstkonzepte im Lesen und in anderen Lernbereichen auf, dies galt auch unter Berücksichtigung von Noten beziehungsweise Verbalbeurteilung (Rosenfeld & Valtin, 1997; Wagner & Valtin, 2003). Kraayenoord und Schneider (1999) berichten für die Klassenstufen 3 und 4 positive Zusammenhänge zwischen Leseselbstkonzept, Leseinteresse, Lesestrategien und den Leseleistungen der Schülerinnen und Schüler, wobei sich gute Leserinnen und Leser von

schlechten durch höhere Ausprägungen in diesen Leistungskovariaten unterscheiden. Positive Zusammenhänge zwischen Leseselbstkonzept und Leseleistung am Ende der Grundschulzeit haben sich auch im Rahmen von IGLU (zuletzt: Bos, Bremerich-Vos, Tarelli & Valtin, 2012) und in der Studie KESS (Mielke, Goy & Pietsch, 2006) gezeigt. Ergebnisse auf Basis von Daten aus den Klassenstufen 4 und 5 weisen darauf hin, dass das Leseselbstkonzept positiv mit der intrinsischen Lesemotivation und der Lesemenge der Schülerinnen und Schüler zusammenhängt und nach den kognitiven Fähigkeiten der stärkste Prädiktor für das Leseverständnis der Schülerinnen und Schüler ist (Retelsdorf & Möller, 2008). Retelsdorf, Köller und Möller (2014) stellten sich wechselseitig bestärkende Effekte zwischen Leseleistung und Leseselbstkonzept von Klassenstufe 5 zu 6 fest. Die hohe Stabilität von Leseleistung und Leseselbstkonzept, die sich in dieser Studie von Klassenstufe 5 bis 9 fand, deutet den Autoren zufolge darauf hin, dass eine Förderung beider Bereiche bereits im Grundschulalter erfolgen sollte, wo sich die Merkmale noch nicht so stark verfestigt haben.

1.2 Lesemotivation

Motivation kann allgemein als zielgerichtete Verhaltensbereitschaft verstanden werden (Deci & Ryan, 1985), wobei Motivation die Zielrichtung, die Ausdauer und die Intensität des Verhaltens bedingt (Schiefele & Schaffner, 2015). In der Motivationsforschung wird zwischen gewohnheitsmäßiger (habituellem) und zeitlich begrenzter (aktueller) Motivation unterschieden. Unter habitueller Lesemotivation kann man entsprechend den stetigen oder wiederkehrenden Wunsch verstehen, etwas zu lesen. Demgegenüber bezeichnet aktuelle Lesemotivation den Wunsch, in einer bestimmten Situation einen bestimmten Text zu lesen (Möller & Schiefele, 2004). Weiterhin wird differenziert zwischen extrinsischer und intrinsischer Motivation. Unter extrinsischer Motivation wird die Absicht verstanden, eine Handlung auf Grund von angestrebten Folgen vorzunehmen, die nicht Teil der Handlung selbst sind. Bei der intrinsischen Motivation wird demgegenüber eine Handlung ausgeführt, weil sie aus sich heraus als interessant oder spannend wahrgenommen wird (Schiefele & Köller, 2010). Extrinsisch lesemotiviert sind Kinder beispielsweise, wenn sie lesen, um gute Noten zu bekommen oder um von den Eltern oder Lehrkräften gelobt zu werden. Die intrinsische Lesemotivation speist sich demgegenüber vor allem aus zwei Quellen: dem Interesse am Thema des Texts oder der Freude an der Tätigkeit des Lesens selbst (Möller & Schiefele, 2004).

In ihrer Expertise zu Leseförderung stellen Artelt et al. (2007) zusammenfassend dar, dass Lesemotivation zentrale Merkmale des Leseverhaltens (wie die Häufigkeit des Lesens und die Ausdauer beim Lesen, insbesondere bei schwierigen Texten) beeinflusst und auf diese Weise Lesekompetenz fördern kann. Insbesondere der positive Zusammenhang von intrinsischer Lesemotivation und Leseleistung ist gut belegt (im Überblick: Schiefele, Schaffner, Möller & Wigfield, 2012; für Ergebnisse am Ende der Grundschulzeit vgl. u. a. Bos et al., 2012; vertiefend: Valtin, Wagner & Schwippert, 2005). Längsschnittliche Studien dokumentieren für die Klassenstufen 2 und 3 (Schiefele, Stutz & Schaffner, 2016) sowie 3 und 4 (McElvany, Kortenbruck & Becker, 2008) wechselseitige Effekte zwischen der intrinsischen Lesemotivation und den Leseleistungen der Schülerinnen und Schüler.

Die Annahmen zu den Zusammenhängen von Selbstkonzept und Motivation mit Verhalten und Leistung lassen sich systematisch in Erwartungs-Wert-Modellen integrieren (Eccles et al., 1983; Wigfield, Rosenzweig & Eccles, 2017). In diesen Modellen werden leistungsbezogene Handlungen und Leistungsergebnisse in einem Fach oder Lernbereich (z.B. Lesen) durch Erwartungen für erfolgreiches Handeln („Ich kann gut lesen.“) und die Wertschätzung diesem Fach oder Lernbereich gegenüber („Ich lese gern.“) erklärt (Möller & Schiefele, 2004). Das bei IGLU erfasste Leseselbstkonzept bildet in diesem Modell die Erwartungskomponente ab, die intrinsische, habituelle Lesemotivation die Wertkomponente.

1.3 Leseverhalten

Den Wirkungsannahmen in Erwartungs-Wert-Modellen zu Determinanten der Lesekompetenz zufolge beeinflussen das Leseselbstkonzept und die intrinsische Lesemotivation das Leseverhalten (im Sinne der Lesemenge und der Lesehäufigkeit) positiv, was zu einem verbesserten Leseverständnis führen kann (Möller & Schiefele, 2004). Für einen Zusammenhang des Leseverhaltens mit den Leseleistungen von Schülerinnen und Schülern sprechen die Befunde zahlreicher Studien (im Überblick: Mol & Bus, 2011). Auch für eine vermittelnde Rolle des Leseverhaltens zwischen motivationalen Faktoren und den Leseleistungen finden sich einige empirische Belege: Bei acht- bis elfjährigen Kindern finden sich positive Zusammenhänge der außerschulischen Lesedauer sowohl mit der intrinsischen Lesemotivation als auch mit den Lesefähigkeiten (McGeown, Osborne, Warhurst, Norgate & Duncan, 2016). Längsschnittlich hat sich zudem gezeigt, dass die intrinsische Lesemotivation in der vierten Klassenstufe das Lesen in der Freizeit in Klasse 10 vorhersagt (Durik, Vida & Eccles, 2006). Eine Studie mit Kindern der Klassenstufen 2 und 3 weist einen vermittelnden Effekt von intrinsischer Lesemotivation auf die Leseleistung über die Lesemenge aus (Stutz, Schaffner & Schiefele, 2016). Für ein Modell sich wechselseitig bedingender Effekte zwischen Lesemotivation, Leseverhalten und Leseleistungen in den Klassenstufen 3 bis 6 sprechen die Befunde von McElvany et al. (2008). Pfof, Dörfler und Artelt (2010) untersuchten bei Kindern der Klassenstufen 3 bis 5 die Zusammenhänge zwischen außerschulischem Leseverhalten und Lesekompetenz und fanden differentielle Effekte nach dem Grad der Bildungsnähe der Elternhäuser: Während sich bei Kindern aus bildungsnahen Familien wechselseitige positive Effekte zwischen Lesekompetenz und außerschulischem Lesen zeigten, war dies bei Kindern aus bildungsfernen Familien nicht der Fall, was auf eine unterschiedliche Qualität des Leseverhaltens und der Leseaktivitäten in den Familien hinweisen kann.

Bei der Betrachtung dieser Zusammenhänge ist zu bedenken, dass sie nicht nur positive Entwicklungen verstärken, sondern gleichermaßen für negative Entwicklungen gelten können. Wenig selbstbewusste und motivierte Leserinnen und Leser lesen weniger und deshalb in Folge auch schlechter als ihre motivierten Mitschülerinnen und Mitschüler. Diese sich positiv und negativ verstärkenden Tendenzen, die zu einer Vergrößerung von Unterschieden führen, werden auch als ‚Matthäus-Effekt‘ bezeichnet (Stanovich, 1986; Rigney, 2010). Ergebnisse von Längsschnittanalysen (Pfof, Dörfler & Artelt, 2012) zeigen beispielsweise, dass leseschwache Schülerinnen und Schüler von Klassenstufe 3 zu 4 signifikant geringere Zuwächse in der Lesekompetenz hatten als ihre leistungsstärkeren

Mitschülerinnen und Mitschüler und dass die Unterschiede in der Lesekompetenzentwicklung der Kinder systematisch mit einem geringeren Umfang des Lesens zum Vergnügen zusammenhängen. Die negativen Wechselwirkungen werden in der Forschung zu lese-rechtschreib-schwachen Schülerinnen und Schülern auch als „Teufelskreis“ (Betz & Breuninger, 1998) beschrieben. Diesen in einen „Engelskreis“ sich positiv verstärkender Mechanismen zu verwandeln, ist eine zentrale pädagogische Herausforderung.

Das Leseverhalten der Schülerinnen und Schüler ist auch im Kontext eines sich insgesamt ändernden Medienverhaltens von Kindern zu betrachten. Aktuellen Befunden der KIM-Studie 2016 (*Kindheit, Internet, Medien*; Feierabend, Plankenhorn & Rathgeb, 2017) zufolge steht Lesen erst weit hinten auf der Liste der Aktivitäten, die von Kindern im Alter von 6 bis 13 Jahren jeden oder fast jeden Tag ausgeübt werden. An erster Stelle steht das Fernsehen (77%), gefolgt von Hausaufgaben/Lernen (70%), drinnen (53%) oder draußen (49%) spielen, Handy/Smartphone nutzen (42%) und Musikhören (35%). Das Lesen von Büchern folgt mit 15 Prozent erst an 13. Stelle. Weitere 9 Prozent der befragten Kinder lesen täglich oder fast täglich Comics, Zeitschriften oder Zeitungen. Nimmt man die Antwortkategorie „ein- oder mehrmals pro Woche“ hinzu, sind es 48 Prozent der Kinder, die regelmäßig in ihrer Freizeit Bücher lesen. Gefragt nach den drei Lieblingstätigkeiten in der Freizeit wird allerdings nur von 5 Prozent der Kinder das Lesen genannt.

1.4 Familiäre und individuelle Bedingungsfaktoren des Risikos schwacher Lesekompetenz

Selbstkonzept, Motivation und Verhaltensweisen stehen ihrerseits in Zusammenhang mit einer Reihe von familiären Merkmalen. Wie auch in erweiterten Erwartungs-Wert-Modellen berücksichtigt liegt eine zentrale Einflussquelle in dem soziokulturellen Milieu, in dem ein Kind aufwächst (Simpkins, Fredricks & Eccles, 2015). In Modellen zum Zusammenhang von familiärer Herkunft und Kompetenzerwerb wird angenommen, dass Strukturmerkmale von Familien (wie sozioökonomischer Status, Bildungsniveau oder Migrationshintergrund) familiäre Prozessmerkmale (wie kulturelle Aktivitäten oder die kommunikative Praxis in Familien) beeinflussen, die ihrerseits für die Ausprägungen individueller Merkmale von Kindern (wie die Motivation) und den Kompetenzerwerb bedeutsam sind (Baumert, Watermann & Schümer, 2003; für den Bereich der Lesekompetenz vgl. McElvany, 2011; Retelsdorf & Möller, 2008).

Der Zusammenhang von strukturellen Herkunftsmerkmalen und der Lesekompetenz von Schülerinnen und Schülern hat sich bei IGLU wiederholt gezeigt (siehe Kapitel 6 und 7 in diesem Band). Vertiefende Analysen auf Basis von Daten aus IGLU 2006 verweisen darauf, dass das relative Risiko, der Gruppe der Leseschwachen anzugehören, für Schülerinnen und Schüler aus bildungsfernen Elternhäusern und solche mit Migrationshintergrund gegenüber Kindern aus bildungsnahen Elternhäusern und Kindern ohne Migrationshintergrund signifikant und substantiell erhöht ist. Zugleich zeigen die Analysen, dass – unter Kontrolle einer Reihe von soziokulturellen, psychologischen und individuellen Merkmalen – für Kinder mit einem positiven Leseselbstkonzept das Risiko sinkt, der Gruppe der Leseschwachen anzugehören, und dies für Kinder mit und Kinder ohne Migrationshintergrund gleichermaßen der Fall ist (Valtin et al., 2010). In vertiefenden Analysen aus Daten von PISA 2000 hat sich das Merkmal der so-

zialen Herkunft bei fünfzehnjährigen Schülerinnen und Schülern ohne und mit Migrationshintergrund als bedeutsam für das Risiko erwiesen, der Gruppe der Leseschwachen anzugehören, wobei hohe Ausprägungen lesebezogener kognitiver Merkmale, wie Dekodierfähigkeit oder metakognitive Lernstrategien, dieses Risiko wiederum reduzierten (Stanat & Schneider, 2004).

In Bezug auf die familiäre Lesesozialisation (Hurrelmann, 2004; Sénéchal, 2012) als Prozessmerkmal zeigen Studien, dass motivationale und verhaltensbezogene Variablen als Prädiktoren von Lesekompetenz mit leseförderlichen Ressourcen und Aktivitäten in den Elternhäusern zusammenhängen. Hierbei sind lesebezogene Merkmale und Verhaltensweisen relevant, beispielsweise gemeinsame sprachliche Aktivitäten von Eltern und Kindern oder die Einstellungen der Eltern zum Lesen (McElvany, Becker & Lüdtke, 2009; Retelsdorf & Möller, 2008; Schaffner, Schiefele & Schmidt, 2013). Die Ausprägungen dieser Zusammenhänge können ihrerseits wiederum von strukturellen Herkunftsmerkmalen der Familien wie den oben benannten bedingt sein (ebd.). Lesespezifische Prozessmerkmale in den Familien und deren Verhältnis zur Lesekompetenz von Schülerinnen und Schülern wurden auch in Analysen auf Basis von Daten aus IGLU betrachtet (Buddeberg, Stubbe & Potthoff, 2008; Stubbe, Buddeberg, Hornberg & McElvany, 2007; Tarelli, 2010), in denen der Fokus auf der Bildung eines Index zur Lesesozialisation und der Analyse der medienrollen Rolle von Lesesozialisation zwischen sozialer Herkunft und Lesekompetenz zur besseren Erklärung primärer Herkunftseffekte lag. Hierbei zeigten sich enge Zusammenhänge zwischen der häuslichen Lesesozialisation und der Lesekompetenz beziehungsweise zwischen der sozialen Herkunft und der häuslichen Lesesozialisation, wobei beide Zusammenhänge in Deutschland – im internationalen Vergleich gesehen – besonders deutlich ausgeprägt sind.

2 Ergebnisse

Die Ergebnisdarstellung in den Abschnitten 2.1 bis 2.3 hat folgenden Aufbau: Für das Leseselbstkonzept (Abschnitt 2.1), die Lesemotivation (Abschnitt 2.2) und das Leseverhalten (Abschnitt 2.3) werden jeweils zunächst Mittelwerte und prozentuale Verteilungen berichtet und im internationalen Vergleich dargestellt. Anschließend werden jeweils auf nationaler Ebene Zusammenhänge der Variablen mit der Leseleistung der Schülerinnen und Schüler fokussiert. Je nach Datenlage werden die Befunde aus 2016 zudem mit denen aus IGLU 2001, 2006 beziehungsweise 2011 verglichen. Beim Leseselbstkonzept und der Lesemotivation wird auf den internationalen Vergleich verzichtet. Es kann vermutet werden, dass sich in diesen Skalen, die auf Selbstberichten beruhen, kulturelle Unterschiede in den Antworttendenzen oder unterschiedliche Bezugsnormen in den Teilnehmerstaaten und -regionen spiegeln, die internationale Vergleiche erschweren. In Abschnitt 2.4 werden dann familiäre und individuelle Bedingungsfaktoren des Risikos schwacher Lesekompetenz regressionsanalytisch betrachtet.

2.1 Leseselbstkonzept

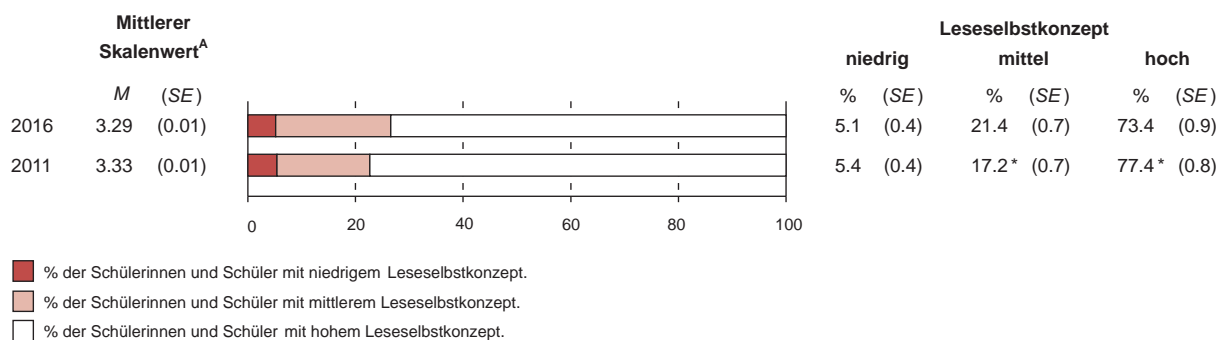
Für die Erfassung des Leseselbstkonzepts wurden den Kindern im Schülerfragebogen Aussagen vorgelegt, zu denen sie gefragt wurden, ob sie ihnen *stark zustimmen*, *einigermaßen zustimmen*, *wenig zustimmen* oder *überhaupt nicht zustimmen*. Folgende Aussagen wurden hierbei herangezogen:

1. Normalerweise bin ich gut im Lesen.
2. Lesen fällt mir sehr leicht.
3. Es fällt mir schwer, Geschichten mit schwierigen Wörtern zu lesen.
4. Lesen fällt mir schwerer als vielen meiner Mitschüler.
5. Lesen fällt mir schwerer als alle anderen Fächer.

Für die Darstellung der Ergebnisse werden die Angaben der Schülerinnen und Schüler bei den Aussagen 1 und 2 mit 1 (*stimme überhaupt nicht zu*) bis 4 (*stimme stark zu*) Punkten kodiert, bei den Aussagen 3, 4 und 5 in umgekehrter Reihenfolge. Anschließend wird für alle Schülerinnen und Schüler der Mittelwert der Angaben gebildet. Die interne Konsistenz der Skala liegt bei einem Cronbachs Alpha von $\alpha = .77$ (in 2016; 2011: $\alpha = .77$). Ergänzend werden Gruppen mit niedrig, mittel und hoch positivem Selbstkonzept gebildet. Kinder mit einem Mittelwert (M) von $1 \leq M < 2$ werden der Kategorie ‚niedrig‘ zugeordnet, Kinder mit einem Mittelwert von $2 \leq M < 3$ der Kategorie ‚mittel‘ und Kinder mit einem Mittelwert von $3 \leq M \leq 4$ der Kategorie ‚hoch‘. Abbildung 4.1 ist zu entnehmen, dass 2016 das Leseselbstkonzept von Schülerinnen und Schülern in Deutschland mit einem Skalenmittelwert von $M = 3.29$ positiv ausgeprägt ist. Die Verteilung auf die Kategorien zeigt, dass 73.4 Prozent der Schülerinnen und Schüler über ein hohes Leseselbstkonzept verfügen. Ein Selbstkonzept mittlerer Höhe liegt bei 21.4 Prozent der Schülerinnen und Schüler vor, ein niedriges Selbstkonzept bei 5.1 Prozent.

Die Abbildung enthält ebenfalls die Werte zum Leseselbstkonzept aus der Erhebung von 2011, bei der den Kindern die identischen Aussagen vorgelegt wurden. Da von Seiten der internationalen Studienleitung die Erhebungsinstrumente im Vergleich zu 2006 und 2001 geändert wurden, können hier für 2006 und 2001 keine Ergebnisse im direkten Vergleich berichtet werden. Vergleicht man die mittleren Skalenwerte im Leseselbstkonzept aus den Jahren 2011 und 2016, so

Abbildung 4.1: Leseselbstkonzept von Schülerinnen und Schülern sowie prozentuale Verteilung nach niedrigem, mittlerem und hohem Selbstkonzept in Deutschland – IGLU 2011 und 2016 im Vergleich



A = Die Skala umfasst fünf Fragen (z.B. Lesen fällt mir sehr leicht.) mit vierstufigem Antwortformat (1 = Stimme überhaupt nicht zu, ..., 4 = Stimme stark zu).

* = Unterschied zu 2016 statistisch signifikant ($p < .05$).

finden sich keine signifikanten Änderungen. Eine differenzierte Betrachtung zeigt jedoch Unterschiede im Anteil der Kinder mit niedrigem, mittlerem und hohem Leseselbstkonzept. 2016 weisen weniger Kinder (73.4 %) ein hohes Selbstkonzept auf als im Jahr 2011 (77.4 %). Zugleich erhöht sich der Anteil an Schülerinnen und Schülern mit einem Leseselbstkonzept in mittlerer Höhe von 17.2 auf 21.4 Prozent. Mit jeweils etwa 5 Prozent sind die Anteile der Schülerinnen und Schüler mit einem niedrigen Selbstkonzept nahezu gleichgeblieben. Der Zusammenhang zwischen Leseselbstkonzept und Leseleistung fällt erwartungsgemäß positiv aus. Die Korrelation zwischen den mittleren Leistungswerten auf der Gesamtskala Lesen und der z-standardisierten Skala ‚Leseselbstkonzept‘ beträgt $r = .41$ ($p < .05$).

Im Vergleich der Befunde von 2011 und 2016 nach Leseleistungen (ohne Abbildung) zeigt sich ein nicht signifikant veränderter, hoher Anteil an Kindern mit hohen Leistungswerten (auf Kompetenzstufe IV und V) und positiven Selbstkonzepten (2011: 88.8 %; 2016: 87.1 %). Hingegen ist der Anteil an Schülerinnen und Schülern mit niedrigen Leistungen (Kompetenzstufen I und II) und hohen Selbstkonzepten signifikant gesunken (2011: 51.6 %; 2016: 43.9 %), was auch bei den Kindern mit mittleren Leistungen (Kompetenzstufe III) und hohen Selbstkonzepten der Fall ist (2011: 72.5 %; 2016: 67.5 %). Im Vergleich zu 2011 fallen die Selbsteinschätzungen in 2016 somit etwas realistischer aus.

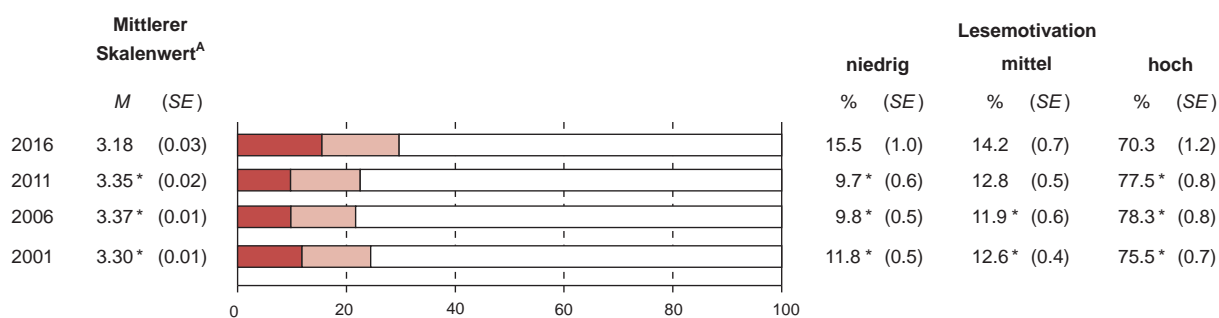
2.2 Lesemotivation

Wie beim Vorgehen für die Erfassung des Leseselbstkonzepts wurden den Kindern auch hinsichtlich der Lesemotivation im Schülerfragebogen Aussagen vorgelegt und der Grad der Zustimmung erfasst. Folgende Aussagen, die allen vier Erhebungszyklen von IGLU verwendet wurden, wurden hierbei herangezogen:

1. Ich würde mich freuen, wenn mir jemand ein Buch schenken würde.
2. Ich finde Lesen langweilig.
3. Ich lese gerne.

Für die Darstellung der Ergebnisse werden die Angaben der Schülerinnen und Schüler zu den Aussagen 1 und 3 mit 1 (*stimme überhaupt nicht zu*) bis 4 Punkten (*stimme stark zu*) kodiert, zu Aussage 2 in umgekehrter Reihenfolge. Aus diesen drei Aussagen wird für jedes Kind der Mittelwert berechnet. Der ermittelte Skalenwert gibt an, inwieweit das jeweilige Kind hoch oder gering lesemotiviert ist. Die interne Konsistenz der Skala liegt bei einem Cronbachs Alpha von $\alpha = .82$. In den vorangehenden Erhebungszyklen ergaben sich für die Skala ähnliche Reliabilitäten (2011: $\alpha = .80$; 2006: $\alpha = .79$; 2001: $\alpha = .77$). Auf Basis dieser Skala werden Gruppen von Schülerinnen und Schülern mit niedriger, mittlerer und hoher Lesemotivation gebildet. Kinder mit einem Mittelwert (M) von $1 \leq M < 2$ werden als ‚niedrig‘ lesemotiviert klassifiziert, bei einem Mittelwert von $2 \leq M < 3$ werden sie als ‚mittel‘ und bei einem Mittelwert von $3 \leq M \leq 4$ als ‚hoch‘ lesemotiviert klassifiziert. In Abbildung 4.2 sind die Ergebnisse im Vergleich der vier Erhebungszyklen von IGLU dargestellt.

Der mittlere Skalenwert von 3.18 im Jahr 2016 verweist bei einer vierstufigen Skala auf eine insgesamt hohe Lesemotivation. Aus der Verteilung auf die Kategorien ist ersichtlich, dass rund 70 Prozent der Schülerinnen und Schüler über eine hohe Lesemotivation verfügen. Eine mittlere Lesemotivation brin-

Abbildung 4.2: Lesemotivation von Schülerinnen und Schülern sowie prozentuale Verteilungen nach niedrigerer, mittlerer und hoher Motivation in Deutschland – IGLU 2001, 2006, 2011 und 2016 im Vergleich¹

■ % der Schülerinnen und Schüler mit niedriger Lesemotivation.

■ % der Schülerinnen und Schüler mit mittlerer Lesemotivation.

□ % der Schülerinnen und Schüler mit hoher Lesemotivation.

A = Die Skala umfasst drei Fragen (z.B. Ich lese gern.) mit vierstufigem Antwortformat (1 = Stimme überhaupt nicht zu, ..., 4 = Stimme stark zu).

* = Unterschied zu 2016 statistisch signifikant ($p < .05$).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

gen rund 14 Prozent und eine niedrige Lesemotivation rund 16 Prozent zum Ausdruck.

Im Trend betrachtet zeigt sich eine signifikante, leichte Abnahme des mittleren Skalenwertes von 3.30 im Jahr 2001 zu 3.18 im Jahr 2016. Beim Vergleich der Jahre 2011 und 2016 ist auffallend, dass sich der Anteil der Schülerinnen und Schüler mit hoher Lesemotivation signifikant um rund 7 Prozentpunkte verringert und gleichzeitig der Anteil der Schülerinnen und Schüler mit niedriger Lesemotivation um rund 6 Prozentpunkte erhöht hat. Bei den Erhebungen in den Jahren 2006 und 2011 hatten sich gegenüber der ersten Erhebung im Jahr 2001 leicht positivere Werte in der Lesemotivation gezeigt. Diese Entwicklung hat sich jedoch nicht fortgesetzt – 2016 hat nun fast jedes sechste Kind am Ende der vierten Jahrgangsstufe eine niedrige Lesemotivation.

Die Lesemotivation steht erwartungsgemäß in Zusammenhang mit der Leseleistung. 2016 beträgt die Korrelation zwischen den mittleren Leistungswerten auf der Gesamtskala Lesen und der z-standardisierten Skala ‚Lesemotivation‘ $r = .33$ ($p < .05$). Eine Betrachtung der Ergebnisse nach Leseleistungen im IGLU-Trend von 2001 bis 2016 (ohne Abbildung) zeigt differentielle, signifikante Entwicklungen für unterschiedliche Leistungsgruppen: Während sich auf den Kompetenzstufen IV und V der Anteil an niedrig lesemotivierten Kindern um rund drei Prozentpunkte erhöht (2001: 5.3%; 2016: 8.4%), erhöht sich der Anteil auf den Kompetenzstufen I und II um rund 11 Prozentpunkte (2001: 18.6%; 2016: 29.6%). Diese Tendenz zeigt sich noch deutlicher beim Anteil an hoch lesemotivierten – hier verringert sich der Anteil bei den Lesestarken um rund 5 Prozentpunkte (2001: 87.2%; 2016: 82.2%), bei den Leseschwachen hingegen um rund 15 Prozentpunkte (2001: 62.6%; 2016: 47.5%). Damit lässt sich festhalten, dass die Abnahme der Lesemotivation von Viertklässlerinnen und Viertklässlern in Deutschland vor allem auf eine geringere Motivation der leistungsschwächeren Schülerinnen und Schüler zurückzuführen ist.

1 Aufgrund von Änderungen in den Items zur Lesemotivation von 2011 zu 2016 umfasst die im IGLU-2011-Band berichtete Skala (Bos et al., S. 124) mehr und andere Items. Entsprechend unterscheiden sich die mittleren Skalenwerte und die prozentualen Verteilungen.

2.3 Leseverhalten

Wie einleitend dargestellt, ist anzunehmen, dass das Leseverhalten eine entscheidende Gelenkstelle zwischen motivationalen Merkmalen und der Leseleistung darstellt. Entsprechend werden den Kindern im IGLU-Schülerfragebogen auch Fragen zu ihrem Leseverhalten gestellt. Betrachtet werden im Folgenden drei Aspekte: das Lesen zum Vergnügen außerhalb der Schule, die Dauer des täglichen Lesens in der Freizeit sowie die Bibliotheksnutzung, um Bücher zu entleihen.

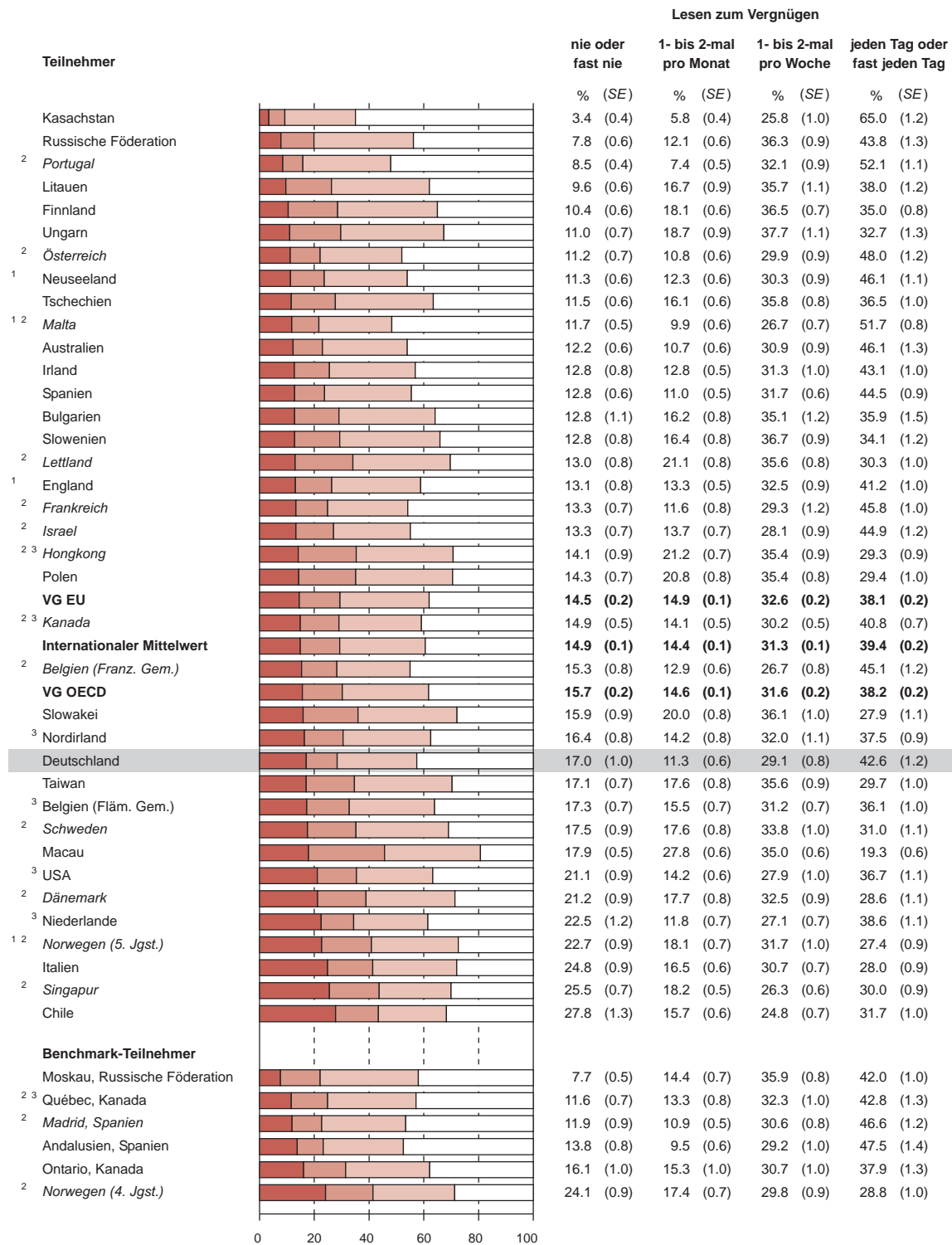
2.3.1 Lesen zum Vergnügen

In allen IGLU-Erhebungen wurden die Schülerinnen und Schüler gefragt, wie häufig sie außerhalb der Schule zu ihrem Vergnügen lesen. Abbildung 4.3 zeigt für IGLU 2016 die Verteilungen der Antworten der Schülerinnen und Schüler im internationalen Vergleich. Die Abbildung ist nach dem Anteil der Schülerinnen und Schüler geordnet, die angeben, nie oder fast nie zum Vergnügen zu lesen.

In Deutschland geben 17 Prozent der Schülerinnen und Schüler an, nie oder fast nie außerhalb der Schule zum Vergnügen zu lesen. Jeden Tag oder fast jeden Tag tun dies nach eigenen Angaben rund 43 Prozent. Die übrigen Schülerinnen und Schüler lesen ein- bis zweimal pro Woche (29.1 %) oder ein- bis zweimal pro Monat (11.3 %) zu ihrem Vergnügen. Im internationalen Vergleich zeigt sich, dass der Anteil an Schülerinnen und Schülern, die nie oder fast nie zu ihrem Vergnügen lesen, in Deutschland vergleichsweise hoch ausfällt. Nur bei sieben Teilnehmern (USA, Dänemark, Niederlande, Norwegen, Italien, Singapur und Chile) sowie Norwegen (4. Jahrgangsstufe) als Benchmark-Teilnehmer fallen die Anteile signifikant höher aus, während in der Mehrzahl der hier betrachteten Bildungssysteme sowie im internationalen Mittel und im Mittel der VG EU signifikant niedrigere Werte zu verzeichnen sind. Zugleich ist in Abbildung 4.3 aber auch ersichtlich, dass der Anteil an Schülerinnen und Schülern, die jeden Tag oder fast jeden Tag zu ihrem Vergnügen lesen, in Deutschland mit 42.6 Prozent ebenfalls vergleichsweise hoch ist – hier liegen die Anteile ebenfalls nur in wenigen Bildungssystemen signifikant höher (in Kasachstan, Portugal, Österreich, Neuseeland, Malta, Australien und Frankreich sowie in den beiden spanischen Benchmark-Teilnehmern Madrid und Andalusien). In den meisten Bildungssystemen sind die Anteile der täglich zum Vergnügen Lesenden signifikant niedriger, was sich auch in den drei betrachteten Mittelwerten (international, EU, OECD) abbildet.

Abbildungen 4.4 und 4.5 veranschaulichen die Angaben zum Leseverhalten der Schülerinnen und Schüler, die nie oder fast nie beziehungsweise jeden Tag oder fast jeden Tag zum Vergnügen lesen, im Trendvergleich von 2001 bis 2016. Im tabellarischen Teil der Abbildungen sind für die Erhebungszyklen die jeweiligen Prozentanteile der Schülerinnen und Schüler ausgewiesen sowie die Veränderungen in den Prozentwerten aufgeführt. Im grafischen Teil der Abbildung sind die Veränderungen zwischen 2001 und 2016 beziehungsweise 2006 und 2016 visualisiert. Für diejenigen Teilnehmerstaaten und -regionen, die nicht an IGLU 2001 teilgenommen haben, wird der Veränderungswert zu 2006 abgebildet.

Abbildung 4.3: Lesen zum Vergnügen im internationalen Vergleich – IGLU 2016 (Anteile der Schülerinnen und Schüler in Prozent)



- % der Schülerinnen und Schüler, die nie oder fast nie außerhalb der Schule zum Vergnügen lesen.
- % der Schülerinnen und Schüler, die 1- bis 2-mal pro Monat außerhalb der Schule zum Vergnügen lesen.
- % der Schülerinnen und Schüler, die 1- bis 2-mal pro Woche außerhalb der Schule zum Vergnügen lesen.
- % der Schülerinnen und Schüler, die jeden Tag oder fast jeden Tag außerhalb der Schule zum Vergnügen lesen.

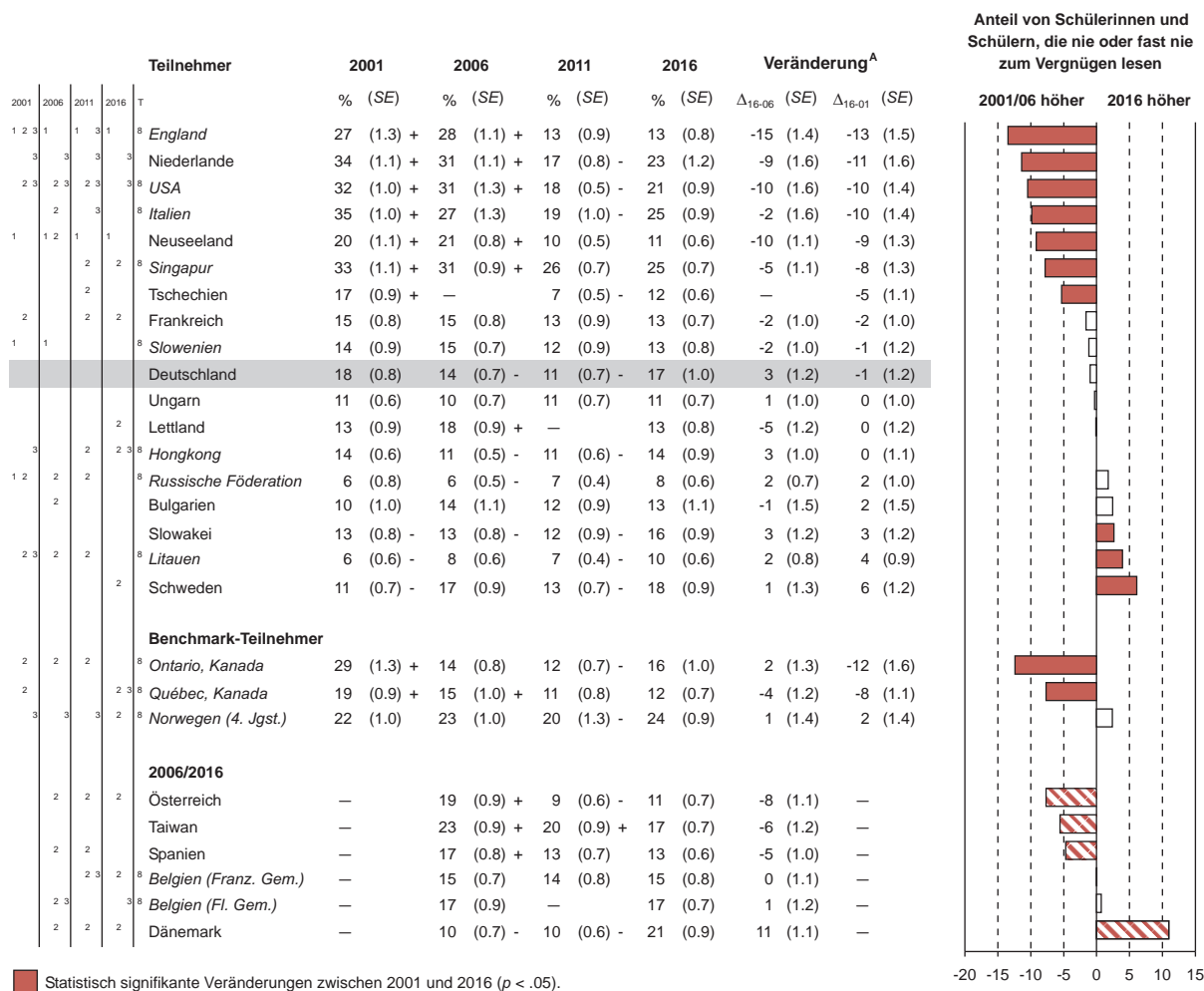
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

Abbildung 4.4: Schülerinnen und Schüler, die angeben, außerhalb der Schule nie oder fast nie zum Vergnügen zu lesen, im internationalen Vergleich – IGLU 2001, 2006, 2011 und 2016 im Vergleich



■ Statistisch signifikante Veränderungen zwischen 2001 und 2016 ($p < .05$).

▨ Statistisch signifikante Veränderungen zwischen 2006 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerseite erreichen nicht die internationalen Vorgaben.

8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

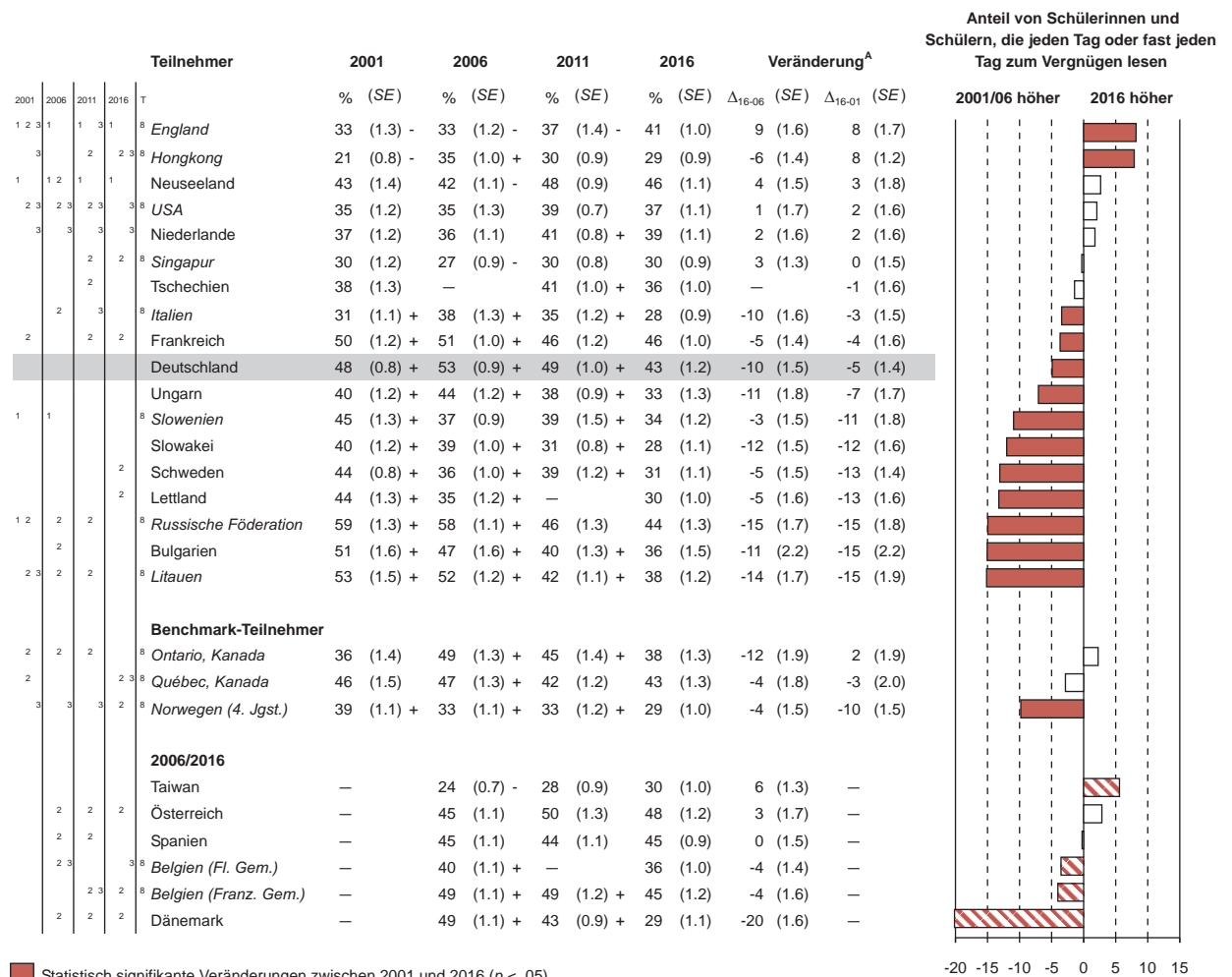
+ = Anteil von Schülerinnen und Schülern, die nie oder fast nie zum Vergnügen lesen, 2001 bzw. 2006 bzw. 2011 statistisch signifikant höher als 2016 ($p < .05$).

- = Anteil von Schülerinnen und Schülern, die nie oder fast nie zum Vergnügen lesen, 2001 bzw. 2006 bzw. 2011 statistisch signifikant niedriger als 2016 ($p < .05$).

$\Delta_{16-01} / \Delta_{16-06}$ = Differenz in den Prozentwerten zwischen 2016 und 2001 bzw. 2006 (bei denjenigen Teilnehmern, die nicht an IGLU 2001 teilgenommen haben).

Für Deutschland zeigt der Trendvergleich, dass sich der Anteil an Kindern, die nie oder fast nie zu ihrem Vergnügen lesen, von 2001 bis 2016 nicht signifikant verändert hat. International ist dies in acht weiteren Bildungssystemen der Fall. In diesem Zeitraum sind in neun Bildungssystemen signifikante Abnahmen dieses Anteils zu beobachten, wobei die Differenzen in England (13 Prozentpunkte), Ontario (12 Punkte), den Niederlanden (11 Punkte), den USA (10 Punkte) und Italien (10 Punkte) am höchsten ausfallen. Eine signifikante Zunahme an ‚Nichtlesern‘ liegt in drei Bildungssystemen vor, wobei sich in Schweden mit 6 Punkten die größte Zunahme zeigt. Die Ergebnisse für Dänemark fallen besonders ins Auge: Hier hat sich von 2006 zu 2016 der Anteil der Kinder, die nie oder fast nie zum Vergnügen lesen, um 11 Prozentpunkte erhöht. In keinem anderen Land ist die Veränderung so deutlich.

Abbildung 4.5: Schülerinnen und Schüler, die angeben, außerhalb der Schule jeden Tag oder fast jeden Tag zum Vergnügen zu lesen, im internationalen Vergleich – IGLU 2001, 2006, 2011 und 2016



■ Statistisch signifikante Veränderungen zwischen 2001 und 2016 ($p < .05$).

▨ Statistisch signifikante Veränderungen zwischen 2006 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

+ = Anteil von Schülerinnen und Schülern, die jeden Tag oder fast jeden Tag zum Vergnügen lesen, 2001 bzw. 2006 bzw. 2011 statistisch signifikant höher als 2016 ($p < .05$).

- = Anteil von Schülerinnen und Schülern, die jeden Tag oder fast jeden Tag zum Vergnügen lesen, 2001 bzw. 2006 bzw. 2011 statistisch signifikant niedriger als 2016 ($p < .05$).

$\Delta_{16-01} / \Delta_{16-06}$ = Differenz in den Prozentwerten zwischen 2016 und 2001 bzw. 2006 (bei denjenigen Teilnehmern, die nicht an IGLU 2001 teilgenommen haben).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Während in Deutschland der Anteil an Kindern, die nie oder fast nie zu ihrem Vergnügen lesen, im Vergleich von 2001 zu 2016 stabil geblieben ist, hat sich der Anteil an Kindern, die jeden Tag oder fast jeden Tag zu ihrem Vergnügen lesen, im selben Zeitraum signifikant um 5 Prozentpunkte verringert, wie in Abbildung 4.5 illustriert ist. Auch in den meisten weiteren der in dieser Abbildung betrachteten Bildungssysteme zeigen sich Abnahmen des Anteils an ‚Vielleasern‘. In insgesamt acht Bildungssystemen liegt die Veränderung von 2001 bis 2016 bei 10 Prozentpunkten oder mehr, unter diesen zeigen sich die größten Veränderungen in Litauen, Bulgarien und Russische Föderation (jeweils 15 Punkte). In Dänemark zeigt sich im Vergleich von 2006 zu 2016 eine deutliche Abnahme des Anteils an Kindern, die täglich oder fast täglich zum Vergnügen lesen, um 20 Prozentpunkte. Nur in wenigen Teilnehmerstaaten ist der Anteil an Kindern,

die täglich oder fast täglich zum Vergnügen lesen, im Trend signifikant gestiegen, und zwar in England und Hongkong (von 2001 bis 2016) sowie in Taiwan, England, Singapur und Neuseeland (von 2006 bis 2016).

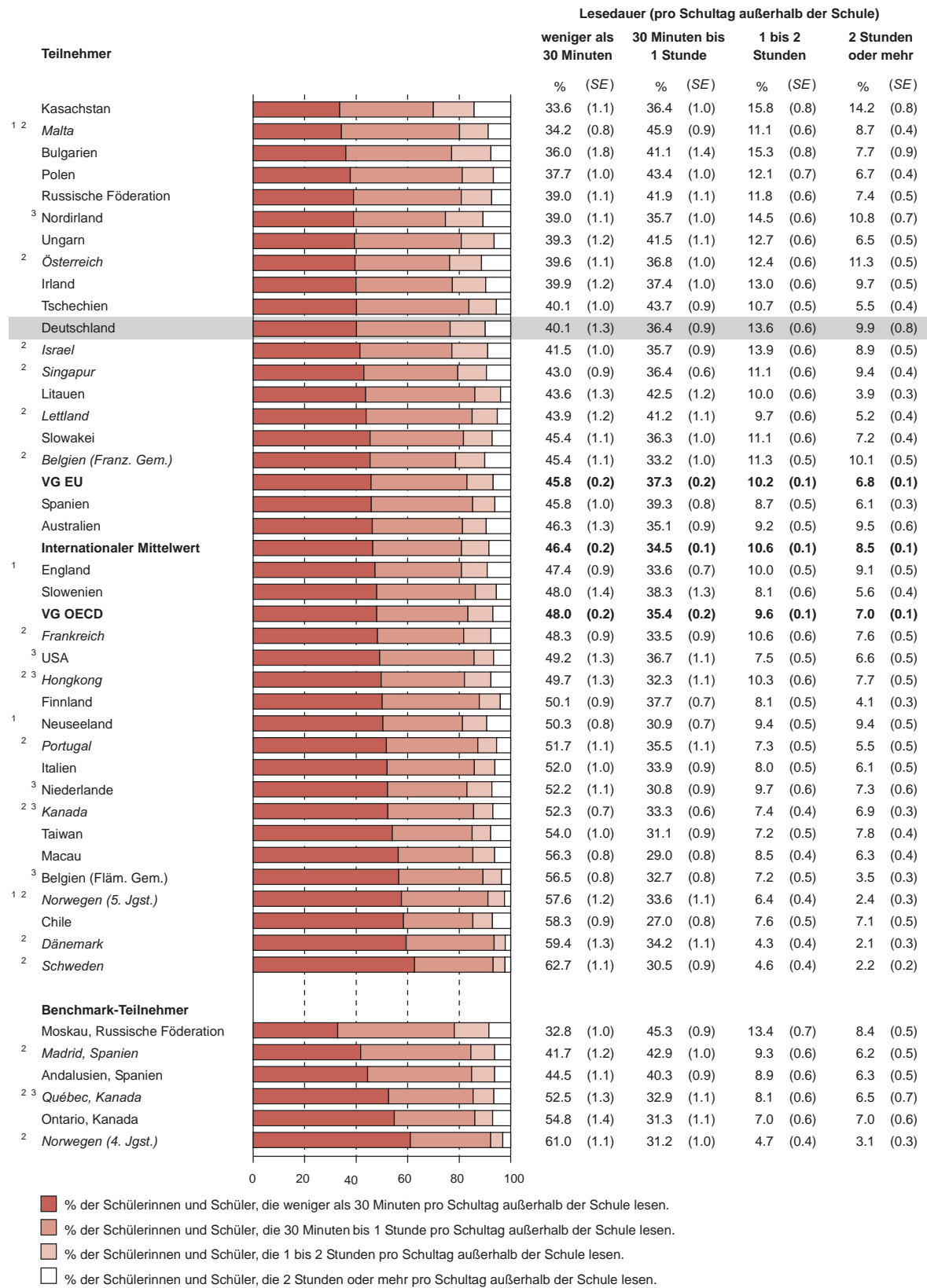
Zwischen dem Lesen zum Vergnügen und der Leseleistung zeigt sich erwartungsgemäß ein signifikanter, positiver Zusammenhang in der Höhe von $r = .32$ ($p < .05$). Signifikante Zusammenhänge ($p < .05$) des Lesens zum Vergnügen zeigen sich ebenfalls mit der Lesemotivation ($r = .66$), dem Leseselbstkonzept ($r = .29$), der Dauer des täglichen Lesens in der Freizeit (siehe Abschnitt 2.3.2; $r = .43$) und der Häufigkeit der Buchausleihe aus der Schulbibliothek oder der örtlichen Bibliothek (siehe Abschnitt 2.3.3; $r = .27$). Differenziert man die Ergebnisse zum außerschulischen Lesen zum Vergnügen in Deutschland nach Leseleistungen im IGLU-Trend von 2001 bis 2016 (ohne Abbildung), lässt sich feststellen, dass sich in allen drei Leseleistungsgruppen (Kompetenzstufen I und II, III sowie IV und V) der Anteil der Kinder, die täglich oder fast täglich zum Vergnügen lesen, signifikant verringert hat. Ähnlich wie bei der Lesemotivation ist auch beim Merkmal ‚Lesen zum Vergnügen‘ die vergleichsweise ungünstigste Entwicklung bei den Schülerinnen und Schülern mit geringen Leseleistungen zu verzeichnen: In dieser Gruppe findet sich die größte (signifikante) Abnahme an Kindern, die täglich oder fast täglich zu ihrem Vergnügen lesen (2001: 35.3%; 2016: 25.7%) und die größte (signifikante) Zunahme des Anteils an Kindern, die dies nie oder fast nie tun (2001: 25.6%; 2016: 33.9%).

Insgesamt zeigen die Trendvergleiche für Deutschland keine deutlichen Veränderungen der Anteile an Kindern, die entweder sehr häufig oder sehr selten bis gar nicht zu ihrem Vergnügen lesen. Auch im internationalen Vergleich lässt sich für Deutschland keine auffällige Entwicklung feststellen. Bei einer Reihe von Teilnehmerstaaten lassen sich aber durchaus positive Veränderungen erkennen (im Sinne einer Verringerung des Anteils der Nichtleser und der Erhöhung des Anteils an Viellesern). Hierbei ist allerdings auch zu beachten, dass in einigen dieser Bildungssysteme die Anteile wenig beziehungsweise viel lesender Kinder zu Beginn des Trends ungünstiger ausgeprägt waren als in Deutschland.

2.3.2 Lesedauer

In Ergänzung zu der – im Hinblick auf die zeitliche Intensität eher allgemeinen – Einschätzung des Lesens zum Vergnügen wird im IGLU-Schülerfragebogen auch die Frage nach der täglichen Lesedauer gestellt (*Wie viel Zeit verbringst du außerhalb der Schule an einem normalen Schultag mit Lesen?*). Die prozentualen Verteilungen der Antworten im internationalen Vergleich von IGLU 2016 sind in Abbildung 4.6 aufgeführt. In Deutschland lesen 40.1 Prozent der Kinder weniger als 30 Minuten pro Tag, 36.4 Prozent tun dies 30 Minuten bis zu einer Stunde, 13.6 Prozent ein bis zwei Stunden und 9.9 Prozent 2 Stunden oder mehr. Die Mehrheit der Kinder in Deutschland liest täglich somit zumindest eine halbe bis eine Stunde außerhalb der Schule. Im internationalen Vergleich zeigt sich, dass der Anteil an Kindern, die weniger als 30 Minuten pro Tag lesen, in Deutschland vergleichsweise gering ausfällt. Nur bei zwei Teilnehmern (Kasachstan mit 33.6% und Malta mit 34.2%) sowie einem Benchmark-Teilnehmer (Moskau mit 32.8%) fällt der Anteil signifikant geringer aus. Demgegenüber liegt er bei 24 Teilnehmern und 4 Benchmark-Teilnehmern – und damit in der Mehrzahl der hier betrachteten 43 Bildungssysteme – sowie im internationalen Mittelwert und im Mittel der Teilnehmer aus der VG EU und VG OECD signifikant höher,

Abbildung 4.6: Lesedauer (Freizeit, täglich) im internationalen Vergleich – IGLU 2016 (Anteile der Schülerinnen und Schüler in Prozent)



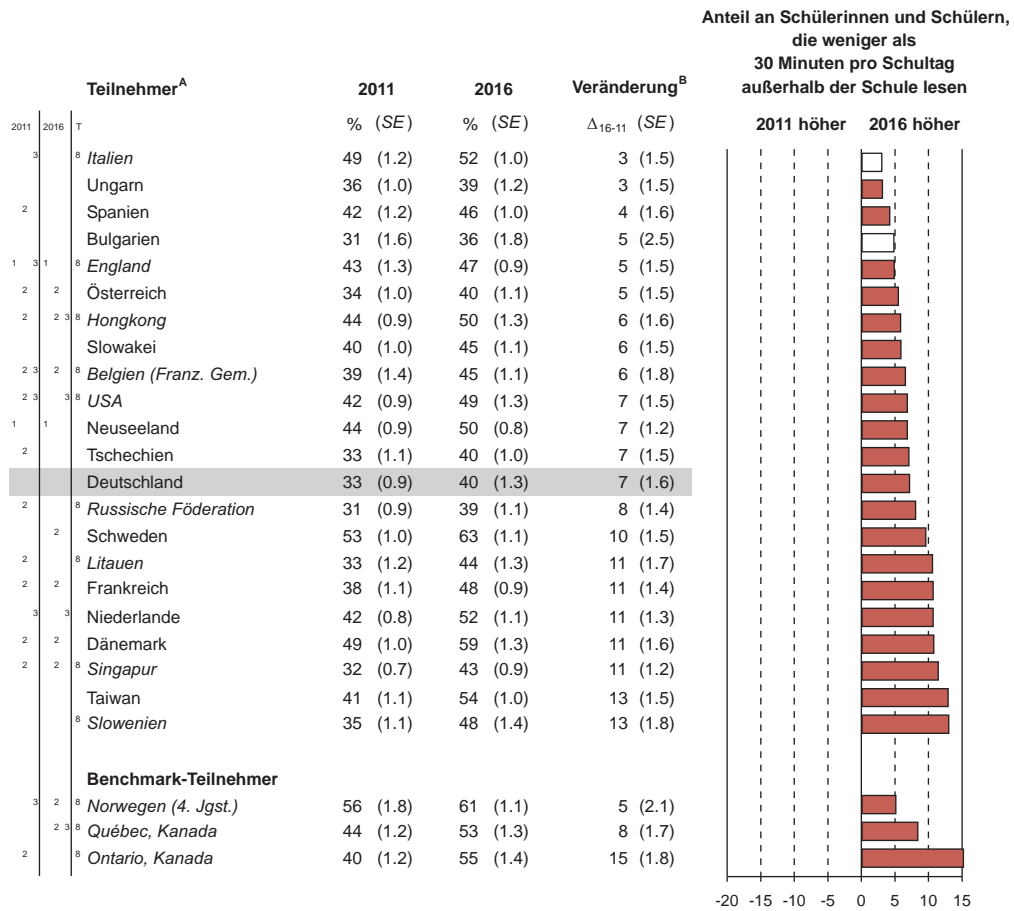
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

Abbildung 4.7: Schülerinnen und Schüler, die angeben, täglich weniger als 30 Minuten außerhalb der Schule zu lesen, im internationalen Vergleich – IGLU 2011 und 2016 im Vergleich



■ Statistisch signifikante Veränderungen zwischen 2011 und 2016 ($p < .05$).

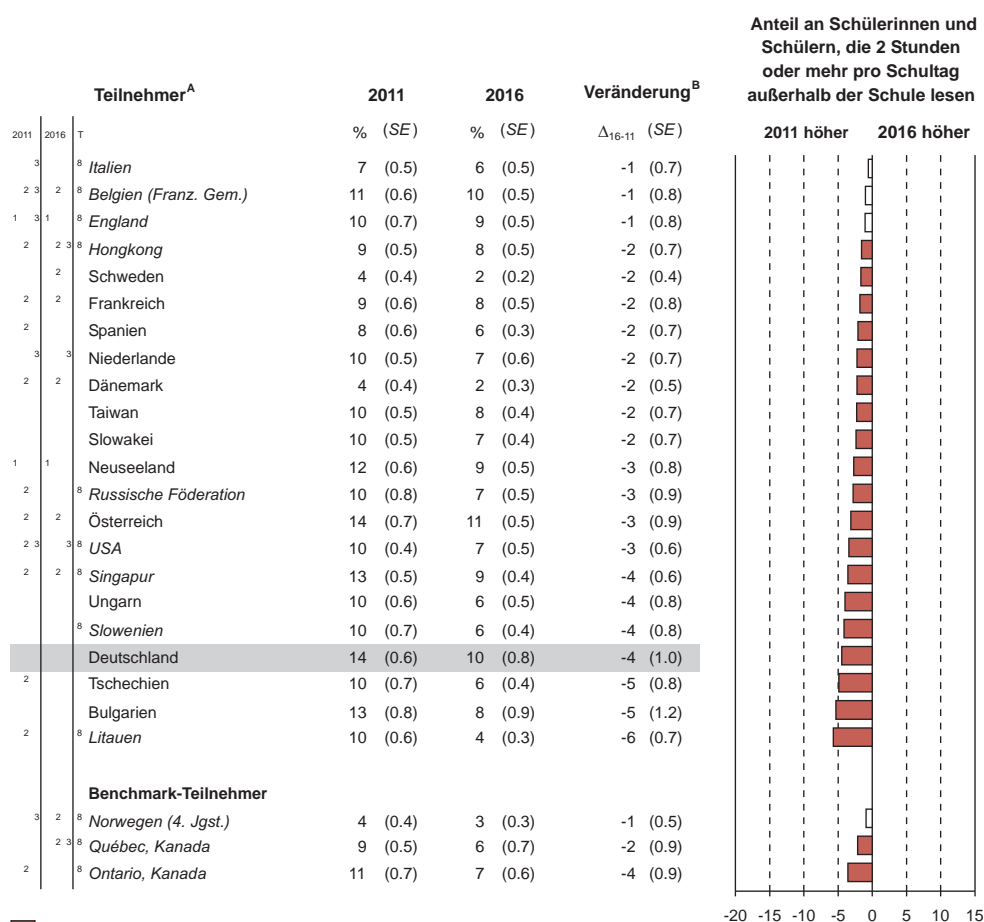
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

- 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
- 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
- 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
- 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.
- A = Da Belgien (Fl. Gem.) und Lettland nicht an IGLU 2011 teilgenommen haben, können hier für diese Teilnehmer keine Ergebnisse berichtet werden.
- B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.
- Δ_{16-11} = Differenz in den Prozentwerten zwischen 2016 und 2011.

unter anderem in den skandinavischen Staaten Schweden (62.7%), Dänemark (59.4%) und Norwegen (5. Jahrgangsstufe, 57.6%), aber auch in anderen europäischen Teilnehmern wie der Flämischen Gemeinschaft in Belgien (56.5%), den Niederlanden (52.2%) oder Italien (52.0%). Betrachtet man den Anteil an Kindern, die täglich zwei Stunden oder mehr lesen, zeigt sich für Deutschland ein im internationalen Vergleich hoher Anteil an Viellesern. Nur in Kasachstan liegt der Wert mit 14.2 Prozent signifikant höher als in Deutschland, signifikant geringer fällt der Anteil hingegen bei 24 Teilnehmern und 5 Benchmark-Teilnehmern sowie im Mittel der Staaten der Vergleichsgruppen EU und OECD aus. In einigen europäischen Ländern liegt der Anteil an Viellesern dabei deutlich unter dem Wert für Deutschland, so beispielsweise in Dänemark (2.1%), Schweden (2.2%), Norwegen (5. Jahrgangsstufe, 2.4%) oder der Flämischen Gemeinschaft in Belgien (3.5%).

Insgesamt gesehen zeigt sich im internationalen Vergleich von IGLU 2016, dass die Schülerinnen und Schüler in Deutschland nach eigenen Angaben re-

Abbildung 4.8: Schülerinnen und Schüler, die angeben, täglich 2 Stunden oder mehr außerhalb der Schule zu lesen, im internationalen Vergleich – IGLU 2011 und 2016 im Vergleich



■ Statistisch signifikante Veränderungen zwischen 2011 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

A = Da Belgien (Fl. Gem.) und Lettland nicht an IGLU 2011 teilgenommen haben, können hier für diese Teilnehmer keine Ergebnisse berichtet werden.

B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

Δ_{16-11} = Differenz in den Prozentwerten zwischen 2016 und 2011.

lativ viel Zeit mit dem täglichen Lesen verbringen. Bei der Einordnung dieses Befunds ist zu berücksichtigen, dass die Frage nach dem täglichen Lesen außerhalb der Schule gestellt wurde und die den Kindern verfügbare Freizeit mit der Schulorganisationsform zusammenhängen kann. So ist beispielsweise anzunehmen, dass in Ländern mit flächendeckenden Ganztagsgrundschulen (etwa in den skandinavischen Ländern) die für außerschulische Aktivitäten verfügbare Freizeit geringer ausfällt als in Deutschland, wo im Schuljahr 2015/2016 rund 56 Prozent aller Grundschulen ein Ganztagsangebot bereitstellten (KMK, 2017).

Für den Vergleich der Erhebungszyklen werden nur Angaben von 2011 und 2016 berücksichtigt, da die Fragen in 2001 und 2006 nicht Teil des internationalen Schülerfragebogens waren. Abbildung 4.7 weist den Anteil an Schülerinnen und Schülern aus, die weniger als 30 Minuten pro Schultag außerhalb der Schule lesen, Abbildung 4.8 veranschaulicht die Anteile für die Kinder, die dies 2 Stunden oder mehr pro Schultag tun. Beim Vergleich der Angaben der Kinder aus 2011 und 2016 zeigt sich, dass sich der Anteil der ‚Wenigleser‘ in den hier

berücksichtigten Teilnehmerstaaten und -regionen mit Ausnahme von Italien und Bulgarien signifikant erhöht hat. Deutschland liegt mit einem Differenzwert von 7 Prozentpunkten im Mittelfeld der im Trend betrachteten Bildungssysteme. Signifikant größere Zunahmen des Anteils an Kindern, die weniger als 30 Minuten pro Tag lesen, finden sich in Singapur, Taiwan und Slowenien sowie in Ontario.

Zugleich hat sich in fast allen Teilnehmerstaaten und -regionen der Anteil der Schülerinnen und Schüler, die angeben, täglich 2 Stunden oder mehr in ihrer Freizeit zu lesen, von 2011 zu 2016 signifikant verringert, wie Abbildung 4.8 verdeutlicht. Die Veränderung in Deutschland in der Höhe von 4 Prozentpunkten unterscheidet sich nicht signifikant von den Veränderungen in den meisten der hier betrachteten Bildungssysteme. Allerdings sind in Italien, Belgien (Französische Gemeinschaft), England und Norwegen (4. Jahrgangsstufe) keine signifikanten Abnahmen des Anteils an Kindern zu beobachten, die täglich in diesem Zeitumfang lesen. In Hongkong, Schweden, Frankreich und Dänemark sind die Veränderungen signifikant geringer als in Deutschland.

Die in den Abbildungen 4.7 und 4.8 dargestellten Veränderungen in der täglichen Lesedauer weisen aus, dass es in einer Reihe von Bildungssystemen gelingt, die Zunahmen an Kindern, die täglich sehr wenig ihrer Freizeit für das Lesen verwenden, und die Abnahmen im Anteil derjenigen, die viel Zeit mit dem täglichen Lesen verbringen, geringer zu halten als in Deutschland. Für die meisten dieser Länder ist allerdings aus deutscher Sicht mit zu berücksichtigen, dass die Ausgangslagen 2011 vergleichsweise ungünstig waren.

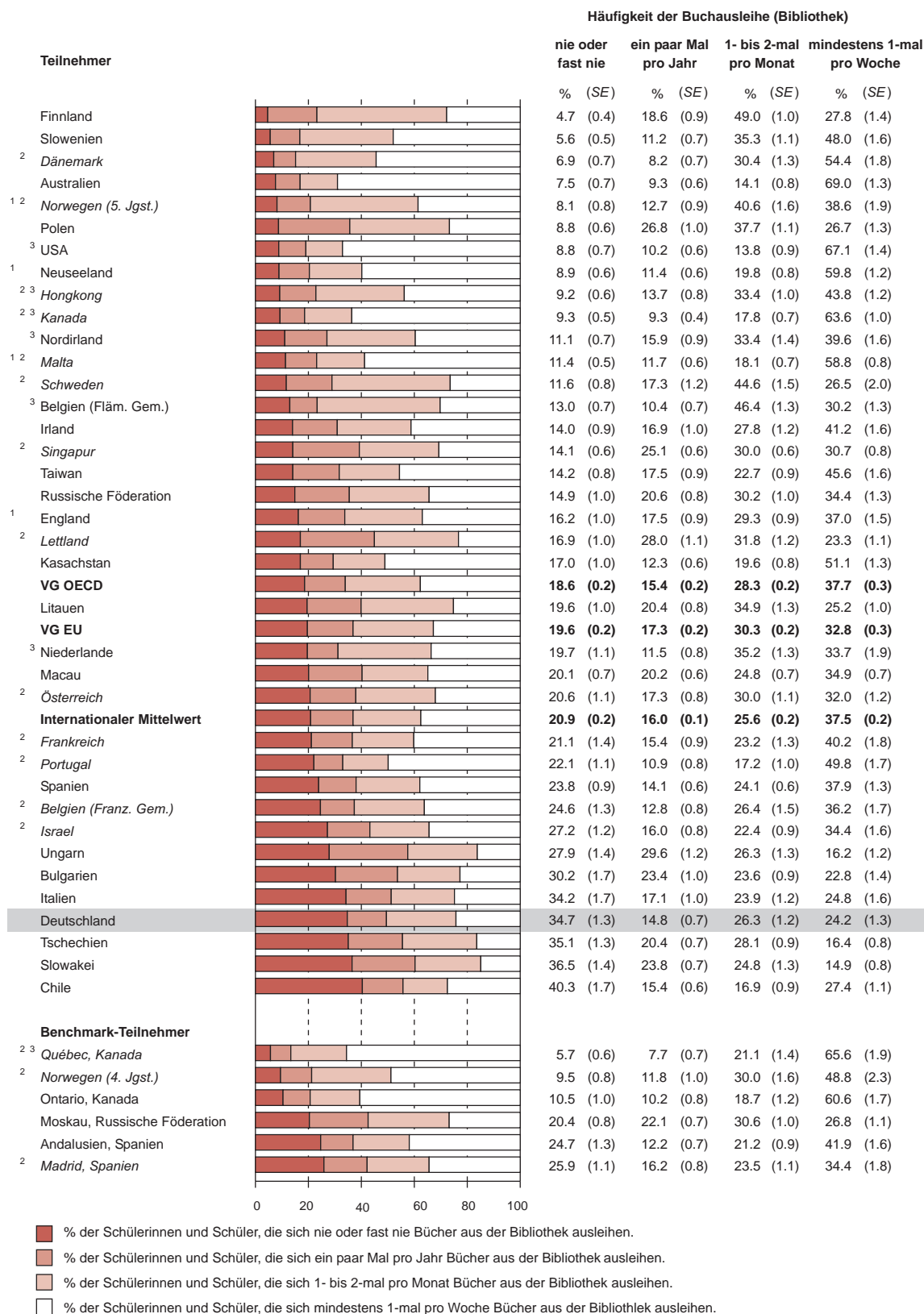
Wie die Häufigkeit des Lesens zum Vergnügen hängt auch die Dauer des täglichen Lesens außerhalb der Schule signifikant mit der Leseleistung der Kinder in Deutschland zusammen ($r = .17$; $p < .05$). Hinsichtlich der Schülerinnen und Schüler mit unterschiedlichem Leistungsniveau im Lesen (ohne Abbildung) lässt sich festhalten, dass sich erwartungskonform die Gruppe der lesestarken Schülerinnen und Schüler (Kompetenzstufen IV und V) in Bezug auf die Dauer ihrer Lesezeit deutlich von der Gruppe der Leseschwachen (Kompetenzstufen I und II) unterscheidet: Während 2016 über die Hälfte der Leseschwachen (51.7%) angibt, weniger als 30 Minuten täglich zu lesen, sind es bei den Lesestarken signifikant geringere 32.4 Prozent. Vergleicht man die Ergebnisse von 2011 und 2016, so fällt auf, dass sich der Anteil der Kinder, die weniger als 30 Minuten pro Tag lesen, in der Gruppe der leistungsstarken Schülerinnen und Schüler signifikant von 24.6 auf 32.4 Prozent erhöht hat. In der Gruppe mit mittleren Leseleistungen (Kompetenzstufe III) ist die Veränderung ebenfalls signifikant (2011: 38.5%; 2016: 45.6%). Zugleich verringert sich auch der Anteil derjenigen, die mehr als 2 Stunden täglich lesen, in den oberen Leistungsgruppen: Bei den Schülerinnen und Schülern auf den Kompetenzstufen IV und V verringert sich der Anteil von 17.0 Prozent im Jahr 2011 auf 11.2 Prozent im Jahr 2016, bei den Kindern auf Kompetenzstufe III von 12.1 Prozent (2011) auf 8.2 Prozent (2016). Bei den Schülerinnen und Schülern auf den Kompetenzstufen I und II zeigen sich demgegenüber keine signifikanten Änderungen bei den täglich wenig und den täglich viel Lesenden. Insgesamt gesehen sind Abnahmen der täglichen Lesedauer zu verzeichnen, die sich bei dieser Variable vor allem auf ein geändertes Leseverhalten der leistungsstärkeren Schülerinnen und Schüler zurückführen lassen.

2.3.3 Bibliotheksnutzung

Bibliotheken können Kinder durch ein vielfältiges Angebot an unterschiedlichen Lesematerialien zum Lesen anregen, indem sie Lesestoff zu altersangemessenen Inhalten oder Interessen von Kindern anbieten. Die Möglichkeit, die Bücher nach Thema oder Umfang selbstständig und nach eigenen Interessen auszuwählen, kann motivationsfördernd wirken. Darüber hinaus sind über eine erhöhte Lesemenge Effekte auf das Leseverständnis zu erwarten. In den IGLU-Erhebungen der Jahre 2006, 2011 und 2016 wurden die Schülerinnen und Schüler danach gefragt, wie häufig sie Bücher aus der Schulbibliothek oder der örtlichen Bibliothek ausleihen, wobei sich im Jahr 2016 die Frage auch auf elektronische Bücher (E-Books) bezieht. Die Ergebnisse im internationalen Vergleich werden in Abbildung 4.9 präsentiert. In Deutschland geben in 2016 34.7 Prozent der Schülerinnen und Schüler an, sich nie oder fast nie Bücher aus der Schulbibliothek oder örtlichen Bibliothek auszuleihen. 14.8 Prozent tun dies ein paar Mal pro Jahr, 26.3 Prozent der Schülerinnen und Schüler ein- bis zweimal pro Monat und 24.2 Prozent mindestens einmal pro Woche. Der internationale Vergleich zeigt, dass der Anteil der Kinder in Deutschland, die nie oder fast nie Bücher aus der Schulbibliothek oder örtlichen Bibliothek ausleihen, hoch ausfällt. Nur in einem Bildungssystem (Chile mit 40.3 %) fällt der Anteil höher aus, in der Slowakei (36.5 %), Tschechien (35.1 %) und Italien (34.2 %) unterscheidet er sich nicht signifikant vom Anteil in Deutschland. In allen übrigen Bildungssystemen sowie in den dargestellten Mittelwerten (international: 20.9 %, EU: 19.6 %, OECD: 18.6 %) gibt es signifikant weniger Kinder, die sich nie oder fast nie Bücher ausleihen. In einigen Bildungssystemen ist der Anteil derjenigen, die nie oder fast nie Bücher entleihen, nur ein Viertel bis ein Fünftel so groß wie in Deutschland (unter den europäischen Teilnehmerstaaten beispielsweise in Dänemark mit 6.9 % und in Slowenien mit 5.6 %). Demgegenüber gibt es in Deutschland vergleichsweise wenige Kinder, die mindestens wöchentlich Bücher entleihen (24.2 %), nur in drei Bildungssystemen fällt dieser Anteil signifikant geringer aus (Slowakei: 14.9 %; Ungarn: 16.2 %; Tschechien: 16.4 %). Demgegenüber liegen in einigen Staaten die Anteile fast doppelt so hoch wie in Deutschland (unter den europäischen Teilnehmern beispielsweise in Dänemark mit 54.5 % oder Slowenien mit 48.0 %).

Insgesamt gesehen ist festzuhalten, dass sich 2016 zumindest gut die Hälfte (50.5 %) aller Viertklässlerinnen und Viertklässler am Ende der Grundschulzeit in Deutschland ein- bis zweimal im Monat Bücher ausleiht, so dass Bibliotheken im Leseverhalten der Kinder durchaus eine Rolle spielen. Im internationalen Vergleich werden Bibliotheken aber in fast allen anderen Teilnehmerstaaten und -regionen deutlich häufiger für die Buchausleihe genutzt. Dies zeigt sich auch, wenn man auf die Gruppe an Kindern schaut, die zumindest monatlich Bücher aus der Bibliothek oder Schulbibliothek entleihen: In dieser Gruppe gibt es mit der Slowakei (39.7 %), Ungarn (42.5 %) und Tschechien (44.5 %) nur drei europäische Bildungssysteme, bei denen der Anteil signifikant geringer ist als in Deutschland (50.5 %). In der weit überwiegenden Mehrheit ist der Anteil signifikant höher, wobei unter den Teilnehmern die skandinavischen Länder Dänemark (84.8 %), Norwegen (5. Jahrgangsstufe, 79.2 %) und Finnland (76.8 %) ins Auge fallen, sowie erneut Slowenien (83.3 %). Bei diesen Befunden ist allerdings auch zu bedenken, dass die Angaben der Schülerinnen und Schüler unter anderem von einem Vorhandensein an Schulbibliotheken abhängig sind. Den bei IGLU 2016 befragten Schulleitungen zufolge stand 2016 für 72 Prozent der

Abbildung 4.9: Häufigkeit der Buchausleihe aus der Schulbibliothek oder örtlichen Bibliothek im internationalen Vergleich – IGLU 2016 (Anteile der Schülerinnen und Schüler in Prozent)



% der Schülerinnen und Schüler, die sich nie oder fast nie Bücher aus der Bibliothek ausleihen.
 % der Schülerinnen und Schüler, die sich ein paar Mal pro Jahr Bücher aus der Bibliothek ausleihen.
 % der Schülerinnen und Schüler, die sich 1- bis 2-mal pro Monat Bücher aus der Bibliothek ausleihen.
 % der Schülerinnen und Schüler, die sich mindestens 1-mal pro Woche Bücher aus der Bibliothek ausleihen.

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

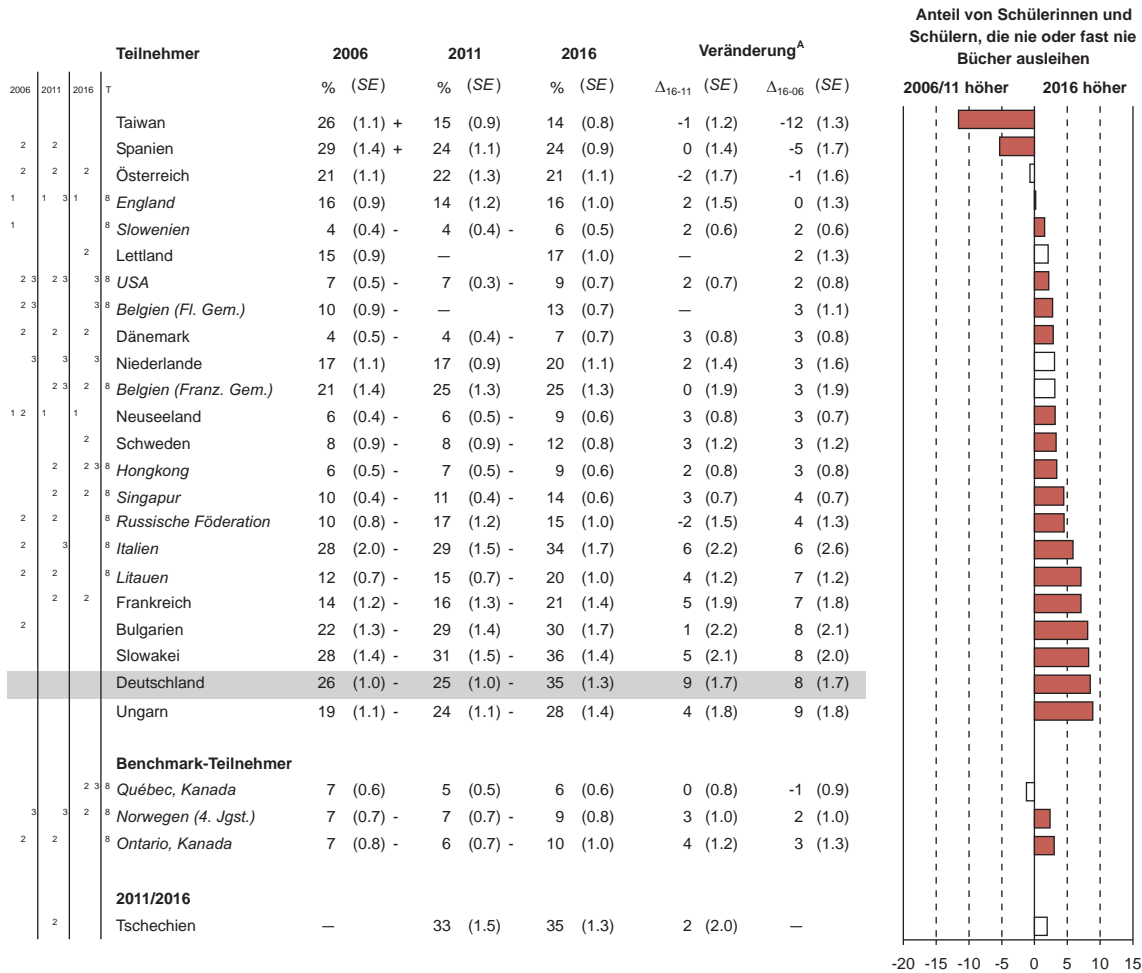
Schülerinnen und Schüler in Deutschland eine Schulbibliothek zur Verfügung, in einer Reihe von Ländern mit intensiverer Bibliotheksnutzung korrespondiert dies auch mit höheren Prozentanteilen an Schulen mit Schulbibliotheken, so zum Beispiel in Slowenien (100 %), Norwegen (5. Jahrgangsstufe, 96 %) oder Dänemark (94 %) (Mullis, Martin, Foy & Hooper, 2017).

In Bezug auf den Trend in der Bibliotheksnutzung werden nun die Anteile der Schülerinnen und Schüler im internationalen Vergleich betrachtet, die nie oder fast nie Bücher aus einer schulischen oder örtlichen Bibliothek entleihen (siehe Abbildung 4.10), und diejenigen, die dies mindestens einmal pro Woche tun (siehe Abbildung 4.11). Für Deutschland zeigt sich von 2006 bis 2016 eine signifikante Zunahme des Anteils an Kindern, die nie oder fast nie Bücher entleihen, um 8 Prozentpunkte. Signifikante Zunahmen zeigen sich in insgesamt 18 Bildungssystemen. Die Veränderung ist in Deutschland nominell ausgeprägter als in den meisten der hier betrachteten Bildungssysteme, wobei sich die Veränderungen in der Russischen Föderation, in Italien, Litauen, Frankreich, Bulgarien, der Slowakei und Ungarn nicht signifikant vom deutschen Veränderungswert unterscheiden. Eine signifikante Verringerung des Anteils an Kindern, die nie oder fast nie Bücher entleihen, ist nur in zwei Staaten zu verzeichnen: In Taiwan beträgt diese Veränderung 12 Prozentpunkte, in Spanien 5 Prozentpunkte.

Aus Abbildung 4.11 ist ersichtlich, dass der Anteil an Kindern, die mindestens einmal pro Woche Bücher entleihen, in Deutschland von 2006 bis 2016 signifikant um 5 Prozentpunkte gesunken ist. Ein entsprechender Trend ist in der Mehrheit der betrachteten Bildungssysteme zu beobachten, wobei sich in Litauen (14 Punkte), in der Russischen Föderation (13 Punkte) sowie in Dänemark, Frankreich, Lettland und Ungarn (jeweils 10 Punkte) die höchsten Veränderungswerte zeigen. Unter diesen ist in Litauen und der Russischen Föderation der Prozentwert der Veränderung signifikant höher als in Deutschland. In zehn Bildungssystemen sind die Veränderungen von 2006 bis 2016 demgegenüber nicht signifikant. Ein signifikanter Anstieg der Bibliotheksnutzung ist nur in Taiwan zu verzeichnen: Von 2006 bis 2016 ist hier der Anteil an Kindern, die mindestens wöchentlich Bücher entleihen, um 8 Prozentpunkte gestiegen.

Wenn man die Buchausleihe für die Schülerinnen und Schüler in Deutschland nach ihrem Leistungsniveau im Lesen differenziert betrachtet (ohne Abbildung), lässt sich feststellen, dass der Anteil der Kinder, die nie oder fast nie Bücher entleihen, in den letzten 10 Jahren in den drei in diesem Kapitel unterschiedenen Leistungsgruppen signifikant gestiegen ist, wobei der höchste Anstieg (um 11 Prozentpunkte) in der Gruppe der schwachen Leserinnen und Leser (Kompetenzstufen I und II) zu beobachten ist. Zugleich ist der Anteil der häufigen Bibliotheksnutzer in allen Leistungsgruppen nominell gesunken. Signifikante Unterschiede zeigen sich allerdings nur bei den leseschwachen Schülerinnen und Schülern (um rund 10 Prozent) sowie bei den Schülerinnen und Schülern auf Kompetenzstufe III (um rund 7 Prozent). Dass sich die größten Veränderungen im Trend in der Gruppe der Leseschwachen zeigen, ist durchaus bedenkenswert.

Abbildung 4.10: Schülerinnen und Schüler, die angeben, nie oder fast nie Bücher aus der Schulbibliothek oder örtlichen Bibliothek auszuleihen, im internationalen Vergleich – IGLU 2006, 2011 und 2016 im Vergleich



Statistisch signifikante Veränderungen zwischen 2006 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

- 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
- 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
- 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
- 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

+ = Anteil von Schülerinnen und Schülern, die nie oder fast nie Bücher ausleihen, 2006 bzw. 2011 statistisch signifikant höher als 2016 ($p < .05$).

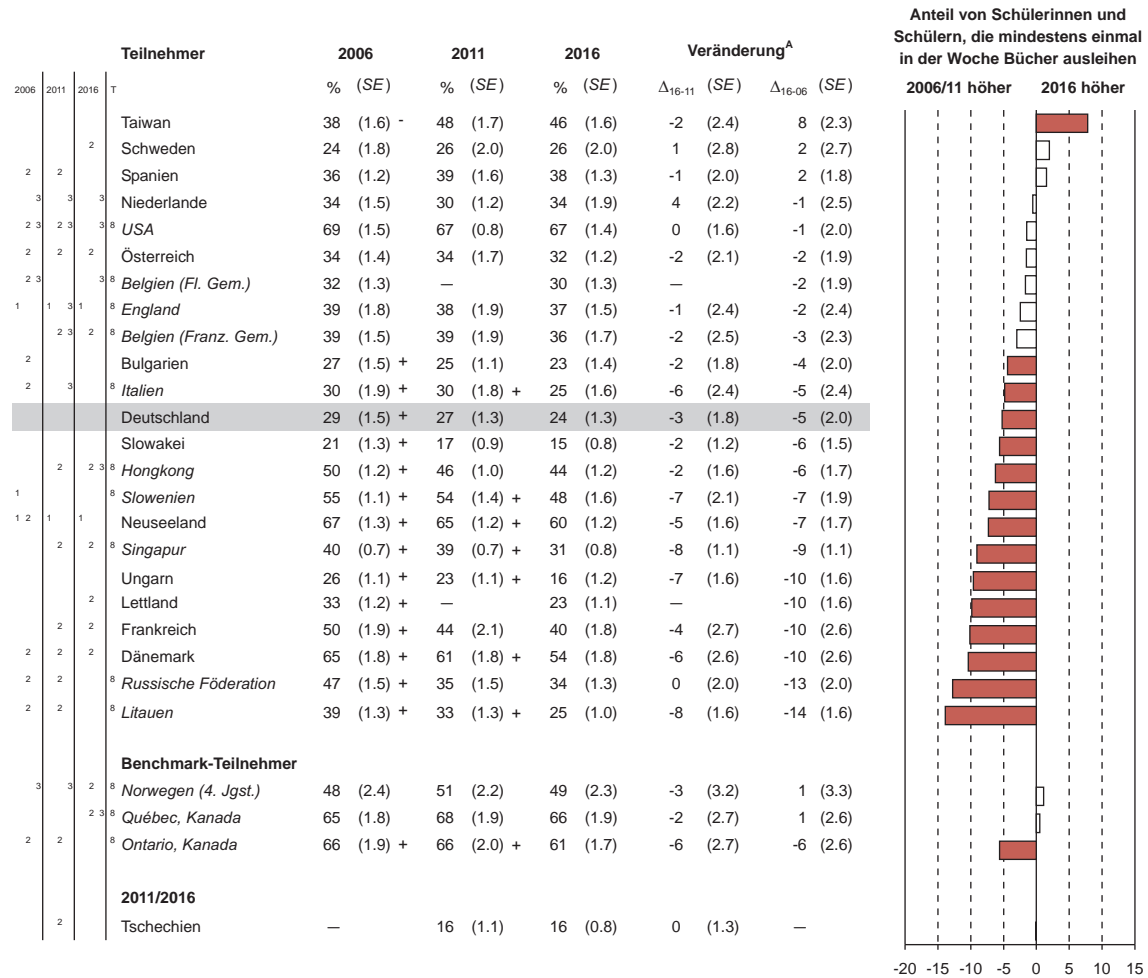
- = Anteil von Schülerinnen und Schülern, die nie oder fast nie Bücher ausleihen, 2006 bzw. 2011 statistisch signifikant niedriger als 2016 ($p < .05$).

$\Delta_{16-06} / \Delta_{16-11}$ = Differenz in den Prozentwerten zwischen 2016 und 2006 bzw. 2011 (bei denjenigen Teilnehmern, die nicht an IGLU 2006 teilgenommen haben).

2.4 Familiäre und individuelle Bedingungsfaktoren des Risikos schwacher Lesekompetenz

Vor dem Hintergrund des Befunds, dass sich die Gruppe der leseschwachen Schülerinnen und Schüler in Deutschland im IGLU-Trend seit 2001 vergrößert hat und 2016 rund 19 Prozent der Kinder nicht die Kompetenzstufe III erreichen (siehe Kapitel 3 in diesem Band), wird im Folgenden betrachtet, in welchem Zusammenhang das Leseselbstkonzept, die Lesemotivation und das Leseverhalten mit einer Verringerung des Risikos stehen, der Gruppe der schwachen Leserinnen und Leser anzugehören, und welche Rolle die Lesesozialisation in der Familie in diesem Kontext spielt. Zugleich werden zentrale strukturelle

Abbildung 4.11: Schülerinnen und Schüler, die angeben, mindestens einmal in der Woche Bücher aus der Schulbibliothek oder örtlichen Bibliothek auszuleihen, im internationalen Vergleich – IGLU 2006, 2011 und 2016 im Vergleich



Statistisch signifikante Veränderungen zwischen 2006 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

+ = Anteil von Schülerinnen und Schülern, die mindestens 1-mal in der Woche Bücher ausleihen, 2006 bzw. 2011 statistisch signifikant höher als 2016 ($p < .05$).

- = Anteil von Schülerinnen und Schülern, die mindestens 1-mal in der Woche Bücher ausleihen, 2006 bzw. 2011 statistisch signifikant niedriger als 2016 ($p < .05$).

$\Delta_{16-06} / \Delta_{16-11}$ = Differenz in den Prozentwerten zwischen 2016 und 2006 bzw. 2011 (bei denjenigen Teilnehmern, die nicht an IGLU 2006 teilgenommen haben).

Herkunftsmerkmale (Migrationshintergrund, Bildungsniveau, Berufsstatus) sowie individuelle Kovariaten (Geschlecht, kognitive Fähigkeiten) berücksichtigt. Wie in den Kapiteln 5, 6 und 7 in diesem Band ausführlicher dargestellt, sind in der Gruppe der leseschwachen Schülerinnen und Schüler, deren Leistungen auf den unteren beiden Kompetenzstufen zu verorten sind, Jungen stärker vertreten als Mädchen und Kinder mit Migrationshintergrund und aus bildungsfernen Elternhäusern stärker als Kinder ohne Migrationshintergrund und aus bildungsnahen Elternhäusern. Im Folgenden werden die Ergebnisse logistischer Mehrebenenregressionen vorgestellt, anhand derer überprüft wurde, inwieweit im Trend aller vier Erhebungszyklen von IGLU das Leseselbstkonzept, die Lesemotivation und das Leseverhalten der Schülerinnen und Schüler über diese

Hintergrundvariablen hinaus prädiktive Kraft für die Zugehörigkeit zur Gruppe der schwachen Leserinnen und Leser haben. Die Ergebnisse werden in zwei Schritten dargestellt.

Zunächst wird in Tabelle 4.1 betrachtet, inwieweit sich das Risiko, zur Gruppe der Leseschwachen zu gehören, durch positive Ausprägungen von Leseselbstkonzept, Lesemotivation, Leseverhalten und der Lesesozialisation im Elternhaus reduziert, wobei Modell 1 dies ohne Kontrollvariablen betrachtet und Modell 2 die oben genannten Merkmale als zentrale Kovariaten der Leseleistung hinzunimmt. Ferner wurden als ergänzende Kontrollvariable die kognitiven Fähigkeiten der Schülerinnen und Schüler in die Analyse einbezogen. Tabelle 4.2 nimmt in der Gegenperspektive in den Blick, inwieweit das Risiko, zur Gruppe der Leseschwachen zu gehören, durch Bildungsniveau, Berufsstatus, Migrationshintergrund und Geschlecht erklärt werden kann (Modell 1) und inwieweit sich diese Effekte verändern, wenn Leseselbstkonzept, Lesemotivation, Leseverhalten, die Lesesozialisation im Elternhaus und die kognitiven Grundfähigkeiten kontrolliert werden (Modell 2).

