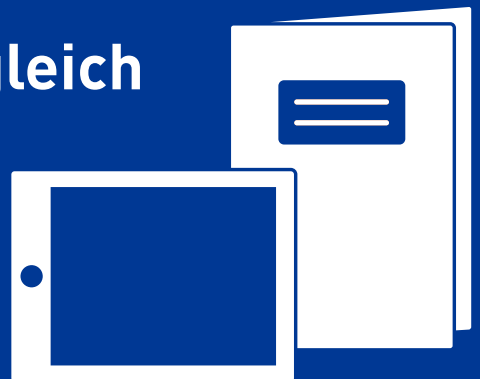


The logo features the text 'PISA' in a large, white, sans-serif font, with '2018' in a smaller font below it. Both are centered within a large, orange speech bubble that has a tail pointing towards the bottom left. The entire graphic is set against a dark blue background.

PISA 2018

Kristina Reiss, Mirjam Weis,
Eckhard Klieme, Olaf Köller (Hrsg.)

Grundbildung im internationalen Vergleich



WAXMANN

Kristina Reiss, Mirjam Weis,
Eckhard Klieme, Olaf Köller (Hrsg.)

PISA 2018

Grundbildung im internationalen Vergleich



Waxmann 2019
Münster · New York

Redaktionsschluss: 05.11.2019

Bibliografische Informationen der Deutschen Nationalbibliothek

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

ISBN 978-3-8309-4100-2

E-Book ISBN 978-8309-9100-7

doi: <https://doi.org/10.31244/9783830991007>

© Waxmann Verlag GmbH, Münster 2019

Steinfurter Straße 555, 48159 Münster

www.waxmann.com

info@waxmann.com

Buchumschlag: Alexandra Goßner, München

Satz: Roger Stoddart, Münster

Druck: mediaprint solutions GmbH, 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
1 PISA 2018 – Ziele und Inhalte der Studie	13
<i>Mirjam Weis & Kristina Reiss</i>	
1.1 Die PISA-Studien und das Bildungssystem	13
1.2 Inhalte der PISA-Studie 2018	15
1.3 Durchführung der PISA-Studie 2018 in Deutschland	16
1.3.1 Stichprobe	16
1.3.2 Prozeduren	18
1.4 Der Rahmen der PISA-Studie 2018: die Balance zwischen den PISA-Zyklen	18
Literatur	19
2 Lesekompetenz heute – eine Schlüsselqualifikation im Wandel	21
<i>Michael Becker-Mrotzek, Thomas Lindauer, Maximilian Pfof, Mirjam Weis, Anselm Strohmaier & Kristina Reiss</i>	
2.1 Lesekompetenz	23
2.1.1 Bedeutung der Lesekompetenz.....	23
2.1.2 Digitaler Wandel.....	23
2.1.3 Stand der Forschung zur Lesekompetenz	26
2.1.4 Die Rahmenkonzeption der Lesekompetenz in PISA 2018.....	28
2.2 Lesekompetenz testen – PISA 2018	33
2.2.1 Situationen des Lesens und Texttypen.....	33
2.2.2 Leseaufgaben und Antwortformate	35
2.3 Fazit	43
Literatur	44
3 Lesekompetenz in PISA 2018 – Ergebnisse in einer digitalen Welt.....	47
<i>Mirjam Weis, Anastasia Doroganova, Carolin Hahnel, Michael Becker-Mrotzek, Thomas Lindauer, Cordula Artelt & Kristina Reiss</i>	
3.1 Lesekompetenz in PISA 2018	47
3.1.1 Der Lesekompetenztest in PISA 2018 und Veränderungen seit 2009.....	49
3.1.2 Lesen in digitalen Medien	50
3.1.3 Prozesse des Lesens und Textquellen.....	51
3.1.4 Kompetenzstufen	51
3.1.5 Fragestellungen aus Sicht des Bildungsmonitorings.....	55
3.2 Lesekompetenz im internationalen Vergleich	56
3.2.1 Gesamtskala Lesekompetenz im internationalen Vergleich	56
3.2.2 Verteilung auf die Stufen der Lesekompetenz	60
3.2.3 Unterschiede in der Lesekompetenz zwischen Mädchen und Jungen	62

3.2.4	Teilskalen der Lesekompetenz im internationalen Vergleich	63
3.3	Vertiefende Analysen der Lesekompetenz in Deutschland	67
3.3.1	Unterschiede zwischen Schularten	68
3.3.2	Unterschiede zwischen Mädchen und Jungen in der Lesekompetenz in Deutschland.....	71
3.3.3	Veränderung der Lesekompetenz im Vergleich zu PISA 2009 und PISA 2015	72
3.4	Zusammenfassung und Diskussion	76
	Literatur	78

4 Lesabezogene Schülermerkmale in PISA 2018: Motivation, Leseverhalten, Selbstkonzept und Lesestrategiewissen..... 81

*Jennifer Diedrich, Anja Schiepe-Tiska, Lisa Ziernwald,
Ana Tupac-Yupanqui, Mirjam Weis, Nele McElvany & Kristina Reiss*

4.1	Lesemotivation, Leseverhalten und Selbstkonzept	83
4.1.1	Erfassung von Lesemotivation, Leseverhalten und Selbstkonzept in PISA 2018.....	84
4.1.2	Ergebnisse des internationalen Vergleichs	85
4.1.3	Unterschiede zwischen Schularten in Deutschland.....	94
4.1.4	Veränderungen zwischen den Erhebungen 2009 und 2018	96
4.2	Lesestrategiewissen	98
4.2.1	Erfassung des Lesestrategiewissens in PISA 2018.....	98
4.2.2	Lesestrategiewissen im internationalen Vergleich	101
4.2.3	Geschlechterunterschiede im Lesestrategiewissen	103
4.2.4	Unterschiede zwischen Schularten in Deutschland.....	103
4.2.5	Veränderungen zwischen den Erhebungen 2009 und 2018	104
4.3	Zusammenfassung und Diskussion	105
	Literatur	107

5 Schulische Lerngelegenheiten zur Sprach- und Leseförderung im Kontext der Digitalisierung 111

*Sarah Hofer, Doris Holzberger, Jörg-Henrik Heine, Frank Reinhold,
Anja Schiepe-Tiska, Mirjam Weis & Kristina Reiss*

5.1	Rahmenbedingungen für das Lesen digitaler Texte	113
5.1.1	Materielle und personelle ICT-Ressourcen.....	114
5.1.2	Integration digitaler Medien in den Schulalltag.....	117
5.1.3	Einsatz digitaler Ressourcen im Unterricht der Landessprache	119
5.2	Extracurriculare Angebote zur Sprach- und Leseförderung.....	121
5.2.1	Außerunterrichtliche Zusatzangebote zur Sprach- und Leseförderung	121
5.2.2	Zusätzlicher Unterricht der Landessprache	123
5.2.3	Angebote für Schülerinnen und Schüler mit einer von der Landessprache abweichenden Herkunftssprache.....	124
5.3	Zusammenfassung und Diskussion	125
	Literatur	127

6	Soziale Herkunft, Zuwanderungshintergrund und Lesekompetenz	129
	<i>Mirjam Weis, Katharina Müller, Julia Mang, Jörg-Henrik Heine, Nicole Mahler & Kristina Reiss</i>	
6.1	Soziale und zugewanderungsbezogene Herkunft	129
6.1.1	Die Bedeutung der sozialen und zugewanderungsbezogenen Herkunft für die Lesekompetenz.....	129
6.1.2	Erfassung der sozialen Herkunft bei PISA 2018	131
6.1.3	Erfassung des Zuwanderungshintergrunds bei PISA 2018	134
6.2	Soziale Herkunft und Lesekompetenz im internationalen Vergleich	136
6.2.1	Zusammenhang zwischen dem sozioökonomischen beruflichen Status und der Lesekompetenz im internationalen Vergleich	136
6.2.2	Zusammenhang zwischen dem sozioökonomischen und soziokulturellen Status und der Lesekompetenz im internationalen Vergleich.....	140
6.3	Soziale Herkunft und Lesekompetenz in Deutschland	143
6.4	Zuwanderungshintergrund und Lesekompetenz im europäischen Vergleich.....	146
6.4.1	Sozioökonomischer beruflicher Status der Eltern von Jugendlichen mit und ohne Zuwanderungshintergrund.....	147
6.4.2	Lesekompetenz von Jugendlichen mit und ohne Zuwanderungshintergrund	149
6.5	Jugendliche mit Zuwanderungshintergrund in Deutschland	151
6.6	Zuwanderungsbezogene und soziale Herkunft im Zusammenhang mit Lesekompetenz in Deutschland.....	154
6.7	Zusammenfassung und Ausblick	158
	Literatur.....	160
7	Vertiefende Analysen zur Umstellung des Modus von Papier auf Computer	163
	<i>Frank Goldhammer, Scott Harrison, Sarah Bürger, Ulf Kroehne, Oliver Lüdtkke, Alexander Robitzsch, Olaf Köller, Jörg-Henrik Heine & Julia Mang</i>	
7.1	Moduswechsel in PISA	165
7.1.1	Moduseffekte.....	165
7.1.2	Umgang mit Moduseffekten	166
7.1.3	Weitere methodische Änderungen	167
7.2	Nationale Ergänzungsstudie zu Moduseffekten in Deutschland	169
7.2.1	Hintergrund und Motivation.....	169
7.2.2	Zentrale Forschungsfragen.....	169
7.3	Datengrundlage der nationalen Ergänzung in PISA 2018.....	170
7.3.1	Untersuchungsdesign und Stichprobe.....	170
7.3.2	Instrumente	172
7.4	Ergebnisse.....	172
7.4.1	Modusunterschiede im erfassten Konstrukt.....	172
7.4.2	Modusunterschiede in den Itemparametern	175
7.4.3	Vergleich originaler und marginaler Trends	178
7.5	Schlussfolgerungen und Diskussion	183
	Literatur	184

8 Mathematische Kompetenz in PISA 2018 – aktueller Stand und Entwicklung 187

Frank Reinhold, Kristina Reiss, Jennifer Diedrich, Sarah Hofer

& Aiso Heinze

8.1	Mathematische Kompetenz in PISA 2018.....	189
8.1.1	Rahmenkonzeption mathematischer Kompetenz.....	189
8.1.2	Erfassung mathematischer Kompetenz und Kompetenzstufen.....	191
8.2	Mathematische Kompetenz im internationalen Vergleich.....	194
8.2.1	Mathematische Kompetenz: Mittelwerte und Streuung.....	195
8.2.2	Verteilung auf Kompetenzstufen.....	197
8.2.3	Geschlechterunterschiede.....	199
8.2.4	Veränderungen der mathematischen Kompetenz.....	201
8.3	Vertiefende Analysen der mathematischen Kompetenz in Deutschland.....	201
8.3.1	Kompetenzstände von Fünfzehnjährigen verschiedener Schularten.....	201
8.3.2	Mädchen und Jungen in Deutschland.....	203
8.3.3	Entwicklung der mathematischen Kompetenz in Deutschland.....	204
8.4	Diskussion.....	205
	Literatur.....	207

9 Naturwissenschaftliche Kompetenz in PISA 2018 – aktueller Stand, Veränderungen und Implikationen für die naturwissenschaftliche Bildung in Deutschland 211

Anja Schiepe-Tiska, Silke Rönnebeck & Knut Neumann

9.1	Naturwissenschaftliche Kompetenz in PISA 2018.....	213
9.1.1	Die Rahmenkonzeption der Domäne Naturwissenschaften.....	213
9.1.2	Der Naturwissenschaftstest in PISA 2018.....	215
9.2	Naturwissenschaftliche Kompetenz im internationalen Vergleich.....	222
9.2.1	Mittelwerte und Streuungen der naturwissenschaftlichen Kompetenz.....	223
9.2.2	Verteilung auf die Stufen der naturwissenschaftlichen Kompetenz.....	226
9.2.3	Unterschiede zwischen Mädchen und Jungen in der naturwissenschaftlichen Kompetenz.....	228
9.3	Vertiefende Analysen der naturwissenschaftlichen Kompetenz in Deutschland.....	230
9.3.1	Unterschiede zwischen Schularten.....	230
9.3.2	Die Veränderung der naturwissenschaftlichen Kompetenz im Vergleich zu PISA 2006 und PISA 2015.....	233
9.4	Zusammenfassung und Diskussion – Stärken und Herausforderungen in Deutschland.....	237
	Literatur.....	238

10	PISA 2018 – die Methodologie	241
	<i>Jörg-Henrik Heine & Kristina Reiss</i>	
10.1	Einführung – Messung von latenten Merkmalen und Kompetenzen	241
10.2	Methodische Veränderungen seit PISA 2012	242
10.2.1	Papier- und computerbasierte Testung.....	243
10.2.2	Mehrstufig adaptives Testdesign.....	244
10.3	Aspekte der technischen Durchführung 2018 in Deutschland.....	247
10.3.1	Datengewinnung – Ziehung der Stichprobe.....	247
10.3.2	Das Testsystem – Student Delivery System (SDS)	248
10.3.3	Prozeduren der Datenaufbereitung.....	250
10.4	PISA-Daten und Methoden für Sekundäranalysen	251
10.4.1	Verfügbarkeit der PISA-Daten.....	251
10.4.2	Praktische Hinweise zur Analyse von PISA-Daten.....	252
10.5	Zusammenfassung.....	254
	Literatur	255
	 Abbildungsverzeichnis.....	 259
	Tabellenverzeichnis	262
	Die Autorinnen und Autoren dieses Berichtsbandes	265

Der Online-Anhang ist unter www.waxmann.com/buch4100 kostenfrei abrufbar.

Vorwort

PISA, das *Programme for International Student Assessment*, ist 2018 in eine neue Runde gegangen. Die Studie wurde in Deutschland – genauso wie alle PISA-Erhebungen seit 2012 – am Zentrum für internationale Vergleichsstudien e.V. (ZIB) durchgeführt, einem An-Institut der Technischen Universität München, zu dem auch das Institut für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel und das Deutsche Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt gehören. Mehr als 600 000 Schülerinnen und Schüler in 79 Staaten weltweit haben sich an der Erhebung beteiligt. Wir wissen nun mehr über ihre Kompetenzen im Lesen, in der Mathematik und in den Naturwissenschaften, aber auch mehr über ihre Motivation, ihr häusliches Umfeld, ihre Möglichkeiten zum Lernen und ihren sozialen Hintergrund. PISA gibt damit ganz wesentlich Auskunft über eine junge Generation fünfzehnjähriger Schülerinnen und Schüler. Anzumerken ist dabei, dass PISA nicht nur die Kompetenzen von Jugendlichen und damit zusammenhängende Variablen benennen kann, sondern aufgrund repräsentativer Stichproben auch den internationalen Vergleich ermöglicht und damit eine Interpretation der Ergebnisse im Kontext der Bildungssysteme unterschiedlicher Staaten.

Entsprechend ist PISA ein zentrales Element, um Stärken und Schwächen der beteiligten Bildungssysteme zu identifizieren, zu beschreiben und im internationalen Kontext zu bewerten. Der vorliegende Band stellt die Ergebnisse von PISA 2018 dar, zeigt den jeweiligen Hintergrund auf und ordnet beides in den internationalen Kontext ein. Berichtet wird an dieser Stelle vor allem auf Grundlage der Ergebnisse in Deutschland. Im Fokus stehen damit die Kompetenzen der Schülerinnen und Schüler sowie die Randbedingungen der Kompetenzentwicklung in Deutschland und der Zusammenhang mit dem deutschen Bildungssystem.

Wie in den früheren PISA-Erhebungen geht es auch 2018 weniger um das schlichte Wissen der Schülerinnen und Schüler als vielmehr um ihre Fähigkeit, dieses Wissen in unterschiedlichen Problemstellungen und Kontexten anwenden zu können. Dabei hat sich die konkrete Umsetzung seit PISA 2000 weiterentwickelt und sich damit an veränderte gesellschaftliche Bedingungen angepasst. Das wird ganz besonders in einer neuen Rahmenkonzeption für die Hauptdomäne Lesen deutlich, die diese Kompetenz breiter als bisher definiert. Lesekompetenz zeigt sich nicht nur im Lesen gedruckter Texte, sondern etwa auch in der Fähigkeit, Foren im Internet zu folgen oder Informationen aus unterschiedlichen Quellen zusammenzutragen. Diese Anforderungen spiegeln sich in neu gestalteten Aufgaben wider, in denen die Möglichkeiten zur Interaktion – bei Bearbeitung der Tests mit dem Computer – explizit genutzt werden. Damit liegen für das Lesen und die Naturwissenschaften neue Aufgabenformate vor, für die Mathematik werden sie im Rahmen der Erhebung PISA 2021 entwickelt.

Fraglos ist PISA ein internationaler Gradmesser für Bildung und kann Trends für die Entwicklung von Bildungssystemen aufzeigen. Dazu braucht es die kontinuierliche Erhebung mit möglichst ähnlichen Instrumenten über die Zeit hinweg. Die Studie mit ihren Tests und Fragebögen kann allerdings nicht permanent auf einem bestimmten Stand verbleiben, sondern muss sich an die Realität, also an das aktuelle schulische und private Umfeld von Fünfzehnjährigen anpassen. Gerade in diesem Spannungsfeld ist der Berichtsband zu lesen und gerade auf diese Weise ist es möglich, nicht nur einen Zeitpunkt abzubilden, sondern Entwicklungen angemessen aufzuzeigen und die notwendigen Änderungen mitzunehmen. So gibt es auf der einen Seite eine wichtige und spannende Kontinuität bei der Interpretation der Daten von PISA 2000 bis zu PISA 2018. Auf der anderen Seite steht aber das wichtige Bemühen, einen neuen Kontext mit veränderten Lebens- und Arbeitsbedingungen zu berücksichtigen.

Systematische Vergleiche sind wichtig, um Informationen über Bildungssysteme zu erhalten. Das gilt gleichermaßen für nationale wie für internationale Vergleichsstudien. Es ist sicherlich nicht vermessen zu sagen, dass PISA dabei eine besondere Rolle spielt. Das Wissen darum, mit welcher Grundbildung Jugendliche in weitere Phasen ihrer Ausbildung gehen, ist unverzichtbar für jede einzelne Gesellschaft. Die internationale Einordnung bietet eine zusätzliche Information über eventuellen Handlungsbedarf. Es wäre wünschenswert, hier noch viel mehr Staaten einzubinden und vielleicht sogar gemeinsame Handlungsoptionen zu entwickeln.

München im November 2019

Kristina Reiss

1 PISA 2018 – Ziele und Inhalte der Studie

Mirjam Weis & Kristina Reiss

In der OECD-Studie PISA werden Kompetenzen im Lesen, in der Mathematik und in den Naturwissenschaften mit standardisierten Tests gemessen. Darüber hinaus werden mit Fragebögen Hintergrundvariablen sowie auch motivationale Orientierungen und Einstellungen erfasst. Die Studie wird alle drei Jahre durchgeführt und jedes Mal steht dabei einer der drei Kompetenzbereiche Lesen, Mathematik oder Naturwissenschaften im Mittelpunkt. Mit der PISA-Studie 2018 begann der dritte Zyklus und zum dritten Mal war Lesen die Hauptdomäne. Insgesamt beteiligten sich 79 Staaten an PISA 2018, davon waren 37 OECD-Staaten. Es wurden international in knapp 22 000 Schulen die Daten von mehr als 600 000 Schülerinnen und Schülern erhoben. In Deutschland nahmen an 223 Schulen insgesamt 5 451 Schülerinnen und Schüler an der PISA-Testung 2018 teil.

1.1 Die PISA-Studien und das Bildungssystem

Seit dem Jahr 2000 werden in den PISA-Studien (*Programme for International Student Assessment*) grundlegende Kompetenzen von Schülerinnen und Schülern gegen Ende der Pflichtschulzeit international verglichen. Entsprechend liegen in den zentralen Bereichen Lesen, Mathematik und Naturwissenschaften empirische Daten vor, die Aussagen über Erträge der Bildungssysteme der beteiligten Staaten ermöglichen (Jude & Klieme, 2010; KMK, 2016; Sälzer & Reiss, 2016). Zudem können durch die langfristige Anlage und die regelmäßigen PISA-Studien im Abstand von drei Jahren Entwicklungen und Veränderungen festgestellt und beschrieben werden. Aus diesen Gründen ist die PISA-Studie in Deutschland Teil der Gesamtstrategie zum Bildungsmonitoring der Kultusministerkonferenz (KMK, 2006; 2016). Umgekehrt spielen die Ergebnisse der PISA-Studien eine wichtige Rolle für Maßnahmen zur Verbesserung des Bildungssystems. So ist es auch eine Reaktion auf die schlechten Ergebnisse der Jugendlichen in Deutschland bei der PISA-Studie 2000, dass Maßnahmen zur Steigerung der Bildungsqualität auf Grundlage von verbindlichen Standards (z. B. KMK, 2004) sowie einer ergebnisorientierten Evaluation durch die Kultusministerkonferenz beschlossen wurden (KMK, 2002).

Die Ergebnisse der PISA-Studie erlauben alle drei Jahre den direkten Vergleich der Kompetenzen von Fünfzehnjährigen in unterschiedlichen Staaten. Somit wird ein

„Benchmarking“ ermöglicht, das Rückschlüsse auf die Leistungsfähigkeit der Bildungssysteme zulässt. Diese „horizontale Einordnung“ bekommt ihren Wert vor allem durch die Teilnahme vieler und unterschiedlicher Staaten. Waren es bei der ersten PISA-Studie im Jahr 2000 nur 32 teilnehmende Staaten, so stieg diese Zahl 2018 auf 79 Staaten¹. Dadurch wird insbesondere eine Einordnung der Ergebnisse aus Deutschland im Vergleich zu einer Vielzahl von Staaten ermöglicht.

Darüber hinaus können die Kompetenzen der Fünfzehnjährigen über die Zeit und damit im Trend betrachtet werden. Die Ergebnisse von mehreren Erhebungszeitpunkten liefern Informationen über Veränderungen und können unter anderem im internationalen Vergleich analysiert werden. So zeigten sich seit der PISA-Studie 2000, als die Lesekompetenz der Jugendlichen in Deutschland unter dem durchschnittlichen Wert der OECD-Staaten lag (Artelt, Stanat, Schneider & Schiefele, 2001), deutliche Verbesserungen, die vermutlich – und zumindest unter anderem – auf gezielte Unterstützungsmaßnahmen zurückzuführen sind (Weis et al., 2016). So wurde in Deutschland die Förderung der Lesekompetenz in den sieben zentralen Handlungsfeldern der KMK betont (KMK, 2002). In der Folge wurden auf Länderebene zahlreiche Projekte zur Verbesserung der Lesekompetenz ergriffen (z. B. *Leseeinitiative Thüringen*, *Leselust in Rheinland-Pfalz* und *Bist du auch lesekalisch*; Artelt et al., 2007). Alle Bundesländer beteiligten sich zwischen 2008 und 2010 am Projekt *ProLesen* (KM Bayern, 2011). Zudem legten die deutschlandweit eingeführten Bildungsstandards einen starken Fokus auf die Lesekompetenz (z. B. KMK, 2004). Seit 2013 wird ein umfassendes deutschlandweites Programm zur Lese- und Sprachförderung durch die BiSS-Initiative (Bildung durch Sprache und Schrift) von Bund und Ländern umgesetzt (Schneider et al., 2012). Aktuelle Informationen zur Leseförderung findet man auf der Webseite „Lesen in Deutschland – Projekte und Initiativen zur Leseförderung“ (<http://www.lesen-in-deutschland.de/html/index.php>).

Insgesamt hat sich PISA als sehr nützliches Instrument erwiesen, mit dem empirisch fundierte Aussagen über die Kompetenzen von Jugendlichen in schulischen Kernbereichen getroffen werden können. Durch die internationale Ausrichtung ist damit auch eine vergleichende Betrachtung von Bildungssystemen verschiedener Staaten möglich. Sicherlich ist bei einer solchen Bewertung Vorsicht geboten. Unterschiede in den Kompetenzen von Fünfzehnjährigen haben ihre Ursache nicht ausschließlich in unterschiedlichen Bildungssystemen, sondern auch wirtschaftliche oder kulturelle Faktoren können eine wichtige Rolle spielen (siehe z. B. Wößmann, 2016). Daher ist es wichtig, nicht nur den rein numerischen Vergleich und damit das „Ranking“ zu betrachten, sondern die Ergebnisse sinnvoll einzuordnen.

1 Im Folgenden wird der Begriff „Staat“ auch für sogenannte „Volkswirtschaften“ benutzt (z. B. OECD, 2016, S. 31). Darunter fallen etwa Baku (Aserbaidschan) oder ein Zusammenschluss von vier chinesischen Verwaltungseinheiten Peking, Shanghai, Jiangsu und Zhejiang (B-S-J-Z; OECD, 2019b). Die Bezeichnungen der Staaten orientieren sich am Verzeichnis der Staatennamen für den amtlichen Gebrauch in der Bundesrepublik Deutschland (Auswärtiges Amt, 2019).

1.2 Inhalte der PISA-Studie 2018

Ziel der PISA-Studie ist es, zentrale und grundlegende Kompetenzen von Fünfzehnjährigen gegen Ende der Pflichtschulzeit zu erfassen, um Aussagen darüber treffen zu können, wie gut die Jugendlichen auf eine erfolgreiche Teilhabe an der modernen Gesellschaft vorbereitet sind (Jude & Klieme, 2010; OECD, 2019a; Sälzer & Reiss, 2016). Dabei werden die zentralen und grundlegenden Kompetenzen im Sinne der sogenannten „Literacy“ als funktionale Grundbildung verstanden. Auf diese Weise werden zum einen die Anwendung von Wissen für das gegenwärtige und zukünftige Leben der Jugendlichen und zum anderen die Anschlussfähigkeit für kontinuierliches Weiterlernen über die Lebensspanne eingeschlossen und betont (OECD, 2019a; Sälzer & Reiss, 2016). Entsprechend liegt ein Fokus darauf, Wissen und Fertigkeiten auf neue Kontexte und Problemstellungen sowohl im Alltag als auch im Berufs- und Bildungskontext anwenden zu können. Diese zentralen und grundlegenden Kompetenzen von Fünfzehnjährigen werden im Lesen, in der Mathematik und in den Naturwissenschaften erhoben. Die drei Bereiche wurden insbesondere gewählt, da sie weltweit in Schulen vermittelt werden und somit als Grundlage eines internationalen Vergleichs geeignet sind. Sie werden außerdem als grundlegend für das persönliche und berufliche Leben angesehen (OECD, 2019a; Sälzer & Reiss, 2016).

Kompetenzen im Lesen, in der Mathematik und in den Naturwissenschaften werden bei PISA mit standardisierten Tests gemessen. Darüber hinaus werden mit Fragebögen, die Schülerinnen und Schüler, ihre Eltern, Schulleitungen und Lehrkräfte ausfüllen, weitere Variablen wie etwa der sozioökonomische Status der Familie, der Zuwanderungshintergrund, Lernkontexte in der jeweiligen Schule, Einstellungen, Motivation und Lernstrategien erfasst. Der aktuellen Forschung und Literatur folgend wird den motivationalen Orientierungen und Einstellungen in den Rahmenmodellen der PISA-Studie eine zunehmend wichtige Rolle für den Erfolg in der Schule und der Arbeitswelt zugeschrieben (Heckman, 2006; Gutman & Schoon, 2013; OECD, 2019a). Daher wird bei PISA 2018 noch stärker als bisher ihre Relevanz und ihr Zusammenhang mit den Leistungsvariablen betont (vgl. Kapitel 4).

Bei jeder PISA-Studie steht einer von drei Kompetenzbereichen im Mittelpunkt und wird als Hauptdomäne untersucht. Mit PISA 2018 beginnt der dritte Zyklus und zum dritten Mal (nach 2000 und 2009) ist Lesen die Hauptdomäne, sodass die Lesekompetenz der Fünfzehnjährigen mit besonders vielen Aufgaben erfasst wird (vgl. Kapitel 2 und 3). Die theoretische Rahmenkonzeption zum Lesen wurde grundlegend überarbeitet, um Texte, die in unterschiedlicher Art digital präsentiert werden, einzu beziehen (OECD, 2019a). Zur Lesekompetenz gehört es, relevante Informationen aus Texten herauszusuchen, Texte zu verstehen, Texte zu nutzen und über sie zu reflektieren. Dazu werden Texte unterschiedlicher Art präsentiert und die Fähigkeit der Jugendlichen untersucht, deren Aussagen, Absichten und Form zu erkennen und zu interpretieren. Es geht genauso darum, sie in einen größeren Zusammenhang einzuordnen und zu bewerten. Lesen im digitalen Zeitalter bedeutet darüber hinaus, in entsprechenden

digitalen Umgebungen zu arbeiten. Der PISA-Test beinhaltet daher interaktive Aufgaben mit mehreren zu lesenden Texten in einer simulierten Web-Umgebung. So wird etwa die Fähigkeit angesprochen, Informationen durch das Navigieren auf Webseiten zu finden, zu vernetzen und zu beurteilen. Sowohl dem standardisierten Test zur Erfassung der Lesekompetenz als auch dem Fragebogen liegt die neue Rahmenkonzeption zur Lesekompetenz zugrunde (OECD, 2019a).

Die Kompetenzen in der Mathematik und in den Naturwissenschaften werden bei der PISA-Studie 2018 als Nebendomänen erfasst. Mathematische Kompetenz wird als Fähigkeit verstanden, die Rolle der Mathematik in der heutigen Welt zu erkennen und zu verstehen, mathematische Urteile treffen zu können sowie mathematisches Wissen und Fertigkeiten nutzen und anwenden zu können (vgl. Kapitel 8; Jude & Klieme, 2010; OECD, 2019a). Dabei geht es sowohl um die Anwendung im privaten und schulischen Bereich als auch um Anwendungen für einfache wissenschaftliche Fragen und Problemstellungen. Naturwissenschaftliche Kompetenz wird als Fähigkeit verstanden, grundlegende Konzepte der Physik, der Chemie, der Biologie und der Geowissenschaften sowie deren Bedeutung in der Welt zu verstehen. Zudem geht es darum, naturwissenschaftliches Wissen bei der Beantwortung wirklichkeitsnaher Fragen zu nutzen. Es wird getestet, inwiefern Jugendliche ihr Wissen anwenden können, um naturwissenschaftliche Fragestellungen zu erkennen, naturwissenschaftliche Phänomene zu beschreiben und Schlussfolgerungen zu ziehen (vgl. Kapitel 9; Jude & Klieme, 2010; OECD, 2019a).

Um die Kompetenzen der Jugendlichen über die Zeit vergleichen zu können, werden Trendaufgaben eingesetzt, die bereits in früheren PISA-Studien verwendet wurden. So umfasste der PISA-Test 2018 zur Messung der Lesekompetenz insgesamt 245 Aufgaben, wovon 72 Trendaufgaben waren (vgl. Kapitel 10). Für die Nebendomänen Mathematik und Naturwissenschaften wurden ausschließlich Trendaufgaben eingesetzt (OECD, 2019b).

1.3 Durchführung der PISA-Studie 2018 in Deutschland

1.3.1 Stichprobe

Die Anzahl der teilnehmenden Staaten hat im Verlauf der vergangenen PISA-Runden deutlich zugenommen (z. B. PISA 2000: 32 Staaten; 2012: 65 Staaten; 2015: 72 Staaten). An der PISA-Studie 2018 nahmen 79 Staaten teil (OECD, 2019b), wovon 37 OECD-Staaten und 42 sogenannte OECD-Partnerstaaten² sind. Insgesamt waren knapp 22 000 Schulen sowie mehr als 600 000 Schülerinnen und Schüler beteiligt. In Deutschland nahmen an 223 Schulen insgesamt 5 451 Schülerinnen und Schüler aller Schular-

2 Von Zypern liegen dem PISA National Center Deutschland keine Daten vor und von Vietnam liegen für die Kompetenzskalen keine Daten vor. Daher werden im vorliegenden Bericht für die drei Kompetenzbereiche Ergebnisse von lediglich 40 OECD-Partnerstaaten einbezogen.

ten an der PISA-Testung 2018 teil. Zusätzlich wurde diese altersbasierte Stichprobe im Rahmen einer internationalen Option um eine klassenbasierte Zusatzstichprobe erweitert. Dabei wurde in den teilnehmenden Schulen (außer in den beruflichen Schulen) die Schülerstichprobe um jeweils 15 Schülerinnen und Schüler der 9. Jahrgangsstufe ergänzt.

Im Fokus von PISA stehen Fünfzehnjährige, da in diesem Alter in vielen Staaten die Pflichtschulzeit endet. Außerdem wird erwartet, dass Schülerinnen und Schüler in diesem Alter grundlegende Fertigkeiten und Wissen erworben haben müssen, um erfolgreich in eine weiterführende Schule oder die Arbeitswelt zu wechseln (OECD, 2019a). PISA adressiert ausschließlich Jugendliche, die eine Schule besuchen, und basiert 2018 auf einer Stichprobe von Jugendlichen, die im Jahr 2002 geboren wurden. Es werden repräsentative Stichproben dieser Jugendlichen von allen teilnehmenden Staaten angestrebt. Eine Voraussetzung zur Teilnahme an der Studie ist, dass die Schülerin oder der Schüler mindestens seit einem Jahr vor der Testung in der jeweiligen Testsprache unterrichtet wurde. Daher werden Jugendliche mit Zuwanderungshintergrund, die weniger als ein Jahr in der Testsprache unterrichtet wurden, von der Erhebung ausgeschlossen. Außerdem werden Schülerinnen und Schüler ausgeschlossen, die aus körperlichen, emotionalen oder geistigen Gründen nicht in der Lage sind, den PISA-Test selbstständig und ohne fremde Hilfe zu bearbeiten. Ein weiterer Ausschlussgrund ist die von der Schulkordinatorin oder vom Schulkoordinator (vgl. Kapitel 10) als unzureichend eingeschätzte Beherrschung der Testsprache. Es gilt allerdings zu berücksichtigen, dass die Anteile der Fünfzehnjährigen, welche von PISA erfasst werden, zwischen den Staaten variieren. Dieser Anteil beträgt in den meisten Staaten etwa 90 Prozent, ist aber in einzelnen Staaten deutlich geringer (z. B. 45.8% in Albanien, 61.9% in Kolumbien, 66.4% in Mexiko, 86.3% in Kanada; OECD, 2019b). Bei PISA 2018 betrug der Anteil in Deutschland 99.3 Prozent, das ist der höchste Wert im Vergleich zu allen anderen PISA-Teilnehmerstaaten (vgl. OECD, 2019b).

Die Ziehung der repräsentativen Stichproben erfolgte in einem mehrstufigen Verfahren, welches von der internationalen Projektleitung vorgegeben und kontrolliert wird. Zunächst wurden per Zufall 234 Schulen in Deutschland für die PISA-Studie 2018 ausgewählt. Von diesen ursprünglich gezogenen Schulen nahmen 223 tatsächlich teil, was einer gewichteten Teilnahmequote von 98 Prozent auf Schulebene entspricht.³ Im nächsten Schritt wurden Schülerinnen und Schüler innerhalb der ausgewählten Schulen ebenfalls zufällig ausgewählt. In Deutschland sind aufgrund von Beschlüssen der KMK und des BMBF alle gezogenen staatlichen Schulen sowie die ausgewählten Schülerinnen und Schüler zur Teilnahme an der PISA-Testung verpflichtet. Bei der PISA-Studie 2018 betrug die gewichtete Teilnahmequote auf Ebene der Schülerinnen und Schüler 90 Prozent. Hinsichtlich der Bearbeitung der Fragebögen gibt es in einigen Bundesländern ebenfalls Verpflichtungen, während diese in anderen Bundesländern freiwillig ist. Insgesamt bearbeiteten in Deutschland 80 Prozent der ausgewählten Jugendlichen den Schü-

3 An zwei Schulen in Deutschland war die Teilnahmequote der Schülerinnen und Schüler gering (50–75%). Die Daten dieser zwei Schulen sind zwar in den Analysen enthalten, gehen aber nicht in die Berechnung des Teilnahmestatus ein.

lerfragebogen. Weitere Informationen zur Beschreibung und Auswahl der Stichprobe bei PISA 2018 finden sich in Mang et al. (2019).

In dem vorliegenden nationalen Bericht wird für die vertiefenden nationalen Analysen in Deutschland analog zum Bericht der PISA-Studie 2015 (vgl. Sälzer & Reiss, 2016) zwischen Gymnasien und nicht gymnasialen Schularten (Hauptschule, Schule mit mehreren Bildungsgängen, Integrierte Gesamtschule und Realschule) unterschieden. Weitere Differenzierungen der nicht gymnasialen Schularten sind nicht sinnvoll, da diese zwischen den verschiedenen Bundesländern nicht mehr vergleichbar sind (Autorengruppe Bildungsberichterstattung, 2018). Bei der PISA-Studie 2018 betrug die Bildungsbeteiligung am Gymnasium 35.5 Prozent und an den nicht gymnasialen Schularten 59.7 Prozent. Berufliche Schulen sowie Sonder- und Förderschulen wurden von 4.8 Prozent der Jugendlichen besucht. Diese werden bei den schulartspezifischen Analysen wegen der geringen Stichprobengröße nicht gesondert betrachtet, sondern gehen ausschließlich in die Analysen der Gesamtstichprobe ein.

1.3.2 Prozeduren

Die Schülerinnen und Schüler bearbeiteten die Testaufgaben und Fragebögen in der Schule am Computer. Die Testungen wurden von Testleiterinnen und Testleitern durchgeführt, die sich mit jeweils einer Lehrkraft pro Schule absprachen und koordinierten. Um die internationale Vergleichbarkeit zu gewährleisten, entsprach die Durchführung dem international vorgegebenen Ablaufplan. Es wurde sichergestellt, dass die internationalen Vorgaben eingehalten wurden, indem geschultes Personal stichprobenartig die Testungen in allen Teilnehmerstaaten besuchte. Während die Eltern die Fragebögen zu Hause auf Papier beantworteten, gaben Lehrkräfte und Schulleitungen ihre Antworten in der Schule in computerbasierte Systeme ein (für Einzelheiten zur Durchführung von PISA 2018 siehe Kapitel 10; für allgemeine Aspekte der Durchführung von PISA in Deutschland sowie der Sicherung der internationalen Vergleichbarkeit siehe Sälzer & Reiss, 2016).

1.4 Der Rahmen der PISA-Studie 2018: die Balance zwischen den PISA-Zyklen

PISA ist immer eine Herausforderung und auch PISA 2018 ist keine Ausnahme. So gilt es einerseits, Kompetenzen zu testen, die dem realen und zeitgemäßen Kontext von Jugendlichen entsprechen. Andererseits möchte PISA auch Entwicklungen berichten, also zeigen, ob und gegebenenfalls wie sich diese Kompetenzen im Laufe der Zeit verändern. Damit geht es beim Design der Testung um eine sinnvolle Balance zwischen Kohärenz und Innovation bei den Testaufgaben und den psychometrischen Methoden. Bei

PISA 2018 wurde diese Herausforderung gesehen und aus inhaltlicher Sicht mit einer innovativen Rahmenkonzeption zur Lesekompetenz beantwortet, die Aspekte des Lesens von digitalen Texten beinhaltet (vgl. Kapitel 2 und 3). Die Entwicklung neuer Aufgaben folgte diesem Konzept. Es ist natürlich unabdingbar, dass weiterhin Aufgaben einbezogen werden, die den Trend erfassen und die Entwicklung von Bildungssystemen widerspiegeln. Ähnlich ist die Methodik von PISA zu sehen. Auch hier gilt es Entwicklungen und Fortschritte zu berücksichtigen, die neue psychometrische Möglichkeiten, wie etwa das adaptive Testen, eröffnen (vgl. Kapitel 10; Mang et al., 2019). Genauso wichtig ist es aber, alte und neue Aspekte so in Einklang zu bringen, dass Entwicklungen, Entwicklungspotenzial und Entwicklungsbedarf gesehen werden können. Mit Sicherheit liegt genau da eine große Herausforderung.

PISA ist aber nicht nur eine Herausforderung, sondern auch eine Erfolgsgeschichte. Den beteiligten Staaten gelang es und gelingt es weiter, sich auf ein gemeinsames Konzept zu einigen und die Ergebnisse vergleichend zu interpretieren. Eine solche Kooperation ist nicht selbstverständlich, sondern ein Ergebnis guter internationaler Zusammenarbeit trotz unterschiedlicher wirtschaftlicher, politischer, gesellschaftlicher und kultureller Voraussetzungen und Ausgangsbedingungen.

Literatur

- Artelt, C., McElevany, 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 (BMBF).
- 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, ... M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_4
- Auswärtiges Amt. (2019). *Verzeichnis der Staatennamen für den amtlichen Gebrauch in der Bundesrepublik Deutschland*. Verfügbar unter <https://www.auswaertiges-amt.de/blob/199312/d472e0efe115296cbc8f90e00a33359b/staatennamen-data.pdf>
- Autorengruppe Bildungsberichterstattung. (2018). *Bildung in Deutschland 2018. Ein indikatorengestützter Bericht mit einer Analyse zu Wirkungen und Erträgen von Bildung*. Bielefeld: Bertelsmann.
- Gutman, L. M., & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people. Literature review*. London: Institute of Education, University of London.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312, 1900–1902. <https://doi.org/10.1126/science.1128898>
- Jude, N. & Klieme, E. (2010). Das Programme for International Student Assessment (PISA). In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73–112). Münster: Waxmann.

- KM Bayern (2011) = Bayerisches Staatsministerium für Unterricht und Kultus, Staatsinstitut für Schulqualität und Bildungsforschung (Hrsg.). (2011). *ProLesen. Auf dem Weg zur Leseschule – Leseförderung in den gesellschaftswissenschaftlichen Fächern. Aufsätze und Materialien aus dem KMK-Projekt „ProLesen“* (2. Aufl.). Donauwörth: Auer.
- KMK (2002) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2002). *PISA 2000 – Zentrale Handlungsfelder. Zusammenfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern. Beschluss der 299. Kultusministerkonferenz vom 17./18.10.2002*. Verfügbar unter <http://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2002/massnahmen.pdf>
- KMK (2004) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.). (2004). *Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. Beschluss vom 4.12.2003* (Beschlüsse der Kultusministerkonferenz). Neuwied: Luchterhand.
- KMK (2006) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Wolters Kluwer.
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring*. München: Wolters Kluwer.
- Mang, J., Wagner, S., Gomolka, J., Schäfer, A., Meinck, S., & Reiss, K. (2019). *Technische Hintergrundinformationen PISA 2018*. München: Technische Universität München. <https://doi.org/10.14459/2019md1518258>
- OECD. (2010). *PISA 2009. Learning trends. Changes in student performance since 2000*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264091580-en>
- OECD. (2016). *PISA 2015 Ergebnisse (Band I): Exzellenz und Chancengerechtigkeit in der Bildung*. Bielefeld: W. Bertelsmann Verlag. <http://dx.doi.org/10.3278/6004573w>
- OECD. (2019a). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- OECD. (2019b). *PISA 2018 results (Volume I): What students know and can do*. Paris: OECD Publishing.
- Sälzer, C. & Reiss, K. (2016). PISA 2015 – die aktuelle Studie. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 13–44). Münster: Waxmann.
- Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, ... Stanat, P. (2012). *Expertise „Bildung durch Sprache und Schrift (BiSS)“: Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung*. Verfügbar unter <http://www.biss-sprachbildung.de/ueber-biss/biss-expertise>
- Weis, M., Zehner, F., Sälzer, C., Strohmaier, A., Artelt, C. & Pfof, M. (2016). Lesekompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 249–283). Münster: Waxmann.
- Wößmann, L. (2016). The importance of school systems: Evidence from international differences in student achievement. *Journal of Economic Perspectives*, 30, 3–32. <https://doi.org/10.1257/jep.30.3.3>

2 Lesekompetenz heute – eine Schlüsselqualifikation im Wandel

Michael Becker-Mrotzek, Thomas Lindauer, Maximilian Pfost,
Mirjam Weis, Anselm Strohmaier & Kristina Reiss

Die Lesekompetenz steht bei der PISA-Studie 2018 im Mittelpunkt. Aufgrund vieler Veränderungen – unter anderem durch die flächendeckende Digitalisierung – wurden die Rahmenkonzeption der Lesekompetenz überarbeitet und neue Aufgaben eingeführt, um auch Lesekompetenz in digitalen Medien zu erfassen. Neben dem Wandel von Wissensquellen und Textformaten ist seit PISA 2000 viel zum Lesen und zur Leseförderung geforscht worden und die neuen Erkenntnisse gingen in die Rahmenkonzeption der Lesekompetenz bei PISA 2018 ein. Beispielsweise wird bei PISA 2018 erstmals die Unterscheidung basaler und komplexer Teilfertigkeiten berücksichtigt, indem die Leseflüssigkeit einbezogen wird. Zudem wird Lesen stärker als bisher als eine zielgerichtete Handlung aufgefasst, deren Gelingen und Ergebnis von den Merkmalen der Lesenden, von Eigenschaften des Textes sowie der Leseaufgabe selbst abhängen. Dazu werden sogenannte Leseszenarien eingesetzt, welche mehrere Texte umfassen, die authentische, ziel- und aufgabenorientierte Lesesituationen simulieren. Insgesamt entsteht ein breiter Blick auf das Konstrukt „Lesekompetenz“ in PISA 2018, welches besser als der Lesekompetenz-Begriff der früheren PISA-Rahmenkonzeptionen zu dem in der Deutschdidaktik breiter angelegten Leseverständnis passt, das immer auch sozial-kulturelle Dimensionen umfasst.

Lesen – das ist spätestens seit PISA 2000 klar – stellt in einer literalen, also schriftbasierten Gesellschaft eine unabdingbare Schlüsselqualifikation dar. Lesen können ist eine unverzichtbare Basis für Lern- und Entwicklungsprozesse sowie für die Teilhabe am kulturellen, beruflichen und sozialen Leben (z. B. UNESCO, 2005). Insofern ist es nicht nur für Individuen erforderlich, lesen zu können, sondern die Gesellschaft ist darauf angewiesen, dass ihre Mitglieder über Lesekompetenz verfügen. Literale Gesellschaften charakterisiert im Gegensatz zu mündlich organisierten Gemeinschaften, dass zentrale Einrichtungen und Institutionen schriftlich organisiert sind. Sie speichern ihr Wissen in Form von Texten, prototypisch stehen dafür Bücher und Bibliotheken. Darüber hinaus ist auch die Kommunikation insbesondere in beruflich-institutionellen Kontexten in weiten Teilen schriftlich organisiert. Hierfür haben sich im Laufe der Zeit eigene Fachsprachen für verschiedene Handlungskontexte mit je spezifischen Textarten herausgebildet, seien es Gesetzestexte, behördliche Bescheide, technische Anleitungen oder Facharti-

kel. Die unterschiedlichen Textarten¹ für verschiedene Kontexte stellen ein „universelles Kulturwerkzeug“ dar (Naumann, Artelt, Schneider & Stanat, 2010, S. 23), nämlich ein Werkzeug für das Wissensmanagement und die Kommunikation.

Es kann offensichtlich nicht allein Aufgabe des Deutschunterrichts sein, Lesekompetenz zu entwickeln und zu vermitteln. Im Deutschunterricht müssen allerdings die dazu notwendigen basalen Fertigkeiten und die grundlegenden, übergreifenden Lesestrategien erlernt werden. Dazu gehören der Schriffterwerb genauso wie die Befähigung zum flüssigen Lesen und zur Nutzung unterschiedlicher Textarten in verschiedenen Kontexten. Auch der kompetente Umgang mit literarischen Texten ist ein wichtiges Ziel des Deutschunterrichts, das jedoch von der Vermittlung einer allgemeinen Lesekompetenz, wie sie für das Lernen in allen Fächern und auch in beruflich-institutionellen Kontexten erforderlich ist, zu unterscheiden ist. Literarische Texte sind ein eigener Gegenstandsbereich, für den zwar auch eine allgemeine Lesekompetenz benötigt wird, der sich darin aber nicht erschöpft. Hier besteht eine Analogie zu anderen Schulfächern (Becker-Mrotzek, Schramm, Thürmann & Vollmer, 2013). So wird für das Verstehen historischer Texte im Geschichtsunterricht Lesekompetenz benötigt, darüber hinaus aber auch historisches Wissen (vgl. Bernhard & Conrad, 2018, Handro & Kilimann, 2019). Genauso ist Lesekompetenz eine Voraussetzung und ein Werkzeug für den erfolgreichen Umgang mit schriftlichen Informationen im Mathematikunterricht (Daroczy, Wolska, Meurers, & Nuerk, 2015; Leiss, Hagen, Neumann & Schwippert, 2017; Leiss, Plath & Schwippert, 2019) und im naturwissenschaftlichen Unterricht (Härtig & Kohnen 2017; Schmellentin, Dittmar, Gilg & Schneider, 2017). Das Lesen, Verstehen und Nutzen fachlicher Texte sowie die Vermittlung der spezifischen Lesestrategien ist daher Aufgabe mehr oder minder aller Schulfächer.

Wegen seiner besonderen Bedeutung ist Lesen alle neun Jahre Hauptdomäne bei PISA, nach 2000 und 2009 nun wieder im Jahr 2018. Allerdings haben sich die grundlegenden Anforderungen und Eigenschaften von Lesen und damit auch die Konzeption von Lesekompetenz in diesem Zeitraum weiterentwickelt. So haben die technologischen Entwicklungen und allem voran die flächendeckende Verfügbarkeit des Internets und seiner zahlreichen Anwendungen (z. B. Webseiten, Suchmaschinen, soziale Medien) dazu beigetragen, dass Texte und Textstrukturen heute ein breiteres Feld umfassen (z. B. OECD, 2019). Ebenso geändert hat sich das Wissen über Lesekompetenz, denn national wie international wurden in den letzten Jahren zahlreiche Studien zum Lesen und vor allem zur Förderung der Lesekompetenz angestoßen. Diese Veränderungen finden ihren Niederschlag bei PISA 2018 in einem modifizierten Konzept von Lesekompetenz, in neuen Aufgabenformaten sowie einem neuen Testbereich zur Leseflüssigkeit (OECD, 2019). Darüber hinaus wird seit PISA 2015 vollständig computerbasiert getestet, das heißt, dass die Schülerinnen und Schüler in Deutschland alle Aufgaben am Computer bearbeiten.

1 Im Folgenden ist von „Textarten“ die Rede, wenn Texte mit einer bestimmten Funktion gemeint sind; in den Fachdidaktiken und der Sprachwissenschaft wird hierfür auch der Ausdruck „Textgenre“ oder „Textsorte“ genutzt.

2.1 Lesekompetenz

2.1.1 Bedeutung der Lesekompetenz

Literale Gesellschaften nutzen Texte als herausragende Mittel des Wissensmanagements *und* der Kommunikation. Daher sind Texte zentrale Träger des gesellschaftlichen Wissens, gespeichert in analogen Printprodukten und digitalen Datenbanken. Texte dienen der Kommunikation und der sozialen Vernetzung sowie dem Austausch von Informationen und tragen so nicht nur im wörtlichen Sinne zur Verständigung zwischen den Menschen bei. Demnach sind Texte ein universell einsetzbares, kulturelles Werkzeug, das sehr flexibel gegenüber Inhalten, Zielen und Nutzern ist. Es können also auf diese Weise fast alle denkbaren Inhalte mit allen möglichen Zielen von jeder Person gespeichert und ausgetauscht werden. Als Texte werden hier in einem weiten Verständnis *kontinuierliche Texte*, bestehend aus Sätzen und Abschnitten, sowie nichtkontinuierliche Darstellungsmittel wie Tabellen, Grafiken, Bilder oder ähnliches mit sprachlichen Anteilen bezeichnet, die zu *gemischten Texten* kombiniert werden können. Dieses relativ breite Verständnis des Textbegriffs erfordert auch einen weiten Begriff der Lesekompetenz, der sich nicht auf kontinuierliche Texte beschränkt. Insbesondere für nichtkontinuierliche und gemischte Texte wird eine umfassende Lesekompetenz benötigt, die über das Verarbeiten von Schriftsprache hinausgeht (Schnotz & Dutke, 2004).

Gemischte (Sach-)Texte kann man als komplexe Werkzeuge verstehen, über die unterschiedlichste inhaltliche Anforderungen einer hoch technologisierten und digitalen Gesellschaft formuliert werden. Sie zu erfassen ist von großer Bedeutung für die Teilhabe an einer literalen Gesellschaft in all ihren Facetten. Aus diesem Grund ist Lesekompetenz in vielen Teilen der Welt ein zentrales Bildungsziel (OECD, 2019); auch in Deutschland spielt Lesekompetenz in den Bildungsstandards für alle Schulstufen eine zentrale Rolle. Bereits in den 2003 formulierten Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss heißt es, dass Lesen den Umgang mit Texten und Medien umfasst (KMK, 2004). In den Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife aus dem Jahr 2012 wird dies so konkretisiert: „Die Schülerinnen und Schüler sind in der Lage, selbstständig Strategien und Techniken zur Erschließung von linearen und nichtlinearen Texten unterschiedlicher medialer Form anzuwenden und zu reflektieren. Lesend erweitern sie ihr kulturhistorisches und domänenspezifisches Orientierungswissen und bewältigen dabei umfangreiche und komplexe Texte“ (KMK, 2014, S. 18).

2.1.2 Digitaler Wandel

Die gesellschaftlichen und technologischen Entwicklungen der letzten zwanzig Jahre haben zu einem Wandel der Rahmenbedingungen von Lesekompetenz geführt. Der Ausbau des Internets sowie die Entwicklung leistungsfähiger und preiswerter digitaler

Endgeräte wie Tablets und Smartphones ermöglicht die breite Nutzung digitaler Medien in praktisch allen Lebensbereichen. Im Rahmen dieser technologischen Entwicklung sind in einem Prozess der gegenseitigen Beeinflussung neue Strukturen und Formen von Texten und neue Kommunikationsmöglichkeiten entstanden. Darüber hinaus beeinflusst der digitale Wandel maßgeblich die Textrezeption sowie die Möglichkeiten zur Erfassung von Lesekompetenz. Zuletzt gewinnt die Bewertung der Glaubwürdigkeit von Texten an Bedeutung.

Die neuen Textarten zeichnen sich unter anderem durch vielfältige Kombinations- und Verknüpfungsmöglichkeiten von Text mit anderen Darstellungsformen wie Bildern, Grafiken, Tabellen, gesprochener Sprache, Film und weiteren Formen aus. Die Verbindung und Verknüpfung unterschiedlicher Formate in Form von gemischten Texten ist zwar nichts prinzipiell Neues, erfordert aber in digitalen Medien jeweils ein angepasstes Vorgehen bei ihrer Nutzung (Becker-Mrotzek, 2003). Schriftlicher Text ist in der Fläche dargestellte Sprache. Er lässt auf den ersten Blick Hinweise auf Inhalt, Texttyp oder Umfang zu und ist damit geeignet, die Rezeption vorab zu strukturieren. Das geht in dieser Form etwa bei Gesprächen nicht, weil diese flüchtig sind und sich über die Zeit erstrecken. Nichtkontinuierliche und gemischte Texte in digitalen Medien greifen das Prinzip der Fläche auf und ergänzen es um vielfältige Möglichkeiten der Darstellung und Rezeption, die sich je nach Medium unterscheiden können. So ist ein kleiner Bildschirm bei einem mobilen Gerät von einem Großbildschirm zu unterscheiden, bei dem etwa das gleichzeitige Öffnen mehrerer Fenster nebeneinander möglich ist. Mit unterschiedlichen Technologien verbinden sich also jeweils spezifische Rezeptionsmöglichkeiten, die gegebenenfalls als Techniken erlernt werden müssen.

Die neuen technischen Möglichkeiten werden auch für die Kommunikation genutzt, wie sich am Beispiel von sozialen Medien zeigt. So ist WhatsApp die meistgenutzte App unter Jugendlichen (Medienpädagogischer Forschungsverbund Südwest, 2018), wobei hier schriftliche Texte insbesondere neben Audionachrichten und Bildern genutzt werden, vielfach durchsetzt mit für Messaging-Dienste typischen Elementen wie etwa sogenannte Textismen, also Änderungen an der konventionellen Orthografie oder Abkürzungen wie „4u“ anstelle von „für dich“ (Kemp, Wood & Waldron, 2014).

Digitale Technologien haben die Vernetzung von Texten untereinander und mit anderen Medien oder Darstellungsformaten erheblich erleichtert. An die Stelle des Hin- und Herblätterns im gedruckten Lexikon, um Verweisen nachzugehen, ist das Anklicken von Links getreten, die beispielsweise ein weiteres Fenster öffnen, sodass beide Textauschnitte nebeneinander stehen und parallel gelesen werden können. Mithilfe von Links lassen sich Verweise auf andere digitale Quellen schnell realisieren. Durch die systematische Nutzung dieser Möglichkeiten bei der Textproduktion entstehen qualitativ neue Textarten, die sogenannten *Texte mit mehreren Quellen*. Diese sind in ihrem Kern nicht mehr linear angelegt, wie das prototypisch etwa bei Erzählungen der Fall ist. Dort legt der Autor beziehungsweise die Autorin den roten Faden durch den Text, dem die Lesenden folgen. Texte mit mehreren Quellen dagegen ermöglichen und erfordern von ihrer Struktur und Anlage her unterschiedliche Lesewege. Je nach Interesse und Zielsetzung

wählt die Leserin oder der Leser aus, was gelesen wird und was nicht, und legt so einen eigenen roten Faden. Insofern erfordert die Nutzung von Texten mit mehreren Quellen neue Lesetechniken und neue Lesestrategien für das selbstständige Navigieren. Dazu gehört es, in Texten mit mehreren Quellen fortwährend Entscheidungen darüber zu treffen, welcher der möglichen Lesewege genutzt werden soll. Die Texte müssen auch in eine für das Verstehen förderliche Reihenfolge gebracht werden. Das Lesen von Texten mit mehreren Quellen erfordert also Entscheidungen, ob ein Informationsangebot für das eigene Leseziel relevant ist oder nicht. Digitale Medien und Texte mit mehreren Quellen erfordern keine qualitativ neue Lesekompetenz, aber eine Ausdifferenzierung und partielle Ergänzungen, etwa um neue Lesestrategien für das Navigieren in diesen Texten. Dies untermauert die Studie von Naumann (2013) basierend auf Daten aus dem PISA 2009 Digital Reading Assessment. Dabei handelt es sich um eine Erweiterung der regulären PISA-Studie, an der 19 Staaten im Jahr 2009 teilnahmen. Die Ergebnisse zeigen, dass die Lesekompetenz im Print-Medium positiv die Effizienz von Navigationsverhalten im Sinne zielgerichteter Auswahl aufgabenrelevanter Textbestandteile im Umgang mit elektronischen Texten vorhersagt. Das Navigationsverhalten sowie die „klassische“ Lesekompetenz im Print-Medium waren beide darüber hinaus prädiktiv für das Textverstehen digitaler Texte (vgl. auch Hahnel, Goldhammer, Naumann & Kröhne, 2016).

Ein interessanter Befund zeigt sich beim direkten Vergleich des Verstehens von linearen Texten, wenn sie den Leserinnen und Lesern entweder in gedruckter Form oder in digitaler Form auf einem Bildschirm (z. B. Computerbildschirm, Tablet) dargeboten werden. So belegt die systematische Übersichtsarbeit von Delgado, Vargas, Ackerman und Salmerón (2018) eine leichte Überlegenheit im Leseverstehen für das gedruckte Medium insbesondere beim Lesen von Sachtexten sowie unter Bedingungen der Zeitrestriktion. Als eine Ursache für diese Diskrepanz wird diskutiert, dass Lesen am Bildschirm einen eher oberflächlichen Verarbeitungsmodus nahelegt und die dargebotene Information weniger tief verarbeitet wird. So zeichnen sich gerade im Bereich der digitalen Medien Interaktionen häufig durch eine schnelle Verarbeitung aus, die bei schwierigeren (Textverstehens-)Aufgaben am Bildschirm zu einer weniger tiefen Verarbeitung führen. Außerdem ist es denkbar, dass für die Bearbeitung von Texten am Bildschirm typische Elemente wie die Notwendigkeit zum Scrollen Aufmerksamkeit binden und eine tiefere Auseinandersetzung erschweren.

Die Verlagerung des Lesens vom traditionellen, gedruckten Text hin zur Rezeption von Information im Internet geht darüber hinaus mit einer weiteren, auch in PISA 2018 abgebildeten Veränderung einher. Während früher häufig lediglich ein einziger Text als Informationsquelle zur Verfügung stand, sind es heute auch Texte mit mehreren Quellen. Diese enthalten oftmals redundante, sich ergänzende oder sich widersprechende Informationen. Demzufolge wird das Verstehen des Textinhalts maßgeblich ergänzt um die Frage nach der Glaubwürdigkeit der dort wiedergegebenen Information, insbesondere der Zuverlässigkeit der Informationsquelle (Bråten & Braasch, 2017; Bråten, Stadler & Salmerón, 2018; Perfetti, Rouet & Britt, 1999). Während im gedruckten Text in der Regel eine Qualitätskontrolle erfolgt, erlauben es Online-Medien, wie zum Beispiel

soziale Medien, Blogs, Foren und private Webseiten, inhaltlich fast uneingeschränkt und gegebenenfalls auch anonym Texte zu verfassen und einem breiten Publikum zur Verfügung zu stellen. Somit ist heute die Fähigkeit, die Glaubwürdigkeit eines Textes richtig einschätzen zu können, ein wichtiger Aspekt von Lesekompetenz.

2.1.3 Stand der Forschung zur Lesekompetenz

Neben dem Wandel der Rahmenbedingungen ist seit PISA 2000 national wie international sehr viel zum Lesen und zur Leseförderung geforscht worden, so dass wir heute über ein verbessertes, weil differenzierteres Verständnis des Leseprozesses, der Lesekompetenz und der Leseförderung verfügen. Ein für den schulischen Alltag zentraler Erkenntnisfortschritt ist die Unterscheidung basaler und komplexer Teilfertigkeiten (Gold, 2018; Lenhard, 2013; Richter & Christmann, 2002).

Basale Leseprozesse

Zu den basalen Leseprozessen gehören das Dekodieren von Wörtern und der Aufbau lokaler Zusammenhänge innerhalb und zwischen Sätzen. Diese Prozesse laufen bei geübten Leserinnen und Lesern weitgehend automatisiert und mühelos ab, weshalb hier auch von flüssigem Lesen oder Leseflüssigkeit gesprochen wird (vgl. Gold, 2018; Kuhn & Stahl, 2003; Lenhard, 2013; Rosebrock & Nix, 2006). Leseflüssigkeit ist die Fertigkeit, Geschriebenes auf Wort- und Satzebene zügig und korrekt zu dekodieren (Nix, 2011; Rosebrock, Nix, Rieckmann & Gold, 2017). Mangelnde Leseflüssigkeit führt dazu, dass im Arbeitsgedächtnis keine oder zu wenig Kapazitäten frei sind, um eine globale Repräsentation des Textinhalts aufzubauen oder um Lesestrategien für das inhaltliche Verständnis des Textes zu nutzen. Leseflüssigkeit wird in der Regel in der Grundschule erworben, aber es mangelt leseschwachen Schülerinnen und Schülern oftmals auch noch in der Sekundarstufe an der nötigen Leseflüssigkeit (Schindler & Richter, 2018). Wesentliche Komponenten der Leseflüssigkeit sind die *Akkuratheit der Wortdekodierung* (Wörter werden richtig erkannt und wiedergegeben) und die *Automatisierung* (das Dekodieren geschieht mühelos und schnell). Akkuratheit und Automatisierung lassen sich auch als Lesegeschwindigkeit zusammenfassen. Dabei ist Lesegeschwindigkeit zwar eine notwendige, aber noch keine hinreichende Bedingung für Leseflüssigkeit. Lesen bedeutet nicht nur, schnell und richtig zu dekodieren, sondern auch, das Gelesene zu verstehen, dazu müssen Sinneinheiten (= Phrasen) als solche erkannt werden.

Diese Aufgabe ist nicht trivial, denn zunächst einmal besteht ein Text aus gleichartigen Buchstabenfolgen, die durch größere Zwischenräume (= Leerzeichen) eine erste Gruppierung erfahren. Sinneinheiten sind z. B. Nominalgruppen wie „*der kleine Hund*“ oder Verbalgruppen wie „*sich schnell verstecken*“. Es geht also darum, diese Sinneinheiten zu erkennen. Dafür werden unter anderem morphosyntaktische Hinweise im Text selbst genutzt, etwa die Wortstellung (Adjektive stehen links vor dem Substantiv „*kal-*

tes Wasser“), gleichartige Flexionsendungen („kaltes, dreieckiges, ungenießbares Wasser“) oder grammatische Kongruenz (z. B. Zusammengehörigkeit von Subjekt und Prädikat: „sie isst gerne Suppe – sie essen gerne Suppe“).

Die Fertigkeit zur sinnvollen Gliederung eines Satzes in zusammenhängende Einheiten zeigt sich auch in einer sinnvollen Betonung. Damit kommt *das sinngestaltende* (= *prosodische*) *Lesen* hinzu. Geübte Leserinnen und Leser dekodieren also nicht nur exakt und schnell, sondern lesen auf Satz- und Textebene auch betont und sinngestaltend. Flüssiges Lesen ist eine zentrale Voraussetzung für textverstehendes Lesen, weil es das Arbeitsgedächtnis entlastet (vgl. Richter & Müller, 2017). Getestet wird die Leseflüssigkeit in PISA 2018 mithilfe von Aufgaben zum Satzverstehen, bei denen schnell zu entscheiden ist, ob ein Satz sinnvoll ist, wie in den drei folgenden Beispielen in Tabelle 2.1 ausgeführt. Dabei wird gemessen, wie viele Sätze in einer gegebenen Zeit, in PISA 2018 sind es drei Minuten, korrekt entschieden werden.

Tabelle 2.1: Beispielaufgaben zur Leseflüssigkeit

Sechs Vögel flogen über die Bäume.	JA	NEIN
Das Fenster sang das Lied lautstark.	JA	NEIN
Der Mann fuhr das Auto zum Geschäft.	JA	NEIN

Komplexe Leseprozesse

Zudem wird Lesen in der aktuellen Leseforschung stärker als bisher als eine zielgerichtete Handlung aufgefasst, deren Gelingen und Ergebnis im Kern von drei Faktoren abhängen, die Snow and the RAND Corporation (2002) in ihrem klassischen Rahmenkonzept beschrieben haben. Danach werden Qualität und Ergebnis des Leseprozesses von Merkmalen der Lesenden, des Textes und der Leseaufgabe beeinflusst. Erst im Zusammenspiel dieser drei Faktoren ergeben sich Art und Ergebnis (= Textverständnis) des Leseprozesses.

Zu den wesentlichen Merkmalen der Lesenden gehören *Weltwissen* und *Sprachwissen*. Weltwissen bezeichnet im Kern das deklarative Wissen über die unterschiedlichen Domänen unserer Wirklichkeit, also das Wissen über Naturphänomene ebenso wie über soziale, historische oder wirtschaftliche Zusammenhänge. Weltwissen spielt beim Lesen deshalb eine wichtige Rolle, weil es als sogenanntes Vorwissen Einfluss nimmt auf das Textverstehen. Je mehr Vorwissen in Bezug auf den Text vorhanden ist, desto leichter können die Textinhalte verstanden und in das vorhandene Wissen integriert werden. Sprachwissen wird hier in einem umfassenden Sinne verstanden; es umfasst neben dem Wortschatz auch das morphosyntaktische Wissen, also die Fertigkeit, Wörter zu Sätzen zusammenzufügen, sowie das pragmatische Wissen, also die Fertigkeit, die kommunikative Absicht einer Äußerung zu verstehen (Nussbaumer, 1991).

Zu den wesentlichen Merkmalen eines Textes gehören seine inhärenten Zwecke beziehungsweise seine kommunikative Funktion, aus denen sich ganz wesentlich Inhalt,

Struktur und Länge herleiten. So soll beispielsweise eine Bedienungsanleitung den Lesenden die Bedienung eines Geräts ermöglichen, was sich wiederum in ihrem Aufbau, ihren sprachlichen Mitteln oder auch der Nutzung von Bildern und Abbildungen niederschlägt (Schnotz, 2002).

Zu den wesentlichen Merkmalen der Leseaufgabe gehören die Ziele, die mit dem Lesen des Textes erreicht werden sollen. Sie ergeben sich in der Regel aus dem größeren Kontext, in dem der Text gelesen wird. Wer beispielsweise ein konkretes Bedienungsproblem mit einem technischen Gerät hat, wird eine Anleitung anders lesen als jemand, der vor der Anschaffung eines Gerätes steht und die Bedienungsfreundlichkeit unterschiedlicher Geräte vergleichen möchte. In dem einen Fall geht es also um konkrete Informationen, im anderen um einen umfassenden Überblick (Rosebrock, 2016).

Neu ist gegenüber den bisherigen Lesemodellen, die bei PISA einbezogen wurden, die ausdrückliche Einbettung des Leseprozesses in übergeordnete Handlungsziele. Das Modell von Snow and the RAND Corporation (2002) nimmt an, dass geübte Leserinnen und Leser Texte lesen, um damit bestimmte Aufgaben zu bewältigen oder eigene Ziele zu erreichen. Diese Aufgaben und Ziele können selbst gestellt sein oder von außen kommen. Zugleich sind Leseaufgaben unterschiedlich komplex und erfordern verschiedene Herangehensweisen. So könnte es in einer Aufgabe lediglich darum gehen, eine einzelne Aussage zu finden, beispielsweise in einer Anleitung den Hinweis auf eine bestimmte Gerätefunktion. Es könnte aber auch notwendig sein, Informationen aus mehreren Texten zusammenzutragen, wenn etwa ein passendes Gerät für bestimmte Bedürfnisse auszuwählen ist. Schließlich kann eine Leseaufgabe auch die kritische Einschätzung eines Textes erfordern, etwa um zu beurteilen, ob ein Text(-Abschnitt) für eine bestimmte Fragestellung relevant ist.

2.1.4 Die Rahmenkonzeption der Lesekompetenz in PISA 2018

Die geänderten Rahmenbedingungen und die damit verbundenen Anforderungen an die Lesekompetenz sowie das erweiterte Wissen darüber haben Anpassungen des theoretischen Lesekonzepts bei PISA erforderlich gemacht. Diese Anpassung erfolgt seit 2009 kontinuierlich und in dreifacher Weise: Das Lesemodell wird modifiziert, es werden neue Textarten berücksichtigt und die Aufgabenformate des Lesetests werden ergänzt. Für die PISA-Studie 2018, in der Lesen zum dritten Mal Hauptdomäne ist, wurde die Rahmenkonzeption darüber hinaus grundlegend überarbeitet. Diese Änderungen werden im Folgenden dargestellt, indem zunächst auf die Rahmenkonzeption bei PISA 2018 und anschließend auf dessen Umsetzung in den Aufgaben des Bereichs Lesen eingegangen wird.

Die Rahmenkonzeption der Lesekompetenz in PISA 2018 berücksichtigt den aktuellen Stand der Forschung. So wird darin erstmals die Leseflüssigkeit als basaler Leseprozess einbezogen. Zudem orientiert sich die Konzeption der komplexen Leseprozesse in der PISA-Rahmenkonzeption an dem Modell von Snow and the RAND Corporation

(2002). Dieses erweiterte Verständnis von Lesekompetenz wird im Modell der Leseprozesse verdeutlicht, welches für PISA 2018 grundlegend neu strukturiert wurde (siehe Abbildung 2.1).

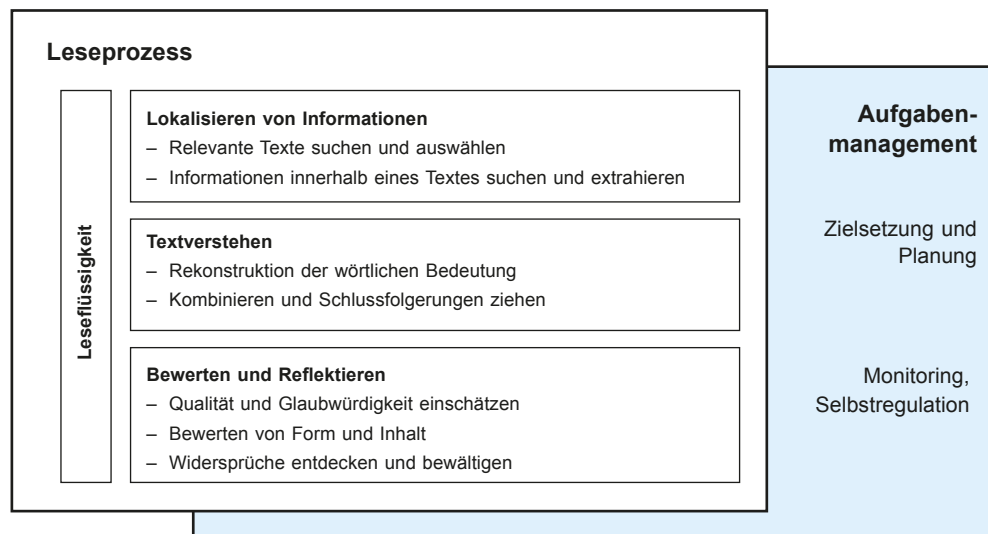


Abbildung 2.1: Modell der Leseprozesse in PISA 2018. Abbildung adaptiert nach OECD, 2019, S. 33

Das Modell der Leseprozesse von PISA 2018 berücksichtigt die Einbettung des Lesens in eine übergeordnete Aufgabe, indem es zwei wesentliche Komponenten unterscheidet, nämlich das *Aufgabenmanagement* und den eigentlichen *Leseprozess* (siehe Abbildung 2.1). Das Aufgabenmanagement bildet – auch graphisch – den Hintergrund des Leseprozesses und umfasst andere Aktivitäten als der eigentliche Leseprozess. Es macht zugleich deutlich, dass der Leseprozess regelmäßig Teil eines größeren Handlungszusammenhangs ist, etwa des schon erwähnten Lösens technischer Probleme durch Lesen von Anleitungen.

Zu den Managementaufgaben gehören das Setzen eigener Ziele und das Planen beziehungsweise Organisieren der Aufgabenbewältigung, denn erst auf der Grundlage von klaren Zielen kann das weitere Vorgehen geplant sowie fortlaufend überwacht und gegebenenfalls korrigiert werden. Aus der Zielsetzung ergibt sich, welche Texte beziehungsweise Textstellen relevant sind und gelesen werden müssen und wann genügend Informationen zur Lösung der Aufgabe vorliegen. Das Aufgabenmanagement verlangt also Prozesse der Zielsetzung, der Selbstregulation und des Monitorings (vgl. Winne & Hadwin, 1998).

Der eigentliche Leseprozess – grafisch im Vordergrund abgebildet – ist in übergeordnete Handlungsziele eingebettet und wird durch das Aufgabenmanagement auf das

Ziel hin orientiert und organisiert. Diese Leseaktivität wird in mehrere Prozesse unterteilt (vgl. Lenhard, 2013; Richter & Christmann, 2002). Grundlegend ist die Leseflüssigkeit, die in der Abbildung links angeordnet ist. Sie ist verantwortlich für ein schnelles, sicheres und akkurates Rekodieren auf Wort- und Satzebene. Darauf baut die komplexe Textverarbeitung mit mehreren Prozessen auf (s. u.), die in der Abbildung rechts angeordnet sind.

Wie oben erläutert macht der aktuelle Stand der Forschung die wichtige Rolle der *Leseflüssigkeit* deutlich (Gold, 2018; Nix, 2011). Die PISA-Studie baut auf diesem aktuellen Stand der Forschung auf, indem bei PISA 2018 erstmals Leseflüssigkeit als basaler Leseprozess zusätzlich erfasst wird. Wer auf Wort- und Satzebene nicht flüssig lesen kann, hat nicht genügend Gedächtniskapazitäten frei für die im Folgenden beschriebenen komplexeren Prozesse der Textverarbeitung. Die komplexeren Prozesse der Textverarbeitung werden in dem neuen PISA-Modell in drei *Prozesse des Lesens* untergliedert, die bisher als *Aspekte* des Lesens bezeichnet wurden. Mit dem Prozessbegriff wird zugleich die Dynamik des Lesens als zielorientierte Handlung verdeutlicht.

Texte lesen ist ein aktiver Prozess der Sinnrekonstruktion mit den drei Prozessen „*Informationen lokalisieren*“, „*Textverstehen*“ und „*Bewerten und Reflektieren*“. Mit diesen Prozessen werden diejenigen kognitiven und sprachlichen Prozesse beschrieben, die Leserinnen und Leser bei der Auseinandersetzung und Nutzung eines Textes ausführen, um ihre Aufgabe zu bewältigen beziehungsweise ihre Ziele zu erreichen. Dabei müssen nicht zwingend bei jedem Leseprozess alle drei Aufgaben bewältigt werden, sondern die Prozesse werden von geübten Leserinnen und Lesern je nach Leseaufgabe in unterschiedlicher Weise kombiniert.

- a) *Lokalisieren von Informationen* besteht aus zwei unterschiedlichen Aktivitäten. Es geht darum, relevante Texte zu suchen und auszuwählen sowie Informationen innerhalb eines Textes zu erkennen und zu extrahieren. In alltäglichen und beruflichen Situationen stellt sich immer wieder die Aufgabe, gezielt in einem oder mehreren Texten beziehungsweise Textabschnitten, Tabellen oder Abbildungen bestimmte Informationen zu finden, ohne den gesamten Text zu lesen. Das verlangt spezifische Suchprozesse wie überfliegendes (scannendes) Lesen mit wechselnder Lesegeschwindigkeit, verbunden mit der Fertigkeit, die Relevanz von Textstellen insbesondere vor dem Hintergrund der Aufgabenstellung einzuschätzen und irrelevante Textstellen im Sinne einer optimalen Ressourcennutzung wieder zu verlassen (McCrudden & Schraw, 2007). So gilt es beispielsweise Bedeutung, Relevanz, Korrektheit oder Genauigkeit einzelner Aussagen einzuschätzen; dafür stehen bisweilen nur wenige Informationen zur Verfügung, wie beispielsweise die Quelle oder die Gesamtqualität eines Textes. Aus diesem Grund stellt das kritische Einschätzen von Texten, Textabschnitten oder anderen Formaten eine zugleich herausfordernde wie wichtige Fertigkeit dar.
- b) *Textverstehen* kann als der Kern des Leseprozesses verstanden werden, weil es hier um die Rekonstruktion der wörtlichen Bedeutung, das Ziehen von Schlussfolgerungen sowie die Verknüpfung mit vorhandenem Wissen geht. Umfassende Leseaufga-

ben, wie das Lesen von Sach- und Fachtexten oder literarischen Texten, erfordern das Verstehen und Kombinieren umfangreicher Textpassagen. Einzelne Informationen aus Sätzen oder Absätzen werden nach und nach zu einem Situationsmodell (Kintsch 1998; Kintsch & Greeno, 1985) zusammengefügt und im Gedächtnis gespeichert. Das erfolgt beispielsweise zu Beginn einer Erzählung, indem die Hinweise und Informationen auf die beteiligten Figuren, die Handlungsorte, die Zeit und weitere Aspekte zu einem ersten Modell der beschriebenen Situation zusammengefügt werden. Beim Lesen von literarischen Texten muss dieses Modell immer wieder revidiert werden, weil etwa aus Gründen der Spannungserzeugung gezielt mehrdeutige Vorstellungen bei den Leserinnen und Lesern hervorgerufen werden. Sachtexte (= faktuale Texte) zeichnen sich demgegenüber eher dadurch aus, dass die gegebenen Informationen zu einem möglichst genauen Aufbau des Situationsmodells führen. Das Situationsmodell speist sich in beiden Fällen aus zwei Quellen, die miteinander verknüpft werden, nämlich zum einen aus der Textbedeutung und zum anderen aus dem Vorwissen. Dabei kommen auch Prozesse des Schlussfolgerns und Ziehens von Inferenzen zum Einsatz. So führt beispielsweise die Erwähnung auch nur einzelner Elemente eines Klassenzimmers oder einer Unterrichtsstunde wie *Tafel, Schüler, Lehrerin, melden* etc. zum Abruf der entsprechenden Situation, falls die Lesenden über das entsprechende Wissen verfügen. Man unterscheidet dabei Schema-Wissen, das eine typische Konstellation wie etwa ein *Klassenzimmer* beschreibt, und Skript-Wissen, das ein typisches Handlungsmuster wie *Unterrichten* umfasst (Nussbaumer, 1991). Eine besondere Herausforderung ergibt sich, wenn ein Situationsmodell auf der Basis von Texten mit mehreren Quellen gebildet werden muss, weil dann unterschiedliche Konzepte integriert werden müssen. Wenn die Texte darüber hinaus inkonsistente oder gar widersprüchliche Informationen enthalten, müssen die unterschiedlichen Informationen bewertet, eingeschätzt und gegebenenfalls verknüpft werden (z. B. Perfetti et al., 1999).

- c) *Bewerten und Reflektieren* geht über das Verstehen der Textbedeutung hinaus, weil es eine kritische Auseinandersetzung mit dem Text erfordert. Dabei geht es vor allem darum, die Qualität und Glaubwürdigkeit eines Textes oder mehrerer Texte anhand unterschiedlicher Merkmale einzuschätzen. Texte können an unterschiedlichen Stellen widersprüchliche Aussagen machen, sie können Sachverhalte ohne weitere Belege oder Argumente behaupten, es können wichtige Angaben fehlen, wie etwa Entstehungsdatum, Autorschaft oder Ort der Veröffentlichung, oder sie können veraltete oder sachlich falsche Informationen enthalten. Kritisches Lesen verlangt zum einen, die textinternen Merkmale auf Konsistenz zu prüfen, und zum anderen einen Vergleich mit dem eigenen Wissen. Das gilt in besonderer Weise für viele digital verfügbare Quellen. Für den Fall, dass Inkonsistenzen festgestellt werden, ist diesen nachzugehen. Ein weiterer Aspekt kritischen Lesens ist das Bewerten von Form und Inhalt. Es gilt zu prüfen, ob der gewählte Schreibstil zu den Inhalten passt, ob die Inhalte verständlich ausgedrückt werden und ob der Stil den Absichten des Textes entspricht.

Aufgrund der sich wandelnden Rahmenbedingungen werden sukzessiv neue Textarten berücksichtigt. Bei PISA 2009 kamen multiple Texte², transaktionale Texte (= veranlassende Texte), sogenannte nachrichtenbasierte Texte (neben autorbasierten) und digitale Texte hinzu. Bei PISA 2015 wurden schließlich für einen Teil der Staaten beziehungsweise bei PISA 2018 für alle Staaten die sogenannten dynamischen, nichtfixierten Texte (z. B. Webseiten) aufgenommen. Außerdem wurde die bislang genutzte Unterscheidung von gedruckten versus digitalen Texten aufgegeben, weil etwa digitale Texte im PDF-Format ohne zusätzliche Werkzeuge gedruckten Texten ähnlicher sind als sogenannte dynamische Formate mit zahlreichen integrierten Werkzeugen wie Links, Suchfunktionen oder Instrumenten zur Hervorhebung oder Kommentierung. Zur Beschreibung und Klassifizierung der Texte werden nun die folgenden vier Dimensionen unterschieden:

- a) *Textquellen*: Hierunter werden zusammengehörige Texteinheiten verstanden. Texte mit einem (oder mehreren) definierten Autorinnen beziehungsweise Autoren, mit einem definierten Veröffentlichungsdatum und einem festen Titel werden als *Texte mit einer Quelle* definiert. Ein typisches Beispiel hierfür ist ein gedrucktes Buch. Genauso gibt es sogenannte *Texte mit mehreren Quellen*, die über diese Merkmale nicht mehr in der Eindeutigkeit verfügen. Diese sind dadurch definiert, dass sie verschiedene Autorinnen oder Autoren haben, zu unterschiedlichen Zeitpunkten veröffentlicht wurden oder verschiedene Titel haben. Hierzu gehören Internet-Foren oder Frage-Antwort-Seiten auf einer Webseite, die von unterschiedlichen Schreibern zu unterschiedlichen Zeiten produziert wurden und auch andere, eingebettete Quellen enthalten können. Weil hier das Prinzip der erkennbaren Autorschaft aufgeweicht wird, unterscheidet man zwischen autorbasierten und nachrichtenbasierten Texten.
- b) *Organisation und Navigation*: Die unterschiedlichen Darstellungsmedien – vom kleinen Display eines Mobiltelefons bis zum 29-Zoll-Monitor – ermöglichen ganz unterschiedliche Formen des Navigierens in Texten, die je spezifische Fertigkeiten erfordern. Vor diesem Hintergrund werden prototypische Formen unterschieden, nämlich statische, weitgehend lineare Texte mit sehr einfachen und wenigen Navigationswerkzeugen einerseits und dynamische Texte andererseits, die mit einer Vielzahl an Instrumenten die Navigation in, aber auch die Neuorganisation von nichtlinearen Textstrukturen ermöglichen, wie sie etwa Nachrichtenseiten großer Medienanbieter enthalten.
- c) *Textformate*: Die Kategorie Textformate aus der bisherigen Rahmenkonzeption wird beibehalten und fasst unterschiedliche Formen zusammen, in denen Texte vorliegen können. Hierzu gehört seit PISA 2000 die Unterscheidung von *kontinuierlichen Texten*, organisiert in Sätzen, Abschnitten und Kapiteln, und *nichtkontinuierlichen Formen* wie Tabellen, Grafiken oder Abbildungen, die oft in kontinuierliche Texte eingebettet sind. Beide Formen finden sich in statischen ebenso wie in dynamischen Texten. Solche *gemischten Textformate* nutzen kontinuierlichen Text und nichtkonti-

2 Das Format „multiple Texte“ aus früheren PISA-Rahmenmodellen wird im Rahmenmodell für PISA 2018 durch die neue Dimension „Textquellen“ repräsentiert (OECD, 2019).

nuierliche Formen, um in einem Zusammenspiel die Inhalte auszudrücken. Sie finden sich vor allem in analogen und digitalen Magazinen, Schulbüchern und Berichten, um die unterschiedlichen Inhalte je passend darstellen zu können.

- d) *Texttypen*: Als vierte Kategorie werden die Texttypen aus den bisherigen Rahmenkonzeptionen übernommen, im Sinne spezifischer Textsorten (siehe Abschnitt 2.2.1).

Damit hat sich das Spektrum der berücksichtigten Textarten seit PISA 2000 erheblich erweitert, weil die schon eingeführten Textformate und -typen aus den vorherigen PISA-Zyklen aus Gründen der Vergleichbarkeit vollständig erhalten wurden.

2.2 Lesekompetenz testen – PISA 2018

Um die geänderten Anforderungen an die Lesekompetenz angemessen zu erfassen, setzt PISA 2018 neue Aufgabenformate ein, die sogenannten *Leseszenarien*. Leseszenarien umfassen mehrere Texte, die authentische, ziel- und aufgabenorientierte Lesesituationen simulieren. In traditionellen Lesetests geht es überwiegend darum, das Textverständnis relativ abstrakt zu erfassen – also zu erheben, ob die wesentlichen Inhalte und Intentionen eines Textes weitgehend kontextfrei verstanden werden. In den Leseszenarien geht es dagegen auch darum, ob mit den verfügbaren Texten eine realitätsnahe (Lese-)Aufgabe bewältigt werden kann. Das Leseszenario beschreibt eine Situation, in der ein vorgegebenes Ziel zu erreichen ist, indem ein oder mehrere Texte gelesen werden. Das verlangt von den Lesenden deutlich mehr, als den Text in seiner wesentlichen Bedeutung zu verstehen. So ist aus der Aufgabe das Leseziel herzuleiten. Dabei sind die erforderlichen Informationen zu bestimmen, zu finden, zusammenzuführen und gegebenenfalls in Bezug auf ihre Relevanz oder andere Aspekte zu bewerten. Die Aufgabenbewältigung kann auf unterschiedlichen Wegen geschehen, so dass der Leseprozess selber geplant und überwacht werden muss, was mehr Freiheiten bei der Bearbeitung eröffnet.

2.2.1 Situationen des Lesens und Texttypen

Den Leseszenarien liegen vier prototypische *Situationen* zugrunde, die die Breite der möglichen Leseanforderungen von Fünfzehnjährigen abdecken sollen (vgl. Naumann et al., 2010, S. 25):

- *Private Situationen* beinhalten Texte, die sich auf persönliche Interessen beziehen, also etwa auf praktische, intellektuelle oder soziale Interessen. Beispiele sind persönliche E-Mails, Briefe oder Blogbeiträge, bei denen sich Schreibende und Lesende in der Regel kennen.
- *Öffentliche Situationen* beinhalten Texte, die an eine breite Öffentlichkeit gerichtet sind. Beispiele sind Zeitungstexte, offizielle Dokumente oder Webseiten, bei denen sich Schreibende und Lesende in der Regel anonym gegenüberstehen.

- *Bildungsbezogene Situationen* beinhalten Texte, die in einem weiten Sinn Wissen vermitteln sollen. Beispiele sind Schulbücher, Aufgabenblätter oder Lernsoftware.
- *Berufsbezogene Situationen* beinhalten Texte, die in beruflichen Zusammenhängen stehen. Beispiele sind Stellenanzeigen oder Arbeitsanweisungen.

Die *Situationen* schaffen einen Rahmen für die Leseaufgaben und beschreiben typische schriftbasierte Handlungskonstellationen mit den je spezifischen Texttypen, Handelnden und Zwecken. Dabei kommen, wie in den bisherigen PISA-Erhebungen, unterschiedliche *Texttypen* zum Einsatz:

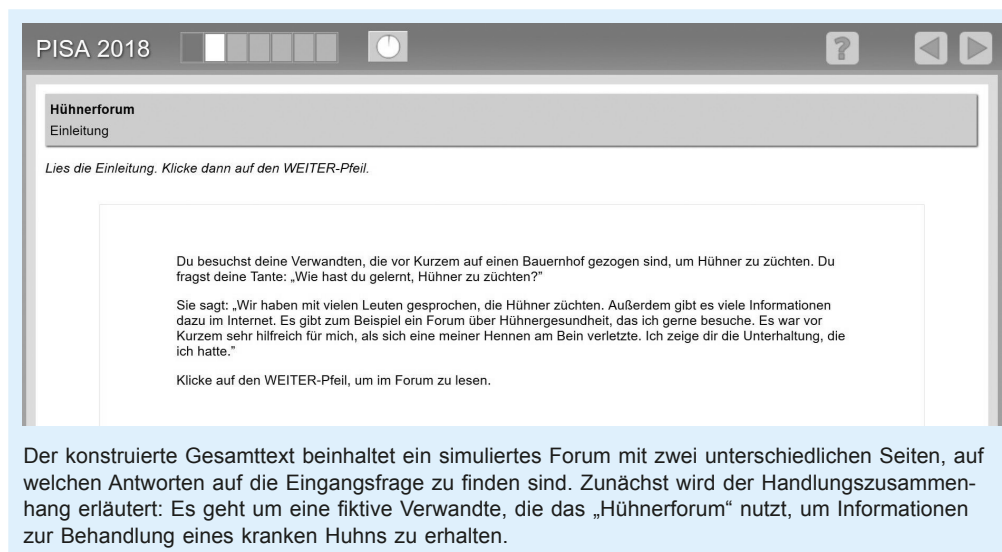
- *Beschreibende Texte* beziehen sich auf wahrnehmbare Objekte und haben insofern einen Raumbezug, zum Beispiel geografische Orte oder Bilder, wie sie in Reiseberichten vorkommen. Sie haben die Funktion, Informationen über das Aussehen der Objekte zu liefern.
- *Narrative Texte* umfassen Erzählungen und Berichte, beziehen sich auf Ereignisse und haben insofern einen Zeitbezug. Sie haben die Funktion, Informationen über Ereignisfolgen und ihre Zusammenhänge zu liefern. Beispiele hierfür sind Romane, Kurzgeschichten, Biografien und Zeitungsberichte von Ereignissen.
- *Darlegende Texte* (Expositionen) beziehen sich auf Gedanken, Konzepte oder Theorien, die ihrerseits üblicherweise in Texten stehen; dazu gehören Erklärungen, Definitionen oder Zusammenfassungen. Typische Beispiele darlegender Texte sind wissenschaftliche Aufsätze, Informationsgrafiken und Einträge in einer Online-Enzyklopädie. Sie haben die Funktion, die textuell dargelegten Gedanken, Konzepte oder Theorien sachlich und verständlich darzustellen.
- *Argumentierende Texte* beziehen sich auf strittige oder unklare Sachverhalte; dazu gehören sowohl meinungsbildende Texte als auch Fachtexte und wissenschaftliche Texte. Sie haben die Funktion, Strittiges zu klären oder Zusammenhänge zwischen unterschiedlichen Sachverhalten herzustellen, indem Gründe für bestimmte Annahmen gegeben werden. Beispiele sind Leserbriefe, Beiträge in Online-Foren oder webbasierte Buch- oder Filmkritiken.
- *Anleitende Texte* beziehen sich auf Handlungen oder Tätigkeiten, beispielsweise in Bedienungsanleitungen, Regelwerken oder Rezepten. Sie haben die Funktion, bestimmte Handlungen oder Tätigkeiten zu ermöglichen.
- *Transaktionale Texte* beziehen sich auf die Interaktion zweier oder mehrerer Personen, die sich in der Regel kennen, beispielsweise in E-Mails oder anderen persönlichen Nachrichten. Sie haben die Funktion, den Leser oder die Leserin zu einer Handlung zu veranlassen, etwa ein Treffen zu organisieren.

Diese Texttypologie unterscheidet nicht primär zwischen faktualen und fiktionalen Texten, weil es nicht das Ziel von PISA ist, Unterschiede im Verstehen dieser beiden Textklassen zu untersuchen. Fiktionale Texte werden hier zu den narrativen (erzählenden) Texten gezählt.

2.2.2 Leseaufgaben und Antwortformate

Leseaufgaben

Alle neuen Leseaufgaben der PISA-Erhebung 2018 werden in einen situativen Handlungskontext gestellt, der den Leseprozess einordnet: Auf diese Weise erhält das Lesen eine Zielorientierung und bis zu einem gewissen Grad wird damit auch die Textfunktion vor der Lektüre erkennbar. Die verschiedenen Handlungskontexte werden als „Aufgabeneinheiten“, die aus einzelnen Aufgaben bestehen, verstanden. Dies soll im Folgenden an der im Feldtest verwendeten Aufgabeneinheit „Hühnerforum“ (Abbildung 2.2) sowie an der in der Hauptstudie eingesetzten Aufgabeneinheit „Die Osterinsel“ (Abbildung 2.3) illustriert werden. Wie oben ausgeführt ist es für die Konstruktion eines Situationsmodells (Kintsch, 1998) zentral, dass der zu bearbeitende Text in einem kommunikativen Handlungskontext situiert ist. In der Aufgabeneinheit „Hühnerforum“ wird dazu eine Internet-Plattform simuliert. Es wird eine soziale Situation konstruiert, in der ein Besuch der Plattform sinnvoll ist beziehungsweise ein Informationsbedürfnis entsteht, so dass der Lektüre eine Zielorientierung zugeschrieben werden kann. Ausgewählte Seiten der Plattform werden im Folgenden erläutert.



Der konstruierte Gesamttext beinhaltet ein simuliertes Forum mit zwei unterschiedlichen Seiten, auf welchen Antworten auf die Eingangsfrage zu finden sind. Zunächst wird der Handlungszusammenhang erläutert: Es geht um eine fiktive Verwandte, die das „Hühnerforum“ nutzt, um Informationen zur Behandlung eines kranken Huhns zu erhalten.

Innerhalb eines solchen (konstruierten) Handlungskontexts werden mehrere Typen von Leseaufgaben (Aufgaben) zum Textverstehen gestellt. Die Bandbreite der Aufgaben reicht vom Auffinden von expliziten Informationen im Text über die Kohärenzbildung zwischen Textteilen – und in anderen PISA-Aufgaben – bis hin zu komplexen Synthesen und der Integration mehrerer Texte beziehungsweise von Text- und Bildteilen sowie dem Evaluieren von Rechercheergebnissen aus dem Internet.

Manche Posts in einem Forum können für das Thema relevant sein, während andere Posts es nicht sind. Klicke entweder **Ja** oder **Nein** an, um anzugeben, ob die Posts in der unten stehenden Tabelle für das Problem von Inge_88 relevant sind.

Ist der Post relevant für das Problem von Inge_88?	Ja	Nein
Post von NellieB79	<input type="radio"/>	<input type="radio"/>
Post von Moni	<input type="radio"/>	<input type="radio"/>
Post von Der_Vogelhandel	<input type="radio"/>	<input type="radio"/>
Post von Bernd	<input type="radio"/>	<input type="radio"/>
Post von Frank	<input type="radio"/>	<input type="radio"/>

In Aufgabe 3 geht es darum, die Relevanz von Textinformationen in Bezug auf das von Inge_88 explizierte Informationsziel „Wie kann ich die Schmerzen meiner Henne lindern“ zu überprüfen. Es sind im Text oder über Textabschnitte verteilte Informationen aufeinander zu beziehen, um die Absicht des Autors zu erschließen beziehungsweise um die Relevanz einer Information einzuschätzen.

Für die Lösung dieser Aufgabe muss der Leseprozess „Bewerten von Form und Inhalt“ eingesetzt werden. Der Jugendliche muss zuerst die wörtliche Bedeutung von jedem Post verstehen und im nächsten Schritt über den Inhalt reflektieren.

Es handelt sich um ein komplexes Multiple-Choice-Antwortformat.

Die richtige Antwort ist Ja, Ja, Nein, Nein, Ja, in dieser Reihenfolge.

(Kompetenzstufe II, vgl. Kapitel 3)

Wer postete die zuverlässigste Antwort auf die Frage von Inge_88?

- NellieB79
- Moni
- Der_Vogelhandel
- Frank

Begründe deine Antwort.

In Aufgabe 6 geht es um die Bewertung von Qualität und Glaubwürdigkeit der Informationen. Diese baut auf der Aufgabe 3 auf, da hier nicht nur die Relevanz, sondern der Grad der Relevanz beurteilt werden muss.

Hier handelt es sich um ein offenes Antwortformat. Die Lösung der Aufgabe wird in den Kodieranweisungen in Tabelle 2.2 dargestellt.

(Kompetenzstufe II, vgl. Kapitel 3)

Abbildung 2.2: Aufgabeneinheit „Hühnerforum“ mit Beispielaufgaben 1, 3 und 6

Tabelle 2.2: Kodieranweisungen zu der in Abbildung 2.2 dargestellten Aufgabe 6 aus der Aufgabeneinheit „Hühnerforum“

Vollständig gelöst:	
Code 1:	<p>Schreibt oder wählt NellieB79 explizit oder implizit aus UND schreibt, dass die Antwort von NellieB79 impliziert, dass Inge_88 ihren Tierarzt fragen sollte, bevor sie der Henne irgendein Medikament gibt.</p> <ul style="list-style-type: none"> • [NellieB79] Nellie sagt, dass sie zuerst ihren Tierarzt fragt. • [keine Auswahl] NellieB79 sagte Inge_88 nicht, was sie tun soll, aber sie sagte, dass sie zuerst ihren Tierarzt fragt, bevor sie Medizin gibt.
ODER	<p>Schreibt oder wählt Moni explizit oder implizit aus UND schreibt, dass Moni ihrer Henne Aspirin gegeben hat und es ihr dann besser ging.</p> <ul style="list-style-type: none"> • [Moni] Moni gab ihrer Henne Aspirin und der Henne ging es dann besser. • [Moni] Moni hat eine Henne, der es besser ging, nachdem sie ihr Aspirin gab.
ODER	<p>Schreibt oder wählt Frank explizit oder implizit aus UND schreibt, dass Frank Tierarzt/ Vogelexperte ist und sich mit der Behandlung von Vögeln auskennt.</p> <ul style="list-style-type: none"> • [Frank] Er ist Tierarzt. • [Frank] Frank ist auf Vögel spezialisiert • [Frank] Frank kennt die Dosieranweisungen für Hühner. • [keine Auswahl] Frank nennt ein Buch zu Vogelmedizin.
Nicht gelöst:	
Code 0:	<p>Andere Antworten, welche falsch, vage, irrelevant oder ungenügend sind.</p> <ul style="list-style-type: none"> • [NellieB79] Es wurde im Forum gepostet. [vage/ungenügend] • [Moni] Moni hat eine Henne. [ungenügend] • [Frank] Frank kennt sich mit Hennen aus. [ungenügend] • [Frank] Frank ist ein Spezialist. [vage: sagt nichts darüber aus was für ein Spezialist] • [Der-Vogelhandel] Sie haben viele Dinge für Hennen. [falsche Auswahl und Begründung]
Code 9:	Keine Antwort

Die folgenden Beispielaufgaben stammen aus der Aufgabeneinheit „Die Osterinsel“, bestehend aus einem Blog einer Forscherin über ihren Aufenthalt auf der Insel. In diesem Blog verweist sie auf eine Buchbesprechung sowie einen wissenschaftlichen Artikel, in denen es um kontroverse Theorien über das Verschwinden der Bäume auf den Osterinsel geht.