Die abhängige Variable bildet in den logistischen Mehrebenenregressionen die Zugehörigkeit der Kinder zu der Gruppe schwacher Leserinnen und Leser auf den Kompetenzstufen I oder II. Die Analysen stellen eine Ergänzung und Spezifizierung des allgemeinen Trendmodells der Leseleistung in diesem Band dar (zu diesem Modell siehe Kapitel 2, Abschnitt 12, zur gemeinsamen Skalierung der Leseleistungswerte im Trend siehe Kapitel 3).

Die Konstrukte ‚Lesesozialisation‘ und ‚Lesemotivation‘ fließen in Form kontinuierlicher Variablen in die Modelle ein. Hierzu wurde auf Basis der in den verschiedenen Erhebungszyklen von IGLU eingesetzten Variablen jeweils ein gemeinsamer Trendindex skaliert. Für die Lesesozialisation im Elternhaus wurde ebenfalls ein Trendindex skaliert. Hierbei wurden Angaben aus dem Elternfragebogen zu vier Aspekten berücksichtigt: die Anzahl der Bücher im Haushalt, die Anzahl der Kinderbücher im Haushalt, die elterliche Lesemotivation sowie leseförderliche Aktivitäten, die Eltern mit ihren Kindern unternehmen (z.B. *Meinem Kind beim Lesen für die Schule helfen*). Die Auswahl der Variablen orientiert sich an der Bildung eines Index zur Lesesozialisation bei Stubbe et al. (2007) und Tarelli (2010), schließt allerdings nicht alle dort berücksichtigten Bereiche mit ein. Details zur Bildung der Variablen zum Bildungsniveau und zum Berufsstatus werden in Kapitel 6 in diesem Band ausgeführt, für die Variablen zum Migrationshintergrund sei auf Kapitel 7 in diesem Band verwiesen.

Zur besseren Veranschaulichung der Bedeutung der Regressionsgewichte wurden die Koeffizienten in sogenannte *odds ratios* (Chancenverhältnisse) transformiert. Diese drücken das mit dem jeweiligen Merkmal verbundene relative Risiko aus, der Gruppe der Leseschwachen anzugehören. Ist ein Wert größer als 1, drückt dies ein erhöhtes Risiko der Gruppenzugehörigkeit aus, ein Wert kleiner als 1 steht für ein verringertes Risiko. Bei dichotomen und dichotomisierten Variablen wird das relative Risiko der genannten Gruppe mit dem der Referenzkategorie verglichen. So geht beispielsweise in Modell 1 in Tabelle 4.1 das Leseverhalten mit einem *odds ratio* von 0.59 einher. Dies bedeutet, dass Kinder, die täglich oder fast täglich zum Vergnügen lesen, ein um den Faktor 0.59 beziehungsweise ein um 41 Prozent vermindertes Risiko haben, der Gruppe der schwachen Leserinnen und Leser anzugehören – und zwar im Vergleich zu Kindern, die nie oder fast nie zu ihrem Vergnügen lesen. Im Fall von kontinuierlichen Variablen stehen die *odds ratios* für die mit einer Standardabweichung

Tabelle 4.1: Prädiktion der Zugehörigkeit zur Gruppe schwacher Leserinnen und Leser (Kompetenzstufen I und II) durch individuelle und familiäre Variablen – Ergebnisse logistischer Mehrebenenregressionen (*odds ratios*)

	Modell 1	Modell 2
Hohes Leseselbstkonzept ^A	0.42 **	0.45 **
Hohe Lesemotivation ^B	— ns	— ns
Intensives Leseverhalten ^C	0.59 **	0.61 *
Hoher Grad an Lesesozialisation im Elternhaus ^D	0.42 **	0.57 **

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

Modell 1: Trendmodell unter Berücksichtigung der Mehrebenenstruktur.

Modell 2: Modell 1 unter Kontrolle von Bildungsniveau, Berufsstatus, Migrationshintergrund und Geschlecht (siehe Tabelle 4.2 in diesem Kapitel) sowie kognitiven Fähigkeiten (siehe Kapitel 2 in diesem Band).

A = Skalierter Index des Leseselbstkonzepts (Messinvarianz über die Zyklen angenommen).

B = Skalierter Index der Lesemotivation (Messinvarianz über die Zyklen angenommen).

C = Lesen zum Vergnügen (0 = Nie oder fast nie; 1 = Jeden Tag oder fast jeden Tag).

D = Skalierter Index der Lesesozialisation im Elternhaus (Messinvarianz über die Zyklen angenommen).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

auf der unabhängigen Variable verbundene Veränderung des relativen Risikos. Ein Anstieg des Index des Leseselbstkonzepts um eine Standardabweichung geht in Modell 1 mit einem *odds ratio* von 0.42 einher, also mit einem um den Faktor 0.42 beziehungsweise 58 Prozent verringerten Risiko. Neben den *odds ratios* sind in den Tabellen 4.1 und 4.2 auch die jeweiligen Signifikanzen dieser Koeffizienten ausgewiesen.

Die Ergebnisse von Modell 1 in Tabelle 4.1 zeigen signifikante Effekte eines hohen Leseselbstkonzepts und des täglichen oder fast täglichen Lesens zum Vergnügen auf die Verminderung des Risikos, zur Gruppe der Leseschwachen zu gehören. Auch der Effekt der Lesesozialisation ist signifikant – ein günstiger Wert reduziert das Risiko, der Gruppe der Leseschwachen anzugehören, um mehr als 50 Prozent. Eine hohe Lesemotivation hat unter gleichzeitiger Berücksichtigung der übrigen Variablen keinen signifikanten Effekt. In einer reduzierten Version von Modell 1 (ohne Abbildung), in der die Lesesozialisation im Elternhaus nicht mit aufgenommen wurde, zeigt sich ein signifikanter Effekt der Lesemotivation (*odds ratio* = 0.85, $p < .05$) unter gleichzeitiger Berücksichtigung von Leseselbstkonzept (*odds ratio* = 0.39, $p < .01$) und Leseverhalten (*odds ratio* = 0.54, $p < .01$). Modell 2 zeigt, dass ein hohes Leseselbstkonzept, das tägliche oder fast tägliche Lesen zum Vergnügen sowie ein hoher Grad an Lesesozialisation im Elternhaus das Risiko, zur Gruppe der schwachen Leserinnen und Leser zu gehören, auch unter Kontrolle der strukturellen Herkunftsmerkmale und der individuellen Kovariaten signifikant reduzieren, wobei die Effekte etwas schwächer ausfallen.

In Tabelle 4.2 zeigen sich in Modell 1 analog zu den Ergebnissen im allgemeinen Trendmodell (siehe Anhang B in diesem Band) signifikante und bedeutsame Effekte des Bildungsniveaus, des Berufsstatus der Eltern und des Migrationshintergrunds der Schülerinnen und Schüler auf das Risiko, der Gruppe der Leseschwachen anzugehören; das Geschlecht der Schülerinnen und Schüler hat unter gleichzeitiger Betrachtung der übrigen Variablen keinen signifikanten Effekt, so dass diese Befunde für Jungen und Mädchen gleichermaßen gelten. Unter Kontrolle von Leseselbstkonzept, Lesemotivation, Leseverhalten und der Lesesozialisation im Elternhaus sowie der kognitiven Grundfähigkeiten (Modell

Tabelle 4.2: Prädiktion der Zugehörigkeit zur Gruppe schwacher Leserinnen und Leser (Kompetenzstufen I und II) durch strukturelle HerkunftsvARIABLEN und Geschlecht – Ergebnisse logistischer Mehrebenenregressionen (*odds ratios*)

	Modell 1	Modell 2
Familie mit hohem Bildungsniveau ^A	0.38 **	— ns
Familie mit hohem Berufsstatus ^B	0.40 **	0.62 *
Migrationshintergrund (ein Elternteil im Ausland geboren) ^C	2.08 **	1.91 **
Migrationshintergrund (beide Elternteile im Ausland geboren) ^D	3.34 **	2.80 **
Geschlecht ^E	— ns	— ns

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

Modell 1: Trendmodell unter Berücksichtigung der Mehrebenenstruktur.

Modell 2: Modell 1 unter Kontrolle von Leseselbstkonzept, Lesemotivation, Leseverhalten und Lesesozialisation im Elternhaus (siehe Tabelle 4.1 in diesem Kapitel) sowie kognitiven Fähigkeiten (siehe Kapitel 2 in diesem Band).

A = Bildungsniveau nach höchstem Bildungsabschluss der Eltern (0 = Kein Elternteil mit (Fach-)Hochschulabschluss; 1 = Mindestens ein Elternteil mit (Fach-)Hochschulabschluss), siehe Kapitel 6 in diesem Band.

B = Berufsstatus nach Angabe der Eltern (0 = Manuell Tätige, Angestellte und kleinere Unternehmer; 1 = Akademiker, Techniker und Führungskräfte), siehe Kapitel 6 in diesem Band.

C = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren; 1 = Ein Elternteil im Ausland geboren), siehe Kapitel 7 in diesem Band.

D = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren; 1 = Beide Elternteile im Ausland geboren), siehe Kapitel 7 in diesem Band.

E = Geschlecht (0 = Mädchen; 1 = Jungen), siehe Kapitel 5 in diesem Band.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

2) zeigt sich, dass das Bildungsniveau in den Familien keinen signifikanten Beitrag zur Erklärung von Leseschwäche leistet und dass sich die Effekte des Berufsstatus und des Migrationshintergrunds leicht reduzieren.

Insgesamt gesehen sprechen die Koeffizienten in Tabelle 4.2 dafür, dass ein positives Leseselbstkonzept und ein häufiges Lesen zum Vergnügen das Risiko, der Gruppe der Leseschwachen anzugehören, ebenso bedeutsam vermindern wie ein hoher Grad an Lesesozialisation im Elternhaus. In der Höhe sind diese Effekte (unter gleichzeitiger Betrachtung aller im Modell aufgenommen Variablen) jeweils denen eines hohen elterlichen Berufsstatus vergleichbar.

Zugleich zeigt Tabelle 4.2 aber auch, dass in der Gruppe der Kinder mit Migrationshintergrund, insbesondere bei beiden im Ausland geborenen Elternteilen, im Vergleich zu Kindern, deren Eltern in Deutschland geboren sind, ein deutlich erhöhtes Risiko besteht, den Leistungsschwachen anzugehören. Auch unter Berücksichtigung der Variablen zu Leseselbstkonzept, Lesemotivation und Leseverhalten sowie der Lesesozialisation im Elternhaus (im Vergleich von Modell 1 und Modell 2 in Tabelle 4.2) verringert sich dieses Risiko nur geringfügig. Zur Erklärung für diesen deutlichen Effekt des Migrationshintergrunds unter Berücksichtigung der übrigen Merkmale im Modell wären weitergehende Analysen und auch zusätzliche Daten notwendig, da bei den Variablen Migrationshintergrund und Berufsstatus Personengruppen zusammengefasst werden, die in Bezug auf ihre kulturellen Praktiken und Wertvorstellungen erheblich voneinander abweichen dürften. Berücksichtigt werden könnte hierbei auch die Häufigkeit, mit der in der Familie Deutsch gesprochen wird.

3 Zusammenfassung

Bei IGLU werden neben den Leseleistungen der Schülerinnen und Schüler auch das Leseselbstkonzept, die Lesemotivation und das Leseverhalten (wie das Lesen zum Vergnügen außerhalb der Schule, die tägliche Lesedauer in der Freizeit sowie das Entleihen von Büchern aus der örtlichen Bibliothek oder Schulbibliothek) erfasst. In dieser Zusammenfassung werden vorwiegend die Veränderungen in diesen Merkmalen – je nach Datenlage – in den letzten 5 bis 15 Jahren dargestellt, Bedingungsfaktoren schwacher Leseleistung fokussiert und die Ergebnisse im Hinblick auf Fördermöglichkeiten diskutiert. Veränderungen in der Stichprobenzusammensetzung, die hinsichtlich der Befunde im Zyklenvergleich zu berücksichtigen sind, sind in Kapitel 2 in diesem Band dargestellt.

Leseselbstkonzept. Die Viertklässlerinnen und Viertklässler in Deutschland schätzen im Jahr 2016 ihre eigenen Lesefähigkeiten überwiegend positiv ein. In den mittleren Skalenwerten zeigen sich keine signifikanten Veränderungen zwischen 2011 und 2016. Im Jahr 2016 weisen etwas weniger Kinder ein hohes Selbstkonzept auf als im Jahr 2011, der Anteil an Kindern mit einem Selbstkonzept mittlerer Höhe hat sich im Gegenzug leicht vergrößert. Mit jeweils 5 Prozent sind die Anteile der Schülerinnen und Schüler mit einem niedrigen Selbstkonzept in etwa gleich geblieben. Eine Betrachtung nach Kompetenzstufen im Lesen zeigt, dass die Schülerinnen und Schüler in 2016 etwas realistischere Selbsteinschätzungen ihrer Lesefähigkeiten haben als in 2011.

Lesemotivation. Insgesamt gesehen sind die Schülerinnen und Schüler in Deutschland weiterhin positiv dem Lesen gegenüber eingestellt. Die Trendbetrachtung zeigt allerdings, dass in 2016 die Werte ungünstiger als 2001 sind – fast jedes sechste Kind hat am Ende der Grundschulzeit nun eine geringe Lesemotivation, eine bedenkenswerte Entwicklung. Im Vergleich der IGLU-Erhebungszyklen zeigen sich Änderungen insbesondere bei den leseschwachen Kindern (Kompetenzstufen I und II): In dieser Leistungsgruppe erhöht sich der Anteil an gering Lesemotivierten um 11 Prozentpunkte, der Anteil an hoch Lesemotivierten sinkt um rund 15 Prozentpunkte.

Lesen zum Vergnügen. 2016 lesen etwas mehr als 70 Prozent der Schülerinnen und Schüler in Deutschland zumindest ein- bis zweimal wöchentlich zu ihrem Vergnügen. Rund 17 Prozent geben an, nie oder fast nie außerhalb der Schule zum Vergnügen zu lesen. Der Anteil hat sich in Deutschland im Vergleich zu 2001 nicht signifikant verändert. In neun Bildungssystemen ist es in diesem Zeitraum gelungen, den Anteil dieser ‚Nichtleser‘ signifikant zu verringern. Zugleich zeigt der internationale Vergleich, dass der Anteil an Schülerinnen und Schülern, die jeden Tag oder fast jeden Tag zu ihrem Vergnügen lesen, in Deutschland mit rund 43 Prozent vergleichsweise hoch ausfällt. Allerdings hat sich dieser Anteil zwischen 2001 und 2016 leicht verringert, was auch für eine Reihe anderer europäischer Staaten zutrifft. Die vergleichsweise ungünstigste Entwicklung betrifft auch hier die Kinder mit geringen Leseleistungen: Im Vergleich zur Gruppe der Leistungsstarken finden sich hier die größte Abnahme an denjenigen, die täglich oder fast täglich zum Vergnügen lesen und die größte Zunahme derjenigen, die dies nie oder fast nie tun.

Tägliche Lesedauer. Rund 60 Prozent der Kinder in Deutschland geben 2016 an, mindestens eine halbe Stunde täglich außerhalb der Schule zu lesen. Der Anteil an Kindern, die angeben, täglich weniger als 30 Minuten außerhalb der Schule zu lesen, fällt im internationalen Vergleich nur in drei Bildungssystemen

signifikant geringer aus. Zwei Stunden und mehr lesen in Deutschland rund 10 Prozent der Schülerinnen und Schüler, nur in einem Teilnehmerstaat liegt der Wert signifikant höher. Im Vergleich zu 2011 zeigt sich, dass sich der Anteil der Wenigleser in fast allen betrachteten Teilnehmerstaaten und -regionen erhöht und der Anteil der Vielleser verringert. Dies ist auch in Deutschland so. Leistungsschwache Kinder lesen erwartungsgemäß weniger in ihrer Freizeit als leistungsstarke. Abnahmen der täglichen Lesedauer, wie sie sich im internationalen Trend im Mittel zeigen, ergeben sich in Deutschland allerdings vor allem aus geringeren Lesezeiten der leistungsstärkeren Schülerinnen und Schüler.

Bibliotheksnutzung. Gut 50 Prozent aller Schülerinnen und Schüler in Deutschland leihen sich zumindest ein- bis zweimal im Monat Bücher aus. Bibliotheken spielen also im Leseverhalten der Kinder am Ende der Grundschulzeit auch in 2016 durchaus eine Rolle. Der internationale Vergleich zeigt aber, dass Bibliotheken in den übrigen Bildungssystemen deutlich häufiger für die Buchausleihe besucht und genutzt werden. Zwischen 2006 und 2016 ist die Anzahl derjenigen, die nie oder fast nie Bücher aus der Bibliothek auszuliehen, in fast allen Ländern gestiegen (in Deutschland 8 Prozentpunkte). Diese Veränderung betrifft in Deutschland alle drei Leseleistungsgruppen, die leseschwachen Schülerinnen und Schüler allerdings am stärksten. Zugleich ist der Anteil der häufigen Bibliotheksnutzer bei den Schülerinnen und Schülern auf den Kompetenzstufen I und II sowie III signifikant gesunken.

Insgesamt gesehen zeigen sich die Schülerinnen und Schüler bei IGLU 2016 als durchaus selbstbewusste und motivierte Leserinnen und Leser. Zugleich zeigt sich im internationalen Trend, dass in Deutschland wie in den meisten Teilnehmerstaaten und -regionen Abnahmen in der Häufigkeit und Intensität des Lesens zu beobachten sind. In Deutschland betrifft dies insbesondere die Gruppe der leseschwachen Schülerinnen und Schüler, bei denen auch vergleichsweise ungünstige Änderungen in der Lesemotivation zu beobachten sind – dieser Befund kann nicht zufriedenstellend sein und sollte zum Anlass gezielter Leseförderung genommen werden.

Bedingungsfaktoren schwacher Leseleistungen. Logistische Mehrebenenregressionen, die den Beitrag familiärer und individueller Variablen zur Vorhersage der Zugehörigkeit zur Gruppe der leseschwachen Kinder modellieren, verdeutlichen, dass in allen vier Erhebungszyklen von IGLU für Kinder aus Akademikerfamilien im Vergleich zu Kindern, deren Eltern einer niedrigeren Berufsstatusgruppe zuzuordnen sind, ein deutlich reduziertes Risiko besteht, den Leistungsschwachen anzugehören. Demgegenüber besteht für Kinder mit Migrationshintergrund, insbesondere bei beiden im Ausland geborenen Elternteilen, im Vergleich zu Kindern, deren Eltern in Deutschland geboren sind, ein deutlich erhöhtes Risiko, den Leistungsschwachen anzugehören. Ein hohes Leseselbstkonzept, häufiges Lesen zum Vergnügen und ein hohes Maß an leseförderlichen Ressourcen und Aktivitäten im Elternhaus mindern demgegenüber beträchtlich das Risiko. Diese Variablen können somit mögliche Ansatzpunkte für eine Förderung leseschwacher Schülerinnen und Schüler darstellen, die sich in ihren Wirkungen ergänzen. Diesem Befund wäre in vertiefenden Analysen und auf Basis längsschnittlicher Daten weiter nachzugehen. Das Ergebnis, dass der Einfluss des Bildungsniveaus in den Familien vollständig über die Lesesozialisation und die individuellen Lesemerkmale zu erklären ist und sich der Effekt des Berufsstatus bei Berücksichtigung dieser Variablen reduziert, lässt sich als Hinweis auf Unterschiede in den kulturellen Praktiken in den Familien lesen. Sozioökonomisch schlechter gestellte und weniger

bildungsnahe Elternhäuser könnten somit von Informationen profitieren, wie sie informelle und formelle Bildungsangebote für die Leseförderung besser nutzen können. Weitere Möglichkeiten hierzu können darin liegen, dass Schulen stärker die Beteiligungen von Eltern ermöglichen und Zugänge zu Wissen und Abstimmungen von schulischen und familiären Fördermöglichkeiten eröffnen. Hiervon können insbesondere Kinder aus sozial benachteiligten Familien und Kinder mit Migrationshintergrund profitieren, wobei die Möglichkeiten, Eltern unterschiedlicher soziokultureller Herkunft entsprechend ihren Bedürfnissen zu beteiligen, mit der Schulform zusammenhängen und vor allem an rhythmisiert organisierten Ganztagsgrundschulen bestehen dürften (Lokhande, Hoefl & Wendt, 2014). Leseförderprogramme, die als Kooperationen zwischen Elternhaus und Schule gestaltet sind, haben sich zudem als effektiver erwiesen als solche, die rein schulisch organisiert waren (Villiger, Niggli, Wandeler & Kutzelmann, 2012).

Die Befunde zu Selbstkonzept, Motivation, und Verhalten und Leistungen verweisen darauf, dass eine umfassende Leseförderung nicht an einem einzelnen dieser Merkmale ansetzen, sondern diese gemeinsam und aufeinander bezogen berücksichtigen sollte, wobei Schulen wie Familien relevante Kontexte darstellen. Hierbei sollte im Blick behalten werden, dass die Förderung der Lesefähigkeiten als solche wiederum positive Wirkungen auf das Selbstkonzept, die Motivation und die Verhaltensweisen der Schülerinnen und Schüler haben kann, insbesondere, wenn diese ihre Leistungen als Erfolge eigener Anstrengungen erleben können. Wie Rosebrock und Nix (2017) aufzeigen, genügt es für schwache Leserinnen und Leser nicht, ihnen Gelegenheiten und Anreize zum intensiven Lesen zu geben, sie brauchen zusätzliche Fördermaßnahmen zur Verbesserung ihrer Lesekompetenz, bei denen folgende Aspekte relevant sein können: (1) auf den individuellen Lernstand im Lesen abgestimmte Aufgaben, (2) Erwerb von Lesestrategien und Arbeitstechniken, (3) Maßnahmen zur Steigerung der Lesemotivation (z.B. Anknüpfen an die Interessen des Kindes, Ermöglichen von Kompetenzerlebnissen, Angebot von reichhaltigen Lesematerialien, kooperative Lernformen), (4) Berücksichtigung des Selbstbilds des Kindes und (5) Zusammenarbeit mit den Eltern (siehe auch Valtin, 2006; zu Fördermöglichkeiten im Leseunterricht siehe auch Kapitel 10 in diesem Band, zur Leseförderung im europäischen Vergleich siehe Kapitel 12). Multidimensionale und dynamische Ansätze zur Förderung des Lesens finden sich auch in der Forschung zum *Reading Engagement*, das affektive (wie die intrinsische Lesemotivation) und verhaltensbezogene Aspekte (wie das Lesen zum Vergnügen) mit kognitiven (wie Lesestrategien) und sozialen Aspekten (wie den Austausch über Gelesenes in der Schule oder Freizeit) zusammenführt und auf dessen Basis Leseförderansätze wie CORI (*Concept-Oriented Reading Instruction*) entwickelt wurden (Guthrie, Wigfield & You, 2012; Wigfield, 2005). Empirische Befunde (ebd.; Wigfield, Gladstone & Turci, 2017) sprechen für die Wirksamkeit dieses Ansatzes.

Literatur

- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., Schneider, W., Stanat, P., Ostermeier, C., Schiefele, U., Valtin, R., Ring, K. & Saalbach, H. (2007). *Förderung von Lesekompetenz – Expertise*. Bonn: Bundesministerium für Bildung und Forschung.
- Baumert, J., Watermann, R. & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und individuelles Mediationsmodell. *Zeitschrift für Erziehungswissenschaft*, 6 (1), 46–71.
- Betz, D. & Breuninger, H. (1998). *Teufelskreis Lernstörungen* (5. Auflage). Weinheim: Beltz.
- 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.
- Buddeberg, I., Stubbe, T. C. & Potthoff, B. (2008). Lesesozialisation im Elternhaus 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-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 127–141). Münster: Waxmann.
- Chapman, J. W. & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educational Psychology*, 67, 279–291.
- Chapman, J. W., Tunmer, W. E. & Prochnow, J. E. (2000). Early reading-related skills and performance, reading self-concept, and the development of academic self-concept: A longitudinal study. *Journal of Educational Psychology*, 92 (4), 703–708.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.
- Durik, A. M., Vida, M. & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology*, 98 (2), 382–393.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives: Psychological and sociological approaches* (S. 75–146). San Francisco, CA: Freeman.
- Feierabend, S., Plankenhorn, T. & Rathgeb, T. (2017). *KIM-Studie 2016. Kindheit, Internet, Medien. Basisstudie zum Medienumgang 6- bis 13-Jähriger in Deutschland*. Stuttgart: Medienpädagogischer Forschungsverbund Südwest.
- Guthrie, J. T., Wigfield, A. & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. L. Christenson, A. L. Reschly & C. Wylie (Hrsg.), *Handbook of research on student engagement* (S. 601–634). New York: Springer.
- Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.
- Hurrelman, B. (2004). Sozialisation der Lesekompetenz. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz: Vertiefende Analysen im Rahmen von PISA 2000* (S. 37–60). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- König, J., Wagner, C. & Valtin, R. (2011). *Jugend-Schule-Zukunft. Psychosoziale Bedingungen der Persönlichkeitsentwicklung. Ergebnisse der Längsschnittstudie AIDA*. Münster: Waxmann.
- KMK – Ständige 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 10.11.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/GTS_2015_Bericht.pdf
- Kraayenoord, C. van & Schneider, W. (1999). Reading achievement, metacognition, reading self-concept and interest: A study of German students in grades 3 and 4. *European Journal of Psychology of Education*, 14 (3), 305–324.

- Lokhande, M., Hoeft, M. & Wendt, H. (2014). *Eltern als Bildungspartner: Wie Beteiligung an Grundschulen gelingen kann*. Berlin: Sachverständigenrat deutscher Stiftungen für Integration und Migration.
- McElvany, N. (2011). Familiäre Bedingungsfaktoren von Lesekompetenz und Effektivität systematischer Förderung. In J. Bothe (Hrsg.), *Funktionaler Analphabetismus im Kontext von Familie und Partnerschaft* (S. 62–71). Münster: Waxmann.
- McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41 (3), 121–131.
- McElvany, N., Kortenbruck, M. & Becker, M. (2008). Lesekompetenz und Lesemotivation. Entwicklung und Mediation des Zusammenhangs durch Leseverhalten. *Zeitschrift für Pädagogische Psychologie*, 22 (3–4), 207–219.
- McGeown, S. P., Osborne, C., Warhurst, A., Norgate, R. & Duncan, L. G. (2016). Understanding children's reading activities: Reading motivation, skill and child characteristics as predictors. *Journal of Research in Reading*, 39 (1), 109–125.
- Mielke, R., Goy, M. & Pietsch, M. (2006). Das Leseselbstkonzept am Ende der Grundschulzeit. In W. Bos & M. Pietsch (Hrsg.), *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen* (S. 87–109). Münster: Waxmann.
- Mol, S. E. & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137 (2), 267–296.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55 (1), 19–27.
- Möller, J. & Trautwein, U. (2015). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (2. Auflage, S. 177–199). Berlin: Springer.
- Möller, J. & Schiefele, U. (2004). Motivationale Grundlagen der Lesekompetenz. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz: Vertiefende Analysen im Rahmen von PISA 2000* (S. 101–124). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mullis, I. V. S. & Martin, M. O. (Hrsg.). (2015). *PIRLS 2016 assessment framework* (2. Auflage). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Zugriff am 23.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/international-results/>
- Pekrun, R. & Helmke, A. (1991). Schule und Persönlichkeitsentwicklung: Theoretische Perspektiven und Forschungsstand. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung* (S. 33–56). Stuttgart: Enke.
- Pfost, M., Dörfler, T. & Artelt, C. (2010). Der Zusammenhang zwischen außerschulischem Lesen und Lesekompetenz. Ergebnisse einer Längsschnittstudie am Übergang von der Grund- in die weiterführende Schule. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42 (3), 167–176.
- Pfost, M., Dörfler, T. & Artelt, C. (2012). Reading competence development of poor readers in a German elementary school sample: An empirical examination of the Matthew effect model. *Journal of Research in Reading*, 35 (4), 411–426.
- Retelsdorf, J., Köller, O. & Möller, J. (2014). Reading achievement and reading self-concept – Testing the reciprocal effects model. *Learning and Instruction*, 29, 21–30.
- Retelsdorf, J. & Möller, J. (2008). Familiäre Bedingungen und individuelle Prädiktoren der Lesekompetenz von Schülerinnen und Schülern. *Psychologie in Erziehung und Unterricht*, 55 (4), 227–237.
- Rigney, D. (2010). *The Matthew effect. How advantage begets further advantage*. New York: Columbia University Press.
- Rosebrock, C. & Nix, D. (2017). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (8., korrigierte Auflage). Baltmannsweiler: Schneider Verlag Hohengehren.
- Rosenfeld, H. & Valtin, R. (1997). Zur Entwicklung schulbezogener Persönlichkeitsmerkmale bei Kindern im Grundschulalter. Erste Ergebnisse aus dem Projekt NOVARA. *Unterrichtswissenschaft*, 25 (4), 316–330.

- Schaffner, E., Schiefele, U. & Schmidt, M. (2013). Die Bedeutung des familiären Hintergrundes für die Lesemotivation und Lesehäufigkeit von Gymnasialschülern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 45 (3), 131–141.
- Schiefele, U. & Köller, O. (2010). Intrinsische und extrinsische Motivation. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Auflage, S. 336–344). Weinheim: Beltz.
- Schiefele, U. & Schaffner, E. (2015). Motivation. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 153–175). Heidelberg: Springer.
- Schiefele, U., Schaffner, E., Möller, J. & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, 47 (4), 427–463.
- Schiefele, U., Stutz, F. & Schaffner, E. (2016). Longitudinal relations between reading motivation and reading comprehension in the early elementary grades. *Learning and Individual Differences*, 51, 49–58.
- Sénéchal, M. (2012). Child language and literacy development at home. In B. H. Wasik (Hrsg.), *Handbook of family literacy* (2. Auflage, S. 38–50). New York: Routledge.
- Simpkins, S. D., Fredricks, J. A. & Eccles, J. S. (2015). The role of parents in the ontogeny of achievement-related motivation and behavioral choices: I. Introduction. *Monographs of the Society for Research in Child Development*, 80 (2), 1–22.
- Stanat, P. & Schneider, W. (2004). Schwache Leser unter 15-jährigen Schülerinnen und Schülern in Deutschland: Beschreibung einer Risikogruppe. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz: Vertiefende Analysen im Rahmen von PISA 2000* (S. 243–273). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21 (4), 360–407.
- Stubbe, T. C., Buddeberg, I., Hornberg, S. & McElvany, N. (2007). Lesesozialisation im Elternhaus im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 299–327). Münster: Waxmann.
- Stutz, F., Schaffner, E. & Schiefele, U. (2016). Relations among reading motivation, reading amount, and reading comprehension in the early elementary grades. *Learning and Individual Differences*, 45, 101–113.
- Tarelli, I. (2010). *Zur Bedeutung und Erfassung der Intensität von häuslicher Lesesozialisation für die Entstehung sozialer Disparitäten im Lesekompetenzerwerb im Grundschulalter*. Dissertation, Technische Universität Dortmund. Zugriff am 23.11.2017 unter https://eldorado.tu-dortmund.de/bitstream/2003/35977/1/Dissertation_Tarelli.pdf
- Valtin, R. (2006). Förderung von Kindern mit Schwierigkeiten beim Schriftspracherwerb (LRS). In B. Hofmann & A. Sasse (Hrsg.), *Legasthenie. Lese-Rechtschreibstörungen oder Lese-Rechtschreibschwierigkeiten? Theoretische Konzepte und praktische Erfahrungen mit Förderprogrammen* (S. 127–148). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- Valtin, R., Hornberg, S., Buddeberg, M., Voss, A., Kowoll, M. E. & Potthoff, B. (2010). Schülerinnen und Schüler mit Leseproblemen – eine ökosystemische Betrachtungsweise. 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. 43–90). Münster: Waxmann.
- Valtin, R., Wagner, C. & Schwippert, K. (2005). Schülerinnen und Schüler am Ende der vierten Klasse – schulische Leistungen, lernbezogene Einstellungen und außerschulische Lernbedingungen. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU – Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 187–238). Münster: Waxmann.
- Villiger, C., Niggli, A., Wandeler, C. & Kutzelnann, S. (2012). Does family make a difference? Mid-term effects of a school/home-based intervention program to enhance reading motivation. *Learning and Instruction*, 22 (2), 79–91.

- Wagner, C. & Valtin, R. (2003). Noten oder Verbalbeurteilungen? Die Wirkung unterschiedlicher Bewertungsformen auf die schulische Entwicklung von Grundschulkindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 45 (1), 27–36.
- Wagner, C. & Valtin, R. (2004). Determinants of self-esteem in young adolescents In J. Baumert, H. W. Marsh, U. Trautwein & G. E. Richards (Hrsg.), *Self-concept, motivation and identity: Where to from here? Proceedings of the Third International Biennial SELF Research Conference [CD-ROM]*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Wigfield, A. (2005). Concept oriented reading instruction – CORI. Ein Programm zur Förderung der Lesemotivation im Unterricht. *Unterrichtswissenschaft*, 33 (2), 106–121.
- Wigfield, A., Gladstone, J. & Turci, L. (2017). Development of children’s ability beliefs and values and school-based interventions to improve them. In F. Guay, H. W. Marsh, D. M. McInerney & R. Craven (Hrsg.), *SELF – Driving positive psychology and well-being* (S. 283–308). Charlotte, NC: Information Age Publishing.
- Wigfield, A., Rosenzweig, E. Q. & Eccles, J. (2017). Achievement values: Interactions, interventions, and future directions. In A. Elliot, C. Dweck & D. Yeager (Hrsg.), *Handbook of competence and motivation: Theory and application* (2. Auflage, S. 116–134). New York: Guilford Press.

Kapitel V

Geschlecht und Lesekompetenz

Nele McElvany, Ursula Kessels, Franziska Schwabe und Daniel Kasper

1 Einleitung

Lesekompetenz ist eine zentrale Voraussetzung für Bildungskarrieren von Jungen und Mädchen. Obwohl die Bildungsbeteiligung der ursprünglich in Bildungsfragen benachteiligten Mädchen in Deutschland im Zuge der Bildungsexpansion stark angestiegen ist, ist die Thematik geschlechtsspezifischer Disparitäten im Zusammenhang mit Bildung auch heute noch unter veränderten Bedingungen ein zentrales Thema der Bildungsforschung (Autorengruppe Bildungsberichterstattung, 2010; Kessels, Heyder & Holder, 2016). Statistiken zeigen, dass heutzutage Jungen an Gymnasien unterrepräsentiert sind, seltener das Abitur machen, häufiger die Schule ohne einen Abschluss verlassen, die Klasse wiederholen und in Förder- und Hauptschulen vertreten sind (zusammenfassend Hannover & Kessels, 2011; Helbig, 2012; Steinmayr & Spinath, 2008). Hinsichtlich konkreter schulischer Kompetenzen berichteten internationale und nationale Schulleistungsuntersuchungen, dass Mädchen im Mittel in allen Teilnehmerstaaten und -regionen deutlich höhere Lese- und Schreibkompetenzen aufweisen als Jungen, wobei Jungen in den Bereichen Naturwissenschaften und Mathematik in vielen Ländern höhere Kompetenzwerte als Mädchen erzielen (Bos, Bremerich-Vos, Tarelli & Valtin, 2012; Hannover, Wolter & Zander, 2017; Wendt, Steinmayr & Kasper, 2016).

Zu den Gründen für die Geschlechterunterschiede im Bildungserfolg existieren unterschiedliche Annahmen (Hannover & Kessels, 2011). Grundsätzlich kann zwischen Theorien, die Geschlechterunterschiede überwiegend auf individuelle und soziale Prozesse zurückführen, und evolutionär-biologischen Theorien, die Geschlechterdifferenzen mit angeborenen Unterschieden begründen, unterschieden werden (vgl. Eagly & Wood, 1999; Hyde, 2014). Zu den bildungsbezogenen Annahmen gehört zudem unter anderem die Argumentation, Ursache für den geringeren Schulerfolg der Jungen sei die „Feminisierung

der Schule“ und das dortige Fehlen männlicher Vorbilder (Neugebauer, 2011). Allerdings zeigten empirische Befunde keine Wirkung des Geschlechts der Lehrperson auf den Schulerfolg von Schülerinnen und Schülern (Hannover et al., 2017; Kuhl & Hannover, 2012). Besser belegt ist der Erklärungsansatz, dass der geringere Schulerfolg von Jungen teilweise ihrem geringeren schulischen Engagement geschuldet ist, das möglicherweise auch in der Stereotypisierung von schulischer Anstrengung als unmännlich begründet ist (Heyder & Kessels, 2017). Verschiedene Studien wiesen darauf hin, dass Mädchen viele positiv mit Schulerfolg korrelierte Merkmale und Verhaltensweisen (z.B. intrinsische Motivation, Wertschätzung von Schule, Selbstdisziplin, Hausaufgabenerledigung) in einem größeren Ausmaß zeigten als Jungen und dass dies ihre besseren Noten teilweise erklären konnte (Hannover & Kessels, 2011; Steinmayr & Spinath, 2008; Valtin, Wagner & Schwippert, 2005).

Ein weiterer möglicher Grund für den insgesamt geringeren Schulerfolg von Jungen könnte auch ihre geringer ausgeprägte Lesekompetenz sein, da der Lesekompetenz eine grundlegende Bedeutung für das weitere Erarbeiten von Inhalten in allen Schulfächern und Domänen zugeschrieben wird (Artelt, Stanat, Schneider & Schiefele, 2001; Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Nicht nur in den sprachlichen, sondern in allen Fächern werden die Wissensbestände zu einem großen Teil über Texte vermittelt. Auch für Geschlechterunterschiede in der Lesekompetenz werden motivationale Ursachen mitverantwortlich gemacht. So deutete bereits eine Vielzahl von Studien auf eine größere Lesemotivation, ein positiveres lesebezogenes Selbstkonzept und ein höheres Leseverhalten der Mädchen im Vergleich zu den Jungen hin (aktuell z.B. Becker & McElvany, 2017).

Im Folgenden werden einleitend vorliegende Untersuchungen zu Geschlechterunterschieden in der Lesekompetenz sowie zu möglichen Ursachen für die geringeren Leistungen von Jungen in diesem Bereich dargestellt. Anschließend folgen die aktuellen Ergebnisse zu Geschlechterunterschieden in IGLU 2016.

2 Forschungsstand

2.1 Geschlechterunterschiede in der Lesekompetenz

Lesekompetenz ist als ein zentrales kulturelles Werkzeug eine entscheidende Voraussetzung für schulischen und beruflichen Erfolg sowie lebenslanges Lernen. Als fächerübergreifende Schlüsselkompetenz ist kompetentes Lesen nicht nur Ziel, sondern auch wichtige Voraussetzung für das Lernen in den meisten Schulfächern. Lesen stellt dabei einen der vier übergeordneten Bereiche von Sprachkompetenzen dar, die sich in rezeptive und produktive sowie schriftliche und mündliche Bereiche systematisieren lassen. Das systematische Lernen des Lesens – und damit des Sprachkompetenzbereichs, der durch Rezeption und Schriftlichkeit gekennzeichnet ist – beginnt für die meisten Kinder in der ersten Schulklasse. Dabei wird aufgrund der regulären Orthographie der deutschen Sprache der Übergang von der alphabetischen zu der orthographischen Stufe des Lesens bei normaler Entwicklung in der dritten Klasse vollzogen (vgl. Stufenmodell von Frith, 1986; Christmann & Groeben, 1999). Nach dem Fokus auf den Erwerb von basalen Lesefähigkeiten in den ersten Schuljahren stehen spätestens ab dem vierten Schuljahr zunehmend komplexere Lese- und

Verstehensprozesse im Mittelpunkt, die schwerpunktmäßig auf den kompetenten Umgang mit Texten abzielen und eine Grundlage für das Lernen in anderen Schulfächern darstellen. Unter Lesekompetenz wird dabei über das Entziffern von Buchstaben, Wörtern und Sätzen hinaus die Fähigkeit der aktiven Auseinandersetzung einschließlich Bedeutungsgenerierung – mit lokaler und globaler Kohärenzbildung – mit Geschriebenem verstanden, wobei die verschiedenen Teilprozesse teilweise automatisiert und teilweise bewusst gesteuert ablaufen (Graesser, Singer & Trabasso, 1994). Das Textverstehen kann mit Kintsch (1998) als Kombination von textbasierten Konstruktions- und vorwissensbasierten Integrationsprozessen und damit als kognitiv-aktive Konstruktionsleistung des Individuums aufgefasst werden (siehe ausführlich Kapitel 3 in diesem Band).

Auch wenn Lesekompetenzen im Regelfall in der Schule erworben werden und ihre Ausbildung somit von Lehrerkompetenzen, Unterrichtsqualität und Klassenkontext abhängt, sind sie zudem durch Merkmale der sozialen und familiären Umwelt sowie durch Merkmale des Individuums mitbestimmt (vgl. Helmke & Weinert, 1997). Hierzu gehört die Geschlechtszugehörigkeit. Für die Untersuchung von Geschlechterunterschieden in der Lesekompetenz kann auf eine längere Forschungstradition zurückgeblickt werden. Hyde (2005, 2014) verweist in ihren Überblicksarbeiten zu Geschlechterähnlichkeiten beziehungsweise -unterschieden darauf, dass sich die Geschlechterunterschiede in der Lesekompetenz substantiell zwischen verschiedenen Ländern unterscheiden und außerdem eine deutliche Entwicklungsdynamik aufzuweisen scheinen. Die IGLU-Ergebnisse zeigten für die Erhebung 2011 für die Mehrheit der teilnehmenden Staaten und Regionen Geschlechterunterschiede und in allen Staaten mit signifikanten Unterschieden fielen diese zugunsten der Mädchen aus (Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Für Deutschland lag der mittlere Vorteil der Mädchen bei acht Punkten und war damit ähnlich hoch wie der mittlere Unterschied von zwölf Punkten in allen Teilnehmerstaaten und -regionen, die zur EU beziehungsweise OECD gehören. Betrachtet man dieses Ergebnis in Relation zu früheren Erhebungen, so ist festzuhalten, dass es keine statistisch signifikante Veränderung der Geschlechterdisparitäten in Deutschland im Laufe der Zeit gab: 2001 hatte der Vorsprung der Mädchen bei 13 Punkten und 2006 bei sieben Punkten gelegen. Dieser Befund entspricht dem internationalen Bild, bei dem sich bei dem Vergleich der Geschlechterdisparitäten 2001 und 2011 nur für zwei Länder Veränderungen zugunsten der Jungen (Niederlande, Schweden) und in einem Land (Russische Föderation) zugunsten der Mädchen und ansonsten keine Veränderungen zeigten (Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Der *IQB-Bildungstrend 2016* zeigte übereinstimmend für Deutschland gleichbleibende Geschlechterdifferenzen im Lesen von 2011 zu 2016 (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017).

Für die im Rahmen der PISA-Studie untersuchten 15-jährigen Schülerinnen und Schüler ergaben sich ebenfalls in vielen Ländern Geschlechterunterschiede in der Lesekompetenz, die überwiegend deutlich höher waren als die bei IGLU berichteten Unterschiede für jüngere Kinder. Für Deutschland lagen die Unterschiede in den verschiedenen Erhebungsjahren (ebenfalls auf einer Skala mit $M = 500$ und $SD = 100$) bei 34 Punkten (2000), 42 Punkten (2003), 42 Punkten (2006), 40 Punkten (2009; vgl. den Mittelwert der teilnehmenden OECD-Länder mit 39 Punkten) sowie 44 Punkten (2012; Hohn, Schiepe-Tiska, Sälzer & Artelt, 2013; Weis, Zehner, Sälzer, Strohmaier, Artelt & Pfof, 2016). Eine substantielle Reduktion des Geschlechterunterschieds in Deutschland auf 21 Punkte wurde für die Erhebung 2015 berichtet (Weis et al., 2016). Diese Reduktion wurde zwar

nicht für alle Länder gleichermaßen beobachtet, insgesamt verringerte sich aber im Vergleich zu 2012 auch der Mittelwert der Geschlechterunterschiede in den teilnehmenden OECD-Ländern von 37 auf 28 Punkte. Eine mögliche Erklärung für diese deutliche Reduktion der Geschlechterunterschiede zwischen den PISA-Erhebungen 2012 und 2015 ist die zwischenzeitliche Umstellung von papierbasierter (2012) zu computerbasierter (2015) Testung, die möglicherweise Vorteile für die Jungen mit sich brachte, deren Leistungen statistisch signifikant anstiegen.

Während demnach bei jüngeren Schülerinnen und Schülern die Geschlechterunterschiede in der Lesekompetenz im Mittel eher gering und in vielen Ländern nicht vorhanden sind (vgl. Mullis et al., 2012), sind bei 15-jährigen Schülerinnen und Schülern, welche sich in der Regel am Ende ihrer Pflichtschulzeit befinden, konstant deutliche Geschlechterunterschiede zugunsten der Mädchen zu finden (vgl. OECD, 2014). Auch wenn die Kompetenztests der verschiedenen Studien nicht direkt vergleichbar sind (vgl. Artelt, Drechsel, Bos & Stubbe, 2008), unterstützt dieses Befundmuster eher Theorien, die Geschlechterunterschiede überwiegend auf individuelle und soziale Prozesse zurückführen als auf evolutionär-biologische Annahmen. Hierzu passt auch, dass die Geschlechterunterschiede in der Lesemotivation bereits am Ende der Grundschulzeit substantiell sind, während die Kompetenzunterschiede erst zeitlich nachgelagert eintreten (siehe Abschnitt 2.2; vgl. Baker & Wigfield, 1999; Logan & Johnston, 2009).

Weiterführende Analysen zu den aktuellen PISA 2015-Daten verdeutlichten, dass der Anteil der Jungen in der Gruppe der schwachen Leserinnen und Leser (PISA-Kompetenzstufe Ia oder darunter), der Anteil der Mädchen hingegen in der Gruppe der starken Leserinnen und Leser (PISA-Kompetenzstufen V und VI) signifikant höher als der Anteil des jeweils anderen Geschlechts war (Weis et al., 2016). Naumann, Artelt, Schneider und Stanat (2010) berichteten auf Basis der PISA-2009-Daten, dass für alle drei untersuchten Bereiche der Lesekompetenz – Suchen und Extrahieren von Informationen; Kombinieren und Interpretieren; Reflektieren und Bewerten – gleichermaßen substantielle Geschlechterunterschiede bestanden (siehe auch Kapitel 3 in diesem Band zu den Bereichen bei IGLU 2016). Zudem zeigten sich die Geschlechterdisparitäten, die bei PISA 2000 bei nichtkontinuierlichen Texten noch deutlich kleiner als bei kontinuierlichen Texten gewesen waren (Stanat & Kunter, 2001), 2009 auch bei nichtkontinuierlichen Texten. Der relative Vorteil der Jungen in diesem Bereich wurde demnach inzwischen aufgehoben.

Weitere Hinweise zu systematischen geschlechtsbezogenen Vorteilen in spezifischen Bereichen lieferten Schwabe, McElvany und Trendtel (2015), indem sie zeigten, dass sowohl bei jüngeren (IGLU 2011) als auch bei älteren (PISA 2009) Schülerinnen und Schülern Mädchen bei Leseaufgaben mit offenem Antwortformat besser abschnitten als Jungen mit vergleichbar guter Gesamtlesekompetenz. Diese Vorteile der Mädchen bei offenen Antwortformataufgaben standen zudem bei den 15-Jährigen, aber nicht bei den jüngeren Kindern mit einem höheren Lesemotivationsniveau in Verbindung. Die Lesemotivation wie auch das lesebezogene Selbstkonzept scheinen demnach ein wesentlicher Aspekt zur Erklärung von Geschlechterunterschieden im Bereich der Lesekompetenz zu sein.

2.2 Mögliche Ursachen für Geschlechterunterschiede in der Lesekompetenz: Leseselbstkonzept, Lesemotivation und Leseverhalten

In der Literatur werden verschiedene Ursachen für Geschlechterunterschiede in fachlichen Kompetenzen diskutiert (Hausmann, 2007; Kessels & Heyder, 2018). Im Rahmen der Untersuchung der Lesekompetenz in IGLU werden vor allem solche Faktoren als relevant angesehen, die auf die Lesemotivation und das Leseverhalten bezogen sind (siehe auch Kapitel 4 in diesem Band). Die motivationalen Faktoren lassen sich gut in das empirisch bewährte Erwartungs-Wert-Modell von Eccles et al. (1983) einordnen, in dem leistungsbezogene Ergebnisse durch domänenspezifische Erfolgserwartungen sowie den Wert, der einer Domäne zugemessen wird, erklärt werden. Als Proxy für die Erfolgserwartung gilt das fähigkeitsbezogene Selbstkonzept (Retelsdorf, Köller & Möller, 2011), welches „generalisierte fachspezifische Fähigkeitseinschätzungen“ (Möller & Köller, 2004, S. 19) enthält. Der einer Domäne zugemessene Wert ist konzeptualisiert als eine Kombination des intrinsischen Wertes, der persönlichen Wichtigkeit, der eingeschätzten Nützlichkeit sowie einer negativen Wertkomponente, den mit der Beschäftigung mit einer Domäne verbundenen Kosten (Steinmayr & Spinath, 2010). Die intrinsische Wertkomponente kann beispielsweise durch das tätigkeitsbezogene positive Erleben und das gegenstandsbezogene Interesse erfasst werden (vgl. McElvany, Kortenbruck & Becker, 2008); die bei IGLU 2016 verwendete Skala „Lesemotivation“ bildet vorrangig diesen intrinsischen Wert ab. Zahlreiche Studien konnten für unterschiedliche Altersgruppen zeigen, dass die Leistungen in einem Schulfach oder einer Kompetenzdomäne umso besser sind, je positiver das Selbstkonzept und je höher der diesem Bereich zugesprochene Wert ist (z.B. Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002). Auch konnten Geschlechterunterschiede in leistungsbezogenen Ergebnissen mit diesem Modell erklärt werden. Dabei sind allerdings viele der Studien auf die Domäne Mathematik bezogen, in der Mädchen im Durchschnitt ein geringeres Selbstkonzept und weniger Interesse äußern (zusammenfassend Heyder, Kessels & Steinmayr, 2017). Inzwischen liegen aber auch Studien vor, die die schlechteren Leistungen von Jungen im Fach Deutsch auf die Wert- und Erwartungskomponenten zurückführen (ebd.).

Die vorliegende Empirie zu Geschlechterunterschieden im lesebezogenen Selbstkonzept im Grundschulalter ist nicht eindeutig. Querschnittlich fand die KESS-4-Studie keine Unterschiede im Selbstkonzept der Hamburger Viertklässlerinnen und Viertklässler, obwohl die Mädchen besser lesen konnten als die Jungen (Mielke, Goy & Pietsch, 2006). Einige längsschnittliche Studien berichteten Ähnlichkeiten zu Beginn der Grundschulzeit, die sich im Laufe der Grundschulzeit zu einem Nachteil der Jungen veränderten, deren Selbsteinschätzung sich stärker verschlechtert als die der Mädchen (Jacobs et al., 2002; Marsh, Craven & Debus, 1998; Upadyaya & Eccles, 2015; für Deutschland: Becker & McElvany, 2017). Andererseits zeigte eine Untersuchung von etwa eintausend Erstklässlerinnen und Erstklässlern (Ehm, Duzy & Hasselhorn, 2011) bereits in diesem frühen Alter ein besseres Leseselbstkonzept der Mädchen, welches nur teilweise auf ihren signifikant besseren Leseleistungen beruhte. Die Geschlechterunterschiede im Leseselbstkonzept zeigten sich auch unter Kontrolle der Leseleistung. Auch Wigfield et al. (1997) fanden in einer US-amerikanischen Kohorten-Längsschnittstudie schon zu Beginn der Grundschulzeit

ein besseres Leseselbstkonzept der Mädchen, wobei sich die Größe dieses Unterschiedes über die Grundschulzeit hinweg nicht veränderte. Diese Ergebnisse stehen im Widerspruch zu den ebenfalls aus den USA stammenden längsschnittlichen Befunden sich erst entwickelnder Unterschiede im Laufe der Grundschul- und Sekundarschulzeit (Jacobs et al., 2002; Marsh et al., 1998; Upadyaya & Eccles, 2015). Diese uneindeutige Befundlage zu Geschlechterunterschieden im Leseselbstkonzept im Grundschulalter verweist auf weiteren Forschungsbedarf.

Bei Aspekten, die der Wertkomponente des Erwartungs-Wert-Modells zuzuordnen sind, fanden sich hingegen konsistent Geschlechterunterschiede zugunsten der Mädchen. Mädchen schrieben dem Lesen während der Grund- und Sekundarschulzeit einen höheren Wert zu als Jungen (z.B. Archambault, Eccles & Vida, 2010; Becker, McElvany & Kortenbruck, 2010; Durik, Vida & Eccles, 2006). Die Daten der IGLU- und PISA-Studien belegten die Größe dieser Geschlechterunterschiede: In Deutschland fand sich in der vierten Klasse eine Effektstärke von $d = 1.50$ und bei den 15-Jährigen von $d = 1.49$ (Schwabe et al., 2015). Längsschnittlich wurde gezeigt, dass der intrinsische Wert des Lesens in der Grundschulzeit das freizeitbezogene Leseverhalten am Ende der Sekundarschulzeit vorhersagte (Durik et al., 2006).

Die Häufigkeit, mit der Kinder in der Freizeit lesen, ist ein weiterer relevanter Prädiktor für die Lesekompetenz. Es wird angenommen, dass mit häufigem Lesen ein Zuwachs an Automatisierung und Leseeffektivität sowie Vorwissen einhergeht (zusammenfassend McElvany et al., 2008). Längsschnittlich konnte für den Grundschulbereich gezeigt werden, dass die intrinsische Lesemotivation nicht direkt, sondern vermittelt über das Leseverhalten die spätere Kompetenz beeinflusst und dass sowohl das Leseverhalten auf die spätere Lesekompetenz als auch andersherum die Kompetenz auf das spätere Leseverhalten wirken (McElvany et al., 2008). Die bereits vorliegenden Erhebungszyklen zu IGLU belegten deutliche Geschlechterunterschiede im Leseverhalten: 2006 gaben in Deutschland nur 11 Prozent der Mädchen, aber 29 Prozent der Jungen an, „nie“ außerschulisch *Bücher oder Zeitschriften* zu lesen; und neun Prozent der Mädchen, aber 19 Prozent der Jungen berichteten „nie“ oder „fast nie“ [*irgendetwas*] außerhalb der Schule zum Vergnügen zu lesen (Hornberg, Valtin, Potthoff, Schwippert & Schulz-Zander, 2007; vgl. auch Valtin, Wagner & Schwippert, 2005). 2011 gaben sechs Prozent der Mädchen, aber 16 Prozent der Jungen an, „nie“ oder „fast nie“ [*irgendetwas*] außerhalb der Schule zum Vergnügen zu lesen (Bos, Bremerich-Vos, Tarelli & Valtin, 2012). Im Längsschnitt zeigte sich, dass Mädchen im Laufe der Grundschule zunehmend mehr lesen, Jungen hingegen nicht (Becker & McElvany, 2017).

Zusammengefasst können als mögliche Ursachen für die Geschlechterunterschiede in den Leseleistungen Unterschiede im Selbstkonzept, in der Lesemotivation und dem daraus resultierenden unterschiedlich ausgeprägten Leseverhalten gelten. Im Folgenden wird dargestellt, ob und inwiefern sich Jungen und Mädchen am Ende der Grundschulzeit in IGLU 2016 in ihrer Lesekompetenz sowie ihrem Leseselbstkonzept, in der Lesemotivation und verschiedenen Indikatoren des Leseverhaltens unterscheiden. Darüber hinaus wird weiterführend geprüft, ob sich Geschlechterunterschiede in der Kompetenz noch finden, wenn Leseselbstkonzept, Lesemotivation und Leseverhalten kontrolliert werden.

3 Ergebnisse

3.1 Geschlechterunterschiede in der Lesekompetenz in den Teilnehmerstaaten

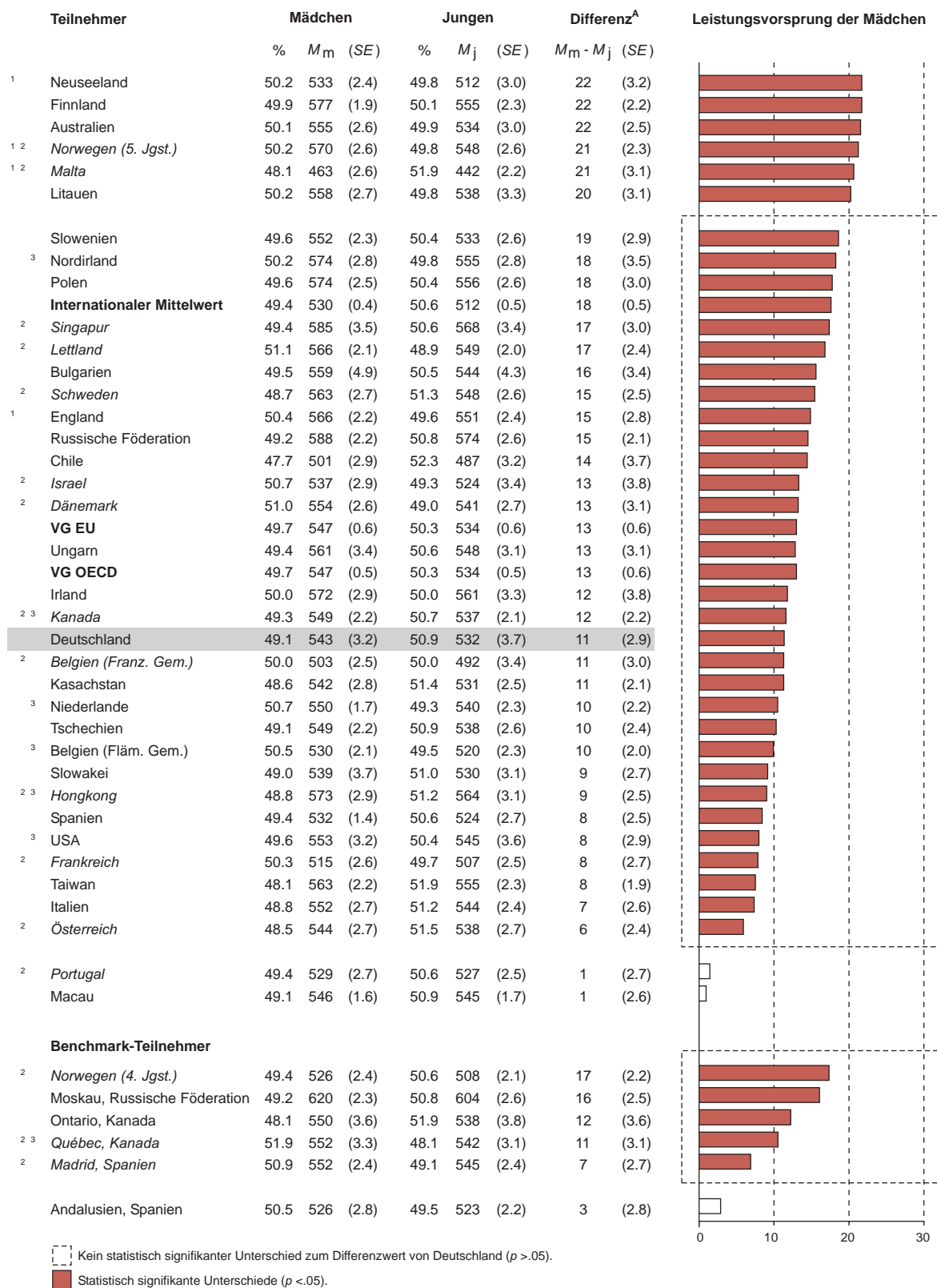
Abbildung 5.1 illustriert die Ergebnisse der Gesamtskala Lesen vergleichend für Mädchen und Jungen. Der linke Teil der Abbildung stellt den Anteil an Mädchen und Jungen, ihre jeweiligen mittleren Lesekompetenzwerte, deren Standardfehler sowie die Differenzwerte mit Standardfehlern für alle Teilnehmerstaaten, Regionen und zentralen Vergleichsgruppen dar. Die Differenzen zwischen Mädchen und Jungen werden zusätzlich graphisch als Balkendiagramm abgebildet. Die rote Färbung eines Balkens weist auf statistisch signifikante Unterschiede zwischen den Geschlechtsgruppen hin. Der schwarz-gestrichelte Rahmen umfasst diejenigen Teilnehmerstaaten und Regionen, in denen die Geschlechterdifferenz nicht statistisch signifikant von dem Differenzwert in Deutschland abweicht.

Die Abbildung verdeutlicht insbesondere das Ergebnis, dass mit Ausnahme von Andalusien (Spanien), Macau und Portugal in allen Teilnehmerstaaten und -regionen ein statistisch signifikanter Unterschied zwischen den Leseleistungen von Mädchen und Jungen vorliegt und dieser bei allen zum Vorteil der Mädchen ausfällt. In Deutschland erreichen Mädchen durchschnittlich 543 und Jungen 532 Punkte. Die Differenz von rund 11 Punkten ist statistisch signifikant, aber als klein zu bewerten (vgl. Böhme & Weirich, 2012, S. 108). Sechs Teilnehmerstaaten und Regionen weisen statistisch signifikant größere Unterschiede als Deutschland auf. In Neuseeland, Finnland, Australien, Norwegen, Malta und Litauen ist der Differenzwert größer als 20 Punkte. Allein zwei Teilnehmer, Portugal und Macau, weisen mit nur einem Punkt Unterschied statistisch signifikant kleinere Unterschiede als Deutschland auf. Der für Deutschland gemessene Unterschied zwischen den Leistungen der Mädchen und Jungen unterscheidet sich weder signifikant von den mittleren Differenzwerten in den Vergleichsgruppen VG EU und VG OECD noch von dem internationalen Mittelwert. Die Leistungsunterschiede zwischen den Geschlechtern in der Grundschule sind insgesamt als klein zu bezeichnen. Allerdings wachsen sie zur Sekundarstufe an. Bei PISA 2009 wurde für Deutschland eine Differenz von 40 Punkten zugunsten der Mädchen berichtet (vgl. Naumann, Artelt, Schneider & Stanat, 2010, S. 52).

3.2 Lesekompetenzen von Mädchen und Jungen nach Textsorten

Abbildung 5.2 stellt die Lesekompetenzen von Mädchen und Jungen nach Textsorten im internationalen Vergleich dar. Auf der linken Seite der Abbildung werden tabellarisch die Differenzen zwischen Mädchen und Jungen getrennt nach literarischem (Erzähltexte) und informierendem (Sachtexte) Lesen sowie die Differenz aus den jeweiligen Vorsprüngen und alle zugehörigen Standardfehler aufgeführt. Rechts wird das Ergebnismuster zu der errechneten Differenz zwischen den Geschlechterunterschieden in informierendem und literarischem Lesen visualisiert. Dort ist dargestellt, bei welcher Textsorte der Vorsprung der Mädchen größer ist.

Abbildung 5.1: Leseleistung nach Geschlecht – Gesamtskala Lesen



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland ($p > .05$).

Statistisch signifikante Unterschiede ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

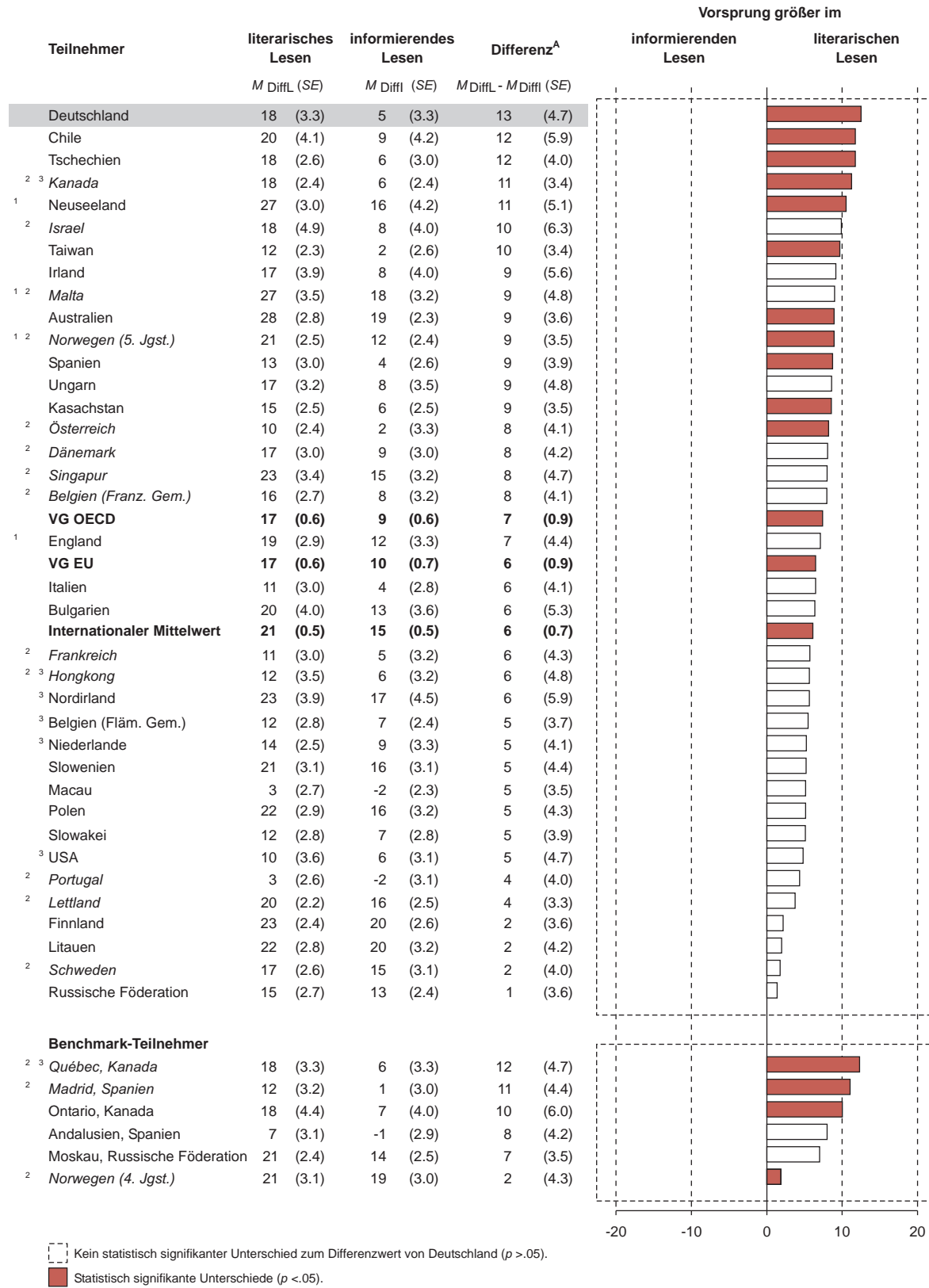
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

Abbildung 5.2: Leistungsvorsprung der Mädchen differenziert nach Textsorten



Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

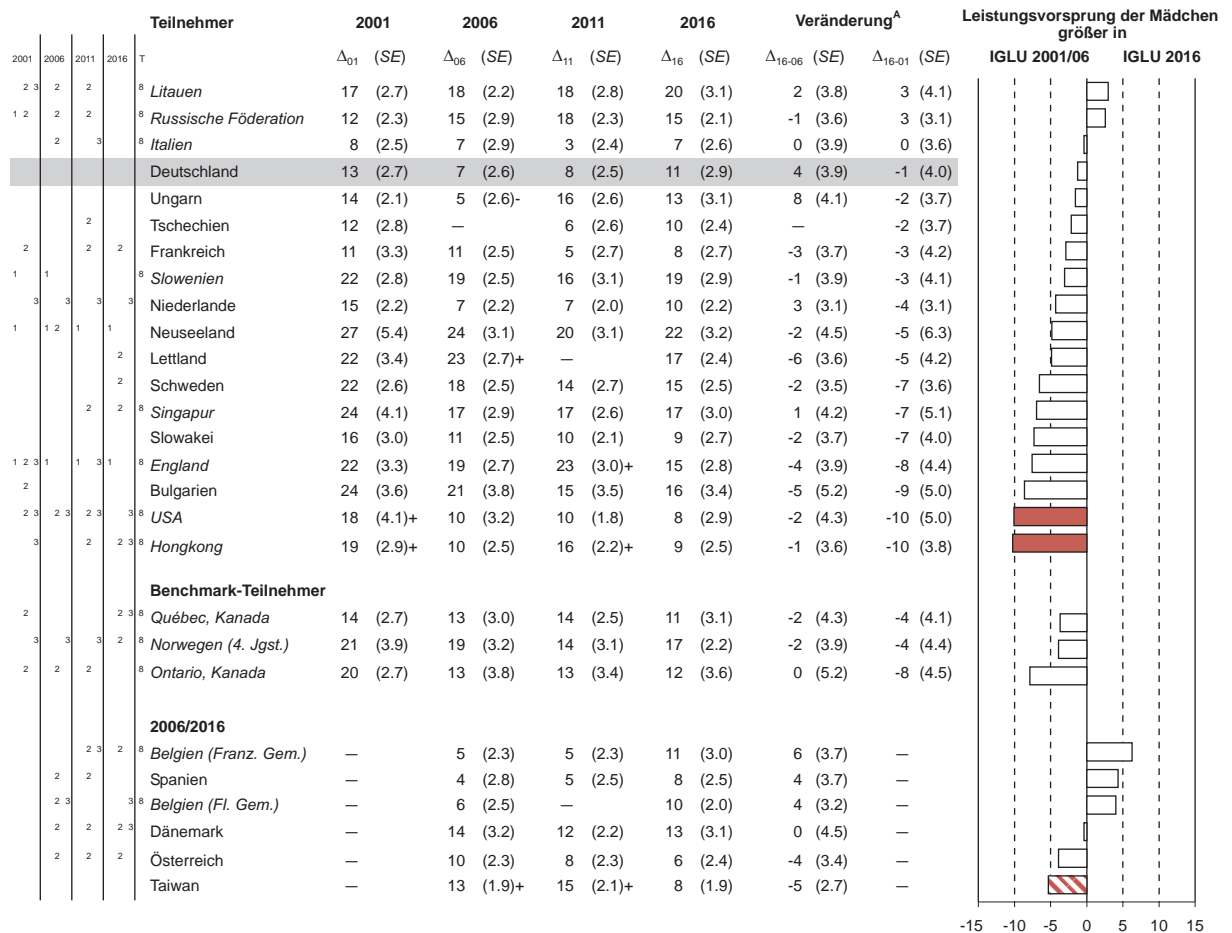
Betrachtet man die Abbildung 5.2, so fällt als erstes auf, dass wenn Geschlechterunterschiede bestehen, diese zugunsten der Mädchen ausfallen. In allen Teilnehmerstaaten und -regionen sind die Geschlechterunterschiede beim Lesen literarischer Texte größer. Auch in Deutschland unterscheiden sich im Jahr 2016 die Leseleistungen von Mädchen und Jungen statistisch bedeutsam nach Textsorten. Der Vorsprung der Mädchen ist beim literarischen Lesen (18 Punkte) deutlich größer als ihr Vorsprung vor den Jungen beim informierenden Lesen (5 Punkte). Dieses Befundmuster unterscheidet sich nicht signifikant von den Geschlechterunterschieden nach Textsorten in allen anderen Teilnehmerstaaten und Regionen, in denen die nominellen Unterschiede ebenfalls auf einen größeren Vorsprung der Mädchen in literarischen Texten hinweisen. In etwa der Hälfte der Teilnehmerstaaten und Regionen fällt dieser Unterschied signifikant aus. Zusammenfassend lässt sich sowohl für Deutschland als auch für die übrigen Teilnehmerstaaten und Regionen schlussfolgern, dass der Vorteil der Mädchen im Lesen insbesondere auf ihre deutlich höheren Leistungen bei literarischen Textsorten zurückzuführen ist, wohingegen sich beim informierenden Lesen kaum Unterschiede zwischen Mädchen und Jungen zeigen.

3.3 Trends in Geschlechterunterschieden in den Leseleistungen zwischen 2001 und 2016

In Abbildung 5.3 sind die Geschlechterunterschiede in der Lesekompetenz im internationalen Vergleich für verschiedene Erhebungszeiträume dargestellt. Der linke Teil der Abbildung zeigt die Geschlechterdifferenz im Lesen nach Jahren für 2001, 2006, 2011 und 2016 getrennt. Berichtet wird die mittlere Differenz in den Teilnehmerstaaten und Regionen sowie der zugehörige Standardfehler. Der rechte Teil der Abbildung visualisiert den Vergleich der Differenzen zwischen 2001 und 2016 beziehungsweise 2006 und 2016. Für diejenigen Teilnehmerstaaten und Regionen, die nicht an IGLU 2001 teilgenommen haben, wird der Differenzwert zu 2006 berechnet. Bei Signifikanz werden diese Fälle rot schraffiert.

Die Abbildung 5.3 veranschaulicht das Ergebnis, dass die Geschlechterunterschiede in der Lesekompetenz bei Viertklässlerinnen und Viertklässlern über 15 Jahre hinweg ähnlich groß geblieben sind. Lediglich in drei Teilnehmerstaaten und Regionen hat sich die Differenz zwischen den Leistungen von Mädchen und Jungen im Lesen statistisch signifikant verändert. In Taiwan, den USA und Hongkong verringerten sich die Unterschiede von 2001 beziehungsweise 2006 zu 2016. Lediglich in fünf Teilnehmerstaaten und Regionen nahmen die Differenzen im Lesen zwischen den Geschlechtern nominell zu. Keine dieser Differenzen ist statistisch signifikant. Auch die Veränderung in Deutschland von IGLU 2001 (13 Punkte) zu IGLU 2016 (11 Punkte) ist statistisch nicht signifikant.

Abbildung 5.3: Vergleich des Leistungsvorsprungs der Mädchen zwischen IGLU 2001, 2006, 2011 und 2016 – Gesamtskala Lesen



■ Statistisch signifikante Veränderungen zwischen 2001 und 2016 ($p < .05$).
 ▨ Statistisch signifikante Veränderungen zwischen 2006 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.

- 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
- 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
- 3 = Die Teilnahmequoten auf Schul- und/oder Schülerbene erreichen nicht die internationalen Vorgaben.
- 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.

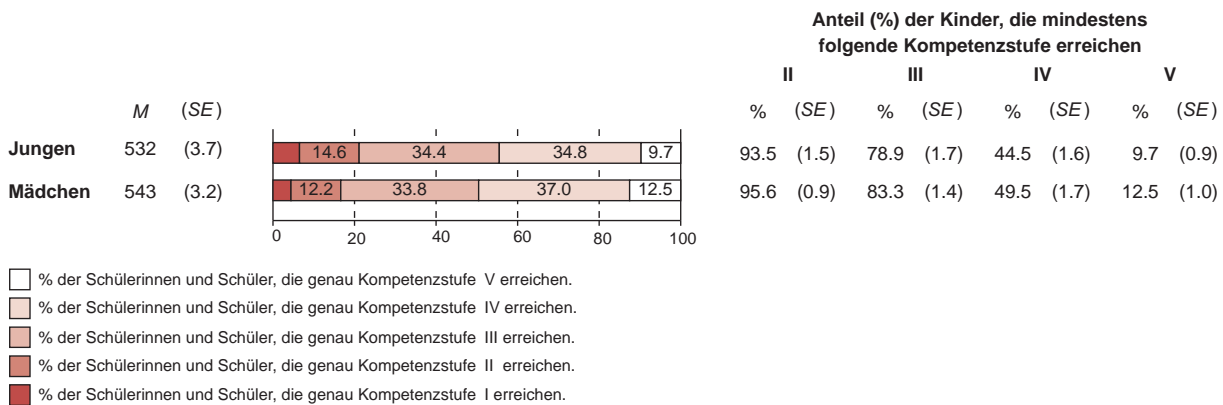
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.
 + = Differenzwert 2001 bzw. 2006 bzw. 2011 statistisch signifikant höher als 2016 ($p < .05$).
 - = Differenzwert 2001 bzw. 2006 bzw. 2011 statistisch signifikant niedriger als 2016 ($p < .05$).

$\Delta_{16-01} / \Delta_{16-06}$ = Differenz in den Mittelwerten zwischen 2016 und 2001 bzw. 2006 (bei denjenigen Teilnehmern, die nicht an IGLU 2001 teilgenommen haben).

3.4 Lesekompetenzen von Mädchen und Jungen in Deutschland nach Kompetenzstufen

In Abbildung 5.4 sind für Deutschland die Anteile von Mädchen und Jungen nach erreichten Kompetenzstufen dargestellt. Im linken Teil der Abbildung sind die Anteile der Mädchen und Jungen illustriert, die jeweils genau eine bestimmte Kompetenzstufe erreicht haben. Im rechten Teil der Abbildung werden zusätzlich die kumulierten Anteile angegeben. Mit kumulierten Anteilen sind die Anteile der Mädchen und Jungen gemeint, die mindestens die entsprechende Kompetenzstufe erreicht haben.

Abbildung 5.4: Prozentuale Verteilung von Mädchen und Jungen in Deutschland auf die fünf Lesekompetenzstufen



IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Insgesamt zeigt sich bei Betrachtung der Abbildung eine relativ ausgewogene Verteilung, wobei nominelle Unterschiede zugunsten der Mädchen bestehen. Vergleicht man beispielsweise die mindestens erreichten Kompetenzstufen von Mädchen und Jungen, so zeigt sich, dass Mädchen in den höheren Kompetenzstufen anteilig stärker vertreten sind. 12.5 Prozent der Mädchen, aber nur 9.7 Prozent der Jungen erreichen die höchste Kompetenzstufe V. Die Kompetenzstufe II erreichen lediglich 4.4 Prozent der Mädchen nicht. Bei Jungen fällt dieser Anteil mit 6.5 Prozent nominell höher aus. Die nominellen Unterschiede in den Verteilungen über die Kompetenzstufen sind nicht statistisch signifikant. Dieses Muster unterscheidet sich von den Ergebnissen bei 15-Jährigen. In der PISA-Stichprobe zeigte sich eine Verteilung zuungunsten der Jungen (siehe Abschnitt 2.1).

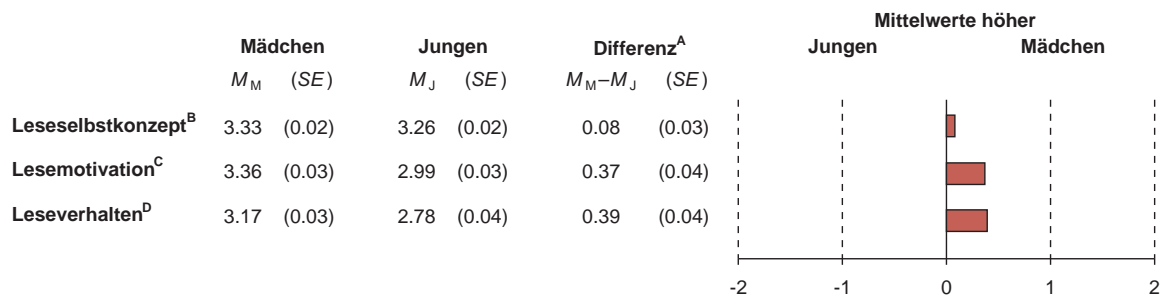
3.5 Leseselbstkonzept, Lesemotivation und Leseverhalten von Mädchen und Jungen in Deutschland

Abbildung 5.5 illustriert für Deutschland Unterschiede in Leseselbstkonzept, Lesemotivation und Leseverhalten im Jahr 2016 nach Geschlecht. Der linke Teil der Abbildung stellt in tabellarischer Form die Mittelwerte zu den drei Konstrukten getrennt für Mädchen und Jungen dar. Außerdem werden die Differenzen dieser Werte sowie alle zugehörigen Standardfehler gezeigt. Den rechten Teil der Abbildung bildet eine graphische Visualisierung der Differenzwerte. Es wird dargestellt, für welche Geschlechtsgruppe der Mittelwert höher ausfällt. Statistisch signifikante Unterschiede sind rot gefärbt gekennzeichnet.

Betrachtet man die Abbildung, ist das deutlichste Ergebnis, dass Mädchen im Vergleich zu Jungen in allen drei Bereichen – Leseselbstkonzept, Lesemotivation und Leseverhalten – höhere Mittelwerte erreichen. Für das Leseselbstkonzept, das mit fünf Items erfragt wurde (z.B. *Lesen fällt mir sehr leicht.*), fällt der Unterschied zwischen Mädchen ($M = 3.33$) und Jungen ($M = 3.26$) nominell am geringsten aus (Effektstärke $d = 0.12$). Obwohl der Unterschied als klein zu bezeichnen ist, ist er statistisch aufgrund der Größe der Stichproben signifikant. Nominell größere Unterschiede bestehen zwischen Mädchen und Jungen in der Lesemotivation, die mit drei Items gemessen wurde (z.B. *Ich lese gerne.*). Mädchen ($M = 3.36$) berichten eine höhere Lesemotivation als Jungen ($M = 2.99$;

$d = 0.42$). Ein analoges Muster zeigt sich auch in Bezug auf das Leseverhalten. Mädchen geben verglichen mit Jungen statistisch signifikant öfter an, dass sie außerhalb der Schule häufig lesen, weil es ihnen Spaß macht (Mädchen: $M = 3.17$; Jungen: $M = 2.78$; $d = 0.54$). In Bezug auf die betrachteten motivationalen und verhaltensbezogenen Maße ist festzuhalten, dass Mädchen und Jungen in allen drei Bereichen, Leseselbstkonzept, Lesemotivation und Leseverhalten am Ende der vierten Klassenstufe in Deutschland im Mittel über ein hohes Niveau verfügen, wobei Mädchen durchschnittlich über höhere Lesemotivation verfügen und öfter in der Freizeit zum Spaß lesen.

Abbildung 5.5: Mittelwerte und Differenzen von Leseselbstkonzept, Lesemotivation und Leseverhalten nach Geschlecht in Deutschland



■ Statistisch signifikante Unterschiede ($p < .05$).

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

B = Die Skala umfasst fünf Fragen (z.B. Lesen fällt mir sehr leicht.) mit vierstufigem Antwortformat (1 = Stimme überhaupt nicht zu, ..., 4 = Stimme stark zu).

C = Die Skala umfasst drei Fragen (z.B. Ich lese gerne.) mit vierstufigem Antwortformat (1 = Stimme überhaupt nicht zu, ..., 4 = Stimme stark zu).

D = Die Frage lautet „Ich lese außerhalb der Schule, weil es mir Spaß macht.“ mit vierstufigem Antwortformat (1 = Nie oder fast nie, ..., 4 = Jeden Tag oder fast jeden Tag).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

3.6 Bedeutung des Geschlechts für die Lesekompetenz unter Berücksichtigung von Leseselbstkonzept, Lesemotivation und Leseverhalten

Tabelle 5.1 gibt eine Übersicht über die Erklärungskraft motivationaler und verhaltensbezogener Variablen für den Geschlechtereffekt im Lesen. Alle dargestellten Modelle beziehen sich als erklärende Variable auf die Gesamtskala Lesen, die über die vier bisherigen Messzeitpunkte von IGLU gemeinsam skaliert wurde (siehe Kapitel 2 in diesem Band). Dargestellt sind unstandardisierte Regressionsgewichte (siehe zum Trendmodell allgemein Kapitel 2, Abschnitt 11 in diesem Band). Alle Modelle enthalten den zentralen Prädiktor Geschlecht (farbig unterlegt). Darauf aufbauend werden verschiedene motivationale und verhaltensbezogene Variablen aufgenommen (farbig unterlegt; Modell 2 bis Modell 4). Modell 3 und Modell 4 enthalten zusätzlich Interaktionsterme sowie das Bildungsniveau der Familie und die kognitiven Fähigkeiten als Kontrollvariablen. Die letzte Zeile der Tabelle gibt den durch die aufgenommenen Prädiktoren erklärten Anteil an der Varianz in den Leseleistungen an.

Tabelle 5.1: Lesekompetenz zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle des Geschlechts und weiterer Merkmale (Mehrebenenmodelle)

	Lineare Regression		Mehrebenenregressionen							
	b	(SE)	Modell 1		Modell 2		Modell 3		Modell 4	
			b	(SE)	b	(SE)	b	(SE)	b	(SE)
Konstante	532.5	(3.3)**	537.9	(4.7)**	524.5	(4.8)**	524.8	(4.7)**	525.0	(5.4)**
Testzeitpunkt										
IGLU 2001	1.7	(3.4) ^{ns}	5.6	(5.0) ^{ns}	1.9	(4.6) ^{ns}	4.5	(3.6) ^{ns}	5.2	(5.2) ^{ns}
IGLU 2006	10.3	(3.6)**	9.4	(5.5)*	7.6	(4.7) ^{ns}	8.6	(3.5)**	7.9	(5.0) ^{ns}
IGLU 2011	3.8	(4.0) ^{ns}	5.3	(5.3) ^{ns}	-0.5	(5.0) ^{ns}	-3.5	(3.9) ^{ns}	-4.2	(5.3) ^{ns}
Merkmale										
Geschlecht ^A	-9.7	(1.4)**	-8.5	(2.4)**	-0.4	(2.3) ^{ns}	2.6	(2.3) ^{ns}	2.2	(4.9) ^{ns}
Familien mit hohem Bildungsniveau ^B							21.2	(2.5)**	21.2	(2.5)**
Kognitive Fähigkeiten ^C							16.0	(0.9)**	16.0	(0.9)**
Interaktionen										
Geschlecht * IGLU 2001									-1.4	(6.4) ^{ns}
Geschlecht * IGLU 2006									1.4	(6.2) ^{ns}
Geschlecht * IGLU 2011									1.4	(6.6) ^{ns}
Kovariaten										
Leseselbstkonzept ^D					27.8	(1.5)**	21.4	(1.6)**	21.4	(1.6)**
Lesemotivation ^E					5.5	(1.7)**	4.7	(2.1)*	4.6	(2.1)*
Leseverhalten ^F					25.5	(3.7)**	21.6	(4.0)**	21.6	(4.0)**
Erklärter Anteil auf Individualebene (70.3%)			0.8		29.6		46.5		46.6	

b = Regressionsgewichte (unstandardisiert).

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

A = Geschlecht (0 = Mädchen; 1 = Jungen).

B = Bildungsniveau nach höchstem Bildungsabschluss der Eltern (0 = Kein Elternteil mit (Fach-)Hochschulabschluss; 1 = Mindestens ein Elternteil mit (Fach-)Hochschulabschluss), siehe Kapitel 6 in diesem Band.

C = Skalierter Index der kognitiven Fähigkeiten (Messinvarianz über die Zyklen angenommen), siehe Kapitel 2 in diesem Band.

D = Skalierter Index des Leseselbstkonzepts (Messinvarianz über die Zyklen angenommen), siehe Kapitel 4 in diesem Band.

E = Skalierter Index der Lesemotivation (Messinvarianz über die Zyklen angenommen), siehe Kapitel 4 in diesem Band.

F = Lesen zum Vergnügen (0 = Nie oder fast nie; 1 = Jeden Tag oder fast jeden Tag), siehe Kapitel 4 in diesem Band.

Das Gesamtmodell zu dem hier dargestellten Ausschnitt ist im Anhang B dargestellt, siehe Modelle 2, 3, 4 und 5.

Die lineare Regression ist nicht Teil des Gesamtmodells.

In der linearen Regression zeigt sich ein statistisch signifikanter Effekt des Geschlechts von -9.7 Punkten zum Nachteil der Jungen. Die Abweichung von den berichteten -11 Punkten Unterschied zum Nachteil der Jungen (siehe Abschnitt 3.1, da berichtet als Vorteil der Mädchen mit 11 Punkten) ergibt sich aus dem Einbezug der weiteren Testzeitpunkte. Berücksichtigt man die Mehrebenenstruktur der Daten im Sinne der Schachtelung in Klassen beziehungsweise Schulen, verringert sich dieser Effekt nominell minimal auf -8.5 Punkte (Modell 1). Der Geschlechtereffekt ist auch unter Berücksichtigung der Mehrebenenstruktur statistisch signifikant. Allerdings klärt er nur einen sehr geringen Anteil der Varianz in den Leseleistungen auf (Varianzanteil innerhalb von Schulen: 70.3%, davon erklärter Varianzanteil von Modell 1: 0.8%). Werden Leseselbstkonzept, Lesemotivation und Leseverhalten aufgenommen (Modell 2), ist der Nachteil der Jungen nicht mehr statistisch bedeutsam. Die Vorteile der Mädchen in der Gesamtskala Lesen können demnach vollständig durch diese Merkmale aufgeklärt werden. Die motivationalen und verhaltensbezogenen

Aspekte erklären Unterschiede in der Schülerschaft in der Gesamtskala Lesen jeweils mit statistisch signifikanten Effekten. Zusätzlich steigt bei Berücksichtigung dieser Aspekte der Anteil der erklärten Varianz auf 29.6 Prozent. Das Ergebnismuster ist auch gegenüber der Hinzunahme wichtiger Kontrollvariablen wie dem Bildungsniveau der Familie und den kognitiven Fähigkeiten (Modell 3) robust. Gleichmaßen gilt das Muster unter Kontrolle der Interaktion zwischen Geschlecht und Testzeitpunkt (Modell 4). Inhaltlich bedeutet dieses Muster, dass unter Annahme eines vergleichbaren Effekts der motivationalen und verhaltensbezogenen Variablen für alle Testzeitpunkte kein nach Testzeitpunkt differenzieller Geschlechterunterschied im Lesen besteht.

4 Zusammenfassung

Zusammenfassend lässt sich festhalten, dass in Deutschland am Ende der vierten Klassenstufe nach wie vor systematische Unterschiede in den Leseleistungen von Mädchen und Jungen bestehen. Die Vorteile der Mädchen sind jedoch mit 11 Punkten eher gering und unterscheiden sich nicht von dem internationalen Mittelwert der Geschlechterdifferenzen. Entsprechend ist auch die Verteilung auf Kompetenzstufen nach Geschlecht relativ gleich, wobei nominell einerseits ein leicht höherer Anteil der Mädchen auf der höchsten Kompetenzstufe und andererseits ein leicht höherer Anteil der Jungen in der Gruppe der schwachen Lesenden zu verzeichnen sind. Ein vertiefender Blick zeigt, dass es insbesondere beim literarischen Lesen im Mittel deutliche Vorteile der Mädchen in den Leseleistungen gibt, während dies beim informierenden Lesen nur minimal der Fall ist. Die Betrachtung über die verschiedenen Erhebungszeitpunkte seit 2001 verweist auf die Konstanz der Geschlechterunterschiede in der Lesekompetenz bei Viertklässlerinnen und Viertklässlern über 15 Jahre.

Die Ergebnisse der Analysen zu Unterschieden in motivationalen und verhaltensbezogenen Merkmalen weisen darauf hin, dass Mädchen in allen untersuchten Bereichen – dem Leseselbstkonzept, der Lesemotivation und dem Leseverhalten – am Ende der vierten Klassenstufe in Deutschland im Mittel über ein höheres Niveau als die Jungen verfügen. Kontrolliert man diese Unterschiede zwischen Mädchen und Jungen, findet sich im Trendmodell kein bedeutsamer Kompetenzunterschied im Lesen zwischen den Geschlechtern mehr. Jungen und Mädchen mit gleichem Niveau im Leseselbstkonzept, in der Lesemotivation und im Leseverhalten verfügen demnach über gleich gute Lesekompetenzen.

Insgesamt verweisen die Ergebnisse zu den Geschlechterunterschieden in den Leseleistungen, dem Leseselbstkonzept, der Lesemotivation und dem Leseverhalten auf mögliche Handlungsansätze in der schulischen und außerschulischen Bildungspraxis. Das differentielle Unterschiedsmuster in Abhängigkeit von den Textsorten verdeutlicht, dass eine stärkere Förderung der Jungen im Bereich des literarischen Lesens den mittleren Unterschied in den Leseleistungen im Vergleich zu den Mädchen reduzieren und damit auch den Gesamtleistungswert der Schülerschaft am Ende der vierten Klassenstufe verbessern könnte. Weitere Ansätze für Wissenschaft und Praxis sind im Bereich der Lesemotivation und des Leseverhaltens zu identifizieren. Die systematische Entwicklung, Überprüfung und Implementation von Fördermaßnahmen, die gezielt gerade auch bei Jungen Lesemotivation und Lesehäufigkeit in der Freizeit verstärken, sind hier vielversprechende Möglichkeiten.

Literatur

- Archambault, I., Eccles, J. S. & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology, 102* (4), 804–816.
- Artelt, C., Drechsel, B., Bos, W. & Stubbe, T. C. (2008). Lesekompetenz in PISA und PIRLS/IGLU – Ein Vergleich. *Zeitschrift für Erziehungswissenschaft, 10, Sonderheft 10/2008*, 35–52.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich.
- Autorengruppe Bildungsberichterstattung (2010). *Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demographischen Wandel*. Bielefeld: W. Bertelsmann Verlag.
- Baker, L. & Wigfield, A. (1999). Dimensions of children's motivation for reading and their relations to reading activity and reading achievement. *Reading Research Quarterly, 34*, 452–477.
- Becker, M. & McElvany, N. (2017). The interplay of gender and social background: a longitudinal study of interaction effects in reading attitudes and behavior. *British Journal of Educational Psychology*. doi:10.1111/bjep.12199
- Becker, M., McElvany, N. & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology, 102* (4), 773–785.
- 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.
- 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.
- Christmann, U. & Groeben, N. (1999). Psychologie des Lesens. In B. Franzmann, K. Hasemann, D. Löffler & E. Schön (Hrsg.), *Handbuch Lesen* (S. 145–223). München: Saur.
- Durik, A. M., Vida, M. & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98* (2), 382–393.
- Eagly, A. H. & Wood, W. (1999). The origins of sex differences in human behavior: evolved dispositions versus social roles. *American Psychologist, 54* (6), 408–423.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives: Psychological and sociological approaches* (S. 75–146). San Francisco: Freeman.
- Ehm, J. H., Duzy, D. & Hasselhorn, M. (2011). Das akademische Selbstkonzept bei Schulanfängern: Spielen Geschlecht und Migrationshintergrund eine Rolle? *Frühe Bildung, 0*, 37–45.
- Frith, U. (1986). Psychologische Aspekte des orthographischen Wissens. In G. Augst (Hrsg.), *New Trends in Graphemics and Orthography* (S. 218–233). Berlin: de Gruyter.
- Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101* (3), 371–395.
- 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.
- Hannover, B., Wolter, I. & Zander, L. (2017). Geschlechtergerechtigkeit im Klassenzimmer. In T. Eckert & B. Gniewosz (Hrsg.), *Bildungsgerechtigkeit* (S. 201–213). Wiesbaden: Springer VS.