PISA 2018

?
◀
▶

Die Osterinsel
Frage 3 / 7

Beziehe dich auf die Buchbesprechung zu *Kollaps* auf der rechten Seite. Klicke die Antworten in der Tabelle an, um die Frage zu beantworten.

Unten stehen Aussagen aus der Buchbesprechung zu „Kollaps“. Sind diese Aussagen Tatsachen oder Meinungen? Klicke für jede Aussage entweder **Tatsache** oder **Meinung** an.

Ist die Aussage eine Tatsache oder eine Meinung?	Tatsache	Meinung
In dem Buch beschreibt der Autor den Kollaps mehrerer Zivilisationen aufgrund der Entscheidungen, die sie trafen, und deren Auswirkung auf die Umwelt.	○	○
Eines der beunruhigendsten Beispiele in dem Buch ist die Osterinsel.	○	○
Sie meißelten die Moai, die berühmten Statuen, und nutzten die verfügbaren natürlichen Ressourcen, um diese riesigen Moai an verschiedene Orte auf der Insel zu transportieren.	○	○
Als die ersten Europäer 1722 auf die Osterinsel kamen, waren die Moai immer noch dort, aber die Bäume waren verschwunden.	○	○
Das Buch ist gut geschrieben und ist es wert, von allen gelesen zu werden, denen die Umwelt am Herzen liegt.	○	○

Blog
Buchbesprechung

www.wissenschaftliche-buchbesprechung.de/Kollaps

Buchbesprechung von „Kollaps“

Jared Diamonds neues Buch „Kollaps“ ist eine eindeutige Warnung vor den Folgen der Zerstörung unserer Umwelt. In dem Buch beschreibt der Autor den Kollaps mehrerer Zivilisationen aufgrund der Entscheidungen, die sie trafen, und deren Auswirkung auf die Umwelt. Eines der beunruhigendsten Beispiele in dem Buch ist die Osterinsel.

Laut dem Autor wurde die Osterinsel irgendwann nach 700 n. Chr. von den Polynesiern besiedelt. Sie entwickelten sich zu einer florierenden Gesellschaft von vielleicht 15.000 Menschen. Sie meißelten die Moai, die berühmten Statuen, und nutzten die verfügbaren natürlichen Ressourcen, um diese riesigen Moai an verschiedene Orte auf der Insel zu transportieren. Als die ersten Europäer 1722 auf die Osterinsel kamen, waren die Moai immer noch dort, aber die Bäume waren verschwunden. Die Bevölkerung war auf wenige Tausend Menschen geschrumpft, die ums Überleben kämpften. Jared Diamond schreibt, dass die Menschen der Osterinsel das Land für landwirtschaftliche und andere Zwecke rodeten und übermäßige Jagd auf die zahlreichen Arten von See- und Landvögeln machten, die auf der Insel lebten. Er vermutet, dass die schwindenden natürlichen Ressourcen zu Stammeskriegen und dem Untergang der Gesellschaft auf der Osterinsel führten.

Die Lektion dieses wundervollen, aber beängstigenden Buches ist, dass die Menschen in der Vergangenheit die Entscheidung trafen, ihre Umwelt zu zerstören, indem sie alle Bäume fällten und Tierarten bis zur Ausrottung jagten. Optimistisch weist der Autor darauf hin, dass wir uns entscheiden können, dieselben Fehler heute **nicht** zu machen. Das Buch ist gut geschrieben und ist es wert, von allen gelesen zu werden, denen die Umwelt am Herzen liegt.

In Aufgabe 3 geht es darum, unter Bezug auf die Buchbesprechung fünf Aussagen entweder als Tatsachenbehauptung oder Meinungsäußerung zu bewerten.

Als Antwortformat dient hier das Multiple-Choice-Format, das heißt, jede Aussage muss bewertet werden.

Die richtige Auswahl lautet: Tatsache, Meinung, Tatsache, Tatsache, Meinung.

(Kompetenzstufe V, vgl. Kapitel 3)

PISA 2018
?

Die Osterinsel
Frage 6 / 7

Beziehe dich auf alle drei Quellen auf der rechten Seite, indem du auf den jeweiligen Tab klickst.

Verwende Drag & Drop, um die Ursachen und die Auswirkung, die sie gemeinsam haben, in die richtigen Zellen der Tabelle zu den Theorien zu ziehen.

Die Theorien

Ursache	Auswirkung	Befürworter der Theorie
		Jared Diamond
		Carl Lipo und Terry Hunt

Die Moai wurden alle in demselben Steinbruch gemeißelt.	Polynesische Ratten fraßen Baumsamen und folglich konnten keine neuen Bäume wachsen.	Siedler benutzten Kanus, um Polynesische Ratten auf die Osterinsel zu bringen.
Die großen Bäume sind von der Osterinsel verschwunden.	Die Bewohner der Osterinsel brauchten natürliche Ressourcen, um die Moai zu transportieren.	Menschen fällten Bäume, um Land für die Landwirtschaft und andere Zwecke zu gewinnen.

Blog
Buchbesprechung
Neues aus der Wissenschaft

www.neuesausderwissenschaft.com/Polynesische_Ratten_Osterinsel

NEUES AUS DER WISSENSCHAFT

Haben Polynesische Ratten die Bäume der Osterinsel zerstört?

Von Michael Kimball, Wissenschaftsjournalist

2005 veröffentlichte Jared Diamond „Kollaps“. In diesem Buch beschrieb er die Besiedlung von Rapa Nui (auch Osterinsel genannt) durch den Menschen.

Das Buch führte bald nach seiner Veröffentlichung zu einer großen Kontroverse. Viele Wissenschaftler stellten Diamonds Theorie über die Geschehnisse auf der Osterinsel in Frage. Sie waren sich einig, dass die riesigen Bäume bereits verschwunden waren, als die Europäer im 18. Jahrhundert zum ersten Mal auf die Insel kamen, aber Jared Diamonds Theorie über den Grund des Verschwindens stimmten sie nicht zu.

Nun haben zwei Wissenschaftler, Carl Lipo und Terry Hunt, eine neue Theorie veröffentlicht. Sie glauben, dass die Polynesische Ratte die Samen der Bäume fraß und so verhinderte, dass neue wuchsen. Die Ratte, so glauben sie, wurde entweder versehentlich oder absichtlich in den Kanus mitgebracht, mit denen die ersten menschlichen Siedler auf die Osterinsel kamen.

Studien haben gezeigt, dass sich eine Rattenpopulation alle 24 Tage verdoppeln kann. So viele Ratten brauchen natürlich auch viel Futter. Um ihre Theorie zu bekräftigen, weisen Lipo und Hunt auf die Überreste von Palmnüssen hin, die Bissspuren von Ratten aufweisen. Zwar räumen sie ein, dass auch Menschen eine Rolle bei der Zerstörung der Wälder auf der Osterinsel spielten. Sie glauben aber, dass die Polynesische Ratte in einer ganzen Reihe von Faktoren ein noch größerer Übeltäter war.

In Aufgabe 6 müssen Informationen von drei verschiedenen Quellen, nämlich „Neues aus der Wissenschaft“ (hier abgebildet), „Buchbesprechung“ (oben abgebildet bei Aufgabe 3) sowie „Blog“ (unten abgebildet bei Aufgabe 7, enthält aber keine für Aufgabe 6 relevanten Informationen) genutzt werden. Zudem müssen die Auswahlantworten nach Ursache und deren Auswirkung differenziert werden sowie die jeweiligen Ursachen dem einen Autor bzw. den beiden anderen zugeordnet werden.

Die Theorien

Ursache	Auswirkung	Befürworter der Theorie
Menschen fällten Bäume, um Land für die Landwirtschaft und andere Zwecke zu gewinnen.	Die großen Bäume sind von der Osterinsel verschwunden.	Jared Diamond
Polynesische Ratten fraßen Baumsamen und folglich konnten keine neuen Bäume wachsen.		Carl Lipo und Terry Hunt
Die Moai wurden alle in demselben Steinbruch gemeißelt.	Die Bewohner der Osterinsel brauchten natürliche Ressourcen, um die Moai zu transportieren.	Siedler benutzten Kanus, um Polynesische Ratten auf die Osterinsel zu bringen.

Für die Lösung müssen Aussagen in die jeweilige Position der Tabelle per Drag-and-Drop gebracht werden. Voraussetzung für die richtige Lösung ist das Verstehen der skizzierten Tabelle.

Die richtige Lösung ist links abgebildet.

(Kompetenzstufe V, vgl. Kapitel 3)

In Aufgabe 7 geht es darum, unter Bezug auf alle drei Quellen (Blog, Abb. hier, Buchbesprechung, Abb. bei Aufgabe 3, wissenschaftlicher Artikel, Abb. bei Aufgabe 6) eine eigene Antwort auf die Frage zu formulieren, warum die großen Bäume verschwunden sind.

Hier handelt es sich um ein offenes Antwortformat. Erwartet wird, dass mindestens einer der folgenden Aspekte benannt wird:

1. Die Menschen fällten oder benutzten die Bäume, um die Moai zu transportieren und/oder rodeten das Land für landwirtschaftliche Zwecke.
2. Die Ratten fraßen die Baumsamen und folglich konnten keine neuen Bäume wachsen.
3. Es ist nicht möglich genau zu erklären, was mit den großen Bäumen passierte, bevor weitere Forschung dazu durchgeführt wird.

(Kompetenzstufe IV, vgl. Kapitel 3)

Abbildung 2.3: Aufgabeneinheit „Die Osterinsel“ mit den Beispielaufgaben 3, 6 und 7

Jede Aufgabe ist so konstruiert, dass sie einen oder mehrere der oben beschriebenen Leseprozesse des Konstrukts „Lesen“ adressiert.

Die Aufgaben, welche bei PISA verwendet werden, unterscheiden sich in den situativen Handlungskontexten, im Textformat, in den Aufgabenvoraussetzungen sowie in der Schwierigkeit. Die Schwierigkeit hängt von der Länge, der Struktur und der Komplexität des Textes ab. Außerdem beeinflussen die Komplexität und Differenziertheit der mentalen Prozesse, die Menge der Information, welche die Leserin oder der Leser zur Bearbeitung braucht, sowie die Spezifität des Wissens, welcher die Leserin oder der Leser von innerhalb und außerhalb des Textes benötigt, die Schwierigkeit einer Aufgabe (OECD, 2019).

Durch die computerbasierte Erhebung der Lesekompetenz kann das gesamte Textkonstrukt in abgestuften Sequenzen präsentiert werden (vgl. „Hühnerforum“). So kann beispielsweise zu einer Einheit oder einem Textverbund als Ergebnis einer Internetrecherche die erste Aufgabe darin bestehen, dass die Schülerinnen und Schüler zuerst ein spezifisches Dokument heraussuchen müssen. Als zweites kann den Schülerinnen und Schülern eine Frage zu einer explizit im Text stehenden Information gestellt werden. Eine folgende Aufgabe besteht dann darin, dass die Aussagen in einem ersten Text mit denjenigen in einem zweiten Text verglichen werden müssen. Die einzelnen Aufgaben sind dabei stets so konstruiert, dass eine folgende Aufgabe zwar unabhängig von einer vorhergehenden gelöst werden kann, aber dennoch in ein globales Textverstehen passt.

Antwortformate

Wie in den Beispielaufgaben (vgl. Abbildungen 2.2 und 2.3) ersichtlich ist, wurden in PISA 2018 wie bisher geschlossene (Multiple-Choice-Aufgaben) und offene Formate (Kurzantworten) eingesetzt. Neu sind durch die computerbasierte Testung offene Formate, bei denen die Testpersonen in Interaktion mit dem digitalen Text treten. Insbesondere sind Markieren und Drag-and-Drop-Formate neue Instrumente, um Lesekompetenzen zu erfassen (siehe z. B. Beispielaufgabe 6 in Abbildung 2.3).

2.3 Fazit

Die gesellschaftlichen und technologischen Entwicklungen der letzten Jahre haben zu einem Wandel an die Anforderungen der Lesekompetenz geführt. Die Angebote des Internets mit einer Vielzahl von Quellen und Formaten sowie die Entwicklung leistungsfähiger Tablets und Smartphones ermöglichen die breite Nutzung digitaler Texte und Informationen in verschiedenen Repräsentationsmodi in praktisch allen Lebensbereichen.

Die Ausweitung des Lesens vom Gedruckten hin zur Rezeption von digital dargebotenen Informationen im Internet geht mit einer weiteren, auch in PISA 2018 abgebildeten Veränderung einher: Während früher in der Schule oft nur mit einem Text als Informationsquelle gearbeitet wurde, sind es heute oft mehrere Texte mit unterschiedlichem Format und verschiedenen Quellen, die bisweilen redundante, sich ergänzende, aber auch sich widersprechende Informationen enthalten.

Neben dem Wandel von Wissensquellen und Textarten ist seit PISA 2000 national wie international sehr viel zum Lesen und zur Leseförderung geforscht worden, sodass wir heute über ein verbessertes, weil differenzierteres Verständnis der Lesekompetenz, der Leseprozesse und der Leseförderung verfügen. Ein für den Unterricht zentraler Erkenntnisfortschritt ist die Unterscheidung basaler und komplexer Teilfertigkeiten (Gold, 2018; Rosebrock et al., 2017). Zudem wird Lesen stärker als bisher als eine zielgerichtete Handlung aufgefasst, deren Gelingen und Ergebnis im Kern von drei Faktoren abhängen (vgl. Snow and the RAND Corporation, 2002), nämlich von den Merkmalen der Lesenden, von Eigenschaften des Textes sowie der Leseaufgabe selbst. Dabei spielt insbesondere die kommunikative Situierung eines Textes mit einem (fiktiven) Informationsziel bei der Konstruktion von Testaufgaben eine wichtige Rolle, wie im oben dargestellten Beispiel „Hühnerforum“ illustriert.

Dieser neue Blick auf das Konstrukt „Lesekompetenz“ in PISA 2018 passt besser als der Lesekompetenz-Begriff der früheren PISA-Rahmenkonzeptionen zu dem in der Deutschdidaktik breiter angelegten Leseverständnis, das immer auch sozial-kulturelle Dimensionen umfasst (vgl. Gold, 2018; Groeben, 2002; Hurrelmann, 2002). Dabei gehört neben den kognitiven und kontextuellen Bedingungen zu einem lesedidaktischen Kompetenzbegriff auch die Einbettung des Lesens beziehungsweise der Leseerfahrung in einen sozialen Kontext. Lesen ist in diesem Verständnis immer Teil einer literalen beziehungsweise kulturellen Praxis. Lesekompetenz ist daher Teil des sozialen Austausches über Autoren- und Leserpositionen beziehungsweise Textinhalte und -interpretationen, eingebettet in die sogenannte Anschlusskommunikation.

Literatur

- Becker-Mrotzek, M. (2003). Mündlichkeit – Schriftlichkeit – Neue Medien. In U. Bredel, H. Günther, P. Klotz, J. Ossner & G. Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache - ein Handbuch. 1. Teilband* (S. 69–89). Paderborn: Schoeningh (UTB).
- Becker-Mrotzek, M., Schramm, K., Thürmann, E. & Vollmer, H. (2013). *Sprache im Fach. Sprachlichkeit und fachliches Lernen*. Münster: Waxmann.
- Bernhardt, M. & Conrad, F. (2018). Sprachsensibler Geschichtsunterricht. Sprachliche Bildung als Aufgabe des Fachs Geschichte. *Geschichte lernen*, (182), 2–9.
- Bräten, I. & Braasch, J. L. G. (2017). Key issues in research on students' critical reading and learning in the 21st century information society. In C. Ng & B. Bartlett (Hrsg.), *Improving reading and reading engagement in the 21st century* (S. 77–98). Singapur: Springer. https://doi.org/10.1007/978-981-10-4331-4_4
- Bräten, I., Stadtler, M. & Salmerón, L. (2018). The role of sourcing in discourse comprehension. In M. F. Schober, D. N. Rapp & M. A. Britt (Hrsg.), *The Routledge handbook of discourse processes* (2. Aufl., S. 141–166). New York, NY: Routledge. <https://doi.org/10.4324/9781315687384-10>
- Daroczy, G., Wolska, M., Meurers, W. D. & Nuerk, H. C. (2015). Word problems: A review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, 6, 348. <https://doi.org/10.3389/fpsyg.2015.00348>
- Delgado, P., Vargas, C., Ackerman, R. & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, 25, 23–38. <https://doi.org/10.1016/j.edurev.2018.09.003>
- Gold, A. (2018). *Lesen kann man lernen – Wie man Lesekompetenz fördern kann* (3. Aufl.). Göttingen: Vandenhoeck & Ruprecht. <https://doi.org/10.13109/9783666310638>
- Groeben, N. (2002). Zur konzeptionellen Struktur des Konstrukts „Lesekompetenz“. In N. Groeben & B. Hurrelmann (Hrsg.), *Lesekompetenz. Bedingungen, Dimensionen, Funktionen* (S. 11–21). Weinheim, München: Juventa.
- Handro, S. & Kilimann, V. (2019): Textverstehen im Geschichtsunterricht. Ein Projekt zur Professionalisierung historischer Leseförderung (ProLeGu). In M. Bönnighausen (Hrsg.), *Praxisprojekte in Kooperationsschulen. Fachdidaktische Modellierung von Lehrkonzepten zur Förderung strategiebasierten Textverstehens in den Fächern Deutsch, Geographie, Geschichte und Mathematik* (S. 165–222). Münster: WTM-Verlag.
- Hahnel, C., Goldhammer, F., Naumann, J. & Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Computers in Human Behavior*, 55, 486–500. <https://doi.org/10.1016/j.chb.2015.09.042>
- Härtig, H. & Kohnen, N. (2017). Die Rolle der Termini beim Lernen mit Physikschulbüchern. In B. Ahrenholz, E. Hövelbrinks & C. Schmellentin (Hrsg.), *Fachunterricht und Sprache in schulischen Lehr-/Lernprozessen* (S. 23–91). Tübingen: Narr.
- Hurrelmann, B. (2002). Zur historischen und kulturellen Relativität des „gesellschaftlich handlungsfähigen Subjekts“ als normativer Rahmenidee für Medienkompetenz. In N. Groeben & B. Hurrelmann (Hrsg.), *Medienkompetenz. Voraussetzungen, Dimensionen, Funktionen* (S. 160–197). Weinheim, München: Juventa.
- Kemp, N., Wood, C. & Waldron, S. (2014). Do i know its wrong: Children's and adults' use of unconventional grammar in text messaging. *Reading and Writing*, 27, 1585–1602. <https://doi.org/10.1007/s11145-014-9508-1>

- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W. & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, 92, 109-129. <https://doi.org/10.1037/0033-295X.92.1.109>
- KMK (2004) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004). *Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss*. Beschluss vom 4.12.2003. Neuwied: Luchterhand.
- KMK (2014) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2014). *Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife*. Beschluss vom 18.10.2012. Köln: Wolters Kluwer.
- Kuhn, M. R. & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95, 3-21. <https://doi.org/10.1037/0022-0663.95.1.3>
- Leiss, D., Hagen, M., Neumann, A. & Schwippert, K. (Hrsg.). (2017). *Mathematik und Sprache. Empirischer Forschungsstand und unterrichtliche Herausforderungen*. Münster: Waxmann.
- Leiss, D., Plath, J. & Schwippert, K. (2019). Language and mathematics – key factors influencing the comprehension process in reality-based tasks. *Mathematical Thinking and Learning*, 21, 131–153. <https://doi.org/10.1080/10986065.2019.1570835>
- Lenhard, W. (2013). *Leseverständnis und Lesekompetenz. Grundlagen – Diagnostik – Förderung*. Stuttgart: Kohlhammer.
- McCrudden, M. T. & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, 19, 113-139. <https://doi.org/10.1007/s10648-006-9010-7>
- Medienpädagogischer Forschungsverbund Südwest (Hrsg.). (2018). *JIM-Studie 2018. Jugend, Information, Medien. Basisuntersuchung zum Medienumgang 12- bis 19-Jähriger*. Stuttgart. Verfügbar unter https://www.mpfs.de/fileadmin/files/Studien/JIM/2018/Studie/JIM_2018_Gesamt.pdf
- Naumann, J. (2013). Online-Leseengagement: Vorhersage von Navigationsverhalten und Textverstehen bei Online-Texten. In C. Rosebrock & A. Bertschi-Kaufmann (Hrsg.), *Literalität erfassen: bildungspolitisch, kulturell, individuell* (S. 188–202). Weinheim: Beltz Juventa.
- 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, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23–71). Münster: Waxmann.
- Nix, D. (2011). *Förderung der Leseflüssigkeit. Theoretische Fundierung und empirische Überprüfung eines kooperativen Lautlese-Verfahrens im Deutschunterricht*. Weinheim: Beltz Juventa.
- Nussbaumer, M. (1991). *Was Texte sind und wie sie sein sollen. Ansätze zu einer sprachwissenschaftlichen Begründung eines Kriterienrasters zur Beurteilung von schriftlichen Schülertexten*. Reihe Germanistische Linguistik, Bd. 119. Tübingen: Niemeyer. <https://doi.org/10.1515/9783111632308>
- OECD. (2019). PISA 2018 reading framework. In OECD (Hrsg.), *PISA 2018 assessment and analytical framework* (S. 21–71). Paris: OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- Perfetti, C. A., Rouet, J.-F. & Britt, M. A. (1999). Toward a theory of documents representation. In H. van Oostendorp & S. R. Goldman (Hrsg.), *The construction of mental representations during reading* (S. 99–122). Mahwah, NJ: Lawrence Erlbaum Associates.

- Richter, T. & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Hrsg.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen* (S. 25–58). Weinheim: Juventa.
- Richter, T. & Müller, B. (2017). Entwicklung hierarchieniedriger Leseprozesse. In M. Philipp (Hrsg.), *Handbuch Schriftspracherwerb und weiterführendes Lesen und Schreiben* (S. 51–66). Weinheim: Beltz Juventa.
- Rosebrock, C. (2016). Anforderungen von Sach- und Informationstexten, Anforderungen literarischer Texte. In A. Bertschi-Kaufmann & T. Graber (Hrsg.), *Lesekompetenz, Leseleistung, Leseförderung – Grundlagen, Modelle und Materialien* (S. 58–75). Zug: Klett und Balmer.
- Rosebrock, C. & Nix, D. (2006). Forschungsüberblick: Leseflüssigkeit (Fluency) in der amerikanischen Leseforschung und -didaktik. *Didaktik Deutsch*, 20, 90–112.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2017). *Leseflüssigkeit fördern. Lautleseverfahren für die Primar- und Sekundarstufe* (5. Aufl.). Seelze: Klett Kallmeyer.
- Schindler, J. & Richter, T. (2018). Reading comprehension: Individual differences, disorders, and underlying cognitive processes. In A. Bar-On & D. Ravid (Hrsg.), *Handbook of communication disorders: theoretical, empirical, and applied linguistic perspectives* (S. 503–524). Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9781614514909-025>
- Schmellentin, C., Dittmar, M., Gilg, E. & Schneider, H. (2017). Sprachliche Anforderungen in Biologielehrmitteln. In B. Ahrenholz, E. Hövelbrinks & C. Schmellentin (Hrsg.), *Fachunterricht und Sprache in schulischen Lehr-/Lernprozessen* (S. 23–91). Tübingen: Narr.
- Schnotz, W. (2002). Wissenserwerb mit Texten, Bildern und Diagrammen. In L. Issing & P. Klima (Hrsg.), *Information und Lernen mit Multimedia und Internet* (S. 65–81). Weinheim: Beltz.
- Schnotz, W. & Dutke, S. (2004). Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat, (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (S. 61–99). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-81031-1_4
- Snow and the Rand Corporation. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Rand Corporation Report. Verfügbar unter https://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1465.pdf
- UNESCO (2005) = United Nations Educational, Scientific and Cultural Organization. (2005). *Education for All Global Monitoring Report: Literacy for life*. Paris: UNESCO Publishing.
- Winne, P. H. & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Metacognition in educational theory and practice* (S. 277–304). New York, NY: Routledge.

3 Lesekompetenz in PISA 2018 – Ergebnisse in einer digitalen Welt

Mirjam Weis, Anastasia Doroganova, Carolin Hahnel,
Michael Becker-Mrotzek, Thomas Lindauer, Cordula Artelt
& Kristina Reiss

Die mittlere Lesekompetenz der Fünfzehnjährigen in Deutschland liegt bei der PISA-Studie 2018 signifikant über dem OECD-Durchschnitt. Der Mittelwert der Lesekompetenz der Fünfzehnjährigen in Deutschland unterscheidet sich 2018 nicht signifikant von den deutschen Durchschnittswerten bei den PISA-Studien 2009 und 2015, ist aber deutlich höher als bei PISA 2000. Die Gruppe der besonders lese schwachen Schülerinnen und Schüler ist im Vergleich zu anderen lesestarken Staaten verhältnismäßig groß und entspricht in etwa dem OECD-Durchschnitt. Die Gruppe der besonders lesestarken Jugendlichen ist in Deutschland hingegen signifikant größer als im Durchschnitt der OECD-Staaten. Die differenzierten Analysen zeigen, dass ein relativ hoher Anteil der Gymnasiastinnen und Gymnasiasten bei PISA 2018 hochkompetent im Lesen ist und diese Gruppe seit PISA 2009 deutlich gewachsen ist. Allerdings gibt es insbesondere an nicht gymnasialen Schularten einen hohen Anteil lese schwacher Jugendlicher, der sowohl im Vergleich zu 2009 als auch im Vergleich zu 2015 bedeutsam angestiegen ist.

3.1 Lesekompetenz in PISA 2018

Lesekompetenz ist nach wie vor eine Schlüsselqualifikation für den Kompetenzerwerb in allen Schulfächern und im beruflichen Bereich sowie ganz allgemein für eine aktive gesellschaftliche Teilhabe. Es stehen immer größer werdende Mengen an Texten und Informationsmaterialien zur Verfügung, die insbesondere über das Internet jederzeit und leicht erreichbar sind und für unterschiedliche Zwecke genutzt werden. Die Digitalisierung ist dabei mit einem Wandel der Anforderungen an Leserinnen und Leser verbunden, da etwa neue Textformate die Gewohnheiten beim Lesen verändern (vgl. Kapitel 2).

Ziel der PISA-Studie ist es, Kompetenzen mit Blick auf eine funktionale Grundbildung zu erfassen – das heißt, mit einem Fokus auf ihre Nutzung im jetzigen und zukünftigen Leben innerhalb und außerhalb von Bildungskontexten (vgl. Kapitel 1). Diese grundlegende Idee von PISA ist seit Beginn unverändert, dennoch gilt es, aktuelle Entwicklungen zu berücksichtigen. In PISA geschieht das regelmäßig für die Bereiche, die als Hauptdomäne gemessen werden. Da Lesen bei PISA 2018 die Hauptdomäne ist,

wurde die Rahmenkonzeption und damit die Definition von Lesekompetenz für diese Studie angepasst.

In der PISA-Rahmenkonzeption 2018 wird Lesekompetenz als die Fähigkeit verstanden „*Texte zu verstehen, zu nutzen, zu bewerten und über sie zu reflektieren sowie bereit zu sein, sich mit ihnen auseinanderzusetzen, um eigene Ziele zu erreichen, eigenes Wissen und Potenzial zu entwickeln und an der Gesellschaft teilzuhaben*“ (OECD, 2019a). Im Vergleich zur Definition von Lesekompetenz bei PISA 2015 (vgl. Weis et al., 2016, S. 252) gab es somit zwei wichtige Erweiterungen. Zum einen wurde im Rahmenmodell für 2018 „*das Bewerten*“ von Textaussagen eingefügt. Damit ist die Fähigkeit gemeint, den Wahrheitsgehalt von Argumenten im Text, den Standpunkt des Autors oder der Autorin sowie die Relevanz des Textes für die eigenen (Lese-)Ziele einzuschätzen und abzuwägen (OECD, 2019a). Mit diesem Zusatz wird unter anderem dem Umstand Rechnung getragen, dass es im digitalen Zeitalter immer wichtiger wird, die Glaubwürdigkeit von Quellen einschätzen zu können, da Online-Medien oft weniger Qualitätskontrollen unterliegen als gedruckte Medien (vgl. Kapitel 2). Im PISA-Lesekompetenztest 2018 wird dem Rechnung getragen, indem durch den Aspekt *Bewerten und Reflektieren* das Einschätzen der Qualität und Glaubwürdigkeit der Textaussagen sowie das Entdecken und Bewältigen von Widersprüchen einbezogen wird.

Zum anderen wurde die Definition von *Texten* erweitert, indem nicht mehr ausschließlich „*geschriebene Texte*“ eingeschlossen werden wie noch 2015. Die neue Definition bei PISA 2018 umfasst ein breites Spektrum an *Texten*, zu denen sowohl kontinuierliche Texte als auch nichtkontinuierliche Formen wie Grafiken, Tabellen und Bilder zählen, die geschriebene Sprache enthalten. Damit werden sowohl handschriftliche und gedruckte als auch bildschirmbasierte Texte erfasst. Letztere schließen auch die sogenannten dynamischen Texte ein, bei denen der Umfang schwieriger abzuschätzen ist, da verschiedene Teile über Hypertext-Links miteinander verbunden sind und so beispielsweise mehrere zusammengefasste Texte als Ergebnisse von Suchmaschinenabfragen zur Verfügung stehen (OECD, 2019a). Die Unterscheidung zwischen dynamischen und statischen (fixierten) Texten schlägt sich unter anderem in der neuen Unterscheidung von Texten mit *einer* versus *mehreren Quellen* im PISA-Lesekompetenztest 2018 sowie in der zusätzlichen Dimension der *Organisation und Navigation* nieder (vgl. Kapitel 2 sowie 3.1.1).

Die beiden Änderungen in der Definition verdeutlichen, wie technische Entwicklungen und ihr Einfluss auf die Lesekompetenz berücksichtigt wurden. Zudem wurde bei PISA 2018 erstmals die Leseflüssigkeit erhoben (vgl. Kapitel 2). Dennoch bleibt das grundlegende Verständnis von Lesekompetenz als *Reading Literacy*, im Sinne einer funktionalen Grundbildung, das bereits bei PISA 2000 konzeptualisiert wurde, erhalten (siehe z. B. Artelt, Stanat, Schneider & Schiefele, 2001; Naumann, Artelt, Schneider & Stanat, 2010). Der Fokus bleibt somit auf Kompetenzen gerichtet, die es ermöglichen, Wissen und Fertigkeiten im Alltag sowie im Berufs- und Bildungskontext zu nutzen (vgl. Kapitel 1). Damit integriert das Lesekonzept bei PISA 2018 Lesen im traditionellen Sinne mit neuen Formen des Lesens, die sich seit den letzten Jahren entwickelten und

derzeit immer häufiger werden (OECD, 2019a). Da wesentliche Aspekte der bisherigen Definition von Lesekompetenz bei PISA 2018 erhalten blieben, ist auf diese Weise ein Vergleich mit den vorherigen PISA-Studien möglich.

3.1.1 Der Lesekompetenztest in PISA 2018 und Veränderungen seit 2009

Neben der Rahmenkonzeption wurde auch der Lesekompetenztest entsprechend der aktuellen Leseforschung sowie gemäß den gesellschaftlichen und technologischen Entwicklungen angepasst (vgl. Kapitel 2). Eine der Neuerungen bei PISA 2018 – im Vergleich zu PISA 2009, als Lesen das letzte Mal Hauptdomäne war – betrifft die Situierung der Texte in einen Handlungskontext: Die neuen Leseaufgaben sind in *Leseszenarien* eingebettet, welche das Lesematerial in Form von Situationen mit realitätsnahen und zielgerichteten Aufgaben präsentieren (Kapitel 2; OECD, 2019a). Zudem ist neu, dass Texte mit *einer* versus *mehreren Quellen* unterschieden werden, und entsprechende Aufgaben entwickelt wurden. Damit wird unter anderem Internet-Foren und Blogs Rechnung getragen, welche oftmals verschiedene Autorinnen oder Autoren haben, zu unterschiedlichen Zeitpunkten veröffentlicht werden oder mehrere Titel aufweisen (vgl. Kapitel 2; OECD, 2019a). Zusätzlich berücksichtigt PISA 2018, dass beim Lesen digitaler Texte auch Fertigkeiten der Navigation eine Rolle spielen können (Hahnel, Goldhammer, Naumann & Kröhne, 2016; Naumann, 2013). Daher wurde die zusätzliche Dimension der *Organisation und Navigation* eingeführt, die zwischen statischen und dynamischen Texten unterscheidet, und es wurden neue Aufgaben mit dynamischen Texten entwickelt, die Navigationsverhalten erfordern (OECD, 2019a). In der PISA-Studie 2018 wird außerdem erstmalig *Leseflüssigkeit* erfasst und einbezogen. Hierbei handelt es sich um basale Prozesse des akkuraten, automatischen und sinngestaltenden Lesens, die den komplexeren Leseprozessen zugrunde liegen (vgl. Kapitel 2; OECD, 2019a).

Weitere wichtige Neuerungen bezüglich der Erfassung der Lesekompetenz sind die *computerbasierte Testung* seit PISA 2015 sowie das *mehrstufig adaptive Testdesign* bei PISA 2018 (vgl. Kapitel 10; Mang et al., 2019). *Adaptives Testen* bedeutet, dass die Schülerin beziehungsweise der Schüler basierend auf der richtigen oder falschen Beantwortung einer Aufgabe im Folgenden eine schwierigere oder einfachere Aufgabe erhält (vgl. Kapitel 10 sowie Mang et al., 2019). Während das bisherige Testdesign von PISA 2000 bis 2015 („rotiertes booklet Design“, vgl. Kapitel 10) sehr gut die Kompetenzen von Schülerinnen und Schülern im mittleren Bereich differenzierte, ermöglicht das adaptive Testen bei PISA 2018 zusätzlich eine genauere Differenzierung der Lesekompetenz von Schülerinnen und Schülern mit sehr niedrigen oder sehr hohen Kompetenzen (OECD, 2019b).

Um eine breite Erfassung der Lesekompetenz zu ermöglichen und die Konsistenz mit den vergangenen PISA-Zyklen zu erhalten, wurden mehrere Dimensionen der PISA-Rahmenkonzeption der Lesekompetenz aufrechterhalten. So wurden die vier

prototypischen Situationen, die als Kontexte des Lesens verstanden werden, beibehalten. Entsprechend sind die Texte beziehungsweise Aufgaben in *private*, *öffentliche*, *bildungsbezogene* sowie *berufsbezogene* Situationen eingebettet. Ebenfalls gleich geblieben sind Textformate (*kontinuierliche*, *nicht-kontinuierliche* und *gemischte* Texte) und Texttypen (*beschreibende*, *narrative*, *darlegende*, *argumentierende*, *anleitende* und *transaktionale* Texte; zur Beschreibung der Situationen, Textformate und Texttypen vgl. Kapitel 2). Zusätzlich wird die Vergleichbarkeit zwischen den einzelnen PISA-Erhebungen gewährleistet, indem sogenannte Trendaufgaben einbezogen werden. Dabei handelt es sich um Aufgaben, die aus vorherigen PISA-Erhebungen übernommen werden.

3.1.2 Lesen in digitalen Medien

Die für PISA 2018 neu entwickelten Leseaufgaben greifen zumeist Textformate und Kommunikationsmöglichkeiten auf, die aufgrund der technologischen Entwicklungen in den letzten Jahren entstanden sind und Eingang in den Lesealltag von Schülerinnen und Schülern gefunden haben (vgl. Kapitel 2). Um dieser veränderten Realität gerecht zu werden, umfassen die PISA-Leseaufgaben als wesentliche Neuerung Texte mit mehreren Quellen und dynamische Texte, die von Leserinnen und Lesern Navigationsentscheidungen erfordern.

Lesen in digitalen Medien erfordert insbesondere, sich aktiv durch digitale Texte zu navigieren, dabei gefundene Textinformationen entsprechend ihrer Quelle zu bewerten und sie mit bisher erlesenen Informationen zu verbinden (Salmerón, Strømsø, Kammerer, Stadler & van den Broek, 2018). Aufschluss hierzu geben Studien, die bei der Untersuchung des Lesens digitaler Medien auch das Lesen von Printprodukten von Schülerinnen und Schülern berücksichtigen und somit Merkmale und Verhaltensweisen identifizieren, die für das digitale Lesen – über Fähigkeiten im linearen Lesen hinaus – relevant sind (z.B. Cheung, Mak & Sit, 2013; Gil-Flores, Torres-Gordillo & Perera-Rodríguez, 2012; Hahnel, Goldhammer, Kröhne & Naumann, 2017; Hahnel et al., 2016; Lee & Wu, 2012; Lim & Jung, 2019; Naumann, 2015; Naumann & Salmerón, 2016; Wu & Peng, 2017). So zeigen Hahnel et al. (2016), dass neben der PISA-Lesekompetenz sowohl grundlegende Fähigkeiten im Umgang mit Computern als auch Fähigkeiten im Einschätzen der Nützlichkeit und Vertrauenswürdigkeit von Online-Informationen für das erfolgreiche Lesen digitaler Texte prädiktiv sind. Der Effekt der basalen Computerfähigkeiten von Schülerinnen und Schülern auf ihre Kompetenz beim Lesen digitaler Texte ließ sich durch ihre Auswahl von Webseiten in digitalen Texten erklären, was bedeutet, dass Jugendliche, die im Umgang mit Maus und Tastatur geübt sind und wissen, wie Informationen digital verwaltet werden, auch eher in der Lage sind, Webseiten mit wichtigen Informationen zu finden. Es sind daher nicht direkt gute Fähigkeiten im Umgang mit Computern, die das Verständnis von digitalen Texten fördern. Vielmehr bedingen diese Fähigkeiten einen sicheren Umgang mit digitalen Medien, der Einfluss auf die Lesekompetenz hat.

Beim Lesen digitaler Texte treffen Leserinnen und Leser häufig auf Informationen, die für eine bestimmte Aufgabe unnötig, nicht hilfreich oder sogar ablenkend sein können. Dem strategischen Auswählen von Seiten in einem digitalen Text – häufig als aufgabenorientierte Navigation bezeichnet – kommt daher eine besondere Rolle zu (Lawless & Schrader, 2008). Wie in Kapitel 2 erläutert, sind sowohl die allgemeine Lesekompetenz als auch das Navigationsverhalten prädiktiv für das Textverstehen digitaler Texte (Hahnel et al., 2016; Naumann, 2013).

Zusammengefasst stellen die neuen PISA-Leseraufgaben eine sinnvolle und wichtige Erweiterung zur zeitgemäßen Erfassung der Lesekompetenz in PISA dar, erfordern jedoch auch eine ergänzende Reflexion der Ergebnisse.

3.1.3 Prozesse des Lesens und Textquellen

Der Leseprozess ist eine aktive zielorientierte Handlung und kann in mehrere Teilprozesse unterteilt werden (vgl. Kapitel 2). Aus der Rahmenkonzeption in PISA 2018 sowie dem angepassten Test ergeben sich drei Teilskalen der Lesekompetenz, für die Ergebnisse berichtet werden (vgl. OECD, 2019a; zur inhaltlichen Beschreibung der Prozesse vgl. Kapitel 2):

- 1) *Lokalisieren von Informationen*
- 2) *Textverstehen*
- 3) *Bewerten und Reflektieren*

Für alle drei Teilskalen wurden sowohl Aufgaben mit einer Textquelle als auch Aufgaben mit mehreren Textquellen eingesetzt. Daraus ergeben sich die zwei zusätzlichen Teilskalen der *Textquellen* (= *Texte mit einer und mehreren Quellen*). Zusätzlich wurde die *Leseflüssigkeit* mittels Satzverifikationsaufgaben mit einer Zeitbegrenzung von drei Minuten gemessen. Die Leseflüssigkeit stellt allerdings keine eigene Teilskala dar, sondern fließt in die Gesamtskala der Lesekompetenz ein. Die Anzahl richtig gelöster Aufgaben der Leseflüssigkeit ergänzt die Informationen der anderen Leseraufgaben zur Einschätzung der Gesamtskala Lesekompetenz, insbesondere auf den unteren Kompetenzstufen (II oder darunter; siehe Abschnitt 3.1.4).

3.1.4 Kompetenzstufen

Die Lesekompetenzen der Schülerinnen und Schüler können bei PISA anhand der Inhalte und Anforderungen der Testaufgaben (Items) interpretiert werden. Personen und Items werden dazu Skalenwerte entsprechend der Kompetenz beziehungsweise der Itemschwierigkeit zugeordnet. Gleiche Skalenwerte entsprechen gleicher Kompetenz oder gleicher Schwierigkeit. Dass Kompetenzen der Jugendlichen und Itemschwierigkeiten auf *einer* Skala abgebildet werden können, basiert auf einer Skalierung auf Grund-

lage der Item-Response-Theorie (2-Parameter-Logistisches-Modell; vgl. Heine et al., 2016). Bei der ersten PISA-Studie im Jahr 2000 wurde der OECD-Mittelwert auf 500 und die Standardabweichung im OECD-Mittel auf 100 festgelegt. Allerdings liegt der OECD-Mittelwert in den späteren Erhebungen nicht mehr genau bei 500, sondern hat sich etwa durch ein anderes Lösungsverhalten der Teilnehmerinnen und Teilnehmer oder den Zuwachs an OECD-Staaten verändert.

Zur genaueren Beschreibung der Schwierigkeiten von Aufgaben sowie der Kompetenzen der Schülerinnen und Schüler werden seit PISA 2000 Kompetenzstufen definiert. In der PISA-Studie 2018 werden acht Kompetenzstufen für die Lesekompetenz unterschieden. Damit wird die Skala der Lesekompetenz im Vergleich zu den vorherigen PISA-Zyklen weiter ausdifferenziert. Im ersten PISA-Zyklus (PISA 2000 bis 2006) waren es für die Lesekompetenz fünf Stufen, im zweiten Zyklus (PISA 2009 bis 2015) sieben Stufen. Im Jahr 2018 wurde nun die Kompetenzstufe Ic zusätzlich eingeführt, um im unteren Kompetenzbereich noch besser differenzieren zu können. Außerdem wurden die Beschreibungen aller Kompetenzstufen überarbeitet, um die neuen Aspekte der Lesekompetenz, die sich durch die neue Rahmenkonzeption und die neuen Aufgaben ergeben, abzubilden (OECD, 2019b).

Die Kompetenzstufen umfassen jeweils etwa 73 Punkte auf der Skala der Lesekompetenz. Je höher die Kompetenzstufe ist, desto schwieriger ist eine Aufgabe und desto kleiner ist die Wahrscheinlichkeit, dass Jugendliche mit einer geringen Lesekompetenz sie korrekt lösen können (vgl. Heine et al., 2016). Tabelle 3.1 gibt einen Überblick über die typischen Anforderungen auf den acht Kompetenzstufen der Lesekompetenz.

Die in Kapitel 2 dargestellten Beispielaufgaben beinhalten sowohl Aufgaben niedrigerer (Ia, II) als auch höherer Kompetenzstufen (IV, V). Das Anforderungsniveau von Beispielaufgabe 1 aus der Aufgabeneinheit „Hühnerforum“ veranschaulicht die Kompetenzstufe Ia. Schülerinnen und Schüler müssen das Hauptthema des Textes sowie die Absicht der Autorin verstehen, um die Aufgabe korrekt zu bearbeiten. Sowohl Aufgabe 3 als auch Aufgabe 6 aus der Aufgabeneinheit „Hühnerforum“ (vgl. Kapitel 2) sind Beispiele für die Kompetenzstufe II. Bei Aufgabe 3 müssen Jugendliche die Grundidee von verschiedenen kurzen Texten erkennen und in einen Zusammenhang setzen. Bei Aufgabe 6 ist es erforderlich, mehrere Behauptungen zu vergleichen und diese im Aufgabenzusammenhang zu bewerten.

Die Schwierigkeit von Aufgabe 7 aus der Aufgabeneinheit „Die Osterinsel“ (vgl. Kapitel 2) fällt in Kompetenzstufe IV. Bei dieser Aufgabe müssen die Jugendlichen umfangreiche Textabschnitte aus mehreren Quellen verstehen, die unterschiedlichen Perspektiven vergleichen sowie deren Zusammenhang bewerten, um Schlüsse zu ziehen, die zur Beantwortung der Frage notwendig sind.

Ein noch schwierigeres Beispiel, das die Kompetenzstufe V veranschaulicht, ist die in Kapitel 2 dargestellte Aufgabe 3 der Aufgabeneinheit „Die Osterinsel“. Zur korrekten Beantwortung müssen die Fünfzehnjährigen hier in der Lage sein, lange Texte zu verstehen und relevante Informationen zu finden, auch wenn diese im Text nicht offensicht-

Tabelle 3.1: Überblick über die typischen Anforderungen der acht Kompetenzstufen (Ic–VI) der Lesekompetenz (Tabelle adaptiert nach OECD, 2019b)

Kompetenzstufe	Wozu die Schülerinnen und Schüler auf der jeweiligen Kompetenzstufe im Allgemeinen in der Lage sind
VI ≥ 698 Punkte	<p>Jugendliche auf dieser Stufe können lange und abstrakte Texte verstehen, in denen die relevanten Informationen tief eingebettet sind und nur indirekt mit der Aufgabe zusammenhängen. Sie können Informationen, die mehrere und potenziell widersprüchliche Perspektiven enthalten, vergleichen, gegenüberstellen und integrieren. Sie können über die Quelle des Textes in Bezug zum Textinhalt reflektieren, indem sie Kriterien nutzen, die außerhalb des Textes liegen (z. B. eigenes Wissen). Sie können Widersprüche und Konflikte zwischen verschiedenen Texten zum selben Thema identifizieren und auflösen, indem sie Schlüsse über die Quelle ziehen. Sie bedenken dabei auch persönliche Interessen der Autorin bzw. des Autors und nutzen andere Hinweise, um die Glaubwürdigkeit der Informationen einschätzen zu können.</p> <p>Aufgaben dieser Stufe beinhalten ein oder mehrere komplexe Texte, die mehrere und möglicherweise widersprüchliche Perspektiven und Positionen enthalten. Die Informationen, die zur Beantwortung der Frage erschlossen werden müssen, können dabei auch in Form von Details vorliegen, die tief eingebettet sind und möglicherweise durch konkurrierende Informationen überdeckt werden.</p>
V 626–697 Punkte	<p>Jugendliche auf dieser Stufe können lange Texte verstehen und herausfinden, welche Informationen im Text relevant sind, auch wenn die relevanten Informationen leicht übersehen werden könnten. Sie können auf der Grundlage eines tiefen Verständnisses textbezogen argumentieren. Sie können den Zusammenhang zwischen einer Frage und einer oder mehreren Informationen, die innerhalb oder über mehrere Texte und Quellen verteilt sind, erschließen.</p> <p>Aufgaben zum Reflektieren erfordern das Aufstellen oder das kritische Bewerten von Hypothesen anhand spezifischer Informationen. Die Jugendlichen können bei komplexen oder abstrakten Aussagen zwischen Inhalt und Zweck sowie zwischen Tatsache und Meinung unterscheiden. Sie können Neutralität und Befangenheit auf der Grundlage expliziter oder impliziter Hinweise beurteilen, welche den Inhalt und/oder die Quelle der Informationen betreffen. Außerdem können sie Schlüsse über die Zuverlässigkeit der Behauptungen oder Folgerungen ziehen, die in einem Text angeboten werden.</p> <p>Aufgaben dieser Stufe beinhalten typischerweise den Umgang mit abstrakten oder kontraintuitiven Konzepten und das Durchlaufen mehrerer Schritte bis zur Zielerreichung. Zudem können Aufgaben dieser Stufe von der Leserin oder dem Leser verlangen, mehrere lange Texte zu bearbeiten und zwischen den Texten hin- und herzuwechseln, um Informationen zu vergleichen und gegenüberzustellen.</p>
IV 553–625 Punkte	<p>Jugendliche auf dieser Stufe können umfangreiche Abschnitte aus Texten mit einer oder mehreren Quellen verstehen. Sie interpretieren die Bedeutung von sprachlichen Feinheiten in einem Textabschnitt, indem sie den gesamten Text berücksichtigen. Sie können Perspektiven vergleichen und Schlüsse auf der Grundlage mehrerer Quellen ziehen. Sie können eingebettete Informationen suchen, lokalisieren und integrieren. Sie sind in der Lage, auf der Grundlage der Aufgabenstellung Schlüsse zu ziehen, um die Relevanz von Zielinformationen zu beurteilen. Sie können Aufgaben bewältigen, die es erfordern, sich den vorherigen Aufgabenkontext zu merken. Außerdem können die Jugendlichen auf dieser Stufe den Zusammenhang zwischen bestimmten Aussagen und der allgemeinen Haltung oder Folgerung einer Person zu einem Thema bewerten.</p> <p>Sie können anhand hervorstechender Textmerkmale wie Überschriften und Abbildungen über die Strategien reflektieren, die eine Autorin oder ein Autor verwendet, um (Schreib-) Ziele zu erreichen. Sie können Behauptungen, die in mehreren Texten explizit vorgebracht werden, vergleichen und gegenüberstellen sowie die Zuverlässigkeit einer Quelle anhand hervorstechender Kriterien beurteilen.</p> <p>Texte dieser Stufe sind oft lang oder komplex und ihr Inhalt oder ihre Form sind möglicherweise unüblich. Die Texte und die Aufgaben enthalten indirekte oder implizite Hinweise.</p>

III 480–552 Punkte	<p>Jugendliche auf dieser Stufe können die wörtliche Bedeutung von Texten mit einer oder mehreren Quellen wiedergeben, ohne dass explizite Inhalte oder organisatorische Hinweise vorhanden sind. Sie können Inhalte integrieren und sowohl grundlegende als auch anspruchsvollere Schlussfolgerungen ziehen. Sie können auch mehrere Teile eines Textes miteinander verknüpfen, um die Grundidee zu identifizieren, einen Zusammenhang zu verstehen oder die Bedeutung eines Wortes oder eines Satzes zu interpretieren, wenn die erforderlichen Informationen auf einer Seite dargestellt sind. Sie können nach Informationen suchen und Zielinformationen lokalisieren, die nicht leicht erkennbar sind und/oder sich in der Nähe von Ablenkungen befinden. In manchen Fällen erkennen sie den Zusammenhang zwischen mehreren Informationen. Jugendliche auf dieser Stufe können über einen Textteil oder über eine begrenzte Anzahl an Texten reflektieren und die Standpunkte mehrerer Autorinnen und Autoren auf der Grundlage expliziter Informationen vergleichen und gegenüberstellen.</p> <p>Aufgaben zum Reflektieren können es erfordern, Vergleiche anzustellen, Erklärungen zu erstellen oder ein Merkmal des Textes zu bewerten. Manche der Aufgaben erfordern ein detailliertes Verständnis eines Textes, der sich mit einem vertrauten Thema beschäftigt, während andere ein globales Verständnis von ungewohnten Inhalten erfordern.</p> <p>Aufgaben dieser Stufe verlangen, dass Jugendliche beim Vergleichen, Gegenüberstellen oder Kategorisieren von Informationen mehrere Merkmale berücksichtigen. Die geforderten Informationen sind oft nicht leicht erkennbar oder es kann eine hohe Anzahl konkurrierender Informationen geben. Die Texte dieser Stufe können außerdem unerwartete oder negativ formulierte Ideen enthalten.</p>
II 408–479 Punkte	<p>Jugendliche auf dieser Stufe können die Grundidee eines Textes von mittlerer Länge erkennen. Sie können Zusammenhänge verstehen oder die Bedeutung eines begrenzten Textteils interpretieren, wenn die Informationen nicht leicht erkennbar sind und/oder wenn einige wenige ablenkende Informationen vorhanden sind, indem sie grundlegende Schlussfolgerungen ziehen. Sie können, wenn sie explizit darauf hingewiesen werden, über den Gesamtzweck oder den Zweck spezifischer Details von Texten mittlerer Länge reflektieren. Sie können über einfache visuelle oder typografische Merkmale reflektieren. Sie können außerdem Behauptungen vergleichen und deren Begründungen anhand kurzer, expliziter Aussagen, bewerten.</p> <p>Bei Aufgaben dieser Stufe müssen unter Umständen auf der Grundlage eines einzigen Textbestandteils Vergleiche und Gegenüberstellungen vorgenommen werden oder es müssen, ausgehend von eigenen Erfahrungen oder Standpunkten, Vergleiche angestellt oder Zusammenhänge zwischen dem Text und nicht im Text enthaltenen Informationen erkannt werden.</p>
Ia 335–407 Punkte	<p>Jugendliche auf dieser Stufe können die wörtliche Bedeutung von Sätzen oder kurzen Abschnitten verstehen. Sie können das Hauptthema oder die Absicht des Autors in einem Textteil zu einem vertrauten Thema erkennen und einen einfachen Zusammenhang zwischen benachbarten Informationsteilen oder zwischen den vorliegenden Informationen und dem eigenen Vorwissen herstellen. Sie können eine relevante Seite aus einer begrenzten Auswahl, auf der Grundlage von einfachen Aufforderungen auswählen sowie eine oder mehrere unabhängige Informationsteile aus kurzen Texten lokalisieren.</p> <p>Lesende auf dieser Stufe können über den Gesamtzweck, den Kernpunkt und die Zusatzinformationen von einfachen Texten mit expliziten Hinweisen reflektieren.</p> <p>Die meisten Aufgaben auf dieser Stufe weisen auf entscheidende Elemente in der Aufgabe und im Text hin.</p>
Ib 262–334 Punkte	<p>Jugendliche auf dieser Stufe können die wörtliche Bedeutung von Sätzen oder kurzen Abschnitten verstehen. Sie können das Hauptthema oder die Absicht des Autors in einem Textteil zu einem vertrauten Thema erkennen und einen einfachen Zusammenhang zwischen benachbarten Informationsteilen oder zwischen den vorliegenden Informationen und dem eigenen Vorwissen herstellen. Sie können auf der Grundlage von einfachen Aufforderungen eine relevante Seite aus einer begrenzten Auswahl auswählen sowie eine oder mehrere unabhängige Informationsteile aus kurzen Texten lokalisieren. Sie können über die gesamte Textfunktion sowie über die relative Wichtigkeit von Informationen (z. B. Hauptidee vs. unwichtige Details) von einfachen Texten mit expliziten Hinweisen reflektieren.</p> <p>Die meisten Aufgaben auf dieser Stufe weisen auf entscheidende Elemente in der Aufgabe und im Text hin.</p>

Ic 189–261 Punkte	Jugendliche auf dieser Stufe können die Bedeutung kurzer, syntaktisch einfacher Sätze auf wörtlicher Ebene verstehen und bestätigen. Sie können für einen eindeutigen und einfachen Zweck innerhalb eines begrenzten Zeitraums lesen. Die Aufgaben auf dieser Stufe umfassen lediglich einfache Wörter und syntaktische Strukturen.
unter Ic ≤ 188 Punkte	

lich sind. Zusätzlich müssen die Jugendlichen Aussagen anhand bestimmter Informationen kritisch bewerten, indem sie zwischen Tatsache und Meinung unterscheiden.

Ein weiteres Beispiel für die Kompetenzstufe V ist Aufgabe 6 aus der Aufgabeneinheit „Die Osterinsel“ (vgl. Kapitel 2). Hier müssen die Schülerinnen und Schüler auf drei Textquellen zurückgreifen und zwischen ihnen hin- und herwechseln, um relevante Informationen zu finden. Es gilt, einen Zusammenhang zwischen den Informationen, die über mehrere Textquellen verteilt sind, zu erschließen. Konkret müssen Unterschiede zwischen zwei Theorien aus zwei der Textquellen und ihre zentrale Gemeinsamkeit benannt werden und gleichzeitig irrelevante Informationen aus der dritten Textquelle als solche erkannt werden.

3.1.5 Fragestellungen aus Sicht des Bildungsmonitorings

Die Fragestellungen dieses Kapitels zielen im Sinne des Bildungsmonitorings darauf, die Lesekompetenz von Fünfzehnjährigen in Deutschland im Jahr 2018 im Vergleich zu Jugendlichen in anderen Staaten sowie im Vergleich zu früheren PISA-Erhebungsrunden darzustellen. Daher werden zum einen Ergebnisse zur Lesekompetenz im internationalen Vergleich berichtet. Zum anderen werden vertiefende nationale Aspekte zu Unterschieden zwischen Schularten sowie Unterschieden zwischen Mädchen und Jungen in Deutschland dargestellt.

Im vorliegenden Kapitel werden die Ergebnisse zur Lesekompetenz der PISA-Studie 2018 im Wesentlichen mit den Ergebnissen aus dem Jahr 2009 verglichen, als Lesen zuletzt Hauptdomäne war (vgl. Heine et al., 2016). Der Grund ist, dass die Hauptdomänen bei PISA jeweils mit einer hinreichend großen Anzahl von Aufgaben erfasst werden (Jude & Klieme, 2010; OECD, 2010). Zusätzlich werden im Folgenden Veränderungen im Vergleich zu PISA 2015 betrachtet, als die Lesekompetenz erstmals computerbasiert erfasst wurde (vgl. Kapitel 7). Allerdings ist zu beachten, dass sich bei PISA 2018 zusätzlich die Rahmenkonzeption, die Leseaufgaben und das Testdesign (hin zum mehrstufig adaptiven Testdesign) verändert haben (siehe Abschnitt 3.1.1). Diese – zeitgemäßen – Veränderungen und Anpassungen sind sinnvoll, da Jugendliche insbesondere digitale Materialien lesen und mit Computern umgehen müssen. Allerdings wird damit auch die Vergleichbarkeit der Befunde über die verschiedenen PISA-Erhebungsrunden hinweg eingeschränkt. Dem wird zunächst begegnet, indem Trendaufgaben eingesetzt werden, die sich über mehrere Runden hinweg nicht unterscheiden. Zudem werden statistische

Verfahren eingesetzt wie die Berücksichtigung des sogenannten Link-Fehlers (vgl. Heine et al., 2016), durch welchen Unsicherheiten des Vergleichs zwischen Erhebungsrunden reduziert werden. Aussagen zu Veränderungen der Lesekompetenz von Fünfzehnjährigen über die Zeit sind zwar unter Vorbehalt zu betrachten, dennoch sind diese Informationen relevant und können im Sinne des Bildungsmonitorings Hinweise auf die Wirksamkeit von ergriffenen Fördermaßnahmen geben.

3.2 Lesekompetenz im internationalen Vergleich

Da in der PISA-Studie der internationale Vergleich im Mittelpunkt steht, werden zunächst die Ergebnisse der Lesekompetenz der Jugendlichen in Deutschland im Vergleich zur Lesekompetenz der Jugendlichen in den OECD-Staaten sowie in den OECD-Partnerstaaten berichtet. Dabei werden Mittelwerte und Streuungen der Lesekompetenz der PISA-Studie 2018 sowie Anteile der Fünfzehnjährigen auf den verschiedenen Kompetenzstufen und Geschlechterunterschiede zwischen den Staaten verglichen. An der PISA-Studie 2018 nahmen 37 OECD-Staaten und 42 sogenannte OECD-Partnerstaaten teil (vgl. Kapitel 1). Der OECD-Mittelwert wird aus den Werten von 36 OECD-Staaten berechnet, da vom OECD-Staat Spanien aus methodischen Gründen keine Ergebnisse zur Lesekompetenz berichtet werden.

3.2.1 Gesamtskala Lesekompetenz im internationalen Vergleich

Mittelwerte der Lesekompetenz

Der Mittelwert der Lesekompetenz in der PISA-Studie 2018 liegt in Deutschland mit 498 Punkten signifikant über dem OECD-Mittelwert von 487 Punkten (siehe auch Abbildung 3.1). Somit verfügen die Jugendlichen in Deutschland im Jahr 2018 über eine mittlere Lesekompetenz, welche signifikant höher ist als die durchschnittliche Lesekompetenz in den OECD-Staaten. Dies war bei der PISA-Studie 2009, als die Lesekompetenz zuletzt Hauptdomäne war, nicht der Fall. Damals erreichten die Fünfzehnjährigen in Deutschland eine durchschnittliche Lesekompetenz, die sich nicht signifikant vom OECD-Mittelwert unterschied. Als die Lesekompetenz bei der ersten PISA-Studie im Jahr 2000 die Hauptdomäne darstellte, lag der Mittelwert der Jugendlichen in Deutschland sogar signifikant unter der durchschnittlichen Lesekompetenz der Jugendlichen in den OECD-Staaten. Somit ist es bei der PISA-Studie 2018 zum ersten Mal der Fall, dass die Lesekompetenz Hauptdomäne ist und die durchschnittliche Lesekompetenz der Fünfzehnjährigen in Deutschland den OECD-Durchschnitt signifikant übertrifft. In den Jahren 2012 und 2015 lag der Mittelwert der Lesekompetenz in Deutschland zwar auch signifikant über dem OECD-Mittelwert, damals wurde die Lesekompetenz allerdings als

Nebendomäne und somit weniger ausführlich erfasst (siehe auch Abschnitt 3.3.3). Die Anzahl und Zusammensetzung der OECD-Staaten, die in den jeweiligen OECD-Durchschnitt von PISA 2000 bis 2018 eingingen, variierte jedoch. Berechnet man den Durchschnitt der 24 OECD-Staaten, für die von 2000 bis 2018 durchgängig Daten verfügbar sind, so zeigt sich ebenfalls, dass die Lesekompetenz in Deutschland bei PISA 2000 signifikant unter diesem OECD-24-Durchschnitt, bei PISA 2009 im OECD-24-Durchschnitt und bei PISA 2012 bis 2018 signifikant über dem OECD-24-Durchschnitt liegt.

Wie in Abbildung 3.1 ersichtlich ist, erreichen die Fünfzehnjährigen in insgesamt 18 OECD-Staaten eine mittlere Lesekompetenz, welche signifikant über dem OECD-Durchschnitt liegt. Zu diesen Staaten gehört Deutschland. Ähnlich wie in Deutschland ist die Lesekompetenz in Dänemark, Norwegen, Belgien und Frankreich, aber auch im Vereinigten Königreich, in Slowenien, Schweden, Portugal, Japan, den Vereinigten Staaten und Australien ausgeprägt. In insgesamt sieben OECD-Staaten ist die mittlere Lesekompetenz der Jugendlichen signifikant höher als in Deutschland, diese sind Estland, Kanada, Finnland, Irland, die Republik Korea, Polen und Neuseeland.

Die Mittelwerte der Lesekompetenz von fünf Staaten unterscheiden sich nicht signifikant vom OECD-Durchschnitt. Zu dieser Gruppe gehören Portugal, Tschechien, die Niederlande, Österreich und die Schweiz. Signifikant niedriger als der OECD-Mittelwert ist die durchschnittliche Lesekompetenz von Jugendlichen in insgesamt 13 OECD-Staaten. Dabei ist zu berücksichtigen, dass es große Unterschiede in der Lesekompetenz zwischen den Staaten dieser Gruppe gibt. Dies spiegelt sich beispielsweise in dem großen Abstand zwischen den Mittelwerten von Kolumbien, dem schwächsten Mitglied dieser Gruppe, und dem stärksten Mitglied Lettland wider, welcher 67 Punkte beträgt. Zwischen dem insgesamt kompetenzstärksten Staat Estland und Kolumbien beträgt der Unterschied sogar 111 Punkte, was mehr als einer Standardabweichung entspricht. Damit ist die Leistungsspanne zwischen den OECD-Staaten und damit die Diversität in der Lesekompetenz sehr groß.

Bezieht man die Ergebnisse der OECD-Partnerstaaten ein, erhöht sich die Bandbreite der Mittelwerte der Lesekompetenz im internationalen Vergleich noch weiter (siehe Tabelle 3.1web). Während die Jugendlichen aus einigen wenigen OECD-Partnerstaaten besonders hohe Mittelwerte der Lesekompetenzen erreichen, zeigen sich auch die niedrigsten Mittelwerte in dieser Gruppe. So erreichen beispielsweise Fünfzehnjährige in vier Verwaltungseinheiten Chinas (Peking, Shanghai, Jiangsu, Zhejiang) mit 555 Punkten sowie Jugendliche in Singapur mit 549 Punkten durchschnittliche Lesekompetenzwerte, die deutlich über dem Mittelwert von Estland liegen. Die durchschnittliche Lesekompetenz der Fünfzehnjährigen auf den Philippinen (340 Punkte) und in der Dominikanischen Republik (342 Punkte) liegt hingegen mehr als zwei Kompetenzstufen unter dem OECD-Mittelwert.

Die Streuung der Lesekompetenz

Neben der Betrachtung der Mittelwerte ist es aufschlussreich, die Streuung der Lesekompetenz zwischen den Staaten zu vergleichen. Diese wird in Form der Standardabweichung (*SD*) angegeben (siehe Abbildung 3.1 sowie Tabelle 3.1web für die OECD-Partnerstaaten) und gibt darüber Auskunft, wie groß die Unterschiede in der Lesekompetenz innerhalb der einzelnen Staaten sind. Im Durchschnitt größere Unterschiede zwischen den lesestärksten und leeschwächsten Jugendlichen innerhalb eines Staates bedeuten eine höhere Streuung (Standardabweichung). Die Standardabweichung der Lesekompetenz in Deutschland beträgt 106 Punkte und ist damit signifikant größer als die Streuung über alle OECD-Staaten hinweg (*SD* = 99). Zudem ist die Streuung in Deutschland damit bei PISA 2018 signifikant größer als dies 2009 und 2015 der Fall war. Eine ähnlich hohe Streuung wie in Deutschland findet sich beispielsweise in der Republik Korea und Neuseeland. Die Ergebnisse in Schweden, den Vereinigten Staaten, Australien, Norwegen und Belgien sind sowohl bezüglich der Mittelwerte als auch bezüglich der Streuungen der Lesekompetenz ähnlich zu den Ergebnissen in Deutschland. In den Niederlanden, der Schweiz, Island und Luxemburg ist zwar die Streuung ähnlich hoch wie in Deutschland, die Mittelwerte der Lesekompetenz sind jedoch deutlich niedriger. Eine besonders hohe Streuung weist Israel (*SD* = 124) auf. Im internationalen Vergleich besonders geringe Streuungen zeigen sich beispielsweise in Estland (*SD* = 93), Irland (*SD* = 91), der Türkei (*SD* = 88), Mexiko (*SD* = 84) und Kolumbien (*SD* = 89). In Estland und Irland sind die Mittelwerte der Lesekompetenz bedeutsam höher als in Deutschland und gleichzeitig die Streuungen signifikant niedriger als im OECD-Mittel. Von den OECD-Partnerstaaten fällt der Zusammenschluss der vier Verwaltungseinheiten Chinas (Peking, Shanghai, Jiangsu, Zhejiang) mit einer hohen mittleren Lesekompetenz bei vergleichsweise geringer Streuung (*SD* = 87) auf (siehe Tabelle 3.1web). Diese Beispiele zeigen, dass eine hohe durchschnittliche Lesekompetenz bei relativ geringen mittleren Unterschieden zwischen den leeschwächsten und lesestärksten Jugendlichen innerhalb eines Staates möglich ist.

Zur Veranschaulichung der Bandbreite der Lesekompetenz innerhalb der Staaten werden in Abbildung 3.1 zusätzlich Perzentilbänder dargestellt (siehe Tabelle 3.1web für die Punktwerte der Perzentile). Hier fällt für Deutschland das relativ niedrige 5-Prozent-Perzentil auf. Das 5-Prozent-Perzentil gibt den Kompetenzwert an, den 95 Prozent der Jugendlichen in einem Staat erreichen. Dieser Punktwert ist für Deutschland (316 Punkte) signifikant niedriger als in manchen anderen OECD-Staaten mit durchschnittlich ähnlich hoher Lesekompetenz, zum Beispiel Dänemark (344 Punkte), Slowenien (335 Punkte) und Japan (337 Punkte). Es gibt allerdings auch OECD-Staaten mit einem ähnlichen Mittelwert wie Deutschland und einem ähnlichen 5-Prozent-Perzentil (z.B. Australien, Belgien, Norwegen). In den OECD-Staaten mit den höchsten mittleren Lesekompetenzen sind auch die Punktwerte der 5-Prozent-Perzentile relativ hoch (z.B. 367 Punkte in Estland; 364 Punkte in Irland), was bedeutet, dass auch die leeschwächste Gruppe in diesen Staaten über eine vergleichsweise hohe Lesekompetenz

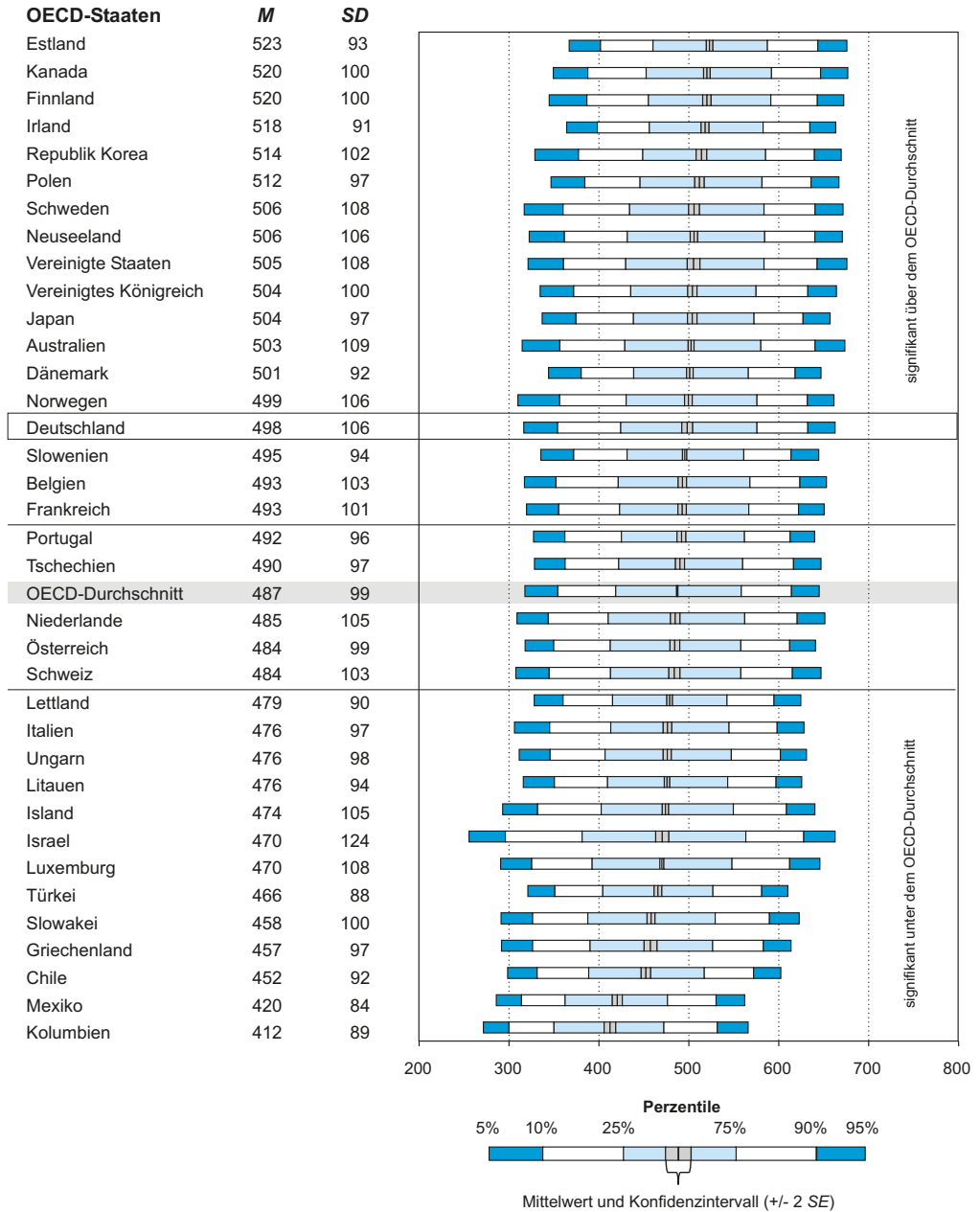


Abbildung 3.1: Mittelwerte, Streuungen und Perzentilbänder der Gesamtskala Lesekompetenz der OECD-Staaten

verfügt. Im oberen Kompetenzbereich grenzt das 95-Prozent-Perzentil entsprechend den Bereich ab, in dem die 5 Prozent der lesestärksten Jugendlichen liegen. Hier fällt positiv auf, dass der Punktwert für das 95-Prozent-Perzentil in Deutschland (663 Punkte) ähnlich hoch oder höher als in den meisten anderen OECD-Staaten mit ähnlich hoher oder sogar höherer mittlerer Lesekompetenz ist (z. B. ähnlich hoch wie in Irland, der Republik Korea, Polen und signifikant höher als in Dänemark, Slowenien, Belgien, Frankreich). Lediglich in insgesamt vier OECD-Staaten ist der Punktwert für das 95-Prozent-Perzentil signifikant höher als in Deutschland (677 Punkte in Kanada, 676 Punkte in Estland und den Vereinigten Staaten, 673 in Australien).

Die Ergebnisse verdeutlichen, dass die Gruppe der besonders lesestarken Jugendlichen in Deutschland auch im internationalen Vergleich sehr hohe Kompetenzen zeigt, die Gruppe der legeschwachen Jugendlichen jedoch vergleichsweise niedrige Kompetenzen aufweist.

3.2.2 Verteilung auf die Stufen der Lesekompetenz

Mehr Informationen über besonders lesestarke beziehungsweise besonders legeschwache Jugendliche bekommt man, wenn man die Anteile von Fünfzehnjährigen auf den untersten (Ia oder darunter) sowie auf den obersten Kompetenzstufen (V und VI) im internationalen Vergleich betrachtet (siehe Abbildung 3.2). Man sieht so insbesondere, wie gut es den Bildungssystemen der verschiedenen Staaten gelingt, den Anteil besonders legeschwacher Schülerinnen und Schüler gering zu halten und gleichzeitig einen möglichst hohen Anteil besonders lesestarker Fünfzehnjähriger aufzubauen.

In Deutschland befinden sich 21 Prozent der Jugendlichen auf den untersten Kompetenzstufen, was leicht – aber nicht signifikant – unter dem OECD-Durchschnitt von 23 Prozent liegt. Damit weist Deutschland im Vergleich zu anderen Staaten, deren Mittelwerte der Lesekompetenz ebenfalls über dem OECD-Mittelwert liegen, einen verhältnismäßig hohen Anteil legeschwacher Schülerinnen und Schüler auf. Von allen OECD-Staaten mit einem Mittelwert der Lesekompetenz über dem OECD-Durchschnitt, sind Deutschland und Belgien die einzigen beiden Staaten, bei welchen der Anteil der Fünfzehnjährigen auf den untersten Kompetenzstufen ähnlich hoch wie im OECD-Mittel ist. Bei allen anderen Staaten mit überdurchschnittlich hohem Mittelwert ist der Anteil der Jugendlichen unterhalb der Kompetenzstufe II signifikant geringer als im OECD-Mittel. Besonders niedrig sind die Anteile der Jugendlichen, die sich auf den untersten Stufen der Lesekompetenz befinden, in den vier OECD-Staaten mit den höchsten Mittelwerten der Lesekompetenz (11 % in Estland, 12 % in Irland, 14 % in Finnland und Kanada). Noch geringer ist dieser Anteil mit fünf Prozent in dem OECD-Partnerstaat des Zusammenschlusses von vier Verwaltungseinheiten Chinas (Peking, Shanghai, Jiangsu, Zhejiang).

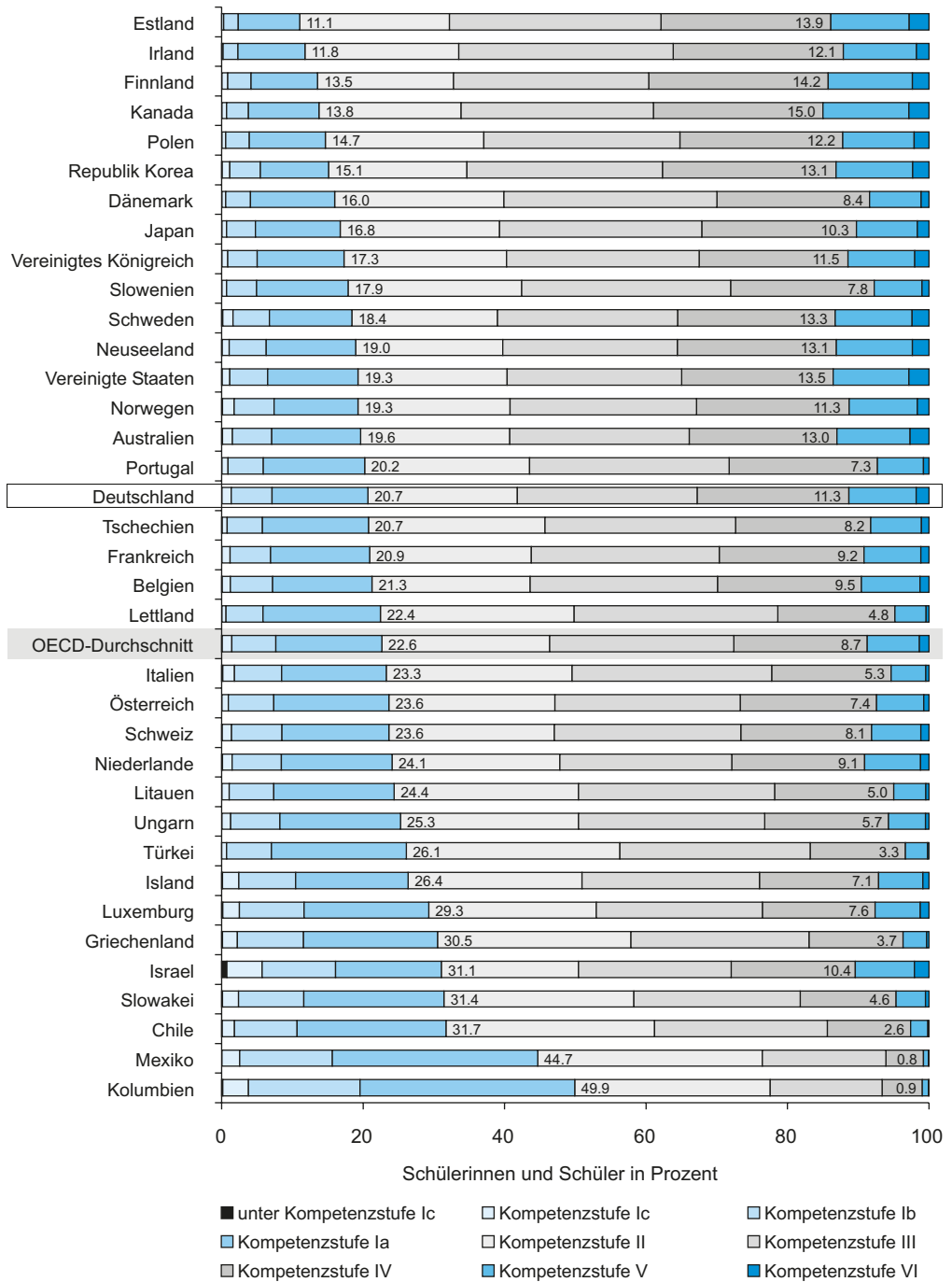


Abbildung 3.2: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe Ia, Ib, Ic oder darunter sowie auf Kompetenzstufe V oder VI für die Gesamtskala Lesekompetenz in den OECD-Staaten

Auf den beiden obersten Kompetenzstufen befinden sich in Deutschland elf Prozent der Jugendlichen. Damit ist der Anteil der besonders lesestarken Fünfzehnjährigen in Deutschland signifikant größer als im OECD-Durchschnitt mit neun Prozent. Nur in sechs der OECD-Staaten sind die Anteile der besonders lesestarken Jugendlichen signifikant größer als in Deutschland, dazu gehören beispielsweise Kanada mit 15 Prozent sowie Estland und Finnland mit jeweils 14 Prozent. Noch größere Anteile von Jugendlichen auf den obersten Kompetenzstufen finden sich in den OECD-Partnerstaaten Singapur mit 26 Prozent und dem Zusammenschluss von vier Verwaltungseinheiten Chinas (Peking, Shanghai, Jiangsu, Zhejiang) mit 22 Prozent.

3.2.3 Unterschiede in der Lesekompetenz zwischen Mädchen und Jungen

In allen Staaten, die an der PISA-Studie 2018 teilnahmen, erreichen Mädchen signifikant höhere Mittelwerte in der Lesekompetenz als Jungen. Allerdings unterscheidet sich das Ausmaß des Geschlechterunterschieds zwischen den Staaten. Während er im OECD-Mittel 30 Punkte beträgt, ist die Spannbreite mit einer Differenz von 10 Punkten in Kolumbien bis zu einer Differenz von 52 Punkten in Finnland zwischen den OECD-Staaten groß. Der größte Unterschied in der Lesekompetenz zwischen Mädchen und Jungen findet sich im OECD-Partnerstaat Katar mit 65 Punkten. In Deutschland beträgt die Geschlechterdifferenz 26 Punkte. Mädchen erreichen im Mittel 512 Punkte, Jungen 486 Punkte, was jeweils signifikant über dem OECD-Mittelwert der Mädchen beziehungsweise Jungen liegt. Deutschland ist damit bezüglich des Geschlechterunterschieds im Vergleich zu den anderen OECD-Staaten in der Lesekompetenz im unteren Mittelfeld zu sehen. Im Vergleich zur PISA-Studie 2009 verringerte sich der Unterschied zwischen Mädchen und Jungen in Deutschland deutlich, damals betrug die Differenz 40 Punkte. Auch im OECD-Durchschnitt war der Geschlechterunterschied im Jahr 2009 mit 39 Punkten signifikant höher als im Jahr 2018. Dieser Trend der deutlich geringeren Geschlechterdifferenzen zeichnete sich erstmals mit der PISA-Studie 2015 sowohl in Deutschland (21 Punkte Differenz) als auch im OECD-Mittel (27 Punkte) ab (vgl. Weis et al., 2016). Allerdings zeigen etwa Finnland, Israel und Norwegen immer noch ähnlich große Unterschiede wie 2009.

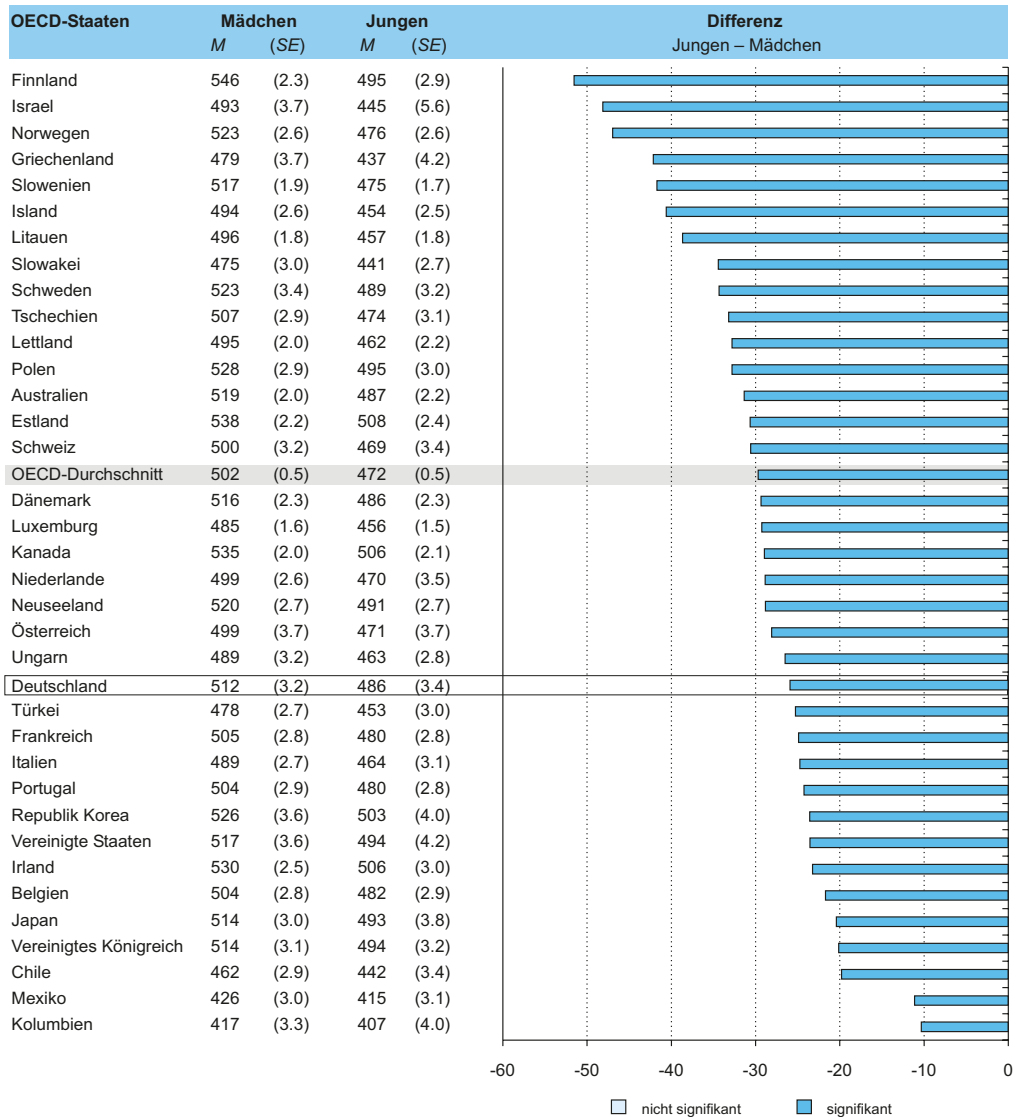


Abbildung 3.3: Mittelwerte der Lesekompetenz nach Geschlecht in den OECD-Staaten

3.2.4 Teilskalen der Lesekompetenz im internationalen Vergleich

Prozesse des Lesens im internationalen Vergleich

Da die Lesekompetenz in der PISA-Studie 2018 als Hauptdomäne ausführlich und mit einer großen Aufgabenmenge erfasst wurde, können Ergebnisse für die drei Prozesse des Lesens auf jeweils eigenen Teilskalen berichtet werden. Daher werden im Folgenden die Kompetenzen der Jugendlichen im *Lokalisieren von Informationen*, im *Textverstehen*

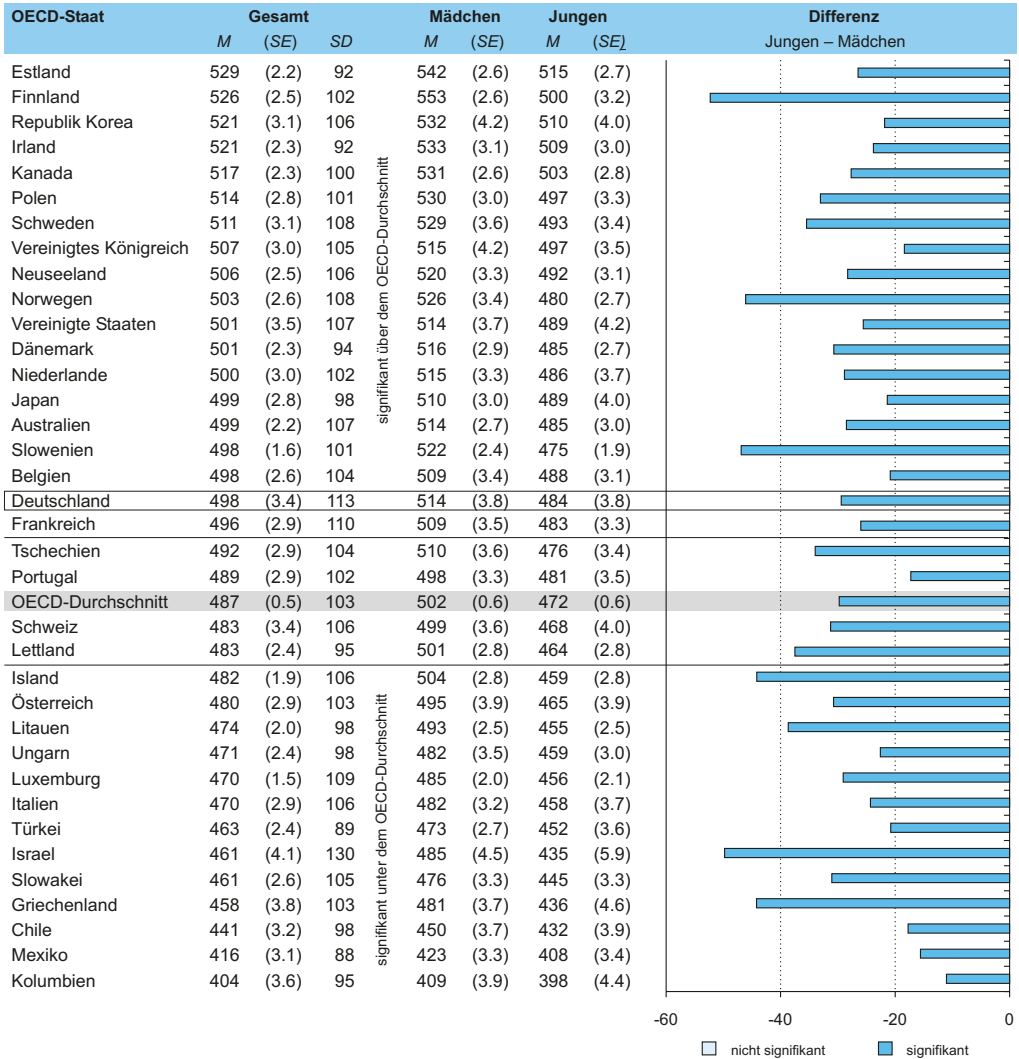


Abbildung 3.4: Mittelwerte und Streuungen der Teilskala *Lokalisieren von Informationen* in den OECD-Staaten

sowie im *Bewerten und Reflektieren* im internationalen Vergleich vorgestellt (siehe Abbildungen 3.4, 3.5, 3.6 für die OECD-Staaten; siehe Tabellen 3.2web, 3.3web, 3.4web für alle teilnehmenden Staaten).

Die Mittelwerte der Jugendlichen in Deutschland liegen für alle drei Teilskalen signifikant über den jeweiligen OECD-Mittelwerten. Insgesamt sind in den meisten Staaten die Mittelwerte der drei Teilskalen relativ ähnlich zum jeweiligen Mittelwert der Gesamtskala der Lesekompetenz. So ist beispielsweise der Mittelwert der Teilskala *Lokalisieren von Informationen* in Deutschland (498) sowie im OECD-Durchschnitt (487) jeweils identisch zum Mittelwert der Gesamtskala der Lesekompetenz. Dies ist zugleich die Teilskala, in welcher die Jugendlichen in Deutschland den höchsten Mittelwert

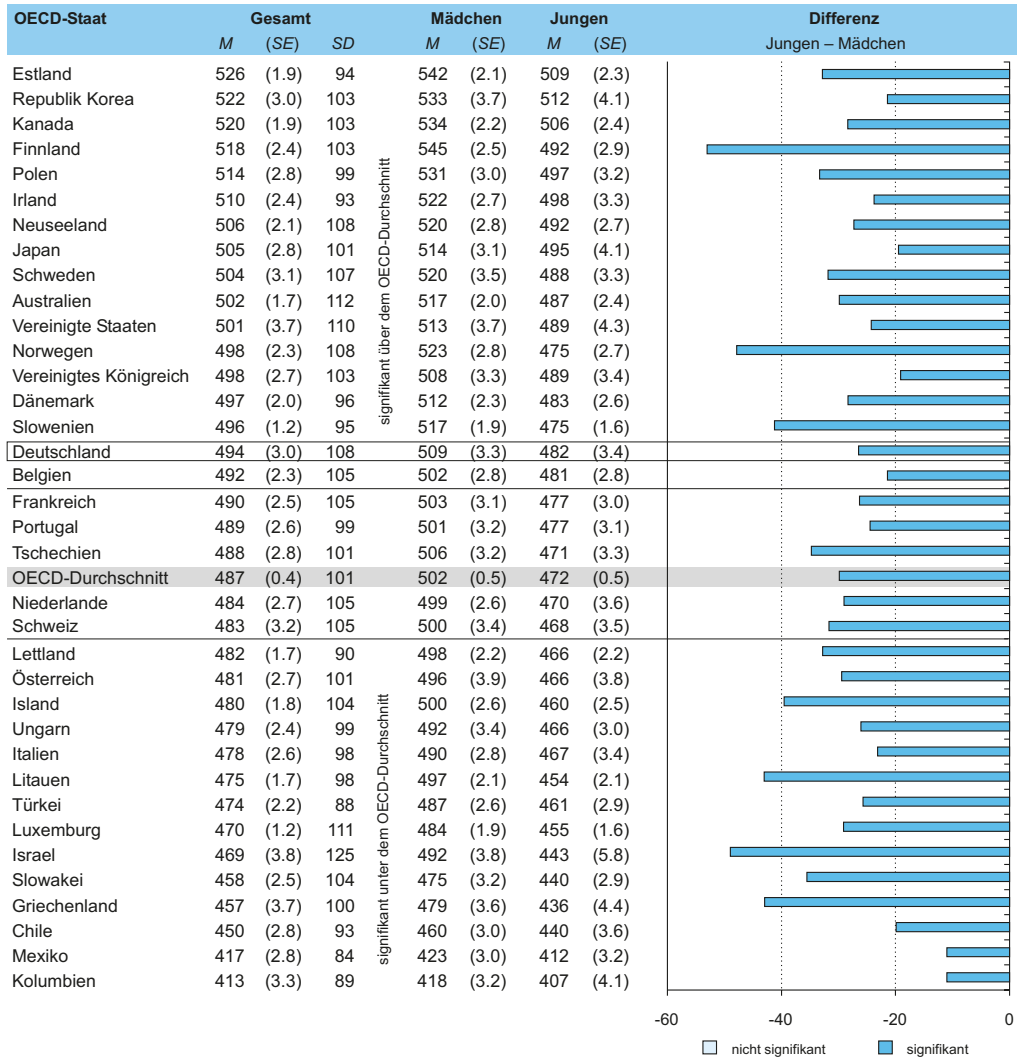


Abbildung 3.5: Mittelwerte und Streuungen der Teilskala *Textverstehen* in den OECD-Staaten

erreichen.¹ Allerdings ist zu beachten, dass die Differenzen zwischen den drei Teilskalen in Deutschland nicht besonders groß sind (1 bis 4 Punkte Differenz), auch wenn sie teilweise signifikant sind. Im OECD-Durchschnitt sind die Unterschiede zwischen den drei Teilskalen ebenfalls eher klein (0 bis 2 Punkte Differenz).

Deutlich größere Unterschiede zwischen den Mittelwerten der Teilskalen *Lokalisieren von Informationen* und *Bewerten und Reflektieren* finden sich beispielsweise in Chile

1 Sowohl die drei Teilskalen der Prozesse des Lesens als auch die zwei Teilskalen der Textquellen können jeweils nur untereinander verglichen werden. So ist aufgrund der vorgenommenen Skalierung beispielsweise kein Vergleich zwischen dem Textverstehen und dem Lesen mehrerer Textquellen möglich.

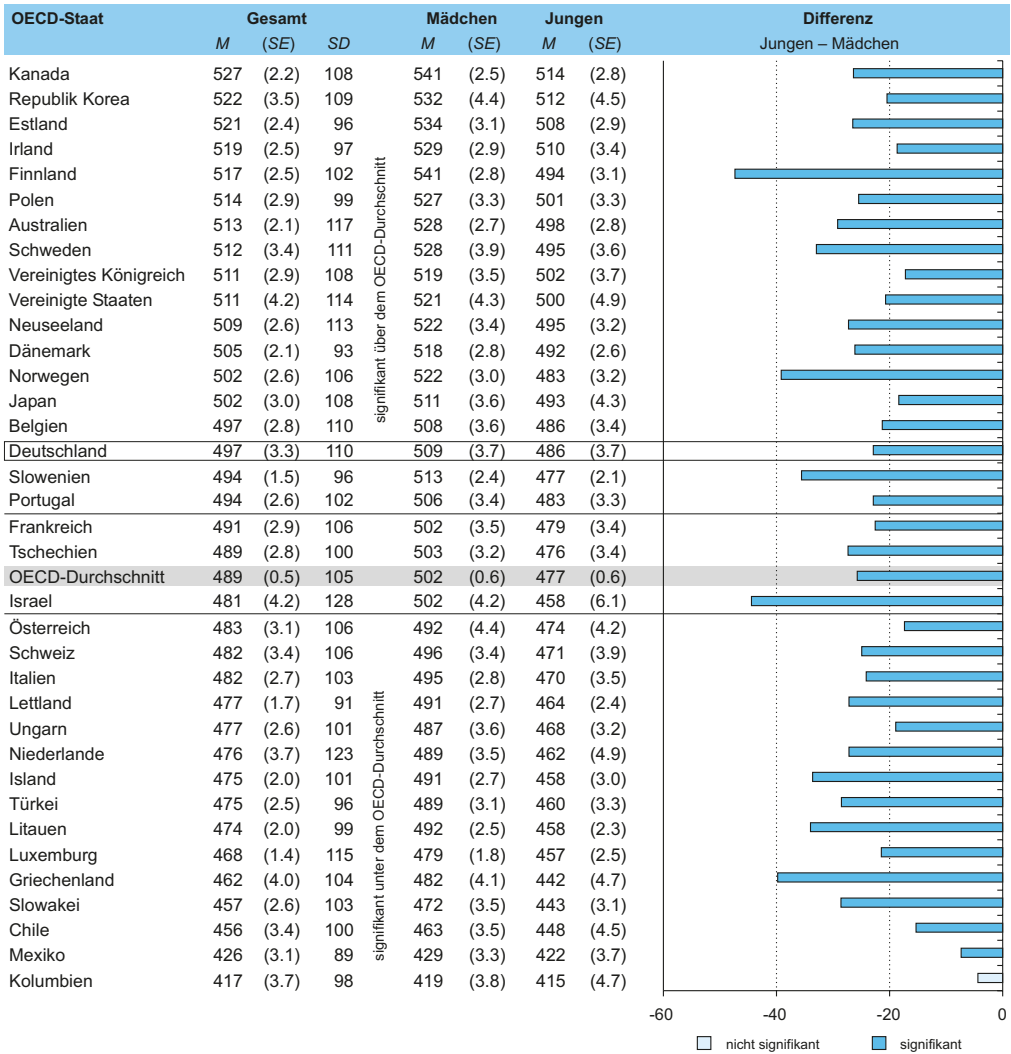


Abbildung 3.6: Mittelwerte und Streuungen der Teilskala *Bewerten und Reflektieren* in den OECD-Staaten

(15 Punkte Differenz), Israel (20 Punkte Differenz) und den Niederlanden (24 Punkte Differenz). Während in Chile und Israel die Jugendlichen deutlich höhere Kompetenzen im *Bewerten und Reflektieren* als im *Lokalisieren von Informationen* zeigen, ist dies in den Niederlanden umgekehrt. In Irland und den Niederlanden sind die Mittelwerte im *Lokalisieren von Informationen* deutlich höher als im *Textverstehen*, wohingegen Jugendliche in der Türkei auf dieser Teilskala im Vergleich zur Teilskala *Textverstehen* niedrigere Mittelwerte erreichen. Keinerlei signifikante Unterschiede zwischen den Teilskalen finden sich etwa im lesestarken OECD-Staat Neuseeland, aber beispielsweise auch in Österreich, wo die Lesekompetenz dem OECD-Mittel entspricht. In diesen Staaten gibt es keine relativen Leseschwächen in einer Teilskala im Vergleich zu den anderen beiden Teilskalen.

In allen OECD-Staaten außer in Kolumbien erreichen die Mädchen in allen Prozessen des Lesens signifikant höhere Mittelwerte als die Jungen. In Kolumbien zeigt sich kein signifikanter Geschlechterunterschied im *Bewerten und Reflektieren*. Wie bei der Gesamtskala der Lesekompetenz, sind die Geschlechterdifferenzen für alle drei Teilskalen in Finnland am größten und in Kolumbien und Mexiko am geringsten. In Deutschland sind die Geschlechterdifferenzen in der Teilskala *Bewerten und Reflektieren* (23 Punkte Differenz) am geringsten und für *Lokalisieren von Informationen* (29 Punkte Differenz) am stärksten ausgeprägt. Auch im OECD-Mittel zeigt sich die Tendenz, dass die Geschlechterunterschiede für die Teilskala *Bewerten und Reflektieren* am geringsten sind (26 Punkte vs. 30 Punkte Differenz).²

Textquellen im internationalen Vergleich: eine versus mehrere Textquellen

Sowohl für die Skala *eine Textquelle* als auch für die Skala *mehrere Textquellen* liegen die Mittelwerte der Jugendlichen in Deutschland signifikant über den jeweiligen OECD-Mittelwerten. Während die Fünfzehnjährigen in Deutschland im Mittel 497 Punkte auf der Skala *mehrere Textquellen* erreichen, sind es 494 Punkte auf der Teilskala *eine Textquelle* (siehe Abbildungen 3.1web, 3.2web für die OECD-Staaten; siehe Tabellen 3.5web, 3.6web für alle teilnehmenden Staaten). Im OECD-Durchschnitt liegt der Mittelwert der Teilskala *mehrere Textquellen* mit 490 Punkten signifikant über dem Mittelwert der Teilskala *eine Textquelle* mit 485 Punkten. Dies zeigt, dass Schülerinnen und Schüler im OECD-Durchschnitt mindestens genauso gut Aufgaben mit mehreren Textquellen lösen können wie Aufgaben mit einer Textquelle. Außerdem fällt auf, dass für die Teilskalen *eine* und *mehrere Textquellen* die Mädchen in allen PISA-Teilnehmerstaaten höhere Mittelwerte erreichen als die Jungen. Dieser Geschlechterunterschied scheint für die Teilskala *eine Textquelle* besonders stark ausgeprägt zu sein (29 Punkte Differenz in Deutschland; 23 Punkte Differenz bei *mehreren Textquellen*).