- Hausmann, M. (2007). Kognitive Geschlechtsunterschiede. In S. Lautenbacher, M. Hausmann & O. Güntürkün (Hrsg.), *Gehirn und Geschlecht – Neurowissenschaft des kleinen Unterschieds zwischen Frau und Mann* (S. 105–124). Heidelberg: Springer.
- Helbig, M. (2012). *Sind Mädchen besser? Der Wandel geschlechtsspezifischen Bildungserfolgs in Deutschland*. Frankfurt: Campus.
- Helmke, A. & Weinert, F. E. (1997). Unterrichtsqualität und Leistungsentwicklung. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241–252). Weinheim: Psychologie Verlags Union.
- Heyder, A. & Kessels, U. (2017). Boys don't work? On the psychological benefits of showing low effort in high school. *Sex Roles*, 77, 72–85.
- Heyder, A., Kessels, U. & Steinmayr, R. (2017). Explaining boys' underachievement in language: Not a lack of aptitude but students' motivational beliefs and parents' perceptions? *British Journal of Educational Psychology*, 87, 205–223.
- 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.
- Hornberg, S., Valtin, R., Potthoff, B., Schwippert, K. & Schulz-Zander, R. (2007). Lesekompetenzen von Mädchen und Jungen im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 195–223). Münster: Waxmann.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60 (6), 581–592. doi: 10.1037/0003-066X.60.6.581
- Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology*, 65, 373–398. doi: 10.1146/annurev-psych-010213-115057
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S. & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, 73, 509–527.
- Kessels, U., Heyder, A. & Holder, K. (2016). Unterricht und Geschlecht aus internationaler Perspektive. In N. McElvany, W. Bos, H.-G. Holtappels, M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts – Aktueller Stand und Perspektiven der Unterrichtsforschung* (S. 89–106). Münster: Waxmann.
- Kessels, U. & Heyder, A. (2018). Geschlechtsunterschiede. In D. Rost, J. Sparfeldt & S. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (5. Auflage) (S. 209–217). Weinheim: Beltz.
- Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. New York: Cambridge University Press.
- Kuhl, P. & Hannover, B. (2012). Differenzielle Benotungen von Mädchen und Jungen. Der Einfluss der von der Lehrkraft eingeschätzten Kompetenz zum selbstgesteuerten Lernen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44 (3), 153–162.
- Logan, S. & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, 32, 199–214. doi: 10.1111/j.1467-9817.2008.01389.x
- Marsh, H. W., Craven, R. & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort–multioccasion study. *Child development*, 69 (4), 1030–1053.
- McElvany, N., Kortenbruck, M. & Becker, M. (2008). Lesekompetenz und Lesemotivation: Entwicklung und Mediation des Zusammenhangs durch Leseverhalten. *Zeitschrift für Pädagogische Psychologie*, 22, 207–219.
- Mielke, R., Goy, M. & Pietsch, M. (2006). Das Leseselbstkonzept am Ende der Grundschulzeit. In W. Bos & M. Pietsch (Hrsg.), *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen* (S. 87–109). Münster: Waxmann.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55, 19–27.
- Mullis, I. V. S., Martin, M. O., Foy, P. & Drucker, K. T. (2012). *PIRLS 2011. International results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

- 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. 23–72). Münster: Waxmann.
- Neugebauer, M. (2011). Werden Jungen von Lehrerinnen bei den Übergangsempfehlungen für das Gymnasium benachteiligt? Eine Analyse auf Basis der IGLU-Daten. In A. Hadjar (Hrsg.), *Geschlechtsspezifische Bildungsungleichheiten* (S. 235–260). Wiesbaden: VS Verlag für Sozialwissenschaften.
- OECD. (2014). *Education at a glance 2014. OECD indicators*. Paris: OECD Publishing. doi: 10.1787/eag-2014-en
- Retelsdorf, J., Köller, O. & Möller, J. (2011). On the effects of motivation on reading performance growth in secondary school. *Learning and Instruction*, 21 (4), 550–559.
- Schwabe, F., McElvany, N. & Trendtel, M. (2015). The school age gender gap in reading achievement: Examining the influences of item format and intrinsic reading motivation. *Reading Research Quarterly*, 50, 219–232.
- Stanat, P. & Kunter, M. (2001). Geschlechterunterschiede in Basiskompetenzen. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 251–270). Opladen: Leske + Budrich.
- Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S. & Haag, N. (2017). *IQB-Bildungstrend 2016: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe*. Münster: Waxmann.
- Steinmayr, R. & Spinath, B. (2008). Sex Differences in School Achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, 22, 185–209.
- Steinmayr, R. & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW). *Diagnostica*, 56, 195–211.
- Upadaya, K. & Eccles, J. S. (2015). Do teachers' perceptions of children's math and reading related ability and effort predict children's self-concept of ability in math and reading? *Educational Psychology*, 35, 110–127.
- Valtin, R., Wagner, C. & Schwippert, K. (2005). Schülerinnen und Schüler am Ende der vierten Klasse – schulische Leistungen, lernbezogene Einstellungen und außerschulische Lernbedingungen. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (S. 187–230). Münster: Waxmann.
- Weis, M., Zehner, F., Sälzer, C., Strohmaier, A., Artelt, C. & Pfost, M. (2016). Lesekompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiefe-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. 257–297). Münster: Waxmann.
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C. & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89, 451–469.

Kapitel VI

Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern

Anke Hußmann, Tobias C. Stubbe und Daniel Kasper

1 Einleitung

Für die Bundesrepublik Deutschland ebenso wie für eine Vielzahl von Staaten ist seit Jahren durch Ergebnisse der empirischen Bildungsforschung belegt, dass eine enge Kopplung zwischen dem sozioökonomischen Status von Familien und dem Bildungserfolg ihrer Kinder besteht. Abhängig von den sozialen Verhältnissen haben Schülerinnen und Schüler unterschiedliche Chancen, am Bildungssystem teilzuhaben. Der Anteil der armutsgefährdeten Bevölkerung in Deutschland, der seit 2005 in Westdeutschland (insbesondere in Nordrhein-Westfalen, Berlin und Bremen) angestiegen ist, verweist auf eine größer werdende soziale Ungleichheit (Statistisches Bundesamt, 2016a). Welche Konsequenzen sich daraus im Hinblick auf Teilhabechancen ergeben, ist breit dokumentiert (zuletzt u. a. Bertelsmann Stiftung, 2016; Laubstein, Holz & Seddig, 2016; Tophoven, Lietzmann, Reiter & Wenzig, 2017).

Die Frage nach Chancengerechtigkeit im Bildungssystem steht schon lange im Fokus der Aufmerksamkeit, nicht erst – wenngleich seither mit breiterem öffentlichen Interesse – mit der Fülle an empirisch fundiertem Wissen fortlaufender und wiederkehrender Schulleistungsstudien wie der *Internationalen Grundschul-Lese-Untersuchung* (IGLU), der *Trends in International Mathematics and Science Study* (TIMSS) und dem *Programme for International Student Assessment* (PISA). Auch zuvor konnten durch international vergleichende Studien Bildungsungleichheiten in verschiedenen Staaten empirisch aufgezeigt werden (Blossfeld & Shavit, 1993). Ein bildungspolitisches Ziel, das durch das regelmäßig stattfindende Bildungsmonitoring und durch sich anschließende konkrete Maßnahmen seither erreicht werden soll, ist, den Bildungserfolg von Schülerinnen und Schülern von ihrer sozialen Herkunft zu entkoppeln und für mehr Chancengerechtigkeit zu sorgen.

Ungleiche Startbedingungen lassen sich bereits mit der Einschulung beobachten (primäre Herkunftseffekte) ebenso wie ihre Fortdauer durch herkunftsbedingte Unterschiede in den familialen und institutionellen Bildungsverlaufentscheidungen (sekundäre Herkunftseffekte). Ein Beispiel für primäre Herkunftseffekte sind die (zumeist auch) an ökonomische Ressourcen, soziale Kontakte oder kulturelle Formen gebundenen Gewohnheiten in Familien wie Theaterbesuche oder gemeinsames Musizieren, von denen auszugehen ist, dass sie das Lernen frühzeitig prägen und zu einem vertrauten Umgang mit Kultur führen, der in der Schule von Bedeutung ist. Ein Beispiel für sekundäre Herkunftseffekte sind Schulformentscheidungen beim Übergang von der Grundschule auf eine weiterführende Schule (siehe Kapitel 8 in diesem Band). Neben dem in den letzten Jahren zunehmenden Wissen über jene Bereiche, die Chancenungleichheiten im Bildungssystem Deutschlands potentiell verstärken, lassen sich Maßnahmen und Programme benennen, die Möglichkeiten zur Förderung, Intervention und Innovation im Bildungssystem hinsichtlich der Schaffung gerechterer Teilhabe eröffnen und eröffnet haben. Im Allgemeinen sind dabei vor allem Veränderungen der letzten Jahre im Bereich frühkindlicher Bildung (z.B. Anders & Roßbach, 2013; Hasselhorn & Kuger, 2014) sowie der Ausbau von Ganztagsangeboten im Primar- und Sekundarschulbereich (z.B. Strietholt, Maniti, Berkemeyer & Bos, 2015; vbw, 2013; Wendt, Goy, Walzebug & Valtin, 2016) hervorzuheben. Konkret in Bezug auf die Entwicklung von Lesekompetenzen ist eine Vielzahl von Sprachförderprogrammen und -angeboten zu nennen, die auf die wirksame Förderung von bildungsbenachteiligten Kindern sowohl mit als auch ohne Migrationshintergrund abzielen (z.B. Stanat, Weirich & Radmann, 2012; Valtin & Tarelli, 2014; zum Thema Leseförderung unter Bedingungen zunehmender Heterogenität siehe auch Kapitel 10 und 11 in diesem Band). Entsprechende Projekte und Programme können praxis- und/oder forschungsorientiert sein, sie finden überregional oder auch regional statt.

Das vorliegende Kapitel stellt zunächst zentrale Befunde zum Zusammenhang von sozialer Herkunft und Bildungserfolg in Deutschland vor und beschreibt Indikatoren, die im Rahmen von IGLU zur Erfassung der sozialen Herkunft berücksichtigt werden (siehe Abschnitte 2 und 3). Gegliedert nach Indikatoren wird anschließend ihr Zusammenhang mit Lesekompetenzen in internationaler (siehe Abschnitt 4) sowie unter Berücksichtigung ausgewählter Indikatoren in nationaler Perspektive dargestellt (siehe Abschnitt 5). Die zentralen Befunde dieses Kapitels werden abschließend zusammenfassend dargestellt (siehe Abschnitt 6).

2 Soziale Herkunft und Bildungserfolg

Die in den vergangenen Jahrzehnten stetig zu beobachtende Bildungsexpansion in Deutschland hat gesamtgesellschaftlich betrachtet eine kontinuierliche Höherqualifizierung der Bevölkerung erbracht (Geißler, 2014). Dass ungeachtet dessen soziale Selektion stattfand und nach wie vor stattfindet, belegen empirisch fundiert vor allem die groß angelegten internationalen Schulleistungstudien wie IGLU, TIMSS und PISA, an denen die Bundesrepublik Deutschland seit Ende der 1990er Jahre regelmäßig teilnimmt (zuletzt Haag, Kocaj, Jansen & Kuhl, 2017; Müller & Ehmke, 2016; Stubbe, Schwippert & Wendt, 2016; Wendt, Stubbe & Schwippert, 2012). Die Ergebnisse zeigen für Deutschland im internationalen Vergleich große soziale Disparitäten; in der ersten PISA-Erhebung 2000 schnitt Deutschland in diesem Aspekt im Vergleich zu allen Teilnehmerstaaten

sogar am ungünstigsten ab (Baumert & Schümer, 2001). Zwischen den Teilnehmerstaaten unterscheidet sich das Ausmaß der sozialen Disparitäten zum Teil deutlich, wie zuletzt im Rahmen von TIMSS 2015 berichtet (Stubbe et al., 2016); in manchen Teilnehmerstaaten, die leistungsstärker als Deutschland sind, gelingt es aber gleichzeitig, die Unterschiede in den Kompetenzen zwischen Kindern und Jugendlichen unterschiedlicher sozialer Herkunft geringer zu halten. Befunde wie diese stützen die Aussagen zahlreicher Autorinnen und Autoren, die Bildungsexpansion in Deutschland habe insgesamt nicht zu einem Abbau an sozialen Disparitäten beigetragen (z.B. Becker & Lauterbach, 2004; Blossfeld, 1993; Schimpl-Neimanns, 2000; Solga, 2005; Vester, 2004, 2005). Das Problem der sozialen Gerechtigkeit und gesellschaftlicher Teilhabechancen ist in Deutschland damit nach wie vor drängend (z.B. vbw, 2017).

War die Kopplung zwischen sozialer Herkunft und den erreichten Kompetenzen im Jahr 2000 noch besonders eng (Baumert & Schümer, 2001), so zeigt sich für Deutschland in der Rückschau der letzten Jahre (2000 bis 2015) zumindest für den Sekundarschulbereich, dass dieser Zusammenhang etwas geringer ausfällt (Müller & Ehmke, 2016). Zugleich aber gibt es weiterhin nur wenige Staaten, in denen das Ausmaß der sozialen Disparitäten in der Sekundarstufe höher ausfällt als in Deutschland (zum Überblick Ehmke & Jude, 2010; Müller & Ehmke, 2016).

Auch für den Grundschulbereich wurde in Deutschland eine eher überdurchschnittliche Kopplung zwischen sozialer Herkunft und dem Bildungserfolg der Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe festgestellt (z.B. Bonsen, Frey & Bos, 2008; Bos, Schwippert & Stubbe, 2007; Wendt et al., 2012). Wie noch im letzten Jahr zu TIMSS 2015 berichtet, waren keine signifikanten Veränderungen zu TIMSS 2007 festzustellen (Stubbe et al., 2016). Ebenso zeigten auch die Ergebnisse zu IGLU 2011 im Vergleich zu den Erhebungszyklen 2001 und 2006 keine signifikanten Veränderungen; soziale Disparitäten blieben somit ausgeprägt bestehen (Wendt et al., 2012). Der Leistungsunterschied, zum Beispiel gemessen daran, ob Viertklässlerinnen und Viertklässler aus armutsgefährdeten Elternhäusern stammen oder nicht, beträgt seither etwa ein Lernjahr.

Der regelmäßig vom *Institut zur Qualitätsentwicklung im Bildungswesen* (IQB) durchgeführte *IQB-Bildungstrend* (bis 2012: *Ländervergleich*) ermöglicht es, den Blick auf soziale Disparitäten im Vergleich der Länder der Bundesrepublik Deutschland zu richten. Bereits im Rahmen von IGLU-E im Jahr 2006 wurde festgestellt, dass zwischen den Ländern zum Teil deutliche Unterschiede bestehen. Soziale Disparitäten in den Stadtstaaten Berlin und Hamburg waren statistisch signifikant höher als im Bundesdurchschnitt und signifikant niedriger als in Bayern (Stubbe, Bos & Hornberg, 2008). Gemessen am sozialen Gradienten, einem statistisch ermittelten Kennwert zur Beschreibung des Zusammenhangs zwischen Merkmalen der sozialen Herkunft und Schülerkompetenz, zeigten die Befunde des *Ländervergleichs 2011* für den Bereich Lesen in Deutschland 36 Punkte (Richter, Kuhl & Pant, 2012). Aktuelle Befunde des *IQB-Bildungstrends 2016* zeigen für den Bereich Lesen, dass auf Bundesebene keine bedeutsamen Veränderungen zu 2011 festzustellen sind (Haag et al., 2017). Im Jahr 2016 beträgt der soziale Gradient 39 Punkte. Dabei bestehen zwischen den Ländern der Bundesrepublik Deutschland signifikante Unterschiede etwa zwischen Schleswig-Holstein (29 Punkte) und Baden-Württemberg (45 Punkte).

3 Indikatoren der sozialen Herkunft

Es gibt unterschiedliche Ansätze und Instrumente, das Konstrukt soziale Herkunft empirisch zu erfassen. Zentral bei der Analyse von sozialen Disparitäten ist die Frage der Validität, also der Gültigkeit dessen, wie der sozioökonomische Status von Familien (oder auch andere Aspekte von sozialer Herkunft) in Deutschland und im internationalen Vergleich erfasst wird. Erhebungsinstrumente, die in der empirischen Bildungsforschung zur Erfassung der sozialen Herkunft von Schülerinnen und Schülern eingesetzt werden, lassen sich theoretisch sehr gut auf der Grundlage der soziologischen Arbeiten von Pierre Bourdieu (1983) beschreiben. Bourdieu unterscheidet ökonomisches, kulturelles und soziales Kapital und beschreibt damit verbundene Ressourcen, auf die Schülerinnen und Schüler zurückgreifen können, um in bestimmten Kontexten, wie dem schulischen, erfolgreich handeln zu können. Ob Schülerinnen und Schüler Zugang zu schulrelevantem Kapital haben und welches dabei als schulrelevant gilt, hängt von gesellschaftlichen (Bourdieu selbst spricht von: Macht-) Verhältnissen ab und von dem Repertoire, das ihnen ihre Familien und ihr soziales Umfeld zur Verfügung stellen.

Als ökonomisches Kapital beschreibt Bourdieu die monetären und materiellen Ressourcen einer Familie. Das kulturelle Kapital beinhaltet verschiedene Aspekte: (a) erworbenes Wissen und Fähigkeiten einer Person (inkorporiertes Kulturkapital), (b) kulturelle Besitztümer wie Musikinstrumente oder Bücher, die nur von Personen adäquat genutzt werden können, die über das notwendige inkorporierte Kulturkapital verfügen (objektiviertes Kulturkapital) und (c) formale schulische und akademische Titel (institutionalisiertes Kulturkapital). Mit sozialem Kapital sind nach Bourdieu die Beziehungsnetzwerke der Menschen beschrieben. Dabei hängt das Ausmaß des sozialen Kapitals, über das eine Person verfügt, zum einen von der Größe eines Netzes ab und zum anderen von dem Kapital, das die übrigen Personen in einem Netzwerk besitzen. Coleman (1988), auf den zur Beschreibung des sozialen Kapitals ergänzend zu Bourdieu aussichtsreich Bezug genommen werden kann, grenzt von dem sozialen Kapital außerhalb der Familie das soziale Kapital innerhalb der Familie ab und betont dessen Wichtigkeit für den Erwerb von schulischen Kompetenzen. Generell gilt: Jede Kapitalform kann in eine andere transformiert werden. Unter Umständen kann dies mit Transformationskosten (z.B. zu investierende Zeit) verbunden sein. Dies ist beispielweise dann der Fall, wenn Familien ihre Kinder im Verlauf ihrer Bildungsbiografien zur Erreichung von (Hoch-)Schulabschlüssen (institutionalisiertes Kulturkapital) unterstützen, indem ökonomisches (z.B. Investition in Nachhilfeunterricht), kulturelles (z.B. Unterstützung durch die Eltern bei der Arbeit für die Schule) und soziales (z.B. Unterstützung durch befreundete Lehrer) Kapital genutzt werden.

Nicht von Bourdieu explizit aufgegriffen, aber eng an die Kapitalformen und Transformationspotentiale von Kapitalformen gebunden, ist der Aspekt der Armut beziehungsweise *Armutgefährdung*. Die Armutgefährdung wird über das Einkommen von Haushalten zusammen mit Informationen zur Haushaltsgröße berechnet. Als armutsgefährdet gelten nach gängiger Definition Haushalte, die zum jeweiligen Zeitpunkt über weniger als 60 Prozent des medianen Nettoäquivalenzeinkommens in einem Staat verfügen. Zunächst wird zu dieser Berechnung die Armutgefährdungsgrenze eines Haushalts ermittelt, die als Produkt aus dem medianen Nettoäquivalenzeinkommen und der gewichteten Haushaltsgröße definiert ist (Haushaltsvorstand: 1.0; jedes weitere Haushaltsmitglied,

das mindestens 14 Jahre alt ist: 0.5; jedes Kind, das maximal 13 Jahre alt ist: 0.3). Liegt das jeweilige Haushaltsnettoeinkommen einer Familie unter der Armutsgefährdungsgrenze, ist dieser Haushalt als armutsgefährdet zu bezeichnen (Bardone & Guio, 2005; nach gleichem Verfahren wie auch in Stubbe et al., 2016 für TIMSS 2015 beschrieben). Nach Angaben des Statistischen Bundesamts, dessen Definition von Armutsgefährdung nicht identisch, aber ähnlich zu der hier beschriebenen ist (vgl. Statistisches Bundesamt, 2016b), liegt der Anteil armutsgefährdeter Personen in Deutschland für das Jahr 2015 bei 16.7 Prozent (ebd.). Sollen Teilhabechancen von Kindern in Deutschland sichergestellt werden, so hängt dies auch mit der finanziellen Situation zusammen, in der Kinder aufwachsen. Dem *fünften Armuts- und Reichtumsbericht* des Bundesministeriums für Arbeit und Soziales (BMAS) zufolge liegt die Armutsgefährdungsquote von Kindern in Deutschland derzeit bei 64 Prozent, wenn beide Eltern nicht erwerbstätig sind; arbeiten beide Elternteile, eines davon in Vollzeit, verringert sich die Armutsgefährdung auf fünf Prozent (BMAS, 2017, S. 25). Empirisch besteht also ein enger Zusammenhang zwischen Armut und anderen Indikatoren der sozialen Herkunft. Zugleich bietet der Aspekt der Armutsgefährdung zusätzliche Informationen zur sozialen Lage von Familien. So ist zum Beispiel bekannt, dass Familien in Deutschland zwar von Armut bedroht sein können, zugleich aber über ein hohes Kultur- und Sozialkapital verfügen können (z. B. Bosch & Kalina, 2007; Butterwegge, 2000).

Zur Erfassung der sozialen Herkunft von Schülerinnen und Schülern haben sich in der empirischen Bildungsforschung in den letzten Jahren verschiedene Indikatoren bewährt. Viele der gängigen Indikatoren liefern relevante Hinweise zur inhaltlichen Beschreibung der Kapitalien. Nicht immer lassen diese sich eindeutig einer Kapitalform zuordnen, wie etwa dem ökonomischen oder dem kulturellen Kapital. Ein Beispiel dafür sind die höchsten Bildungsabschlüsse der Eltern. Sie gelten als ein Indikator für das institutionalisierte und inkorporierte Kulturkapital, sagen aber zugleich etwas über die mögliche finanzielle Lage einer Familie (ökonomisches Kapital) aus, weil höhere Bildungsabschlüsse tendenziell mit einem höheren Einkommen verbunden sind. Auch ist festzustellen, dass das soziale Kapital – wenn überhaupt – zumeist nur rudimentär erfasst wird; so etwa hinsichtlich der Eltern-Kind-Interaktionen im Elternhaus. Dass dies aus theoretischer Perspektive nicht ausreichend ist, wurde bereits vielfach kritisiert (z. B. Bos, Stubbe & Buddeberg, 2010; Walzebug, 2015). In IGLU 2016 (und ebenso TIMSS 2015) wurden daher weitere Instrumente zur Erfassung des sozialen Kapitals im Rahmen der nationalen Ergänzung der Elternfragebogenerhebung eingesetzt, Ergebnisse hierzu werden an anderer Stelle veröffentlicht (Lorenz & Stubbe, 2017).

Viele Indikatoren zur Erfassung der sozialen Herkunft haben sich in den vergangenen Jahren bewährt, darunter die nachfolgend vorgestellten: Geradezu als „Klassiker“ im Rahmen der empirischen Bildungsforschung zu bezeichnen ist die Anzahl der im Haushalt vorhandenen *Bücher*. Die Variable erfasst primär das objektivierte Kulturkapital und liefert damit auch einen Hinweis auf das in einer Familie zur Verfügung stehende ökonomische Kapital. Indirekt misst die Variable zudem das inkorporierte Kulturkapital, da ein Zusammenhang zwischen dem Bildungsniveau einer Familie und der Anzahl der Bücher, die diese besitzt (Bildungsnähe des Elternhauses), besteht. Die Frage nach den Büchern im Haushalt weist für Schüler- und Elternbefragungen gute Messeigenschaften auf und ist einfach und ökonomisch in der Durchführung. Auch Angaben zu weiteren *Besitztiteln im Haushalt* liefern Informationen zum ökonomischen beziehungs-

weise kulturellen Kapital von Familien (z.B. Anzahl an Autos beziehungsweise Vorhandensein von Musikinstrumenten oder Tageszeitungen).

Direkter Indikator zur Erfassung des kulturellen Kapitals in einer Familie ist die Frage nach dem *Bildungsniveau der Eltern*, das heißt dem höchsten Bildungsabschluss der Mutter und des Vaters. In IGLU wird dieser Indikator mit Hilfe der internationalen Bildungsskala (*International Standard Classification of Education – ISCED*) erhoben, die von der *United Nations Educational, Scientific and Cultural Organization* (UNESCO) entwickelt wurde (Schroedter, Lechert & Lüttinger, 2006; UNESCO, 2003).

Neben dem Bildungsniveau der Eltern ist ihre berufliche Tätigkeit von Bedeutung. Der *Berufsstatus von Personen* liefert Informationen über das kulturelle und ökonomische Kapital, da einerseits ein bestimmtes Bildungsniveau erreicht sein muss, um einen bestimmten Beruf ausüben zu können, andererseits durch bestimmte Berufe Einkommensverhältnisse klar definiert sind. Der Beruf wird offen erfragt und mit Hilfe der Nominalskala *International Standard Classification of Occupations* (ISCO) kodiert. Dabei wird jedem Beruf ein vierstelliger Code zugewiesen, der allerdings noch keine Aussagen über den Status eines Berufes erlaubt (Hoffmann, 2003). Diese Codes lassen sich dann in die verschiedenen Indizes des Berufsstatus überführen. Alle zwanzig Jahre werden die ISCO-Codes aktualisiert (zuletzt 2008). In den IGLU-Erhebungszyklen 2001 und 2006 wurde der ISCO-88 genutzt, seit IGLU 2011 der ISCO-08.

Des Weiteren werden zwei Indizes des Berufsstatus der Eltern gebildet. Sie basieren auf unterschiedlichen theoretischen Annahmen und weisen unterschiedliche statistische Eigenschaften auf:

- Der *International Socio-Economic Index of Occupational Status* (ISEI) gibt an, wie gut ein Beruf geeignet ist, um die Ausbildung einer Person in Einkommen umzuwandeln. Die Skala ist eindimensional und nimmt ganzzahlige Werte zwischen 10 und 90 an (Ganzeboom, de Graaf & Treiman, 1992; Ganzeboom & Treiman, 1996). In der empirischen Bildungsforschung wird häufig der höchste ISEI im Haushalt (HISEI) genutzt.
- Erikson, Goldthorpe und Portocarero (1979) teilen Berufe in distinkte Gruppen ein, die sich in Hinblick auf unterschiedliche Dimensionen durch hohe interne Homogenität und hohe externe Heterogenität auszeichnen. Bei diesen sogenannten EGP-Klassen handelt es sich also um eine Nominalskala.

Schließlich liefert das *Einkommen* von Haushalten direkte Angaben über das ökonomische Kapital. Eine weitere Variable, die aus dem Haushaltseinkommen und der Haushaltsgröße abgeleitet werden kann, ist die Einkommensarmut von Familien (s.o.).

Die in diesem Abschnitt vorgestellten Indikatoren wurden im Rahmen von IGLU 2016 in Deutschland eingesetzt. International kamen deutlich weniger Variablen für den sozioökonomischen Status von Familien zum Einsatz. Die nachfolgenden Befunde gliedern sich daher in internationale und nationale Analysen: In Abschnitt 4 werden zunächst Befunde im internationalen Vergleich berichtet, anschließend im Abschnitt 5 vertiefend soziale Disparitäten im deutschen Bildungssystem.

4 Zum Zusammenhang von sozialer Herkunft und Lesekompetenzen im internationalen Vergleich

Für internationale Vergleiche lassen sich im Trend seit 2001 drei Indikatoren der sozialen Herkunft nutzen (Bos et al., 2007):

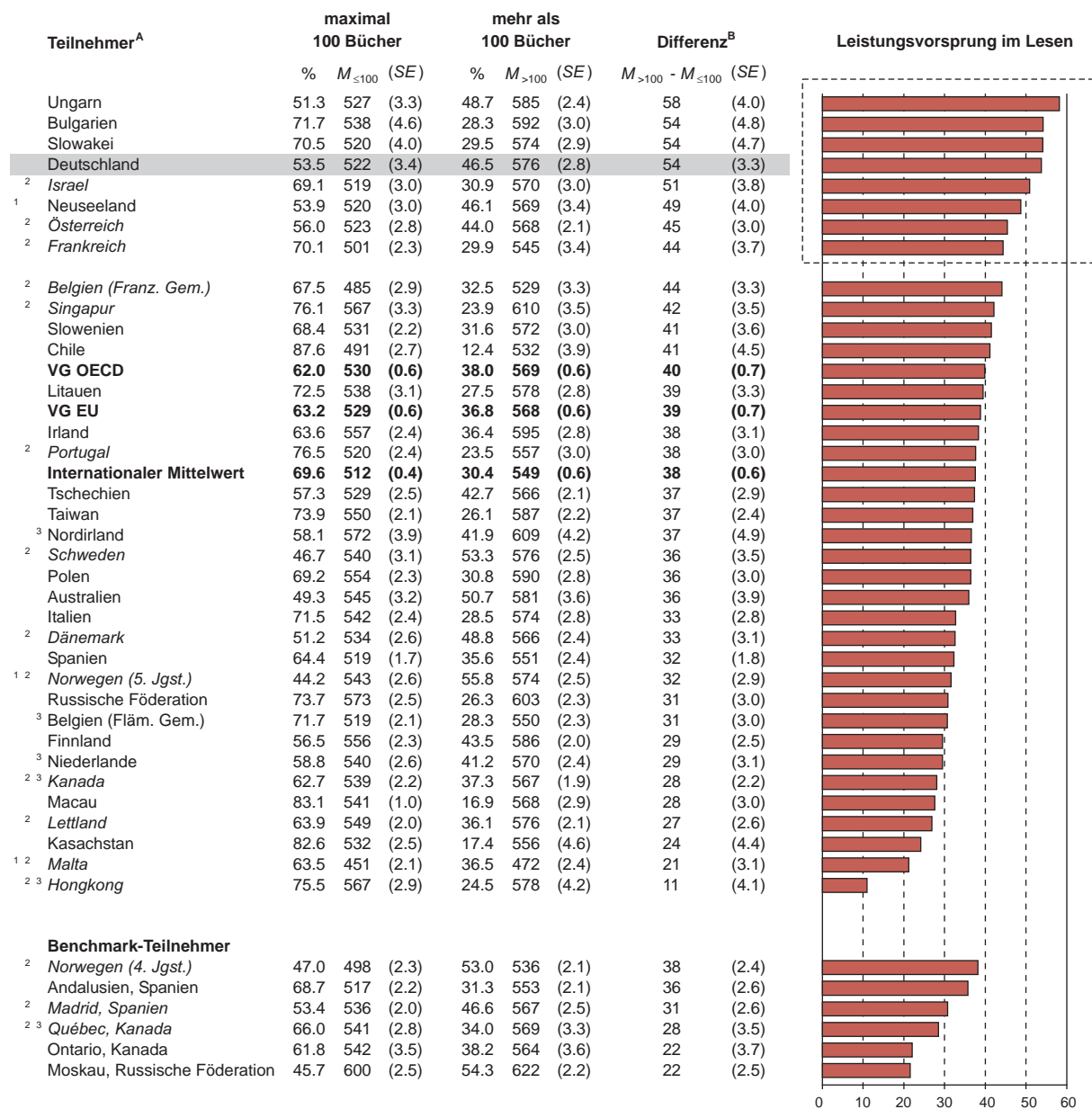
- Die Anzahl der *Bücher im Haushalt* wird mit einer fünfstufigen Ordinalskala erfasst, die in diesem Kapitel dichotomisiert mit den Gruppen „maximal 100 Bücher“ und „mehr als 100 Bücher“ genutzt wird.
- Der *Berufsstatus* der Eltern ist ein weiterer Indikator. Er basiert auf einer zwölfstufigen Nominalskala, die in eine dreistufige Ordinalskala rekodiert werden kann. Hieraus zusammengefasste Berufsgruppen bieten die Möglichkeit, eine einfache Hierarchisierung zum Beruf der Eltern für internationale Vergleiche vorzunehmen. Unterschieden werden drei Berufsgruppen, die in vergangenen Berichterstattungen zu IGLU noch als „Sozialschicht“ bezeichnet wurden (vgl. Bos et al., 2007; S. 231; Wendt et al., 2012, S. 180). Es besteht uneingeschränkte Vergleichbarkeit zu diesen Befunden mit den Befunden aus IGLU 2016. Die erste Berufsgruppe (manuelle Tätigkeiten) fasst Fachkräfte in der Landwirtschaft und Fischerei, Handwerker, Anlagen- und Maschinenbediener sowie angelernte Arbeiter und Hilfsarbeiter zusammen. Zur zweiten Berufsgruppe (Angestellte und kleinere Unternehmer) zählen Eigentümer eines kleinen Betriebes und Angestellte. Leitungen eines großen Unternehmens, leitende Bedienstete, Wissenschaftler sowie in einem verwandten Beruf Tätige, Techniker und Personen, die in einem gleichrangigen nichttechnischen Beruf tätig sind, bilden die dritte Berufsgruppe (Akademiker, Techniker und Führungskräfte). In diesem Beitrag werden die erste und die dritte Berufsgruppe gegenübergestellt.
- Das *Bildungsniveau der Eltern* wird wie in Abschnitt 3 beschrieben mit Hilfe der ISCED-Level erfasst. Die Ergebnisse dieser Variable werden dichotomisiert berichtet: „kein Elternteil mit (Fach-)Hochschulabschluss“ gegenüber „mindestens ein Elternteil mit (Fach-)Hochschulabschluss“.

Zum Indikator *Bücher im Haushalt* zeigt Abbildung 6.1 für die Staaten und Regionen, die an IGLU 2016 teilgenommen haben (siehe Kapitel 2, Abschnitt 4.2.1 in diesem Band), in wie viel Prozent der Familien jeweils maximal 100 beziehungsweise mehr als 100 Bücher vorhanden sind, wie hoch die Lesekompetenz in beiden Gruppen und wie groß die Leistungsdifferenz ist. Zudem ist durch das gestrichelte Kästchen hervorgehoben, welche Teilnehmer sich in der Leistungsdifferenz nicht signifikant von Deutschland unterscheiden.

In Deutschland sind die beiden hier verglichenen Gruppen ungefähr gleich groß: in 46,5 Prozent der Familien gibt es mehr als 100 Bücher. Nur in wenigen Bildungssystemen liegt dieser Wert bei über 50 Prozent (Norwegen, Schweden, Australien und Moskau). Hingegen finden sich viele Teilnehmer, bei denen deutlich mehr als die Hälfte der Familien maximal 100 Bücher besitzt. Mit über 75 Prozent ist dieser Anteil in Chile, Macau, Kasachstan, Portugal, Singapur und Hongkong am höchsten.

Der Leistungsvorsprung von Kindern aus Familien mit mehr als 100 Büchern beträgt in Deutschland 54 Punkte – und somit etwas mehr als ein Lernjahr (vgl. Bos et al., 2007). Bei keinem Teilnehmer ist dieser Wert signifikant größer. In Ungarn, Bulgarien, der Slowakei, Israel, Neuseeland, Österreich und Frankreich sind die Leistungsunterschiede ebenfalls deutlich ausgeprägt und nicht signifikant vom deutschen Wert verschieden. Zwar lässt sich in allen Bildungssystemen ein

Abbildung 6.1: Leistungsvorsprung im Leseverständnis von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland ($p > .05$).

Statistisch signifikante Unterschiede ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

A = Da in England und in den USA kein Elternfragebogen administriert wurde, können für diese Teilnehmer keine Ergebnisse berichtet werden.

B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

signifikanter Vorsprung der bildungsnäheren Schülerinnen und Schüler zeigen, in Hongkong, Malta, Moskau, Ontario und Kasachstan sind diese Differenzen aber weniger als halb so groß wie in Deutschland.

Die Abbildung 6.2 zeigt den Trend dieser Leistungsdifferenzen seit 2001 für diejenigen Teilnehmer, die an mindestens zwei Zyklen neben IGLU 2016 teilgenommen haben, signifikant besser oder nicht signifikant verschieden zu Deutschland im Lesen in IGLU 2016 abgeschnitten haben oder der EU beziehungsweise OECD angehören (siehe Kapitel 2, Abschnitt 4.2.1 in diesem Band). Deutschland gehört neben der Slowakei, Ungarn und Slowenien zu den vier Staaten, in denen seit 2001 eine signifikante Vergrößerung der sozialen Disparitäten stattgefunden hat. Während der Anstieg der Leistungsunterschiede in den drei osteuropäischen Staaten während der 15 Jahre relativ konstant war, muss für Deutschland festgehalten werden, dass es von 2011 zu 2016 einen Anstieg um 13 Punkte gegeben hat, während die sozialen Disparitäten zuvor ohne nennenswerte Schwankungen etwa 40 Punkte betragen. Somit kann an dieser Stelle zwar nicht von einem Trend gesprochen werden, dennoch zeigt sich in 2016 ein Effekt in die ungünstige Richtung.

Positive Entwicklungen zeigen sich in Spanien, wo sich die Leistungsdifferenzen seit 2006 signifikant verringert haben. Auch Hongkong ist an dieser Stelle zu nennen, da sich die Werte seit 2001 von 21 auf 11 Punkte nahezu halbiert haben – auch wenn dies zufallskritisch nicht abgesichert ist.

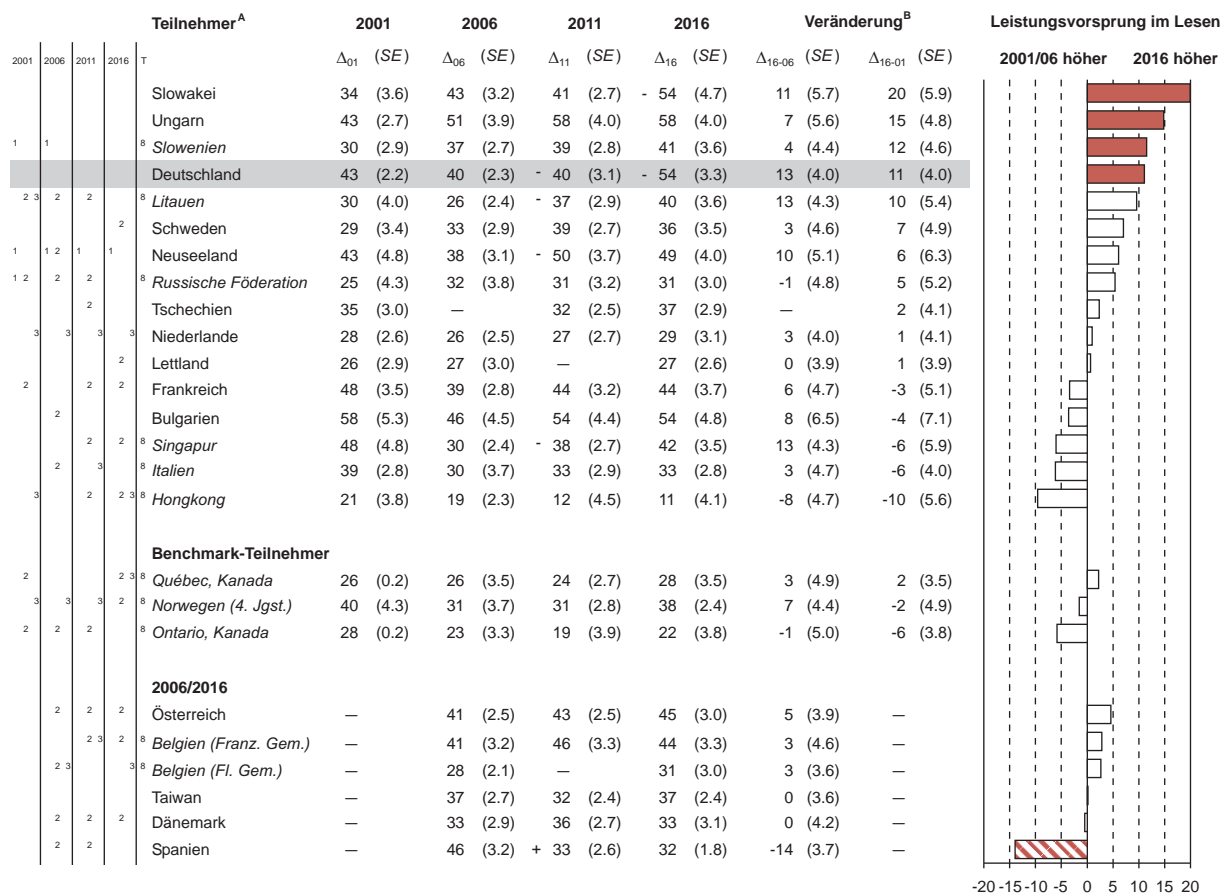
Noch größere Unterschiede zwischen den Teilnehmerstaaten und -regionen als bei der Büchervariablen zeigen sich hinsichtlich des *Berufsstatus* (siehe Abbildung 6.3). Der Anteil der ersten Berufsgruppe (manuelle Tätigkeiten) liegt zwischen 3.2 Prozent in Moskau und 30.1 Prozent in Bulgarien. Beträgt der Anteil einer Gruppe weniger als 10 Prozent, wie dies neben Moskau bei zehn Teilnehmern der Fall ist, sind Aussagen zur Leistungsdifferenz nur eingeschränkt möglich. In Deutschland entspricht der Wert mit 13.2 Prozent ungefähr dem Durchschnitt der OECD-Staaten. Der Anteil der dritten Berufsgruppe (Akademiker, Techniker und Führungskräfte) ist mit 29.2 Prozent in keinem Bildungssystem so gering wie in Deutschland. Ähnliche Werte finden sich nur in Italien, Österreich und Chile. Betrachtet man die internationalen Referenzwerte, so liegen diese mit 46.8 Prozent (VG EU) bis 49.1 Prozent (VG OECD) deutlich höher. Mehr als zwei Drittel der Familien gehören in Singapur, Norwegen, Schweden und Moskau zu dieser Gruppe.

Hinsichtlich der Leistungsdifferenzen zwischen den beiden dargestellten Gruppen findet sich wie bei der Büchervariable kein Teilnehmer, bei dem diese Differenz signifikant größer ausfällt als in Deutschland. Nominell größer ist diese Differenz in Israel, Bulgarien, Ungarn, Singapur und der Slowakei. Hierzulande erzielen Kinder von Akademikern, Technikern und Führungskräften 72 Punkte (ungefähr anderthalb Lernjahre) höhere Leistungen als Kinder von manuell tätigen Personen. In Hongkong beträgt der Leistungsvorsprung lediglich 8 Punkte und ist damit nicht signifikant. Ebenfalls relativ geringe soziale Disparitäten finden sich in Macau und Kasachstan.

Im Trend (ohne Abbildung) zeigt sich für Deutschland ein Anstieg dieser Differenz um ungefähr 10 Punkte (2011: 61 Punkte, 2006: 60 Punkte), der allerdings nicht statistisch signifikant ist.

Als dritte Indikatorvariable für den sozialen Status wird in Abbildung 6.4 das *Bildungsniveau der Eltern* betrachtet. Der Anteil von Familien, in denen kein Elternteil einen tertiären Bildungsabschluss erreicht hat, liegt im internationalen Vergleich bei rund 58 Prozent und der Anteil von Familien mit tertiärem

Abbildung 6.2: Unterschiede im Leistungsvorsprung im Leseverständnis von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern im Vergleich von IGLU 2001, 2006, 2011 und 2016



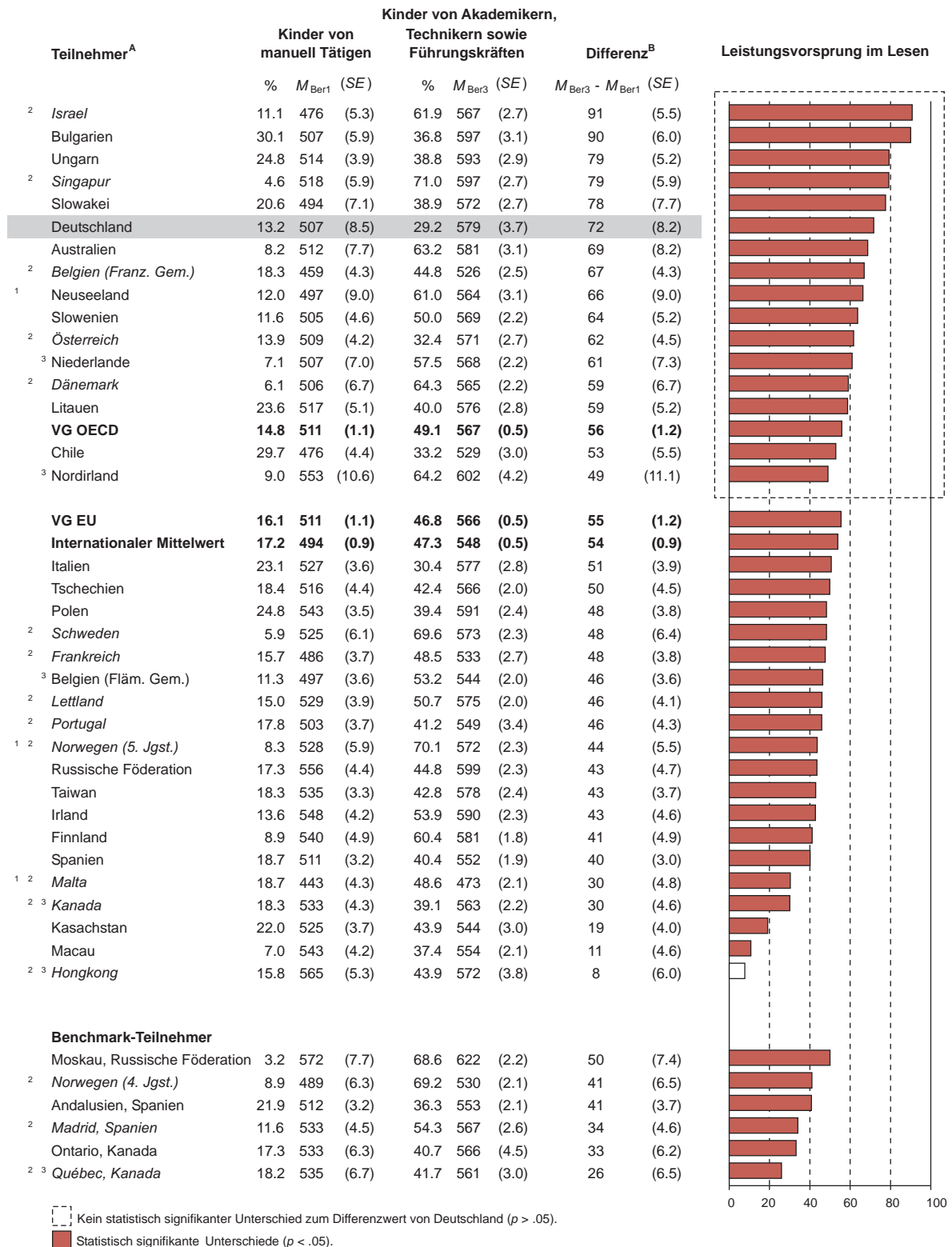
■ Statistisch signifikante Veränderungen zwischen 2001 und 2016 ($p < .05$).
 ▨ Statistisch signifikante Veränderungen zwischen 2006 und 2016 ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse zwischen den Studienzyklen ausgegangen werden muss.
 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
 8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen zwischen den Studienzyklen.
 A = Da in England und in den USA kein Elternfragebogen administriert wurde, können hier für diese Teilnehmer keine Ergebnisse berichtet werden.
 B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.
 + = Differenzwert 2006 bzw. 2011 statistisch signifikant höher als 2016 ($p < .05$).
 - = Differenzwert 2006 bzw. 2011 statistisch signifikant niedriger als 2016 ($p < .05$).
 $\Delta_{16-01} / \Delta_{16-06}$ = Differenz in den Mittelwerten zwischen 2016 und 2001 bzw. 2006 (bei denjenigen Teilnehmern, die nicht an IGLU 2001 teilgenommen haben).

Bildungsabschluss entsprechend bei rund 42 Prozent. Zwischen den Teilnehmern zeigen sich wiederum große Unterschiede. So gehören in Italien nur 18.3 Prozent der Familien zu der Gruppe mit tertiärem Bildungsabschluss und auch in Chile, Hongkong, Macau, Malta, Österreich und Tschechien liegt dieser Wert unter oder nur knapp über 30 Prozent. Hoch ist der Anteil hingegen in Moskau (76.7%), Dänemark (63.9%) und Norwegen (4. Jgst.: 61.9%; 5. Jgst.: 61.1%).

In Deutschland erzielen Kinder aus Familien, in denen kein Elternteil einen (Fach-)Hochschulabschluss besitzt, 48 Punkte weniger im Leseverständnis als Kinder aus Familien, in denen dies der Fall ist. Nur in Bulgarien, Israel, Ungarn und Singapur fallen die sozialen Disparitäten in diesem Vergleich signifikant größer aus. Wie schon bei den anderen Indikatoren zeigen sich in Hongkong,

Abbildung 6.3: Leistungsvorsprung im Leseverständnis von Kindern aus Familien der dritten Berufsgruppe (Akademiker, Techniker und Führungskräfte) vor denen der ersten Berufsgruppe (manuelle Tätigkeiten)



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland ($p > .05$).

Statistisch signifikante Unterschiede ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

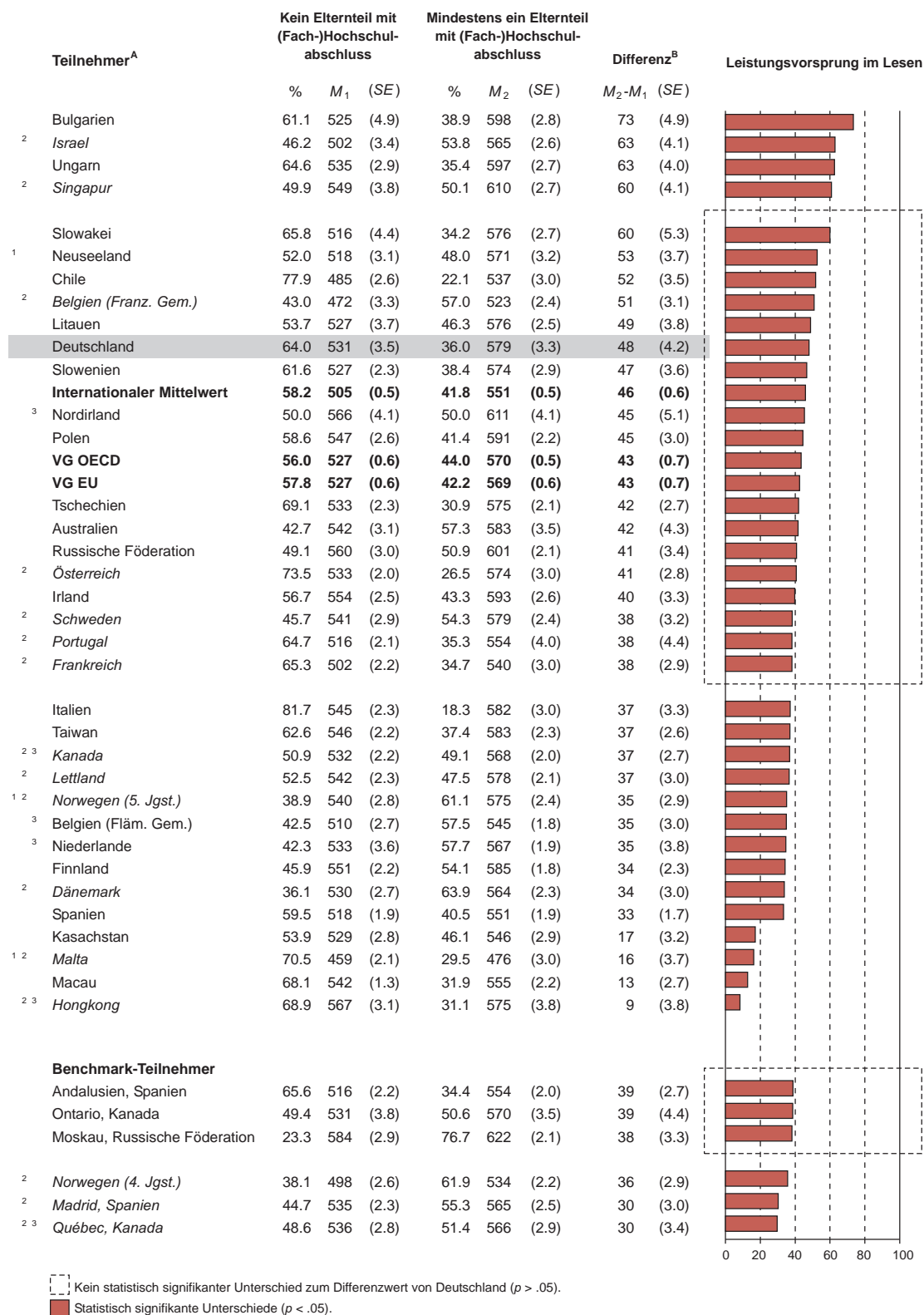
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

A = Da in England und in den USA kein Elternfragebogen administriert wurde, können für diese Teilnehmer keine Ergebnisse berichtet werden.

B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

Abbildung 6.4: Leistungsvorsprung im Leseverständnis von Kindern aus Familien, in denen mindestens ein Elternteil einen tertiären Bildungsabschluss erreicht hat (ISCED-Level 6, 7 und 8) vor Kindern aus Familien, in denen dies nicht der Fall ist



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland (*p* > .05).

Statistisch signifikante Unterschiede (*p* < .05).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülererebene erreichen nicht die internationalen Vorgaben.

A = Da in England und in den USA kein Elternfragebogen administriert wurde, können für diese Teilnehmer keine Ergebnisse berichtet werden.

B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

Macau, Malta und Kasachstan die geringsten Leistungsunterschiede zwischen den Gruppen.

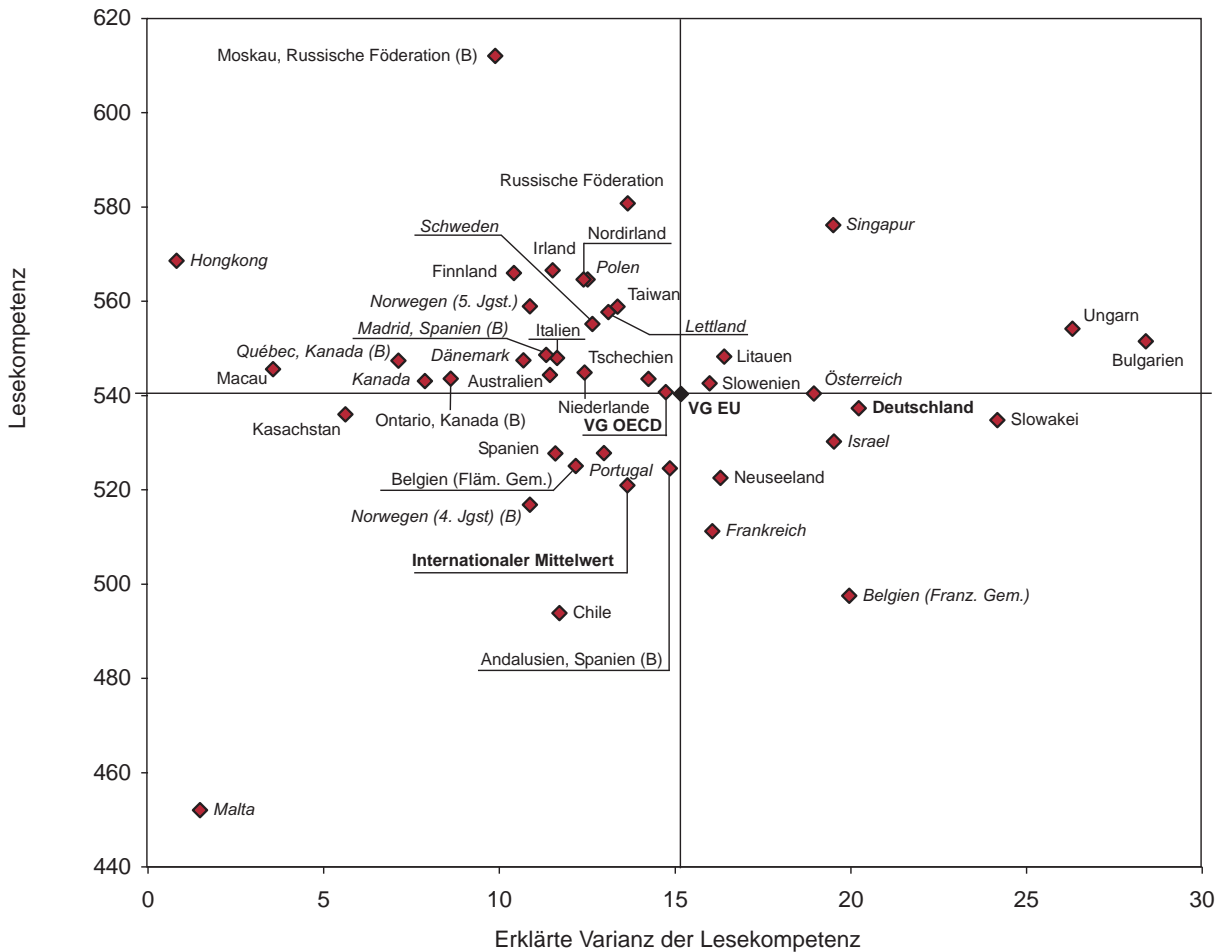
Im Trend (ohne Abbildung) zeigen sich für Deutschland im Vergleich zu 2011 (51 Punkte) und 2006 (43 Punkte) nur minimale – nicht signifikante – Unterschiede. Zu beachten ist im Trendvergleich auch, dass die Kodierung des ISCED von 2011 zu 2016 im oberen Skalenbereich weiter ausdifferenziert wurde, so dass direkte Vergleiche nur eingeschränkt möglich sind.

Mit Hilfe von Regressionsanalysen kann für alle Teilnehmer der Varianzanteil der Lesekompetenz bestimmt werden, der durch die drei zuvor beschriebenen Variablen (Bücher im Haushalt, Berufsstatus, Bildungsniveau) erklärt wird. Die Abbildung 6.5 zeigt den Zusammenhang zwischen dieser Varianzaufklärung und der Höhe der Lesekompetenz in den verschiedenen Bildungssystemen. Die mittlere erklärte Varianz beträgt 15.1 Prozent und die mittlere Lesekompetenz 541 Punkte.

Im oberen linken Quadranten finden sich diejenigen Teilnehmer, denen es gelingt, an ihren Grundschulen ein hohes Leseniveau zu erreichen und zugleich die Unterschiede in der Lesekompetenz – gemessen an den drei Indikatoren der sozialen Herkunft – gering zu halten. In Abhängigkeit von der jeweiligen Situation in den Bildungssystemen kann dies als ein Hinweis auf geringe soziale Disparitäten in den Schülerleistungen interpretiert werden. Neben Hongkong, wo das Ausmaß der sozialen Disparitäten besonders gering ist, gehören beispielsweise auch Macau, die skandinavischen Staaten Finnland, Schweden und Dänemark sowie Kanada zu dieser Gruppe. Einen ebenfalls unterdurchschnittlichen Zusammenhang zwischen den drei Indikatoren der sozialen Herkunft der Schülerfamilien und der individuellen Lesekompetenz, aber gleichzeitig auch unterdurchschnittliche mittlere Leseleistungen erreichen beispielsweise Malta, Kasachstan, Spanien, Portugal und Chile (unterer linker Quadrant).

In den beiden rechten Quadranten befinden sich die Staaten und Regionen mit einem überdurchschnittlichen Anteil an erklärter Varianz durch die Indikatoren der sozialen Herkunft und der Leseleistung. Dies kann auf eine enge Koppelung zwischen sozialer Herkunft und Leseleistung hinweisen. Deutschland gehört neben der Slowakei, Israel, Neuseeland, Frankreich und der Französischen Gemeinschaft in Belgien zu den Teilnehmern, die dabei eine nominell unterdurchschnittliche Lesekompetenz erreichen (unterer rechter Quadrant), während beispielsweise in Singapur, Ungarn und Bulgarien die Lesekompetenz überdurchschnittlich ist (oberer rechter Quadrant). In IGLU 2011 gehörte Deutschland bei dieser Betrachtung noch jener Gruppe von Teilnehmern an, in denen zwar eine relativ enge Kopplung von sozialer Herkunft und Leseleistung bestand, die Lesekompetenz aber überdurchschnittlich ausfiel (oberer rechter Quadrant; vgl. Wendt et al., 2012, Abbildung 6.5).

Abbildung 6.5: Zusammenhang zwischen der erklärten Varianz der Lesekompetenz auf Individualebene durch die Indikatoren der sozialen Stratifikation (ISCED, Berufsstatus, Bücher im Haushalt) und den mittleren Lesekompetenzen in den Teilnehmerstaaten



Da in England und den USA kein Elternfragebogen administriert wurde, können für diese Teilnehmer keine Ergebnisse berichtet werden. Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss. B = Benchmark-Teilnehmer

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

5 Der Zusammenhang zwischen sozialer Herkunft und Lesekompetenzen in Deutschland

International wurden in den IGLU/PIRLS-Fragebögen nur sehr wenige Indikatoren der sozialen Herkunft eingesetzt (siehe Abschnitt 3), so dass die IGLU-Fragebögen in Deutschland um weitere Fragen zu diesem Thema ergänzt wurden. In diesem Abschnitt werden ausschließlich nationale Befunde berichtet.

Wie bereits dargestellt sind die unterschiedlichen Variablen zum Berufsstatus von Personen besonders geeignete Indikatoren für den sozialen Status von Familien. In allen bisherigen IGLU-Erhebungen wurden die Eltern offen nach ihren Berufen gefragt, um anschließend EGP-Klassen und den ISEI berechnen zu können.

Tabelle 6.1 zeigt die mittlere Lesekompetenz für die einzelnen EGP-Klassen im Zeitverlauf. Schwankungen zwischen den Erhebungszeitpunkten lassen sich teilweise durch Veränderungen in der Lesekompetenz über alle Schülerinnen und Schüler hinweg erklären. Im Rahmen dieses Kapitels sind allerdings insbesondere

Tabelle 6.1: Mittlere Lesekompetenz pro EGP-Klasse in Deutschland bei IGLU 2001, 2006, 2011 und 2016

Sozioökonomische Stellung der Familie (EGP-Klassen) ^A	2001 ^B				2006 ^B				2011 ^B				2016			
	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)
Obere Dienstklasse (I)	575	(2.4)	57	(2.1)	581	(2.9)	56	(2.9)	575	(3.1)	60	(1.8)	579	(3.3)	61	(2.6)
Untere Dienstklasse (II)	564	(2.9)	62	(2.3)	571	(3.5)	56	(2.5)	563	(3.4)	65	(3.2)	572	(3.6)	65	(2.8)
Routinedienstleistungen (III)	543	(3.4)	61	(2.6)	551	(3.8)	61	(2.8)	546	(5.6)	63	(5.3)	545	(5.9)	65	(5.5)
Selbstständige (IV)	549	(4.0)	62	(2.4)	557	(3.7)	56	(2.8)	540	(4.4)	60	(4.2)	543	(7.7)	77	(4.3)
(Fach-)Arbeiter (V, VI)	532	(2.4)	62	(2.1)	541	(3.1)	61	(2.3)	533	(3.1)	62	(2.7)	537	(4.1)	70	(3.8)
Un- und angelernte Arbeiter (VII)	519	(2.7)	67	(1.8)	529	(2.5)	64	(2.1)	522	(3.5)	62	(2.4)	510	(5.1)	77	(6.8)
Gesamt	539	(1.9)	67	(1.0)	548	(2.2)	67	(1.2)	541	(2.2)	66	(1.4)	537	(3.2)	78	(3.2)

A = Die EGP-Klassifikation erfolgte für die Studienzyklen 2001 und 2006 auf Basis von ISCO-88-Kodes (ILO, 1990), für die Studienzyklen 2011 und 2016 auf Basis von ISCO-08-Kodes (ILO, 2012).

B = Abweichungen in den berichteten Standardfehlern zur nationalen Berichtslegung von IGLU 2011 (Wendt et al., 2012, S. 185) sind in einem differenten Berechnungsverfahren begründet.

IEA: Progress in International Reading Literacy Study (PIRLS)

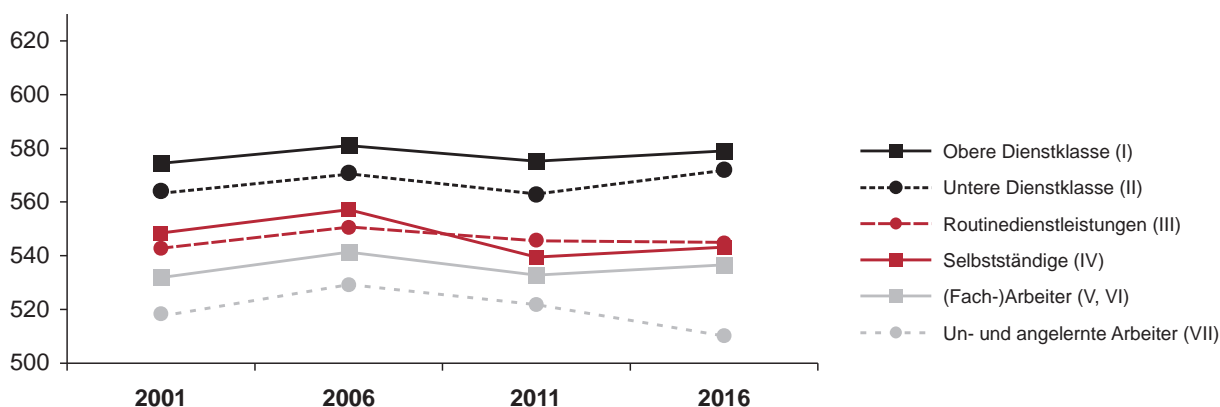
© IGLU 2016

die Unterschiede zwischen den verschiedenen sozialen Lagen relevant. Betrachtet man als Referenzgruppen die obere Dienstklasse sowie die (Fach-)Arbeiter, so zeigt sich im Jahr 2016 für erstere ein Leistungsvorsprung von 42 Punkten (ungefähr ein Lernjahr). In den vergangenen 15 Jahren hat es bei diesem Wert nahezu keine Veränderungen gegeben (2011: 42 Punkte, 2006: 40 Punkte, 2001: 43 Punkte). Etwas anders sieht es aus, wenn man die obere Dienstklasse mit den un- und angelernten Arbeitern vergleicht. Bei den ersten drei Erhebungen in den Jahren 2001, 2006 und 2011 lag dieser Wert zwischen 52 und 56 Punkten, während er in der aktuellen Erhebung 69 Punkte beträgt. Diese (nicht signifikante) Veränderung ist insbesondere auf die Verringerung der Lesekompetenz in der Gruppe der un- und angelernten Arbeiter um 12 Punkte seit IGLU 2011 zurückzuführen.

Diese relative Konsistenz der sozialen Disparitäten in den Schülerleistungen im Zeitverlauf ist auch den relativ parallelen Kurvenverläufen in der Abbildung 6.6 zu entnehmen. Lediglich zwischen den Klassen „Routinedienstleistungen“ und „Selbstständige“ gibt es eine Überschneidung vom Erhebungszyklus 2006 zu 2011, wobei zu betonen ist, dass die Differenzen zwischen diesen beiden Gruppen zu keinem Erhebungszeitpunkt signifikant ausfallen.

Der Zusammenhang zwischen der Lesekompetenz und der zweiten Variable für den Berufsstatus – dem ISEI – lässt sich am besten mit einer Regressionsanalyse darstellen (ohne Abbildung). Im Jahr 2016 liegt der soziale Gradient (Steigung der Regressionsgeraden auf der Individualebene) des höchsten ISEI im Haushalt (HISEI) bei 1.37 und unterscheidet sich damit nicht signifikant von den entsprechenden Werten aus IGLU 2011 (1.23), IGLU 2006 (1.38) und IGLU 2001 (1.24). Das Modell erklärt 14.4 Prozent der Varianz der Lesekompetenz (IGLU 2016) und somit etwa genau so viel wie fünf Jahre zuvor, aber knapp fünf Prozentpunkte mehr als noch im Jahr 2001 (2011: 14.7%, 2006: 12.7%, 2001: 9.7%).

Ein weiterer Indikator für den sozialen Status, der ausschließlich für Deutschland vorliegt, ist die Armutsgefährdung der Schülerfamilien. Nach der

Abbildung 6.6: Mittlere Lesekompetenz pro EGP-Klasse in Deutschland bei IGLU 2001, 2006, 2011 und 2016

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

offiziellen Definition der EU (Bardone & Guio, 2005) gelten 24.5 Prozent der Familien, deren Kinder sich an IGLU 2016 beteiligt haben, als armutsgefährdet (2011: 25.4%; vgl. Stubbe, Tarelli & Wendt, 2012, Abb. 7.5). Wie Abbildung 6.7 zu entnehmen ist, variiert diese Quote stark zwischen gesellschaftlichen Gruppen. Von den Familien in Deutschland, in denen beide Elternteile im Ausland geboren wurden, gelten 38.6 Prozent als armutsgefährdet, während dies nur für 16.8 Prozent der Familien zutrifft, in denen beide Elternteile in Deutschland geboren wurden. Im Jahr 2011 lag die Armutsgefährdungsquote in Familien, in denen beide Elternteile im Ausland geboren wurden, allerdings bei 49.4 Prozent, wobei beachtet werden muss, dass diese Differenz von 10.8 Prozentpunkten statistisch nicht signifikant ist.

Ein noch deutlicherer Zusammenhang zeigt sich zwischen dem Bildungsniveau der Eltern und der Armutsgefährdung der Familien. Verfügen die Eltern über keinen beruflichen Abschluss und haben maximal einen Realschulabschluss erreicht, beträgt die Quote 76.4 Prozent. Hingegen sind nur 9.5 Prozent der Familien armutsgefährdet, in denen mindestens ein Elternteil über einen (Fach-) Hochschulabschluss verfügt. Ähnliche Ergebnisse zeigen sich für die EGP-Klassen: Dort variieren die Quoten zwischen 5.7 Prozent in der oberen Dienstklasse und 41.9 Prozent in der Gruppe der un- und angelernten Arbeiter.

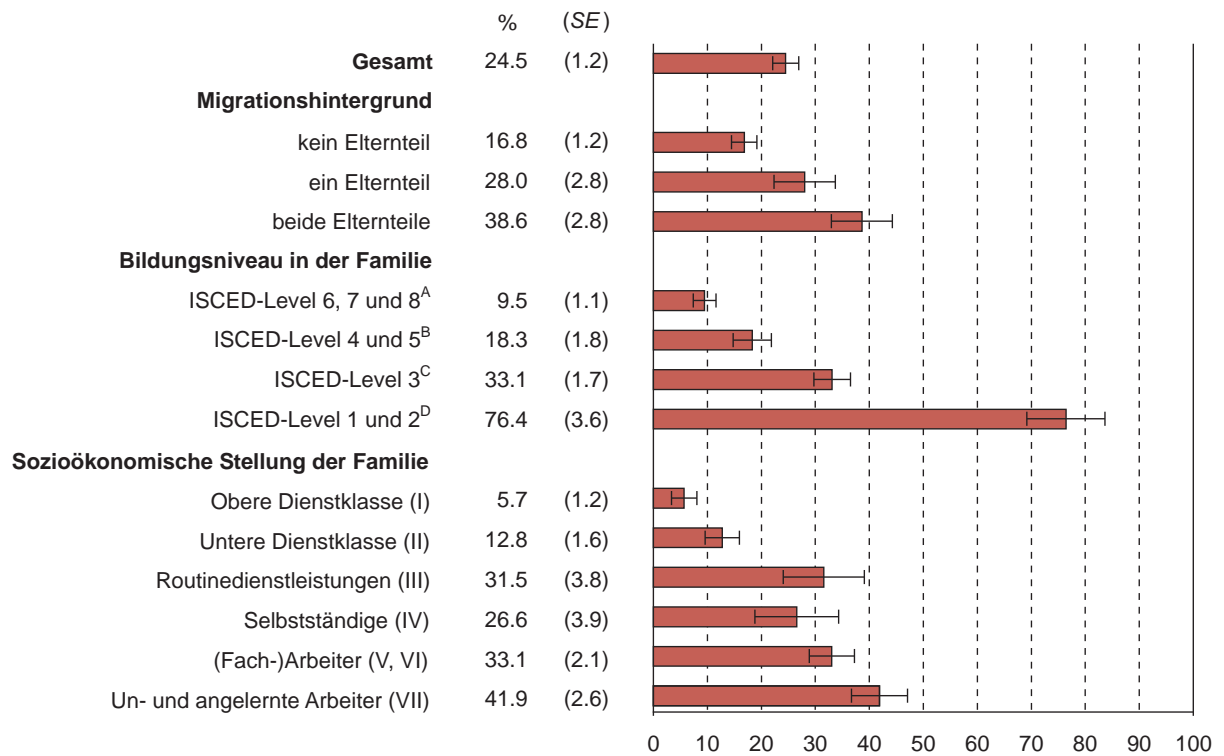
Wie Tabelle 6.2 entnommen werden kann, liegt die Lesekompetenz von Kindern aus nicht armutsgefährdeten Elternhäusern 52 Punkte über der Lesekompetenz von armutsgefährdeten Kindern. Im Jahr 2011 (als erstmalig im Rahmen von IGLU die Armutsgefährdung erfasst wurde) betrug dieser Wert 40 Punkte. Die Differenz von zwölf Punkten ist statistisch jedoch nicht signifikant.

Innerhalb der einzelnen EGP-Klassen liegt die Leistungsdifferenz zwischen armutsgefährdeten und nicht armutsgefährdeten Schülerinnen und Schülern zwischen 20 und 39 Punkten. Bedingt durch die geringe Größe der Teilstichproben erlauben die vorliegenden Daten allerdings kaum generalisierbare Aussagen zu Unterschieden zwischen den einzelnen EGP-Klassen.

Tabelle 6.3 zeigt, wie sich die Leistungstrends von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in Deutschland unter Kontrolle von Merkmalen der sozialen Herkunft darstellen. Die hier vorgenommene Trendperspektive ermöglicht es, Entwicklungstendenzen zu beschreiben.

Die lineare Regression dient als Referenzmodell; dargestellt sind die Koeffizienten, die nicht die Mehrebenenstruktur der Daten berücksichtigen. Hierbei

Abbildung 6.7: Armutsgefährdungsquoten nach Migrationshintergrund, höchstem Bildungsabschluss und EGP-Klasse (in Prozent)



A = (Fach-)Hochschulabschluss.

B = (Fach-)Hochschulreife und Abschluss einer Lehrausbildung (Berufsschule, Lehre) bzw. berufsqualifizierender Abschluss an einer Berufsfachschule (oder gleichwertiger Abschluss). Alternativ dazu: Meister-/Techniker- oder gleichwertiger Fachschulabschluss.

C = (Fach-)Hochschulreife oder Abschluss einer Lehrausbildung (Berufsschule, Lehre) bzw. berufsqualifizierender Abschluss an einer Berufsfachschule (oder gleichwertiger Abschluss).

D = Maximal Realschulabschluss (kein beruflicher Abschluss).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Tabelle 6.2: Mittlere Lesekompetenz nach Armutsgefährdung und EGP-Klasse

Sozioökonomische Stellung der Familie (EGP-Klassen)	nicht armutsgefährdet				armutsgefährdet				Differenz ^A	
	M_n	(SE)	SD	(SE)	M_a	(SE)	SD	(SE)	$M_n - M_a$	(SE)
Obere Dienstklasse (I)	583	(3.5)	59	(2.5)	546	(11.9)	57	(11.4)	37	(12.4)
Untere Dienstklasse (II)	576	(3.8)	65	(3.2)	556	(8.7)	65	(6.9)	20	(9.5)
Routinedienstleistungen (III)	557	(5.3)	56	(4.0)	517	(15.5)	78	(12.6)	39	(16.4)
Selbstständige (IV)	554	(8.1)	70	(5.4)	527	(16.8)	83	(10.2)	27	(18.6)
(Fach-)Arbeiter (V, VI)	550	(5.2)	68	(4.0)	513	(6.8)	68	(7.2)	37	(8.6)
Un- und angelernte Arbeiter (VII)	525	(6.3)	69	(7.1)	490	(9.9)	87	(12.6)	36	(11.7)
Gesamt	563	(2.6)	68	(2.3)	511	(6.3)	80	(6.7)	52	(6.8)

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Tabelle 6.3: Lesekompetenzen zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle von Indikatoren der sozialen Herkunft (Mehrebenenmodell)

	Lineare Regression		Mehrebenenregressionen							
	b	(SE)	Modell 1		Modell 2		Modell 3		Modell 4	
	b	(SE)	b	(SE)	b	(SE)	b	(SE)	b	(SE)
Konstante	531.8	(2.9) **	528.8	(5.4) **	527.9	(5.3) **	545.1	(4.6) **	544.6	(5.1) **
Testzeitpunkt										
IGLU 2001	6.4	(3.4) *	9.3	(6.0) ns	7.5	(5.8) ns	2.5	(4.6) ns	2.3	(5.6) ns
IGLU 2006	14.9	(3.4) **	15.0	(5.6) **	14.7	(5.2) **	8.5	(4.4) *	10.5	(5.3) *
IGLU 2011	4.2	(3.3) ns	6.0	(5.8) ns	5.2	(5.7) ns	-2.5	(4.7) ns	-2.4	(5.5) ns
Merkmale										
Familie mit hohem Bildungsniveau ^A	44.5	(1.5) **	38.7	(3.0) **	25.4	(3.6) **	18.2	(3.3) **	18.1	(3.3) **
Familie mit hohem Berufsstatus ^B					24.3	(3.3) **	15.2	(3.1) **	17.0	(6.5) **
Migrationshintergrund (ein Elternteil) ^C							-12.7	(4.0) **	-12.7	(4.0) **
Migrationshintergrund (beide Elternteile) ^D							-26.5	(3.8) **	-26.6	(3.8) **
Interaktionen										
Hoher Berufsstatus ^B * IGLU 2001									0.5	(8.1) ns
Hoher Berufsstatus ^B * IGLU 2006									-6.4	(8.1) ns
Hoher Berufsstatus ^B * IGLU 2011									-0.5	(7.5) ns
Kovariaten										
Kognitive Fähigkeiten ^E							17.3	(1.0) **	17.3	(1.0) **
Erklärter Anteil der Gesamtvarianz	7.7									
Erklärter Anteil auf Individualebene (70.3%)			15.9		23.6		44.7		44.7	

b = Regressionsgewichte (unstandardisiert).

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

A = Bildungsniveau nach höchstem Bildungsabschluss der Eltern (0 = Kein Elternteil mit (Fach-)Hochschulabschluss;

1 = Mindestens ein Elternteil mit (Fach-)Hochschulabschluss).

B = Berufsstatus nach Angabe der Eltern (0 = Manuell Tätige, Angestellte und kleinere Unternehmer;

1 = Akademiker, Techniker und Führungskräfte).

C = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren;

1 = Ein Elternteil im Ausland geboren), siehe Kapitel 7 in diesem Band.

D = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren;

1 = Beide Elternteile im Ausland geboren), siehe Kapitel 7 in diesem Band.

E = Skalierter Index der kognitiven Fähigkeiten (Messinvarianz über die Zyklen angenommen), siehe Kapitel 2 in diesem Band.

Das Gesamtmodell zu dem hier dargestellten Ausschnitt ist im Anhang B dargestellt, siehe Modelle 6, 7, 8 und 9. Die lineare Regression ist nicht Teil des Gesamtmodells.

zeigt sich, dass die Lesekompetenz unter Kontrolle des Bildungsniveaus der Familien im Jahr 2001 6 Punkte und im Jahr 2006 15 Punkte höher lag als in der aktuellen Erhebung. Theoriekonform zeigt sich zudem ein deutlicher Effekt des Bildungsniveaus des Elternhauses. Über alle Erhebungszeitpunkte hinweg erreichen Kinder aus Familien, in denen mindestens ein Elternteil einen (Fach-) Hochschulabschluss besitzt, 45 Punkte (ungefähr ein Lernjahr) mehr als Kinder aus Familien, in denen dies nicht der Fall ist. Insgesamt erklärt dieses Modell 8 Prozent der Gesamtvarianz.

In den Modellen 1 bis 4 werden die Ergebnisse von Mehrebenenregressionen berichtet (siehe Kapitel 2, Abschnitt 11 in diesem Band). Damit wird der Tatsache Rechnung getragen, dass die in IGLU getesteten Viertklässlerinnen

und Viertklässler im Klassenverbund unterrichtet werden und insofern mehr gemeinsame Charakteristika (dieselben Lehrerinnen und Lehrer, vergleichbare Lerngelegenheiten usw.) haben als Kinder aus verschiedenen Klassen.

Analysen der Mehrebenenstruktur haben gezeigt, dass sich 70.3 Prozent der Varianz der Lesekompetenz über alle Erhebungszeitpunkte hinweg durch Merkmale auf der Individualebene (Level 1, *within*) erklären lassen und entsprechend 29.7 Prozent durch Merkmale auf der Klassenebene (Level 2, *between*). In den in diesem Kapitel dargestellten Mehrebenenregressionen werden zur Erklärung der Lesekompetenz neben der Berücksichtigung des Studienzyklus ausschließlich Variablen auf Individualebene aufgenommen, so dass in den Modellen 1 bis 4 angegeben wird, welcher Prozentsatz des Varianzanteils auf Individualebene jeweils durch das Modell erklärt wird.

Hinsichtlich der berücksichtigten Variablen unterscheidet sich Modell 1 nicht vom Referenzmodell. Entsprechend zeigen sich auch nur geringfügige Änderungen in den Koeffizienten. Berücksichtigt man also die Mehrebenenstruktur, dann reduzieren sich die 45 Punkte auf rund 40 Punkte. Unter Berücksichtigung der Mehrebenenstruktur bleibt der Effekt des Bildungsniveaus hochsignifikant. Bei den Erhebungszeitpunkten ist nur die Differenz zwischen 2016 und 2006 signifikant. In Modell 2 wird ein weiterer Indikator für den sozialen Status aufgenommen. Der bereits aus Abbildung 6.3 bekannte Berufsstatus wird in dieser Tabelle in dichotomisierter Form dargestellt, wobei die ersten beiden Berufsgruppen (manuelle Tätigkeiten sowie Angestellte und kleinere Unternehmer) zusammengefasst und der dritten Berufsgruppe (Akademiker, Techniker und Führungskräfte) gegenübergestellt werden. Die Effekte der beiden Indikatoren für den sozioökonomischen Status sind etwa gleich groß (25 bzw. 24 Punkte). Hinsichtlich der erklärten Varianz auf Individualebene ist das umfangreichere Modell 2 mit 23.6 Prozent etwas erklärungsstärker als das Modell 1 (15.9%).

In den Modellen 3 und 4 werden weitere Variablen aufgenommen, deren Zusammenhang mit schulischen Kompetenzen gut dokumentiert ist. Entsprechend erklären diese beiden Modelle fast die Hälfte der auf Individualebene vorhandenen Varianz in der Lesekompetenz.

Als Kovariate werden die kognitiven Fähigkeiten der Viertklässlerinnen und Viertklässler berücksichtigt (siehe Kapitel 2 in diesem Band). Eine Zunahme dieses Wertes um eine Standardabweichung geht mit einer um 17 Punkte höheren Lesekompetenz einher.

Zudem wird in den Modellen 3 und 4 der Migrationshintergrund der Kinder aufgenommen – ein Merkmal, das häufig im Zusammenhang mit dem Thema „soziale Gerechtigkeit“ aufgegriffen wird, zumal bekannt ist, dass ein enger Zusammenhang zwischen diesen beiden Variablen besteht (siehe Kapitel 7 in diesem Band). Verglichen mit Schülerinnen und Schülern, deren Eltern beide in Deutschland geboren wurden, schneiden diejenigen Viertklässlerinnen und Viertklässler, bei denen beide Elternteile im Ausland geboren wurden, 27 Punkte schlechter ab. Wurde nur ein Elternteil im Ausland geboren, beträgt dieser Wert 13 Punkte. Der Einfluss der beiden Indikatoren zum sozioökonomischen Status fällt bei Berücksichtigung des Migrationshintergrundes zwar etwas geringer aus, bleibt jedoch deutlich signifikant.

Das Modell 4 unterscheidet sich von dem vorangegangenen Modell dadurch, dass die Interaktion zwischen Erhebungszyklus und Berufsstatus kontrolliert wurde, um Veränderungen der sozialen Disparitäten im Zeitverlauf aufzeigen zu können. Keiner dieser Interaktionsterme nimmt allerdings signifikante Werte an.

6 Zusammenfassung

Die Befunde aus IGLU 2016 bestätigen insgesamt die aus den vergangenen Studien bekannten und theoriekonformen Erkenntnisse zum Zusammenhang von Lesekompetenzen und sozialer Herkunft: In nahezu allen teilnehmenden Staaten und Regionen (Ausnahme Hongkong bei Betrachtung des Indikators Berufsgruppen, siehe Abbildung 6.3) sind signifikante Unterschiede in der Lesekompetenz zwischen Schülerinnen und Schülern oberer und unterer sozialer Lagen zu beobachten. Deutlich zeigt sich in international vergleichender Perspektive dabei auch, dass sich das Ausmaß zwischen den Teilnehmern erheblich voneinander unterscheidet. Einigen gelingt es entscheidend besser als Deutschland, Lösungen für den Abbau von sozialen Disparitäten in ihrem Bildungssystem umzusetzen. Zum Beispiel gemessen am Indikator der Büchervariable zeigen sich in der Trendperspektive insbesondere in Spanien positive Entwicklungen seit 2006. Deutschland hingegen gehört neben der Slowakei, Slowenien und Ungarn zu den vier Staaten, in denen seit 2001 eine signifikante Vergrößerung der sozialen Disparitäten stattgefunden hat. Die Notwendigkeit, soziale Disparitäten bei der Beschreibung von Bildungssystemen im Rahmen von Schulleistungsstudien zu berücksichtigen, ist aufgrund der aktuellen Befunde damit weiterhin gegeben.

In nationaler Perspektive sind vertiefende Analysen zum Zusammenhang von Lesekompetenzen und sozialer Herkunft auf Basis der IGLU-Daten 2016 möglich, da hierzulande mehr Indikatoren der sozialen Lage durch die schriftliche Befragung der Eltern erfasst wurden als international. Die internationalen Befunde bestätigen sich aber auch hier, das heißt, es zeigt sich deutlich, dass ausgeprägte soziale Disparitäten im deutschen Bildungssystem vorhanden sind. Trotz verbreiteter, programmatischer Bemühungen in den vergangenen 15 Jahren haben sich somit keine signifikanten Veränderungen der sozialen Disparitäten ergeben. Nach wie vor sind die Leistungsunterschiede zwischen Kindern, die in armutsgefährdeten Elternhäusern aufwachsen, und ihren Mitschülerinnen und Mitschülern, die nicht armutsgefährdet sind, sehr groß. Sie entsprechen umgerechnet etwa einem Lernjahr. Selbiges gilt für Kinder, deren Eltern den Dienstklassen angehören im Vergleich zu jenen, deren Eltern den Gruppen der Arbeiter angehören. Im Hinblick auf die bereits einleitend genannten Teilhabechancen ist damit festzustellen, dass es Deutschland noch nicht gelungen ist, einen allgemeinen Anspruch auf Chancengleichheit im Bildungssystem durch- beziehungsweise umzusetzen.

Literatur

- Anders, Y. & Roßbach, H. G. (2013). Frühkindliche Bildungsforschung in Deutschland. In M. Stamm & D. Edelmann (Hrsg.), *Handbuch frühkindliche Bildungsforschung* (S. 183–195). Wiesbaden: Springer. doi: 10.1007/978-3-531-19066-2_13
- Bardone, L. & Guio, A.-C. (2005). In-work poverty. New commonly agreed indicators at EU level. *Statistics in focus*, 5, 1–11.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiss (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–407). Opladen: Leske + Budrich.

- Becker, R. & Lauterbach, W. (2004). Dauerhafte Bildungsungleichheiten – Ursachen, Mechanismen, Prozesse und Wirkungen. In R. Becker & W. Lauterbach (Hrsg.), *Bildung als Privileg?* (S. 9–40). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bertelsmann Stiftung. (2016). *Steigende Kinderarmut beeinträchtigt Chancen fürs ganze Leben. Pressemitteilung vom 12.09.2016*. Zugriff am 17.07.2017 unter <http://www.bertelsmann-stiftung.de/de/presse/pressemitteilungen/pressemitteilung/pid/steigende-kinderarmut-beeintraechtigt-chancen-fuers-ganze-leben/>
- Blossfeld, H.-P. (1993). Changes in educational opportunities in the Federal Republic of Germany. A longitudinal study of cohorts born between 1916 and 1965. In Y. Shavit & H.-P. Blossfeld (Hrsg.), *Persistent inequality. Changing educational attainment in thirteen countries* (S. 51–74). Boulder: Westview Press.
- Blossfeld, H.-P. & Shavit, Y. (1993). Dauerhafte Ungleichheiten: zur Veränderung des Einflusses der sozialen Herkunft auf die Bildungschancen in dreizehn industrialisierten Ländern. *Zeitschrift für Pädagogik*, 39 (1), 25–52.
- BMAS – Bundesministerium für Arbeit und Soziales. (2017). *Lebenslagen in Deutschland. Armuts- und Reichtumsberichterstattung der Bundesregierung. Der fünfte Armuts- und Reichtumsbericht der Bundesregierung*. Zugriff am 17.07.2017 unter http://www.bmas.de/SharedDocs/Downloads/DE/PDF-Pressemitteilungen/2017/5-arb-langfassung.pdf?__blob=publicationFile&v=4
- Bonsen, M., Frey, K. A. & Bos, W. (2008). Soziale Herkunft. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 141–156). Münster: Waxmann.
- Bos, W., Schwippert, K. & Stubbe, T. C. (2007). Die Koppelung von sozialer Herkunft und Schülerleistung im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 225–247). Münster: Waxmann.
- Bos, W., Stubbe, T. C. & Buddeberg, M. (2010). Gibt es eine armutsbedingte Bildungsbenachteiligung? Die Operationalisierung verschiedener Indikatoren der sozialen Herkunft in der empirischen Bildungsforschung. In D. H. Rost (Hrsg.), *Intelligenz, Hochbegabung, Vorschulerziehung, Bildungsbenachteiligung* (S. 165–208). Münster: Waxmann.
- Bosch, G. & Kalina, T. (2007). Niedriglöhne in Deutschland – Zahlen, Fakten, Ursachen. In G. Bosch & C. Weinkopf (Hrsg.), *Arbeiten für wenig Geld. Niedriglohnbeschäftigung in Deutschland* (S. 20–105). Frankfurt a. M.: Campus.
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (Soziale Welt, Sonderband 2, S. 183–198). Göttingen: Schwartz.
- Butterwegge, C. (Hrsg.). (2000). *Kinderarmut in Deutschland. Ursachen, Erscheinungsformen und Gegenmaßnahmen*. Frankfurt a. M.: Campus.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94 (1), 95–120.
- Ehmke, T. & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. 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. 231–254). Münster: Waxmann.
- Erikson, R., Goldthorpe, H. J. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30 (4), 415–441.
- Ganzeboom, H. B. G., Graaf, P. M. de & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21 (1), 1–56.
- Ganzeboom, H. B. G. & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research*, 25 (3), 201–239.
- Geißler, R. (2014). Bildungsexpansion und Bildungschancen. In Bundeszentrale für politische Bildung (Hrsg.), *Sozialer Wandel in Deutschland* (Informationen zur politischen Bildung, Heft 324). Zugriff am 17.07.2017 unter <http://www.bpb.de/izpb/198031/bildungsexpansion-und-bildungschancen?p=all>
- Haag, N., Kocaj, A., Jansen, M. & Kuhl, P. (2017). Soziale Disparitäten. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016*.

- Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 213–235). Münster: Waxmann.
- Hasselhorn, M. & Kuger, S. (2014). Wirksamkeit schulrelevanter Förderung in Kindertagesstätten. *Zeitschrift für Erziehungswissenschaft*, 17 (2), 299–314. doi:10.1007/s11618-013-0473-2
- Hoffmann, E. (2003). International statistical comparisons of occupational and social structures. Problems, possibilities and the role of ISCO-88. In J. H. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in cross-national comparison. A European working book for demographic and socio-economic variables* (S. 137–158). New York: Plenum Press.
- ILO – International Labour Office. (1990). *International Standards Classification of Occupations. ISCO-88*. Genf: International Labour Office.
- ILO – International Labour Office. (2012). *International Standards Classification of Occupations. ISCO-08*. Genf: International Labour Office.
- Laubstein, C., Holz, G. & Seddig, N. (2016). *Armutsfolgen für Kinder und Jugendliche. Erkenntnisse aus empirischen Studien in Deutschland*. Gütersloh: Bertelsmann Stiftung.
- Lorenz, J. & Stubbe, T. C. (2017). *Families' social capital and students' achievement in Mathematics and Science. Findings from TIMSS 2015 Germany*. Manuskript in Vorbereitung.
- 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.
- 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.
- Schimpl-Neimanns, B. (2000). Soziale Herkunft und Bildungsbeteiligung. Empirische Analysen zu herkunftsspezifischen Bildungsungleichheiten zwischen 1950 und 1989. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 52 (4), 636–669.
- Schroedter, J. H., Lechert, Y. & Lüttinger, P. (2006). *Die Umsetzung der Bildungsskala ISCED – 1997 für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976–2004. ZUMA-Methodenbericht 2006/08*. Zugriff am 23.11.2017 unter https://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_methodenberichte/2006/06_08_Schroedter.pdf
- Solga, H. (2005). Meritokratie – die moderne Legitimation ungleicher Bildungschancen. In P. A. Berger & H. Kahlert (Hrsg.), *Institutionalisierte Ungleichheiten. Wie das Bildungswesen Chancen blockiert* (S. 19–38). Weinheim: Juventa.
- Stanat, P., Weirich, S. & Radmann, S. (2012). Sprach- und Leseförderung. 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. 251–276). Münster: Waxmann.
- Statistisches Bundesamt. (2016a). *Armutsgefährdung in Westdeutschland im 10-Jahres-Vergleich gestiegen* (Pressemitteilung vom 22. September 2016 – 334/16). Wiesbaden: Statistisches Bundesamt.
- Statistisches Bundesamt. (2016b). *Anteil der von Armut oder sozialer Ausgrenzung bedrohten Bevölkerung in Deutschland nahezu unverändert* (Pressemitteilung vom 3. November 2016 – 391/16). Wiesbaden: Statistisches Bundesamt.
- Strietholt, R., Manitius, V., Berkemeyer, N. & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagschulen. *Zeitschrift für Erziehungswissenschaft*, 18 (4), 737–761. doi: 10.1007/s11618-015-0634-6
- 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-E 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.), *Mathematische und naturwissenschaftli-*

- che 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*. Münster: Waxmann.
- Tophoven, S., Lietzmann, T., Reiter, S. & Wenzig, C. (2017). *Armutsmuster in Kindheit und Jugend. Längsschnittbetrachtungen von Kinderarmut*. Zugriff am 27.19.2017 unter https://www.bertelsmann-stiftung.de/fileadmin/files/Projekte/Familie_und_Bildung/Studie_WB_Armutsmuster_in_Kindheit_und_Jugend_2017.pdf
- UNESCO. (2003). International Standard Classification of Education, ISCED 1997. In J. H. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in cross-national comparison. A European working book for demographic and socio-economic variables*. New York: Plenum Press.
- Valtin, R. & Tarelli, I. (2014). *Lesekompetenz nachhaltig stärken. Evidenzbasierte Maßnahmen und Programme*. Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- vbw – Vereinigung der Bayerischen Wirtschaft e. V. (Hrsg.). (2013). *Zwischenbilanz Ganztagsgrundschulen: Betreuung oder Rhythmisierung?* Münster: Waxmann.
- vbw – Vereinigung der Bayerischen Wirtschaft e. V. (Hrsg.). (2017). *Bildung 2030 – veränderte Welt. Fragen an die Bildungspolitik*. Münster: Waxmann.
- Vester, M. (2004). Die Illusion der Bildungsexpansion. Bildungsöffnungen und soziale Segregation in der Bundesrepublik Deutschland. In S. Engler & B. Kraus (Hrsg.), *Das kulturelle Kapital und die Macht der Klassenstrukturen. Sozialstrukturelle Verschiebungen und Wandlungsprozesse des Habitus* (S. 13–54). Weinheim: Juventa.
- Vester, M. (2005). Die selektive Bildungsexpansion. Die ständische Regulierung der Bildungschancen in Deutschland. In P. A. Berger & H. Kahlert (Hrsg.), *Institutionalisierte Ungleichheiten. Wie das Bildungswesen Chancen blockiert* (S. 39–70). Weinheim: Juventa.
- Walzebug, A. (2015). *Sprachlich bedingte soziale Ungleichheit. Theoretische und empirische Betrachtungen am Beispiel mathematischer Testaufgaben und ihrer Bearbeitung*. Münster: Waxmann.
- Wendt, H., Goy, M., Walzebug, A. & Valtin, R. (2016). Bildungsangebote an Ganz- und Halbtagsgrundschulen in Deutschland. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 225–245). 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.), *Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 175–190). Münster: Waxmann.