3.3 Vertiefende Analysen der Lesekompetenz in Deutschland

Zusätzlich zu den Ergebnissen im internationalen Vergleich, ermöglicht die PISA-Studie differenzierte Analysen zur Lesekompetenz der Jugendlichen in Deutschland. Dabei wird im Folgenden insbesondere auf Unterschiede zwischen Schularten, auf Geschlechterunterschiede sowie auf Veränderungen der Lesekompetenz in Deutschland seit PISA 2000 eingegangen.

2 Die angegebenen Differenzen wurden mit Dezimalstellen berechnet. Daher können beim Vergleich mit den Tabellen vermeintliche Inkonsistenzen aufgrund von Rundungen vorkommen.

3.3.1 Unterschiede zwischen Schularten

Da die Aufgliederung in verschiedene weiterführende Schularten nach der Grundschule eine Besonderheit des Schulsystems in Deutschland darstellt, ist es sinnvoll, Unterschiede zwischen den Schularten in vertiefenden nationalen PISA-Analysen zu untersuchen. Analog zum nationalen Bericht zur PISA-Studie 2015 (Sälzer & Reiss, 2016) wird hier zwischen Gymnasien und nicht gymnasialen Schularten (Hauptschule, Schule mit mehreren Bildungsgängen, Integrierte Gesamtschule und Realschule) unterschieden. Die Gesamtstichprobe in Deutschland umfasst 5451 Schülerinnen und Schüler. Davon besuchen 35.5 Prozent das Gymnasium und 59.7 Prozent nicht gymnasiale Schularten. Die restlichen Jugendlichen besuchen berufliche Schulen (3.0%) sowie Sonder- und Förderschulen (1.8%). Diese werden aufgrund der geringen Stichprobengröße an dieser Stelle nicht gesondert betrachtet.

Lesekompetenz: Mittelwerte und Streuungen im Vergleich

Wie in Abbildung 3.7 ersichtlich, unterscheidet sich die durchschnittliche Lesekompetenz der Fünfzehnjährigen an nicht gymnasialen Schularten deutlich von der durchschnittlichen Lesekompetenz der Gleichaltrigen an Gymnasien. Während der Mittelwert der Jugendlichen an nicht gymnasialen Schularten 458 Punkte beträgt, erreichen die Gymnasiastinnen und Gymnasiasten eine durchschnittliche Lesekompetenz von 578 Punkten. Somit liegt die durchschnittliche Lesekompetenz von Jugendlichen an Gymnasien mehr als eine Standardabweichung (120 Punkte) über dem Mittelwert der Fünfzehnjährigen an nicht gymnasialen Schularten sowie etwa eine Kompetenzstufe (80 Punkte) über dem Gesamtmittelwert. Zudem zeigt die Abbildung die großen Streuungen innerhalb des Gymnasiums sowie innerhalb der nicht gymnasialen Schularten. Wie zu erwarten, findet sich innerhalb der nicht gymnasialen Schularten eine größere Streuung der Lesekompetenz ($SD = 93$) als innerhalb des Gymnasiums ($SD = 76$). Dies ist nicht verwunderlich, da bei den nicht gymnasialen Schularten mehrere Bildungsgänge zusammengefasst werden. Aber auch zwischen den Jugendlichen an Gymnasien gibt es beträchtliche Unterschiede. Das kommt beispielsweise dadurch zum Ausdruck, dass sich die mittleren 50 Prozent auf die Kompetenzstufen III bis IV verteilen. Zudem überschneiden sich die Verteilungen der Lesekompetenz der Fünfzehnjährigen an nicht gymnasialen Schularten und dem Gymnasium (siehe Abbildung 3.7).

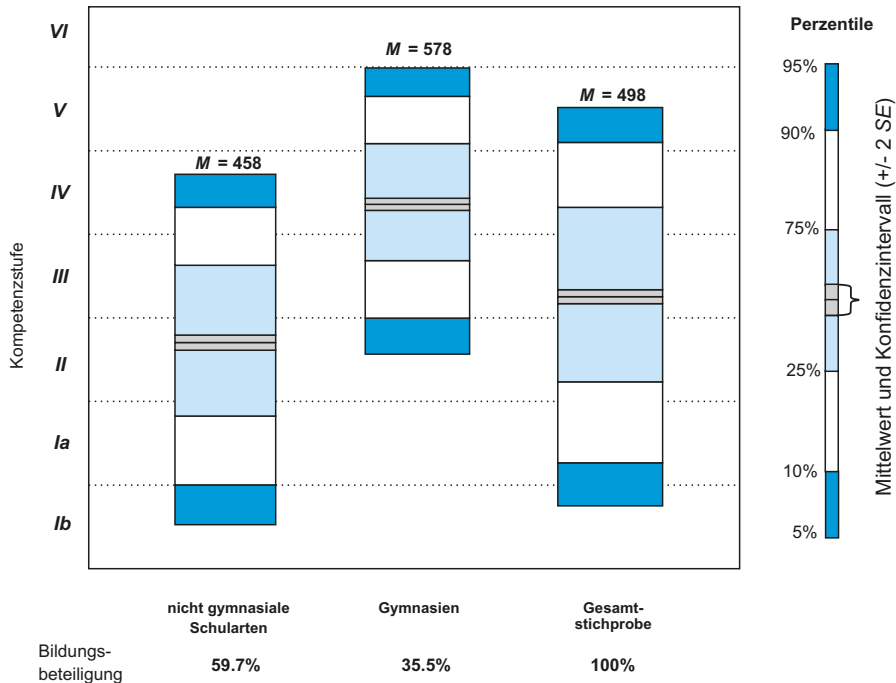


Abbildung 3.7: Perzentilbänder für die Lesekompetenz für die Gesamtstichprobe und nach Schulart in Deutschland

Teilskalen der Lesekompetenz: Prozesse des Lesens und Textquellen

Auch in den einzelnen Teilskalen der Lesekompetenz unterscheiden sich sowohl die mittleren Leistungen als auch die Streuungen stark zwischen Jugendlichen an nicht gymnasialen Schularten und Jugendlichen am Gymnasium. Zu den Teilskalen zählen hier sowohl die *Prozesse des Lesens* (*Lokalisieren von Informationen, Textverstehen, Bewerten und Reflektieren*) als auch die *Textquellen* (*eine Textquelle, mehrere Textquellen*). Während die Fünfzehnjährigen an nicht gymnasialen Schulen in allen fünf Teilskalen mittlere Kompetenzen zwischen 453 und 456 Punkten erreichen, liegen die Mittelwerte der Gymnasiastinnen und Gymnasiasten zwischen 576 und 582 Punkten (siehe Tabelle 3.2). Die Streuungen sind wie zu erwarten für alle fünf Teilskalen der Lesekompetenz an nicht gymnasialen Schulen deutlich größer als am Gymnasium.

Tabelle 3.2: Mittelwerte und Streuungen der Teilskalen der Lesekompetenz nach Schulart in Deutschland

	nicht gymnasiale Schularten			Gymnasium		
	M	(SE)	SD	M	(SE)	SD
Prozesse des Lesens						
Lokalisieren von Informationen	456	(3,9)	100	582	(3,2)	83
Textverstehen	454	(3,3)	96	576	(3,0)	78
Bewerten und Reflektieren	456	(3,6)	96	580	(3,1)	79
Textquellen						
Eine Textquelle	453	(3,6)	98	577	(3,0)	79
Mehrere Textquellen	455	(3,4)	94	579	(3,0)	78
Gesamtskala	458	(3,3)	93	578	(2,7)	76

Verteilung auf die Stufen der Lesekompetenz

Um ein vollständigeres Bild über die Unterschiede zwischen den Schularten zu bekommen, ist es aufschlussreich die prozentualen Anteile der Jugendlichen auf den einzelnen Kompetenzstufen getrennt nach Schulart zu betrachten (siehe Abbildung 3.8). Dabei zeigen sich deutlich unterschiedliche Anteile auf den verschiedenen Kompetenzstufen. So befinden sich beispielsweise lediglich zwei Prozent der Gymnasiastinnen und Gymnasiasten auf den untersten Kompetenzstufen (Ia oder darunter), aber 29 Prozent der Fünfzehnjährigen an nicht gymnasialen Schularten. Dies ist ein relativ hoher Anteil von Jugendlichen, der über sehr eingeschränkte Lesekompetenzen verfügt und somit

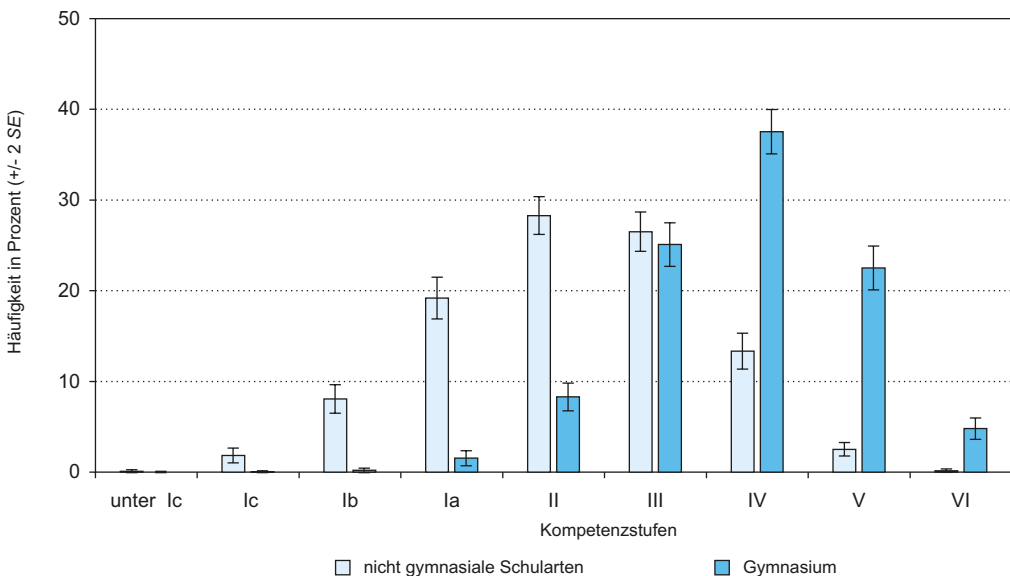


Abbildung 3.8: Prozentuale Anteile der Fünfzehnjährigen auf den Stufen der Lesekompetenz nach Schulart in Deutschland

schlechte Voraussetzungen im Hinblick auf die zukünftige Schul- und Berufslaufbahn mitbringt. Die Gruppe der besonders lesestarken Jugendlichen ist hingegen an nicht gymnasialen Schularten im Vergleich zum Anteil an den Gymnasien relativ klein. Während sich 27 Prozent der Gymnasiastinnen und Gymnasiasten auf den höchsten Kompetenzstufen (V und VI) befinden, sind es nur drei Prozent der Jugendlichen an nicht gymnasialen Schularten. Damit verfügt gut ein Viertel der Fünfzehnjährigen an Gymnasien über sehr hohe Kompetenzen, um Texte verschiedener Medien zu verstehen, zu nutzen, zu bewerten und über sie zu reflektieren. Diese Jugendlichen bringen somit ausgezeichnete Voraussetzungen für die weitere schulische oder berufliche Ausbildung mit.

3.3.2 Unterschiede zwischen Mädchen und Jungen in der Lesekompetenz in Deutschland

Da in Deutschland Mädchen mit durchschnittlich 512 Punkten eine deutlich höhere Lesekompetenz zeigen als Jungen mit durchschnittlich 486 Punkten (siehe auch Absatz 3.2.4), wird im Folgenden dieser Geschlechterunterschied für Deutschland differenziert aufgeschlüsselt. Abbildung 3.9 stellt die prozentualen Anteile auf den Kompetenzstufen getrennt für Mädchen und Jungen dar. Es fällt auf, dass der Anteil leseschwacher Mädchen deutlich kleiner ist als der Anteil leseschwacher Jungen. Während sich 16 Prozent der Mädchen auf den untersten Kompetenzstufen (Ia oder darunter) befinden, sind es 24 Prozent der Jungen. Genauso ist bei den lesestarken Jugendlichen der Anteil der

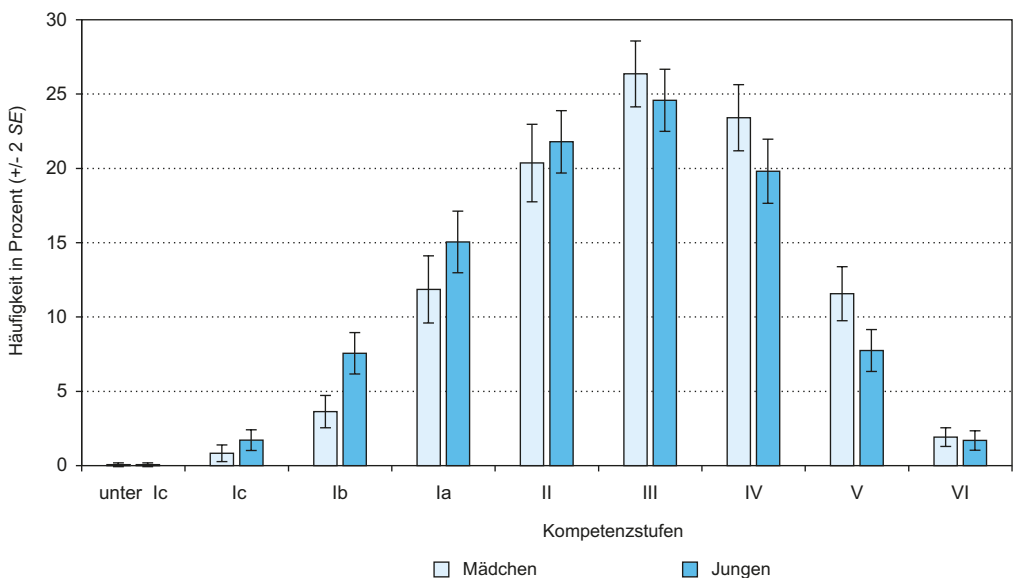


Abbildung 3.9: Prozentuale Anteile auf den Stufen der Lesekompetenz nach Geschlecht in Deutschland

Mädchen auf den obersten Kompetenzstufen V und VI mit 13 Prozent deutlich höher als der Anteil der Jungen mit neun Prozent.

3.3.3 Veränderung der Lesekompetenz im Vergleich zu PISA 2009 und PISA 2015

Wie man in Abbildung 3.10 sieht, unterscheidet sich die mittlere Lesekompetenz der Fünfzehnjährigen in Deutschland bei der PISA-Studie 2018 nicht signifikant von den Durchschnittswerten der Lesekompetenz in Deutschland bei den PISA-Studien 2009 und 2015. Damit gehört Deutschland zu einer Gruppe von elf Staaten, bei denen keine statistisch bedeutsamen Veränderungen der Lesekompetenz zwischen PISA 2009 und PISA 2018 sowie zwischen PISA 2015 und PISA 2018 bestehen. Zu dieser Gruppe zählen sowohl lesestarke (z. B. Kanada, Schweden) als auch leseschwache Staaten (z. B. Mexiko, Chile). Wie in Deutschland liegt die mittlere Lesekompetenz der Jugendlichen in Kanada, Schweden, den Vereinigten Staaten, Dänemark und Frankreich bei PISA 2018 signifikant über dem OECD-Durchschnitt, wobei es in diesen Staaten keine signifikanten Änderungen der Mittelwerte im Vergleich zu 2009 und 2015 gibt. Für Deutschland stellt dies einen positiven Verlauf dar, denn die mittlere Lesekompetenz der Jugendlichen in Deutschland lag bei der PISA-Studie 2000 noch signifikant unter dem OECD-Mittelwert. Im Vergleich zu PISA 2000 ist der Mittelwert der Lesekompetenz in Deutschland von 484 Punkten auf 498 Punkte im Jahr 2018 signifikant gestiegen. Die OECD ordnet daher Deutschland neben Chile, Israel, Polen und Kolumbien einer Gruppe von OECD-Staaten mit einem positiven durchschnittlichen Langzeittrend zu, der allerdings abflachend ist (OECD, 2019b). Die Ergebnisse liefern Hinweise darauf, dass sich die durchschnittliche Lesekompetenz der Jugendlichen in Deutschland auf einem hohen Niveau über dem OECD-Mittelwert etabliert hat (siehe Abbildung 3.11).

Dass dies keinesfalls selbstverständlich ist, zeigt der Blick auf andere OECD-Staaten. In 13 OECD-Staaten sind die Mittelwerte der Lesekompetenz bei PISA 2018 im Vergleich zu 2009 signifikant gesunken. Die deutlichsten Rückgänge der mittleren Lesekompetenz zeigen sich in Island, Griechenland, der Republik Korea und den Niederlanden. Während der Mittelwert der Lesekompetenz in den Niederlanden 2009 noch über dem OECD-Durchschnitt lag, unterscheidet er sich 2018 nicht mehr signifikant vom OECD-Durchschnitt. Auch in der Schweiz und in Belgien fiel die mittlere Lesekompetenz der Jugendlichen im Vergleich zu 2009 signifikant ab.

In sieben OECD-Staaten stieg die Lesekompetenz der Fünfzehnjährigen seit 2009 hingegen an. Dabei ist der Zuwachs um 22 Punkte in Irland und Estland am höchsten. In Österreich verbesserte sich die mittlere Lesekompetenz der Jugendlichen im Vergleich zu 2009 um 14 Punkte. Die Türkei ist der einzige OECD-Staat, in dem sich die Lesekompetenz bei PISA 2018 im Vergleich zu 2015 verbesserte.

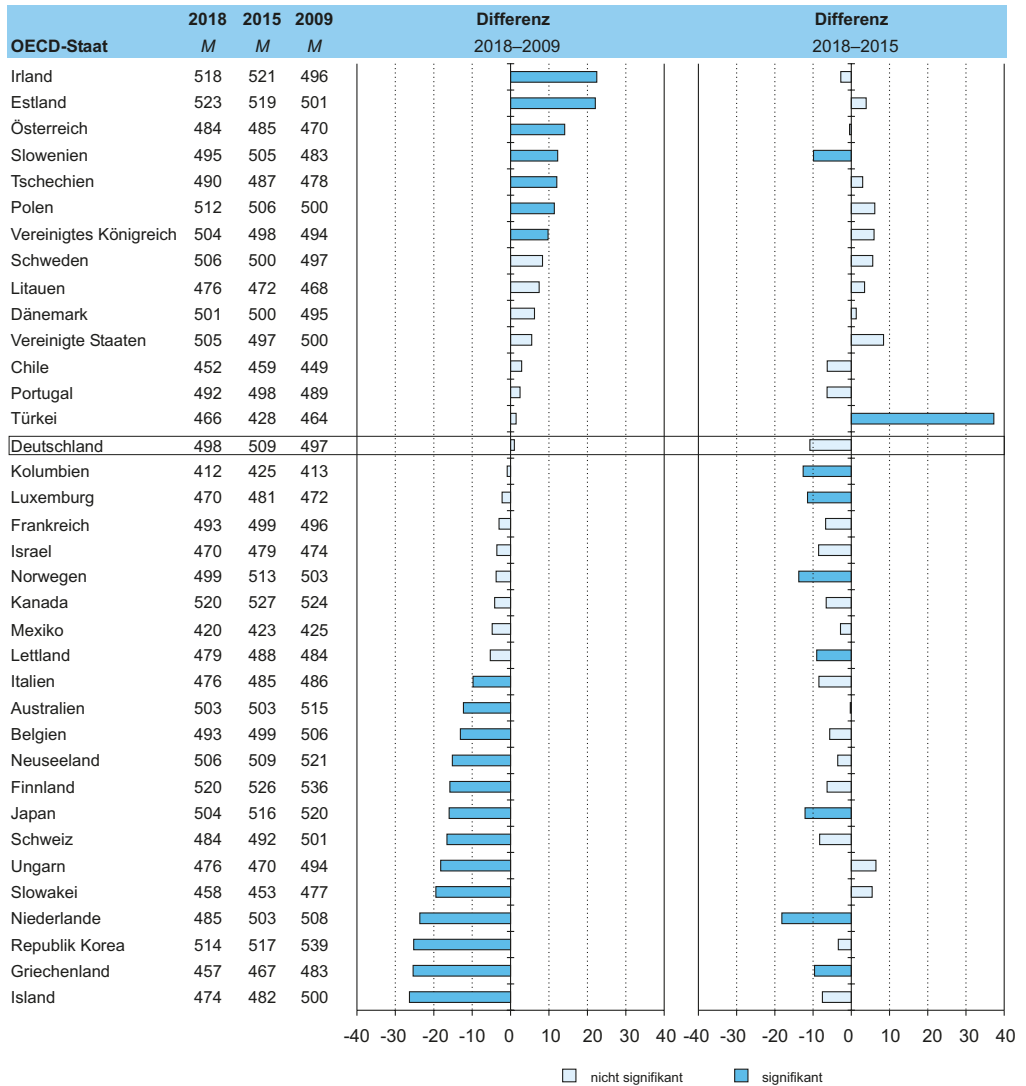


Abbildung 3.10: Mittelwerte der Lesekompetenz in PISA 2009, 2015 und 2018 in den OECD-Staaten

Anmerkung: Es werden hier die Staaten einbezogen, die bei PISA 2018 OECD-Staaten sind. Für die Standardfehler siehe Tabelle 3.7.web.

Bei der Betrachtung der Kompetenzstufen fällt auf, dass sich in Deutschland der Anteil der Jugendlichen auf den obersten Kompetenzstufen (V und VI) bei PISA 2018 im Vergleich zum Jahr 2009 signifikant vergrößert hat (um 4 Prozentpunkte). Allerdings ist im Vergleich zu 2015 auch der Anteil auf den untersten Kompetenzstufen (Ia oder darunter) bedeutsam angestiegen (um 4 Prozentpunkte).

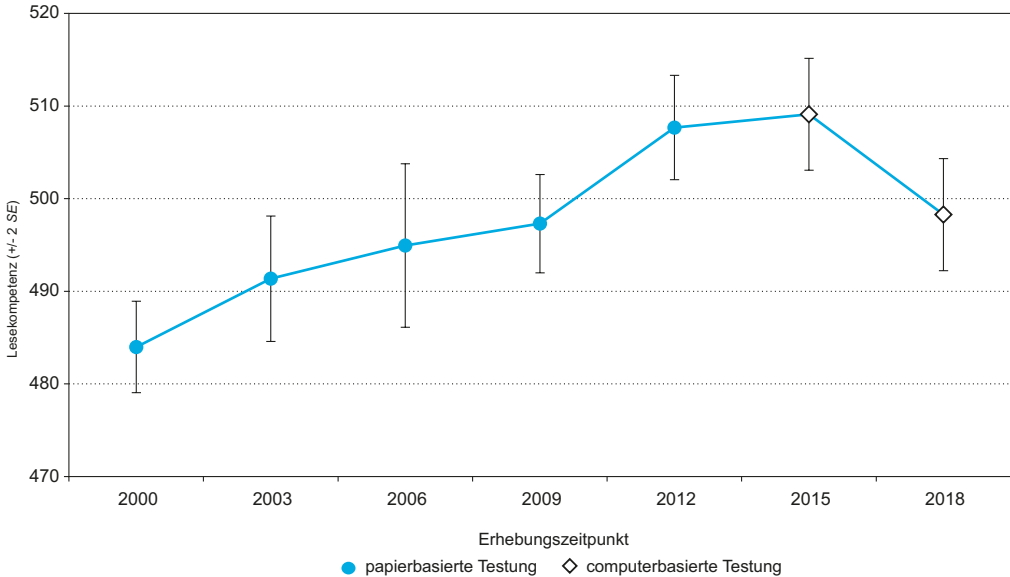


Abbildung 3.11: Veränderung der mittleren Lesekompetenz von 2000 bis 2018 in Deutschland

Betrachtet man die Lesekompetenz bei PISA 2018 im Vergleich zu 2015 differenziert nach Schulart, zeigt sich eine signifikante Verringerung der mittleren Lesekompetenz der Jugendlichen an nicht gymnasialen Schularten (siehe Tabelle 3.3). Im Vergleich zu 2009 gibt es bei den Schülerinnen und Schülern an nicht gymnasialen Schularten keine signifikante Änderung. Für die Fünfzehnjährigen an Gymnasien zeigen sich weder im Vergleich zu 2009 noch zu 2015 signifikante Veränderungen in der mittleren Lesekompetenz. Auch hat sich der Anteil der Gymnasiastinnen und Gymnasiasten auf den untersten Kompetenzstufen (Ia oder darunter) weder im Vergleich zu 2009 noch zu 2015 signifikant verändert (siehe Abbildung 3.12). Allerdings hat sich dieser Anteil an nicht gymnasialen Schularten sowohl im Vergleich zum Jahr 2009 als auch zum Jahr 2015 signifikant erhöht (zu 2009 um 6 Prozentpunkte, zu 2015 um 8 Prozentpunkte). Das bedeutet konkret, dass insbesondere an nicht gymnasialen Schularten mehr Jugendliche nur über sehr eingeschränkte Lesekompetenzen verfügen und nicht hinreichend für eine weitere Ausbildung vorbereitet sind. Umgekehrt ist die Spitzengruppe der Gymnasiastinnen und Gymnasiasten (Kompetenzstufen V und VI) im Vergleich zu 2009 signifikant gewachsen (um 7 Prozentpunkte). Diese Gruppe der besonders lesestarken Jugendlichen an Gymnasien verfügt über ein sehr hohes Niveau der Lesekompetenz und somit über optimale Voraussetzungen für zukünftiges selbstständiges Lernen mittels Lesen sowie für eine gelingende gesellschaftliche Teilhabe.

Die mittlere Lesekompetenz der Mädchen ist sowohl bei PISA 2018 als auch bei PISA 2009 und 2015 deutlich höher als die mittlere Lesekompetenz der Jungen (siehe auch Abschnitt 3.2.3). Diese Geschlechterdifferenz hat sich im Jahr 2018 im Vergleich zu PISA 2009 zwar verringert, aber im Vergleich zu 2015 nicht signifikant verändert.

Tabelle 3.3: Mittelwerte und Streuungen der Gesamtskala Lesekompetenz nach Geschlecht und Schulart bei PISA 2018, 2015 und 2009 in Deutschland

	2018				2015				2009			
	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)
Mädchen	512	(3,2)	101	(1,9)	520	(3,1)	98	(1,9)	518	(2,9)	90	(1,9)
Jungen	486	(3,4)	108	(1,8)	499	(3,7)	101	(1,9)	478	(3,6)	95	(2,1)
Nicht gymnasiale Schularten	458	(3,3)	93	(1,6)	478	(3,4)	88	(1,8)	467	(2,9)	81	(1,7)
Gymnasium	578	(2,7)	76	(1,6)	583	(3,0)	73	(1,8)	575	(2,9)	61	(1,5)
Gesamtstichprobe	498	(3,0)	106	(1,5)	509	(3,0)	100	(1,6)	497	(2,7)	95	(1,8)

Bei separater Betrachtung der Lesekompetenz der Mädchen und Jungen zeigt sich allerdings 2018 eine signifikante Verringerung der mittleren Lesekompetenz der Jungen um 13 Punkte im Vergleich zu 2015. Im Vergleich zu 2009 hat sich die Lesekompetenz der Jungen nicht signifikant verändert. Dies spricht gegen die Vermutung, welche nach der PISA-Erhebungsrunde 2015 nahelag, dass die Änderung vom papierbasierten auf das computerbasierte Testen zur Steigerung der Lesekompetenz der Jungen geführt haben könnte (vgl. Weis et al., 2016). Andererseits besteht im Vergleich zu den Erhebungsrounden 2000 bis 2012 weiterhin eine verringerte Geschlechterdifferenz seit PISA 2015. Zwischen 2000 und 2012 war die Geschlechterdifferenz im Lesen in Deutschland durchwegs sehr hoch ausgeprägt (35 und 44 Punkte) und bei PISA 2015 hat sich diese Differenz

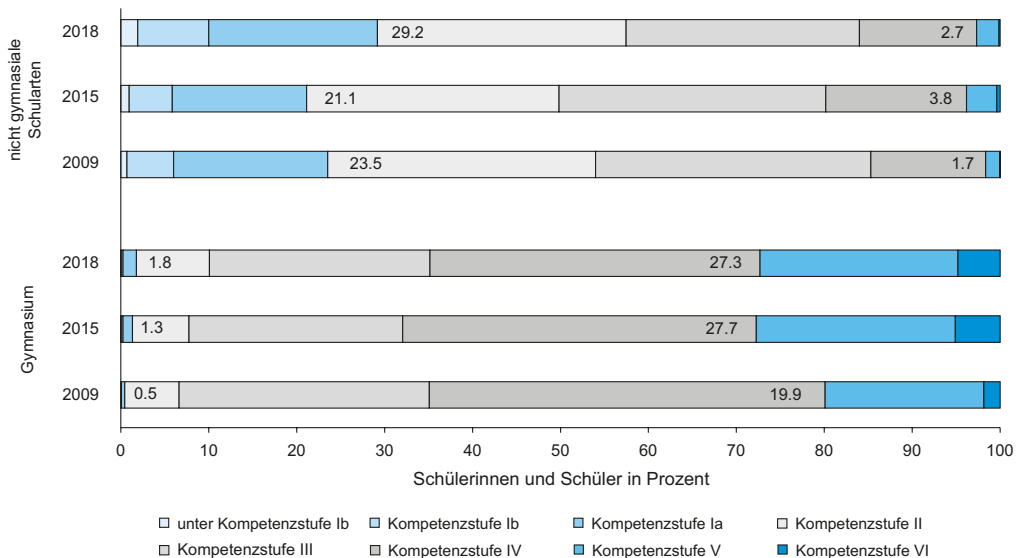


Abbildung 3.12: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe Ia, Ib, Ic oder darunter sowie auf Kompetenzstufe V oder VI für die Gesamtskala Lesekompetenz bei PISA 2018, 2015 und 2009 in Deutschland

in etwa halbiert (21 Punkte). Die mittlere Lesekompetenz der Mädchen bei PISA 2018 unterscheidet sich weder signifikant zu den Mittelwerten von 2009 noch von 2015.

Analog hat sich der Anteil der besonders legeschwachen Jungen (Kompetenzstufen Ia oder darunter) seit dem Jahr 2009 nicht verändert (24% bei PISA 2018 und 2009), und es gibt somit weiterhin einen relativ hohen Anteil an Jungen, die nur über äußerst eingeschränkte Lesekompetenz verfügen. Im Vergleich zu PISA 2015 hat sich der Anteil der Jungen auf den untersten Kompetenzstufen 2018 sogar erhöht (19% in 2015). Der Anteil der Mädchen auf den untersten Kompetenzstufen hat sich hingegen weder im Vergleich zu 2009 noch zu 2015 signifikant verändert. Auf den obersten Kompetenzstufen hat sich der Geschlechterunterschied seit 2009 beträchtlich verringert. Bei PISA 2018 ist der Anteil der lesestarken Jungen (9%) mehr als doppelt so hoch wie im Jahr 2009 (4%). Bei den Mädchen hat sich dieser Anteil weder im Vergleich zu 2009 noch zu 2015 signifikant verändert.

3.4 Zusammenfassung und Diskussion

PISA 2018 hat für Deutschland die gute Nachricht, dass die Lesekompetenz der Fünfzehnjährigen im Vergleich zum OECD-Mittelwert überdurchschnittlich ausgeprägt ist. Die Leistungen sind signifikant besser als im Mittel der OECD-Staaten und in lediglich sieben Staaten liegt der Mittelwert der Lesekompetenz bedeutsam über dem Mittelwert von Deutschland (z.B. in Estland, Kanada, Finnland und Irland). Allerdings setzt sich der kontinuierliche Aufwärtstrend der Lesekompetenz der Jugendlichen in Deutschland, der sich im Zeitraum zwischen 2000 und 2015 zeigte (vgl. OECD, 2016, Weis et al., 2016), nicht mehr in gleicher Weise fort. Beschränkt man sich auf Deutschland, so unterscheidet sich die mittlere Lesekompetenz der Fünfzehnjährigen 2018 nicht signifikant von den Mittelwerten aus 2009 und 2015. Sie liegt allerdings deutlich höher als bei PISA 2000. Somit lässt sich der Verlauf von 2000 bis 2018 für Deutschland als positiver durchschnittlicher Langzeittrend beschreiben (OECD, 2019b).

Erfreulich ist darüber hinaus, dass sich der Anteil der besonders lesestarken Jugendlichen in Deutschland bei PISA 2018 im Vergleich zum Jahr 2009 signifikant vergrößert hat. Es zeigt sich, dass insbesondere die Spitzengruppe der Gymnasiastinnen und Gymnasiasten signifikant gewachsen ist. Es gibt in Deutschland auch im Vergleich zum OECD-Durchschnitt einen relativ großen Anteil an Fünfzehnjährigen, die hochkompetent im Lesen sind und optimale Voraussetzungen für ein selbstständiges Lernen mittels Lesen mitbringen. Es gilt hier, in der Förderung dieser Spitzengruppe nicht nachzulassen und so hoffentlich ihren Anteil aufrechterhalten und ausbauen zu können.

Eine nicht so gute Nachricht ist die besonders große Streuung der Lesekompetenz in Deutschland und zwar sowohl im internationalen Vergleich als auch im Vergleich zur Streuung in Deutschland in den Jahren 2009 und 2015. Sie weist auf eine höhere Diversität der Lesekompetenz zwischen den Jugendlichen in Deutschland hin. Sicherlich ist es erfreulich, dass es eine relativ große Gruppe lesestarker Schülerinnen und Schüler gibt.

Wenig erfreulich ist aber der große Anteil sehr leseschwacher Schülerinnen und Schüler. Mehr als ein Fünftel der Fünfzehnjährigen zeigt Kompetenzen unterhalb der Kompetenzstufe II und ist damit kaum in der Lage, sinnerfassend mit Texten umzugehen. Dieser Anteil ist seit PISA 2015 signifikant gestiegen, was als ein ernstes Signal verstanden werden muss. Verbesserungen sollten hier möglich sein, denn gerade in den sehr lesestarken Staaten gibt es weniger Schülerinnen und Schüler mit ähnlich schwachen Leistungen.

Insbesondere an nicht gymnasialen Schularten ist der Anteil besonders leseschwacher Jugendlicher (auf den Kompetenzstufen Ia oder darunter) im Vergleich zu 2009 und zu 2015 in Deutschland bedeutsam angestiegen. Dass sich 29 Prozent der Jugendlichen an nicht gymnasialen Schularten bei PISA 2018 auf den untersten Kompetenzstufen befinden und demnach nur über sehr eingeschränkte Lesekompetenzen verfügen, ist ein besorgniserregender Befund. Die relativ hohen Anteile an besonders leseschwachen Jugendlichen in Deutschland, die sich insbesondere an nicht gymnasialen Schularten befinden, sollten zur Konsequenz haben, dass leseschwache Kinder und Jugendliche in Deutschland noch stärker gefördert werden. Besonders wirksam scheint die durchgängige Leseförderung vom Vorschulbereich bis zum Ende der Schulzeit zu sein (Artelt et al., 2007; Beck, von Dewitz & Titz, 2015). Zudem zeigt die Leseforschung, dass beispielsweise die Leseflüssigkeit durch den Einsatz von Lautleseverfahren und das Textverstehen durch die Vermittlung von Lesestrategien gefördert werden können (Gold, 2018; Lenhard, 2013). Es sind weitere Anstrengungen notwendig, um insbesondere Risikogruppen, wozu alle Kinder und Jugendlichen mit einer festgestellten Leseschwäche zählen, gezielt in ihrer Sprach- und Lesekompetenz zu fördern (Gold & Dubowy, 2013; Kreyenfeld & Krapf, 2013; Schröder, Keller, Dintsioudi & List, 2013).

In Deutschland sind somit insgesamt weiterhin Bemühungen notwendig, um die Leseförderung sowie die Förderung der sprachlichen Bildung insgesamt zu intensivieren. Dies machen insbesondere die Befunde zum hohen Anteil der leseschwachen Jugendlichen an nicht gymnasialen Schularten in Deutschland deutlich. Es ist auch weitere empirische Forschung notwendig, sowohl in Form von Grundlagenforschung als auch zur Entwicklung, zur Evaluation und zum Transfer wirksamer Interventionen zur Förderung der Sprach- und Lesekompetenz, um eine evidenzbasierte Schul- und Unterrichtsentwicklung zu ermöglichen. Ein positives Beispiel ist die Bund-Länder-Initiative BiSS (Bildung durch Sprache und Schrift), die als deutschlandweites Programm zur Lese- und Sprachförderung in allen Ländern von der frühkindlichen Bildung bis zur Sekundarstufe umgesetzt wird (Schneider et al., 2012). Derartige Initiativen, Programme und Forschungsprojekte sind notwendig, um wirksame Förderkonzepte (weiter) zu entwickeln und in die Fläche zu transferieren.

Die Ergebnisse legen nahe, dass eine Förderung der besonders leseschwachen Schülerinnen und Schüler über alle Schularten hinweg in den Blick genommen werden muss. Dabei darf nicht in Vergessenheit geraten, dass eine Unterstützung der Leistungsspitze eine zentrale Aufgabe bleibt.

Literatur

- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., ... Saalbach, H. (2007). *Förderung von Lesekompetenz – Expertise*. Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- 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, ... M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_4
- Beck, L., von Dewitz, N. & Titz, C. (2015). Durchgängige Leseförderung. *BiSS-Journal, Bildung durch Sprache und Schrift*, 3, 4–11. Verfügbar unter <https://biss-sprachbildung.de/pdf/biss-journal-3-november-2015-1.pdf>
- Cheung, K., Mak, S. & Sit, P. (2013). Online reading activities and ICT use as mediating variables in explaining the gender difference in digital reading literacy: Comparing Hong Kong and Korea. *The Asia-Pacific Education Researcher*, 22, 709–720. <https://doi.org/10.1007/s40299-013-0077-x>
- Gil-Flores, J., Torres-Gordillo, J.-J. & Perera-Rodríguez, V.-H. (2012). The role of online reader experience in explaining students' performance in digital reading. *Computers & Education*, 59, 653–660. <https://doi.org/10.1016/j.compedu.2012.03.014>
- Gold, A. (2018). *Lesen kann man lernen. Wie man Lesekompetenz fördern kann* (3. Aufl.). Göttingen: Vandenhoeck & Ruprecht. <https://doi.org/10.13109/9783666310638>
- Gold, A. & Dubowy, M. (2013). *Frühe Bildung. Lernförderung im Elementarbereich*. Stuttgart: Kohlhammer.
- Hahnel, C., Goldhammer, F., Kröhne, U. & Naumann, J. (2017). Reading digital text involves working memory updating based on task characteristics and reader behavior. *Learning and Individual Differences*, 59, 149–157. <https://doi.org/10.1016/j.lindif.2017.09.001>
- Hahnel, C., Goldhammer, F., Naumann, J. & Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Computers in Human Behavior*, 55, 486–500. <https://doi.org/10.1016/j.chb.2015.09.042>
- Heine, J.-H., Mang, J., Borchert, L., Gomolka, J., Kröhne, U., Goldhammer, F. & Sälzer, C. (2016). Kompetenzmessung in PISA 2015. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 383–430). Münster: Waxmann.
- Jude, N. & Klieme, E. (2010). Das Programme for International Student Assessment (PISA). In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73–112). Münster: Waxmann.
- Kreyenfeld, M. & Krapf, S. (2013). Soziale Ungleichheit und Kinderbetreuung – Eine Analyse der sozialen und ökonomischen Determinanten der Nutzung von Kindertageseinrichtungen. In R. Becker & W. Lauterbach (Hrsg.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (4. Aufl., S. 107–128). Wiesbaden: Springer. https://doi.org/10.1007/978-3-531-92484-7_4
- Lawless, K. A. & Schrader, P. G. (2008). Where do we go now? Understanding research on navigation in complex digital environments. In D. J. Leu & J. Coiro (Hrsg.), *Handbook of New Literacies* (S. 267–296). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lee, Y.-H. & Wu, J.-Y. (2012). The effect of individual differences in the inner and outer states of ICT on engagement in online reading activities and PISA 2009 reading literacy:

- Exploring the relationship between the old and new reading literacy. *Learning and Individual Differences*, 22, 336–342. <https://doi.org/10.1016/j.lindif.2012.01.007>
- Lenhard, W. (2013). *Leseverständnis und Lesekompetenz. Grundlagen – Diagnostik – Förderung*. Stuttgart: Kohlhammer.
- Lim, H. J. & Jung, H. (2019). Factors related to digital reading achievement: A multi-level analysis using international large scale data. *Computers & Education*, 133, 82–93. <https://doi.org/10.1016/j.compedu.2019.01.007>
- Mang, J., Wagner, S., Gomolka, J., Schäfer, A., Meinck, S. & Reiss, K. (2019). *Technische Hintergrundinformationen PISA 2018*. München: Technische Universität München. <https://doi.org/10.14459/2019md1518258>
- Naumann, J. (2013). Online-Leseengagement: Vorhersage von Navigationsverhalten und Textverstehen bei Online-Texten. In C. Rosebrock & A. Bertschi-Kaufmann (Hrsg.), *Literalität erfassen: bildungspolitisch, kulturell, individuell* (S. 188–202). Weinheim: Beltz Juventa.
- Naumann, J. (2015). A model of online reading engagement: Linking engagement, navigation, and performance in digital reading. *Computers in Human Behavior*, 53, 263–277. <https://doi.org/10.1016/j.chb.2015.06.051>
- 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, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23–71). Münster: Waxmann.
- Naumann, J. & Salmerón, L. (2016). Does Navigation Always Predict Performance? Effects of Navigation on Digital Reading are Moderated by Comprehension Skills. *The International Review of Research in Open and Distributed Learning*, 17, 42–59. <https://doi.org/10.19173/irrodl.v17i1.2113>
- OECD. (2010). *PISA 2009. Learning trends. Changes in student performance since 2000*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264091580-en>
- OECD. (2016). *PISA 2015 Results (Volume I): Excellence and equity in education*. Paris: OECD Publishing.
- OECD. (2019a). PISA 2018 Reading framework. In OECD (Hrsg.), *PISA 2018 assessment and analytical framework* (S. 21–71). Paris: OECD Publishing. <https://doi.org/10.1787/b25e-fab8-en>
- OECD. (2019b). *PISA 2018 Results (Volume I): What students know and can do*. Paris: OECD Publishing.
- Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, M. & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S. Schroeder & P. van den Broek (Hrsg.), *Learning to read in a digital world* (S. 91–120). Amsterdam: John Benjamins. <https://doi.org/10.1075/swll.17.04sal>
- Sälzer, C. & Reiss, K. (2016). PISA 2015 – die aktuelle Studie. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 13–44). Münster: Waxmann.
- Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, ... Stanat, P. (2012). *Expertise „Bildung durch Sprache und Schrift (BiSS)“: Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung*. Verfügbar unter https://www.bmbf.de/files/BISS_Expertise.pdf
- Schröder, L., Keller, H., Dintsioudi, A. & List, M. (2013). *Sprachliche Bildung im Kita-Alltag. Gespräche mit Kindern anregen und lebendig gestalten*. Berlin: Cornelsen.

- Weis, M., Zehner, F., Sälzer, C., Strohmaier, A., Artelt, C. & Pfof, M. (2016). Lesekompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 249–283). Münster: Waxmann.
- Wu, J. Y. & Peng, Y.-C. (2017). The modality effect on reading literacy: Perspectives from students' online reading habits, cognitive and metacognitive strategies, and web navigation skills across regions. *Interactive Learning Environments*, 25, 859–876. <https://doi.org/10.1080/10494820.2016.1224251>

4 Lesebezogene Schülermerkmale in PISA 2018: Motivation, Leseverhalten, Selbstkonzept und Lesestrategiewissen

Jennifer Diedrich, Anja Schiepe-Tiska, Lisa Ziernwald,
Ana Tupac-Yupanqui, Mirjam Weis, Nele McElvany & Kristina Reiss

Lesemotivation, Leseverhalten, lesebezogenes Selbstkonzept und Lesestrategiewissen sind neben der Lesekompetenz wichtige Indikatoren einer erfolgreichen schulischen Leseförderung. PISA 2018 zeigt, dass im Vergleich zum OECD-Durchschnitt die Lesefreude und das Leseverhalten in Deutschland geringer ausgeprägt sind. Schülerinnen und Schüler geben an, dass sie vor allem lesen, um Informationen zu finden. Das lesebezogene Selbstkonzept ist in Deutschland hingegen überdurchschnittlich ausgeprägt, was in der Zusammenschau mit der Lesekompetenz eine realistische Einschätzung der eigenen Fähigkeiten widerspiegelt. Darüber hinaus zeichnen sich die Jugendlichen in Deutschland im internationalen Vergleich durch das höchste Lesestrategiewissen aus. Mädchen berichten positivere Ausprägungen in der Lesemotivation, dem Leseverhalten und der wahrgenommenen Kompetenz sowie im Lesestrategiewissen als Jungen, ein Muster, welches sich im Großteil der OECD-Staaten zeigt. Lediglich in der wahrgenommenen Schwierigkeit beim Lesen lassen sich in Deutschland keine Geschlechterunterschiede finden. Darüber hinaus lesen Jugendliche an Gymnasien lieber und häufiger zum Vergnügen und berichten ein höheres Selbstkonzept als Fünfzehnjährige nicht gymnasialer Schularten. Betrachtet man die Entwicklung der Lesemotivation und der Lesemenge im Vergleich zu PISA 2009, haben sich diese, ebenso wie im OECD-Durchschnitt, signifikant verringert. Nahezu alle Jugendlichen (96.3%) nutzen aber mehrmals wöchentlich oder täglich Chats und auch Nachrichten werden von mehr Jugendlichen online als papierbasiert gelesen. Das Lesestrategiewissen beim Erinnern von Texten hat im Vergleich zu PISA 2009 abgenommen. Beim Zusammenfassen von Texten hat es sich jedoch nicht verändert und auf hohem Niveau stabilisiert. Für die Mädchen ist das Wissen um erfolgreiche Lesestrategien beim Zusammenfassen von Texten sogar angestiegen.

In unserer schriftsprachlich geprägten Gesellschaft bildet Lesen ein bedeutsames Fundament, um Informationen aufzunehmen, zu verarbeiten und weiterzugeben (OECD, 2019b). Eine ausgeprägte Lesemotivation und ein ausgeprägtes Leseverhalten sowie positive Selbstbilder und Strategien zum Lesen stellen neben der Lesekompetenz wichtige Ziele schulischer Bildung dar. Gemeinsam beeinflussen diese mehrdimensionalen Bildungsziele die Bereitschaft, lebenslang zu lernen, und ermöglichen eine fundierte, aktive

Teilhabe an gesellschaftlichen Diskursen (Blossfeld et al., 2015; Schiepe-Tiska, in Druck; UNESCO, 2006).

Viele der lesebezogenen Schülermerkmale weisen wechselseitige Zusammenhänge mit der Lesekompetenz sowie untereinander auf. Eine hohe Lesemotivation führt in der Regel zu häufigerem Lesen, was wiederum zu einer höheren Lesekompetenz führt (Klauda & Guthrie, 2015). Der Zusammenhang zwischen Lesekompetenz und Interesse am Lesen wird zudem durch das lesebezogene Selbstkonzept vermittelt (Viljaranta, Tolvanen, Aunola & Nurmi, 2014). Jugendliche, die gerne lesen, trauen sich beim Lesen mehr zu und wer sich mehr zutraut, zeigt zumeist eine höhere Lesekompetenz. In der Konsequenz werden über die Zeit hinweg gute Leserinnen und Leser noch besser, während schwächere Leserinnen und Leser noch schwächer werden (Pfof, Hattie, Dörfler & Artelt, 2014). Aus diesen wechselseitigen Zusammenhängen ergibt sich somit entweder eine negative (Teufelskreis) oder positive Rückkoppelung zwischen der Lesekompetenz und den lesebezogenen Schülermerkmalen (Pfof, Dörfler & Artelt, 2013).

Bereits der ersten PISA-Erhebung im Jahr 2000 lag ein mehrdimensionales Bildungsverständnis zugrunde, welches neben kognitiven Aspekten auch motivationale Orientierungen, Selbstbilder und Strategiewissen umfasste (Baumert, 2001). Diesen lesebezogenen Schülermerkmalen wurde mit jeder überarbeiteten Rahmenkonzeption der Lesekompetenz mehr Bedeutung zugemessen; in PISA 2018 sind sie ein wichtiger Bestandteil der Definition der Lesekompetenz (OECD, 2019b). Bei PISA 2000 berichteten Jugendliche in Deutschland einerseits ein niedrigeres Leseinteresse als Jugendliche in anderen OECD-Staaten, andererseits wendeten sie häufiger erfolgreiche Lernstrategien beim Lesen an (Artelt, Demmrich & Baumert, 2001). Auch in PISA 2009, als Lesen zum zweiten Mal Hauptdomäne war, zeigten Fünfzehnjährige in Deutschland eine im internationalen Vergleich unterdurchschnittliche Lesemotivation bei gleichzeitig hohem Lesestrategiewissen (Artelt, Naumann & Schneider, 2010). Ein wichtiges Anliegen dieses Kapitels ist es deshalb, diese Variablen weiter in ihrer Entwicklung zu betrachten.

Im Folgenden werden die empirischen Befunde zu den lesebezogenen Schülermerkmalen aus PISA 2018 dargestellt. Für Lesemotivation, Leseverhalten – insbesondere die Lesemenge, das Lesen zum Vergnügen und das Online-Lesen – und lesebezogenes Selbstbild sowie Lesestrategiewissen werden die Ergebnisse der Fünfzehnjährigen in Deutschland betrachtet und in den internationalen Vergleich eingeordnet. Für den Vergleich werden vor allem Staaten einbezogen, die bei PISA 2018 eine hohe Lesekompetenz gezeigt haben (vgl. Kapitel 3). Darüber hinaus werden Staaten berücksichtigt, die mit Deutschland kulturell oder sprachlich vergleichbar sind. Ein wesentlicher Grund dafür ist, dass die berichteten Schülermerkmale auf Selbsteinschätzungen der Jugendlichen beruhen, die mit Hilfe des Schülerfragebogens erfasst wurden. Diese Einschätzungen können durch den kulturellen Hintergrund der Schülerinnen und Schüler mit beeinflusst werden, und es kann etwa sein, dass Fünfzehnjährige grundsätzlich Aussagen eher zustimmen als sie abzulehnen, generell extreme Antwortkategorien vermeiden oder so antworten, wie sie meinen, dass es von ihnen erwartet wird (vgl. van de Vijver & He, 2017). Diese Aspekte sind vor allem dann wichtig, wenn Selbsteinschätzungen über

verschiedene Staaten hinweg verglichen werden. Aus deutscher Sicht ist darüber hinaus der Vergleich Fünfzehnjähriger nicht gymnasialer Schularten mit Jugendlichen des Gymnasiums interessant. Auch hier gilt es zu berücksichtigen, dass Jugendliche die Aussagen in der Regel in Abhängigkeit von ihrer Bezugsgruppe einschätzen (Heine, Lehman, Peng & Greenholtz, 2002). Fünfzehnjährige am Gymnasium vergleichen sich demnach eher mit ihren Schulkameradinnen und Schulkameraden am Gymnasium als mit Jugendlichen nicht gymnasialer Schularten. Schließlich konzentrieren sich die Analysen auf den Vergleich mit PISA 2009, um Ansatzpunkte für eine gezielte Förderung der lesebezogenen Schülermerkmale zu finden. Zur leichteren Orientierung in Bezug auf die betrachteten Konstrukte zeigt Tabelle 4.1 eine Übersicht.

Tabelle 4.1: Übersicht der lesebezogenen Schülermerkmale

Übergeordnetes Konstrukt	Merkmal	Beispielitem
Lesemotivation	Lesefreude	Wie sehr stimmst du diesen Aussagen über das Lesen zu? [...] „Lesen ist eines meiner liebsten Hobbys.“
	Lesemenge	Wie viel Zeit verbringst du normalerweise damit, zu deinem Vergnügen zu lesen? „Bis zu 30 Minuten täglich“
Leseverhalten	Lesen zum Vergnügen	Wie oft liest du zu deinem Vergnügen ... „Zeitschriften“
	Online-Lesen	Wie oft beschäftigst du dich mit den folgenden Leseaktivitäten? „Chatten (z. B. Whatsapp®)“
Lesebezogenes Selbstkonzept	Wahrgenommene Kompetenz	Wie sehr stimmst du folgenden Aussagen zu? „Ich lese flüssig“
	Wahrgenommene Schwierigkeit	Wie sehr stimmst du folgenden Aussagen zu? „Ich hatte schon immer Mühe mit dem Lesen“
Lesestrategie-wissen	Zusammenfassen von Informationen (<i>Zusammenfassen</i>)	siehe Abbildung 4.11
	Verstehen und Erinnern eines Textes (<i>Erinnern</i>)	siehe Abbildung 4.12
	Beurteilung der Qualität und Vertrauenswürdigkeit einer Online-Quelle (<i>Beurteilen</i>)	siehe Abbildung 4.13
	Gesamtwissensindikator	Summenwert der Paarvergleiche der drei Lernszenarien <i>Zusammenfassen</i> , <i>Erinnern</i> und <i>Beurteilen</i>

4.1 Lesemotivation, Leseverhalten und Selbstkonzept

Eine Facette der intrinsischen Motivation ist, ob eine Schülerin oder ein Schüler freiwillig liest und Freude daran hat (Wigfield, 1997). Wie andere Komponenten der Lesemotivation auch, hängt Lesefreude positiv mit der Lesekompetenz zusammen (Guthrie, Klauda & Ho, 2013): Umso mehr Freude Jugendliche am Lesen haben, desto höher ist

ihre Lesekompetenz und umgekehrt. Ob jedoch eine erhöhte Motivation allgemein mit einer besseren Kompetenz einhergeht, hängt wesentlich davon ab, wie viel gelesen wird. Außerdem zählt, was (z. B. Prosa, Sachbuch, Online-Nachrichten oder E-Mails) beziehungsweise warum (als Hausaufgabe oder zum Vergnügen; OECD, 2019a) gelesen wird. Wenig überraschend zeigt vor allem das Lesen zum Vergnügen, also außerhalb von Schulaufgaben, positive Zusammenhänge mit intrinsischer Motivation (Schiefele, Schaffner, Möller & Wigfield, 2012) und Lesekompetenz. Dabei spielt das Lesen von Zeitungen oder Comics eine geringere Rolle bei der Entwicklung der Lesekompetenz als das Lesen von Prosa (Pfost et al., 2013).

Auch ein positives Selbstkonzept beeinflusst die Bereitschaft, sich innerhalb und außerhalb der Schule mit unterschiedlichen Texten auseinanderzusetzen. Das Leseselbstkonzept ist Teil des akademischen Selbstkonzepts (Lafontaine, Dupont, Jaegers & Schillings, 2019) und setzt sich aus drei Facetten zusammen: lesebezogene Einstellungen, die wahrgenommene eigene Kompetenz sowie die wahrgenommenen Schwierigkeiten mit dem Lesen (Chapman & Tunmer, 1995). Insbesondere die selbsteingeschätzte wahrgenommene Schwierigkeit hängt besonders stark mit der Lesekompetenz zusammen (Chapman & Tunmer, 1995; Lafontaine et al., 2019). Außerdem ist diese Facette weniger anfällig für sozial erwünschtes Antwortverhalten als die wahrgenommene Kompetenz (Lafontaine et al., 2019).

4.1.1 Erfassung von Lesemotivation, Leseverhalten und Selbstkonzept in PISA 2018

Zur Erfassung der lesebezogenen Schülermerkmale bei PISA 2018 beantworteten die Schülerinnen und Schüler einen Fragebogen. Zumeist stufen die Fünfzehnjährigen ihre Wahrnehmung in Fragen wie „Wie sehr stimmst du den folgenden Aussagen zu?“ beziehungsweise Items wie „Für mich ist Lesen Zeitverschwendung“ auf Skalen von „stimme überhaupt nicht zu“ bis „stimme völlig zu“ ein. Wie eingangs erwähnt sind diese Einstufungen anfällig für kulturelle Einflüsse wie Antworttendenzen. Im Rahmen von PISA wird versucht, dieser Problematik auf der Ebene der Fragebogenkonstruktion, der Ebene der nationalen, sprachlichen Adaptierung und der Ebene der statistischen Auswertung zu begegnen. Dennoch sind kulturell bedingte Verzerrungen nicht vollkommen auszuschließen (OECD, 2019a). Im technischen Bericht zu PISA 2018 (OECD, in Vorbereitung) werden, wie bereits bei der PISA-Erhebung 2015 (OECD, 2017), ausführliche Informationen zu den verwendeten Fragen und ihrer kulturellen Vergleichbarkeit enthalten sein.

Erhoben wurde die *Lesefreude* bei PISA 2018 über fünf Aussagen (z. B. „Lesen ist eines meiner liebsten Hobbys“; Cronbachs $\alpha = .79^1$). Diesen Aussagen sollten die Fünf-

1 Die berichteten internen Konsistenzen beziehen sich auf die deutsche Stichprobe. Genauere Analysen der Skalen werden national (ZIB, in Vorbereitung) bzw. international (OECD, in Vorbereitung) in Zukunft berichtet werden.

zehnjährigen auf einer vierstufigen Skala von „stimme überhaupt nicht zu“ bis „stimme völlig zu“ bewerten.

Die *Lesemenge* wurde mit der Frage erfasst „Wie viel Zeit verbringst du normalerweise damit, zu deinem Vergnügen zu lesen?“. Die Jugendlichen konnten aus den folgenden Antwortoptionen eine auswählen: „Ich lese nicht zum Vergnügen“, „Bis zu 30 Minuten täglich“, „Mehr als 30 Minuten, aber weniger als 60 Minuten täglich“, „1 bis 2 Stunden täglich“ oder „Mehr als 2 Stunden täglich“. Für weiterführende Analysen wurden die Fünfzehnjährigen in drei Gruppen eingeteilt: „Viel-Lesende“ (lesen mehr als zwei Stunden täglich zum Vergnügen), „Normal-Lesende“ (lesen bis zu zwei Stunden am Tag zum Vergnügen) sowie „Lesemuffel“ (lesen nicht zum Vergnügen).

Welche *Inhalte* Schülerinnen und Schüler *zum Vergnügen* lesen, wurde mit fünf Items erhoben. Die Fünfzehnjährigen gaben an, wie häufig sie bestimmte Textsorten wie Zeitschriften, Comics, Fiktion, Sachbücher und Tageszeitungen zum Vergnügen lesen. Als Antwortoptionen konnten sie zwischen „ich weiß nicht, was das ist“, „nie oder fast nie“, „mehrmals pro Monat“, „mehrmals die Woche“ und „mehrmals pro Tag“ wählen.

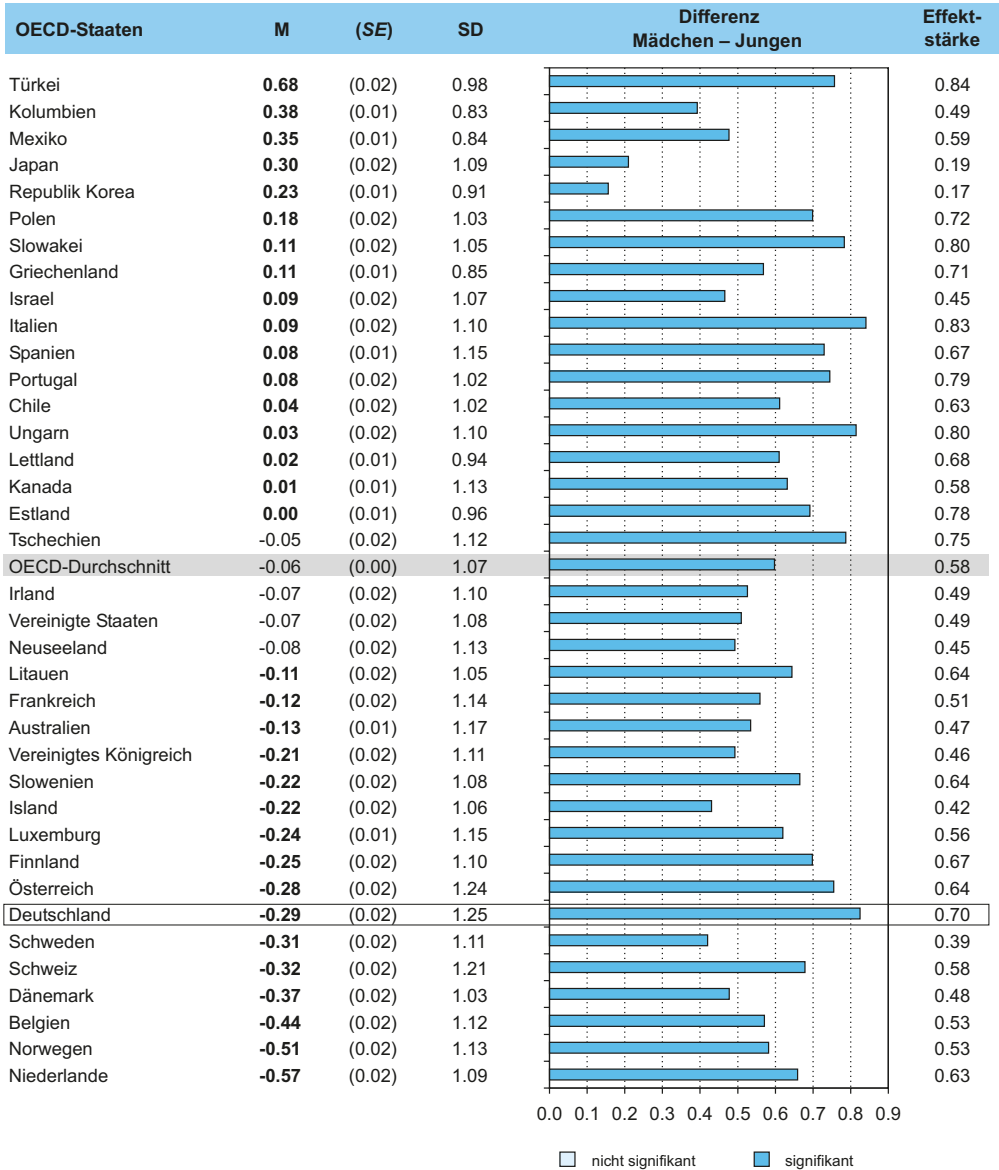
Das *Online-Lesen* wurde mit der Frage erfasst, wie häufig die Jugendlichen sich mit bestimmten Leseaktivitäten beschäftigen. Die Items umfassten das Lesen von E-Mails (kurz *E-Mails*), das Chatten wie zum Beispiel über Whatsapp®, Messenger® (kurz *Chat*), das Lesen von Online-Nachrichten (kurz *Online-Nachrichten*), das Suchen von Informationen im Internet über ein bestimmtes Thema (kurz *Online-Themen*), die Teilnahme an Online-Diskussionen und -Foren (kurz *Foren*), das Suchen im Internet nach praktischen Informationen etwa zu Fahrplänen, Veranstaltungen, Hinweisen, Rezepten (kurz *Online-Informationen*). Von der fünfstufigen Skala – „Ich weiß nicht, was das ist“, „Nie oder fast nie“, „Mehrmals pro Monat“, „Mehrmals pro Woche“, „Mehrmals pro Tag“ – werden für die hier berichteten Auswertungen die letzten beiden Kategorien zusammengefasst als Zustimmung gewertet.

Die *wahrgenommene Kompetenz* sowie die *wahrgenommene Schwierigkeit* beim Lesen wurden mit jeweils drei Items auf dergleichen Skala wie die Lesefreude erhoben. Ein Beispielitem für die *wahrgenommene Kompetenz* (Cronbachs $\alpha = .77$) ist „Ich bin ein guter Leser/eine gute Leserin“; ein Beispielitem für die *wahrgenommene Schwierigkeit* (Cronbachs $\alpha = .72$) ist „Ich finde es schwierig, Fragen zu einem Text zu beantworten“.

4.1.2 Ergebnisse des internationalen Vergleichs

Abbildung 4.1 zeigt die Ausprägung der *Lesefreude* in den OECD-Staaten. Die präsentierten Skalenwerte sind am OECD-Mittelwert standardisiert, was neben dem Vergleich mit dem OECD-Durchschnitt auch einen Vergleich mit der Lesefreude bei PISA 2009 erlaubt. Diese Skalenwerte ermöglichen, dass Staaten untereinander verglichen werden können, sie lassen aber keine Aussagen über die absolute Höhe der Lesefreude zu.

Im Vergleich zum OECD-Durchschnitt berichten Schülerinnen und Schüler in Deutschland eine signifikant geringere Lesefreude. Auch in den anderen deutschspra-



Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt ($p < .05$) sind **fett** hervorgehoben.

Abbildung 4.1: Lesefreude im internationalen Vergleich

chigen Staaten liegt die Lesefreude unter dem Durchschnitt der OECD-Staaten. Anders sieht es in den Staaten aus, die sich durch eine besonders hohe Lesekompetenz auszeichnen, darunter Estland, Kanada und Polen (vgl. Kapitel 3). Jugendliche in diesen Staaten berichten durchweg eine deutlich höhere Lesefreude als die Jugendlichen in Deutschland. Gleiches gilt für einen Großteil jener Staaten, die eine ähnlich hohe Lesekompetenz erreichen wie Deutschland, wie zum Beispiel Frankreich, Australien und Slowenien, aber auch für Staaten mit deutlich geringerer mittlerer Kompetenz (z. B. Türkei, Kolumbien, Mexiko). Sie zeigen ebenfalls eine signifikant höhere Lesefreude als die Fünfzehnjährigen in Deutschland. Eine geringere Freude muss demnach nicht zwangsläufig das Ergebnis einer geringeren Kompetenz sein.

Betrachtet man Unterschiede zwischen Mädchen und Jungen, berichten Mädchen in Deutschland ebenso wie in allen OECD-Staaten signifikant höhere Lesefreude als Jungen ($d = 0.70$). Dieser Unterschied ($M_{Diff} = 0.82$) ist auch in der Türkei, Italien, Ungarn und Tschechien sehr stark ausgeprägt. In Japan und in der Republik Korea, die sich durch eine hohe Lesekompetenz auszeichnen, ist die Lesefreude der Jungen hingegen deutlich höher.

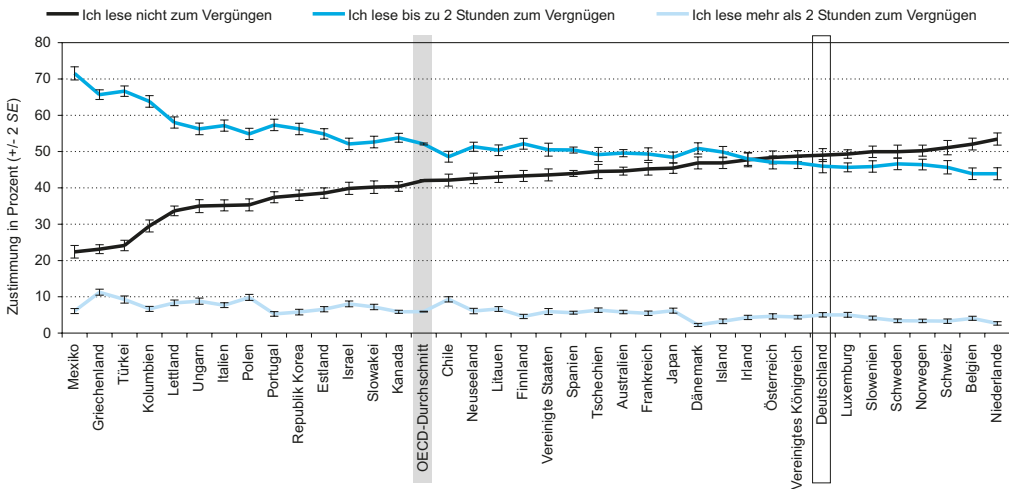
Betrachtet man die Lesefreude differenzierter auf Ebene einzelner Items (siehe Tabelle 4.2), um Aussagen über die absolute Höhe machen zu können, so zeigt sich ein deutliches Bild: In Deutschland stimmen, ähnlich wie im OECD-Durchschnitt, die Hälfte der Fünfzehnjährigen der Aussage „Ich lese nur, wenn ich muss“ zu. Etwas mehr als die Hälfte der Jugendlichen gibt an nur zu lesen, um gezielte Informationen zu bekommen. Dieser Anteil von 54.5 Prozent liegt signifikant über dem OECD-Durchschnitt von 49.7 Prozent. Gleichzeitig gibt nur rund ein Viertel der Jugendlichen an, dass Lesen eines ihrer liebsten Hobbys ist. Das sind etwa sieben Prozentpunkte weniger als im OECD-Durchschnitt. Dieses Muster deutet daraufhin, dass Schülerinnen und Schüler in Deutschland eine eher instrumentelle Lesemotivation aufweisen. Lesen dient ihnen als Mittel zum Zweck und es wird nur dann gelesen, wenn es notwendig ist, also etwa gezielt nach Informationen gesucht wird.

Tabelle 4.2: Prozent der Zustimmung zu Einzelitems der Lesefreude im Vergleich zum OECD-Durchschnitt

Lesefreude Einzelitems	Deutschland	OECD-Durchschnitt
	%	%
Ich lese nur, wenn ich muss	50.3	49.1
Lesen ist eines meiner liebsten Hobbys	26.5	33.7
Ich rede gern mit anderen Leuten über Bücher	23.9	36.6
Für mich ist Lesen Zeitverschwendung	34.2	28.4
Ich lese nur, um Informationen zu bekommen, die ich brauche	54.5	49.7

Anmerkung: Zustimmung entspricht der Zusammenfassung der beiden Kategorien „Stimme eher zu“ und „Stimme völlig zu“.

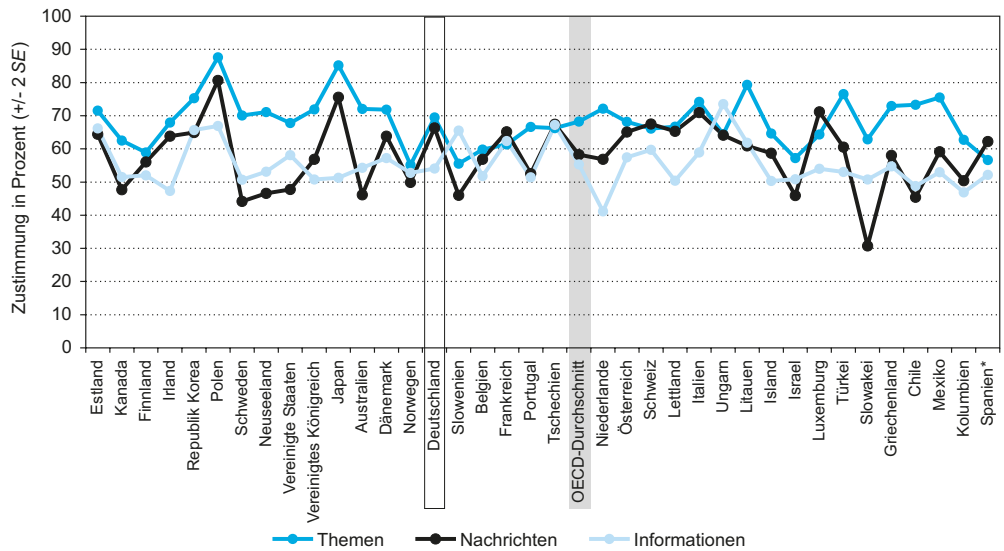
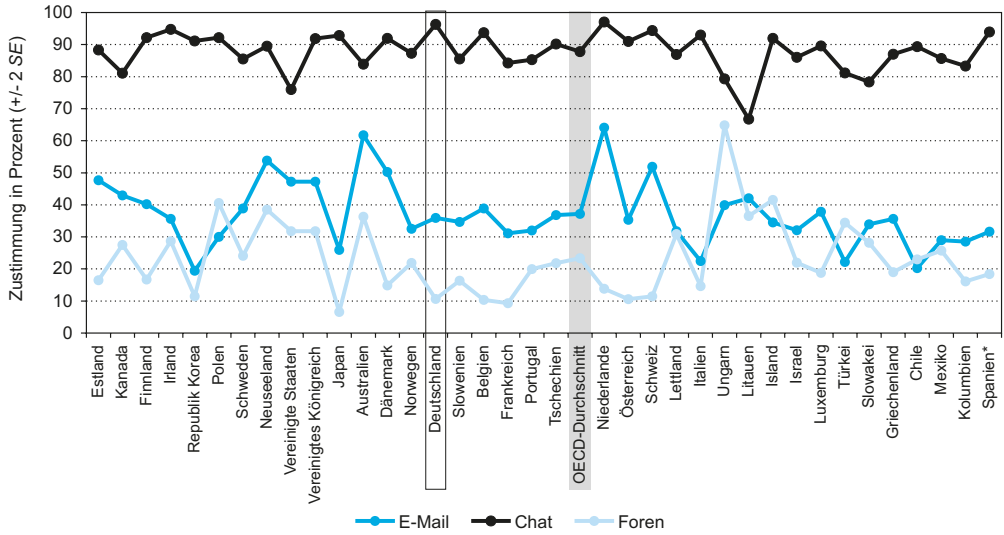
Neben der Lesefreude ist auch die *Lesemenge* bedeutsam, also der Umfang dessen, was Jugendliche *zum Vergnügen* lesen. Dabei handelt es sich um die Manifestation der Motivation im Verhalten. Abbildung 4.2 zeigt, wie häufig Schülerinnen und Schüler in den OECD-Staaten zum Vergnügen lesen. In Deutschland geben die Hälfte der Fünfzehnjährigen an, dass sie nicht zum Vergnügen lesen („Lesemuffel“). Dieser Wert liegt signifikant über dem OECD-Durchschnitt (42.0 %). Insgesamt betrachtet finden sich mit Ausnahme von Mexiko, Griechenland, Türkei und Kolumbien in nahezu allen OECD-Staaten hohe Anteile von Jugendlichen, die nicht zum Vergnügen lesen. Die lesekompetenzstarken Staaten reihen sich dort ebenfalls ein, auch wenn die durchschnittliche Lesemenge in diesen im Vergleich zu Deutschland etwas höher ist. Trotzdem gibt es auch in Deutschland eine kleine Gruppe von Jugendlichen (5.0 %), die sogar mehr als zwei Stunden täglich zum Vergnügen liest („Viel-Lesende“).



Anmerkung: Anteil der Schülerinnen und Schüler, welche eine bestimmte Lesemenge angeben. Die Staaten sind nach dem relativen Anteil der Antwortkategorie „Ich lese nicht zum Vergnügen“ geordnet.

Abbildung 4.2: Lesemenge im internationalen Vergleich

Lesen findet heutzutage nicht mehr nur auf Papier statt, sondern verstärkt auch in digitalen Medien (vgl. Kapitel 2 und 3). Abbildung 4.3 zeigt die prozentualen Anteile der Schülerinnen und Schüler, welche einen bestimmten Texttyp mehrmals wöchentlich oder mehrmals täglich lesen. Im *Online-Lesen* findet die Nutzung von Chats wie WhatsApp® die höchste Zustimmung. In Deutschland gaben 96.3 Prozent der Jugendlichen an, dies mehrmals in der Woche oder mehrmals täglich zu tun. Dieser Wert liegt zwar signifikant über dem OECD-Durchschnitt, folgt aber dennoch dem internationalen Trend: Mit Ausnahme von Litauen gaben Fünfzehnjährige aller OECD-Staaten dies als häufigste Form digitalen Lesens an. Die wöchentliche oder tägliche Nutzung von Foren wurde in Deutschland (10.7%) wie auch im OECD-Durchschnitt (23.4%) am seltensten berichtet. Diese beiden Formen digitalen Lesens bilden mit wenigen Aus-



Anmerkung: Prozent der Zustimmung („mehrmals pro Woche“/„mehrmals pro Tag“) zu verschiedenen Formen des Online-Lesens. Die Staaten sind nach ihrer Lesekompetenz geordnet. *Für Spanien sind aus methodischen Gründen keine Aussagen über die Lesekompetenz möglich, daher ist es an dieser Stelle gelistet. Dies ist aber keine Aussage über die tatsächliche Lesekompetenz Spaniens.

Abbildung 4.3: Online-Lesen im internationalen Vergleich

nahmen, in denen E-Mails noch seltener als Foren genutzt wurden (z. B. in der Türkei oder Chile), die Eckpunkte des Online-Lesens der meisten OECD-Staaten. Interessant ist das Online-Lesen von Nachrichten. In den OECD-Staaten taten dies zwischen 30.8 Prozent (Slowakei) und 80.6 Prozent (Polen) der Jugendlichen mehrmals wöchentlich oder täglich. In Deutschland gaben 66.4 Prozent der Schülerinnen und Schüler an, wöchentlich oder täglich Nachrichten online zu lesen, wohingegen nur 28.8 Prozent der Jugendlichen angaben, papierbasierte Tageszeitungen zu lesen.

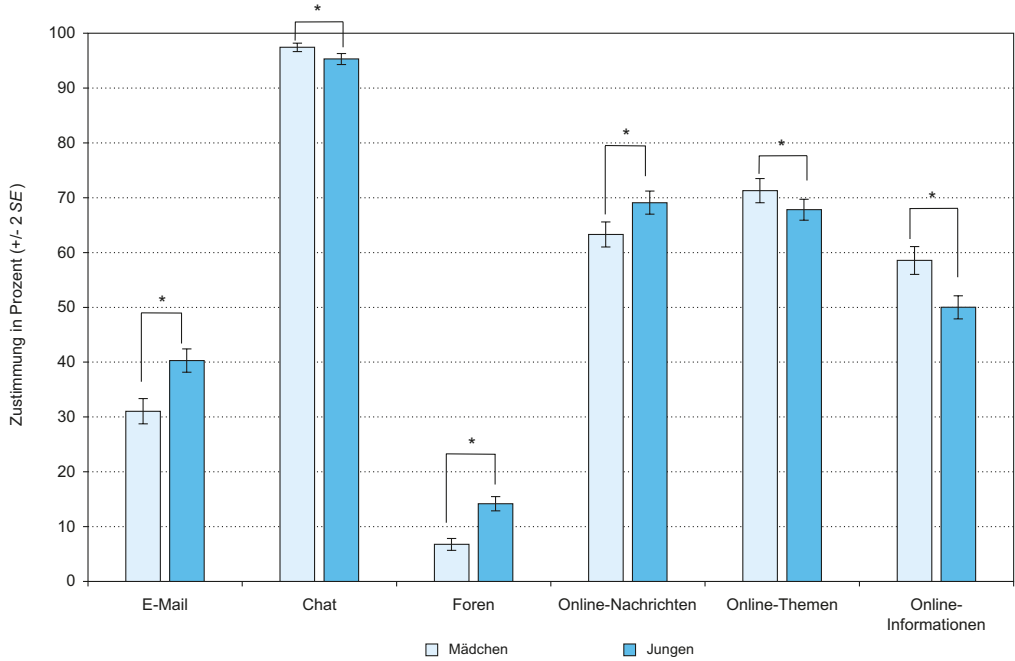
Mädchen und Jungen unterscheiden sich ähnlich wie in der *Lese Freude* auch in ihrem *Leseverhalten*. In Deutschland geben Jungen (60.6 %) deutlich häufiger als Mädchen (36.0 %) an, nicht zum Vergnügen zu lesen. Ebenso berichten signifikant weniger Jungen (2.8 %) als Mädchen (7.6 %), mehr als zwei Stunden täglich zum Vergnügen zu lesen. Dies ist in allen OECD-Staaten der Fall außer in der Republik Korea – nur dort gibt es keinen Geschlechterunterschied bei den „Viel-Lesenden“.

In Deutschland zeigt sich der größte Geschlechterunterschied im *Online-Lesen* in der Nutzung von E-Mails (Mädchen: 31.1 %; Jungen: 40.3 %; Abbildung 4.4), gefolgt von Foren (Mädchen: 6.8 %; Jungen: 14.1 %), welche an sich am seltensten genutzt werden. Die Mädchen recherchierten häufiger im Internet, sowohl nach bestimmten Themen als auch nach praktischen Informationen. Die geringsten Geschlechterunterschiede zeigten sich beim gleichermaßen hoch ausgeprägten Chatten.

Betrachtet man differenzierter, welche Art von Texten die unterschiedlichen Gruppen von Fünfzehnjährigen lesen (Abbildung 4.5), zeigt sich ein eindeutiges Bild: „Viel-Lesende“ lesen vorwiegend Fiktion, während „Lesemuffel“ allenfalls Tageszeitungen lesen. Letzteres taten sie durchschnittlich nur „ein paar Mal im Jahr“.

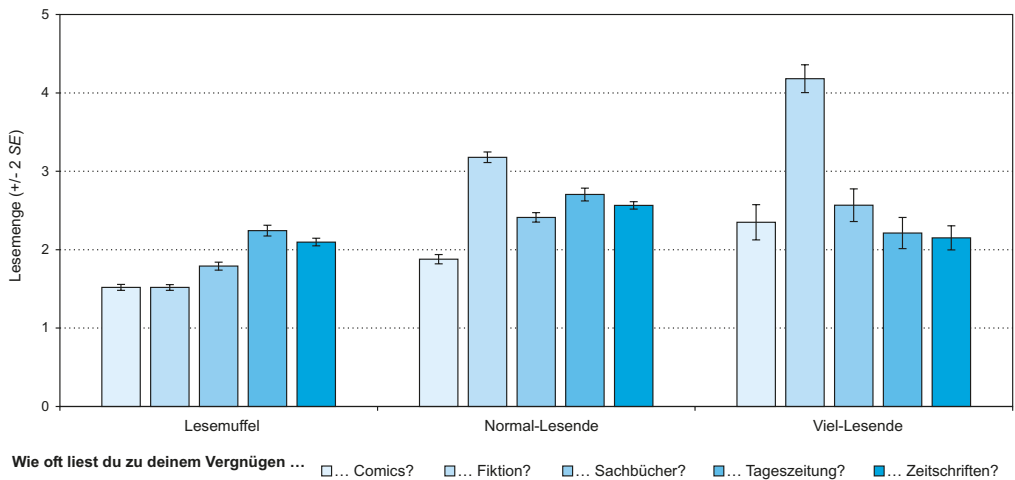
Die Ergebnisse zu den beiden Facetten des lesebezogenen Selbstkonzepts ergänzen sich in Deutschland. Während die persönlich *wahrgenommene Lesekompetenz* über dem OECD-Durchschnitt liegt (Abbildung 4.6), sind die *wahrgenommenen Schwierigkeiten* beim Lesen geringer als im OECD-Durchschnitt (Abbildung 4.7). Diese Kongruenz findet sich jedoch nicht in allen Staaten. Zum Beispiel in Norwegen, mit einer vergleichbar ausgeprägten Lesekompetenz wie in Deutschland, schätzen die Fünfzehnjährigen ihre *wahrgenommene Kompetenz* im Lesen im Vergleich zum OECD-Durchschnitt ebenfalls höher ein. Sie geben aber gleichzeitig an, dass sie im Vergleich zum OECD-Mittelwert Lesen schwieriger finden. Es ist also nicht zwangsläufig so, dass Schülerinnen und Schüler, die der Aussage „Ich kann schwierige Texte verstehen“ zustimmen, auch die Aussage „Ich finde es schwierig, Fragen zu einem Text zu beantworten“ ablehnen. Betrachtet man andererseits mit Estland und Polen Staaten, die sich durch eine überdurchschnittliche Lesekompetenz auszeichnen, nehmen die Fünfzehnjährigen Lesen zwar als wenig schwierig wahr, schätzen ihre Kompetenz jedoch im Vergleich zum OECD-Mittelwert unterdurchschnittlich ein.

In Deutschland berichten Mädchen ähnlich wie im OECD-Durchschnitt eine signifikant höhere *wahrgenommene Kompetenz* im Lesen als die Jungen (Abbildung 4.6). Dieses Muster zeigt sich in den meisten OECD-Staaten. Die größten Unterschiede zwischen Mädchen und Jungen zeigen sich, wie auch bei der Lesemotivation, in Italien und der



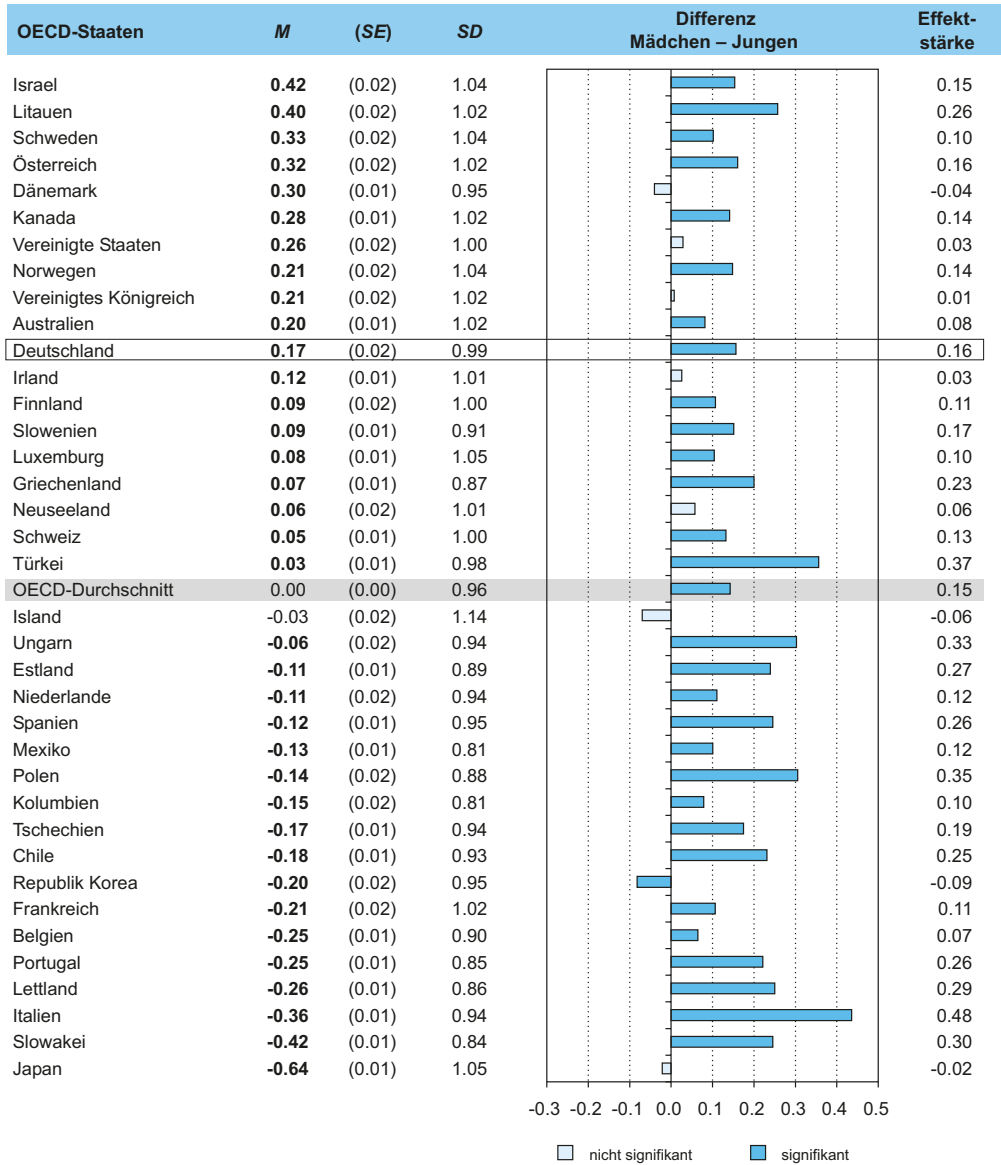
Anmerkung: Prozent der Zustimmung („mehrmals pro Woche“/ „mehrmals pro Tag“) zu verschiedenen Formen des Online-Lesens differenziert nach Geschlecht. * Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p < .05$).

Abbildung 4.4: Online-Lesen differenziert nach Geschlecht



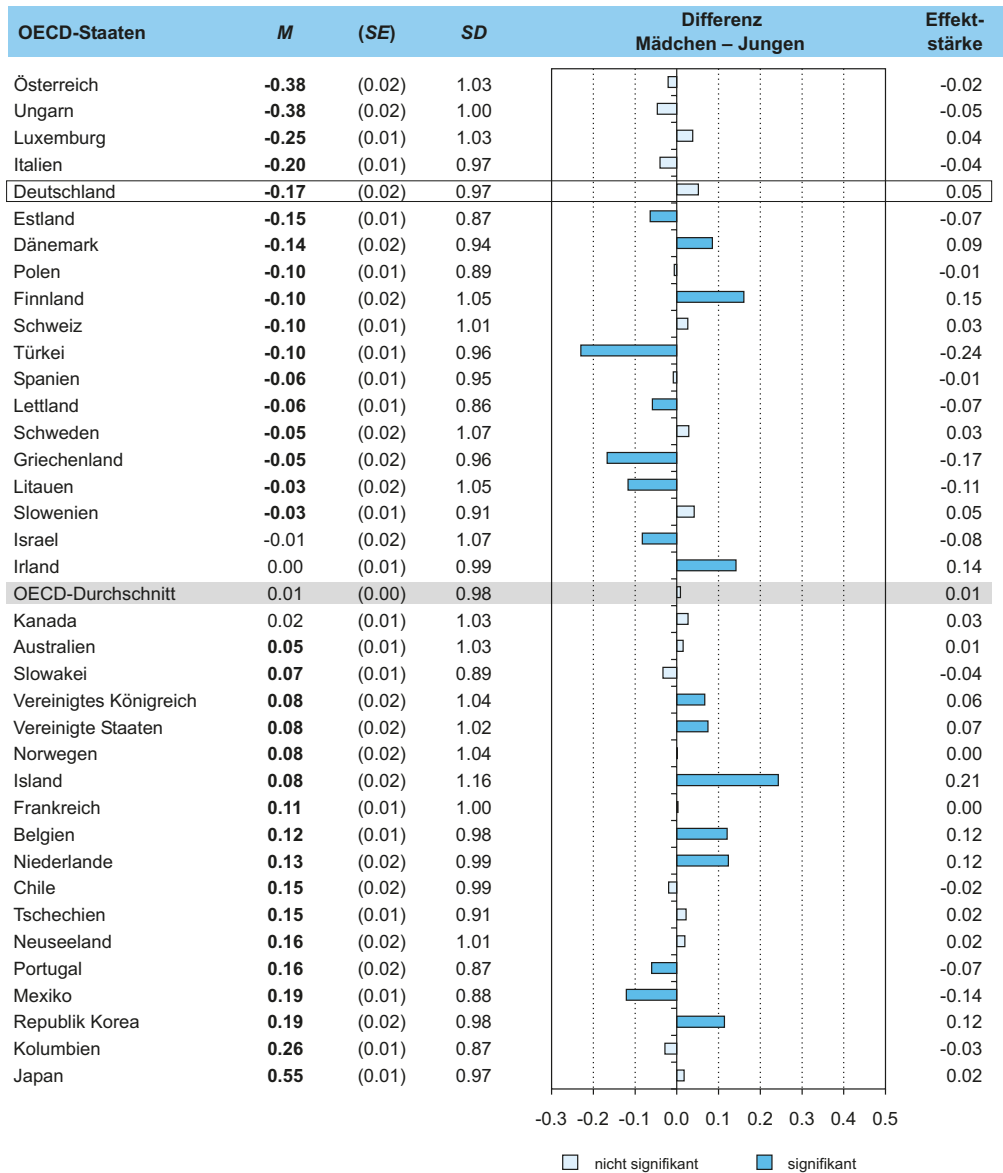
Anmerkung: Lesen zum Vergnügen in Abhängigkeit von der Lesemenge. Schülerinnen und Schüler bewerteten die Häufigkeit, mit der sie diese Textarten lesen auf einer Skala von 1 („nie oder fast nie“) bis 5 („mehrmals in der Woche“).

Abbildung 4.5: Lesen zum Vergnügen in Abhängigkeit von der Lesemenge in Deutschland



Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt ($p < .05$) sind **fett** hervorgehoben.

Abbildung 4.6: Wahrgenommene Kompetenz im internationalen Vergleich



Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt ($p < .05$) sind **fett** hervorgehoben.

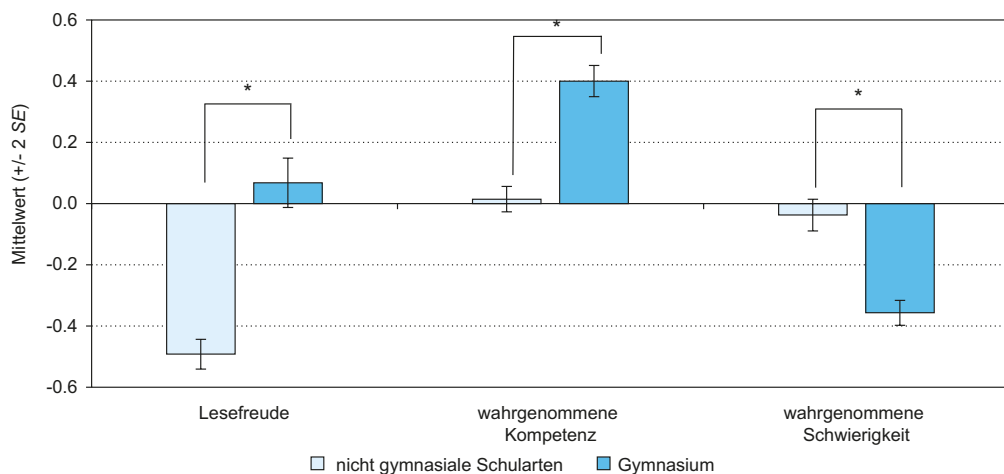
Abbildung 4.7: Wahrgenommene Schwierigkeit im internationalen Vergleich

Türkei. In Dänemark, Irland, Island, Japan, Neuseeland, den Vereinigten Staaten und dem Vereinigten Königreich zeigen sich keine signifikanten Unterschiede zwischen Mädchen und Jungen. Die Republik Korea ist der einzige OECD-Staat, in dem Jungen eine signifikant höhere *wahrgenommene Kompetenz* angeben als Mädchen. Für die *wahrgenommene Schwierigkeit* beim Lesen lassen sich in Deutschland hingegen ähnlich wie im OECD-Durchschnitt keine signifikanten Unterschiede zwischen Jungen und Mädchen finden. Insgesamt ist das Bild aber weniger einheitlich als bei der *wahrgenommenen Kompetenz*. In neun OECD-Staaten schätzen Mädchen die *wahrgenommene Schwierigkeit* im Lesen höher ein, wohingegen in acht OECD-Staaten Jungen die *wahrgenommene Schwierigkeit* im Lesen höher einschätzen. Dies ist insbesondere deshalb interessant, da die Geschlechterunterschiede in der Lesekompetenz in allen OECD-Staaten zugunsten der Mädchen ausfallen (vgl. Kapitel 3), die *wahrgenommene Schwierigkeit* dies aber nur zum Teil widerspiegelt.

4.1.3 Unterschiede zwischen Schularten in Deutschland

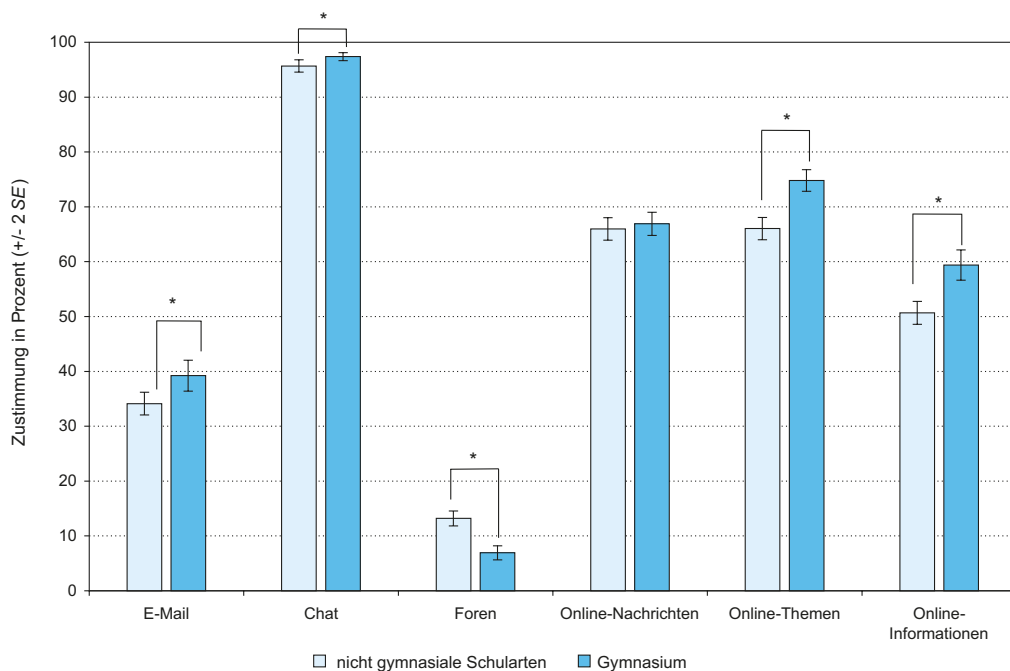
Zusätzlich zum internationalen Vergleich lassen sich in Deutschland Unterschiede zwischen Schularten untersuchen. Fünfzehnjährige an Gymnasien und nicht gymnasialen Schularten unterscheiden sich signifikant hinsichtlich der *Lesefreude* und dem Selbstkonzept (*wahrgenommene Kompetenz* und *wahrgenommene Schwierigkeit*; Abbildung 4.8). Schülerinnen und Schüler nicht gymnasialer Schularten haben im Mittel weniger Freude am Lesen als Gymnasiastinnen und Gymnasiasten. Auch geben 57,0 Prozent der Jugendlichen in nicht gymnasialen Schularten an, dass sie nicht zum Vergnügen lesen und somit „Lesemuffel“ sind. An den Gymnasien sind es zwar deutlich weniger, dennoch berichten etwas mehr als ein Drittel der Gymnasiastinnen und Gymnasiasten (35,7%), dass sie nicht zum Vergnügen lesen. In der Gruppe der „Viel-Lesenden“ (mehr als zwei Stunden täglich) zeigen sich hingegen keine signifikanten Unterschiede zwischen den Schularten. Bezüglich der beiden Facetten des Selbstkonzepts schätzen Gymnasiastinnen und Gymnasiasten ihre Lesekompetenz höher ein und finden Lesen weniger schwierig als Jugendliche an nicht gymnasialen Schularten.

Untersucht man die Frage, ob es Schulartunterschiede gibt, welche Texte Jugendliche *online lesen*, so nutzen Schülerinnen und Schüler an nicht gymnasialen Schularten mehr Foren (Zustimmung: 6,9% Gymnasium vs. 13,2% nicht gymnasiale Schularten; siehe Abbildung 4.9), während gleichzeitig 14,9 Prozent der gleichen Schulart angeben nicht zu wissen, was Foren sind. Das Nachrichtenlesen im Internet ist bei Fünfzehnjährigen aller Schularten gleichermaßen beliebt, während Gymnasiastinnen und Gymnasiasten signifikant häufiger im Internet Themen und Informationen recherchieren.



Anmerkung: Mittelwerte und Standardabweichungen der am OECD-Durchschnitt standardisierten Merkmale Lese Freude, wahrgenommene Kompetenz und wahrgenommene Schwierigkeit. * Statistisch signifikante Unterschiede zwischen nicht gymnasialen Schularten und dem Gymnasium ($p < .05$).

Abbildung 4.8: Lesemotivation und lesebezogenes Selbstkonzept differenziert nach Schulart

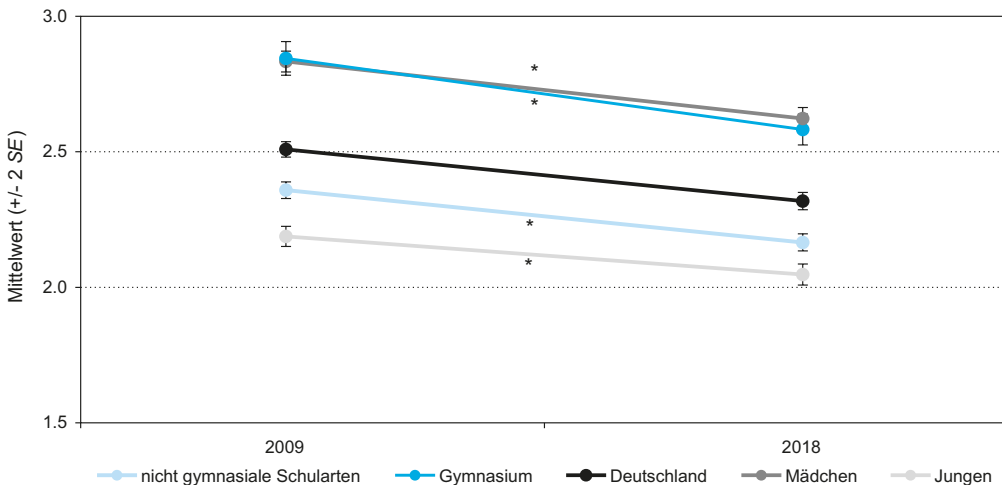


Anmerkung: Prozent der Zustimmung („mehrmals pro Woche“/ „mehrmals pro Tag“) zu verschiedenen Formen des Online-Lesens differenziert nach Schularten. * Statistisch signifikante Unterschiede zwischen nicht gymnasialen Schularten und dem Gymnasium ($p < .05$).

Abbildung 4.9: Online-Lesen differenziert nach Schulart

4.1.4 Veränderungen zwischen den Erhebungen 2009 und 2018

Lesefreude, *Lesemenge* und *Online-Lesen* wurden bereits bei PISA 2009 erhoben, als Lesekompetenz zum letzten Mal die Hauptdomäne war.² Deshalb kann für diese Merkmale untersucht werden, wie sie sich in Deutschland entwickelt haben. Betrachtet man die Veränderung der *Lesefreude* von PISA 2009 bis PISA 2018³, zeigt sich für Deutschland, dass diese signifikant abgenommen hat. Der Anteil der Jugendlichen, die lesen, weil sie müssen oder nach Informationen suchen, hat seit der PISA-Erhebung 2009 signifikant zugenommen. Die Abnahme der *Lesefreude* in Deutschland ist höher als im Durchschnitt der OECD-Staaten. Dieser Trend ist sowohl für Mädchen ($M_{Diff} = -0.21$) als auch für Jungen ($M_{Diff} = -0.14$) in Deutschland zu beobachten (Abbildung 4.10). Differenziert nach Schularten ging die *Lesefreude* in den vergangenen neun Jahren an den Gymnasien ($M_{Diff} = -0.26$) und den nicht gymnasialen Schularten ($M_{Diff} = -0.19$) vergleichbar zurück.



Anmerkung: Einfacher Mittelwert der Skala Lesefreude² * Statistisch signifikante Unterschiede zwischen den Erhebungen 2009 und 2018

Abbildung 4.10: Veränderung der Lesefreude in Deutschland differenziert nach Schulart und Geschlecht

Auch in Bezug auf die *Lesemenge zum Vergnügen* zeigt sich ein problematisches Bild. In der PISA-Erhebung 2018 vergrößerte sich der Anteil der „Lesemuffel“, welche nicht zum

2 In PISA 2018 kamen drei weitere OECD-Staaten hinzu: Litauen, Lettland und Kolumbien. Diese Staaten wurden zur Vergleichbarkeit von PISA 2018 und PISA 2009 in die Trendanalysen in PISA 2009 aufgenommen. Dies könnte zu Abweichungen beim Vergleich zur Berichterstattung der Ergebnisse aus 2009 führen.

3 Für die Berechnung der Veränderung zwischen 2009 und 2018 wurde für beide Erhebungen nach Umpolung der negativ gepolten Items ein einfacher Mittelwert berechnet. Hier muss vom am OECD-Durchschnitt standardisierten Index, welcher bis hierher verwendet wurde, abgewichen werden, da mit diesem keine Schulartvergleiche für die Erhebung 2009 möglich sind.

Vergnügen lesen, in Deutschland im Vergleich zu 2009 um 7.7 Prozentpunkte (2009: 41.2 %; 2018: 49.1 %). Gleichzeitig veränderte sich der Anteil der Jugendlichen, die mehr als zwei Stunden täglich lesen („Viel-Lesende“) nicht signifikant (2009: 5.9 %; 2018: 5 %). Diese Entwicklung zeigt sich sowohl für Mädchen und für Jungen als auch in beiden Schularten (Tabelle 4.3). Deskriptiv ist in den beiden Gruppen, welche häufiger zum Vergnügen lasen (Mädchen bzw. Gymnasiastinnen und Gymnasiasten), eine etwas stärkere Abnahme des Merkmals als in ihren Vergleichsgruppen (Jungen bzw. Jugendliche nicht gymnasialer Schularten) festzustellen.

Tabelle 4.3: Veränderung der Anteile der „Lesemuffel“, „Normal-Lesenden“ und „Viel-Lesenden“ zwischen 2009 und 2018 nach Schulart und Geschlecht

	Differenz 2018–2009			Differenz 2018–2009	
	Prozentpunkte	(SE)		Prozentpunkte	(SE)
Gymnasium			Mädchen		
Lesemuffel	10.8	(2.2)	Lesemuffel	8.5	(1.6)
Normal-Lesende	-9.5	(2.2)	Normal-Lesende	-7.7	(1.7)
Viel-Lesende	-1.2	(0.9)	Viel-Lesende	-0.8	(0.8)
Nicht Gymnasium			Jungen		
Lesemuffel	8.2	(1.4)	Lesemuffel	5.8	(1.6)
Normal-Lesende	-7.6	(1.5)	Normal-Lesende	-5.1	(1.7)
Viel-Lesende	-0.6	(0.7)	Viel-Lesende	-0.7	(0.5)

Anmerkung: Veränderung der Anteile der Schülerinnen und Schüler, welche eine bestimmte Lesemenge angeben. Statistisch signifikante Unterschiede zwischen den Erhebungen 2009 und 2018 ($p < .05$) sind **fett** hervorgehoben.

Betrachtet man die Entwicklung in den anderen OECD-Staaten, so nimmt lediglich in der Republik Korea, Mexiko, Österreich und der Slowakei der Anteil der Lesemuffel signifikant ab. Alle anderen OECD-Staaten verzeichnen im *Lesen zum Vergnügen* ebenfalls einen negativen Trend im Sinne einer Zunahme der Gruppe der „Lesemuffel“, die in den meisten OECD-Staaten signifikant ist.

Allgemein hat es beim *Online-Lesen* in den letzten neun Jahren Veränderungen gegeben (Tabelle 4.4). In Deutschland hat Chatten noch einmal deutlich zugelegt (82.6 % im Jahr 2009; 96.3 % im Jahr 2018) und übertrifft die an Bedeutung verlierende E-Mail (65.3 % im Jahr 2009; 35.9 % im Jahr 2018) bei weitem. Beide Veränderungen sind auch im OECD-Durchschnitt zu beobachten (Differenz Chat: 9.8 Prozentpunkte; Differenz E-Mails: -18.0 Prozentpunkte).

Tabelle 4.4: Veränderung des Online-Lesens zwischen 2009 und 2018

	2018		2009		Differenz 2018–2009	
	%	(SE)	%	(SE)	Prozentpunkte	(SE)
E-Mail	35.9	(0.8)	65.3	(0.9)	-29.4	(1.2)
Chat	96.3	(0.4)	82.6	(0.6)	13.7	(0.7)
Foren	10.6	(0.5)	15.7	(0.6)	-5.1	(0.8)
Nachrichten	66.4	(0.7)	62.0	(0.8)	4.3	(1.1)
Themen	69.5	(0.8)	52.0	(0.9)	17.5	(1.2)
Informationen	54.1	(0.8)	36.1	(0.9)	18.0	(1.2)

Anmerkungen: Prozent der Zustimmung („mehrmals pro Woche“/ „mehrmals pro Tag“) zu verschiedenen Formen des Online-Lesens in den PISA Erhebungen 2009 und 2018, sowie Veränderung in diesem Zeitraum. Statistisch Signifikante Unterschiede zwischen den Erhebungen 2009 und 2018 ($p < .05$) sind **fett** hervorgehoben.

4.2 Lesestrategiewissen

Neben dem Leseprozess ist der Umgang mit den Leseaufgaben eine zentrale Komponente der aktuellen PISA-Rahmenkonzeption (vgl. Kapitel 2). Jugendliche benötigen Fähigkeiten, den Leseprozess ihren persönlichen Zielen entsprechend zu gestalten und bei Verständnisschwierigkeiten ihre Ziele anzupassen. Zum Beispiel kann es bei der Zusammenfassung eines langen und schwierigen Texts hilfreich sein, zunächst die wichtigsten Sätze im Text zu unterstreichen und diese anschließend in eigenen Worten zusammenzufassen. Um erfolgreich mit verschiedenen Texten umgehen zu können, müssen Lernende unterschiedliche Strategien (Zielsetzungs- und Planungsstrategien sowie Monitoring- und Selbstregulationsstrategien) kennen, aus diesen eine Strategie als geeignet für die vorliegende Aufgabe auswählen und diese gezielt anwenden (Boekaerts, 1999). Lesestrategiewissen beschreibt dabei konkret die Anpassung des eigenen Lesens an die gegebene Aufgabe unter Beachtung der eigenen Fähigkeiten und entsprechend der Angemessenheit geeigneter Strategien (Flavell, 1979). Lesestrategiewissen ist eine zentrale metakognitive Lernstrategie und erklärt wesentliche Anteile der Unterschiede in der Lesekompetenz (Callan, Marchant, Finch & German, 2016).

4.2.1 Erfassung des Lesestrategiewissens in PISA 2018

PISA erhebt, inwieweit Jugendliche über lern- und verstehensrelevante Strategien beim Lesen und Verstehen von Texten verfügen. Dazu wurde bei PISA 2009 ein Wissenstest entwickelt (Artelt et al., 2010), der für die aktuelle Erhebungsrunde erweitert wurde. Die Schülerinnen und Schüler bearbeiteten im Rahmen des Schülerfragebogens drei Lernszenarien. Beim *Zusammenfassen von Informationen* (kurz: *Zusammenfassen*, Abbildung 4.11) soll der Inhalt eines Textes verstanden und behalten werden. Das zweite Lernsze-

nario, *Verstehen und Erinnern eines Textes* (kurz: *Erinnern*, Abbildung 4.12), beschreibt die Aufgabe, einen schwierigen Sachtext zusammenzufassen. Die Jugendlichen sind bei beiden Szenarien aufgefordert, auf einer sechsstufigen Skala einzuschätzen, wie nützlich die dargebotenen Strategien für das jeweilige Ziel sind. Lernszenario 3, *Beurteilung der Qualität und Vertrauenswürdigkeit einer Online-Quelle* (kurz: *Beurteilen*, Abbildung 4.13), berücksichtigt die Weiterentwicklung der Rahmenkonzeption in Bezug auf das digitale Lesen (vgl. Kapitel 2 und 3). In diesem Szenario sollen sich die Schülerinnen und Schüler vorstellen, dass in ihrem E-Mail-Eingang eine Nachricht von einem bekannten Mobilfunkanbieter eingegangen ist und sie über den Gewinn eines Smartphones informiert. Damit das Smartphone zugeschickt werden kann, sollen sie einen Link anklicken und ihre Adressdaten eingeben. Die Jugendlichen sollen einschätzen, wie angemessen die dargebotenen Reaktionen sind, also zum Beispiel das Überprüfen der E-Mail-Adresse des Absenders.

Leseaufgabe: Du hast gerade einen zwei Seiten langen und ziemlich schwierigen Text über die Veränderungen des Wasserspiegels eines Sees in Afrika gelesen. Du sollst eine Zusammenfassung schreiben.

Wie beurteilst du den Nutzen der folgenden Strategien für das Schreiben einer Zusammenfassung dieses zweiseitigen Textes?

(Bitte eine Antwort pro Zeile auswählen.)

	Überhaupt nicht nützlich (1)	(2)	(3)	(4)	(5)	Sehr nützlich (6)
Ich schreibe eine Zusammenfassung. Danach prüfe ich, ob die einzelnen Abschnitte des Textes in der Zusammenfassung enthalten sind, da der Inhalt eines jeden Abschnitts darin vorkommen sollte.	ST165Q01IA01 <input type="radio"/>	ST165Q01IA02 <input type="radio"/>	ST165Q01IA03 <input type="radio"/>	ST165Q01IA04 <input type="radio"/>	ST165Q01IA05 <input type="radio"/>	ST165Q01IA06 <input type="radio"/>
Ich versuche, möglichst viele Sätze aus dem Text wörtlich zu übernehmen.	ST165Q02IA01 <input type="radio"/>	ST165Q02IA02 <input type="radio"/>	ST165Q02IA03 <input type="radio"/>	ST165Q02IA04 <input type="radio"/>	ST165Q02IA05 <input type="radio"/>	ST165Q02IA06 <input type="radio"/>
Vor dem Schreiben der Zusammenfassung lese ich den Text so oft wie möglich.	ST165Q03IA01 <input type="radio"/>	ST165Q03IA02 <input type="radio"/>	ST165Q03IA03 <input type="radio"/>	ST165Q03IA04 <input type="radio"/>	ST165Q03IA05 <input type="radio"/>	ST165Q03IA06 <input type="radio"/>
Ich überprüfe sorgfältig, ob die wichtigsten Fakten des Textes in der Zusammenfassung enthalten sind.	ST165Q04IA01 <input type="radio"/>	ST165Q04IA02 <input type="radio"/>	ST165Q04IA03 <input type="radio"/>	ST165Q04IA04 <input type="radio"/>	ST165Q04IA05 <input type="radio"/>	ST165Q04IA06 <input type="radio"/>
Ich lese den Text und unterstreiche dabei die wichtigsten Sätze. Dann schreibe ich diese mit eigenen Worten als Zusammenfassung.	ST165Q05IA01 <input type="radio"/>	ST165Q05IA02 <input type="radio"/>	ST165Q05IA03 <input type="radio"/>	ST165Q05IA04 <input type="radio"/>	ST165Q05IA05 <input type="radio"/>	ST165Q05IA06 <input type="radio"/>

Abbildung 4.11: Lernszenarien Lesestrategiewissen – Zusammenfassen von Informationen

Wie beurteilst du den Nutzen der folgenden Strategien für das Verstehen und Behalten des Textes?

(Bitte eine Antwort pro Zeile auswählen.)

	Überhaupt nicht nützlich (1)	(2)	(3)	(4)	(5)	Sehr nützlich (6)
Ich konzentriere mich auf die Teile des Textes, die leicht zu verstehen sind.	ST164Q01IA01 <input type="radio"/>	ST164Q01IA02 <input type="radio"/>	ST164Q01IA03 <input type="radio"/>	ST164Q01IA04 <input type="radio"/>	ST164Q01IA05 <input type="radio"/>	ST164Q01IA06 <input type="radio"/>
Ich lese den Text zweimal schnell durch.	ST164Q02IA01 <input type="radio"/>	ST164Q02IA02 <input type="radio"/>	ST164Q02IA03 <input type="radio"/>	ST164Q02IA04 <input type="radio"/>	ST164Q02IA05 <input type="radio"/>	ST164Q02IA06 <input type="radio"/>
Nachdem ich den Text gelesen habe, bespreche ich den Inhalt mit anderen Personen.	ST164Q03IA01 <input type="radio"/>	ST164Q03IA02 <input type="radio"/>	ST164Q03IA03 <input type="radio"/>	ST164Q03IA04 <input type="radio"/>	ST164Q03IA05 <input type="radio"/>	ST164Q03IA06 <input type="radio"/>
Ich unterstreiche wichtige Textpassagen.	ST164Q04IA01 <input type="radio"/>	ST164Q04IA02 <input type="radio"/>	ST164Q04IA03 <input type="radio"/>	ST164Q04IA04 <input type="radio"/>	ST164Q04IA05 <input type="radio"/>	ST164Q04IA06 <input type="radio"/>
Ich fasse den Text mit eigenen Worten zusammen.	ST164Q05IA01 <input type="radio"/>	ST164Q05IA02 <input type="radio"/>	ST164Q05IA03 <input type="radio"/>	ST164Q05IA04 <input type="radio"/>	ST164Q05IA05 <input type="radio"/>	ST164Q05IA06 <input type="radio"/>
Ich lese jemandem den Text laut vor.	ST164Q06IA01 <input type="radio"/>	ST164Q06IA02 <input type="radio"/>	ST164Q06IA03 <input type="radio"/>	ST164Q06IA04 <input type="radio"/>	ST164Q06IA05 <input type="radio"/>	ST164Q06IA06 <input type="radio"/>

Abbildung 4.12: Lernszenarien Lesestrategiewissen – Verstehen und Erinnern eines Textes

Leseaufgabe: Du hast in deinem Postfach eine Nachricht von einem bekannten Mobilfunkanbieter, in der dir mitgeteilt wird, dass du ein Smartphone gewonnen hast. Der Absender fordert dich auf, den Link anzuklicken und ein Formular mit deinen Daten auszufüllen, damit er dir das Smartphone zuschicken kann.

Wie angemessen sind deiner Meinung nach die folgenden Reaktionen auf diese Mitteilung?

(Bitte eine Antwort pro Zeile auswählen.)

	Überhaupt nicht angemessen (1)	(2)	(3)	(4)	(5)	Sehr angemessen (6)
Die E-Mail beantworten und weitere Informationen zum Smartphone erfragen.	ST166Q01HA01 <input type="radio"/>	ST166Q01HA02 <input type="radio"/>	ST166Q01HA03 <input type="radio"/>	ST166Q01HA04 <input type="radio"/>	ST166Q01HA05 <input type="radio"/>	ST166Q01HA06 <input type="radio"/>
Die E-Mail-Adresse des Absenders überprüfen.	ST166Q02HA01 <input type="radio"/>	ST166Q02HA02 <input type="radio"/>	ST166Q02HA03 <input type="radio"/>	ST166Q02HA04 <input type="radio"/>	ST166Q02HA05 <input type="radio"/>	ST166Q02HA06 <input type="radio"/>
Auf den Link klicken und das Formular so schnell wie möglich ausfüllen.	ST166Q03HA01 <input type="radio"/>	ST166Q03HA02 <input type="radio"/>	ST166Q03HA03 <input type="radio"/>	ST166Q03HA04 <input type="radio"/>	ST166Q03HA05 <input type="radio"/>	ST166Q03HA06 <input type="radio"/>
Die E-Mail löschen, ohne vorher auf den Link zu klicken.	ST166Q04HA01 <input type="radio"/>	ST166Q04HA02 <input type="radio"/>	ST166Q04HA03 <input type="radio"/>	ST166Q04HA04 <input type="radio"/>	ST166Q04HA05 <input type="radio"/>	ST166Q04HA06 <input type="radio"/>
Auf der Webseite des Mobilnetzanbieters überprüfen, ob das Smartphone-Angebot dort erwähnt wird.	ST166Q05HA01 <input type="radio"/>	ST166Q05HA02 <input type="radio"/>	ST166Q05HA03 <input type="radio"/>	ST166Q05HA04 <input type="radio"/>	ST166Q05HA05 <input type="radio"/>	ST166Q05HA06 <input type="radio"/>

Abbildung 4.13: Lernszenarien Lesestrategiewissen – Beurteilung der Qualität und Vertrauenswürdigkeit einer Online-Quelle

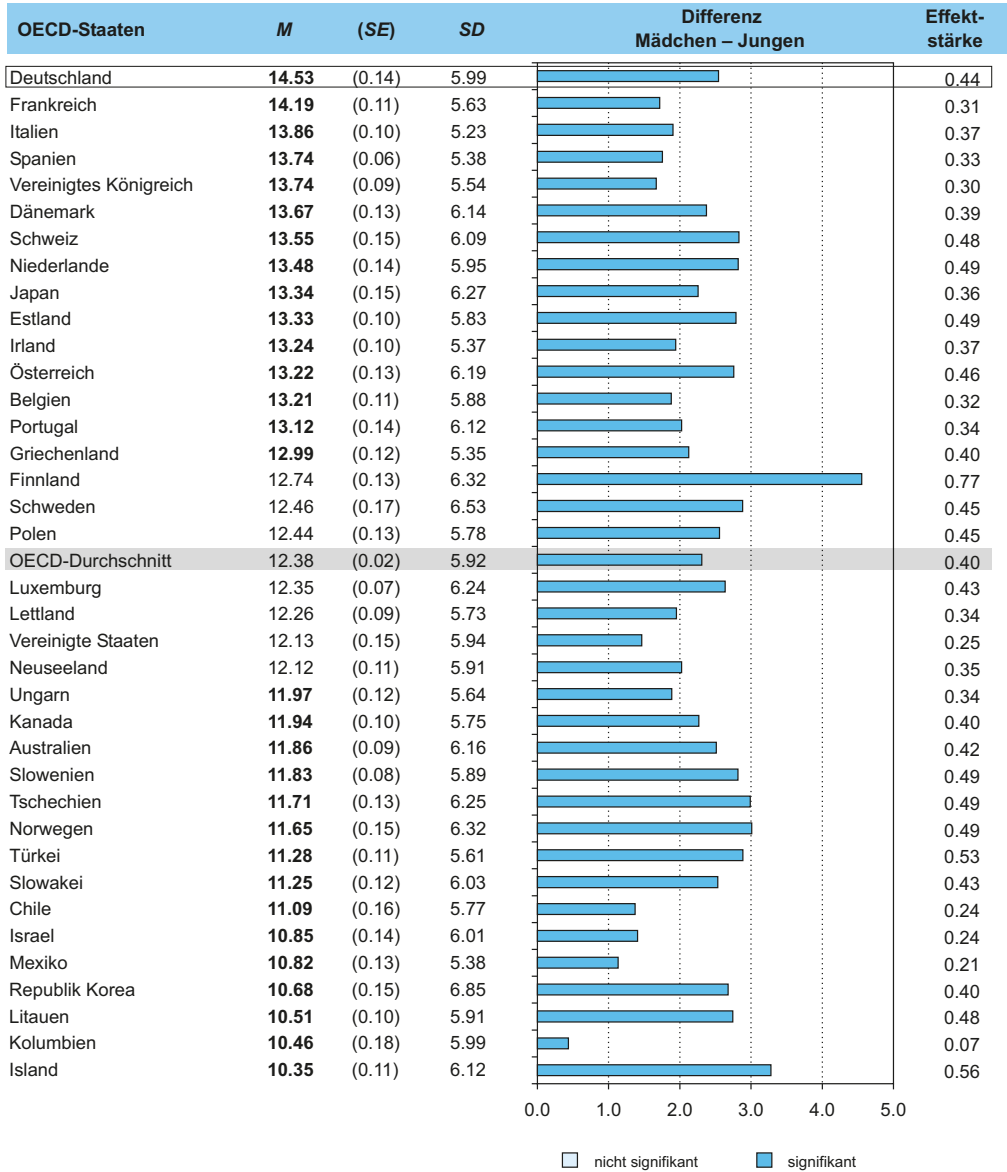
Um die Szenarien auszuwerten und dementsprechend das Lesestrategiewissen der Jugendlichen zu erfassen, wurde die Nützlichkeit und Angemessenheit der Lesestrategien in dem jeweiligen Szenario von internationalen Expertinnen und Experten im Vorfeld bewertet. Hierbei wurden die Strategien paarweise miteinander verglichen (Paarvergleiche). Die Lösung des Szenarios wurde als richtig bewertet, wenn die Nützlichkeitsbewertung der Strategien durch die Schülerin oder den Schüler mit denen der Expertinnen und Experten übereinstimmten (Artelt, Beinicke, Schlagmüller & Schneider, 2009;

OECD, 2017). Um nicht nur relative Vergleiche der Staaten untereinander zu ermöglichen, sondern auch Aussagen darüber machen zu können, wie gut Jugendliche eines Staates tatsächlich die Nützlichkeit verschiedener Lesestrategien in den Szenarien einschätzen konnten, wurde in diesem Beitrag ein Gesamtindex des Lesestrategiewissens (kurz: *Gesamtwissensindikator*; Cronbachs $\alpha = .89$) als Summenwert über die insgesamt 23 Paarvergleiche (OECD, 2017) der drei Szenarien gebildet. Insgesamt konnten somit 23 Punkte erreicht werden. Die Erfassung des Lesestrategiewissens beruht also nicht auf einer reinen Selbsteinschätzung, sondern erfolgt nach einem Gütemaßstab (vgl. Artelt, Neuenhaus, Lingel & Schneider, 2012). Somit haben kulturelle Unterschiede im Antwortverhalten einen geringeren Einfluss, als dies bei der Erfassung der Lesemotivation, des Leseverhaltens und des Selbstkonzepts der Fall ist.

4.2.2 Lesestrategiewissen im internationalen Vergleich

Das Lesestrategiewissen ist bei den fünfzehnjährigen Schülerinnen und Schülern in Deutschland mit 14,5 Punkten im Vergleich zum OECD-Durchschnitt (12,4 Punkte) überdurchschnittlich hoch ausgeprägt (Abbildung 4.14). Deutschland liegt mit diesem Wert sogar an der Spitze der 37 OECD-Staaten. Außer in Frankreich verfügen die Jugendlichen der anderen OECD-Staaten über ein signifikant geringeres Lesestrategiewissen als die Fünfzehnjährigen in Deutschland. Dazu gehören auch jene Staaten, deren Lesekompetenz im oberen Drittel liegt, darunter Estland, Kanada, Finnland und Irland. In Kanada liegt das Strategiewissen der Jugendlichen sogar signifikant unter dem Durchschnitt der OECD-Staaten. Die niedrigsten Werte erreichen Jugendliche in Litauen, Kolumbien und Island. Berücksichtigt man jedoch zusätzlich die absolute erreichte Punktzahl, so relativiert sich das Bild etwas: Auch in Deutschland werden im Durchschnitt nur etwas mehr als die Hälfte der Paarvergleiche (63 %) richtig gelöst.

Neben der Betrachtung des *Gesamtwissensindikators* ermöglichen es die Ergebnisse der einzelnen Leseszenarien, relative Stärken und Schwächen im Lesestrategiewissen zu identifizieren. Im OECD-Durchschnitt besteht eine relative Stärke beim *Zusammenfassen*. Hier werden 56 Prozent der Paarvergleiche richtig gelöst. Beim *Erinnern* sind es 52 Prozent und beim *Beurteilen* 51 Prozent. Für Deutschland zeigt sich eine ähnliche Tendenz (*Zusammenfassen* 64 %, *Erinnern* 62 %, *Beurteilen* 60 %), wobei alle Werte signifikant über dem OECD-Durchschnitt liegen. Auch in 20 weiteren OECD-Staaten zeigt sich die relative Stärke vor allem beim *Zusammenfassen* von Texten darunter Kanada, Polen sowie die Republik Korea als lesekompetenzstarke Staaten. Estland, Schweiz, Slowakei und die Türkei weisen zusätzlich ähnlich hohe Werte beim *Erinnern* auf. Wie auch in Deutschland gelingt es den Jugendlichen in den meisten Staaten (18) weniger, erfolgreiche Strategien beim *Beurteilen* von E-Mail-Inhalten auszuwählen. Ausnahmen sind Australien, Finnland, Japan, Litauen, Neuseeland und das Vereinigte Königreich, die in dieser Strategie ihre relative Stärke haben.



Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt ($p < .05$) sind **fett** hervorgehoben.

Abbildung 4.14: Lesestrategiewissen im internationalen Vergleich

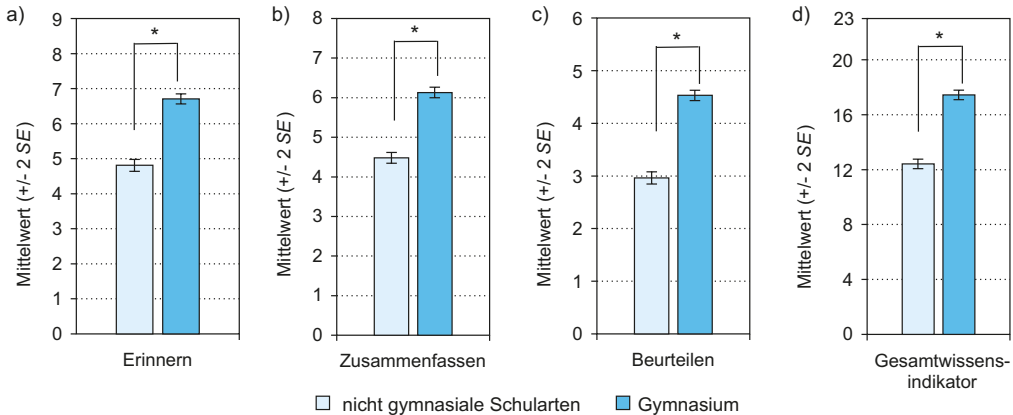
Betrachtet man die Streuung des Lesestrategiewissens in Deutschland (Abbildung 4.14), so ist sie ähnlich stark wie im Durchschnitt der OECD-Staaten. Diese durchschnittliche Streuung von knapp sechs Punkten umfasst rund ein Viertel der gesamten Paarvergleiche beziehungsweise die Hälfte der durchschnittlich erreichten Punkte im Mittel aller OECD-Staaten. Es gibt entsprechend beachtliche Unterschiede zwischen den Fünfzehnjährigen innerhalb der Staaten. Auch in jenen Staaten, welche eine signifikant niedrigere Streuung als der OECD-Durchschnitt aufweisen, darunter Frankreich, Italien und Irland, betrug diese noch rund fünf Punkte. Die Streuung der anderen deutschsprachigen Staaten erweist sich als signifikant größer als der OECD-Durchschnitt mit über sechs Punkten.

4.2.3 Geschlechterunterschiede im Lesestrategiewissen

Geschlechterunterschiede im Lesestrategiewissen (Abbildung 4.14) zeigen sich in Deutschland sowie in allen OECD-Staaten zugunsten der Mädchen. Mädchen lösen in Deutschland im Schnitt 2.56 Paarvergleiche mehr als Jungen ($d = 0.44$). Dies ist vergleichbar mit der Geschlechterdifferenz in der Schweiz ($d = 0.48$) sowie in Estland ($d = 0.49$). Deutlich höher ausgeprägt ist der Geschlechterunterschied zum Beispiel in Finnland ($d = 0.77$). Betrachtet man die Szenarien einzeln, um Geschlechterunterschiede in den relativen Stärken und Schwächen zu identifizieren, so ist in Deutschland der Unterschied zwischen Mädchen und Jungen am geringsten für das Szenario *Beurteilen* von E-Mail-Inhalten ausgeprägt ($d = 0.59$; $d_{\text{Erinnern}} = 0.36$, $d_{\text{Zusammenfassen}} = 0.46$).

4.2.4 Unterschiede zwischen Schularten in Deutschland

Im *Gesamtwissensindikator* (Abbildung 4.15) erreichen Gymnasiastinnen und Gymnasiasten im Schnitt 17.44 Punkte ($SD = 4.72$), Schülerinnen und Schüler nicht gymnasialer Schularten erreichen 12.41 Punkte ($SD = 5.85$). An den Gymnasien ist die Identifizierung erfolgreicher Strategien über die drei Szenarien hinweg ausgeglichen: Die Jugendlichen lösen im Schnitt drei Viertel der Paarvergleiche richtig (*Erinnern* 74.5%, *Zusammenfassen* 76.6%, *Beurteilen* 75.5%). Jugendliche nicht gymnasialer Schularten zeigen hingegen eine relative Stärke beim *Zusammenfassen* von Texten (56.0%) und eine Schwäche beim *Beurteilen* von E-Mail Inhalten (49.3%). Beim *Erinnern* lösen sie 53.4 Prozent der Paarvergleiche richtig.



Anmerkung: Mittelwert der gelösten Paarvergleiche im Lesestrategiewissen differenziert nach Schularten. * Statistisch signifikante Unterschiede zwischen nicht gymnasialen Schularten und dem Gymnasium ($p < .05$).

Abbildung 4.15: Lesestrategiewissen nach Schularten

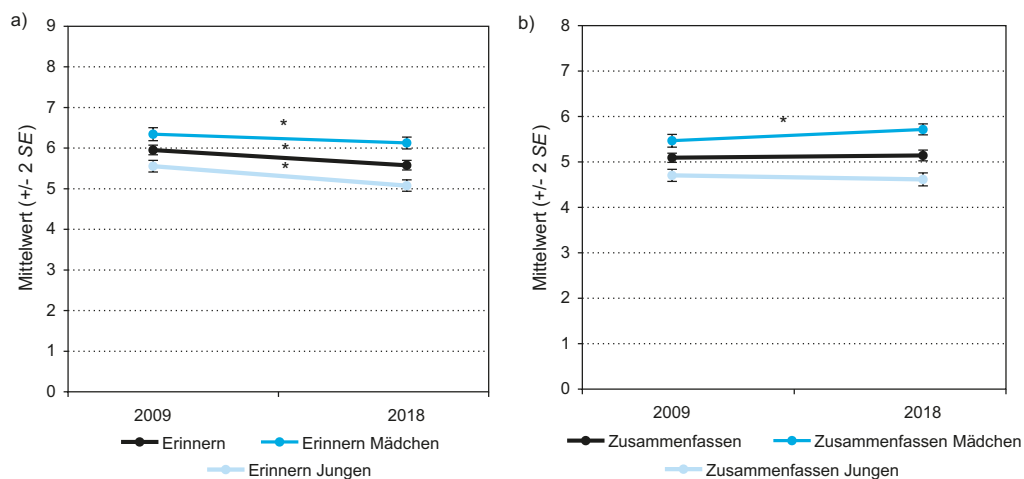
4.2.5 Veränderungen zwischen den Erhebungen 2009 und 2018

Da die beiden Szenarien *Erinnern* und *Zusammenfassen* bereits bei PISA 2009 zum Einsatz kamen, als Lesen das letzte Mal Hauptdomäne war, kann auch ihre Veränderung betrachtet werden. Die Abbildung 4.16 zeigt die Entwicklung des Lesestrategiewissens in den beiden Szenarien zwischen 2009 und 2018.

Die Ausprägung des Lesestrategiewissens im Szenario *Erinnern* hat sich von PISA 2009 zu PISA 2018 in Deutschland signifikant verringert (-4 Prozentpunkte). Ähnliches lässt sich im OECD-Durchschnitt beobachten (-5 Prozentpunkte). Ein Rückgang zeigt sich auch in den lesestarken Staaten Estland (-5 Prozentpunkte), Finnland (-7 Prozentpunkte) und Irland (-8 Prozentpunkte) sowie in den Staaten, deren Lesekompetenz sich nicht signifikant vom OECD-Durchschnitt unterscheidet, wie zum Beispiel Tschechien (-11 Prozentpunkte), Österreich (-8 Prozentpunkte) und die Schweiz (-4 Prozentpunkte). Einzig im Vereinigten Königreich (7 Prozentpunkte) zeigten die Fünfzehnjährigen bei PISA 2018 ein signifikant höheres Lesestrategiewissen beim *Erinnern* von Texten.

Die Ausprägung des Lesestrategiewissens im Szenario *Zusammenfassen* von Texten hat sich im OECD-Durchschnitt (-4 Prozentpunkte) im Vergleich zu PISA 2009 ebenfalls signifikant verringert. Für die deutschen Schülerinnen und Schüler gilt das jedoch nicht, ihr Wissen blieb stabil. Das trifft auch auf ein Drittel der OECD-Staaten zu, darunter Luxemburg und das Vereinigte Königreich. Eine Verbesserung des Lesestrategiewissens beim *Zusammenfassen* von Texten zeigt sich in keinem der OECD-Staaten.

Betrachtet man die Entwicklung getrennt für Mädchen und Jungen, so nahm die erfolgreiche Anwendung von Strategien im Szenario *Erinnern* in Deutschland sowohl bei Mädchen (von $M_{2009} = 6.34$ auf $M_{2018} = 6.13$) als auch bei Jungen (von $M_{2009} = 5.56$ auf $M_{2018} = 5.08$) ab, wobei die Differenz von -0.21 bei den Mädchen deutlich kleiner ist als die Differenz von -0.48 bei den Jungen. Von allen OECD-Staaten kam es lediglich



Anmerkung: * Statistisch signifikante Unterschiede zwischen den Erhebungen 2009 und 2018 ($p < .05$).

Abbildung 4.16: Veränderungen des Lesestrategiewissens in den Szenarien *Erinnern* und *Zusammenfassen* in Deutschland zwischen PISA 2009 und PISA 2018

im Vereinigten Königreich zu einer Verbesserung des Lesestrategiewissens im Szenario *Erinnern* für beide Geschlechter. In der Türkei, Schweden und Polen verbesserten sich nur die Mädchen signifikant von 2009 zu 2018.

Für das Szenario *Zusammenfassen* verschlechterte sich das Wissen der Mädchen und Jungen im OECD-Durchschnitt gleichermaßen. In Deutschland zeigt sich jedoch ein anderes Bild: Während die Identifizierung erfolgreicher Strategien bei den Jungen stabil blieb ($M_{2009_Ju} = 4.70$, $M_{2018_Ju} = 4.62$), verbesserte sich dieses Lesestrategiewissen bei den Mädchen sogar (von $M_{2009_Mä} = 5.47$ auf $M_{2018_Mä} = 5.72$). Ein ähnliches Bild wie in Deutschland zeigte sich auch in den Staaten Griechenland, Niederlande, Spanien und Türkei. Auch hier verbesserten sich die Mädchen signifikant von 2009 zu 2018, wohingegen kein signifikanter Unterschied bei den Jungen vorlag.

Betrachtet man die Entwicklung des Lesestrategiewissens getrennt nach Schularten, nahm das Wissen im Szenario *Erinnern* in beiden Schularten ab; an den nicht gymnasialen Schularten ($M_{Diff_nongy} = -0.52$) noch deutlicher als an den Gymnasien ($M_{Diff_Gy} = -0.40$). Das Strategiewissen bezüglich des *Zusammenfassens* von Texten blieb erfreulicherweise in beiden Schularten stabil.

4.3 Zusammenfassung und Diskussion

Lesebezogene Schülermerkmale wie Lesefreude, Selbstkonzept oder Lesestrategiewissen sind neben der Lesekompetenz wichtige Aspekte des Lesens, ihre Entwicklung und Förderung ein wichtiges Ziel schulischer Bildung. Entwicklung und Förderung brauchen deshalb besondere Aufmerksamkeit. Die Ergebnisse der aktuellen PISA-Erhebungsrunde

machen für Deutschland Stärken und Schwächen deutlich. Fünfzehnjährige Schülerinnen und Schüler in Deutschland verfügen im Vergleich zu den anderen OECD-Staaten über das höchste Lesestrategiewissen. Betrachtet man die absolute Anzahl der durchschnittlich gelösten Aufgaben, so werden in Deutschland rund 60 Prozent der Strategien richtig erkannt. Getrennt für die drei Szenarien zeigt sich, dass das *Zusammenfassen* von Texten Schülerinnen und Schülern in Deutschland tendenziell leichter fällt, als den richtigen Umgang mit einer Spam-E-Mail zu erkennen. Jungen verfügen in allen Szenarien über ein geringeres Lesestrategiewissen als Mädchen. Diese Ergebnisse zur Ausprägung des Lesestrategiewissens bei deutschen Schülerinnen und Schülern scheinen überraschend, berücksichtigt man die hohen Zusammenhänge des Lesestrategiewissens mit der Lesekompetenz (Artelt et al., 2009). Jugendliche in Deutschland sind zwar Spitzenreiter im internationalen Vergleich im Lesestrategiewissen, nicht aber in der Lesekompetenz. Das Wissen um erfolgreiche Strategien scheint nicht automatisch zu bedeuten, dass Fünfzehnjährige dieses Wissen auch anwenden. Eine Förderung sollte demnach nicht nur an der Vermittlung von Strategien ansetzen, sondern auch stärker deren Anwendung berücksichtigen.

Schwächen zeigen sich in Deutschland bei der Freude am Lesen. Diese ist im Vergleich zum OECD-Durchschnitt geringer ausgeprägt und betont in Deutschland vor allem Aspekte einer instrumentellen Motivation. Dies scheint sich auch auf das Leseverhalten auszuwirken, denn knapp die Hälfte der Schülerinnen und Schüler gibt an, nicht oder kaum zum Vergnügen zu lesen. Dabei ist Deutschland in der gleichen Situation wie ein Großteil der OECD-Staaten. Im Vergleich der Geschlechter sind mehr Jungen als Mädchen der Gruppe der „Lesemuffel“ zuzuordnen. Betrachtet man das Leseverhalten der „Lesemuffel“ genauer, so berichten sie, wenn überhaupt informative Textarten wie Tageszeitungen oder Zeitschriften zu lesen, während Jugendliche, die häufiger lesen, auch Prosa lesen. Auch beim Lesen digitaler Quellen überwiegen Nutzungsarten, die an Informationsaustausch geknüpft sind (z. B. Chatten). Betrachtet man darüber hinaus die Veränderungen in der Lesefreude und dem Leseverhalten seit der PISA-Erhebung 2009, so lassen sich diese durch eine Abnahme der Lesefreude und des Anteils des Lesens zum Vergnügen charakterisieren. Diese Entwicklung zeigt sich für beide Geschlechter und über die Schularten hinweg.

Im lesebezogenen Selbstkonzept zeigte sich ein komplexeres Bild. In Deutschland, wie auch in den anderen deutschsprachigen Staaten, nehmen Jugendliche ihre Kompetenz höher wahr, als es Schülerinnen und Schüler anderer OECD-Staaten im Durchschnitt berichten. Gleichzeitig nahmen die Fünfzehnjährigen die Schwierigkeit des Lesens nicht zwangsläufig kongruent wahr. Diese Inkongruenz ist besonders augenscheinlich bei den Mädchen: Zwar nahmen sie ihre Kompetenz im Lesen signifikant höher wahr als die Jungen, aber unterschieden sich nicht signifikant in ihrer Wahrnehmung der Schwierigkeit des Lesens, obschon Mädchen in Deutschland im Allgemeinen höhere Lesekompetenzwerte aufweisen.

Insgesamt zeigen die Befunde, dass Schülerinnen und Schüler in Deutschland für die gesellschaftliche Teilhabe über ein ausreichendes Wissen in Bezug auf den situationsad-

äquaten Umgang mit verschiedenen Leseaufgaben sowie Lesestrategien verfügen. Darüber hinaus liegt, auf nationaler Ebene, der Durchschnitt der wahrgenommenen Kompetenz im Lesen genauso wie auch die durchschnittliche Lesekompetenz der Jugendlichen in Deutschland über dem OECD-Durchschnitt. Entsprechend kann das Bildungsziel einer realistischen Selbsteinschätzung (Blossfeld et al., 2015) als erreicht angesehen werden. Allerdings ist in Deutschland – genau wie in anderen deutschsprachigen Staaten – die Motivation, außerhalb der Schule zu lesen, nur gering ausgeprägt. Da die lesebezogenen Schülermerkmale wechselseitig mit der Lesekompetenz zusammenhängen (Pfof et al., 2013), bedarf es einer verstärkten Förderung der Lesemotivation sowohl der Fünfzehnjährigen im Allgemeinen als auch leseschwacher Schülerinnen und Schüler im Speziellen (Mol & Bus, 2011). Die Förderung von Lesekompetenz, Lesemotivation und Leseverhalten ist als ganzheitlicher Prozess zu betrachten, welcher im Vorschulbereich beginnen und bis zum Ende der Schulzeit fortgeführt werden sollte (Artelt et al., 2007; Beck, Dewitz & Titz; Kapitel 3). Da die Lesemotivation im Jugendalter im Vergleich zur Grundschule abnimmt (Artelt et al., 2007), ist es gerade die durchgängige Förderung, die erfolgreich sein könnte. Die vorliegenden Ergebnisse verdeutlichen, dass weitere Bemühungen im schulischen und außerschulischen Bereich in Deutschland nötig sind, um die Lesemotivation sowie die Lesemenge der Kinder und Jugendlichen gezielt zu fördern.

Literatur

- Artelt, C., Beinicke, A., Schlagmüller, M. & Schneider, W. (2009). Diagnose von Strategiewissen beim Textverstehen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(2), 96–103. <https://doi.org/10.1026/0049-8637.41.2.96>
- Artelt, C., Demmrich, A. & Baumert, J. (2001). Selbstreguliertes Lernen. In J. Baumert (Hrsg.), *PISA 2000* (S. 271–297). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_8
- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., ... Saalbach, H. (2007). *Förderung von Lesekompetenz – Expertise*. Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- Artelt, C., Naumann, J. & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009: Bilanz nach einem Jahrzehnt* (S. 73–112). Münster: Waxmann.
- Artelt, C., Neuenhaus, N., Lingel, K. & Schneider, W. (2012). Entwicklung und wechselseitige Effekte von metakognitiven und bereichsspezifischen Wissenskomponenten in der Sekundarstufe. *Psychologische Rundschau*, 63(1), 18–25. <https://doi.org/10.1026/0033-3042/a000106>
- Baumert, J. (Hrsg.). (2001). *PISA 2000*. Opladen: Leske + Budrich. <https://doi.org/10.1007/978-3-322-83412-6>
- Beck, L., von Dewitz, N. & Titz, C. (2015) Durchgängige Leseförderung. *BiSS-Journal Bildung durch Sprache und Schrift*, (3), 4–11. Verfügbar unter <https://biss-sprachbildung.de/pdf/biss-journal-3-november-2015-1.pdf>

- Blossfeld, H.-P., Bos, W., Daniel, H.-D., Hannover, B., Köller, O., Lenzen, D., ... Wößmann, L. (2015). *Bildung. Mehr als Fachlichkeit. Gutachten*. Münster: Waxmann.
- Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research*, 31(6), 445–457. [https://doi.org/10.1016/S0883-0355\(99\)00014-2](https://doi.org/10.1016/S0883-0355(99)00014-2)
- Callan, G. L., Marchant, G. J., Finch, W. H. & German, R. L. (2016). Metacognition, strategies, achievement, and demographics: Relationships across countries. *Educational Sciences: Theory & Practice*, 16(5). <https://doi.org/10.12738/estp.2016.5.0137>
- Chapman, J. W. & Tunmer, W. E. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, 87(1), 154–167. <https://doi.org/10.1037/0022-0663.87.1.154>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911. <https://doi.org/10.1037//0003-066X.34.10.906>
- Guthrie, J. T., Klauda, S. L. & Ho, A. N. (2013). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly*, 48(1), 9–26. <https://doi.org/10.1002/rrq.035>
- Heine, S. J., Lehman, D. R., Peng, K. & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology*, 82(6), 903–918. <https://doi.org/10.1037//0022-3514.82.6.903>
- Klauda, S. L. & Guthrie, J. T. (2015). Comparing relations of motivation, engagement, and achievement among struggling and advanced adolescent readers. *Reading and Writing*, 28(2), 239–269. <https://doi.org/10.1007/s11145-014-9523-2>
- Lafontaine, D., Dupont, V., Jaegers, D. & Schillings, P. (2019). Self-concept in reading: Factor structure, cross-cultural invariance and relationships with reading achievement in an international context (PIRLS 2011). *Studies in Educational Evaluation*, 60, 78–89. <https://doi.org/10.1016/j.stueduc.2018.11.005>
- 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. <https://doi.org/10.1037/a0021890>.
- OECD. (in Vorbereitung). *PISA 2018 technical report*. Paris: OECD Publishing.
- OECD. (2017). *PISA 2015 technical report*. Verfügbar unter <http://www.oecd.org/pisa/data/2015-technical-report/>
- OECD. (2019a). PISA 2018 questionnaire framework. In OECD (Hrsg.), *PISA 2018 assessment and analytical framework* (S. 217–256). Paris: OECD Publishing. <https://doi.org/10.1787/850d0ef8-en>
- OECD. (2019b). PISA 2018 reading framework. In OECD (Hrsg.), *PISA 2018 assessment and analytical framework* (pp. 21–71). Paris: OECD Publishing. <https://doi.org/10.1787/5c07e4f1-en>
- Pfost, M., Dörfler, T. & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, 26, 89–102. <https://doi.org/10.1016/j.lindif.2013.04.008>
- Pfost, M., Hattie, J., Dörfler, T. & Artelt, C. (2014). Individual differences in reading development. *Review of Educational Research*, 84(2), 203–244. <https://doi.org/10.3102/0034654313509492>

- 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-Tiska, A. (in Druck). *Konzeptionierung und Erfassung mehrdimensionaler Bildungsziele sowie deren schulische und unterrichtliche Rahmenbedingungen in internationalen Large-Scale Assessments am Beispiel der MINT-Fächer*. (Habilitationsschrift). Technische Universität München, München.
- UNESCO. (2006). *Education for all: Literacy for life*. (EFA Global Monitoring Report, Vol. 4). Paris: Unesco.
- Van de Vijver, F. J. R. & He, J. (2017). Bias assessment and prevention in noncognitive outcome measures in context assessments. In N. Jude, D. Kaplan, E. Kileme & S. Kuger (Hrsg.), *Assessing contexts of learning: An international perspective* (Methodology of educational measurement and assessment, Bd. 36, S. 229–253). Cham: Springer. https://doi.org/10.1007/978-3-319-45357-6_9
- Viljaranta, J., Tolvanen, A., Aunola, K. & Nurmi, J.-E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research*, 58(6), 734–756. <https://doi.org/10.1080/00313831.2014.904419>
- Wigfield, A. (1997). Reading motivation: A domain-specific approach to motivation. *Educational Psychologist*, 32(2), 59–68. https://doi.org/10.1207/s15326985ep3202_1
- ZIB. (in Vorbereitung). *PISA 2018 Skalenhandbuch*. Münster: Waxmann.

5 Schulische Lerngelegenheiten zur Sprach- und Leseförderung im Kontext der Digitalisierung

Sarah Hofer, Doris Holzberger, Jörg-Henrik Heine, Frank Reinhold, Anja Schiepe-Tiska, Mirjam Weis & Kristina Reiss

Dieses Kapitel beschreibt ausgewählte schulische Lerngelegenheiten zur Sprach- und Leseförderung und fokussiert dabei auf den Stand der Digitalisierung an Schulen als Rahmenbedingung für das Lesen digitaler Texte sowie auf Lernangebote, die zusätzlich zum Regelunterricht zur Verfügung stehen. Schulen in Deutschland sind im internationalen Vergleich hinsichtlich der Anzahl und Qualität digitaler Ressourcen weniger gut ausgestattet. Auch in Bezug auf personelle Ressourcen, die Integration digitaler Medien in den Schulalltag sowie deren Einsatz im Deutschunterricht werden für Deutschland größtenteils unterdurchschnittliche Werte berichtet. Allerdings liegen an Schulen in Deutschland signifikant häufiger als im OECD-Durchschnitt theoretisch ausgearbeitete Konzepte zum fachspezifischen Einsatz digitaler Medien im Unterricht vor. Außerunterrichtliche Lerngelegenheiten der Landessprache, wie Debattierklubs oder Projekte mit Bibliotheken, werden an deutschen Schulen im internationalen Vergleich seltener angeboten. Allerdings besteht in Deutschland und auch im Mittel der OECD-Staaten an etwa der Hälfte der Schulen die Möglichkeit, sich durch die Beteiligung an einem Schultheater oder Musical sowie die Erstellung eines Schuljahrbuchs, einer Schüler- oder Klassenzeitung vertieft mit der jeweiligen Landessprache auseinanderzusetzen. Der Schwerpunkt bei der Ausrichtung zusätzlicher Deutschstunden liegt in Deutschland auf Stütz- und Förderunterricht. Für Jugendliche, deren Herkunftssprache nicht der Landessprache entspricht, werden in Deutschland insbesondere ergänzender Förderunterricht und auf den Regelunterricht vorbereitende Kurse zum Spracherwerb angeboten – die Häufigkeit dieser Maßnahmen liegt über dem OECD-Durchschnitt und hat seit PISA 2009 deutlich zugenommen. Beim Vergleich gymnasialer und nicht gymnasialer Schulen zeigt sich eine intensivere Nutzung digitaler Ressourcen im Deutschunterricht an nicht gymnasialen Schulen. Es wird aber auch deutlich, dass den Schülerinnen und Schülern an nicht gymnasialen Schulen weniger Möglichkeiten zur Sprach- und Leseförderung zur Verfügung stehen, die über den Regelunterricht hinausgehen.

Kompetenzunterschiede von Schülerinnen und Schülern können teilweise auf Unterschiede in Schulmerkmalen zurückgeführt werden. Dazu werden vielfältige schulische Rahmenbedingungen betrachtet, welche zum Teil von den Schulen selbst gestaltet und zum Teil vom Schulsystem vorgegeben sind (z. B. Klieme, 2013; Reynolds, 2014; Scheerens, 2004). Ziel des vorliegenden Kapitels ist es, insbesondere auf diejenigen Rahmenbedingungen näher einzugehen, welche als Lerngelegenheiten zur Entwicklung der Lesekompetenz aufgefasst werden können. Schulische Lerngelegenheiten zur Sprach- und Leseförderung bieten sich hier nicht nur im Fachunterricht (Shanahan & Shanahan, 2008). Eine tiefe Auseinandersetzung mit der gesprochenen und geschriebenen Sprache – gerade auch im Kontext des Verarbeitens online verfügbarer Informationen – kann ebenso außerhalb des Deutschunterrichts erfolgen. Entsprechend können an den Schulen vielfältige Möglichkeiten zur Förderung der Sprach- und Lesekompetenz realisiert werden. Diese umfassen etwa die Bereitstellung und Nutzung digitaler Technologien zur Förderung des Lesens digitaler Texte (vgl. Kapitel 2) genauso wie nicht durch den Lehrplan vorgeschriebene extracurriculare Lernangebote, die eine vertiefte Auseinandersetzung mit der Landessprache erlauben.

So liefert das Kapitel ein empirisch fundiertes Bild der aktuellen Situation der Digitalisierung an Schulen in Deutschland. Das Angebot und die Nutzung digitaler Ressourcen (häufig auch als „Information and Communication Technology“ bzw. ICT bezeichnet) spielen eine zunehmend wichtige Rolle bei der Gestaltung schulischer Lerngelegenheiten und sind Voraussetzung für die Kompetenzentwicklung in Bezug auf das Lesen, Verstehen und Bewerten digitaler Texte (vgl. Gorski, 2009; vgl. Kapitel 2 und 3). Das Kapitel gibt zudem einen Überblick über Angebote und Lerngelegenheiten, die an den Schulen außerhalb des Regelunterrichts der Landessprache bereitgestellt werden – von Zusatzangeboten wie dem Schultheater bis hin zum Förderunterricht (vgl. Helmke & Schrader, 2014; Linberg, Struck & Bäumer, 2018). Ein zentraler Aspekt ist die Darstellung des Istzustands an Schulen in Deutschland. Bei PISA wird für den internationalen Vergleich der Durchschnitt der OECD-Staaten als Vergleichsgröße genutzt. So kann abgeschätzt werden, wo Deutschland relativ zu den anderen OECD-Staaten einzuordnen ist. Um mögliche Hinweise auf besonders förderliche schulische Rahmenbedingungen im Kontext der Lesekompetenz zu erkennen, findet zum Teil ein Vergleich mit solchen Staaten statt, in denen die mittlere Lesekompetenz signifikant höher ist als in Deutschland (d. h. mit den besonders lesestarken Staaten Estland, Kanada, Finnland, Irland, der Republik Korea, Polen und Neuseeland; vgl. Kapitel 3). Zum Teil erfolgt eine Gegenüberstellung mit den deutschsprachigen Nachbarstaaten, für die aufgrund des gemeinsamen Sprachraums und der Nähe der Bildungssysteme eine relativ hohe Vergleichbarkeit gegeben ist. Zugleich legt dieses Kapitel einen speziellen Fokus auf den Vergleich der schulischen Lerngelegenheiten zur Sprach- und Leseförderung verschiedener Schularten in Deutschland (vgl. Kapitel 1). Der erste Teil des Kapitels widmet sich den Rahmenbedingungen für das Lesen digitaler Texte an Schulen und geht dabei sowohl auf ICT-Ressourcen als auch spezifisch auf den Einsatz digitaler Medien im Unterricht der Landessprache ein.

Im zweiten Teil wird die Vielfalt und Ausgestaltung des Lernangebots im Bereich der Sprach- und Leseförderung außerhalb des Regelunterrichts in den Blick genommen.

5.1 Rahmenbedingungen für das Lesen digitaler Texte

Metaanalysen zeigen positive Effekte des Einsatzes digitaler Medien im Unterricht der Sekundarstufe auf die Leistung und Motivation (Hillmayr, Reinhold, Ziernwald & Reiss, 2017) sowie speziell auf das Leseverständnis von Schülerinnen und Schülern (Moran, Ferdig, Pearson, Wardrop & Blomeyer, 2008). Digital unterstütztes Lehren und Lernen hat das Potenzial die Entwicklung diverser Kompetenzen und insbesondere die Entwicklung der Kompetenz des Lesens in digitalen Medien zu fördern. Für das erfolgreiche Lesen digitaler Texte spielen sowohl die allgemeine Lesekompetenz als auch grundlegende Fähigkeiten im Umgang mit Computern und im Einschätzen der Nützlichkeit und Vertrauenswürdigkeit von Online-Informationen eine Rolle. Fähigkeiten im Umgang mit Computern bedingen einen sicheren Umgang mit digitalen Medien, welcher die Lesekompetenz positiv beeinflusst (vgl. Kapitel 3; Hahnel, Goldhammer, Naumann & Kröhne, 2016). Die erfolgreiche Integration digitaler Medien in schulischen Kontexten stellt daher eine zentrale Herausforderung im heutigen Bildungswesen dar (OECD, 2018).

Dabei erscheint insbesondere die Frage nach einem gleichberechtigten Zugang zu digitalen Medien relevant, denn es können nur diejenigen Schülerinnen und Schüler von Bildungstechnologien profitieren, die auch Zugang dazu haben (Gorski, 2009). Die in PISA erhobenen Aussagen der Schulleitungen sowie der Schülerinnen und Schüler können als Indikator für den Stand der Digitalisierung im schulischen Bereich gesehen werden. Sie ermöglichen durch den internationalen Vergleich – etwa mit den strukturell ähnlichen deutschsprachigen Nachbarstaaten – die Identifikation möglicher Problem-bereiche. Im Folgenden sollen zunächst materielle und personelle Ressourcen im Kontext der Digitalisierung dargestellt werden. Anschließend wird auf die Integration digitaler Ressourcen in den Schulalltag eingegangen. Diese genannten Aspekte beziehen sich zunächst nicht spezifisch auf den Unterricht der Landessprache, konkret also den Deutschunterricht. Sie geben einen Einblick in den Istzustand der Digitalisierung an Schulen, die fächerübergreifend neue Lerngelegenheiten zur Sprach- und insbesondere Leseförderung ermöglichen kann (vgl. Shanahan & Shanahan, 2008). Der fachspezifische Einsatz digitaler Ressourcen im Unterricht der Landessprache wird im letzten Abschnitt näher betrachtet.

5.1.1 Materielle und personelle ICT-Ressourcen

Die Angaben zur Anzahl digitaler Ressourcen pro Schülerin beziehungsweise Schüler in der neunten Klassenstufe beruhen auf den Einschätzungen der Schulleitungen (siehe Tabelle 5.1). Es zeigt sich, dass die Schulen in Deutschland hinsichtlich der Ausstattung mit digitalen Geräten signifikant unter dem OECD-Durchschnitt liegen. Die einzige Ausnahme stellt die Anzahl der Computer mit Internetanschluss dar, die für Lehrkräfte zur Verfügung stehen. Hier berichten die Schulleitungen in Deutschland einen Wert, der sich nicht signifikant vom OECD-Durchschnitt unterscheidet. Schulen in den deutschsprachigen Nachbarstaaten Luxemburg und Österreich sind in Bezug auf die Anzahl der Computer, die Anzahl der Computer mit Internetanschluss sowie die Anzahl tragbarer Computer pro Neuntklässlerin beziehungsweise Neuntklässler laut Angaben der Schulleitungen deutlich besser ausgestattet als Schulen in Deutschland.

Betrachtet man die Entwicklung seit PISA 2015, so hat in Deutschland nur die Anzahl tragbarer Computer pro Schülerin beziehungsweise Schüler in der neunten Klassenstufe signifikant zugenommen (siehe Abbildung 5.1). In Bezug auf die anderen in Tabelle 5.1 angegebenen ICT-Ressourcen lassen sich keine signifikanten Veränderungen seit PISA 2015 feststellen. Es zeigt sich allerdings eine positive – wenn auch nicht signifikante – Veränderung in der an der Schülerzahl relativierten Anzahl der Computer mit Internetanschluss, die für Lehrkräfte im Mittel zur Verfügung stehen ($M = 0.24$ in 2015 und $M = 0.49$ in 2018). In PISA 2018 variieren dabei die Angaben der Schulleitungen deutlich zwischen den Schulen ($SD = 1.73$ in 2018 verglichen mit $SD = 0.53$ in 2015).

Beim Vergleich von Schularten zeigt sich nur bei der Ausstattung mit tragbaren Computern ein signifikant geringerer Wert bei nicht gymnasialen Schulen (siehe Abbildung 5.1). Hinsichtlich der Anzahl der Computer im Allgemeinen, der Anzahl der Computer mit Internetanschluss, der Anzahl interaktiver Whiteboards und der Anzahl der Computer mit Internetanschluss für Lehrkräfte (jeweils pro Neuntklässler/-in) werden von Schulleitungen nicht gymnasialer Schulen zwar durchwegs niedrigere Werte berichtet, die Unterschiede sind allerdings nicht signifikant. Betrachtet man die Entwicklung seit PISA 2015, so hat sich die Situation offenbar nur an den Gymnasien merklich verbessert. Die Anzahl der tragbaren Computer pro Neuntklässlerin beziehungsweise

Tabelle 5.1: Ausstattung mit materiellen ICT-Ressourcen pro Schülerin beziehungsweise Schüler in der neunten Klassenstufe

Pro Schülerin bzw. Schüler in der 9. Klassenstufe	Deutschland		OECD-Durchschnitt	
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)
Anzahl Computer	0.61	(0.05)	0.90	(0.02)
Anzahl Computer mit Internetanschluss	0.57	(0.05)	0.88	(0.03)
Anzahl tragbarer Computer	0.17	(0.02)	0.42	(0.02)
Anzahl interaktiver Whiteboards	0.10	(0.01)	0.14	(0.00)
Anzahl Computer mit Internetanschluss für Lehrkräfte	0.49	(0.13)	0.57	(0.01)

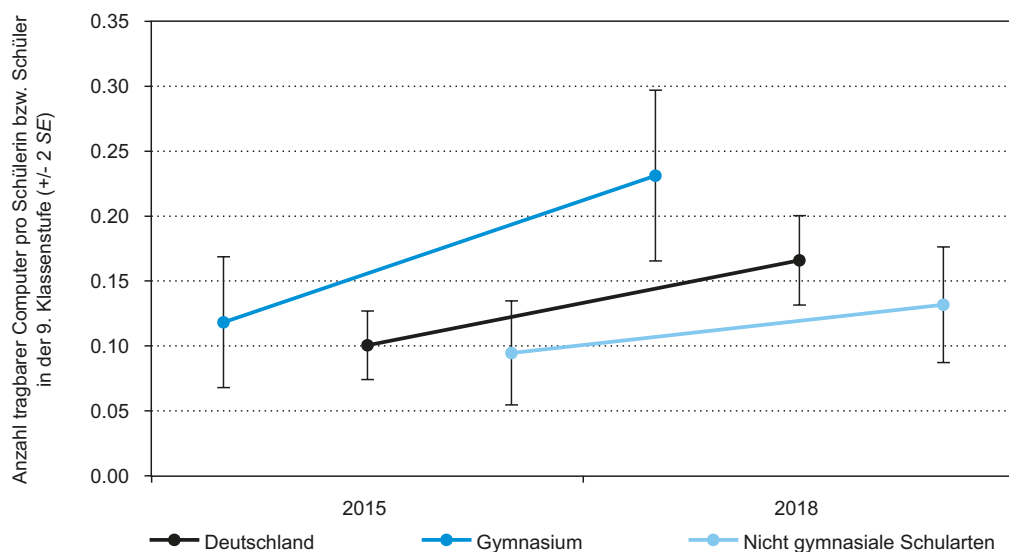


Abbildung 5.1: Entwicklung der Anzahl tragbarer Computer pro Schülerin beziehungsweise Schüler in der neunten Klassenstufe zwischen PISA 2015 und PISA 2018 nach Schulart

Neuntklässler ist an den Gymnasien von $M = 0.12$ in 2015 auf $M = 0.23$ in 2018 angestiegen. Ebenso nahm die Anzahl interaktiver Whiteboards pro Neuntklässlerin beziehungsweise Neuntklässler zwischen PISA 2015 ($M = 0.06$) und PISA 2018 ($M = 0.10$) an Gymnasien signifikant zu. Wie für Deutschland insgesamt so finden sich auch innerhalb der Schularten auffällig hohe Streuungen in den Angaben der Schulleitungen zur Anzahl der Computer mit Internetanschluss speziell für Lehrkräfte, die auf deutliche Unterschiede zwischen den Schulen hinweisen.

In PISA 2018 beantworteten die Schulleitungen weitere Fragen zur Qualität der materiellen ICT-Ressourcen sowie zu personellen ICT-Ressourcen, die in PISA 2015 nicht enthalten waren. Aus der Perspektive der Schulleitungen zeichnet sich auch hier ein sehr eindeutiges, durchweg wenig positives Bild ab (Tabelle 5.2). Schulleitungen in Deutschland nehmen sowohl die Qualität der Internetverbindung und die Rechenleistung der digitalen Geräte als auch die Verfügbarkeit geeigneter Software und einer effektiven Online-Lernplattform als signifikant schlechter wahr als Schulleitungen im OECD-Durchschnitt. Bei der Qualität der Internetanbindung weist Deutschland fast die niedrigsten Werte aller OECD-Staaten auf.

Auch hinsichtlich personeller ICT-Ressourcen liegt Deutschland nach Einschätzung der Schulleitungen signifikant unter dem OECD-Durchschnitt. So zeigen sich Defizite in der technischen und pädagogischen Kompetenz der Lehrkräfte für den Einsatz digitaler Geräte, beim Zugang zu effektiven professionellen Weiterbildungsmaterialien zur Nutzung digitaler Geräte sowie in der Verfügbarkeit von qualifiziertem Personal zur technischen Unterstützung. Deutschland liegt bei letzterem in der Wahrnehmung der Schulleitungen im unteren Drittel der OECD-Staaten.

Tabelle 5.2: Qualität materieller sowie personelle ICT-Ressourcen

Prozentualer Anteil von Schülerinnen und Schülern, deren Schulleitung den folgenden Aussagen zustimmte														
	Bandbreite und Geschwindigkeit der Internetverbindung ausreichend		Rechenleistung der digitalen Geräte ausreichend		Ausreichende Verfügbarkeit geeigneter Software		Online-Lernplattform zur Unterstützung verfügbar		Technische und pädagogische ICT-Kompetenz der Lehrkräfte ausreichend		Zugang zu Lernmaterialien für Lehrkräfte zur Nutzung digitaler Geräte		Qualifiziertes Personal zur technischen Unterstützung ausreichend	
OECD-Staaten	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Estland	74.8	(1.6)	77.3	(1.8)	82.0	(1.6)	66.5	(2.0)	63.6	(2.1)	79.3	(1.8)	63.3	(2.0)
Kanada	81.4	(2.1)	84.2	(1.7)	87.3	(1.4)	65.1	(1.7)	69.0	(2.3)	77.6	(1.9)	61.9	(2.6)
Finnland	72.9	(3.5)	77.9	(3.0)	75.3	(3.1)	80.0	(2.6)	50.1	(3.1)	61.9	(3.5)	64.8	(3.3)
Irland	75.9	(3.8)	73.2	(3.7)	71.9	(3.9)	45.4	(4.0)	49.3	(4.4)	47.4	(4.1)	20.7	(3.1)
Republik Korea	83.4	(2.9)	75.6	(3.3)	69.7	(3.5)	55.8	(3.7)	83.2	(3.0)	51.9	(3.8)	36.9	(3.6)
Polen	58.9	(3.4)	51.1	(3.7)	54.4	(3.3)	34.7	(3.0)	78.5	(2.5)	66.9	(2.9)	31.4	(3.0)
Schweden	89.1	(2.3)	93.0	(1.8)	87.2	(2.4)	80.0	(2.8)	72.0	(3.1)	83.2	(2.6)	76.6	(2.9)
Neuseeland	87.9	(2.4)	94.3	(1.6)	93.1	(1.4)	76.5	(2.9)	60.5	(3.2)	73.9	(3.5)	77.2	(2.6)
Vereinigte Staaten	82.4	(3.3)	89.3	(2.7)	86.7	(2.9)	77.1	(2.9)	73.6	(3.5)	77.5	(3.0)	60.4	(3.7)
Vereinigtes Königreich	75.2	(2.7)	66.8	(3.2)	80.8	(2.4)	65.9	(3.8)	72.3	(3.3)	64.5	(3.5)	69.3	(3.2)
Japan	45.2	(3.8)	46.1	(4.2)	39.7	(3.9)	24.0	(3.5)	27.3	(3.5)	19.1	(2.8)	10.2	(2.3)
Australien	72.4	(1.8)	84.3	(1.6)	92.6	(1.1)	75.9	(1.7)	67.7	(1.8)	72.3	(1.9)	76.2	(1.7)
Dänemark	89.9	(2.0)	83.7	(2.4)	83.5	(2.3)	90.9	(2.1)	80.2	(2.6)	82.2	(2.6)	79.5	(2.6)
Norwegen	79.9	(2.5)	81.0	(2.5)	82.3	(2.6)	76.1	(2.8)	74.5	(2.6)	72.7	(2.7)	90.1	(2.1)
Deutschland	31.7	(3.7)	58.8	(3.8)	59.2	(4.0)	32.7	(3.7)	56.7	(3.8)	40.7	(3.8)	34.4	(3.3)
Slowenien	90.0	(0.3)	81.6	(0.4)	86.4	(0.5)	77.4	(0.5)	77.2	(0.6)	78.0	(0.4)	61.1	(0.6)
Belgien	69.3	(2.4)	71.8	(2.8)	75.1	(2.2)	46.9	(2.7)	55.1	(2.8)	65.4	(2.8)	54.3	(2.8)
Frankreich	56.6	(3.4)	73.2	(3.2)	79.0	(2.7)	35.2	(3.3)	56.5	(3.2)	70.9	(3.3)	49.1	(3.6)
Portugal	32.0	(3.4)	30.7	(2.8)	44.3	(3.2)	34.9	(3.4)	62.5	(3.2)	53.4	(3.4)	27.9	(2.7)
Tschechien	71.6	(2.6)	54.3	(3.2)	68.6	(2.7)	57.0	(3.1)	63.4	(2.9)	88.6	(2.1)	48.2	(3.3)
OECD-Durchschnitt	67.5	(0.5)	68.5	(0.5)	71.3	(0.5)	54.1	(0.5)	64.6	(0.5)	64.7	(0.5)	54.1	(0.5)
Niederlande	87.1	(3.0)	84.1	(3.0)	87.3	(2.7)	50.4	(4.5)	51.6	(4.3)	70.7	(4.1)	78.1	(3.5)
Österreich	67.9	(3.0)	79.1	(3.1)	85.3	(2.7)	67.3	(2.9)	83.3	(2.6)	79.3	(2.7)	72.9	(2.9)
Schweiz	73.8	(3.6)	86.0	(2.4)	89.7	(2.2)	48.5	(3.8)	70.0	(3.0)	72.6	(3.8)	75.0	(3.2)
Lettland	79.1	(1.4)	65.5	(2.2)	80.7	(1.7)	51.3	(2.0)	79.8	(1.7)	57.3	(2.1)	42.7	(2.3)
Spanien	52.9	(2.3)	49.4	(1.9)	53.5	(2.1)	51.5	(2.6)	53.3	(2.3)	55.2	(2.0)	38.2	(2.2)
Italien	60.4	(3.1)	70.8	(3.0)	70.7	(3.0)	46.3	(3.3)	50.2	(3.4)	74.5	(3.0)	44.0	(3.5)
Ungarn	48.0	(3.7)	42.2	(3.5)	62.9	(3.8)	35.4	(3.0)	55.5	(4.0)	28.8	(3.4)	66.7	(3.5)
Litauen	91.3	(0.9)	77.3	(1.6)	71.5	(1.7)	66.8	(1.8)	84.1	(1.4)	77.8	(1.3)	71.1	(1.4)
Island	78.1	(0.2)	84.1	(0.2)	78.4	(0.2)	42.8	(0.3)	40.8	(0.3)	63.7	(0.2)	75.9	(0.2)
Israel	45.6	(3.8)	48.1	(3.6)	55.7	(3.8)	68.2	(2.9)	55.6	(3.5)	51.3	(3.7)	55.7	(3.8)
Luxemburg	78.8	(0.1)	89.4	(0.1)	89.6	(0.1)	23.9	(0.1)	59.3	(0.1)	65.4	(0.1)	46.9	(0.1)
Türkei	76.6	(3.0)	82.1	(2.9)	68.1	(3.4)	65.5	(3.3)	74.9	(3.2)	75.7	(3.0)	42.5	(3.2)
Slowakei	61.0	(2.7)	58.9	(2.8)	76.4	(2.6)	41.5	(3.0)	82.6	(2.0)	77.1	(2.4)	45.3	(3.0)
Griechenland	62.7	(2.7)	47.4	(3.0)	49.7	(3.1)	34.2	(3.6)	62.8	(3.1)	44.0	(2.9)	13.8	(2.2)
Chile	57.7	(3.6)	55.5	(3.8)	50.7	(4.1)	38.7	(3.7)	62.2	(3.7)	67.6	(3.7)	68.3	(3.6)
Mexiko	31.7	(2.9)	37.1	(3.3)	41.9	(3.3)	33.8	(3.0)	76.5	(3.0)	55.3	(3.0)	44.7	(3.4)
Kolumbien	25.2	(2.8)	28.4	(3.1)	27.8	(2.9)	36.2	(3.5)	55.5	(3.7)	49.8	(3.8)	27.1	(3.0)

In den drei deutschsprachigen Nachbarstaaten Luxemburg, Österreich und Schweiz nehmen die Schulleitungen den Stand der Digitalisierung insgesamt als deutlich besser wahr als Schulleitungen in Deutschland. Dasselbe gilt für die lesestarken Staaten Kanada, Estland und die Republik Korea. Damit zeigt sich ein konsistentes Bild: Aus Sicht der Schulleitungen haben Schulen in Deutschland Defizite in Bezug auf die Anzahl digitaler Ressourcen, ihre Qualität, die ICT-Kompetenz des pädagogischen und die Verfügbarkeit technischen Personals.

Zwischen Schulleitungen an Gymnasien und nicht gymnasialen Schulen zeigen sich keine signifikanten Unterschiede bezüglich der Angaben zur Qualität der materiellen ICT-Ressourcen sowie zu personellen ICT-Ressourcen. Schulleitungen an Gymnasien schätzen allerdings die technische und pädagogische Kompetenz der Lehrkräfte für den Einsatz digitaler Geräte etwas häufiger als ausreichend ein, als dies an nicht gymnasialen Schulen der Fall ist (60 % gegenüber 53 %). Ebenso stehen Online-Lernplattformen eher an Gymnasien (38 %) als an nicht gymnasialen Schularten (27 %) zur Verfügung. Schulleitungen an nicht gymnasialen Schularten berichten tendenziell von einer stärkeren technischen Unterstützung durch qualifiziertes Personal (37 %) als Schulleitungen an Gymnasien (29 %).

5.1.2 Integration digitaler Medien in den Schulalltag

Schulen können auf unterschiedliche Art und Weise die Integration digitaler Technologien in den Schulalltag fördern – etwa durch Anreize zur ICT-Nutzung für Lehrkräfte, durch zeitliche Ressourcen zur Vorbereitung von digital-unterstützten Unterrichtsstunden, durch regelmäßige Kooperation und den Austausch bezüglich der Verwendung digitaler Medien innerhalb des Kollegiums sowie mit der Schulleitung und durch Konzepte zum fachspezifischen Einsatz digitaler Geräte. Auch hierzu wurden bei PISA 2018 die Schulleitungen in Deutschland befragt (Tabelle 5.3). Sie berichten signifikant seltener von Anreizen für Lehrkräfte zur Integration digitaler Medien in den Unterricht, als dies im OECD-Durchschnitt der Fall ist. Nur 13 Prozent der Schulleitungen in Deutschland planen für ihre Lehrkräfte zeitliche Ressourcen zur Vorbereitung des Einsatzes digitaler Medien im Unterricht fest in den Schulalltag ein. Dies ist der zweitniedrigste Wert unter den OECD-Staaten. Nur jede fünfte Schule in Deutschland verfügt über ein Konzept zur Kooperation von Lehrkräften beim Einsatz digitaler Medien. Auch hier liegt Deutschland signifikant unter dem OECD-Durchschnitt. Regelmäßige Gespräche der Schulleitung mit Lehrkräften über die pädagogische Verwendung digitaler Medien finden an fast der Hälfte der Schulen in Deutschland statt, allerdings immer noch signifikant seltener als im OECD-Durchschnitt (63 %). Positiv ist anzumerken, dass an Schulen in Deutschland signifikant häufiger als im OECD-Durchschnitt theoretisch ausgearbeitete Konzepte zum fachspezifischen Einsatz digitaler Medien im Unterricht vorliegen. Die dargestellten Befunde treffen für Gymnasien und nicht gymnasiale Schularten gleichermaßen zu.

Tabelle 5.3: Integration digitaler Medien in den Schulalltag

Prozentualer Anteil von Schülerinnen und Schülern, deren Schulleitung den folgenden Aussagen zustimmt										
Meine Schule verfügt über die folgenden Dinge ...										
	Anreize für Lehrkräfte zur Integration digitaler Medien in ihren Unterricht		Geplante zeitliche Ressourcen für Lehrkräfte zur Vorbereitung des Einsatzes digitaler Medien		Konzept zur Förderung der Zusammenarbeit bei der Verwendung digitaler Medien		Regelmäßige Gespräche mit Lehrkräften über die pädagogische Verwendung digitaler Medien		Konzept zum fachspezifischen Einsatz digitaler Medien	
OECD-Staaten	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Estland	64.1	(1.9)	63.9	(2.0)	45.1	(2.1)	74.6	(2.1)	26.4	(1.8)
Kanada	34.5	(2.4)	44.7	(2.4)	40.0	(2.4)	72.0	(2.1)	55.2	(2.2)
Finnland	37.4	(2.9)	51.6	(3.1)	42.0	(3.7)	67.4	(3.4)	45.7	(3.3)
Irland	36.2	(4.2)	43.1	(4.0)	39.2	(4.0)	72.9	(3.6)	38.2	(4.0)
Republik Korea	19.1	(3.3)	36.9	(4.2)	31.4	(4.1)	31.8	(3.5)	75.5	(3.6)
Polen	95.1	(1.4)	40.1	(3.1)	26.5	(2.7)	78.9	(2.3)	15.2	(2.3)
Schweden	89.0	(2.2)	81.0	(2.7)	38.6	(3.3)	94.5	(1.6)	27.5	(3.5)
Neuseeland	31.3	(2.7)	64.5	(3.4)	61.7	(3.5)	87.7	(2.1)	75.8	(2.8)
Vereinigte Staaten	30.6	(4.3)	67.6	(3.8)	53.7	(3.8)	73.8	(3.9)	69.6	(4.0)
Vereinigtes Königreich	26.0	(3.4)	46.7	(3.6)	38.3	(3.7)	61.5	(3.0)	49.5	(3.9)
Japan	43.5	(3.6)	13.9	(2.7)	17.3	(3.0)	42.3	(3.9)	42.6	(3.8)
Australien	35.9	(1.8)	56.2	(1.8)	52.9	(2.2)	80.1	(1.6)	73.4	(1.9)
Dänemark	88.8	(2.2)	76.8	(3.3)	43.0	(3.6)	88.5	(2.1)	43.1	(3.5)
Norwegen	72.7	(3.1)	71.8	(2.9)	54.3	(3.4)	76.6	(2.3)	60.3	(3.2)
Deutschland	45.4	(4.1)	12.9	(2.5)	20.2	(3.1)	45.2	(4.1)	54.0	(3.9)
Slowenien	95.7	(0.3)	27.9	(0.6)	33.8	(0.5)	59.8	(0.4)	62.4	(0.6)
Belgien	59.6	(3.3)	36.2	(3.0)	35.1	(3.1)	61.9	(3.0)	39.4	(3.0)
Frankreich	71.8	(3.1)	22.7	(3.1)	43.8	(3.3)	69.8	(3.5)	32.9	(3.7)
Portugal	41.6	(3.3)	37.3	(3.2)	41.3	(3.2)	35.2	(2.5)	48.8	(3.5)
Tschechien	82.6	(2.2)	30.5	(2.6)	17.2	(2.5)	48.5	(2.9)	51.4	(3.2)
OECD-Durchschnitt	56.7	(0.5)	43.6	(0.5)	35.8	(0.5)	62.9	(0.5)	47.9	(0.5)
Niederlande	90.5	(2.6)	50.9	(4.2)	29.8	(3.4)	65.8	(4.6)	66.3	(4.3)
Österreich	72.0	(3.2)	11.1	(2.4)	34.9	(3.7)	59.4	(3.8)	54.6	(3.5)
Schweiz	65.2	(3.5)	35.3	(3.5)	22.4	(3.1)	50.1	(3.5)	41.2	(4.0)
Lettland	69.5	(2.0)	55.4	(2.1)	17.1	(1.6)	81.0	(1.7)	35.6	(2.2)
Spanien	11.5	(1.3)	18.6	(1.6)	31.4	(1.9)	30.8	(1.8)	36.3	(2.2)
Italien	49.4	(3.3)	59.7	(3.1)	47.9	(3.4)	36.8	(3.3)	40.6	(3.6)
Ungarn	78.9	(2.8)	25.2	(2.9)	32.7	(3.5)	49.0	(3.4)	30.5	(3.4)
Litauen	96.0	(0.7)	19.3	(1.4)	14.6	(1.4)	93.2	(0.9)	29.6	(1.8)
Island	90.6	(0.1)	55.7	(0.3)	36.7	(0.3)	74.8	(0.2)	52.7	(0.2)
Israel	27.8	(3.8)	39.5	(3.6)	43.3	(4.0)	39.2	(3.9)	62.6	(3.9)
Luxemburg	88.3	(0.1)	36.7	(0.1)	60.5	(0.1)	62.5	(0.1)	60.0	(0.1)
Türkei	91.9	(2.2)	57.0	(4.2)	42.9	(4.0)	67.6	(3.6)	60.3	(3.7)
Slowakei	67.5	(2.4)	63.3	(2.9)	21.3	(2.5)	75.8	(2.8)	33.5	(3.1)
Griechenland	33.8	(3.3)	24.6	(3.0)	23.0	(3.0)	44.1	(3.4)	60.9	(3.3)
Chile	21.7	(2.9)	40.6	(4.1)	21.5	(3.1)	50.0	(3.4)	32.5	(3.4)
Mexiko	18.8	(2.5)	53.0	(3.1)	34.6	(3.3)	61.5	(2.7)	45.4	(3.3)
Kolumbien	23.5	(2.9)	40.3	(3.5)	33.8	(3.2)	64.5	(3.3)	41.5	(3.6)

Nimmt man für den internationalen Vergleich speziell die lesestarken Staaten (z. B. Kanada, Estland, Finnland, Polen oder die Republik Korea) in den Blick, so zeigt sich kein einheitliches Bild hinsichtlich der Integration digitaler Medien in den Schulalltag (siehe Tabelle 5.3). Allerdings scheinen hier, speziell was die zeitlichen Ressourcen für Lehrkräfte sowie den gezielten Austausch im Kollegium und mit der Schulleitung betrifft, die Bedingungen für den tatsächlichen Einsatz digitaler Geräte im Unterricht günstiger zu sein.

5.1.3 Einsatz digitaler Ressourcen im Unterricht der Landessprache

Die Ausstattung einer Schule mit materiellen und personellen ICT-Ressourcen sowie Ansätze zur Integration digitaler Ressourcen in den Schulalltag stellen eine wichtige Voraussetzung für den tatsächlichen Einsatz digitaler Medien als Lerngelegenheit dar (vgl. Bos et al., 2014). Zur Frage, ob digitale Medien im Unterricht der Landessprache genutzt werden, liegen in PISA 2018 Angaben der Schülerinnen und Schüler vor. Dabei berichten etwa zwei Drittel der Jugendlichen in Deutschland, dass in ihrem Deutschunterricht in der Regel keine digitalen Geräte verwendet werden. Damit liegt der Anteil an Schülerinnen und Schülern, deren Deutschunterricht ohne regelmäßigen Einsatz digitaler Medien erfolgt, signifikant über dem OECD-Durchschnitt (51 %). Im deutschsprachigen Nachbarstaat Luxemburg gibt allerdings trotz besserer Voraussetzungen für den Einsatz digitaler Medien an Schulen (z. B. eine größere Anzahl an Computern; siehe Abschnitt 5.1.1) ein ähnlich großer Anteil an Jugendlichen an, in einer typischen Schulwoche im Deutschunterricht keine Zeit mit digitalen Medien zu verbringen. Auch im lesestarken Staat Irland scheint eine regelmäßige Nutzung nicht etabliert zu sein. Anders sieht es in Finnland, der Republik Korea und Estland aus, wo man sich bereits Mitte der 1990er Jahre intensiv mit der Nutzung digitaler Medien für das Lernen in der Schule auseinandersetzte (z. B. „Tiger-Leap-Program“ und „E-Estonia“). Hier berichten die Schülerinnen und Schüler von einem regelmäßigeren Einsatz digitaler Ressourcen im Unterricht der Landessprache (siehe Tabelle 5.4).

Betrachtet man die Situation in Deutschland, so scheinen Fünfzehnjährige an nicht gymnasialen Schulen in einer normalen Schulwoche im Deutschunterricht mehr Zeit mit digitalen Medien zu verbringen als Gleichaltrige am Gymnasium. Signifikant mehr Jugendliche an nicht gymnasialen Schulen als an Gymnasien geben die Kategorien 31 bis 60 Minuten und mehr als 60 Minuten digitale Mediennutzung pro Woche an. Zugleich ist der Anteil an Schülerinnen und Schülern, die in einer typischen Schulwoche im Deutschunterricht keine digitalen Lerngelegenheiten nutzen, an nicht gymnasialen Schulen signifikant geringer als am Gymnasium. Die Umsetzung digitaler Lerngelegenheiten scheint an nicht gymnasialen Schulen häufiger zu gelingen, als dies an Gymnasien der Fall ist.

Tabelle 5.4: Zeit, die in einer normalen Schulwoche während des Unterrichts der Landessprache mit der Nutzung digitaler Geräte verbracht wird

OECD-Staaten	Aufgewandte Zeit pro Woche							
	Keine Zeit		1–30 Minuten		31–60 Minuten		Mehr als 60 Minuten	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Estland	50.1	(1.0)	35.1	(0.9)	9.4	(0.5)	5.1	(0.4)
Finnland	33.9	(1.1)	41.5	(0.9)	15.8	(0.6)	8.3	(0.8)
Irland	66.6	(1.3)	19.6	(0.8)	7.9	(0.5)	5.6	(0.6)
Republik Korea	54.4	(1.3)	14.8	(0.7)	11.1	(0.5)	19.4	(1.1)
Polen	59.3	(1.0)	25.7	(0.7)	8.2	(0.4)	5.7	(0.3)
Schweden	11.4	(1.2)	19.3	(1.1)	24.2	(0.8)	44.9	(2.1)
Neuseeland	14.7	(0.8)	23.4	(0.9)	21.1	(0.7)	40.1	(1.4)
Vereinigte Staaten	25.6	(1.6)	32.3	(0.9)	20.5	(0.8)	19.6	(1.2)
Vereinigtes Königreich	66.1	(1.1)	18.9	(0.9)	6.9	(0.4)	7.0	(0.5)
Japan	84.9	(1.2)	8.8	(0.8)	2.4	(0.3)	3.1	(0.5)
Australien	17.5	(0.6)	25.5	(0.5)	23.8	(0.5)	32.7	(0.8)
Dänemark	1.6	(0.2)	7.1	(0.5)	14.1	(0.6)	76.8	(0.9)
Deutschland	65.7	(1.1)	22.2	(0.8)	5.8	(0.4)	6.0	(0.5)
Slowenien	68.4	(0.7)	20.8	(0.6)	5.6	(0.4)	4.8	(0.3)
Belgien	57.6	(0.9)	21.9	(0.6)	9.4	(0.5)	9.2	(0.5)
Frankreich	71.7	(0.9)	16.4	(0.6)	6.3	(0.4)	4.7	(0.4)
Tschechien	62.6	(1.0)	21.0	(0.7)	8.6	(0.5)	6.6	(0.4)
OECD-Durchschnitt	51.4	(0.2)	23.4	(0.1)	11.1	(0.1)	13.2	(0.1)
Schweiz	56.7	(1.4)	25.9	(1.0)	9.2	(0.6)	5.6	(0.6)
Lettland	58.7	(0.8)	27.4	(0.6)	7.2	(0.4)	6.2	(0.3)
Spanien	63.9	(1.0)	16.5	(0.5)	9.8	(0.5)	9.4	(0.6)
Italien	54.7	(1.1)	23.3	(0.6)	10.8	(0.5)	10.1	(0.6)
Ungarn	62.6	(1.2)	24.4	(0.9)	7.1	(0.4)	4.6	(0.3)
Litauen	49.1	(0.9)	32.2	(0.7)	11.4	(0.5)	6.3	(0.3)
Island	33.1	(0.8)	32.4	(0.8)	17.4	(0.7)	16.6	(0.6)
Israel	55.2	(1.4)	24.2	(0.9)	11.8	(0.7)	7.5	(0.8)
Luxemburg	68.1	(0.6)	16.8	(0.5)	6.9	(0.4)	7.3	(0.4)
Türkei	45.7	(1.2)	28.7	(0.9)	13.1	(0.5)	11.6	(0.9)
Slowakei	50.8	(1.0)	29.8	(0.8)	11.7	(0.5)	6.5	(0.4)
Griechenland	70.4	(1.1)	16.4	(0.7)	7.3	(0.4)	4.4	(0.4)
Chile	50.1	(1.2)	30.8	(0.8)	10.9	(0.6)	7.8	(0.5)
Mexiko	62.4	(1.1)	22.7	(0.9)	7.7	(0.4)	5.6	(0.4)

5.2 Extracurriculare Angebote zur Sprach- und Leseförderung

Die Bedeutung der Leseförderung ist insbesondere nach PISA 2000 nicht nur erkannt (Artelt, Stanat, Schneider & Schiefele, 2001), sondern sowohl im Rahmen der sieben zentralen Handlungsfelder der KMK (KMK, 2002) als auch in vielen Projekten umgesetzt worden (vgl. Kapitel 1). Ob und wie diese Maßnahmen die einzelne Schule erreichen und insbesondere welche Lerngelegenheiten im Bereich der Sprach- und Leseförderung bereitgestellt werden (vgl. Bischof, Hochweber, Hartig & Klieme, 2013), wurde in PISA 2018 mithilfe von Fragebögen an die Schulleitungen erfasst. Im Folgenden werden zunächst außerunterrichtliche Zusatzangebote (im Sinne von Wahlkursen, Projekten und speziellen Schulveranstaltungen) für Schülerinnen und Schüler näher betrachtet, bevor die Ausgestaltung zusätzlichen Unterrichts der Landessprache thematisiert wird. Zuletzt werden die schulischen Angebote für Lernende mit einer von der Landessprache abweichenden Herkunftssprache dargestellt, wobei vor dem Hintergrund des zunehmenden Anteils von Jugendlichen mit Zuwanderungshintergrund (vgl. Kapitel 6) auch auf die Entwicklung dieser Angebote seit PISA 2009 eingegangen wird.

5.2.1 Außerunterrichtliche Zusatzangebote zur Sprach- und Leseförderung

Schulen bieten unterschiedliche Möglichkeiten sich außerhalb des Unterrichts vertieft mit der Landessprache auseinanderzusetzen. Abbildung 5.2 gibt einen Überblick über die von den Schulleitungen berichteten Häufigkeiten verschiedener Lerngelegenheiten außerhalb des regulären Unterrichts.

In Deutschland und auch im Mittel der OECD-Staaten bietet etwa die Hälfte der Schulen die Möglichkeit, sich durch die Beteiligung an einem Schultheater oder Musical sowie die Erstellung eines Schuljahrbuchs, einer Schüler- oder Klassenzeitung vertieft mit der Landessprache auseinanderzusetzen. An etwa einem Drittel der Schulen in Deutschland und damit signifikant seltener als im OECD-Durchschnitt wird von Projekten mit örtlichen Bibliotheken berichtet. Deutschland erreicht auch in Bezug auf das Angebot an Buch- oder Leseklubs sowie Debattierklubs signifikant niedrigere Werte als der OECD-Durchschnitt. An weniger als jeder vierten Schule werden nach Angaben der Schulleitungen derartige Lerngelegenheiten angeboten. Deutlich fällt auch der Unterschied zwischen Deutschland und dem Mittel der OECD-Staaten hinsichtlich des Angebots an Vorträgen oder Seminaren etwa von Autorinnen und Autoren oder Journalistinnen und Journalisten aus. In den OECD-Staaten finden derartige Veranstaltungen an etwa drei Viertel der Schulen statt, während dieser Anteil in Deutschland unter 50 Prozent liegt. Signifikant über dem OECD-Durchschnitt liegen Schulen in Deutschland nur bei Projekten mit lokalen Zeitungen. In den lesestarken Staaten Estland, Kanada, Finnland, Irland und der Republik Korea bietet sich ein anderes Bild. Hier finden nach Angaben der Schulleitungen an mehr als drei Viertel der Schulen regelmäßig Vorträge

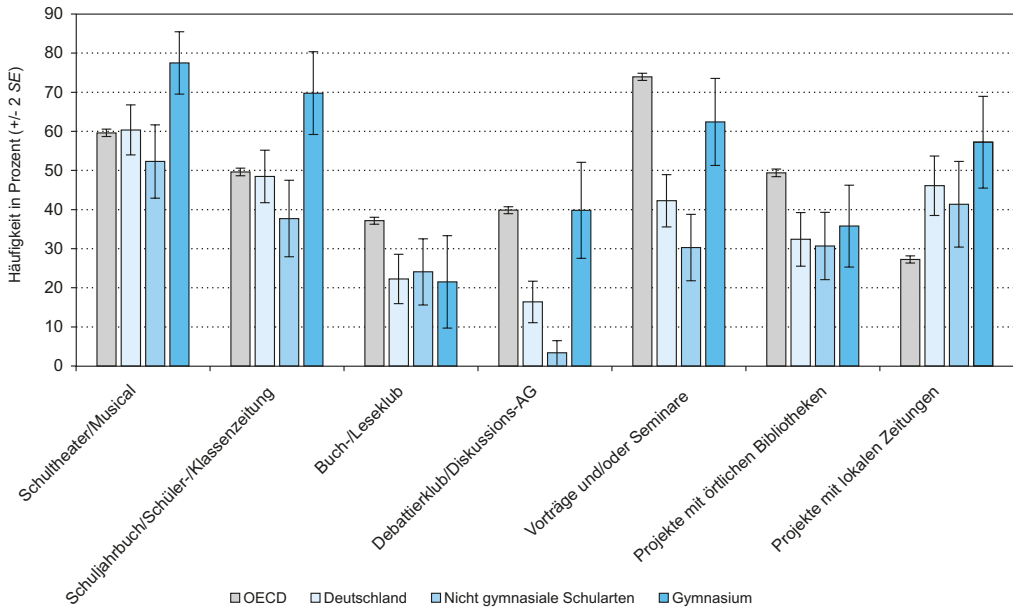


Abbildung 5.2: Zusatzangebote für den Unterricht der Landessprache im OECD-Durchschnitt, in Deutschland insgesamt und nach Schularten

oder Seminare mit Lesebezug statt. Auch hinsichtlich der Zusammenarbeit mit örtlichen Bibliotheken zeigen Staaten wie Estland sowie insbesondere Finnland und Polen deutlich mehr Aktivität als Schulen in Deutschland. Die im Vergleich zu den anderen Zusatzangeboten geringste Verbreitung haben in Deutschland (über alle Schularten hinweg betrachtet) offenbar die in Irland und auch in der Republik Korea weit verbreiteten Debattierklubs. Allerdings handelt es sich hier um das Debattieren in der Art eines sportlichen Wettkampfes, das vor allem im angelsächsischen Raum etabliert ist, in Deutschland jedoch erst langsam an Popularität gewinnt. Insgesamt weist die Befundlage auf ein im internationalen Vergleich eher geringes Angebot an außerunterrichtlichen Aktivitäten zur Sprach- und Leseförderung an Schulen in Deutschland hin.

Die differenzierten Analysen für Deutschland zeigen, dass nicht nur Angebote wie Schultheater oder Musical, das Erstellen eines Schuljahrbuchs, einer Schüler- oder Klassenzeitung sowie Projekte mit Zeitungen, sondern auch die von den lesestarken Staaten genannten Debattierklubs und Experten-Vorträge an Gymnasien signifikant häufiger vorkommen als an nicht gymnasialen Schularten (siehe Abbildung 5.2). Beim Angebot von Buch- oder Leseklubs sowie in der Zusammenarbeit mit örtlichen Bibliotheken zeigen sich keine signifikanten Unterschiede zwischen den Schularten. Gerade den Schülerinnen und Schülern an nicht gymnasialen Schulen, von denen ein relativ hoher Anteil nur über sehr niedrige Lesekompetenzen verfügt (vgl. Kapitel 3), werden damit insgesamt eher wenige außerunterrichtliche Lerngelegenheiten angeboten.

5.2.2 Zusätzlicher Unterricht der Landessprache

In Deutschland berichten 64 Prozent der Schulleitungen, dass sie Zusatzunterricht neben den regulären Deutschstunden anbieten. Damit liegt Deutschland signifikant über dem OECD-Durchschnitt von 46 Prozent. Besonders lesestarke Staaten – Estland, Kanada, Finnland und Irland – scheinen auffällig selten (signifikant seltener als im OECD-Durchschnitt) zusätzlichen Unterricht der Landessprache umzusetzen. Estland, Finnland und Irland weichen auch hinsichtlich der regulären Unterrichtszeit, die in der neunten Klassenstufe für den Unterricht der Landessprache zur Verfügung steht, signifikant nach unten vom OECD-Durchschnitt ab (Estland, Finnland und Irland: < 3.1 Zeitstunden pro Woche; OECD-Durchschnitt: 3.7 Zeitstunden). Auch in Deutschland liegt dieser Wert mit 3.3 Zeitstunden unter dem OECD-Durchschnitt.

Der Zusatzunterricht fokussiert in Deutschland vor allem auf Stütz- und Förderunterricht (59%), gefolgt von einem Angebot für sowohl Leistungsstarke als auch Leistungsschwache (37%). Zusätzliche Unterrichtsstunden ausschließlich zur Begabungsförderung werden in Deutschland offenbar nicht ausgewiesen (0%). Zusatzunterricht findet darüber hinaus äußerst selten ohne eine vorherige Leistungsdifferenzierung statt (3%). Anders als in Deutschland dominieren im Durchschnitt der OECD-Staaten Zusatzangebote, die sowohl die Begabungsförderung als auch Stütz- und Förderunterricht umfassen (52%) – gefolgt von ausschließlichem Stütz- und Förderunterricht (31%). Schulleitungen in Deutschland berichten signifikant niedrigere Werte für die Förderung in beide Richtungen und signifikant höhere Werte für ausschließlichen Stütz- und Förderunterricht. Betrachtet man die Art des Zusatzangebots in den besonders lesestarken OECD-Staaten, erscheint bemerkenswert, dass weder in Estland noch in Finnland, der Republik Korea oder Polen der Schwerpunkt ausschließlich auf Stütz- und Förderunterricht liegt. Vielmehr wird er in allen vier Staaten signifikant seltener als in Deutschland angeboten. Es findet signifikant häufiger eine Förderung in beide Richtungen oder ausschließliche Begabungsförderung statt. Entsprechend scheinen die zusätzlichen schulischen Unterrichtsangebote für die Landessprache in diesen Staaten gezielter gestaltet zu sein.

Bezüglich der Häufigkeit des Angebots von zusätzlichem Deutschunterricht zeigen sich zwischen den Schularten in Deutschland keine Unterschiede, wohl aber sind die Schwerpunkte jeweils etwas anders gelagert. Während der zusätzliche Deutschunterricht an etwa der Hälfte der Gymnasien ausschließlich für die Förderung leistungsschwächer Schülerinnen und Schüler und die andere Hälfte für Förderangebote für Leistungsstarke und -schwache verwendet wird, überwiegt an nicht gymnasialen Schularten ausschließlicher Stütz- und Förderunterricht (65%).