Kapitel VII

Lesekompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund

Heike Wendt und Knut Schwippert

1 Einleitung

Migration hat viele Facetten, die bedeutend für die Bildungsintegration sind. Modifizierte Einbürgerungsrechte und sich verändernde Migrationsbewegungen stellen besondere Herausforderungen an das Bildungssystem dar (z. B. Autorengruppe Bildungsberichterstattung, 2016). Neben der Multikulturalität und der damit einhergehenden Bereicherung des Schulalltags stellen die Integration und Förderung von Kindern und Jugendlichen, die noch nicht lange in Deutschland leben oder deren Eltern nach Deutschland zugewandert sind, spezielle Anforderungen an die Gestaltung von Schule und Unterricht.

Die Grundschule hat den Auftrag, Kindern grundlegende Kompetenzen zu vermitteln, und es gilt ausdrücklich als bildungspolitisches Ziel, sowohl zuwanderungsbezogene Disparitäten in der Bildungsbeteiligung als auch im Kompetenzerwerb möglichst gering zu halten und zu reduzieren (KMK, 2002).

Nationale und internationale Schulleistungsstudien belegen seit Jahrzehnten, dass sich die Leistungen von Schülerinnen und Schülern derselben Altersgruppe aufgrund unterschiedlicher Hintergrundmerkmale voneinander unterscheiden. Spätestens mit der *Internationalen Reading Literacy Study* (Elley, 1992; Lehmann, Peek, Pieper & von Stritzky, 1995) wurde der Nachweis erbracht, dass Viertklässlerinnen und Viertklässler mit Migrationshintergrund im Vergleich zu ihren Mitschülerinnen und Mitschülern ohne Migrationshintergrund im Durchschnitt niedrigere Leseleistungen erzielen. Mit der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) als Folgestudie kann dieser Befund seit 2001 bestätigt werden (Schwippert, Bos & Lankes, 2003; Schwippert, Hornberg, Freiberg & Stubbe, 2007; Schwippert, Wendt & Tarelli, 2012). Die Leistungsunterschiede liegen in der Größenordnung dessen, was Kinder in einem Schuljahr durchschnittlich dazulernen (Hornberg, Bos, Buddeberg, Potthoff & Stubbe, 2007), teilweise sogar deutlich darüber. Zudem zeigt sich, dass

Viertklässlerinnen und Viertklässler mit Migrationshintergrund in Deutschland im Lesen in den niedrigen Kompetenzstufen überrepräsentiert, in den höchsten Kompetenzstufen hingegen deutlich unterrepräsentiert sind. Der *IQB-Bildungstrend 2016* sowie bereits vorangehende, bis 2006 im Rahmen von IGLU realisierte Ländervergleiche (Schwippert, Bos & Lankes, 2004; Schwippert, Hornberg & Goy, 2008; Haag, Böhme & Stanat, 2012) zeigen zudem deutliche regionale Unterschiede auf (Rjosk, Haag, Heppt & Stanat, 2017). Besonders hoch fallen dabei zuwanderungsbezogene Disparitäten in den drei Stadtstaaten aus. Für einige Länder der Bundesrepublik Deutschland, wie beispielsweise Bayern oder Brandenburg, zeigen sich hingegen keine signifikanten Unterschiede in den Lesekompetenzen zwischen Kindern mit einem im Ausland geborenen Elternteil im Vergleich zu Kindern ohne Migrationshintergrund. Auch unterscheiden sich die Disparitäten deutlich zwischen unterschiedlichen Geburtsländern der Kinder oder Eltern (ebd.).

Leistungsunterschiede, die zuungunsten von Schülerinnen und Schülern aus Familien mit jüngerer Migrationsgeschichte ausfallen, werden im Wesentlichen auf zwei Ebenen diskutiert:

Aus *individueller Perspektive* werden familiäre und individuelle Eigenschaften von Schülerinnen und Schülern mit Migrationshintergrund betrachtet. Diskutiert wird hier der Unterschied zwischen Familien mit und ohne Migrationshintergrund im Hinblick auf verschiedene sozioökonomische und soziokulturelle Merkmale, wie beispielsweise die Einkommensverhältnisse oder das Bildungsniveau der Eltern, die bei ungünstiger Ausprägung nachteilige Einflüsse auf den Kompetenzerwerb ausüben können. Es zeigt sich, dass ein nicht unerheblicher Anteil der Leistungsdisparitäten auf Unterschiede in den sozioökonomischen Lebensverhältnissen zurückzuführen ist (z.B. Autorengruppe Bildungsberichterstattung, 2016; siehe auch Kapitel 6 in diesem Band). Erweiterten Erwartungswert-Modellen zufolge kann angenommen werden, dass das soziokulturelle Milieu, das auch durch den Zuwanderungsstatus geprägt sein kann, in dem ein Kind aufwächst, familiäre Prozessmerkmale (wie kulturelle Aktivitäten oder die kommunikative Praxis in Familien) beeinflusst, die ihrerseits wiederum für die Ausprägungen individueller Merkmale von Kindern (wie Einstellungen und Verhaltensweisen) und den Kompetenzerwerb bedeutsam sind (siehe Kapitel 4 in diesem Band). Belegt ist, dass Schülerinnen und Schüler mit Migrationshintergrund, deren Eltern mit ihnen zu Hause Deutsch sprechen, bessere Leistungen aufweisen als jene, deren Eltern nicht mit ihren Kindern zu Hause Deutsch sprechen (z.B. Schwippert et al., 2012). Auch spielt die Art, wie eine Sprache oder gar mehrere Sprachen in der Familie erworben und gefördert werden, eine Rolle für den Kompetenzerwerb (Kempert et al., 2016). Während sich Kinder mit und ohne Migrationshintergrund von ihren Mitschülerinnen und Mitschülern in den meisten Studien eher nicht oder nicht bedeutsam hinsichtlich ihrer Lesemotivation (Kigel, McElvany & Becker, 2015; Miyamoto, Pfof & Artelt, 2017) unterscheiden, lassen sich kleine Effekte hinsichtlich etwas geringer ausgeprägter Leseselbstkonzepte zuungunsten von Kindern mit Migrationshintergrund finden (Mielke, Goy & Pietsch, 2006).

Laut Miyamoto et al. (2017) gibt es Hinweise darauf, dass Kinder mit Migrationshintergrund seltener lesen als ihre Mitschülerinnen und Mitschüler ohne Migrationshintergrund. Zudem kann vermutet werden, dass die Sprache und Verfügbarkeit des genutzten Lesematerials entscheidend dafür sind, inwieweit sich eine positive Lesemotivation förderlich auf die Lesekompetenzentwicklung auswirkt.

Aus *institutioneller Perspektive* lässt sich betrachten, inwieweit sich Merkmale des Bildungssystems oder der Schule auf den Kompetenzerwerb auswirken. Im Fokus stehen hier neben der Qualität der pädagogischen Arbeit und spezifischen Unterstützungsangeboten für Schülerinnen und Schüler mit Migrationshintergrund auch Fragen der Bildungsentscheidungen in Kombination mit verschiedenen Bildungsabschlussmöglichkeiten, der Durchlässigkeit des Schulsystems sowie zentralen Kontextfaktoren des Schulbesuchs (siehe Kapitel 11 in diesem Band; vgl. auch Dronkers, Levels & de Heus, 2014; Morris-Lange, Wendt & Wohlfarth, 2013; Rjosk et al., 2014; Siewert, 2013; Stanat, Schwippert & Gröhlich, 2010).

IGLU erlaubt es, auf repräsentativer Datenbasis Leistungsdisparitäten in ihren Größenordnungen sowohl im internationalen Vergleich als auch im nationalen Trendvergleich von 2001 über 2006 und 2011 bis 2016 einzuordnen. Damit können Veränderungen in den Entwicklungen von unterschiedlichen Leistungsständen von Kindern aus Familien mit und ohne Migrationshintergrund in Deutschland über einen Zeitraum von 15 Jahren untersucht werden. Im aktuellen IGLU-Erhebungszyklus ist dabei insbesondere interessant, inwieweit sich Ergebnisse, die in IGLU 2011 auf eine Verringerung von Leistungsunterschieden zwischen Kindern mit und ohne Migrationshintergrund hindeuteten (Schwippert et al., 2012), im Trend fortgesetzt haben.

Als Schülerinnen und Schüler mit Migrationshintergrund klassifizierte Kinder sind sehr unterschiedlich. Zudem verändern weltpolitische Ereignisse sowie politische Entscheidungen stetig die Gruppe derer, die als Schülerinnen und Schüler mit Migrationshintergrund betrachtet werden könnten. In IGLU wurde auf internationaler Ebene zwischen den Teilnehmerländern über viele Jahre um die Auswahl geeigneter Indikatoren gerungen. So wurde der Sprachgebrauch innerhalb der Familie als Indikator kontinuierlich erhoben, allerdings in unterschiedlicher Operationalisierung (Wendt, Bos, Tarelli, Vaskova & Walzebug, 2016). Das bedeutet, dass sich migrationsbezogene Leistungsdisparitäten unter Nutzung dieses Indikators zwar für den jeweiligen Erhebungszyklus international vergleichend einordnen lassen, dass sich Entwicklungen in den Disparitäten über verschiedene Erhebungszyklen hinweg allerdings nicht im internationalen Vergleich betrachten lassen. Für Deutschland aber ist, aufgrund von ergänzenden Zusatzerhebungen, eine differenzierte Betrachtung von migrationsbezogenen Leistungsdisparitäten auch seit 2001 möglich. Für diese Analysen werden Schülerinnen und Schüler dann als Kinder mit Migrationshintergrund beschrieben, wenn ein Elternteil oder beide Elternteile außerhalb Deutschlands geboren wurden. Aufgrund des Erhebungsdesigns ist eine darüber hinaus gehende Differenzierung nach Herkunftsgruppen oder Generationen jedoch nicht möglich. Ebenso lassen sich mit IGLU keine verlässlichen Aussagen zur Integration von Kindern treffen, die im Zuge der seit 2015 zu beobachtenden angestiegenen Zuwanderung von Geflüchteten und anderen Personengruppen nach Deutschland gekommen sind. Dies hat zwei Gründe: Erstens ist das Stichprobendesign der Studie hierfür nicht ausgelegt (vgl. Wendt, Schwippert & Stubbe, 2016), zweitens gelten für die Teilnahme an IGLU sogenannte Ausschlusskriterien (siehe Kapitel 2 in diesem Band), die regeln, dass nur diejenigen Schülerinnen und Schüler an der Testung teilnehmen, die mindestens seit einem Jahr in der Testsprache des teilnehmenden Landes unterrichtet werden. Hiervon waren zum Zeitpunkt der Testung von IGLU 2016 im Mai des letzten Jahres die meisten neu zugewanderten Kinder betroffen (2.8% der Schülerinnen und Schüler in der Stichprobe). Zum anderen liegen die jüngeren größeren fluchtbedingten

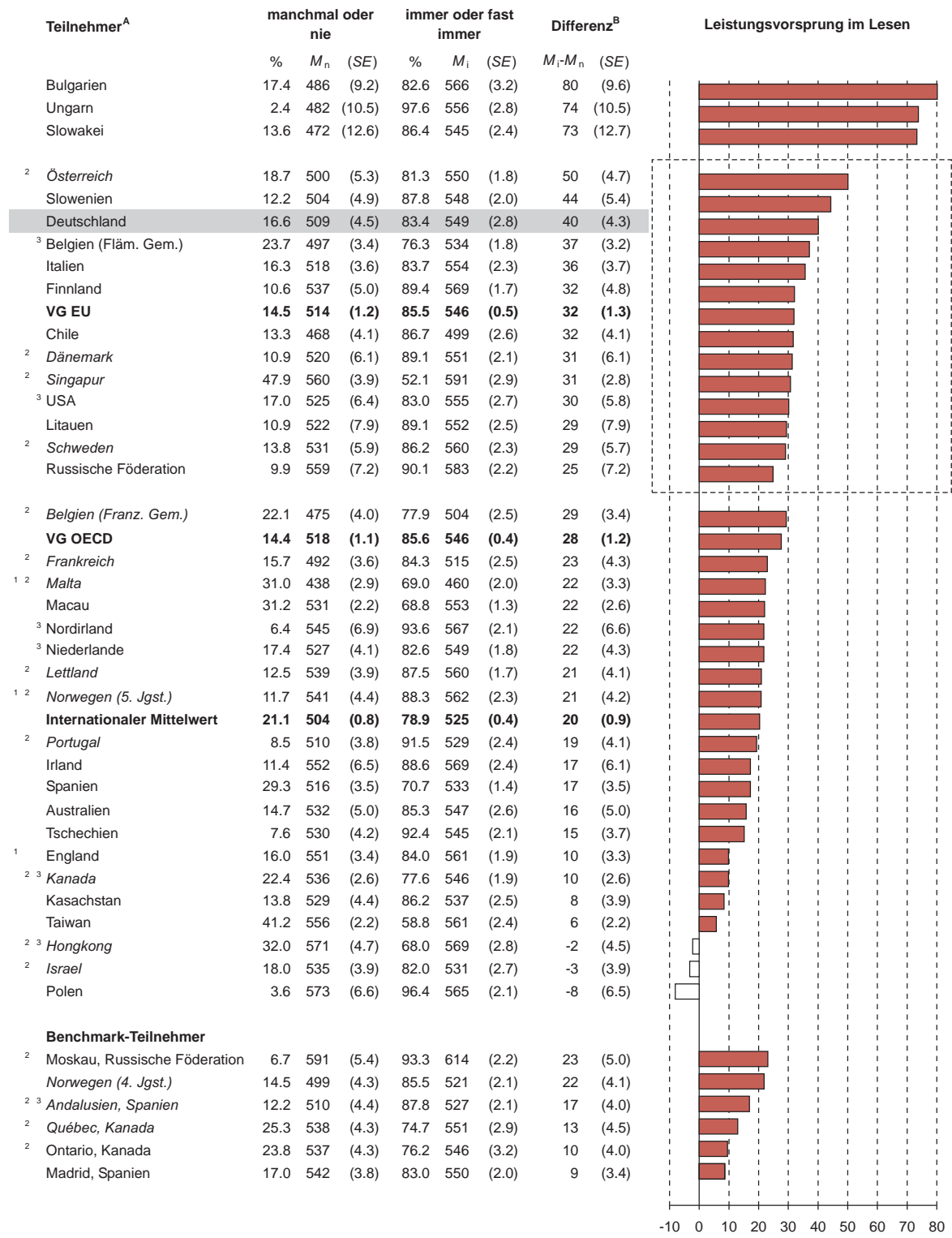
Zuwanderungen nach Deutschland zeitlich nicht weit genug vor der Erhebung von IGLU im Mai/Juni 2016, so dass nur sehr wenige Kinder von neu zugewanderten Geflüchteten zum Testzeitpunkt bereits als anerkannte Geflüchtete die vierte Jahrgangsstufe einer Grundschule in Deutschland besuchten. Diese erst kürzlich in Deutschland zugewanderten Schulkinder machen etwa 3.4 Prozent aller Viertklässlerinnen und Viertklässler in Deutschland aus. Insofern sie sich am Test beteiligt haben, zählen sie damit zu den Schülerinnen und Schülern mit Migrationshintergrund, die Ergebnisse für diese Gruppe werden aber nicht gesondert ausgewiesen.

In diesem Kapitel werden Analysen zu den Lesekompetenzen von Schülerinnen und Schülern der vierten Jahrgangsstufe aus Familien mit und ohne Migrationshintergrund vorgestellt. Im ersten Teil des Kapitels werden zunächst durch familiären Sprachgebrauch bedingte Leistungsdisparitäten von Kindern mit und ohne Migrationshintergrund international verglichen. Für eine Einordnung, inwieweit eine Operationalisierung die Bevölkerungssituationen aller Teilnehmerstaaten und -regionen angemessen repräsentiert, wird auf die IGLU/PIRLS-Enzyklopädie verwiesen (Mullis, Martin, Goh & Prendergast, 2017). Für den internationalen Vergleich werden wie im gesamten Berichtsband IGLU-Teilnehmerstaaten berücksichtigt, die Mitglieder der *Europäischen Union* (EU) sind, die der *Organisation for Economic Co-operation and Development* (OECD) angehören, deren Leistungswerte auf der Gesamtskala Lesen signifikant über dem für Deutschland berichteten Mittelwert liegen oder sich von diesem nicht signifikant unterscheiden. Im Anschluss an die Analyse im internationalen Vergleich werden die Befunde um eine Betrachtung von Unterschieden auf nationaler Ebene zwischen den Erhebungszyklen von 2001, 2006, 2011 und 2016 ergänzt. Zudem werden Unterschiede in verschiedenen motivationalen und verhaltensbezogenen Merkmalen betrachtet.

2 Lesekompetenzen von Schülerinnen und Schülern im internationalen Vergleich nach Sprachgebrauch in der Familie

Die Beherrschung der Amts- oder Verkehrssprache eines Staates gilt als eine wichtige Voraussetzung für den Integrationserfolg von zugewanderten Menschen. Da die Unterrichtssprache hierzulande in der Regel Deutsch ist, sind Kenntnisse der deutschen Sprache für erfolgreiche Bildungskarrieren zentral (Gogolin, 1994; Gogolin, Neumann & Roth, 2003). In Abbildung 7.1 sind die Leistungsunterschiede in der Lesekompetenz zwischen Schülerinnen und Schülern in Abhängigkeit vom familialen Sprachgebrauch im internationalen Vergleich dargestellt. Die Leistungsunterschiede sind im graphischen Teil der Abbildung illustriert, wobei signifikante Unterschiede durch die Färbung der Differenzbalken kenntlich gemacht sind. Im tabellarischen Teil der Abbildung sind neben den Leistungswerten auch die Anteile der Kinder in zwei Gruppen gegenübergestellt: Schülerinnen und Schüler, die die Testsprache zu Hause nie oder manchmal beziehungsweise immer oder fast immer sprechen. Für Deutschland zeigt sich, dass 83.4 Prozent aller Schülerinnen und Schüler angeben, zu Hause immer oder fast immer Deutsch zu sprechen. Im europäischen Vergleich liegt Deutschland mit einem Anteil von 16.6 Prozent aller Schülerinnen und Schüler, bei denen die Testsprache nur manchmal oder nie in der Familie gesprochen

Abbildung 7.1: Lesekompetenzen von Schülerinnen und Schülern nach zu Hause gesprochener Sprache im internationalen Vergleich



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland ($p > .05$).

Statistisch signifikante Unterschiede ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

A = Da in Neuseeland die Testsprache nicht erhoben wurde, können hier für diesen Teilnehmer keine Ergebnisse berichtet werden.

B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

wird, im Mittelfeld. In vielen europäischen Ländern gilt also, dass für ein Viertel bis ein Sechstel aller Grundschul Kinder Mehrsprachigkeit zur Alltagsrealität gehört.

In Deutschland erzielen Viertklässlerinnen und Viertklässler, die angeben, immer oder fast immer zu Hause Deutsch zu sprechen, einen Leistungsmittelwert von 549 Punkten. Ihre Mitschülerinnen und Mitschüler, die angeben, nie oder nur manchmal zu Hause Deutsch zu sprechen, erreichen lediglich 509 Punkte. Die Leistungsdifferenz in Deutschland beträgt 40 Punkte und somit fast eine halbe Standardabweichung vom Skalenmittelwert. Dieser Unterschied entspricht in etwa dem Leistungszuwachs von einem Schuljahr (Hornberg et al., 2007). In drei Teilnehmerstaaten (Bulgarien, Ungarn und der Slowakei) fällt der Leistungsunterschied signifikant größer aus als in Deutschland. Es zeigt sich aber auch, dass in 20 Staaten (ohne *Benchmark*-Teilnehmer), darunter beispielsweise auch Frankreich oder die Niederlande, sowie in der VG OECD die sprachbedingte Leistungsdisparität signifikant niedriger ausfällt als in Deutschland. Auffällig sind in Polen und Ungarn die besonders kleinen Gruppen von Schülerinnen und Schülern, die manchmal oder nie die Testsprache zu Hause sprechen. Die für diese Staaten berichteten Differenzen sollten daher als Phänomen gewürdigt, aber hinsichtlich ihrer numerischen Ausprägung nicht überinterpretiert werden.

3 Lesekompetenzen von Schülerinnen und Schülern in Deutschland im Trend

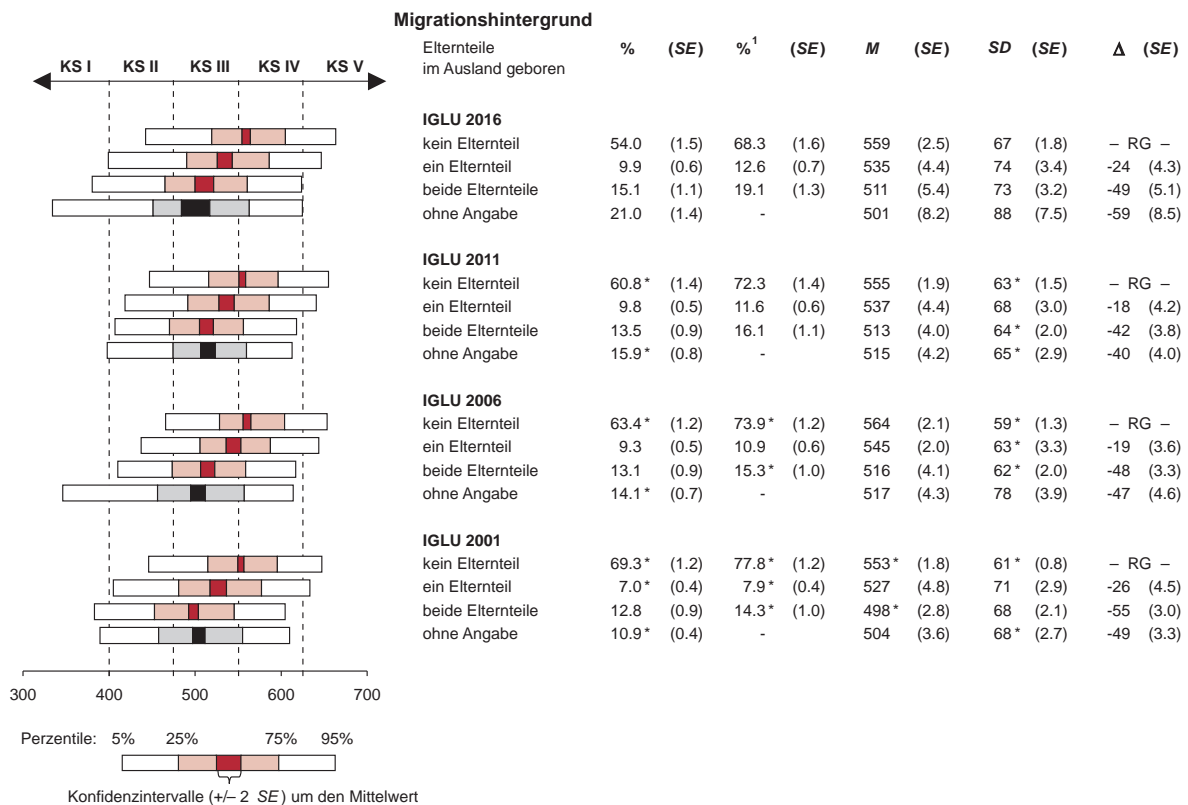
Mit Blick auf eine Unterscheidung von Schülerinnen und Schülern mit und ohne Migrationshintergrund erscheint international die Erfassung des familiären Sprachgebrauchs des Kindes als Minimalinformation akzeptabel (Martin, Mullis & Kennedy, 2003, 2007). Um vertiefende Trendanalysen für Deutschland zu ermöglichen, wurde der in Deutschland eingesetzte Schülerfragebogen in Anlehnung an den vorangegangenen Erhebungszyklus um Fragen zum Geburtsland der Eltern ergänzt. Damit können folgende Gruppen von Schülerinnen und Schülern unterschieden werden:

- die Eltern sind beide in Deutschland geboren;
- ein Elternteil ist in Deutschland und ein Elternteil im Ausland geboren;
- die Eltern sind beide im Ausland geboren.

In Abbildung 7.2 sind die Lesekompetenzen von Schülerinnen und Schülern auf der Gesamtskala aus Familien mit unterschiedlichem Migrationshintergrund in Deutschland für 2001, 2006, 2011 und 2016 dargestellt. Auch werden die Veränderungen in den prozentualen Anteilen von Schülerinnen und Schülern nach dem Geburtsland der Eltern berichtet. Der Anteil von Kindern mit fehlenden Elternangaben und damit unklarem Migrationsstatus hat sich seit 2001 fast verdoppelt und beträgt 2016 21 Prozent. Es ist daher schwierig, verlässliche Daten über den Anteil von Kindern mit Migrationshintergrund in der IGLU-Stichprobe zu gewinnen.

Da sich die prozentualen Verteilungen und Zusammensetzungen von Schülerinnen und Schülern, deren Eltern keine Angaben gemacht haben, zwischen den Erhebungszyklen deutlich unterscheiden, werden in dieser Abbildung zusätzlich Werte für diese Schülergruppe berichtet.

Abbildung 7.2: Lesekompetenzen von Schülerinnen und Schülern in Deutschland nach im Ausland geborenen Elternteilen – IGLU 2001, 2006, 2011 und 2016



1 = Unter Ausschluss der Fälle ‚ohne Angabe‘.
 Δ = Differenz zum Leistungsmittelwert der Schülerinnen und Schüler ohne Migrationshintergrund.
 * = Unterschied zu 2016 statistisch signifikant (p < .05).
 RG = Referenzgruppe.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Laut Angaben der Eltern, die den Fragebogen ausgefüllt haben, liegt der Anteil von Schülerinnen und Schülern mit Migrationshintergrund (mindestens ein Elternteil im Ausland geboren) bei rund 32 Prozent. Im Vergleich zu IGLU 2001 und 2006 sind 2016 signifikant höhere Anteile zu beobachten. Entsprechend ist von 2001 zu 2016 der Anteil an Kindern, deren Eltern beide in Deutschland geboren wurden, von 77.8 Prozent auf 68.3 Prozent signifikant zurückgegangen.

Im Vergleich zu 2001 lassen sich 2016 sowohl für Kinder, deren Eltern beide in Deutschland geboren wurden, als auch für ihre Mitschülerinnen und Mitschüler, deren Eltern beide im Ausland geboren wurden, signifikant höhere Leseleistungen beobachten. Auch für die Gruppe der Viertklässlerinnen und Viertklässler mit einem im Ausland geborenen Elternteil lässt sich ein nomineller Leistungszuwachs beobachten, der allerdings nicht signifikant ist. Dass dieser Unterschied nicht zufallskritisch abgesichert werden kann, liegt an der relativ kleinen Gruppengröße insbesondere in 2001. Für die Gruppe der Schülerinnen und Schüler ohne Elternangaben zum Migrationsstatus ist eine geringe Abnahme des mittleren Leistungswerts zu beobachten, die allerdings nicht signifikant ausfällt.

Für 2016 ist auffallend, dass bei allen Schülergruppen im Vergleich zu den vorangegangenen Erhebungszyklen eine größere Standardabweichung besteht. Für die Kinder, deren Eltern beide in Deutschland geboren wurden, ist die Zunahme der Heterogenität im Vergleich von 2016 zu 2001, 2006 und 2011 signifikant. Für Schülerinnen und Schüler, deren Elternteile beide im Ausland geboren wurden, fällt dieser Vergleich von 2016 zu 2006 und 2011 ebenfalls signifikant größer aus.

Die Leistungsdisparitäten zwischen Kindern mit und ohne Migrationshintergrund erweisen sich in der Perspektive der 15 Jahre von 2001 zu 2016 als weitestgehend konstant. Schülerinnen und Schüler mit Migrationshintergrund erzielen 2016 im Vergleich zu ihren Mitschülerinnen und Mitschülern, deren Eltern beide in Deutschland geboren wurden, Leseleistungen, die mit 24 (ein Elternteil im Ausland geboren) beziehungsweise 49 (beide Elternteile) Leistungspunkten geringer ausfallen und damit im Rahmen dessen liegen, was Kinder in einem halben beziehungsweise ganzen Schuljahr dazulernen.

Der von 2001 zu 2011 zu konstatierende Befund einer signifikanten Reduktion von migrationsbezogenen Leistungsdisparitäten ist im Vergleich von 2001 zu 2016 nicht mehr festzustellen, da sich nun nicht mehr nur ein Leistungszuwachs für die Kinder mit Migrationshintergrund, sondern auch für Kinder, deren Eltern in Deutschland geboren wurden, bei gleichzeitiger nomineller Zunahme der Heterogenität innerhalb der Gruppen, beobachten lässt.

Tabelle 7.1 zeigt, wie sich die Leistungstrends von Viertklässlerinnen und Viertklässlern in Deutschland unter Kontrolle des Migrationsstatus sowie weiterer Schülermerkmale darstellen. Die links in der Tabelle dargestellte lineare Regression dient als Referenzmodell und zeigt Koeffizienten, die sich ohne Berücksichtigung der Mehrebenenstruktur der Daten ergeben. In dieser Modellierung zeigt sich, dass in Deutschland generell die Lesekompetenz von Viertklässlerinnen und Viertklässlern unter Kontrolle des Migrationshintergrundes in der aktuellen Erhebung 7 Punkte höher lag als im Jahr 2001 und 5 Punkte niedriger als im Jahr 2006. Zudem zeigt sich ein deutlicher Effekt des Migrationshintergrundes. Über alle Erhebungszeitpunkte hinweg fallen die Leseleistungen von Kindern aus Familien mit Migrationshintergrund (beide Elternteile) 49 Punkte geringer aus als die Leseleistungen ihrer Mitschülerinnen und Mitschüler ohne Migrationshintergrund. Der Leistungsrückstand von Kindern mit einem im Ausland geborenen Elternteil liegt bei 22 Punkten. Insgesamt erklärt dieses Modell 8.2 Prozent der Gesamtvarianz.

In den Modellen 1 bis 4 werden die Ergebnisse von Mehrebenenregressionen berichtet (siehe Kapitel 2, Abschnitt 11 in diesem Band), somit sind die Parameter als durchschnittliche Klasseneffekte zu interpretieren. Allgemeine Analysen zur Mehrebenenstruktur in IGLU 2016 zeigen, dass sich 70.3 Prozent der Varianz der Lesekompetenz über alle Erhebungszeitpunkte hinweg durch Merkmale auf der Individualebene (Level 1, *within*) erklären lassen und entsprechend 29.7 Prozent durch Merkmale auf der Klassenebene (Level 2, *between*). In den in diesem Kapitel dargestellten Mehrebenenregressionen werden zur Erklärung der Lesekompetenz neben der Berücksichtigung des Erhebungszyklus ausschließlich Variablen auf Individualebene aufgenommen, so dass in den Modellen 1 bis 4 angegeben wird, welcher Prozentsatz des Varianzanteils auf Individualebene jeweils durch das Modell erklärt wird. Hinsichtlich der berücksichtigten Variablen unterscheidet sich Modell 1 nicht vom Referenzmodell. Unter Berücksichtigung der Mehrebenenstruktur fallen die allgemeinen Unterschiede zwischen den Erhebungszyklen nicht signifikant aus. Die Leistungsunterschiede zuungunsten der Kinder aus Familien mit Zuwanderungsgeschichte sind in diesen Modellen jedoch ebenfalls signifikant.

Für Kinder mit einem im Ausland geborenen Elternteil reduziert sich der Leistungsunterschied zu Kindern ohne Migrationshintergrund nur geringfügig. Für Kinder, deren Eltern beide im Ausland geboren wurden, reduziert sich der Leistungsunterschied von 49 auf rund 41 Punkte. Dieser Befund lässt sich als Hinweis darauf interpretieren, dass ein Teil der migrationsbezogenen

Tabelle 7.1: Lesekompetenzen zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle von Indikatoren zu Migration und sozialer Herkunft (Mehrebenenmodell)

	Lineare Regression		Mehrebenenregressionen			
	b	(SE)	Modell 1 b (SE)	Modell 2 b (SE)	Modell 3 b (SE)	Modell 4 b (SE)
Konstante	559.2	(2.5)**	554.2 (4.3)**	547.5 (5.1)**	551.8 (4.9)**	550.8 (5.7)**
Testzeitpunkt						
IGLU 2001	-7.1	(3.0)**	-4.1 (4.9) ^{ns}	-2.3 (5.3) ^{ns}	-0.6 (4.7) ^{ns}	1.8 (5.9) ^{ns}
IGLU 2006	5.5	(3.3)*	5.6 (4.9) ^{ns}	6.6 (5.0) ^{ns}	4.9 (4.4) ^{ns}	7.5 (5.7) ^{ns}
IGLU 2011	-2.6	(3.4) ^{ns}	-1.4 (5.2) ^{ns}	-1.1 (5.3) ^{ns}	-4.1 (4.9) ^{ns}	-5.9 (5.9) ^{ns}
Merkmale						
Geschlecht ^A					-5.7 (2.4)**	-5.7 (2.4)**
Familie mit hohem sozioökonomischen Status ^B				31.7 (2.7)**	22.4 (2.5)**	26.9 (6.5)**
Migrationshintergrund (ein Elternteil) ^C	-22.0	(1.8)**	-19.6 (3.7)**	-16.9 (4.1)**	-12.0 (3.7)**	-13.7 (9.2) ^{ns}
Migrationshintergrund (beide Elternteile) ^D	-49.3	(1.7)**	-40.7 (3.6)**	-32.6 (4.1)**	-27.2 (3.8)**	-27.6 (9.1)**
Interaktionen						
Familie mit hohem sozioökonomischen Status ^B * IGLU 2001						-6.0 (7.8) ^{ns}
Familie mit hohem sozioökonomischen Status ^B * IGLU 2006						-9.8 (8.3) ^{ns}
Familie mit hohem sozioökonomischen Status ^B * IGLU 2011						0.4 (7.4) ^{ns}
Migrationshintergrund (ein Elternteil) ^C * IGLU 2001						-0.1 (12.2) ^{ns}
Migrationshintergrund (beide Elternteile) ^D * IGLU 2001						3.1 (11.4) ^{ns}
Migrationshintergrund (ein Elternteil) ^C * IGLU 2006						3.2 (12.8) ^{ns}
Migrationshintergrund (beide Elternteile) ^D * IGLU 2006						-6.7 (11.3) ^{ns}
Migrationshintergrund (ein Elternteil) ^C * IGLU 2011						0.2 (11.8) ^{ns}
Migrationshintergrund (beide Elternteile) ^D * IGLU 2011						8.9 (11.8) ^{ns}
Kovariaten						
Kognitive Fähigkeiten ^E					17.7 (1.0)**	17.7 (0.9)**
Erklärter Anteil der Gesamtvarianz	8.2					
Erklärter Anteil auf Individualebene (70.3%)			18.9	30.2	44.0	44.1

b = Regressionsgewichte (unstandardisiert).

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

A = Geschlecht (0 = Mädchen; 1 = Jungen), siehe Kapitel 5 in diesem Band.

B = Sozioökonomischer Status nach Angabe der Eltern (0 = Manuell Tätige, Angestellte und kleinere Unternehmer; 1 = Akademiker, Techniker und Führungskräfte), siehe Kapitel 6 in diesem Band.

C = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren; 1 = Ein Elternteil im Ausland geboren).

D = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren; 1 = Beide Elternteile im Ausland geboren).

E = Skalierter Index der kognitiven Fähigkeiten (Messinvarianz über die Zyklen angenommen), siehe Kapitel 2 in diesem Band.

Das Gesamtmodell zu dem hier dargestellten Ausschnitt ist im Anhang B dargestellt, siehe Modelle 10,11,12 und 13. Die lineare Regression ist nicht Teil des Gesamtmodells.

Leistungsunterschiede womöglich über ungünstige segregationsbedingte Merkmale der Klassenkomposition vermittelt sein könnte (siehe hierzu Kapitel 11 in diesem Band).

In Modell 2 wird ein Indikator für den sozialen Status aufgenommen. In Anlehnung an das Vorgehen aus Kapitel 6 in diesem Band wird hier der Berufsstatus in dichotomisierter Form herangezogen, wobei die ersten beiden Berufsgruppen (manuelle Tätigkeiten sowie Angestellte und kleinere Unternehmer) zusammengefasst und der dritten Berufsgruppe (Akademiker, Techniker und Führungskräfte) gegenübergestellt werden. Der Einfluss des Migrationshintergrundes fällt unter Berücksichtigung dieses Indikators der sozialen Herkunft etwas geringer aus, bleibt jedoch deutlich signifikant. Dies kann als Hinweis auf den bekannten Befund gedeutet werden, wonach sich ein Teil der migrationsbezogenen Leistungsdisparitäten auf Unterschiede in den sozioökonomischen Lebensverhältnissen von Familien mit unterschiedlichem Migrationsstatus zurückführen lässt (siehe hierzu Kapitel 6, Abbildung 6.7 in diesem Band).

Im Modell 3 werden weitere Variablen aufgenommen, deren Zusammenhang mit schulischen Kompetenzen gut dokumentiert ist. Entsprechend erklärt dieses Modell mit 44 Prozent fast die Hälfte der auf Individualebene vorhandenen Varianz in der Lesekompetenz. Als Kovariaten werden neben dem Geschlecht (siehe Kapitel 5 in diesem Band) auch kognitive Grundfähigkeiten der Viertklässlerinnen und Viertklässler berücksichtigt (siehe Kapitel 2 in diesem Band). Eine Zunahme des Wertes auf dieser Skala um eine Standardabweichung geht mit einer um 18 Punkte höheren Lesekompetenz einher. Der Einfluss des Migrationshintergrundes fällt bei zusätzlicher Berücksichtigung dieser beiden Indikatoren etwas geringer aus, bleibt jedoch weiterhin deutlich signifikant. Das Modell 4 unterscheidet sich von dem vorangegangenen Modell dadurch, dass auch die Interaktion zwischen Erhebungszyklus und Migrationsstatus kontrolliert wird. Dies erlaubt es, mögliche Veränderungen der migrationsbezogenen Disparitäten im Zeitverlauf aufzuzeigen, wobei die Interaktionsterme nicht signifikant sind. In diesem Modell lassen sich bedingt durch die großen Standardfehler keine signifikanten Unterschiede in der Leseleistung zwischen Kindern ohne Migrationshintergrund und denen mit einem im Ausland geborenen Elternteil nachweisen. Die deutlich geringeren Leistungsdisparitäten zuungunsten von Viertklässlerinnen und Viertklässlern, deren Elternteile beide im Ausland geboren wurden, bleiben jedoch bestehen. Die in Abbildung 7.2 berichteten positiven Leistungstrends von Kindern mit und ohne Migrationshintergrund von 2001 zu 2016 werden durch die mehrebenenanalytische Modellierung sowie multivariate Betrachtung unter Berücksichtigung von Veränderungen in der sozialen Zusammensetzung von Schülerschaften zwischen den Erhebungszyklen relativiert.

4 Leseleistungen auf unterschiedlichen Kompetenzstufen im Trend

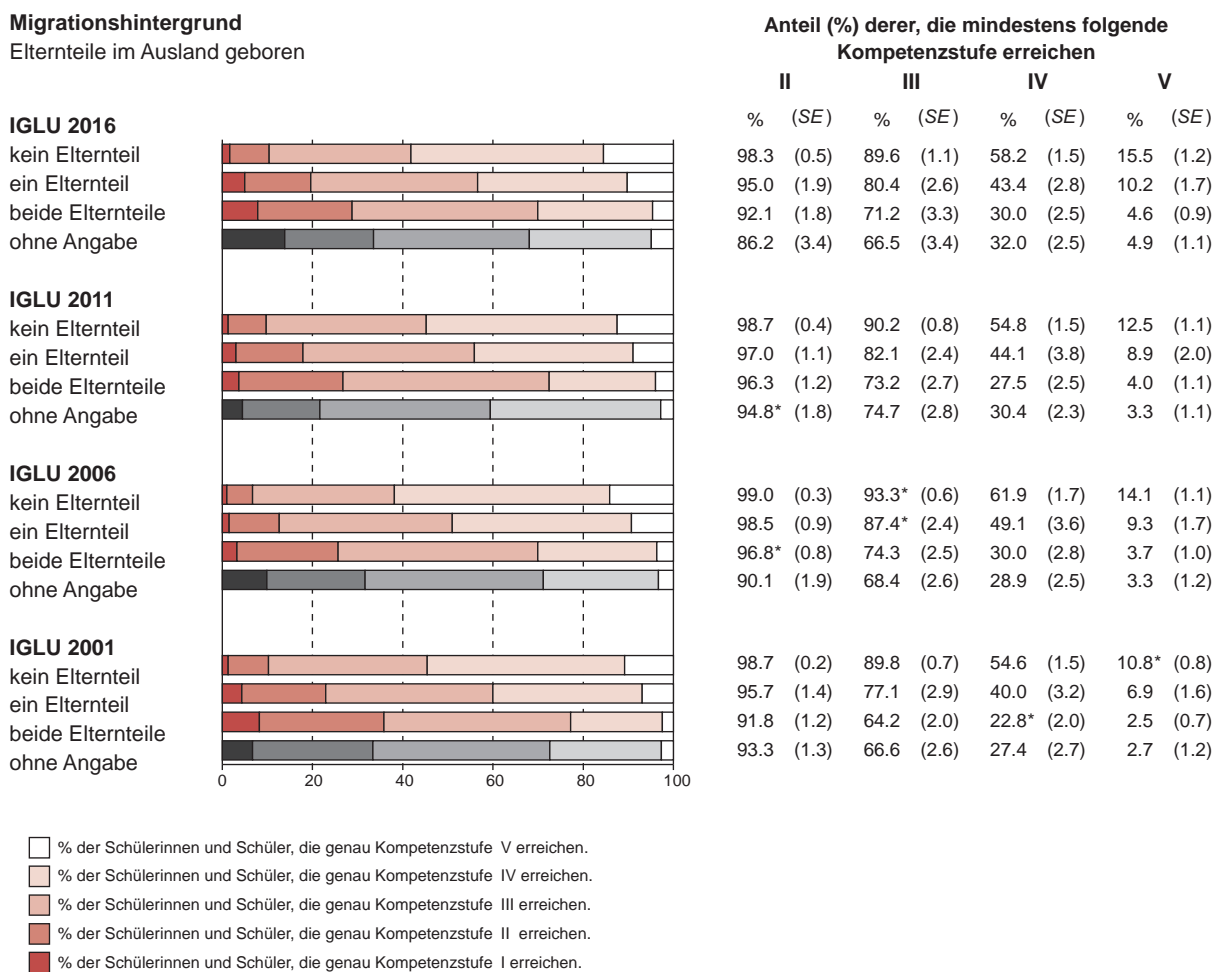
In IGLU werden fünf Stufen der Lesekompetenz unterschieden, die in Kapitel 3 beschrieben sind. Mit Hilfe der Kompetenzstufen können die Leistungen der Schülerinnen und Schüler inhaltlich beschrieben werden. Zudem lässt sich analysieren, wie viele Schülerinnen und Schüler eines jeden Staates oder einer Region unterschiedlich hohe Leistungsanforderungen bewältigen, wie groß

also zum Beispiel die Gruppe von Schülerinnen und Schülern ist, die eine hohe Lesekompetenz aufweisen (Kompetenzstufen IV und V), und wie groß die Gruppe derer ist, die nur über eine niedrige Lesekompetenz verfügen, die es ihnen lediglich in Ansätzen erlaubt, dieses Wissen produktiv einzusetzen (Kompetenzstufen I und II). Bildungspolitisches Ziel ist es, möglichst allen Schülerinnen und Schülern diese grundlegenden Kompetenzen auf dem Niveau der Kompetenzstufe III zu vermitteln und gleichzeitig möglichst viele Schülerinnen und Schüler mit höheren Kompetenzen auszustatten.

Abbildung 7.3 ist die prozentuale Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen zu entnehmen. Für alle Erhebungen zeigt sich, dass sich Kinder aus Familien ohne Migrationshintergrund systematisch häufiger auf den höheren Lesekompetenzstufen verorten lassen. Kinder aus Familien mit Migrationshintergrund hingegen zeigen häufiger Leistungen auf rudimentärem beziehungsweise niedrigerem Niveau.

In IGLU 2016 erreichen lediglich 4.6 Prozent der Schülerinnen und Schüler, deren Eltern beide im Ausland geboren wurden, die Kompetenzstufe V, während es 15.5 Prozent der Schülerinnen und Schüler sind, deren Eltern beide in Deutschland geboren wurden.

Abbildung 7.3: Anteile von Kindern auf den unterschiedlichen Kompetenzstufen in Abhängigkeit von im Ausland geborenen Elternteilen – IGLU 2001, 2006, 2011 und 2016 im Vergleich (Angaben in Prozent)



- % der Schülerinnen und Schüler, die genau Kompetenzstufe V erreichen.
- % der Schülerinnen und Schüler, die genau Kompetenzstufe IV erreichen.
- % der Schülerinnen und Schüler, die genau Kompetenzstufe III erreichen.
- % der Schülerinnen und Schüler, die genau Kompetenzstufe II erreichen.
- % der Schülerinnen und Schüler, die genau Kompetenzstufe I erreichen.

* = Unterschied zu 2016 statistisch signifikant ($p < .05$).

Während mit 10.4 Prozent nur jede zehnte Schülerin beziehungsweise jeder zehnte Schüler ohne Migrationshintergrund Leseleistungen auf einem Niveau unter der Kompetenzstufe III zeigt, sind es bei Kindern, deren Eltern beide im Ausland geboren sind, 28.8 Prozent, also mehr als jedes vierte Kind dieser Gruppe.

Im Vergleich von 2001 und 2016 ist für alle drei Gruppen ein signifikant gesteigener Anteil an starken Leserinnen und Lesern festzustellen. Für Kinder mit Migrationshintergrund (beide Elternteile im Ausland geboren) ist zudem ein gesunkener Anteil von schwachen Leserinnen und Lesern zu beobachten: Zeigten 2001 noch rund 36 Prozent von ihnen Leseleistungen auf dem Niveau der Kompetenzstufen I und II, sind es 2016 mit 29 Prozent signifikant weniger Schülerinnen und Schüler.

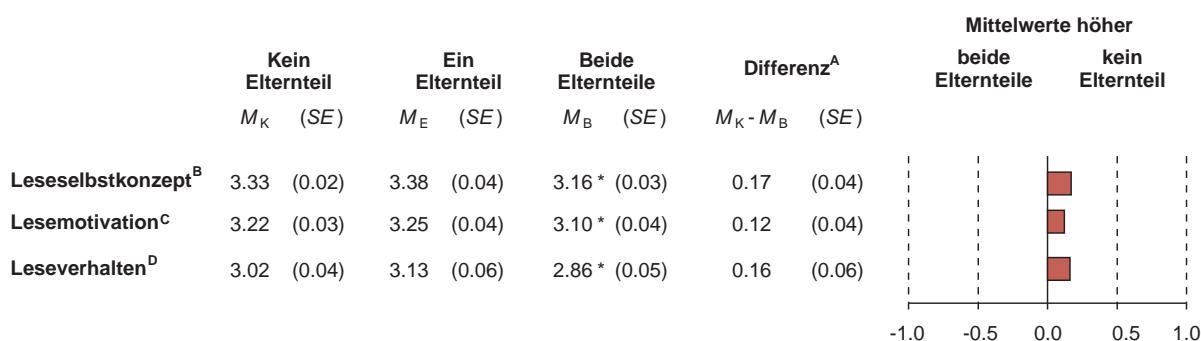
5 Leseselbstkonzept, Lesemotivation und Leseverhalten

Neben den in der IGLU-Konzeption beschriebenen Bereichen lesebezogener kognitiver Kompetenzen stellen auch positive Einstellungen gegenüber dem Lesen und ein positives Selbstkonzept eigener Fähigkeiten wichtige Zielbereiche des Unterrichts dar (siehe Kapitel 4 in diesem Band).

In Abbildung 7.4 sind Unterschiede in Leseselbstkonzept, Lesemotivation und Leseverhalten nach Migrationsstatus der Schülerfamilien dargestellt. Im linken Teil der Abbildung sind in tabellarischer Form die Mittelwerte sowie zugehörige Standardfehler zu den drei Konstrukten getrennt nach Migrationsstatus dargestellt. Der Differenzwert für Kinder, deren Eltern beide in Deutschland beziehungsweise beide im Ausland geboren sind, wird zudem im rechten Teil der Abbildung graphisch illustriert. Statistisch signifikante Unterschiede sind rot gefärbt.

In Bezug auf die betrachteten motivationalen und verhaltensbezogenen Maße ist festzuhalten, dass ein Großteil der Viertklässlerinnen und Viertklässler in Deutschland unabhängig von ihrem Migrationsstatus in allen drei betrachteten Indikatoren (Leseselbstkonzept, Lesemotivation und Leseverhalten) hohe positive Selbsteinschätzungen zum Ausdruck bringt. Wie Abbildung 7.4 verdeutlicht, haben Kinder aus Familien ohne Migrationshintergrund positivere Werte in den drei Merkmalen. Sie berichten, dass sie über eine höhere Lesemotivation verfügen (0.12), häufiger in der Freizeit zum Spaß lesen (0.16) und auch selbstbewusster hinsichtlich ihrer Leseselbstkonzepte sind (0.17) als ihre Mitschülerinnen und Mitschüler, deren Eltern beide im Ausland geboren sind. Die Unterschiede zwischen den Gruppen sind zwar statistisch signifikant, dennoch handelt es sich um sehr kleine Effekte.

Eine Betrachtung von Veränderungen zu 2001 ist in IGLU für die Lesemotivation möglich. Wie in Kapitel 4 in diesem Band gezeigt, lässt sich für Viertklässlerinnen und Viertklässler in Deutschland eine geringe Abnahme der Lesemotivation von 2001 zu 2016 beobachten, die vor allem auf eine geringere Motivation der leseschwächeren Schülerinnen und Schüler zurückzuführen ist. Betrachtet man differentielle Veränderungen nach Migrationsstatus, zeigen sich für die Gruppen von Kindern mit einem im Ausland geborenen Elternteil keine signifikanten Veränderungen, während sich für die beiden Gruppen der Schülerinnen und Schüler mit beiden in Deutschland sowie im Ausland geborenen Elternteilen signifikant geringere Motivationen für 2016 feststellen lassen. Mit Differenzen von $d = 0.11$ beziehungsweise $d = 0.16$ sind diese zwar

Abbildung 7.4: Mittelwerte und Differenzen von Leseselbstkonzept, Lesemotivation und Leseverhalten in Abhängigkeit von im Ausland geborenen Elternteilen in Deutschland


■ Statistisch signifikante Unterschiede ($p < .05$).

* = Unterschied zu kein Elternteil statistisch signifikant ($p < .05$).

A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

B = Die Skala umfasst fünf Fragen (z.B. Lesen fällt mir sehr leicht.) mit vierstufigem Antwortformat (1 = Stimme überhaupt nicht zu, ..., 4 = Stimme stark zu).

C = Die Skala umfasst drei Fragen (z.B. Ich lese gerne.) mit vierstufigem Antwortformat (1 = Stimme überhaupt nicht zu, ..., 4 = Stimme stark zu).

D = Die Frage lautet „Ich lese außerhalb der Schule, weil es mir Spaß macht.“ mit vierstufigem Antwortformat (1 = Nie oder fast nie, ..., 4 = Jeden Tag oder fast jeden Tag).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

gering, im Vergleich zu den hier dargestellten migrationsbezogenen Disparitäten jedoch beachtlich. Nominell ist eine marginale Zunahme der migrationsbezogenen Disparitäten ($d = 0.04$) zu beobachten, die allerdings nicht signifikant ausfällt.

6 Zusammenfassung

Die Befunde aus IGLU 2016 bestätigen, dass Schülerinnen und Schüler, die zu Hause nicht die Testsprache sprechen, in fast allen Teilnehmerstaaten und -regionen niedrigere Leseleistungen aufweisen als ihre Mitschülerinnen und Mitschüler, die zu Hause die Testsprache sprechen. Dies ist auch in Deutschland nicht anders. Hier liegt der Unterschied in der Leseleistung bei 40 Punkten und damit in der Größenordnung dessen, was Kinder im Laufe eines Jahres durchschnittlich dazulernen. Deutlich zeigt sich in international vergleichender Perspektive, dass die Leistungsunterschiede erheblich variieren. Einige Teilnehmerstaaten und -regionen weisen im Vergleich zu Deutschland deutlich niedrigere migrationsbezogene Disparitäten in ihrem Bildungssystem auf. Weitergehend zu prüfen wäre, ob dies durch Maßnahmen in den Bildungssystemen oder durch die Zusammensetzung der Migrantengruppen in den jeweiligen Staaten zu erklären ist.

Veränderungen in den Leseleistungen zwischen 2001 und 2016 können in IGLU für Deutschland anhand von Angaben zum Geburtsland der Eltern betrachtet werden. Im Vergleich zu 2001 erreichen in IGLU 2016 sowohl Kinder, deren Eltern beide in Deutschland geboren wurden, als auch ihre Mitschülerinnen und Mitschüler, deren Eltern beide im Ausland geboren wurden, signifikant höhere Leseleistungen. Die Veränderungen in den Leistungsmittelwerten gehen dabei sowohl mit gestiegenen Anteilen an Kindern mit sehr guten Leseleistungen als auch mit reduzierten Anteilen an Kindern mit geringeren Leseleistungen einher. Allerdings zeigt sich auch eine Vergrößerung der Streuung und der Leistungsheterogenität. Zudem ist mit 29 Prozent der Anteil an schwachen

Leserinnen und Lesern aus Familien mit Migrationshintergrund (beide Elternteile), verglichen mit Schülerinnen und Schülern ohne Migrationshintergrund, immer noch eklatant hoch. Nur etwa jedes zehnte der Kinder, deren Eltern in Deutschland geboren wurden, zeigt vergleichbar schwache Leseleistungen auf dem Niveau der Kompetenzstufen I und II.

Die migrationsbezogenen Leistungsdisparitäten sind seit 2001 praktisch unverändert. Ergebnisse aus IGLU 2011, die auf eine Reduktion der Leistungsdisparitäten hindeuteten, lassen sich 2016 nicht nachweisen. Komplexere Trendanalysen, die die Veränderungen in den Leseleistungen von Kindern mit und ohne Migrationshintergrund von 2001 zu 2016 bei mehrbenenanalytischer Modellierung, unter Kontrolle von zentralen Kovariaten sowie unter Berücksichtigung von Veränderungen in der sozialen und migrationsbezogenen Zusammensetzung der Schülerschaften, betrachten, geben keine Hinweise darauf, ob Veränderungen in der Schülerschaft den Erfolg von verbessertem Unterricht, inner- oder außerschulischer Lesefördermaßnahmen überschatten.

Die dargestellten Befunde erlauben es auch für 2016 nicht, ein positives Fazit zu ziehen. Auf Basis der IGLU-Ergebnisse lassen sich keine Hinweise darauf erkennen, dass es dem deutschen Schulsystem, trotz vielfältiger Bemühungen, bisher gelungen ist, dem bildungspolitischen Ziel, der systematischen Reduktion von zuwanderungsbezogenen Disparitäten, näher zu kommen. Allerdings muss hierbei auch berücksichtigt werden, dass sich in Deutschland die Zusammensetzung der Schülerinnen und Schüler aus Familien mit Migrationshintergrund in 15 Jahren verändert hat. Die Förderung von Kindern entsprechend ihren unterschiedlichen individuellen, sprachlichen sowie soziokulturellen Voraussetzungen sollte noch stärker als bisher in den Fokus bildungspolitischer sowie praktischer Bemühungen rücken. Hierbei scheint es sinnvoll, die unterschiedlichen Voraussetzungen von Kindern aus verschiedenen sozialen und kulturellen Milieus auch im zeitlichen Wandel differenzierter in den Blick zu nehmen, als dies bisher getan wurde.

Literatur

- Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann. Zugriff am 03.10.2016 unter <http://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/bildungsbericht-2016>
- Dronkers, J., Levels, M. & Heus, M. de (2014). Migrant pupils' scientific performance: The influence of educational system features of origin and destination countries. *Large-scale Assessments in Education*, 2 (3), 1–28. doi: 10.1186/2196-0739-2-3
- Elley, W. B. (1992). *The IEA Study of reading literacy: Achievement and instruction in thirty-two school systems*. Exeter: Pergamon.
- Gogolin, I. (1994). *Der monolinguale Habitus der multilingualen Schule*. Münster: Waxmann.
- Gogolin, I., Neumann, U. & Roth, H.-J. (2003). *Förderung von Kindern und Jugendlichen mit Migrationshintergrund*. BLK-Gutachten. Bonn: Bundesministerium für Bildung und Forschung.
- 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.
- Hornberg, S., Bos, W., Buddeberg, I., Potthoff, B. & Stubbe, T. C. (2007). Anlage und Durchführung von IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lese-*

- kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 21–46). Münster: Waxmann.
- 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 VS.
- Kigel, R. M., McElvany, N. & Becker, M. (2015). Effects of immigrant background on text comprehension, vocabulary, and reading motivation: A longitudinal study. *Learning and Instruction*, 35, 73–84.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2002). *Bericht „Zuwanderung“. Beschluss der Kultusministerkonferenz vom 24.05.2002 i. d. F. vom 16.11.2006*. Zugriff am 07.11.2017 unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_05_24-Zuwanderung.pdf
- Lehmann, R. H., Peek, R., Pieper, I. & von Stritzky, R. (1995). *Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen*. Weinheim: Beltz.
- Martin, M. O., Mullis, I. V. S. & Kennedy, A. M. (Hrsg.). (2003). *PIRLS 2001 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S. & Kennedy, A. M. (Hrsg.). (2007). *PIRLS 2006 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mielke, R., Goy, M. & Pietsch, M. (2006). Das Leseselbstkonzept am Ende der Grundschulzeit. In W. Bos & M. Pietsch (Hrsg.), *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen* (S. 87–109). Münster: Waxmann.
- Miyamoto, A., Pfost, M. & Artelt, C. (2017). Reciprocal relations between intrinsic reading motivation and reading competence: A comparison between native and immigrant students in Germany. *Journal of Research in Reading*. doi: 10.1111/1467-9817.12113
- Morris-Lange, S., Wendt, H. & Wohlfarth, C. (2013). *Segregation an deutschen Schulen. Ausmaß, Folgen und Handlungsempfehlungen für bessere Bildungschancen*. Berlin: SVR-Forschungsbereich.
- Mullis, I. V. S., Martin, M. O., Goh, S. & Prendergast, C. (Hrsg.). (2017). *PIRLS 2016 Encyclopedia: Education policy and curriculum in reading*. Zugriff am 07.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/encyclopedia/>
- Rjosk, C., Haag, N., Heppt, B. & Stanat, P. (2017). Zuwanderungsbezogene Disparitäten. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 237–275). Münster: Waxmann.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E. & Stanat, P. (2014). Socio-economic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, 32, 63–72.
- Schwippert, K., Bos, W. & Lankes, E.-M. (2003). Heterogenität und Chancengleichheit am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 265–302). Münster: Waxmann.
- Schwippert, K., Bos, W. & Lankes, E.-M. (2004). Heterogenität und Chancengleichheit am Ende der vierten Jahrgangsstufe in den Ländern der Bundesrepublik Deutschland und im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 165–190). Münster: Waxmann.
- Schwippert, K., Hornberg, S., Freiberg, M. & Stubbe, T. C. (2007). Lesekompetenzen von Kindern mit Migrationshintergrund im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 249–269). Münster: Waxmann.

- Schwippert, K., Hornberg, S. & Goy, M. (2008). Lesekompetenzen von Kindern mit Migrationshintergrund im nationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU-E 2006. Die Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 111–125). Münster: Waxmann.
- 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. 191–207). Münster: Waxmann.
- Siewert, J. (2013). *Herkunftsspezifische Unterschiede in der Kompetenzentwicklung: Weil die Schule versagt? Untersuchungen zum Ferieneffekt in Deutschland*. Münster: Waxmann.
- Stanat, P., Schwippert, K. & Gröhlich, C. (2010). Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb: Längsschnittliche Überprüfung eines umstrittenen Effekts. In C. P. Allemann-Ghionda, P. Stanat, K. Göbel & C. Röhner (Hrsg.), *Migration, Sprache, Identität. Zeitschrift für Pädagogik, 55. Beiheft* (S. 147–164). Weinheim: Beltz.
- Wendt, H., Bos, W., Tarelli, I., Vaskova, A. & Walzebug, A. (Hrsg.). (2016). *IGLU & TIMSS 2011. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. 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. Selzer, 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 VIII

Der Übergang von der Primar- in die Sekundarstufe

Tobias C. Stubbe, Wilfried Bos und Michael Schurig

1 Einleitung

Der Übergang von der Grundschule auf eine Schulform der Sekundarstufe I hat für den zukünftigen Bildungserfolg von Schülerinnen und Schülern in Deutschland eine zentrale Bedeutung (Bellenberg & Klemm, 2000). Diese im internationalen Vergleich frühe Verteilung von Schülerinnen und Schülern auf unterschiedliche Bildungsgänge wird von vielen Expertinnen und Experten kritisiert. Zum einen ist eine verlässliche Vorhersage der zukünftigen Leistungsentwicklung von Zehnjährigen problematisch. Zum anderen weisen sowohl die Empfehlungen von Grundschulen als auch die Entscheidungen von Eltern deutliche Zusammenhänge mit dem sozialen Hintergrund und einer möglichen Migrationsgeschichte der Familie auf (im Überblick Maaz & Nagy, 2009).

Von Befürwortern dieser frühen Trennung wird häufig die Möglichkeit der nachträglichen Korrektur der ursprünglichen Schullaufbahnentscheidung im Verlauf der Sekundarstufe angeführt. Schulformwechsel werden dabei als horizontale Durchlässigkeit bezeichnet. Die Möglichkeit, einen höherwertigen Schulabschluss auf einer Schulart zu erwerben, die herkömmlich nicht zu diesem Abschluss führt (z. B. Realschulabschluss an einer Hauptschule), wird vertikale Durchlässigkeit genannt. Empirische Daten zeigen, dass Schulformwechsel vor allem in Form von Abstiegen vorkommen und dass diese darüber hinaus – wie auch die ursprünglichen Schullaufbahnentscheidungen – vom sozialen Hintergrund der Eltern abhängen (im Überblick Stubbe, 2009). Die vertikale Durchlässigkeit des deutschen Sekundarschulsystems wird zwar in den vergangenen Jahren verstärkt genutzt, es muss aber bezweifelt werden, ob Abschlüsse, die an Schulen erworben werden, die eigentlich nicht zu diesem Abschluss führen, auf dem Arbeitsmarkt denselben Wert besitzen wie auf der dafür vorgesehenen Schulform erworbene Abschlüsse (Schuchart, 2007, 2013).

Traditionell werden in der Bundesrepublik Deutschland in der Sekundarstufe I drei Schulformen unterschieden: Hauptschule, Realschule und Gymnasium. In den 1970er Jahren wurden in einigen Ländern zudem Gesamtschulen eingeführt, die diese drei Schulformen in einer integrieren (Köller, 2008). Dieses Schulsystem erfuhr in den vergangenen Jahren in vielen Ländern zahlreiche Reformen, durch die bestehende Schulformen abgeschafft und neue Schulformen eingeführt wurden. Tabelle 8.1 gibt einen Überblick über die Schulformen in den einzelnen Ländern für das Schuljahr der Erhebung der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) 2016 (Schuljahr 2015/16). Während Haupt-

Tabelle 8.1: Übersicht über die Schulformen der Sekundarstufe I in den Ländern der Bundesrepublik Deutschland (Schuljahr 2015/16)

Land	Schulen mit einem Bildungsgang			Schulen mit mehreren Bildungsgängen		
	Hauptschule	Realschule	Gymnasium	ohne Gymnasialbildung		mit Gymnasialbildung
				ohne gymn. Oberstufe	mit gymn. Oberstufe	
Baden-Württemberg	Hauptschule Werkrealschule ^A	Realschule	Gymnasium	–	–	Gemeinschaftsschule*
Bayern	Mittelschule ^A	Realschule	Gymnasium	–	(Gesamtschule) ^B	–
Berlin	–	–	Gymnasium	–	–	Integrierte Sekundarschule* Gemeinschaftsschule*
Brandenburg	–	–	Gymnasium	Oberschule ^{CD}	–	Gesamtschule
Bremen	–	–	Gymnasium	–	–	Oberschule
Hamburg	–	–	Gymnasium	–	–	Stadtteilschule
Hessen	Hauptschule	Realschule	Gymnasium	Mittelstufenschule	–	Gesamtschule*
Meckl.-Vorpommern	–	–	Gymnasium	Regionale Schule ^D	–	Gesamtschule*
Niedersachsen	Hauptschule	Realschule	Gymnasium	Oberschule ^E	Oberschule ^E	Gesamtschule*
Nordrhein-Westfalen	Hauptschule	Realschule	Gymnasium	–	Sekundarschule	Gesamtschule*
Rheinland-Pfalz	–	–	Gymnasium	Realschule plus ^{CD}	–	Gesamtschule*
Saarland	–	–	Gymnasium	–	–	Gemeinschaftsschule*
Sachsen	–	–	Gymnasium	Oberschule ^{CD}	–	–
Sachsen-Anhalt	–	–	Gymnasium	Sekundarschule	–	Gesamtschule Gemeinschaftsschule*
Schleswig-Holstein	–	–	Gymnasium	Regionalschule ^{DF}	–	Gemeinschaftsschule
Thüringen	–	–	Gymnasium	Regelschule ^D	–	Gesamtschule* Gemeinschaftsschule

* = Eine gymnasiale Oberstufe kann an der jeweiligen Schule oder als Kooperation mit anderen Schulen angeboten werden.

A = An dieser Schulform kann neben dem Hauptschulabschluss auch ein mittlerer Schulabschluss erworben werden, der jedoch nicht dem Realschulabschluss entspricht.

B = Insgesamt 5 Schulen, die als ‚Schulen besonderer Art‘ geführt werden.

C = An dieser Schulform ist der Erwerb der Fachoberschulreife möglich.

D = Wechsel an ein Gymnasium / Übertritt in die gymnasiale Oberstufe bei entsprechendem Notendurchschnitt oder dem Besuch spezieller Kurse (Sachsen) explizit möglich.

E = Kann als Oberschule ohne Gymnasialbildung oder als Oberschule mit Gymnasialbildung geführt werden.

F = Schulform läuft zum Schuljahr 2018/2019 aus. Bestehende Regionalschulen werden zu Gemeinschaftsschulen umgebaut.

Die Angaben basieren auf verfügbaren Veröffentlichungen der Kultusministerien der Länder der Bundesrepublik Deutschland (Stand: Schuljahr 2015/2016).

schulen zunehmend abgeschafft werden, finden sich in fast allen Ländern neue Schulformen, die mindestens den Haupt- und Realschulbildungsgang anbieten oder aber alle drei traditionellen Bildungsgänge umfassen. Bei letzteren kann weiterhin unterschieden werden zwischen Schulformen, die eine eigene gymnasiale Oberstufe führen, und Schulformen, von denen Schülerinnen und Schüler nach der zehnten Jahrgangsstufe auf eine andere Schule mit diesem Angebot wechseln müssen, um das Abitur zu erwerben.

Nur das Gymnasium findet sich in allen 16 Ländern und ist somit die einzige bundesweit einheitliche Schulform im Sekundarschulsystem. Die ursprüngliche Gliederung in Hauptschule, Realschule und Gymnasium ist in keinem Land des deutschen Schulsystems mehr vorhanden. Zwar gibt es in Baden-Württemberg, Hessen, Niedersachsen und Nordrhein-Westfalen noch Hauptschulen, diese werden jedoch durch ältere (z.B. Gesamtschule) und neuere integrierte Schulformen (z.B. Sekundarschule) ergänzt. Eine vergleichsweise übersichtliche Schullandschaft findet sich in den Stadtstaaten Bremen und Hamburg sowie in Sachsen. In Bremen und Hamburg besteht neben dem Gymnasium nur eine weitere Schulform, auf der Schülerinnen und Schüler alle Schulabschlüsse erreichen können (Oberschule bzw. Stadtteilschule). In Sachsen wird das Gymnasium ebenfalls nur durch eine weitere Schulform ergänzt, diese bietet jedoch keine Gymnasialbildung und damit keine Möglichkeit zum Erwerb des Abiturs (Oberschule).

Die Benennung neuer (zum Teil auch alter) Schulformen verläuft in den Ländern weitgehend autonom und folgt keiner bundesweit einheitlichen Regelung. So finden sich für neue Schulformen ohne Gymnasialbildung in den neun Ländern, in denen diese Schulform existiert, sieben unterschiedliche Bezeichnungen. In Mecklenburg-Vorpommern zum Beispiel heißt sie *Regionale Schule*, in Rheinland-Pfalz hingegen *Realschule plus*.

Eine weitere Herausforderung bei der Analyse des Übergangs von der Primar- zur Sekundarstufe stellen die gesetzlichen Vorgaben zur Übergangsempfehlung dar. Während in einigen Ländern von der Grundschule eine verbindliche Empfehlung für eine weiterführende Schule ausgesprochen wird (z.B. Bayern), werden von den Grundschulen anderer Länder zwar Empfehlungen gegeben, die endgültige Entscheidung liegt jedoch bei den Eltern (z.B. Nordrhein-Westfalen) (zum Überblick siehe Tabelle 8.2; auch Arnold, Bos, Richert & Stubbe, 2010). Um diesen Unterschieden zu entsprechen, müssten Analysen zum Thema Grundschulübergang eigentlich auf Länderebene durchgeführt werden. Da dies mit den Daten aus IGLU 2016 jedoch nicht möglich ist, werden in diesem Kapitel alle Befunde auf Bundesebene dargestellt.

Die Übersicht verdeutlicht, dass trotz der Unterschiede die abgebenden Grundschulen in allen Ländern eine Art von Übergangsempfehlung aussprechen. Die Ergebnisse der Studie *Der Übergang von der Grundschule in die weiterführende Schule – Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (ÜBERGANG), die als deutsche Erweiterung der *Trends in International Mathematics and Science Study* (TIMSS) 2007 durchgeführt wurde, verdeutlichen, dass diese Empfehlungen in erster Linie auf den Leistungen und Noten der Schülerinnen und Schüler basieren (Maaz, Baumert, Gresch & McElvany, 2010). Von zentraler Bedeutung sind hierbei die Leistungen und Noten in Deutsch und Mathematik, sowie – wenn auch in geringerem Maße – die Leistungen in der Domäne Naturwissenschaften beziehungsweise die Sachkundenote (Maaz & Nagy, 2009). Zudem wird in der Untersuchung beschrieben, dass Lehrkräfte ihre Empfehlungen zwar in erster Linie auf Grundlage

Tabelle 8.2: Übersicht über die Empfehlungspraktiken in den Ländern der Bundesrepublik Deutschland (Schuljahr 2015/16)

Land	Übergangsempfehlung der Grundschule	Endgültige Übergangsentscheidung liegt ...	
		... bei den Eltern	... bei der Grundschule
Baden-Württemberg	schriftlich	✓	-
Bayern	schriftlich	-	✓
Berlin	mündlich	✓	-
Brandenburg	schriftlich	-	✓
Bremen	(mündlich) ^A	✓	(✓) ^B
Hamburg	schriftlich	✓	-
Hessen	schriftlich	✓	-
Mecklenburg-Vorpommern	schriftlich	✓	(✓) ^C
Niedersachsen	(mündlich) ^D	✓	-
Nordrhein-Westfalen	schriftlich	✓	-
Rheinland-Pfalz	schriftlich	✓	-
Saarland	schriftlich	✓	-
Sachsen	schriftlich	-	✓
Sachsen-Anhalt	schriftlich	✓	-
Schleswig-Holstein	schriftlich	✓ ^E	(✓)
Thüringen	schriftlich	✓	-

A = Die Eltern werden im Rahmen eines verpflichtenden Beratungsgesprächs über den Leistungsstand ihrer Kinder informiert. Eine Übergangsempfehlung wird von der Grundschule nicht explizit ausgesprochen.

B = Nehmen Eltern nicht am verpflichtenden Beratungsgespräch zum Übergang teil, entscheidet die Grundschule.

C = Entscheiden sich die Eltern abweichend von der Empfehlung der Grundschule für ein Gymnasium oder einen gymnasialen Bildungsgang, besteht dort eine halbjährige Probezeit. Wird diese nicht bestanden, muss die Schülerin oder der Schüler das Gymnasium oder den gymnasialen Bildungsgang verlassen.

D = Die Eltern werden im Rahmen von mindestens zwei Beratungsgesprächen über die Lernentwicklung ihrer Kinder informiert. Eine Übergangsempfehlung wird von der Grundschule nicht explizit ausgesprochen.

E = Bei einer von der Empfehlung abweichenden Schulformentscheidung sind Eltern zur Teilnahme an einer Beratung an der weiterführenden Schule verpflichtet.

Die Angaben basieren auf verfügbaren Veröffentlichungen der Kultusministerien der Länder der Bundesrepublik Deutschland und der KMK (2015) (Stand: Schuljahr 2015/2016).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

von Noten und der von ihnen wahrgenommenen Begabung und Leistungen der Schülerinnen und Schüler formulieren, dass aber auch Merkmale wie Sozialverhalten oder Motivation von Bedeutung sind (Anders, McElvany & Baumert, 2010).

Neben dieser Studie wurden zum Thema ‚Grundschulübergang‘ in den vergangenen Jahren zahlreiche weitere Studien durchgeführt. In den drei bisherigen Zyklen von IGLU sowie im Rahmen von TIMSS 2011 und TIMSS 2015 wurde dieses Thema beispielsweise – auch unter Berücksichtigung von Länderunterschieden – ausführlich und hinsichtlich verschiedener Aspekte behandelt (Arnold, Bos, Richert & Stubbe, 2007, 2010; Bos et al., 2004; Milek, Lütke, Trautwein, Maaz & Stubbe, 2009; Milek, Stubbe, Trautwein, Lütke & Maaz, 2010; Stubbe & Bos, 2008; Stubbe, Bos & Euen, 2012; Stubbe, Lorenz, Bos & Kasper, 2016).

Zahlreiche regionale Studien haben sich in den vergangenen Jahren ebenfalls mit dem Übergang in die Sekundarstufe (und auch mit anderen Übergängen im Bildungsverlauf) befasst. Hierzu sind beispielsweise die Arbeiten der Forschergruppe *Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen*

im *Vorschul- und Schulalter* (BiKS) in Bayern und Hessen zu nennen (Kleine, Paulus & Blossfeld, 2009).

Interessante Möglichkeiten zur Analyse des Übergangs mit Längsschnittdaten werden sich in den nächsten Jahren durch das *Nationale Bildungspanel* (NEPS) eröffnen. Im Rahmen einer Teilstudie (Startkohorte 2) befragt das NEPS bereits seit 2010 eine Kohorte von Kindern seit deren Kindergartenzeit (LifBi, 2015). Mit den Daten der sechsten und siebten Erhebungswelle wird es in den nächsten Jahren möglich sein, für diese Kohorte den Übergang von der Primar- in die Sekundarstufe unter Heranziehung von Kompetenzdaten längsschnittlich zu beobachten.

2 Schullaufbahnpräferenzen von Lehrkräften und Eltern: Verteilung und Ausmaß der Übereinstimmung

Lehrkräfte wurden in IGLU 2016 wie schon in den vorangegangenen Zyklen nach dem erwarteten Schulabschluss der einzelnen Viertklässlerinnen und Viertklässler gefragt. Dieses Vorgehen ist der Abfrage der ausgesprochenen Übergangsempfehlung – wie es zum Teil im Rahmen von TIMSS gemacht wurde (Stubbe et al., 2012) – vorzuziehen, da die Empfehlungspraktiken und deren Verbindlichkeit zwischen den Ländern erheblich variieren (siehe Abschnitt 1). Mit der Frage nach dem erwarteten Schulabschluss (im Folgenden als ‚Schullaufbahnpräferenzen der Lehrkräfte‘ bezeichnet) erhält man einen guten Hinweis auf die prognostizierte Entwicklung der einzelnen Schülerinnen und Schüler aus Sicht ihrer Grundschullehrkräfte.

Wie in den vergangenen Erhebungen zu IGLU und TIMSS wurden die Schullaufbahnpräferenzen der Eltern über deren Angabe zu der Schulform erfasst, die ihre Kinder im kommenden Schuljahr voraussichtlich besuchen werden. Da die Eltern in fast allen Ländern die Schulanmeldungen ihrer Kinder vornehmen, kann ihre Präferenz als Indikator für die im Anschluss an die Grundschule besuchte Schule interpretiert werden. In Tabelle 8.3 finden sich die relativen Häufigkeiten der Schullaufbahnpräferenzen der Lehrkräfte und der Eltern für die vier Erhebungszeitpunkte von IGLU.

Für jeweils rund 40 Prozent der Schülerinnen und Schüler erwarten die in 2016 befragten Lehrkräfte einen Realschulabschluss beziehungsweise das Abitur als Schulabschluss. Für knapp 20 Prozent wird ein Hauptschulabschluss erwartet. Im Trend seit 2001 wird deutlich, dass die Lehrkräfte für immer weniger Kinder einen Hauptschulabschluss erwarten (Rückgang um 11.9 Prozentpunkte) und für immer mehr Kinder das Abitur (Anstieg um 8.1 Prozentpunkte).

Entsprechend der zunehmenden Verbreitung von Schulen mit mehreren Bildungsgängen findet sich bei den Elternpräferenzen ein deutlicher Anstieg zugunsten dieser Schulen. Nachdem sich der Anteil der Elternpräferenzen in den zehn Jahren von IGLU 2001 bis zu IGLU 2011 von 7.8 Prozent auf 14.3 Prozent fast verdoppelt hatte, zeigt sich in den fünf Jahren von 2011 bis 2016 erneut nahezu eine Verdoppelung auf inzwischen 25.0 Prozent.

Realschulen und insbesondere Hauptschulen verlieren nach Angabe der Eltern entsprechend Anteile der Schülerschaft. Seit 2001 ist der Anteil von Hauptschülerinnen und Hauptschülern von 22.1 auf 6.1 Prozent eingebrochen. Eine Realschule besuchen in der aktuellen Studie 21.3 Prozent der Fünftklässlerinnen und Fünftklässler gegenüber 29.2 Prozent im Jahr 2001.

Tabelle 8.3: Schullaufbahnpräferenzen der Lehrkräfte und Eltern für Deutschland 2001, 2006, 2011 und 2016 in Zeilenprozent

	Hauptschule	Realschule	Gymnasium	Schulen mit mehreren Bildungsgängen
Lehrkräfte				
2016	17.4	39.6	43.0	-
2011	21.6	36.6	41.8	-
2006	24.8	35.5	39.7	-
2001	29.3	35.7	34.9	-
Eltern				
2016	6.1	21.3	47.6	25.0
2011	11.9	28.3	45.5	14.3
2006	14.7	27.0	47.3	11.0
2001	22.1	29.2	40.8	7.8

Differenzen zu 100 Prozent ergeben sich durch Rundungsfehler.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Die Kreuztabelle 8.4 zeigt für die drei Präferenzen der Lehrkräfte den jeweiligen Anteil der Schülerinnen und Schüler nach Elternpräferenz. Vergleicht man diese Werte mit jenen aus dem Jahr 2011 (Stubbe et al., 2012, Tabelle 8.4), zeigen sich relativ große Unterschiede. Diese Veränderungen ergeben sich insbesondere durch die steigende Bedeutung der Schulen mit mehreren Bildungsgängen (siehe Abschnitt 1). Nur 34.9 Prozent der Viertklässlerinnen und Viertklässler, deren Lehrkräfte für sie einen Hauptschulabschluss erwarten, werden den Elternangaben zufolge tatsächlich auf diese Schulform wechseln (2011: 54.5%). Praktisch unverändert sind die Prozentwerte für den Wechsel auf eine Realschule (2016: 23.0%; 2011: 22.1%) beziehungsweise auf ein Gymnasium (2016: 1.7%; 2011: 1.2%). Der Wert für Schulen mit mehreren Bildungsgängen ist 2016 erwartungskonform deutlich höher als noch fünf Jahre zuvor (2016: 40.6%; 2011: 22.2%).

Auch bei den Schülerinnen und Schülern, die eine Lehrkräftepräferenz für die Realschule aufweisen, fällt der Anteil derer, die auf eine Realschule wechseln werden, mit 44.3 Prozent geringer aus als 2011 (59.5%). Besonders deutlich hat sich auch in dieser Gruppe der Wert für die Schulen mit mehreren Bildungsgängen verändert (36.0% vs. 18.8%).

Für Kinder, für die die Grundschullehrkräfte das Abitur erwarten, sind Schulen mit mehreren Bildungsgängen hingegen deutlich weniger bedeutsam. Wie schon vor fünf Jahren besuchen knapp 90 Prozent von ihnen ein Gymnasium. Der Anteil, der auf eine Schule mit mehreren Bildungsgängen wechselt, ist moderat um 1.8 Prozentpunkte auf 8.6 Prozent gestiegen.

Tabelle 8.4: Übereinstimmungen der Schullaufbahnpräferenzen von Lehrkräften und Eltern in Zeilenprozent

		Schullaufbahnpräferenz der Eltern				<i>n</i>
		Hauptschule	Realschule	Gymnasium	Schule mit mehreren Bildungsgängen	
Schullaufbahnpräferenz der Lehrkräfte	Hauptschule	34.9	23.0	1.7	40.6	381
	Realschule	2.1	44.3	17.7	36.0	928
	Gymnasium	0.0	4.3	87.1	8.6	1237
<i>n</i>		153	555	1247	591	2546

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

3 Leistungsrelevante Merkmale der Schülerinnen und Schüler sowie Schullaufbahnpräferenzen der Lehrkräfte

In allen Ländern sollen die Noten der Schülerinnen und Schüler das wichtigste Kriterium für deren weitere Schullaufbahn sein (KMK, 2015). Entsprechend zeigt sich auch für IGLU 2016 ein enger Zusammenhang zwischen den Noten in den Hauptfächern und der Schullaufbahnpräferenz der Grundschullehrkräfte (siehe Abbildung 8.1).

Die Varianz der Präferenz kann zu 73 Prozent durch die Noten in Deutsch und in Mathematik erklärt werden. Wie auch in anderen Studien zeigt sich zudem, dass die Deutschnote einen engeren Zusammenhang mit der Schullaufbahnpräferenz der Lehrkräfte aufweist als die Mathematiknote.

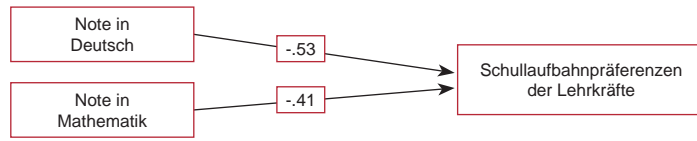
Da aus unterschiedlichen Studien bekannt ist, dass die Korrelation zwischen Schulnoten und standardisierten Leistungstests relativ gering ausfällt (im Überblick Lintorf, 2012), wird dieser Zusammenhang in Abbildung 8.2 visualisiert. Die deutlichen Überlappungen der vier Kurven zeigen, dass Schülerinnen und Schüler mit mittleren Testleistungen je nach Lehrkraft alle Noten von ausreichend (oder schlechter) bis sehr gut erhalten.

Entsprechend zeigt die Abbildung 8.3 den Zusammenhang zwischen der Lesekompetenz und den Schullaufbahnpräferenzen der Lehrkräfte. Auch in dieser Abbildung erkennt man die deutlichen Überlappungen der drei Kurven, was darauf hinweist, dass die Lehrkräfte von Schülerinnen und Schülern mit mittleren Testleistungen – insbesondere für diejenigen der Kompetenzstufe III – Präferenzen für alle drei Schulformen aufweisen.

Neben den Fachleistungen sollen nach Vorgabe der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) bei der Festlegung der Übergangsempfehlung auch allgemeine leistungsrelevante Eigenschaften der Schülerinnen und Schüler Berücksichtigung finden (KMK, 2015). In Tabelle 8.5 ist der Zusammenhang zwischen ausgewählten motivationalen und volitionalen Persönlichkeitsmerkmalen der Schülerinnen und Schüler sowie der Schullaufbahnpräferenz der Lehrkräfte dargestellt.

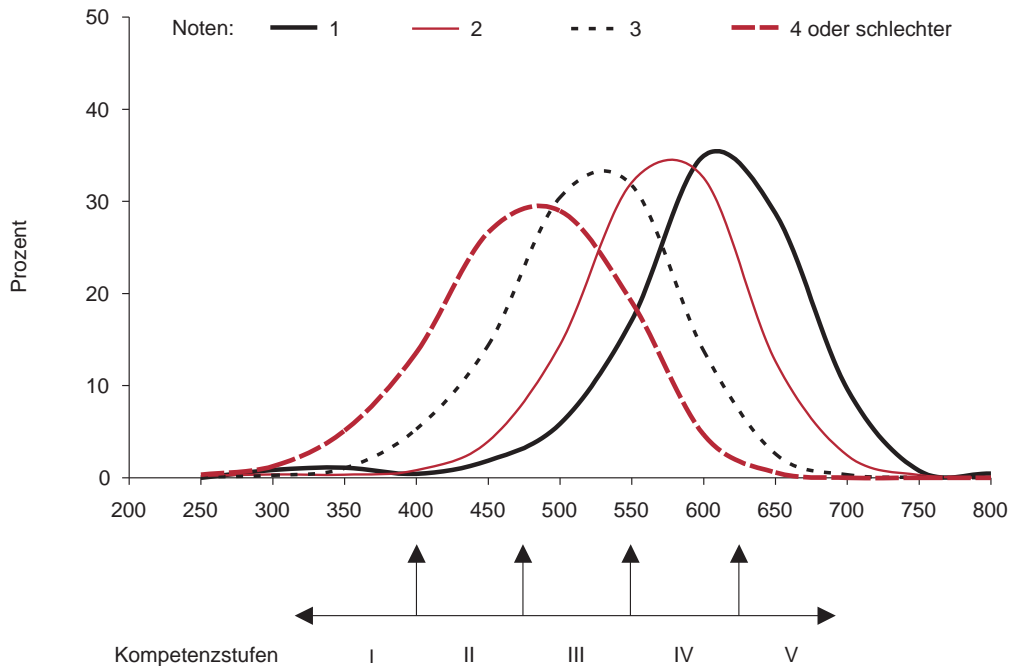
Entsprechend der bildungspolitischen Vorgaben zeigen sich positive Zusammenhänge zwischen den Schullaufbahnpräferenzen der Lehrkräfte und dem Leseselbstkonzept (erhoben über den Schülerfragebogen) sowie der Anstren-

Abbildung 8.1: Zusammenhang der Schulnoten mit der Schullaufbahnpräferenz der Lehrkräfte (standardisierte Regressionskoeffizienten)



$R^2 = .73$
 Alle berichteten Pfadkoeffizienten sind signifikant ($p < .05$).

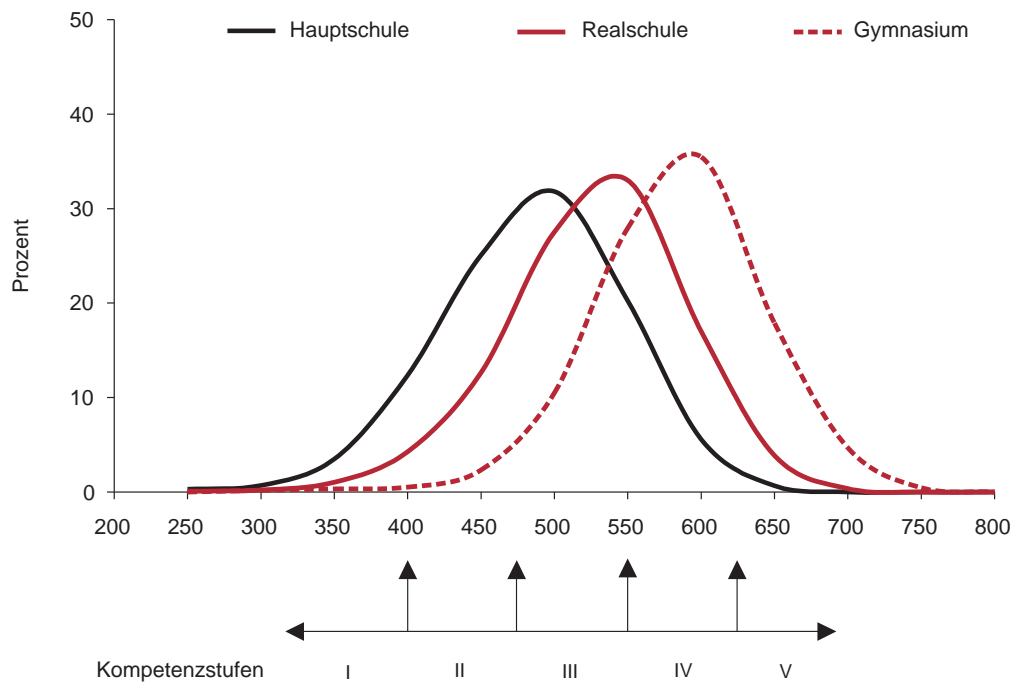
Abbildung 8.2: Testleistungen differenziert nach Deutschnoten – Gesamtskala Lesen



gungsbereitschaft (erhoben über den Elternfragebogen) der Viertklässlerinnen und Viertklässler beziehungsweise ein negativer Zusammenhang zwischen der Lehrkräftepräferenz und der Leistungsangst (erhoben über den Elternfragebogen) der Kinder. Im Trend seit IGLU 2006 (Arnold et al., 2007, Tabelle X.6; Stubbe et al., 2012, Tabelle 8.7) zeigen sich für diese Werte praktisch keine Veränderungen – lediglich für die Anstrengungsbereitschaft fiel die Korrelation im Jahr 2006 mit .25 substantiell geringer aus als in den Folgestudien (2011: .42, 2016: .46).

Auch unter Kontrolle der Lesekompetenz und sogar unter Kontrolle der Deutschnote lassen sich signifikante Zusammenhänge zwischen den Persönlichkeitsmerkmalen und den Schullaufbahnpräferenzen feststellen. Während diese Berücksichtigung leistungsrelevanter Merkmale der Schülerinnen und Schüler

Abbildung 8.3: Testleistungen differenziert nach Schullaufbahnpräferenzen der Lehrkräfte – Gesamtskala Lesen



IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

beim Übergang in die Sekundarstufe erwünscht sein mag, wird der folgende Abschnitt zeigen, dass auch Hintergrundmerkmale der Schülerinnen und Schüler sowie ihrer Familien, die nicht unmittelbar mit Leistung in Verbindung stehen, einen Zusammenhang mit den Präferenzen der Grundschullehrkräfte aufweisen. Dies widerspricht dem gesellschaftlichen Anspruch auf gleiche Chancen, höhere Bildungsabschlüsse erreichen zu können – unabhängig vom sozialen Status des Elternhauses.

Tabelle 8.5: Leistungsrelevante Merkmale von Schülerinnen und Schülern und Schullaufbahnpräferenz der Lehrkräfte (ohne und mit Berücksichtigung der Lesekompetenz beziehungsweise der Deutschnote)

	Korrelation mit der Schullaufbahnpräferenz der Lehrkräfte	Korrelation mit der Schullaufbahnpräferenz der Lehrkräfte unter Kontrolle der Lesekompetenz	Korrelation mit der Schullaufbahnpräferenz der Lehrkräfte unter Kontrolle der Deutschnote
Selbstkonzept Lesen	0.36	0.19	0.09
Leistungsangst	-0.33	-0.25	-0.19
Anstrengungsbereitschaft	0.46	0.35	0.20

Alle Korrelationen sind signifikant auf einem Niveau von .001.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

4 Soziale Merkmale der Schülerinnen und Schüler und Schullaufbahnpräferenzen der Lehrkräfte und der Eltern

Die soziale Lage von Familien wird im Folgenden mit Hilfe der EGP-Klassen (Erikson, Goldthorpe & Portocarero, 1979) beschrieben. Dabei werden Personen ausgehend von Angaben zu ihrer beruflichen Stellung distinkten Gruppen zugeordnet (siehe Kapitel 6 in diesem Band).

In Tabelle 8.6 sind die relativen Chancen (odds ratios) für eine Gymnasialpräferenz der Lehrkräfte beziehungsweise der Eltern in Abhängigkeit von der sozialen Lage der Familie dargestellt. Als Referenzgruppe dienen jeweils Kinder, deren Lehrkräfte beziehungsweise Eltern eine Präferenz für eine andere Schulform als das Gymnasium haben und deren Eltern der Gruppe der (Fach-)Arbeiter angehören.

Ohne Berücksichtigung weiterer Einflussfaktoren (Modell I) ist die Chance einer Schülerin beziehungsweise eines Schülers aus der oberen Dienstklasse, eine Gymnasialpräferenz der Lehrkraft zu erhalten, 5.18-mal (2011: 4.71-mal, 2006: 3.97, 2001: 4.42-mal) so hoch wie die eines (Fach-)Arbeiterkindes. Da ein Zusammenhang zwischen der EGP-Klasse der Familie und den schulischen Leistungen der Kinder besteht (siehe Kapitel 6 in diesem Band), werden in Modell II zunächst die kognitiven Fähigkeiten (siehe Kapitel 2 in diesem Band) und in Modell III zudem die Lesekompetenz der Viertklässlerinnen und Viertklässler kontrolliert, um zu prüfen, ob ungleiche Chancen in ungleichen Leistungen begründet liegen. Zwar sind die odds ratios in diesen Modellen geringer; zumindest für die obere und untere Dienstklasse bleiben sie jedoch signifikant und fallen tendenziell sogar etwas höher aus als in den vorangegangenen Studienzyklen (Arnold et al., 2007, Tabelle X.7; Stubbe et al., 2012, Tabelle 8.8).