5.2.3 Angebote für Schülerinnen und Schüler mit einer von der Landessprache abweichenden Herkunftssprache

Kinder und Jugendliche mit Defiziten in der Kenntnis der Unterrichtssprache haben wesentliche Schwierigkeiten bei der Partizipation an schulischen Lernangeboten und bleiben in ihren Leistungen oft hinter ihren Klassenkameradinnen und -kameraden zurück (Heinze, Herwartz-Emden, Braun & Reiss, 2011). Auch Deutschland sieht sich mit dieser Herausforderung konfrontiert. Bei PISA 2018 ist der Anteil der Fünfzehnjährigen mit Zuwanderungshintergrund in Deutschland (36 %) im Vergleich zu PISA 2009 (26 %) noch einmal gestiegen (vgl. Kapitel 6). Daher werden im Folgenden Zusatzangebote zur Sprach- und Leseförderung betrachtet, die sich speziell an Schülerinnen und Schüler wenden, deren Herkunftssprache von der Unterrichtssprache abweicht. Die Frage, was sich hier seit PISA 2009 getan hat, erscheint vor dem Hintergrund der aktuellen Entwicklungen von besonderer Relevanz.

Die Schulleitungen gaben an, mit welchen Maßnahmen Jugendliche beim Erwerb der Landessprache unterstützt werden, deren Herkunftssprache nicht die jeweilige Landessprache ist. Als mögliche Unterstützungsmaßnahmen wurden Förderunterricht parallel zum Regelunterricht, Kurse in Vorbereitung auf den Regelunterricht, in der Herkunftssprache durchgeführter Unterricht und eine Reduktion der Klassengröße, um besser auf die Bedürfnisse der jeweiligen Schülerinnen und Schüler eingehen zu können, aufgeführt.

PISA 2018 zeigt, dass etwa zwei Drittel der Schulen in Deutschland zusätzlichen Förderunterricht für Jugendliche mit Sprachdefiziten anbieten, was signifikant über dem OECD-Durchschnitt von 58 Prozent liegt (Tabelle 5.5). Knapp die Hälfte der Schulleitungen in Deutschland (43 %) gibt an, dass diese Jugendlichen einen Vorbereitungskurs vor dem Regelunterricht besuchen. Auch dieser Anteil liegt signifikant über dem Mittelwert der OECD-Staaten (25 %). Unterricht in der von der Unterrichtssprache abweichenden Herkunftssprache wird in den OECD-Staaten genauso wie in Deutschland eher selten realisiert. Bezüglich der Reduktion der Klassengröße zeigt sich, dass an 17 Prozent der Schulen im OECD-Durchschnitt laut Aussage der Schulleitungen die Klassengröße reduziert wird, um den besonderen Bedürfnissen dieser Jugendlichen gerecht zu werden. Dies trifft auf acht Prozent der Schulen in Deutschland zu.

Obwohl der Anteil der Fünfzehnjährigen mit Zuwanderungshintergrund an nicht gymnasialen Schulen höher ist als an Gymnasien (vgl. Kapitel 6), unterscheiden sich die Schularten nicht bezüglich der Häufigkeiten der Maßnahmen zur Förderung von Jugendlichen mit anderer Herkunftssprache.

Tabelle 5.5: Maßnahmen zur Förderung von Schülerinnen und Schülern mit anderer Herkunftssprache in PISA 2009 und PISA 2018

	PISA 2009				PISA 2018			
	Deutschland		OECD-Durchschnitt		Deutschland		OECD-Durchschnitt	
	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Zusätzlicher Förderunterricht in der Landessprache	29.5	(3.2)	57.4	(0.6)	67.1	(3.3)	58.5	(0.5)
Vorbereitungskurs in der Landessprache	11.3	(2.1)	22.7	(0.5)	42.6	(4.0)	25.2	(0.4)
Unterricht in der Herkunftssprache in einigen Fächern	1.2	(0.8)	6.4	(0.3)	3.4	(1.4)	9.3	(0.3)
Unterricht in der Herkunftssprache zu einem beträchtlichen Teil	0.6	(0.6)	5.0	(0.3)	1.4	(1.0)	7.2	(0.3)
Reduktion der Klassengröße	6.2	(1.3)	16.6	(0.4)	7.9	(1.7)	17.0	(0.4)

Beim Vergleich der in PISA 2009 und PISA 2018 angegebenen Maßnahmen, zeigt sich ein signifikanter Zuwachs des zusätzlichen Förderunterrichts parallel zum Regelunterricht sowie der Kurse in Vorbereitung auf den Regelunterricht. Dies sind zugleich die beiden Maßnahmen, die sowohl 2009 als auch 2018 in Deutschland am häufigsten vorkommen (siehe Tabelle 5.5). Zusätzlicher Förderunterricht im Hinblick auf die Landessprache, wie etwa Deutsch als Zweitsprache (DaZ) beziehungsweise Deutsch als Fremdsprache (DaF), hat seit 2009 deutlich zugenommen und wird nach den Ergebnissen von PISA 2018 nun nicht mehr nur an etwa einem Drittel sondern an zwei Drittel der Schu- len angeboten. Auch das Angebot an Deutsch-Vorbereitungskursen scheint sich knapp vervierfacht zu haben.

5.3 Zusammenfassung und Diskussion

Dieses Kapitel beschreibt den aktuellen Stand schulischer Lerngelegenheiten zur Sprach- und Leseförderung im Kontext der Digitalisierung in Deutschland. Es zeigt sich, dass die sachlichen und personellen Voraussetzungen für das Lesen digitaler Texte an Schulen in Deutschland im internationalen Vergleich eher unterdurchschnittlich sind. In Bezug auf die Digitalisierung an deutschen Schulen hat sich insbesondere seit PISA 2015 wenig getan. Auch wenn es sich bei den berichteten Aussagen um Einschätzungen der Schulleitungen beziehungsweise der Schülerinnen und Schüler und nicht um eine tatsächliche Bestandsaufnahme handelt, stellt die Wahrnehmung der unmittelbar Betroffenen einen wichtigen und praxisrelevanten Indikator für den Stand der Digitalisierung im Bildungsbereich dar, der Deutschland im internationalen Vergleich einen Platz im unteren Drittel attestiert. Dieses Ergebnis steht im Einklang mit den Befunden anderer großer Studien, wie etwa dem Monitor Digitale Bildung der Bertelsmann Stiftung (Schmid, Goertz, Behrens & Bertelsmann Stiftung, 2017) sowie ICILS 2013 und ICILS 2018 (*International*

Computer and Information Literacy Study; Bos et al., 2014; Eickelmann, Gerick, Labusch & Vennemann, 2019). Kürzlich beschlossene Maßnahmen sollten hier allerdings Wirkung zeigen. So kann der DigitalPakt Schule insbesondere im Hinblick auf die Ausstattung der Schulen in Deutschland ein wichtiger Schritt sein. Daneben sollten die Aus- und Weiterbildung von Lehrkräften, die Verfügbarkeit technisch-qualifizierten Personals, die Bereitstellung adäquater Software und zeitlicher Ressourcen sowie die Ausarbeitung von konkreten Vorschlägen für den Einsatz digitaler Geräte insbesondere im Deutschunterricht weiter vorangetrieben werden (vgl. Hillmayr et al., 2017). Entscheidend ist, ob digitale Ressourcen auch tatsächlich effektiv im Unterricht verwendet werden (vgl. McFarlane, 2019).

Hinsichtlich extracurricularer Angebote zur Sprach- und Leseförderung zeigen sich deutliche Unterschiede zwischen den Staaten und auch zwischen den Schularten innerhalb Deutschlands. Insbesondere die nicht gymnasialen Schulen bieten ihren Schülerinnen und Schülern nur wenige über den Regelunterricht hinausgehende Angebote zur Sprach- und Leseförderung an. Dabei steht die Bedeutung und das Potenzial außerunterrichtlicher Lerngelegenheiten außer Frage: Je diverser die Lernangebote, umso besser stehen die Chancen für die Entwicklung transferierbarer Kompetenzen (vgl. Barnett & Ceci, 2002; McLellan, 1996). Der Schwerpunkt bei der Ausrichtung zusätzlicher Unterrichtsstunden der Landessprache liegt in Deutschland – anders als in einigen besonders lesestarken OECD-Staaten – auf Stütz- und Förderunterricht. Der in Deutschland im Vergleich zu anderen lesestarken Staaten große Anteil an leseschwachen Jugendlichen (vgl. Kapitel 3) mag darauf hinweisen, dass in Deutschland weitere Bemühungen notwendig sind, um leseschwache Jugendliche noch stärker zu unterstützen – möglicherweise auch durch einen Ausbau extracurricularer Angebote gerade an nicht gymnasialen Schulen. Auch die Begabungsförderung sollte im Kontext des Deutschunterrichts stärker adressiert werden. Ein wichtiger Schritt ist hier sicherlich die gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler aus dem Jahr 2016 (KMK, 2016), die aktuell in Form des BMBF-Projektes „Leistung macht Schule“ (LemaS) umgesetzt wird. Schulen in Deutschland scheinen auf den seit PISA 2009 zunehmenden Anteil an Jugendlichen mit Zuwanderungshintergrund mit einem verstärkten Angebot an zusätzlichem Unterricht (z. B. DaZ), der auf eine Verbesserung ihrer Deutschkenntnisse abzielt, zu reagieren. Ebenso werden mittlerweile an fast der Hälfte der Schulen in Deutschland spezielle Kurse angeboten, die diese Schülerinnen und Schüler auf den Regelunterricht vorbereiten sollen. Angesichts der demographischen Entwicklung sind diese Angebote notwendig, sinnvoll und sicherlich ausbaufähig.

Literatur

- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, ... M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–137). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_4
- Barnett, S. M. & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://dx.doi.org/10.1037/0033-2909.128.4.612>
- Bischof, L. M., Hochweber, J., Hartig, J. & Klieme, E. (2013). Schulentwicklung im Verlauf eines Jahrzehnts: Erste Ergebnisse des PISA-Schulpanels. In N. Jude & E. Klieme (Hrsg.), *Zeitschrift für Pädagogik Beiheft, vol 59. PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (S. 172–199). Weinheim: Beltz Juventa.
- Bos, W., Eickelmann, B., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., ... Wendt, H. (2014). *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Eickelmann, B., Gerick, J., Labusch, A. & Vennemann, M. (2019). Schulische Voraussetzungen als Lern- und Lehrbedingungen in den ICILS-2018-Teilnehmerländern. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, ... J. Vahrenhold (Hrsg.), *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking* (S. 141–174). Münster: Waxmann.
- Gorski, P. C. (2009). Insisting on digital equity: Reframing the dominant discourse on multicultural education and technology. *Urban Education*, 44(3), 348–364. <https://doi.org/10.1177/0042085908318712>
- Hahnel, C., Goldhammer, F., Naumann, J. & Kröhne, U. (2016). Effects of linear reading, basic computer skills, evaluating online information, and navigation on reading digital text. *Computers in Human Behavior*, 55, 486–500. <https://doi.org/10.1016/j.chb.2015.09.042>
- Heinze, A., Herwartz-Emden, L., Braun, C. & Reiss, K. (2011). Die Rolle von Kenntnissen der Unterrichtssprache beim Mathematiklernen. Ergebnisse einer quantitativen Längsschnittstudie in der Grundschule. In S. Prediger & E. Özdil (Hrsg.), *Mathematiklernen unter Bedingungen der Mehrsprachigkeit. Stand und Perspektiven der Forschung und Entwicklung in Deutschland* (S. 11–34). Münster: Waxmann.
- Helmke, A. & Schrader, F.-W. (2014). Angebots-Nutzungs-Modell. In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (17. akt. u. erw. Neuaufl., S. 149–150). Bern: Huber.
- Hillmayr, D., Reinhold, F., Ziernwald, L. & Reiss, K. (2017). *Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe. Einsatzmöglichkeiten, Umsetzung und Wirksamkeit*. Münster: Waxmann.
- Klieme, E. (2013). The role of large-scale assessments in research on educational effectiveness and school development. In M. von Davier, E. Gonzalez, I. Kirsch & K. Yamamoto (Hrsg.), *The role of international large-scale assessments: perspectives from technology, economy, and educational research* (S. 115–148). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-4629-9_7
- KMK (2002) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2002). *PISA 2000 – Zentrale Handlungsfelder. Zusammenfassung*. Münster: Waxmann.

- menfassende Darstellung der laufenden und geplanten Maßnahmen in den Ländern. Beschluss der 299. Kultusministerkonferenz vom 17./18.10.2002. Verfügbar unter https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf
- KMK (2016) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2016). *Gemeinsame Initiative von Bund und Ländern zur Förderung leistungsstarker und potenziell besonders leistungsfähiger Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 10.11.2016*. Verfügbar unter https://www.bmbf.de/files/Initiative_Leistungsstarke_Beschluss.pdf
- Linberg, T., Struck, O. & Bäumer, T. (2018). Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft*, 21(6), 1205–1227. <https://doi.org/10.1007/s11618-018-0830-2>
- McFarlane, A. (2019). *Growing up digital: What do we really need to know about educating the digital generation?* Nuffield Foundation.
- McLellan, H. (Hrsg.). (1996). *Situated learning perspectives*. Englewood Cliffs, NJ: Educational Technology Publications.
- Moran, J., Ferdig, R. E., Pearson, P. D., Wardrop, J. & Blomeyer, R. L. (2008). Technology and reading performance in the middle-school grades: A meta-analysis with recommendations for policy and practice. *Journal of Literacy Research*, 40(1), 6–58. <https://doi.org/10.1080/10862960802070483>
- OECD. (2018). *The Future of education and skills: Education 2030. The future we want*. Paris: OECD Publishing.
- Reynolds, D. (2014). *Advances in school effectiveness research and practice*. Oxford, England: Pergamon.
- Scheerens, J. (2004). *Review of school and instructional effectiveness research. Paper commissioned for the EFA Global Monitoring Report 2005, The Quality Imperative*. Paris: Unesco.
- Schmid, U., Goertz, L., Behrens, J. & Bertelsmann Stiftung (2017). *Monitor Digitale Bildung: Die Schulen im digitalen Zeitalter*. <https://doi.org/10.11586/2017041>
- Shanahan, T. & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40–59. <https://doi.org/10.17763/haer.78.1.v62444321p602101>

6 Soziale Herkunft, Zuwanderungshintergrund und Lesekompetenz

Mirjam Weis, Katharina Müller, Julia Mang, Jörg-Henrik Heine,
Nicole Mahler & Kristina Reiss

Die Daten von PISA 2018 zeigen, dass der Zusammenhang zwischen sozialer Herkunft (z. B. sozioökonomischer beruflicher Status) und Lesekompetenz in Deutschland im Vergleich zum OECD-Mittel überdurchschnittlich ausgeprägt ist. Zudem ist im Vergleich zu anderen europäischen Staaten in Deutschland der Unterschied in der Lesekompetenz zwischen Jugendlichen ohne Zuwanderungshintergrund und Jugendlichen mit Zuwanderungshintergrund relativ groß und der Zuwanderungshintergrund besonders stark mit dem sozioökonomischen beruflichen Status der Eltern assoziiert. Die Indikatoren der sozialen Herkunft tragen damit deutlich zur Erklärung von Disparitäten in der Lesekompetenz zwischen Jugendlichen mit und ohne Zuwanderungshintergrund bei. Bei PISA 2018 hat sich die mittlere Lesekompetenz der ersten Generation in Deutschland im Vergleich zu PISA 2009 signifikant verschlechtert, während sich die Lesekompetenz der zweiten Generation signifikant verbessert hat. Der Anteil der besonders leseschwachen Jugendlichen der ersten Generation ist sehr hoch und seit 2009 deutlich angestiegen. Damit verfügen über die Hälfte der Jugendlichen der ersten Generation über nur eingeschränkte Lesekompetenzen. Insgesamt zeigen die Ergebnisse von PISA 2018, dass in Deutschland weiterhin Handlungsbedarf besteht, um die Leistungsdisparitäten, die bezüglich zuwanderungsbezogener und sozialer Herkunft bestehen, zu verringern.

6.1 Soziale und zuwanderungsbezogene Herkunft

6.1.1 Die Bedeutung der sozialen und zuwanderungsbezogenen Herkunft für die Lesekompetenz

Die soziale Herkunft ist wahrscheinlich die Kontextvariable, die am häufigsten in der Bildungsforschung untersucht wird (Sirin, 2005). Inzwischen verweisen zahlreiche Studien auf den Zusammenhang zwischen Merkmalen der sozialen Herkunft und den Schulleistungen allgemein sowie domänenspezifischen Schulleistungen wie Lesekompetenz (z. B. Aikens & Barbarin, 2008; Davis-Kean, 2005; Magnuson, 2007; Sirin, 2005).

Aufgrund der hohen Relevanz der sozialen Herkunft für die Schulleistungen spielt sie auch in der PISA-Studie eine wichtige Rolle.

In Deutschland kam es aufgrund der Ergebnisse der PISA-Studie 2000, welche einen im internationalen Vergleich sehr hohen Zusammenhang zwischen der sozialen Herkunft und der Lesekompetenz von Jugendlichen in Deutschland zeigten (Baumert & Schümer, 2001), sowohl öffentlich als auch innerhalb der Bildungsforschung zu vielen Diskursen. Während bei PISA 2000 der Zusammenhang zwischen der sozialen Herkunft und der Lesekompetenz der Jugendlichen in nur wenigen OECD-Staaten so hoch war wie in Deutschland (Baumert & Schümer, 2001), war dieser Zusammenhang bei PISA 2009, als die Lesekompetenz das nächste Mal die Hauptdomäne darstellte, bedeutsam geringer und unterschied sich nicht mehr signifikant vom OECD-Durchschnitt (Ehmke & Jude, 2010).

Zudem zeigte sich in den vergangenen PISA-Studien, dass Familien mit Zuwanderungshintergrund in Deutschland häufig über einen geringeren sozialen Status verfügen. Es wurde deutlich, dass diese sozialen Disparitäten in zugewanderten Familien in Deutschland höher sind als in anderen OECD-Staaten (z.B. im Vergleich mit dem Vereinigten Königreich bei PISA 2015; vgl. Rauch, Mang, Härtig & Haag, 2016; Weis, Mang, Baumann & Reiss, 2018). In diesem Zusammenhang ist relevant, ob die Jugendlichen selbst zugewandert sind oder, ob sie in Deutschland geboren wurden, aber ihre Eltern zugewandert sind (*Zuwanderungsstatus*). Hinweise dafür, dass Schülerinnen und Schüler mit einem Zuwanderungshintergrund deutlich niedrigere Kompetenzen im Lesen, in Mathematik und in den Naturwissenschaften in Deutschland aufweisen, bestanden auch schon vor der ersten PISA-Studie (Alba, Handl & Müller, 1994; Esser 1990; Nauck, Diefenbach & Petri, 1998). Mit den PISA-Daten konnten diese Befunde bestätigt und im internationalen Vergleich betrachtet werden. Obwohl die soziale Herkunft eine große Rolle bei der Erklärung der Leistungsdisparitäten von Jugendlichen mit Zuwanderungshintergrund spielt, kann sie diese aber dennoch nicht vollständig erklären (z.B. Gebhardt, Rauch, Mang, Sälzer & Stanat, 2013; Kiemer, Haag, Müller & Ehmke, 2017). Daher wird im vorliegenden Kapitel sowohl die Rolle verschiedener Indikatoren zur sozialen Herkunft als auch die Rolle von Zuwanderung im Zusammenhang mit der Lesekompetenz betrachtet.

Die Ergebnisse von PISA 2000 und PISA 2009 zeigten, dass in Deutschland neben dem Zuwanderungsstatus und Indikatoren zur sozialen Herkunft auch die *zu Hause gesprochene Sprache* eine Rolle spielt (Baumert & Schümer, 2001; Stanat, Rauch & Segritz, 2010). Es wird angenommen, dass sich geringere Sprachkenntnisse in der Instruktionssprache von Jugendlichen mit Zuwanderungshintergrund negativ auf deren Lesekompetenzen auswirken (Baumert & Schümer, 2001; Kiemer et al., 2017). Bei PISA 2000 zeigte sich eine niedrigere Lesekompetenz von Fünfzehnjährigen mit Zuwanderungshintergrund, die hauptsächlich ihre Herkunftssprache mit ihrer Familie sprachen, im Vergleich zu Fünfzehnjährigen, die hauptsächlich Deutsch mit ihrer Familie sprachen. Der Einfluss der zu Hause gesprochenen Sprache war bei der PISA-Studie 2009 zwar immer noch bedeutsam, aber deutlich geringer als im Jahre 2000.

Diese PISA-Befunde leisten einen wichtigen Beitrag für Bildungsforschung, Politik und Gesellschaft, indem soziale Disparitäten in den Schulleistungen und mit dem Zuwanderungshintergrund assoziierte Disparitäten im internationalen Vergleich sowie über die Zeit hinweg in Deutschland abgebildet werden können. Insgesamt machen die bisherigen Befunde deutlich, dass eine gemeinsame Betrachtung der sozialen und zuwanderungsbezogenen Disparitäten sinnvoll ist (Kierner et al., 2017). Im vorliegenden Kapitel wird zunächst untersucht, wie hoch der Zusammenhang zwischen der sozialen Herkunft der Jugendlichen und ihrer Lesekompetenz in Deutschland im Jahr 2018 im Vergleich zu den OECD-Staaten ist. Zudem werden differenzierte Ergebnisse zu spezifischen Gruppen von Schülerinnen und Schülern in Deutschland berichtet. So werden zusätzlich soziale Klassen betrachtet (definiert durch die Berufe der Eltern), deren Merkmale der sozialen Herkunft spezifisch für Gruppen leseschwacher und lesestarker Schülerinnen und Schüler beschrieben werden. In einem zweiten Teil des Kapitels liegt der Fokus auf Jugendlichen mit Zuwanderungshintergrund. Hier wird zunächst auf den sozioökonomischen beruflichen Status der Eltern von Jugendlichen mit und ohne Zuwanderungshintergrund in ausgewählten europäischen Staaten sowie auf die Lesekompetenz der Jugendlichen nach Zuwanderungsstatus eingegangen. Anschließend werden die Veränderungen in Deutschland zwischen den Jahren 2009 und 2018 beschrieben. Neben der Veränderung der Anteile von Jugendlichen mit Zuwanderungshintergrund wird auch auf Merkmale der sozialen Herkunft der Familien sowie auf die zu Hause gesprochene Sprache eingegangen. Zudem wird die Verteilung des Zuwanderungshintergrunds in verschiedenen Schularten in Deutschland berichtet. Am Schluss des Kapitels werden soziale und zuwanderungsbezogene Disparitäten verbunden und untersucht, in welchem Maße die Kompetenzunterschiede im Lesen zwischen Fünfzehnjährigen mit und ohne Zuwanderungshintergrund durch Unterschiede in der sozialen Herkunft erklärt werden können.

6.1.2 Erfassung der sozialen Herkunft bei PISA 2018

Berufe der Eltern

Ein wichtiger Aspekt der sozialen Herkunft sind die *Berufe der Eltern*, da diese sowohl Informationen über das Einkommen als auch über Prestige, Bildungsniveau und den Besitz von Kulturgütern der Eltern geben können (Baumert & Maaz, 2006; Ehmke & Jude, 2010; Sirin, 2005). In der PISA-Studie werden die Jugendlichen im Fragebogen daher nach den Berufen der Eltern befragt. Die Antworten der Schülerinnen und Schüler werden nach der internationalen Standardklassifikation der Berufe (International Standard Classification of Occupations, ISCO-08; ILO, 2012) kodiert und anhand des internationalen sozioökonomischen Index des beruflichen Status (International Socio-Economic Index of Occupational Status, ISEI; Ganzeboom & Treimann, 2012) in eine Rangreihe gebracht, wobei die Werte von 11 bis 90 Punkten reichen und höhere

Werte einen höheren sozioökonomischen beruflichen Status bedeuten. Da es sich beim ISEI um den Status der beruflichen Tätigkeit unter Berücksichtigung des Bildungsniveaus und des Einkommens handelt, kann er als Maß für den sozioökonomischen Status verstanden werden (Ganzeboom, de Graaf & Treiman, 1992). Der höhere ISEI-Wert beider Elternteile (Highest International Socio-Economic Index of Occupational Status, HISEI) liefert einen Index für den höchsten *sozioökonomischen beruflichen Status der Eltern*¹. Aufgrund der Umstellung des internationalen Klassifikationssystems von ISCO-88 zu ISCO-08 ist seit PISA 2012 und so auch bei PISA 2018 die Vergleichbarkeit des sozioökonomischen beruflichen Status (HISEI) mit vorherigen PISA-Erhebungsrounden (2000 bis 2009) eingeschränkt.

Zusätzlich wird im vorliegenden Kapitel in den nationalen Analysen ein kategoriales Differenzierungsschema eingesetzt, welches Berufe in soziale Schichten einteilt. Es beruht auf der Annahme, dass die Lebensbedingungen in Abhängigkeit des Einkommens, der Bildung, der gesellschaftlichen Anerkennung und der beruflichen Gestaltungsmöglichkeiten variieren (Hradil, 2005), und kann damit Unterschiede im sozialen Entwicklungsumfeld der Schülerinnen und Schüler sichtbar machen. Bei diesem sogenannten *Erikson-Goldthorpe-Portocarero-(Klassen)-Schema (EGP-Klassenschema*; Erikson, Goldthorpe & Portocarero, 1979; Erikson & Goldthorpe, 2002) werden Berufe, die sich vergleichsweise ähnlich sind, in diskrete Klassen zusammengefasst. Grundlage des EGP-Klassifikationsschemas ist ein Kategoriensystem, bei dem Berufe nach der Art der Tätigkeit, der Stellung im Beruf sowie der Weisungsbefugnis geordnet werden. Anhand dieser soziologisch fundierten Typologie können Berufsgruppen anschaulich beschrieben werden. Verwendet wird das Klassifikationsschema mit den sechs Klassen *obere Dienstklasse (I), untere Dienstklasse (II), Routinedienstleistungen Handel und Verwaltung (III), Selbstständige (IV), Facharbeiter und Arbeiter mit Leitungsfunktion (V, VI) sowie un- und angelernte Arbeiter, Landarbeiter (VII)* (Erikson et al., 1979; siehe Tabelle 6.1web für beispielhafte Beschreibungen der EGP-Klassen).

Für die hier vorgenommene operationale Einordnung in die EGP-Klassen sind die Angaben zum Beruf der Eltern, die Art des Beschäftigungsverhältnisses sowie die Weisungsbefugnis maßgeblich. Die Informationen dazu basieren auf den Angaben der Eltern (Elternfragebogen) sowie der Schülerinnen und Schüler (Schülerfragebogen) und beziehen sich auf die Bezugsperson². Fehlende Werte wurden mit der aktuellen Version des Pakets *mice* (vgl. van Buuren & Groothuis-Oudshoorn, 2011) für die freie Statistikumgebung *R* (R Core Team, 2019) nach dem Prinzip der multiplen Imputation geschätzt. Die resultierenden zehn Schätzungen für jeden fehlenden Wert basieren auf einem Imputationsmodell, das neben den Kompetenzschätzern für die drei PISA-Domä-

1 In vorherigen nationalen PISA-Berichten wurde der HISEI als *sozioökonomischer Status* bezeichnet (vgl. z. B. Müller & Ehmke, 2016). Um Verwechslungen mit den OECD-Berichten zu vermeiden, welche den ESCS als *sozioökonomischen Status* berichten (z. B. OECD, 2016), wird im vorliegenden Kapitel der HISEI als *sozioökonomischer beruflicher Status* bezeichnet. Der ESCS wird in diesem Kapitel als *sozioökonomischer und soziokultureller Status* bezeichnet.

2 Die Angaben zur Bezugsperson beziehen sich auf den Vater beziehungsweise wenn diese Angaben fehlen, auf die Mutter.

nen unterschiedliche Variablen (z. B. Bildungsdauer der Eltern) beinhaltet, welche mit der EGP-Klassifikation assoziiert sind.³ Die imputierten Werte wurden ausschließlich für die Analysen zu den EGP-Klassen verwendet.

Sozioökonomischer und soziokultureller Status

Weitere wichtige Aspekte der sozialen Herkunft sind das Bildungsniveau der Eltern und der Besitz von Wohlstandsgütern (vgl. z. B. Sirin, 2005; OECD, 2019). Daher wird in der PISA-Studie mit dem *sozioökonomischen und soziokulturellen Status* (Index of Economic, Social, and Cultural Status, ESCS) ein Index verwendet, welcher sich aus dem *sozioökonomischen beruflichen Status* (HISEI), der *Bildungsdauer der Eltern* (PARED) und dem *Besitz von Wohlstandsgütern* (HOMEPOS) zusammensetzt. Der *sozioökonomische und soziokulturelle Status* (ESCS) wird über alle OECD-Staaten hinweg standardisiert und weist einen Mittelwert von 0 und eine Standardabweichung von 1 für den OECD-Durchschnitt auf (OECD, 2019).

Das Bildungsniveau der Eltern wird bei PISA durch die höchste *Bildungsdauer der Eltern* (PARED) erfasst. Hierzu werden die Fünfzehnjährigen im Fragebogen nach den höchsten Bildungsabschlüssen ihrer Eltern befragt. Diese werden anschließend nach der internationalen Standardklassifikation des Bildungswesens (International Standard Classification of Education, ISCED-97; OECD, 1999) eingeordnet. Aus diesem ISCED-Level wird die höchste Bildungsdauer in Jahren der Eltern abgeleitet. Der internationale Index des Bildungsniveaus hat sich seit PISA 2000 nicht verändert.

Zur Erfassung der (*Kultur- und*) *Wohlstandsgüter eines Jugendlichen* (HOMEPOS) werden sowohl klassisch-kulturelle als auch lernbezogene und computerbezogene Besitztümer in der Familie einbezogen, nach deren Verfügbarkeit im familiären Umfeld die Jugendlichen ebenfalls im Fragebogen befragt werden („Wie viele/Welche der folgenden Dinge gibt es bei dir zu Hause?“). Bei PISA 2015 gab es Änderungen im Vergleich zu 2012 in diesem Index, beispielsweise kamen Fragen zu weiteren Besitztümern wie Tablet-Computer oder E-Book-Reader hinzu. Bei PISA 2018 wurden die Fragen zu den (*Kultur- und*) *Wohlstandsgütern* aus PISA 2015 weitestgehend übernommen (vgl. Mang, Ustjanzew, Lefke, Schiepe-Tiska & Reiss, 2019); es ergaben sich somit nur geringfügige Änderungen im Index der (*Kultur- und*) *Wohlstandsgüter* im Vergleich zu PISA 2015.

Da der *sozioökonomische und soziokulturelle Status* (ESCS) sowohl den sozioökonomischen beruflichen Status als auch das Bildungsniveau und den Besitz von Wohlstandsgütern beinhaltet, können damit Unterschiede in den Kompetenzen relativ umfassend mit Aspekten der sozialen Herkunft in Verbindung gebracht werden (Müller & Ehmke,

3 Das Imputationsmodell umfasst neben den Kompetenzschätzern für die drei PISA-Domänen und dem Geschlecht der Jugendlichen sechs Variablen zur gesprochenen Sprache im Umfeld der Familie, Schule und Freunde sowie die Variablen „ESCS“, „HISEI“, „PARED“, „HOMEPOS“, „CULT-POSS“, „HEDRES“, „WEALTH“, „JOYREAD“, „SCHULART“, „IMMIG“. Zu den Inhalten dieser Variablen (soweit sie nicht hier im Kapitel beschrieben sind) wird auf die internationalen Codebücher verwiesen, die verfügbar sind unter: <http://www.oecd.org/pisa/data/2018database/>.

2016). Im Vergleich zu den einzelnen Prädiktoren der sozialen Herkunft (HISEI, PARED oder HOMEPOS) ist die Varianzaufklärung des ESCS mit Blick auf die in PISA getesteten Kompetenzen höher (vgl. Ehmke & Siegle, 2005). Allerdings erlauben Analysen mit dem ESCS keine differentiellen Aussagen über die Wirkung einzelner Faktoren der sozialen Herkunft (Caro & Cortés, 2012; Müller & Ehmke, 2016). So zeigte die Metaanalyse von Sirin (2005), dass der Einbezug verschiedener Komponenten der sozialen Herkunft zu unterschiedlichen Ergebnissen im Zusammenhang mit schulischen Kompetenzen führt.

Daher wird in dem vorliegenden Bericht zusätzlich zu dem übergreifenden Faktor *sozioökonomischer und soziokultureller Status* (ESCS) der *sozioökonomische berufliche Status* (HISEI und EGP-Klassen) als separater Faktor betrachtet, um differenzierte Aussagen zu ermöglichen. Zudem werden für Jugendliche aus zugewanderten Familien die einzelnen Aspekte der sozialen Herkunft *sozioökonomischer beruflicher Status* (HISEI), *Bildungsdauer der Eltern* (PARED) und *Besitz von Wohlstandsgütern* (HOMEPOS) getrennt betrachtet.

6.1.3 Erfassung des Zuwanderungshintergrunds bei PISA 2018

Definition des Zuwanderungshintergrunds

Zur Definition des Zuwanderungsstatus wird in der PISA-Studie 2018 das Geburtsland der Jugendlichen und ihrer Eltern herangezogen (analog zu den Erhebungsrounden seit 2006; vgl. Rauch et al., 2016; Stanat & Christensen, 2006; Stanat et al., 2010). Die Informationen dazu basieren auf den Angaben der Schülerinnen und Schüler (Schülerfragebogen). Dabei wurden folgende Kategorien unterschieden:

Ohne Zuwanderungshintergrund:	Kein Elternteil im Ausland geboren.
Ein Elternteil im Ausland geboren:	Ein Elternteil im Ausland, ein Elternteil in Deutschland (bzw. im jeweiligen OECD-Teilnehmerstaat) geboren.
Zweite Generation:	Beide Elternteile im Ausland geboren, Jugendlicher in Deutschland (bzw. im jeweiligen OECD-Teilnehmerstaat) geboren.
Erste Generation:	Beide Elternteile und Jugendlicher im Ausland geboren.

Jugendliche mit nur einem in Deutschland geborenen Elternteil werden in diesem Bericht separat betrachtet. Diese Differenzierung entspricht den nationalen Berichten seit PISA 2003 (Ramm, Prenzel, Heidemeier & Walter, 2004; Rauch et al., 2016), unterscheidet sich aber vom internationalen Vorgehen der OECD, bei welchem diese Jugendliche den Schülerinnen und Schülern ohne Zuwanderungshintergrund zugeordnet werden.

Umgang mit fehlenden Werten

Da für einige Jugendliche aufgrund fehlender Werte keine eindeutige Zuordnung des Zuwanderungsstatus möglich ist, werden die Kategorien für den Zuwanderungsstatus im vorliegenden Bericht um die Kategorie „nicht zuzuordnen“ erweitert. Dies entspricht früheren nationalen Berichten (vgl. Rauch et al., 2016, Gebhardt et al., 2013; Stanat et al., 2010), unterscheidet sich aber vom Vorgehen der OECD. Da die OECD diese Schülerinnen und Schüler aus den Analysen ausschließt beziehungsweise mit der Information nur eines Elternteiles den Zuwanderungskategorien zuordnet, wird die Vergleichbarkeit mit den Analysen der OECD erschwert (OECD, 2019).

Differenzierung verschiedener Herkunftsgruppen

In früheren nationalen PISA-Berichten (2000 bis 2012) wurden Gruppen von Jugendlichen aus den häufigsten Herkunftsländern separat betrachtet, da mit dem Herkunftsland Leistungsdisparitäten teilweise erklärt werden konnten (vgl. z. B. Stanat et al., 2010). Differenziert man die in PISA 2018 getesteten Jugendlichen mit Zuwanderungshintergrund in Deutschland entsprechend, können folgende Herkunftsländer beziehungsweise Kategorien unterschieden werden:

Ehemalige UdSSR:	Mindestens ein Elternteil in Russland, Kasachstan oder einer anderen ehemaligen Sowjetrepublik geboren.
Türkei:	Mindestens ein Elternteil in der Türkei geboren.
Polen:	Mindestens ein Elternteil in Polen geboren.
Anderes Land:	Zusammengefasste Kategorie mehrerer Staaten, aus denen jeweils ein geringer Anteil an Jugendlichen stammt.

Die Kategorie „Anderes Land“ umfasst alle Jugendlichen, deren Eltern nicht aus einem der Herkunftsländer ehemalige UdSSR, Türkei und Polen stammen. Die meisten dieser Kategorie zugeordneten Familien bei PISA 2018 sind aus Syrien, dem Kosovo und Rumänien zugewandert. Diese Gruppen sind jedoch für separate Betrachtungen zu klein.

Jugendliche, deren Eltern in unterschiedlichen Ländern geboren wurden, wurden aus den Analysen ausgeschlossen. Jugendliche mit einem im Ausland und einem in Deutschland geborenen Elternteil wurden dem Herkunftsland des zugewanderten Elternteils zugeordnet.

Wie in PISA 2015 (vgl. Rauch et al., 2016) wird auch hier für die Herkunftsländer ehemalige UdSSR, Türkei und Polen auf Analysen verzichtet, die Zuwanderungsstatus und Herkunftsland zugleich berücksichtigen. Der Grund ist die selbst für diese Länder mit hoher Zuwanderungsquote zu geringe Anzahl von Jugendlichen in der Stichprobe.

Jugendliche mit Zuwanderungshintergrund, die weniger als ein Jahr in der Testsprache unterrichtet wurden, waren von der Erhebung ausgeschlossen (vgl. Kapitel 1).

Die zu Hause gesprochene Sprache

Die *zu Hause gesprochene Sprache* wird von den Jugendlichen im Fragebogen erfragt („Welche Sprache sprichst du am häufigsten zu Hause?“). Hierbei standen die zehn Sprachen zur Auswahl, die in früheren PISA-Erhebungsrounden am häufigsten genannt wurden. Zusätzlich wurde ein offenes Antwortfeld angeboten, in das weitere Sprachen eingetragen werden konnten.

6.2 Soziale Herkunft und Lesekompetenz im internationalen Vergleich

6.2.1 Zusammenhang zwischen dem sozioökonomischen beruflichen Status und der Lesekompetenz im internationalen Vergleich

Mittelwerte und Streuung des sozioökonomischen beruflichen Status im internationalen Vergleich

Der mittlere sozioökonomische berufliche Status (HISEI) liegt in Deutschland mit 51.8 Punkten in der Nähe des OECD-Durchschnitts⁴ von 52.4 Punkten (siehe Abbildung 6.1). In insgesamt neun Staaten unterscheidet sich der mittlere sozioökonomische berufliche Status nicht vom OECD-Mittelwert. Neben Deutschland zählen zu diesen Staaten Litauen, Irland, die Schweiz, die Republik Korea, Belgien, Slowenien, Japan und Frankreich. Besonders hoch und signifikant über dem OECD-Mittelwert liegt der mittlere sozioökonomische berufliche Status in den skandinavischen Staaten (Norwegen, Island, Dänemark, Schweden und Finnland) sowie in Israel. Der mittlere HISEI Norwegens stellt mit 61.6 Punkten den höchsten der OECD-Staaten dar. Ebenfalls sehr hohe Mittelwerte finden sich im Vereinigten Königreich, in Kanada, Neuseeland und Australien. In 13 OECD-Staaten liegt der mittlere sozioökonomische berufliche Status signifikant unter dem OECD-Durchschnitt, beispielsweise in Tschechien, in der Slowakei, in Chile, Kolumbien und Mexiko. Der niedrigste Wert der OECD-Staaten im sozioökonomischen beruflichen Status zeigt sich in der Türkei mit 37.7 Punkten.

Im Vergleich zu PISA 2015 hat sich der mittlere sozioökonomische berufliche Status in Deutschland im Jahr 2018 nicht bedeutsam verändert. Der OECD-Mittelwert ist hingegen seit 2015 signifikant angestiegen (von 51.8 auf 52.4 Punkte). Besonders große Anstiege im mittleren sozioökonomischen beruflichen Status finden sich beispielsweise in Dänemark (von 56.2 auf 60.0 Punkte), Luxemburg (von 48.8 auf 51.4 Punkte) und Polen (von 45.7 auf 49.8 Punkte). Eine signifikante Verringerung im sozioökonomischen

4 Im vorliegenden Kapitel werden alle Ergebnisse im Zusammenhang mit der Lesekompetenz von lediglich 36 OECD-Staaten berichtet, da vom OECD-Staat Spanien aus methodischen Gründen keine Ergebnisse zur Lesekompetenz berichtet werden.

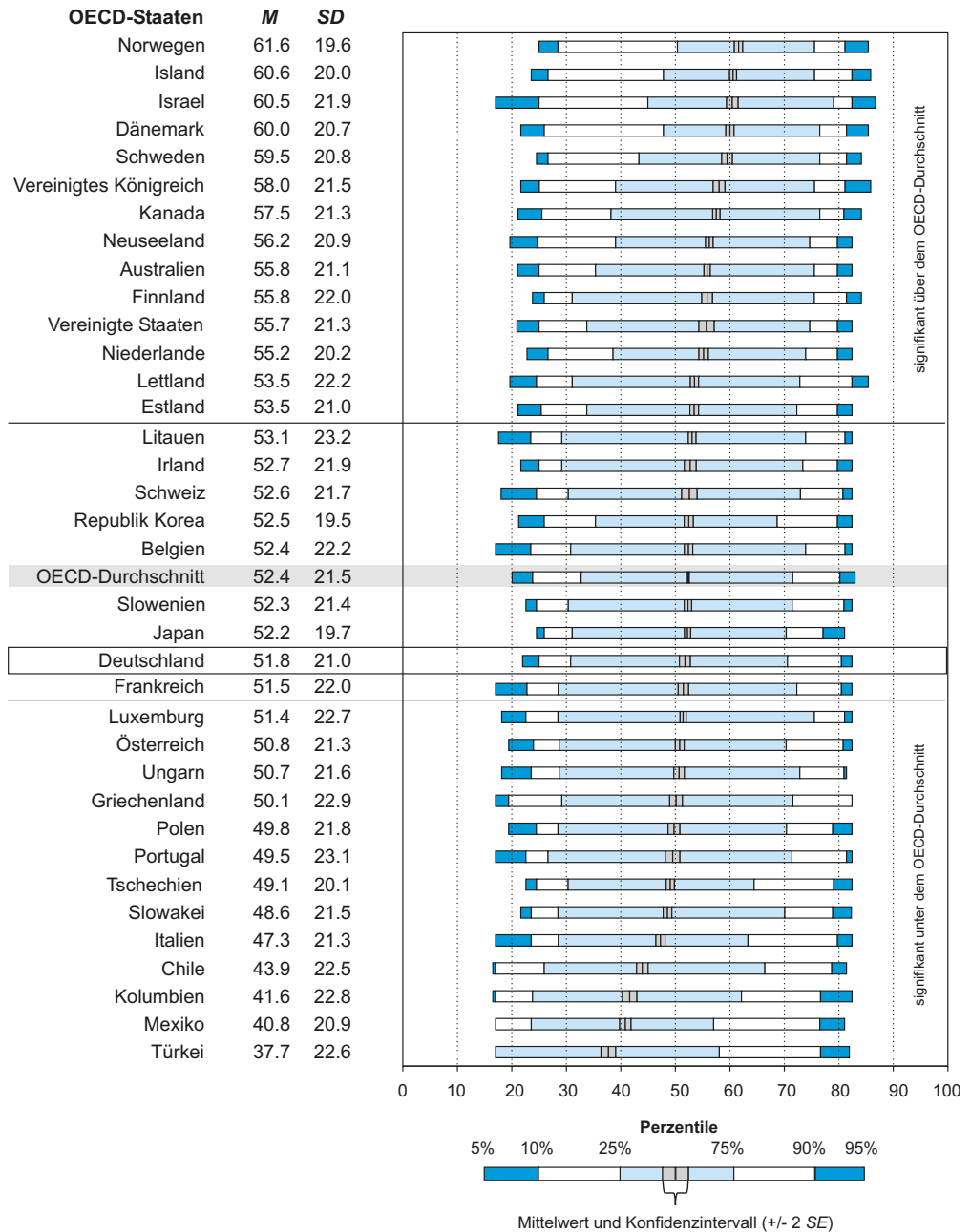


Abbildung 6.1: Mittelwerte, Streuungen und Perzentilbänder des sozioökonomischen beruflichen Status (HISEI) der OECD-Staaten

beruflichen Status zeigte sich lediglich in den drei Staaten Australien (von 57.5 zu 55.8), Italien (von 49.7 zu 47.3) und Neuseeland (von 57.4 zu 56.2).

Neben dem mittleren sozioökonomischen beruflichen Status ist seine Standardabweichung als Maß für die Streuung von besonderem Interesse. Diese sagt aus, wie groß die Unterschiede im sozioökonomischen beruflichen Status innerhalb der jeweiligen Staaten sind. In Deutschland liegt die Streuung bei PISA 2018 signifikant unter dem OECD-Durchschnitt, wie es auch bei PISA 2015 der Fall war. Noch geringere Unterschiede im sozioökonomischen beruflichen Status zeigen sich in Norwegen, Island, den Niederlanden, der Republik Korea, Japan und Tschechien. Besonders hohe Streuungen, welche signifikant über dem OECD-Durchschnitt liegen, zeigen sich beispielsweise in Litauen, Luxemburg, Griechenland, Portugal und Kolumbien.

Sowohl in Deutschland als auch im OECD-Durchschnitt hat sich die Streuung zwischen den Jahren 2015 und 2018 signifikant vergrößert. In Deutschland stieg die Streuung von 20.4 auf 21.0 Punkte und im OECD-Mittel von 21.1 auf 21.5 Punkte. Besonders deutlich vergrößert hat sich die Streuung im Vergleich zu 2015 etwa in Kanada, Neuseeland und in der Türkei. Signifikant verringert hat sich die Streuung lediglich in Dänemark und Polen.

Zusammenhang zwischen dem sozioökonomischen beruflichen Status und der Lesekompetenz im internationalen Vergleich

Um Aussagen über den Zusammenhang zwischen sozioökonomischem beruflichem Status (HISEI) und Lesekompetenz treffen zu können, wurde eine lineare Regression berechnet. Die Ergebnisse werden anhand von drei Kennwerten, nämlich dem *Achsenabschnitt*, der *Steigung* der Regressionsgeraden (sog. sozialer Gradient) und der *Stärke des Zusammenhangs* (Varianzaufklärung R^2), in Tabelle 6.1 berichtet. Der *Achsenabschnitt* gibt die Lesekompetenz in Punkten an, die eine Schülerin oder ein Schüler im Mittel bei einem durchschnittlichen sozioökonomischen beruflichen Status der Eltern erreichen würde. Die *Steigung des sozialen Gradienten* sagt aus, um welchen Wert sich die Lesekompetenz der Jugendlichen vergrößern würde, wenn sich der sozioökonomische berufliche Status der Eltern um eine Standardabweichung erhöhte. Im OECD-Durchschnitt läge der Wert der Lesekompetenz um 31 Punkte höher, wenn der sozioökonomische berufliche Status der Eltern um eine Standardabweichung (21.5 Punkte, siehe Abbildung 6.1) höher ausfiele. In Deutschland liegt der soziale Gradient mit rund 40 Punkten signifikant über dem OECD-Durchschnitt. Auch die *Stärke des Zusammenhangs*, die den Anteil der Unterschiede in der Lesekompetenz angibt, der durch den sozioökonomischen beruflichen Status erklärt wird, ist in Deutschland mit rund 13 Prozent signifikant höher ausgeprägt als im OECD-Durchschnitt mit 10 Prozent. Die farblichen Markierungen in Tabelle 6.1 geben an, in welchen Staaten sich dieser Anteil der Varianzaufklärung bedeutsam vom OECD-Durchschnitt unterscheidet. Neben Deutschland ist die Stärke des Zusammenhangs in zehn weiteren OECD-Staaten – beispielsweise in Österreich, der Schweiz und Belgien – höher als im OECD-Durchschnitt. In insgesamt 15 Staaten liegt

Tabelle 6.1: Zusammenhang zwischen der Lesekompetenz und dem sozioökonomischen beruflichen Status in den OECD-Staaten (Prädiktorvariable: HISEI)

OECD-Staaten	Lesekompetenz		Steigung des sozialen Gradienten		Stärke des Zusammenhangs	
	Achsenabschnitt	(SE)	Steigung	(SE)	Varianzaufklärung R ²	(SE)
Japan	510	(2.4)	20.7	(2.0)	3.8	(0.7)
Republik Korea	517	(2.6)	25.3	(2.0)	5.0	(0.8)
Kanada	522	(1.5)	24.5	(1.2)	5.8	(0.5)
Island	470	(2.0)	27.2	(2.0)	5.9	(0.8)
Lettland	484	(1.5)	23.0	(1.3)	6.9	(0.8)
Finnland	522	(1.7)	25.5	(1.6)	7.0	(0.9)
Schweden	505	(2.6)	29.4	(2.3)	7.1	(1.1)
Norwegen	492	(2.1)	31.2	(1.8)	7.1	(0.8)
Estland	527	(1.8)	25.8	(1.8)	7.2	(0.9)
Vereinigtes Königreich	505	(2.0)	27.2	(1.5)	7.2	(0.8)
Irland	522	(1.8)	24.2	(1.4)	7.2	(0.8)
Australien	504	(1.6)	30.7	(1.1)	7.5	(0.5)
Dänemark	496	(1.6)	26.5	(1.5)	7.5	(0.9)
Italien	486	(2.4)	28.1	(1.8)	8.0	(1.0)
Niederlande	489	(2.5)	31.7	(2.1)	8.1	(1.0)
Vereinigte Staaten	506	(2.8)	32.4	(2.1)	8.8	(1.0)
Polen	520	(2.4)	29.4	(1.9)	9.2	(1.1)
Kolumbien	427	(3.1)	27.0	(2.1)	9.7	(1.4)
Griechenland	465	(3.0)	29.3	(1.8)	10.1	(1.1)
Litauen	482	(1.4)	28.5	(1.4)	10.5	(1.0)
Neuseeland	505	(1.8)	36.1	(1.9)	10.8	(1.1)
Slowenien	499	(1.3)	31.3	(1.4)	10.8	(0.9)
Türkei	485	(2.3)	28.3	(2.0)	10.9	(1.6)
Mexiko	438	(3.0)	29.6	(2.1)	11.0	(1.4)
Chile	469	(2.5)	29.7	(1.5)	11.0	(1.0)
Slowakei	471	(2.0)	35.5	(1.8)	12.3	(1.1)
Österreich	493	(2.3)	35.1	(1.7)	12.3	(1.1)
Tschechien	499	(2.1)	37.5	(2.2)	12.6	(1.3)
Deutschland	512	(2.6)	39.7	(1.7)	13.4	(1.1)
Schweiz	489	(2.4)	37.5	(2.4)	13.6	(1.6)
Portugal	500	(1.9)	34.0	(1.5)	13.8	(1.1)
Israel	470	(2.9)	44.5	(2.1)	14.2	(1.2)
Frankreich	501	(2.1)	37.3	(1.8)	14.2	(1.2)
Belgien	498	(2.0)	39.6	(1.5)	15.6	(1.0)
Ungarn	483	(2.0)	41.5	(1.9)	17.7	(1.5)
Luxemburg	478	(1.2)	44.8	(1.5)	18.7	(1.1)
OECD-Durchschnitt	493	(0.4)	31.4	(0.3)	10.1	(0.2)

Anmerkungen: Prädiktorvariable ist der sozioökonomische berufliche Status (HISEI). Der farblich gekennzeichnete Unterschied zum OECD-Durchschnitt bezieht sich auf die Stärke des Zusammenhangs.



signifikant unter dem OECD-Durchschnitt



nicht signifikant verschieden vom OECD-Durchschnitt



signifikant über dem OECD-Durchschnitt

der Anteil der Unterschiede in der Lesekompetenz, der durch den sozioökonomischen beruflichen Status der Eltern vorhergesagt wird, bedeutsam unter dem OECD-Durchschnitt. Zu diesen Staaten zählen unter anderem Japan, Kanada, Finnland, Schweden und Estland.

6.2.2 Zusammenhang zwischen dem sozioökonomischen und soziokulturellen Status und der Lesekompetenz im internationalen Vergleich

Äquivalent wurde zum Zusammenhang zwischen sozioökonomischem und soziokulturellem Status (ESCS) und der Lesekompetenz eine lineare Regression berechnet (siehe Tabelle 6.2). Erwartungsgemäß fällt die Steigung des sozialen Gradienten sowie die Stärke des Zusammenhangs mit diesem globaleren Indikator als Prädiktor höher aus als dies für den sozioökonomischen beruflichen Status der Fall ist. So liegt die Steigung des sozialen Gradienten für den sozioökonomischen und soziokulturellen Status als Prädiktor für die Lesekompetenz im OECD-Durchschnitt bei rund 37 Punkten. Bei einem um eine Standardabweichung höheren sozioökonomischen und -kulturellen Status ($SD = 0.93$) würde die Lesekompetenz also im Schnitt um 37 Punkte höher ausfallen. In Deutschland beträgt dieser Wert rund 42 Punkte und liegt damit signifikant über dem OECD-Durchschnitt. Ebenfalls höher als im OECD-Durchschnitt ist die Steigung des sozialen Gradienten in neun weiteren OECD-Staaten ausgeprägt, beispielsweise in der Schweiz, Belgien, Ungarn und Frankreich. In keinem OECD-Staat ist dieser Wert signifikant höher als in Deutschland. Bedeutsam niedriger als im OECD-Durchschnitt ist die Steigung des sozialen Gradienten in einer Gruppe von zehn Staaten, wozu unter anderem Mexiko, die Türkei, Lettland, Estland und Kanada gehören.

Der Anteil der Unterschiede in der Lesekompetenz, der durch den sozioökonomischen und -kulturellen Status erklärt wird, beträgt im OECD-Durchschnitt rund zwölf Prozent. Anhand der farblichen Markierungen in Tabelle 6.2 ist ersichtlich, in welchen Staaten sich diese Stärke des Zusammenhangs bedeutsam vom OECD-Durchschnitt unterscheidet. In Deutschland beträgt der Wert rund 17 Prozent und liegt damit – genau wie etwa in der Schweiz, in Belgien, Frankreich und Luxemburg – signifikant über dem OECD-Mittel. In insgesamt zwölf Staaten ist der Anteil der Unterschiede in der Lesekompetenz, der durch den sozioökonomischen und -kulturellen Status erklärt wird, signifikant kleiner als im OECD-Mittel. Dazu zählen unter anderem Estland, Kanada, Norwegen, Finnland und Dänemark.

Tabelle 6.2: Zusammenhang zwischen der Lesekompetenz und dem sozioökonomischen und -kulturellen Status in den OECD-Staaten (Prädiktorvariable: ESCS)

OECD-Staaten	Lesekompetenz		Steigung des sozialen Gradienten		Stärke des Zusammenhangs	
	Achsenabschnitt	(SE)	Steigung	(SE)	Varianzaufklärung R ²	(SE)
Estland	522	(1.8)	28.7	(2.1)	6.2	(0.8)
Island	458	(2.3)	32.8	(2.7)	6.6	(1.0)
Kanada	509	(1.6)	31.6	(1.6)	6.7	(0.6)
Lettland	480	(1.6)	28.6	(1.7)	7.2	(0.8)
Norwegen	483	(2.2)	35.1	(2.0)	7.5	(0.9)
Japan	508	(2.4)	37.6	(2.8)	8.0	(1.2)
Republik Korea	512	(2.6)	37.4	(2.8)	8.0	(1.1)
Italien	484	(2.3)	31.6	(1.9)	8.9	(1.0)
Finnland	510	(1.8)	37.7	(2.2)	9.2	(1.0)
Vereinigtes Königreich	498	(2.1)	33.4	(1.8)	9.3	(1.0)
Dänemark	483	(1.7)	37.9	(1.8)	9.9	(0.9)
Australien	493	(1.6)	38.2	(1.2)	10.1	(0.6)
Niederlande	476	(2.7)	38.7	(2.5)	10.5	(1.3)
Schweden	495	(2.6)	38.9	(2.2)	10.7	(1.2)
Irland	515	(1.8)	34.0	(1.7)	10.7	(1.1)
Griechenland	462	(3.1)	34.7	(2.1)	10.9	(1.2)
Türkei	494	(2.7)	25.0	(1.8)	11.4	(1.8)
Polen	518	(2.4)	39.0	(2.6)	11.6	(1.4)
Vereinigte Staaten	502	(2.7)	36.4	(2.1)	12.0	(1.4)
Slowenien	493	(1.2)	40.5	(1.8)	12.1	(1.0)
Chile	472	(2.5)	31.9	(1.5)	12.7	(1.1)
Neuseeland	502	(1.8)	39.0	(1.6)	12.9	(1.0)
Österreich	485	(2.2)	40.1	(1.9)	13.0	(1.2)
Litauen	476	(1.4)	39.5	(1.6)	13.2	(1.0)
Portugal	505	(2.0)	30.5	(1.4)	13.5	(1.2)
Mexiko	451	(3.7)	24.9	(1.7)	13.7	(1.7)
Kolumbien	444	(3.9)	26.2	(1.8)	13.7	(1.8)
Israel	458	(3.1)	46.6	(1.9)	14.0	(1.0)
Schweiz	486	(2.5)	43.4	(2.3)	15.6	(1.6)
Tschechien	500	(2.1)	44.9	(2.1)	16.5	(1.4)
Deutschland	510	(2.7)	41.9	(1.7)	17.2	(1.4)
Belgien	491	(2.0)	45.5	(1.3)	17.2	(0.8)
Slowakei	468	(1.9)	45.6	(2.0)	17.5	(1.5)
Frankreich	496	(2.1)	46.9	(2.0)	17.5	(1.3)
Luxemburg	471	(1.2)	39.8	(1.2)	17.8	(1.0)
Ungarn	482	(2.1)	45.8	(2.2)	19.1	(1.7)
OECD-Durchschnitt	489	(0.4)	37.0	(0.3)	12.0	(0.2)

Anmerkungen: Prädiktorvariable ist der ESCS. Der farblich gekennzeichnete Unterschied zum OECD-Durchschnitt bezieht sich auf die Stärke des Zusammenhangs.



signifikant unter dem OECD-Durchschnitt



nicht signifikant verschieden vom OECD-Durchschnitt



signifikant über dem OECD-Durchschnitt

Die Frage, ob eine hohe Varianzaufklärung durch den sozioökonomischen und -kulturellen Status mit einer hohen mittleren Lesekompetenz einhergeht, vertieft den Aspekt der Bildungsgerechtigkeit. Abbildung 6.2 zeigt, dass es eine Reihe von Staaten gibt, in denen die mittlere Lesekompetenz über dem OECD-Durchschnitt liegt, der Zusammenhang zwischen dem sozioökonomischen und -kulturellen Status (ESCS) und der Lesekompetenz aber unter dem OECD-Durchschnitt. Dazu gehören alle Staaten, die in dem Quadranten links oben abgebildet sind. In diesen Staaten liegt somit eine – im Hinblick auf den Ausgleich sozialer Ungleichheit – insgesamt wünschenswerte Konstellation vor. Am Beispiel von Estland, Kanada und Finnland wird deutlich, dass eine überdurchschnittliche mittlere Lesekompetenz auch mit einem geringen Zusammenhang zwischen sozialer Herkunft und Lesekompetenz einhergehen kann.

Deutschland – aber beispielsweise auch Belgien und Frankreich – gehört zu den Staaten, die sowohl in der mittleren Lesekompetenz als auch im Zusammenhang der Lesekompetenz mit dem sozioökonomischen und -kulturellen Status signifikant über dem OECD-Durchschnitt liegen. In diesen Staaten, die im Quadranten rechts oben zu finden sind, hängen im Mittel überdurchschnittlich hohe Werte in der Lesekompetenz stark mit der sozialen Herkunft der Jugendlichen zusammen.

Im Quadranten rechts unten befinden sich Staaten, in denen die mittlere Lesekompetenz der Jugendlichen unter dem OECD-Mittelwert liegt und diese gleichzeitig überdurchschnittlich stark durch die soziale Herkunft erklärt wird. Zu diesen Staaten gehören unter anderem die Slowakei, Luxemburg und Ungarn. Im Quadranten links unten befinden sich ebenfalls Staaten mit einer unterdurchschnittlichen Lesekompetenz, die allerdings unterdurchschnittlich stark mit der sozialen Herkunft zusammenhängt. Dies ist beispielsweise in Lettland, Italien und Island der Fall.

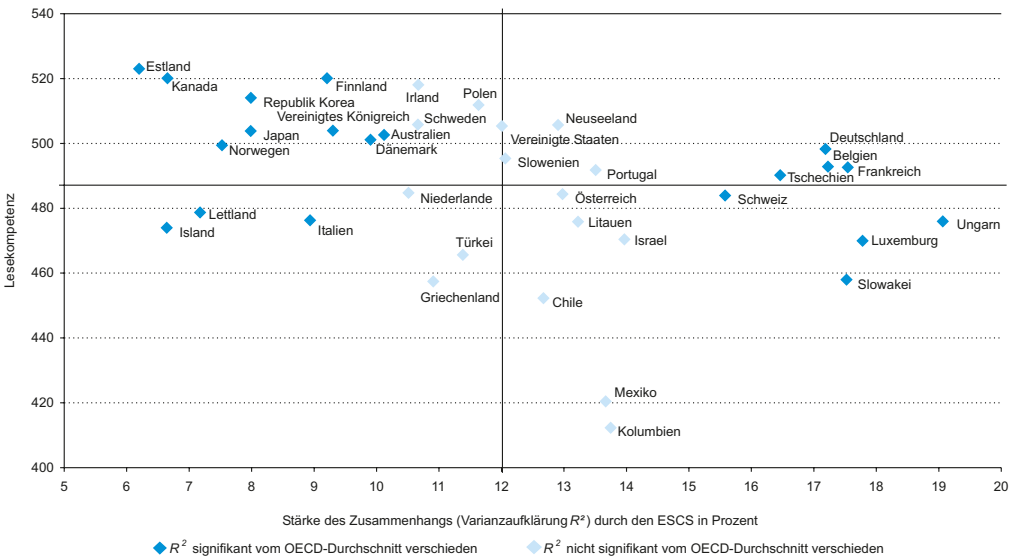


Abbildung 6.2: Lesekompetenz und Varianzaufklärung (R^2) durch den sozioökonomischen und -kulturellen Status (ESCS) in den OECD-Staaten

6.3 Soziale Herkunft und Lesekompetenz in Deutschland

Um die soziale Herkunft der Fünfzehnjährigen differenzierter beschreiben zu können, kommt in den nationalen Analysen mit den EGP-Klassen ein kategoriales Klassifikationsschema zur anschaulichen Beschreibung von Berufsgruppen zum Einsatz (siehe Abschnitt 6.1.2). In Tabelle 6.3 werden die Lebensbedingungen der Familien in Deutschland differenziert nach den EGP-Klassen dargestellt. Zunächst zeigt sich, dass den beiden Dienstklassen (EGP-Klassen I und II) rund 47 Prozent der Bezugspersonen zugeordnet werden können. Während diese Gruppe in der ersten PISA-Erhebung im Jahr 2000 mit 35 Prozent etwa ein Drittel der Gesamtgruppe ausmachte (Baumert & Schümer, 2001, S. 336), sind es aktuell beinahe die Hälfte. In fast allen Merkmalen, anhand derer mit den verfügbaren Daten die soziale Herkunft beschrieben wird, zeigt diese größte Gruppe überdurchschnittlich hohe Werte.

Der Blick auf den international klassifizierten sozioökonomischen beruflichen Status (HISEI) zeigt, dass sich dieser in allen Klassen jeweils deutlich vom Mittelwert in Deutschland unterscheidet (siehe Tabelle 6.3).⁵ Während die Werte des sozioökonomischen beruflichen Status in den beiden Dienstklassen (EGP-Klassen I und II) signifikant über dem nationalen Durchschnitt liegen, sind die Werte aller anderen EGP-Klassen unterdurchschnittlich ausgeprägt. Das heißt, die Werte des sozioökonomischen beruflichen Status der Bezugspersonen, deren Beruf der Gruppe der Routinedienstleistungen (EGP-Klasse III), der Selbständigen und Landwirte (EGP-Klasse IV), der Facharbeiter und leitenden Angestellten (EGP-Klassen V und VI) sowie der un- und angelernten Arbeiter und Landarbeiter (EGP-Klasse VII) zugeordnet wird, liegen signifikant unter dem nationalen Mittelwert. Was die Ausstattung mit Kultur- und Wohlstandsgütern (HOMEPOS) angeht zeigt sich ein vergleichbares Bild. Schülerinnen und Schüler, deren Bezugspersonen Tätigkeiten in den Dienstklassen (EGP-Klassen I und II) ausüben, verfügen über überdurchschnittlich viele Ressourcen. Die Ausstattung mit Wohlstandsgütern ist hingegen unterdurchschnittlich ausgeprägt, wenn die Bezugspersonen im Bereich der Routinedienstleistungen (EGP-Klasse III) beziehungsweise der EGP-Klassen V, VI und VII tätig sind. Die Ausstattung mit Kultur- und Wohlstandsgütern in der EGP-Klasse der Selbständigen (IV) liegt im durchschnittlichen Bereich. Eltern, die Berufe ausüben, die sich den beiden Dienstklassen (EGP-Klassen I und II) zuordnen lassen, weisen zudem deutlich höhere Bildungsabschlüsse auf, was die in Bildungsjahren ausgedrückten Werte illustrieren. Diese liegen in den beiden höchsten EGP-Klassen im Mittel bis zu zwei Jahre über dem nationalen Durchschnitt von 14,3 Jahren.

Hinsichtlich des sozioökonomischen beruflichen Status (HISEI), der Ausstattung mit Kultur- und Wohlstandsgütern sowie dem Bildungshintergrund der Eltern schreibt sich damit ein Muster in den EGP-Klassen fort, das sich bereits bei PISA 2015 (Müller

5 Aufgrund der verwendeten statistischen Verfahren (Imputation fehlender Werte; siehe Abschnitt 6.1.2) unterscheiden sich die Angaben des mittleren HISEI der Gesamtgruppe in Deutschland in Tabelle 6.3 geringfügig von den Werten in Abbildung 6.1.

& Ehmke, 2016) zeigte. Auffallend hingegen ist der Blick auf den Kindergartenbesuch als Annäherung an den frühen Bildungsweg der Jugendlichen. Während in PISA 2015 im Mittel rund 26 Prozent der befragten Jugendlichen in Deutschland angaben, erst im Alter von vier Jahren oder später in den Kindergarten gekommen zu sein, sind dies in der aktuellen Befragung nur noch rund 19 Prozent. Das heißt, die meisten der bei PISA 2018 fünfzehnjährigen Jugendlichen haben bereits im Alter von drei Jahren oder früher eine Einrichtung im elementaren Bildungsbereich besucht. Dabei ist der Anteil der Jugendlichen, die erst im Alter von vier Jahren oder später einen Kindergarten besuchten, bei Jugendlichen mit Eltern in den Dienstklassen (EGP-Klassen I und II) besonders gering und signifikant niedriger als im nationalen Durchschnitt. Bei PISA 2015 war dies bereits für die obere Dienstklasse (EGP-Klasse I) der Fall.

Tabelle 6.3: Merkmale der sozialen Herkunft und der Bildungswege differenziert nach EGP-Klassen

EGP-Klassen	Bezugsperson		Sozioökonomischer beruflicher Status (HISEI)		Besitz Wohlstandsgüter		Bildungsdauer der Eltern		Kindergartenbesuch mit vier Jahren oder älter	
	%	(SE)	M	(SE)	M	(SE)	M	(SE)	%	(SE)
Obere Dienstklasse (I)	28.4	(0.8)	66.9	(0.6)	0.4	(0.0)	15.8	(0.1)	14.2	(1.1)
Untere Dienstklasse (II)	18.3	(0.6)	57.4	(0.7)	0.1	(0.0)	14.8	(0.1)	14.8	(1.6)
Routinedienstleistungen Handel und Verwaltung (III)	6.0	(0.4)	43.0	(1.3)	-0.4	(0.1)	12.9	(0.3)	22.8	(3.8)
Selbständige (IV)	9.6	(0.5)	43.6	(0.9)	0.0	(0.1)	13.9	(0.2)	18.0	(2.2)
Facharbeiter und Arbeiter mit Leitungsfunktion (V-VI)	19.8	(0.7)	42.4	(0.6)	-0.2	(0.0)	13.6	(0.1)	21.9	(1.8)
Un- und angelernte Arbeiter, Landarbeiter (VII)	17.9	(0.7)	36.4	(0.7)	-0.4	(0.0)	13.1	(0.1)	20.0	(2.0)
Gesamt	100		51.2	(0.4)	0.0	(0.0)	14.3	(0.1)	18.6	(0.9)

Anmerkungen: $n = 5451$ (nach der Imputation fehlender Werte); signifikante Unterschiede ($p < .05$) in den Kennwerten für die EGP-Klassen zum Gesamtmittelwert sind fett gedruckt.

Die nach Kompetenzstufen differenzierte Betrachtung der sozialen Herkunft in den EGP-Klassen in Tabelle 6.4 zeigt, dass die Ausstattung mit Kultur- und Wohlstandsgütern (HOMEPOS) in der Gruppe der besonders lesestarken Jugendlichen (Kompetenzstufe V und VI) verglichen mit allen Schülerinnen und Schülern (siehe Tabelle 6.3) insgesamt hoch ist. Deutlich über dem gruppenspezifischen Durchschnitt der lesestarken Jugendlichen (Kompetenzstufe V und VI) liegen jene, deren Bezugspersonen in der oberen Dienstklasse (EGP-Klasse I) tätig sind. Aber auch bei lesestarken Jugendlichen, deren Eltern einer anderen EGP-Klassen zugeordnet wurden, finden sich bei den Kultur- und Wohlstandsgütern verglichen mit dem (z -standardisierten) Gesamtmittelwert (siehe Tabelle 6.3) teilweise überdurchschnittliche Werte.

Dass aber nicht alle lesestarken Schülerinnen und Schüler aus gut situierten Elternhäusern stammen, zeigt der Blick auf lesestarke Jugendliche (Kompetenzstufe V und VI) deren Bezugspersonen in einem un- oder angelernten Beruf aus dem manuellen Bereich (EGP-Klasse VII) tätig sind. Der durchschnittliche sozioökonomische berufliche Status (HISEI) dieser Bezugspersonen liegt mit rund 43 Punkten deutlich unter dem gruppenspezifischen Mittelwert ($M = 64.8$) der lesestarken Jugendlichen sowie dem der Gesamtgruppe ($M = 51.2$; siehe Tabelle 6.3). Die Ausstattung mit Kultur- und Wohlstandsgütern liegt zwar auch unter dem Gruppendurchschnitt der lesestarken Jugendlichen, ist jedoch vergleichbar mit dem nationalen Gesamtmittelwert. Im Gegensatz dazu findet sich in der Gruppe der leseschwachen Jugendlichen (unter Kompetenzstufe II), mit Ausnahme der Jugendlichen deren Eltern im Bereich der beiden Dienstklassen tätig sind, insgesamt eine unterdurchschnittliche sozioökonomische Gesamtsituation, die sich auch in der Ausstattung mit Kultur- und Wohlstandsgütern widerspiegelt.

Tabelle 6.4: Merkmale der sozialen Herkunft differenziert nach EGP-Klassen in den Gruppen der lesestarken (Kompetenzstufe V und VI) und leseschwachen (unter Kompetenzstufe II) Schülerinnen und Schüler

Lesekompetenz Kompetenzstufe V und VI						
EGP-Klassen	Sozioökonomischer beruflicher Status (HISEI)		Besitz Wohlstandsgüter		Bildungsdauer der Eltern	
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)
Obere Dienstklasse (I)	74.6	(0.9)	0.8	(0.0)	16.9	(0.2)
Untere Dienstklasse (II)	66.5	(1.5)	0.6	(0.1)	15.9	(0.3)
Routinedienstleistungen Handel und Verwaltung (III)	51.3	(5.3)	0.5	(0.2)	14.2	(0.7)
Selbständige (IV)	51.7	(3.3)	0.5	(0.2)	14.5	(0.7)
Facharbeiter und Arbeiter mit Leitungsfunktion (V-VI)	49.2	(2.5)	0.3	(0.1)	14.3	(0.5)
Un- und angelernte Arbeiter, Landarbeiter (VII)	43.3	(3.4)	0.2	(0.1)	13.7	(0.5)
Gesamt	64.8	(1.0)	0.6	(0.0)	15.8	(0.2)

Lesekompetenz unter Kompetenzstufe II						
EGP-Klassen	Sozioökonomischer beruflicher Status (HISEI)		Besitz Wohlstandsgüter		Bildungsdauer der Eltern	
	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)	<i>M</i>	(<i>SE</i>)
Obere Dienstklasse (I)	60.6	(1.8)	0.0	(0.1)	14.7	(0.3)
Untere Dienstklasse (II)	49.8	(1.9)	-0.3	(0.1)	13.9	(0.4)
Routinedienstleistungen Handel und Verwaltung (III)	39.5	(2.3)	-0.7	(0.2)	12.2	(0.7)
Selbständige (IV)	37.9	(2.0)	-0.3	(0.1)	12.6	(0.5)
Facharbeiter und Arbeiter mit Leitungsfunktion (V-VI)	38.4	(1.3)	-0.3	(0.1)	13.3	(0.3)
Un- und angelernte Arbeiter, Landarbeiter (VII)	34.0	(1.0)	-0.4	(0.1)	12.6	(0.3)
Gesamt	42.8	(0.7)	-0.3	(0.0)	13.3	(0.1)

Anmerkung: Signifikante Unterschiede ($p < .05$) in den Kennwerten für die EGP-Klassen zum gruppenspezifischen Mittelwert sind **fett** gedruckt.

Angesichts der dargestellten Disparitäten im sozioökonomischen beruflichen Status, in der Ausstattung mit Kultur- und Wohlstandsgütern und den Befunden zum sozialen Gradienten (siehe Abschnitt 6.2) zeigt sich wie erwartet, dass sich die den verschiedenen EGP-Klassen zugeordneten Jugendlichen in der Lesekompetenz unterscheiden. In Abbildung 6.3 werden diese Unterschiede quantifiziert und zudem mit den Befunden der Lesekompetenz bei PISA 2000 und 2009 verglichen. Bemerkenswert ist das deutlich höhere Niveau der Lesekompetenz von Jugendlichen mit Eltern in den EGP-Klassen V, VI und VII bei PISA 2018 im Vergleich zum Jahr 2000. Das heißt, dass sich insbesondere die Lesekompetenz von Jugendlichen, deren Bezugspersonen als Facharbeiter und leitende Angestellte (EGP-Klassen V und VI) oder als un- und angelernte Arbeiter und Landarbeiter (EGP-Klasse VII) tätig sind, seit PISA 2000 verbessert hat. In den anderen EGP-Klassen zeigen sich keine signifikanten Veränderungen bei PISA 2018 im Vergleich zu PISA 2000.

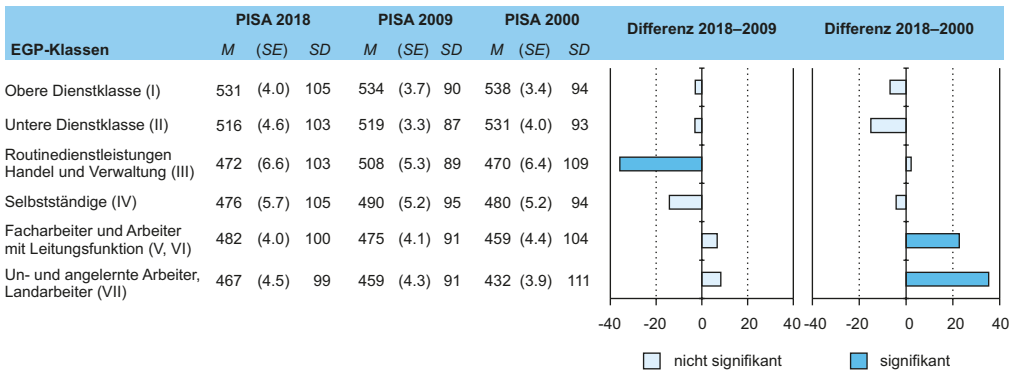


Abbildung 6.3: Mittelwerte und Streuungen der Lesekompetenz differenziert nach EGP-Klassen (Bezugsperson) bei PISA 2000, PISA 2009 und 2018 in Deutschland

6.4 Zuwanderungshintergrund und Lesekompetenz im europäischen Vergleich

Im Folgenden werden Unterschiede im sozioökonomischen beruflichen Status der Eltern (HISEI) zwischen Jugendlichen mit und ohne Zuwanderungshintergrund in 22 ausgewählten europäischen Staaten dargestellt. Zudem wird auf die Lesekompetenz der Jugendlichen mit und ohne Zuwanderungshintergrund in diesen Staaten eingegangen. Die Anteile der Fünfzehnjährigen mit Zuwanderungshintergrund in den ausgewählten Staaten bei PISA 2018 sowie deren Entwicklungen im Vergleich zu PISA 2009 sind in Tabelle 6.2web dargestellt. Die hier ausgewählten Staaten beschränken sich auf europäische Staaten, da sich diese aufgrund politischer Rahmenvorgaben durch europäische Institutionen bezüglich Migrationsbewegungen sinnvoll vergleichen lassen (Rauch et al., 2016). So wird auch im Migrationsbericht 2016/17 vom Bundesministerium des

Innern, für Bau und Heimat (2019) die Zu- und Abwanderung im europäischen Vergleich betrachtet und erläutert, dass die Vergleichbarkeit seit dem Jahr 2009 aufgrund entsprechender UN-Empfehlungen gegeben ist. Zudem zeigen sich ähnliche demografische Entwicklungen in den europäischen Staaten (vgl. Rauch et al., 2016). Polen und Rumänien konnten aufgrund zu geringer Anteile von Jugendlichen mit Zuwanderungshintergrund nicht für vergleichende Analysen berücksichtigt werden.

6.4.1 Sozioökonomischer beruflicher Status der Eltern von Jugendlichen mit und ohne Zuwanderungshintergrund

Die Eltern von Jugendlichen mit Zuwanderungshintergrund haben in nahezu allen europäischen Staaten einen geringeren sozioökonomischen beruflichen Status (HISEI) als Eltern von Jugendlichen ohne Zuwanderungshintergrund (siehe Tabelle 6.5). Dieser Unterschied ist in Deutschland im Vergleich zu anderen europäischen Staaten besonders groß. Ein höherer sozioökonomischer beruflicher Status in zugewanderten Familien als in Familien ohne Zuwanderungshintergrund findet sich in Portugal und Ungarn. In Bulgarien und der Slowakei finden sich keine signifikanten Unterschiede im sozioökonomischen beruflichen Status zwischen Eltern von Jugendlichen mit und ohne Zuwanderungshintergrund. Ein detaillierterer Blick auf die Zuwanderungsgenerationen zeigt über nahezu alle europäischen Staaten hinweg, dass vor allem bei Jugendlichen der ersten und zweiten Generation ein geringerer elterlicher sozioökonomischer beruflicher Status vorliegt. Dies ist auch für Deutschland der Fall. Der sozioökonomische berufliche Status (HISEI) der Familien mit einem im Ausland geborenen Elternteil unterscheidet sich hingegen in den meisten europäischen Staaten kaum von Familien, in denen beide Elternteile im jeweiligen europäischen Staat geboren sind. Deutschland gehört allerdings neben Belgien und Frankreich zur Gruppe der Staaten, in denen der sozioökonomische berufliche Status von Familien mit einem im Ausland geborenen Elternteil signifikant geringer ist als der von Familien ohne Zuwanderungshintergrund.

Tabelle 6.5: Unterschiede im sozioökonomischen beruflichen Status der Eltern (HISEI) zwischen Jugendlichen mit und ohne Zuwanderungshintergrund in ausgewählten europäischen Staaten

Staat	Ohne Zuwanderungshintergrund		Mit Zuwanderungshintergrund							
	<i>M</i>	(<i>SE</i>)	Insgesamt Zuwanderung		Ein Elternteil im Ausland geboren		Zweite Generation		Erste Generation	
			<i>DIFF</i>	(<i>SE</i>)	<i>DIFF</i>	(<i>SE</i>)	<i>DIFF</i>	(<i>SE</i>)	<i>DIFF</i>	(<i>SE</i>)
Westeuropa										
Belgien	55.7	(0.4)	-9.6	(0.7)	-4.3	(0.8)	-14.2	(0.9)	-14.3	(1.6)
Deutschland	55.5	(0.6)	-10.6	(0.8)	-3.1	(1.1)	-16.1	(1.0)	-12.9	(1.6)
Frankreich	53.6	(0.5)	-7.0	(0.9)	-2.7	(1.1)	-10.9	(1.2)	-12.5	(2.0)
Niederlande	56.7	(0.5)	-5.7	(0.8)	-0.3	(1.0)	-11.4	(1.2)	-10.6	(2.7)
Österreich	53.6	(0.4)	-8.6	(0.8)	1.4	(1.0)	-14.3	(0.7)	-13.1	(1.3)
Schweiz	56.8	(0.7)	-7.5	(0.8)	1.6	(0.9)	-14.3	(0.9)	-11.0	(1.4)
Vereinigtes Königreich	58.9	(0.6)	-2.0	(0.8)	1.3	(0.9)	-2.9	(1.2)	-6.7	(1.5)
Südeuropa										
Griechenland	52.0	(0.6)	-8.2	(0.7)	-0.8	(1.0)	-15.6	(1.2)	-18.4	(1.8)
Italien	48.8	(0.5)	-8.1	(0.7)	-0.8	(1.0)	-13.6	(1.3)	-16.4	(1.3)
Kroatien	48.0	(0.4)	-2.2	(0.7)	0.3	(0.7)	-7.0	(1.1)	-4.9	(2.6)
Portugal	48.3	(0.7)	5.5	(1.0)	8.8	(0.9)	-2.0	(2.1)	-3.9	(2.6)
Spanien	50.2	(0.4)	-7.3	(0.6)	-0.2	(0.7)	-12.7	(0.9)	-11.3	(0.8)
Nordeuropa										
Dänemark	61.8	(0.4)	-8.4	(0.8)	-1.2	(1.0)	-16.6	(1.0)	-15.2	(2.7)
Finnland	56.4	(0.5)	-3.9	(1.0)	0.6	(1.2)	-8.5	(2.1)	-11.3	(2.2)
Norwegen	62.7	(0.4)	-4.5	(0.8)	0.6	(0.9)	-7.3	(1.2)	-12.3	(1.8)
Schweden	61.4	(0.5)	-5.9	(0.9)	0.4	(0.9)	-9.3	(1.3)	-11.7	(1.5)
Osteuropa										
Bulgarien	49.2	(0.7)	0.3	(2.0)	2.3	(2.0)	-11.9	(5.4)	-3.0	(8.1)
Russische Föderation	56.1	(0.7)	-1.8	(0.8)	-0.9	(0.9)	-2.2	(1.7)	-5.0	(2.0)
Slowakei	48.6	(0.4)	1.1	(1.5)	0.9	(1.5)	4.4	(5.6)	0.5	(6.2)
Slowenien	53.7	(0.4)	-7.6	(1.1)	-1.4	(1.3)	-13.8	(1.8)	-15.6	(1.5)
Tschechien	49.5	(0.4)	-3.2	(1.2)	-0.9	(1.2)	-11.5	(2.1)	-4.8	(2.7)
Ungarn	50.4	(0.5)	4.6	(1.4)	4.5	(1.7)	8.5	(2.6)	0.6	(2.7)

Anmerkung: **fett**: signifikante Differenz zu Familien ohne Zuwanderungshintergrund ($p < .05$)

6.4.2 Lesekompetenz von Jugendlichen mit und ohne Zuwanderungshintergrund

Betrachtet man die Lesekompetenzen im Zusammenhang mit dem Zuwanderungshintergrund, so zeigen sich in fast allen der ausgewählten europäischen Staaten geringere Kompetenzwerte der Jugendlichen aus zugewanderten Familien im Vergleich zu Jugendlichen ohne Zuwanderungshintergrund (siehe Tabelle 6.6), wobei sich die Höhe dieser Disparitäten zum Teil deutlich unterscheidet. Nur in der Russischen Föderation und in Ungarn zeigen die Jugendlichen mit Zuwanderungshintergrund eine höhere mittlere Lesekompetenz als Jugendliche ohne Zuwanderungshintergrund. In Kroatien und Portugal gibt es keine signifikanten Unterschiede.

In Deutschland ist der Unterschied mit 52 Punkten zwischen Jugendlichen ohne Zuwanderungshintergrund (524 Punkte) und Jugendlichen mit Zuwanderungshintergrund (472 Punkte) relativ groß. Die Lesekompetenz der Gesamtgruppe der Jugendlichen mit Zuwanderungshintergrund veränderte sich bei PISA 2018 im Vergleich zu PISA 2009 in Deutschland allerdings nicht signifikant. In fünf europäischen Staaten zeigen sich im Vergleich zum Jahr 2009 deutliche Verschlechterungen der Lesekompetenz von Jugendlichen mit Zuwanderungshintergrund (z.B. in den Niederlanden, der Schweiz und Finnland), während sich lediglich in Österreich und der Russischen Föderation Verbesserungen beobachten lassen.

In den meisten westeuropäischen Staaten zeigt sich das Muster, dass Jugendliche der ersten Generation die geringsten Kompetenzwerte erreichen, gefolgt von Jugendlichen der zweiten Generation und Jugendlichen mit einem im Ausland geborenen Elternteil. Für Deutschland fällt im Vergleich zu den meisten anderen europäischen Staaten die besonders geringe Lesekompetenz von Jugendlichen der ersten Generation (405 Punkte) auf. Ähnlich niedrige Werte finden sich für die erste Generation lediglich in den Niederlanden, in Griechenland und in der Slowakei. Zudem hat sich in Deutschland die mittlere Lesekompetenz der ersten Generation bei PISA 2018 im Vergleich zu PISA 2009 deutlich verschlechtert, während sich die Lesekompetenz der zweiten Generation verbessert hat. Zu der Gruppe von Staaten, in welchen sich die mittlere Lesekompetenz der Jugendlichen der ersten Generation signifikant verringerte, gehören neben Deutschland auch Belgien, die Niederlande und Tschechien. In Österreich, dem Vereinigten Königreich und Italien nahm die mittlere Lesekompetenz der Jugendlichen der ersten Generation seit 2009 hingegen signifikant zu. Wie in Deutschland verbesserte sich auch in Österreich und der Russischen Föderation die mittlere Lesekompetenz der Jugendlichen der zweiten Generation. In einer Gruppe von fünf Staaten zeigen sich hingegen teilweise starke Verringerungen der Kompetenzwerte der Jugendlichen der zweiten Generation (Niederlande, Schweiz, Griechenland, Finnland, Slowakei). Die Schweiz und das Vereinigte Königreich sind die einzigen beiden westeuropäischen Staaten, in welchen sich die Jugendlichen mit einem im Ausland geborenen Elternteil hinsichtlich ihrer Lesekompetenz nicht signifikant von den Jugendlichen ohne Zuwanderungshintergrund

unterscheiden. Dies ist ebenfalls der Fall in insgesamt sieben der süd-, nord- und ost-europäischen Staaten (z. B. Italien, Norwegen, Schweden).

Tabelle 6.6: Lesekompetenz und Zuwanderungshintergrund in ausgewählten europäischen Staaten

Staat	Ohne Zuwanderungshintergrund			Mit Zuwanderungshintergrund											
				Insgesamt Zuwanderung			Ein Elternteil im Ausland geboren			Zweite Generation			Erste Generation		
	M	(SE)	+/-	M	(SE)	+/-	M	(SE)	+/-	M	(SE)	+/-	M	(SE)	+/-
Westeuropa															
Belgien	511	(2.6)	-11.7	462 ^a	(3.2)	-11.0	482 ^a	(3.9)	-14.4	459 ^a	(4.6)	3.8	427 ^a	(5.3)	-21.0
Deutschland	524	(3.5)	10.3	472 ^a	(4.9)	2.0	497 ^a	(5.2)	-2.7	477 ^a	(6.6)	19.8	405 ^a	(11.7)	-47.4
Frankreich	504	(3.0)	-3.2	472 ^a	(4.3)	2.2	493 ^a	(4.8)	-2.4	463 ^a	(5.8)	12.7	427 ^a	(7.5)	-0.5
Niederlande	501	(3.1)	-14.9	450 ^a	(4.6)	-37.5	477 ^a	(5.5)	-36.6	433 ^a	(6.7)	-36.4	399 ^a	(12.9)	-73.3
Österreich	501	(2.7)	18.7	455 ^a	(4.0)	21.8	490 ^a	(5.0)	15.1	446 ^a	(4.3)	19.0	421 ^a	(5.5)	36.8
Schweiz	506	(3.3)	-6.4	469 ^a	(3.9)	-18.0	498	(4.8)	-18.1	455 ^a	(4.6)	-16.1	448 ^a	(6.1)	-5.9
Vereinigtes Königreich	511	(2.6)	11.8	500 ^a	(3.9)	10.2	512	(5.2)	7.7	493 ^a	(5.7)	0.8	488 ^a	(7.0)	29.8
Südeuropa															
Griechenland	465	(3.4)	-24.1	442 ^a	(4.9)	-19.5	469	(5.0)	-23.8	421 ^a	(6.8)	-35.2	396 ^a	(9.1)	-24.3
Italien	483	(2.6)	-8.4	460 ^a	(3.8)	3.0	482	(5.2)	-7.2	445 ^a	(5.9)	-2.7	433 ^a	(7.3)	23.2
Kroatien	481	(2.8)	2.4	476	(3.6)	5.5	479	(3.8)	1.8	473	(5.7)	7.4	464	(11.8)	12.1
Portugal	492	(2.8)	2.5	496	(4.3)	1.6	509 ^a	(4.1)	2.9	485	(10.5)	12.0	436 ^a	(9.1)	-21.3
Nordeuropa															
Dänemark	511	(2.0)	9.4	472 ^a	(3.1)	1.4	501 ^a	(4.7)	-5.5	446 ^a	(3.7)	0.0	434 ^a	(7.5)	12.2
Finnland	529	(2.1)	-10.3	478 ^a	(5.5)	-25.6	513 ^a	(5.6)	-9.4	458 ^a	(10.7)	-36.1	420 ^a	(8.9)	-29.1
Norwegen	508	(2.3)	0.3	483 ^a	(4.0)	0.0	510	(4.7)	5.9	464 ^a	(7.2)	-0.4	451 ^a	(5.5)	2.4
Schweden	526	(2.8)	18.9	472 ^a	(4.9)	-0.9	519	(4.3)	13.4	472 ^a	(6.3)	16.4	410 ^a	(7.0)	-5.8
Osteuropa															
Bulgarien	426	(3.8)	-8.4	406 ^a	(8.2)	1.1	409	(9.5)	-4.6	376 ^a	(19.6)	42.7	412	(18.1)	46.8
Russische Föderation	478	(3.1)	15.8	487 ^a	(4.7)	30.4	491 ^a	(4.9)	13.6	492 ^a	(7.3)	58.7	456 ^a	(8.5)	12.4
Slowakei	460	(2.2)	-18.4	446 ^a	(7.3)	-22.9	452	(8.1)	-14.9	437	(17.6)	-92.2	391 ^a	(19.4)	-61.1
Slowenien	504	(1.4)	15.1	464 ^a	(3.2)	3.1	486 ^a	(5.0)	8.0	464 ^a	(7.3)	17.7	423 ^a	(8.2)	7.2
Tschechien	495	(2.4)	15.0	461 ^a	(6.4)	-5.9	473 ^a	(6.3)	2.1	459 ^a	(10.9)	13.3	421 ^a	(14.6)	-53.0
Ungarn	476	(2.3)	-18.5	494 ^a	(6.3)	-17.2	497 ^a	(7.4)	-18.4	510 ^a	(11.1)	-16.4	469	(16.8)	-22.7

Anmerkungen: **fett**: signifikante Unterschiede zwischen 2009 und 2018 ($p < .05$)

^a signifikante Unterschiede zu Schülerinnen und Schülern ohne Zuwanderungshintergrund ($p < .05$)

+/- Mittelwertsdifferenz 2018–2009.

6.5 Jugendliche mit Zuwanderungshintergrund in Deutschland

Bei der PISA-Studie 2018 beträgt der Anteil von Jugendlichen mit Zuwanderungshintergrund 36 Prozent. Damit ist der Anteil der Fünfzehnjährigen mit Zuwanderungshintergrund im Vergleich zu PISA 2009 um 10 Prozentpunkte gestiegen. Betrachtet man die prozentualen Anteile für die verschiedenen Gruppen des Zuwanderungsstatus getrennt, zeigt sich, dass insbesondere die Gruppe der Jugendlichen mit einem im Ausland geborenen Elternteil und die Gruppe der Fünfzehnjährigen der zweiten Generation im Vergleich zu PISA 2009 deutlich gewachsen sind (siehe Tabelle 6.7). Der Anteil der Fünfzehnjährigen der ersten Generation stieg nicht signifikant an.

Bemerkenswert ist bei PISA 2018 die sehr hohe Diversität der Herkunftsstaaten der Jugendlichen mit Zuwanderungshintergrund in Deutschland. Es stammen sechs Prozent aller Fünfzehnjährigen in Deutschland aus der ehemaligen UdSSR, fünf Prozent aus der Türkei, etwa drei Prozent aus Polen, darüber hinaus aber auch insgesamt 22 Prozent – und damit 10 Prozentpunkte mehr als 2009 – mit jeweils kleinen Anteilen aus anderen Ländern. Relativ viele der Familien dieser Gruppe stammen aus Syrien, dem Kosovo

Tabelle 6.7: Prozentuale Anteile der Fünfzehnjährigen mit Zuwanderungshintergrund in Deutschland

Zuwanderungsstatus	2009		2018	
	%	(SE)	%	(SE)
Zuwanderungsstatus eindeutig zuzuordnen?				
Nicht zuzuordnen	11.0	(0.8)	14.6	(1.1)
Zuzuordnen	89.0	(0.8)	85.4	(1.1)
	<i>gültige %*</i>	(SE)	<i>gültige %*</i>	(SE)
Ohne Zuwanderungshintergrund	74.3	(1.1)	64.4	(1.3)
Mit Zuwanderungshintergrund	25.7	(1.1)	35.6	(1.3)
Generationsstatus				
Ein Elternteil im Ausland geboren	8.1	(0.4)	13.4	(0.5)
Zweite Generation	11.7	(0.8)	15.8	(0.9)
Erste Generation	5.8	(0.4)	6.4	(0.5)
Herkunftsländer				
Ehemalige UdSSR	5.3	(0.5)	6.1	(0.5)
Türkei	5.7	(0.5)	5.1	(0.4)
Polen	2.8	(0.3)	2.6	(0.3)
Anderes Land	11.9	(0.6)	21.8	(1.1)

Anmerkungen: In der Tabelle werden gerundete Werte angegeben. Es können auf die Rundungen zurückzuführende vermeintliche Inkonsistenzen vorkommen.

Daten beruhen auf Schülerangaben.

fett: signifikante Unterschiede zwischen 2009 und 2018 ($p < .05$)

* gültige % = Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind.

und Rumänien. Insbesondere sind Jugendliche, deren Familien aus Italien, Bosnien und Herzegowina oder Griechenland stammen, nicht mehr am häufigsten anteilig in der Kategorie „anderes Land“ vertreten, wie dies bei PISA 2015 der Fall war (Rauch et al., 2016). Diese Veränderung lässt sich durch die gestiegenen Zuzüge aus Rumänien, Syrien und dem Kosovo nach Deutschland in den Jahren 2015 bis 2017 erklären (Bundesministerium des Innern, 2016; Bundesministerium des Innern, für Bau und Heimat, 2019).

Die soziale Herkunft und die zu Hause gesprochene Sprache von Jugendlichen mit Zuwanderungshintergrund

Im Folgenden werden Unterschiede in der sozialen Herkunft zwischen Jugendlichen mit und ohne Zuwanderungshintergrund analysiert. Wie in Tabelle 6.8 ersichtlich, ist der sozioökonomische berufliche Status der Eltern (HISEI) von Jugendlichen mit Zuwanderungshintergrund signifikant geringer als der von Jugendlichen ohne Zuwanderungshintergrund. Dies gilt auch, wenn die drei Kategorien des Zuwanderungsstatus jeweils separat betrachtet werden. Bei der differenzierten Analyse nach Herkunftsländern fällt auf, dass der elterliche sozioökonomische berufliche Status von Jugendlichen, deren Eltern aus der Türkei stammen, besonders niedrig ist.

Auch der Besitz von Wohlstandsgütern (HOMEPOS) von Fünfzehnjährigen mit Zuwanderungshintergrund ist signifikant geringer als von Gleichaltrigen ohne Zuwanderungshintergrund. Besonders gering ist der Besitz von Wohlstandsgütern von Jugendlichen der ersten Generation. Im Vergleich zu PISA 2009 hat der Besitz von Wohlstandsgütern von Jugendlichen der ersten Generation signifikant abgenommen. In Bezug auf die unterschiedlichen Herkunftsländer hat der Besitz von Wohlstandsgütern der Jugendlichen, deren Familien aus der ehemaligen UdSSR zugewandert sind, im Vergleich zu 2009 zugenommen, wohingegen dieser bei Jugendlichen, deren Eltern aus Polen stammen, abnahm. Besonders wenige Wohlstandsgüter stehen Fünfzehnjährigen, deren Eltern aus der Türkei stammen, zur Verfügung.

Während sich die elterliche Bildungsdauer (PARED) der Jugendlichen mit Zuwanderungshintergrund bei PISA 2018 im Vergleich zu PISA 2009 nicht verändert hat, ist die Bildungsdauer der Eltern von Jugendlichen ohne Zuwanderungshintergrund im Vergleich zu 2009 signifikant gestiegen und damit weiterhin signifikant höher als die der Eltern von Jugendlichen mit Zuwanderungshintergrund.

Im Folgenden wird die zu Hause gesprochene Sprache differenziert für die verschiedenen Gruppen des Zuwanderungsstatus betrachtet. Insgesamt hat sich der Anteil an Jugendlichen mit Zuwanderungshintergrund, die zu Hause hauptsächlich Deutsch sprechen, im Vergleich zu PISA 2009 nicht signifikant verändert (siehe Tabelle 6.8). Betrachtet man allerdings die einzelnen Gruppen nach Generationsstatus getrennt, zeigt sich, dass bei PISA 2018 deutlich weniger Jugendliche der ersten Generation hauptsächlich Deutsch zu Hause sprechen als dies 2009 noch der Fall war. Dasselbe gilt – wenngleich auf insgesamt deutlich höherem Niveau – für die Jugendlichen mit einem im Ausland geborenen Elternteil. Bei Fünfzehnjährigen der ersten Generation ist der Anteil

der Jugendlichen, die hauptsächlich zu Hause Deutsch sprechen, mit Abstand am niedrigsten. Der Anteil der Familien, die zu Hause Deutsch sprechen, ist bei aus der Türkei stammenden Familien geringer als in den anderen Herkunftsgruppen.

Tabelle 6.8: Unterschiede in der zu Hause gesprochenen Sprache sowie in der sozialen Herkunft zwischen Jugendlichen mit und Jugendlichen ohne Zuwanderungshintergrund

Zuwanderungsstatus	Sozioökonomischer beruflicher Status		Besitz Wohlstandsgüter ¹		Bildungsdauer der Eltern		Deutsch als zu Hause gesprochene Sprache							
	2018		2009		2018		2009							
	gültige %*	(SE)	gültige %*	(SE)	gültige %*	(SE)	gültige %*	(SE)						
Ohne Zuwanderung	55.5	(0.6)	0.3	(0.0)	0.3	(0.0)	14.5	(0.1)	14.8	(0.1)	99.3	(0.1)	98.7	(0.2)
Mit Zuwanderung	44.9 ^a	(0.7)	-0.2 ^a	(0.0)	-0.2 ^a	(0.0)	13.6 ^a	(0.1)	13.5 ^a	(0.1)	58.0 ^a	(1.9)	53.6 ^a	(1.4)
Generationsstatus														
Ein Elternteil im Ausland geboren	52.3 ^a	(1.0)	0.1 ^a	(0.1)	0.1 ^a	(0.0)	14.6	(0.2)	14.4 ^a	(0.2)	89.0 ^a	(1.7)	80.4^a	(1.7)
Zweite Generation	39.4 ^a	(0.9)	-0.3 ^a	(0.0)	-0.3 ^a	(0.0)	12.5 ^a	(0.2)	12.7 ^a	(0.2)	49.8 ^a	(3.1)	47.0 ^a	(2.0)
Erste Generation	42.6 ^a	(1.5)	-0.4 ^a	(0.0)	-0.6^a	(0.1)	13.9 ^a	(0.3)	13.4 ^a	(0.3)	26.3 ^a	(3.5)	13.5^a	(2.5)
Herkunftsländer														
Ehem. UdSSR	43.6 ^a	(1.4)	-0.3 ^a	(0.0)	0.0^a	(0.1)	14.3	(0.2)	14.5	(0.3)	55.9 ^a	(4.6)	56.3 ^a	(3.0)
Türkei	35.5 ^a	(1.5)	-0.5 ^a	(0.1)	-0.6 ^a	(0.1)	11.5 ^a	(0.3)	11.0 ^a	(0.3)	42.0 ^a	(3.7)	38.2 ^a	(3.4)
Polen	43.6 ^a	(1.8)	0.0 ^a	(0.1)	-0.3^a	(0.1)	14.3	(0.2)	14.3	(0.3)	60.3 ^a	(5.2)	53.4 ^a	(5.1)
Anderes Herkunftsland	47.6 ^a	(0.8)	-0.1 ^a	(0.0)	-0.2 ^a	(0.0)	14.1 ^a	(0.2)	13.7 ^a	(0.1)	66.1 ^a	(2.1)	56.2^a	(1.8)

Anmerkungen: Die Angaben beruhen auf Schülerangaben.

fett: signifikante Unterschiede zwischen 2009 und 2018 ($p < .05$)

^a signifikante Unterschiede zu Schülerinnen und Schülern ohne Zuwanderungshintergrund ($p < .05$)

¹ z-standardisiert

* gültige % = Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind. Da die Vergleichbarkeit des sozioökonomischen beruflichen Status zwischen PISA 2009 und PISA 2018 durch die Überarbeitung der Berufsklassifizierung nicht eindeutig ist, wurde auf eine Analyse der Veränderung des sozioökonomischen beruflichen Status (HISEI) seit PISA 2009 verzichtet.

Die Ergebnisse zu den kulturellen Ressourcen (CULTPOSS) finden sich in Tabelle 6.4web.

Bildungsbeteiligung von Jugendlichen mit Zuwanderungshintergrund

Da im deutschen Schulsystem die Aufgliederung in verschiedene weiterführende Schularten nach der Grundschule eine Besonderheit darstellt, wird im Folgenden berichtet, welche Anteile der Fünfzehnjährigen mit und ohne Zuwanderungshintergrund ein Gymnasium oder eine nicht gymnasiale Schule besuchen. Der Anteil Fünfzehnjähriger ohne Zuwanderungshintergrund, die ein Gymnasium besuchen, ist um 13 Prozentpunkte höher als der Anteil der Gleichaltrigen mit Zuwanderungshintergrund. An den nicht gymnasialen Schulen ist der Anteil der Fünfzehnjährigen mit Zuwanderungshintergrund um 14 Prozentpunkte höher als der Anteil der Fünfzehnjäh-

rigen ohne Zuwanderungshintergrund (siehe Tabelle 6.9). Dieser Unterschied ist besonders deutlich für Jugendliche der ersten Generation, von denen insgesamt nur 16 Prozent ein Gymnasium besuchen. Erfreulich ist, dass der Anteil der Jugendlichen der zweiten Generation, die ein Gymnasium besuchen, im Jahr 2018 im Vergleich zu PISA 2009 um zehn Prozentpunkte angestiegen ist.

Tabelle 6.9: Prozentuale Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund an Gymnasien und nicht gymnasialen Schularten

	Gymnasium		nicht gymnasiale Schularten*	
	%	(SE)	%	(SE)
Ohne Zuwanderungshintergrund	43.0	(2.9)	52.8	(2.9)
Mit Zuwanderungshintergrund	29.8	(2.4)	66.4	(2.5)
Generationsstatus				
Ein Elternteil im Ausland geboren	35.7	(2.7)	61.7	(2.7)
Zweite Generation	30.3	(3.3)	65.8	(3.3)
Erste Generation	16.1	(2.9)	77.9	(3.3)
Nicht zuzuordnen	19.7	(3.3)	71.9	(4.3)

Anmerkungen: * Nicht gymnasiale Schularten umfassen Hauptschule, Integrierte Gesamtschule, Realschule und Schulen mit mehreren Bildungsgängen.

fett: signifikante Unterschiede zwischen 2009 und 2018 ($p < .05$)

6.6 Zuwanderungsbezogene und soziale Herkunft im Zusammenhang mit Lesekompetenz in Deutschland

Es stellt sich die Frage, ob beziehungsweise in welchem Maße die Kompetenzunterschiede im Lesen zwischen Fünfzehnjährigen mit und ohne Zuwanderungshintergrund durch die Unterschiede in der sozialen Herkunft erklärt werden können. Außerdem ist von Interesse, welche Rolle die zu Hause gesprochene Sprache spielt. Diese Fragen wurden mittels multivariater Regressionsanalysen für PISA 2018 und 2009 im Vergleich untersucht. Mit den Regressionsmodellen wurde geprüft, ob Unterschiede in der Lesekompetenz zwischen Jugendlichen mit und ohne Zuwanderungshintergrund bestehen bleiben oder signifikant geringer werden, wenn der Generationsstatus (im ersten Schritt, Modell I), Merkmale der sozialen Herkunft (im zweiten Schritt, Modell II) und die zu Hause gesprochene Sprache (im dritten Schritt, Modell III) schrittweise einbezogen werden. Außerdem wurde gezeigt, welches der Modelle die größte Varianzaufklärung der Lesekompetenz hat.

Für PISA 2018 zeigt sich im ersten Modell, dass die Jugendlichen mit Zuwanderungshintergrund in allen Generationsstatusgruppen signifikant niedrigere Lesekompetenzen aufweisen als Jugendliche ohne Zuwanderungshintergrund (siehe Spalte Modell I in

Tabelle 6.10). Bei PISA 2009 war dies für Jugendliche mit einem im Ausland geborenen Elternteil nicht der Fall, diese Gruppe unterschied sich damals nicht signifikant von der Gruppe der Jugendlichen ohne Zuwanderungshintergrund. Zugleich zeigt sich, dass die mittlere Lesekompetenz der Jugendlichen mit einem im Ausland geborenen Elternteil bei PISA 2018 signifikant niedriger ist als im Jahr 2009. Darüber hinaus ist sowohl bei Jugendlichen der ersten Generation als auch bei Jugendlichen, die keine Angaben zum Zuwanderungsstatus gemacht haben, bei PISA 2018 eine geringere Lesekompetenz als im Jahr 2009 festzustellen. Ein weiteres interessantes Ergebnis ist, dass sich die Gruppe der Jugendlichen ohne Zuwanderungshintergrund bei PISA 2018 im Vergleich zum Jahr 2009 signifikant verbessert hat. Damit zeigt diese Gruppe seit Beginn der PISA-Studien erstmals eine signifikant positive Entwicklung (vgl. Weis et al., 2018).

Die Ergebnisse des zweiten Modells zeigen, dass die Aspekte der sozialen Herkunft die Unterschiede in der Lesekompetenz zwischen Jugendlichen mit und ohne Zuwanderungshintergrund zu einem relativ hohen Anteil erklären können (siehe Spalte Modell II in Tabelle 6.10). Sowohl der sozioökonomische berufliche Status als auch der Besitz von Wohlstandsgütern und die elterliche Bildungsdauer hängen positiv mit der Lesekompetenz zusammen. Mit der Aufnahme dieser drei Merkmale der sozialen Herkunft in das Regressionsmodell verringern sich die Regressionskoeffizienten der Generationsstatusgruppen signifikant und das Modell hat eine deutlich höhere Varianzaufklärung. Wenn die soziale Herkunft einbezogen wird, gibt es zwischen Jugendlichen der zweiten Generation und Jugendlichen ohne Zuwanderungshintergrund keinen signifikanten Unterschied in der Lesekompetenz. Dies war bei PISA 2009 nicht der Fall. Dennoch erklären die einbezogenen Indikatoren der sozialen Herkunft die Kompetenzunterschiede zwischen Jugendlichen mit und ohne Zuwanderungshintergrund nicht vollständig. Im Vergleich zu 2009 haben sozioökonomischer beruflicher Status und Besitz von Wohlstandsgütern bei PISA 2018 einen größeren Einfluss, während der Einfluss der Bildungsdauer der Eltern abnahm. Hierbei ist allerdings zu beachten, dass die Änderung in der Klassifikation des sozioökonomischen beruflichen Status (siehe Abschnitt 6.1.2) sowie leichte Veränderungen in der Definition der Wohlstandsgüter (HOMEPOS) diese Veränderungen der Koeffizienten im Vergleich zum Jahr 2009 (mit)erklären könnten. Auch im zweiten Modell hat die Lesekompetenz der Jugendlichen der ersten Generation im Vergleich zu PISA 2009 abgenommen.

Im dritten Modell wird deutlich, dass die zu Hause gesprochene Sprache signifikant zur Vorhersage der Lesekompetenz beiträgt, selbst wenn Zusammenhänge zum Zuwanderungshintergrund und zur sozialen Herkunft einbezogen wurden (siehe Spalte Modell III in Tabelle 6.10). Wenn zu Hause hauptsächlich eine andere Sprache als Deutsch gesprochen wird, ist die durchschnittliche Lesekompetenz der Jugendlichen um 45 Punkte niedriger. Auch bei PISA 2009 hatte die zu Hause gesprochene Sprache einen signifikanten Einfluss, wobei sich dieser bei PISA 2018 noch signifikant vergrößert hat. Wie im Jahr 2009 sind auch bei PISA 2018 die Unterschiede in der Lesekompetenz zwischen Jugendlichen ohne Zuwanderungshintergrund und Jugendlichen der zweiten Generation sowie Jugendlichen mit einem im Ausland geborenen Elternteil nicht mehr

signifikant, wenn soziale Herkunft und zu Hause gesprochene Sprache einbezogen werden. Für die Gruppe der Jugendlichen der zweiten Generation zeigt sich (bei Einbezug von sozialer Herkunft und Deutsch als zu Hause gesprochener Sprache) zudem eine signifikante Verbesserung der mittleren Lesekompetenz im Vergleich zu PISA 2009.

Tabelle 6.10: Regressionsmodelle zu Zuwanderungshintergrund, sozialer Herkunft und Lesekompetenz für PISA 2018 und 2009

	Modell I		Modell II		Modell III	
	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)
2018						
Ohne Zuwanderung	529*	(3.5)	520* ^a	(2.9)	521* ^a	(2.9)
Ein Elternteil im Ausland geboren	-22^a	(5.2)	-13 ^a	(4.7)	-5	(4.6)
Zweite Generation	-43 ^a	(7.2)	-8	(6.8)	14	(7.4)
Erste Generation	-108^a	(12.6)	-70^a	(12.2)	-34 ^a	(12.7)
Nicht zuzuordnen	-105^a	(15.8)	-72 ^a	(15.4)	-62 ^a	(15.2)
Sozioökonomischer beruflicher Status ¹			24^{ab}	(1.8)	24^{ab}	(1.8)
Besitz Wohlstandsgüter ¹			21^a	(2.0)	20^a	(2.0)
Bildungsdauer der Eltern ¹			4^a	(1.7)	4^a	(1.7)
Sprachgebrauch ²					-45^a	(6.0)
<i>n</i>	4256		4256		4256	
<i>R</i> ²	0.08		0.22		0.23	
	Modell I		Modell II		Modell III	
	<i>b</i>	(SE)	<i>b</i>	(SE)	<i>b</i>	(SE)
2009						
Ohne Zuwanderung	519*	(2.5)	515*	(2.2)	515*	(2.2)
Ein Elternteil im Ausland geboren	-2	(6.5)	1	(5.6)	3	(5.4)
Zweite Generation	-46 ^a	(7.3)	-17 ^a	(6.5)	-8	(6.3)
Erste Generation	-53 ^a	(7.6)	-31 ^a	(7.0)	-18 ^a	(7.7)
Nicht zuzuordnen	-59 ^a	(10.8)	-41 ^a	(9.9)	-38 ^a	(9.6)
Sozioökonomischer beruflicher Status ¹			18 ^a	(1.8)	18 ^a	(1.8)
Besitz Wohlstandsgüter ¹			13 ^a	(1.8)	13 ^a	(1.7)
Bildungsdauer der Eltern ¹			14 ^a	(1.8)	14 ^a	(1.8)
Sprachgebrauch ²					-18 ^a	(7.2)
<i>n</i>	3900		3900		3900	
<i>R</i> ²	0.04		0.18		0.18	

Anmerkungen: Daten beruhen auf Schülerangaben.

* Die hier angegebene Regressionskonstante ist der geschätzte Mittelwert in der Gruppe der Schülerinnen und Schüler ohne Migrationshintergrund. Für die Schülerinnen und Schüler mit Migrationshintergrund gibt der Regressionskoeffizient jeweils die Abweichung von der Regressionskonstanten an.

fett: Unterschied zum Regressionskoeffizienten für 2009 statistisch signifikant ($p < .05$)

^a signifikante Partialregressionskoeffizienten ($p < .05$)

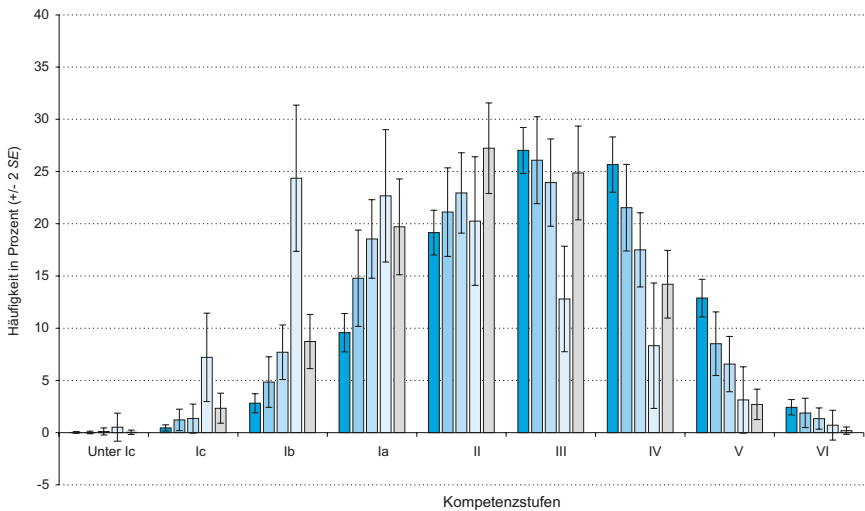
¹ z-standardisiert

² Referenzgruppe: Deutsch als Familiensprache

^b eingeschränkte Vergleichbarkeit der Signifikanz zwischen den Erhebungen aufgrund der Überarbeitung der Berufsklassifikation von ISCO-88 zu ISCO-08

Verteilung auf die Stufen der Lesekompetenz

Um die Lesekompetenz der Jugendlichen mit Zuwanderungshintergrund in Deutschland noch differenzierter zu betrachten, werden in Abbildung 6.4 die prozentualen Anteile auf den Kompetenzstufen der Lesekompetenz für die verschiedenen Zuwanderungsgruppen dargestellt (vgl. Kapitel 3 für die inhaltliche Beschreibung der Stufen). Bei der Betrachtung der Verteilungen auf den Kompetenzstufen fällt auf, dass mehr Schülerinnen und Schüler ohne Zuwanderungshintergrund die obersten Stufen (V und VI) erreichen als Jugendliche mit Zuwanderungshintergrund. Umgekehrt befinden sich mehr Schülerinnen und Schüler mit Zuwanderungshintergrund auf den untersten Kompetenzstufen (unter II). Insbesondere ist der Anteil von Jugendlichen der ersten Generation auf den untersten Kompetenzstufen (unter II) besonders hoch (55 %). Im Vergleich zu PISA 2009 (34 %) ist dieser Anteil bedeutsam gewachsen. Andererseits hat der Anteil der besonders lesestarken Jugendlichen der zweiten Generation (8 % auf den obersten Kompetenzstufen V und VI) im Vergleich zu PISA 2009 (3 %) bedeutsam zugenommen. Der Anteil der Jugendlichen ohne Zuwanderungshintergrund auf den obersten Kompetenzstufen (V und VI) hat bei PISA 2018 (15 %) im Vergleich zu PISA 2009 (10 %) ebenfalls signifikant zugenommen.



	Unter Ic	Ic	Ib	Ia	II	III	IV	V	VI
	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)	% (SE)
Ohne Zuwanderungshintergrund	0.0 (0.0)	0.4 (0.1)	2.8 (0.5)	9.6 (0.9)	19.2 (1.1)	27.0 (1.1)	25.7 (1.3)	12.9 (0.9)	2.4 (0.4)
<i>Generationsstatus</i>									
Ein Elternteil im Ausland geboren	0.0 (0.1)	1.2 (0.5)	4.8 (1.2)	14.8 (2.3)	21.1 (2.1)	26.1 (2.1)	21.5 (2.1)	8.5 (1.5)	1.9 (0.7)
Zweite Generation	0.1 (0.2)	1.3 (0.7)	7.7 (1.3)	18.5 (1.9)	23.0 (1.9)	23.9 (2.1)	17.5 (1.8)	6.6 (1.3)	1.3 (0.5)
Erste Generation	0.5 (0.7)	7.2 (2.1)	24.4 (3.5)	22.7 (3.2)	20.3 (3.1)	12.8 (2.5)	8.3 (3.0)	3.1 (1.6)	0.7 (0.7)
Nicht zuzuordnen	0.0 (0.1)	2.3 (0.7)	8.7 (1.3)	19.7 (2.3)	27.2 (2.2)	24.9 (2.2)	14.2 (1.6)	2.7 (0.7)	0.2 (0.2)

Abbildung 6.4: Prozentuale Anteile auf den Stufen der Lesekompetenz nach Zuwanderungsstatus in Deutschland

6.7 Zusammenfassung und Ausblick

Die Ergebnisse der PISA-Studie 2018 zeigen erneut, dass im Vergleich zu anderen OECD-Staaten der Zusammenhang zwischen sozialer Herkunft und Lesekompetenz in Deutschland besonders stark ausgeprägt ist. Dieser wurde anhand zweier Indikatoren (ESCS und HISEI) regressionsanalytisch überprüft. Dabei zeigte sich jeweils, dass die Steigung des sozialen Gradienten in Deutschland signifikant höher ausgeprägt ist als im OECD-Durchschnitt. Zudem ist der Anteil der Unterschiede in der Lesekompetenz, der durch die soziale Herkunft (HISEI bzw. ESCS) erklärt wird (Stärke des Zusammenhangs, Varianzaufklärung), in Deutschland ebenfalls deutlich höher ausgeprägt als im OECD-Durchschnitt.

Die Ergebnisse belegen, dass es nach wie vor sehr starke soziale Disparitäten in der Lesekompetenz in Deutschland gibt. Sowohl die mittlere Lesekompetenz als auch die Vorhersagbarkeit der Lesekompetenz durch den sozioökonomischen und -kulturellen Status liegen signifikant über dem OECD-Durchschnitt. Andere Staaten wie beispielsweise Estland, Kanada und Finnland zeigen eine überdurchschnittliche mittlere Lesekompetenz bei einem geringen Zusammenhang zwischen sozialer Herkunft und Lesekompetenz. Somit bestehen in Deutschland nach wie vor Herausforderungen, den Zusammenhang zwischen der Lesekompetenz und der sozialen Herkunft zu entkoppeln.

Trotz der zahlreichen Belege für den starken Zusammenhang zwischen sozioökonomischem Status und schulischen Kompetenzen wie Lesekompetenz (z. B. Davis-Kean, 2005; Magnuson, 2007; Sirin, 2005), ist noch nicht hinreichend erforscht, wie er erklärt werden kann. Es ist anzunehmen, dass der Zusammenhang über multiple Faktoren vermittelt wird und durch verschiedene und miteinander verknüpfte Kontexte (z. B. Schule, Familie, Nachbarschaft) bedingt ist (vgl. z. B. Aikens & Barbarin, 2008; Bronfenbrenner, 1979). So gibt es Hinweise dafür, dass kontextuelle Schulvariablen wie die Wirksamkeit der Lehrkräfte und die Qualität der Schule den Zusammenhang teilweise erklären können (z. B. Martins & Veiga, 2010). Allerdings scheinen auch die häusliche Umgebung (z. B. Magnuson, 2007), die Stimulation des Lesens (Aikens & Barbarin, 2008) und die kindliche Selbstregulation (Sektan, McClelland, Acock & Morrison, 2010; Størksen, Ellingsen, Wanless & McClelland, 2014) eine Rolle beim Einfluss der sozialen Herkunft auf die Schulleistungen zu spielen. Da der Zusammenhang zwischen sozialer Herkunft und Lesekompetenz multiple Ursachen zu haben scheint und nicht auf einen Kontext beschränkt werden kann, sollten zielgerichtete Interventionen entwickelt werden, welche die wechselseitigen Verknüpfungen zwischen den Kontexten sowie multiple Risikofaktoren berücksichtigen (Aikens & Barbarin, 2008).

In Deutschland ist bei PISA 2018 der Zuwanderungshintergrund der Jugendlichen besonders stark mit dem sozioökonomischen beruflichen Status der Eltern assoziiert, auch wenn in nahezu allen europäischen Staaten Eltern von Jugendlichen mit Zuwanderungshintergrund einen geringeren sozioökonomischen beruflichen Status (HISEI) aufweisen als solche ohne Zuwanderungshintergrund. Die OECD ordnet Deutschland einer Gruppe von Staaten zu, in denen ein besonders hoher Anteil der Jugendlichen mit

Zuwanderungshintergrund über vergleichsweise wenig sozioökonomische Ressourcen verfügt (OECD, 2019).

Die Ergebnisse der Regressionsmodelle zeigen, dass die Indikatoren der sozialen Herkunft deutlich zur Erklärung von Disparitäten in der Lesekompetenz zwischen Jugendlichen mit und ohne Zuwanderungshintergrund beitragen. Allerdings können sie die Disparitäten nicht vollständig erklären. Die zu Hause gesprochene Sprache leistet einen zusätzlichen Beitrag zur Aufklärung der Unterschiede. Bei der Interpretation der Ergebnisse bezüglich der zu Hause gesprochenen Sprache ist einschränkend zu beachten, dass bei den PISA-Berichten bisher lediglich berücksichtigt wird, welche Sprache hauptsächlich im Familienkontext gesprochen wird (Rauch et al., 2016). Agirdag und Vanlaar (2018) zeigten jedoch mittels Reanalysen der Daten von PISA 2012, dass der Zusammenhang zwischen zu Hause gesprochener Sprache und Lesekompetenz teilweise über Hintergrundmerkmale der Jugendlichen, Schulmerkmale sowie weitere Kontexte des Sprachgebrauchs erklärt werden kann. Zudem hängt der Gebrauch der Instruktionssprache im Schulkontext positiv mit der Lesekompetenz zusammen (so z.B. das Sprechen mit Mitschülerinnen und Mitschülern in der Instruktionssprache). Zukünftige Forschung könnte differenziert untersuchen, welche Gesprächspartner, Kommunikationszwecke oder Kontexte eine Rolle für den Zusammenhang zwischen der Nutzung der Instruktionssprache und der Lesekompetenz spielen.

Im Vergleich zu anderen europäischen Staaten ist in Deutschland der Unterschied in der Lesekompetenz zwischen Jugendlichen mit und ohne Zuwanderungshintergrund bei PISA 2018 relativ groß. Betrachtet man die Lesekompetenz der Jugendlichen mit Zuwanderungshintergrund in Deutschland im Vergleich zu PISA 2009, so hat sich der Mittelwert nicht signifikant verändert. Bei der differenzierten Betrachtung der Generationsgruppen wird hingegen ersichtlich, dass bei PISA 2018 im Vergleich zu PISA 2009 die mittlere Lesekompetenz der ersten Generation in Deutschland deutlich abgenommen hat, während die Lesekompetenz der zweiten Generation signifikant zugenommen hat. Bei PISA 2018 ist insbesondere der Anteil der besonders leseschwachen Jugendlichen (unter Kompetenzstufe II) der ersten Generation sehr hoch und seit 2009 deutlich angestiegen. Damit verfügt über die Hälfte der Jugendlichen der ersten Generation über nur eingeschränkte Lesekompetenzen. Bei der zweiten Generation hat hingegen der Anteil besonders lesestarker Jugendlicher im Vergleich zu PISA 2009 bedeutsam zugenommen.

Zudem zeigt sich bei PISA 2018 insgesamt ein Anstieg des Anteils der Fünfzehnjährigen mit Zuwanderungshintergrund sowie eine sehr hohe Diversität der Herkunftsländern der Jugendlichen mit Zuwanderungshintergrund in Deutschland. Im Jahr 2009 stammte noch ein verhältnismäßig großer Anteil der Jugendlichen mit Zuwanderungshintergrund aus der ehemaligen UdSSR, der Türkei und Polen und ein kleinerer Anteil aus „anderen Herkunftsländern“. Somit ist die Zusammensetzung der Gruppe der Schülerinnen und Schüler mit Zuwanderungshintergrund im Jahr 2018 heterogener geworden.

Die Ergebnisse von PISA 2018 zeigen, dass in Deutschland weiterhin Handlungsbedarf besteht, um die Integrationskraft des Bildungssystems zu verbessern. Es sind evi-

denzbasierte Interventionsansätze nötig, um Leistungsdisparitäten zu verringern, die bezüglich zuwanderungsbezogener und sozialer Herkunft bestehen.

Literatur

- Agirdag, O. & Vanlaar, G. (2018). Does more exposure to the language of instruction lead to higher academic achievement? A cross-national examination. *International Journal of Bilingualism*, 22, 123–137. <http://dx.doi.org/10.1177/1367006916658711>
- Aikens, N. L. & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of educational psychology*, 100(2), 235. <https://doi.org/10.1037/0022-0663.100.2.235>
- Alba, R. D., Handl, J. & Müller, W. (1994). Ethnische Ungleichheit im deutschen Bildungssystem. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 46, 209–237. <http://dx.doi.org/10.1007/s11577-017-0421-x>
- Baumert, J. & Maaz, K. (2006). Das theoretische und methodische Konzept von PISA zur Erfassung sozialer und kultureller Ressourcen der Herkunftsfamilie: Internationale und nationale Rahmenkonzeption. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000* (S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90082-7_1
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–410). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_10
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Bundesministerium des Innern. (2016). *Migrationsbericht des Bundesamtes für Migration und Flüchtlinge im Auftrag der Bundesregierung. Migrationsbericht 2015*. Verfügbar unter http://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/Migrationsberichte/migrationsbericht-2015.pdf?__blob=publicationFile
- Bundesministerium des Innern, für Bau und Heimat. (2019). *Migrationsbericht der Bundesregierung. Migrationsbericht 2016/17*. Verfügbar unter <http://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/Migrationsberichte/migrationsbericht-2016-2017.html?nn=1663558>
- Caro, D. H. & Cortés, D. (2012). Measuring family socioeconomic status: An illustration using data from PIRLS 2006. *Issues and Methodologies in Large-Scale Assessments*, 5, 9–33.
- Davis-Kean, P. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19, 294–304. <https://doi.org/10.1037/0893-3200.19.2.294>
- Ehmke, T. & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 231–254). Münster: Waxmann.
- Ehmke, T. & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. *Zeitschrift für Erziehungswissenschaft*, 8, 521–540. <https://doi.org/10.1007/s11618-005-0157-7>

- Erikson, R. & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives*, 16, 31–44. <https://doi.org/10.1257/089533002760278695>
- Erikson, R., Goldthorpe, J. H. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30, 341–415. <https://doi.org/10.1111/j.1468-4446.2009.01246.x>
- Esser, H. (1990). Familienmigration und Schulkarriere ausländischer Kinder und Jugendlicher. In H. Esser & J. Friedrichs (Hrsg.), *Generation und Identität. Theoretische und empirische Beiträge zur Migrationssoziologie* (S. 127–146). Opladen: Westdeutscher Verlag. http://dx.doi.org/10.1007/978-3-322-91777-5_6
- Ganzeboom, H. B. G., de Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1–56. [https://doi.org/10.1016/0049-089X\(92\)90017-B](https://doi.org/10.1016/0049-089X(92)90017-B)
- Ganzeboom, H. B. G. & Treiman, D. (2012). *International stratification and mobility file: Conversion tools*. Amsterdam: Department of Social Research Methodology. Verfügbar unter <http://www.harryganzeboom.nl/ismf/index.htm>.
- Gebhardt, M., Rauch, D., Mang, J., Sälzer, C. & Stanat, P. (2013). Mathematische Kompetenzen von Schülerinnen und Schülern mit Zuwanderungshintergrund. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 275–308). Münster: Waxmann.
- Hradil, S. (2005). *Soziale Ungleichheit in Deutschland* (8. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-663-11417-8>
- ILO (2012) = International Labour Office. (2012). *ISCO-08 International Standard Classification of Occupations. Structure, group definitions and correspondence tables*. Vol I. ILO: Genf.
- Kierner, K., Haag, N., Müller, K. & Ehmke, T. (2017). Einfluss sozialer und zugewanderungsbezogener Disparitäten, sowie der Klassenkomposition auf die Veränderung der mathematischen Kompetenz von der neunten zur zehnten Klassenstufe. In K. Reiss, E. Klieme, O. Köller & P. Stanat (Hrsg.), *PISA Plus 2012–2013. Kompetenzentwicklung im Verlauf eines Schuljahres*. Zeitschrift für Erziehungswissenschaft, Sonderband 33 (S. 125–149). Wiesbaden: Springer VS. <https://doi.org/10.1007/s11618-017-0753-3>
- Magnuson, K. (2007). Maternal education and children's academic achievement during middle childhood. *Developmental Psychology*, 43, 1497–1512. <http://dx.doi.org/10.1037/0012-1649.43.6.1497>
- Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A. & Reiss, K. (2019). *PISA 2015 Skalendokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Martins, L. & Veiga, P. (2010). Do inequalities in parents' education play an important role in PISA students' mathematics achievement test score disparities? *Economics of Education Review*, 29, 1016–1033. <http://dx.doi.org/10.1016/j.econedurev.2010.05.001>
- 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. 317–347). Münster: Waxmann.
- Nauck, B., Diefenbach, H. & Petri, K. (1998). Intergenerationale Transmission von kulturellem Kapital unter Migrationsbedingungen: Zum Bildungserfolg von Kindern und Jugendlichen aus Migrantenfamilien in Deutschland. *Zeitschrift für Pädagogik*, 44, 701–722.
- OECD. (1999). *Classifying educational programmes: Manual for ISCED-97 implementation in OECD countries*. Verfügbar unter <http://www.oecd.org/dataoecd/7/2/1962350.pdf>.

- OECD. (2016). *PISA 2015 Ergebnisse (Band I): Exzellenz und Chancengerechtigkeit in der Bildung*. Bielefeld: W. Bertelsmann Verlag. <http://dx.doi.org/10.3278/6004573w>
- OECD. (2019). *PISA 2018 results (Volume II): Where all students can succeed*. Paris: OECD.
- R Core Team. (2019). *R: A Language and Environment for Statistical Computing* (Version 3.6.1). <https://doi.org/10.1787/1d0bc92a-en>
- Ramm, G., Prenzel, M., Heidemeier, H. & Walter, O. (2004). Soziokulturelle Herkunft: Migration. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, ... U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 254–272). Münster: Waxmann.
- Rauch, D., Mang, J., Härtig, H. & Haag, N. (2016). Naturwissenschaftliche Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 317–347). Münster: Waxmann.
- Sektnan, M., McClelland, M. M., Acock, A. & Morrison, F. J. (2010). Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Childhood Research Quarterly*, 25, 464–479. <http://dx.doi.org/10.1016/j.ecresq.2010.02.005>
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75, 417–453. <http://dx.doi.org/10.3102/00346543075003417>
- Stanat, P. & Christensen, G. S. (2006). *Where immigrant students succeed. A comparative review of performances and engagement in PISA 2003*. Paris: OECD.
- Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel & S. Wolfgang (Hrsg.), *PISA 2009: Bilanz nach einem Jahrzehnt* (S. 200–230). Münster: Waxmann.
- Størksen, I., Ellingsen, I. T., Wanless, S. B. & McClelland, M. M. (2014). The influence of parental socioeconomic background and gender on self-regulation among 5-year-old children in Norway. *Early Education and Development*, 0, 1–22. <http://dx.doi.org/10.1080/10409289.2014.932238>
- van Buuren, S. & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1–67. Verfügbar unter <https://www.jstatsoft.org/v45/i03/>.
- Weis, M., Mang, J., Baumann, B. & Reiss, K. (2018). Zuwanderung und Erfolg aus Sicht der PISA-Studie: Ein Gesamtüberblick von 2000 bis 2015. In P. Genkova & A. Riecken (Hrsg.), *Handbuch Migration und Erfolg*. Springer, Wiesbaden. https://doi.org/10.1007/978-3-658-18403-2_27-1

7 Vertiefende Analysen zur Umstellung des Modus von Papier auf Computer

Frank Goldhammer, Scott Harrison, Sarah Bürger, Ulf Kroehne, Oliver Lüdtke, Alexander Robitzsch, Olaf Köller, Jörg-Henrik Heine & Julia Mang

In PISA 2015 wurde der Erhebungsmodus von Papier zu Computer umgestellt. Eine nationale Ergänzungsstudie im Rahmen von PISA 2018 hatte entsprechend das Ziel, vertiefende Analysen zu möglichen Unterschieden papierbasierter und computerbasierter Messungen durchzuführen. Im Fokus standen die Vergleichbarkeit des gemessenen Konstrukts und der einzelnen Aufgaben (Items), beispielsweise hinsichtlich ihrer Schwierigkeit. Darüber hinaus wurden die Auswirkungen des Moduswechsels auf die Vergleichbarkeit mit den Ergebnissen früherer PISA-Erhebungen in Deutschland untersucht. Als empirische Basis wurden Daten aus dem PISA-2015-Feldtest genutzt sowie Daten, die im Rahmen der nationalen PISA-Haupterhebung 2018 an einem zweiten Testtag mit papierbasierten Testheften aus PISA 2009 zusätzlich erhoben wurden. Erste Ergebnisse der Ergänzungsstudie liefern Belege für die Konstruktäquivalenz zwischen papier- und computerbasierten Messungen. Zudem weisen die Daten der Ergänzungsstudie darauf hin, dass die computerbasierten Items im Mittel etwas schwieriger sind als die papierbasierten Items. Hinsichtlich der Veränderungen zwischen 2015 und 2018 zeigt sich eine hohe Übereinstimmung von international berichtetem (originalem) und nationalem (marginalem) Trend. Die Veränderungen zwischen 2009 und 2018 fallen für den nationalen Trend, der allein auf papierbasierten Messungen beruht, insgesamt etwas günstiger aus als für den originalen Trend.

Studien wie PISA zeigen durch die wiederkehrenden Erhebungen nicht nur Momentaufnahmen zur Leistungsfähigkeit von Bildungssystemen. Vielmehr generieren sie auch Daten über ihre Entwicklung und geben insbesondere Hinweise, ob nachfolgende Generationen von Fünfzehnjährigen im Vergleich zu früheren Generationen eine höhere oder niedrigere Kompetenz aufweisen. Allerdings müssen Studien eine Reihe von Voraussetzungen erfüllen, damit diese Trendinformationen zuverlässig sind (vgl. Mazzeo & von Davier, 2008). So muss in den Erhebungssituationen sichergestellt werden, dass die jeweils eingesetzten Messinstrumente dasselbe Merkmal (Konstrukt) mit derselben Metrik und mit gleicher Präzision erfassen. Sicherlich wäre es am einfachsten, Testinstrumente gar nicht zu wechseln und sie von Runde zu Runde immer wieder zu verwenden („When measuring change, do not change the measure“, Beaton, 1990, S. 10), aber das ist angesichts einer sich stetig verändernden Welt gerade bei Kompetenztests kaum möglich.

PISA geht dieses Problem grundsätzlich dadurch an, dass zumindest Teilmengen von Items (d. h. einzelne Aufgaben), so genannte Anker- oder Trenditems, wiederholt eingesetzt werden und somit eine Verankerung der Skalen über Erhebungsrunden möglich ist. Diese Trenditems dürfen ihre so genannten Itemparameter wie Schwierigkeit und Trennschärfe über die Zeit nicht verändern. Konkret darf die Lösung dieser Items nicht durch Faktoren beeinflusst werden, die von Erhebung zu Erhebung variieren. In der psychometrischen Literatur (z. B. Meredith, 1993) wird diese Herausforderung im Rahmen von invarianten Messmodellen diskutiert. Für PISA bedeutet dies, dass von Runde zu Runde empirisch geprüft werden muss, ob die Trenditems in ihren Itemparametern stabil sind. Gleichzeitig müssen vor den Erhebungen Maßnahmen beziehungsweise Strategien im Testdesign geplant werden, durch die man sicherstellen kann, dass bei hinreichend vielen dieser Items invariante Itemparameter gegeben sind. Die OECD selbst hat dies nach PISA 2006 als zentrales Problem erkannt, da die Teilnehmerstaaten valide Trendschätzungen wünschen. In der Folge haben Mazzeo und von Davier (2008) in einer von der OECD initiierten Expertise folgende Empfehlungen für eine hohe Präzision bei der Trendschätzung – vor allem basierend auf Erkenntnissen aus dem National Assessment of Educational Progress (NAEP) in den USA – gegeben:

- **Möglichst große Anzahlen von Trenditems:** Dies war angesichts der PISA-Strategie von Haupt- und Nebendomänen bei einer moderaten Anzahl von Testheften (bis 2012 Beschränkung auf 13 Testhefte) nicht einfach umzusetzen. Diese Schwierigkeit bezieht sich vor allem auf die Trenditems der Nebendomänen einer Erhebungsrunde, für die im Vergleich zur Hauptdomäne in einem Testheft weniger Platz zur Verfügung stand.
- **Beachtung – besser Kontrolle – von Kontexteffekten:** Items können ihre Eigenschaften abhängig von den Items ändern, die von den Schülerinnen und Schülern zuvor bearbeitet wurden. In diesem Zusammenhang ist etwa die Variation von Testdomänen (Lesen, Mathematik, Naturwissenschaften) innerhalb einer gemeinsamen Testung kritisch.
- **Beachtung und Kontrolle von Positionseffekten:** Itemparameter (z. B. Schwierigkeit) schwanken in Abhängigkeit von ihrer Position in einem Testheft. Von Runde zu Runde sollten also die Positionen von Itemgruppen (*Testlets*, vgl. Kapitel 10) in den Heften konstant gehalten werden.
- **Beachtung von Testleteffekten:** In PISA wird in der Regel ein relativ langer Aufgabenstamm mit relativ vielen zu lösenden Items präsentiert. Dies erhöht die Gefahr von Itemabhängigkeiten, sodass ein Testlet eher wenige Items enthalten sollte.
- **Hinreichend viele Items mit Kurztext- und geschlossenem Antwortformat (z. B. Multiple-Choice-Items):** Dies sichert im Gegensatz zu offenen Antwortformaten eine höhere Objektivität bei der Kodierung in richtige und falsche Lösungen.
- **Nutzung alternativer Item-Response-Theorie (IRT)-Modelle:** Mit der Verwendung des 2- oder 3-Parameter-Modells im Vergleich zum 1-Parameter-Modell sollte die Gefahr, dass Items in unterschiedlichen Teilnehmerstaaten unterschiedlich ausgeprägte Parameter haben, reduziert werden.

Was diese Maßnahmen allerdings nicht verhindern können ist, dass Items beziehungsweise ihre Parameter epochalen Veränderungen unterliegen können, indem beispielsweise ihre Inhalte über die Zeit aus Curricula und Unterricht genommen werden beziehungsweise Themen ihre Relevanz verlieren. So wurde in PISA 2012 ein Lesetext über die idyllische Natur des Golfs von Mexiko und die dazu formulierten Items nicht verwendet (OECD, 2014), da die Ölpest im Golf im Jahr 2010 ihn unbrauchbar gemacht hatte. Es ist somit selbst bei der Verwendung identischer Testhefte in verschiedenen Erhebungsrounden damit zu rechnen, dass einzelne Items oder ganze Itemgruppen zu den jeweiligen Zeitpunkten nicht identische Parameter aufweisen. Das Spannungsfeld zwischen Fortführung und Änderung eines Messinstrumentes wird weiter dadurch bestimmt, dass mit der Änderung der Rahmenkonzeption einer Hauptdomäne planmäßig neue Items zur vollständigen Konstruktdeckung erforderlich werden. So kamen in PISA 2015 simulationsbasierte Items zur Erfassung der naturwissenschaftlichen Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* hinzu.

7.1 Moduswechsel in PISA

Seit der Expertise von Mazzeo und von Davier (2008) stellt sicherlich die Umstellung von Papier-Bleistift-Tests (*paper-based assessment*, PBA) zu computerbasierten Tests (*computer-based assessment*, CBA) im Jahr 2015 die offensichtlichste Änderung in PISA dar. Entsprechend akut stellte sich die Frage der Trendfortsetzung. Hinzu kam 2018 als weitere wesentliche Änderung am Testdesign das adaptive Testen in der Hauptdomäne (zu den methodischen Grundlagen vgl. Kapitel 10 sowie Mang et al., 2019).

Welche Gründe sprachen dafür, den Modus von PBA auf CBA zu wechseln? Aus diagnostischer Sicht können mit CBA die Messungen in mehrfacher Hinsicht verbessert werden. Beispielsweise werden neue, interaktive und simulationsbasierte Itemformate realisierbar (Parshall, Harmes, Davey & Pashley, 2010). Die Messgenauigkeit lässt sich erhöhen, indem Items orientiert an den Antworten auf vorherige Items adaptiv vorgegeben werden (van der Linden, 2005). Zudem können Log-Daten (z. B. Bearbeitungszeiten) gesammelt werden, die einen Einblick in das individuelle Lösungsverhalten erlauben (Goldhammer & Zehner, 2017; Kroehne & Goldhammer, 2018).

7.1.1 Moduseffekte

Wechselt man den Erhebungsmodus, indem man Items, die bisher in Testheften vorgelegt wurden, auf dem Computer darbietet, dann ändert man Merkmale dieser Items, also etwa bei einem Leseverständnisitem das Format des zu lesenden Textes oder auch das Antwortformat. In der Konsequenz können sich die psychometrischen Eigenschaften der Messung ändern. Solche *Moduseffekte* können nicht nur Schwierigkeiten und Trennschärfen (*Diskriminationen*) einzelner Items betreffen, sondern damit zusammenhän-

gend auch Eigenschaften einer ganzen Skala ändern (z. B. die Metrik, Messgenauigkeit oder Interpretation des Testwertes). Aus Metaanalysen ist bekannt, dass Richtung und Stärke von Moduseffekten von unterschiedlichen Faktoren abhängen können (Kingston, 2008; Randy Elliot et al., 2008; Wang, Jiao, Young, Brooks & Olson, 2007). Moduseffekte können entsprechend als das Ergebnis einer (unbekannten) Mischung von Einzeleffekten geänderter Merkmale der Messung verstanden werden (Kroehne & Martens, 2011). Im Grunde bedeutet das, dass in jeder Studie eine eigene empirische Überprüfung von Moduseffekten zu leisten ist.

Für PISA 2015 wurden im 2014 durchgeführten Feldtest Moduseffekte geprüft (OECD, 2017a), wobei Schülerinnen und Schüler einer Schule zufällig einer Bedingung zugewiesen wurden, in der sie entweder Items auf Papier vorgelegt bekamen oder die gleichen Items am Computer bearbeiteten. Wenn die zufällige Zuweisung erfolgreich umgesetzt wird, entstehen vergleichbare Gruppen und die gemessenen Unterschiede in den Testleistungen zwischen den beiden Gruppen können auf Unterschiede in den Modi (CBA vs. PBA) zurückgeführt werden. Die Analyse der Feldtestdaten aller PISA-2015-Teilnehmerstaaten ergab Moduseffekte bezüglich der Itemschwierigkeiten (vgl. OECD, 2016, Annex A6). Insgesamt erwiesen sich die Items am Computer als schwieriger, wobei einzelne Items mehr oder weniger stark davon betroffen waren. Zusätzliche Auswertungen der deutschen Feldtestdaten (Robitzsch et al., 2017) konnten diesen Anstieg der Schwierigkeiten der CBA-Items auch für Fünfzehnjährige in deutschen Schulen nachweisen. Aus den internationalen Analysen wurde zudem geschlossen, dass sich mit dem Wechsel des Modus das erfasste Konstrukt, was also die Items konkret inhaltlich messen (Testwertinterpretation), nicht geändert hatte.

7.1.2 Umgang mit Moduseffekten

Ein essenzielles Interesse der PISA-Teilnehmerstaaten ist die Fortschreibung des Trends über Erhebungsrunden hinweg sowohl auf internationaler Ebene und auf Grundlage aller Items (*originaler* Trend) als auch auf nationaler Ebene und auf Grundlage immer wieder verwendeter Items (*marginaler* Trend). Beim Moduswechsel in PISA 2015 zu CBA war daher eine zentrale Frage, wie für eine stabile Trendfortsetzung die Vergleichbarkeit mit den PBA-Messungen der vergangenen Erhebungsrunden sichergestellt werden kann (Mazzeo & von Davier, 2008). Moduseffekte stellen hierbei kein prinzipielles Problem dar, da Messungen unter bestimmten Voraussetzungen vergleichbar gemacht werden können (allgemein bezeichnet als *test equating*, Kolen & Brennan, 2014; bzw. im Falle von PISA 2015 *concurrent calibration*, z. B. von Davier et al., 2019). Es ist damit möglich, die gegebenenfalls durch den CBA-Modus geänderten Itemschwierigkeiten zu adjustieren, sodass die Messungen vergleichbar sind. Diese statistischen Verfahren setzen *Konstruktäquivalenz* voraus, was bedeutet, dass PBA- und CBA-Messung dasselbe Konstrukt erfassen beziehungsweise sich die Testwertinterpretation mit dem Moduswechsel nicht ändert.

In PISA 2015 war gemäß den durchgeführten empirischen Analysen (OECD, 2017a) lediglich eine Teilmenge der Items von Moduseffekten betroffen. Die Vergleichbarkeit der Itemschwierigkeiten bei CBA- und PBA-Messung wurde dadurch erreicht, dass sich anhand der CBA-Items ohne Schwierigkeitsänderung gegenüber PBA (*invariante Items*) eine vergleichbare Metrik herstellen ließ. CBA-Items mit Schwierigkeitsänderung (*nicht invariante Items*) durften sich dagegen hinsichtlich ihrer Schwierigkeit von den entsprechenden PBA-Items unterscheiden. Für die Trendbestimmung war zentral, dass in der internationalen Auswertung des Feldtests und Haupttests mit den als invariant angenommenen Items der Einfluss des Modus vollständig berücksichtigt werden konnte. Hinzu kam für die Interpretation des nationalen Trends die in den internationalen Analysen des Feldtests nicht weitergehend untersuchte Fragestellung, ob sich die Moduseffekte zwischen Teilnehmerstaaten unterschieden (d. h., ob es eine sogenannte Interaktion des Modus mit Teilnehmerstaat gab, OECD, 2016). Indirekt wurde dieser Frage bei der Analyse der Haupttestdaten nachgegangen, indem bei fehlender Modellpassung Itemschwierigkeiten spezifisch für Teilnehmerstaaten bestimmt wurden. Für Deutschland deuten die von Robitzsch et al. (2017) vorgenommenen Zusatzanalysen der deutschen PISA-2015-Daten darauf hin, dass Moduseffekte die nationale Trendschätzung trotz der vorgenommenen Korrekturen moderiert haben könnten. Die Analysen der deutschen Daten bei Berücksichtigung der nationalen Feldtestdaten aus 2014 zeigten beispielsweise für die Hauptdomäne Naturwissenschaften, dass je nach Annahmen über den Moduseffekt und das Skalierungsmodell der deutsche Mittelwert zwischen 501 und 531 Punkten betragen konnte. Insgesamt konnten diese Analysen zeigen, dass mit dem Moduswechsel eine gewisse Unsicherheit bei der Trendschätzung einherging (vgl. auch Feskens, Fox & Zwitser, 2019).

7.1.3 Weitere methodische Änderungen

Bei einem Vergleich der Messungen von PISA 2018 beziehungsweise PISA 2015 mit Messungen von PISA 2012 und früheren Erhebungsrunden ist generell zu beachten, dass der Wechsel von PBA nach CBA nur eine von mehreren methodischen Änderungen war, die in PISA 2015 vollzogen und in PISA 2018 fortgeführt wurden (vgl. OECD, 2016, Annex A5 und Tabelle 7.1). Es kann also potentiell weitere Effekte geben, die die Vergleichbarkeit der Messungen beeinflussen.

In PISA 2015 wurde neben dem Modus das Testdesign dahingehend geändert, dass Unterschiede zwischen Haupt- und Nebendomäne in der Administrierung reduziert wurden und insbesondere für die Nebendomänen die Menge an Trenditems erhöht wurde. Des Weiteren wurden im Rahmen der Datenanalyse für die PISA-Studie 2015 die Daten der vorherigen Erhebungsrunden bis einschließlich PISA 2006 zur Bestimmung der Schwierigkeiten der Items einbezogen (im Rahmen einer konkurrenten, d.h. gemeinsamen Itemkalibrierung; OECD, 2017a, S. 149). Mit PISA 2015 wurde zudem ein komplexeres Skalierungsmodell eingeführt, in dem sich Items darin unterscheiden

dürfen, wie gut sie zwischen kompetenten und weniger kompetenten Personen differenzieren (Itemdiskrimination; d. h. 2-Parameter logistisches bzw. 2PL-Modell, vgl. OECD, 2017a, S. 141). Auch die Behandlung von Items, die in verschiedenen Teilnehmerstaaten unterschiedlich ausgeprägte Itemparameter aufweisen (z. B. unterschiedliche Schwierigkeiten; *differential item functioning*, DIF), wurde in PISA 2015 geändert. Bis einschließlich 2012 wurden Items mit DIF, welches auf Übersetzungsfehler oder andere technische Fehler zurückgeführt werden konnte, im jeweiligen Skalierungsmodell für einen Staat ausgeschlossen. Für alle anderen Items wurden trotz des Vorliegens von DIF internationale Itemparameter bestimmt. Seit PISA 2015 werden hingegen für Items mit DIF jeweils staatspezifische Itemschwierigkeiten ermittelt. Schließlich änderte sich auch der Umgang mit Items, die am Ende eines Testhefts nicht mehr rechtzeitig bearbeitet werden konnten (vgl. OECD, 2016, Annex A5, S. 306). Während diese fehlenden Werte bei der Bestimmung der Kompetenz in früheren Erhebungsrunden als falsch gewertet wurden, behandelt man diese Items seit 2015 so, als seien sie nie vorgelegt worden. Die Anzahl der nicht erreichten Items geht außerdem ins Hintergrundmodell ein, sodass bei der Bestimmung des Kompetenzwertes (*Plausible Values*) einer Schülerin beziehungsweise eines Schülers dem potentiellen Zusammenhang zwischen Kompetenz und der Anzahl nicht erreichter Items Rechnung getragen wird.

Was die Änderungen bei der Skalierung betrifft (konkurrente Itemkalibrierung, Skalierungsmodell mit Diskrimination, spezifische Itemparameter bei DIF, Behandlung nicht erreichter Items) zeigte sich in bisherigen Studien kaum ein Einfluss auf relative Unterschiede zwischen Staaten und Trendschätzungen (Jerrim, Parker, et al., 2018; OECD, 2016; Robitzsch et al., 2017).

Tabelle 7.1: Haupt-/Nebendomäne sowie wesentliche Änderungen nach Erhebungsrunde.

	2009	2012	2015	2018
Lesen	Hauptdomäne	Nebendomäne	Nebendomäne – Mehr Trenditems	Hauptdomäne – Neue innovative CBA-Items
Mathematik	Nebendomäne	Hauptdomäne	Nebendomäne – Mehr Trenditems	Nebendomäne
Naturwissenschaften	Nebendomäne	Nebendomäne	Hauptdomäne – Neue innovative CBA-Items	Nebendomäne – Neue Trenditems aus 2015
Methodische Änderungen			– Moduswechsel zu CBA – Konkurrente Itemkalibrierung – Skalierungsmodell mit Diskrimination – Spezifische Itemparameter bei DIF – Behandlung nicht erreichter Items	– Adaptive Testung in Hauptdomäne

Mit PISA 2018 erfolgte gegenüber PISA 2015 eine weitere methodische Umstellung, da die Hauptdomäne Lesen nun adaptiv erfasst wird, das heißt, eine Testperson bekommt kein bestimmtes Testheft mehr vorgelegt, sondern in Abhängigkeit des vorausgehenden Lösungserfolgs mehr oder weniger schwierige Itemgruppen (*Testlets*; nähere Erläuterungen finden sich in Kapitel 10 sowie Mang et al., 2019). Eine weitere Änderung in PISA 2018 betrifft die Testzusammenstellung in den Domänen Lesen und Naturwissenschaften mit neuen, innovativen CBA-Items, womit sich der Kontext bisheriger Trenditems ändert. Die CBA-Items für Lesen in PISA 2018 beinhalten Trenditems, die auch in früheren Erhebungsrunden als Trenditems dienten sowie zusätzlich neue innovative CBA-Items (z. B. Items mit mehreren Textquellen). Die CBA-Trenditems für Naturwissenschaften in PISA 2018 enthalten sowohl innovative CBA-Items (u. a. interaktive Items zur Messung der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren*), die für die Erhebungsrunde 2015 entwickelt wurden (etwa 2/3 der Items), als auch computerisierte PBA-Items aus früheren Erhebungsrunden (etwa 1/3 der Items).

7.2 Nationale Ergänzungsstudie zu Moduseffekten in Deutschland

7.2.1 Hintergrund und Motivation

Die nationale Ergänzungsstudie, die vom Zentrum für internationale Vergleichsstudien (ZIB) in Abstimmung mit der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) durchgeführt wird, hat das Ziel, vertiefende Analysen zur Vergleichbarkeit papierbasierter und computerbasierter Messungen durchzuführen. Untersucht werden Fragen der Vergleichbarkeit des gemessenen Konstrukts (*Konstruktäquivalenz*) und der Vergleichbarkeit der Items (*Invarianz der Itemparameter*). Darüber hinaus werden insbesondere die Auswirkungen des Moduswechsels auf den Trend in Deutschland untersucht. Als empirische Basis werden zum einen Daten aus dem Feldtest von PISA 2015 genutzt. Zum anderen werden Daten einbezogen, die im Rahmen der nationalen PISA-Haupterhebung 2018 mit PBA-Instrumenten zusätzlich erhoben wurden.

7.2.2 Zentrale Forschungsfragen

Modusunterschiede im erfassten Konstrukt

Zentrale Voraussetzung für die Vergleichbarkeit von CBA-Messung und PBA-Messung und somit die Trendbestimmung ist, dass die CBA-Messung das gleiche Konstrukt wie die bisherige PBA-Messung erfasst. Diese Frage der Konstruktäquivalenz soll über die internationalen PISA-2015-Analysen hinausgehend untersucht werden, indem

der Zusammenhang (die Korrelation) zwischen der CBA-Messung und der PBA-Messung berechnet wird (vgl. Kroehne, Buerger, Hahnel & Goldhammer, 2019). Zudem soll geprüft werden, ob der Lösungserfolg in Items der CBA-Messung in gleicher Weise von konstruktrelevanten Itemeigenschaften (z. B. kognitive Anforderung) abhängt wie in der PBA-Messung.

Modusunterschiede in Itemparametern

Die Vergleichbarkeit (Invarianz) der Items wird im Hinblick auf ihre Schwierigkeit betrachtet. Darüber hinaus interessiert der Moduseffekt auf die Itemdiskrimination (Trennschärfe – d.h. das Ausmaß, in dem die einzelnen Items zur Unterscheidung von Schülerinnen und Schülern beitragen), der in PISA 2015 im Rahmen der Modellvergleiche nicht thematisiert wurde (OECD, 2016, Annex A6). Dies betrifft die Frage, ob CBA-Items in gleicher Weise zwischen mehr oder weniger kompetenten Schülerinnen und Schülern unterscheiden wie die PBA-Items (vgl. Buerger, Kroehne, Koehler & Goldhammer, 2019).

Vergleich originaler und marginaler Trends

Je Domäne soll die Frage untersucht werden, inwiefern sich die international berichtete Trendschätzung (*originaler Trend*) zwischen PISA 2009 und PISA 2018 von einer nationalen Trendschätzung (*marginaler Trend*) unterscheidet, die nur auf den Beobachtungen innerhalb einer Nation beruht (Robitzsch et al., 2017). Für die Interpretation der originalen Trendschätzung muss der in PISA 2015 durchgeführte Wechsel des Testmodus von Papier auf Computer berücksichtigt werden. Mit Hilfe der nationalen Ergänzungsstudie ist es hingegen möglich, die nationale Trendschätzung von 2009 nach 2018 ausschließlich auf der Basis papierbasierter Messungen vorzunehmen und damit das methodische Vorgehen möglichst identisch zu PISA 2009 umzusetzen. Dieser Vergleich informiert darüber, wie sich die jeweilige Kompetenz in Deutschland bei vollständiger Konstanzhaltung der Testinstrumente und des Testdesigns über unterschiedliche Kohorten der Fünfzehnjährigen hinweg entwickelt hat.

7.3 Datengrundlage der nationalen Ergänzung in PISA 2018

7.3.1 Untersuchungsdesign und Stichprobe

Um die Moduseffekt- und Trendfragestellungen eingehender zu untersuchen, wurde im Jahr 2018 ein aufwändiges Untersuchungsdesign umgesetzt (siehe Abbildung 7.1), das über das Design des Feldtests in PISA 2015 hinausgeht. An allen 223 teilnehmenden PISA-Schulen wurde ein zweiter Testtag durchgeführt, an dem je Schule bis zu 30 fünfzehnjährige Schülerinnen und Schüler PBA-Instrumente bearbeiteten. In der einen Hälfte

der PISA-Schulen (sog. Typ-A-Schulen) wurden Schülerinnen und Schülern, die am ersten Testtag an der PISA-Testung teilnahmen, am zweiten Testtag zusätzlich Testhefte vorgelegt, das heißt diese Schülerinnen und Schüler nahmen am CBA von PISA 2018 sowie am PBA der nationalen Ergänzung teil. In der anderen Hälfte der PISA-Schulen (sog. Typ-B-Schulen) wurden fünfzehnjährigen Schülerinnen und Schülern, die *nicht* am ersten Testtag an der PISA-Testung teilnahmen, Testhefte vorgelegt, das heißt, diese Schülerinnen und Schüler nahmen nur am PBA der nationalen Ergänzung 2018 teil. Diese Schulstichprobe wurde um zehn kleine Schulen erweitert, da die Realisierung einer Gruppe von weiteren 30 fünfzehnjährigen Schülerinnen und Schülern an kleinen PISA-Schulen nicht möglich war. Für diese Erweiterung wurden die zehn kleinsten Typ-B-Schulen ausgewählt, wobei jeweils die zweite Ersatzschule (vgl. Mang et al., 2019) gezogen wurde. Die Zuweisung einer Schule zu Typ A oder Typ B erfolgte zufällig.

Insgesamt nahmen 111 Typ-A-Schulen mit insgesamt 2361 Schülerinnen und Schülern an der Ergänzungsstudie teil sowie 104 Typ-B-Schulen mit insgesamt 2368 Schülerinnen und Schülern. Nichtteilnahmen von Schulen kamen dadurch zustande, dass auf der Ebene der Schülerinnen und Schüler innerhalb einer Schule der Grenzwert von 50 Prozent bei der Teilnahme nicht erreicht wurde (Typ-A-Schulen: 11, Typ-B-Schulen: 3). Eine Nichtteilnahme von Schülerinnen und Schülern war durch Abwesenheit der Jugendlichen aufgrund von Krankheit oder Ähnlichem bedingt. Eine entsprechende Nichtteilnahme-Adjustierung wurde bei der Bestimmung der Schul- beziehungsweise Schüलगewichte vorgenommen.

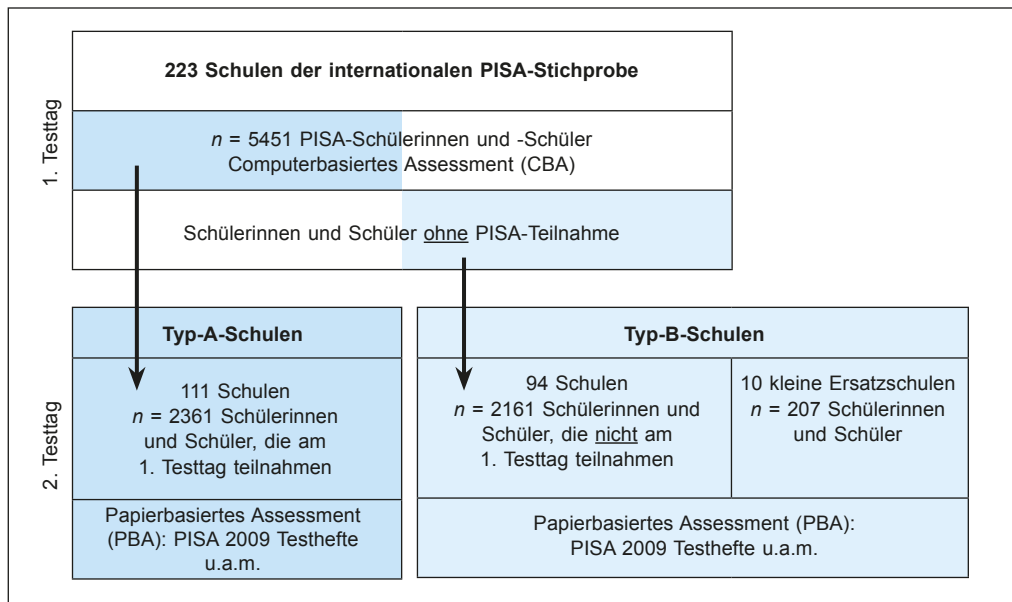


Abbildung 7.1: Untersuchungsdesign der nationalen Ergänzung mit Typ-A- und Typ-B-Schulen (2. Testtag) in Anknüpfung an das Design von PISA 2018 (1. Testtag)

Messungen in den Typ-A-Schulen boten den Mehrwert, die Konstruktäquivalenz anhand des Zusammenhangs (Korrelation) zwischen PBA-Messung und entsprechender CBA-Messung zu prüfen. Messungen in den Typ-B-Schulen dienten dazu, Moduseffekte auf Itemparameter zu untersuchen und die Auswirkungen des Wechsels von PBA zu CBA auf den Trend zu betrachten.

7.3.2 Instrumente

Die eingesetzten papierbasierten Testhefte mit Items zu Lesen, Mathematik und Naturwissenschaften einschließlich der Durchführungsprozeduren entsprachen genau den Testheften von PISA 2009. Die Testhefte dieser Erhebungsrunde wurden ausgewählt, da Lesen im Jahr 2009 wie bei PISA 2018 die Hauptdomäne war. Auf diese Weise lässt sich der reine PBA-Trend zwischen 2009 und 2018 insbesondere für Lesen bestimmen. An Typ-A- und Typ-B-Schulen wurde zudem ein kurzer Fragebogen in Papierform zur Nutzungsmotivation von Computern vorgelegt, sowie an Typ-B-Schulen zusätzlich der PISA 2018 Schülerfragebogen und ein Test zum schlussfolgernden Denken.

An Typ-A-Schulen wurden die PISA-2009-Testhefte den Schülerinnen und Schülern so zugewiesen, dass Items der Mathematik und der Naturwissenschaften nicht wiederholt zu bearbeiten waren. Für Lesen konnten solche Itemwiederholungen nicht vollständig vermieden werden, da Lesekompetenzitems im PISA-Test 2018 adaptiv administriert wurden und somit im Vorhinein nicht feststand, welche Itemgruppen von einer Person zu bearbeiten waren.

Aus dem Testdesign von PISA 2018 folgt, dass die in der nationalen Ergänzung 2018 ermittelten Unterschiede zwischen CBA und PBA (Moduseffekte) nicht nur Differenzen infolge des Modus (PBA vs. CBA), sondern auch infolge des Testdesigns (nicht adaptiv vs. adaptiv) beinhalten. Für die Interpretation der Unterschiede muss zusätzlich berücksichtigt werden, dass in PISA 2018 für Lesen und Naturwissenschaften nicht nur computerisierte (Trend-)Items dargeboten wurden, sondern auch neuartige Itemformate zum Einsatz kamen. Auch wenn diese neuen Items für die Analysen nicht direkt berücksichtigt werden, unterscheidet sich dadurch der Kontext der Items zwischen PBA und CBA.

7.4 Ergebnisse

7.4.1 Modusunterschiede im erfassten Konstrukt

Ein erprobter Test in Papierform, beispielsweise zur Messung von Leseverständnis, sollte nach Übertragung in ein computerbasiertes Format Leseverständnis in äquivalenter Weise messen, das heißt, Testwertunterschiede sollten in gleicher Weise durch Unterschiede im interessierenden Konstrukt Leseverständnis bedingt sein (Konstrukt-

äquivalenz) und nicht durch konstrukt fremde Einflüsse. Ein empirischer Beleg für Konstruktäquivalenz wäre, dass die Korrelation der PBA- und der korrespondierenden CBA-Messung mehr oder weniger perfekt ausfällt beziehungsweise sich die Unterscheidung von Personen oder Gruppen nicht ändert. Eine weitere Evidenz bestünde darin, dass solche Itemeigenschaften, die in der PBA-Messung Herausforderungen beim Lesen darstellen, in gleicher Weise Herausforderungen in der CBA-Messung sind. Wenn die Art der Herausforderungen zwischen PBA-Messung und CBA-Messung unterschiedlich wäre, würden die Messungen unterschiedliche Kompetenzkonstrukte erfassen.

Korrelation zwischen PBA und CBA

Basierend auf Daten der PISA-Studie 2018 und der nationalen Ergänzungsstudie (Typ-A-Schulen; Lesen: $n = 2\,348$, Mathematik: $n = 1\,790$, Naturwissenschaften: $n = 2\,348$) wurde unter Berücksichtigung der geschachtelten Datenstruktur (d. h. Schülerinnen und Schüler in Schulen) und Verwendung der Schülergewichte für jede Domäne ein konfirmatorisches Zwei-Faktoren-Modell mit korrelierenden Faktoren geschätzt. Die beiden Faktoren wurden durch den Erhebungsmodus definiert (Faktor 1: CBA, Faktor 2: PBA). Itemschwierigkeiten und Itemdiskriminationen wurden für diese Analyse zwischen und auch innerhalb der Modi frei geschätzt.

Die ermittelten messfehlerkorrigierten Korrelationen zeigten, dass die CBA-Messung einer Domäne sehr hoch mit der korrespondierenden PBA-Messung korreliert. Für Naturwissenschaften fiel die latente Korrelation mit $r = .96$ am höchsten aus und wich statistisch nicht signifikant von einer perfekten Korrelation von $r = 1.00$ ab. Für Mathematik war die Korrelation mit $r = .90$ etwas niedriger, dicht gefolgt von Lesen mit $r = .89$, wobei die Abweichung von der perfekten Korrelation jeweils signifikant ausfiel. Bei der Bewertung der Höhe der Korrelation ist zu beachten, dass in den Analysen mögliche Effekte der Position von Items und Modus aufgrund des Erhebungsdesigns nicht berücksichtigt werden können, was zu einer Unterschätzung der Korrelation führen kann (Kroehne et al., 2019). Zu einer Unterschätzung können auch individuell unterschiedliche Schwankungen der Kompetenzen vom ersten auf den zweiten Testtag beitragen. Die hier gefundenen hohen signifikanten Zusammenhänge bedeuten, dass die CBA-Messung einer Domäne weitestgehend wie die PBA-Messung Schülerinnen und Schüler differenziert, das heißt, dass die Schülerinnen und Schüler, die in der PBA-Messung am besten waren, auch in der CBA-Messung die besten Leistungen zeigten und Schülerinnen und Schüler mit relativ schwachen Leistungen diese auch bei beiden Messungen zeigten.

Einfluss von Itemeigenschaften in PBA und CBA

Die Validierung der Testwertinterpretation durch die Analyse der Effekte schwierigkeitsgenerierender Itemeigenschaften ist ein in der psychometrischen Literatur übliches Vorgehen (z. B. Embretson, 1983). Im Sinne der Konstruktäquivalenz von PBA und CBA

wird erwartet, dass solche Effekte nicht durch den Modus moderiert werden. Basierend auf den PISA-2015-Feldtestdaten aus Deutschland (PBA: $n = 340$ bis 345 , CBA: $n = 652$ bis 664 , je nach Domäne) wurde untersucht, ob konstruktrelevante Eigenschaften der Testitems deren Schwierigkeit in beiden Modi in gleicher Weise bestimmen. Dazu wurden je Itemeigenschaft linear logistische Testmodelle mit Fehlerkomponente (De Boeck et al., 2011) im Hinblick auf Haupt- und Interaktionseffekte von Itemeigenschaft und Modus getestet. Als zufällige Effekte wurden Interzepte spezifiziert, die über Items, Personen und Schulen variierten, sowie ein zufälliger Effekt für den Modus, der über Items variierte. Damit enthält das Modell die Annahme, dass der Lösungserfolg zusätzlich von Items, Personen und Schulen abhängig ist und der Moduseffekt zwischen Items unterschiedlich stark ausfallen kann.

Im Fall von Leseitems ist nach der Rahmenkonzeption von PISA 2015 beispielsweise die Art der kognitiven Anforderung beziehungsweise der konkrete *Prozess des Lesens* eine konstruktrelevante Itemeigenschaft. Items, bei denen im Text enthaltene Information zu entnehmen ist (*Suchen und Extrahieren*), sind im Durchschnitt einfacher als Items, bei denen über den Text hinausgehend zu reflektieren ist (*Reflektieren und Bewerten*) (OECD, 2017b). Finden sich bei der CBA-Messung diese Schwierigkeitsunterschiede so wieder, wie sie für die PBA-Messung zwischen kognitiven Prozessen gezeigt werden, dann weist dies auf die Messung äquivalenter Konstrukte hin. Im Hinblick auf die in 2015 eingesetzten Leseitems zeigt Abbildung 7.2 für drei Itemeigenschaften (Situation, Textformat und Prozesse des Lesens), wie sich je Modus der erwartete Lösungserfolg zwischen Ausprägungen der jeweiligen Itemeigenschaft unterscheidet. Die horizontale Achse zeigt die Stufen der jeweiligen Itemeigenschaft, die vertikale Achse den erwarteten Lösungserfolg beziehungsweise wie leicht Items mit dieser Stufe im Mittel zu lösen sind (z. B. sind bzgl. *Prozesse des Lesens* die Items mit der Anforderung *Suchen und Extrahieren* am einfachsten). Von Interesse sind hier vor allem die Itemeigenschaften Textformat und Prozesse des Lesens, da sie aus theoretischer Sicht die Itemschwierigkeit beziehungsweise den erwarteten Lösungserfolg eines Leseverständnisitems beeinflussen.

Über alle Ausprägungen von Situation, Textformat und Prozesse des Lesens hinweg erwiesen sich die PBA-Items (gestrichelte Linie) einfacher als die CBA-Items (durchgezogenen Linie). Zentral für die Frage der Konstruktäquivalenz ist, dass das Muster der Schwierigkeitsvariation für CBA weitgehend konsistent zum Muster für PBA ausfiel. Auffällig ist, dass sich das Textformat *Mehrere* (d.h. mehrere unabhängige Texte zum gleichen Inhalt sind zu lesen) bei der CBA-Messung relativ zu den anderen Textformaten als weniger leicht erwies, als dies bei der PBA-Messung der Fall war; zu berücksichtigen ist hier allerdings die kleine Anzahl von Items, was zu recht hoher Unsicherheit führt. Bei den kognitiven Anforderungen fällt auf, dass sich Items zum *Reflektieren und Bewerten* bei der CBA-Messung relativ zu den anderen Aspekten im Vergleich mit der PBA-Messung als weniger schwer erwiesen.

Wie für Lesen zeigte sich für Mathematik und Naturwissenschaften zwischen PBA und CBA insgesamt ein konsistentes Bild, wie Itemeigenschaften die Itemschwierigkeit bestimmen. Zusätzlich ließ sich zeigen, dass der Moduseffekt auf die Itemschwierigkeit

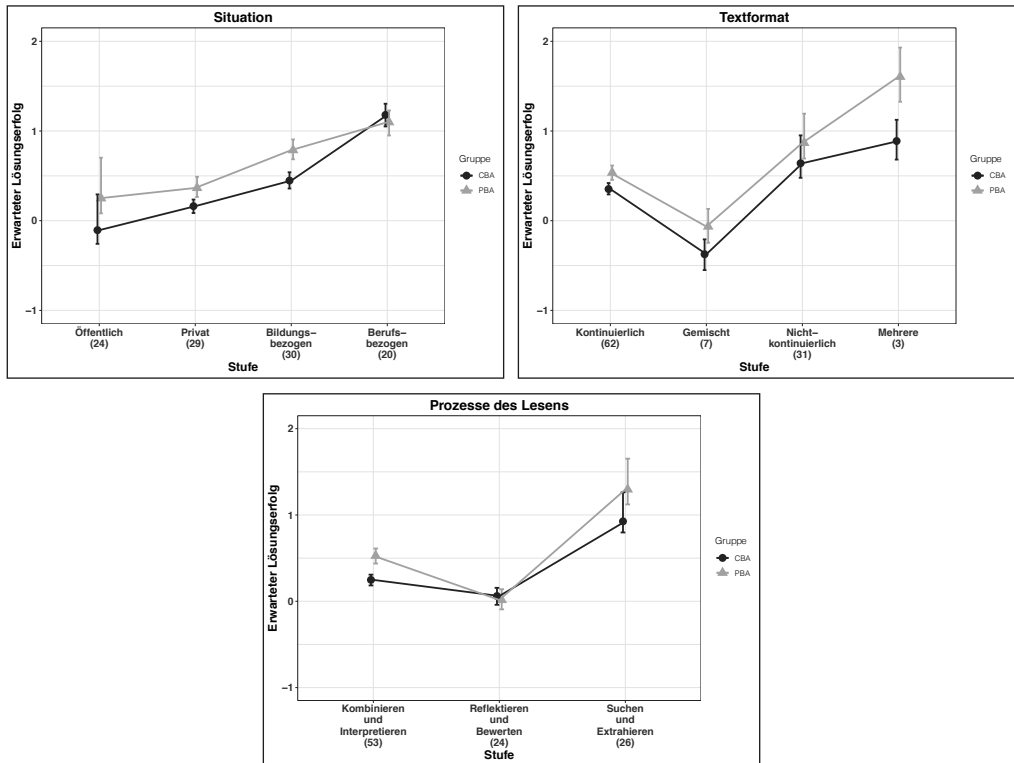


Abbildung 7.2: Erwarteter Lösungserfolg bei Leseitems in Abhängigkeit der Ausprägung einer Itemeigenschaft und des Modus; die jeweilige Anzahl der Items ist in Klammern angegeben

über Items hinweg unterschiedlich stark ausfällt (siehe Abschnitt 7.4.2), aber diese Variation nicht von den betrachteten Itemeigenschaften abhängt, was für die Konstruktäquivalenz der Messungen spricht.

7.4.2 Modusunterschiede in den Itemparametern

Basierend auf den Daten der PISA-Studie 2018 (Lesen: $n = 5451$, Mathematik: $n = 2043$, Naturwissenschaften: $n = 2963$) und der nationalen Ergänzungsstudie (Typ-B-Schulen; Lesen: $n = 2053$, Mathematik: $n = 1425$, Naturwissenschaften: $n = 1423$) konnte in ersten Analysen gezeigt werden, wie sich Itemparameter, das heißt Schwierigkeit und Diskrimination, infolge eines Wechsels von PBA nach CBA ändern. Dazu wurde je Domäne ein Zwei-Gruppen-Modell mit den Gruppen PBA (Typ-B-Schulen) und CBA (PISA-2018-Hauptstudie) unter Berücksichtigung der geschachtelten Datenstruktur und Verwendung der Schüलगewichte getestet. Je Gruppe wurde ein konfirmatorisches Ein-Faktoren-Modell spezifiziert, wobei Itemschwierigkeiten und Itemdiskriminationen

sowohl je Item als auch je Gruppe spezifisch geschätzt wurden.¹ Als abgeleitete Modellparameter wurden in diesen Analysen die Modusunterschiede in den Itemschwierigkeiten und Itemdiskriminationen auf statistische Signifikanz geprüft; es wurde also geprüft, ob sich Itemparameter überzufällig zwischen den Modi unterscheiden.

Abbildung 7.3 zeigt je Domäne die Diskriminationen der CBA-Messung (vertikale Achse) in Relation zu den Diskriminationen der PBA-Messung (horizontale Achse), wobei ein höherer Wert stärkere Diskrimination anzeigt. Jeder Punkt repräsentiert ein Item und der vertikale Abstand zur diagonalen Linie zeigt an, wie stark jeweils die Diskrimination in der CBA-Messung von der in der PBA-Messung abweicht. Liegt ein Item oberhalb der Diagonalen ist es in der CBA-Messung stärker diskriminierend als in der

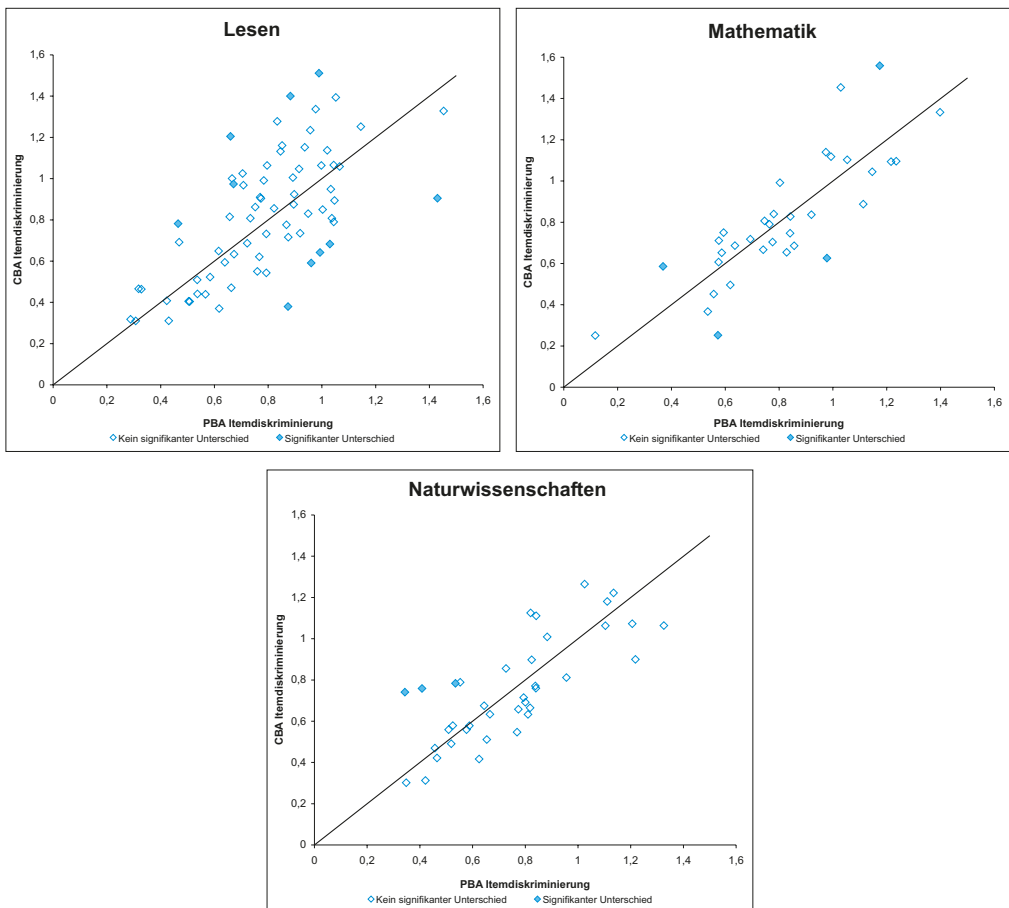


Abbildung 7.3: Zusammenhang der Itemdiskriminationen der CBA-Messung und der PBA-Messung je Domäne. Die gefüllten Punkte zeigen Items mit signifikanten Unterschieden zwischen CBA und PBA an

1 In der gewählten Parametrisierung entsprechen die Faktorladung der Itemdiskrimination und der itemspezifische Schwellenparameter der (mit der Faktorladung gewichteten) Itemschwierigkeit.

PBA-Messung und umgekehrt. Items mit statistisch signifikant unterschiedlichen Diskriminationen sind dabei mit einem gefüllten Punkt repräsentiert. Nur bei einem relativ kleinen Teil von Items unterscheiden sich die Diskriminationen zwischen CBA und PBA (signifikante Unterschiede: 11 von 75 Items in Lesen, 4 von 34 Items in Mathematik, 3 von 38 Items in Naturwissenschaften). Entsprechend hoch fallen die Korrelationen zwischen den PBA- und CBA-Diskriminationen aus ($r = .75$ für Lesen, $r = .83$ für Mathematik und $r = .77$ für Naturwissenschaften).

Abbildung 7.4 stellt entsprechend je Domäne die Schwierigkeiten der CBA-Messung (vertikale Achse) in Relation zu den Schwierigkeiten der PBA-Messung (horizontale Achse) dar, wobei ein höherer Wert eine höhere Schwierigkeit bedeutet. Items mit mehr als zwei Antwortkategorien wurden auf zwei Antwortkategorien reduziert (‚richtig‘ mit voller Punktzahl, andernfalls ‚falsch‘). Konsistent zu Analysen der PISA-2015-Feld-

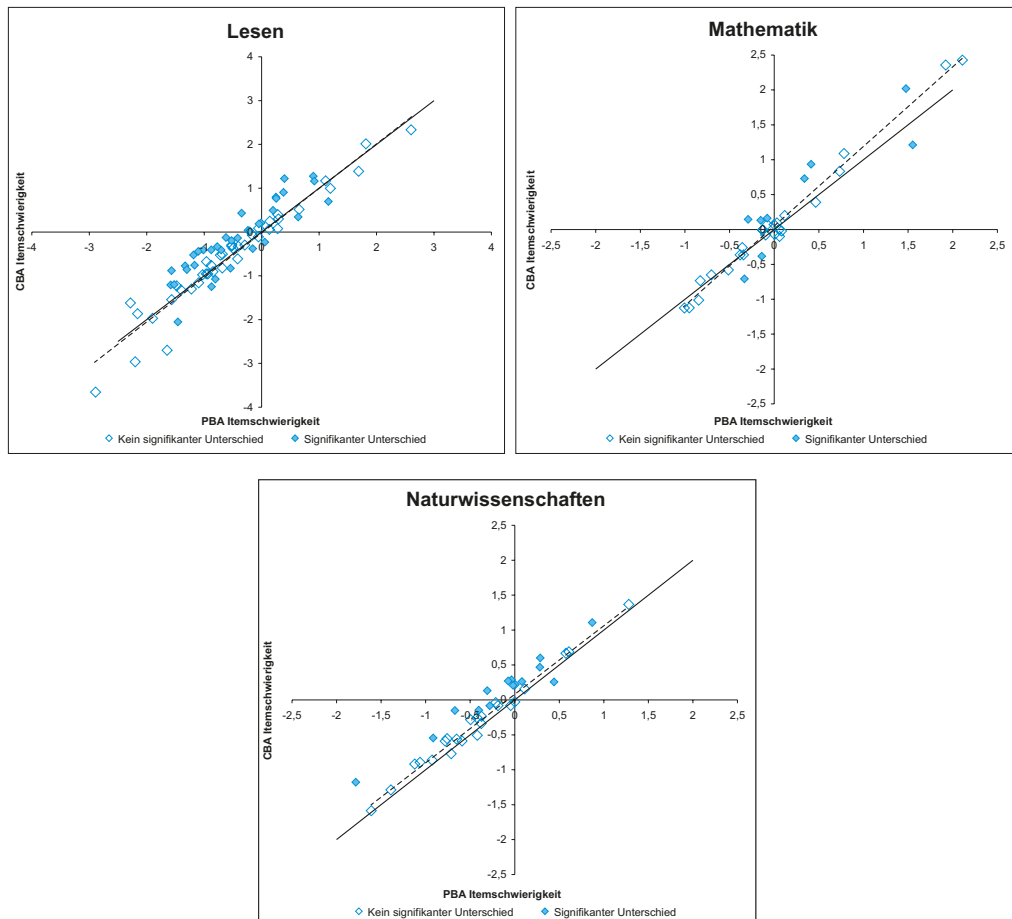


Abbildung 7.4: Zusammenhang der Itemschwierigkeiten der CBA-Messung und der PBA-Messung je Domäne. Die gefüllten Punkte zeigen Items mit signifikanten Unterschieden zwischen CBA und PBA an. Die gestrichelte Linie stellt die Trendlinie der invariant klassifizierten Items dar

testdaten (vgl. OECD, 2016, Annex A6; Robitzsch et al., 2017), erwiesen sich in allen Domänen die CBA-Items eher als schwieriger, wenige Items waren einfacher und ein beträchtlicher Anteil der Items zeigte keinen signifikanten Unterschied (signifikante Unterschiede: 35 von 75 Items bei Lesen, 9 von 34 Items bei Mathematik, 15 von 38 Items bei Naturwissenschaften). Der mittlere Moduseffekt auf die Schwierigkeit war in allen drei Domänen negativ, das heißt, CBA-Items waren im Mittel schwieriger, jedoch war der mittlere Effekt nur bei Naturwissenschaften und Lesen signifikant. Die Korrelationen zwischen den PBA- und CBA-Schwierigkeiten fielen mit $r = .94$ für Lesen, $r = .97$ für Mathematik und $r = .97$ für Naturwissenschaften sehr hoch aus. Wie aus der Signifikanzbeurteilung jedoch ersichtlich ist, unterscheiden sich die Moduseffekte dennoch zwischen den Items (siehe Abschnitt 7.4.1).

Bei der Gruppe der in der Ergänzungsstudie als invariant klassifizierten Items ist für die Kompetenzdomäne Naturwissenschaften zu beachten, dass diese Items, wie durch die Lage der gestrichelten Linie in Abbildung 7.4 angezeigt wird, im Mittel auch eher schwieriger ausfielen. Das bedeutet, dass mit den als invariant angenommenen Items der Einfluss des Modus auf die Itemschwierigkeit nicht vollständig ausgeschlossen werden kann.

7.4.3 Vergleich originaler und marginaler Trends

In der Literatur zu Trendanalysen wird – wie oben bereits erwähnt – der originale Trend von einem marginalen Trend unterschieden (Carstensen, Prenzel & Baumert, 2009; Gebhardt & Adams, 2007; Robitzsch et al., 2017). Die Interpretation des für PISA international berichteten (originalen) Trends ist vor allem dadurch erschwert, dass 2015 ein Wechsel des Testmodus von PBA auf CBA vorgenommen wurde (siehe Tabelle 7.1). Mit Hilfe der nationalen Ergänzungsstudie (Typ-B-Schulen) ist es möglich zu untersuchen, wie sich die nationalen (marginalen) Trends in Deutschland in den drei Domänen bei Konstanzhaltung des Testdesigns von 2009 entwickelt haben. Dies erlaubt es abzuschätzen, inwiefern die auf der internationalen Metrik berichteten Leistungstrends von der Änderung des Testmodus und Testdesigns abhängen.

In der originalen Trendschätzung werden die Mittelwerte für Deutschland aus den querschnittlichen PISA-Berichterstattungen subtrahiert. Die originale Trendschätzung beruht somit auf den mit Hilfe aller Teilnehmerstaaten gewonnenen internationalen Itemparametern. In der marginalen Trendschätzung wird dagegen die Trendschätzung für Deutschland unabhängig von den internationalen Itemparametern nur auf Basis der nationalen Daten vorgenommen. Dies hat zur Folge, dass der marginale Trend für Deutschland nur auf Basis der über die Erhebungen gemeinsam vorgelegten Trenditems bestimmt wird. Im Gegensatz dazu werden in der originalen Trendschätzung auch die nur zu einer Erhebung eingesetzten Items (Nicht-Trenditems) berücksichtigt. Wenn Schülerinnen und Schüler in Deutschland bei den Trenditems (relativ zu den internationalen Itemparametern) bessere oder schlechtere Leistungen zeigen würden als bei den

Nicht-Trenditems, so würde dies zu einem Unterschied zwischen originalelem und marginalem Trend führen (Carstensen et al., 2009). Die marginale Trendschätzung kann deshalb als robuste Alternative zur Bestimmung von nationalen Trends angesehen werden, die im Vergleich zur originalen Trendschätzung weniger sensitiv gegenüber der Menge der Nicht-Trenditems sowie möglicher Item-Staaten-Interaktionen in Itemparametern ist (Robitzsch & Lüdtke, 2019). In den nachfolgenden Analysen wird die Skala der marginalen Trendschätzung so transformiert, dass Mittelwert und Streuung (Standardabweichung) für den ersten Zeitpunkt (PISA 2009) mit denen der originalen Trendschätzung übereinstimmen.² Dadurch wird eine vergleichbare Metrik etabliert und es lassen sich originale und marginale Trendschätzung miteinander vergleichen.

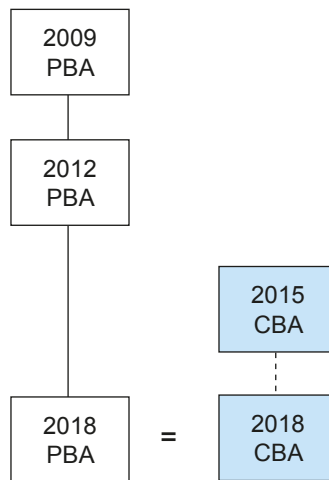


Abbildung 7.5: Schematische Darstellung des Vorgehens bei der marginalen Trendschätzung. Die Stichproben der Typ-B-Schulen der nationalen Ergänzungsstudie 2018 („2018 PBA“) und der internationalen PISA-Studie („2018 CBA“) werden als zufallsäquivalent angesehen

Abbildung 7.5 veranschaulicht die Analysestrategie für die Bestimmung der marginalen Trends von 2009 nach 2018 unter Berücksichtigung der Ergänzungsstudie. Aus den PISA-Erhebungen 2009 und 2012 sowie den Typ-B-Schulen der nationalen Ergänzungsstudie für 2018 kann eine marginale Trendschätzung gewonnen werden, die nur auf papierbasierter Messung beruht und damit die eingangs erwähnte Forderung von Beaton (1990) nach einer Konstanthaltung des Messinstrumentes aufnimmt. Des Weiteren ermöglicht das Untersuchungsdesign der Ergänzungsstudie, marginale Trendschätzungen für mögliche Effekte des Testmodus (Papier vs. Computer) zu adjustieren. Aufgrund der zufälligen Zuweisung kann davon ausgegangen werden, dass die Schülerinnen und Schüler, die 2018 papierbasiert getestet wurden (Stichprobe „2018 PBA“ in Abbildung

2 Als Referenz für die marginale Trendschätzung in einer Domäne wurde der Mittelwert und die Standardabweichung von PISA 2009 verwendet (z. B. für Lesen $M = 497$, $SD = 95$).

7.5), sich nicht systematisch von den Schülerinnen und Schülern unterscheiden, die an der computerbasierten PISA-Erhebung teilgenommen haben (Stichprobe „2018 CBA“). Unterschiede in den Testleistungen zwischen den Stichproben „2018 PBA“ und „2018 CBA“ können somit auf Unterschiede im Testmodus zurückgeführt werden. Die Befunde zu diesen Unterschieden erlauben es zusätzlich, die 2015 erstmalig auf dem Computer durchgeführte Erhebung („2015 CBA“) im Rahmen einer marginalen Trendschätzung für mögliche Moduseffekte zu adjustieren.

Die Schätzung der marginalen Trends wurde mit drei verschiedenen Skalierungsverfahren vorgenommen. Für die technische Durchführung der Skalierung kann zwischen einer separaten Skalierung mit anschließendem Linking und einer konkurrenten Itemkalibrierung unterschieden werden (Kolen & Brennan, 2014). In einer separaten Skalierung wurden die einzelnen Erhebungen zunächst separat skaliert und anschließend die Itemparameter in einem simultanen Linking nach der *Haebara*-Methode (vgl. Kolen & Brennan, 2014) auf eine gemeinsame Metrik transformiert. Dabei wurde sowohl eine Skalierung nach dem 1PL- (Methode L1) als auch dem 2PL-Modell (Methode L2) vorgenommen. Da die PISA-Studie bis 2012 ein 1PL-Modell als Skalierungsmodell verwendete und dann auf ein 2PL-Modell wechselte, schien es angebracht zu überprüfen, wie sensitiv die Trendschätzung gegenüber der Wahl des Skalierungsmodells ist. Zusätzlich wurde eine konkurrente Itemkalibrierung auf Basis des 2PL-Modells vorgenommen, in der die einzelnen Erhebungen als Gruppen in einem IRT-Mehrgruppenmodell behandelt wurden (Methode C2). Für alle drei Methoden wurden für den (computerbasierten) Link von 2015 auf 2018 nur die von der OECD als invariant (zwischen PISA 2015 und 2018) ausgewiesenen Items (Trenditems) verwendet. Alle Skalierungen wurden unter Verwendung der Schülergewichte durchgeführt. Es wurden für jede Domäne eindimensionale Modelle verwendet, wobei nur die Schülerinnen und Schüler einbezogen wurden, denen Items in der jeweiligen Domäne vorgelegt wurden.³

In einem ersten Schritt wurden anhand der Ergebnisse der Typ-B-Schulen der nationalen Ergänzungsstudie für jede Domäne die mittleren Effekte des Testmodus auf die Itemschwierigkeiten (in einem 1PL-Modell) bestimmt. Insgesamt zeigte sich, dass die Items im CBA-Modus im Mittel schwieriger ausfielen (siehe Abschnitt 7.4.2), wobei der Unterschied 15 Punkte in Lesen, 6 Punkte in Mathematik und 12 Punkte in Naturwissenschaften betrug.⁴ Diese Befunde stehen im Einklang mit der Analyse der deutschen Feldtestdaten von PISA 2015, wobei die Unterschiede zwischen CBA- und PBA-Modus in den Feldtestdaten etwas größer ausfielen (Robitzsch et al., 2017). Da die nationale Ergänzungsstudie zu PISA 2018 auf einer erheblich größeren und repräsentativen Stichprobe beruht, wurden diese mittleren Differenzen in den Itemschwierigkeiten in den folgenden Analysen für jede Domäne verwendet, um die marginalen Trendschätzungen für

3 Innerhalb jeder Schule wurden die Schülergewichte pro Domäne so angepasst, dass die Summe der adjustierten Gewichte mit der Summe der Gewichte in der originalen Stichprobe übereinstimmt.

4 Für die Trendanalyse wurden nicht erreichte Items bei PBA gemäß den vorausgehenden PBA-Erhebungsrunden kodiert („falsch“), während für CBA die in PISA 2015 eingeführte Kodiervorschrift verwendet wurde („fehlend“).

den Wechsel des Testmodus in PISA 2015 zu adjustieren (siehe kursiv gedruckte Leistungswerte für PISA 2015 in Tabelle 7.2).

Tabelle 7.2: Originale und marginale Trendschätzungen von 2009 nach 2018 in Lesen, Mathematik und Naturwissenschaften

Jahr	n_{orig}	n_{marg}		orig	marg		
		PBA	CBA		L1	L2	C2
<i>Lesen</i>							
2009	4979	4975	—	497	497*	497*	497*
2012	5001	3470	—	508	502	501	502
2015	6504	—	2746	509	512	505	504
2018	5451	2346	5451	498	491	491	492
<i>Mathematik</i>							
2009	4979	3503	—	513	513*	513*	513*
2012	5001	4971	—	514	516	517	516
2015	6504	—	2739	506	511	513	512
2018	5451	1635	3000	500	507	507	507
<i>Naturwissenschaften</i>							
2009	4979	3477	—	520	520*	520*	520*
2012	5001	3505	—	524	527	528	526
2015	6504	—	6501	509	522	522	523
2018	5451	1622	2969	503	513	515	514

Anmerkung: n_{orig} = Stichprobengrößen für originale Trendschätzung; n_{marg} = Stichprobengrößen für marginale Trendschätzung; orig = Mittelwerte originale Trendschätzung; marg = Mittelwerte marginale Trendschätzung; L1 = Separate Skalierung mit 1PL-Modell und Haebara-Linking; L2 = Separate Skalierung mit 2PL-Modell und Haebara-Linking; C2 = konkurrente Itemkalibrierung mit 2PL-Modell. * = Fixierung auf Mittelwert und Standardabweichung von 2009 für Deutschland. Die Standardfehler für die Mittelwerte von 2009 bis 2018 in der originalen Trendschätzung sowie der Mittelwerte von 2009 bis 2015 in den marginalen Trendschätzungen betragen 2.5 bis 3 Punkte. Die Standardfehler der Mittelwerte in der marginalen Trendschätzung in 2018 betragen 4.5 bis 5 Punkte. Die Link-Fehler für den marginalen Trend von 2009 nach 2018 betragen etwa 2.0 Punkte in Lesen, 3.5 Punkte in Mathematik und 3.5 Punkte in Naturwissenschaften.

Tabelle 7.2 zeigt die Ergebnisse der marginalen Trendschätzungen von 2009 nach 2018 für die drei Domänen. In der Kompetenzdomäne Lesen zeigt sich relativ unabhängig vom Skalierungsverfahren ein Abfall der Testleistungen von 497 Punkten in PISA 2009 auf 491 (Methode L1 und L2) beziehungsweise 492 Punkte (Methode C2) in PISA 2018, der allerdings nicht statistisch signifikant ist. In der international berichteten originalen Trendschätzung besteht praktisch keine Veränderung der mittleren Leseleistung zwischen 2009 (497 Punkte) und 2018 (498 Punkte). Allerdings ist sowohl im international berichteten originalen Trend als auch im nationalen marginalen Trend ein deutlicher Abfall der Leistungswerte von 2012 beziehungsweise 2015 nach 2018 zu verzeichnen, der

den insgesamt positiven Trend von 2009 nach 2012 (bzw. 2009 nach 2015) wieder aufwiegt.

Für Mathematik zeigt sich im nationalen Trend von 2009 nach 2018 konsistent über die drei Skalierungsmethoden ein statistisch nicht signifikanter Rückgang von 6 Punkten in den Leistungswerten. Im originalen Trend besteht ein deutlicher und statistisch signifikanter Abfall von 13 Punkten.

Die Befunde für Naturwissenschaften zeigen im originalen Trend von 2009 nach 2018 eine signifikante negative Veränderung um 17 Punkte, die für den marginalen Trend nicht statistisch signifikant ausfällt (zwischen 5 und 7 Punkten). Zusätzlich wird auch der Unterschied zwischen der originalen und der marginalen Trendschätzung (zwischen 10 und 12 Punkten) statistisch signifikant. Diese Ergebnisse stehen im Einklang mit bereits durchgeführten Analysen des Trends von 2009 nach 2015, in denen der im originalen Trend sichtbare Abfall in Naturwissenschaften nicht für den marginalen Trend auftrat (Robitzsch et al., 2017).

Zusammenfassend zeigt die Abbildung 7.6 die Veränderungen in den drei Domänen zwischen 2009 und 2018 bei Berücksichtigung des originalen und marginalen (C2) Trends. Es wird deutlich, dass über alle drei Domänen im marginalen Trend ein leichter Leistungsabfall von 5 bis 6 Punkten zu beobachten ist. Die Leistungsveränderungen im originalen Trend sind dagegen nicht konsistent über die Domänen ausgeprägt und könnten auf die inhaltlichen sowie methodischen Änderungen zwischen den Erhebungsrunden zurückzuführen sein (siehe Tabelle 7.1).

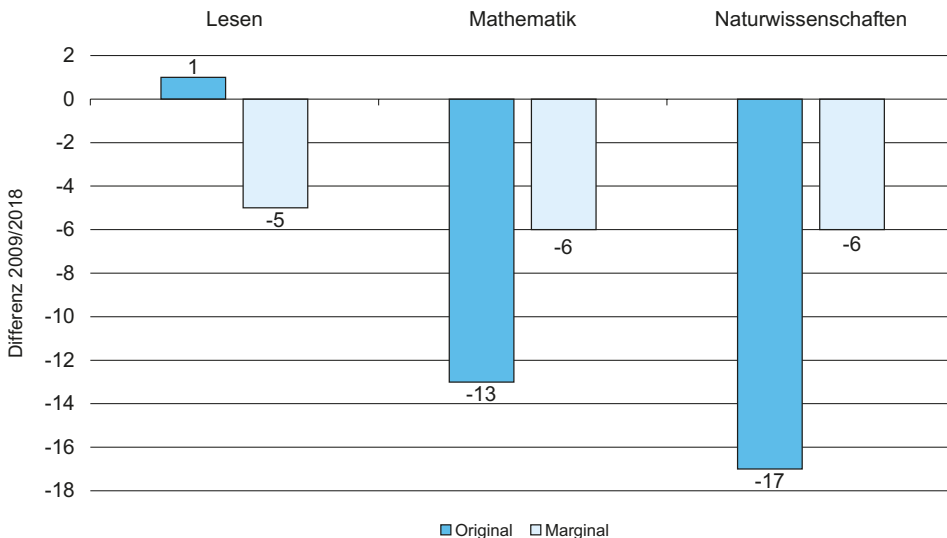


Abbildung 7.6: Originaler und marginaler Trend (C2) der Leistungen in den drei Domänen zwischen 2009 und 2018

7.5 Schlussfolgerungen und Diskussion

Die ersten Ergebnisse aus der nationalen Ergänzungsstudie des ZIB zu Moduseffekten in Deutschland liefern für jede Domäne empirische Belege, dass sich mit dem Wechsel von papierbasiertem Assessment (PBA) zu computerbasiertem Assessment (CBA) das erfasste Konstrukt nicht geändert hat (Konstruktäquivalenz). Dies zeigt sich einerseits an den sehr hohen Korrelationen zwischen CBA- und PBA-Messung als auch an dem zwischen CBA und PBA konsistenten Einfluss von Itemeigenschaften auf die Schwierigkeit. Somit ist eine wichtige Voraussetzung für die Vergleichbarkeit von CBA- und PBA-Messung erfüllt.