Der Trend seit IGLU 2001 ist für die Präferenz der Lehrkräfte in Tabelle 8.7 komprimiert dargestellt. Es werden die odds ratios für die sogenannte *service class* (EGP I und II) im Vergleich zur *working class* (EGP V, VI und VII)

Tabelle 8.6: Relative Chancen (odds ratios) für eine Gymnasialpräferenz der Lehrkräfte beziehungsweise der Eltern nach sozialer Lage (EGP) der Familien

Sozioökonomische Stellung der Familie (EGP-Klassen)	Gymnasialpräferenz der Lehrkräfte			Gymnasialpräferenz der Eltern		
	Modell I	Modell II	Modell III	Modell I	Modell II	Modell III
Obere Dienstklasse (I)	5.18**	4.79**	3.81**	4.84**	4.39**	3.44**
Untere Dienstklasse (II)	3.39**	3.26**	2.68**	2.76**	2.54**	2.02*
Routinedienstklasse (III)	2.02**	2.01**	— ^{ns}	1.63**	1.53*	— ^{ns}
Selbstständige (IV)	1.65**	1.67**	— ^{ns}	— ^{ns}	— ^{ns}	— ^{ns}
(Fach-)Arbeiter (V, VI)	Referenzgruppe (odds ratio = 1)					
Un- und angelernte Arbeiter (VII)	— ^{ns}	— ^{ns}	— ^{ns}	0.66*	— ^{ns}	— ^{ns}
Nagelkerkes R^2	0.12	0.22	0.32	0.11	0.19	0.27

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

Modell I: Ohne Kontrolle von Kovariaten.

Modell II: Kontrolle der kognitiven Fähigkeiten.

Modell III: Kontrolle der kognitiven Fähigkeiten und der Lesekompetenz (internationale Skalierung).

Tabelle 8.7: Relative Chancen (odds ratios) für eine Gymnasialpräferenz der Lehrkräfte für Kinder aus der *service class* (EGP I und II) im Vergleich mit Kindern aus der *working class* (EGP V, VI und VII) bei IGLU 2001, 2006, 2011 und 2016

	Modell I	Modell II	Modell III
2001	4.18**	3.49**	2.63**
2006	4.06**	3.40**	2.72**
2011	4.48**	4.07**	3.14**
2016	5.13**	4.76**	3.37**

Signifikanzniveau: ** = signifikant ($p < .01$).

Modell I: Ohne Kontrolle von Kovariaten.

Modell II: Kontrolle der kognitiven Fähigkeiten.

Modell III: Kontrolle der kognitiven Fähigkeiten und der Lesekompetenz (internationale Skalierung).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

für eine Gymnasialpräferenz der Lehrkräfte berichtet. In allen drei Modellen lässt sich im Zeitverlauf eine Zunahme der sozialen Disparitäten bei den Schullaufbahnpräferenzen feststellen.

Ein umgekehrter Trend lässt sich seit 2006 für die Schullaufbahnpräferenz der Eltern beobachten (Arnold et al., 2007, Tabelle X.7; Stubbe et al., 2012, Tabelle 8.8). In Modell I (siehe Tabelle 8.6) beträgt der Wert für die obere Dienstklasse in der aktuellen Studie 4.84 (2011: 5.21, 2006: 5.51). Unter Kontrolle der kognitiven Fähigkeiten und der Lesekompetenz ist die Chance für eine Gymnasialpräferenz der Eltern bei IGLU 2016 in der oberen Dienstklasse 3.44-mal (2011: 3.76, 2006: 3.83) so hoch wie in der Klasse der (Fach-)Arbeiter. Auch wenn die Gründe für diese Entwicklung ohne weitere Daten und Analysen nicht verlässlich identifiziert werden können, liegt die Vermutung nahe, dass die veränderte Schulstruktur in praktisch allen Ländern hierfür verantwortlich ist. Auch für Eltern, die sich für ihr Kind das Abitur wünschen, stellt das Gymnasium nicht mehr die einzige Option dar.

Mädchen haben ohne Berücksichtigung weiterer Merkmale eine größere Chance auf eine Gymnasialpräferenz als Jungen (ohne Tabelle) – sowohl bei den Lehrkräften (1.26-mal so hoch, $p < .05$) als auch bei den Eltern (1.40-mal so hoch, $p < .01$). Im Rahmen von IGLU 2001 und IGLU 2006 konnten entsprechende Effekte gefunden werden. Bei IGLU 2011 gab es hingegen keine unterschiedlichen Chancen von Mädchen und Jungen. Unter Kontrolle der kognitiven Fähigkeiten beziehungsweise der kognitiven Fähigkeiten und der Lesekompetenz lassen sich keine signifikanten geschlechtsspezifischen Ungleichheiten bei der Schullaufbahnpräferenz der Lehrkräfte feststellen. Anders sieht es bei den Eltern aus, die auch unter Kontrolle dieser Merkmale für Mädchen häufiger das Gymnasium präferieren als für Jungen.

Ein Zusammenhang mit dem Migrationshintergrund der Schülerinnen und Schüler lässt sich weder für die Präferenz der Lehrkräfte noch für die Präferenz der Eltern feststellen (ohne Tabelle). In vorangegangenen Zyklen konnten zum Teil entsprechende Effekte identifiziert werden, wobei sich bereits eine Tendenz zur Reduzierung dieser Disparitäten angedeutet hat (Stubbe et al., 2012, siehe vergleichend für TIMSS 2015 auch Stubbe et al., 2016).

Eine alternative und anschauliche Möglichkeit, den Zusammenhang zwischen Herkunft und Schullaufbahnentscheidung darzustellen, ist die Berechnung der-

jenigen Kompetenzwerte, die notwendig sind, damit Schülerinnen und Schüler mit einer Wahrscheinlichkeit von mehr als 50 Prozent eine Empfehlung für das Gymnasium erhalten. Diese Kompetenzwerte werden im Folgenden als ‚kritische Werte‘ beziehungsweise gruppenspezifische Standards bezeichnet. In Tabelle 8.8 sind für die Gesamtpopulation, die sechs EGP-Klassen, die drei Gruppen nach Migrationsstatus der Eltern sowie für Mädchen und Jungen die jeweiligen ‚kritischen Werte‘ (gruppenspezifische Standards) für die Lesekompetenz sowohl hinsichtlich der Lehrerpräferenz als auch hinsichtlich der Elternpräferenz dargestellt. Abbildung 8.4 zeigt zudem für die Werte der sechs EGP-Klassen den Trend seit 2001.

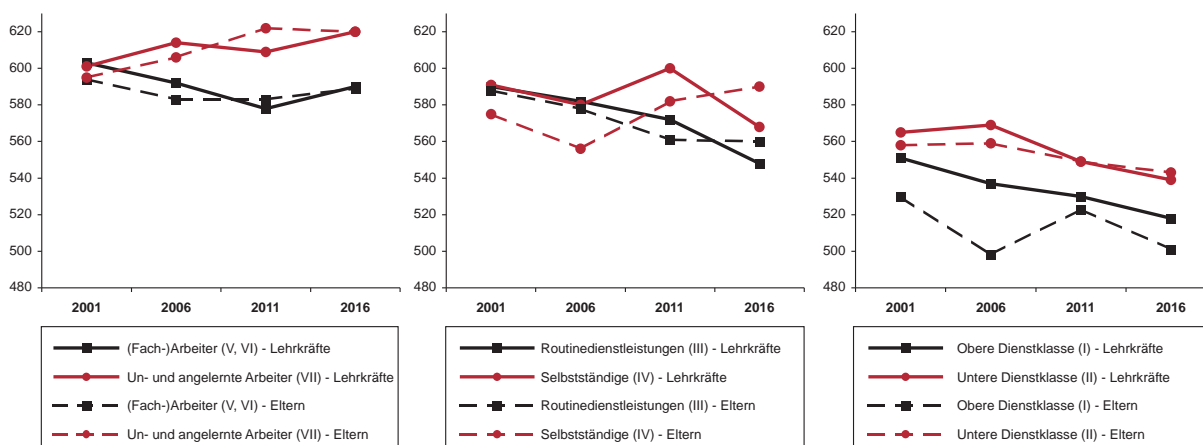
Insgesamt wird eine Gymnasialpräferenz der Lehrkräfte ab einem Lesekompetenzwert von 562 Punkten (25 Punkte oberhalb des deutschen Mittelwertes) hinreichend wahrscheinlich. Zwischen den EGP-Klassen unterscheiden sich die gruppenspezifischen Standards zum Teil erheblich. Während Kinder aus der oberen Dienstklasse bereits mit einer Kompetenz, die 19 Punkte unterhalb des deutschen Mittelwertes liegt, gute Chancen auf eine Gymnasialpräferenz ihrer Lehrkräfte haben, benötigen Kinder von (Fach-)Arbeitern eine Kompetenz, die 53 Punkte oberhalb des deutschen Mittelwertes liegt. Die Differenz zwischen diesen beiden EGP-Klassen beträgt mit 72 Punkten somit etwa anderthalb Lernjahre (Bos et al., 2007).

Beim Migrationshintergrund und beim Geschlecht der Kinder finden sich hingegen keine substantiell unterschiedlichen gruppenspezifischen Standards. Tendenziell benötigen Kinder, deren Eltern beide im Ausland geboren wurden, eine geringfügig höhere Lesekompetenz als Kinder ohne Migrationshintergrund, damit ihre Lehrkräfte das Abitur für wahrscheinlich halten. Selbiges gilt für Jungen im Vergleich mit Mädchen.

Tabelle 8.8: Gruppenspezifische Standards („kritische Werte“) für eine Gymnasialpräferenz der Lehrkräfte und der Eltern – Gesamtskala Lesen

	Gruppenspezifischer Standard („kritischer Wert“) für eine Gymnasialpräferenz der Lehrkräfte Lesen	Gruppenspezifischer Standard („kritischer Wert“) für eine Gymnasialpräferenz der Eltern Lesen
Gesamt	562	558
Obere Dienstklasse (I)	518	501
Untere Dienstklasse (II)	539	543
Routinedienstleistungen (III)	548	560
Selbstständige (IV)	568	590
(Fach-)Arbeiter (V,VI)	590	589
Un- und angelernte Arbeiter (VII)	620	620
kein Elternteil im Ausland geboren	559	562
ein Elternteil im Ausland geboren	553	544
beide Elternteile im Ausland geboren	571	534
Mädchen	559	550
Jungen	566	568

Abbildung 8.4: Gruppenspezifische Standards („kritische Werte“) für eine Gymnasialpräferenz der Lehrkräfte und der Eltern – Gesamtskala Lesen 2001, 2006, 2011 und 2016



IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Hinsichtlich der Elternpräferenzen ist insbesondere der Trend seit 2001 interessant (Arnold et al., 2007, Tabelle X.8; Stubbe et al., 2012, Tabelle 8.10). In den ersten beiden IGLU-Zyklen lagen für den Kompetenzbereich Lesen die ‚kritischen Werte‘ der Eltern unterhalb der entsprechenden Werte für die Lehrkräfte (2001: 9 Punkte Differenz; 2006: 15 Punkte Differenz). Im Jahr 2011 gab es sowohl im Lesen (IGLU) als auch in Mathematik und Naturwissenschaften (TIMSS) kaum nennenswerte Unterschiede zwischen Eltern und Lehrkräften (jeweils 3 Punkte Differenz). Selbiges gilt für die aktuelle Studie, in der der Wert der Eltern nominell vier Punkte niedriger ist als jener der Lehrkräfte. Diese Befunde entsprechen denen der Regressionsanalysen (Tabelle 8.6) und lassen sich wie dort ausgeführt wahrscheinlich auf die veränderte Schullandschaft zurückführen, da das Gymnasium nicht mehr die einzige Option für Eltern darstellt, die sich für ihr Kind das Abitur wünschen.

Betrachtet man die ‚kritischen Werte‘ nach dem sozialen Status der Elternhäuser, zeigen sich für die EGP-Klassen II, V/VI und VII praktisch keine Unterschiede zwischen Lehrer- und Elternpräferenz. Bei Kindern der EGP-Klassen III und IV sind die Eltern sogar etwas weniger zuversichtlich als die Grundschullehrkräfte. Lediglich in der oberen Dienstklasse fällt der ‚kritische Wert‘ bei den Eltern deutlich geringer als bei den Lehrkräften aus. Bereits ab einer Leistung, die 37 Punkte unterhalb des deutschen Mittelwertes liegt, wird in diesen Familien die Entscheidung für ein Gymnasium wahrscheinlich. Kinder von (Fach-)Arbeitern benötigen für dieselbe Schulwahl eine Leseleistung, die einem Vorsprung von knapp zwei Lernjahren (89 Punkte) entspricht.

Hinsichtlich des Migrationshintergrundes zeigt sich die aus anderen Studien bekannte relativ hohe Bildungsaspiration von Familien mit Migrationshintergrund (z. B. Ditton et al., 2005; Dumont, Maaz, Neumann & Becker, 2014; Gresch, 2012). Während bei Kindern, von denen mindestens ein Elternteil im Ausland geboren wurde, der Übergang aufs Gymnasium bereits mit durchschnittlichen Lesekompetenzen wahrscheinlich ist, benötigen Kinder ohne Migrationshintergrund nach Einschätzung der Eltern einen Wert, der 25 Punkte oberhalb des deutschen Mittelwertes liegt.

Bei dem Geschlecht der Schülerinnen und Schüler zeigt sich für die Elternpräferenz ein etwas stärkerer Vorteil für Mädchen als bei der Lehrerpräferenz:

Für einen wahrscheinlichen Wechsel auf ein Gymnasium benötigen Jungen eine um 18 Punkte höhere Lesekompetenz als Mädchen.

5 Zusammenfassung

Die in diesem Beitrag vorgelegten Trendanalysen zum Übergang von der Grundschule in die Sekundarstufe bei IGLU 2001, 2006, 2011 und 2016 zeigen ein insgesamt relativ konstantes Bild. Veränderungen ergeben sich insbesondere durch die veränderte Schulstruktur. Jedes vierte Kind wechselt den Elternangaben zufolge 2016 auf eine Schule mit mehreren Bildungsgängen – damit ist dieser Anteil mehr als dreimal so hoch wie 15 Jahre zuvor. Knapp die Hälfte der Viertklässlerinnen und Viertklässler wechselt auf ein Gymnasium. Seit 2006 ist dieser Wert relativ konstant. Hinsichtlich der Schullaufbahnpräferenzen der Lehrkräfte zeigen sich eine stetige Abnahme der Werte für die Hauptschule und ein entsprechender Anstieg für das Gymnasium.

Wie schon in den vorangegangenen Zyklen ergibt sich auch für IGLU 2016 ein relativ geringer Zusammenhang zwischen den Ergebnissen des Lesetests und der Deutschnote. Aus der zugleich engen Kopplung von Noten und Schullaufbahnpräferenzen der Lehrkräfte ergibt sich ein ebenfalls relativ geringer Zusammenhang zwischen den Ergebnissen des Lesetests und den von den Lehrkräften angegebenen Präferenzen. Dabei muss aber berücksichtigt werden, dass die Lesekompetenz nur eines von verschiedenen Kriterien (z. B. Kompetenzen in anderen Fächern oder Anstrengungsbereitschaft) ist, die Grundschullehrkräfte bei ihrer Empfehlung berücksichtigen sollen.

Bedenklich ist hingegen der konstant hohe Zusammenhang zwischen der sozialen Herkunft der Schülerinnen und Schüler und den Schullaufbahnpräferenzen. Auch unter Kontrolle der Lesekompetenz und der Deutschnote haben Kinder aus bildungsnahen Elternhäusern eine deutlich höhere Chance auf eine Gymnasialpräferenz als Kinder aus bildungsfernen Familien.

Literatur

- Anders, Y., McElvany, N. & Baumert, J. (2010). Die Einschätzung lernrelevanter Schülermerkmale zum Zeitpunkt des Übergangs von der Grundschule auf die weiterführende Schule: Wie differenziert urteilen Lehrkräfte? In K. Maaz, J. Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 313–330). Bonn: Bundesministerium für Bildung und Forschung.
- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T. C. (2007). Schullaufbahnpräferenzen am Ende der vierten Klassenstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 271–297). Münster: Waxmann.
- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T. C. (2010). Der Übergang von der Grundschule in die Sekundarstufe: Schullaufbahnpräferenzen von Lehrkräften und Eltern im Ländervergleich. 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. 13–32). Münster: Waxmann.
- Bellenberg, G. & Klemm, K. (2000). Scheitern im System, Scheitern des Systems? Ein etwas anderer Blick auf Schulqualität. In H.-G. Rolff, W. Bos, K. Klemm, H. Pfeiffer

- & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung* (S. 51–75). Weinheim: Juventa.
- Bos, W., Valtin, R., Hornberg, S., Buddeberg, I., Goy, M. & Voss, A. (2007). Internationaler Vergleich 2006: Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 109–160). Münster: Waxmann.
- Bos, W., Voss, A., Lankes, E.-M., Schwippert, K., Thiel, O. & Valtin, R. (2004). Schullaufbahneempfehlungen von Lehrkräften für Kinder am Ende der vierten Jahrgangsstufe. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 191–228). Münster: Waxmann.
- Ditton, H., Krüskens, J. & Schauenberg, M. (2005). Bildungsungleichheit – der Beitrag von Familie und Schule. *Zeitschrift für Erziehungswissenschaft*, 8 (2), 285–304.
- Dumont, H., Maaz, K., Neuman, M. & Becker, M. (2014). Soziale Ungleichheiten beim Übergang von der Grundschule in die Sekundarstufe I. Theorie, Forschungsstand, Interventions- und Fördermöglichkeiten. *Zeitschrift für Erziehungswissenschaft*, 17 (24), 141–165.
- Erikson, R., Goldthorpe, H. J. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30 (4), 415–441.
- Gresch, C. (2012). *Der Übergang in die Sekundarstufe I. Leistungsbeurteilung, Bildungsaspiration und rechtlicher Kontext bei Kindern mit Migrationshintergrund*. Wiesbaden: Springer VS.
- Kleine, L., Paulus, W. & Blossfeld, H.-P. (2009). Die Formation elterlicher Bildungsentscheidungen beim Übergang von der Grundschule in die Sekundarstufe I. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft, Sonderheft 12, S. 103–125). Wiesbaden: VS Verlag für Sozialwissenschaften.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe)*. Zugriff am 01.10.2016 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf
- Köller, O. (2008). Gesamtschule – Erweiterung statt Alternative. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 437–465). Reinbek: Rowohlt.
- LifBi – Leibniz-Institut für Bildungsverläufe e.V. (2015). *Startkohorte 2: Kindergarten (SC2). Studienübersicht Wellen 1 bis 5*. Zugriff am 01.10.2016 unter https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC2/SC2_Studien_W1-5.pdf
- Lintorf, K. (2012). *Wie vorhersagbar sind Grundschulnoten? Prädiktionskraft individueller und kontextspezifischer Merkmale*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Maaz, K., Baumert, J., Gresch, C. & McElvany, N. (Hrsg.). (2010). *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Berlin: Bundesministerium für Bildung und Forschung.
- Maaz, K. & Nagy, G. (2009). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft, Sonderheft 12, S. 153–182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Milek, A., Lüdtke, O., Trautwein, U., Maaz, K. & Stubbe, T. C. (2009). Wie konsistent sind Referenzgruppeneffekte bei der Vergabe von Schulformempfehlungen? Bundesland-spezifische Analysen von Daten der IGLU-Studie. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft, Sonderheft 12, S. 282–301). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Milek, A., Stubbe, T. C., Trautwein, U., Lüdtke, O. & Maaz, K. (2010). *Reference group effects on teachers' school track recommendations: Results from PIRLS 2006 Germany*. Vortrag auf der 4th IEA International Research Conference, Göteborg. Zugriff am 01.10.2016 unter http://www.iea.nl/fileadmin/user_upload/IRC/IRC_2010/Papers/IRC2010_Milek_Stubbe_etal.pdf
- Schuchart, C. (2007). Schulabschluss und Ausbildungsberuf. Zur Bedeutung der schulartbezogenen Bildungsbiografie. *Zeitschrift für Erziehungswissenschaft*, 10 (3), 381–398.
- Schuchart, C. (2013). Kein Abschluss ohne Anschluss? Durchlässigkeit und Vergleichbarkeit von Bildungswegen in der Sekundarstufe II. *Die Deutsche Schule*, 105 (4), 345–363.
- Stubbe, T. C. (2009). *Bildungsentscheidungen und sekundäre Herkunftseffekte. Soziale Disparitäten bei Hamburger Schülerinnen und Schülern der Sekundarstufe I*. Münster: Waxmann.
- Stubbe, T. C. & Bos, W. (2008). Schullaufbahneempfehlungen von Lehrkräften und Schullaufbahnentscheidungen von Eltern am Ende der vierten Jahrgangsstufe. *Empirische Pädagogik*, 22 (1), 49–63.
- Stubbe, T. C., Bos, W. & Euen, B. (2012). Der Übergang von der Primar- in die Sekundarstufe. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 209–226). Münster: Waxmann.
- Stubbe, T. C., Lorenz, J., Bos, W. & Kasper, D. (2016). Der Übergang von der Primar- in die Sekundarstufe. 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. 351–365). Münster: Waxmann.

Kapitel IX

Unterrichtsführung, Sozialklima und kognitive Aktivierung im Deutschunterricht in vierten Klassen

Ruven Stahns, Svenja Rieser und Eva-Maria Lankes

1 Einleitung

Sowohl in Deutschland als auch international wurden in den letzten Jahren und Jahrzehnten Untersuchungen mit dem Ziel durchgeführt, Unterrichtsmerkmale zu ermitteln, die mit fachlichem Lernzuwachs, aber auch mit dem Aufbau eines positiven domänenspezifischen Fähigkeitsselbstkonzepts, der Entwicklung von Lernmotivation und dem Fachinteresse zusammenhängen (zusammenfassend Wang, Haertel & Walberg, 1993; Seidel & Shavelson, 2007; König, Wagner & Valtin, 2011; Hattie, 2014, 2015; Helmke, 2015). Lipowsky (2007) hat auf Befunde aus dem US-amerikanischen Raum aufmerksam gemacht, die darauf hinweisen, dass die Unterrichtsqualität und Merkmale der Lehrkräfte für die Qualität des Unterrichts in der Grundschule und insbesondere für die Lernentwicklung von leistungsschwächeren Schülerinnen und Schülern mit ungünstigen Lernvoraussetzungen von Bedeutung sind. Nachdem in Deutschland im Anschluss an die Veröffentlichung der Ergebnisse des *Programme for International Student Assessment* (PISA) vor allem Untersuchungen zum Deutschunterricht in der Sekundarstufe I realisiert wurden, wurden in den letzten Jahren zunehmend Studien durchgeführt, in denen Merkmale des Deutsch- oder auch im Speziellen des Leseunterrichts in der Grundschule im Mittelpunkt standen (z.B. Lankes, 2004; Lankes & Carstensen, 2007; Kleinbub, 2010; Gabriel, 2014; Lotz, 2016).

Merkmale des Leseunterrichts sind Gegenstand von Kapitel 10 in diesem Band. Im Mittelpunkt des vorliegenden Kapitels steht der Deutschunterricht am Ende der vierten Klasse. Ausgehend von der Unterscheidung von Unterrichtsquantität und -qualität in einem „Angebots-Nutzungs-Modell“ (Helmke, 2015) (siehe Abschnitt 2) werden im ersten Teil des vorliegenden Kapitels einige Befunde zu der zur Verfügung stehenden Unterrichtszeit vorgestellt, die in den Teilnehmerstaaten und -regionen der *Internationalen Grundschul-Lese-Untersuchung* (IGLU 2016) für den Deutsch- beziehungsweise Sprach-

unterricht aufgebracht wird (siehe Abschnitt 3). Hierzu werden Angaben aus dem Fragebogen für Lehrkräfte genutzt, die in allen Teilnehmerstaaten und -regionen von IGLU 2016 erfasst wurden. Im zweiten Teil des Kapitels wird die Gestaltung des Deutschunterrichts am Ende der vierten Klasse in Deutschland fokussiert. In den Kontextfragebögen, die den Schülerinnen und Schülern sowie den Lehrkräften vorgelegt wurden, die in Deutschland an IGLU 2016 teilgenommen haben, finden sich Items, mit denen Unterrichtsmerkmale eingeschätzt werden können, die den Dimensionen „Unterrichtsführung“, „Sozialklima“ und „kognitive Aktivierung“ zuzuordnen sind. Diese drei „Grunddimensionen der Unterrichtsqualität“ (Klieme, Schümer & Knoll, 2001, S. 51) gelten für die Entwicklung von Leistung, Lernmotivation und fachspezifischem Interesse als bedeutsame Größen. Daher sind sie in der jüngeren Vergangenheit zunehmend Gegenstand von Studien zur Unterrichtsqualität in der Grundschule geworden (Fauth, Decristan, Rieser, Klieme & Büttner, 2014a, 2014b; Gabriel, 2014; Lotz, 2016; Rieser, Stahns, Walzebug & Wendt, 2016) (siehe Abschnitte 4, 5 und 6). Im vorliegenden Kapitel wird berichtet, wie sich der Deutschunterricht am Ende der Grundschulzeit in Deutschland aus Sicht von Schülerinnen und Schülern sowie Lehrkräften hinsichtlich der drei Dimensionen darstellt (siehe Abschnitt 7.1). Zudem wird untersucht, ob die Einschätzung der Unterrichtsqualität aus Schülersicht mit dem Leistungsstand der Schülerinnen und Schüler im Lesen und mit der Lesemotivation zusammenhängt (siehe Abschnitt 7.2). Die Daten, die die Grundlage dieses Teils des Kapitels darstellen, stammen aus einer nationalen Ergänzung der Kontextfragebögen von IGLU – eine Einordnung der Befunde aus Deutschland im internationalen Vergleich ist deswegen nicht möglich.

2 Unterrichtsquantität und -qualität als Variablen in einem Angebots-Nutzungs-Modell des Unterrichts

In dem in Kapitel 2 vorgestellten Rahmenmodell der Studie IGLU 2016 werden Annahmen darüber zusammengefasst, in welchem Zusammenhang Leistungen im Lesen und zentrale Kontextbedingungen stehen. Zu diesen Faktoren gehört auch der Unterricht. Um die diesem Rahmenmodell zugrunde liegenden Annahmen über Merkmale erfolgreichen Unterrichts und deren Auswirkungen auf Lernprozesse zu konkretisieren, lohnt die Berücksichtigung sogenannter „Prozess-Mediator-Produkt-Modelle“, die den theoretischen Rahmen jüngerer Untersuchungen zur Qualität des Unterrichts darstellen. Das „Angebots-Nutzungs-Modell“ von Helmke (z.B. Helmke, 2015, S. 69 ff.) ist eines der populärsten dieser Modelle.

Die lerntheoretische Grundlage von Helmkes Modell stellen gemäßigt konstruktivistische Vorstellungen des Lernprozesses dar. Demnach stellt schulisches Lernen einen domänenspezifischen, kumulativen, eigenaktiven, individuell verlaufenden, aber häufig im Austausch mit anderen stattfindenden und institutionell verankerten Konstruktionsprozess dar, der in hohem Maße von der Lernausgangslage der Schülerinnen und Schüler beeinflusst wird – dieser Prozess kann beziehungsweise muss von Lehrkräften unterstützt werden (z.B. Lipowsky, 2009, S. 78 ff.; Helmke, 2015, S. 59 ff.). Wie in dem Rahmenmodell von IGLU werden im Angebots-Nutzungs-Modell auf einer Vielzahl von Ebenen Variablen berücksichtigt, von denen davon auszugehen ist, dass sie für den Erfolg schulischer Lernprozesse von Bedeutung sind (Individual-

ebene, Klassenebene, Unterrichtsebene, Schulebene, Merkmale der Familien der Schülerinnen und Schüler, Merkmale der Lehrperson) (vgl. dazu und zu den folgenden Ausführungen Helmke, 2015, S. 71 ff.). Den gemäßigt konstruktivistischen Vorstellungen des Lernprozesses entsprechend, denen das Modell verpflichtet ist, wird angenommen, dass Lehrkräfte Lernerfolge von Schülerinnen und Schülern nicht unmittelbar herbeiführen können, sondern lediglich die Lernprozesse unterstützen können, indem sie eine möglichst lernwirksame Umgebung schaffen. Der Unterricht stellt eine Lerngelegenheit dar, die von den Schülerinnen und Schülern genutzt werden muss, damit sich Lernerfolge einstellen. Die erfolgreiche Nutzung des Unterrichtsangebots auf Seiten der Schülerinnen und Schüler hängt demnach von „*Mediationsprozessen*“ (ebd., S. 71) ab. Von Bedeutung seien, so Helmke, zunächst die Wahrnehmung und Interpretation des Unterrichtsgeschehens durch die Schülerinnen und Schüler. In der Folge sei entscheidend, was dadurch an „motivationalen, emotionalen und volitionalen [...] Prozessen“ (ebd.) bei den Schülerinnen und Schülern ausgelöst werde. Das Ergebnis der Mediationsprozesse sind im besten Fall Lernaktivitäten auf Schülerseite und daraus resultierend nachhaltige Lernerfolge. Die Nutzung des Lernangebots wird potentiell durch eine Reihe von Variablen auf der Individual- oder Klassenebene beeinflusst, die von Lehrkräften nicht verändert werden können (z.B. die kognitiven Grundfähigkeiten einzelner Schülerinnen und Schüler oder die Klassenzusammensetzung). Auf die Unterrichtsgestaltung hingegen haben Lehrkräfte Einfluss. Im Unterricht können sie Anpassungen an die Lernvoraussetzungen der Schülerinnen und Schüler vornehmen, die die Wahrscheinlichkeit erhöhen, dass das Lernangebot genutzt wird.

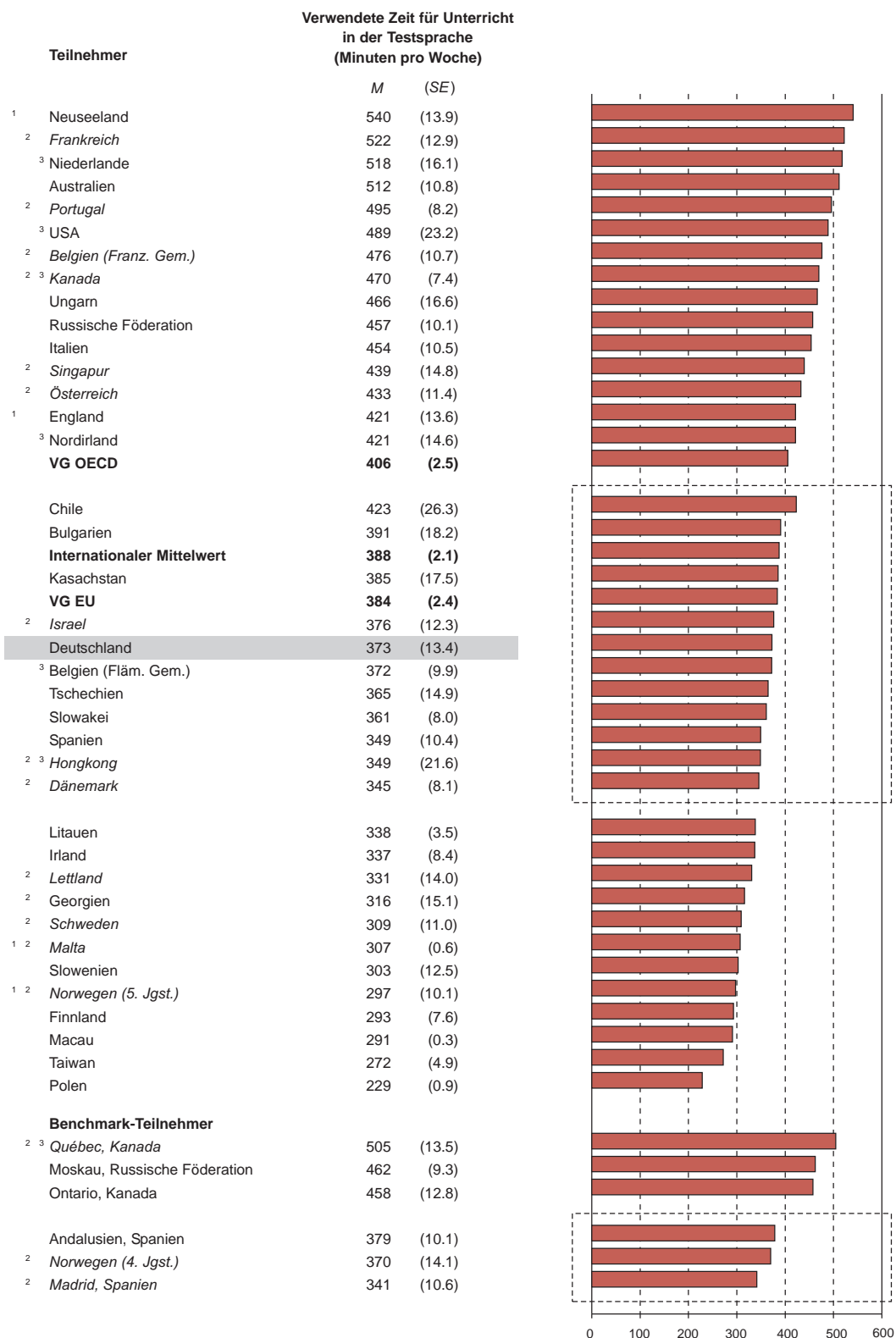
Um Lernprozesse im Unterricht zu unterstützen, empfiehlt Helmke die Orientierung an empirisch abgesicherten Qualitätsmerkmalen des Unterrichts – zum Beispiel denen, die er in einem eigenen Katalog vorstellt (z.B. „Klassenführung“ oder „Aktivierung“; Helmke, 2015, S. 168 f.). Die dort berücksichtigten Merkmale lassen sich zu einem großen Teil den drei oben genannten Basisdimensionen der Unterrichtsqualität zuordnen. Neben Merkmalen der „Prozessqualität des Unterrichts“ (ebd., S. 76) ist die „Unterrichtsquantität“ (ebd.) eine der im Angebots-Nutzungs-Modell berücksichtigten Variablen, denn eine Voraussetzung für Lernerfolge ist, dass überhaupt ein Unterrichtsangebot bereitgestellt wird.

Im Folgenden werden zunächst einige Befunde zur Quantität des Deutsch- beziehungsweise Sprachunterrichts am Ende der Grundschulzeit in Deutschland im internationalen Vergleich vorgestellt, bevor auf nationaler Ebene der Fokus auf Aspekte der Unterrichtsqualität gerichtet wird.

3 Unterrichtszeit für Deutsch- beziehungsweise Sprachunterricht im internationalen Vergleich

Zur differenzierten Beschreibung der Unterrichtszeit unterscheidet Helmke in Anlehnung an Treiber (1982) unter anderem „*nominale Unterrichtszeit*“ und „*tatsächliche Unterrichtszeit*“ (Helmke, 2015, S. 78). Die nominale Unterrichtszeit ist die für den Unterricht in einem Fach oder Lernbereich in einem bestimmten Zeitraum zum Beispiel in Curricula vorgegebene Zeit (ebd.). Davon zu unterscheiden ist die tatsächliche Unterrichtszeit, die in einem Fach oder Lernbereich in einem bestimmten Zeitraum realisiert wird (ebd.). Die Datenerhebungen im

Abbildung 9.1: Unterrichtszeit für die Testsprache als Fach beziehungsweise für Sprachunterricht in der Testsprache im internationalen Vergleich (Angaben der Lehrkräfte in Minuten pro Woche)



Kein statistisch signifikanter Unterschied zum Differenzwert von Deutschland ($p > .05$).

Statistisch signifikante Abweichung von $M = 0$ ($p < .05$).

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

Rahmen von IGLU 2016 ermöglichen Aussagen über die Unterrichtszeit, die nach Angaben der Lehrkräfte pro Woche im Durchschnitt für den Deutsch- beziehungsweise Sprachunterricht aufgebracht wird (Frage: *Wie viel Zeit verwenden Sie in einer normalen Schulwoche für Deutsch-/Sprachunterricht und/oder sprachbezogene Unterrichtsaktivitäten mit den Schülern?*) Die Angaben waren in Minuten pro Woche zu machen. Es handelt sich dabei um Angaben über die tatsächliche Unterrichtszeit im Sinne Treibers und Helmkes.

Im Durchschnitt werden in Deutschland laut Auskünften der Lehrkräfte in der Woche 373 Minuten der Unterrichtszeit mit Deutsch- beziehungsweise Sprachunterricht zugebracht (siehe Abbildung 9.1). Damit liegt die Unterrichtszeit in der Woche in Deutschland unter dem Wert der wichtigen Vergleichsgruppen (VG OECD: 406 Minuten, VG EU: 384 Minuten) und unter dem internationalen Mittelwert der Teilnehmerländer von IGLU 2016 (388 Minuten). Allerdings erreicht davon nur der Unterschied zu dem Mittelwert der Vergleichsgruppenländer der OECD die Signifikanzgrenze. In einigen Nachbarländern Deutschlands wird mit Sprachunterricht signifikant mehr Unterrichtszeit verbracht, vor allem in Frankreich (522 Minuten) und den Niederlanden (518 Minuten). Dagegen fällt der Unterschied zu anderen Nachbarländern wie Dänemark (345 Minuten) oder Tschechien (365 Minuten) nicht signifikant aus – dort wird laut Auskünften der Lehrkräfte weniger Zeit mit Sprachunterricht verbracht als in Deutschland.

Aufgrund der Ergebnisse empirischer Studien folgert Helmke, dass die tatsächliche Unterrichtszeit und der Leistungszuwachs von Schülerinnen und Schülern statistisch linear positiv zusammenhängen, „doch von einer bestimmten Zone an führt zusätzliche Unterrichtszeit nur noch zu minimalen Verbesserungen auf der Leistungsseite“ (Helmke, 2015, S. 79).

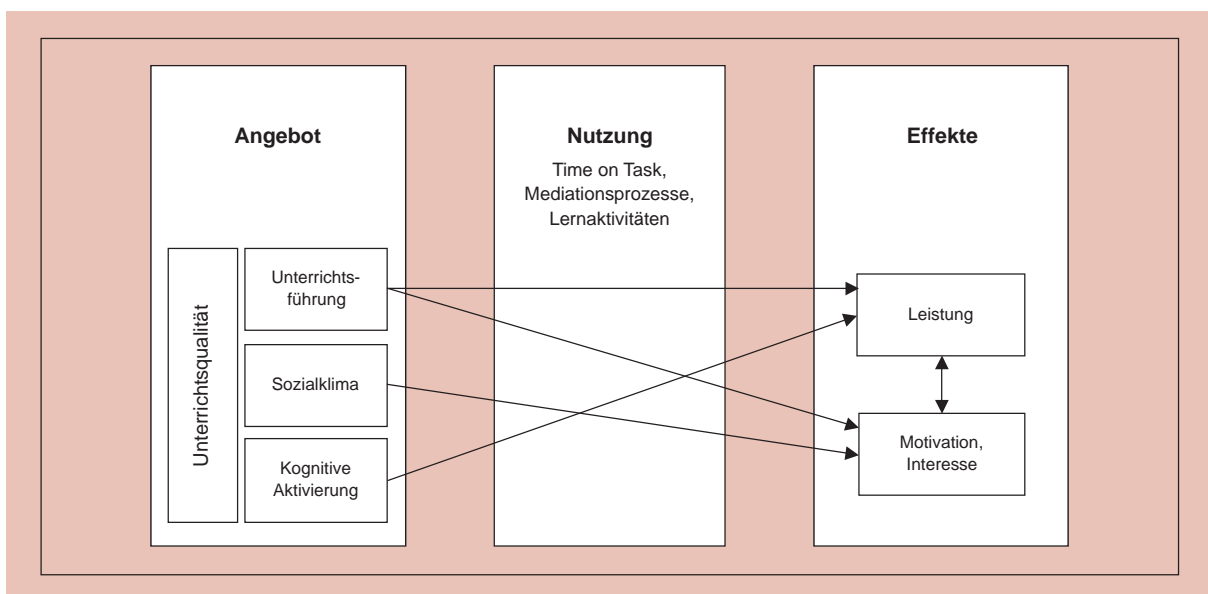
Unterrichtszeit – im Sinne der tatsächlichen Unterrichtszeit – mit Deutsch- beziehungsweise Sprachunterricht zu verbringen, kann wohl nicht als hinreichende Bedingung für die Entwicklung von Lesekompetenzen betrachtet werden. Von der nominalen und der tatsächlichen Unterrichtszeit grenzt Helmke in Anlehnung an Treiber (1982) unter anderem „*nutzbare Instruktionszeit*“ (Helmke, 2015, S. 78) und „*aktive Lernzeit*“ (ebd.) ab. Nutzbare Instruktionszeit ist die Zeit, die tatsächlich mit dem Unterrichten von Fachinhalten zugebracht wird – und nicht zum Beispiel mit organisatorischen Aspekten (ebd.). Die aktive Lernzeit (*time on task*) ist die Unterrichtszeit, die von den Schülerinnen und Schülern aktiv für die Auseinandersetzung mit Lerninhalten genutzt wird (ebd., S. 78 f.). Mit der nutzbaren Instruktionszeit und der aktiven Lernzeit ist der Übergang zwischen Unterrichtsquantität und Unterrichtsqualität markiert: Maßnahmen, die die nutzbare Instruktionszeit erhöhen, sind vor allem der Qualitätsdimension „Unterrichtsführung“ zuzuordnen, Maßnahmen, die die aktive Lernzeit erhöhen, den Dimensionen „Unterrichtsführung“ und „kognitive Aktivierung“.

4 Merkmale lernwirksamen Unterrichts: Befunde zur Unterrichtsqualität

Im Folgenden wird der Forschungsstand zu einigen Qualitätsmerkmalen des Unterrichts dargestellt. Da der Deutschunterricht in der Grundschule im Rahmen quantitativer Studien erst seit wenigen Jahren intensiver empirisch erforscht wird, erfolgt die Zusammenfassung unter Rückgriff auf Befunde zu anderen Fächern und Klassenstufen, die die Grundlage für die Auseinandersetzung mit dem Deutschunterricht in der Grundschule in Studien aus der jüngeren Vergangenheit darstellen. Da nicht davon auszugehen ist, dass die Qualität des Unterrichts in unterschiedlichen Fächern ohne Berücksichtigung von fachspezifischen Besonderheiten zu beurteilen ist, werden im nächsten Abschnitt (siehe Abschnitt 5) einige Forschungsergebnisse zum Deutschunterricht in der Grundschule vorgestellt.

Die Ergebnisse zahlreicher Studien zum Unterricht in der Sekundarstufe I und der Grundschule haben dazu beigetragen, dass sich im deutschsprachigen Raum ein Modell der Unterrichtsqualität etabliert hat, in dem die drei oben genannten Basisdimensionen angesetzt werden (z. B. Klieme et al., 2001; Klieme & Rakoczy, 2003; Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Helmke & Klieme, 2008; Klieme et al., 2008; Klieme, Steinert & Hochweber, 2010; Decristan et al., 2015; Rieser et al., 2016). Klieme et al. (2006) bezeichnen diese Dimensionen als „strukturierte, klare und störungspräventive Unterrichtsführung“, „unterstützendes, schülerorientiertes Sozialklima“ und „kognitive Aktivierung“ (ebd., S. 131). In Abbildung 9.2 wird dargestellt, von welchen Zusammenhängen zwischen den drei Basisdimensionen und Lernerträgen auf Schülerseite ausgegangen werden kann.

Abbildung 9.2: Die drei Basisdimensionen der Unterrichtsqualität (Darstellung in Anlehnung an Klieme et al., 2006, S. 131 und Klieme & Rakoczy, 2008, S. 228)



Für die Lernwirksamkeit des Unterrichts dürfte „Klassenführung“ das bedeutendste Merkmal der Dimension „klare, strukturierte und störungspräventive Unterrichtsführung“ sein. Für das vorliegende Kapitel werden unter der Bezeichnung „Klassenführung“ Maßnahmen zusammengefasst, die dazu dienen, Unterrichtsverzögerungen oder Störungen zu verhindern oder ihnen zu begegnen. Das Ziel ist, dass sich die Lernenden über einen möglichst großen Anteil der Unterrichtszeit mit den Unterrichtsinhalten auseinandersetzen (Helmke, 2015, S. 173). Es kann erwartet werden, dass eine effiziente Klassenführung den Anteil der nutzbaren Instruktionszeit erhöht und Auswirkungen auf die aktive Lernzeit der Schülerinnen und Schüler hat. Effiziente Klassenführung führt zu einem weitgehend störungsfreien Unterrichtsverlauf ohne oder nur mit kurzen Phasen, in denen die Schülerinnen und Schüler sich nicht mit den Fachinhalten auseinandersetzen (können oder müssen) – sei es, weil Störungen auftreten, denen nicht effizient begegnet wird, sei es, weil sich der Unterrichtsverlauf durch Mängel bei der Planung oder Unklarheiten bei den organisatorischen Abläufen verzögert. Neben dem Merkmal „Klassenführung“ ist auch das Merkmal „Strukturiertheit“ der Dimension „Unterrichtsführung“ zuzuordnen. Lipowsky (2009) weist unter anderem auf eine kognitionspsychologische Fassung von Strukturiertheit hin, die für die Einordnung der im Folgenden präsentierten Ergebnisse relevant ist: Diese Art von Strukturiertheit im Unterricht sei von Bedeutung für die Verbindung des Vorwissens mit dem zu erwerbenden Wissen auf Seiten der Schülerinnen und Schüler. Auch trage Strukturiertheit in diesem Sinne dazu bei, „den Aufbau einer komplexen und geordneten Wissensstruktur beim Lernenden zu erleichtern“ (ebd., S. 83). Wichtig dafür sei unter anderem die inhaltliche Kohärenz des Unterrichts, die den Lernenden auch explizit vermittelt werde, „indem Zusammenhänge zwischen verschiedenen Aspekten des Unterrichtsinhalts hergestellt werden, indem die Übersicht und Einordnung neuer Informationen zum Beispiel mittels *Advance Organizers* erleichtert wird und wichtige Unterrichtsergebnisse zusammengefasst werden“ (ebd.).

Der Zusammenhang zwischen einer effizienten Klassenführung und der Leistungsentwicklung dürfte einer der am besten dokumentierten in der Unterrichtsforschung sein (zusammenfassend Lipowsky, 2009; Hattie, 2015; Helmke, 2015). Auf den Unterricht in der Grundschule beziehen sich Ergebnisse der Studie SCHOLASTIK (Helmke & Weinert, 1997): Den Unterricht erfolgreicher Grundschulklassen – das Kriterium dafür ist die Leistungsentwicklung in Mathematik – zeichnet aus Beobachtersicht unter anderem eine effiziente Klassenführung und Zeitnutzung aus. Zudem ist die Strukturiertheit der Lehrerausführungen für den Mathematikunterricht in der Grundschule eine wichtige Größe. Für den Sachunterricht in der Grundschule liegen Befunde aus der Studie *Individuelle Förderung und adaptive Lern-Gelegenheiten in der Grundschule* (IGEL) vor: Eine effiziente Klassenführung aus Schüler-, Lehrkraft- und Beobachtersicht hängt den Studienergebnissen zufolge mit der Leistungsentwicklung zusammen (Fauth et al., 2014a, 2014b). Eine Auswertung der Schülerbefragungen, die im Rahmen der *Trends in International Mathematics and Science Study* (TIMSS 2015) durchgeführt wurden, zeigt sowohl für den Mathematik- als auch für den Sachunterricht, dass leistungstärkere Schülerinnen und Schüler die Klassenführung als effizienter einschätzen als leistungsschwächere Schülerinnen und Schüler (Rieser et al., 2016).

Im Mittelpunkt der Diskussion um die Dimension „unterstützendes, schülerorientiertes Sozialklima“ steht der Zusammenhang zwischen Unterrichtsgestaltung und Lehrerverhalten einerseits sowie der Entwicklung der Lernmotivation

und des Interesses der Schülerinnen und Schüler andererseits. Motivation und Interesse sind nicht nur Unterrichtsziele eigenen Werts, sondern können als Voraussetzungen erfolgreicher Lernprozesse gesehen werden. In den Arbeiten zur Unterrichtsqualität, in denen Unterrichtsaspekte aus dieser Dimension untersucht werden, bezieht man sich meist auf die Selbstbestimmungstheorie nach Deci und Ryan (1993). Um die Entwicklung der Motivation zu fördern, müssen den Lernenden im Unterricht dieser Theorie zufolge Autonomie- und Kompetenzerleben sowie das Erfahren sozialer Einbindung ermöglicht werden. Operationalisiert wird ein positives, unterstützendes Sozialklima, das das ermöglicht, zum Beispiel durch die Beurteilung der „Fehlerkultur“ im Unterricht, die Art der Lehrer-Schüler- beziehungsweise der Schüler-Schüler-Beziehung oder das Lob durch die Lehrkraft für die von Schülerinnen und Schülern gezeigten Leistungen (Lotz, 2016, S. 85).

Merkmale der Dimension „Sozialklima“ können auf Grundlage des aktuellen Forschungsstands als zentral für die Entwicklung der Lernmotivation und des Interesses angesehen werden (z. B. Klieme et al., 2001, 2006; König et al., 2011; Kunter & Voss, 2011). Ergebnisse von Fauth et al. (2014a), die auf der Grundlage von Schülerbefragungen gewonnen wurden, legen nahe, dass der Zusammenhang zwischen Merkmalen, die der zweiten Grunddimension zuzuordnen sind, und der Entwicklung des Interesses auch für den Sachunterricht in der Grundschule bestehen dürfte. Ein Zusammenhang mit der Leistung kann theoretisch ebenfalls hergestellt werden, allerdings sind die Ergebnisse weniger eindeutig: Fauth et al. (2014b, S. 133) gelingt es, einen signifikanten Zusammenhang zwischen Beobachtereinschätzungen eines unterstützenden Sozialklimas und der Leistungsentwicklung nachzuweisen. Hingegen können Rieser et al. (2016) auf Grundlage der Daten aus TIMSS 2015 weder für den Mathematik- noch für den Sachunterricht eine Beziehung zwischen der Leistung der Schülerinnen und Schüler und ihrer Einschätzung des Sozialklimas aufzeigen.

Die Diskussion um die kognitive Aktivierung als Qualitätsdimension des Unterrichts gründet auf konstruktivistischen Vorstellungen des Lernprozesses. Mayer (2004) hat sich mit der Forderung nach Schüleraktivität als notwendiger Voraussetzung für Lernerfolge auseinandergesetzt und die Unterscheidung von „behavioral activity“ und „cognitive activity“ (ebd., S. 14) vorgenommen. Erfolgreiches Lernen, so Mayer, setze kognitive Aktivität voraus, nicht aber zwangsläufig „physische“ und „beobachtbare“ Aktivität. Das Potential des Unterrichts, die kognitive Aktivität anzuregen, steht im Mittelpunkt der Diskussion um die Qualitätsdimension „kognitive Aktivierung“. Helmke (2015, S. 205 f.) bezeichnet als kognitive Aktivierung eine tiefgehende Verarbeitung von Informationen durch den Einsatz anspruchsvoller Lernstrategien, die quantitativ und qualitativ zum Lernfortschritt von Lernenden beiträgt. Dafür werden Prozesse des Aufbaus und der Modifikation (elaborative Prozesse) und Prozesse der Organisation und Strukturierung von Informationen und des Wissens (organisierende Prozesse) als zentral angesehen.

Lipowsky (2009, S. 93) weist darauf hin, dass das Ausmaß der kognitiven Aktivität von Schülerinnen und Schülern nicht unmittelbar zu beobachten ist. Um es einzuschätzen, seien Indikatoren zu bestimmen. Kognitiv aktivierender Unterricht zeichnet sich zum Beispiel aus durch

anspruchsvolle Aufgabenstellungen, herausfordernde Fragen, Konfrontation der Lernenden mit scheinbaren Widersprüchen oder durch Klassengespräche, während derer die Schüler(innen) über den Unter-

richtsgegenstand nachdenken und sprechen, Argumente austauschen und von der Lehrperson nicht sofort eine Rückmeldung zu ihren Antworten erhalten, sondern hin und wieder ihre Fehler selbst bemerken sollen (Faust, Lipowsky & Gleich, 2011, S. 50).

Das Lehrerverhalten dürfte maßgeblich für das Potential zur kognitiven Aktivierung des Unterrichtsangebots sein – dem steht Schülerverhalten auf der Nutzungsseite des Unterrichts gegenüber (Lipowsky, 2009). Hinweise auf die kognitive Aktivierung beziehungsweise die kognitiven Aktivitäten von Lernenden können darin gesehen werden, dass sie

z.B. Argumente austauschen, Querverbindungen zu anderen Themen oder Konzepten herstellen, Lösungswege erläutern, vergleichen und beurteilen, Vermutungen formulieren, Fragen stellen, Antworten und Lösungen hinterfragen und ihr Wissen auf andere Situationen übertragen (ebd., S. 93).

Die kognitive Aktivierung hat sich in einer Reihe von Untersuchungen zum Unterricht in der Sekundarstufe I als Prädiktor für den Leistungsstand und/oder die Entwicklung der fachlichen Lernleistungen erwiesen (z.B. Klieme et al., 2001, 2006; Klieme & Rakoczy, 2008; Baumert & Kunter, 2011; Kunter & Voss, 2011). In der Studie IGEL zeigt sich ein Zusammenhang zwischen der Entwicklung des Interesses und den Einschätzungen der kognitiven Aktivierung des Sachunterrichts durch Drittklässlerinnen und Drittklässler (Fauth et al., 2014a, S. 4 ff.), ein Zusammenhang zwischen der kognitiven Aktivierung und der Entwicklung der fachlichen Leistung kann hier jedoch nicht nachgewiesen werden (ebd.). Fauth et al. (2014b) weisen allerdings einen positiven Zusammenhang von kognitiver Aktivierung aus Beobachterperspektive und der Leistung nach. Rieser et al. (2016) finden in den Daten von TIMSS 2015 einen negativen Zusammenhang zwischen der wahrgenommenen kognitiven Aktivierung und dem Leistungsstand von Schülerinnen und Schülern: Leistungsstärkere nehmen sich sowohl im Mathematik- als auch im Sachunterricht weniger stark kognitiv aktiviert wahr als Leistungsschwächere. Möglicherweise gelingt es im Mathematik- und Sachunterricht nicht, die Leistungsstärkeren ihren Möglichkeiten entsprechend kognitiv herauszufordern.

5 Befunde zur Qualität des Deutschunterrichts in der Grundschule

„Unterrichtsführung“, „Sozialklima“ und „kognitive Aktivierung“ sind aufgrund der Ergebnisse zahlreicher Studien als Qualitätsdimensionen des Unterrichts etabliert. Die Befunde zum Zusammenhang von Merkmalen der Dimensionen „Unterrichtsführung“ und „Sozialklima“ sowie der Leistung, der Motivation und des Interesses können wohl auf den Deutsch- und Leseunterricht in der Grundschule übertragen werden, auch wenn in den meisten der oben genannten Studien naturwissenschaftlicher Unterricht und Mathematikunterricht in der Sekundarstufe I analysiert wurden. Diese beiden Dimensionen dürften fach- und klassenstufenübergreifend für die Qualität des Unterrichts von Bedeutung sein. Hinsichtlich der kognitiven Aktivierung stellt sich die Situation anders dar: In

zahlreichen aktuellen Studien werden bei der Operationalisierung der kognitiven Aktivierung fachspezifische Merkmale berücksichtigt – das erschwert die Übertragung der Ergebnisse dieser Studien auf den Deutschunterricht. Zudem dürfte es auch Unterschiede zwischen dem Deutschunterricht in der Grundschule und dem in den weiterführenden Schulen geben: Kognitive Aktivierung im Leseunterricht in der Grundschule (vgl. Lotz, 2016) wird auf andere Weise zu gewährleisten sein als kognitive Aktivierung im Literaturunterricht in der Sekundarstufe (vgl. Winkler & Steinmetz, 2016).

Zur Ergänzung der Zusammenfassung des Forschungsstands werden im Folgenden einige Forschungsergebnisse zur Qualität des Deutschunterrichts in der Grundschule vorgestellt. Die meisten der vorliegenden Studien zum Deutschunterricht in der Grundschule beziehen sich auf den Leseunterricht. Dazu gehören die Studien *Persönlichkeits- und Lernentwicklung von Grundschulkindern* (PERLE) (Lotz, 2016) und *VERA – Gute Unterrichtspraxis* (Kleinbub, 2010) sowie Arbeiten, die im Rahmen von IGLU entstanden sind (Lankes, 2004; Lankes & Carstensen, 2007; Tarelli, Lankes, Drossel & Gegenfurtner, 2012; Nortvedt, Gustafsson & Lehre, 2016; Stahns & Rieser, 2018). Im Folgenden werden exemplarisch für den Deutschunterricht in der Grundschule einige Befunde zum Leseunterricht zusammenfassend dargestellt. Um ein breites Bild des Forschungsstands zum Deutschunterricht zu ermöglichen, werden neben den Ergebnissen von Fragebogenerhebungen auch Ergebnisse von Videostudien einbezogen, auch wenn im Rahmen von IGLU 2016 keine Unterrichtsaufzeichnungen durchgeführt wurden.

In der Studie *VERA – Gute Unterrichtspraxis* wurden Videoaufzeichnungen des Leseunterrichts in vierten Klassen der Grundschule durchgeführt und Lesekompetenzen zu zwei Messzeitpunkten erhoben (Kleinbub, 2010; Heyne, 2014). Kleinbub hat eine Untersuchung vorgelegt, in der der videographierte Unterricht unter anderem hinsichtlich der Merkmale „Kompetenzorientierung“ und „Aktivierung“ durch Rater eingeschätzt wurde (Kleinbub, 2016, S. 17 ff.). Mit dem Merkmal „Kompetenzorientierung“ werden Aspekte des Unterrichtsgeschehens beurteilt, die mit der kognitiven Aktivierung der Schülerinnen und Schüler in Verbindung stehen dürften, zum Beispiel die Art der Vermittlung von Lesestrategien. Von der „Kompetenzorientierung“ grenzt Kleinbub das Merkmal „Aktivierung“ ab: Hier steht die „Aktivierung im Sinne einer aktiven Teilhabe der Schülerinnen und Schüler am Unterrichtsgeschehen“ (ebd., S. 19) im Mittelpunkt. „Aktivierung“ ist das einzige in der Studie berücksichtigte Merkmal, für das ein signifikanter Zusammenhang mit der Leistungsentwicklung nachgewiesen werden kann (Kleinbub, 2016, S. 29). Kleinbub berücksichtigt keine Merkmale, die sich den Dimensionen „Unterrichtsführung“ und „Sozialklima“ zuordnen lassen (ebd., S. 17).

Auf Grundlage der Ergebnisse einer Videostudie, die im Rahmen der Studie PERLE durchgeführt wurde, können in einer Stichprobe erster Klassen die Merkmale „Klassenführung“, „Unterrichtsklima“ sowie „kognitive Aktivierung“ im Leseanfangsunterricht beurteilt werden (Gabriel, 2014; Lotz, 2016). Für die Einschätzung der kognitiven Aktivierung durch Rater wurden unter anderem die Art und das kognitive Niveau der Lehrerfragen und Aspekte des Lehrerfeedbacks zu Schüleräußerungen sowie die Vermittlung von Lesestrategien in Leseübungen erfasst (Lotz, 2016, S. 75 ff.). Aspekte der Klassenführung und des Unterrichtsklimas werden in den videographierten Unterrichtseinheiten durch die Rater insgesamt positiv bewertet (Gabriel, 2014). In einer Arbeit von Lotz (2016) wird die Ausprägung der kognitiven Aktivierung in Unterrichtsphasen beurteilt, in denen

in den aufgezeichneten Unterrichtseinheiten eine Leseübung realisiert wurde. Die Einschätzung fiel weniger positiv aus als die des Unterrichtsklimas und der Klassenführung. Allerdings bleibt die Frage offen, wie die Ergebnisse für die Entwicklung der Leseleistung zu bewerten sind. Der Zusammenhang zwischen der Ausprägung der kognitiven Aktivierung während der Leseübung und den Lesekompetenzen der Schülerinnen und Schüler wurde nicht ermittelt.

Exemplarisch für Arbeiten auf der Grundlage der Befragung von Lehrkräften stehen Befunde zum Unterricht aus IGLU 2011. Der Fragebogen für Lehrkräfte von IGLU 2011 enthielt Fragen zur Unterrichtsgestaltung, die unter anderem Aufschluss über das „kognitive Anregungspotential des Unterrichts“ (Tarelli et al., 2012, S. 157) erlauben. Die Schülerinnen und Schüler kamen nach Einschätzung der Lehrkräfte im Unterricht relativ häufig dazu, sich über Gelesenes auszutauschen und über die Texte nachzudenken. Wo es um die Vermittlung von Lesestrategien und von für das Textverständnis relevantem Wissen geht, zeigt sich im Vergleich zu einer Reihe anderer europäischer Staaten Entwicklungspotential (ebd., S. 158). Stahns und Rieser (2018) untersuchen auf Grundlage dieser Daten Zusammenhänge zwischen der Unterrichtsgestaltung und der Lesekompetenz. Unter anderem kann eine negative Korrelation zwischen der Einschätzung der Aspekte kognitiv aktivierenden Unterrichts und der Lesekompetenz der Schülerinnen und Schüler nachgewiesen werden: Je häufiger die dieser Dimension zugeordneten Aspekte laut Auskunft der Lehrkräfte im Unterricht realisiert wurden (z.B. die Instruktion von Lesestrategien), desto geringer war die Lesekompetenz der Schülerinnen und Schüler. Stahns und Rieser gehen davon aus, dass das unter anderem damit zusammenhängen könnte, dass das tatsächliche Potential kognitiver Aktivierung mit den in IGLU 2011 zur Verfügung stehenden Fragebogen-Items nicht optimal zu ermitteln ist.

Ebenfalls auf Grundlage von Daten der Studie IGLU/PIRLS 2011 untersuchen Nortvedt et al. (2016) den Zusammenhang zwischen Unterrichtsgestaltung, Lesekompetenz und Mathematikkompetenz. Angenommen wird, dass die Instruktionsqualität sich moderierend auf den Zusammenhang von Leseverstehen und Mathematikkompetenz auswirkt (ebd., S. 99). Dabei wird vor allem der Zusammenhang zwischen einem (kognitiv herausfordernden) Erwerb von Lesestrategien im Unterricht und der Mathematikkompetenz betont – für die Bearbeitung von textbasierten Mathematikaufgaben sei der Einsatz von Lesestrategien notwendig, der im Unterricht erlernt werde (ebd., S. 99 ff.). Um nachzuweisen, dass sich die Instruktionsqualität moderierend auf den Zusammenhang von Lesekompetenz und Mathematikkompetenz auswirkt, wird unter anderem der Zusammenhang der wahrgenommenen Instruktionsqualität im Lese- und Mathematikunterricht und den Lese- und Mathematikkompetenzen berechnet. Um die Instruktionsqualität abzubilden, werden Items aus dem Schülerfragebogen von IGLU eingesetzt, zum Beispiel *„Ich weiß, was meine Lehrerin/mein Lehrer von mir erwartet.“* (Wendt, Bos, Tarelli, Vaskova, & Walzebug, 2016, S. 61) Zwischen der Einschätzung des Unterrichts und der Lesekompetenz zeigt sich für Deutschland ein nicht signifikanter negativer Zusammenhang (Nortvedt et al., 2016, S. 106 f.). Dieses Ergebnis könnte mit der Auswahl der Items zu tun haben, mit denen die Instruktionsqualität erfasst wurde: Zwar wird die Bedeutung der Vermittlung von Lesestrategien betont, erfasst werden konnte dieser Aspekt der Unterrichtsgestaltung mit den eingesetzten Fragebogen-Items aber nicht (ebd., S. 109).

Was die Befunde zur Gestaltung des Deutschunterrichts in der Grundschule betrifft, so ist zu bemerken, dass nur in wenigen Studien die Gestaltung des gesamten Deutschunterrichts im Mittelpunkt steht. Zumeist werden einzelne Kompetenzbereiche wie Rechtschreiben oder Lesen fokussiert. Der Zusammenhang zwischen der Unterrichtsgestaltung und den Kompetenzen der Schülerinnen und Schüler wurde nur in einem Teil der genannten Studien betrachtet und nachgewiesen. Insbesondere hinsichtlich der kognitiven Aktivierung sind die Ergebnisse nicht durchgängig in Übereinstimmung mit den Erwartungen in Bezug auf die Lernwirksamkeit (siehe Abbildung 9.2). Unter anderem fehlt es an Nachweisen für (positive) Zusammenhänge zwischen fachlichen Leistungen und der kognitiven Aktivierung der Schülerinnen und Schüler. Von Bedeutung für die Einordnung der Ergebnisse ist, dass sie aus Studien stammen, in denen methodisch unterschiedliche Zugänge gewählt wurden. Mit der Wahl der Erhebungsinstrumente und der Perspektive, aus der Daten erhoben werden (Lernende, Lehrende oder externe Beobachter), können Einschränkungen bezüglich der Validität und Objektivität der Einschätzungen einhergehen (Clausen, 2002, S. 45 ff.). Die Befunde zum Deutschunterricht in der Grundschule im deutschsprachigen Raum stammen aus Studien, in denen die Daten primär auf Grundlage von Videoaufzeichnungen des Unterrichts (z. B. Kleinbub, 2016; Lotz, 2016) oder Befragungen von Lehrkräften mit standardisierten Fragebögen (z. B. Tarelli et al., 2012; Nortvedt et al., 2016) gewonnen wurden. Die Ergebnisse einiger Studien zeigen, dass Einschätzungen desselben Unterrichts aus verschiedenen Perspektiven sehr unterschiedlich ausfallen können (Clausen, 2002; Fauth et al., 2014b), insbesondere hinsichtlich der kognitiven Aktivierung (ebd.). Zudem ist mit Einschätzungen aus den drei Perspektiven in unterschiedlichem Maße die Vorhersage von Lernergebnissen möglich (ebd.). Im Rahmen von IGLU 2016 können Einschätzungen der Qualität des Unterrichts aus Sicht von Schülerinnen und Schülern sowie Lehrkräften vorgenommen werden. Dieser Zugang ist kennzeichnend für zahlreiche *Large-Scale*-Studien (z. B. Klieme et al., 2008; Rieser et al., 2016).

6 Schüler- und Lehrkräfteangaben zum Deutschunterricht im Rahmen von IGLU 2016

Im Folgenden werden Ergebnisse der Befragung der Schülerinnen und Schüler sowie der Lehrkräfte vorgestellt, die in Deutschland an IGLU 2016 teilgenommen haben. Die Fragebögen enthalten Items, mit denen einige Aspekte der Unterrichtsführung (Klassenführung und Strukturiertheit), des Sozialklimas und der kognitiven Aktivierung im Deutschunterricht eingeschätzt werden können.

Um die Wahrnehmung des Unterrichts aus Sicht der Schülerinnen und Schüler zu erheben, wurde auf Items zurückgegriffen, deren Einsatz im Rahmen der Studie IGEL erprobt wurde (Fauth et al., 2014a) und die auch im Rahmen von TIMSS 2015 eingesetzt wurden (Rieser et al., 2016). Für den Einsatz in IGLU 2016 wurden die Items adaptiert und inhaltliche und sprachliche Anpassungen vorgenommen. Die Items, die in die Fragebögen für Lehrkräfte aufgenommen wurden, wurden in Anlehnung an die entsprechenden Items in den Schülerfragebögen formuliert. Der Fragebogen für Lehrkräfte enthält weniger Fragen zu den vier Qualitätsmerkmalen als der Schülerfragebogen, weil eine Adaption nicht bei allen Fragen aus dem Schülerfragebogen möglich war.

Zur Einschätzung der Klassenführung wird unter anderem erhoben, welche Angaben die Lernenden und Lehrenden zum Auftreten von Störungen oder Verzögerungen im Deutschunterricht machen (Beispielitem aus dem Schülerfragebogen: *„Im Deutschunterricht muss unsere Lehrerin/ unser Lehrer lange warten, bis alle Schüler leise sind.“* Beispielitem aus dem Fragebogen für Lehrkräfte: *„Im Deutschunterricht muss ich lange warten, bis alle Schüler leise sind.“*). Aus dem Schülerfragebogen wurden sechs Items verwendet, die sich auf Aspekte effizienter Klassenführung beziehen, und aus dem Fragebogen für Lehrkräfte fünf Items. Mit weiteren Fragen wurde die Einschätzung eines zweiten Merkmals der Dimension „Unterrichtsführung“ – die (inhaltliche) Strukturiertheit des Unterrichts – ermittelt. Die Items zur Einschätzung der Strukturiertheit beziehen sich auf Maßnahmen, die es erleichtern können, die inhaltliche Struktur des Deutschunterrichts nachzuvollziehen, indem der „rote Faden“ der Unterrichtsstunden verdeutlicht wird (Beispielitem aus dem Schülerfragebogen: *„Unsere Deutschlehrerin/ unser Deutschlehrer sagt uns, worauf es in der Unterrichtsstunde besonders ankommt.“* Beispielitem aus dem Fragebogen für Lehrkräfte: *„Im Deutschunterricht mache ich deutlich, worauf es in der Stunde besonders ankommt.“*). Der Schülerfragebogen und der Fragebogen für Lehrkräfte enthalten jeweils drei Items, die sich auf die Strukturiertheit beziehen. Die kognitive Aktivierung ist die Qualitätsdimension, bei deren Operationalisierung fachspezifische Aspekte des Unterrichts am stärksten Berücksichtigung finden können (siehe Abschnitt 5). Einige Unterrichtsmerkmale, die der kognitiven Aktivierung zugeordnet werden, sind fachübergreifend für eine tiefe Verarbeitung von Unterrichtsinhalten relevant (u. a. anspruchsvolle Fragen, der Umgang mit Schülerfehlern und sachlich-konstruktive Rückmeldungen), andere sind als fach- beziehungsweise domänenspezifisch anzusehen (z. B. die Vermittlung von Lesestrategien im Leseunterricht, der Einsatz von grammatischen Proben im Grammatikunterricht). Wegen der Unterscheidung von vier Lernbereichen im Deutschunterricht in der Grundschule kann es sinnvoll sein, nicht nur eine fachspezifische, sondern auch eine lernbereichsspezifische Operationalisierung der kognitiven Aktivierung vorzunehmen. Bei den im Folgenden vorgestellten Analysen wird allerdings nicht ein bestimmter Lernbereich fokussiert, sondern der Deutschunterricht im Allgemeinen. Daher werden für die Einschätzung der kognitiven Aktivierung Merkmale berücksichtigt, die in allen Lernbereichen des Deutschunterrichts für die kognitive Aktivierung der Schülerinnen und Schüler von Bedeutung sein dürften. Bei der Auswahl der Items wurde berücksichtigt, dass die Schülerinnen und Schüler vermutlich nicht über die notwendigen Kompetenzen verfügen, um fachliche und fachdidaktische Aspekte des Unterrichts differenziert beurteilen zu können. Die Einschätzung der kognitiven Aktivierung im Deutschunterricht wird über Urteile einiger Aspekte des Lehrerverhaltens und der Unterrichtsgestaltung vorgenommen, die unter anderem die Aktivierung von Vorwissen, den Lebensweltbezug der Lerninhalte und das Verständnis von Lerninhalten betreffen (Beispielitem aus dem Schülerfragebogen: *„Unsere Deutschlehrerin/ unser Deutschlehrer möchte, dass wir unsere Antworten erklären.“* Beispielitem aus dem Fragebogen für Lehrkräfte: *„Im Deutschunterricht bitte ich die Schüler oft, ihre Antworten zu erklären.“*). Aus dem Schülerfragebogen wurden sechs Items ausgewählt, um die Wahrnehmung der kognitiven Aktivierung einzuschätzen, aus dem Fragebogen für Lehrkräfte fünf. Über die Wahrnehmung des Sozialklimas im Deutschunterricht geben Fragen Auskunft, die sich unter anderem auf den sachlichen Umgang mit Schülerfehlern, Lob und weitere Aspekte der Schüler-Lehrer-

Beziehung beziehungsweise auf Aspekte der Interaktion im Klassenzimmer beziehen (Beispielitem aus dem Schülerfragebogen: *„Unsere Deutschlehrerin/ unser Deutschlehrer ist auch dann nett zu mir, wenn ich einen Fehler mache.“* Beispielitem aus dem Fragebogen für Lehrkräfte: *„Im Deutschunterricht versuche ich, den Schülern das Gefühl zu geben, dass Fehler nicht schlimm sind.“*). Aus dem Schülerfragebogen gehen sieben Items in die Analysen ein, aus dem Fragebogen für Lehrkräfte sechs. Das Antwortformat ist bei allen Items vierstufig. Erfragt werden entweder Häufigkeiten (von 1 = *nie* bis 4 = *in jeder Stunde*) oder der Grad der Zustimmung (von 1 = *stimme völlig zu* bis 4 = *stimme überhaupt nicht zu*).

7 Ergebnisse

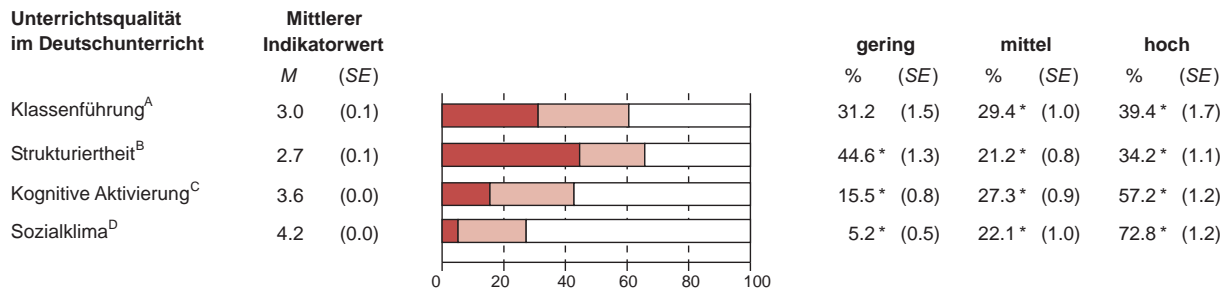
7.1 Einschätzung des Deutschunterrichts durch die Lernenden und die Lehrkräfte

In diesem Abschnitt wird berichtet, wie die Viertklässlerinnen und Viertklässler sowie die Lehrkräfte, die an IGLU 2016 teilgenommen haben, den Deutschunterricht hinsichtlich der Merkmale „Klassenführung“, „Strukturiertheit“, „kognitive Aktivierung“ und „Sozialklima“ einschätzen.

Für die Darstellung der Ergebnisse der Schüler- und Lehrerbefragungen werden Indikatoren gebildet. Wenn bei der Beantwortung der Items aus den Fragebögen eine Antwort gewählt wurde, die für eine positive Einschätzung der Unterrichtsqualität spricht, wird ein Punkt vergeben. Gibt ein Kind beispielsweise an, der Aussage *„Unsere Deutschlehrerin/unser Deutschlehrer ist auch dann nett zu mir, wenn ich einen Fehler mache“* völlig oder zumindest eher zuzustimmen, wird dafür ein Punkt vergeben. Dieses Verfahren wird für alle Items angewendet. Anschließend werden die Punkte der Items, die dasselbe Merkmal erfassen, zusammengerechnet (z. B. Sozialklima aus Schülerperspektive: 7 Items). Auf diese Weise ergeben sich – in Abhängigkeit von der Anzahl der Fragen zu den vier Merkmalen in den Fragebögen – Werte zwischen 0 und 7 Punkten. Je höher dieser Wert ist, desto positiver ist das jeweilige Merkmal ausgeprägt. Um die Werte aller Merkmale miteinander vergleichen zu können, findet anschließend eine Umrechnung in ein Format statt, in dem die Werte Ausprägungen zwischen 0 und 5 Punkten annehmen können. Inhaltlich können die Indikatoren so interpretiert werden, dass bei Punktwerten unter 2 von einer schwachen Ausprägung des jeweiligen Merkmals ausgegangen wird. Werte von 2 bis unter 4 Punkten werden als mittlere Ausprägungen interpretiert. Werte ab 4 Punkten stehen für starke Ausprägungen der Unterrichtsmerkmale.

In Abbildung 9.3 wird dargestellt, wie die Schülerinnen und Schüler die Qualität des Deutschunterrichts einschätzen.

Die Fragen, die sich auf die unterschiedlichen Aspekte der Klassenführung beziehen, beantworten etwa 39 Prozent der Schülerinnen und Schüler so, dass von einer sehr positiven Wahrnehmung der Klassenführung gesprochen werden kann. Allerdings erlauben die Angaben von etwas mehr als 31 Prozent der Lernenden die Einschätzung, dass die Klassenführung aus ihrer Perspektive wenig effizient erfolgt. Etwas weniger positiv als die Klassenführung wird die Strukturiertheit des Unterrichts bewertet. Die Einschätzungen von nur etwa 34 Prozent der Schülerinnen und Schüler lassen auf eine für die Lernenden

Abbildung 9.3: Durch Schülerinnen und Schüler wahrgenommene Unterrichtsqualität im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)


■ % der Schülerinnen und Schüler, die die Unterrichtsqualität als gering ausgeprägt einschätzen.

■ % der Schülerinnen und Schüler, die die Unterrichtsqualität als mittelstark ausgeprägt einschätzen.

□ % der Schülerinnen und Schüler, die die Unterrichtsqualität als hoch ausgeprägt einschätzen.

* = Unterschied in den Prozentwerten zwischen den Kategorien statistisch signifikant ($p < .05$).

Geprüft wurden die Kategorien gering gegen mittel, mittel gegen hoch und hoch gegen gering.

A = Das Merkmal ‚Klassenführung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

B = Das Merkmal ‚Strukturiertheit‘ wurde erhoben mit drei Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

C = Das Merkmal ‚Kognitive Aktivierung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

D = Das Merkmal ‚Sozialklima‘ wurde erhoben mit sieben Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

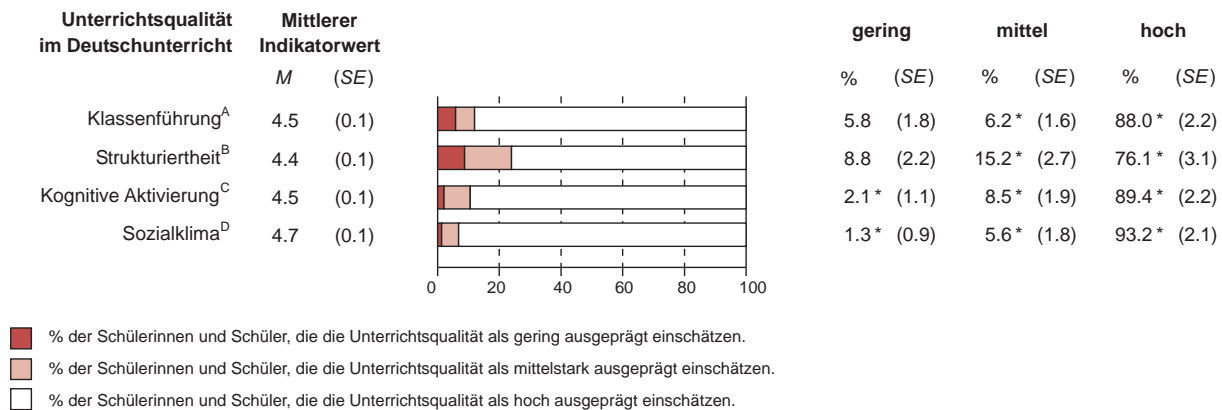
IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

deutlich erkennbare inhaltliche Struktur des Unterrichts schließen. Von etwa 45 Prozent der Lernenden wird der Unterricht als wenig strukturiert wahrgenommen. Diese Schülerinnen und Schüler nehmen wahr, dass nur in wenigen Stunden oder nie am Anfang der Stunde erläutert wird, worum es im Unterricht gehen soll, am Ende der Stunde eine Zusammenfassung erfolgt oder erläutert wird, worauf es in den Unterrichtsstunden inhaltlich ankommt. Die Einschätzung der kognitiven Aktivierung fällt insgesamt positiv aus: Etwa 57 Prozent der Lernenden nehmen in ihrem Deutschunterricht in hohem Maße Merkmale wahr, die dieser Dimension zugeordnet werden. Ungefähr 16 Prozent der Lernenden berichten von einer geringen Ausprägung der Merkmale kognitiv aktivierenden Unterrichts. Aus Schülersicht am positivsten von allen Gestaltungsmerkmalen wird das Sozialklima wahrgenommen. Ungefähr 73 Prozent der Schülerinnen und Schüler haben die Fragen dazu so beantwortet, dass von der Wahrnehmung eines positiv unterstützenden Sozialklimas ausgegangen werden kann. Die Antworten von etwa 5 Prozent der Kinder lassen auf die Wahrnehmung eines zumindest mitunter ungünstigen Sozialklimas schließen.

In Abbildung 9.4 wird dargestellt, wie die Lehrkräfte ihren Deutschunterricht hinsichtlich der vier Qualitätsmerkmale einschätzen.

88 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die die Fragen so beantworten, dass von einer positiven Einschätzung der eigenen Klassenführung in den meisten oder in allen Unterrichtsstunden ausgegangen werden kann. Nur etwa 6 Prozent der Kinder werden von Lehrkräften unterrichtet, deren Antwortverhalten darauf schließen lässt, dass die Klassenführung ihnen nach eigener Einschätzung hinsichtlich einiger oder aller Aspekte nicht immer gelingt. Hinsichtlich der Strukturiertheit des Deutschunterrichts zeigt sich ebenfalls die Tendenz zu einer positiven Einschätzung des eigenen Unterrichts. Etwa 76 Prozent der Lernenden werden von Lehrkräften unterrichtet, die zu einer sehr positiven Einschätzung hinsichtlich der Strukturiertheit des eigenen Deutschunterrichts kommen. Etwa 9 Prozent der Kinder werden von Lehrkräften unterrichtet, die ihre Unterrichtsgestaltung diesbezüglich nicht oder zumindest

Abbildung 9.4: Durch Lehrkräfte wahrgenommene Unterrichtsqualität im Deutschunterricht (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent)

* = Unterschied in den Prozentwerten zwischen den Kategorien statistisch signifikant ($p < .05$).
 Geprüft wurden die Kategorien gering gegen mittel, mittel gegen hoch und hoch gegen gering.

A = Das Merkmal ‚Klassenführung‘ wurde erhoben mit fünf Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

B = Das Merkmal ‚Strukturiertheit‘ wurde erhoben mit drei Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

C = Das Merkmal ‚Kognitive Aktivierung‘ wurde erhoben mit fünf Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

D = Das Merkmal ‚Sozialklima‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

nicht durchgehend positiv beurteilen. Hinsichtlich der Einschätzung der kognitiven Aktivierung ähneln die Werte denen der Klassenführung. Ungefähr 89 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die ihren Deutschunterricht in diesem Sinne als anforderungsreich einschätzen. Nur etwa 2 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, deren Angaben es zulassen, den Deutschunterricht aus Sicht der Lehrenden als wenig kognitiv aktivierend zu bezeichnen. Die positivsten Einschätzungen der Unterrichtsgestaltung aus Sicht der Lehrkräfte finden sich für das Sozialklima. 93 Prozent der Lernenden werden von Lehrkräften unterrichtet, die diesbezüglich zu einer weitgehend positiven Einschätzung kommen. Lediglich etwa 1 Prozent der Lernenden wird von Deutschlehrkräften unterrichtet, deren Angaben eher nicht zu der Einschätzung kommen lassen, dass das Sozialklima durchweg unterstützend ist. In den Ergebnissen zu allen vier Qualitätsmerkmalen wird eine sehr positive Einschätzung des eigenen Deutschunterrichts durch die Lehrkräfte in Deutschland deutlich.

7.2 Die Wahrnehmung des Deutschunterrichts und der Zusammenhang mit Leseleistung und Lesemotivation

Im Mittelpunkt der Diskussion um die Qualität von Unterricht steht der Zusammenhang zwischen der Unterrichtsgestaltung und dem Lernerfolg von Schülerinnen und Schülern (siehe Abbildung 9.2). Lernerfolge im Deutschunterricht zeigen sich nicht alleine in einer ausgeprägten fachlichen Kompetenz, sondern auch in der Entwicklung anderer unterrichtlicher Zielkriterien (z. B. der Entwicklung der fachspezifischen Motivation).

Im Folgenden werden Befunde zum Zusammenhang zwischen der Wahrnehmung des Unterrichts aus Sicht der Schülerinnen und Schüler und der Leseleistung sowie der Lesemotivation berichtet. Zum einen werden die Unter-

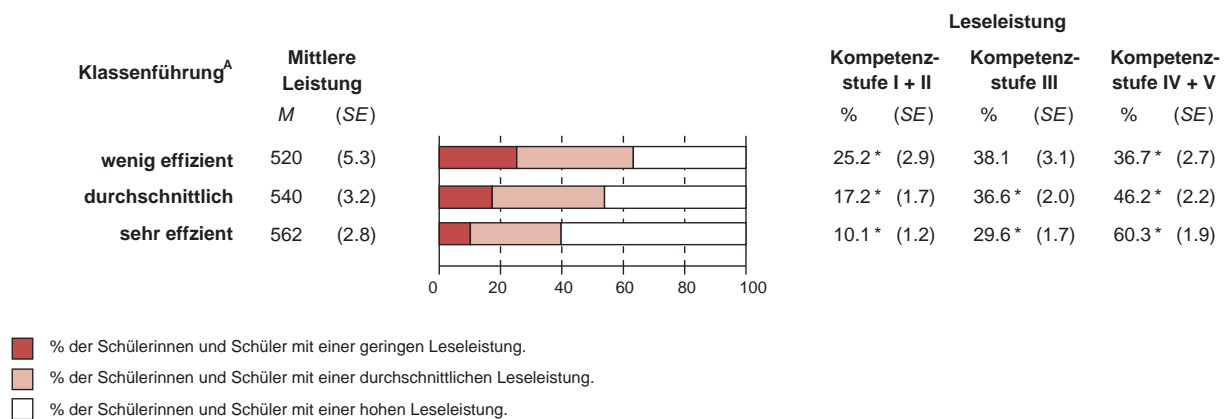
richtseinschätzungen der Lernenden mit dem Kompetenzniveau im Lesen in Verbindung gesetzt, das die Schülerinnen und Schüler bei IGLU 2016 erreicht haben (siehe dazu Kapitel 3 in diesem Band). Zum anderen wird der Zusammenhang zwischen der Wahrnehmung der Unterrichtsgestaltung und der Zuordnung der Lernenden zur Gruppe der gering, durchschnittlich und hoch motivierten Leserinnen und Leser dargestellt, die in Kapitel 4 eingeführt wurde. Schließlich werden die Ergebnisse der Berechnung von bivariaten Korrelationen zwischen der Einschätzung des Unterrichts durch die Schülerinnen und Schüler und der Leseleistung sowie der Lesemotivation berichtet.

Dass im Folgenden die Zusammenhänge zwischen den Einschätzungen der Lehrkräfte und den Leseleistungen sowie der Lesemotivation der Viertklässlerinnen und Viertklässler ausgeblendet bleiben, hat mit dem besonderen Status der Schülerinnen und Schüler und ihrer Wahrnehmung zu tun (siehe Abschnitt 2). Auch wenn es der Lehrkraft gelingt, ein aus ihrer Perspektive hochwertiges Unterrichtsangebot bereitzustellen, bedeutet das nicht zwangsläufig, dass sich positive Effekte bei den Lernenden einstellen. Die Nutzung des Unterrichtsangebots durch die Schülerinnen und Schüler dürfte unter anderem durch deren Wahrnehmung der Qualität des Unterrichts beeinflusst werden.

In der Berichtslegung von IGLU 2016 werden fünf Kompetenzniveaus unterschieden, denen die Schülerinnen und Schüler auf Grundlage ihrer Testergebnisse zugeordnet werden. Im Folgenden werden die Schülerinnen und Schüler, die aufgrund ihrer Leseleistungen den Kompetenzniveaus IV und V zugeordnet werden, als Gruppe leistungsstarker Leserinnen und Leser betrachtet, Schülerinnen und Schüler auf Kompetenzniveau III werden als durchschnittliche Leserinnen und Leser angesehen und Schülerinnen und Schüler, deren Leistungen auf den Kompetenzniveaus I und II liegen, werden als leistungsschwache Leserinnen und Leser betrachtet.

Die Ergebnisse, die in Abbildung 9.5 dargestellt werden, können als Hinweis auf einen Zusammenhang zwischen der Wahrnehmung effizienter Klassenführung und den Leseleistungen gelesen werden. Etwa 60 Prozent der Lernenden, die die Klassenführung im Deutschunterricht als sehr effizient wahrnehmen, gehören zu den leistungsstarken Leserinnen und Lesern. Das gilt nur für etwa 46 be-

Abbildung 9.5: Leseleistungen nach wahrgenommener Klassenführung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



* = Unterschied in den Prozentwerten zwischen den Leistungsgruppen statistisch signifikant ($p < .05$).

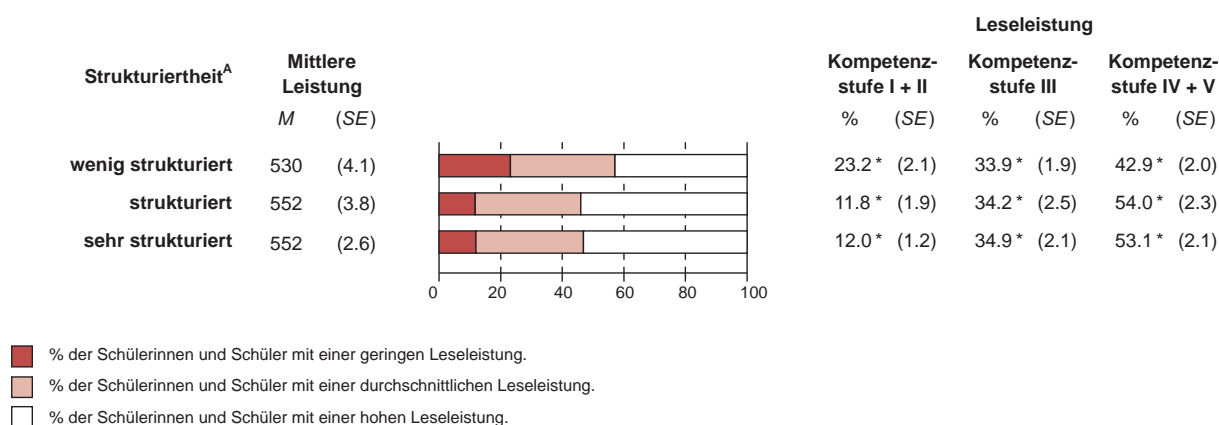
Geprüft wurden die Kompetenzstufen I + II gegen III, III gegen IV + V und IV + V gegen I + II.

A = Das Merkmal ‚Klassenführung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

ziehungsweise 37 Prozent der Lernenden, die den Unterricht als weniger (d.h. als durchschnittlich) effizient oder wenig effizient geführt wahrnehmen. Von den Schülerinnen und Schülern, die den Unterricht als wenig effizient geführt wahrnehmen, gehört ein größerer prozentualer Anteil zur Gruppe der leistungsschwachen Leserinnen und Leser (25.2 %) als von denjenigen, die den Unterricht hinsichtlich der Klassenführung als sehr effizient (10.1 %) oder durchschnittlich effizient (17.2 %) geführt einschätzen.

Hinsichtlich der Strukturiertheit des Deutschunterrichts aus Schülersicht (siehe Abbildung 9.6) zeigen sich keine eindeutigen Tendenzen, wenn es um einen Zusammenhang zwischen der Unterrichtsgestaltung und den Leseleistungen geht. Etwa 53 Prozent der Schülerinnen und Schüler, die den Unterricht als sehr strukturiert wahrnehmen, gehören zu den leistungsstarken Leserinnen und Lesern. Ein ähnliches Bild ergibt sich bei den Schülerinnen und Schülern, die den Unterricht als etwas weniger strukturiert einschätzen. Von diesen gehören sogar 54 Prozent zur Gruppe der leistungsstarken Leserinnen und Leser. Auch immerhin noch etwa 43 Prozent der Schülerinnen und Schüler, die den Unterricht als wenig strukturiert wahrnehmen, gehören zu dieser Gruppe.

Abbildung 9.6: Leseleistungen nach wahrgenommener Strukturiertheit im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



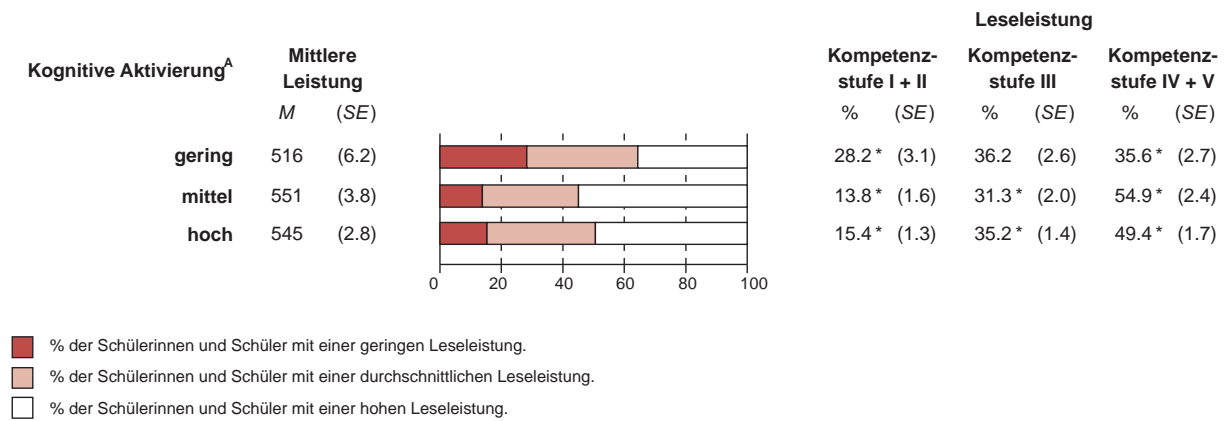
* = Unterschied in den Prozentwerten zwischen den Leistungsgruppen statistisch signifikant ($p < .05$).

Geprüft wurden die Kompetenzstufen I + II gegen III, III gegen IV + V und IV + V gegen I + II.

A = Das Merkmal „Strukturiertheit“ wurde erhoben mit drei Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

Etwa 49 Prozent der Schülerinnen und Schüler, die den Deutschunterricht als in hohem Maße kognitiv aktivierend einschätzen, gehören zur Gruppe der leistungsstarken Leserinnen und Leser (siehe Abbildung 9.7). Unter den Lernenden, die den Unterricht als etwas weniger kognitiv aktivierend einschätzen, fällt der Anteil der Leistungsstarken allerdings mit 54.9 Prozent noch höher aus. Immerhin noch etwa 36 Prozent der Lernenden, die den Unterricht in geringem Maße als kognitiv aktivierend wahrnehmen, gehören zu den leistungsstarken Leserinnen und Lesern. Diese Ergebnisse sprechen eher gegen einen positiven Zusammenhang zwischen der kognitiven Aktivierung, wie sie hier operationalisiert wurde, und den Testleistungen im Lesen. Bei den leistungsschwachen Lernenden zeigt sich eine erwartete Tendenz: Ein größerer Teil der Schülerinnen und Schüler, der den Unterricht in geringem Maße als kognitiv aktivierend beschreibt, gehört zu den leistungsschwachen Leserinnen und Lesern (28.2 %), als das bei den Lernenden der Fall ist, die den Unterricht als mittelstark beziehungsweise in hohem Maße kognitiv aktivierend einschätzen. Nicht in Einklang mit den Erwartungen ist

Abbildung 9.7: Leseleistungen nach wahrgenommener kognitiver Aktivierung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



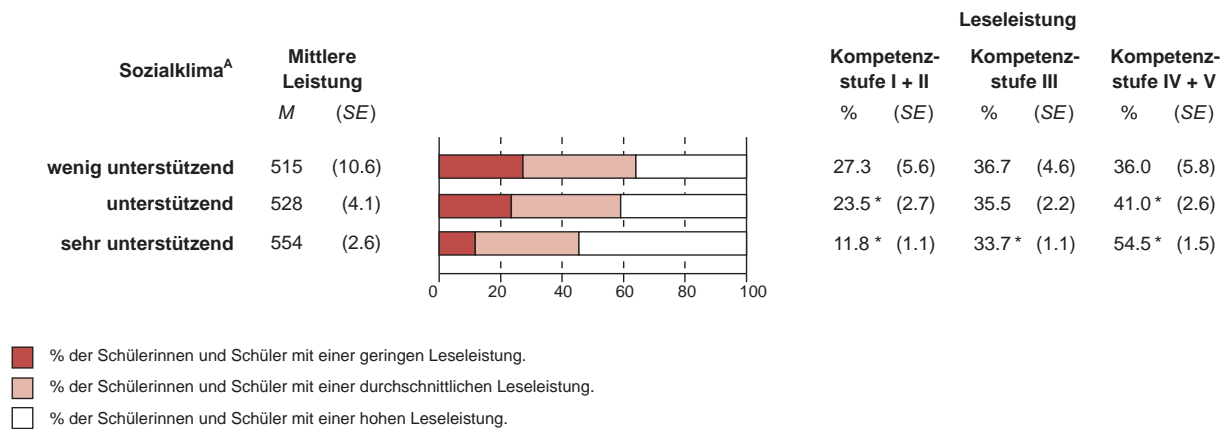
* = Unterschied in den Prozentwerten zwischen den Leistungsgruppen statistisch signifikant ($p < .05$).
 Geprüft wurden die Kompetenzstufen I + II gegen III, III gegen IV + V und IV + V gegen I + II.

A = Das Merkmal ‚Kognitive Aktivierung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

ein Befund zur Gruppe derjenigen, die von geringer kognitiver Aktivierung berichten. Jeweils mehr als ein Drittel dieser Lernenden gehört zur Gruppe der durchschnittlichen (36.2 %) beziehungsweise leistungsstarken Leserinnen und Leser (35.6 %).

Etwa 55 Prozent der Schülerinnen und Schüler, die das Sozialklima im Deutschunterricht als sehr unterstützend wahrnehmen, gehören zur Gruppe der leistungsstarken Leserinnen und Leser (siehe Abbildung 9.8). Das gilt auch für 41 beziehungsweise 36 Prozent der Schülerinnen und Schüler, die das Sozialklima als etwas weniger (aber immer noch) unterstützend oder wenig unterstützend wahrnehmen. Das ist in der Tendenz erwartungskonform. Auch bei den leistungsschwachen Leserinnen und Lesern zeigt sich in der Tendenz das erwartete Bild. Während nur etwa 12 Prozent derjenigen, die den Unterricht hinsichtlich des Sozialklimas sehr positiv einschätzen, zu dieser Gruppe gehören, gilt das für

Abbildung 9.8: Leseleistungen nach wahrgenommenem Sozialklima im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



* = Unterschied in den Prozentwerten zwischen den Leistungsgruppen statistisch signifikant ($p < .05$).
 Geprüft wurden die Kompetenzstufen I + II gegen III, III gegen IV + V und IV + V gegen I + II.

A = Das Merkmal ‚Sozialklima‘ wurde erhoben mit sieben Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

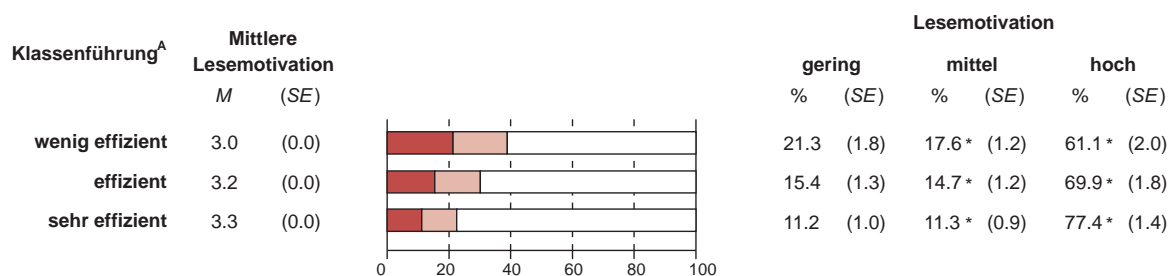
einen größeren Anteil derjenigen, die das Klima weniger positiv (23.5 %) oder zum Teil wohl als problematisch wahrnehmen (27.3 %).

In Kapitel 4 in diesem Band wurde die Operationalisierung der Lesemotivation für IGLU 2016 erläutert. Zudem wurden Ergebnisse zur Lesemotivation der Schülerinnen und Schüler in Deutschland vorgestellt. Dabei wurden Schülerinnen und Schüler mit geringer, mittlerer und hoher Lesemotivation unterschieden. Diese drei Gruppen werden auch im Folgenden unterschieden, wenn es um den Zusammenhang zwischen der Wahrnehmung des Unterrichts aus Schülersicht und der Lesemotivation geht.

Hinsichtlich des Zusammenhangs zwischen der Einschätzung der Klassenführung und der Lesemotivation zeigen sich in der Tendenz die erwarteten Ergebnisse (siehe Abbildung 9.9). Etwa 77 Prozent derjenigen, die den Unterricht in Bezug auf die Klassenführung sehr positiv einschätzen, gehören zur Gruppe derjenigen, die als hoch motiviert angesehen werden können. Das gilt für etwas geringere Anteile derjenigen, die den Unterricht weniger positiv (69.9 %) beziehungsweise zum Teil als wenig effizient geführt (61.1 %) einschätzen. Bei den Schülerinnen und Schülern mit geringer Lesemotivation zeigt sich ein Befund mit umgekehrten Vorzeichen. Der Anteil dieser Schülerinnen und Schüler wird in den Gruppen geringer, bei denen die Einschätzung des Unterrichts positiver ausfällt. Mit etwa 11 Prozent ist dieser Anteil in der Gruppe derjenigen am geringsten, die den Unterricht als sehr effizient geführt wahrnehmen.

Hinsichtlich des Zusammenhangs zwischen der Strukturiertheit des Deutschunterrichts und der Lesemotivation zeigen sich in der Tendenz ähnliche Befunde wie bei der Klassenführung (siehe Abbildung 9.10). In der Gruppe derjenigen, die den Unterricht am positivsten wahrnehmen, ist der Anteil der hoch motivierten Leserinnen und Leser mit etwa 77 Prozent größer als bei denjenigen, die den Unterricht als weniger (72.9 %) oder wenig strukturiert wahrnehmen (62.6 %). Ähnlich wie bei der Klassenführung zeigt sich auch bei der Strukturiertheit des Unterrichts die umgekehrte Tendenz bei den gering motivierten Leserinnen und Lesern. Der Anteil dieser Schülerinnen und Schüler fällt geringer unter denjenigen Lernenden aus, die den Unterricht positiver wahrnehmen.

Abbildung 9.9: Lesemotivation nach wahrgenommener Klassenführung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



 % der Schülerinnen und Schüler mit einer geringen Lesemotivation.

 % der Schülerinnen und Schüler mit einer durchschnittlichen Lesemotivation.

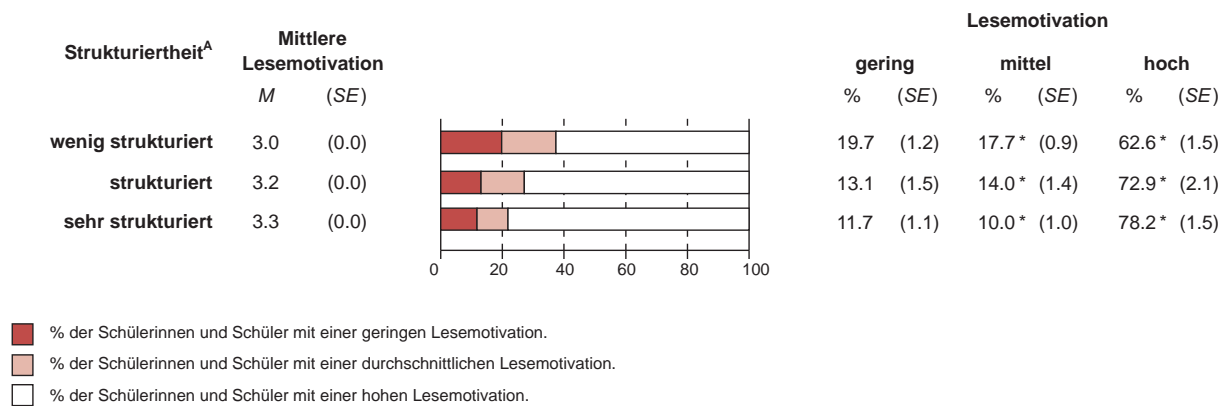
 % der Schülerinnen und Schüler mit einer hohen Lesemotivation.

* = Unterschied in den Prozentwerten zwischen den Kategorien statistisch signifikant ($p < .05$).

Geprüft wurden die Kategorien gering gegen mittel, mittel gegen hoch und hoch gegen gering.

A = Das Merkmal ‚Klassenführung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

Abbildung 9.10: Lesemotivation nach wahrgenommener Strukturiertheit im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



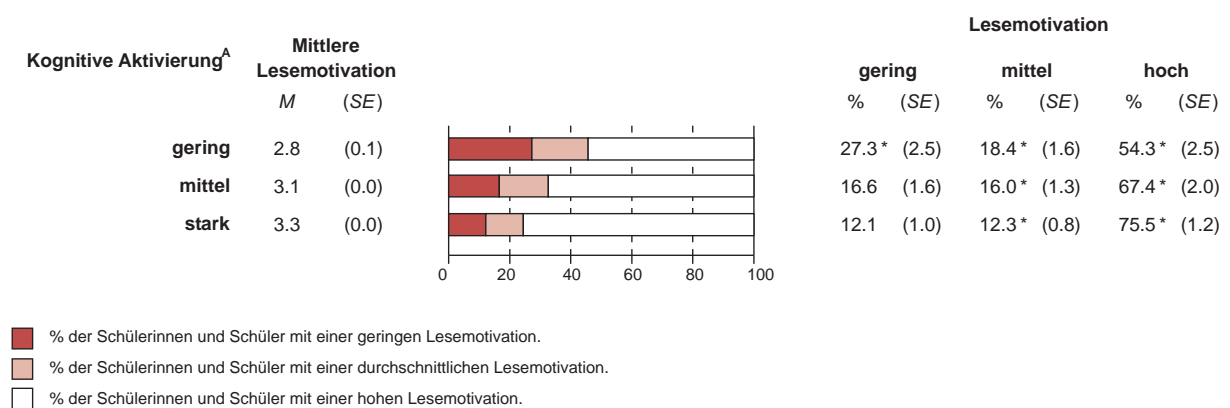
■ % der Schülerinnen und Schüler mit einer geringen Lesemotivation.
 ■ % der Schülerinnen und Schüler mit einer durchschnittlichen Lesemotivation.
 □ % der Schülerinnen und Schüler mit einer hohen Lesemotivation.

* = Unterschied in den Prozentwerten zwischen den Kategorien statistisch signifikant ($p < .05$).
 Geprüft wurden die Kategorien gering gegen mittel, mittel gegen hoch und hoch gegen gering .
 A = Das Merkmal ‚Strukturiertheit‘ wurde erhoben mit drei Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

IEA: Progress in International Reading Literacy Study (PIRLS) © IGLU 2016

Die Ergebnisse zum Zusammenhang der Wahrnehmung der kognitiven Aktivierung und der Lesemotivation können als Hinweise darauf gelesen werden, dass es einen Zusammenhang zwischen der Gestaltung des Unterrichts und der Lesemotivation geben könnte (siehe Abbildung 9.11). Wie bei den beiden anderen Merkmalen ist der Anteil der hoch motivierten Leserinnen und Leser in der Gruppe der Lernenden, die den Unterricht in hohem Maße als kognitiv aktivierend wahrnehmen, höher als in den beiden anderen Gruppen. Zudem ist auch bei den Lernenden der Anteil der gering motivierten Leserinnen und Leser niedriger, die den Unterricht als mittelstark (16.6 %) und stark (12.1 %) kognitiv aktivierend beschreiben, im Vergleich zu dem Anteil der in geringem Maße kognitiv herausgeforderten Schülerinnen und Schüler (27.3 %).

Abbildung 9.11: Lesemotivation nach wahrgenommener kognitiver Aktivierung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



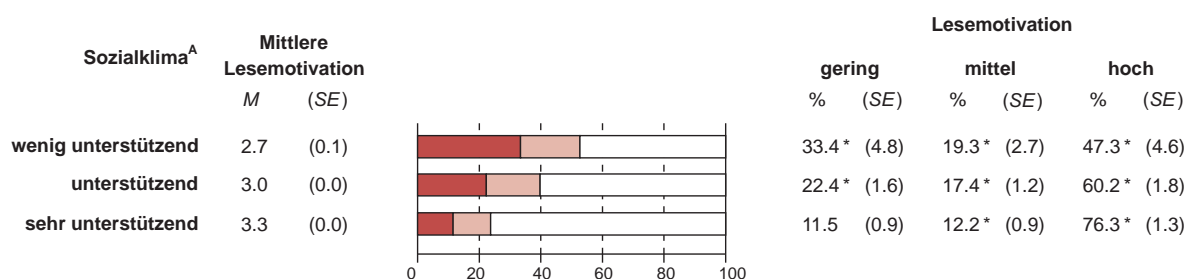
■ % der Schülerinnen und Schüler mit einer geringen Lesemotivation.
 ■ % der Schülerinnen und Schüler mit einer durchschnittlichen Lesemotivation.
 □ % der Schülerinnen und Schüler mit einer hohen Lesemotivation.

* = Unterschied in den Prozentwerten zwischen den Kategorien statistisch signifikant ($p < .05$).
 Geprüft wurden die Kategorien gering gegen mittel, mittel gegen hoch und hoch gegen gering .
 A = Das Merkmal ‚Kognitive Aktivierung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

IEA: Progress in International Reading Literacy Study (PIRLS) © IGLU 2016

Die aus den Ergebnissen anderer Studien in den letzten Jahren gewonnenen Befunde zur Unterrichtsqualität lassen insbesondere Zusammenhänge zwischen der Wahrnehmung eines unterstützenden Sozialklimas und der Lesemotivation erwarten. Es zeigt sich, dass der Anteil derjenigen, die als hoch motiviert angesehen werden können, unter denjenigen Viertklässlerinnen und Viertklässlern am höchsten ausfällt, die das Sozialklima als sehr unterstützend wahrnehmen (76.3 %; siehe Abbildung 9.12). Die Anteile dieser Kinder werden unter denjenigen geringer, die das Sozialklima unterstützend oder wenig unterstützend empfinden (60.2 bzw. 47.3 %). Der Anteil der gering Motivierten fällt kleiner bei den Schülerinnen und Schülern aus, die das Sozialklima als unterstützend (22.4 %) oder sehr unterstützend (11.5 %) wahrnehmen. Von denjenigen, die das Sozialklima als wenig unterstützend wahrnehmen, liegt der Anteil derjenigen, die als gering motiviert angesehen werden können, bei über 33 Prozent.

Abbildung 9.12: Lesemotivation nach wahrgenommenem Sozialklima im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent)



- % der Schülerinnen und Schüler mit einer geringen Lesemotivation.
- % der Schülerinnen und Schüler mit einer durchschnittlichen Lesemotivation.
- % der Schülerinnen und Schüler mit einer hohen Lesemotivation.

* = Unterschied in den Prozentwerten zwischen den Kategorien statistisch signifikant ($p < .05$).
Geprüft wurden die Kategorien gering gegen mittel, mittel gegen hoch und hoch gegen gering.

A = Das Merkmal „Sozialklima“ wurde erhoben mit sieben Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Bivariate Korrelationen zwischen der Wahrnehmung des Unterrichts durch die Schülerinnen und Schüler und der Leseleistung sowie der Lesemotivation können das Bild, das sich durch die beschriebenen Ergebnisse ergeben hat, ergänzen und Hinweise auf näher zu betrachtende Zusammenhänge zwischen der Unterrichtsgestaltung und den Lernerfolgen erbringen.

In Tabelle 9.1 sind die Korrelationen zwischen der durch die Lernenden wahrgenommenen Unterrichtsqualität sowie der Leseleistung und der Lesemotivation eingetragen. Abgesehen von der Korrelation zwischen der kognitiven Aktivierung und der Leseleistung sind alle Zusammenhänge signifikant. Das dürfte mit der Stichprobengröße zusammenhängen. Auch die höchsten Korrelationen, die zwischen der Einschätzung des Sozialklimas und der Leseleistung (.23) sowie der Klassenführung und der Leseleistung (.19) bestehen, sind eher schwach. Kein Zusammenhang ist zwischen der Wahrnehmung der kognitiven Aktivierung und der Leseleistung zu erkennen. Das ist vor dem Hintergrund der Annahmen zur Rolle der kognitiven Aktivierung erklärungsbedürftig (siehe Abschnitt 8). Wenn es um den Zusammenhang zwischen der Unterrichtswahrnehmung durch die Schülerinnen und Schüler und der Lesemotivation geht, ergibt sich hinsichtlich des Sozialklimas ein ähnliches Bild wie bei der Leistung. Die Korrelation

Tabelle 9.1: Bivariate Korrelationen zwischen der Wahrnehmung der Unterrichtsqualität und den Testleistungen im Lesen und der Lesemotivation

	Leseleistung	Lesemotivation
Klassenführung ^A	.19*	.16*
Strukturiertheit ^B	.10*	.15*
Kognitive Aktivierung ^C	– ns	.17*
Sozialklima ^D	.23*	.24*

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$)

A = Das Merkmal ‚Klassenführung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

B = Das Merkmal ‚Strukturiertheit‘ wurde erhoben mit drei Fragen mit vierstufigem Antwortformat (1 = Nie, ..., 4 = In jeder Stunde).

C = Das Merkmal ‚Kognitive Aktivierung‘ wurde erhoben mit sechs Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

D = Das Merkmal ‚Sozialklima‘ wurde erhoben mit sieben Fragen mit vierstufigem Antwortformat (1 = Stimme völlig zu, ..., 4 = Stimme überhaupt nicht zu).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

zwischen der Einschätzung des Sozialklimas und der Lesemotivation fällt am höchsten aus (.24). Am zweithöchsten ist die Korrelation zwischen der Wahrnehmung der kognitiven Aktivierung und der Lesemotivation (.17). Auch diese Korrelationen sind eher schwach.

8 Zusammenfassung und Ausblick

Mit der Unterrichtszeit, die im Durchschnitt mit Deutsch- beziehungsweise Sprachunterricht zugebracht wird, bleibt man in den Grundschulklassen in Deutschland im internationalen Vergleich hinter dem internationalen Mittelwert und dem Wert der Vergleichsgruppe der Staaten der Europäischen Union zurück. Diese Unterschiede sind allerdings nicht signifikant. Signifikant unterscheidet sich die Unterrichtszeit, die in Deutschland mit Deutsch- beziehungsweise Sprachunterricht zugebracht wird, vom Mittelwert der OECD-Staaten. Mit der tatsächlich realisierten Unterrichtszeit für den Deutsch- und Sprachunterricht ist eine zentrale Variable eines Angebots-Nutzungs-Modells des Deutschunterrichts beschrieben. Allerdings kann während der Unterrichtszeit ein unterschiedlich lernwirksames Unterrichtsangebot realisiert werden. Damit rückt die Qualität des Unterrichts in den Fokus.

Helmke (2015) hat in der Auseinandersetzung mit der Rolle der Lehrenden in seinem Angebots-Nutzungs-Modell darauf hingewiesen, dass das Verhalten der Lehrkräfte und ihre Unterrichtsgestaltung entscheidende Faktoren dafür sind, dass die Lernenden das Unterrichtsangebot positiv wahrnehmen und nutzen. Für die Gestaltung des Unterrichts sei für Lehrkräfte eine Orientierung an empirischen Befunden zur Unterrichtsqualität sinnvoll (ebd., S. 84). Mit der Klassenführung, der Strukturiertheit, der kognitiven Aktivierung und dem Sozialklima wurden im Rahmen einer nationalen Ergänzung der Kontextfragebögen von IGLU 2016 einige Merkmale potentiell lernwirksamen Deutschunterrichts aus Sicht der Schülerinnen und Schüler sowie der Lehrkräfte erhoben. Es zeigt sich, dass die über die vier Merkmale insgesamt positive Wahrnehmung der Unterrichtsqualität durch die Schülerinnen und Schülern durch die Einschätzung der Lehrkräfte noch übertroffen wird. In einer sehr positiven Einschätzung der eigenen

Unterrichtsgestaltung wird in der Literatur häufig ein Hinweis auf eine selbstwertdienliche Verzerrung und mangelnde Objektivität der Lehrkräfte gesehen (z.B. Clausen, 2002). Um die Qualität der Einschätzungen der Lehrkräfte abschließend zu beurteilen, kann deren prädiktive Validität ermittelt werden. Zudem kann ein Vergleich der Einschätzungen der Lehrkräfte und der Schülerinnen und Schüler auf Klassenebene vorgenommen werden. Beides wird Gegenstand zukünftiger Analysen sein.

Die Wahrnehmung des Deutschunterrichts durch die Schülerinnen und Schüler ist zwar insgesamt ebenfalls positiv, allerdings fallen ihre Einschätzungen durchaus differenziert aus – die deskriptiven Befunde weisen also auf Unterschiede in der Wahrnehmung der vier Merkmale hin. Insbesondere bei Einschätzungen der Unterrichtsqualität durch Grundschülerinnen und Grundschüler ist zu befürchten, dass der sogenannte „Halo-Effekt“ auftritt: „Differenzielle Angaben zu einzelnen Facetten der Unterrichtsqualität können überlagert werden durch die allgemeine Beliebtheit und Wertschätzung der Lehrkräfte [...]“ (Helmke, 2015, S. 286) Es bleibt zukünftigen Analysen auf Klassenebene vorbehalten, ein drei- beziehungsweise vierdimensionales Modell der Unterrichtsqualität, das sich in den Befunden andeutet, in den Daten nachzuweisen.

Hinsichtlich des Zusammenhangs zwischen den Einschätzungen der Unterrichtsqualität und der Leseleistung sowie der Lesemotivation zeigt sich die Notwendigkeit, die einzelnen Merkmale getrennt voneinander zu betrachten. In den Daten finden sich einige Hinweise auf Zusammenhänge in erwarteter Richtung, zum Beispiel zwischen der Klassenführung sowie dem Sozialklima und der Leistung und zwischen der Wahrnehmung aller vier Unterrichtsmerkmale und der Lesemotivation. Die prädiktive Kraft der Schülereinschätzungen, die sich so andeutet, ist in folgenden Untersuchungen in Mehrebenenanalysen zu überprüfen.

Sowohl bei dem Versuch, die Unterrichtswahrnehmung und die Leseleistung mittels des erreichten Leistungsstandes in Verbindung zu setzen, als auch bei den Ergebnissen zu den bivariaten Korrelationen zwischen Unterrichtseinschätzungen aus Schülersicht und der Leseleistung zeigt sich der bei Berücksichtigung der Forschungslage (siehe Abschnitte 4 und 5) nicht unerwartete, aber doch erklärungsbedürftige Befund, dass die Wahrnehmung der kognitiven Aktivierung aus Sicht der Schülerinnen und Schüler nicht mit der Leseleistung in einem Zusammenhang steht. Eine Erklärung dafür kann darin gesehen werden, dass für IGLU 2016 auf den Versuch verzichtet wurde, die kognitive Aktivierung unter Berücksichtigung lernbereichsspezifischer Besonderheiten der vier Kompetenzbereiche des Deutschunterrichts zu operationalisieren. Die innerhalb der Deutschdidaktik zu führende Diskussion darüber, was kognitive Aktivierung in Bezug auf den Deutschunterricht bedeuten kann und wie die kognitive Aktivierung zu operationalisieren ist, steht noch am Anfang. Eine stärker auf den Gegenstand „Lesen“ fokussierte Operationalisierung wäre möglich, würde aber einen Fokus auf den gesamten Deutschunterricht nicht mehr ermöglichen. Weitere Analysen zur Gestaltung des Leseunterrichts, die in Kapitel 10 begonnen werden, könnten zumindest Aufklärung über einige fachspezifische Aspekte des Leseunterrichts leisten (vgl. dazu auch Lotz, 2016). Zudem ist zu überprüfen, ob die kognitive Aktivierung durch die Einschätzung der Unterrichtsgestaltung und des Lehrerverhaltens aus Sicht der Schülerinnen und Schüler bestmöglich zu erfassen ist. Möglich wäre es, diese Einschätzungen dadurch zu ergänzen, dass die Lernenden um Auskünfte über ihre kognitiven Lernaktivitäten während des Unterrichts gebeten werden. Um die kognitive Aktivierung als Aspekt der

Unterrichtsgestaltung zu erfassen, können Schülerinnen und Schüler zum Beispiel gefragt werden, ob ihre Lehrkräfte sich im Unterricht nach dem Verständnis von Lerninhalten erkundigen. Es ist aber auch möglich, die Schülerinnen und Schüler zu fragen, ob sie ihr Verständnis von Lerninhalten im Unterricht tatsächlich überprüfen mussten und damit die Art von Tiefenlernaktivitäten zu realisieren hatten, die für die kognitive Aktivierung entscheidend sind (siehe Abschnitt 4). Hinsichtlich der kognitiven Aktivierung zeigt sich ein weiterer Befund. Große Anteile der Schülerinnen und Schüler, die den Unterricht als in geringem Maße kognitiv aktivierend wahrnehmen, gehören zu den durchschnittlichen (36.2 %) und leistungsstarken (35.6 %) Leserinnen und Lesern (siehe Abbildung 9.7). Eine kausale Interpretation der Daten ist zwar nicht möglich, aber womöglich sind diese Schülerinnen und Schüler nicht wegen ihres Unterrichts im gezeigten Maße leistungsfähig, sondern trotz ihres Unterrichts. Auch für diese Schülerinnen und Schüler muss ein herausforderndes Angebot geschaffen werden, was sich dann auch in ihren Einschätzungen des Unterrichts widerspiegeln müsste.

Zwischen der Einschätzung des Sozialklimas und der Lesemotivation sind Zusammenhänge zu erkennen. Allerdings stellt sich die Frage, ob mit den berücksichtigten Items tatsächlich den für die Dimension „Sozialklima“ bedeutenden Annahmen der Selbstbestimmungstheorie der Motivation entsprochen wird. Dort werden die drei Grundbedürfnisse Autonomie- und Kompetenzerleben sowie das Erfahren sozialer Einbindung in den Mittelpunkt gestellt (Deci & Ryan, 1993). Es wäre wünschenswert, ein positiv unterstützendes Sozialklima mit stärkerem Fokus auf diese Bedürfnisse und zudem wohl auch mit Berücksichtigung fachspezifischer Besonderheiten zu operationalisieren, als das bei den im Rahmen von IGLU 2016 durchgeführten Befragungen der Fall war.

Literatur

- Baumert, J. & Kunter, M. (2011). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 163–192). Münster: Waxmann.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität*. Münster: Waxmann.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39 (2), 224–238.
- Decristan, J., Hondrich, L. A., Büttner, G., Hertel, S., Klieme, E., Kunter, M., Lühken, A., Adl-Amini, K., Djakovic, S.-K., Mannel, S., Naumann, A. & Hardy, I. (2015). Impact of additional guidance in science education on primary students' conceptual understanding. *Journal of Educational Research*, 108 (5), 358–370.
- Faust, G., Lipowsky, F. & Gleich, A.-K. (2011). Unterrichtsqualität in der Grundschule. Kognitive Aktivierung in der PERLE-Videostudie „Sprache“. *Die Grundschulzeitschrift*, 245/246, 48–52.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014a). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014b). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28 (3), 127–137.
- Gabriel, K. (2014). *Videobasierte Erfassung von Unterrichtsqualität im Anfangsunterricht der Grundschule – Klassenführung und Unterrichtsklima in Deutsch und Mathematik*. Kassel: University Press.

- Hattie, J. (2014). *Lernen sichtbar machen für Lehrpersonen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning for Teachers“ besorgt von Wolfgang Beywl und Klaus Zierer*. Baltmannsweiler: Schneider.
- Hattie, J. (2015). *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“ besorgt von Wolfgang Beywl und Klaus Zierer*. Baltmannsweiler: Schneider.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (6. Auflage). Seelze: Klett-Kallmeyer.
- Helmke, A. & Klieme, E. (2008). Unterricht und Entwicklung sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 301–312). Weinheim: Beltz.
- Helmke, A. & Weinert, F. E. (1997). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241–251). Weinheim: Beltz.
- Heyne, N. (2014). *Merkmale des Leseunterrichts im vierten Schuljahr im Fokus videobasierter Analysen*. Münster: Waxmann.
- Kleinbub, I. (2010). *Unterrichtsqualität im Leseunterricht. Eine videobasierte Analyse in vierten Klassen*. Trier: Wissenschaftlicher Verlag.
- Kleinbub, I. (2016). Ein-Blick ins Klassenzimmer: deutschdidaktische Unterrichtsforschung am Beispiel der Videostudie VERA – Gute Unterrichtspraxis. In M. Krelle & W. Senn (Hrsg.), *Qualitäten von Deutschunterricht* (S. 15–36). Stuttgart: Fillibach bei Klett.
- Klieme, E., Jude, N., Rauch, D., Ehlers, H., Helmke, A., Eichler, W., Thomé, G. & Willenberg, H. (2008). Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 319–344). Weinheim: Beltz.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms BiQua* (S. 127–146). Münster: Waxmann.
- Klieme, E. & Rakoczy, K. (2003). Unterrichtsqualität aus Schülerperspektive: Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider & K.-J. Tillmann (Hrsg.), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 334–359). Opladen: Leske + Budrich.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54 (2), 222–237.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In E. Klieme & J. Baumert (Hrsg.), *TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–57). Bonn: BMBF.
- Klieme, E., Steinert, B. & Hochweber, J. (2010). Zur Bedeutung der Schulqualität für Unterricht und Lernergebnisse. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S. 231–255). Münster: Waxmann.
- König, J., Wagner, C. & Valtin, R. (2011). *Jugend-Schule-Zukunft. Psychosoziale Bedingungen der Persönlichkeitsentwicklung. Ergebnisse der Längsschnittstudie AIDA*. Münster: Waxmann.
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multi-kriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Münster: Waxmann.
- Lankes, E.-M. (2004). Leseunterricht in der Grundschule. Unterschiede zwischen Lehrkräften im internationalen Vergleich. *Zeitschrift für Erziehungswissenschaft*, 7 (4), 551–568.
- Lankes, E.-M. & Carstensen, C. H. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 161–193). Münster: Waxmann.

- Lipowsky, F. (2007). Unterrichtsqualität in der Grundschule – Ansätze und Befunde der nationalen und internationalen Forschung. In K. Möller., P. Hanke, C. Beinbrech, A. K. Hein, T. Kleickmann & R. Schages (Hrsg.), *Qualität von Grundschulunterricht entwickeln, erfassen und bewerten* (Jahrbuch Grundschulforschung, Bd. 11, S. 35–49). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lipowsky, F. (2009). Unterricht. In E. Wild, & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 73–101). Heidelberg: Springer Medizin Verlag.
- Lotz, M. (2016). *Kognitive Aktivierung im Leseunterricht in der Grundschule. Eine Videostudie zur Gestaltung und Qualität von Leseübungen im ersten Schuljahr*. Wiesbaden: Springer Fachmedien.
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guides methods of instruction. *American Psychologist*, 59 (1), 14–19.
- Nortvedt, G. A., Gustafsson, J.-E. & Lehre, A.-C. W. (2016). The importance of instructional quality for den relation between achievement in reading and mathematics. In T. Nilssen & J.-E. Gustafsson (Hrsg.), *Teacher quality, instructional quality and student outcomes. Relationships across countries, cohorts and time* (S. 97–113). Charn: Springer.
- Rieser, S., Stahns, R., Walzebug, A. & Wendt, H. (2016). Einblicke in die Gestaltung des Mathematik- und Sachunterrichts. 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. 203–220) Münster: Waxmann.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454–499.
- Stahns, R. & Rieser, S. (2018, i. Vorb.). Qualität des Leseunterrichts in vierten Klassen in der Grundschule unter den Bedingungen von Mehrsprachigkeit. Ergebnisse der Lehrkräftebefragung von IGLU 2011.
- Tarelli, I., Lankes, E.-M., Drossel, K. & Gegenfurtner, A. (2012). Lehr- und Lernbedingungen an Grundschulen 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. 137–173). Münster: Waxmann.
- Treiber, B. (1982). Lehr- und Lernzeiten im Unterricht. In B. Treiber & F. E. Weinert (Hrsg.), *Lehr-Lern-Forschung. Ein Überblick in Einzeldarstellungen* (S. 12–36). München: Urban & Schwarzenberg.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63 (3), 249–294.
- Wendt, H., Bos, W., Tarelli, I., Vaskova, A. & Walzebug, A. (2016). *IGLU & TIMSS 2011. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster: Waxmann.
- Winkler, I. & Steinmetz, M. (2016). Zum Spannungsverhältnis von deutschdidaktischen Fragestellungen und empirischen Erkenntnismöglichkeiten am Beispiel des Projekts KoALa. In M. Krelle & W. Senn (Hrsg.), *Qualitäten von Deutschunterricht* (S. 37–56). Stuttgart: Fillibach bei Klett.

Kapitel X

Förderung von Leseflüssigkeit und Leseverstehen im Leseunterricht

Albert Bremerich-Vos, Ruven Stahns, Anke Hußmann und Michael Schurig

1 Einleitung

Im internationalen Vergleich ist die Heterogenität der Leseleistungen in Deutschland ausgesprochen groß. Fast 6 Prozent der Viertklässlerinnen und Viertklässler können nur rudimentär lesen, 19 Prozent erreichen nicht die Kompetenzstufe III. 34 beziehungsweise 36 Prozent lesen auf den Stufen III und IV, der Anteil der leistungsstärksten Leserinnen und Leser auf Stufe V beträgt 11 Prozent (siehe Kapitel 3 in diesem Band). Erhebliche Heterogenität im Allgemeinen schließt relative Homogenität auf der Ebene einzelner Klassen nicht aus. Gleichwohl ist es sinnvoll, genauer zu untersuchen, wie Lehrkräfte unter Bedingungen von (gewachsener) Heterogenität den Leseunterricht gestalten. Während in Kapitel 9 die Gestaltung des Deutschunterrichts im Allgemeinen hinsichtlich der Dimensionen „Unterrichtsführung“, „Sozialklima“ und „kognitive Aktivierung“ im Zentrum steht, werden in diesem Beitrag Aspekte der Förderung von Lesekompetenz im Leseunterricht fokussiert.

In fast allen Modellen, die sich auf die Entwicklung des Lesens vom Ende des Kindergartens bis zum Ende der Grundschulzeit beziehen, stehen die „Etappen“ Erwerb der Graphem-Phonem-Korrespondenzregeln, zunehmend flüssiges Lesen zunächst auf Wort- und Satzebene und schließlich Leseverständnis auf der Ebene des Textes im Zentrum. Insofern geht es in der vierten Klasse in erster Linie um die Förderung des Leseverstehens, bei leseschwachen Kindern kann aber auch noch die Förderung der Leseflüssigkeit angezeigt sein. Sie haben womöglich immer noch Schwierigkeiten, Wörter „auf einen Blick“ zu erfassen, und sie lesen in der Regel ausgesprochen langsam (siehe Abschnitt 3). Flüssig Lesende können ihre Aufmerksamkeit voll und ganz dem Textverstehen widmen. Dabei kann insbesondere der flexible Einsatz einzelner Lesestrategien beziehungsweise einer Kombination solcher Strategien hilfreich sein. „Die übereinstimmende Befundlage lässt sich [...] so zusammenfassen, dass insbesondere

Lautlese-Verfahren als empfehlenswert zur Verbesserung der Leseflüssigkeit eingeordnet werden, während zur Förderung des Leseverständnisses vor allem lesestrategieorientierte Ansätze als wirksam gelten.“ (Souvignier, Förster & Kawohl, 2016, S. 79)

In diesem Kapitel wird dargestellt, wie die im Rahmen der *Internationalen Grundschul-Lese-Untersuchung* (IGLU 2016) befragten Lehrkräfte die damit bezeichneten Lehraufgaben angehen. Dabei wird im Folgenden auf Antworten der Lehrkräfte auf Items in einem Fragebogen Bezug genommen. Ein Teil der Items wurde international eingesetzt, weitere wurden auf nationaler Ebene ergnzt.

Das Kapitel ist wie folgt gegliedert: Zunchst wird auf den zeitlichen Umfang speziell des Leseunterrichts in Deutschland eingegangen, so dass deutlich wird, welche Zeitrume mit Angaben wie „*immer oder fast immer*“, „*oft*“ und „*manchmal*“ in etwa gemeint sind (siehe Abschnitt 2). Dann wird dargelegt, wie gro die Anteile besonders schwacher und leistungsstarker Schulerinnen und Schuler nach Auffassung der Lehrkrafte sind und auf welche Quellen sie sich bei ihren Einschatzungen stutzen (siehe Abschnitt 3). In Abschnitt 4 stehen, vornehmlich unter dem Aspekt der Forderung der Leseschwachen, Praktiken des Lautlesens und des stillen Lesens im Zentrum. Dann wird – auch mit Blick auf die Forderung der leistungsstarkeren Kinder – danach gefragt, welchen Stellenwert nach Auskunft der Lehrkrafte Lesestrategien haben beziehungsweise welche Lesestrategien in erster Linie gelehrt und geubt werden (siehe Abschnitt 5). Die Befunde werden unter anderem unter Bezug auf die Testergebnisse und aktuelle Forschungsergebnisse zur Forderung basaler Fahigkeiten, vor allem der Leseflussigkeit, und des strategischen Lesens kommentiert.

2 Zum zeitlichen Umfang des Leseunterrichts

Zum Leseunterricht von Schulerinnen und Schulern in Deutschland liegen Angaben von circa 200 Lehrkraften vor. In der vierten Klasse, in der getestet wurde, unterrichten sie im Mittel 21 Kinder. Von den knapp 30 Wochenpflichtstunden, die im Mittel in den Landern der Bundesrepublik Deutschland zu unterrichten sind, werden im Durchschnitt 8 (373 Minuten, $SD = 170$ Minuten) fur Unterricht im Lesen, Schreiben, Sprechen, im Umgang mit Literatur und in weiteren sprachbezogenen Bereichen verwendet. Hochgerechnet auf das Schuljahr ist der Befund fur Deutschland damit nahezu identisch mit dem internationalen Mittelwert, in den die Ergebnisse aller Teilnehmer an IGLU 2016 eingegangen sind (vgl. Mullis, Martin, Foy & Hooper, 2017). Gefragt, wie viel Zeit sie in einer normalen Schulwoche speziell fur Leseunterricht und/oder Leseaktivitaten der Schulerinnen und Schuler verwenden, geben die Lehrkrafte im Mittel drei Schulstunden an (136 Minuten), wobei die Streuung betrachtlich ist (126 Minuten). Die Schatzungen der einzelnen Lehrkrafte durften zwar recht unsicher sein. Gleichwohl ist bemerkenswert, dass dieser Anteil, international betrachtet, ausgesprochen gering ist. Umgerechnet auf das Schuljahr ergeben sich, absolut betrachtet, etwa 90 Stunden fur expliziten Leseunterricht, auch uber Fachergrenzen hinweg. Der internationale Mittelwert liegt weit daruber, namlich bei knapp 160 Stunden (vgl. ebd.). Berucksichtigt man auch noch das jahrliche Unterrichtsvolumen insgesamt, das in Deutschland bei etwa 840 Zeitstunden liegt, dann ergibt sich speziell fur den Leseunterricht jahresbezogen eine Quote von 11 Prozent. Nur in sechs Teilnehmerstaaten und -regionen ist sie geringer, der internationale Mittelwert betragt 18 Prozent. Es kann also geschlossen wer-

den, dass der Anteil des Leseunterrichts am Gesamtunterricht in Deutschland im internationalen Vergleich ausgesprochen gering ist.

3 Auskünfte zu den Anteilen besonders schwacher und leistungsstarker Schülerinnen und Schüler

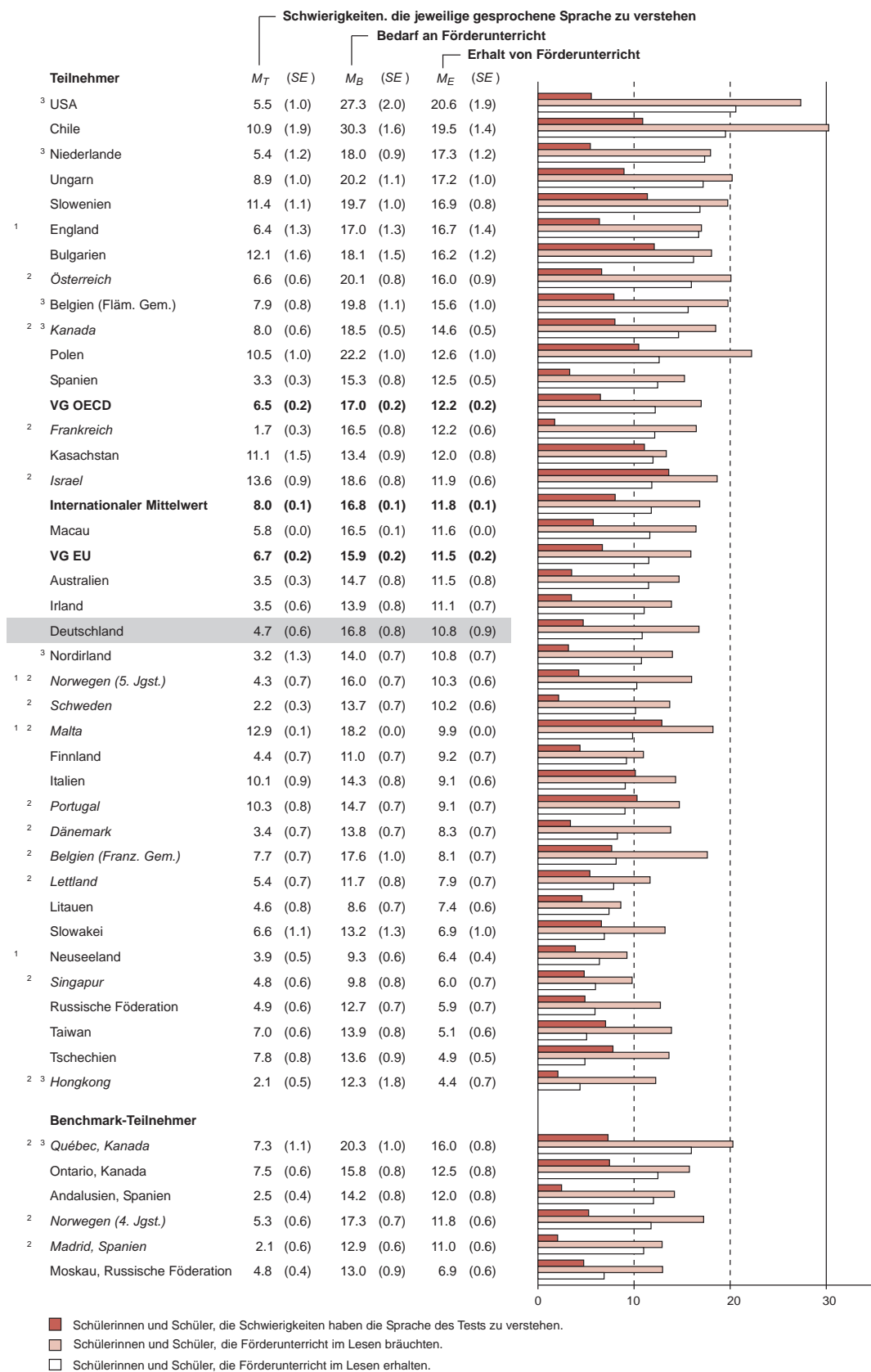
Die Lehrkräfte wurden auch gefragt, wie viele Schülerinnen und Schüler Förderunterricht im Lesen benötigen und wie viele ihn auch tatsächlich bekommen. Es ergibt sich folgender Befund: Laut ihren Angaben besteht bei rund jeder sechsten Schülerin beziehungsweise jedem sechsten Schüler in Deutschland Förderbedarf; faktisch gefördert allerdings wird etwa jede beziehungsweise jeder zehnte. Bezogen auf den Anteil leseschwacher Schülerinnen und Schüler (Kompetenzstufe I und II) in Deutschland (18.9%) zeigt sich, dass insgesamt nur 34.1 Prozent dieser Kinder eine schulische Förderung im Lesen erhalten. Dies sind diejenigen Kinder, für die die Lehrkräfte angeben, dass sie eine Leseförderung bräuchten, als auch jene, für die Lehrkräfte keinen Bedarf an Leseförderung sehen.

International kann dieser Befund nur globaler, das heißt nicht differenziert nach Leistungsgruppen, betrachtet werden. Dabei zeigt sich, dass das einleitend berichtete Ergebnis von Förderbedarf und Fördererhalt in Bezug auf alle Schülerinnen und Schüler in Form von Prozentangaben wie folgt einzuordnen ist (siehe Abbildung 10.1): Die Lehrkräfte in Deutschland sehen bei knapp 17 Prozent der Kinder einen Bedarf an Förderunterricht, was dem Anteil der schwachen Leserinnen und Leser auf den Kompetenzstufen I und II nahekommt, der bei knapp 19 Prozent liegt. Diese Quote entspricht dem internationalen Mittelwert und ist auch nicht signifikant vom Ergebnis auf den Ebenen EU und OECD verschieden. Das gilt ebenso für die Angaben zum tatsächlichen Erhalt von Förderunterricht. In einer größeren Zahl von EU-Ländern ist das Verhältnis von Soll und Ist in der Perspektive der Lehrenden deutlich günstiger als in Deutschland, so in den Niederlanden und in England. Auffällig ist auch, dass es in einigen Teilnehmerstaaten und -regionen mit signifikant leistungsstärkeren Schülerinnen und Schülern wie in der Russischen Föderation, Taiwan und Hongkong markant ungünstiger ausfällt. Speziell für Deutschland ist ein weiterer Befund bemerkenswert. Im Jahr 2016 erreichten 18.9 Prozent der Kinder nicht die Kompetenzstufe III (siehe Kapitel 3 in diesem Band), im Jahr 2011 waren es noch 15.4 Prozent. Während die Lehrkräfte 2011 aber bei 23.1 Prozent der Viertklässlerinnen und Viertklässler Förderbedarf sahen (Tarelli, Lankes, Drossel & Gegenfurtner, 2012, S. 156), ist dieser Anteil 2016 nicht gestiegen, sondern um mehr als 6 Prozent gesunken.

Gefragt, wie viele Kinder in der Klasse sie als fortgeschrittene Leserinnen und Leser einstufen, geben die Lehrkräfte im Mittel rund 8 (genauer: 8.4, $SD = 4.7$) an. Versteht man die Etikettierung „fortgeschritten“ im IGLU-Rahmen so, dass es sich um Kinder handelt, die mindestens Kompetenzstufe IV erreicht haben, das heißt 47 Prozent (siehe Kapitel 3 in diesem Band), dann sind die Angaben der Lehrkräfte im Mittel ($MW = 8.4$ von 21.2, das heißt 40%) in etwa vergleichbar. Ob es sich aber um dieselben Kinder handelt, ist nicht zu ermitteln, weil die Lehrkräfte nur gebeten waren, summarische Angaben zu machen.

Der Fragebogen enthält zwar keine Frage dazu, auf welche Quellen sich die Lehrkräfte bei ihren Schätzungen stützen. Es lag ihnen aber das in Tabelle 10.1 dargestellte Item vor.

Abbildung 10.1: Förderbedarf und Förderangebot (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent)



- Schülerinnen und Schüler, die Schwierigkeiten haben die Sprache des Tests zu verstehen.
- Schülerinnen und Schüler, die Förderunterricht im Lesen bräuchten.
- Schülerinnen und Schüler, die Förderunterricht im Lesen erhalten.

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.

2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.

3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

Tabelle 10.1: Quellen zur Beurteilung von Fortschritten im Lesen der Schülerinnen und Schüler (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent)

Welches Gewicht messen Sie den folgenden Quellen bei, um den Fortschritt der Schüler im Lesen zu beurteilen?	großes Gewicht		einiges Gewicht		wenig oder kein Gewicht	
	%	(SE)	%	(SE)	%	(SE)
Beurteilung der laufenden Arbeit der Schüler	80.3	(2.4)	17.4	(2.3)	0.2	(0.2)
Klassenarbeiten (z. B. von der Lehrkraft entworfene oder dem Lehrbuch entnommene Arbeiten)	41.2	(3.4)	51.5	(3.3)	5.1	(1.6)
Nationale, bundeslandspezifische oder regionale Lernstandserhebungen	3.6	(1.2)	40.1	(3.7)	54.0	(3.8)

Differenzen zu 100 Prozent ergeben sich durch Rundungsfehler.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Es wird deutlich, dass vor allem kontinuierliche Beobachtungen und auch Klassenarbeiten die Informationsquellen erster Wahl sind. Auf Lernstandserhebungen und andere objektive Tests setzt die Mehrheit der Lehrkräfte nicht. Angesichts dieses Befunds ist zu vermuten, dass sie sich, wenn sie Kinder als schwache beziehungsweise fortgeschrittene Leserinnen und Leser einstufen, an einer Sozialnorm orientieren, wobei die jeweilige Klasse die Referenzgruppe sein dürfte. Eine inhaltsbezogene beziehungsweise kriteriale Norm, wie sie etwa der Rückmeldung zu den Leseleistungen im Rahmen der Vergleichsarbeiten in dritten Klassen (VerA 3) zugrunde liegt, dürfte kaum eine Rolle gespielt haben.

Daraus zu schließen, dass die diagnostischen Urteile der Lehrkräfte nicht akkurat sind, wäre aber verfehlt. Folgt man Schrader (2011), dann kann man drei Facetten von Urteilsgenauigkeit unterscheiden. Die Urteile betreffen demnach die mittlere Leseleistung in der jeweiligen Klasse, die Rangplätze der einzelnen Schülerinnen und Schüler und die Leistungsstreuung. Versteht man die den Lehrkräften vorgelegten Fragen so, dass sie auf Angaben zu Rangplätzen und zur Streuung zielen, dann zeigt sich, dass sie mit den Testergebnissen kompatibel sind. So ist zum Beispiel in den Klassen der Lehrkräfte, die überdurchschnittlich hohe Zahlen förderbedürftiger Kinder nennen, auch ein überdurchschnittlich hoher Anteil der Schülerinnen und Schüler auf den Kompetenzstufen I und II anzutreffen (ohne Abbildung).

4 Förderung der Leseflüssigkeit

In den Klassen mit relevanten Anteilen von Kindern, die auf den Kompetenzstufen I und II lesen beziehungsweise von den Lehrkräften als förderbedürftig angesehen werden, dürfte – auch noch im vierten Schuljahr – unter anderem die Förderung der Leseflüssigkeit angezeigt sein.

Als zentrale Komponenten der Leseflüssigkeit sind Dekodiergenauigkeit, Automatisierung, Lesegeschwindigkeit und prosodisch phrasiertes Lesen anzusehen (z. B. Rosebrock, Nix, Rieckmann & Gold, 2013, S. 15 ff.). Dem Zwei-Wege-Modell der Worterkennung zufolge (vgl. z. B. Dehaene, 2010, S. 50 ff.) können Leserinnen und Leser Graphemfolgen auf direktem Weg über das ortho-

graphische Lexikon („Sichtwortschatz“) Bedeutungen zuweisen und/oder sie sind auf den indirekten Weg über eine phonologische Rekodierung angewiesen. Beim indirekten Weg werden Einheiten unterhalb der Wortebene (einzelne Grapheme, Cluster, Schreibsilben) zunächst phonetisch beziehungsweise phonologisch repräsentiert, bevor die Bedeutung des Zielworts erfasst werden kann. Auch gute Leserinnen und Leser sind von Fall zu Fall noch auf diese indirekte Route angewiesen. Steht aber nur dieser prälexikalische Weg zur Verfügung, erfordert der Abgleich der erlesenen Klanggestalt mit den verfügbaren semantischen Einträgen im mentalen Lexikon viel Zeit, zumal dann, wenn er etappenweise über verschiedene „Wortvorformen“ erfolgt. So mag zwar Dekodiergenauigkeit erreicht werden, der Grad der dafür erforderlichen Aufmerksamkeit ist aber hoch, das Ausmaß der Automatisierung demnach gering. Wenn vermittelt über einen umfangreichen Sichtwortschatz die Wortbedeutungen im semantischen Lexikon schnell aktiviert werden können, also vor allem bei frequenten Wörtern der direkte lexikalische Pfad zur Verfügung steht, ist das Arbeitsgedächtnis weniger beansprucht als zum Beispiel dann, wenn Wörter noch lautierend erlesen werden müssen. In vielen Studien hat sich gezeigt, dass die Lesegeschwindigkeit, die allerdings unter anderem vom Leseinteresse, von der Komplexität des jeweiligen Textes und vom Umfang des passiven Wortschatzes beeinflusst ist, und Leistungen in standardisierten Leseverständnistests hoch korreliert sind (z.B. Gasteiger & Gasteiger-Klicpera, 1993, S. 88). Über die Dekodiergenauigkeit und automatisiertes, schnelles Lesen hinaus wird die Fähigkeit zum betonten und sinn gestaltenden Lesen als Komponente von Leseflüssigkeit verstanden. Hierbei geht es um Aspekte wie den nicht auf einzelne Wörter, sondern auf Wortgruppen (z.B. Phrasen) und Sätze bezogenen Akzent in Form einer Variation der Lautstärke und/oder der Tonhöhe und um die Pausengliederung. Zentrales Kriterium ist die semantisch-syntaktische Plausibilität der Phrasierungen beim (lauten) Lesen. Dieses Kriterium wird zum Beispiel verfehlt, wenn monoton Wort für Wort gelesen wird.

Die Komponenten der Leseflüssigkeit sind nicht modular, sondern interdependent. Dabei hat die Dekodiergenauigkeit eher basalen Charakter; gelingt die Dekodierung auf Wortebene allerdings nur bei erheblicher Aufmerksamkeit und langsam, ist das kompetente sinnbetonende Lesen nicht möglich. Diagnostisch einschlägige Fragen sind etwa: Wie synthetisiert ein Kind ein schwieriges Wort, nutzt es die silbische oder morphematische Gliederung, wiederholt es den Wortanfang? Kommen Wortersetzungen, Hinzufügungen und Auslassungen vor? Korrigiert es sich bei Verlesungen, erfolgreich oder nicht? Werden Satzzeichen genutzt, an welchen Stellen wird pausiert (von Wedel-Wolff, 2002)? Haben Kinder noch Probleme bei der schnellen Erkennung von Wörtern, können Übungen zum Blitzlesen beziehungsweise zum blitzschnellen Worterkennen angezeigt sein (Mayer, 2012; Valtin, Naegelé & Sasse, 2013).

Für die Förderung der Leseflüssigkeit können verschiedene Verfahren in Betracht gezogen werden, vor allem stilles Lesen und Varianten lauten Lesens. Die Wirksamkeit von ausgedehntem stillen Lesen (*sustained silent reading*) im Unterricht ist derzeit noch nicht hinreichend belegt. Wer auf stilles Lesen setzt, hofft nicht nur, dass die Kinder allmählich motivierter lesen und ein zunehmend positiveres Selbstkonzept entwickeln. Mit ausgedehntem stillen Lesen sollten auch – so die Erwartung – Verbesserungen der Leseflüssigkeit und des Leseverstehens einhergehen (zusammenfassend Rosebrock & Nix, 2014, S. 57 ff.). Die Erwartung scheint plausibel zu sein, ist doch vielfach belegt, dass gute Leserinnen und Leser in der Regel mehr lesen als weniger kompetente. Hier

ist aber einzuwenden, dass diese Korrelation nicht als kausale Relation verstanden werden sollte. So ist zwar denkbar, dass im Gefolge der Lektüre größerer Textmengen die Lesekompetenz steigt. Es ist aber auch plausibel, dass man mehr liest, weil man kompetenter ist. Insbesondere für Kinder im Grundschulalter stehen nach unserer Kenntnis Langzeitstudien noch aus. Es gibt aber Hinweise darauf, dass gerade für schwache Leserinnen und Leser ausgedehntes leises Lesen größerer Textmengen eine Überforderung ist (NICHD, 2000, S. 13). „Für sie können die Vielleseverfahren zum gegenwärtigen Zeitpunkt deshalb nicht als erstes Mittel der Wahl empfohlen werden [...].“ (Rosebrock et al., 2013, S. 22)

Metaanalysen englischsprachiger Forschung zur Förderung der Leseflüssigkeit haben gezeigt, dass lautes Reihum-Lesen in der Klasse gar keine oder nur geringe Effekte hat (NICHD, 2000, S. 3 ff.). Kritisch wird unter anderem eingewandt, dass die Vorlesenden sich nicht vorbereiten können (Spinner, 2004, S. 133). Es komme auch häufig vor, dass mitlesende Schülerinnen und Schüler schneller lesen als stockend und langsam Vortragende (Rosebrock & Nix, 2014, S. 45). Zudem werde im Zuge der mitunter zahlreichen Wechsel der Vorleser-Rolle der Leseprozess oft unterbrochen, weil die jeweiligen „Nachfolger“ Mühe haben, an der ihnen zugewiesenen Textstelle einzusetzen. Überdies könne das Verstehen dadurch beeinträchtigt werden, dass sich Schülerinnen und Schüler nach getaner Vorlesungsarbeit zurücklehnen und nicht mehr aufmerksam sind. Aufgrund dieser und weiterer Schwächen (vgl. Nix, 2011) ist lautes Reihum-Lesen für die Förderung von Leseflüssigkeit und Textverstehen nicht geeignet.

Liest aber die Lehrkraft der Klasse laut vor, werden meist positive Effekte auf das Textverstehen angenommen (Spinner, 2004, S. 133). Das Verständnis des Textes wird nicht dadurch beeinträchtigt, dass Passagen unflüssig gelesen werden. Hinzu kommt, dass die Kinder es hier mit einem positiven Modell- beziehungsweise Experten-Leser zu tun haben, der ihnen zum Vorbild werden kann.

Weder das Reihum-Lesen noch der Vortrag durch die Lehrkraft stehen aber im Mittelpunkt der Diskussion um die Förderung der Lesekompetenz durch lautes Lesen. In zahlreichen Studien erwiesen sich nämlich Ansätze als effektiv, bei denen auf begleitetes, wiederholtes lautes Lesen gesetzt wird. Bei verschiedenen Varianten dieses Konzepts – Mitschülerinnen und Mitschüler, Eltern oder Lehrpersonen als laut Mitlesende – zeigten sich in den USA substantielle Wirkungen. Beim Vergleich von Experimental- und Kontrollgruppen ergaben sich moderate ($d = 0.50$) Effekte (NICHD, 2000).

Anders als in angelsächsischen Ländern ist die Förderung der Leseflüssigkeit in Deutschland erst in den letzten Jahren in den Fokus gerückt (Rosebrock et al., 2013). Sieht man von Studien ab, die sich vornehmlich auf Kinder mit diagnostiziertem Förderbedarf beziehen, dann sind einschlägige Projekte allerdings noch rar und die Ergebnisse darüber hinaus nicht eindeutig, wie anhand von zwei Studien gezeigt werden soll.

Lauer-Schmaltz, Rosebrock und Gold (2014) untersuchten die Wirksamkeit von zwei Varianten wiederholten lauten Lesens. Im Rahmen einer Klassen-Lautleseroutine (KLL) liest die Lehrkraft den Text zunächst zweimal vor, die Kinder fahren die Wörter mit dem Finger nach. Es folgt ein Gespräch über Inhalte und Stolpersteine. Anschließend wird der Text chorisch gelesen, bevor die Kinder abschließend in Tandems lesen und sich wechselseitig Feedback geben. Die zweite Variante besteht darin, dass in Tandems aus jeweils einem schwächeren und einem stärkeren Leser halblaut synchron gelesen wird, wobei der stärkere Modell- und Kontrollfunktion übernimmt. Circa 800 Drittklässlerinnen und Drittklässler nahmen an Vor-, Haupt- und Follow-up-Tests teil, verteilt auf

die beiden Versuchsgruppen und eine Kontrollgruppe mit regulärem Deutschunterricht, über dessen Merkmale allerdings nichts weiter ausgesagt wird. Für die Leseflüssigkeit, gemessen mit den auf Wort und Satz bezogenen Teilen des *Salzburger Lese-Screenings SLS 1–4* (Mayringer & Wimmer, 2003) beziehungsweise von ELFE 1–6 (Lenhard & Schneider, 2006), ergaben sich zwar kurzfristig signifikant größere Zuwächse als für die Kontrollgruppe, längerfristig waren sie aber nur für KLL geringfügig größer ($d = 0.12$), nicht für die Tandems. Im Hinblick auf das Textverstehen konnte kurzzeitig vor allem bei den Tandems ein größerer Zuwachs als bei der Kontrollgruppe festgestellt werden ($d = 0.28$); längerfristig gab es aber keine signifikanten Differenzen zwischen den drei Gruppen. Angesichts dessen, dass die Kinder der Kontrollgruppe im Lauf eines Schulhalbjahrs bei der Leseflüssigkeit einen Zuwachs von $d = 0.80$ erreichten, werfen Lauer-Schmaltz et al. die Frage auf, ob eine gezielte Förderung der Leseflüssigkeit in dritten Klassen überhaupt noch sinnvoll ist. „Es scheint [...] plausibel, dass die rasche Steigerung der Leseflüssigkeit durch den ‚normalen‘ Unterricht in der Grundschule befördert wird“ (ebd., S. 57). Es gab auch Indizien dafür, dass die Arbeit in Tandems für schwächere Kinder zu anspruchsvoll war, wofür vor allem Wortschatzprobleme verantwortlich sein könnten. Auf der anderen Seite stufte etwa ein Drittel der Kinder das Tandem-Training als zu leicht ein. Insofern sei zu prüfen, ob für schwache Leserinnen und Leser eine Förderung der Leseflüssigkeit außerhalb des Regelunterrichts ins Auge zu fassen ist.

Munser-Kiefer (2014) verglich unter anderem die Dekodierleistungen (gemessen mit Hilfe von ELFE 1–6 auf Wortebene) von jeweils etwa 130 Drittklässlerinnen und Drittklässlern. Im Lauf eines Halbjahrs arbeitete die Versuchsgruppe in Form von Tandems. Texte wurden abschnittsweise abwechselnd gelesen, das nicht lesende Kind verwies das andere auf seine Verlesungen, daraufhin segmentierte das lesende Kind die Fehlerwörter in Silben und las Wort und Satz erneut. Im Anschluss las jedes Kind für sich die Fehlerwörter erneut und weitere komplexe Wörter nach freier Wahl, zum Schluss las man sich wechselseitig die einschlägigen Sätze erneut vor. Die gesamte Prozedur dauerte 20 Minuten, in der restlichen Zeit arbeiteten die Kinder wie die Kontrollgruppe im normalen Leseunterricht. Es zeigte sich, dass die Kinder in den Tandemgruppen, je nach ihren Leistungen Zwillingen in der Kontrollgruppe zugeordnet, im Hinblick auf die zentrale Facette der Leseflüssigkeit, die Dekodierfähigkeit, signifikant und in mittlerer Größenordnung ($d = 0.40$ bis 0.50) überlegen waren (vgl. ebd., S. 307). Beim Leseverständnis gab es keine größeren Zuwächse als in der Vergleichsgruppe.

Anders als für das laute Reihum-Lesen und das stille Viellesen sind die Befunde für Varianten des wiederholten und begleiteten (halb-)lauten Lesens demnach international betrachtet ermutigend, wobei die Ergebnisse für Deutschland noch rar und auch nicht eindeutig sind.

Vor diesem Hintergrund sollen im Folgenden die Auskünfte der Lehrkräfte im Rahmen von IGLU 2016 betrachtet werden. Pro Schultag erleben Schülerinnen und Schüler in Deutschland im Mittel etwa eine halbe Stunde Leseunterricht und/oder Leseaktivitäten wobei die Streuung beträchtlich ist. Die Lehrkräfte sollten unter anderem angeben, wie oft sie selbst Schülerinnen und Schülern vorlesen beziehungsweise sie auffordern, laut vorzulesen. Die Antwortoptionen waren (1) „jeden oder fast jeden Tag“, (2) „ein- bis zweimal pro Woche“, (3) „ein- bis zweimal pro Monat“ und (4) „nie oder fast nie“. 71 Prozent der Viertklässlerinnen und Viertklässler werden von Lehrkräften unterrichtet, die angeben, dass

sie ihnen mindestens ein- bis zweimal pro Woche vorlesen. Nur bei 6 Prozent kommt das nie oder fast nie vor. Noch häufiger sollen die Kinder laut vorlesen. 67 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die das jeden oder fast jeden Tag verlangen, weitere 29 Prozent von Lehrkräften, die dies ein- bis zweimal pro Woche einfordern. Die Lehrkräfte wurden nicht gefragt, welche Varianten des klassenweisen lauten Vorlesens sie im Einzelnen praktizieren beziehungsweise favorisieren. Nimmt man an, dass Auskünfte der Schülerinnen und Schüler zu diesem Thema valide sind, dann dürfte die dominierende Variante das laute Reihum-Vorlesen sein: 67 Prozent der Viertklässlerinnen und Viertklässler zufolge kommt in jeder oder in den meisten Stunden das Reihum-Lesen vor. Unterschiedliche Schülerinnen und Schüler lesen nacheinander Teile eines Textes vor, bis der ganze Text gelesen ist. Die chorische Version, also das gemeinsame laute Lesen eines Textes, wird den Kindern zufolge vergleichsweise selten praktiziert. Ein Drittel kreuzt an, das sei in jeder oder den meisten Stunden der Fall, 29 Prozent votieren für „in einigen Stunden“, 37 Prozent entscheiden sich für die Option „nie oder fast nie“.

Nach eigenem Bekunden ermuntern die Lehrkräfte die Kinder in noch größerem Maß als zum lauten Vorlesen dazu, im Unterricht leise für sich selbst zu lesen. Das sollen fast alle Viertklässlerinnen und Viertklässler jeden Tag oder ein- bis zweimal pro Woche tun.

Nach dem lauten oder leisen Lesen kommt es in der Regel zu einem mündlichen Diskurs über den jeweiligen Text. Dabei haben mehr als 90 Prozent der Kinder jeden oder fast jeden Tag beziehungsweise ein- bis zweimal pro Woche Fragen der Lehrkräfte zu beantworten. Deutlich seltener kommt es vor, dass sie nach dem Lesen in schriftlicher Form reagieren müssen. Diese Option kommt den Angaben der Lehrkräfte zufolge nur bei 5 Prozent der Kinder jeden oder fast jeden Tag zum Tragen, bei 40 Prozent ein- bis zweimal pro Woche. Dabei hat ein Setting, in das Schreiben einbezogen ist, im Vergleich mit einem nur auf Mündlichkeit setzenden Arrangement einen wesentlichen Vorteil: Jede einzelne Schülerin beziehungsweise jeder einzelne Schüler ist gehalten, im Anschluss an eine das Nachdenken befördernde, „verzögerte“ Planung etwas zu Papier zu bringen. Wird dagegen ausschließlich auf das Gespräch gesetzt, wird die Neigung gerade schwacher Leserinnen und Leser begünstigt, sich gar nicht zu beteiligen beziehungsweise die Beiträge anderer zu wiederholen.

Dem gegenwärtigen Forschungsstand zufolge ist der am ehesten Erfolg versprechende Weg, um besonders leseschwache, das heißt noch weitgehend unflüssig lesende Kinder, zu fördern, das wiederholte (halb-)laute Lesen im Team. Den Lehrkräften wurde im Rahmen der nationalen Ergänzung des internationalen Fragebogens die folgende Aussage vorgelegt: „In meinem Leseunterricht lasse ich bessere und schwächere Leser gemeinsam einen Text halblaut lesen“. 20 Prozent der Kinder in Deutschland werden von Lehrkräften unterrichtet, die angeben, dass sie diesen Ansatz fast jeden Tag beziehungsweise ein- bis zweimal pro Woche praktizieren. 44 Prozent werden von Lehrkräften unterrichtet, für die der Ansatz überhaupt nicht in Frage kommt, 19 Prozent von Lehrkräften, die ihn ein- bis zweimal pro Monat realisieren. 17 Prozent der Kinder erleben Lehrkräfte, die (halb-)lautes Lesen im Team noch seltener einsetzen. Die Lehrkräfte, die nicht für gänzliche Ablehnung votierten, wurden zusätzlich nach einigen Merkmalen des von ihnen favorisierten Verfahrens gefragt (siehe Tabelle 10.2).

Tabelle 10.2: Verfahren zum halblauten Lesen im Unterricht (Anteile der Schülerinnen und Schüler nach Angabe der Lehrkräfte in Prozent)

Verfahren zum halblauten Lesen	ja		nein	
	%	(SE)	%	(SE)
Die Schüler müssen einen Text gemeinsam und wiederholt lesen.	87.0	(3.7)	13.0	(3.7)
Der bessere Leser achtet auf Lesefehler und korrigiert sie.	77.2	(4.2)	22.8	(4.2)
Die beiden Schüler lesen den Text so oft, bis der schwächere keine Fehler mehr macht.	24.8	(4.8)	75.2	(4.8)
Ich überprüfe, ob der schwächere Schüler den Text fehlerfrei lesen kann.	40.6	(5.5)	59.4	(5.5)

IEA: Progress in International Reading Literacy Study (PIRLS)

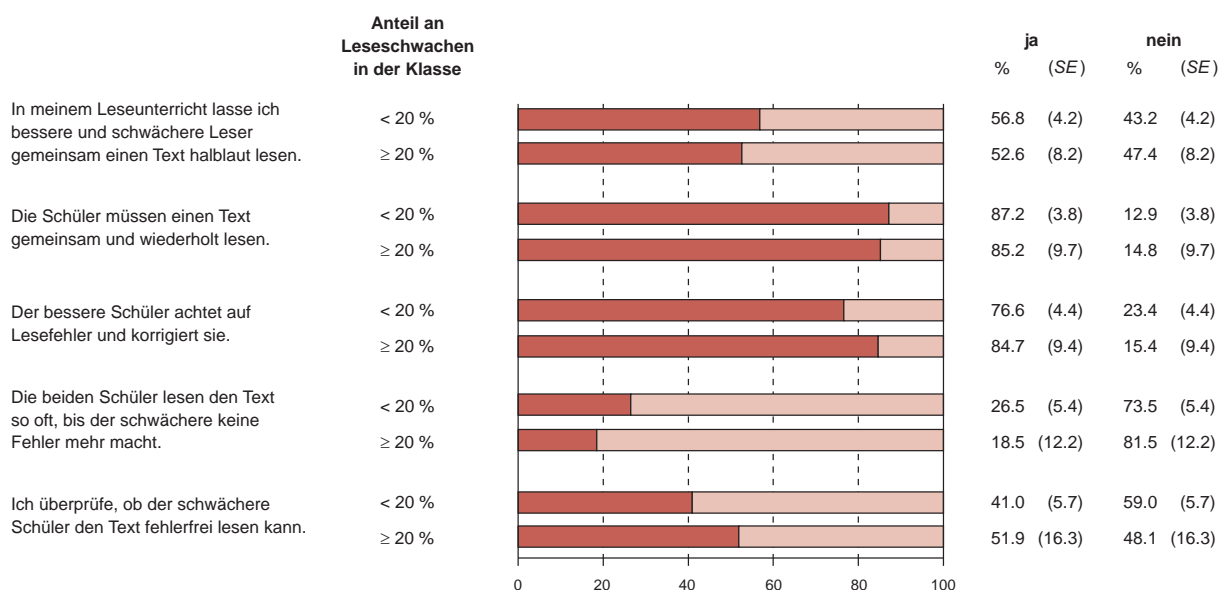
© IGLU 2016

Mehr als die Hälfte der Schülerinnen und Schüler in Deutschland hat also den befragten Lehrkräften zufolge im Unterricht Gelegenheit, im Rahmen von Lautlese-Tandems zu üben, wenn auch in der Regel nur selten. 87 Prozent dieser Kinder werden von Lehrkräften unterrichtet, die angeben, dass die Schülerinnen und Schüler in diesem Setting einen Text gemeinsam und wiederholt lesen müssen. Dabei hat – so die Lehrkräfte von drei Vierteln der Schülerinnen und Schüler – die bessere Leserin beziehungsweise der bessere Leser auf Lesefehler zu achten und sie zu korrigieren. Bei nur noch einem Viertel der Schülerinnen und Schüler müssen beide Kinder den Text den Angaben der Lehrkräfte zufolge so oft lesen, bis das schwächere keine Fehler mehr macht. Weniger als die Hälfte der Kinder wird von Lehrkräften unterrichtet, die überprüfen, ob der schwächere Schüler beziehungsweise die schwächere Schülerin den Text am Ende fehlerfrei lesen kann. Auf das zentrale Erfolgskriterium, nämlich das fehlerfreie Lesen des schwächeren Partners, wird allerdings nur bei einer Minderheit geachtet.

Die Hypothese liegt nahe, dass diese Prozedur vor allem von den Lehrenden ins Auge gefasst wird, in deren Klassen relativ viele Kinder schwach oder sehr schwach Lesende sind. Differenziert man zwischen den Klassen, in denen weniger als 20 Prozent der Schülerinnen und Schüler allenfalls auf den Kompetenzstufen I und II lesen können, und den Klassen, bei denen das auf 20 und mehr Prozent zutrifft, dann ergibt sich folgendes Bild (siehe Abbildung 10.2).

Demnach trifft die Hypothese nicht zu. Die Lehrkräfte, die überhaupt die Arbeit in Lautlese-Tandems vorsehen, tun das unabhängig davon, wie groß der Anteil schwach Lesender in ihren Klassen ist.

Abbildung 10.2: Verfahren zum halblauten Lesen differenziert nach hohem und niedrigem Anteil an leseschwachen Schülerinnen und Schülern in der Klasse (Anteile der Schülerinnen und Schüler nach Angaben nach Lehrkräften in Prozent)



IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

5 Förderung von Lesestrategien

In Klassen, in denen (fast) alle Schülerinnen und Schüler als fortgeschrittene Leserinnen und Leser angesehen werden können, ist die Förderung der Leseflüssigkeit gar nicht mehr beziehungsweise nur noch für wenige angezeigt. Für die Kinder wäre sie eine Unterforderung (Lauer-Schmaltz et al., 2014). Anders verhält es sich mit der Förderung des Wissens über Lesestrategien und ihren flexiblen Einsatz. Sie kann bereits in der Grundschule mit Aussicht auf Erfolg betrieben werden (z. B. Munser-Kiefer, 2014; Rosebrock & Nix, 2014).

Lesestrategien sind lehr- und lernbare Sequenzen von mehr oder weniger komplexen Operationen zur Erreichung von (Lern-)Zielen, die sich auf Bereiche beziehungsweise Domänen beziehen, nicht auf einzelne Aufgaben. Es gibt zahlreiche Vorschläge zur Klassifikation von Strategien (Artelt, 2000). Beschränkt man sich auf den kognitiven Aspekt, dann bietet sich eine Unterscheidung an zwischen

- Wiederholungsstrategien (z. B. Aufzeichnungen mehrmals durchlesen, Auswendiglernen),
- Elaborationsstrategien (Verknüpfung mit Vorwissen, Konstruktion von Analogien, Beispielen, kritischen Fragen),
- reduktiven Strategien und Organisationsstrategien (z. B. Unterstreichen von Wesentlichem, Zusammenfassen von verstreuten Informationen, Darstellung in veränderter Form, z. B. als Netz von Begriffen),
- metakognitiven Strategien (Planen, Überwachen und Regulation des Leseprozesses und Bewertung des Resultats).

Von Fall zu Fall sind Zuordnungen zu einzelnen Strategien erschwert beziehungsweise nicht möglich. So kann die Zusammenfassung eines Textes oder eines Abschnitts mit eigenen Worten sowohl als reduktiv-organisierend als auch als elaborativ anzusehen sein. Denn es werden als unwichtig angese-

hene Teile des Textes getilgt (Reduktion), es kann nötig sein, ursprünglich an verschiedenen Stellen platzierte Informationen zusammenzuführen, das heißt, Information neu zu organisieren. Wenn schließlich die komprimierte und reorganisierte Information mit eigenen Worten reformuliert werden soll, muss elaborativ verarbeitet, das heißt Vorwissen aktiviert werden, zum Beispiel Wissen um Bedeutungsbeziehungen auf Ebenen wie Wort und Satz.

Dass Lesestrategien nicht erst in den Schulformen der Sekundarstufe I, sondern bereits in den höheren Jahrgangsstufen der Grundschule Lerngegenstand sein sollten, wird in den Deutsch-Lehrplänen aller Länder der Bundesrepublik Deutschland herausgestellt.

Als Beispiele für die Anwendung primär reduktiv-organisierender Strategien sind dort etwa genannt: Zwischenüberschriften formulieren, zusammenfassen (Berlin, Brandenburg, Hamburg, Niedersachsen, Saarland, Sachsen, Thüringen), Fragen an den Text stellen und beantworten (Bremen, Rheinland-Pfalz, Mecklenburg-Vorpommern, Niedersachsen), zentrale Aussagen erfassen (Hamburg, Hessen, Nordrhein-Westfalen, Rheinland-Pfalz, Niedersachsen, Saarland, Thüringen). Leistungserwartungen, die primär mit elaborativen Strategien zu tun haben, sind zum Beispiel die folgenden: über Aspekte von Texten nachdenken und eigene Vorstellungen und Gedanken entwickeln (Baden-Württemberg, Hamburg, Hessen, Nordrhein-Westfalen, Saarland), aus Texten Schlussfolgerungen ziehen (Bremen, Mecklenburg-Vorpommern, Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz, Sachsen, Thüringen), Gedanken und Meinungen zu Texten äußern und Stellung beziehen (Bremen, Mecklenburg-Vorpommern, Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz, Niedersachsen, Saarland, Thüringen), Vermutungen zum Inhalt anstellen (Rheinland-Pfalz, Niedersachsen).

Schließlich sind metakognitive Strategien im Spiel, wenn unbekannte Wörter in Texten identifiziert beziehungsweise Unverstandenes durch Nachdenken, Nachschlagen, Nachfragen geklärt oder Verstehenshilfen genutzt werden sollen (Berlin, Brandenburg, Bremen, Mecklenburg-Vorpommern, Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz, Niedersachsen, Sachsen-Anhalt).

International betrachtet kann als Konsens angesehen werden, dass Lesestrategien direkt und explizit gelehrt werden sollten (Kamil et al., 2008; Slavin, Lake, Chambers, Cheung & Davis, 2009). Die systematische Sichtung von mehr als 200 empirischen Studien im Rahmen von NICHD (2000) ergab, dass sich insbesondere ein Training der Kombination von Strategien bewährte. In kooperativen Settings übernehmen die Schülerinnen und Schüler abwechselnd die Rollen von Lehrenden und Lernenden (ebd., S. 4–46, vgl. Artelt et al., 2007). Hattie (2013, S. 420) zufolge ergaben zwei Metaanalysen von 16 beziehungsweise 22 Studien im Mittel Effekte auf das Leseverständnis in Höhe von 0.74. Die bekannteste Version stammt von Palincsar und Brown (1984). In 15 halbstündigen Sitzungen hatten 117 Siebt- und Achtklässlerinnen und -klässler die Aufgabe, abwechselnd in der Rolle von Lehrenden und von Lernenden vier Strategien anzuwenden: Wortbedeutungen klären, Fragen generieren, voraussagen, wie es im jeweiligen Text weitergehen wird, und zusammenfassen. Es gibt diverse Varianten und Weiterentwicklungen dieses kooperativen Ansatzes. Die Lehrperson fungiert als Modell, das die Anwendung der Strategien mit Hilfe von Lautem Denken anhand von prototypischen Texten beziehungsweise Textabschnitten vorführt. In der Folge üben die Kinder in kleinen Gruppen, wobei jedes Kind im Verlauf der gemeinsamen Arbeit die Rolle des Experten für die verschiedenen Strategien übernimmt.

In Deutschland wurden Lesestrategietrainings bislang vornehmlich für die ersten Jahrgangsstufen der Sekundarstufe I konzipiert. Am bekanntesten ist „Wir werden Textdetektive“ (Gold, Mokhlesgerami, Rühl, Schreblowski & Souvignier, 2004). Hier werden sieben Detektivmethoden (bzw. Strategien und Techniken) eingeübt, und zwar die Organisationsstrategien „*Wichtiges unterstreichen*“ und „*Wichtiges zusammenfassen*“, die elaborativen Strategien „*Überschrift beachten*“ und „*Bildlich vorstellen*“ und die metakognitiven Strategien „*Umgang mit Textschwierigkeiten*“, „*Verstehen überprüfen*“ und „*Behalten überprüfen*“. Dieses Förderprogramm, ein Wiederholungsprogramm und eine auf lernschwache Kinder zugeschnittene Version wurden in mehreren Studien evaluiert. Es ergaben sich nachhaltige Zuwächse beim Strategiewissen, Transfereffekte im Hinblick auf das Leseverstehen konnten aber nicht durchgängig nachgewiesen werden (Gold, 2010). Bei einem kleinen Effekt ($d = 0.25$) profitierten Gymnasiastinnen und Gymnasiasten am meisten ($d = 0.34$), Haupt-, Real- und Gesamtschülerinnen und -schüler dagegen gewannen nicht hinzu. Am stärksten ($d = 0.44$) waren die Effekte bei leseschwachen Kindern, die mit dem reduzierten Programm „Wir werden Lesedetektive“ gefördert worden waren (Rühl & Souvignier, 2006).

Es liegen auch einige Programme zur Förderung von Lesestrategien in der Grundschule vor, zum Beispiel *Lesen im Leseteam trainieren* (Kirschhock & Munser-Kiefer, 2012), *Der Lese-Sportler* (Schulte & Souvignier, 2013) und *Käpt'n Carlo sucht den Leseschatz* (Spörer, Koch, Schünemann & Völlinger, 2016). Munser-Kiefer (2014) berichtet, dass eine Intervention auf der Basis des Ansatzes von Palincsar und Brown, das heißt des wechselseitigen Lehrens und Lernens der Strategien vorhersagen, Fragen stellen und beantworten, schwierige Wörter klären und zusammenfassen, bei der Versuchsgruppe zu einem deutlich größeren Zuwachs im Leseverständnis führte als bei der Vergleichsgruppe ($d = 0.42$). Es profitierten aber im Wesentlichen die mittleren Leistungsgruppen. Die Zuwächse unterschieden sich bei der schwächsten Gruppe weder bei einem literarischen noch bei einem Sachtext signifikant von denen in der Kontrollgruppe; die stärksten Drittklässlerinnen und Drittklässler gewannen nur bei einem Sachtext signifikant hinzu. Eine detailliertere Analyse mit Hilfe des IGLU-Modells des Leseverstehens ergab, dass die Vorteile vor allem die Verstehensprozesse betrafen, für die auf externes Wissen zurückzugreifen ist (komplexe Schlussfolgerungen, Interpretieren, Prüfen und Bewerten) (ebd., S. 324 ff.). Souvignier et al. (2016) befragten Lehrende und Kinder in 40 dritten und vierten Klassen unter anderem, ob ein Training anhand der drei Strategien Umgang mit Textschwierigkeiten, Vorhersagen und Zusammenfassen ein besseres Textverstehen bewirkt habe. Auf einer Skala von 1 (= stimmt gar nicht) bis 4 (= stimmt genau) ergab sich bei den Lehrenden ein Mittelwert von 3.2 ($SD = 0.7$), bei den Schülerinnen und Schülern von 3.4 ($SD = 0.5$) (ebd., S. 92). Koch und Spörer (2016, S. 112) berichten – allerdings nur für Fünftklässlerinnen und Fünftklässler (165 Schülerinnen und Schüler aus acht Klassen) – einen nachhaltigen Effekt ($d = 0.40$) zugunsten der Versuchsgruppe, die mit den vier von Palincsar und Brown (1984) vorgeschlagenen Strategien gearbeitet hatte. Befunde wie diese fügen sich ein in ein zentrales Ergebnis der Recherchen von Slavin et al. (2009). Sie fanden „extensive evidence supporting forms of cooperative learning in which students work in small groups to help one another master reading skills and in which the success of the team depends on the individual learning of each team member“ (ebd., S. 1453).

Kooperatives Lernen in kleinen Gruppen ist aber weder in Deutschland noch in den anderen an IGLU 2016 teilnehmenden Staaten und Regionen vorherr-

schend. In Deutschland werden 26 Prozent der Schülerinnen und Schüler von Lehrkräften unterrichtet, die angeben, „*immer*“ oder „*fast immer*“ die ganze Klasse gemeinsam im Lesen zu unterrichten (internationaler MW = 32%). Nimmt man diejenigen hinzu, die „*oft*“ ankreuzen, kommt man auf 71 Prozent der Schülerinnen und Schüler. Auf die Arbeit in homogenen oder heterogenen Gruppen setzt eine deutliche Mehrheit der Befragten allenfalls „*manchmal*“ oder aber „*nie*“: 69 beziehungsweise 72 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die entsprechend antworten.

Die Lehrkräfte in Deutschland legen wie die in den anderen Teilnehmerstaaten und -regionen vor allem Wert darauf, dass die Kinder Informationen im Text lokalisieren können. Ihren Angaben zufolge setzen sich 96 Prozent der Schülerinnen und Schüler jeden oder fast jeden Tag beziehungsweise ein- bis zweimal pro Woche mit einschlägigen Aufgaben auseinander. Diese Quote ist identisch mit dem internationalen Mittelwert. Erhebliche Divergenzen werden aber erkennbar, wenn es um die Förderung wissensbasierter Verstehensleistungen und damit auch um den Einsatz elaborativer Strategien geht. So wurde international gefragt, wie oft es vorkommt, dass die Kinder das Gelesene mit eigenen Erfahrungen oder auch mit anderweitig Gelesenem vergleichen müssen. Die Optionen „*jeden oder fast jeden Tag*“ und „*ein- bis zweimal pro Woche*“ kreuzten im internationalen Mittel 83 ($SE = 0.4$) beziehungsweise 75 ($SE = 0.4$) Prozent der befragten Lehrkräfte an, in Deutschland aber nur 69 ($SE = 3.1$) beziehungsweise 45 Prozent ($SE = 3.8$). Noch deutlicher sind die Diskrepanzen bei Aufgaben, die nicht in erster Linie auf Textinhalte, sondern auf den Stil und die Struktur von Texten zielen. In internationaler Perspektive bekommen solche Aufgaben 69 Prozent ($SE = 0.4$) der Schülerinnen und Schüler einmal pro Woche gestellt; in Deutschland sind es nur 24 Prozent ($SE = 2.9$). Auch Aufgaben, in deren Rahmen die Perspektive oder Absicht des Textes beziehungsweise des Autors zu ermitteln sind, werden von Schülerinnen und Schülern im Deutschunterricht mit wenigstens einmal wöchentlich deutlich seltener bearbeitet als im internationalen Durchschnitt. Der internationale Mittelwert liegt hier bei 66 Prozent ($SE = 0.4$), in Deutschland beträgt er die Hälfte (32%, $SE = 3.0$) (vgl. Mullis et al., 2017, Exhibit 9.2).

Die Items im internationalen Fragebogen für Lehrkräfte, die sich darauf beziehen, wie häufig sie sich im Unterricht um die Entwicklung von Lesefertigkeiten und die Anwendung von Lesestrategien kümmern, wurden auf nationaler Ebene ergänzt. Es wurden mehrere Items zur Vermittlung von Lesestrategien hinzugefügt (siehe Tabelle 10.3).

Wenn eine Lehrkraft immer wieder zeigt, wie man verschiedene Lesestrategien einsetzt, präsentiert sie sich als Modell-Leserin, an der sich die Schülerinnen und Schüler orientieren können. Ihre Erklärungen gelten vor allem dem Nutzen, den die Anwendung von Strategien von Fall zu Fall haben kann. Dabei ist nicht die gleichsam mechanische Abarbeitung eines Sets von Strategien intendiert, sondern es soll die Erkenntnis angebahnt werden, dass sie flexibel eingesetzt werden sollten. Lautet zum Beispiel die Überschrift eines Textes „Ein Unglück“, ist es nicht sinnvoll, sich zunächst an Unglücke aller Art zu erinnern. Geht es aber um Texte mit Titeln wie „Das Dromedar“ oder „Die mittelalterliche Burg“, ist es angebracht, vor dem Lesen das, was man dazu bereits weiß, zu notieren und im Anschluss an die Lektüre zu bedenken, ob der Text Neues enthält, ob Elemente des Vorwissens zu korrigieren sind und so weiter.

Tabelle 10.3: Entwicklung von Lesefertigkeiten im Unterricht (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent)

Den Schülern erklären und zeigen, ...	jeden oder fast jeden Tag		ein- bis zweimal die Woche		ein- bis zweimal pro Monat		nie oder fast nie	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)
wie man verschiedene Lesestrategien einsetzt.	7.7	(2.4)	36.7	(4.0)	48.2	(4.2)	7.4	(2.1)
wann man verschiedene Lesestrategien einsetzt.	5.1	(1.8)	32.4	(3.8)	52.9	(4.4)	9.5	(2.3)
wie man Fragen zum Text entwickelt, während man liest.	3.8	(1.4)	28.6	(3.6)	50.4	(4.4)	17.3	(2.8)
wie man zentrale Ideen eines Textes zusammenfasst.	10.2	(2.4)	46.5	(3.9)	39.5	(3.8)	3.7	(1.4)
wie sie den Textinhalt mit ihrem Vorwissen verbinden können.	7.5	(2.2)	39.3	(3.9)	47.7	(3.4)	5.5	(1.6)
wie man wichtige Textstellen markiert (z. B. durch Unterstreichen).	21.6	(3.2)	55.7	(4.3)	22.1	(3.4)	0.6	(0.4)
dass man wichtige Textstellen manchmal wörtlich abschreiben sollte.	3.7	(1.5)	29.4	(3.5)	44.3	(3.9)	22.6	(3.2)
wie sie während des Lesens Voraussagen über den weiteren Verlauf des Textes anstellen.	1.8	(1.1)	25.4	(3.0)	53.8	(4.0)	19.0	(2.9)
wie sie überprüfen, ob ihnen alle Wortbedeutungen bekannt sind.	23.7	(3.7)	27.0	(3.4)	39.7	(3.9)	9.6	(2.2)
wie sie den Textaufbau mit dem von anderen Texten vergleichen.	1.8	(1.1)	12.8	(2.6)	50.5	(4.0)	34.8	(3.9)
wie sie Informationen im Text über Abschnittsgrenzen hinweg verknüpfen.	4.8	(1.8)	14.3	(2.7)	44.5	(4.2)	36.4	(3.8)
wie sie Überschriften zu Textabschnitten formulieren.	4.2	(1.6)	33.6	(3.6)	53.3	(4.3)	8.9	(2.2)
wie sie wichtige Informationen in ein anderes Darstellungsformat überführen (z. B. Mind-Map, Tabelle, Diagramm).	1.9	(1.1)	15.9	(3.0)	56.3	(3.6)	25.9	(3.2)

Differenzen zu 100 Prozent ergeben sich durch Rundungsfehler.

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Der Tabelle 10.3 ist zu entnehmen, dass nur eine Minderheit der Kinder von Lehrkräften unterrichtet wird, die in nennenswertem Umfang explizit lehren, wie und wann Lesestrategien einzusetzen sind. 56 beziehungsweise 62 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die die Optionen „ein- bis zweimal pro Monat“ beziehungsweise „nie“ angekreuzt haben. Die im Modell des reziproken Lehrens und Lernens von Palincsar und Brown (1984) zentralen prozessbezogenen Strategien des Fragenstellens und des Voraussagens, wie es im Text wohl weitergehen wird, werden den Angaben der Lehrkräfte zufolge nur für 32 beziehungsweise 27 Prozent der Viertklässlerinnen und Viertklässler mindestens ein- bis zweimal pro Woche thematisch. Auch die primär reduktiv-organisierenden Strategien der Verknüpfung von Informationen über Abschnittsgrenzen hinweg, der Formulierung von Teilüberschriften und der Transformation in ein anderes Darstellungsformat werden nur einer Minderheit der Kinder in beachtlicherem Umgang – mindestens einmal wöchentlich – vermittelt, nämlich 14, 34 und 16 Prozent. Nur das Zusammenfassen, eine ebenfalls primär reduktiv-organisierende Strategie, und die damit verbundene Technik des Unterstreichens werden häufiger praktiziert. 57 beziehungsweise 77 Prozent der Schülerinnen und Schüler in Deutschland werden von Lehrkräften unterrichtet,

die diese Tätigkeiten jeden oder fast jeden Tag beziehungsweise ein- bis zweimal pro Woche im Unterricht einfordern. Die Differenzen deuten darauf hin, dass das Markieren nicht immer im Dienst des Zusammenfassens steht. Die elaborative Strategie, Textinhalte mit dem Vorwissen zu verknüpfen, wird von fast der Hälfte der Unterrichtenden (47%) in nennenswertem Umfang explizit vermittelt. Die für die elaborative Verarbeitung von Lesetexten konstitutive Verknüpfung von Textinhalt und Vorwissen wird demnach in mehr als der Hälfte der Klassen allenfalls ein- oder zweimal im Monat thematisch. Die Viertklässlerinnen und Viertklässler müssen auch lernen, sich beim Lesen gleichsam über die Schulter zu schauen. Es sind metakognitive Strategien im Spiel, wenn während des Leseprozesses lexikalische Lücken erkannt und gegebenenfalls mit Hilfe von Lexika, Mitschülerinnen und Mitschülern oder auch der Lehrkraft geschlossen werden sollen. Die Hälfte der Schülerinnen und Schüler wird von Lehrkräften unterrichtet, die explizit mindestens einmal pro Woche vermitteln, wie man lexikalische Hürden erkennen und überwinden kann. Bei der anderen Hälfte kommt das deutlich seltener oder auch gar nicht vor. Bereits der internationale Vergleich zeigt, dass die Kinder in Deutschland selten mit der anspruchsvollen, elaborativen Strategien erfordernden Aufgabe konfrontiert sind, nicht den Inhalt, sondern die Struktur von Texten zu ermitteln. Gelegenheiten, Strukturen miteinander zu vergleichen und etwa zu lernen, dass zwei narrative Texte sehr verschiedenen Inhalts in struktureller Hinsicht (nahezu) identisch sein können, haben sie erst recht nicht: 85 Prozent der Schülerinnen und Schüler wird nach Angaben der Lehrkräfte nicht explizit vermittelt, wie man den Aufbau eines Textes mit dem von anderen vergleichen kann.