Aus den Ergebnissen geht außerdem hervor, dass die Itemdiskriminationen zwischen CBA und PBA etwas schwanken, sich größtenteils aber nicht signifikant unterscheiden. Dagegen fallen die Unterschiede bei der Itemschwierigkeit deutlicher aus und mit der Tendenz, dass CBA-Items im Vergleich zu den PBA-Items schwieriger sind. Für die Naturwissenschaften zeigt sich, dass die in der Ergänzungsstudie als invariant klassifizierten Items (d. h. kein signifikanter Schwierigkeitsunterschied) im Mittel auch etwas schwieriger sind. Dies deckt sich mit früheren Befunden basierend auf den PISA-2015-Feldtestdaten aus Deutschland (Jerrim, Micklewright, Heine, Sälzer & McKeown, 2018; Robitzsch et al., 2017) und bedeutet, dass mit den als invariant angenommenen Items der Einfluss des Modus auf die Schwierigkeit nicht vollständig ausgeschlossen werden kann.

Die durchgeführten marginalen Trendanalysen zeigen für alle drei Domänen leichte Rückgänge in den Leistungen der Schülerinnen und Schüler in Deutschland zwischen 2009 und 2018, die allerdings durchgängig nicht signifikant ausfallen. Anders ist es beim originalen Trend für Mathematik und Naturwissenschaften, der signifikant schwächere Leistungen belegt; im Falle von Lesen wird das Niveau jedoch gehalten. Möglicherweise ist der negative originale Trend in den Naturwissenschaften durch den Umgang der OECD mit dem Wechsel des Testmodus von 2012 nach 2015 begründet, der auch auf Ebene aller OECD-Staaten zu einem mittleren Leistungsabfall von 8 Punkten führte. So zeigen die vorliegenden marginalen Trendschätzungen, die auf Grundlage der nationalen Ergänzungsstudie für den Wechsel des Testmodus adjustieren, einen deutlich geringeren Rückgang in der Leistung in Naturwissenschaften (vgl. Robitzsch et al., 2017). Die Trendanalysen zeigen für die Veränderung von 2015 nach 2018 für alle drei Kompetenzdomänen eine hohe Übereinstimmung zwischen originalem und marginalem Trend (mit Ausnahme von Methode L1 bei Lesen). Damit fällt der jeweilige Leistungsrückgang zwischen den alternativen Trendbestimmungen weitgehend vergleichbar aus. Beim Vergleich von marginalem und originalem Trend ist zu beachten, dass Unterschiede auch dadurch zustande kommen können, dass der marginale Trend nur mit den Trenditems gebildet wird.

Stärken der Ergänzungsstudie 2018 sind die im Vergleich zum Feldtest von PISA 2015 deutlich größere und repräsentative Stichprobe sowie die Variation des Modus innerhalb der Person (an Typ-A-Schulen), was je Kompetenzdomäne die Berechnung

der Korrelation zwischen CBA- und PBA-Messung erlaubte. Die Verwendung der Testhefte aus PISA 2009 ermöglichte die optimale Ermittlung eines PBA-Trends von 2009 nach 2018 insbesondere für Lesen. Bei der Interpretation der Befunde sind jedoch auch einige Einschränkungen zu beachten. An den Typ-A-Schulen war es aufgrund der PISA-Durchführungsstandards nicht möglich, die Reihenfolge des Modus zu variieren. Durch die konstante Reihenfolge, CBA gefolgt von PBA, kann die Bestimmung der Korrelation zwischen CBA- und PBA-Messung durch Effekte des Testtags (d.h. 1. versus 2. Testtag) beeinträchtigt sein. Darüber hinaus sind im Falle von Lesen Effekte der wiederholten Bearbeitung desselben Items nicht auszuschließen. Bei der Stichprobenziehung für die Typ-B-Schulen konnten keine Förderschulen realisiert werden. Um die Definition der PISA-Population beibehalten zu können, wurden ersatzweise Schülerinnen und Schüler der Förderschulen von Typ-A-Schulen verwendet, die bereits am 1. Testtag teilgenommen hatten.

Bei der Interpretation von Moduseffekten ist zu beachten, dass sich die für die PBA-Messung verwendeten Testhefte aus PISA 2009 von der CBA-Messung in PISA 2018 nicht nur im Modus (Papier versus Computer) unterscheiden, sondern auch hinsichtlich der Zusammenstellung der Testitems, da in PISA 2018 Lesen adaptiv getestet wurde sowie bei Lesen und Naturwissenschaften zusätzliche neue, innovative Items den Itemkontext geändert haben (siehe Tabelle 7.1). Für die Frage der Vergleichbarkeit der PISA 2018 CBA-Messung mit der PBA-Messung aus PISA 2009 interessiert jedoch gerade der Gesamteffekt aller Merkmalsunterschiede zwischen PBA und CBA.

Insgesamt zeigen die vorliegenden Befunde, dass eine Veränderung des eingesetzten Testinstrumentes (Modus, Testdesign, Itemkontext) einen Einfluss auf die Trendschätzung haben kann, das heißt, dass Leistungsänderungen unterschiedlich stark ausfallen (vgl. die Änderungen von 2009 nach 2018). Entsprechend sollten quantitative Änderungen im Trend mit Vorsicht interpretiert werden. Nach der Befundlage erweisen sich jedoch die Richtungen der Änderung zwischen den alternativen Methoden der Trendbestimmungen als konsistent.

Literatur

- Beaton, A. E. (1990). Introduction. In A. E. Beaton & R. Zwick (Hrsg.), *The effect of changes in the national assessment: Disentangling the NAEP 1985–86 reading anomaly (Report No. 17-TR-21)* (S. 1–13). Princeton, NJ: National Assessment of Educational Progress.
- Buerger, S., Kroehne, U., Koehler, C. & Goldhammer, F. (2019). What makes the difference? The impact of item properties on mode effects in reading assessments. *Studies in Educational Evaluation*, 62, 1–9. <https://doi.org/10.1016/j.stueduc.2019.04.005>
- Carstensen, C., Prenzel, M. & Baumert, J. (2009). Trendanalysen in PISA: Wie haben sich die Kompetenzen in Deutschland zwischen PISA 2000 und PISA 2006 entwickelt? *Zeitschrift für Erziehungswissenschaft*, 11–34. https://doi.org/10.1007/978-3-531-91815-0_2

- De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., ... Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software*, 39(12), 1–28. <https://doi.org/10.18637/jss.v039.i12>
- Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93(1), 179–197. <https://doi.org/10.1037/0033-2909.93.1.179>
- Feskens, R., Fox, J.-P. & Zwitser, R. (2019). Differential item functioning in PISA due to mode effects. In B. P. Veldkamp & C. Sluijter (Hrsg.), *Theoretical and practical advances in computer-based educational measurement* (S. 231–247). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-18480-3_12
- Gebhardt, E. & Adams, R. J. (2007). The influence of equating methodology on reported trends in PISA. *Journal of Applied Measurement*, 8(3), 305–322.
- Goldhammer, F. & Zehner, F. (2017). What to make of and how to interpret process data. *Measurement: Interdisciplinary Research and Perspectives*, 15(3–4), 128–132. <https://doi.org/10.1080/15366367.2017.1411651>
- Jerrim, J., Micklewright, J., Heine, J.-H., Sälzer, C. & McKeown, C. (2018). PISA 2015: how big is the ‘mode effect’ and what has been done about it? *Oxford Review of Education*, 44. <https://doi.org/10.1080/03054985.2018.1430025>
- Jerrim, J., Parker, P., Choi, A., Chmielewski, A. K., Sälzer, C. & Shure, N. (2018). How robust are cross-country comparisons of PISA scores to the scaling model used? *Educational Measurement: Issues and Practice*, 37(4), 28–39. <https://doi.org/10.1111/emip.12211>
- Kingston, N. M. (2008). Comparability of computer- and paper-administered multiple-choice tests for K–12 populations: A synthesis. *Applied Measurement in Education*, 22(1), 22–37. <https://doi.org/10.1080/08957340802558326>
- Kolen, M. J. & Brennan, R. L. (2014). *Test equating, scaling, and linking*. New York, NY: Springer. <https://doi.org/10.1007/978-1-4939-0317-7>
- Kroehne, U., Buerger, S., Hahnel, C. & Goldhammer, F. (2019). Construct equivalence of PISA reading comprehension measured with paper-based and computer-based assessment. *Educational Measurement: Issues and Practice*, 38(3), 97–111. <https://doi.org/10.1111/emip.12280>
- Kroehne, U. & Goldhammer, F. (2018). How to conceptualize, represent, and analyze log data from technology-based assessments? A generic framework and an application to questionnaire items. *Behaviormetrika*. <https://doi.org/10.1007/s41237-018-0063-y>
- Kroehne, U. & Martens, T. (2011). 11 Computer-based competence tests in the national educational panel study: The challenge of mode effects. *Zeitschrift für Erziehungswissenschaft*, 14(2), 169. <https://doi.org/10.1007/s11618-011-0185-4>
- Mang, J., Wagner, S., Gomolka, J., Schäfer, A., Meinck, S. & Reiss, K. (2019). *Technische Hintergrundinformationen PISA 2018*. München: Technische Universität München.
- Mazzeo, J. & von Davier, M. (2008). Review of the Programme for International Student Assessment (PISA) test design: Recommendations for fostering stability in assessment results. *Education Working Papers EDU/PISA/GB (2008)*, 28, 23–24.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>
- OECD. (2014). *PISA 2012 technical report*. Paris: OECD Publishing.
- OECD. (2016). *PISA 2015 results (Volume I): Excellence and equity in education*. Paris: OECD Publishing.
- OECD. (2017a). *PISA 2015 technical report*. Paris: OECD Publishing.
- OECD. (2017b). *PISA 2018 reading literacy framework*. Paris: OECD Publishing.

- Parshall, C. G., Harmes, J. C., Davey, T. & Pashley, P. J. (2010). Innovative items for computerized testing. In W. J. van der Linden & C. A. W. Glas (Hrsg.), *Elements of adaptive testing* (S. 215–230). New York, NY: Springer. https://doi.org/10.1007/978-0-387-85461-8_11
- Randy Elliot, B., James, B., Andreas, O., Brent, S., Bruce, K. & Fred, Y. (2008). Does it matter if I take my mathematics test on computer? A second empirical study of mode effects in NAEP. *The Journal of Technology, Learning and Assessment*, 6(9). Verfügbar unter <https://ejournals.bc.edu/index.php/jtla/article/view/1639>.
- Robitzsch, A. & Lüdtke, O. (2019). Linking errors in international large-scale assessments: calculation of standard errors for trend estimation. *Assessment in Education: Principles, Policy & Practice*, 26(4), 444–465. <https://doi.org/10.1080/0969594X.2018.1433633>
- Robitzsch, A., Lüdtke, O., Köller, O., Kröhne, U., Goldhammer, F. & Heine, J.-H. (2017). Herausforderungen bei der Schätzung von Trends in Schulleistungsstudien. *Diagnostica*, 63(2), 148–165. <https://doi.org/10.1026/0012-1924/a000177>
- van der Linden, W. J. (2005). *Linear models for optimal test design*. New York: Springer. <https://doi.org/10.1007/0-387-29054-0>
- von Davier, M., Yamamoto, K., Shin, H. J., Chen, H., Khorramdel, L., Weeks, J., ... Kandathil, M. (2019). Evaluating item response theory linking and model fit for data from PISA 2000–2012. *Assessment in Education: Principles, Policy & Practice*, 1–23. <https://doi.org/10.1080/0969594X.2019.1586642>
- Wang, S., Jiao, H., Young, M. J., Brooks, T. & Olson, J. (2007). Comparability of computer-based and paper-and-pencil testing in K–12 reading assessments: A meta-analysis of testing mode effects. *Educational and Psychological Measurement*, 68(1), 5–24. <https://doi.org/10.1177/0013164407305592>

8

Mathematische Kompetenz in PISA 2018 – aktueller Stand und Entwicklung

Frank Reinhold, Kristina Reiss, Jennifer Diedrich, Sarah Hofer & Aiso Heinze

Schülerinnen und Schüler in Deutschland erreichen bei PISA 2018 in Bezug auf ihre mathematische Kompetenz einen Wert von 500 Punkten und liegen damit signifikant über dem OECD-Durchschnitt. Im Vergleich zur PISA-Erhebung 2012, in der Mathematik als Hauptdomäne geprüft wurde, hat sich der Abstand zum OECD-Durchschnitt leicht verringert. Es bestehen weiterhin signifikante Unterschiede in der mathematischen Kompetenz zwischen fünfzehnjährigen Mädchen und Jungen. Der Leistungsvorsprung der Jungen ist im Vergleich zu PISA 2015 zurückgegangen, was auf einen signifikanten Rückgang der mittleren mathematischen Kompetenz der Jungen – im Vergleich zu 2012 – bei weitgehend gleichbleibender Kompetenz der Mädchen zurückzuführen ist. Im Vergleich zu PISA 2012 hat der Anteil leistungsschwacher Jugendlicher, die Kompetenzstufe II nicht erreichen, zugenommen. Mehr als ein Fünftel der Fünfzehnjährigen in Deutschland verfügt lediglich über rudimentäre mathematische Kenntnisse, die mit Problemen an der gesellschaftlichen Teilhabe einhergehen können. Insbesondere an nicht gymnasialen Schularten ist der Anteil dieser Fünfzehnjährigen signifikant größer geworden. Etwa 13 Prozent der Jugendlichen in Deutschland befinden sich auf den höchsten Kompetenzstufen V und VI, ein Anteil, der deutlich über dem OECD-Durchschnitt liegt, auch wenn hier ein Rückgang im Vergleich zu PISA 2012 zu verzeichnen ist.

Ziele und Inhalte des Mathematikunterrichts in Deutschland waren über einen langen Zeitraum hinweg wesentlich durch eine eher fachliche Perspektive geprägt und weniger auf alltägliche Problemstellungen und Anwendungen in Alltagssituationen ausgerichtet (Reiss, Reinhold & Strohmaier, in Druck). Bereits vor der ersten PISA-Erhebungsrunde im Jahr 2000 äußerten sich Wissenschaftlerinnen und Wissenschaftler sowohl aus der Mathematikdidaktik als auch aus der Erziehungswissenschaft diesem Umstand gegenüber kritisch und sprachen sich für einen stärkeren Bezug des Mathematikunterrichts zur Lebensumwelt der Kinder und Jugendlichen, aber auch zur Allgemeinbildung aus (Heymann, 1996; Klafki, 1985; Winter, 1995). Breite öffentliche Aufmerksamkeit erreichte die Debatte um Aufgaben und Ziele des Mathematikunterrichts insbesondere durch die beiden internationalen Vergleichsstudien TIMSS und PISA (Baumert et al., 1997; Blum et al., 2004; Klieme, Neubrand & Lüdtke, 2001), die – entgegen den vor-

herrschenden Erwartungen – Schülerinnen und Schülern in Deutschland mittelmäßige Leistungen attestierten, die signifikant unter dem internationalen Durchschnitt lagen. Das Bildungssystem in Deutschland, das maßgeblich durch Inhalte – also den *Input* – bestimmt war, erwies sich im internationalen Vergleich als weniger effizient und insbesondere auch sozial ungerechter als erwartet.

Durch die anschließende Diskussion (z. B. Klieme et al., 2003) kam es zu einem Paradigmenwechsel, der Basis für die durch die Kultusministerkonferenz beschlossenen und für alle Bundesländer gültigen Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (KMK, 2003) war. Sie charakterisieren die Ziele des Mathematikunterrichts – also den *Output* – durch die explizite Benennung von Kompetenzen, die Schülerinnen und Schüler am Ende ihrer Pflichtschulzeit in Deutschland erreichen sollen. Der Paradigmenwechsel wurde auf schulpraktischer Ebene durch eine neue Aufgabekultur, konkret durch das Einbeziehen kognitiv aktivierender und alltagsbezogener Mathematikaufgaben unterstützt (etwa im SINUS-Programm; Prenzel, Friedrich & Stadler, 2009). Weiter wurde der Umgang mit leistungsheterogenen Gruppen, insbesondere die Förderung leistungsschwacher (KMK, 2010) und leistungsstarker (KMK, 2015) Schülerinnen und Schüler, explizit als Zielsetzung eines zeitgemäßen Mathematikunterrichts in Deutschland formuliert.

Diese Ausrichtung von Unterricht passt in das Konzept der regelmäßig durchgeführten PISA-Erhebungen. Sie haben zum Ziel, das mathematische Kompetenzniveau von Fünfzehnjährigen zu beschreiben und international einzuordnen (OECD, 2019). Dabei wird in PISA unter mathematischer Kompetenz die Fähigkeit verstanden, mathematisch zu denken und mathematische Konzepte, Verfahren, Fakten und Werkzeuge zur Beschreibung, Erklärung und Vorhersage von Phänomenen in für Fünfzehnjährige relevanten Kontexten einzusetzen (OECD, 2019; siehe auch Abschnitt 8.1.1). Die Rahmenkonzeption mathematischer Kompetenz in PISA betont den Anwendungscharakter mathematischen Wissens und umfasst insbesondere – im Einklang mit dem Kompetenzbegriff nach Weinert (2001) – neben kognitiven Elementen auch affektive Elemente, konkret motivationale und emotionale Einstellungen und Verhaltensweisen in Bezug auf Mathematik (OECD, 2013a, 2019; vgl. auch Schiepe-Tiska & Schmidtner, 2013).

PISA bietet die Möglichkeit, Stärken und Schwächen nationaler Bildungssysteme zu identifizieren, zu erklären und gegebenenfalls Handlungsempfehlungen abzuleiten. Dabei geht es nicht nur allgemein um die mathematischen Kompetenzen von Schülerinnen und Schülern, sondern auch speziell um mögliche Disparitäten in Bezug auf das Geschlecht oder die besuchte Schulart sowie die Anteile an leistungsschwachen und leistungsstarken Jugendlichen. Betrachtet man die bisherige Entwicklung der in PISA gemessenen mathematischen Kompetenz von Schülerinnen und Schülern in Deutschland, so lässt sich von 2000 (Klieme et al., 2001) und 2003 (Blum et al., 2004) über 2006 (Frey, Asseburg, Carstensen, Ehmke & Blum, 2007) und 2009 (Frey, Heinze, Mildner, Hochweber & Asseburg, 2010) bis 2012 (Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013) ein positiver Trend feststellen. Insbesondere war zwischen 2003 und 2012 ein signifikanter Anstieg verzeichnet worden. Dieser Trend setzte sich 2015 (Hammer et al.,

2016) nicht in gleicher Weise fort. Es war keine signifikante Veränderung in den Kompetenzen zu beobachten, allerdings lagen sie weiterhin signifikant über dem OECD-Durchschnitt. Das vorliegende Kapitel schließt an diese Untersuchungen an und geht im Detail auf folgende Fragen ein:

- Wo positionieren sich Fünfzehnjährige in Deutschland bezogen auf ihre mathematische Kompetenz im internationalen Vergleich?
- Welche Veränderung der durchschnittlichen mathematischen Kompetenz von Fünfzehnjährigen in Deutschland ist im Vergleich zu PISA 2012 beziehungsweise PISA 2015 zu verzeichnen?
- Welche Kompetenzunterschiede in Mathematik zeigen sich in Deutschland zwischen fünfzehnjährigen Mädchen und Jungen sowie zwischen verschiedenen Schularten?

8.1 Mathematische Kompetenz in PISA 2018

Der nachfolgende Abschnitt gibt einen Überblick über die Rahmenkonzeption mathematischer Kompetenz in PISA 2018, über die Erfassung dieser Kompetenz in der aktuellen Erhebungsrunde sowie über die Kompetenzstufen.

8.1.1 Rahmenkonzeption mathematischer Kompetenz

In PISA 2018 wird mathematische Kompetenz weiterhin im Sinne des Grundbildungskonzeptes (*mathematical literacy*) verstanden als die „Fähigkeit einer Person, Mathematik in vielfältigen Kontexten zu formulieren, anzuwenden und zu interpretieren. Sie beinhaltet außerdem mathematisches Schlussfolgern und die Anwendung mathematischer Konzepte, Prozeduren, Fakten und Werkzeuge, um Phänomene zu beschreiben, zu erklären und vorherzusagen. Mathematische Grundbildung unterstützt Personen zu erkennen und zu verstehen, welche Rolle Mathematik in der Welt spielt, sowie fundierte Urteile und Entscheidungen zu treffen, die den Anforderungen des Lebens dieser Person als konstruktivem, engagiertem und reflektiertem Bürger entsprechen“ (OECD, 2019, S. 75). Zur mathematischen Grundbildung gehört es danach, mathematische Werkzeuge zur Bearbeitung von Problemen in einem realen Kontext anwenden zu können. Aufgaben mit einem rein innermathematischen Kontext ohne Realitätsbezug stehen nicht im Fokus der PISA-Testung. Da in PISA 2018 Mathematik als Nebendomäne geprüft wird, liegt dieser Erhebungsrunde wie bereits bei PISA 2015 die Rahmenkonzeption von PISA 2012 (OECD, 2013a) zugrunde. Insbesondere wird mathematische Grundbildung in den drei Dimensionen *Prozesse*, *Inhalte* und *Kontexte* strukturiert.

Mathematische Prozesse und fundamentale mathematische Fähigkeiten

Aufgaben in PISA 2018 lassen sich jeweils mathematischen Prozessen zuordnen, die die zur Lösung notwendigen mathematischen Aktivitäten beschreiben und konkretisieren (OECD, 2019). So beinhaltet der Prozess *Situationen mathematisch formulieren* die Übersetzung eines Problems aus der realen Welt in ein mathematisches Problem. Dazu ist es notwendig, Möglichkeiten zur Nutzung von Mathematik zu erkennen und darauf aufbauend ein mathematisches Modell für ein Problem bereitstellen zu können. Dies umfasst etwa die mathematische Darstellung einer realen Situation unter Verwendung geeigneter Variablen, Symbole, Diagramme und generischer Modelle. Die Nutzung der Mathematik zur Problemlösung beschreibt der Prozess *mathematische Konzepte, Fakten, Prozeduren und Schlussfolgerungen anwenden*. Dies umfasst etwa arithmetische Berechnungen, das Lösen von Gleichungen, das Extrahieren mathematischer Informationen aus Tabellen und Grafiken, das Darstellen von Formen im Raum sowie die Analyse von Daten. Der Prozess *mathematische Ergebnisse interpretieren, anwenden und bewerten* umfasst schließlich den Umgang mit mathematischen Resultaten. Dazu müssen insbesondere mathematische Ergebnisse zurück in den realen Problemkontext überführt werden, indem sie zu interpretieren, auf Plausibilität zu prüfen und zu nutzen sind.

Den drei mathematischen Prozessen liegen sieben fundamentale mathematische Fähigkeiten (*capabilities*) zugrunde – *Kommunizieren, Mathematisieren, Repräsentieren, Argumentieren, Problemlösestrategien entwickeln, mit Mathematik symbolisch, formal und technisch umgehen* sowie *mathematische Hilfsmittel verwenden* – die zur Lösung der gestellten Aufgaben notwendig sind (OECD, 2019; vgl. auch Sälzer et al., 2013, S. 55 für eine detaillierte Beschreibung des Zusammenhangs von Prozessen und Fähigkeiten). Diese sind von zentralen mathematischen Teilkompetenzen abgeleitet (Niss, 2003; Niss & Højgaard, 2011) und weisen Ähnlichkeit mit den in den Bildungsstandards formulierten allgemeinen mathematischen Kompetenzen auf (KMK, 2003).

Mathematische Inhaltsbereiche

Die in PISA 2018 gestellten Aufgaben können vier Inhaltsbereichen zugeordnet werden (OECD, 2019): Im Fokus des Inhaltsbereiches *Veränderung und Beziehungen* stehen funktionale Zusammenhänge. Dabei geht es etwa um Veränderungen im zeitlichen Verlauf (z.B. Wachstumsprozesse) genauso wie um kausale Zusammenhänge. Inhalte der mathematischen Teilgebiete Algebra und Funktionen können angewendet werden, um diese Prozesse mathematisch zu modellieren oder zwischen symbolischen und ikonischen Repräsentationen funktionaler Zusammenhänge zu übersetzen. Der Inhaltsbereich *Raum und Form* umfasst Phänomene, die in der Umwelt der Schülerinnen und Schüler physikalisch oder visuell präsentiert sind – etwa räumliche Szenarien aus unterschiedlichen Perspektiven. Dabei spielen geometrische Kenntnisse, räumliches Vorstellungsvermögen sowie das Messen eine entscheidende Rolle. Im Bereich *Quantität* sind verschiedene Arten der Quantifizierung von Objekten zu verorten – etwa das Zählen,

Größenordnungsvorstellungen, unterschiedliche Repräsentationen von Zahlen, Arithmetik sowie der Umgang mit Anteilen. Der Inhaltsbereich *Unsicherheit und Daten* umfasst schließlich Situationen, in denen statistische Phänomene oder Wahrscheinlichkeiten im Fokus stehen – etwa wissenschaftliche Vorhersagen oder wirtschaftliche Modelle. Grundlegende statistische Kenntnisse können angewendet werden, um Abweichungen in Messungen zu interpretieren oder Schlussfolgerungen in Situationen zu ziehen, in denen die Datengrundlage gewisse Variationen aufweist.

Die Inhaltsbereiche kategorisieren die in PISA adressierten mathematischen Konzepte, Aussagen und Verfahren und können als Schnittstelle zu den Schulcurricula gesehen werden. Es ist aber zu erwähnen, dass der *literacy*-basierte PISA-Test keine curriculare Validität im Sinne einer Überprüfung etwas des Stoffes der neunten Jahrgangsstufe anstrebt, sondern Schülerinnen und Schüler mit für sie zum Teil neuen oder ungewohnten Aufgabenstellungen konfrontiert, die Kreativität und Transferfähigkeiten erfordern. Unbestritten ist aber, dass die fachlichen Voraussetzungen zur Lösung der Aufgaben Ergebnisse schulischer Lerngelegenheiten sind.

Kontexte

Zur Passung der Aufgaben an die Lebensumwelt von Fünfzehnjährigen werden die PISA-Aufgaben einem von vier für Jugendliche relevanten Kontexten zugeordnet, die ein möglichst breites Spektrum unterschiedlicher Interessensgebiete abdecken sollen (OECD, 2019): Aufgaben mit einem *persönlichen* Kontext fokussieren auf Aktivitäten die Schülerinnen und Schüler selbst sowie ihre Familien betreffend, also etwa Reisen oder Sport. *Berufliche* Kontexte umfassen Situationen aus der Arbeitswelt wie etwa Qualitätskontrollen oder Rechnungswesen. *Gesellschaftliche* Kontexte zielen auf Probleme der Allgemeinheit ab wie etwa Demografie oder öffentlicher Personenverkehr. Aufgaben mit einem *wissenschaftlichen* Kontext können beispielsweise das Klima oder medizinische Probleme wie die Genetik betreffen.

8.1.2 Erfassung mathematischer Kompetenz und Kompetenzstufen

Zur Erfassung der mathematischen Kompetenz in PISA 2018 wurden Aufgaben aus vorhergehenden Erhebungsrunden verwendet, da Mathematik in dieser Erhebungsrunde als Nebendomäne geprüft wurde. Eine Aufgabeneinheit besteht aus unterschiedlichen Aufgaben, die auf einem gemeinsamen Stimulus aufbauen. Die Aufgaben beinhalten zu etwa gleichen Teilen alle drei Prozesse, alle vier Inhalte sowie alle vier Kontexte sowie unterschiedliche Schwierigkeitsgrade.

PISA verwendet eine gemeinsame Skalierung von Personenfähigkeiten und Aufgabenschwierigkeiten (vgl. Kapitel 10). Dies erlaubt eine kriteriumsorientierte Interpretation der erreichten mathematischen Kompetenz im Kontext der Anforderungen, die für Aufgaben entsprechender Kompetenzstufe typisch sind (siehe Tabelle 8.1). Diese Skala

Tabelle 8.1: Stufen mathematischer Kompetenz in PISA

Kompetenzstufe	Wozu die Schülerinnen und Schüler auf der jeweiligen Kompetenzstufe im Allgemeinen in der Lage sind
VI ≥ 669 Punkte	Schülerinnen und Schüler auf dieser Stufe können Informationen, die sie aus der Untersuchung und Modellierung komplexer Problemsituationen erhalten, konzeptualisieren, verallgemeinern und auf neue Situationen anwenden. Sie können verschiedene Informationsquellen und Darstellungen miteinander verknüpfen und flexibel zwischen diesen hin und her wechseln. Schülerinnen und Schüler auf dieser Stufe besitzen die Fähigkeit zu anspruchsvollem mathematischem Denken und Argumentieren. Sie können dieses mathematische Verständnis und ihre Beherrschung symbolischer und formaler mathematischer Operationen und Beziehungen nutzen, um Ansätze und Strategien zum Umgang mit neuartigen Problemsituationen zu entwickeln. Schülerinnen und Schüler auf dieser Stufe können ihr Tun und ihre Überlegungen, die zu ihren Erkenntnissen, Interpretationen und Argumentationen geführt haben, präzise beschreiben und kommunizieren, einschließlich der Beurteilung von deren Angemessenheit für die jeweilige Ausgangssituation.
V 607–668 Punkte	Schülerinnen und Schüler auf dieser Stufe können Modelle für komplexe Situationen konzipieren und mit ihnen arbeiten, einschränkende Bedingungen identifizieren und Annahmen spezifizieren. Sie können im Zusammenhang mit diesen Modellen geeignete Strategien für die Lösung komplexer Probleme auswählen, sie miteinander vergleichen und bewerten. Schülerinnen und Schüler auf dieser Stufe können strategisch vorgehen, indem sie sich auf breit gefächerte, gut entwickelte Denk- und Argumentationsfähigkeiten, passende Darstellungen, symbolische und formale Beschreibungen und für diese Situationen relevante Einsichten stützen. Sie sind imstande, über ihr Tun zu reflektieren und ihre Interpretationen und Überlegungen zu formulieren und zu kommunizieren.
IV 545–606 Punkte	Schülerinnen und Schüler auf dieser Stufe können effektiv mit expliziten Modellen komplexer konkreter Situationen arbeiten, auch wenn sie einschränkende Bedingungen enthalten oder die Aufstellung von Annahmen erfordern. Sie können verschiedene Darstellungsformen, darunter auch symbolische, auswählen und zusammenführen, indem sie sie direkt zu Aspekten von Realsituationen in Beziehung setzen. Schülerinnen und Schüler auf dieser Stufe können in diesen Kontexten gut ausgebildete Fertigkeiten anwenden und mit einem gewissen mathematischen Verständnis flexibel argumentieren. Sie können Erklärungen und Begründungen für ihre Interpretationen, Argumentationen und Handlungen geben und sie anderen mitteilen.
III 483–544 Punkte	Schülerinnen und Schüler auf dieser Stufe können klar beschriebene Verfahren durchführen, auch solche, die sequenzielle Entscheidungen erfordern. Sie können einfache Problemlösungsstrategien auswählen und anwenden. Schülerinnen und Schüler auf dieser Stufe können Darstellungen interpretieren und nutzen, die aus verschiedenen Informationsquellen stammen, und hieraus unmittelbare Schlüsse ableiten. Sie können kurze Berichte zu ihren Interpretationen, Ergebnissen und Überlegungen geben.
II 421–482 Punkte	Schülerinnen und Schüler auf dieser Stufe können Situationen in Kontexten interpretieren und erkennen, die nicht mehr als direkte Schlussfolgerungen erfordern. Sie können relevante Informationen einer einzigen Quelle entnehmen und eine einzige Darstellungsform benutzen. Schülerinnen und Schüler auf dieser Stufe können elementare Algorithmen, Formeln, Verfahren oder Regeln anwenden. Sie sind zu direkten Schlussfolgerungen und wörtlichen Interpretationen der Ergebnisse imstande.
I 358–420 Punkte	Schülerinnen und Schüler auf dieser Stufe können auf Fragen zu vertrauten Kontexten antworten, bei denen alle relevanten Informationen gegeben und die Fragen klar definiert sind. Sie können Informationen identifizieren und Routineverfahren gemäß direkter Instruktionen in expliziten Situationen anwenden. Sie können Handlungen ausführen, die klar ersichtlich sind und sich unmittelbar aus den jeweiligen Situationen ergeben.
unter I ≤ 357 Punkte	

wurde 2003 normiert ($M = 500$, $SD = 100$, vgl. OECD, 2003) und besteht seitdem unverändert. Zur Interpretation dienen sechs Kompetenzstufen, die jeweils 62 Punkte auf der Skala umfassen, wobei Kompetenzstufe VI lediglich nach unten beschränkt ist.

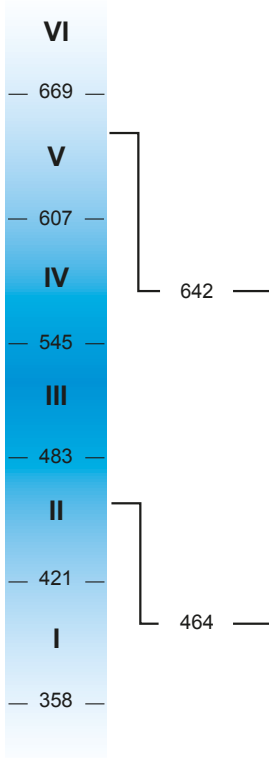
Von besonderem Interesse sind Schülerinnen und Schüler mit sehr geringen Kompetenzen (Kompetenzstufe I oder darunter) oder sehr hohen Kompetenzen (Kompetenzstufen V und VI). Erreichen Fünfzehnjährige nicht die Kompetenzstufe II, so kann davon ausgegangen werden, dass sie bestenfalls über grundlegende mathematische Fähigkeiten verfügen, die mit erhöhtem Risiko erschwerter Bedingungen in Ausbildung, Beruf und gesellschaftlicher Teilhabe einhergehen. Verfügen Fünfzehnjährige über eine mathematische Kompetenz der Stufe V oder höher, so sollten sie über sehr elaborierte mathematische Fähigkeiten verfügen, die eine ausgezeichnete Voraussetzung für das Leben in einer hochtechnologisierten Informations- und Kommunikationsgesellschaft sind.

Als Illustration der Kompetenzstufen können Beispielaufgaben dienen. An dieser Stelle werden zwei Aufgaben der Aufgabeneinheit *Bergsteigen am Mount Fuji* (Abbildung 8.1, vgl. OECD, 2013b) dargestellt. Weitere ausführliche Aufgabenbeispiele können dem Kapitel zur mathematischen Kompetenz im PISA-Berichtsband 2012 (Sälzer et al., 2013, S. 60ff.) entnommen werden.

Aufgabe 1 stellt mit einer Schwierigkeit von 464 Punkten eine tendenziell einfache Aufgabe da, die der Kompetenzstufe II zugeordnet ist. In einem geschlossenen Antwortformat soll die mittlere Zahl an Personen bestimmt werden, die den Berg pro Tag besteigen. Die für die Berechnung notwendigen Informationen müssen dem Text entnommen werden, sodass die Aufgabe dem Prozess *Situationen mathematisch formulieren* und dem Inhaltsbereich *Quantität* zugeordnet wird. Aufgabe 2 stellt mit einer Schwierigkeit von 642 Punkten eine vergleichsweise schwierige Aufgabe da, die der Kompetenzstufe V zugeordnet ist. In einem offenen Antwortformat soll der Zeitpunkt bestimmt werden, zu dem der Protagonist Toshi spätestens aufbrechen muss, um Auf- und Abstieg bis 20 Uhr zu meistern. Auch hier müssen die Informationen dem Text entnommen und zunächst mathematisiert werden. Erneut steht der Prozess *Situationen mathematisch formulieren* im Vordergrund und die Aufgabe gehört zum Inhaltsbereich *Veränderungen und Beziehungen*. Zudem wird konzeptuelles Verständnis des Geschwindigkeitsbegriffes benötigt, um die Aufgabe zu lösen. Die Zeitspannen von sechs Stunden für den Aufstieg und drei Stunden für den Abstieg führen zur Antwort 11 Uhr (vgl. auch Hammer et al., 2016, S. 227ff.).

Bergsteigen am Mount Fuji

Der Mount Fuji ist ein berühmter schlafender Vulkan in Japan.



Aufgabe 2:
 Der Gotemba-Wanderweg auf den Mount Fuji hinauf ist ungefähr 9 Kilometer (km) lang. Die Wanderer müssen von der 18 km langen Wanderung bis 20:00 Uhr zurück sein. Toshi schätzt, dass er den Berg mit durchschnittlich 1,5 Kilometer pro Stunde hinaufsteigen kann und mit der doppelten Geschwindigkeit absteigen kann. Diese Geschwindigkeiten berücksichtigen Essens- und Ruhepausen.
 Wenn man Toshis geschätzte Geschwindigkeiten zugrunde legt: Wann muss er seine Wanderung spätestens beginnen, damit er bis 20:00 Uhr zurück ist?

Aufgabe 1:
 Der Mount Fuji ist für die Öffentlichkeit jedes Jahr nur vom 1. Juli bis 27. August zur Besteigung freigegeben. Ungefähr 200.000 Menschen besteigen den Mount Fuji während dieser Zeit. Wie viele Menschen besteigen den Mount Fuji durchschnittlich pro Tag?
 A. 340
 B. 710
 C. 3 400
 D. 7 100
 E. 7 400

Abbildung 8.1: Beispielaufgaben der Aufgabeneinheit „Bergsteigen am Mount Fuji“ (mit Genehmigung entnommen aus Hammer et al., 2016, S. 228)

8.2 Mathematische Kompetenz im internationalen Vergleich

In diesem Abschnitt wird die mathematische Kompetenz von Fünfzehnjährigen aus Deutschland im internationalen Vergleich betrachtet. Es werden dabei auch die jeweiligen Anteile von Schülerinnen und Schüler auf den verschiedenen Kompetenzstufen, Geschlechterunterschiede sowie Trendentwicklungen berichtet.

8.2.1 Mathematische Kompetenz: Mittelwerte und Streuung

Der Mittelwert (M) der mathematischen Kompetenz der OECD-Staaten in PISA 2018 beträgt 489 Punkten, die Standardabweichung (SD) liegt bei 91 Punkten. Schülerinnen und Schüler in Deutschland erreichen mit 500 Punkten weiterhin ein Ergebnis, das signifikant über dem OECD-Durchschnitt liegt. Abbildung 8.2 stellt mit Mittelwerten, Standardabweichungen und Perzentilbändern der mathematischen Kompetenz die zentralen Befunde – die OECD-Teilnehmerstaaten betreffend – dar. Die Mittelwerte können dabei als Maß für die mittlere mathematische Kompetenz in einem Staat verstanden werden, die Standardabweichungen als Maß für die Streuung. Die Perzentilbänder stellen zudem graphisch Informationen über die Streuung sowie die Verteilung innerhalb eines Staates dar. In Abbildung 8.2 werden die OECD-Staaten in drei Gruppen gegliedert: Staaten, deren Mittelwerte signifikant über dem OECD-Durchschnitt liegen, Staaten, die sich nicht signifikant vom OECD-Durchschnitt unterscheiden, sowie Staaten, deren Mittelwerte signifikant unter dem OECD-Durchschnitt liegen.

In 22 OECD-Staaten – darunter auch Deutschland – erreichen Fünfzehnjährige eine durchschnittliche mathematische Kompetenz, die signifikant über dem OECD-Mittelwert liegt. Diese Gruppe wird angeführt von Japan (527 Punkte), der Republik Korea (526 Punkte), Estland (523 Punkte) und den Niederlanden (519 Punkte). Frankreich (495 Punkte), Island (495 Punkte) und Neuseeland (494 Punkte) grenzen diese Gruppe nach unten zum Durchschnitt ab. In dieser Spitzengruppe liegt die mittlere mathematische Kompetenz etwa ein Drittel einer Kompetenzstufe über dem OECD-Mittelwert. Bei vier OECD-Staaten – darunter Italien (487 Punkte) und Australien (491 Punkte) – unterscheidet sich der Mittelwert nicht vom OECD-Mittelwert. In elf OECD-Staaten wird ein Mittelwert erreicht, der signifikant unter dem OECD-Mittelwert liegt. Das Kompetenzniveau dieser Staaten reicht von Luxemburg (483 Punkte) über Chile (417 Punkte), Mexiko (409 Punkte) bis Kolumbien (391 Punkte), wobei Chile, Mexiko und Kolumbien jeweils einen Unterschied von über einer Kompetenzstufe zum OECD-Mittelwert aufweisen.

Elf OECD-Staaten – darunter Finnland, Slowenien, die Schweiz, Polen und die Niederlande – erreichen signifikant höhere Punktzahlen als Deutschland. Zehn OECD-Staaten – darunter Frankreich, Österreich, Norwegen und Schweden – unterscheiden sich in ihren Mittelwerten nicht von Deutschland, während 15 OECD-Staaten – darunter Griechenland, die Türkei, die Vereinigten Staaten, Italien, Australien und Kolumbien – signifikant niedrigere mittlere Kompetenzwerte als Deutschland aufweisen.

Erweitert man den Fokus, so lässt sich festhalten, dass sowohl die höchsten als auch die niedrigsten mittleren mathematischen Kompetenzen nicht in OECD-Staaten selbst, sondern außerhalb der OECD erreicht werden (siehe Tabelle 8.1web). Spitzenwerte erzielen B-S-J-Z (China)¹ mit 591 Punkten und Singapur mit 569 Punkten, also einem

1 B-S-J-Z (China) bezieht sich auf vier Verwaltungseinheiten Chinas, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Zhejiang.

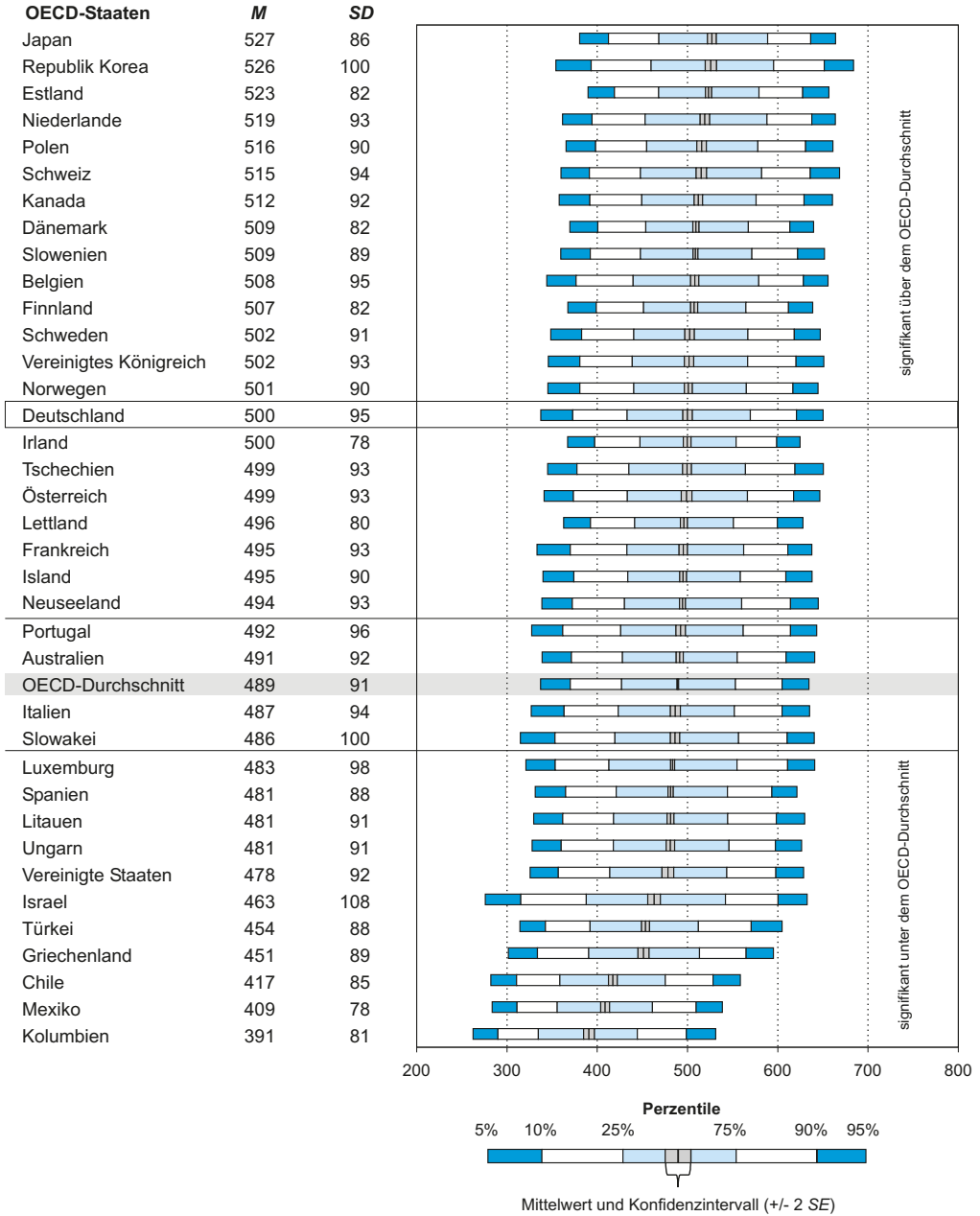


Abbildung 8.2: Perzentilbänder mathematischer Kompetenz in den OECD-Staaten

Unterschied von über einer Kompetenzstufe zu Deutschland und zum OECD-Mittelwert.

Maße für die Streuung der mathematischen Kompetenz um den Mittelwert sind die Breite der Perzentilbänder sowie die Standardabweichung. Dabei geht eine große Streuung mit einem großen Unterschied in der mathematischen Kompetenz leistungsschwacher und leistungsstarker fünfzehnjähriger Schülerinnen und Schüler einher und kann als Indikator dafür interpretiert werden, dass entsprechenden Bildungssystemen eine gleichmäßige Förderung leistungsschwacher und leistungsstarker Jugendlicher nur unzureichend gelingt. Mit Blick auf die Perzentilbänder zeigt sich, dass sich die OECD-Staaten nicht nur in der erreichten mittleren mathematischen Kompetenz, sondern auch in der Streuung dieser Kompetenz unterscheiden (Abbildung 8.2). Mit einer Standardabweichung von 95 Punkten weist Deutschland eine signifikant höhere Streuung auf als der OECD-Durchschnitt ($SD = 91$). Deutschland gehört – mit der Schweiz, Belgien, Portugal, Luxemburg, der Slowakei, der Republik Korea und Israel – zu einer Gruppe von OECD-Staaten mit einer überdurchschnittlich hohen Leistungsheterogenität. Nur Israel ($SD = 108$) sowie die Republik Korea ($SD = 100$) weisen eine signifikant höhere Streuung auf als Deutschland. Bemerkenswert ist, dass sechs OECD-Staaten – darunter Estland, Finnland, Dänemark und Japan – eine vergleichsweise homogene Kompetenzverteilung ihrer Fünfzehnjährigen aufweisen und sie gleichzeitig überdurchschnittliche Kompetenzwerte – insbesondere höhere als Deutschland – erreichen. In diesen Bildungssystemen gelingt demnach eine Förderung der leistungsschwachen Schülerinnen und Schüler, ohne die Förderung der Leistungsspitze zu vernachlässigen.

8.2.2 Verteilung auf Kompetenzstufen

Die in PISA verwendete gemeinsame Skalierung von Personenfähigkeiten und Aufgabenschwierigkeiten lässt eine kriteriale Interpretation der erreichten mathematischen Kompetenz im Kontext konkreter Aufgabenanforderungen zu. Von besonderem Interesse sind hierbei diejenigen Gruppen von Schülerinnen und Schülern, die lediglich über rudimentäre mathematische Kompetenzen verfügen (unterste Kompetenzstufen, d. h. Kompetenzstufe I oder darunter) sowie die mathematisch besonders leistungsstarken Fünfzehnjährigen (oberste Kompetenzstufen, d. h. Kompetenzstufe V oder VI). Die Anteile der Schülerinnen und Schüler auf den untersten sowie den obersten Kompetenzstufen sind für die OECD-Staaten in Abbildung 8.3 dargestellt.

Im Durchschnitt der OECD-Staaten erreichen 23 Prozent der Fünfzehnjährigen lediglich die untersten Kompetenzstufen. In Deutschland beträgt dieser Anteil 21 Prozent und ist damit zwar geringer, unterscheidet sich aber nicht signifikant vom OECD-Durchschnitt. Unter den OECD-Staaten weisen Estland (10 Prozent) und Japan (12 Prozent) einen deutlich geringeren Anteil an Schülerinnen und Schülern der untersten Kompetenzstufen auf. In vier OECD-Partnerstaaten, nämlich B-S-J-Z (China),

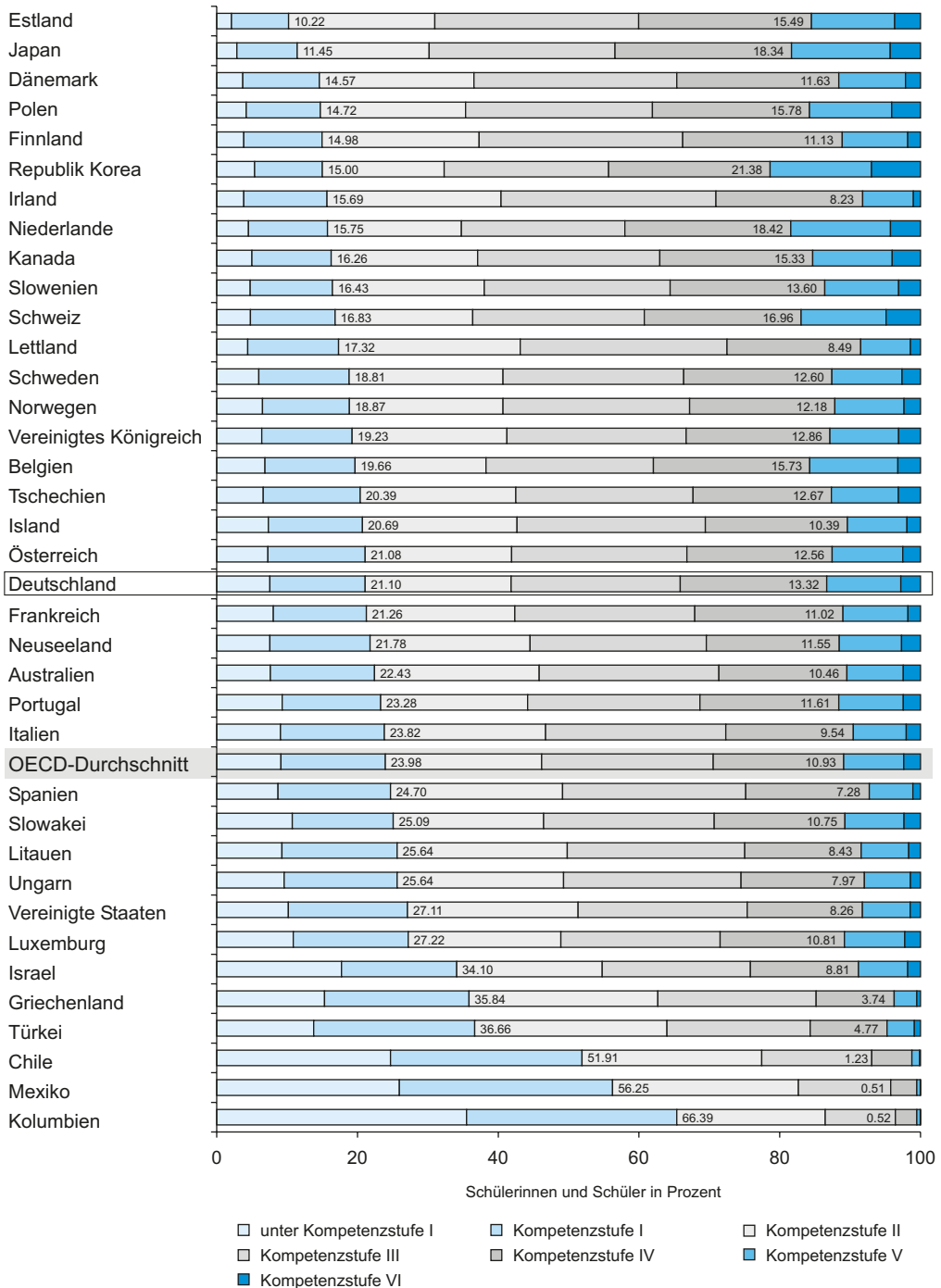


Abbildung 8.3: Prozentuale Anteile der Schülerinnen und Schüler auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI

Macau (China), Singapur und Hongkong (China), beträgt dieser Anteil weniger als zehn Prozent.

Die obersten Stufen mathematischer Kompetenz erreichen im Durchschnitt der OECD-Staaten elf Prozent der Schülerinnen und Schüler. Dieser Anteil beträgt in Deutschland 13 Prozent und liegt damit signifikant über dem OECD-Durchschnitt. In vier OECD-Staaten – der Republik Korea, Japan, Tschechien und den Niederlanden – liegt der Anteil besonders leistungsstarker Fünfzehnjähriger bei über 17 Prozent. In fünf OECD-Partnerstaaten übersteigt der Anteil auch diese Grenze, darunter Singapur mit 37 Prozent und B-S-J-Z (China) mit 46 Prozent.

8.2.3 Geschlechterunterschiede

Die geringe Beteiligung von Frauen in mathematisch-naturwissenschaftlichen Bereichen (MINT) ist wohl eines der meist diskutierten Themen der letzten Jahrzehnte, wenn es um Geschlecht und Bildung geht (Ceci, Ginther, Kahn & Williams, 2014). Die Studienlage in Bezug auf tatsächliche Geschlechterunterschiede in der Mathematikleistung ist allerdings als uneinheitlich zu bezeichnen. In mehreren metaanalytischen Studien wurde wiederholt eine mittlere Effektgröße nahe Null für geschlechtsspezifische Unterschiede in der mathematischen Leistung berichtet (Else-Quest, Hyde & Linn, 2010; Hyde, Lindberg, Linn, Ellis & Williams, 2008). Ob Geschlechterunterschiede gefunden werden oder nicht, scheint insbesondere von sozialen und kulturellen sowie methodischen Bedingungen abzuhängen (Halpern et al., 2007; Penner, 2008; Voyer & Voyer, 2014). Die in PISA erfasste mathematische Kompetenz weist für Deutschland sowohl 2012 als auch 2015 signifikante Geschlechterunterschiede zum Nachteil der Schülerinnen aus. Um Bildungsgerechtigkeit zu gewährleisten, sollte ein Bildungssystem allen Kindern und Jugendlichen unabhängig ihres Geschlechts ermöglichen, entsprechend ihres Potenzials zu lernen. Vor diesem Hintergrund werden mögliche Unterschiede zwischen fünfzehnjährigen Mädchen und Jungen in PISA 2018, die in Abbildung 8.4 für die OECD-Staaten dargestellt sind, interpretiert.

Betrachtet man den Durchschnitt aller OECD-Staaten, so weisen Jungen mit 492 Punkten auf der Skala mathematischer Kompetenz einen signifikant höheren Wert auf als Mädchen mit 487 Punkten. Signifikante Unterschiede zugunsten der Jungen zeigen sich in 21 der 37 OECD-Staaten, darunter auch in Deutschland mit einer Differenz von sieben Punkten. Österreich, Mexiko, und Japan zählen zu den sieben Staaten, in denen diese Differenz mehr als zehn Punkte beträgt. Der Unterschied zugunsten der Jungen scheint daher sowohl ein Phänomen in leistungsstarken wie auch in leistungsschwachen Staaten zu sein.

In 13 OECD-Staaten unterscheiden sich Mädchen und Jungen in ihrer mathematischen Kompetenz nicht signifikant, darunter die Türkei, Polen, Slowenien und Schweden sowie die Republik Korea und die Niederlande, die beide zur Spitzengruppe der OECD-Staaten gehören. In diesen Staaten scheint die Förderung von Mädchen und

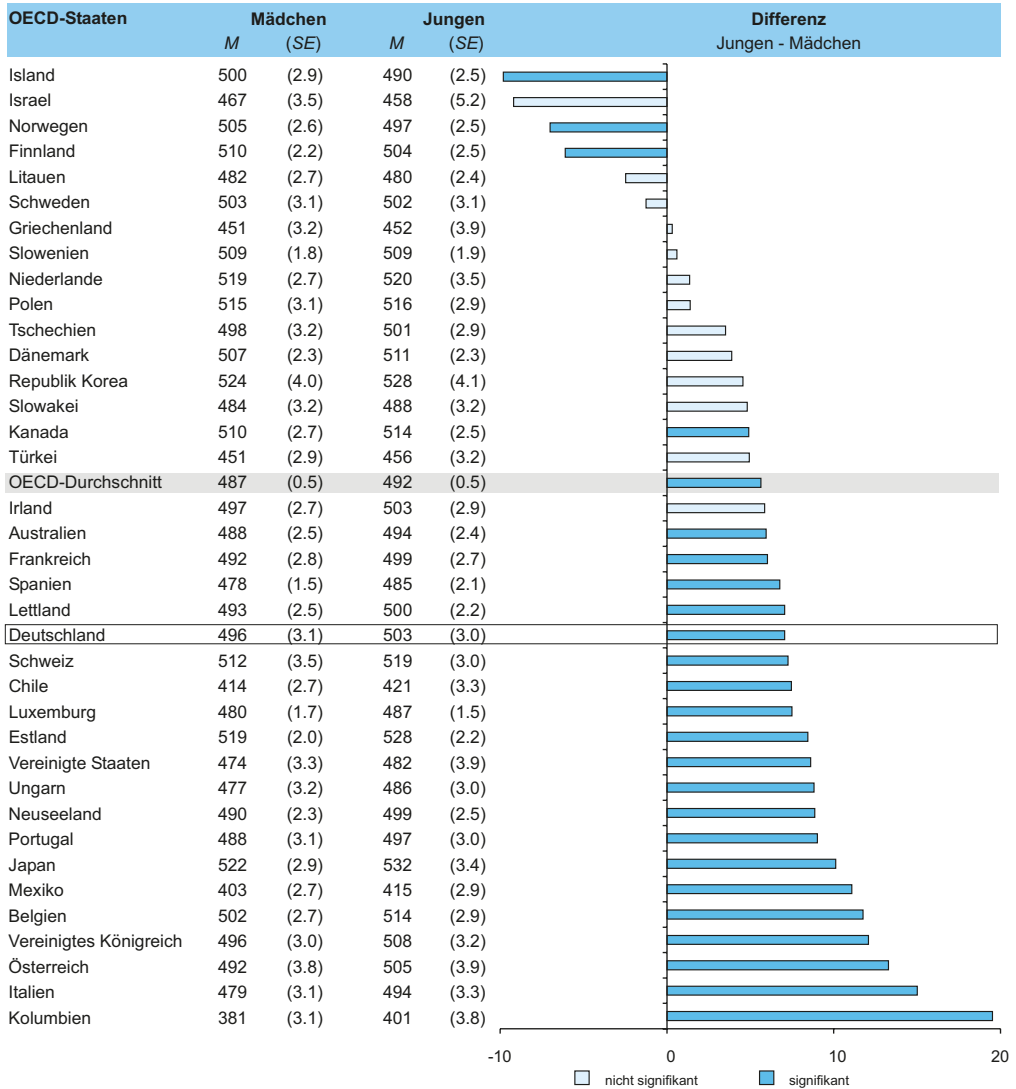


Abbildung 8.4: Mittelwerte mathematischer Kompetenz von Mädchen und Jungen in den OECD-Staaten

Jungen besonders gut zu gelingen. In drei OECD-Staaten – Finnland, Norwegen und Island – fallen Unterschiede in der mathematischen Kompetenz zugunsten der Mädchen aus. In vier der fünf OECD-Partnerstaaten der Spitzengruppe – Singapur, Macau (China), Chinesisch Taipeh und Hongkong (China) – zeigen sich keine signifikanten Unterschiede zwischen Mädchen und Jungen.

8.2.4 Veränderungen der mathematischen Kompetenz

PISA ermöglicht Trendergebnisse, die die Entwicklung der Effektivität der Bildungssysteme hinsichtlich der mathematischen Kompetenz ihrer Schülerinnen und Schüler sowohl international wie auch national beschreiben kann. In diesem Kapitel werden als Referenzwerte für diese Trends die letzten beiden Erhebungsrunden 2012 und 2015 betrachtet. PISA 2012 ist dabei von besonderem Interesse, weil Mathematik als Hauptdomäne geprüft wurde. PISA 2015 wird berücksichtigt, da sowohl Aufgaben als auch Erhebungsmodus – insbesondere die computerbasierte Testung – nahezu identisch zu PISA 2018 waren.

Deutschland zählt zu den Staaten, die im Vergleich zu 2012 einen leichten Rückgang bei der Entwicklung mathematischer Kompetenz aufweisen. Einen ähnlichen Trend zeigen auch andere OECD-Staaten, etwa Island und Frankreich, in denen die Mathematikleistung vergleichbar mit Deutschland ist. Mit den Niederlanden und der Schweiz sind darunter aber auch Staaten, die der Spitzengruppe der OECD-Mitglieder zuzurechnen sind. In mehreren Staaten mit zu Deutschland vergleichbarer Mathematikkompetenz hat die Differenz zwischen leistungsstarken und leistungsschwachen Jugendlichen abgenommen. Dies ist nicht zwingend positiv zu interpretieren: So hat sich etwa in Irland und Neuseeland der Anteil leistungsstarker Jugendlicher tendenziell verringert, während beispielsweise Lettland einen Rückgang leistungsschwacher Schülerinnen und Schüler zu verzeichnen hat. Besonders positiv kann erneut Estland hervorgehoben werden: Neben dem bereits beschriebenen geringen Anteil an Jugendlichen auf der untersten Kompetenzstufe verzeichnet dieser Staat auch einen insgesamt positiven Trend.

8.3 Vertiefende Analysen der mathematischen Kompetenz in Deutschland

Nachfolgend werden detailliertere Ergebnisse zur mathematischen Kompetenz in PISA 2018 von Schülerinnen und Schülern in Deutschland – insbesondere im Hinblick auf Schularten und das Geschlecht – dargestellt. Die zeitliche Entwicklung wird für die verschiedenen Gruppen mit Bezug zu PISA 2012 (Sälzer et al., 2013) sowie auch zu PISA 2015 (Hammer et al., 2016) analysiert.

8.3.1 Kompetenzstände von Fünfzehnjährigen verschiedener Schularten

Die Mittelwerte und Standardabweichungen der mathematischen Kompetenz in PISA 2018 für gymnasiale und nicht gymnasiale Schularten sind in Tabelle 8.2 dargestellt. Fünfzehnjährige an Gymnasien zeigen mit 570 Punkten demnach eine signifikant höhere mathematische Kompetenz als Fünfzehnjährige an nicht gymnasialen Schularten

Tabelle 8.2: Mittelwerte und Standardabweichungen für die Gesamtstichprobe und nach Schulart

Schulart	<i>n</i>	<i>M</i>	(<i>SE</i>)	<i>SD</i>	(<i>SE</i>)
Nicht gymnasiale Schularten	3 168	465	(2.9)	83	(1.7)
Gymnasium	2 042	570	(2.3)	71	(1.5)
Gesamtstichprobe	5 451	500	(2.6)	95	(1.5)

mit 465 Punkten. Dies entspricht einem mittleren Leistungsunterschied von über einhalb Kompetenzstufen (105 Punkte).

Auch die Streuung der mathematischen Kompetenz an Gymnasien (*SD* = 71) und an nicht gymnasialen Schularten (*SD* = 83) unterscheidet sich signifikant. Die in Abbildung 8.5 dargestellten Perzentilbänder für beide Schularten sowie den Durchschnitt in Deutschland zeigen, dass sich die Kompetenzverteilungen zum Teil überlappen. Dies zeigt wie bereits in allen vorausgegangenen PISA-Runden, dass durchaus leistungsstarke Schülerinnen und Schüler an nicht gymnasialen Schularten sowie leistungsschwache Jugendliche an Gymnasien zu finden sind. Gleichzeitig ist bei der Verteilung der Jugendlichen auf die Schularten festzustellen, dass – wie zu erwarten – im Wesentlichen doch

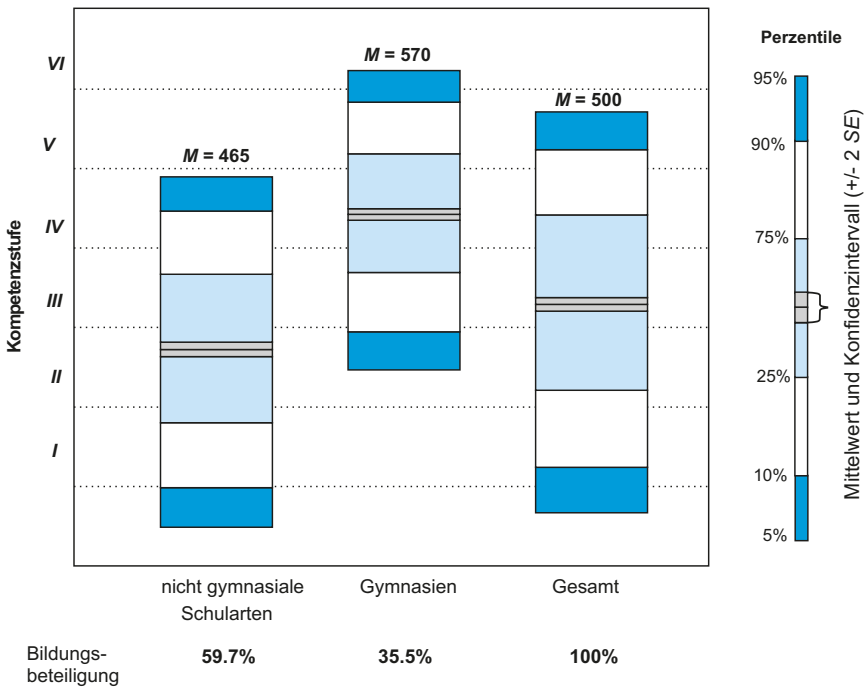


Abbildung 8.5: Perzentilbänder mathematischer Kompetenz in Deutschland nach Schulart und für die Gesamtstichprobe

die Schülerinnen und Schüler aus dem leistungsstärkeren Bereich ein Gymnasium besuchen und in nicht gymnasialen Schularten die Leistungsspitze kaum vertreten ist. Dies erscheint insbesondere bei solchen nicht gymnasialen Schularten bemerkenswert, die zum Abitur führen und an denen offenbar eine adäquate Förderung leistungsstärkerer Schülerinnen und Schüler weniger gelingt.

Diese Analyse kann mit Hilfe der in Abbildung 8.6 dargestellten prozentualen Anteile an Schülerinnen und Schülern auf den einzelnen Kompetenzstufen weiter ausgeführt werden. Über 20 Prozent der Schülerinnen und Schüler in Deutschland erfüllen in PISA 2018 nicht die Anforderungen der Kompetenzstufe II. Am Gymnasium sind dies etwas über zwei Prozent, ihr Anteil bei den nicht gymnasialen Schularten beträgt knapp 30 Prozent. Im Vergleich dazu erreichen in Deutschland knapp 14 Prozent der Fünfzehnjährigen die obersten Kompetenzstufen V und VI und zwar 30 Prozent der Gymnasiastinnen und Gymnasiasten, aber lediglich vier Prozent der Schülerinnen und Schüler an nicht gymnasialen Schulen.

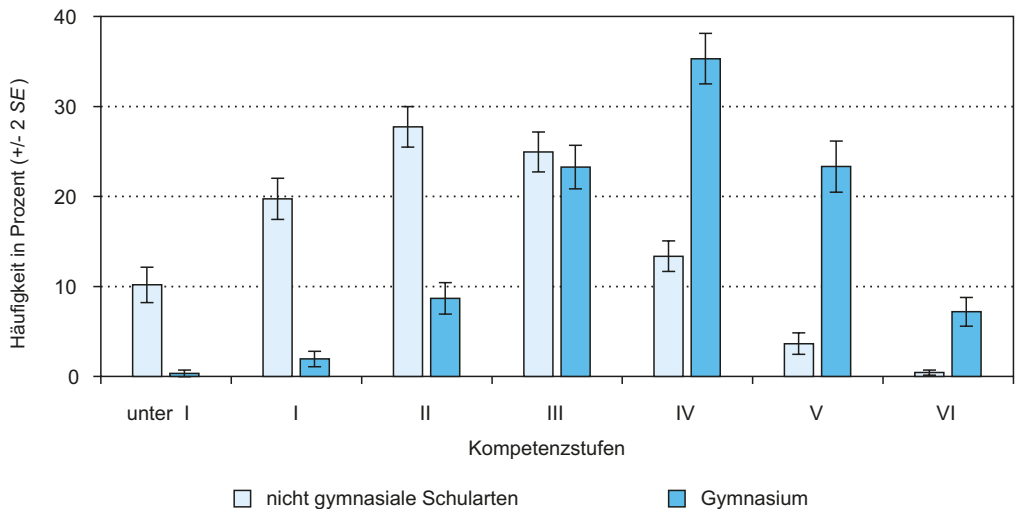


Abbildung 8.6: Prozentuale Anteile von Schülerinnen und Schülern auf den Stufen mathematischer Kompetenz in Deutschland nach Schulart

8.3.2 Mädchen und Jungen in Deutschland

In Deutschland zeigen Jungen mit 503 Punkten ($SD = 98$) eine signifikant höhere mathematische Kompetenz als Mädchen mit 496 Punkten ($SD = 92$). Jedoch unterscheiden sich die Anteile von Mädchen und Jungen auf den untersten Kompetenzstufen kaum (siehe Abbildung 8.7). Der – in allen PISA-Runden bestehende – Kompetenzunterschied zwischen Mädchen und Jungen ist daher nicht auf eine unterschiedliche Verteilung auf die unteren Kompetenzstufen, sondern auf die oberen Kompetenzstufen

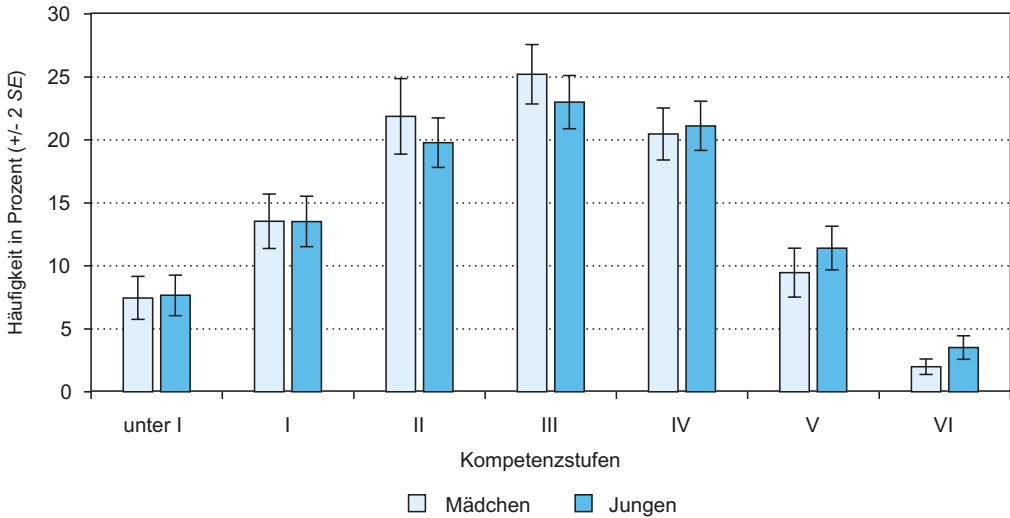


Abbildung 8.7: Prozentualer Anteil der Mädchen und Jungen auf den Stufen der mathematischen Kompetenz in Deutschland

zurückzuführen. Insbesondere ist auf den obersten Kompetenzstufen der Anteil der Jungen mit 15 Prozent signifikant höher als der Anteil der Mädchen mit elf Prozent.

8.3.3 Entwicklung der mathematischen Kompetenz in Deutschland

In PISA 2018 liegt die mittlere mathematische Kompetenz der Fünfzehnjährigen in Deutschland mit 500 Punkten knapp sechs Punkte unter dem Wert in PISA 2015 und 13 Punkte unter dem Wert von PISA 2012. Im Vergleich zu dem bei PISA 2012 erreichten Höchstwert ergibt sich bei PISA 2018 ein signifikanter Rückgang der mathematischen Kompetenz. Trotzdem liegt sie auch in PISA 2018 – wie seit 2006 konsequent – signifikant über dem OECD-Mittelwert. Abbildung 8.8 zeigt die Entwicklung der mathematischen Kompetenz von Fünfzehnjährigen in Deutschland seit 2003.

Es erscheint allerdings bemerkenswert, dass seit PISA 2003 die Mittelwerte an Gymnasien fast kontinuierlich (mit Ausnahme von PISA 2009) geringer werden. In PISA 2015 erreichten Schülerinnen und Schüler an Gymnasien bereits schwächere Leistungen als in den Erhebungen 2003 und 2012, in denen Mathematik als Hauptdomäne geprüft wurde. Gleiches gilt nun für PISA 2018.

Weiter ist anzumerken, dass die Streuung der mathematischen Kompetenz in Deutschland bei PISA 2018 ($SD = 95$) genauso wie bei PISA 2012 ($SD = 96$) über dem OECD-Durchschnitt liegt. Die Verteilung mathematischer Kompetenz ist in Deutschland breit und der Unterschied zwischen leistungsschwachen und leistungsstarken Schülerinnen und Schülern nach wie vor groß. Es ist anzunehmen, dass die hohe Hetero-

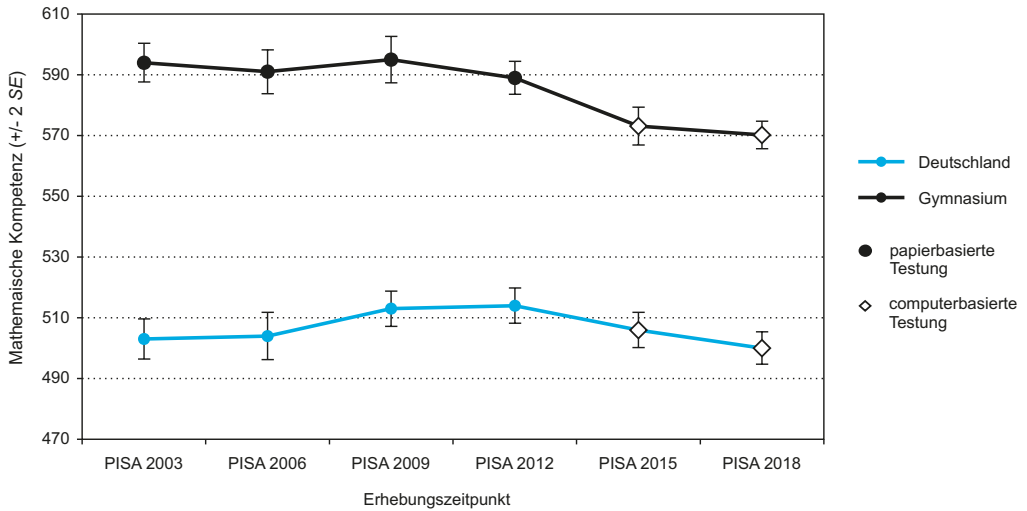


Abbildung 8.8: Mathematische Kompetenz in Deutschland und an Gymnasien von PISA 2003 bis PISA 2018

genität auf einen Anstieg derjenigen Schülerinnen und Schüler zurückzuführen ist, die die Kompetenzstufe II nicht erreichen. Mit 21 Prozent ist dieser Anteil signifikant größer als 2012. Insbesondere ist der Anteil der Schülerinnen und Schüler an nicht gymnasialen Schulen, die nur die untersten Kompetenzstufen erreichen, mit knapp 30 Prozent sehr hoch. Gleichzeitig ist der Anteil sehr leistungsstarker Schülerinnen und Schüler in Deutschland, die die Kompetenzstufen V oder VI erreichen, mit 13 Prozent zwar identisch zu PISA 2015, aber signifikant kleiner als bei PISA 2012. Auch hier ist allerdings der OECD-Trend zu sehen, denn dieser Anteil ist im Durchschnitt der Staaten ebenfalls gesunken.

Der in PISA 2012 festgestellte Unterschied zwischen Mädchen und Jungen, der 2015 mit knapp 17 Punkten seinen höchsten Wert erreichte, ist bei PISA 2018 auf lediglich sieben Punkte signifikant zurückgegangen. Grund für diese Änderung ist insbesondere eine geringere mittlere Kompetenz bei den Jungen: Während die mathematische Kompetenz bei Mädchen in Deutschland im Vergleich zu PISA 2012 weitgehend unverändert ist, hat diese bei den Jungen seit 2012 um knapp 17 Punkte signifikant abgenommen.

8.4 Diskussion

In Deutschland erreichen Schülerinnen und Schüler eine durchschnittliche mathematische Kompetenz von 500 Punkten und liegen damit signifikant über dem OECD-Durchschnitt. Die Entwicklung der mathematischen Kompetenz in PISA – über alle OECD-Staaten gesehen – erscheint weitgehend konstant. Allerdings hat sich die mathe-

matische Kompetenz der Fünfzehnjährigen in Deutschland im Vergleich zur Erhebung PISA 2012, in der Mathematik als Hauptdomäne geprüft wurde, signifikant verringert.

Etwa ein Fünftel aller Fünfzehnjährigen – sowohl im OECD-Durchschnitt als auch in Deutschland – verfügen lediglich über rudimentäre mathematische Kenntnisse, die mit Problemen an der gesellschaftlichen Teilhabe einhergehen können. In Deutschland hat der Anteil leistungsschwacher Jugendlicher, die Kompetenzstufe II nicht erreichen, seit 2012 zugenommen. Zudem zeigt sich seit 2012 hier wie auch im nationalen Bildungstrend (Stanat, Schipolowski, Mahler, Weirich & Henschel, 2019) ein tendenzieller Rückgang der Leistungen am Gymnasium. Eine wesentliche Frage, die hieraus abgeleitet werden kann, ist wie Schülerinnen und Schüler mit unterschiedlichen Voraussetzungen in leistungsheterogenen Klassenverbänden im Bereich mathematischer Kompetenz adäquat gefördert werden können. Eine mögliche Antwort hierauf bieten adaptive und interaktive Lernumgebungen, die Lehrkräfte im Umgang mit Leistungsheterogenität im Klassenraum professionell unterstützen können und sich empirisch als besonders wirksam für leistungsschwache Lernende erweisen (Reinhold, Hoch, Werner, Richter-Gebert & Reiss, 2020).

Weiterhin bestehen über die OECD-Staaten hinweg signifikante Unterschiede in der mathematischen Kompetenz von Jungen und Mädchen zugunsten der Jungen. In Deutschland ist hier ein signifikanter Rückgang des in PISA 2015 stark ausgeprägten Geschlechterunterschieds zu verzeichnen. Dieser lässt sich jedoch nicht auf eine Leistungssteigerung bei den Mädchen, sondern auf einen Rückgang der Leistung bei den Jungen zurückführen. Es gilt zu analysieren, welche Ursachen dieser einseitige Rückgang bei den Jungen haben kann, der insbesondere auch im aktuellen nationalen Bildungstrend zu beobachten ist (Stanat et al., 2019). Im Einklang mit der Idee, dass es soziale und kulturelle Gegebenheiten innerhalb eines Bildungssystems sind, die zu Geschlechterunterschieden beitragen (Halpern et al., 2007), stellen etwa Schweden und die Niederlande Beispiele dafür dar, dass Geschlechterunterschiede im Bereich Mathematik durchaus vermieden werden können. Eine genauere Betrachtung der Unterschiede zwischen den Lernumwelten dieser Staaten und denjenigen Bildungssystemen, die signifikante Geschlechterunterschiede zu verzeichnen haben – wie etwa weiterhin auch Deutschland – kann hier wertvolle Erkenntnisse für konkrete Maßnahmen für einen geschlechtergerechten Mathematikunterricht liefern.

Die Ergebnisse aus PISA 2018 zeigen, dass sich Deutschland in Bezug auf die mathematische Kompetenz in einer Gruppe von Staaten wiederfindet, in denen die Leistungen signifikant über dem OECD-Durchschnitt liegen. Dies gelingt nur wenigen Staaten und ist gerade angesichts einer heterogener werdenden Gesellschaft und damit einer sich wandelnden Zusammensetzung der Schülerinnen und Schüler in Deutschland bemerkenswert. Allerdings bleiben Probleme, die sich bereits über viele PISA-Runden hinweg gezeigt haben. An erster Stelle ist hier der relativ große Anteil von Jugendlichen zu nennen, die Leistungen nur auf den unteren Kompetenzstufen zeigen und damit nicht hinreichend für ihren weiteren Weg speziell im Beruf und allgemein in der Gesellschaft ausgebildet sind. Ihr Anteil variiert zwar von PISA-Runde zu PISA-Runde, ist aber mit

etwa einem Fünftel der Schülerinnen und Schüler in Deutschland deutlich zu hoch. Es gilt, sich auf diese Gruppe noch stärker zu konzentrieren und über gezieltere Unterstützungsmaßnahmen nachzudenken. Die Einführung von Bildungsstandards war ein wichtiger Schritt und die Kontrolle ihrer Implementation ist gelungen (vgl. Stanat et al., 2019). Allerdings forderten Klieme et al. (2003) darüber hinaus, dass Schülerinnen und Schüler, Lehrkräfte, Eltern und Politik jeweils Verantwortung für den Erfolg des Lernens übernehmen müssten. Dieser Aspekt ist bisher nicht systematisch in den Blick genommen worden. Das gilt genauso auch für die Förderung von besonders begabten oder auch leistungsfähigen Jugendlichen. Auch hier gilt es zu sehen, dass Lehren und Lernen in die Mitte der Gesellschaft genommen werden muss und für alle ihre Mitglieder eine zentrale Verantwortung bedeutet.

Literatur

- Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., ... Neubrand, J. (1997). *TIMSS – mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich: Deskriptive Befunde*. Opladen: Leske + Budrich. <https://doi.org/10.1007/978-3-322-95096-3>
- Blum, W., Neubrand, M., Ehmke, T., Senkbeil, M., Jordan, A., Ulfig, F. & Carstensen, C. H. (2004). Mathematische Kompetenz. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, ... U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs* (S. 47–92). Münster: Waxmann.
- Ceci, S. J., Ginther, D. K., Kahn, S. & Williams, W. M. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, 15(3), 75–141. <https://doi.org/10.1177/1529100614541236>
- Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103–127. <https://doi.org/10.1037/a0018053>
- Frey, A., Asseburg, R., Carstensen, C. H., Ehmke, T. & Blum, W. (2007). Mathematische Kompetenz. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 249–276). Münster: Waxmann.
- Frey, A., Heinze, A., Mildner, D., Hochweber, J. & Asseburg, R. (2010). Mathematische Kompetenz von PISA 2003 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 153–176). Münster: Waxmann.
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S. & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, 8(1), 1–51. <https://doi.org/10.1111/j.1529-1006.2007.00032.x>
- Hammer, S., Reiss, K., Lehner, M. C., Heine, J.-H., Sälzer, C. & Heinze, A. (2016). Mathematische Kompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015: Eine Studie zwischen Kontinuität und Innovation* (S. 219–248). Münster: Waxmann.

- Heymann, H. W. (1996). *Allgemeinbildung und Mathematik*. Weinheim: Beltz.
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B. & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, 321(5888), 494–495. <https://doi.org/10.1126/science.1160364>
- Klafki, W. (1985). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Weinheim: Beltz.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., ... Vollmer, H. J. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bonn: Bundesministerium für Bildung und Forschung (BMBF).
- Klieme, E., Neubrand, M. & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, ... M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 141–191). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_5
- KMK (2003) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Hrsg.). (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss*. München: Luchterhand.
- KMK (2010) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2010). *Förderstrategie für leistungsschwächere Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 04.03.2010*. Verfügbar unter https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_03_04-Foerderstrategie-Leistungsschwaechere.pdf
- KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Förderstrategie für leistungsstarke Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 11.06.2015*. Verfügbar unter https://www.kmk.org/fileadmin/Dateien/pdf/350-KMK-TOP-011-Fu-Leistungsstarke_-_neu.pdf
- Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatsis & S. Papastavridis (Hrsg.), *3rd Mediterranean Conference on Mathematics Education* (S. 115–124). Athens, Greece: Hellenic Mathematical Society and Cyprus Mathematical Society.
- Niss, M. & Højgaard, T. (2011). *Competencies and mathematical learning – Ideas and inspiration for the development of mathematics teaching and learning in Denmark*. Roskilde: Roskilde University.
- OECD. (2003). *The PISA 2003 Assessment framework—Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD Publishing.
- OECD. (2013a). *PISA 2012 assessment and analytical framework*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264190511-en>
- OECD. (2013b). *PISA 2012 released mathematics items*. Paris: OECD Publishing. Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf>
- OECD. (2019). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- Penner, A. M. (2008). Gender differences in extreme mathematical achievement: An international perspective on biological and social factors. *American Journal of Sociology*, 114(S1), S138–S170. <https://doi.org/10.1086/589252>
- Prenzel, M., Friedrich, A. & Stadler, M. (Hrsg.). (2009). *Von SINUS lernen. Wie Unterrichtsentwicklung gelingt*. Seelze: Kallmeyer.

- Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J. & Reiss, K. (2020). Learning fractions with and without educational technology: What matters for high-achieving and low-achieving students? *Learning and Instruction*, 65, 101264. <https://doi.org/10.1016/j.learninstruc.2019.101264>
- Reiss, K., Reinhold, F. & Strohmaier, A. (in Druck). Mathematikdidaktik. In M. Rothgangel, U. Abraham, H. Bayrhuber, V. Frederking, W. Jank & H. J. Vollmer (Hrsg.), *Lernen im Fach und über das Fach hinaus. Bestandsaufnahmen und Forschungsperspektiven aus 17 Fachdidaktiken im Vergleich. Allgemeine Fachdidaktik, Band 2* (S. 234–259). Münster: Waxmann.
- Sälzer, C., Reiss, K., Schiepe-Tiska, A., Prenzel, M. & Heinze, A. (2013). Zwischen Grundlagenwissen und Anwendungsbezug: Mathematische Kompetenz im internationalen Vergleich. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 47–98). Münster: Waxmann.
- Schiepe-Tiska, A. & Schmidtner, S. (2013). Mathematikbezogene emotionale und motivationale Orientierungen, Einstellungen und Verhaltensweisen von Jugendlichen in PISA 2012. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 99–122). Münster: Waxmann.
- Stanat, P., Schipolowski, S., Mahler, N., Weirich, S. & Henschel, S. (Hrsg.). (2019). *IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich*. Münster: Waxmann.
- Voyer, D. & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Weinert, F. E. (Hrsg.). (2001). *Leistungsmessungen in Schulen* (2. Aufl.). Weinheim: Beltz.
- Winter, H. (1995). Mathematikunterricht und Allgemeinbildung. *Mitteilungen der Gesellschaft für Didaktik der Mathematik*, 61, 37–46.

9 Naturwissenschaftliche Kompetenz in PISA 2018 – aktueller Stand, Veränderungen und Implikationen für die naturwissenschaftliche Bildung in Deutschland

Anja Schiepe-Tiska, Silke Rönnebeck & Knut Neumann

Fünfzehnjährige in Deutschland erreichen bei PISA 2018 in den Naturwissenschaften 503 Punkte und liegen abermals über dem OECD-Durchschnitt. Deutschland gehört damit zu einer Gruppe von zwölf Staaten, die auf diesem Niveau den OECD-Spitzenstaaten Estland und Japan sowie Finnland, der Republik Korea und Kanada folgt. Im Vergleich zu PISA 2015, der als belastbarer Referenzpunkt gesehen werden kann, bleibt die naturwissenschaftliche Kompetenz weitgehend stabil. Die Ergebnisse legen nahe, dass in Deutschland die Spitzenförderung vergleichsweise gut gelingt. Die Förderung schwächerer Schülerinnen und Schüler scheint jedoch verstärkte Aufmerksamkeit zu erfordern, denn ein Fünftel der Jugendlichen erreicht nicht die Kompetenzstufe II. Gerade in den nicht gymnasialen Schularten hat sich die naturwissenschaftliche Kompetenz im Vergleich zu PISA 2015 signifikant verringert. Mädchen und Jungen unterscheiden sich in Deutschland bei PISA 2018 ähnlich wie im OECD-Durchschnitt nicht in ihrer naturwissenschaftlichen Kompetenz. Im Vergleich zu PISA 2015 ist aber insbesondere bei den Jungen eine bedeutsame Abnahme der naturwissenschaftlichen Kompetenz zu beobachten, während sich gleichzeitig die Streuung vergrößert hat. Die Kompetenz der Mädchen bleibt weitgehend stabil.

Die großen Herausforderungen des 21. Jahrhunderts, wie zum Beispiel der Klimawandel, erfordern mehr denn je eine fundierte naturwissenschaftliche Bildung. Diese muss es Jugendlichen ermöglichen, naturwissenschaftliche Phänomene und technische Prozesse im Alltag zu erkennen, zu verstehen und zu erklären. Mit Blick darauf, dass ständig neue Erkenntnisse gewonnen werden und immer neue Technologien Einzug in den Alltag halten, müssen Jugendliche lernen, sich neue Informationen zu erschließen. Diese Informationen können sowohl aus naturwissenschaftlichen oder technischen Publikationen als auch aus eigenen naturwissenschaftlichen Untersuchungen stammen. Außerdem muss naturwissenschaftliche Bildung Schülerinnen und Schüler befähigen, die gewonnenen Informationen kritisch zu bewerten und daraus Schlussfolgerungen für das eigene Handeln zu ziehen (OECD, 2019a; siehe auch NRC, 2012). Dabei reicht es nicht, wenn

Schülerinnen und Schüler über spezifisches Wissen verfügen. Sie müssen vielmehr auch in der Lage sein, sich neues Wissen eigenständig anzueignen. Dies setzt neben einem fundierten Inhaltswissen Prozess- und Überblickswissen voraus (vgl. Shavelson, Ruiz-Primo & Wiley, 2005). Nicht zuletzt erfordern die beschriebenen Kompetenzen ein gehöriges Maß an Motivation (vgl. Schiepe-Tiska, Roczen, Müller, Prenzel & Osborne, 2016; Schiepe-Tiska, Simm & Schmidtner, 2016). Eine naturwissenschaftliche Bildung, die Jugendliche im beschriebenen Sinne gleichermaßen auf die Begegnung mit aktuellen und zukünftigen Herausforderungen vorbereitet, wird auch als naturwissenschaftliche Grundbildung bezeichnet (vgl. deBoer, 2000).

Die Vermittlung einer naturwissenschaftlichen Grundbildung ist zentrale Aufgabe der allgemeinbildenden Schule. Inwieweit es im internationalen Vergleich gelingt, Schülerinnen und Schülern eine entsprechende Grundbildung zu vermitteln, untersucht die PISA-Studie seit fast zwei Jahrzehnten. In den ersten Erhebungsrunden zeigte sich für Deutschland, dass die Leistungen der Fünfzehnjährigen im Bereich der Naturwissenschaften nicht den Erwartungen entsprachen (Prenzel, Rost, Senkbeil, Häußler & Klopp, 2001; Rost, Walter, Carstensen, Senkbeil & Prenzel, 2004). In der Folge wurden zahlreiche Maßnahmen zur Verbesserung des naturwissenschaftlichen Unterrichts initiiert. An erster Stelle ist die Einführung von Bildungsstandards für den Mittleren Schulabschluss für die naturwissenschaftlichen Fächer zu nennen (KMK, 2005a, 2005b, 2005c). Außerdem wurde eine Reihe von Programmen zur Verbesserung des naturwissenschaftlichen Unterrichts aufgelegt. Beispiele sind etwa SINUS, das Programm zur Effizienzsteigerung des mathematisch-naturwissenschaftlichen Unterrichts (Prenzel, Friedrich & Stadler, 2009), sowie die Kontextprojekte in der Biologie, Chemie und Physik (Bayrhuber et al., 2007; Demuth, Gräsel, Parchmann & Ralle, 2008; Mikelskis-Seifert & Duit, 2010). Darüber hinaus gibt es in den letzten Jahren zunehmend Angebote, die sich direkt an Schülerinnen und Schüler richten, wie zum Beispiel naturwissenschaftliche Schülerwettbewerbe (Petersen, Blankenburg & Höffler, 2017).

Die wiederholte Durchführung der PISA-Studie erlaubt mit aller gebotenen Vorsicht Rückschlüsse über die Wirkungen dieser Maßnahmen zu ziehen. So lässt die Verbesserung und Stabilisierung der naturwissenschaftlichen Kompetenz in Deutschland vermuten, dass die initiierten Maßnahmen in ihrer Gesamtheit wirksam waren. Seit PISA 2006 liegen die Leistungen der Schülerinnen und Schüler in Deutschland signifikant über dem OECD-Durchschnitt (Prenzel et al., 2007; Rönnebeck, Schöps, Prenzel, Mildner & Hochweber, 2010; Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2013). In der letzten Erhebungsrunde stagnierte dieser Aufwärtstrend (Schiepe-Tiska, Rönnebeck et al., 2016), was jedoch möglicherweise durch Veränderungen in der PISA-Rahmenkonzeption, die Umstellung auf computerbasiertes Testen sowie die in diesem Zuge neu entwickelten Aufgaben begründet war (Robitzsch et al., 2016). Umso interessanter ist die Frage, wie die Ergebnisse der Erhebungsrunde 2018 ausfallen, der dieselbe Rahmenkonzeption zugrunde liegt und die die gleichen Testprozeduren nutzt wie PISA 2015.

9.1 Naturwissenschaftliche Kompetenz in PISA 2018

9.1.1 Die Rahmenkonzeption der Domäne Naturwissenschaften

PISA untersucht, in welchem Maße Fünfzehnjährige über naturwissenschaftliche Grundbildung (*Scientific Literacy*) am Ende ihrer Pflichtschulzeit verfügen (OECD, 2019a). Dabei stehen drei Fragen im Vordergrund: (a) Was wissen und können Jugendliche, wenn sie naturwissenschaftlichen und technischen Fragen und Problemen gegenüberstehen? (b) Sind sie in der Lage, dieses Wissen flexibel in unterschiedlichen Situationen anzuwenden? (c) Welche Bedeutung haben Naturwissenschaften und Technik für sie? Dieses Verständnis naturwissenschaftlicher Grundbildung zielt nicht nur auf das Wiedergeben des in der Schule erworbenen Wissens. Vielmehr steht im Mittelpunkt, inwieweit Jugendliche dieses Wissen in unterschiedlichen, alltagsnahen Kontexten, in denen Naturwissenschaften und Technik eine Rolle spielen, sinnvoll und problemlösend anwenden können (Bybee & McCrae, 2011). Eine solche Grundbildung befähigt Schülerinnen und Schüler, aktiv an gesellschaftlichen Prozessen teilzuhaben und ihre Lebensumwelt mitzugestalten (Prenzel et al., 2001). Sie bildet außerdem das Fundament für lebenslanges Lernen.

Das Kompetenzmodell in PISA 2018 wurde von PISA 2015 übernommen. In PISA 2015 waren die Naturwissenschaften das letzte Mal Hauptdomäne (OECD, 2016; Schiepe-Tiska, Rönnebeck et al., 2016). Damals wurde die Rahmenkonzeption überarbeitet, aktualisiert und mit Blick auf das erstmals eingesetzte, computerbasierte Erhebungsverfahren weiterentwickelt (Abbildung 9.1). Sie beschreibt *Teilkompetenzen*, *Kontexte* und *Wissensbereiche* sowie *motivationale Orientierungen und Einstellungen*.

Im Mittelpunkt der Rahmenkonzeption steht die Unterscheidung dreier Teilkompetenzen. *Phänomene naturwissenschaftlich erklären* beschreibt, inwieweit Fünfzehnjährige naturwissenschaftliche Erklärungen für natürliche und technische Phänomene erken-

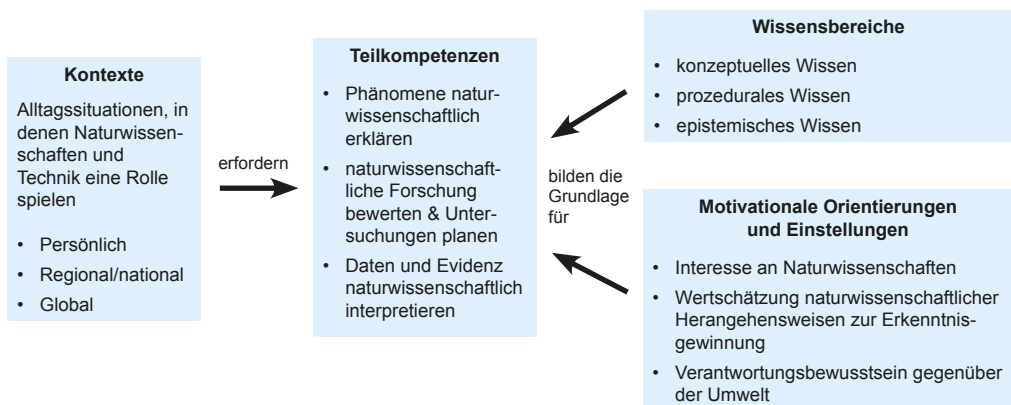


Abbildung 9.1: Die PISA-2018-Rahmenkonzeption naturwissenschaftlicher Grundbildung

nen, entwickeln und bewerten können. *Naturwissenschaftliche Forschung bewerten und Untersuchungen planen* umfasst die Fähigkeit, naturwissenschaftliche Fragestellungen zu beschreiben und zu beurteilen sowie Vorschläge zu entwickeln, wie diese naturwissenschaftlich untersucht werden können. *Daten und Evidenz naturwissenschaftlich interpretieren* bezieht sich darauf, unterschiedlich präsentierte naturwissenschaftliche Daten zu analysieren und zu interpretieren, aus diesen angemessene Schlussfolgerungen zu ziehen sowie Argumente und Evidenz kritisch zu hinterfragen.

Die Teilkompetenzen sollen Jugendliche in die Lage versetzen, naturwissenschaftliche Fragestellungen und Probleme, denen sie im Alltag begegnen, zu lösen. Damit dies auch in der PISA-Erhebung möglichst realitätsnah abgebildet werden kann, sind die Testaufgaben in Kontexte eingebettet. Die Kontexte sind so gewählt, dass sie den Jugendlichen möglichst vertraut sind und viele Interessen und Erfahrungsmöglichkeiten Fünfzehnjähriger berücksichtigen. Der Kontext kann sich dabei entweder auf die Jugendlichen selbst, ihre Familie oder Freunde beziehen (*persönlicher Kontext*, z. B. der persönliche Verbrauch von Materialien und Energie), die sie umgebende Gesellschaft betreffen (*regionaler/nationaler Kontext*, z. B. die Entsorgung von Abfällen) oder von weltweiter Relevanz sein (*globaler Kontext*, z. B. Klimaveränderung oder das Ausbreiten infektiöser Krankheiten).

Drei Wissensbereiche bilden die Grundlage, um die Teilkompetenzen in konkreten Situationen anwenden zu können. *Konzeptuelles Wissen* bezieht sich auf deklaratives, eher objektorientiertes naturwissenschaftliches Wissen („Wissen, dass“), welches in Deutschland meist schulfachbezogen vermittelt wird. *Prozedurales Wissen* („Wissen, wie“) beschreibt das fächerübergreifende Prozess- und Methodenwissen, welches sich auf den zentralen Prozess der naturwissenschaftlichen Erkenntnisgewinnung bezieht. Es befähigt Jugendliche, naturwissenschaftliche Fragen zu stellen und Daten zu gewinnen, die reliable und valide Antworten auf diese Fragen ermöglichen. Darüber hinaus soll es ihr Bewusstsein dafür schärfen, dass naturwissenschaftliches Wissen – wissenschaftstheoretisch betrachtet – vorläufig und von der Art, wie die Erkenntnisse generiert werden, abhängig ist. Schließlich umfasst *epistemisches Wissen* das Verständnis der Bedeutung wichtiger charakteristischer Eigenschaften der Naturwissenschaften und deren Funktionen beim Generieren naturwissenschaftlichen Wissens (Duschl, 2008). Dazu gehören zum Beispiel das Wissen um die Bedeutung von Modellen in den Naturwissenschaften, die als formale, abstrakte Repräsentationen und nicht als exakte Abbildungen der materiellen Welt verstanden werden, sowie das Verständnis der zentralen Rolle des wissenschaftlichen Austauschs (vgl. Schiepe-Tiska, Rönnebeck et al., 2016). Um diese Wissensbereiche auf Frage- und Problemstellungen im Alltag anwenden zu können, müssen Jugendliche häufig Wissen aus verschiedenen Disziplinen miteinander verknüpfen. Deshalb unterscheidet PISA nicht herkömmliche Schulfächer, sondern die Wissenssysteme *physikalische Systeme*, *lebende Systeme* sowie *Erd- und Weltraumsysteme*.