6 Zusammenfassung und Bewertung

Den Angaben der Lehrkräfte zufolge sind etwa 17 Prozent der Viertklässlerinnen und Viertklässler förderbedürftig, im Mittel pro Klasse drei bis vier Kinder. Bedenkt man, dass – neben einem hinreichend großen passiven Wortschatz (siehe Kapitel 11 in diesem Band) – flüssiges Lesen eine notwendige Bedingung für Textverstehen ist, dann liegt es nahe, bei diesen Kindern zunächst auf die Förderung der Komponenten der Leseflüssigkeit – Dekodiergenauigkeit, Geschwindigkeit, Automatisierung und prosodisch phrasiertes Lesen – zu setzen. Berücksichtigt man die Resultate der internationalen Forschung, dann dürfte das begleitete und wiederholte laute Lesen die am besten geeignete methodische Variante sein. Sie ist dem nach wie vor sehr verbreiteten lauten Reihum-Lesen und dem stillen Viellesen überlegen, wird in Deutschland aber nur von etwa der Hälfte der Lehrkräfte praktiziert, wobei auf Erfolgskontrolle verzichtet wird. Ob beziehungsweise in welchem Umfang die Leseflüssigkeit gefördert werden sollte, hängt von der Komposition der jeweiligen Klasse ab. Dass für die Förderung je nach den Bedingungen vor Ort nicht der Regelunterricht, sondern separate Settings geeignet sein können (Lauer-Schmaltz et al., 2014), liegt auf der Hand.

Für fortgeschrittene Leserinnen und Leser ist, so der Tenor der einschlägigen Forschung, die mehr und mehr automatisierte, flexible Anwendung von Lesestrategien ein zentrales Lernziel. In den länderübergreifenden Bildungsstandards und in den Curricula aller Länder der Bundesrepublik Deutschland sind sie verankert. Es zeigt sich aber, dass die Lehrkräfte, summarisch gesagt, der expliziten Lehre sowohl der reduktiv-organisierenden als auch der

elaborativen Lesestrategien wenig Aufmerksamkeit widmen. Insofern gibt es hier eine erhebliche Kluft zwischen dem intendierten und dem realisierten Curriculum. Es liegt nahe, diesen Befund mit dem vergleichsweise schwachen Abschneiden der Viertklässlerinnen und Viertklässler in Deutschland im Rahmen von IGLU 2016 in Verbindung zu bringen. Für die Lösung von Aufgaben auf den Kompetenzstufen III bis V, die verständige Verknüpfung von verstreuten Informationen und die Integration von Textinformation und Vorwissen, sind diese Strategien ja hilfreich.

Was das „Schicksal“ von in der Forschung vielfach positiv evaluierten Konzepten der Leseförderung angeht, so resümieren Souvignier et al. (2016) „Betrachtet man die unterrichtliche Praxis, dann weisen Unterrichtsbeobachtungen über die vergangenen 30 Jahre konsistent darauf hin, dass die genannten Methoden – Lautes Lesen [wie ‚paired reading‘, ‚repeated reading‘, d. Verf.] und strategieorientierter Unterricht – kaum Gegenstand alltäglicher Unterrichtspraxis sind [...]“ (S. 79). Es gibt keinen Anlass, diese Schlussfolgerung im Licht der Ergebnisse von IGLU 2016 zu revidieren.

Literatur

- Artelt, C. (2000). *Strategisches Lernen*. Münster: Waxmann.
- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., Schneider, W., Stanat, P., Ostermeier, C., Schiefele, U., Valtin, R., Ring, K. & Saalbach, H. (2007). *Förderung von Lesekompetenz. Expertise* (Bildungsforschung Bd. 17). Bonn: Bundesministerium für Bildung und Forschung.
- Dehaene, S. (2010). *Lesen – Die größte Erfindung der Menschheit und was dabei in unseren Köpfen passiert*. München: Knaus.
- Gasteiger, C. & Gasteiger-Klicpera, B. (1993). *Lesen und Schreiben – Entwicklung und Schwierigkeiten: Die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit*. Bern: Huber.
- Gold, A. (2010). *Lesen kann man lernen. Lesestrategien für das 5. und 6. Schuljahr*. Göttingen: Vandenhoeck & Ruprecht.
- Gold, A., Mokhlesgerami, J., Rühl, K., Schreblowski, S. & Souvignier, E. (2004). *Wir werden Textdetektive [Lehrermanual und Arbeitsheft]*. Göttingen: Vandenhoeck & Ruprecht.
- Hattie, J. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider Hohengehren.
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T. & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide*. Washington, D.C: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.
- Kirschhock, E.-M. & Munser-Kiefer, M. (2012). *Lesen im Leseteam trainieren. Das Materialpaket zum Programm*. Donauwörth: Auer.
- Koch, H. & Spörer, N. (2016). Effekte des Reziproken Lehrens im Vergleich mit einer von Lehrkräften konzipierten Unterrichtseinheit zur Förderung der Lesekompetenz. In M. Philipp & E. Souvignier (Hrsg.), *Implementation von Lesefördermaßnahmen* (S. 99–121). Münster: Waxmann.
- Lauer-Schmaltz, M., Rosebrock, C. & Gold, A. (2014). Lautlesetandems in der Grundschule. *Didaktik Deutsch* 19 (37), 45–61.
- Lenhard, W. & Schneider, W. (2006). *ELFE 1–6. Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Mayer, A. (2012). *Blitzschnelle Worterkennung*. Dortmund: Borgmann Media.
- Mayringer, H. & Wimmer, H. (2003). *SLS 1–4. Salzburger Lese-Screening für die Klassenstufen 1–4*. Bern: Huber.

- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Boston College, TIMSS & PIRLS International Study Center. Zugriff am 23.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/international-results/>
- Munser-Kiefer, M. (2014). *Leseförderung im Leseteam in der Grundschule. Eine Interventionsstudie zur Förderung von basaler Lesefertigkeit und (meta-)kognitiven Lesestrategien*. Münster: Waxmann.
- NICHD – National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: U.S. Government Printing Office.
- Nix, D. (2011). *Förderung der Leseflüssigkeit. Theoretische Fundierung und empirische Überprüfung eines kooperativen Lautlese-Verfahrens im Deutschunterricht*. Weinheim: Juventa.
- Palincsar, A. & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1 (2), 117–175.
- Rosebrock, C. & Nix, D. (2014). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (7., überarbeitete und erweiterte Auflage). Baltmannsweiler: Schneider Hohengehren.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2013). *Leseflüssigkeit fördern. Lautleseverfahren für die Primar- und Sekundarstufe* (2. Auflage). Seelze: Kallmeyer/Klett.
- Rühl, K. & Souvignier, E. (2006). *Wir werden Lesedetektive – Lehrermanual und Arbeitsheft*. Göttingen: Vandenhoeck & Ruprecht.
- Schrader, F.-W. (2011). Lehrer als Diagnostiker. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 683–698). Münster: Waxmann.
- Schulte, E. & Souvignier, E. (2013). Der Lese-Sportler. Ein Programm für individuelle Leseförderung. *Die Grundschulzeitschrift*, 27 (7), 58–61.
- Slavin, R. E., Lake, C., Chambers, B., Cheung, A. & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research*, 79 (4), 1391–1466.
- Souvignier, E., Förster, N. & Kawohl, E. (2016). Implementation eines Förderkonzeptes zur diagnosebasierten individuellen Leseförderung in der Grundschule. In M. Philipp & E. Souvignier (Hrsg.), *Implementation von Lesefördermaßnahmen* (S. 77–98). Münster: Waxmann.
- Spinner, K. H. (2004). Lesekompetenz in der Schule. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (S. 125–138). Wiesbaden: Verlag für Sozialwissenschaften.
- Spörer, N., Koch, H., Schünemann, N. & Völlinger, V. (2016). *Das Lesetraining mit Käpt'n Carlo für 4. und 5. Klassen. Ein Lehrermanual mit Unterrichtsmaterialien zur Förderung des verstehenden und motivierten Lesens*. Göttingen: Hogrefe.
- Tarelli, I., Lankes, E.-M., Drossel, K. & Gegenfurtner, A. (2012). Lehr- und Lernbedingungen an Grundschulen 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. 137–173). Münster: Waxmann.
- Valtin, R., Naegele, I. & Sasse, A. (2013). Wie sollte ein erfolgreiches Material für Kinder mit LRS aussehen? Theoretische Grundlagen und ein praktisches Beispiel: „Das schaffe ich!“ In F. Hellmich & K. Siekmann (Hrsg.), *Sprechen, Lesen und Schreibenlernen. Erfolgreiche Konzepte der Sprachförderung* (S. 200–215). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.
- von Wedel-Wolff, A. (2002). *Üben im Leseunterricht der Grundschule*. Braunschweig: Westermann.

Kapitel XI

Bausteine adaptiven Leseunterrichts angesichts gewachsener Heterogenität

Albert Bremerich-Vos, Heike Wendt und Anke Hußmann

1 Einleitung

Wie in den Kapiteln zuvor berichtet, besteht in Deutschland am Ende der Grundschulzeit eine noch größere Leistungsheterogenität im Lesen, als dies in den 15 Jahren zuvor der Fall war. Die Anteile sowohl der leseschwachen als auch der lesestarken Kinder sind größer geworden (siehe Kapitel 3 in diesem Band). Zu denken gibt insbesondere der nun auf 18,9 Prozent gestiegene Anteil an Schülerinnen und Schülern, die am Ende der Grundschulzeit keine Leistungen auf Kompetenzstufe III erreichen. Es ist davon auszugehen, dass sie in der Sekundarstufe I in allen Fächern mit erheblichen Schwierigkeiten konfrontiert sein werden. Insbesondere im Hinblick auf die Förderung dieser Kinder besteht somit dringender Handlungsbedarf.

Soll der gewachsenen Heterogenität Rechnung getragen werden, ist adaptives Unterrichten angezeigt. Wer adaptiv lehrt, ist fähig, den „Unterricht so auf die individuellen Voraussetzungen der Schülerinnen und Schüler auszurichten und während des Unterrichts laufend anzupassen, dass für möglichst viele Schülerinnen und Schüler günstige Bedingungen für das Erreichen der Lernziele geschaffen werden [...]“ (Brühwiler, 2014, S. 74) Der damit formulierte Anspruch ist hoch, und ob beziehungsweise in welchem Maß er einzulösen ist, hängt nicht zuletzt von der Zusammensetzung der jeweiligen Klasse ab.

Ein bereits mehr oder weniger ausgeprägtes Heterogenitätsspektrum in einer Klasse wird in der Regel noch breiter, wenn Kinder mit sonderpädagogischem Förderbedarf (SPF) hinzukommen. 2009 ratifizierte die Bundesrepublik Deutschland die UN-Konvention über die Rechte von Menschen mit Behinderungen aus dem Jahr 2006. Bund und Länder verpflichteten sich, „den Grundsatz der Inklusion von Menschen mit Behinderungen auch im Schulbereich umzusetzen und das allgemeine Schulsystem in Deutschland zu einer inklusiven Schullandschaft weiterzuentwickeln“ (KMK, 2015, S. 6). In allen

Ländern der Bundesrepublik Deutschland werden seither Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) zunehmend an allgemeinbildenden Schulen unterrichtet (KMK, 2011). Bei der Einstufung der Kinder spielen amtliche Feststellungsverfahren eine wesentliche Rolle, die je nach Land zum Beispiel im Hinblick auf die daran beteiligten Akteure und die Methoden der Diagnostik unterschiedlich gestaltet sein können. Teilweise wird von einem solchen Verfahren aber auch abgesehen, und zwar dann, wenn es um die Förderschwerpunkte „Lernen“, „Sprache“ und „emotionale und soziale Entwicklung“ geht. Anders als bei den Förderschwerpunkten „Sehen“, „Hören“ und „körperliche und motorische Entwicklung“ ist hier die Abgrenzung von „normaler“ und „beeinträchtigter“ Entwicklung schwierig beziehungsweise problematisch. Die Etikettierung „sonderpädagogischer Förderbedarf“ ist also mit Vorsicht zu interpretieren, insofern „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.“ (Gresch, Kölm & Kocaj, 2017, S. 288).

Vorsicht ist auch geboten, wenn es um die Attestierung von Lese- und Rechtschreibschwierigkeiten (LRS) geht. Schülerinnen und Schüler mit LRS gelten nicht als (allgemein) „lernbehindert“. Bei ihnen liegt „nur“ eine „Teilleistungsstörung“ in den Bereichen Lesen und/oder Rechtschreiben vor. Seit Jahrzehnten wird über „Legasthenie“, „Lese-Rechtschreib-Störung“, „Lese-Rechtschreib-Schwäche“, „Lese-Rechtschreib-Schwierigkeiten“ kontrovers diskutiert. Die Differenzen sind nicht nur terminologischer, sondern auch konzeptioneller Natur. Unter Vernachlässigung einer Reihe von Details kann man zwei Grundmodelle unterscheiden. Einem älteren „medizinischen“ Modell zufolge ist bei Kindern nur dann von einer „Legasthenie“ zu sprechen, wenn eine Diskrepanz zwischen (sehr) schwachen Lese- und/oder Rechtschreibleistungen auf der einen und mindestens durchschnittlichen Leistungen bei einem (nicht auf Verbales bezogenen) Intelligenztest auf der anderen Seite konstatiert werden kann. In einer Reihe von Ländern – Bayern, Berlin, Hamburg, Mecklenburg-Vorpommern, Niedersachsen, Schleswig-Holstein und Sachsen – sind für die Diagnose Ärzte zu konsultieren beziehungsweise ist eine Testung der kognitiven Leistungsfähigkeit obligatorisch.

Allerdings wird selbst in der primär medizinisch ausgerichteten Forschung mittlerweile nicht mehr (ausschließlich) auf das IQ-Diskrepanz-Modell gesetzt. Davon zeugen die 2015 von der *Deutschen Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie e.V.* (DGKJP) erarbeiteten Leitlinien zur „Diagnostik und Behandlung von Kindern und Jugendlichen mit Lese- und/oder Rechtschreibstörung“. Expertinnen und Experten stimmten zu etwa 60 Prozent der folgenden Empfehlung zu: „Zur Diagnostik der Lese und/oder Rechtschreibstörung soll auf das Kriterium der Alters- oder Klassennormdiskrepanz oder (sic) auf das Kriterium der IQ-Diskrepanz zurückgegriffen werden“ (ebd., S. 25). Hier gehen also auch „nur“ leistungsschwache Kinder in die Definition ein (zur Problematik der Leitlinien aus pädagogischer Sicht vgl. Valtin, 2015).

Dem wird in anderen Ländern (Baden-Württemberg, Brandenburg, Hessen, Nordrhein-Westfalen, Rheinland-Pfalz, Thüringen) Rechnung getragen, in denen LRS nur leistungsbezogen definiert wird. Es handelt sich demnach um Schülerinnen und Schüler, denen in den ersten beiden Klassen die grundlegenden Voraussetzungen für den Schriftspracherwerb noch fehlen beziehungsweise deren (Lese-)Leistungen später über einen längeren Zeitraum den Anforderungen nicht

entsprechen. Für die Diagnostik kommen prozessbegleitende Beobachtungen, aber auch normierte Tests in Frage.

Die Vielzahl der Definitionen und der Operationalisierungen der Konstrukte in Form von Angaben zu Prozentranggrenzen bei normorientierten Tests beleuchten Valtin, Voss und Bos (2015) kritisch. Sie kommen – wie vermehrt auch andere Autorinnen und Autoren auf nationaler und internationaler Ebene – zu dem Ergebnis, „dass die Leseleistungen der Kinder auf einem Kontinuum zu verorten sind und die Abgrenzungen verschiedener Gruppen höchst willkürlich sind [...]“ (ebd., S. 55, vgl. z.B. auch Scheerer-Neumann, 2015).

Es bleibt der Anspruch, alle Kinder je nach ihren Lernvoraussetzungen zu fördern, das heißt adaptiv zu unterrichten.

Im Folgenden werden zunächst einige Befunde aus IGLU 2016 zur Heterogenität vierter Klassen (siehe Abschnitt 2) und Reaktionen in Form von äußerer und innerer Differenzierung in Deutschland berichtet (siehe Abschnitt 3). Mit dem Fokus auf adaptives Lehren in besonders herausfordernden Klassen folgen Ausführungen, die sich auf die Förderung des phonologischen Rekodierens, die Arbeit am Wortschatz, den Umgang mit Texten in Leichter Sprache und die Diagnose von Lernverläufen beziehen (siehe Abschnitt 4). Ein Fazit schließt das Kapitel ab (siehe Abschnitt 5).

2 Aspekte von Heterogenität an Grundschulen in Deutschland

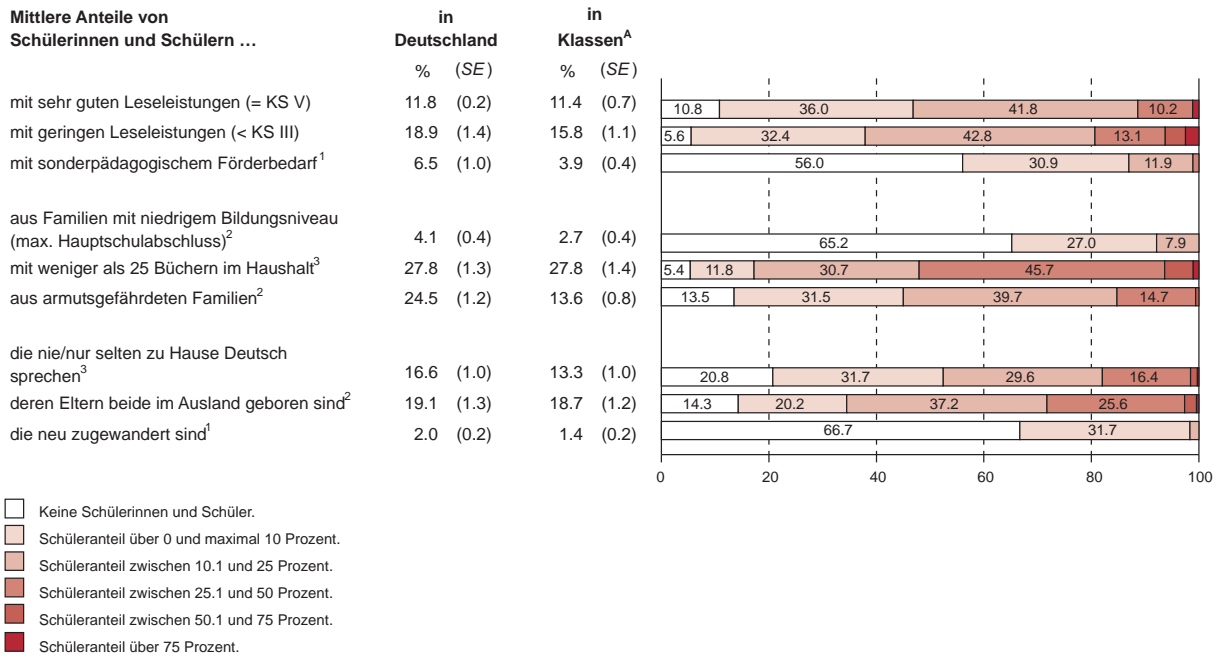
Heterogenität an Grundschulen in Deutschland ist je nach Schulstandort unterschiedlich ausgeprägt. Vierte Klassen unterscheiden sich erheblich hinsichtlich ihrer Zusammensetzung. Dies betrifft sowohl die leistungsbezogene Heterogenität, diagnostizierte sonderpädagogische Förderbedarfe als auch soziale oder migrationsbedingte Schülermerkmale (siehe Abbildung 11.1).

Was die Leseleistungen angeht, so liegt der Anteil der Kinder auf der höchsten Kompetenzstufe V in 42 Prozent aller Grundschulen zwischen 10 und 25 Prozent, in weiteren 12 Prozent sogar darüber. In mehr als jeder dritten Grundschule macht er maximal 10 Prozent aus, in etwa jeder zehnten Grundschule aber erreicht überhaupt kein Schüler beziehungsweise keine Schülerin dieses Leistungsniveau.

In 6 Prozent der Schulen finden sich keine Kinder mit Leseleistungen auf den Kompetenzstufen I und II. In 32 Prozent aller Grundschulen beträgt ihr Anteil maximal 10 Prozent, in 43 Prozent zwischen 10 und 25 Prozent, in weiteren 13 Prozent zwischen 25 und 50 Prozent. In etwas mehr als 6 Prozent aller Grundschulen kommt mindestens die Hälfte der Viertklässlerinnen und -klässler über die beiden unteren Kompetenzstufen nicht hinaus.

Kinder mit diagnostiziertem sonderpädagogischem Förderbedarf finden sich in mehr als der Hälfte (56 %) der Klassen gar nicht, in 12 Prozent ist ihr Anteil hingegen bedeutsam (10 bis 25 %). Laut Schulstatistik der Kultusministerkonferenz (KMK, 2017a) wurden im Schuljahr 2015/2016 in Deutschland insgesamt 517 000 Schülerinnen und Schüler sonderpädagogisch gefördert, unabhängig davon, ob ein Förderbedarf förmlich festgestellt wurde oder nicht. Damit betrug die Förderquote 7.1 Prozent, wovon auf Förderschulen 4.4 und auf allgemeine Schulen 2.7 Prozent entfielen. Die Hälfte aller in allgemeinen Schulen geförderten Schülerinnen und Schüler besuchte die Grundschule, in absoluten

Abbildung 11.1: Mittlere Anteile von Schülerinnen und Schülern nach unterschiedlichen Diversitätsmerkmalen in Klassen der vierten Jahrgangsstufe an Regelschulen in Deutschland (Schulen nach Zusammensetzungsmerkmalen in Prozent)



A = Aggregierte Schülermerkmale (mittlerer Schüleranteil in Klassen der 4. Jahrgangsstufe in Deutschland).
 1 = Schülerspezifische Angabe von Lehrkräften in der „Klassenliste“ zu diagnostizierten sonderpädagogischen Förderbedarfen.
 2 = Elternangabe.
 3 = Schülerangabe.

Zahlen 99 000 von 195 000 (KMK, 2017b). Für sie waren am häufigsten die Förderschwerpunkte Lernen (41 000 Kinder), emotionale und soziale Entwicklung (21 000 Kinder) und Sprache (17 000 Kinder) angegeben. Etwa 80 Prozent der Kinder wurden demnach in einem dieser drei Bereiche gefördert. Dem *IQB-Bildungstrend 2016* zufolge machen die Kinder, die in diesen drei Bereichen einen amtlich festgestellten Förderbedarf haben, in vierten Klassen insgesamt 5.6 Prozent aller Schülerinnen und Schüler aus. 54 Prozent davon besuchen eine allgemeine Schule, 46 Prozent eine Förderschule (Gresch et al., 2017, S. 285 f.). Aufgrund der geringeren Stichprobe und besonderen Studienteilnahmebedingungen ist zu vermuten, dass der in Abbildung 11.1 berichtete Anteil etwas unterschätzt wird. Hinzu kommen Kinder, die aufgrund pauschaler Ressourcenzuweisung sonderpädagogisch gefördert werden, ohne dass für sie ein entsprechendes Gutachten vorliegt. Sie sind, folgt man den Angaben der Schulleitungen im Rahmen von IGLU 2016, in zwei Dritteln der Grundschulen (64 %) in Deutschland anzutreffen.

Differenziert nach Merkmalen der sozialen Herkunft der Schülerinnen und Schüler (siehe auch Kapitel 6 in diesem Band) zeigt sich, dass Kinder aus Familien mit niedrigem Bildungsniveau in knapp zwei Drittel aller Klassen in Deutschland nicht vertreten sind, in 27 Prozent mit einem Anteil von maximal 10 Prozent, in 8 Prozent mit einer Quote zwischen 10 und 25 Prozent. In fast der Hälfte der Klassen sind zwischen einem Viertel und der Hälfte Kinder anzutreffen, bei denen es zuhause weniger als 25 Bücher gibt, in 17 Prozent macht deren Anteil höchstens 10 Prozent der Kinder aus. Substantielle Anteile – das heißt Anteile im Umfang von 10 bis 25 beziehungsweise von 25 bis 50 Prozent

– von Kindern aus armutsgefährdeten Familien gibt es in knapp 40 beziehungsweise 15 Prozent der Klassen.

Was Kinder mit Zuwanderungshintergrund angeht, so interessieren im Hinblick auf den Leseunterricht unter anderem diejenigen, die angeben, zuhause nie oder fast nie Deutsch zu sprechen (siehe Kapitel 7 in diesem Band). Aggregiert man dieses Schülermerkmal auf Klassenebene, sind solche Kinder in etwa einem Fünftel der Klassen gar nicht zu finden. In einem weiteren Drittel der Grundschulen macht ihr Anteil höchstens 10 Prozent aus. In 30 Prozent aller Grundschulen liegt ihr Anteil zwischen 10 Prozent und einem Viertel der Schülerschaft. In jeder sechsten Grundschule in Deutschland (16 %) spricht etwa jedes vierte bis zweite Kind zu Hause vornehmlich eine andere Sprache als Deutsch. Vereinzelt trifft dies sogar auf mindestens jedes zweite Kind zu.

Betrachtet man Subgruppen der Kinder mit Zuwanderungshintergrund, dann sind weitere Disparitäten zu berichten: Der Anteil der Kinder, deren Eltern beide im Ausland geboren wurden, liegt an Grundschulen in Deutschland in etwa einem Drittel der Klassen bei maximal 10 Prozent, in etwas mehr als einem weiteren Drittel zwischen 10 und 25 Prozent, bei einem Viertel zwischen 25 und 50 Prozent. In wenigen Grundschulen gibt es sogar Klassen, in denen die Quote dieser Kinder zwischen 50 und 75 Prozent liegt.

63 Prozent aller Schulleitungen an Grundschulen in Deutschland gaben im Sommer 2016 an, dass an ihrer Schule neu zugewanderte Kinder beziehungsweise Kinder aus Flüchtlingsfamilien unterrichtet werden. Es handelt sich um durchschnittlich 11 Kinder pro Schule. Im bundesdeutschen Mittel betrug ihr Anteil 2 Prozent, aggregiert auf Klassenebene 1.4 Prozent. Für gut zwei Drittel aller Grundschulen ist festzuhalten, dass in der vierten Jahrgangsstufe keine Integration von neu zugewanderten Schülerinnen und Schülern erfolgt war. An etwa jeder dritten Schule wurden neu zugewanderte Kinder unterrichtet, hier betrug ihr Anteil maximal 10 Prozent, ausgesprochen selten war er höher.

Berücksichtigt man die hier genannten Dimensionen von Heterogenität, wird deutlich, dass die Herausforderungen im Hinblick auf adaptives Unterrichten je nach Zusammensetzung der Klasse vor Ort stark variieren können. Internationale Befunde sprechen dafür, dass die Komposition der Klasse nicht nur die Leistungen der einzelnen Schülerinnen und Schüler (Dumont, Neumann, Maaz & Trautwein, 2013) oder Übergangentscheidungen beeinflusst (Lintorf, Guill & Wendt, 2017), sondern auch Auswirkungen zum Beispiel auf das Selbstkonzept, die Motivation und die Lernfreude hat. Intensiv untersucht wurde beispielsweise der sogenannte *Big-Fish-Little-Pond-Effekt*. Er besteht darin, dass bei Konstanzhaltung der individuellen Leistung in leistungsstarken Klassen das individuelle Fähigkeitsselbstkonzept sinkt. Die Schülerinnen und Schüler vergleichen ihre Leistungen mit denen ihrer Mitschülerinnen und -schüler in der jeweiligen Klasse. In einer leistungsstarken Klasse sehen sie sich dann als eher „kleiner Fisch“. In Deutschland beziehen sich Studien zu Effekten der Klassenzusammensetzung auf Schulleistungen bislang vor allem auf Schülerinnen und Schüler in der Sekundarstufe I (Dumont, Neumann, Maaz & Trautwein, 2013). Hier ist zusätzlich der Einfluss der Schulform zu bedenken. Auch für den Grundschulbereich liegen mittlerweile einige Studien zu Zusammenhängen zwischen Merkmalen der Klassenzusammensetzung und der Leseleistung am Ende der vierten Jahrgangsstufe vor (Ammermüller & Pischke, 2009; Bellin, Dunge & Gunzenhauser, 2010; Martin, Foy, Mullis & O'Dwyer, 2013; Kristen, 2008; Smith, Wendt & Kasper, 2016). Aufgrund des querschnittlichen Designs besitzen sie allerdings nur begrenzte Aussagekraft, können

aber bei mehrebenenanalytischer Modellierung bedeutsame Zusammenhänge zwischen unterschiedlichen Merkmalen der Komposition und Leseleistungen feststellen. Was die Entwicklung der Leseleistung angeht, so kommen zum Beispiel Köller, Schütte, Zimmermann, Retelsdorf und Leucht (2013) für die Sekundarstufe zu dem Ergebnis, dass diese durch die Leistungen der Mitschülerinnen und -schüler wenig beeinflusst wird. Bellin (2009) untersuchte die Wirkung der Klassenkomposition auf Leistungen unter anderem im Lesen in Berliner Grundschulen. Es zeigte sich, dass die sozialen und migrationsbedingten Disparitäten in den Schuleinzugsbereichen, die sich in den Zusammensetzungen der Klassen widerspiegeln, die Leistungsentwicklung der Kinder zusätzlich beeinflussen. Auch hier handelt es sich allerdings um kleine Effekte. Längsschnittuntersuchungen, die zu anderen Leistungsdomänen für den Grundschulbereich vorliegen und zusätzlich auch für die Betrachtung von Kompositionseffekten Prozessmerkmale (z. B. Unterrichtsmerkmale) berücksichtigen (van Ophuysen & Wendt, 2009; Vennemann, im Erscheinen), zeigen, dass die Leistungsunterschiede zwischen Schülerinnen und Schülern zu deutlich größeren Anteilen durch Prozessmerkmale erklärt werden können. Die Befunde verdeutlichen damit, dass es auf einen angemessenen Umgang mit Heterogenität ankommt (Morris-Lange, Wendt & Wohlfahrt, 2013).

3 Zur Sicht der Schulleitungen auf die Integration neu zugewanderter Schülerinnen und Schüler und Reaktionen in Form von äußerer und innerer Differenzierung

Die Schulleitungen wurden gefragt, inwieweit sie den folgenden vier Aussagen zur Integration von neu zugewanderten Schülerinnen und Schülern in ihrer Schule zustimmen:

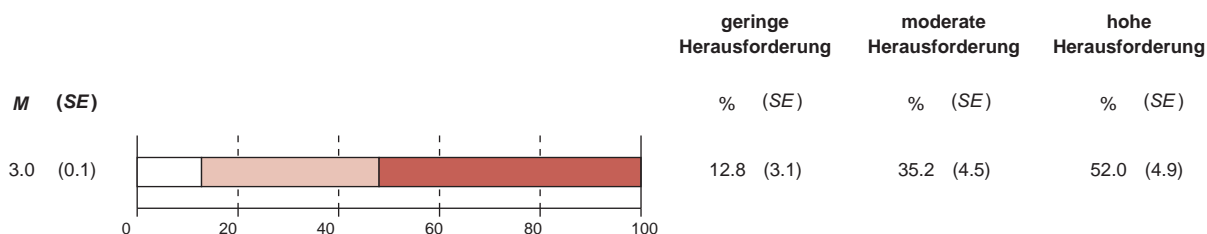
Die Integration von neu zugewanderten Schülerinnen und Schülern stellt ...

1. unsere Schule vor besondere organisatorische Herausforderungen.
2. unsere Schule vor besondere räumliche Herausforderungen.
3. einen Großteil des Lehrerkollegiums vor besondere pädagogische Herausforderungen.
4. eine große Mehrbelastung dar.

Für die Darstellung der Ergebnisse wurden die Angaben der Schulleitungen für alle Aussagen mit 1 („*stimme überhaupt nicht zu*“) bis 4 („*stimme stark zu*“) Punkten kodiert. Daraufhin wurde der Mittelwert für diese vier Aussagen für jede Schulleitung berechnet. Der daraus resultierende Skalenwert gibt an, inwieweit die jeweilige Leitung Aussagen zur Integration von neu zugewanderten Schülerinnen und Schülern positiv oder negativ bewertet. Die interne Konsistenz der Skala liegt bei einem Cronbachs Alpha von $\alpha = .819$. Anschließend wurden verschiedene Gruppen von Schulleitungen gebildet. Leitungen mit einem Mittelwert (M) von $1 \leq M < 2$ werden unter ‚geringe Herausforderung‘, mit einem Mittelwert von $2 \leq M < 3$ unter ‚moderate‘ und Leitungen mit einem Mittelwert von $3 \leq M \leq 4$ unter ‚hohe Herausforderung‘ eingestuft. Berücksichtigt wurden nur Angaben der Leitungen von Schulen, in denen diese Kinder auch anzutreffen sind, das heißt von 63 Prozent.

Aus Abbildung 11.2 ist ersichtlich, dass mehr als die Hälfte der neu zugewanderten Schülerinnen und Schüler in Schulen unterrichtet wird, deren Leitungen die mit ihrer Integration verbundenen Ansprüche als hoch ansehen. Demgegenüber ist nur ein Achtel dieser Kinder in Schulen, deren Leitungen diese Herausforderung als eher gering betrachten.

Abbildung 11.2: Integration von neu zugewanderten Schülerinnen und Schülern aus Sicht der Schulleitungen an Schulen mit entsprechender Schülerschaft (Skalenmittelwert und Anteil der Schulen in %)



Die Skala umfasst vier Fragen (z. B. *Die Integration von neu zugewanderten Schülern stellt einen Großteil des Lehrerkollegiums vor besondere pädagogische Herausforderungen.*) mit vierstufigem Antwortformat (4 = Stimme stark zu, ..., 1 = Stimme überhaupt nicht zu).

IEA: Progress in International Reading Literacy Study (PIRLS)

© IGLU 2016

Die Integration von neu zugewanderten Schülerinnen und Schülern angemessen zu gestalten ist nur eine der Aufgaben, die im Rahmen der Bemühungen zu einem kompetenten Umgang mit Heterogenität anstehen. Auch andere Facetten von Heterogenität sind zu beachten. Auf die damit verbundenen Herausforderungen kann in verschiedenen Formen reagiert werden. Grob unterscheiden lassen sich Varianten äußerer und innerer Differenzierung.

Äußere Differenzierung liegt zum Beispiel vor, wenn in einer Schule alle Kinder mit diagnostiziertem Förderbedarf in einer einzigen „Integrationsklasse“ unterrichtet werden. In Deutschland liegt der Anteil solcher Klassen bei 25.3 Prozent. Auch für die Unterrichtung neu zugewanderter Kinder gibt es in Deutschland Modelle äußerer Differenzierung (vgl. Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, 2015, S. 45). So können sie etwa in allen Fächern in einer speziell für sie eingerichteten Klasse unterrichtet werden. In sogenannten „Willkommensklassen“ sollen insbesondere ihre sprachlichen Fähigkeiten so gefördert werden, dass sie nach relativ kurzer Zeit am Unterricht in einer Regelklasse teilnehmen können. In Deutschland liegt der Anteil solcher Klassen bei 7.7 Prozent. Darüber hinaus ist der jahrgangsübergreifende Unterricht eine Möglichkeit, Kinder in alters- und leistungsgemischten Gruppen individuell zu fördern. Diese Form äußerer Differenzierung besteht zum Beispiel, wenn an Grundschulen statt des ersten und zweiten Schuljahres eine ein- bis dreijährige flexible Schuleingangsphase angeboten wird. Der Anteil an jahrgangsübergreifendem Unterricht liegt bei Grundschulen in Deutschland bei 12.2 Prozent.

Um der Leistungsheterogenität Rechnung zu tragen, werden an vielen Grundschulen in Deutschland weitere Lerngelegenheiten jenseits der Klassengrenzen angeboten. Dabei handelt es sich um Förderangebote in Deutsch für leistungsschwache Schülerinnen und Schüler (81.7%), Förderkurse für lese- und rechtschreibschwache Schülerinnen und Schüler (77.4%), Lesekurse beziehungsweise Lesestunden außerhalb des Unterrichts (77.4%), Zusatzunterricht für Schülerinnen und Schüler nichtdeutscher Muttersprache (66.6%), Zusatzangebote

für leistungsstarke Schülerinnen und Schüler (61.8%), Leseförderung in kleinen Gruppen durch „Lesemütter“ (ehrenamtliche Eltern oder Großeltern, 59.1%), Zusatzunterricht in Deutsch (53.5%) oder auch jahrgangübergreifende Leseprojekte (45.9%). Bilingualer Unterricht wird an 3.6 Prozent aller Schulen angeboten. Den Angaben der Schulleitungen zufolge befinden sich nur 5.3 Prozent der Grundschulen in Deutschland im voll gebundenen Ganztage, der besonders günstige Voraussetzungen für die Schaffung weiterer Lerngelegenheiten bietet.

Innere Differenzierung beschreibt Maßnahmen zur Leistungsdifferenzierung. Hier haben Lehrkräfte in erster Linie die Option, für Subgruppen von Schülerinnen und Schülern die Quantität und/oder Qualität der Aufgaben zu variieren und mit gestuften Hilfen zu arbeiten. Nach Angaben der Lehrkräfte erhält etwas weniger als die Hälfte der Schülerinnen und Schüler in Deutschland (45.7%) je nach Leistung unterschiedlich schwere Aufgaben und damit verbundene Zeitfenster, um diese Aufgaben zu bearbeiten. 86.2 Prozent beziehungsweise 83.2 Prozent der Schülerinnen und Schüler werden von Lehrkräften unterrichtet, die leistungsschwachen Kindern zusätzliche Unterstützung bieten beziehungsweise leistungsstarken Kindern Extraaufgaben geben. Kooperative Lernformen erleben 75.6 Prozent der Schülerinnen und Schüler in ihrem Lesebeziehungsweise Deutschunterricht, Angebote zur Freiarbeit hingegen 45.7 Prozent.

4 Vorschläge zum adaptiven Lehren beziehungsweise „kompetenten Umgang mit Heterogenität“ im Leseunterricht der Grundschule

In diesem Abschnitt wird exemplarisch beschrieben, wie im Hinblick auf adaptives Lehren in besonders herausfordernden Klassen im Leseunterricht vorgefahren werden kann. Die Ausführungen beziehen sich zunächst auf die Förderung des *phonologischen Rekodierens*. Sie kann auch in den höheren Jahrgangsstufen der Grundschule bei Kindern mit erheblichen Lernschwierigkeiten beziehungsweise mit einem Förderschwerpunkt im Bereich der geistigen Entwicklung angezeigt sein. Sind sie in inklusiven vierten Klassen zu finden, hat man sich als Lehrkraft auf eine Förderung basaler Lesefähigkeiten einzustellen (siehe Abschnitt 4.1).

Erhebliche Heterogenität ist auch im Bereich des *Wortschatzes* auszumachen, dem in der Deutschdidaktik in den letzten Jahren mehr und mehr Aufmerksamkeit gewidmet wird. Ein zentrales – und nur schwer erreichbares – Ziel der Arbeit in diesem Feld sollte sein, dass Kinder ihre Wortschatzlücken mehr und mehr selbstständig erkennen und Strategien entwickeln, sie mit Hilfe anderer, vor allem auch kompetenterer Mitschülerinnen und -schüler, zu schließen (siehe Abschnitt 4.2).

Eine Möglichkeit, den divergenten Lern- und Sprachvoraussetzungen der Kinder im Bereich Lesen gerecht zu werden, wird darin gesehen, Subgruppen oder zunächst auch der ganzen Klasse *Texte in Leichter Sprache* zu präsentieren. Wenn im Unterricht Texte gelesen werden, die in Leichter Sprache transformiert wurden oder bereits in diesem Format vorliegen, stellt sich unter anderem die Frage, wie auch Kinder zu ihrem Recht kommen können, die keine Schwierigkeiten mit dem Verständnis von Texten haben, die nicht in Leichter Sprache verfasst sind (siehe Abschnitt 4.3).

Adaptiver Leseunterricht beruht, so der Anspruch, auf einer zuverlässigen *Diagnostik*. Mittlerweile liegen Verfahren vor, mit denen der individuelle *Lernverlauf* der Kinder erfasst und dokumentiert werden kann. Hieraus ergeben sich Ansatzpunkte für gezielte Förderung. Es werden zwei Verfahren vorgestellt. Insbesondere die zweite Variante, das „Rügener Inklusionsmodell“, in dem drei Förderstufen unterschieden werden, ist im Kontext der Debatten um inklusiven Grundschulunterricht von Interesse (siehe Abschnitt 4.4).

4.1 Förderung des phonologischen Rekodierens

1989 legte Günther ein Modell von Stufen des Schriftspracherwerbs vor, das als eine Erweiterung eines Modells von Frith (1985) konzipiert war. Frith hatte zwischen einer logographemischen, einer alphabetischen und einer orthographischen Stufe unterschieden, Günther fügte unter anderem eine basale Stufe hinzu, die er als „präliteral-symbolisch“ bezeichnete. Auf Günthers Modell wird aktuell immer wieder Bezug genommen, wenn es um Schülerinnen und Schüler im Förderschwerpunkt geistige Entwicklung geht. Hier wird im Bemühen, deren Fähigkeiten differenziert zu erfassen, auf einen „erweiterten Lesebegriff“ beziehungsweise auf die Unterscheidung von „Lesen im weiteren Sinn“ und „Lesen im engeren Sinn“ gesetzt. Bei Ersterem geht es um das „Lesen“ von ikonischen Abbildungen, Symbolen (z. B. Piktogrammen) und Wörtern, die anhand hervorstechender visueller Merkmale erkannt werden. Von Lesen im engeren Sinn wird gesprochen, wenn auf der alphabetischen Stufe Phonem-Graphem-Korrespondenzen erfasst werden können, Voraussetzung für das Leseverständnis auf Wort-, Satz- und Textebene. Ratz (2013) befragte Lehrkräfte von mehr als 1600 Schülerinnen und Schülern mit dem Förderschwerpunkt Geistige Entwicklung, darunter 550 Grundschülerinnen und -schüler. Nach ihrer Auskunft konnten 14 Prozent noch gar nicht „lesen“, 18 Prozent verstanden Bilder und ikonische Zeichen, 11 Prozent Symbole. 39 beziehungsweise 18 Prozent wurde die Fähigkeit attestiert, alphabetisch beziehungsweise orthographisch zu lesen (vgl. ebd., S. 355). Nach Jahrgangsstufen wurde nicht differenziert. Auf der alphabetischen Stufe besteht die Herausforderung darin, distinkte Grapheme mit Phonemen zu verknüpfen, die ihrerseits keine „natürlichen“ Einheiten des Gesprochenen sind. Die Aussprache eines Konsonanten in einer Silbe ist von der Koartikulation mit dem vokalischen Kern abhängig und Plosive wie /d/, /t/, /b/ können noch nicht einmal „unsilbisch“ ausgesprochen werden. Sie erscheinen immer in Kombination mit dem Schwa-Laut /ə/. „Es geht beim alphabetischen Lesen also nicht darum, isolierte Phoneme miteinander zu verschleifen (zu synthetisieren), sondern um die Fähigkeit, Gruppen von Graphemen (zu Beginn meist Verbindungen aus Konsonant und Vokal) direkt gemeinsam zu erlesen.“ (Kuhl, Euker & Ennemoser, 2015, S. 42 f.) Kuhl et al. (ebd.) testeten Schülerinnen und Schüler im Alter zwischen 10 und 17 Jahren mithilfe des Kieler Leseaufbaus (Dummer-Smoch & Hackethal, 2002). Es ergaben sich zwar positive Effekte im Hinblick auf die Fähigkeit, geschriebene Wörter phonologisch zu rekodieren. Die Dekodierleistung, das heißt die Bedeutungserfassung auf Wortebene, und das Satzverständnis verbesserten sich aber nicht (vgl. ebd., S. 50).

Nehmen also Kinder mit intellektueller Beeinträchtigung und andere Kinder mit erheblichen Lernproblemen am inklusiven Leseunterricht teil, kann man als Lehrkraft damit rechnen, dass einige auch noch in der zweiten Klasse und

später nur selten den direkten Weg von Graphemfolgen über das orthographische Lexikon (den Sichtwortschatz) zum Eintrag im semantischen Lexikon gehen können. Sie sind noch darauf angewiesen, Einheiten unterhalb der Wortebene, das heißt einzelne Grapheme, Cluster, Schreibsilben, zunächst phonologisch zu rekodieren, bevor – im günstigen Fall – die Bedeutung des Zielworts erfasst werden kann. Das kann auch für Kinder gelten, denen LRS attestiert wurde. Übungen zur auditiven beziehungsweise visuellen Wahrnehmung und Verarbeitung im Umgang mit nichtsprachlichem Material sind hier nicht angezeigt. Förderlich ist vielmehr, „symptomnah“ anzusetzen und anhand von strukturell einfachen ein- und zweisilbigen Wörtern an der Lautanalyse und -synthese zu arbeiten. Gelingt das, kann zur Gliederung von schriftlich vorliegenden Wörtern übergegangen werden. Die Kinder sollen lernen, Silbengrenzen in gedruckten Wörtern zu markieren und dabei der Regel zu folgen, dass dann, wenn zwischen zwei Vokalbuchstaben mehrere Buchstaben für Konsonanten stehen, der letzte zur jeweils folgenden Silbe gehört. Die visuelle Gliederung von Wörtern in Silben ist zum Beispiel Kern des Potsdamer Lesetrainings *PotsBlitz* (Ritter & Scheerer-Neumann, 2009). Sukzessive werden die zu erlesenden Wörter silbenstrukturell komplexer und länger und in der Folge lesen die Kinder am PC „Blitzwörter“, wobei die Dauer ihrer Präsenz variiert werden kann. Dann werden kurze Texte wiederholt gelesen, bis ein Erfolgskriterium erreicht ist (85 Wörter pro Minute). So soll der Schritt vom indirekten, das heißt phonologisch rekodierenden Erlesen zur automatisierten Worterkennung und darüber hinaus zum Lesen von Texten mit einer kleinen Zahl von Wörtern gelingen. Gerade beim Übergang vom rekodierenden zum automatisierten Wortlesen haben leseschwache Kinder massive Schwierigkeiten. *PotsBlitz* wurde, anders als viele andere Programme, im Hinblick auf seine Wirksamkeit evaluiert. Nachweislich führt es zur Erhöhung der Lesegeschwindigkeit, ohne dass die Lesegenauigkeit leidet.

4.2 Arbeit am Wortschatz

Ein Kind mag zwar mit *Wortformen* kaum noch Schwierigkeiten haben, Wörter und Sätze also weitgehend korrekt phonologisch repräsentieren und dabei silbische und auch morphembezogene Informationen nutzen können. Dennoch kann es auf der *inhaltlichen Ebene* scheitern, weil der Wortschatz nicht breit beziehungsweise umfangreich genug ist oder weil es ihm an Qualität beziehungsweise Tiefe mangelt. Im ersten Fall ist zum Beispiel gar nicht bekannt, dass „Pfersich“ Teil eines semantischen Netzes mit „Obst“ beziehungsweise „Früchte“ als Oberbegriff ist, im zweiten ist das zwar präsent, es wird aber zum Beispiel nicht gewusst, anhand welchen semantischen Merkmals Pfirsiche von Nektarinen zu unterscheiden sind (Motsch, Marks & Ulrich, 2015, S. 29). Im Deutschunterricht in der Grundschule hat die Arbeit am Wortschatz eine lange Tradition. Hier werden vor allem Wortfamilien und Wortfelder thematisiert. Man sammelt Wörter mit demselben Stamm-Morphem (*lesen, ablesen, [sich] verlesen, unlesbar, Lesung* usw.) beziehungsweise lotet ein Wortfeld wie *gehen* aus (*schreiten, schlendern, stolzieren, hasten* usw.). Insbesondere im Umgang mit Kindern, denen man besondere Schwierigkeiten im semantisch-lexikalischen Bereich beziehungsweise lexikalische Störungen attestiert hat, aber auch mit neu zugewanderten Kindern wird am basalen Wortschatz in lebensweltlich relevanten thematischen Feldern wie zum Beispiel „menschlicher Körper und Gesundheit“, „Kleidung“, „Lebensmittel“, „Haustiere“ gearbeitet. Ziel ist nicht nur, dass die

Kinder die Bedeutungen der wenigen neuen Wörter kennen lernen und mit diesen Wörtern zukünftig aktiv und passiv kompetent umgehen. Vor allem ist intendiert, auf diese Weise selbstreguliertes Wortschatzlernen anzuregen. „Mit diesem intensiven exemplarischen Arbeiten an wenigen Wortfeldern war die Hoffnung verknüpft, dass das Kind auch in allen anderen Erfahrungsräumen des Alltags vermehrt neue Wörter einzuspeichern beginnt.“ (Motsch et al., 2015, S. 71) Nicht selten aber trotz diese Hoffnung. Seifert, Schwab und Gasteiger-Klicpera (2014) etwa boten Kindern in zweiten Klassen, darunter 71 Prozent mit einer anderen Herkunftssprache als Deutsch, einen Lernwortschatz im Umfang von 135 Wörtern an, aufgeteilt auf neun Themenkomplexe, zum Beispiel „Advent und Weihnachten“ und „Winter und Fasching“ (respektive „Karneval“). Zum Thema „Winter und Fasching“ zum Beispiel lernten sie 15 Wörter, in der Mehrzahl Substantive, jeweils mit bestimmtem Artikel und Pluralform, aber auch einige Verben (mit Formen in der 3. Person Singular und in der 1. Person Plural) und Adjektive (einschließlich Komparativformen). Das Vokabular wurde eingeführt und geübt, bevor die Kinder mit thematisch einschlägigen Texten konfrontiert wurden. Es zeigte sich, dass die Kinder in der Versuchsgruppe zwar im Hinblick auf das Leseverständnis bedeutsam mehr profitierten als die in einer Kontrollgruppe. Effekte im Bereich des Wortschatzes, gemessen mit einer Kurzform des *Wortschatz- und Wortfindungstests für 6- bis 10-Jährige* (WWT 6–10; Glück, 2007), blieben aber aus. Die Autorinnen überprüften nicht separat, ob die gelernten Wörter aktiv und passiv beherrscht wurden, vermuten das aber: „It is likely that the children learned the vocabulary offered, but then again they did not transfer their knowledge to untrained vocabulary.“ (Schwab, Seifert & Gasteiger-Klicpera, 2014, S. 218) Sie sind der Auffassung, dass – anders als von ihnen praktiziert – neue Wörter oft und in verschiedenen Kontexten zu präsentieren seien und dass es auch hilfreich sei, Kinder nach bedeutungsähnlichen Wörtern in ihrer jeweiligen Herkunftssprache zu fragen. Wie, bezogen auf das als Konkretum gebrauchte Wort *Strand*, eine Wortschatzarbeit beschaffen sein könnte, bei der alle für den Eintrag in das mentale Lexikon relevanten Aspekte berücksichtigt sind, verdeutlicht Alber (2016, S. 125, vgl. auch Juska-Bacher & Jakob, 2014). Hier sind auf einer Wortschatzkarte Reimwörter (*Hand, Wand, Sand*) angegeben (phonologischer Aspekt), Mitglieder der Wortfamilie (*Strandtuch, Strandbad* usw.) (morphologischer Aspekt), Teile eines semantischen Netzwerks (*Strand – Boot – baden – Meer – Sand – Steg*), bedeutungsähnliche Wörter in der Zielsprache (*Ufer, Küste*) und in anderen (Herkunfts-)Sprachen, Beispielsätze zur Erläuterung des syntaktischen Aspekts und eine bildliche Darstellung. Es liegt auf der Hand, dass größere Teile des Wortschatzes auf diese sehr detaillierte Weise nicht vermittelt werden können.

Insbesondere für Kinder, bei denen emotional-soziale Auffälligkeiten diagnostiziert beziehungsweise vermutet werden, kommen Wörter in Betracht, mit denen Emotionen bezeichnet werden. Sind sie verfügbar, kann das dabei helfen, innere Zustände und Prozesse angemessener wahrzunehmen. Verknüpfungen von Emotionen und prototypischen Verhaltensweisen, die sich bildhaft darstellen lassen, sind problematisch. So kann ein mimischer Ausdruck mit verschiedenen Emotionen korrespondieren, diverse Emotionen können auf identische Weise ausgedrückt oder auch maskiert werden. Insofern liegt es nahe, Elemente des Gefühlswortschatzes (*Furcht, Ärger, Freude, Traurigkeit, Ekel* usw.) in Form von Rollenspielen, vor allem aber im Umgang mit narrativen Texten zu thematisieren (z. B. Petermann, Koglin, Natzke & von Marées, 2013).

Angesichts dessen, dass Kindern mit Schwierigkeiten im semantisch-lexikalischen Bereich im Vergleich mit Gleichaltrigen mehrere tausend Wörter „fehlen“ können, erscheint es als besonders aussichtsreich, ihnen Fragestrategien zu vermitteln. Sie können als „wichtigster Motor für die lexikalische Entwicklung“ (Motsch et al., 2015, S. 40) angesehen werden. Dabei geht es zunächst darum, dass die Kinder während des Leseprozesses lexikalische Lücken überhaupt bemerken. Dann sollte man immer wieder erfahren können, dass man nicht bloßgestellt wird, wenn man Mitschülerinnen und Mitschüler oder die Lehrkraft nach Wortbedeutungen fragt. Wenn es sich anbietet, kann gefragt werden, wie neue Wörter in bereits bestehende, einmal mehr, einmal weniger „dichte“ semantische Netze zu integrieren sind, indem sie zum Beispiel Oberbegriffen, Wörtern auf derselben hierarchischen Ebene oder Wörtern mit gegensätzlicher Bedeutung zugeordnet werden (zu weiteren Strategien Motsch et al., 2015, S. 120 ff.).

4.3 Texte in Leichter Sprache

Insbesondere für Kinder mit Lern- und Sprachproblemen kann es förderlich sein, dass ihnen Texte mit veränderter typographischer Gestalt, das heißt mit vergrößerter Schrift, größeren Buchstaben-, Wort- und Zeilenabständen und markanter Absatzgliederung angeboten werden. Als den Leseprozess erleichternd wird zum Beispiel auch angesehen, dass Satz- und Zeilenende übereinstimmen.

Vorschläge, Texte auf diese Weise graphisch zu gestalten, sind Teil des Konzepts „Leichte Sprache“, das in den letzten Jahren nicht zuletzt im Kontext von Inklusion (einschließlich Zuwanderung) propagiert und kontrovers erörtert wurde. Den „Regeln für Leichte Sprache“ zufolge (Netzwerk Leichte Sprache, 2013) sollten zum Beispiel einfache, häufig vorkommende, kurze Wörter verwendet werden, wobei bei der wiederholten Bezeichnung desselben Gegenstands oder derselben Handlung auf lexikalische Konstanz zu achten sei. Seien morphologisch komplexere Wörter nicht zu vermeiden, sollten Trennungsmarker genutzt werden, insbesondere Mediopunkte (*Leichte · Sprache · Regel · werk*). Verben sollten nur im Aktiv und Indikativ gebraucht werden, Redewendungen und Metaphern seien zu vermeiden. (Haupt-)Sätze sollten kurz sein und der Struktur Subjekt-Verb-Objekt (bzw. Adverbial) entsprechen, kausale und konditionale Nebensätze sind aber „zugelassen“. Verortet man das Lesen von Texten in Leichter Sprache im *Gemeinsamen Europäischen Referenzrahmen für Sprachen* (GER, 2001), in dem 6 Niveaus (A1, A2, B1, B2, C1 und C2) unterschieden werden, dann handelt es sich etwa um eine Anforderung auf dem Niveau A2. Auf dieser Stufe können „kurze, einfache Texte zu vertrauten, konkreten Themen“ gelesen und verstanden werden, „die einen sehr frequenten Wortschatz [...] enthalten“ (ebd., S. 75). Die „Regeln“ für die Produktion „leichter“ Texte lassen sich verknüpfen mit Resultaten älterer Forschung zur Textverständlichkeit. Langer, Schulz von Thun und Tausch (2011) zum Beispiel identifizierten faktorenanalytisch vier Dimensionen der Verständlichkeit von Texten, Einfachheit (einfache, kurze Sätze, geläufige, nicht mehr als dreisilbige Wörter), Kürze/Prägnanz (wobei sowohl „Gedrängtheit“ als auch „Weitschweifigkeit“ zu vermeiden seien), Gliederung/Ordnung (leichte Erkennbarkeit eines „roten Fadens“), schließlich anregende Zusätze wie Beispiele und Illustrationen in mittlerer Dosierung. Merkmale Leichter Sprache korrespondieren insbesondere mit den ersten drei Dimensionen. Die Regeln für die Gestaltung von Texten in Leichter Sprache sind

aus der Praxis heraus erwachsen. Die Forschung dazu entwickelt sich erst in den letzten Jahren. Zwar kann auf Befunde aus einer Reihe von Disziplinen zurückgegriffen werden. Sie müssen aber gebündelt werden (Bredel & Maaß, 2016). Leichte Sprache lässt sich beschreiben als ein Register, als eine funktionale Varietät des Deutschen. Die Aufgabe, Texte in dieser Varietät zu formulieren beziehungsweise Ausgangstexte in solche in Leichter Sprache zu transformieren, ist durchaus anspruchsvoll, zumal dann, wenn wesentliche „Verluste“ auf inhaltlicher Ebene vermieden werden sollen. Es stellt sich unter anderem immer wieder „die Frage, welche Mittel eingesetzt werden können, um Ersatz für standardsprachliche Strukturen zu finden, die nicht zum System der Leichten Sprache gehören“ (ebd., S. 15). Den Leserinnen und Lesern soll der „Einstieg“ in die Schriftsprachlichkeit erleichtert werden. Unter dem Signum „barrierefreier Zugang zur Schriftkultur“ soll aber keine neue Barriere errichtet werden. Leichte Sprache ist keine neue Norm, sondern als eine Übergangsvarietät zu verstehen, als Hilfe auf dem Weg zum Verstehen bildungssprachlicher Texte (zu Versionen des Konstrukts „Bildungssprache“ siehe auch Stahns, 2016).

Mittlerweile werden zahlreiche Texte in Leichter oder auch in „einfacher“ Sprache angeboten, darunter auch literarische, zum Beispiel Märchen, Fabeln und Balladen, und Sachtexte für Kinder. Genannt seien nur die Reihe *einfach lesen!* bei Cornelsen für Leseinsteiger in den Klassen 2 und 3 und für Lesefortgeschrittene auf Niveau 1 ab Klasse 3 und das Online-Portal zur Leseförderung *Antolin* mit der neuen Rubrik „Einfache Sprache“. Werden solche Texte eingesetzt, spricht viel dafür, adaptiven Unterricht so zu gestalten, dass sie von allen Kindern in der Klasse gelesen werden und den lesekundigeren Kindern darüber hinaus weitere Texte angeboten werden.

4.4 Prozessdiagnostik und der „Response-to-Intervention“-Ansatz

Diagnostische Verfahren zur Ermittlung des Förderbedarfs wurden und werden immer wieder kritisiert, vor allem in den Bereichen „Emotionale und soziale Entwicklung“, aber auch in den Feldern „Lernen“ und „Sprache“ und LRS. Es komme zu „zweifelhaften Gruppenbildungen von Schülerinnen und Schülern aufgrund weniger und prognostisch nicht valider Ist-Standmessungen von Schülermerkmalen“ (Voß et al., 2016, S. 25). Eine Alternative besteht darin, nicht mehr (nur) auf „klassifizierende“ Statusdiagnostik, sondern (darüber hinaus) auf eine entwicklungsorientierte, den schulischen Lernprozess begleitende Diagnostik zu setzen. Lehrkräfte diagnostizieren seit jeher prozessbegleitend, und zwar sowohl eher informell, insofern sie Schülerinnen und Schüler beobachten, Schlüsse daraus ziehen, wie Kinder Leseaufgaben beantworten oder Hausaufgaben machen, als auch formell, indem sie Tests und Klassenarbeiten auswerten.

Mittlerweile stehen Testverfahren beziehungsweise Testreihen zur Verfügung, mit denen der Lernverlauf einzelner Kinder und der ganzen Klasse im Bereich Lesen objektiv, zuverlässig und valide erfasst werden kann. Ein Beispiel ist der internetbasierte *quop*-Ansatz (Souvignier, Förster & Salaschek, 2014; vgl. auch Walter, 2013). Die Kinder bearbeiten im Verlauf eines Schuljahrs acht Tests einer Reihe im relativ kurzen zeitlichen Abstand von drei Wochen. Jeder Test ist kurz, die Bearbeitungszeit beträgt etwa 10 Minuten, und alle Tests sind, obwohl die Leseaufgaben und Texte variieren, gleich schwierig, wie umfangreiche Analysen im Rahmen von mehreren Studien mit mehr als 3000 Kindern ergeben haben.

Zugleich sind sie so sensitiv, dass auch über die kurzen Zeitintervalle hinweg Leistungsveränderungen reliabel erfasst werden können. So gestatten sie insbesondere Schlüsse im Hinblick auf individuellen Förderbedarf. Die Lernverläufe der einzelnen Kinder werden anschaulich visualisiert, so dass auf einen Blick zu erkennen ist, ob die Entwicklung vom ersten bis zum achten Test „aufwärts“ geht, stagniert oder „nach unten“ weist. Bislang liegen Testreihen für alle vier Jahrgangsstufen der Grundschule vor. In Klasse 1 geht es um phonologische Bewusstheit und Wort- und Satzverständnis, in Klasse 2 um Wort-, Satz- und Textverständnis und in den Klassen 3 und 4 jeweils um Lesegeschwindigkeit, -genauigkeit und um das Leseverständnis, wobei hier differenziert wird nach textbasiertem und vorwissensbasiertem Verständnis. Die Auswertung der Tests und die Dokumentation der Leistungsentwicklungen geschehen automatisiert. Die Lehrkräfte, die *quop* einsetzen, hielten das Verfahren für sehr praktikabel; sie gaben an, dass sie die Testergebnisse für die weitere Planung des Leseunterrichts im Allgemeinen und im Hinblick auf die Identifizierung von Kindern mit Förderbedarf im Besonderen nutzten. Im Rahmen eines Versuchs-Kontrollgruppen-Designs wurde in mehr als 150 Klassen auch die Wirksamkeit des Ansatzes überprüft. Kontrollklassen waren solche, in denen zu Beginn des Schuljahrs standardisierte Lesetests eingesetzt worden waren, in denen den Lehrkräften also diagnostische Informationen zum Leistungsstand aller Kinder zur Verfügung standen. In mehreren Studien ergab der am Ende des Schuljahrs unternommene Vergleich zwischen diesen Klassen und den Versuchsklassen, in denen systematisch Informationen zum Lernverlauf gewonnen und dokumentiert worden waren, signifikant und auch praktisch bedeutsame höhere Lernzuwächse, gemessen mit ELFE 1–6, in den Versuchsklassen (z. B. Souvignier & Förster, 2011; Souvignier, Förster & Schulte, 2014; Hinweise zur Teilnahme und zu den Kosten unter <https://www.quop.de/de/mitmachen/teilnahme/>).

quop ist ein Beispiel eines curriculumbasierten Instruments (*curriculum-based measurement*), mit dem sensitiv und kurzfristig Leistungsveränderungen erfasst werden können. Eine für die Gestaltung inklusiven (Lese-)Unterrichts bedeutsame Variante ist unter der Bezeichnung *Response to Intervention* in den Jahren um 1980 in den USA entwickelt worden (National Center on Response to Intervention, 2010; Burns, Appleton & Stehouwer, 2005). In Deutschland wurde sie im Rahmen des groß angelegten „Rügener Inklusionsmodells“ praktiziert. Hier wurden drei Förderebenen unterschieden. Auf einer ersten Ebene geht es um die ganze Klasse. Zeigt sich, dass bei Testungen einzelne Kinder zum Beispiel im Leistungsbereich Lesen deutlich unterdurchschnittlich abschneiden, werden sie – Förderebene II – von der Grundschullehrkraft in Kleingruppen über mehrere Wochen intensiv zusätzlich gefördert, wobei in ersten Klassen 3 Förderstunden und in dritten und vierten Klassen jeweils 1 Förderstunde zur Verfügung stehen. Gelingt am Ende der Anschluss an das Leistungsvermögen der anderen Kinder nicht, wird auf Förderebene III individuell oder in sehr kleinen Gruppen über mehrere Wochen weiter gefördert, nun aber nicht mehr von der Grundschullehrkraft, sondern von einer Sonderschullehrerin beziehungsweise einem Sonderschullehrer. Können die Leistungen hier nicht so gesteigert werden, dass ein Wechsel auf Ebene II oder sogar I möglich ist, soll das Kind, gegebenenfalls nach Feststellung eines sonderpädagogischen Förderbedarfs, im Klassenkontext zieldifferent unterrichtet, aber auch weiterhin auf den Ebenen II und III gefördert werden. Zwischen 2010 und 2014 wurden circa 400 Kinder auf Rügen von der ersten bis zur vierten Klasse inklusiv unterrichtet und im Lesen, Rechtschreiben und in Mathematik und darüber hinaus in den

Förderbereichen Lernen, Sprache und emotional-soziale Entwicklung zu verschiedenen Zeitpunkten getestet. In der Domäne Lesen kamen unter anderem der *Kieler Leseaufbau* (Dummer-Smoch & Hackethal, 2002), die *Verlaufsdagnostik sinnerfassendes Lesen* (Walter, 2013) und ELFE 1–6 (Lenhard & Schneider, 2006) zum Einsatz, für den Klassenunterricht auf Ebene I waren mehrere Lehrgänge (z.B. „Piri“) und das Förderprogramm *Wir werden Lesedetektive* (Rühl & Souvignier, 2006) vorgesehen. Die Rügener Versuchsgruppe wurde unter anderem mit etwa gleich vielen Kindern verglichen, die in Stralsund Grundschul-, Diagnoseförder- und Sprachheilklassen besuchten. Voß et al. (2016) präsentieren eine Fülle von Teilergebnissen. Summarisch betrachtet fallen die Leistungen der Viertklässlerinnen und Viertklässler auf Rügen und in Stralsund im Lesen gleich aus. Differenzen bei Schülerinnen und Schülern ohne Lernrisiken beziehungsweise mit günstigen Voraussetzungen sind nicht auszumachen. Rügener „Zwillinge“ der Stralsunder Kinder in Diagnoseförderklassen (Schwerpunkt Lernen) erreichten bereits nach drei Schuljahren deren Leistungsstand am Ende des vierten Schuljahrs. In ihrer Gesamtheit betrachtet entwickelten sich die Rügener Kinder im Bereich der emotional-sozialen Entwicklung positiver, beim Vergleich der „Zwillinge“ mit Risiken in diesem Feld ergeben sich keine Unterschiede. Das gilt auch für die Kinder mit spezifischen Sprachentwicklungsproblemen. Es zeigten sich allerdings Defizite der Förderung auf Ebene III: „So ist die Anzahl der Kinder, die [hier] eine Förderung [...] in den Bereichen Sprache oder emotional-soziale Entwicklung brauchen und nicht erhalten, zu hoch“ (ebd., S. 277). Bei den Grundschullehrpersonen war die Akzeptanz des Rügener Inklusionsmodells deutlich geringer als bei den Sonderschullehrerinnen und -lehrern. Zwar standen die Grundschullehrkräfte Inklusion im Allgemeinen nicht negativ gegenüber, aber nur etwa die Hälfte war davon überzeugt, dass für förderbedürftige Kinder der gemeinsame Unterricht das angemessene Konzept ist. „Zwei Drittel [...] möchten nicht dazu verpflichtet werden, förderungsbedürftige Kinder zu unterrichten [...]“ (ebd., S. 265). Voß et al. vermuten, dass die „Systemumstellung“ auf Rügen wahrscheinlich zu abrupt war, dass es sinnvoll gewesen wäre, die Lehrkräfte stärker zu beteiligen, und dass zu wenig auf Praxisbegleitung, Fortbildung und Auffrischveranstaltungen geachtet wurde. Vor allem aber: Die Umstellung auf inklusiven Unterricht brauche sehr viel mehr Zeit und sei sukzessive in kleineren Schritten anzugehen.

5 Fazit

Kompetenter Umgang mit Heterogenität ist leicht zu postulieren, aber nur schwer ins Werk zu setzen. Wie kann man zum Beispiel als Lehrkraft vor Ort kompetent adaptiv unterrichten, wenn die einen viele Wörter noch mühsam und schrittweise erlesen müssen, während die anderen bereits über einen großen Sichtwortschatz verfügen? Wenn es für die einen angezeigt ist, Texte in Leichter Sprache vorzusehen, während anderen inhaltlich, strukturell und sprachlich recht komplexe Texte kaum noch größere Schwierigkeiten bereiten? Insbesondere die Ergebnisse der Befragung der Grundschullehrkräfte im Rahmen des ambitionierten Projekts „Rügener Inklusionsmodell“ sollten zu denken geben. Der Weg hin zu einer inklusiven Grundschule ist weit und sollte in kleinen, die Akteure nicht überfordernden Schritten gegangen werden. Der Fortbildungsbedarf ist groß und es sollten Maßnahmen zur Reduzierung der Mehrbelastung der Grundschullehrkräfte vorgesehen werden. Dass die Arbeit auf der im Rügener Modell vorgesehenen

Förderebene III, die den Lehrkräften für Sonderpädagogik oblag, zum Teil unbefriedigend war, hat mit deren Arbeitsbedingungen zu tun. Auch um deren Verbesserung sollte man sich bemühen, so dass sie nicht nur sporadisch, sondern kontinuierlich und „auf Augenhöhe“ mit den Grundschullehrkräften kooperieren können.

Literatur

- Alber, K. (2016). Wortschatzumfang, Wortschatztiefe und Verarbeitungsgeschwindigkeit – Analyse der lexikalischen Kompetenz. *Zeitschrift für angewandte Linguistik*, 65 (1), 107–128.
- Ammermüller, A. & Pischke, J.-S. (2009). Peer effects in European primary schools: Evidence from PIRLS. *Journal of Labor Economics*, 27 (3), 315–348.
- Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: Springer.
- Bellin, N., Dunge, O. & Gunzenhauser, C. (2010). The importance of class composition for reading achievement: Migration background, social composition, and instructional practices. An analysis of the German 2006 PIRLS data. *IERI Monograph Series, Issues and Methodologies in Large-Scale Assessments*, 3, 9–34.
- Bredel, U. & Maaß, C. (2016). *Leichte Sprache: Theoretische Grundlagen. Orientierung für die Praxis*. Berlin: Duden.
- Brühwiler, C. (2014). *Adaptive Lehrkompetenz und schulisches Lernen*. Münster: Waxmann.
- Burns, M. K., Appleton, J. J. & Stehouwer, J. D. (2005). Meta-analysis of response-to-intervention research: Examining field-based and research-implemented models. *Journal of Psychoeducational Assessment*, 23, 381–394.
- Deutsche Gesellschaft für Kinder und Jugendpsychiatrie, Psychosomatik und Psychotherapie. (Hrsg.). (2015). *Diagnostik und Behandlung von Kindern und Jugendlichen mit Lese- und/oder Rechtschreibstörung*. Zugriff am 18.11.2017 unter www.kjp.med.uni-muenchen.de/download/leitlinie_lrs_kjp_langfassung.pdf
- Dummer-Smoch, L. & Hackethal, R. (2002). *Kieler Leseaufbau. Handbuch* (6. Auflage). Kiel: Veris.
- Dumont, H., Neumann, N., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen: Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60 (3), 163–183.
- Frith, U. (1985). Beneath the surface of developmental dyslexia. Are comparisons between developmental and acquired disorders meaningful? In K. Patterson, J. C. Marshall & M. Coltheart (Hrsg.), *Surface dyslexia. Neuropsychological and cognitive studies of phonological reading* (S. 301–330). London: Erlbaum.
- GER – Europarat/Rat für kulturelle Zusammenarbeit. (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin: Langenscheidt.
- Glück, C. W. (2007). *Wortschatz- und Wortfindungstest für 6–10-Jährige (WWW 6–10)*. München: Elsevier.
- Gresch, C., Kölm, J. & Kocaj, A. (2017). Amtlich festgestellter sonderpädagogischer Förderbedarf und sonderpädagogische Förderung. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 282–290). Münster: Waxmann.
- Günther, K.-B. (1989). Ontogenese, Entwicklungsprozeß und Störungen beim Schriftspracherwerb unter besonderer Berücksichtigung der Schwierigkeiten von lern- und sprachbehinderten Kindern. In K.-B. Günther (Hrsg.), *Ontogenese, Entwicklungsprozeß und Störungen beim Schriftspracherwerb* (S. 12–33). Heidelberg: Schindele.
- Juska-Bacher, B. & Jakob, S. (2014). Wortschatzumfang und Wortschatzqualität und ihre Bedeutung im fortgesetzten Spracherwerb. *Zeitschrift für angewandte Linguistik*, 61 (1), 49–75.
- KMK – 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 vom*

- 20.10.2011). Zugriff am 23.11.2017 unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2011/2011_10_20-Inklusive-Bildung.pdf
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Empfehlungen zur Arbeit in der Grundschule (Beschluss der Kultusministerkonferenz vom 02.07.1970 i.d.F. vom 11.06.2015)*. Zugriff am 18.11.2017 unter http://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2015/Empfehlung_350_KMK_Arbeit_Grundschule_01.pdf
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017a). *Sonderpädagogische Förderung in Förderschulen (Sonderschulen) 2015/2016*. Zugriff am 23.11.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus_SoPae_2015.pdf
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2017b). *Sonderpädagogische Förderung in allgemeinen Schulen (ohne Förderschulen) 2015/2016*. Zugriff am 23.11.2017 unter https://www.kmk.org/fileadmin/Dateien/pdf/Statistik/Dokumentationen/Aus_SoPae_Int_2015.pdf
- Köller, O., Schütte, K., Zimmermann, F., Retelsdorf, J. & Leucht, M. (2013). Starke Klasse, hohe Leistungen? Die Rolle der Leistungsstärke der Klasse für die individuellen Mathematik- und Leseleistungen in der Sekundarstufe I. *Psychologie in Erziehung und Unterricht*, 60, 184–197.
- Kristen, C. (2008). Schulische Leistungen von Kindern aus türkischen Familien am Ende der Grundschulzeit: Befunde aus der IGLU-Studie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Special Issue 48*, 230–251.
- Kuhl, J., Euker, N. & Ennemoser, M. (2015). Förderung des lautorientierten Lesens bei Schülerinnen und Schülern mit intellektueller Beeinträchtigung. *Empirische Sonderpädagogik*, 7 (1), S. 41–55.
- Langer, I., Schulz von Thun, F. & Tausch, R. (2011). *Sich verständlich ausdrücken* (9. Auflage). München: Ernst Reinhardt.
- Lenhard, W. & Schneider, W. (2006). *ELFE 1–6. Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe.
- Lintorf, K., Guill, K. & Wendt, H. (2017). „Was dein Kind kann, kann meins schon lange!“ Effekte der sozialen Komposition von Grundschulklassen auf den Übergang. *Zeitschrift für Erziehungswissenschaft*. doi: 10.1007/s11618-017-0762-2
- Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache (Hrsg.). (2015). *Neu zugewanderte Kinder und Jugendliche im deutschen Schulsystem*. Köln: Druckhaus Süd.
- Martin, M., Foy, P., Mullis I. & O’Dwyer, L. (2013). Effective schools in reading, mathematics, and science at the fourth grade. In M. Martin & I. Mullis (Hrsg.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade – implications for early learning* (S. 109–180). Chestnut Hill, MA: IEA Publishing.
- Morris-Lange, S., Wendt, H. & Wohlfarth, C. (2013). *Segregation an deutschen Schulen. Ausmaß, Folgen und Handlungsempfehlungen für bessere Bildungschancen* (Sachverständigenrat deutscher Stiftungen für die Integration und Migration). Berlin: SVR GmbH.
- Motsch, H.-J., Marks, D.-K. & Ulrich, T. (2015). *Wortschatzsammler – Evidenzbasierte Strategitherapie lexikalischer Störungen im Kindesalter*. München: Ernst Reinhardt.
- National Center on Response to Intervention. (2010). *Essential Components of RTI – A Closer Look at Response to Intervention*. Zugriff am 20.11.2017 unter http://www.rti4success.org/sites/default/files/rtiessentialcomponents_042710.pdf
- Netzwerk Leichte Sprache. (2013). *Die Regeln für Leichte Sprache*. Zugriff am 23.11.2017 unter www.leichte-sprache.de/dokumente/upload/21dba_regeln_fuer_leichte_sprache.pdf
- Petermann, F., Koglin, U., Natzke, H. & Marées, N. von (2013). *Verhaltenstraining in der Grundschule: ein Programm zur Förderung emotionaler und sozialer Kompetenzen*. Göttingen: Hogrefe.
- Ratz, C. (2013). Zur aktuellen Diskussion und Relevanz des erweiterten Lesebegriffs. *Empirische Sonderpädagogik*, 3 (4), 343–360.
- Ritter, C. & Scheerer-Neumann, G. (2009). *PotsBlitz – Das Potsdamer Lesetraining. Förderung der basalen Lesefähigkeiten*. Köln: ProLog.

- Rühl, K. & Souvignier, E. (2006). *Wir werden Lesedetektive – Lehrermanual und Arbeitsheft*. Göttingen: Vandenhoeck & Ruprecht.
- Scheerer-Neumann, G. (2015). *Lese-Rechtschreib-Schwäche und Legasthenie – Grundlagen, Diagnostik und Förderung*. Stuttgart: Kohlhammer.
- Schwab, S., Seifert, S. & Gasteiger-Klicpera, B. (2014). Improving reading in children with German as a first or second language. *International Journal of Early Years Education*, 22 (2), 210–222.
- Seifert, S., Schwab, S. & Gasteiger-Klicpera, B. (2014). Ein Konzept für differenzierten Leseunterricht in heterogenen Klassen. *Sprache-Stimme-Gehör*, 38 (2), 21–27.
- Smith, D. S., Wendt, H., & Kasper, D. (2016). Social reproduction and sex in German primary schools. *Compare: Journal of Comparative and International Education*, 47 (2), 240–256. doi: 10.1080/03057925.2016.1158643
- Souvignier, E. & Förster, N. (2011). Effekte prozessorientierter Diagnostik auf die Entwicklung der Lesekompetenz leseschwacher Viertklässler. *Empirische Sonderpädagogik*, 3 (3), 243–255.
- Souvignier, E., Förster, N. & Salaschek, M. (2014). Quop: ein Ansatz internet-basierter Lernverlaufsdiagnostik und Testkonzepte für Mathematik und Lesen. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (Tests und Trends N. F., Bd. 12, S. 239–256). Göttingen: Hogrefe.
- Souvignier, E., Förster, N. & Schulte, E. (2014). Wirksamkeit formativen Assessments – Evaluation des Ansatzes der Lernverlaufsdiagnostik. In M. Hasselhorn, W. Schneider & U. Trautwein (Hrsg.), *Lernverlaufsdiagnostik* (Tests und Trends N. F., Bd. 12, S. 221–237). Göttingen: Hogrefe.
- Stahns, R. (2016). Bildungssprachliche Merkmale von Texten und Items. *Didaktik Deutsch*, 41, 44–55.
- van Ophuysen, S. & Wendt, H. (2009). Zur Veränderung der Mathematikleistung von Klasse 4 bis 6. Welchen Einfluss haben Kompositions- und Unterrichtsmerkmale? *Zeitschrift für Erziehungswissenschaft*, 12, 302–327.
- Valtin, R. (2015), *Stellungnahme zu: Diagnostik und Behandlung von Kindern und Jugendlichen mit Lese und/oder Rechtschreibstörung. Evidenz- und konsensbasierte Leitlinie*. Zugriff am 27.11.2017 unter urn:nbn:de:0111-pedocs-124112
- Valtin, R., Voss, A. & Bos, W. (2015). Zur Diagnose von isolierten und kombinierten Leseproblemen. *Didaktik Deutsch*, 20 (38), 40–59.
- Vennemann, M. (2017, im Erscheinen). *Individual- und Kompositionseffekte beim Kompetenzzuwachs in Mathematik und Naturwissenschaft am Ende der Grundschulzeit*. Dissertation, Universität Paderborn.
- Voß, S., Blumenthal, Y., Mahlau, K., Marten, K., Diehl, K., Sikora, S. & Hartke, B. (2016). *Der Response-to-Intervention-Ansatz in der Praxis. Evaluationsergebnisse zum Rügener Inklusionsmodell*. Münster: Waxmann.
- Walter, J. (2013). *Verlaufsdiagnostik sinnerfassenden Lesens*. Göttingen: Hogrefe.

Kapitel XII

Einordnung der IGLU-2016-Befunde in das europäische Rahmenkonzept für gute Leseförderung

Renate Valtin

1 Einleitung

Schriftsprachkompetenzen sind eine grundlegende Voraussetzung für Bildung und persönliche Entwicklung, da sie es dem Einzelnen ermöglichen, ein erfülltes und sinnstiftendes Leben zu führen und aktiv am gesellschaftlichen Leben teilzunehmen. Vor diesem Hintergrund ist es alarmierend, dass in Europa eine/r von fünf Fünfzehnjährigen und nahezu 55 Millionen Erwachsene nicht über grundlegende Lese- und Schreibkompetenzen verfügen (European Commission, 2012). Dies mindert ihre Chancen, einen angemessenen Arbeitsplatz zu finden, und birgt damit ein höheres Armutsrisiko. Mangelnde Lese- und Schreibkompetenzen begrenzen auch die Möglichkeiten zu kultureller Teilhabe und lebenslangem Lernen. Beunruhigend ist die Tatsache, dass sich seit PISA 2000 die Anzahl der Kinder und Jugendlichen mit schwachen Lesekompetenzen im europäischen Durchschnitt nicht verringert hat, wie die internationalen Studien *Internationale Grundschul-Lese-Untersuchung (IGLU)/Progress in International Reading Literacy Study (PIRLS)* und *Programme for International Student Assessment (PISA)* zeigen. So ist in PISA der Anteil der leseschwachen Jugendlichen von 17,8 Prozent im Jahr 2012 auf 19,7 Prozent 2015 gestiegen. Nur in vier EU-Staaten betrug dieser Anteil 2015 weniger als das von der EU sich selbst gestellte Ziel von 15 Prozent: Irland, Finnland, Estland und Polen (European Commission, 2016). Die Europäische Kommission hat auf diese Herausforderung in jüngster Vergangenheit mit zwei Maßnahmen reagiert:

Im Jahr 2011 wurde eine „High Level Group of Experts on Literacy“ eingesetzt, die ein umfassendes Aktionsprogramm zur Verbesserung der schriftsprachlichen Kompetenzen aller europäischen Bürgerinnen und Bürger vorlegte (European Commission, 2012).¹ Die Förderung schriftsprachlicher Kompetenzen

¹ Die Autorin gehörte dem Gremium als Sachverständige für Deutschland an.

wird als eine gesamtgesellschaftliche Aufgabe gesehen, wobei vor allem drei Handlungsfelder wichtig sind: (1) die Schaffung einer schriftreichen Umgebung, (2) die Verbesserung des Unterrichts und (3) die Steigerung von Partizipation, Inklusion und Chancengerechtigkeit.

Ferner finanzierte die EU von 2014 bis 2016 den Aufbau eines Europäischen Netzwerks von Organisationen, die sich der Förderung von Lesen und Schreiben widmen. Im Projekt ELINET (*European Literacy Policy Network*) arbeiteten 78 Partnerorganisationen aus 28 europäischen Ländern zusammen. Koordiniert wurde das Projekt von Prof. Dr. Christine Garbe, Universität Köln. Die Autorin war Leiterin des Teams „Family literacy, preschool and primary school“. Übergeordnetes Ziel von ELINET ist es, die Förderung von Schriftsprachkompetenzen als zentralen Bestandteil der Bildungspolitik in seinen Mitgliedsländern zu verankern mit dem Ziel, die Anzahl der Kinder, Jugendlichen und Erwachsenen mit geringen Schriftsprachkompetenzen in Europa zu senken. ELINET hat bei allen Aktivitäten einen altersübergreifenden Ansatz gewählt und versteht Lese- und Schreibförderung als lebenslange Aufgabe, die nicht nur von der Schule, sondern auch in der Familie und im Kindergarten, in Ausbildung und Beruf, in Gemeinden und kulturellen und sozialen Einrichtungen wahrgenommen werden muss. Zu den zahlreichen Instrumenten, die von ELINET entwickelt wurden, gehören die Entwicklung eines Europäischen Rahmenkonzepts guter Praxis der Lese- und Schreibförderung (Garbe, Mallows & Valtin, 2016) und eine Sammlung von Beispielen guter Praxis aus ganz Europa sowie die Erarbeitung von *Literacy-Länderberichten* für alle an ELINET beteiligten Mitgliedsländer. Die von der High Level Group herausgestellten drei Handlungsfelder stehen auch bei ELINET im Zentrum. Die auf der ELINET-Netzseite veröffentlichten Länderberichte folgen einem einheitlichen theoretischen Rahmenmodell, um eine Vergleichbarkeit der Leistungsdaten, Indikatoren und politischen Maßnahmen zwischen den Ländern zu gewährleisten (Brooks, Garbe, Mallows & Valtin, 2015). Sie stellen Informationen bereit über ausgewählte Ergebnisse aus den internationalen Lesestudien, wobei erstmals Ergebnisse aus IGLU, PISA und dem *Programme for the International Assessment of Adult Competencies* (PIACC) gemeinsam berichtet werden. Ferner werden Daten und Informationen zu Lernbedingungen sowie Maßnahmen und Projekte zur Lese- und Schreibförderung dargestellt. Die Berichte enthalten auch Hinweise auf die spezifischen Problembereiche jedes Landes, in denen der dringendste Handlungsbedarf besteht. Es gibt für Deutschland einen umfassenden, über 100 Seiten langen Bericht (Garbe et al., 2016) sowie einen kürzeren Bericht (Garbe, Lafontaine, Mallows & Valtin, 2016).

Der folgende Beitrag bezieht sich auf leseförderliche Rahmenbedingungen, Maßnahmen und Projekte sowie auf Beispiele guter Praxis, wie sie in den erfolgreichen europäischen Ländern aufzufinden sind. Als erfolgreich anzusehen sind Länder, die in den IGLU-Erhebungen bedeutsame Leistungszuwächse zeigen (Slowenien, Slowakei und Norwegen in der 4. Jgst.), sowie Länder, die 2016 in IGLU signifikant besser abschneiden als Deutschland (Mittelwert der Leseleistungen: 537), nämlich die Russische Föderation (581), Irland (567), Finnland (566), Polen (565), Nordirland (565), Norwegen (5. Jgst.) (559), England (559), Lettland (558), Schweden (555), Ungarn (554), Bulgarien (552), Litauen (548), Italien (548), Dänemark (547), und die Niederlande (545). In drei dieser europäischen Länder ist es zwischen 2001 und 2016 gelungen, den Anteil an Kindern, die nie oder fast nie zum Vergnügen lesen, substantiell zu reduzieren (England, Italien, Niederlande). Im Fokus dieses Beitrags steht Irland, das erfolg-

reichste EU-Land neben Finnland, über dessen Bildungssystem bereits zahlreiche Veröffentlichungen vorliegen (u. a. Sahlberg, 2015).

Da Lesekompetenz das kumulative Ergebnis eines Lernprozesses ist, der beeinflusst ist durch zahlreiche gesellschaftliche, schulische und individuelle Faktoren, die wiederum in Wechselwirkung zueinander stehen, können keine einzelnen Maßnahmen für den Lernerfolg oder -misserfolg verantwortlich gemacht werden, aber es lassen sich Rahmenbedingungen und Aktivitäten benennen, die in erfolgreichen Staaten realisiert wurden. Das soll in der Folge geschehen, wobei drei Handlungsfelder unterschieden werden, die sich allerdings überschneiden können. Vorauszuschicken ist, dass die in diesem Beitrag berichteten IGLU-Befunde sich auf Gesamtdeutschland beziehen. Zwischen den Ländern bestehen jedoch teilweise erhebliche Unterschiede in den Leseleistungen und auch den bildungspolitischen Rahmenbedingungen und Maßnahmen (u. a. Berkemeyer, Bos, Hermstein, Abendroth & Semper, 2017).

2 Schaffung einer schriftreichen Umgebung

Da das Lesenkönnen stark von der Lesemotivation abhängt, sind in Familie, Kindergarten und Schule Lernumgebungen zu gestalten, die Kinder zum Lesen lernen und zum Lesen anregen.

Gezielte Elternarbeit (Family Literacy Programme). Obwohl der Unterricht im Lesen und Schreiben eine schulische Angelegenheit darstellt, werden im Elternhaus wichtige Fundamente für den Schriftspracherwerb gelegt. Eltern sind für den frühkindlichen Spracherwerb und die Förderung der Begegnung mit Schriftsprache wichtig. Sie stellen kulturelle Ressourcen und Aktivitäten bereit und sind mit ihrem eigenen Leseverhalten Vorbilder. Die in diesem Band vorgelegten Daten verweisen auf die Wichtigkeit der Lesesozialisation in Deutschland, allerdings wurden nicht wie in früheren IGLU-Erhebungen die Indikatoren der Lesesozialisation im internationalen Vergleich dargestellt. Bei der Auswertung der IGLU-2011-Daten (Hoeft, Wendt & Kasper, 2015) zeigte sich, dass Deutschland bei der Häufigkeit leseförderlicher Aktivitäten, die Eltern mit ihren Kindern im Vorschulalter betreiben, im europäischen Mittelfeld liegt. Besonders interessant ist der Indikator „Bücher vorlesen“, der in der Forschung als wichtiger Prädiktor für den Leseerfolg gilt. Deutschland liegt mit einem Anteil von 70 Prozent der Eltern, die ihren Kindern im Vorschulalter oft Bücher vorgelesen haben, im europäischen Vergleich an der Spitze, gleich hinter Finnland mit 73,4 Prozent. Gleichwohl bedeutet dieses Ergebnis auch, dass mit 30 Prozent der Kinder nur wenig gelesen wird. Dies entspricht auch den Daten der Vorlesestudie (Stiftung Lesen, DIE ZEIT & Deutsche Bahn, 2017): 55 Prozent aller Eltern lesen ihren Kindern in den ersten 12 Monaten nicht regelmäßig vor. In 28 Prozent der Familien ist das sogar innerhalb der ersten drei Jahre nicht der Fall. Zu fordern ist deshalb gezielte Elternarbeit, um die Eltern vom Wert des Vorlesens und auch der Bedeutung der sprachlichen Förderung im Vorschulalter zu überzeugen und ihnen dabei Hilfen anzubieten.

Unter den Beispielen guter Praxis, die von ELINET ausgewählt wurden, befinden sich auch für den Bereich der *Family Literacy* einige Projekte, so das italienische Programm *Nati per Leggere* und aus England die *Bookstart*-Programme. *Nati per Leggere*, an dem auch Kinderärzte mitwirken, ist ein landesweites Projekt, das die Entwicklung schriftsprachlicher Voraussetzungen im Kindesalter, die Verbesserung der Eltern-Kind-Beziehungen und das Lese-

interesse fördert (Balbinot et al., 2016). In England gibt es eine Reihe landesweiter Programme (u.a. *Bookstart* und *Bookbuzz*) mit dem Ziel, schon bei kleinen Kindern das Vergnügen am Lesen durch das Bereitstellen kostenlosen Lesematerials zu fördern (Dombey, 2016). Sowohl in Italien als auch in England wird der Elternarbeit große Aufmerksamkeit gewidmet. Ob dies zu den Erfolgen beider Länder bei der Verminderung des Anteils von Kindern, die nicht zum Vergnügen lesen, zusammenhängt, kann nur vermutet werden.

Evaluiert wird in Deutschland gegenwärtig *Lesestart – Drei Meilensteine für das Lesen*, das größte bundesweite Leseförderprogramm, das Familien mit kleinen Kindern von Anfang an bis zu ihrem Eintritt in die Schule begleitet und vom *Bundesministerium für Bildung und Forschung* (BMBF) finanziert wird (<http://www.lesestart.de>). Ergebnisse wurden noch nicht veröffentlicht.

Gute Ausstattung von Schulbibliotheken

Für Schulen beinhaltet die Schaffung einer schriftreichen Umgebung die Bereitstellung reichhaltigen Lesematerials in gedruckter und in elektronischer Form, beispielsweise in einer Schulbibliothek. Nicht alle Schulen halten jedoch Schulbibliotheken vor. Den bei IGLU 2016 befragten Schulleitungen zufolge stand 2016 für 72.4 Prozent der Schülerinnen und Schüler in Deutschland eine Schulbibliothek zur Verfügung, in einer Reihe von Ländern sind die Prozentzahlen deutlich höher, so in Slowenien (100%) oder Dänemark (94%) (Mullis, Martin, Foy & Hooper, 2017). Der Deutsche Bibliotheksverband hat kürzlich Deutschland in Sachen Schulbibliotheken als Notstandsgebiet bezeichnet und folgende Ursachen für den unzureichenden Ausbaustand von Schulbibliotheken aller Schularten genannt: das Fehlen klar definierter rechtlicher und finanzieller Zuständigkeiten und verbindlicher Organisationsstrukturen in Bund, Ländern und Kommunen (Deutscher Bibliotheksverband, 2016).

Ausbau öffentlicher Bibliotheken. Auch ein verstärktes Angebot an öffentlichen Bibliotheken wäre wünschenswert, vor allem angesichts der im internationalen Vergleich niedrigen Quote bei der Bibliotheksnutzung. Es gibt zahlreiche interessante Projekte örtlicher Büchereien, die Kinder im Vor- und Grundschulalter und ihre Eltern zum Besuch von Büchereien und zum Lesen verlocken. Es handelt sich in Deutschland jedoch nur um nicht koordinierte, vereinzelte Initiativen.

Reichhaltige Angebote von motivierenden Leseanlässen. Ein Beispiel für die Zusammenarbeit von Eltern, Bibliothekaren, Verlagen, Autoren, Lehrkräften und Freiwilligen ist das Programm *Reading Badge* aus Slowenien, ein landesweites Leseanimationsprogramm, an dem 99 Prozent der Grundschulen beteiligt sind und dessen Ziel es ist, die Freude am Lesen zu wecken und zu erhalten. Die zahlreichen Initiativen, die mit diesem Projekt verbunden sind, haben in der Öffentlichkeit in Slowenien das Bewusstsein von der Bedeutung des Lesenkönnens geschärft, was möglicherweise zu der Verringerung der Anzahl der Kinder, die nicht zum Vergnügen lesen, beigetragen haben mag.

Auch in Deutschland gibt es in diesem Bereich wiederum nur einzelne Projekte, wobei beispielhaft verwiesen werden soll auf das Netzwerk *Leseförderung Lüneburg e.V.*, eine lokale Initiative, in der die Universität und Schulen, die Stadtbibliothek und Buchhandlungen, das örtliche Literaturhaus, der Verein Mentor e.V., die Landeszeitung Lüneburg und andere Akteure der Leseförderung für Kinder und Jugendliche sich zusammengeschlossen haben

und gemeinsame Projekte entwickeln und durchführen. Dazu gehören *Buchstart Lüneburg* und *Gedichte für Wichte* zur vorschulischen Förderung, die jährliche Aktionswoche *Lüneburg Liest* und das Projekt *Zeitungsleser – Weltentdecker* (Netzwerk Leseförderung Lüneburg e. V., 2017).

2.1 Verbesserung der Lernbedingungen und der Qualität des Unterrichts

Das Rahmenkonzept, das in ELINET für die Leseförderung entwickelt wurde, nennt für das Handlungsfeld „Verbesserung des Unterrichts“ zahlreiche Aspekte, wie die Qualität der Vorschule, ein kohärentes Schriftsprachcurriculum, anspruchsvoller Leseunterricht, frühes Erkennen und Fördern von leseschwachen Schülerinnen und Schülern sowie gut ausgebildete Lehrkräfte (Brooks et al., 2015). Betrachtet man die in IGLU erfolgreichen Staaten, so sind in vielen von ihnen die im Folgenden aufgeführten Maßnahmen realisiert beziehungsweise es sind Reformbemühungen erkennbar.

Sicherung einer qualitativ hochwertigen Bildung und Erziehung im Elementarbereich. In allen IGLU-Erhebungen zeigt sich sowohl international als auch in Deutschland, dass Kinder, die drei Jahre oder länger einen Kindergarten besuchten, bessere Leseleistungen im vierten Schuljahr aufweisen (Mullis, Martin, Foy & Drucker, 2012; Mullis et al., 2017, Exhibit 4.6). Für Deutschland zeigt sich allerdings, dass dieser Effekt bei Kontrolle des sozialen Hintergrunds gering ist (Steinmann, Zieger, Hogrebe & Strietholt, 2016). In einigen Staaten gibt es eine kostenfreie Elementarerziehung für alle Kinder, Sprachstandserhebungen und eine obligatorische Sprachförderung für Kinder mit entsprechendem Bedarf. Viele Staaten haben auch bessere Rahmenbedingungen im Vorschulbereich, beispielsweise bei der Höhe der Bildungsausgaben, die in Deutschland im internationalen Vergleich sehr gering ist, sowie der Qualifikation der Erzieherinnen und Erzieher (European Commission, EACEA, Eurydice & Eurostat, 2014).

In Finnland wurde kürzlich der Elementarbereich auch administrativ zum Bestandteil des Bildungssystems erklärt, um eine bessere Übereinstimmung der Curricula zu erzielen (Sahlberg, 2015, S. 51). Mehr als ein Drittel der Personen, die in vorschulischen Einrichtungen arbeiten, haben eine universitäre Ausbildung, gemeinsam mit den zukünftigen Lehrkräften.

Sowohl Finnland als auch Irland haben ein nationales Curriculum für den vorschulischen Bereich. An erster Stelle in beiden Curricula stehen Bildungsziele wie Wohlbefinden, Kommunikation, soziales Verhalten und Selbstständigkeit. Die Vorbereitung auf das schulische Lesen und Schreibenlernen findet in kommunikativ bedeutsamen Kontexten statt, damit Kinder die Funktion von geschriebener Sprache erfahren können (Finnish National Board of Education, 2010, S. 10–11). Aistear (NCCA, 2009), der irische Rahmenplan, setzt auf einen ganzheitlichen Ansatz, bei dem allgemeine Sprachförderung, das gemeinsame Lesen von Bilderbüchern und das Sprechen über das Gelesene im Mittelpunkt stehen. In Deutschland wird als Vorbereitung für das schulische Lesen- und Schreibenlernen vielfach der Einsatz von Trainingsprogrammen zur Förderung der „phonologischen Bewusstheit“ empfohlen, ein Konzept, das aus linguistischer, fachdidaktischer und lernpsychologischer Sicht fragwürdig ist (Valtin, 2012) und zudem keine langfristigen Effekte hat (u. a. Ennemoser & Hartung, 2017). Einen ähnlichen Ansatz zur Förderung schriftsprachlicher

Aktivitäten wie im finnischen und irischen Curriculum bietet ein Projekt der Kita Frankfurt, mit zahlreichen Anregungen, wie man Kinder im Vorschulalter durch eine schriftreiche Umgebung zum Lesen- und Schreibenlernen motivieren kann (Lenel, 2011).

Verbesserung der Rahmenbedingungen und des Unterrichts in der Grundschule. In diesem Abschnitt sollen nur solche Merkmale betrachtet werden, die in den IGLU-Erhebungen in Deutschland besonders problematische Befunde ergeben haben. Dazu gehören die für den Leseunterricht aufgewendete Zeit, die Qualität des Deutschunterrichts, die Förderung von Schülerinnen und Schülern mit Leseschwäche sowie die Lehrerfortbildung.

Für den Leseunterricht aufgewendete Zeit. Eine der zentralen Forderungen aus ELINET bezieht sich auf ausreichende Zeit für das Lesen, nicht nur im Unterricht der jeweiligen Landessprache, sondern auch fächerübergreifend. Schülerinnen und Schüler des vierten Schuljahrs in Deutschland verbringen laut Lehrerangaben 87 Stunden pro Jahr speziell mit Leseunterricht und/oder Leseaktivitäten, der internationale Mittelwert ist 156 (Mullis et al., 2017, Exhibit 9.1). Die in IGLU erfolgreichen Staaten bieten erheblich mehr Stunden an, so die Niederlande (205), die Russische Föderation (171), Irland (150) und England (125). Da in den anderen Staaten Ganztagschulsysteme bestehen, gibt es vermutlich auch mehr Zeit für Leseunterricht. Dieses Defizit an schulischer Unterrichtszeit wird in Deutschland nicht kompensiert durch häufiges außerschulisches Lesen, denn drei Viertel der befragten Schülerinnen und Schüler geben an, weniger als eine Stunde pro Schultag außerhalb der Schule zu lesen.

Qualitativ hochwertiger Leseunterricht. Ein guter Leseunterricht ist durch ein hohes kognitives Anregungspotential, Methodenvielfalt und individualisierte Lernangebote gekennzeichnet. Die Analyse der Lehrerangaben aus IGLU 2006 von Lankes und Carstensen (2007) verweist auf kulturelle Unterschiede in der Unterrichtsführung zwischen den europäischen Ländern. In Deutschland waren zwei Typen vorherrschend: Eng geführter Unterricht im Klassenverband ohne individuelle Unterstützung sowie wenig anregender, wenig abwechslungsreicher Klassenunterricht ohne individuelle Förderung. Abwechslungsreicher Unterricht mit ausgeprägter Förderkultur war in England und Schottland häufig verbreitet.

In Staaten wie Schweden, England, Finnland oder Dänemark enthält der Unterricht neben eng geführten Phasen der Instruktion und Anleitung häufig Phasen, in denen Kinder selbstständig viele unterschiedliche Texte, vor allem Bücher eigener Wahl, mit wiederum unterschiedlichen Intentionen lesen. Darüber hinaus gibt es in vielen Staaten eine ausgeprägte Förderkultur, um Kinder in ihrem individuellen Lernprozess zu unterstützen.

In den erfolgreichen Ländern ist der Anteil der Schülerinnen und Schüler, denen anspruchsvolle Lesestrategien vermittelt werden, deutlich höher als in Deutschland, in der Russischen Föderation sind es mindestens 90, aber zu meist 100 Prozent aller Kinder der vierten Klasse (Mullis et al., 2017, Exhibit 9.2). Obwohl seit Jahren in der fachdidaktischen Diskussion die Bedeutung von Lesestrategien betont wird und die Vermittlung von Lesestrategien in den länderübergreifenden Bildungsstandards und in den Curricula aller Länder der Bundesrepublik Deutschland verankert ist, wird dies von den Lehrkräften vernachlässigt. Dieses Thema sollte verstärkt in der Lehreraus- und -fortbildung aufgegriffen werden.

Realisierung des Förderanspruchs von Kindern mit Leseproblemen. In einigen europäischen Ländern haben Kinder mit Leseproblemen entweder einen verbrieften Anspruch auf schulische Förderung, wie zum Beispiel in Irland, oder es ist selbstverständlich, dass sie in der Schule besondere, individuelle Hilfen erhalten, wie beispielsweise beim *Reading-Recovery*-Programm in England, das auch zu den Gute-Praxis-Beispielen von ELINET gehört. In Deutschland besteht ein derartiger Anspruch nur auf dem Papier. Alle Länder der Bundesrepublik Deutschland haben Richtlinien, in denen lese-rechtschreibschwachen Schülerinnen und Schülern ein Anspruch auf Förderung zugesichert wird. Auch die „Grundsätze zur Förderung von Schülerinnen und Schülern mit besonderen Schwierigkeiten im Lesen und Rechtschreiben oder im Rechnen“ (KMK, 2007) sehen „individuelle Unterstützungsprogramme wie Intervallförderung oder Förderung in Zusatzkursen“ vor. Die Ergebnisse des *IQB-Ländervergleichs 2011* zur Prüfung der Bildungsstandards zeigten jedoch, dass nur eine Minderheit der Schülerinnen und Schüler, die die Mindeststandards verfehlen, tatsächlich eine besondere schulische Förderung erhält (Stanat, Weirich & Radmann, 2012). Das belegen auch die Befunde aus IGLU 2006 (Valtin, Hornberg, Buddeberg, Kowoll & Potthoff, 2010) und die in diesem Band berichteten Ergebnisse von 2016. Von den leseschwachen Schülerinnen und Schülern (Kompetenzstufe I und II) erhält nur jedes dritte Kind eine besondere schulische Förderung. Auch 2016 besteht also eine alarmierende Diskrepanz zwischen Förderbedarf und tatsächlicher Förderung (siehe auch Kapitel 10 in diesem Band). Zu dieser Diskrepanz tragen möglicherweise zwei Faktoren bei: erstens der Sachverhalt, dass nicht einmal jedes dritte Kind auf den beiden unteren Lesekompetenzstufen von den Lehrkräften als förderbedürftig eingeschätzt wird, und zweitens das Fehlen von zusätzlichem Personal. In erfolgreichen Ländern stehen Fachspezialisten (z.B. Leseexperten, Sprachtherapeuten, Beratungslehrer) oder zusätzliches pädagogisches Personal, wie Hilfsassistenten, zur Verfügung, mit dem Ziel, die individuelle Förderung von Schülerinnen und Schülern zu optimieren und die Belastung für Lehrkräfte zu kompensieren, was auch im Hinblick auf die angestrebte Inklusion von Bedeutung ist. Deutschland hat EU-weit die ungünstigsten Verhältnisse: Für 84 Prozent der Schülerinnen und Schüler (2011 waren es noch 78 %) stehen keine Fachspezialisten und für 66,6 Prozent keine Hilfslehrkräfte im Leseunterricht zur Verfügung. Nur Italien hat einen ungünstigeren Anteil an Fachspezialisten (87 %), allerdings stehen mehr Hilfslehrkräfte (44 %) zur Verfügung. Besonders günstige Werte erzielen Länder wie Finnland (2,5 % bzw. 22,5 %), Irland (2,2 % bzw. 65 %), Dänemark (7,1 % bzw. 55 %) und England (30,4 % und 5,5 %).

Verbesserung der Lehreraus- und -fortbildung. In ELINET wird die Bedeutung der Rolle der Lehrkraft für Lernerfolge, aber auch für das Wohlbefinden der Schülerinnen und Schüler herausgestellt. Eine qualitativ hochwertige Aus- und Fortbildung ist deshalb zu fordern, auch für Deutschland. Die IGLU-Ergebnisse weisen darauf hin, dass der Schulalltag durch starke Heterogenität in den Lernvoraussetzungen und Leistungen der Schülerinnen und Schüler gekennzeichnet ist. Dies ist allerdings kein spezifisch deutsches Problem, zumal in Ländern wie Skandinavien oder Italien, die signifikant bessere Leseleistungen als Deutschland erreichen, die Inklusion von Kindern mit Behinderungen schon seit Jahrzehnten realisiert ist (Valtin, Sander & Reinartz, 1990). In Deutschland müssen Lehrkräfte erst lernen, sich auf diese Heterogenität einzustellen. Insofern ist es wichtig, dieser Thematik einen bedeutenden Stellenwert nicht nur in

der Lehrerausbildung zu geben, sondern auch in der Fortbildung. In diesem Zusammenhang ist es bedenklich, dass es keine einheitlichen Länderregelungen zu Fortbildungsverpflichtungen oder Inhalten gibt und auch in Ländern der Bundesrepublik Deutschland mit einer Verpflichtung zur Fortbildung „ein substantieller Anteil der Lehrkräfte“ in den letzten beiden Schuljahren nicht an Fortbildungsveranstaltungen teilgenommen hat, wie die Ergebnisse des IQB-Ländervergleichs 2011 zeigen (Richter, Kuhl, Reimers & Pant, 2012). Im IQB-Bildungstrend 2016 wurden fachliche Weiterbildungen von Lehrkräften „nicht in die Analyse einbezogen“ (Rjosk, Hoffmann, Richter, Marx & Gresch, 2017, S. 350). Daten zur Lehrerfortbildung liegen jedoch in IGLU 2016, wie schon in den Vorjahren, vor. Die Lehrkräfte wurden gefragt, wie viel Zeit sie in den vorangehenden beiden Jahren in einer auf Lesen bezogenen Lehrerfortbildung verbracht haben. Hier die Angaben (in Klammern die Zahlen für die Russische Föderation, das Land, in dem die größte Verbesserung der Leseleistung zu verzeichnen ist): mehr als 16 Stunden 5 Prozent (59%), 6–15 Stunden 26 Prozent (16%), weniger als 6 Stunden 41 Prozent (15%), keine 27 Prozent (10%) (Mullis et al., 2017, Exhibit 8.4).

Umfassende Reformen im Bildungssystem. Wurde bislang auf einzelne Maßnahmen verwiesen, so sollen hier zwei Beispiele umfassender Reformen im Bildungs- beziehungsweise Schulsystem genannt werden. In Irland wurde 2011 eine Nationale Strategie zur Verbesserung der schriftsprachlichen und mathematischen Leistungen entwickelt, die sich auf alle Bereiche des Bildungssystems, auf die Einbeziehung von Eltern, den vorschulischen und schulischen Bereich erstreckt (Department of Education and Skills, 2011). Weitere Merkmale sind die Erhöhung des Anteils an Leseunterricht, die Betonung von Lesen als klassenübergreifende Aufgabe, eine stärkere Gewichtung von Lesen bei der Schulevaluation, die jährliche Durchführung von standardisierten Erhebungsverfahren in der zweiten, vierten und sechsten Klasse, eine bessere Lehrerausbildung und -fortbildung, die Revision von Curricula sowie umfassende Maßnahmen zur Sprachförderung im Vorschulalter, wie sie oben schon angesprochen wurden. Die Nationale Strategie wurde 2011 beschlossen, als die Viertklässlerinnen und Viertklässler, die bei IGLU 2016 so erfolgreich abgeschnitten haben, eingeschult wurden.

Das *Hamburger Sprachförderkonzept* (May & Berger, 2014; Behörde für Schule und Berufsbildung der Freien und Hansestadt Hamburg, 2015) richtet sich auf alle sprachbildenden Maßnahmen einer Schule. Sprachbildung wird als durchgängige Aufgabe des Unterrichts in allen Fächern angesehen. Der spezifische Sprachförderbedarf eines Kindes wird jährlich auf der Grundlage von standardisierten Testverfahren in den Bereichen Allgemeine Sprachentwicklung, Lesen und Rechtschreibung festgestellt. Laut Hamburgischem Schulgesetz sind Schülerinnen und Schüler, deren Sprachkenntnisse nicht ausreichen, um erfolgreich am Unterricht teilzunehmen, verpflichtet, zum Erwerb der deutschen Sprache oder zur Verbesserung ihrer Sprachkenntnisse an zusätzlichem Unterricht teilzunehmen. Die individuelle Förderung ist kontinuierlich und erfolgt auf diagnostischer Grundlage. Schulen erhalten besondere Ressourcen für die Sprachförderung von Kindern mit einem ausgeprägten Sprachförderbedarf. Jede Hamburger Schule hat zur Umsetzung des Sprachförderkonzepts eine qualifizierte Sprachlernberaterin oder einen qualifizierten Sprachlernberater, welche ein schulspezifisches Sprachförderkonzept erstellen und die Umsetzung

begleiten, evaluieren und weiterentwickeln. Im Rahmen des *Monitoring* berichten die Schulen jährlich über die Umsetzung und den Erfolg ihrer Sprachfördermaßnahmen. Nach Einführung des Sprachförderkonzepts hat Hamburg beachtliche Fortschritte gemacht. Laut *IQB-Bildungstrend* (Stanat, Schipolowski, Rjosk, Weirich & Haag, 2017) ist Hamburg das einzige Land, in dem sich der Anteil der Kinder, die den Regelstandard erreichen oder übertreffen, von 2011 zu 2016 signifikant erhöht hat. Der Anteil der Schülerinnen und Schüler, die nur den Mindeststandard im Lesen erfüllen, liegt in Hamburg bei 14,2 Prozent, in den anderen Stadtstaaten erheblich höher (Bremen 20 %, Berlin 25 %).

2.2 Steigerung von Teilhabe, Inklusion und Chancengerechtigkeit

Deutschland hat besondere Probleme mit der Chancengerechtigkeit, da es nach wie vor große Disparitäten in den Leistungen verschiedener sozialer Gruppen gibt. In manchen Ländern gelingt es wesentlich besser, diese Disparitäten gering zu halten, wie in Kapitel 6 in diesem Band gezeigt wurde. In der wissenschaftlichen und bildungspolitischen Diskussion besteht weitgehend Einigkeit, welche Instrumente erforderlich sind, um eine Bildungsbenachteiligung von Kindern aus bildungsfernen Elternhäusern möglichst kleinzuhalten (siehe auch OECD, 2016). Dazu gehören verschiedene, sehr früh einsetzende Präventionsmaßnahmen, die Unterstützung von Lernenden mit besonderem Förderbedarf (z. B. Zweitsprachenlerner, Migranten, Kinder aus benachteiligten Familien, Jungen) sowie curriculare und schulstrukturelle Maßnahmen.

Präventive Maßnahmen enthalten auch eine systematische Gesundheitsvorsorge. Besonders beeindruckend ist hier Finnland mit dem umfassenden Gesundheitsprogramm NEUVOLA (deutsch: Beratung), an dem so gut wie alle Kinder und ihre Mütter bereits während der Schwangerschaft, bis zum Schulbeginn teilnehmen. Bevor sie mit sieben Jahren in die Schule kommen, haben die Kinder mehrere Dutzend Gesundheitschecks hinter sich, die alle Bereiche von den Zähnen bis zu Augen, Ohren, Motorik, Sprache und psychosoziales Verhalten betreffen. Bei Problemen wird sofort therapeutisch reagiert, auch mit dem Ziel, in der Schule Lernschwierigkeiten zu vermeiden. Die enormen Kosten dieses Unterfangens lohnen sich, so hat Finnland den geringsten Anteil an Kindern mit sehr schwacher Lesekompetenz.

Zu präventiven Maßnahmen gehören Anreize, dass vorschulische Einrichtungen auch von bildungsbenachteiligten Kindern besucht werden, sowie die Bereitstellung besonderer Förderangebote, wie es im *Early-Start-Programm* in Irland der Fall ist, das sich an Kinder und Eltern in ungünstigen sozialen Lagen richtet (Curtin, Hall, Kennedy & Shiel, 2016).

Förderung von Schülerinnen und Schülern mit Migrationshintergrund. In fast allen an IGLU beteiligten Staaten gibt es zwischen Kindern mit und ohne Migrationshintergrund Leistungsunterschiede im Lesen, die sich in Deutschland vorwiegend auf Unterschiede in den sozioökonomischen und soziokulturellen Lebensverhältnissen der Schülerfamilien zurückführen lassen (siehe Kapitel 7 in diesem Band). Förderprogramme sollten entsprechend – auch um Stigmatisierungen zu vermeiden – konzeptionell primär am konkreten Förderbedarf der individuellen Schülerinnen und Schüler orientiert sein. Auf zwei vielversprechende, allerdings lokale Projekte sei hier verwiesen, die sich auch in der ELINET-Sammlung guter Praxisbeispiele finden:

Vorlesen in Familien ist ein sozialpräventives Projekt der *Phantastischen Bibliothek Wetzlar* zur Unterstützung der gesellschaftlichen Integration von Kindern aus bildungsbenachteiligten Familien. *FLY – ein Family-Literacy-Projekt in Hamburg* ist ein integrativer Ansatz zur aktiven Eltern- und Kindarbeit im Rahmen der Sprachförderung.

Derartige wichtige und begrüßenswerte Initiativen werden in den Ländern in Deutschland aber nur selten koordiniert oder flächendeckend gefördert. Ein Beispiel für einen landesweiten Aktionsplan liefert Irland: *Delivering Equality of Opportunity in Schools* (DEIS) ermöglicht frühkindliche Erziehung für bildungsbenachteiligte Kinder, identifiziert Schulen in bildungsbenachteiligten Gebieten und bietet – je nach Bedarf – verschiedene Unterstützungsmaßnahmen an und gibt Geld für Bücher und Schulbibliotheken. Die bisherigen Evaluationen zeigen positive Ergebnisse (Department of Education and Skills, 2017).

Schulstrukturelle Faktoren. Deutschland zeichnet sich im europäischen Vergleich durch zwei schulstrukturelle Besonderheiten aus: das weitgehende Fehlen von Ganztagsangeboten sowie die frühe Aufteilung von Schülerinnen und Schülern in mehr oder weniger gegliederte Bildungsgänge in der Sekundarstufe. Beide Faktoren werden als Hindernis für die Herstellung von Chancengleichheit angesehen.

Wie Daten aus TIMSS 2015 zeigen, besuchen 43 Prozent der Viertklässlerinnen und Viertklässler in Deutschland eine Ganztagsgrundschule. Davon besuchen nur 6 Prozent eine Ganztagschule mit rhythmisiertem Angebot, also einen Schultyp, von dem besonders günstige Wirkungen auf die Lern- und Persönlichkeitsentwicklung zu erwarten sind (Wendt, Goy, Walzebug & Valtin, 2016). Immerhin wünschen sich laut JAKO-O-Elternbefragung 70 Prozent der Eltern eine Ganztagschule und nur 28 Prozent eine Halbtagsschule für ihr Kind (Valtin, 2012). Eltern haben aber auch noch in einem anderen Bereich Wünsche, die nicht dem gegenwärtigen Stand entsprechen. Die JAKO-O-Elternbefragung von 2014 ergab – wie schon 2012 – eine große Mehrheit der Eltern, die eine frühe Aufteilung der Schülerinnen und Schüler nach dem vierten Schuljahr ablehnen. 58 Prozent sprechen sich für eine Aufteilung nach Klasse 6, 17 Prozent für eine Aufteilung nach Klasse 9 aus (Killus & Tillmann, 2014). IGLU hat wiederholt gezeigt, dass bei der Aufteilung der Schülerinnen und Schüler nach Klasse 4 eine erhebliche Benachteiligung von Kindern aus bildungsfernen Elternhäusern festzustellen ist (Arnold, Bos, Richert & Stubbe, 2007, 2010, Kapitel 8 in diesem Band).

3 Bilanz

Eine ELINET-Arbeitsgruppe hat eine „Europäische Erklärung des Grundrechts auf Lese- und Schreibkompetenz“ formuliert, die 2016 von allen an ELINET beteiligten Partnerorganisationen verabschiedet und in 24 Sprachen übersetzt wurde (Valtin et al., 2016). Lesenkönnen wird schon seit Jahrzehnten als Menschenrecht angesehen, erstmals werden hier aber Voraussetzungen beschrieben, die erfüllt sein müssen, damit ein derartiges Recht auch realisiert werden kann. Viele dieser Bedingungen wurden im vorliegenden Beitrag behandelt. Nicht thematisiert wurde die digitale Kompetenz, die zukünftig für das lebenslange Lernen, die Berufsausübung und die gesellschaftliche Teilhabe unverzichtbar ist. Die Kultusministerkonferenz hat mit ihrer 2016 beschlossenen Strategie „Bildung in

der Digitalen Welt“ anspruchsvolle Ziele gesetzt: Schon in den nächsten zwei Jahren soll der vorgeschlagene Kompetenzrahmen in die Lehrpläne integriert und Bestandteil des Curriculums für jedes Fach sein. Ferner sollen für alle Schülerinnen und Schüler digitale Lernumgebungen geschaffen werden. Wie Schülerinnen und Schüler beim Lesen elektronischer Texte abschneiden, wurde zeitgleich mit IGLU 2016 bei ePIRLS (Mullis & Martin, 2015) untersucht, leider hat Deutschland nicht an dieser Studie teilgenommen.

Im europäischen Vergleich der hier betrachteten bildungspolitischen, curricularen und unterrichtlichen Faktoren wird deutlich, dass es für Deutschland noch ein weiter Weg ist, bis alle Voraussetzungen geschaffen sind zur Erfüllung des Grundrechts auf Lese- und Schreibkompetenz:

Jede Person in Europa hat das Recht, angemessene Lese- und Schreibkompetenz zu erwerben. Die EU-Mitgliedsstaaten gewährleisten, dass alle Menschen ungeachtet ihres Alters oder Geschlechts, ihrer sozialen und ethnischen Herkunft sowie ihrer religiösen Orientierung über die nötigen Ressourcen und Möglichkeiten verfügen, um gute Lese- und Schreibfähigkeiten zu erwerben, damit sie geschriebene Kommunikation in gedruckter und auch digitaler Form wirksam verstehen und verwenden können. (Valtin et al., 2016)

Literatur

- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T. (2007). Schullaufbahnpräferenzen am Ende der vierten Jahrgangsstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 271–297). Münster: Waxmann.
- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T. (2010). Der Übergang von der Grundschule in die Sekundarstufe: Schullaufbahnpräferenzen von Lehrkräften und Eltern im Ländervergleich. 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. 13–32). Münster: Waxmann.
- Balbinot, V., Cunha, J., Garbe, C., Lafontaine, D., Diónisio, M. L., Shiel, G., Tamburlini, G. & Valtin, R. (2016). *Literacy in Italy. Country report. Children and adolescents*. Zugriff am 20.11.2017 unter <http://www.eli-net.eu/research/country-reports/>
- Behörde für Schule und Berufsbildung der Freien und Hansestadt Hamburg. (2015). *Handreichung. Inklusive Bildung und sonderpädagogische Förderung 4. Baustein. Integriertes Förderkonzept*. Zugriff am 20.11.2017 unter <http://li.hamburg.de/content-blob/7569306/bd0b4d6341fd011881077b7719267506/data/download-pdf-bsb-handreichung-integriertes-foerderkonzept.pdf>
- Berkemeyer, N., Bos, W., Hermstein, B., Abendroth, S. & Semper, I. (2017). *Chancenspiegel – eine Zwischenbilanz. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002*. Gütersloh: Bertelsmann Stiftung.
- Brooks, G., Garbe, C., Mallows, D. & Valtin, R. (2015). *ELINET country reports – Frame of reference*. Zugriff am 20.11.2017 unter <http://www.eli-net.eu/research/country-reports/>
- Curtin, A., Hall, K., Kennedy, E. & Shiel, G. (2016). *Literacy in Ireland. Country report short version*. Zugriff am 20.11.2017 unter http://www.eli-net.eu/fileadmin/ELINET/Redaktion/user_upload/Ireland_Short_Report2.pdf
- Department of Education and Skills. (2011). *Literacy and numeracy for learning and life. The national strategy to improve literacy and numeracy among children and young people 2011–2020*. Zugriff am 20.11.2017 unter https://www.education.ie/en/Publications/Policy-Reports/lit_num_strategy_full.pdf

- Department of Education and Skills. (2017). *20.11.2017*. Zugriff am 20.11.2017 unter <https://www.education.ie/en/Schools-Colleges/Services/DEIS-Delivering-Equality-of-Opportunity-in-Schools/>
- Deutscher Bibliotheksverband. (2016). *Bibliothek und Schule – zwei starke Partner*. Zugriff am 20.11.2017 unter <http://www.bibliotheksportal.de/themen/bibliothek-und-bildung/bibliothek-und-schule/schulbibliotheken.html>
- Dombey, H. (2016). *Literacy in England. Country report short version*. Zugriff am 20.11.2017 unter http://www.eli-net.eu/fileadmin/ELINET/Redaktion/user_upload/England_Short_Report1.pdf
- Ennemoser, M. & Hartung, N. (2017). Wirksamkeit verschiedener Sprachfördermaßnahmen bei Risikokindern im Vorschulalter. *Unterrichtswissenschaft*, 45 (3), 198–219.
- European Commission. (2012). *EU high level group of experts on literacy. Final report, September 2012*. Zugriff am 20.11.2017 unter http://ec.europa.eu/dgs/education_culture/repository/education/policy/school/doc/literacy-report_en.pdf
- European Commission. (2016). *PISA 2015 – EU performance and initial conclusions regarding education policies in Europe*. Zugriff am 20.11.2017 unter https://ec.europa.eu/education/sites/education/files/pisa-2015-eu-policy-note_en.pdf
- European Commission, EACEA, Eurydice & Eurostat. (2014). *Key data on early childhood education and care in Europe. 2014 Edition*. Eurydice and Eurostat Report. Luxembourg: Publications Office of the European Union.
- Finnish National Board of Education. (2010). *National Core Curriculum for Pre-primary Education*. Zugriff am 20.11.2017 unter http://www.oph.fi/english/curricula_and_qualifications/pre-primary%20education
- Garbe, C., Grotluschen, A., Lafontaine, D., Shiel, G., Tröster, M. & Valtin, R. (2016). *Literacy in Germany. Country report. Children, adolescents and adults*. Zugriff am 20.11.2017 unter http://www.eli-net.eu/fileadmin/ELINET/Redaktion/user_upload/Germany_Comprehensive_Report1.pdf
- Garbe, C., Lafontaine, D., Mallows, D. & Valtin, R. (2016). *Literacy in Germany. Country report short version*. Zugriff am 20.11.2017 unter http://www.eli-net.eu/fileadmin/ELINET/Redaktion/user_upload/Germany_Short_Report1.pdf
- Garbe, C., Lafontaine, D., Shiel, G., Sulkunen, S. & Valtin, R. (2016). *Literacy in Finland. Country report. Children and adolescents*. Zugriff am 18.11.2017 unter <http://www.eli-net.eu/research/country-reports/>
- Garbe, C., Mallows, D. & Valtin, R. (2016). *European framework of good practice in raising literacy levels of children, adolescents and adults*. Zugriff am 20.11.2017 unter <http://www.eli-net.eu/good-practice/europeanframeworkofgoodpractice/>
- Hoefl, M., Wendt, H. & Kasper, D. (2015). Familiäre Lernumwelten in Europa – Zusammenhänge formeller und informeller häuslicher Aktivitäten zwischen Eltern und Kindern im Vorschulalter in den Förderdimensionen Early Literacy und Early Numeracy. 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. 135–160). Münster: Waxmann.
- Killus, D. & Tillmann, K.-J. (2014). *Eltern zwischen Erwartungen, Kritik und Engagement: Ein Trendbericht zu Schule und Bildungspolitik in Deutschland*. Waxmann: Münster.
- KMK – Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2007). *Grundsätze zur Förderung von Schülerinnen und Schülern mit besonderen Schwierigkeiten im Lesen und Rechtschreiben oder im Rechnen. Beschluss der Kultusministerkonferenz vom 04.12.2003 i.d.F. vom 15.11.2007*. Zugriff am 20.11.2017 unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschlusse/2003/2003_12_04-Lese-Rechtschreibschwaecher.pdf
- Lankes, E.-M. & Carstensen, C. H. (2007). Der Leseunterricht aus der Sicht der Lehrkräfte. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006 Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 161–193). Münster: Waxmann.
- Lenel, A. (2011). *Sprechen, Schreiben, Lesen – Kinder auf dem Weg zur Schrift, Kita Frankfurt*. Zugriff am 20.11.2017 unter <http://www.kitafrankfurt.de/publikationen/broschueren/>
- May, P. & Berger, C. (2014). Monitoring des Hamburger Sprachförderkonzepts. In R. Valtin & I. Tarelli (Hrsg.), *Lesekompetenz nachhaltig stärken. Evidenzbasierte Maßnahmen und Programme* (S. 225–246). Berlin: Deutsche Gesellschaft für Lesen und Schreiben.

- Mullis, I. V. S., Martin, M. O., Foy, P. & Drucker, K. T. (2012). *PIRLS 2011 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S. & Martin, M. O. (Hrsg.). (2015). *PIRLS 2016 assessment framework*. Zugriff am 19.11.2017 unter <http://timssandpirls.bc.edu/pirls2016/framework.html>
- Mullis, I. V. S., Martin, M. O., Foy, P. & Hooper, M. (2017). *PIRLS 2016 international results in reading*. Zugriff am 05.12.2017 unter <http://timssandpirls.bc.edu/pirls2016/international-results/>
- NCCA – National Council for Curriculum and Assessment. (2009). *Aistear: The early childhood curriculum framework: Principles and themes*. Zugriff am 20.11.2017 unter http://www.ncca.ie/en/Curriculum_and_Assessment/Early_Childhood_and_Primary_Education/Early_Childhood_Education/Aistear_Toolkit/Aistear_and_PSC_audit.pdf
- Netzwerk Leseförderung Lüneburg e. V. (2017). Zugriff am 20.11.2017 unter <http://www.netzwerk-lesefoerderung.de/>
- OECD – Organisation for Economic Co-operation and Development. (2016). *Low-performing students: Why they fall behind and how to help them succeed, PISA*. Paris: OECD Publishing. doi: 10.1787/9789264250246-en.
- 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.
- Rjosk, C., Hoffmann, L., Richter, D., Marx, A. & Gresch, C. (2017). Qualifikation von Lehrkräften und Einschätzungen zum gemeinsamen Unterricht von Kindern mit und Kindern ohne sonderpädagogischen Förderbedarf. In P. Stanat, S. Schipolowski, C. Rjosk, S. Weirich & N. Haag (Hrsg.), *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich* (S. 335–353). Münster: Waxmann.
- Sahlberg, P. (2015). *Finnish Lessons 2.0. What can we learn from educational change in Finland?* (2. Auflage). New York: Teachers College Press.
- Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S. & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.
- Stanat, P., Weirich, S. & Radmann, S. (2012). Sprach- und Leseförderung. 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. 251–276). Münster: Waxmann.
- Stiftung Lesen, DIE ZEIT & Deutsche Bahn. (2017). *Vorlesestudie Vorlesen – aber ab wann? Vorlesestudie 2017 – Vorlesen und Erzählen als sprachliche Impulse in den ersten Lebensjahren*. Zugriff am 20.11.2017 unter <https://www.stiftunglesen.de/forschung/forschungsprojekte/vorlesestudie>
- Steinmann, I., Zieger, L. R., Hogrebe, N. & Strietholt, R. (2016). Lesen Kinder, die nicht in der Kita waren, am Ende der Grundschule schlechter? In R. Strietholt, W. Bos, H. G. Holtappels & N. McElvany (Hrsg.), *Jahrbuch der Schulentwicklung* (Daten, Beispiele und Perspektiven, Bd. 19, S. 161–185). Weinheim: Beltz/Juventa.
- Valtin, R. (2012). Phonologische Bewusstheit: Ein kritischer Blick auf ein modisches Konstrukt. *Frühe Bildung*, 1 (4), 223–225.
- Valtin, R., Bird, V., Brooks, G., Brozo, B., Clement, C., Ehmig, S., Garbe, C., Greef, M. de, Hanemann, U., Hammink, K., Mallows, D., Nascimbeni, F., Sulkunen, S. & Tamburlini, G. (2016). *Europäische Erklärung des Grundrechts auf Lese- und Schreibkompetenz*. Zugriff am 20.11.2017 unter <http://www.eli-net.eu/about-us/literacy-declaration/>
- Valtin, R., Hornberg, S., Buddeberg, M., Kowoll, M. E. & Potthoff, B. (2010). Schülerinnen und Schüler mit Leseproblemen – eine ökosystemische Betrachtungsweise. 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. 43–90). Münster: Waxmann.
- Valtin, R., Sander, A. & Reinartz, A. (Hrsg.). (1990). *Gemeinsam leben – gemeinsam lernen. Zur Integration behinderter Kinder in der Grundschule* (2. Auflage). Frankfurt: Arbeitskreis Grundschule.

Wendt, H., Goy, A., Walzebug, A. & Valtin, R. (2016). Bildungsangebote an Ganz- und Halbtagsgrundschulen in Deutschland. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland* (S. 225–246). Münster: Waxmann.

Anhang A

Tabelle A.1: Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2001

IGLU/PIRLS 2001		Teilnahme an IGLU/PIRLS†			Besonderheiten bzgl. der nationalen Zielpopulation				Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten				
		2016	2011	2006	Getestete Jahrgangsstufe	Durchschnittsalter*	Ausschöpfungsgrad**	Ausschlüsse***	Schulteilnahmequote in %		Schülerteilnahmequote in %	Gesamteilnahmequote in %	
	Teilnehmer							ohne Ersatzschulen	mit Ersatzschulen		ohne Ersatzschulen	mit Ersatzschulen	
	Argentinien				4	10.2	100	4.1	89	92	91	81	84
³	Belize				4	9.8	100	0.8	80	80	94	75	75
	Bulgarien	✓	✓	✓	4	10.9	100	2.7	97	97	97	93	93
	Deutschland	✓	✓	✓	4	10.5	100	1.8	98	98	88	86	86
^{1 2 3}	England ^A	✓	✓	✓	5	10.2	100	5.7	57	87	94	54	82
²	Frankreich	✓	✓	✓	4	10.1	100	5.3	93	97	97	90	94
^{2 3}	Griechenland				4	9.9	100	7.3	78	85	97	76	82
³	Hongkong	✓	✓	✓	4	10.2	100	2.8	73	98	99	72	97
	Iran	✓	✓	✓	4	10.4	100	0.5	97	100	98	95	98
	Island			✓	4	9.7	100	3.1	95	95	87	82	82
²	Israel ^B	✓	✓	✓	4	10.0	100	22.4	96	98	96	92	94
	Italien	✓	✓	✓	4	9.8	100	2.9	90	100	98	88	98
³	Kolumbien		✓		4	10.5	100	3.3	80	98	96	76	94
	Kuwait		(✓)	✓	4	9.9	100	0.0	87	89	91	80	81
	Lettland	✓		✓	4	11.0	100	4.6	89	96	93	83	89
^{2 3}	Litauen	✓	✓	✓	4	10.9	90	3.8	56	97	85	47	83
³	Marokko	✓	✓	✓	4	11.2	100	1.0	74	74	93	69	69
	Mazedonien			✓	4	10.7	100	4.2	97	97	97	94	94
¹	Neuseeland ^A	✓	✓	✓	5	10.1	100	3.2	94	100	96	90	96
³	Niederlande	✓	✓	✓	4	10.3	100	3.7	53	89	98	52	87
³	Norwegen	(✓)	✓	✓	4	10.0	100	2.8	82	89	92	76	82
²	Ontario, Kanada ^D	(✓)	(✓)	✓	4	10.0	60	5.4	90	97	94	85	91
²	Québec, Kanada ^D	(✓)	(✓)	✓	4	10.0	60	5.4	90	97	94	85	91
³	Republik Moldau			✓	4	10.8	100	0.5	84	100	96	81	96
	Rumänien		✓	✓	4	11.1	100	4.5	96	96	97	93	93
^{1 2}	Russische Föderation ^C	✓	✓	✓	3 oder 4	10.3	100	6.6	100	100	97	97	97
^{1 3}	Schottland ^A			✓	5	9.8	100	4.7	76	79	95	72	74
	Schweden	✓	✓	✓	4	10.8	100	5.0	97	99	93	90	92
	Singapur	✓	✓	✓	4	10.1	100	1.4	100	100	98	98	98
	Slowakei	✓	✓	✓	4	10.3	100	2.0	88	100	96	84	96
¹	Slowenien ^A	✓	✓	✓	3	9.8	100	0.3	98	99	95	94	94
	Tschechische Republik	✓	✓		4	10.5	100	5.0	90	95	94	85	90
	Türkei				4	10.2	100	3.9	100	100	97	97	97
	Ungarn	✓	✓	✓	4	10.7	100	2.1	98	98	97	95	95
^{2 3}	USA	✓	✓	✓	4	10.2	100	5.3	61	86	96	59	83
	Zypern				4	9.7	100	2.0	98	100	97	95	97

Tabelle A.2: Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2006

IGLU/PIRLS 2006		Teilnahme an IGLU/PIRLS ¹			Besonderheiten bzgl. der nationalen Zielpopulation				Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten				
Teilnehmer		2016	2011	2001	Getestete Jahrgangsstufe	Durchschnittsalter*	Ausschöpfungsgrad**	Ausschlüsse***	Schulteilnahmequote in %		Schülerteilnahmequote in %	Gesamteilnahmequote in %	
									ohne Ersatzschulen	mit Ersatzschulen		ohne Ersatzschulen	mit Ersatzschulen
² Alberta, Kanada			(✓)		4	9.9	100	7.1	100	100	96	96	96
^{2 3} Belgien (Fläm. Gem.)	✓				4	10.0	100	7.1	69	92	99	68	91
Belgien (Franz. Gem.)	✓	✓			4	9.9	100	3.9	85	100	95	81	95
² Britisch-Kolumbien, Kanada					4	9.8	100	7.6	98	99	95	93	94
² Bulgarien	✓	✓	✓		4	10.9	100	6.4	88	97	97	85	94
² Dänemark	✓	✓			4	10.9	100	6.2	89	99	97	86	96
Deutschland	✓	✓	✓		4	10.5	100	0.7	97	99	94	90	92
¹ England ^E	✓	✓	✓		5	10.3	100	2.4	86	99	93	80	92
Frankreich	✓	✓	✓		4	10.0	100	3.8	94	97	98	92	95
² Georgien	✓	✓			4	10.1	80	7.3	94	100	98	93	98
Hongkong	✓	✓	✓		4	10.0	100	3.9	91	100	97	89	97
Indonesien		✓			4	10.4	100	3.2	99	100	98	97	98
Iran	✓	✓	✓		4	10.2	100	3.8	100	100	99	99	99
Island			✓		4	9.8	100	3.8	99	99	91	90	90
² Israel ^F	✓	✓	✓		4	10.1	100	22.5	98	100	93	91	93
² Italien	✓	✓	✓		4	9.7	100	5.3	91	100	97	88	97
Katar	✓	✓			4	9.8	100	1.4	100	100	94	94	94
Kuwait		(✓)	✓		4	9.8	100	0.3	99	99	89	88	88
Lettland	✓		✓		4	11.0	100	4.7	97	98	94	91	92
² Litauen	✓	✓	✓		4	10.7	93	5.1	99	100	92	90	92
¹ Luxemburg ^G					5	11.4	100	3.9	100	100	99	99	99
Marokko	✓	✓	✓		4	10.8	100	1.1	98	99	95	93	94
Mazedonien			✓		4	10.6	100	4.9	100	100	96	96	96
Neuschottland, Kanada					4	10.0	100	4.0	99	100	96	96	96
^{1 2} Neuseeland ^E	✓	✓	✓		4.5–5.5	10.0	100	5.3	92	99	96	88	95
³ Niederlande	✓	✓	✓		4	10.3	100	3.6	70	93	97	67	90
³ Norwegen	(✓)	✓	✓		4	9.8	100	3.8	68	82	87	58	71
² Ontario, Kanada	(✓)	(✓)	✓		4	9.8	100	8.3	88	90	97	85	87
² Österreich	✓	✓			4	10.3	100	5.1	100	100	98	97	97
² Polen	✓	✓			4	9.9	100	5.1	99	100	95	94	95
Québec, Kanada	(✓)	(✓)	✓		4	10.1	100	3.6	96	96	84	81	81
² Republik Moldau			✓		4	10.9	91	0.6	98	100	95	93	95
Rumänien		✓	✓		4	10.9	100	2.4	99	99	98	97	97
² Russische Föderation	✓	✓	✓		4	10.8	100	7.7	100	100	97	97	97
^{1 3} Schottland ^E			✓		5	9.9	100	2.3	69	87	94	65	81
Schweden	✓	✓	✓		4	10.9	100	3.9	100	100	96	96	96
Singapur	✓	✓	✓		4	10.4	100	0.9	100	100	95	95	95
Slowakei	✓	✓	✓		4	10.4	100	3.6	93	98	96	89	94
¹ Slowenien ^H	✓	✓	✓		3 oder 4	9.9	100	0.8	93	97	96	90	93
² Spanien	✓	✓			4	9.9	100	5.3	99	100	97	95	97
¹ Südafrika ^G	(✓)	(✓)			5	11.9	100	4.3	94	96	92	86	88
Taiwan	✓	✓			4	10.1	100	2.9	98	100	99	97	99
¹ Trinidad & Tobago ^E	✓	✓			5	10.1	100	0.7	99	99	95	94	94
Ungarn	✓	✓	✓		4	10.7	100	3.7	99	100	97	96	97
^{2 3} USA	✓	✓	✓		4	10.1	100	5.9	57	86	96	54	82

Tabelle A.3: Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2011

IGLU/PIRLS 2011		Teilnahme an IGLU/ PIRLS ¹			Besonderheiten bzgl. der nationalen Zielpopulation				Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten				Besonderheiten bzgl. der Leistungs- messung	
		2016	2006	2001	Getestete Jahr- gangs- stufe	Durch- schnitts- alter	Aus- schöp- fungs- grad**	Aus- schlüs- se***	Schulteilnahme- quote in %		Schü- lerteil- nahme- quote in %	Gesamtteil- nahmequote in %		Anteil der Schülerinnen und Schüler ohne gültige Leistungswerte in %
Teilnehmer								ohne Ersatz- schu- len	mit Ersatz- schu- len			ohne Ersatz- schu- len	mit Ersatz- schu- len	
^{2 3}	Aserbaidshon	✓			4	10.2	100	7.2	84	100	100	84	100	3
	Australien	✓			4	10.0	100	4.4	96	98	95	91	93	2
^{2 3}	Belgien (Franz. Gem.)	✓	✓		4	10.1	100	5.6	77	85	97	74	82	1
	Bulgarien	✓	✓	✓	4	10.7	100	2.5	97	100	95	92	95	2
²	Dänemark	✓	✓		4	10.9	100	7.3	87	98	97	84	95	0
	Deutschland	✓	✓	✓	4	10.4	100	1.9	96	99	96	92	95	1
^{1 3}	England ^l	✓	✓	✓	5	10.3	100	2.4	73	87	94	69	82	2
	Finnland	✓	✓		4	10.8	100	3.1	97	99	96	93	95	0
²	Frankreich	✓	✓	✓	4	10.0	100	5.2	98	100	98	96	97	1
²	Georgien	✓	✓		4	10.0	92	4.9	97	98	98	95	96	3
²	Hongkong	✓	✓	✓	4	10.1	100	11.8	86	88	94	81	83	0
	Indonesien	✓	✓		4	10.4	100	2.5	100	100	97	97	97	7
	Iran	✓	✓	✓	4	10.2	100	4.5	100	100	99	99	99	6
	Irland	✓			4	10.3	100	2.5	98	100	95	93	95	1
²	Israel ^l	✓	✓	✓	4	10.1	100	24.6	98	99	94	92	93	2
³	Italien	✓	✓		4	9.7	100	3.7	81	98	96	78	95	1
²	Kanada	✓			4	9.9	100	9.9	98	98	96	94	94	1
²	Katar	✓	✓		4	10.0	100	6.2	100	100	99	99	99	11
	Kolumbien			✓	4	10.4	100	1.5	89	99	97	86	95	7
²	Kroatien				4	10.7	100	7.9	99	100	95	94	95	0
²	Litauen	✓	✓	✓	4	10.7	93	5.6	94	100	94	89	94	1
¹	Malta ^l	✓			5	9.8	100	3.6	100	100	95	95	95	6
⁴	Marokko	✓	✓	✓	4	10.5	100	2.0	99	99	96	95	95	33
¹	Neuseeland ^l	✓	✓	✓	4.5-5.5	10.1	100	3.3	93	99	94	87	93	2
³	Niederlande	✓	✓	✓	4	10.2	100	3.7	68	92	97	66	89	0
³	Nordirland	✓			4	10.4	100	3.5	62	85	93	58	79	1
³	Norwegen	(✓)	✓	✓	4	9.7	100	4.2	57	83	86	49	71	1
⁴	Oman	✓			4	9.9	100	1.5	98	98	98	96	96	16
²	Österreich	✓	✓		4	10.3	100	5.1	100	100	98	98	98	1
	Polen	✓	✓		4	9.9	100	3.8	100	100	96	96	96	1
	Portugal	✓			4	10.0	100	2.5	87	99	95	83	93	1
	Rumänien		✓	✓	4	10.9	100	4.0	99	100	97	96	97	4
²	Russische Föderation	✓	✓	✓	4	10.8	100	5.3	100	100	98	98	98	0
	Saudi-Arabien	✓			4	10.0	100	1.6	95	100	98	94	98	9
	Schweden	✓	✓	✓	4	10.7	100	4.1	97	99	92	88	91	1
²	Singapur	✓	✓	✓	4	10.4	100	6.3	100	100	96	96	96	1
	Slowakei	✓	✓	✓	4	10.4	100	4.6	95	99	97	92	96	1
	Slowenien	✓	✓	✓	4	9.9	100	2.6	96	97	97	94	95	1
²	Spanien	✓	✓		4	9.8	100	5.4	96	99	97	93	96	1
	Taiwan	✓	✓		4	10.2	100	1.4	100	100	99	99	99	1
¹	Trinidad & Tobago ^l	✓	✓		5	10.3	100	0.9	99	99	96	95	95	5
²	Tschechische Republik	✓	✓	✓	4	10.4	100	5.1	90	99	94	85	94	0
	Ungarn	✓	✓	✓	4	10.7	100	4.2	98	99	97	94	96	2
^{2 3}	USA	✓	✓	✓	4	10.2	100	7.2	80	85	96	77	81	1
	Vereinigte Arabische Emirate (VAE)	✓			4	9.8	100	3.3	100	100	97	97	97	10
	Teilnahme mit Jgst. 6													
¹	Botsuana				6	12.8	100	0.3	100	100	99	99	99	9
¹	Honduras				6	12.7	100	4.5	91	100	97	88	97	7
^{1 2 3}	Kuwait		(✓)	(✓)	6	11.9	78	0.5	88	88	82	72	72	12
¹	Marokko				6	12.7	100	2.0	99	99	95	94	94	10
	Benchmark-Teilnehmer													
	Abu Dhabi, VAE	✓			4	9.7	100	2.7	99	99	97	96	96	11
²	Andalusien, Spanien	✓			4	9.9	100	5.1	99	99	97	96	96	1
²	Alberta, Kanada		(✓)		4	9.9	100	6.8	97	99	95	93	94	1
²	Dubai, VAE	✓			4	9.8	100	5.1	99	99	96	94	94	7
²	Florida, USA				4	10.4	89	12.9	96	96	95	91	91	1
¹	Malta (Maltesisch)				5	9.8	100	4.1	100	100	94	94	94	6
²	Ontario, Kanada	✓	(✓)	(✓)	4	9.8	100	7.9	99	99	96	95	95	1
	Québec, Kanada	✓	(✓)	(✓)	4	10.1	100	3.7	95	96	96	90	92	0
^{1 4}	Südafrika (5. Jgst.; Englisch, Afrikaans) ^k	✓	(✓)		5	11.4	100	1.9	98	100	94	92	94	16

Tabelle A.4: Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2016

IGLU/PIRLS 2016		Teilnahme an IGLU/PIRLS†			Besonderheiten bzgl. der nationalen Zielpopulation‡				Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten				Besonderheiten bzgl. der Leistungsmessung	
Teilnehmer		2011	2006	2001	Getestete Jahrgangsstufe	Durchschnittsalter*	Aus-schöpfungsgrad**	Aus-schlüsse***	Schulerteilnahmequote in %		Schülerteilnahmequote in %	Gesamterteilnahmequote in %		Anteil der Schülerinnen und Schüler ohne gültige Leistungswerte in %
									ohne Ersatz-schulen	mit Ersatz-schulen		ohne Ersatz-schulen	mit Ersatz-schulen	
Aserbaidschan		✓			4	10.1	100	2.1	100	100	96	96	96	3
Australien		✓			4	10.0	100	4.8	97	100	95	92	94	1
Bahrain					4	9.9	100	2.7	99	99	98	98	98	5
³ ⁸ Belgien (Fläm. Gem.)			✓		4	10.1	100	1.6	79	94	98	77	92	0
² ⁸ Belgien (Franz. Gem.)		✓	✓		4	10.0	100	6.0	96	100	97	93	97	1
Bulgarien		✓	✓	✓	4	10.8	100	4.3	100	100	95	95	95	1
Chile					4	10.1	100	4.0	92	100	96	88	96	2
² Dänemark		✓	✓		4	10.8	100	9.8	87	96	94	82	90	1
Deutschland		✓	✓	✓	4	10.3	100	4.2	97	100	96	93	95	1
¹ ⁸ England ^L		✓	✓	✓	5	10.3	100	3.7	99	100	96	95	96	1
Finnland		✓			4	10.8	100	2.4	98	99	96	95	96	0
² Frankreich		✓	✓	✓	4	9.8	100	5.4	99	100	96	95	96	1
² ⁸ Georgien ^M		✓	✓		4	9.7	96	3.8	98	99	97	95	96	2
² ³ ⁸ Hongkong		✓	✓	✓	4	9.9	100	10.1	74	91	87	64	79	0
⁷ ⁸ Iran ^N		✓	✓	✓	4	10.2	100	4.1	100	100	99	99	99	5
Irland		✓			4	10.5	100	3.1	100	100	96	96	96	1
² ⁸ Israel		✓	✓	✓	4	10.0	100	24.9	98	99	95	93	94	2
⁸ Italien		✓	✓	✓	4	9.7	100	4.9	89	99	96	85	95	0
² ³ Kanada ^M		✓			4	9.9	97	7.5	81	90	96	77	86	1
Kasachstan					4	10.3	100	4.9	100	100	99	99	99	0
⁸ Katar		✓	✓		4	10.0	100	3.9	100	100	97	97	97	7
² Lettland			✓	✓	4	10.9	100	7.9	95	97	94	89	91	0
⁸ Litauen		✓	✓	✓	4	10.8	100	4.2	100	100	95	95	95	0
Macau					4	10.0	100	3.6	100	100	98	98	98	0
¹ ² Malta ^L		✓			5	9.7	100	7.9	100	100	96	96	96	5
⁷ ⁸ Marokko ^N		✓	✓	✓	4	10.2	100	1.7	100	100	99	99	99	10
¹ Neuseeland ^L		✓	✓	✓	4.5–5.5	10.1	100	3.7	85	97	96	81	92	2
³ Niederlande		✓	✓	✓	4	10.1	100	3.1	69	90	96	66	86	0
³ Nordirland		✓			4	10.4	100	3.0	84	88	96	81	84	1
Oman		✓			4	9.7	100	0.6	99	100	99	98	98	8
² Österreich		✓	✓		4	10.3	100	5.6	100	100	98	98	98	0
⁸ Polen		✓	✓		4	10.7	100	3.9	95	99	91	86	90	0
² Portugal		✓			4	9.8	100	7.5	97	99	94	91	93	0
⁸ Russische Föderation		✓	✓	✓	4	10.8	100	3.8	100	100	98	98	98	0
Saudi-Arabien		✓			4	9.9	100	2.3	92	100	96	88	96	7
² Schweden		✓	✓	✓	4	10.7	100	5.2	99	100	95	94	95	0
² ⁸ Singapur		✓	✓	✓	4	10.4	100	11.1	100	100	97	97	97	1
Slowakei		✓	✓	✓	4	10.4	100	4.8	94	100	97	92	97	1
⁸ Slowenien		✓	✓	✓	4	9.9	100	2.4	94	94	96	90	90	1
Spanien		✓	✓		4	9.9	100	4.8	99	100	97	95	97	0
Taiwan		✓	✓		4	10.1	100	0.9	100	100	98	98	98	0
¹ ⁸ Trinidad & Tobago ^L		✓	✓		5	10.2	100	1.3	100	100	96	96	96	4
Tschechien		✓		✓	4	10.3	100	3.4	100	100	95	95	95	1
Ungarn		✓	✓	✓	4	10.6	100	4.5	98	100	97	95	97	0
³ ⁸ USA		✓	✓	✓	4	10.1	100	4.8	75	92	94	71	86	1
Vereinigte Arabische Emirate (VAE)		✓			4	9.8	100	3.3	98	99	96	95	95	6
Teilnahme mit Jgst. 5														
¹ ² Norwegen ^o					5	10.8	100	5.3	95	99	96	91	95	0
Benchmark-Teilnehmer														
Abu Dhabi, VAE		✓			4	9.7	100	3.9	100	100	96	96	96	10
Andalusien, Spanien		✓			4	9.8	100	4.2	99	100	96	96	96	0
Buenos Aires, Argentinien					4	10.0	100	2.8	88	100	92	81	92	3
Dubai, VAE		✓			4	9.9	100	3.2	99	99	96	95	95	2
² Madrid, Spanien					4	9.9	100	6.5	100	100	97	97	97	0
Moskau, Russische Föderation					4	10.8	100	3.3	100	100	97	97	97	0
² ⁸ Norwegen (4. Jgst.) ^o		(✓)	(✓)	(✓)	4	9.8	100	5.1	95	99	96	91	95	1
⁸ Ontario, Kanada		✓	(✓)	(✓)	4	9.8	100	4.1	96	97	96	92	93	1
² ³ ⁸ Québec, Kanada		✓	(✓)	(✓)	4	10.1	100	5.1	39	67	96	37	64	0
¹ ³ ⁸ Südafrika (5. Jgst.; Englisch, Afrikaans, Zulu) ^p		✓	(✓)		5	11.6	100	1.1	84	89	96	81	86	9
PIRLS Literacy-Teilnehmer														
⁷ Ägypten					4	10.0	100	1.2	100	100	97	97	97	10
⁷ Kuwait					4	9.6	100	4.0	98	98	93	91	91	3
⁷ Südafrika (4. Jgst.)					4	10.6	100	2.5	92	97	96	88	94	8
Benchmark-Teilnehmer														
¹ ² ⁷ Dänemark (3. Jgst.)					3	9.8	100	9.3	88	97	95	83	92	0

Anmerkungen.

Die Tabellen A.1, A.2 und A.3 wurden aus dem Anhang des deutschen Berichtsbandes zu IGLU 2011 übernommen (Bos, Tarelli, Bremerich-Vos & Schwippert, 2012, S. 261–263) und jeweils um eine Spalte zur Kennzeichnung der Teilnahme an IGLU 2016 ergänzt.

Die Fußnotennummerierung mit Ziffern erfolgt konsistent mit der deutschen Berichterstattung zu IGLU 2011 (Bos et al., 2012), TIMSS 2011 (Bos, Wendt, Köller & Selter, 2012) sowie TIMSS 2015 (Wendt, Bos, Selter, Köller, Schwippert & Kasper, 2016). Die Fußnotenziffern 4, 5 und 6 werden in diesem Band zu IGLU 2016 nicht genutzt und entsprechend nicht erläutert.

- = Als Abweichung von der internationalen Vorgabe werden folgende Fälle gekennzeichnet: Getestete Jahrgangsstufe entspricht nicht der vierten Jahrgangsstufe (vgl. zugleich Fußnoten A, E, I und L); Ausschöpfungsgrad der nationalen Zielpopulation kleiner 100%; Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) größer 5%, aber kleiner oder gleich 10%; Schulteilnahmequote kleiner 85%, aber größer oder gleich 75%; Gesamteilnahmequote kleiner 75%, aber größer oder gleich 60%; Anteil der Schülerinnen und Schüler ohne gültige Leistungswerte größer 15%, aber kleiner oder gleich 25%.
- = Als starke Abweichung von der internationalen Vorgabe werden folgende Fälle gekennzeichnet: Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) größer 10%; Schulteilnahmequote kleiner 75%; Gesamteilnahmequote kleiner 60%; Anteil der Schülerinnen und Schüler ohne gültige Leistungswerte größer 25%.
- 1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
- 2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
- 3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
- 4 = Sehr hoher Anteil an Schülerinnen und Schülern mit nicht skalierbaren Leistungswerten.
- 7 = Die beiden Teilnehmer Iran und Marokko haben sowohl an IGLU 2016 als auch an PIRLS *Literacy* teilgenommen. Die zugehörigen Leistungswerte im Lesen basieren in Anlehnung an die internationale Berichterstattung auf einer Zusammenführung der Ergebnisse beider Studienkomponenten (Mullis, Martin, Foy & Hooper, 2017). Die Teilnehmer Ägypten, Kuwait und Südafrika (4. Jahrgangsstufe) sowie der Benchmark-Teilnehmer Dänemark (3. Jahrgangsstufe) haben sich ausschließlich an PIRLS *Literacy* beteiligt (ebd.). Ägypten, Kuwait und Südafrika sind deshalb nicht Teil dieser Berichterstattung zu IGLU 2016. Dänemark hat sich neben der regulären Teilnahme mit der 4. Jahrgangsstufe zusätzlich als Benchmark-Teilnehmer mit der 3. Jahrgangsstufe und hierbei ausschließlich an PIRLS *Literacy* beteiligt; um diesen Kohortenvergleich im Rahmen des IGLU 2016-Berichtes berücksichtigen zu können, wird Dänemark auf der Gesamtskala Lesen (und nur dort) zusätzlich mit der 3. Jahrgangsstufe berichtet.
- 8 = Bei Trendvergleichen werden nur die Teilnehmer berücksichtigt, die zusätzlich zu IGLU 2016 an mindestens zwei weiteren IGLU-Zyklen teilgenommen haben und zugleich Mitglied der EU und/oder OECD sind und/oder auf der Gesamtskala Lesen bei IGLU 2016 signifikant besser als oder nicht signifikant verschieden von Deutschland abgeschnitten haben. Somit werden im Trendvergleich folgende Teilnehmer nicht berücksichtigt: Georgien, Iran, Katar, Marokko, Trinidad und Tobago sowie Südafrika (5. Jahrgangsstufe). Zudem werden Israel und Polen aufgrund von Änderungen in den Übersetzungen der Testinstrumente oder in den Erhebungsbedingungen im Trendvergleich nicht berücksichtigt. Belgien (Flämische Gemeinschaft) wird berücksichtigt, da es für diesen Berichtsband aus deutscher Perspektive relevante Trendvergleiche erlaubt. Darüber hinaus ist bei Trendvergleichen zu beachten, dass in Litauen die Ergebnisse für 2016 keine Schülerinnen und Schüler berücksichtigen, die in Polnisch oder Russisch unterrichtet wurden. Weiterhin sollten bei folgenden Teilnehmern und Benchmark-Teilnehmern Änderungen in den Ausschluss- und Teilnahmequoten über die drei Studienzyklen bei Trendvergleichen berücksichtigt werden: Belgien (Flämische Gemeinschaft), Belgien (Französische Gemeinschaft), England, Hongkong, Italien, Litauen, Russische Föderation, Singapur, Slowenien, USA sowie Norwegen (4. Jahrgangsstufe), Ontario (Kanada) und Québec (Kanada) (Mullis, Martin, Foy & Hooper, 2017).
- * = Durchschnittsalter berechnet als: Anzahl der Jahre + Anzahl der Monate/12 (z. B. 9 Jahre, 3 Monate = $9 + 3/12 = 9.25$ Jahre).
- ** = Ausschöpfungsgrad der nationalen Zielpopulation in Prozent bezogen auf die internationale Vorgabe (100%).
- *** = Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) in Prozent.
- † Klammern werden gesetzt, wenn die Teilnahme mit der entsprechenden Jahrgangsstufe im betreffenden Zyklus nicht als Benchmark-Teilnehmer sondern als (regulärer) Teilnehmer oder umgekehrt erfolgte. Im Fall von Kuwait kennzeichnen die Klammern die Teilnahme an IGLU mit Jahrgangsstufe 6 in 2011 im Vergleich zur Jahrgangsstufe 4 in 2001 und 2006.
- ‡ Erläuterungen von Trendentwicklungen im Ausschöpfungsgrad oder der Ausschlussquote in Georgien, Hongkong, Kanada, Österreich und Singapur werden in Appendix C.5 (Mullis, Martin, Foy & Hooper, 2017) aufgeführt.
- A = In England, Neuseeland, Schottland und Slowenien wurde gemäß der internationalen Vorgabe (Campbell et al., 2001, S. 7) ein anderer Jahrgang als Stufe 4 als äquivalente nationale Zielpopulation interpretiert (vgl. Mullis, Martin, Gonzalez & Kennedy, 2003, S. 38).
- B = Die Ergebnisse von Israel wurden nicht in die Berichterstattung zu IGLU 2001 mit aufgenommen, da dort die angestrebte nationale Zielpopulation unterhalb von 80 Prozent der Gesamtpopulation lag (vgl. Bos et al., 2003, S. 11).
- C = Vor der Transformation des Schulsystems im Jahr 2001 existierten in der Russischen Föderation zwei Schulsysteme parallel („Grundschule Klassen 1 bis 3“ und „Grundschule Klassen 1 bis 4“), in denen entweder die dritte oder die vierte Jahrgangsstufe der PIRLS-Zielpopulation entsprach. Für die Stichprobenziehung wurde nach beiden Systemen stratifiziert, so dass die Vergleichbarkeit mit den Stichproben der anderen Teilnehmerländer gegeben ist (vgl. Schwippert & Lenkeit, 2012, S. 190–191).
- D = Kanada wird 2001 ausschließlich durch die Provinzen Ontario und Québec repräsentiert, die teilweise für die Berichterstattung als ein Land (Kanada) berichtet werden (vgl. Martin, Mullis & Kennedy, 2003, S. 115–117).
- E = Da das Durchschnittsalter der Viertklässlerinnen und Viertklässler in England, Neuseeland, Schottland sowie Trinidad und Tobago bedeutsam unter der internationalen Vorgabe (9.5 Jahre) lag, wurde gemäß der Kombination der Kriterien ‚formale Beschulungszeit‘ und ‚Durchschnittsalter des Schülerjahrgangs‘ eine höhere Jahrgangsstufe als äquivalente nationale Zielpopulation interpretiert (vgl. Mullis, Martin, Kennedy & Foy, 2007, S. 37–38).

- F = Die Ergebnisse von Israel wurden nicht in die Berichterstattung zu IGLU 2006 mit aufgenommen, da dort die angestrebte nationale Zielpopulation unterhalb von 80 Prozent der Gesamtpopulation lag (vgl. Bos et al., 2007, S. 40).
- G = Luxemburg und Südafrika wählten aufgrund von Herausforderungen bezüglich der Testsprache für die Erhebung eine höhere Jahrgangsstufe (vgl. Mullis, Martin, Kennedy & Foy, 2007, S. 38).
- H = Infolge der Transformation des Schulsystems existierten in Slowenien zum Zeitpunkt der Testung mehrere Schulsysteme parallel, in denen entweder die dritte oder die vierte Jahrgangsstufe der PIRLS-Zielpopulation entsprach. Für die Stichprobenziehung wurde nach beiden Systemen stratifiziert, so dass eine Vergleichbarkeit mit den Stichproben der anderen Teilnehmerländer gegeben ist (vgl. Kennedy, Mullis, Martin & Trong, 2007, S. 374).
- I = Da das Durchschnittsalter der Viertklässlerinnen und Viertklässler in England, Malta, Neuseeland sowie Trinidad und Tobago bedeutsam unter der internationalen Vorgabe (9.5 Jahre) lag, wurde gemäß der Kombination der Kriterien ‚formale Beschulungszeit‘ und ‚Durchschnittsalter des Schülerjahrgangs‘ eine höhere Jahrgangsstufe als äquivalente nationale Zielpopulation interpretiert (vgl. Mullis, Martin, Foy & Drucker, 2012).
- J = Die Ergebnisse von Israel wurden nicht in die Berichterstattung zu IGLU 2011 mit aufgenommen, da dort die angestrebte nationale Zielpopulation unterhalb von 80 Prozent der Gesamtpopulation lag.
- K = Südafrika testete ausschließlich Schülerinnen und Schüler der Jahrgangsstufe 5, für die Englisch oder Afrikaans als Unterrichtssprache vorgesehen ist (vgl. Mullis, Martin, Foy & Drucker, 2012).
- L = Da das Durchschnittsalter der Viertklässlerinnen und Viertklässler in England, Malta, Neuseeland sowie Trinidad und Tobago bedeutsam unter der internationalen Vorgabe (9.5 Jahre) lag, wurde gemäß der Kombination der Kriterien ‚formale Beschulungszeit‘ und ‚Durchschnittsalter des Schülerjahrgangs‘ eine höhere Jahrgangsstufe als äquivalente nationale Zielpopulation interpretiert (Mullis, Martin, Foy & Hooper, 2017).
- M = Hinweise zum geringen Ausschöpfungsgrad: In Georgien wurden nur in Georgisch und Aserbaidschanisch unterrichtete Schülerinnen und Schüler getestet. In Kanada wurden nur Schülerinnen und Schüler aus den Provinzen Alberta, Britisch Kolumbien, Manitoba, Neubraunschweig, Neufundland, Ontario, Québec und Saskatchewan getestet (Mullis, Martin, Foy & Hooper, 2017).
- N = Für Iran und Marokko werden die Teilnahmequoten sowie die Anteile von Schülerinnen und Schülern ohne gültige Leistungswerte für die Zusammenführung von IGLU 2016 und PIRLS *Literacy* berichtet (Mullis, Martin, Foy & Hooper, 2017).
- O = Norwegen entschied sich der internationalen Berichterstattung zu PIRLS 2016 zufolge (Mullis, Martin, Foy & Hooper, 2017) zur Teilnahme mit der Jahrgangsstufe 5, um bessere Vergleiche mit Schweden und Finnland zu erreichen und zugleich zur Teilnahme als Benchmark-Teilnehmer mit Jahrgangsstufe 4, um die Trenderhebung für diese Jahrgangsstufe fortzusetzen.
- P = Südafrika testete ausschließlich Schülerinnen und Schüler der Jahrgangsstufe 5, für die Englisch, Afrikaans oder Zulu als Unterrichtssprache vorgesehen ist (Mullis, Martin, Foy & Hooper, 2017).

Literatur

- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Valtin, R., Hornberg, S., Buddeberg, I., Goy, M. & Voss, A. (2007). Internationaler Vergleich 2006: Lesekompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 109–160). Münster: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.). (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.). (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Campbell, J. R., Kelly, D. L., Mullis, I. V. S., Martin, M. O. & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001* (2. Aufl.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Kennedy, A. M., Mullis, I. V. S., Martin, M. O. & Trong, K. L. (Hrsg.). (2007). *PIRLS 2006 encyclopedia: A guide to reading education in the forty PIRLS 2006 countries*. Chestnut Hill, MA: Boston College.
- Martin, M. O., Mullis, I. V. S. & Kennedy, A. M. (Hrsg.). (2003). *PIRLS 2001 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Zugriff am 05.12.2017: <http://timssandpirls.bc.edu/pirls2016/international-results/>
- Mullis, I. V. S., Martin, M. O., Foy, P. & Drucker, K. T. (2012). *PIRLS 2011 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., Gonzalez, E. J. & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEA'S study of reading literacy achievement in primary schools*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center Boston College.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M. & Foy, P. (2007). *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary school in 40 countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Schwippert, K. & Lenkeit, J. (Hrsg.). (2012). *Progress in reading literacy in national and international context. The impact of PIRLS 2006 in 12 countries*. Münster: Waxmann.
- Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K. & Kasper, D. (Hrsg.). (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann

Anhang B

Tabelle B.1: Veränderungen der Leseleistungen zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle von Schülermerkmalen (Mehrebenenmodelle)

	Modell 1 <i>b</i> (SE)	Modell 2 <i>b</i> (SE)	Modell 3 <i>b</i> (SE)	Modell 4 <i>b</i> (SE)	Modell 5 <i>b</i> (SE)
Konstante	533.5 (4.4)**	537.9 (4.7)**	524.5 (4.8)**	524.8 (4.7)**	525.0 (5.4)**
Testzeitpunkt					
IGLU 2001	5.7 (5.0) ^{ns}	5.6 (5.0) ^{ns}	1.9 (4.6) ^{ns}	4.5 (3.6) ^{ns}	5.2 (5.2) ^{ns}
IGLU 2006	9.4 (5.6)*	9.4 (5.5)*	7.6 (4.7) ^{ns}	8.6 (3.5)**	7.9 (5.0) ^{ns}
IGLU 2011	5.4 (5.3) ^{ns}	5.3 (5.3) ^{ns}	-0.5 (5.0) ^{ns}	-3.5 (3.9) ^{ns}	-4.2 (5.3) ^{ns}
Merkmale					
Geschlecht ^A		-8.5 (2.4)**	-0.4 (2.3) ^{ns}	2.6 (2.3) ^{ns}	2.2 (4.9) ^{ns}
Familie mit hohem Bildungsniveau ^B				21.2 (2.5)**	21.2 (2.5)**
Familie mit hohem Berufsstatus ^C					
Migrationshintergrund (ein Elternteil) ^D					
Migrationshintergrund (beide Elternteile) ^E					
Interaktionen					
Geschlecht * IGLU 2001					-1.4 (6.4) ^{ns}
Geschlecht * IGLU 2006					1.4 (6.2) ^{ns}
Geschlecht * IGLU 2011					1.4 (6.6) ^{ns}
Hoher Berufsstatus ^C * IGLU 2001					
Hoher Berufsstatus ^C * IGLU 2006					
Hoher Berufsstatus ^C * IGLU 2011					
Migrationshintergrund (ein Elternteil) ^D * IGLU 2001					
Migrationshintergrund (beide Elternteile) ^E * IGLU 2001					
Migrationshintergrund (ein Elternteil) ^D * IGLU 2006					
Migrationshintergrund (beide Elternteile) ^E * IGLU 2006					
Migrationshintergrund (ein Elternteil) ^D * IGLU 2011					
Migrationshintergrund (beide Elternteile) ^E * IGLU 2011					
Kovariaten					
Leseselbstkonzept ^F			27.8 (1.5)**	21.4 (1.6)**	21.4 (1.6)**
Lesemotivation ^G			5.5 (1.7)**	4.7 (2.1)*	4.6 (2.1)*
Leseverhalten ^H			25.5 (3.7)**	21.6 (4.0)**	21.6 (4.0)**
Kognitive Fähigkeiten ^I				16.0 (0.9)**	16.0 (0.9)**
Varianzanteile					
Erklärter Anteil der Gesamtvarianz (29.7%)					
Erklärter Anteil der Gesamtvarianz (70.3%)	0.2	0.8	29.6	46.5	46.6

b = Regressionsgewichte (unstandardisiert).

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p < .05$); ** = signifikant ($p < .01$).

A = Geschlecht (0 = Mädchen; 1 = Jungen), siehe Kapitel 5 in diesem Band.

B = Bildungsniveau nach höchstem Bildungsabschluss der Eltern (0 = Kein Elternteil mit (Fach-)Hochschulabschluss;

1 = Mindestens ein Elternteil mit (Fach-)Hochschulabschluss), siehe Kapitel 6 in diesem Band.

C = Berufsstatus nach Angabe der Eltern (0 = Manuell Tätige, Angestellte und kleinere Unternehmer;

1 = Akademiker, Techniker und Führungskräfte), siehe Kapitel 6 in diesem Band. In Kapitel 3 als „Sozioökonomischer Status“ bezeichnet, da als exemplarische Variable stellvertretend für den sozioökonomischen Status aufgenommen.

Modell 6	Modell 7	Modell 8	Modell 9	Modell 10	Modell 11	Modell 12	Modell 13
<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)
528.8 (5.4)**	527.9 (5.3)**	545.1 (4.6)**	544.6 (5.1)**	554.2 (4.3)**	547.5 (5.1)**	551.8 (4.9)**	550.8 (5.7)**
9.3 (6.0) ^{ns}	7.5 (5.8) ^{ns}	2.5 (4.6) ^{ns}	2.3 (5.6) ^{ns}	-4.1 (4.9) ^{ns}	-2.3 (5.3) ^{ns}	-0.6 (4.7) ^{ns}	1.8 (5.9) ^{ns}
15.0 (5.6)**	14.7 (5.2)**	8.5 (4.4)*	10.5 (5.3)*	5.6 (4.9) ^{ns}	6.6 (5.0) ^{ns}	4.9 (4.4) ^{ns}	7.5 (5.7) ^{ns}
6.0 (5.8) ^{ns}	5.2 (5.7) ^{ns}	-2.5 (4.7) ^{ns}	-2.4 (5.5) ^{ns}	-1.4 (5.2) ^{ns}	-1.1 (5.3) ^{ns}	-4.1 (4.9) ^{ns}	-5.9 (5.9) ^{ns}
						-5.7 (2.4)**	-5.7 (2.4)**
38.7 (3.0)**	25.4 (3.6)**	18.2 (3.3)**	18.1 (3.3)**				
	24.3 (3.3)**	15.2 (3.1)**	17.0 (6.5)**		31.7 (2.7)**	22.4 (2.5)**	26.9 (6.5)**
		-12.7 (4.0)**	-12.7 (4.0)**	-19.6 (3.7)**	-16.9 (4.1)**	-12.0 (3.7)**	-13.7 (9.2) ^{ns}
		-26.5 (3.8)**	-26.6 (3.8)**	-40.7 (3.6)**	-32.6 (4.1)**	-27.2 (3.8)**	-27.6 (9.1)**
			0.5 (8.1) ^{ns}				-6.0 (7.8) ^{ns}
			-6.4 (8.1) ^{ns}				-9.8 (8.3) ^{ns}
			-0.5 (7.5) ^{ns}				0.4 (7.4) ^{ns}
							-0.1 (12.2) ^{ns}
							3.1 (11.4) ^{ns}
							3.2 (12.8) ^{ns}
							-6.7 (11.3) ^{ns}
							0.2 (11.8) ^{ns}
							8.9 (11.8) ^{ns}
		17.3 (1.0)**	17.3 (1.0)**			17.7 (1.0)**	17.7 (0.9)**
15.9	23.6	44.7	44.7	18.9	30.2	44.0	44.1

D = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren;

1 = Ein Elternteil im Ausland geboren), siehe Kapitel 7 in diesem Band.

E = Migrationshintergrund nach Geburtsland der Eltern (0 = Kein Elternteil im Ausland geboren;

1 = Beide Elternteile im Ausland geboren), siehe Kapitel 7 in diesem Band.

F = Skalierter Index des Leseselbstkonzepts (Messvarianz über die Zyklen angenommen), siehe Kapitel 4 in diesem Band.

G = Skalierter Index der Lesemotivation (Messvarianz über die Zyklen angenommen), siehe Kapitel 4 in diesem Band.

H = Lesen zum Vergnügen (0 = Nie oder fast nie; 1 = Jeden Tag oder fast jeden Tag), siehe Kapitel 4 in diesem Band.

I = Skalierter Index der kognitiven Fähigkeiten (Messvarianz über die Zyklen angenommen), siehe Kapitel 2 in diesem Band.

Abbildungsverzeichnis

Abbildung 2.1:	Rahmenmodell zur Betrachtung des Zusammenhangs von Schülerleistungen und deren Bedingungen.....	38
Abbildung 2.2:	Staaten und Regionen, die an IGLU 2016 teilgenommen haben.....	40
Abbildung 2.3:	Teilnehmer der Trendvergleiche zu IGLU 2001, 2006, 2011 und 2016	42
Abbildung 2.4:	Veränderung des Mittelwerts von Viertklässlerinnen und Viertklässlern in Deutschland beim Vergleich der Leistung auf der Gesamtskala Lesen in Abhängigkeit von der Ausschlussquote	50
Abbildung 2.5:	Normalverteilung mit Perzentilen	68
Abbildung 2.6:	Beziehung von Benchmarks und Kompetenzstufen.....	69
Abbildung 2.7:	Darstellung von Perzentilbändern mit Konfidenzintervallen am Beispiel der Leseleistung in IGLU 2016	70
Abbildung 3.1:	Die theoretische Struktur der Lesekompetenz in IGLU.....	84
Abbildung 3.2:	Aufgaben zum Erzähltext <i>Marie und das rote Huhn</i>	100
Abbildung 3.3:	Kompetenzstufen und Beispielaufgaben (literarischer Text).....	102
Abbildung 3.4:	Aufgaben zum Sachtext <i>Die Grüne Meeresschildkröte und die Reise ihres Lebens</i>	108
Abbildung 3.5:	Kompetenzstufen und Beispielaufgaben (Sachtext).....	109
Abbildung 3.6:	Testleistungen der Schülerinnen und Schüler im internationalen Vergleich – Gesamtskala Lesen.....	114
Abbildung 3.7:	Relative Positionierung zu Deutschland und Vergleich der Testleistungen zwischen IGLU 2001, 2006, 2011 und 2016 – Gesamtskala Lesen.....	116
Abbildung 3.8:	Prozentuale Verteilung der Schülerinnen und Schüler auf die fünf Kompetenzstufen (Lesen) im internationalen Vergleich	126
Abbildung 3.9:	Differenz zwischen den Testleistungen nach Textsorten im internationalen Vergleich.....	131
Abbildung 3.10:	Veränderung der Differenz zwischen den Testleistungen nach Textsorten in IGLU 2001, 2006, 2011 und 2016 im internationalen Vergleich	132
Abbildung 3.11:	Differenzen zwischen den Testleistungen nach Verstehensprozessen im internationalen Vergleich	135
Abbildung 3.12:	Veränderung der Differenz zwischen den Testleistungen nach Textsorten in IGLU 2001, 2006, 2011 und 2016 im internationalen Vergleich	136
Abbildung 4.1:	Leseselbstkonzept von Schülerinnen und Schülern sowie prozentuale Verteilung nach niedrigem, mittlerem und hohem Selbstkonzept in Deutschland – IGLU 2011 und 2016 im Vergleich.....	149
Abbildung 4.2:	Lesemotivation von Schülerinnen und Schülern sowie prozentuale Verteilungen nach niedrigerer, mittlerer und hoher Motivation in Deutschland – IGLU 2001, 2006, 2011 und 2016 im Vergleich.....	151
Abbildung 4.3:	Lesen zum Vergnügen im internationalen Vergleich – IGLU 2016 (Anteile der Schülerinnen und Schüler in Prozent)	153
Abbildung 4.4:	Schülerinnen und Schüler, die angeben, außerhalb der Schule nie oder fast nie zum Vergnügen zu lesen, im internationalen Vergleich – IGLU 2001, 2006, 2011 und 2016 im Vergleich.....	154
Abbildung 4.5:	Schülerinnen und Schüler, die angeben, außerhalb der Schule jeden Tag oder fast jeden Tag zum Vergnügen zu lesen, im internationalen Vergleich – IGLU 2001, 2006, 2011 und 2016.....	155
Abbildung 4.6:	Lesedauer (Freizeit, täglich) im internationalen Vergleich – IGLU 2016 (Anteile der Schülerinnen und Schüler in Prozent).....	157
Abbildung 4.7:	Schülerinnen und Schüler, die angeben, täglich weniger als 30 Minuten außerhalb der Schule zu lesen, im internationalen Vergleich – IGLU 2011 und 2016 im Vergleich.....	158
Abbildung 4.8:	Schülerinnen und Schüler, die angeben, täglich 2 Stunden oder mehr außerhalb der Schule zu lesen, im internationalen Vergleich – IGLU 2011 und 2016 im Vergleich.....	159

Abbildung 4.9:	Häufigkeit der Buchausleihe aus der Schulbibliothek oder örtlichen Bibliothek im internationalen Vergleich – IGLU 2016 (Anteile der Schülerinnen und Schüler in Prozent)	162
Abbildung 4.10:	Schülerinnen und Schüler, die angeben, nie oder fast nie Bücher aus der Schulbibliothek oder örtlichen Bibliothek auszuleihen, im internationalen Vergleich – IGLU 2006, 2011 und 2016 im Vergleich	164
Abbildung 4.11:	Schülerinnen und Schüler, die angeben, mindestens einmal in der Woche Bücher aus der Schulbibliothek oder örtlichen Bibliothek auszuleihen, im internationalen Vergleich – IGLU 2006, 2011 und 2016 im Vergleich	165
Abbildung 5.1:	Leseleistung nach Geschlecht – Gesamtskala Lesen	184
Abbildung 5.2:	Leistungsvorsprung der Mädchen differenziert nach Textsorten	185
Abbildung 5.3:	Vergleich des Leistungsvorsprungs der Mädchen zwischen IGLU 2001, 2006, 2011 und 2016 – Gesamtskala Lesen	187
Abbildung 5.4:	Prozentuale Verteilung von Mädchen und Jungen in Deutschland auf die fünf Lesekompetenzstufen	188
Abbildung 5.5:	Mittelwerte und Differenzen von Leseselbstkonzept, Lesemotivation und Leseverhalten nach Geschlecht in Deutschland	189
Abbildung 6.1:	Leistungsvorsprung im Leseverständnis von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern.....	202
Abbildung 6.2:	Unterschiede im Leistungsvorsprung im Leseverständnis von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern im Vergleich von IGLU 2001, 2006, 2011 und 2016	204
Abbildung 6.3:	Leistungsvorsprung im Leseverständnis von Kindern aus Familien der dritten Berufsgruppe (Akademiker, Techniker und Führungskräfte) vor denen der ersten Berufsgruppe (manuelle Tätigkeiten).....	205
Abbildung 6.4:	Leistungsvorsprung im Leseverständnis von Kindern aus Familien, in denen mindestens ein Elternteil einen tertiären Bildungsabschluss erreicht hat (ISCED-Level 6, 7 und 8) vor Kindern aus Familien, in denen dies nicht der Fall ist	206
Abbildung 6.5:	Zusammenhang zwischen der erklärten Varianz der Lesekompetenz auf Individualebene durch die Indikatoren der sozialen Stratifikation (ISCED, Berufsstatus, Bücher im Haushalt) und den mittleren Lesekompetenzen in den Teilnehmerstaaten	208
Abbildung 6.6:	Mittlere Lesekompetenz pro EGP-Klasse in Deutschland bei IGLU 2001, 2006, 2011 und 2016	210
Abbildung 6.7:	Armutsgefährdungsquoten nach Migrationshintergrund, höchstem Bildungsabschluss und EGP-Klasse (in Prozent).....	211
Abbildung 7.1:	Lesekompetenzen von Schülerinnen und Schülern nach zu Hause gesprochener Sprache im internationalen Vergleich	223
Abbildung 7.2:	Lesekompetenzen von Schülerinnen und Schülern in Deutschland nach im Ausland geborenen Elternteilen – IGLU 2001, 2006, 2011 und 2016	225
Abbildung 7.3:	Anteile von Kindern auf den unterschiedlichen Kompetenzstufen in Abhängigkeit von im Ausland geborenen Elternteilen – IGLU 2001, 2006, 2011 und 2016 im Vergleich (Angaben in Prozent)	229
Abbildung 7.4:	Mittelwerte und Differenzen von Leseselbstkonzept, Lesemotivation und Leseverhalten in Abhängigkeit von im Ausland geborenen Elternteilen in Deutschland	231
Abbildung 8.1:	Zusammenhang der Schulnoten mit der Schullaufbahnpräferenz der Lehrkräfte (standardisierte Regressionskoeffizienten)	242
Abbildung 8.2:	Testleistungen differenziert nach Deutschnoten – Gesamtskala Lesen	242
Abbildung 8.3:	Testleistungen differenziert nach Schullaufbahnpräferenzen der Lehrkräfte – Gesamtskala Lesen.....	243
Abbildung 8.4:	Gruppenspezifische Standards („kritische Werte“) für eine Gymnasialpräferenz der Lehrkräfte und der Eltern – Gesamtskala Lesen 2001, 2006, 2011 und 2016	247
Abbildung 9.1:	Unterrichtszeit für die Testsprache als Fach beziehungsweise für Sprachunterricht in der Testsprache im internationalen Vergleich (Angaben der Lehrkräfte in Minuten pro Woche).....	254

Abbildung 9.2:	Die drei Basisdimensionen der Unterrichtsqualität (Darstellung in Anlehnung an Klieme et al., 2006, S. 131 und Klieme & Rakoczy, 2008, S. 228).....	256
Abbildung 9.3:	Durch Schülerinnen und Schüler wahrgenommene Unterrichtsqualität im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	265
Abbildung 9.4:	Durch Lehrkräfte wahrgenommene Unterrichtsqualität im Deutschunterricht (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent).....	266
Abbildung 9.5:	Leseleistungen nach wahrgenommener Klassenführung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	267
Abbildung 9.6:	Leseleistungen nach wahrgenommener Strukturiertheit im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	268
Abbildung 9.7:	Leseleistungen nach wahrgenommener kognitiver Aktivierung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	269
Abbildung 9.8:	Leseleistungen nach wahrgenommenem Sozialklima im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	269
Abbildung 9.9:	Lesemotivation nach wahrgenommener Klassenführung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	270
Abbildung 9.10:	Lesemotivation nach wahrgenommener Strukturiertheit im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	271
Abbildung 9.11:	Lesemotivation nach wahrgenommener kognitiver Aktivierung im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	271
Abbildung 9.12:	Lesemotivation nach wahrgenommenem Sozialklima im Deutschunterricht (Anteile der Schülerinnen und Schüler in Prozent).....	272
Abbildung 10.1:	Förderbedarf und Förderangebot (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent).....	282
Abbildung 10.2:	Verfahren zum halblauten Lesen differenziert nach hohem und niedrigem Anteil an leseschwachen Schülerinnen und Schülern in der Klasse (Anteile der Schülerinnen und Schüler nach Angaben nach Lehrkräften in Prozent).....	289
Abbildung 11.1:	Mittlere Anteile von Schülerinnen und Schülern nach unterschiedlichen Diversitätsmerkmalen in Klassen der vierten Jahrgangsstufe an Regelschulen in Deutschland (Schulen nach Zusammensetzungsmerkmalen in Prozent).....	300
Abbildung 11.2:	Integration von neu zugewanderten Schülerinnen und Schülern aus Sicht der Schulleitungen an Schulen mit entsprechender Schülerschaft (Skalenmittelwert und Anteil der Schulen in %).....	303

Tabellenverzeichnis

Tabelle 2.1:	Vergleichsgruppen in IGLU 2016	41
Tabelle 2.2:	Länderspezifische Modalitäten der Teilnahme an IGLU 2001, 2006, 2011 und 2016	45
Tabelle 2.3:	Testheftdesign in IGLU 2016	54
Tabelle 2.4:	Untersuchungsablauf von IGLU 2016	59
Tabelle 2.5:	Beteiligungszahlen, Rücklaufquoten und zentrale Charakteristika zu IGLU 2001, 2006, 2011 und 2016	62
Tabelle 3.1:	Sach- und Erzähltexte in IGLU 2016	86
Tabelle 3.2:	Prozentuale Anteile der Verstehensprozesse bei IGLU 2016 in Prozent	87
Tabelle 3.3:	Kompetenzbereiche des Faches Deutsch der KMK-Bildungsstandards	89
Tabelle 3.4:	Lehrpläne für das Fach Deutsch in der Grundschule nach Ländern der Bundesrepublik Deutschland und Veröffentlichungszeiträumen vor und nach IGLU 2001, 2006, 2011 und 2016	90
Tabelle 3.5:	Kompetenzstufen und Skalenwerte IGLU 2016	92
Tabelle 3.6:	Mittlere Lösungshäufigkeiten in den Teilen des Leseverständnistests ELFE II nach IGLU-Kompetenzstufen	94
Tabelle 3.7:	Veränderung der Schülerleistungen zwischen IGLU 2001 und 2016 unter Kontrolle von Schülermerkmalen – Gesamtskala Lesen	120
Tabelle 3.8:	Lesekompetenzen nach Perzentilen im Vergleich: IGLU 2001, 2006, 2011 und 2016 – Gesamtskala Lesen	122
Tabelle 3.9:	Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen in IGLU 2001, 2006, 2011 und 2016 (Angaben in Prozent)	128
Tabelle 4.1:	Prädiktion der Zugehörigkeit zur Gruppe schwacher Leserinnen und Leser (Kompetenzstufen I und II) durch individuelle und familiäre Variablen – Ergebnisse logistischer Mehrebenenregressionen (<i>odds ratios</i>)	167
Tabelle 4.2:	Prädiktion der Zugehörigkeit zur Gruppe schwacher Leserinnen und Leser (Kompetenzstufen I und II) durch strukturelle Herkunftsvariablen und Geschlecht – Ergebnisse logistischer Mehrebenenregressionen (<i>odds ratios</i>)	168
Tabelle 5.1:	Lesekompetenz zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle des Geschlechts und weiterer Merkmale (Mehrebenenmodelle)	190
Tabelle 6.1:	Mittlere Lesekompetenz pro EGP-Klasse in Deutschland bei IGLU 2001, 2006, 2011 und 2016	209
Tabelle 6.2:	Mittlere Lesekompetenz nach Armutsgefährdung und EGP-Klasse	211
Tabelle 6.3:	Lesekompetenzen zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle von Indikatoren der sozialen Herkunft (Mehrebenenmodell)	212
Tabelle 7.1:	Lesekompetenzen zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle von Indikatoren zu Migration und sozialer Herkunft (Mehrebenenmodell)	227
Tabelle 8.1:	Übersicht über die Schulformen der Sekundarstufe I in den Ländern der Bundesrepublik Deutschland (Schuljahr 2015/16)	236
Tabelle 8.2:	Übersicht über die Empfehlungspraktiken in den Ländern der Bundesrepublik Deutschland (Schuljahr 2015/16)	238
Tabelle 8.3:	Schullaufbahnpräferenzen der Lehrkräfte und Eltern für Deutschland 2001, 2006, 2011 und 2016 in Zeilenprozent	240
Tabelle 8.4:	Übereinstimmungen der Schullaufbahnpräferenzen von Lehrkräften und Eltern in Zeilenprozent	241
Tabelle 8.5:	Leistungsrelevante Merkmale von Schülerinnen und Schülern und Schullaufbahnpräferenz der Lehrkräfte (ohne und mit Berücksichtigung der Lesekompetenz beziehungsweise der Deutschnote)	243
Tabelle 8.6:	Relative Chancen (<i>odds ratios</i>) für eine Gymnasialpräferenz der Lehrkräfte beziehungsweise der Eltern nach sozialer Lage (EGP) der Familien	244

Tabelle 8.7:	Relative Chancen (odds ratios) für eine Gymnasialpräferenz der Lehrkräfte für Kinder aus der <i>service class</i> (EGP I und II) im Vergleich mit Kindern aus der <i>working class</i> (EGP V, VI und VII) bei IGLU 2001, 2006, 2011 und 2016	245
Tabelle 8.8:	Gruppenspezifische Standards („kritische Werte“) für eine Gymnasialpräferenz der Lehrkräfte und der Eltern – Gesamtskala Lesen.....	246
Tabelle 9.1:	Bivariate Korrelationen zwischen der Wahrnehmung der Unterrichtsqualität und den Testleistungen im Lesen und der Lesemotivation	273
Tabelle 10.1:	Quellen zur Beurteilung von Fortschritten im Lesen der Schülerinnen und Schüler (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent)	283
Tabelle 10.2:	Verfahren zum halblauten Lesen im Unterricht (Anteile der Schülerinnen und Schüler nach Angabe der Lehrkräfte in Prozent).....	288
Tabelle 10.3:	Entwicklung von Lesefertigkeiten im Unterricht (Anteile der Schülerinnen und Schüler nach Angaben der Lehrkräfte in Prozent).....	293
Tabelle A.1:	Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2001	329
Tabelle A.2:	Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2006	330
Tabelle A.3:	Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2011	331
Tabelle A.4:	Weiterführende Informationen zu den Teilnehmerstaaten – IGLU/PIRLS 2016	332
Tabelle B.1:	Veränderungen der Leseleistungen zwischen IGLU 2001, 2006, 2011 und 2016 unter Kontrolle von Schülermerkmalen (Mehrebenenmodelle).....	336