Neben dem Wissen, über das Schülerinnen und Schüler verfügen, beeinflussen auch ihre motivationalen Orientierungen und Einstellungen, ob und wie sie ihre Teilkompetenzen in einer naturwissenschaftlichen Problemsituation nutzen. *Interesse an Naturwis-*

senschaften, die Wertschätzung naturwissenschaftlicher Herangehensweisen zur Erkenntnisgewinnung sowie das Verantwortungsbewusstsein gegenüber der Umwelt sind wichtige Bildungsziele, die – im Sinne einer mehrdimensionalen Bildung – eine eng gefasste, rein kognitive Leistungsorientierung überwinden, im Schulalltag aber leicht aus dem Blick geraten können (vgl. Schiepe-Tiska, Roczen et al., 2016; Schiepe-Tiska, Simm & Schmidtner, 2016).

9.1.2 Der Naturwissenschaftstest in PISA 2018

Durch den Status als Nebendomäne umfasste der Naturwissenschaftstest bei PISA 2018 nur Aufgaben, die bereits in früheren Erhebungsrounden eingesetzt wurden. Diese Trendaufgaben werden geheim gehalten, damit Veränderungen der Kompetenz über die Erhebungsrounden hinweg untersucht werden können. Die Aufgaben wurden wie bereits bei PISA 2015 computerbasiert administriert (vgl. Schiepe-Tiska, Rönnebeck et al., 2016). In diesem Kontext wurden für PISA 2015 auch neue Aufgaben entwickelt, welche die zusätzlichen Möglichkeiten des Darbietungsmodus am Computer nutzen (sogenannte interaktive Aufgaben). Ein Beispiel ist die Simulation von Einflüssen verschiedener Variablen (siehe Beispielaufgaben *Energieeffiziente Häuser*, Abbildungen 9.2 bis 9.6).

Jede Aufgabeneinheit besteht aus einem Stimulusmaterial, welches den Kontext beschreibt, und bis zu fünf Aufgaben (auch Items genannt). Bei PISA 2018 wurden insgesamt 34 Aufgabeneinheiten zu den Naturwissenschaften eingesetzt, die 115 Aufgaben umfassten. Zwei Drittel davon waren für PISA 2015 neu entwickelt worden, ein Drittel stammte aus früheren Erhebungsrounden. Alle Aufgaben lassen sich den Teilkompetenzen, Wissensbereichen und Wissenssystemen zuordnen (Tabelle 9.1). Darüber hinaus bilden sie drei kognitive Anforderungsniveaus ab (niedrig – 31 Aufgaben, mittel – 74 Aufgaben, hoch – 10 Aufgaben, vgl. Schiepe-Tiska, Rönnebeck et al., 2016). Motivationale Orientierungen und Einstellungen wurden diesmal wegen des Status als Nebendomäne nicht erhoben.

Es werden drei mögliche Antwortformate unterschieden. Bei einfachen Auswahlaufgaben (33 Aufgaben) muss aus vier vorgegebenen Antwortoptionen die richtige Antwort ausgewählt werden (siehe Aufgabeneinheit *Energieeffiziente Häuser*, Aufgabe 3, Abbildung 9.6). Antwortoptionen können auch Elemente aus einer Grafik oder einem Text sein. Bei komplexen Mehrfachwahl-Antworten (47 Aufgaben) gibt es unterschiedliche Antwortmöglichkeiten: Die Schülerinnen und Schüler beurteilen für jede der dargebotenen Aussagen, ob diese richtig oder falsch ist, sie vervollständigen (mehrere) Lücken eines Satzes durch das Auswählen einer Antwort im Drop-down-Menü, sie wählen aus einer Liste von Antworten eine oder mehrere aus oder sie geben ihre Antworten durch Drag-and-Drop, um Antworten einander zuzuordnen, sie in eine Rangfolge zu bringen oder zu kategorisieren. Offene Antwortformate (35 Aufgaben) erfordern das Eingeben einer Wortgruppe, eines Satzes oder kurzen Absatzes (siehe Aufgabeneinheit *Energieeffiziente Häuser*, Aufgabe 2, Frage 2, Abbildung 9.5) oder es muss ein Graph oder

Diagramm gezeichnet werden. Je nach Komplexität der offenen Antworten werden diese entweder computerbasiert (3 Aufgaben) oder nachträglich anhand ausführlicher Anweisungen von geschulten Kodierern und Kodierern (32 Aufgaben) ausgewertet. Die Antworten werden den Kategorien richtig, teilweise richtig oder falsch zugeordnet (vgl. Kapitel 10 sowie Mang et al., 2019).

Tabelle 9.1: Verteilung der Aufgaben des Naturwissenschaftstests auf die drei Teilkompetenzen, Wissensbereiche und Wissenssysteme

Bereich der naturwissenschaftlichen Kompetenz	Anzahl Aufgaben
<i>Teilkompetenzen</i>	115
Phänomene naturwissenschaftlich erklären	49
Naturwissenschaftliche Forschung bewerten und Untersuchungen planen	30
Daten und Evidenz naturwissenschaftlich interpretieren	36
<i>Wissensbereiche</i>	115
konzeptuelles Wissen	49
prozedurales Wissen	47
epistemisches Wissen	19
<i>Wissenssysteme</i>	115
Physikalische Systeme	38
Lebende Systeme	47
Erd- und Weltraumssysteme	30

Die Abbildungen 9.2 bis 9.6 zeigen interaktive Beispielaufgaben einer Aufgabeneinheit, die in PISA 2015 im Feldtest eingesetzt wurde (vgl. Schiepe-Tiska, Rönnebeck et al., 2016). Ähnliche Aufgaben wurden auch 2018 von den Schülerinnen und Schülern bearbeitet. Weitere Beispielaufgaben, auch aus anderen Erhebungsrounden, sind hier zu finden: www.pisa.tum.de/beispielaufgaben/.

Bei der Aufgabeneinheit *Energieeffiziente Häuser* sollen die Jugendlichen mithilfe einer Simulation herausfinden, wie sich unterschiedliche Dachfarben unter Berücksichtigung der Außentemperatur auf den Energieverbrauch¹ eines Hauses auswirken. Die Aufgabeneinheit ist dem Wissenssystem *physikalische Systeme* zugeordnet. Ihre Lösung erfordert unter anderem konzeptuelles Wissen zu Strahlungsenergie Absorption, Reflexion und Wärme. Die Aufgabeneinheit beginnt mit einer Einleitung, die den Kontext (*regional/national – natürliche Ressourcen*) beschreibt (Abbildung 9.2) und die Simulation erläutert, welche die Jugendlichen an dieser Stelle ausprobieren sollen (Abbildung

1 Der Begriff „Energieverbrauch“ wird in dieser Aufgabe, obwohl aus physikalischer Sicht nicht angemessen, bewusst verwendet, da es ein gängiger Begriff der Alltagssprache ist. PISA zielt explizit darauf ab, zu erfassen, inwieweit Jugendliche in der Lage sind, ihre Kompetenzen in Alltagskontexten anzuwenden, das heißt, inwieweit Schülerinnen und Schüler die Aufgabe trotz Verwendung des Begriffs „Energieverbrauch“ korrekt lösen können oder ob durch die Verwendung des Begriffs möglicherweise Alltagsvorstellungen von Energie ausgelöst werden, die Schülerinnen und Schüler in der korrekten Bearbeitung der Aufgabe behindern.

9.3). Die drei folgenden Aufgaben, erfordern unterschiedliche Teilkompetenzen sowie konzeptuelles und prozedurales Wissen und sind verschiedenen kognitiven Anforderungsniveaus zugeordnet (Tabelle 9.2).

Tabelle 9.2: Einordnung der Beispielaufgaben der Aufgabeneinheit *Energieeffiziente Häuser*

	Aufgabe 1	Aufgabe 2		Aufgabe 3
		Frage 1	Frage 2	
Teilkompetenz	Daten und Evidenz naturwissenschaftlich interpretieren	Daten und Evidenz naturwissenschaftlich interpretieren	Phänomene naturwissenschaftlich erklären	Daten und Evidenz naturwissenschaftlich interpretieren
Wissensbereich	prozedurales Wissen	prozedurales Wissen	konzeptuelles Wissen	konzeptuelles Wissen
Kognitives Anforderungsniveau	niedrig	mittel	mittel	hoch
Antwortformat	offene Antwort	komplexe Mehrfachwahl-Antwort	offene Antwort	einfache Auswahl-Antwort

PISA 2015

Energieeffiziente Häuser
Einleitung

Lies die Einleitung. Klicke dann auf den WEITER-Pfeil.

ENERGIEEFFIZIENTE HÄUSER

Es besteht ein weltweit steigendes Interesse, energieeffiziente Häuser zu bauen. Durch eine Verringerung des Energieverbrauchs können Eigentümer Geld sparen und die Emission von Treibhausgasen in die Atmosphäre kann verringert werden. Architekten können Simulationen verwenden, um die Auswirkungen zu untersuchen, die unterschiedliche Entscheidungen beim Entwurf eines Hauses auf den Energieverbrauch haben.

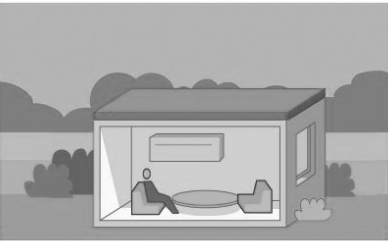


Abbildung 9.2: Aufgabeneinheit *Energieeffiziente Häuser*, Einleitung

PISA 2015
☐☐☐☐☐☐
🕒
?
⏪
⏩

Energieeffiziente Häuser

Einleitung

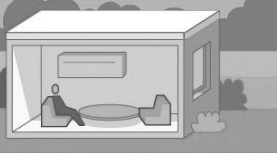
Mit dieser Simulation kannst du untersuchen, wie verschiedene Dachfarben den Energieverbrauch beeinflussen. Ein Teil der Sonneneinstrahlung, die auf das Dach trifft, wird reflektiert. Ein Teil der Sonneneinstrahlung wird absorbiert und erwärmt das Haus.

Das Haus in der Simulation verbraucht Energie sowohl zum Heizen als auch zum Kühlen, um bei unterschiedlichen Außentemperaturen immer eine angenehme Innentemperatur von 23 °C im Haus aufrechtzuerhalten.


Um zu sehen, wie die Steuerelemente in dieser Simulation funktionieren, folge diesen Schritten:

1. Klicke auf eine **Dachfarbe**.
2. Klicke auf eine **Außentemperatur**.
3. Klicke auf „Ausführen“, um zu sehen, was mit dem Energieverbrauch passiert. Die Ergebnisse werden in der Tabelle angezeigt.

Hinweis: Die verbrauchte Energie wird in Wattstunden gemessen. Eine Wattstunde entspricht der Energiemenge, die bei einem Watt Leistung in einer Stunde verbraucht wird.




Energieverbrauch



Wattstunden

Dachfarbe



Innentemperatur 23 °C

Außentemperatur (C°)

● 0 ○ 10 ○ 20 ○ 30 ○ 40

Ausführen

Außentemperatur (C°)	Dachfarbe	Energieverbrauch (Wattstunden)

Abbildung 9.3: Aufgabeneinheit *Energieeffiziente Häuser*, Erläuterung der Simulation

PISA 2015
☐☐☐☐☐☐
🕒
?
⏪
⏩

Energieeffiziente Häuser


Frage 1 / 4


► So führst du die Simulation aus


Führe die Simulation aus, um Daten anhand der Informationen unten zu erhalten. Verwende Drag & Drop und wähle dann Daten in der Tabelle aus, um die Frage zu beantworten.

Einige Häuser werden in einer Gegend mit sehr heißem Klima, häufig mit Außentemperaturen von 40 °C und mehr, gebaut. Man hat dich darum gebeten, bei der Entscheidung zu helfen, welche Dachfarbe am besten für diese Häuser verwendet werden soll.

Ordne die drei Dachfarben nach **sinkendem** Energieverbrauch für ein Haus, das bei sehr heißem Klima auf 23 °C abgekühlt wird.










Energieverbrauch

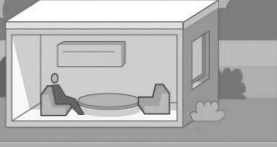
Höchster
→
Niedrigster








✳ Wähle drei Zeilen mit Daten in der Tabelle aus, um deine Antwort zu stützen.




Energieverbrauch



Wattstunden

Dachfarbe



Innentemperatur 23 °C

Außentemperatur (C°)

● 0 ○ 10 ○ 20 ○ 30 ○ 40

Ausführen

Außentemperatur (C°)	Dachfarbe	Energieverbrauch (Wattstunden)

Abbildung 9.4: Aufgabeneinheit *Energieeffiziente Häuser*, Aufgabe 1

PISA 2015

Energieeffiziente Häuser

Frage 2 / 4

► So führst du die Simulation aus

Führe die Simulation aus, um Daten anhand der Informationen unten zu erhalten. Wähle aus den Drop-down-Menüs aus, wähle Daten in der Tabelle aus und gib dann eine Erklärung ein, um die Frage zu beantworten.

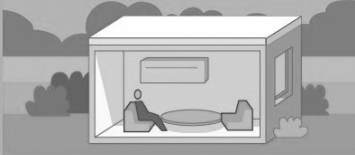
Wenn die Außentemperatur bei 10 °C liegt, was ist dann der Unterschied beim Energieverbrauch zwischen einem Haus mit einem weißen Dach und einem Haus mit einem schwarzen Dach?


Bei 10 °C verbraucht ein Haus mit einem weißen Dach

Wähle Energie als ein Haus mit einem schwarzen Dach.

★ Wähle zwei Zeilen mit Daten in der Tabelle aus, um deine Antwort zu stützen.

Erkläre den Unterschied beim Energieverbrauch, indem du beschreibst, was mit der Sonneneinstrahlung passiert, wenn sie auf Dächer mit diesen zwei unterschiedlichen Farben trifft.






Dachfarbe

Innentemperatur 23 °C

Außentemperatur (C°)

● 0 ○ 10 ○ 20 ○ 30 ○ 40



Ausführen

Außentemperatur (C°)	Dachfarbe	Energieverbrauch (Wattstunden)

Abbildung 9.5: Aufgabeneinheit *Energieeffiziente Häuser*, Aufgabe 2

Bei Aufgabe 1 (Abbildung 9.4) soll ein Haus in einer Gegend mit sehr hohen Außentemperaturen auf 23°C abgekühlt werden. Je nach Dachfarbe (weiß, rot, schwarz) wird dafür unterschiedlich viel Energie benötigt. Die Jugendlichen sind aufgefordert, die Dachfarben nach dem Energieverbrauch zu ordnen. Dazu können sie sich, nachdem sie eine Außentemperatur von 40°C ausgewählt haben, für jede Dachfarbe den Energieverbrauch anzeigen lassen. Anschließend sollen sie die Dachfarben in eine Reihenfolge bringen, und drei Zeilen aus der Tabelle auswählen, um ihre Auswahl zu unterstützen. Die korrekte Reihenfolge ist schwarz (höchster Energieverbrauch bei 40°C), rot, weiß.

Für die Beantwortung von Aufgabe 2 (Abbildung 9.5) sollen die Schülerinnen und Schüler mithilfe der Simulation den Energieverbrauch eines Hauses mit einem weißen und einem schwarzen Dach bei einer gewünschten Innentemperatur von 23°C und einer Außentemperatur von 10°C vergleichen und ihre Antwort erneut mit Daten aus der Simulation belegen. In der ersten Frage sollen sie den Satz vervollständigen „Bei 10 °C verbraucht ein Haus mit einem weißen Dach ___ Energie als ein Haus mit einem schwarzen Dach“, indem sie aus einem Drop-down-Menü „mehr“ (richtige Antwort) oder „weniger“ auswählen. Zeilen mit der Außentemperatur 10°C sowie den Dachfarben weiß und schwarz würden diese Antwort belegen. Anschließend sollen die Jugendlichen in Frage 2 mit eigenen Worten den Unterschied im Energieverbrauch bei weißer und schwarzer Dachfarbe erläutern und dabei die Wirkung der Sonneneinstrahlung berücksichtigen.

sichtigen. Korrekt wäre zum Beispiel die Antwort, dass die Strahlungsenergie der Sonne von schwarzen Dächern in höherem Maße absorbiert wird als von weißen Dächern.

In Aufgabe 3 (Abbildung 9.6) soll die Abhängigkeit des Energieverbrauchs von der Außentemperatur über alle drei Dachfarben hinweg auf den gesamten Temperaturbereich verallgemeinert werden. Von den vier präsentierten möglichen Antworten ist Option 3 richtig: „Wenn sich der Unterschied zwischen der Außentemperatur und der Innentemperatur vergrößert, steigt der Energieverbrauch.“

Energieeffiziente Häuser
Frage 4 / 4

► So führst du die Simulation aus

Führe die Simulation aus, um Daten anhand der Informationen unten zu erhalten. Klicke eine Antwort an, um die Frage zu beantworten.

Welchen Schluss kannst du, anhand der Simulation, hinsichtlich der Beziehung zwischen Außentemperatur und Energieverbrauch für den gesamten Temperaturbereich und alle drei Dachfarben ziehen?

- Wenn die Außentemperatur steigt, steigt der Energieverbrauch.
- Wenn die Außentemperatur sinkt, steigt der Energieverbrauch.
- Wenn sich der Unterschied zwischen der Außentemperatur und der Innentemperatur vergrößert, steigt der Energieverbrauch.
- Wenn sich der Unterschied zwischen der Außentemperatur und der Innentemperatur verringert, steigt der Energieverbrauch.

Dachfarbe:

Innentemperatur 23 °C
Außentemperatur (C°) 0 10 20 30 40

Energieverbrauch
Wattstunden

Außentemperatur (C°)	Dachfarbe	Energieverbrauch (Wattstunden)

Ausführen

Abbildung 9.6: Aufgabeneinheit *Energieeffiziente Häuser*, Aufgabe 3

Auswertung des Tests und Kompetenzstufen

Bei PISA 2018 wird aufgrund des Status als Nebendomäne nur eine Gesamtskala für die naturwissenschaftliche Kompetenz gebildet. Als die Naturwissenschaften bei PISA 2006 erstmals Hauptdomäne waren, wurde der OECD-Mittelwert auf 500 und die Standardabweichung auf 100 Punkte festgelegt. Diese Skala bildet nach wie vor den Bezugspunkt. Die Antworten der Schülerinnen und Schüler werden mithilfe psychometrischer Modelle der Item-Response-Theorie skaliert (vgl. Kapitel 10). Dies ermöglicht die Abbildung der Schwierigkeit einer Aufgabe und der Fähigkeit eines Jugendlichen auf einer gemeinsamen Skala. Diese Skala wird in Abschnitte unterteilt, die Aufgaben mit vergleichbaren Schwierigkeiten zusammenfassen.

Tabelle 9.3: Stufen naturwissenschaftlicher Kompetenz in PISA 2018

Kompetenzstufe	Wozu die Schülerinnen und Schüler auf der jeweiligen Kompetenzstufe im Allgemeinen in der Lage sind
VI ≥ 708 Punkte	Auf Kompetenzstufe VI können Schülerinnen und Schüler auf vielfältige, miteinander in Beziehung stehende Ideen und Konzepte aus den Naturwissenschaften zurückgreifen und konzeptuelles, prozedurales und epistemisches Wissen anwenden, um erklärende Hypothesen für unbekannte naturwissenschaftliche Phänomene, Ereignisse und Prozesse aufzustellen oder Vorhersagen zu treffen. Beim Interpretieren von Daten und Evidenz sind sie in der Lage, zwischen relevanten und irrelevanten Informationen zu unterscheiden, und können auf Wissen, welches sie außerhalb des regulären Schulcurriculums erworben haben, zurückgreifen. Sie können Argumente, die auf naturwissenschaftlicher Evidenz und naturwissenschaftlichen Theorien beruhen, von solchen, die auf anderen Annahmen beruhen, unterscheiden. Schülerinnen und Schüler auf Kompetenzstufe VI können alternative Designs komplexer Experimente, Feldstudien oder Simulationen bewerten und ihre Entscheidung für ein bestimmtes Design begründen.
V 634–707 Punkte	Auf Kompetenzstufe V können Schülerinnen und Schüler abstrakte naturwissenschaftliche Ideen und Konzepte verwenden, um unbekannte und komplexere Phänomene, Ereignisse und Prozesse, die mehrere kausale Verknüpfungen beinhalten, zu erklären. Sie sind in der Lage, anspruchsvolleres epistemisches Wissen anzuwenden, um alternative experimentelle Designs zu bewerten und ihre Entscheidung zu begründen, sowie theoretisches Wissen zu nutzen, um Informationen zu interpretieren oder Vorhersagen zu treffen. Schülerinnen und Schüler auf Kompetenzstufe V können verschiedene Ansätze einer gegebenen Fragestellung naturwissenschaftlich bearbeiten, bewerten sowie die Grenzen der Interpretation von Daten und Quellen und die Effekte von Unsicherheiten in naturwissenschaftlichen Daten erkennen.
IV 559–633 Punkte	Auf Kompetenzstufe IV können Schülerinnen und Schüler komplexeres oder abstrakteres konzeptuelles Wissen, das entweder vorgegeben ist oder abgerufen werden muss, verwenden, um Erklärungen für komplexe oder weniger bekannte Ereignisse und Prozesse zu entwickeln. Sie können Experimente mit zwei oder mehr unabhängigen Variablen in einem beschränkten Kontext durchführen. Sie sind in der Lage, ein experimentelles Design auf Basis ihres prozeduralen und epistemischen Wissen zu begründen. Schülerinnen und Schüler auf Kompetenzstufe IV können Daten aus mäßig komplexen Datensätzen oder weniger bekannten Kontexten interpretieren, über die Daten hinausgehende angemessene Schlussfolgerungen ziehen und ihre Entscheidungen begründen.
III 485–558 Punkte	Auf Kompetenzstufe III können Schülerinnen und Schüler auf mäßig komplexes konzeptuelles Wissen zurückgreifen, um Erklärungen für bekannte Phänomene zu erkennen oder zu entwickeln. In weniger bekannten oder komplexeren Situationen können sie Erklärungen mithilfe relevanter Hinweise oder Unterstützung entwickeln. Sie können auf Aspekte von prozeduralem oder epistemischem Wissen zurückgreifen, um einfache Experimente in einem beschränkten Kontext durchzuführen. Schülerinnen und Schüler auf Kompetenzstufe III sind in der Lage, zwischen naturwissenschaftlichen und nicht-naturwissenschaftlichen Inhalten zu unterscheiden und Evidenz, die eine naturwissenschaftliche Aussage unterstützt, zu erkennen.
II 410–484 Punkte	Schülerinnen und Schüler auf Kompetenzstufe II können auf alltagsrelevantes konzeptuelles Wissen sowie grundlegendes prozedurales Wissen zurückgreifen, um eine angemessene naturwissenschaftliche Erklärung zu identifizieren, Daten zu interpretieren und die Fragestellung zu erkennen, die mit einem einfachen Experiment beantwortet werden soll. Sie können grundlegendes oder alltägliches naturwissenschaftliches Wissen nutzen, um valide Schlussfolgerungen, die auf einfachen Datensätzen beruhen, zu erkennen. Schülerinnen und Schüler auf Kompetenzstufe II zeigen grundlegendes epistemisches Wissen indem sie in der Lage sind, Fragestellungen zu erkennen, die naturwissenschaftlich untersucht werden können.
Ia 335–409 Punkte	Schülerinnen und Schüler auf Kompetenzstufe Ia sind in der Lage, grundlegendes oder alltagsrelevantes konzeptuelles Wissen und prozedurales Wissen zu nutzen, um Erklärungen für einfache naturwissenschaftliche Phänomene zu erkennen. Mit Unterstützung können sie vorstrukturierte naturwissenschaftliche Untersuchungen mit maximal zwei Variablen durchführen. Sie sind in der Lage, einfache kausale oder korrelative Zusammenhänge zu erkennen und graphische oder visuelle Daten, die einem geringen kognitiven Anforderungsniveau entsprechen, zu interpretieren. Schülerinnen und Schüler auf Kompetenzstufe Ia können die beste naturwissenschaftliche Erklärung für vorgegebene Daten in bekannten persönlichen, regionalen/nationalen und globalen Kontexten auswählen.
Ib 260–334 Punkte	Schülerinnen und Schüler auf Kompetenzstufe Ib können sehr grundlegendes naturwissenschaftliches Wissen beziehungsweise naturwissenschaftliches Alltagswissen nutzen, um Aspekte bekannter oder sehr einfacher naturwissenschaftlicher Phänomene zu erkennen. Sie sind in der Lage einfache Muster in Daten sowie grundlegende naturwissenschaftliche Begriffe zu erkennen und expliziten Instruktionen zu folgen, um eine wissenschaftliche Arbeitsweise auszuführen.

Diese Abschnitte werden als Kompetenzstufen bezeichnet und aufbauend auf der überarbeiteten Rahmenkonzeption von PISA 2015 inhaltlich charakterisiert (Tabelle 9.3). PISA unterscheidet sechs Kompetenzstufen. Seit PISA 2015 wird außerdem die unterste Kompetenzstufe I noch einmal in Ia und Ib unterteilt, um im unteren Leistungsbereich besser differenzieren zu können. Mit steigender Kompetenzstufe werden die Anforderungen, welche die Aufgaben an die Fähigkeiten der Jugendlichen stellen, komplexer. Auf den unteren Kompetenzstufen erfordern die Aufgaben grundlegendes Wissen in bekannten Kontexten. Zur Lösung von Aufgaben auf den oberen Kompetenzstufen müssen Jugendliche hingegen auf abstrakte Konzepte und Ideen zurückgreifen, um mit deren Hilfe unbekannte, komplexe naturwissenschaftliche Phänomene, Prozesse und Ereignisse zu erklären.

Das Mindestniveau, das alle fünfzehnjährigen Schülerinnen und Schüler erreichen sollten (Breitenförderung im Sinne von *Naturwissenschaften für alle*), entspricht Kompetenzstufe II. Jugendlichen mit geringerer Kompetenz fehlen grundlegende naturwissenschaftliche Kompetenzen, um im Alltag einfache naturwissenschaftliche Probleme und Zusammenhänge zu verstehen oder aus Daten und Informationen Schlussfolgerungen zu ziehen. Dies ist in einer Welt, die zunehmend Informationen zur Verfügung stellt, deren Verlässlichkeit selbstständig eingeschätzt werden muss, kritisch zu betrachten. Jugendliche, die die Kompetenzstufen V und VI erreichen, zeichnen sich hingegen durch eine hoch ausgeprägte naturwissenschaftliche Kompetenz aus. Sie sind in der Lage, auch Aufgaben zu lösen, die besonders hohe fachliche Anforderungen stellen. Sie gelten als leistungsstark und bringen hervorragende Voraussetzungen für naturwissenschaftlich-technische Studiengänge mit (Spitzenförderung).

9.2 Naturwissenschaftliche Kompetenz im internationalen Vergleich

Im Folgenden werden zunächst die Mittelwerte und Streuungsmaße der PISA 2018 Teilnehmerstaaten berichtet, um einen ersten Eindruck über die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler in Deutschland im internationalen Vergleich zu gewinnen. Anschließend wird anhand der Verteilung der Jugendlichen auf die Kompetenzstufen differenziert analysiert, inwieweit es in Deutschland gelingt, Jugendliche in den Naturwissenschaften in der Breite (das Erreichen der Kompetenzstufe II) sowie in der Spitze (Kompetenzstufen V und VI) zu fördern.

9.2.1 Mittelwerte und Streuungen der naturwissenschaftlichen Kompetenz

Abbildung 9.7 gibt einen Überblick über die Mittelwerte und Streuungen der naturwissenschaftlichen Kompetenz in den 37 OECD-Staaten (siehe auch Tabelle 9.1.web im Online-Anhang für die OECD-Partnerstaaten, vgl. OECD 2019b). Bei PISA 2018 beträgt die durchschnittliche naturwissenschaftliche Kompetenz über alle OECD-Staaten hinweg 489 Punkte. Mit Bezug auf diesen Mittelwert lassen sich in Abbildung 9.7 drei Gruppen unterscheiden: Staaten, deren Mittelwerte sich nicht signifikant vom OECD-Durchschnitt unterscheiden, finden sich im mittleren Bereich der Tabelle (zwischen den beiden fettgedruckten Linien), Staaten im unteren Tabellenbereich (unterhalb der unteren fettgedruckten Linie) zeigen signifikant niedrigere Kompetenzwerte als der OECD-Durchschnitt, Staaten im oberen Bereich (oberhalb der oberen fettgedruckten Linie) erreichen mittlere Kompetenzwerte signifikant über dem Durchschnitt aller OECD-Staaten. Zu dieser letztgenannten Gruppe zählt auch Deutschland mit einem Mittelwert von 503 Punkten.

Insgesamt erreichen 20 OECD-Staaten mittlere Kompetenzwerte signifikant über dem OECD-Durchschnitt. Dabei ist jedoch zu beachten, dass die Rangfolge der Staaten nur bedingt aussagekräftig ist, da sich die Punktwerte einzelner Staaten oft nicht statistisch signifikant voneinander unterscheiden. Wie schon bei PISA 2015 ist es daher sinnvoller, Gruppen von Staaten mit statistisch nicht signifikant voneinander abweichenden Mittelwerten zu betrachten. Im oberen Bereich der Tabelle bilden Estland (530 Punkte) und Japan (529 Punkte) die Spitzengruppe, gefolgt von einer Gruppe mit Finnland (522 Punkte), der Republik Korea (519 Punkte) und Kanada (518 Punkte). Deutschland (503 Punkte) ist Teil einer weiteren, vergleichsweise homogenen Gruppe, die zwölf Staaten umfasst. Sie beginnt mit Neuseeland (508 Punkte) und endet mit der Schweiz (495 Punkte). Vor dieser Gruppe findet sich Polen (511 Punkte), was sich zwar nicht signifikant von Neuseeland (508 Punkte), Slowenien (507) und dem Vereinigten Königreich (505 Punkte) unterscheidet, wohl aber von den anderen Staaten dieser Gruppe. Den Abschluss der Staaten oberhalb des OECD-Durchschnitts bilden Frankreich und Dänemark mit je 493 Punkten.

Die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler in Deutschland zeigt sich damit bei PISA 2018 international anschlussfähig, erreicht aber nicht die Ergebnisse der beiden Spitzengruppen. Das Bild verschärft sich, wenn man zusätzlich die Ergebnisse der OECD-Partnerstaaten betrachtet (siehe Tabelle 9.1.web). Der Zusammenschluss der vier chinesischen Verwaltungseinheiten Peking, Shanghai, Jiangsu und Zhejiang (im Folgenden B-S-J-Z China) als kompetenzstärkstes Bildungssystem liegt mit einem mittleren Kompetenzwert von 590 Punkten noch einmal 60 Punkte, das heißt mehr als eine halbe Standardabweichung oder nahezu eine Kompetenzstufe, über dem OECD-Spitzenreiter Estland. Es folgen Singapur und Macau (China), wo mit 551 beziehungsweise 544 Punkten mittlere Kompetenzwerte erreicht werden, die den Spitzenwert der OECD-Staaten ebenfalls deutlich übertreffen. Auch Hongkong (China) und Chine-

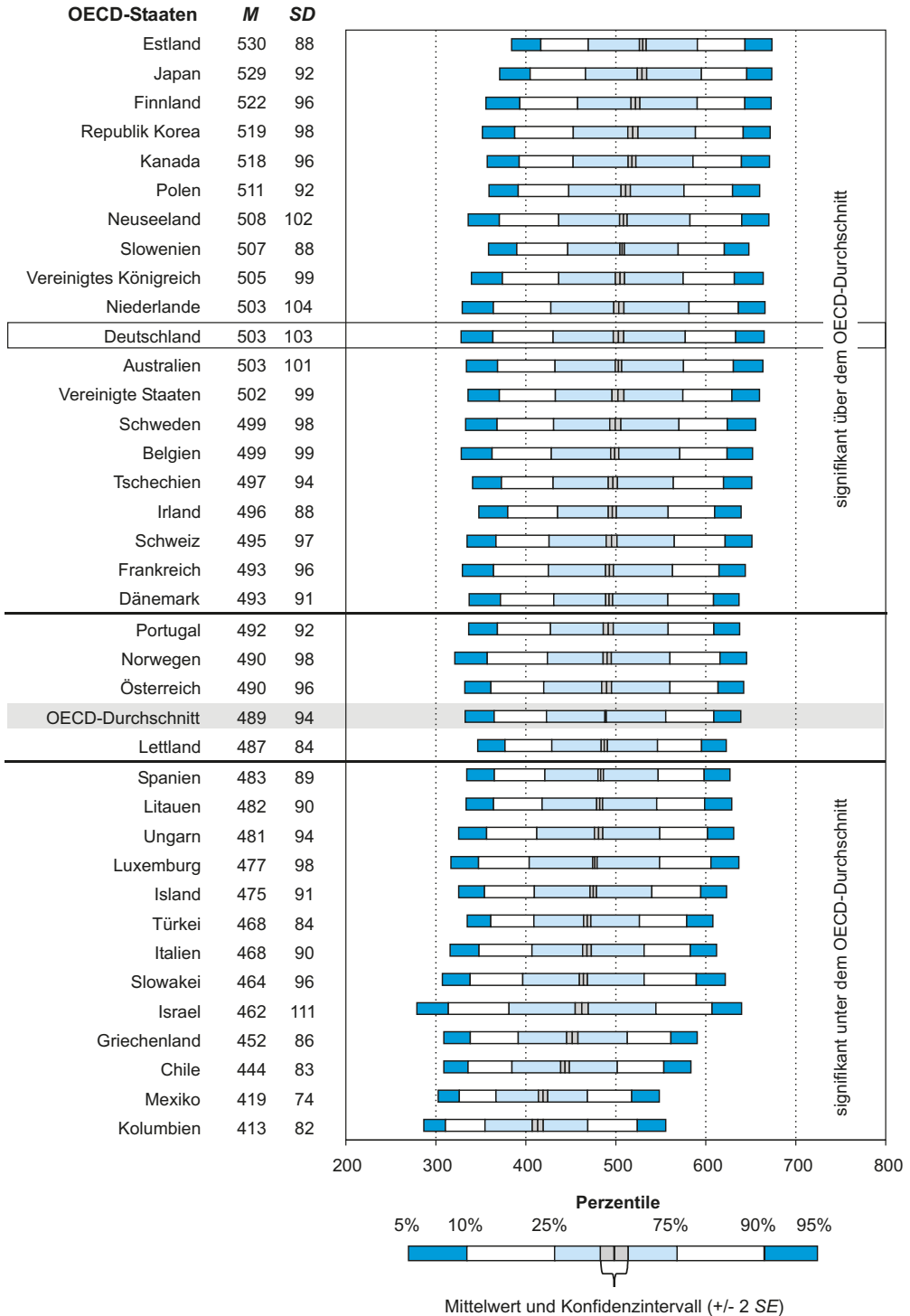


Abbildung 9.7: Mittelwerte, Streuungen und Perzentilbänder der naturwissenschaftlichen Kompetenz der OECD-Staaten

sisch Taipeh zeigen mit 517 beziehungsweise 516 Punkten Ergebnisse, die ähnlich zu denen der Spitzengruppen der OECD-Staaten sind.

Die Gruppe der Staaten, deren Mittelwerte sich nicht signifikant vom OECD-Mittelwert unterscheiden, besteht aus Portugal, Norwegen, Österreich und Lettland. Kompetenzwerte signifikant unterhalb des OECD-Durchschnitts werden in 13 Staaten erreicht, darunter Spanien (483 Punkte), Ungarn (481 Punkte) und Israel (462 Punkte). Schlusslicht der Kompetenzverteilung ist Kolumbien mit 413 Punkten. Innerhalb der OECD-Staaten beträgt der Unterschied zwischen dem kompetenzstärksten und -schwächsten Staat somit gut eineinhalb Kompetenzstufen. Mit Ausnahme der oben genannten asiatischen Staaten befinden sich alle OECD-Partnerstaaten im Bereich unterhalb des OECD-Durchschnitts – viele davon sehr deutlich (siehe Tabelle 9.1.web). Schlusslichter sind hier der Kosovo und Panama (je 365 Punkte), die Philippinen (357 Punkte) und mit noch einmal 21 Punkten Abstand die Dominikanische Republik (336 Punkte).

Über die Heterogenität der Kompetenzverteilung in einem Staat gibt der Kennwert der Standardabweichung Aufschluss. Die größte Streubreite findet sich innerhalb der OECD-Staaten in Israel ($SD = 111$), gefolgt von den Niederlanden ($SD = 104$), Deutschland ($SD = 103$), Neuseeland ($SD = 102$) und Australien ($SD = 101$). In all diesen Staaten liegt die Streuung signifikant über dem OECD-Mittelwert ($SD = 94$). Eine deutlich größere Homogenität in der Kompetenzverteilung als in Deutschland findet sich in der Gruppe der kompetenzstarken Staaten insbesondere in Estland und Slowenien (jeweils $SD = 88$) sowie Japan und Polen (jeweils $SD = 92$). In diesen Staaten scheint eine Breitenförderung in den Naturwissenschaften auch auf hohem Niveau besser zu gelingen. Gleiches gilt in noch höherem Maße für die OECD-Partnerstaaten B-S-J-Z (China), Macau (China) und Hongkong (China) mit Standardabweichungen zwischen 83 und 86 Punkten (siehe Tabelle 9.1.web).

Eine grafische Darstellung der Streuung zeigen die Perzentilbänder in Abbildung 9.7 (Punktwerte in Tabelle 9.1.web). Sie teilen die Kompetenzverteilung in definierte Prozentbereiche auf. So beschreibt der Punktwert für das 95. Perzentil den Kompetenzwert, bei dem die Gruppe der fünf Prozent kompetenzstärksten Jugendlichen eines Staates beginnt, und liefert somit einen Hinweis, wie gut in dem betreffenden Staat die Spitzenförderung gelingt. In den Staaten an der Spitze der OECD-Kompetenzverteilung zeigt sich ein sehr einheitliches Bild. In Estland beginnt diese Gruppe bei 674 Punkten, es folgen Japan und Finnland mit 673 Punkten und die Republik Korea mit 672 Punkten. In Deutschland liegt der Wert mit 665 Punkten nur etwas niedriger als in den Spitzenstaaten. Er ist vergleichbar mit dem Punktwert, der in der Rangfolge benachbarten Staaten Niederlande (666 Punkte), Vereinigtes Königreich und Australien (jeweils 664 Punkte). Im Bereich der Spitzenförderung sind die Unterschiede zwischen den kompetenzstärksten OECD-Staaten also vergleichsweise gering. In einer eigenen Liga spielt allerdings wiederum der OECD-Partnerstaat B-S-J-Z (China) mit 721 Punkten.

Deutlichere Unterschiede zwischen den OECD-Staaten finden sich im unteren Bereich der Kompetenzverteilung. Dies zeigt der Punktwert für das fünfte Perzentil an,

das heißt der Kompetenzwert, den 95 Prozent der Schülerinnen und Schüler erreichen. Hier zeigt insbesondere Estland Stärken (384 Punkte), gefolgt von Japan (371 Punkte) und, mit schon deutlich größerem Abstand, Polen und Slowenien (je 359 Punkte). Ein nochmal deutlich höheres Niveau erreichen wiederum 95 Prozent der Fünfzehnjährigen in den beiden OECD-Partnerstaaten B-S-J-Z (China) mit 448 Punkten und Macau (China) mit 402 Punkten. In der Gruppe der zwölf von Deutschland im Mittelwert nicht signifikant verschiedenen OECD-Staaten erreicht Deutschland zusammen mit Belgien (je 328 Punkte) und den Niederlanden (329 Punkte) die niedrigsten Werte. Ein deutlich höheres Niveau, das 95 Prozent der Jugendlichen erreichen, findet sich neben Slowenien im Vereinigten Königreich (340 Punkte), in Tschechien (341 Punkte) und insbesondere auch in Irland (348 Punkte). In diesen Staaten scheint die Förderung kompetenzschwacher Jugendlicher vergleichsweise besser zu gelingen als in Deutschland.

9.2.2 Verteilung auf die Stufen der naturwissenschaftlichen Kompetenz

Ein weiteres Maß, das die Streuung der Kompetenz innerhalb eines Staates verdeutlicht, ist die Verteilung der Schülerinnen und Schüler auf Kompetenzstufen. Diese erlaubt, neben der Vergleichsperspektive mit anderen Staaten, auch eine kriteriumsorientierte Interpretation der im Test erzielten Kompetenzwerte. Sie beschreibt inhaltlich, über welche Kompetenzen Jugendliche auf einzelnen Kompetenzstufen verfügen (siehe Abschnitt 9.1.2). Die Extrembereiche der Verteilung – das heißt die prozentualen Anteile der Schülerinnen und Schüler unterhalb der Kompetenzstufe II beziehungsweise auf den Kompetenzstufen V und VI – geben Auskunft über die Anteile besonders kompetenzschwacher beziehungsweise -starker Jugendlicher in den Naturwissenschaften. Während letztere als potenzieller Nachwuchs für Karrieren im naturwissenschaftlich-technischen Bereich betrachtet werden können, erfordert erstere Gruppe besondere Aufmerksamkeit. Ihre naturwissenschaftlichen Kompetenzen sind gering und es ist zu befürchten, dass sie nicht ausreichen werden, die heutigen Anforderungen in Beruf, Gesellschaft und Alltag zu bewältigen. Schülerinnen und Schülern unterhalb der Kompetenzstufe II fehlen die Kompetenzen, alltägliches Wissen zu nutzen, um einfache Daten zu interpretieren, Erklärungen zu entwickeln oder einfache, valide Schlussfolgerungen zu ziehen. Sie können somit weder aktuelle naturwissenschaftliche Diskussionen verfolgen und bewerten noch Konsequenzen bestimmter Handlungsoptionen einschätzen (siehe Abschnitt 9.1.2).

Die Verteilung der Schülerinnen und Schüler in den OECD-Staaten auf die Stufen der naturwissenschaftlichen Kompetenz ist in Abbildung 9.8 dargestellt. Betrachtet man zunächst den unteren Kompetenzbereich, so befinden sich im OECD-Durchschnitt 22.0 Prozent der Fünfzehnjährigen unterhalb der Kompetenzstufe II. Für Deutschland ist dieser Anteil mit 19.6 Prozent signifikant kleiner. In der Gruppe von Staaten, deren Mittelwerte sich von Deutschland nicht signifikant unterscheiden, variiert der Anteil zwischen 14.6 Prozent in Slowenien und 20.2 Prozent in der Schweiz. Slowenien weist damit, ebenso wie Irland (17.0 %) sowie die sechs Staaten an der Spitze der Kompetenz-

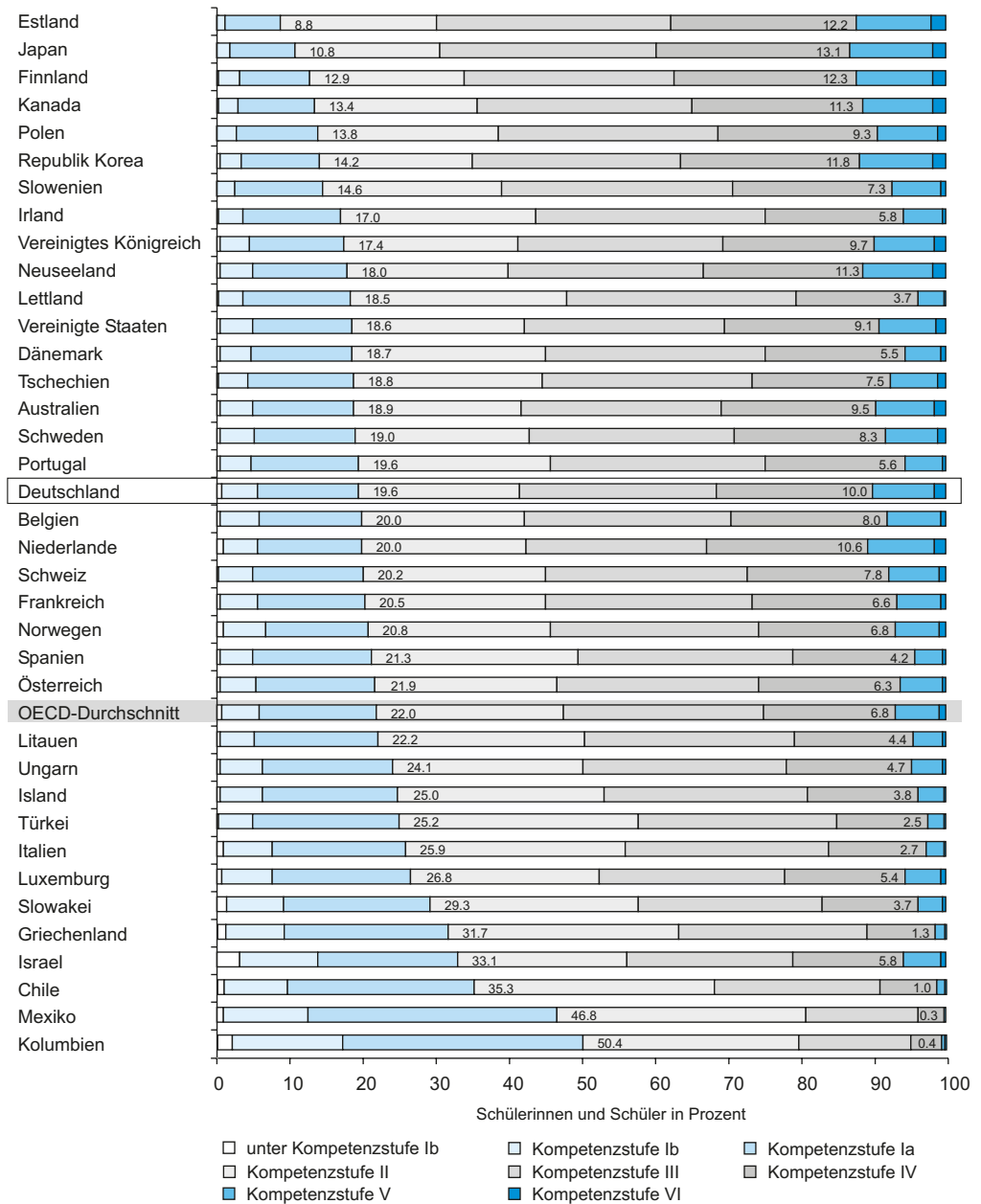


Abbildung 9.8: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI

verteilung, einen signifikant niedrigeren Anteil kompetenzschwacher Jugendlicher auf als Deutschland. Eine besonders erfolgreiche Förderung kompetenzschwacher Jugendlicher scheint – wie schon die Betrachtung der Streuung andeutete – in Estland (8.8 %) und Japan (10.8 %) zu gelingen.

Im oberen Kompetenzbereich, das heißt auf den Kompetenzstufen V und VI, beträgt der Anteil der Schülerinnen und Schüler im OECD-Durchschnitt 6.8 Prozent. Der Anteil in Deutschland ist hier mit 10.0 Prozent signifikant größer. Betrachtet man wiederum die Gruppe der Staaten, deren Mittelwerte sich von Deutschland nicht signifikant unterscheiden, so reichen die Anteile von 5.8 Prozent in Irland bis 11.3 Prozent in Neuseeland. Signifikant höhere Anteile auf den Kompetenzstufen V und VI als in Deutschland finden sich lediglich in den drei Spitzenstaaten Japan (13.1 %), Finnland (12.3 %) und Estland (12.2 %), in denen die höchsten Anteile innerhalb der OECD-Staaten zu beobachten sind.

Betrachtet man die Befunde zu Mittelwerten, Streuungen und Kompetenzstufen in der Zusammenschau, hat sich Deutschland im Kreis der Staaten etabliert, denen eine Förderung der naturwissenschaftlichen Kompetenz auf hohem Niveau gelingt. Eine Herausforderung bleibt nach wie vor die Förderung kompetenzschwacher Jugendlicher. Bildungssysteme wie Estland oder Japan zeigen, dass eine erfolgreiche Breitenförderung im Sinne von *Naturwissenschaften für alle*, die auch hoch kompetente Jugendliche nicht vernachlässigt, besser gelingen kann. Hier scheinen in Deutschland weitere Anstrengungen nötig, um zur internationalen Spitzengruppe aufzuschließen.

9.2.3 Unterschiede zwischen Mädchen und Jungen in der naturwissenschaftlichen Kompetenz

Damit Bildungssysteme ihre Potenziale vollständig ausschöpfen, ist es wichtig, Mädchen und Jungen gleichermaßen zu fördern. Eine zentrale Frage von PISA ist daher, ob sich Mädchen und Jungen in ihrer naturwissenschaftlichen Kompetenz unterscheiden. Abbildung 9.9 zeigt die naturwissenschaftlichen Kompetenzen von Mädchen und Jungen in den OECD-Staaten. In der Mehrheit der OECD-Staaten findet sich kein signifikanter Unterschied in der naturwissenschaftlichen Kompetenz von Mädchen und Jungen. Die gilt auch für Deutschland und die Mehrheit der Staaten aus der Gruppe mit von Deutschland nicht signifikant verschiedenen Kompetenzmittelwerten sowie für Japan, die Republik Korea und Polen. In zwölf Staaten erreichen die Mädchen wie auch im OECD-Durchschnitt signifikant höhere mittlere Kompetenzwerte. Die größten Differenzen zugunsten der Mädchen finden sich in Finnland (24 Punkte) und Israel (19 Punkte), aber auch in Estland, Schweden, den Niederlanden und Slowenien beträgt der Vorsprung der Mädchen zwischen fünf und zehn Punkten. Lediglich in Kolumbien und Mexiko weisen die Jungen signifikant höhere mittlere Kompetenzwerte auf.

Neben den Mittelwerten ist auch die Verteilung von Mädchen und Jungen auf die Randbereiche der Kompetenzstufenverteilung von Interesse, da sie Hinweise auf einen

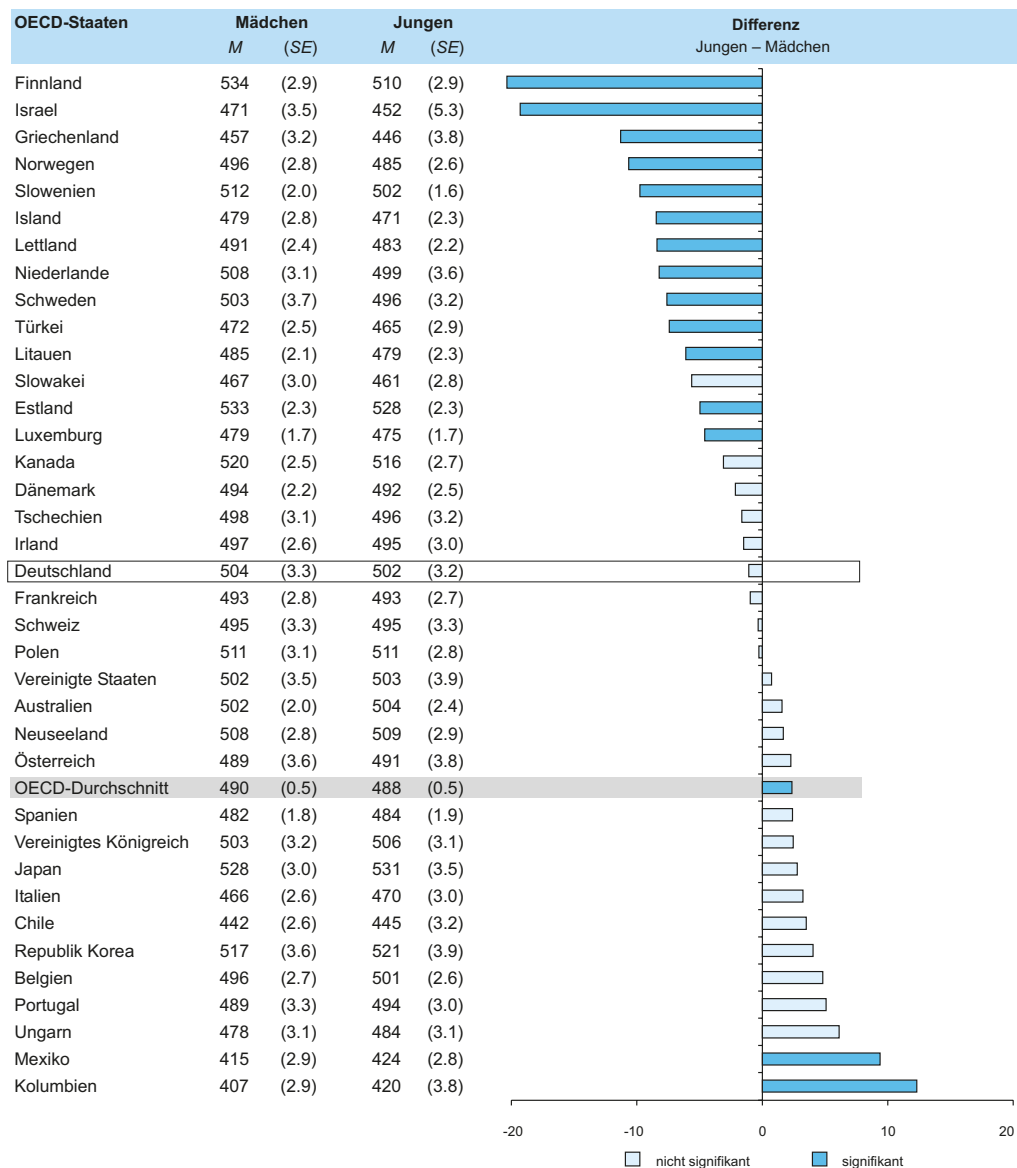


Abbildung 9.9: Mittelwerte naturwissenschaftlicher Kompetenz nach Geschlecht in den OECD-Staaten

geschlechtsspezifischen Förderbedarf liefern kann. Auf den unteren Kompetenzstufen (Stufe Ia und darunter) befinden sich im OECD-Durchschnitt signifikant weniger Mädchen (20.8 %) als Jungen (23.2 %). Dieses Bild zeigt sich in einem Großteil der OECD-Staaten, darunter auch in Deutschland mit einem Anteil von 18.2 Prozent für die Mädchen und 20.8 Prozent für die Jungen. Unter den Spitzenstaaten sind die Unterschiede für Finnland, Kanada und Polen statistisch signifikant, in Finnland ist der Unterschied mit fast 8.0 Prozentpunkte zugunsten der Mädchen beträchtlich. In Estland, Japan

und der Republik Korea ist hingegen kein signifikanter Unterschied zu sehen. In der Gruppe der von Deutschland im Mittelwert der naturwissenschaftlichen Kompetenz nicht signifikant verschiedenen Staaten sind signifikant niedrigere Anteile der Mädchen auf den unteren Kompetenzstufen neben Deutschland für vier weitere Staaten (Neuseeland, Niederlande, Schweden und Slowenien) zu beobachten. Nur in zwei Staaten, Kolumbien und Mexiko, ist der Anteil der Mädchen unter Kompetenzstufe II höher als der der Jungen.

Betrachtet man das obere Ende der Kompetenzstufenverteilung, also die Anteile auf den Kompetenzstufen V und VI, so finden sich dort in 14 OECD-Staaten signifikant mehr Jungen als Mädchen. Die gilt auch für Deutschland mit einem Mädchenanteil von 8.7 Prozent gegenüber einem Jungenanteil von 11.1 Prozent. Besonders große Unterschiede finden sich in Neuseeland (3.3 Prozentpunkte mehr Jungen) sowie der Republik Korea und Japan (je 2.7 Prozentpunkte mehr Jungen). Nur in Finnland sind auf den oberen Kompetenzstufen statistisch signifikant mehr Mädchen vertreten als Jungen. Zusammenfassend lässt sich sagen, dass die Mädchen in der Gruppe der kompetenzstärksten Jugendlichen in fast allen OECD-Staaten unterrepräsentiert sind, sich in der Regel aber auch mehr Jungen als Mädchen auf den unteren Kompetenzstufen finden.

9.3 Vertiefende Analysen der naturwissenschaftlichen Kompetenz in Deutschland

Neben dem internationalen Vergleich erlaubt die Stichprobe bei PISA 2018 – wie auch schon in vorangegangenen Zyklen – einen vertieften Blick auf die Situation in Deutschland. Im Folgenden werden Analysen zu Kompetenzunterschieden zwischen den Schularten vorgestellt. Im Anschluss wird berichtet, wie sich die Kompetenz im Vergleich zu den Zyklen, als die Naturwissenschaften Hauptdomäne waren (PISA 2006 und 2015), verändert hat.

9.3.1 Unterschiede zwischen Schularten

Analog zu PISA 2015 wird im Folgenden nur zwischen dem Gymnasium und nicht gymnasialen Schularten unterschieden (vgl. Kapitel 1, Schiepe-Tiska, Rönnebeck et al., 2016). Bei PISA 2018 betrug die Bildungsbeteiligung am Gymnasium 35.5 Prozent, an den nicht gymnasialen Schularten 59.7 Prozent. Der Rest verteilte sich auf Berufliche Schulen (3.0%) sowie Sonder- und Förderschulen (1.8%).

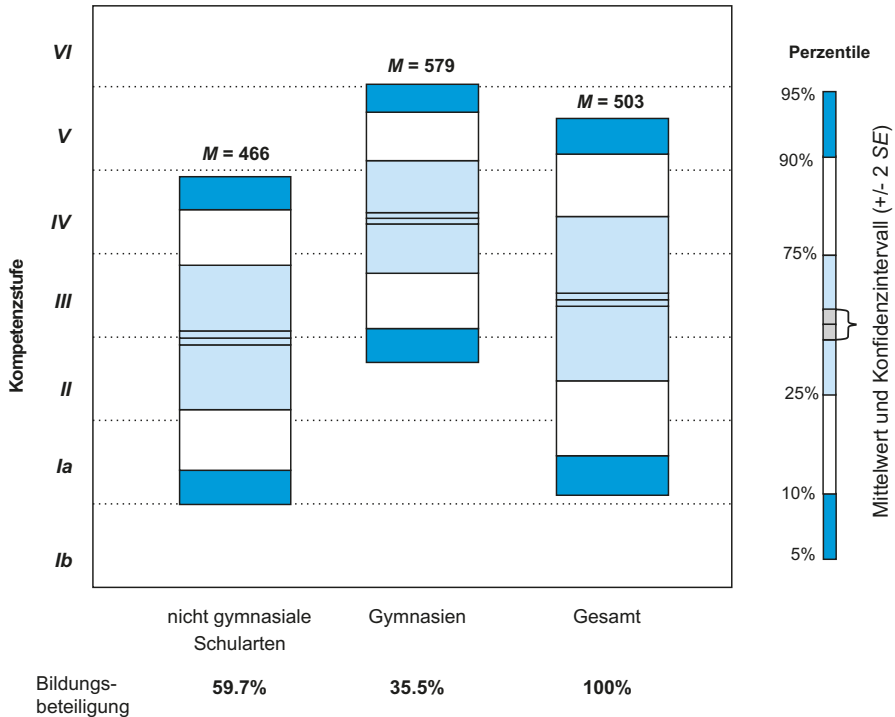


Abbildung 9.10: Perzentilbänder naturwissenschaftlicher Kompetenz in Deutschland für die Gesamtstichprobe und nach Schulart

Mittelwerte und Streuungen der naturwissenschaftlichen Kompetenz

Die Mittelwerte und Streuungen der naturwissenschaftlichen Kompetenz in Deutschland für die Gesamtstichprobe sowie getrennt nach Gymnasien und nicht gymnasialen Schularten sind in Abbildung 9.10 dargestellt. Jugendliche an Gymnasien erreichen mit 579 Punkten einen Mittelwert, der über eine Standardabweichung beziehungsweise knapp eineinhalb Kompetenzstufen höher liegt als der Mittelwert der Schülerinnen und Schüler an nicht gymnasialen Schularten, der 466 Punkte beträgt.

Die Schularten unterscheiden sich jedoch nicht nur in ihrer mittleren Kompetenz, sondern auch in der Streuung. Diese ist für die nicht gymnasialen Schularten ($SD = 90$ Punkte) höher als für die Gymnasien ($SD = 76$ Punkte), was allerdings vor dem Hintergrund unterschiedlicher Bildungsgänge plausibel erscheint.

Trotz der Unterschiede in der mittleren naturwissenschaftlichen Kompetenz zwischen den Schularten lassen sich Überlappungen der Kompetenzverteilungen beobachten. Knapp zehn Prozent der Schülerinnen und Schüler nicht gymnasialer Schularten erreichen ein Kompetenzniveau, das über dem Mittelwert der Gymnasien liegt. Umgekehrt liegt das Kompetenzniveau von knapp zehn Prozent der Jugendlichen an Gymnasien unterhalb des Mittelwertes der nicht gymnasialen Schularten.

Verteilung auf die Stufen der naturwissenschaftlichen Kompetenz

Die bisher präsentierten Ergebnisse spiegeln sich auch in den Häufigkeitsverteilungen der Schülerinnen und Schüler auf die Stufen der naturwissenschaftlichen Kompetenz für beide Schularten wider, die in Abbildung 9.11 dargestellt sind. Diese Häufigkeitsverteilungen sind, wie schon in früheren PISA-Erhebungen, gegeneinander verschoben. An den nicht gymnasialen Schularten befindet sich der Großteil der Jugendlichen auf den Kompetenzstufen II und III (56.8%). Mit 27.5 Prozent erreicht über ein Viertel der Schülerinnen und Schüler nicht Kompetenzstufe II, während der Anteil auf den Kompetenzstufen V und VI mit 2.6 Prozent gering ist. Im Vergleich dazu erreichen Jugendliche an Gymnasien am häufigsten Kompetenzstufe IV (38.0%). Fünfzehnjährige mit sehr niedrigen Kompetenzen, das heißt unter Kompetenzstufe II, sind hier praktisch nicht zu finden, wohingegen sich knapp ein Viertel der Schülerinnen und Schüler (23.7%) auf den höchsten Kompetenzstufen V und VI befindet und somit ausgezeichnete Zukunftsperspektiven beispielsweise für Berufswege im naturwissenschaftlich-technischen Bereich mitbringt.

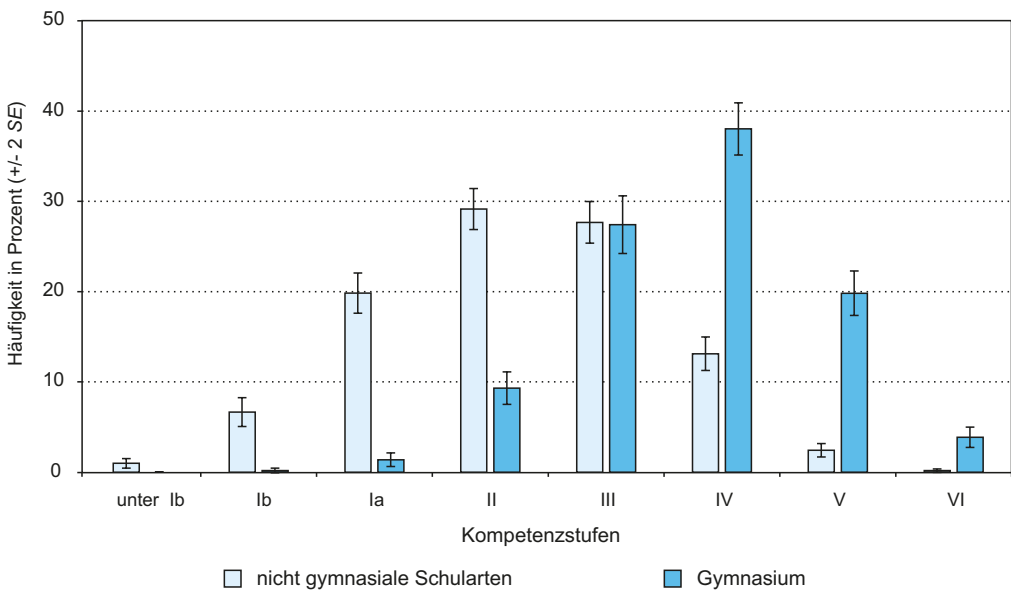


Abbildung 9.11: Prozentuale Anteile Fünfzehnjähriger auf den Stufen der naturwissenschaftlichen Kompetenz in Deutschland nach Schulart

9.3.2 Die Veränderung der naturwissenschaftlichen Kompetenz im Vergleich zu PISA 2006 und PISA 2015

Aussagen zu Trends sind bei einem langfristig angelegten Bildungsmonitoring wie PISA von besonderem Interesse, da sie Hinweise zu Wirkungen von Maßnahmen der Kompetenzförderung auf der Ebene eines Bildungssystems liefern können. Die Naturwissenschaften waren bei PISA 2006 erstmals Hauptdomäne. Zu diesem Zeitpunkt wurden eine elaborierte Rahmenkonzeption sowie ein ausführlicher Naturwissenschaftstest entwickelt, die noch immer einen wichtigen Bezugspunkt bilden. Bei PISA 2015, als die Naturwissenschaften zum zweiten Mal im Fokus der Erhebung standen, wurde dieses Rahmenkonzept ebenso wie die Testaufgaben und -prozeduren modifiziert. Diese Modifikationen waren nötig, um neuen Entwicklungen, wie zum Beispiel der Umstellung auf computerbasiertes Testen, Rechnung zu tragen (vgl. Schiepe-Tiska, Rönnebeck et al., 2016). Gleichzeitig schränkt dies jedoch die Vergleichbarkeit der Befunde und damit Aussagen über Veränderungen von PISA 2006 zu PISA 2018 ein. Sowohl eine Reanalyse der deutschen PISA-2015-Feldtestdaten als auch eine an PISA 2018 angekoppelte Begleitforschungsstudie kommen zu dem Schluss, dass einige der Aufgaben am Computer im Mittel etwas schwieriger zu lösen waren, als wenn sie auf Papier dargeboten wurden (Robitzsch et al., 2017; vgl. Kapitel 7). Die Veränderungen der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015 – und damit auch die Veränderungen zwischen PISA 2006 und PISA 2018 – sind daher mit entsprechenden Vorbehalten zu betrachten (vgl. Schiepe-Tiska, Rönnebeck et al., 2016). Dies gilt nicht für die Veränderungen von PISA 2015 zu PISA 2018, da PISA 2018 die gleiche Rahmenkonzeption, die gleichen Aufgaben (wenn auch durch den Status als Nebendomäne bedingt eine geringere Anzahl) und die gleichen Testprozeduren nutzt wie PISA 2015.

Tabelle 9.4 zeigt die Veränderung der naturwissenschaftlichen Kompetenz von PISA 2006 bis PISA 2018 für Deutschland. Im Vergleich zu PISA 2006 verringert sich diese signifikant um 13 Punkte von 516 auf 503 Punkte. Bei PISA 2015 war diese Entwicklung bereits zu beobachten, sie ließ sich jedoch nicht als statistisch bedeutsam absichern und es blieb offen, inwieweit dieses Ergebnis auf die beschriebenen Veränderungen zurückzuführen war (vgl. Schiepe-Tiska, Rönnebeck et al., 2016).

Tabelle 9.4: Naturwissenschaftliche Kompetenz in Deutschland bei PISA 2018, 2015 und 2006 für die Gesamtstichprobe und getrennt nach Geschlecht und Schulart

	2018				2015				2006			
	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)	M	(SE)	SD	(SE)
Mädchen	504	(3.3)	98	(2.1)	504	(3.2)	97	(1.6)	513	(4.5)	97	(2.0)
Jungen	502	(3.2)	107	(1.8)	514	(2.8)	93	(1.6)	517	(5.6)	103	(2.5)
nicht gymnasiale Schularten	466	(3.1)	90	(1.9)	476	(3.0)	86	(1.5)	486	(4.3)	87	(2.3)
Gymnasium	579	(2.5)	76	(1.6)	585	(2.6)	75	(2.0)	598	(2.9)	68	(1.7)
Gesamtstichprobe	503	(2.9)	103	(1.6)	509	(2.7)	99	(1.5)	516	(3.8)	100	(2.0)

Auch in 17 weiteren OECD-Staaten findet sich eine signifikante Abnahme der naturwissenschaftlichen Kompetenz (siehe Abbildung 9.12). Dies betrifft als Teile der Spitzengruppe auch Kanada (-16 Punkte) und vor allem Finnland (-41 Punkte). In der Gruppe der im Mittelwert von Deutschland nicht signifikant verschiedenen Staaten (siehe Abschnitt 9.2.1) sind die größten Verluste in der Schweiz², Tschechien (je -16 Punkte), den Niederlanden (-21 Punkte), Neuseeland (-22 Punkte) und Australien (-24 Punkte) zu beobachten. Vergleichbare Verluste wie in Deutschland finden sich in den europäischen Staaten Belgien, Irland und Slowenien (je -12 Punkte). Signifikante Zuwächse zwischen PISA 2006 und PISA 2018 verzeichnen lediglich fünf Staaten: Polen, Portugal, Kolumbien und die Vereinigten Staaten mit Gewinnen zwischen 13 und 25 Punkten sowie die Türkei mit einem Kompetenzzuwachs von 44 Punkten. Insgesamt bleibt der OECD-Mittelwert der aktuellen OECD-Staaten jedoch relativ stabil.

Betrachtet man die als verlässlicher einzuschätzenden Veränderungen zwischen PISA 2015 und 2018, zeigt sich ein etwas anderes Bild. Die naturwissenschaftliche Kompetenz in Deutschland zeigt keine signifikante Veränderung. Ebenso wie in Deutschland lassen sich in 22 weiteren OECD-Staaten keine statistisch abgesicherten Veränderungen in der naturwissenschaftlichen Kompetenz beobachten. Darunter befinden sich auch die meisten Staaten, die sich nicht signifikant vom Mittelwert 2018 in Deutschland unterscheiden. Auch der Mittelwert der aktuellen OECD-Staaten verändert sich nicht signifikant. Signifikant niedrigere Mittelwerte sind in zwölf OECD-Staaten zu beobachten. Dazu gehören erneut die Spitzenstaaten Kanada (-10 Punkte) und Finnland (-9 Punkte) sowie Japan (-9 Punkte). Unter den von Deutschland im Mittelwert nicht signifikant verschiedenen Staaten sind statistisch bedeutsame Abnahmen in der Schweiz (-10 Punkte), Australien (-7 Punkte) und Slowenien (-6 Punkte) zu verzeichnen. Eine signifikante Steigerung der naturwissenschaftlichen Kompetenz gelingt nur in zwei Staaten, nämlich in Polen (+10 Punkte) und insbesondere in der Türkei mit einem Zuwachs von 43 Punkten.

Eine Betrachtung der Veränderung in den Verteilungen auf die Stufen der naturwissenschaftlichen Kompetenz kann Hinweise auf potenzielle Förderbedarfe, aber gegebenenfalls auch auf Effekte erfolgter Fördermaßnahmen geben. Im Vergleich zu PISA 2006 nimmt in Deutschland der Anteil der Jugendlichen, deren naturwissenschaftliche Kompetenz unterhalb der Kompetenzstufe II liegt, statistisch bedeutsam um 4.2 Prozentpunkte auf fast 20 Prozent zu. Damit ist knapp ein Fünftel der Schülerinnen und Schüler in Deutschland nicht in der Lage, im Alltag einfache naturwissenschaftliche Probleme und Zusammenhänge zu verstehen oder aus Daten und Informationen Schlussfolgerungen zu ziehen. Der Anteil kompetenzstarker Jugendlicher, welche die Kompetenzstufen V und VI erreichen, geht leicht von 11.8 auf 10.0 Prozent zurück. Dieser Rückgang ist jedoch nicht statistisch signifikant. Auch im Vergleich zu PISA 2015 steigt der Anteil der Jugendlichen in Deutschland, deren naturwissenschaftliche Kompetenz unterhalb der

2 Die angegebenen Differenzen wurden mit Dezimalstellen berechnet. Daher können beim Vergleich mit den Tabellen vermeintliche Inkonsistenzen aufgrund von Rundungen vorkommen.

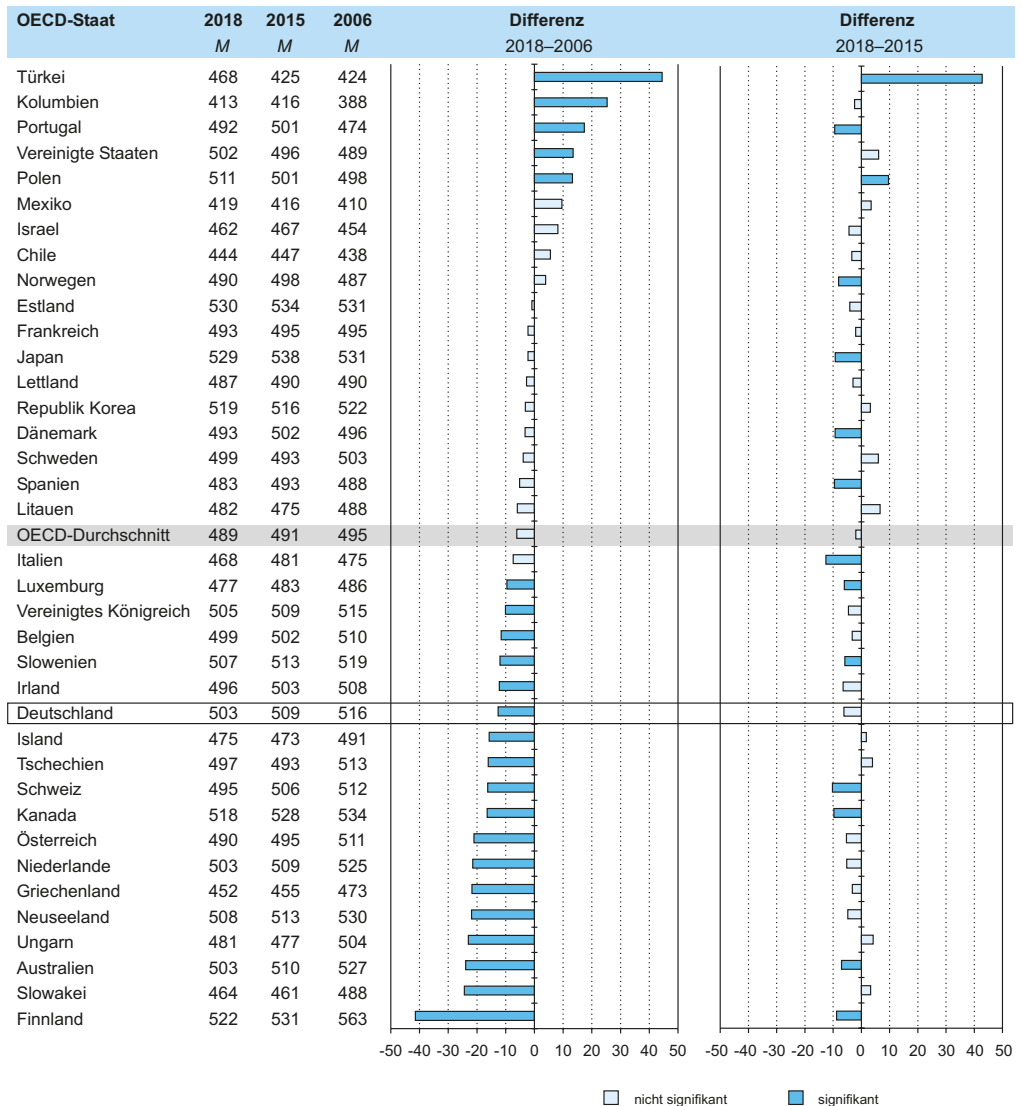


Abbildung 9.12: Mittelwerte naturwissenschaftlicher Kompetenz in den OECD-Staaten in PISA 2018, PISA 2015 und PISA 2006

Kompetenzstufe II liegt, von 17.0 auf 19.6 Prozent an, lässt sich aber nicht als statistisch bedeutsam absichern. Der Anteil der Fünfzehnjährigen auf den Kompetenzstufen V und VI bleibt mit 10.0 Prozent vergleichbar zu PISA 2015. Damit scheint die Spitzenförderung in Deutschland über die Erhebungsrounden hinweg vergleichsweise gut zu gelingen, während sich in der Breitenförderung klarer Handlungsbedarf abzeichnet.

Neben den Veränderungen auf der Gesamtskala der naturwissenschaftlichen Kompetenz lassen sich für Deutschland auch Veränderungen der mittleren Kompetenz der Gymnasien und nicht gymnasialen Schularten untersuchen. Die Entwicklung der Kom-

petenzwerte in den verschiedenen Schularten zeigt Tabelle 9.4. Im Vergleich zu PISA 2006 ist der Mittelwert der naturwissenschaftlichen Kompetenz sowohl an den Gymnasien (-19 Punkte) als auch an den nicht gymnasialen Schularten (-20 Punkte) signifikant gesunken. Dabei steigt an den nicht gymnasialen Schularten der Anteil an Schülerinnen und Schülern, welche die Kompetenzstufe II nicht erreichen (+8.3 Prozentpunkte). Gleichzeitig sinkt an den Gymnasien der Anteil kompetenzstarker Jugendlicher (-6.7 Prozentpunkte). Zudem ist die Streuung an den Gymnasien zwischen 2006 und 2018 signifikant gestiegen, was darauf hindeutet, dass es manchen Gymnasien besser als anderen gelingt, die naturwissenschaftliche Kompetenz ihrer Schülerinnen und Schüler zu fördern.

Im Vergleich zwischen den PISA-Studien 2018 und 2015, denen die gleiche Rahmen- und Testkonzeption zugrunde liegt, verringert sich an den Gymnasien die naturwissenschaftliche Kompetenz um sechs Punkte, wobei diese Veränderung nicht statistisch signifikant ist. Auch die Anteile auf den unteren und oberen Kompetenzstufen sind vergleichbar zwischen PISA 2018 und PISA 2015. Hingegen ist die Abnahme der naturwissenschaftlichen Kompetenz an nicht gymnasialen Schularten um zehn Punkte statistisch bedeutsam und manifestiert sich auch durch eine signifikante Zunahme der Jugendlichen, welche die Kompetenzstufe II nicht erreichen (+4.8 Prozentpunkte).

Betrachtet man die Unterschiede zwischen Mädchen und Jungen, so nimmt in Deutschland zwischen PISA 2006 und PISA 2018 der Mittelwert für die Jungen signifikant um 17 Punkte ab, während die Veränderung bei den Mädchen statistisch nicht bedeutsam ist. Der gleiche Effekt zeigt sich zwischen PISA 2015 und PISA 2018. Während sich die mittlere Kompetenz der Jungen statistisch signifikant um zwölf Punkte verringert, bleibt der Mittelwert der Mädchen unverändert bei 504 Punkten. Damit ist die 2015 erstmals aufgetretene Geschlechterdifferenz zugunsten der Jungen wieder verschwunden. Zudem hat seit 2015 die Streuung für die Jungen signifikant um 14 Punkte zugenommen (auf $SD = 107$). Dies deutet auf eine Zunahme der Leistungsheterogenität innerhalb der Gruppe der Jungen hin, was ein Blick auf die Verteilung von Mädchen und Jungen auf die Stufen naturwissenschaftlicher Kompetenz verdeutlicht. Im unteren Kompetenzbereich (das heißt auf Kompetenzstufe Ia und darunter) bleibt der Anteil der Mädchen bei PISA 2018 mit 18.2 Prozent praktisch unverändert im Vergleich zur Erhebung vor drei Jahren und ist auch nicht bedeutsam verschieden im Vergleich zu PISA 2006. Bei den Jungen hingegen steigt der Anteil signifikant von 14.9 Prozent bei PISA 2006 beziehungsweise 15.9 Prozent bei PISA 2015 auf 20.8 Prozent bei PISA 2018 an. Damit befinden sich bei PISA 2018 nun signifikant mehr Jungen auf den unteren Kompetenzstufen als Mädchen (siehe Abschnitt 9.2.3). Im oberen Kompetenzbereich, das heißt auf den Kompetenzstufen V und VI, bleibt in Deutschland auch bei PISA 2018 ein statistisch signifikant höherer Anteil an Jungen bestehen. Allerdings geht dieser Anteil bei PISA 2018 leicht auf 11.1 Prozent (gegenüber 12.4% bei PISA 2015) zurück, während der Anteil der Mädchen mit 8.7 Prozent gegenüber PISA 2015 stabil bleibt.

9.4 Zusammenfassung und Diskussion – Stärken und Herausforderungen in Deutschland

Schülerinnen und Schüler in Deutschland zeigen auch bei PISA 2018 eine über dem OECD-Durchschnitt liegende naturwissenschaftliche Kompetenz. Diese ist nach wie vor international anschlussfähig, auch wenn es noch immer nicht gelungen ist, zu den Spitzengruppen Estland und Japan sowie Finnland, der Republik Korea und Kanada aufzuschließen. Im Vergleich zu PISA 2015 als belastbaren Referenzpunkt bleibt die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler in Deutschland stabil. Dass dies nicht selbstverständlich ist, zeigt zum Beispiel der Rückgang der naturwissenschaftlichen Kompetenz in den Spitzenstaaten Japan, Finnland und Kanada.

Die Förderung leistungsstarker Jugendlicher scheint im internationalen Vergleich noch vergleichsweise gut zu gelingen, während sich in der Breitenförderung inzwischen Schwächen andeuten: Ein Fünftel der Schülerinnen und Schüler erreicht Kompetenzstufe II und damit die Mindestanforderungen an eine naturwissenschaftliche Grundbildung, nicht. Diesen Jugendlichen fehlen grundlegende naturwissenschaftliche Kompetenzen, um im Alltag einfache naturwissenschaftliche Zusammenhänge oder Probleme zu verstehen. Sie sind nicht in der Lage, aus verfügbaren Daten und Informationen angemessene Schlussfolgerungen zu ziehen. Gerade mit Blick auf aktuelle Herausforderungen wie Globalisierung, Digitalisierung und Klimawandel fehlt ihnen ein naturwissenschaftliches Verständnis, um sich informiert und reflektiert an gesellschaftlichen Debatten und Entscheidungen zu beteiligen. Vor allem an nicht gymnasialen Schularten hat der Anteil der Fünfzehnjährigen, welche die Kompetenzstufe II nicht erreichen, im Vergleich zu PISA 2015 um knapp fünf Prozentpunkte zugenommen. Demnach hat sich auch insgesamt gesehen die naturwissenschaftliche Kompetenz an nicht gymnasialen Schularten verringert. Betrachtet man dies in Zusammenhang mit den Ergebnissen der PISA-2012-Längsschnittstudie, die zeigen konnte, dass an nicht gymnasialen Schularten die in PISA erfasste naturwissenschaftliche Kompetenz von der neunten zur zehnten Klasse noch einmal abnimmt (Schiepe-Tiska et al., 2017), gibt diese Entwicklung Anlass zur Sorge. Auch an den Gymnasien zeichnet sich ein solcher Trend ab. Diese Befunde überraschen wenig, wenn man berücksichtigt, dass sich der naturwissenschaftliche Unterricht zwischen PISA 2006 und PISA 2015 kaum verändert hat (Schiepe-Tiska, Schmidner et al., 2016). Groß angelegte Unterrichtsprogramme zur Weiterentwicklung des mathematisch-naturwissenschaftlichen Unterrichts wie SINUS oder die Kontextprojekte fehlen, beziehungsweise wurden nicht systematisch weiterentwickelt. Dazu passt auch, dass sich das Erreichen der Bildungsstandards in den Fächern Biologie, Physik und Chemie seit 2012 insgesamt in Deutschland kaum verändert hat (Stanat, Schipolowski, Mahler, Weirich & Henschel, 2019). Hier scheint eine deutliche Intensivierung der Bemühungen um eine breite Förderung angezeigt zu sein.

Mädchen und Jungen unterscheiden sich in Deutschland, ähnlich wie im OECD-Durchschnitt und der Mehrheit der OECD-Staaten, nicht in ihrer naturwissen-

schaftlichen Kompetenz. Die erstmals bei PISA 2015 aufgetretene Geschlechterdifferenz lässt sich somit nicht replizieren (Schiepe-Tiska, Rönnebeck et al., 2016). Während die Kompetenz der Mädchen weitgehend stabil bleibt, ist bei den Jungen eine bedeutsame Abnahme zu beobachten. Auch hat sich die Streuung der Kompetenz bei den Jungen im Vergleich zu PISA 2015 vergrößert. Vor allem auf den unteren Kompetenzstufen hat ihr Anteil deutlich zugenommen. Offen bleibt, worauf sich dieser Rückgang bei den Jungen zurückführen lässt. Mögliche Ursachen könnten die bei PISA 2015 berichtete Abnahme von Interesse an den Naturwissenschaften sowie die weniger positiv ausgeprägten motivationalen Selbstbilder sein (Schiepe-Tiska, Simm & Schmidtner, 2016), die auch der aktuelle Bildungstrend vor allem für die Jungen bestätigt (Stanat et al., 2019). Maßnahmen zur Förderung der naturwissenschaftlichen Kompetenz sollten deshalb im Sinne einer mehrdimensionalen Bildung auch motivationale Orientierungen und Einstellungen als wichtige Bildungsziele adressieren (vgl. Schiepe-Tiska, im Druck). Diese sind nicht nur für die spätere Berufswahl und die Sicherung des MINT-Nachwuchses von Bedeutung, sondern beeinflussen auch die Bereitschaft, sich mit Naturwissenschaften auseinanderzusetzen, sich auf neue Sachverhalte einzulassen sowie lebenslang zu lernen. Damit bilden sie eine entscheidende Grundlage für eine aktive gesellschaftliche Teilhabe.

Literatur

- Bayrhuber, H., Bögeholz, S., Elster, D., Hammann, M., Hößle, C., Lücken, M., ... Sandmann, A. (2007). Biologie im Kontext – Ein Programm zur Kompetenzförderung durch Kontextorientierung im Biologieunterricht und zur Unterstützung von Lehrerprofessionalisierung. *MNU*, 60, 282–286.
- Bybee, R. & McCrae, B. (2011). Scientific literacy and student attitudes: Perspectives from PISA 2006 science. *International Journal of Science Education*, 33, 7–26. <https://doi.org/10.1080/09500693.2010.518644>
- DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching*, 37, 582–601. [https://doi.org/10.1002/1098-2736\(200008\)37:6<582::AID-TEA5>3.0.CO;2-L](https://doi.org/10.1002/1098-2736(200008)37:6<582::AID-TEA5>3.0.CO;2-L)
- Demuth, R., Gräsel, C., Parchmann, I. & Ralle, B. (Hrsg.). (2008). *Chemie im Kontext: Von der Innovation zur nachhaltigen Verbreitung eines Unterrichtskonzepts*. Münster: Waxmann.
- Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, 32, 268–291. <https://doi.org/10.3102/0091732X07309371>
- KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Biologie für den mittleren Bildungsabschluss (Beschluss vom 16. Dezember 2004)*. München: Wolters Kluwer.
- KMK (2005b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Beschlüsse der Kultusministerkonferenz – Bildungs-*

- standards im Fach Chemie für den mittleren Bildungsabschluss (Beschluss vom 16. Dezember 2004)*. München: Wolters Kluwer.
- KMK (2005c) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den mittleren Bildungsabschluss (Beschluss vom 16. Dezember 2004)*. München: Wolters Kluwer.
- Mang, J., Wagner, S., Gomolka, J., Schäfer, A., Meinck, S. & Reiss, K. (2019). *Technische Hintergrundinformationen PISA 2018*. München: Technische Universität München. <https://doi.org/10.14459/2019md1518258>
- Mikelskis-Seifert, S. & Duit, R. (2010). *Physik im Kontext – Konzepte, Ideen, Materialien für effizienten Physikunterricht*. Seelze: Friedrich Verlag.
- NRC (2012) = National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C.: The National Academies Press.
- OECD. (2019a). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing.
- OECD. (2019b). *PISA 2018 results (Volume I): What students know and can do*. Paris: OECD Publishing.
- Petersen, S., Blankenburg, J. & Höffler, T. (2017). Challenging gifted students in science: The German Science Olympiads. In K. S. Taber, M. Sumida & L. McClure (Hrsg.), *Teaching gifted learners in STEM subjects: Developing talent in science, technology, engineering and mathematics* (S. 157–170). Abingdon: Routledge. <https://doi.org/10.4324/9781315697147-11>
- Prenzel, M., Friedrich, A. & Stadler, M. A. (2009). *Von SINUS lernen: Wie Unterrichtsentwicklung gelingt*. Seelze-Velber: Klett/Kallmeyer.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P. & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, ... M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 192–250). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_6
- Prenzel, M., Schöps, K., Rönnebeck, S., Senkbeil, M., Walter, O., Carstensen, C. & Hammann, M. (2007). Naturwissenschaftliche Kompetenz im internationalen Vergleich. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (S. 61–124). Münster: Waxmann.
- Robitzsch, A., Lüdtke, O., Köller, O., Kröhne, U., Goldhammer, F. & Heine, J.-H. (2017). Herausforderungen bei der Schätzung von Trends in Schulleistungsstudien. *Diagnostica*, 63, 148–165. <https://doi.org/10.1026/0012-1924/a000177>
- Rönnebeck, S., Schöps, K., Prenzel, M., Mildner, D. & Hochweber, J. (2010). Naturwissenschaftliche Kompetenz von PISA 2006 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, ... P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 177–198). Münster: Waxmann.
- Rost, J., Walter, O., Carstensen, C., Senkbeil, M. & Prenzel, M. (2004). Naturwissenschaftliche Kompetenz. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, ... U. Schiefele (Hrsg.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland: Ergebnisse des zweiten internationalen Vergleichs* (S. 111–146). Münster: Waxmann.
- Schiepe-Tiska, A. (im Druck). *Konzeptionierung und Erfassung mehrdimensionaler Bildungsziele sowie deren schulische und unterrichtliche Rahmenbedingungen in internationalen*

- Large-Scale Assessments am Beispiel der MINT-Fächer*. Habilitationsschrift. München: Technische Universität München.
- Schiepe-Tiska, A., Roczen, N., Müller, K., Prenzel, M. & Osborne, J. (2016). Science-related outcomes: Attitudes, motivation, value beliefs, strategies. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Hrsg.), *Assessing contexts of learning: An international perspective* (S. 301–329). Berlin: Springer. https://doi.org/10.1007/978-3-319-45357-6_12
- Schiepe-Tiska, A., Rönnebeck, S., Heitmann, P., Schöps, K., Prenzel, M. & Nagy, G. (2017). Die Veränderung der naturwissenschaftlichen Kompetenz von der 9. zur 10. Klasse bei PISA und den Bildungsstandards unter Berücksichtigung geschlechts- und schulartspezifischer Unterschiede sowie der Zusammensetzung der Schülerschaft. *Zeitschrift für Erziehungswissenschaft*, 20(S2), 151–176. <https://doi.org/10.1007/s11618-017-0754-2>
- Schiepe-Tiska, A., Schmidtner, S., Müller, K., Heine, J.-H., Neumann, K. & Lüdtke, O. (2016). Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 133–176). Münster: Waxmann.
- Schiepe-Tiska, A., Schöps, K., Rönnebeck, S., Köller, O. & Prenzel, M. (2013). Naturwissenschaftliche Kompetenz in PISA 2012: Ergebnisse und Herausforderungen. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 189–216). Münster: Waxmann.
- Schiepe-Tiska, A., Simm, I. & Schmidtner, S. (2016). Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 99–132). Münster: Waxmann.
- Shavelson, R. J., Ruiz-Primo, M. A. & Wiley, E. W. (2005). Windows into the mind. *Higher Education*, 49, 413–430. <https://doi.org/10.1007/s10734-004-9448-9>
- Stanat, P., Schipolowski, S., Mahler, N., Weirich, S. & Henschel, S. (2019). Zusammenfassung und Einordnung der Befunde. In P. Stanat, S. Schipolowski, N. Mahler, S. Weirich & S. Henschel. (Hrsg.), *IQB-Bildungstrend 2018. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I im zweiten Ländervergleich* (S. 427–450). Münster: Waxmann.

10 PISA 2018 – die Methodologie

Jörg-Henrik Heine & Kristina Reiss

Das vorliegende Kapitel widmet sich der mit den PISA-Studien verbundenen Methodologie. Nach einer kurzen Einführung zu grundlegenden Aspekten der Messung geht es im Kern des Kapitels um die Veränderungen in den eingesetzten Methoden und Verfahrensweisen seit der PISA-Erhebung im Jahre 2012. Die Veränderungen betreffen die Testadministration und das Testdesign. Es werden zentrale Aspekte der technischen Durchführung der PISA-Studie in Deutschland skizziert. Ein abschließender Abschnitt umfasst praktische Hinweise zur Verfügbarkeit der PISA-Daten und zu „Werkzeugen“ für replizierende und weiterführende Sekundäranalysen. In einem elektronischen Anhang zu diesem Kapitel (vgl. www.waxmann.com/buch4100) werden zentrale Begriffe und Methoden, welche für das Verständnis der PISA zugrundeliegenden Methodik zentral sind, erläutert.

10.1 Einführung – Messung von latenten Merkmalen und Kompetenzen

In den Sozialwissenschaften allgemein und bei internationalen Vergleichsstudien wie PISA speziell, stehen theoretische Konzepte in der Regel im Zentrum der jeweiligen Fragestellungen. Diese als *Konstrukte* bezeichneten theoretischen Konzepte beziehen sich beispielsweise auf unterschiedliche Merkmale von Bildungssystemen, Schulsystemen oder Schularten oder auch einzelne Gruppen von Lehrenden, Lernenden und deren Eltern. Eine übergeordnete Gemeinsamkeit solcher auch als *Dimensionen* oder *Domänen* bezeichneten Merkmale ist der Umstand, dass sich diese typischerweise der unmittelbar sinnlichen Erfahrung entziehen. So weisen die *Dimensionen* von *Konstrukten* wie „Intelligenz“, „Kompetenz“ und „Motivation“ im Gegensatz zu physikalischen Eigenschaften (wie Größe und Gewicht) einen lediglich *indirekten empirischen Bezug* auf (Browne, 2000). Diese werden daher im methodologischen Kontext der Sozialwissenschaften als *latente (verborgene) Variablen* bezeichnet, welche erst über eine geeignete *Operationalisierung* empirisch zugänglich und damit messbar gemacht werden (Narens, 1981; Steyer & Eid, 2000). Das in PISA als Schwerpunkt thematisierte Konstrukt der *Kompetenz* (vgl. BMBF, Jude, Hartig & Klieme, 2008) mit einzelnen Dimensionen wie *Lesekompetenz*,

(vgl. Kapitel 2 und Kapitel 3), *Mathematische Kompetenz* (vgl. Kapitel 8), und *Naturwissenschaftliche Kompetenz* (vgl. Kapitel 9) wird so über einzelne Testaufgaben operationalisiert. Die Basis für eine angemessene Operationalisierung ist in den *Rahmenkonzeptionen* gegeben (vgl. OECD, 2019), welche über grundlegende *Begriffsdefinitionen* die Inhalte der zu entwickelnden Testaufgaben (vgl. ETS, 2018a) vorgeben. Neben solchen inhaltlichen und definitorischen Fragestellungen müssen im Rahmen der Operationalisierung vor allem auch methodologische Aspekte wie die Form der Testgestaltung (siehe Abschnitt 10.2.1 und Abschnitt 10.2.2) und die anschließende Verbindung der Lösungen von einzelnen Testaufgaben zu einem Messwert im Rahmen der *psychometrischen Skalierung* (Browne, 2000, Torgerson, 1961) angemessen umgesetzt werden.

Die methodologischen Möglichkeiten der Skalierung und Messung von Kompetenzen sind nach Klieme und Hartig (2007) ein zentraler Aspekt des Kompetenzkonzepts im Bereich der Bildungswissenschaft. Die Relevanz des Konzepts ist vor allem dadurch gegeben, dass Kompetenzen als Ergebnisse von Bildungsprozessen verstanden werden (Hartig, 2008).

Internationale Schulleistungsstudien wie PISA fokussieren auf den Bereich von funktionalen und fachbezogenen Kompetenzen. In Abgrenzung zu curricular eng definierten Begriffen wie „Schülerleistung“, „Bildungsergebnis“ oder auch „Intelligenz“ bezieht sich der Kompetenzbegriff im Rahmen der PISA-Studie auf das breitere Potenzial, Gelerntes zum Zweck des Problemlösens in unterschiedlichen Situationen anwenden zu können (Prenzel, Gogolin & Krüger, 2007). Daneben werden im Rahmen der PISA-Studien auch weitere Merkmale in Fragebögen für Lehrende, Lernende und Eltern erfasst. Diese beziehen sich auf theoretische Konzepte, welche potenziell mit Dimensionen der Kompetenz assoziiert sind wie etwa Motivation oder allgemeine Einstellungen (vgl. z. B. Kapitel 4).

10.2 Methodische Veränderungen seit PISA 2012

Die methodologischen Aspekte weisen bei PISA über die einzelnen Erhebungsrunden hinweg einerseits kontinuierliche und andererseits innovative Elemente auf. Kontinuierlich ist beispielsweise die Erhebung der drei „klassischen“ Domänen Lesen, Mathematik und Naturwissenschaften, welche mit wechselnder Schwerpunktsetzung von Beginn an erhoben werden. Neben dieser Kontinuität gibt es aber auch fortwährend Veränderungen. So werden die drei Domänen von Runde zu Runde jeweils durch optionale innovative Kompetenzdomänen ergänzt. Weitere Veränderungen beziehen sich auf technische Aspekte wie den *Testmodus*, der in den meisten Staaten von einer auf Papier basierenden zu einer computerbasierten Umgebung wechselte (Heine et al., 2016, S. 407 ff.; vgl. dazu auch Kapitel 7), oder die Form der Darbietung der Testaufgaben und Fragebögen, also auf das *Testdesign*. Daneben bestehen inhaltlich begründete Veränderungen, die am deutlichsten an der Weiterentwicklung der Rahmenkonzeptionen für den Kompetenzbereich *Lesen* (vgl. OECD, 2019) und der entsprechenden *Operationalisierung*

der PISA-Domänen, sichtbar werden. So wurden für die *Lesekompetenz* in PISA 2018 neben 72 bereits 2015 bestehenden Trenaufgaben (die im Wesentlichen auch bereits in früheren Runden verwendeten wurden) auch 173¹ neu entwickelte Aufgaben eingesetzt. Diese *Operationalisierung* mit neuen PISA-Aufgabenformaten bezieht sich auf neue Formen des Lesens, welche eine „*Analyse, Synthese, Integration und Interpretation relevanter Informationen aus mehreren Text- oder Informationsquellen*“ (OECD, 2019, S. 23) erfordern. Eingesetzt werden beispielsweise „simulierte“ Beiträge in einem Online-Forum (vgl. ETS, 2018a, S. 5). Dabei müssen relevante Texte gefunden (19 Aufgaben – 11 %), Informationen aus mehreren Textquellen kombiniert werden, um Schlussfolgerungen zu ziehen (15 Aufgaben – 9 %) oder Widersprüche entdeckt und bewältigt werden (vgl. OECD, 2019, S. 33; ETS, 2018a, S. 3 f.; Kapitel 2). Insgesamt besteht bei PISA ein Veränderungstrend dahingehend, dass immer mehr neue Domänen zusätzlich angeboten und erhoben werden und als langfristige Perspektive das Prinzip der spezifischen Schwerpunktsetzung in einer PISA-Runde aufgehoben werden soll (ETS, 2018b; ETS, 2016a).

10.2.1 Papier- und computerbasierte Testung

Die Messung *latenter Variablen* ist in den Sozialwissenschaften mit dem Einsatz psychometrischer Tests und Fragebögen assoziiert (z.B. Mummendey & Grau, 2014). Traditionellerweise werden Fragebögen und auch Kompetenz- und Leistungstests dabei in Papierform an die Teilnehmenden der Untersuchung ausgegeben. Allerdings zeigt sich zunehmend, dass sich durch den Einsatz computerbasierter Testszenarien ein systematischer Informations- und Effizienzgewinn bei der Datenerhebung erzielen lässt. Beginnend mit der internationalen Erhebung 2015 wurde die computerbasierte Datenerhebung bei PISA als Standard (mit einer optionalen Beibehaltung der papierbasierten Testung) eingeführt. Während bei der Erhebung im Jahre 2015 von 73 teilnehmenden Staaten noch 14 Staaten die papierbasierte Testung beibehalten haben (OECD, 2018), hat sich die Anzahl der papierbasiert testenden Staaten für die PISA-Runde 2018 auf neun von insgesamt 79 teilnehmenden Staaten reduziert. Allgemein lassen sich für eine computerbasierte Testung vier grundlegende Aspekte anführen. So kann aus messtheoretischer Perspektive angeführt werden, dass die Genauigkeit der Messung durch eine computerbasierte Erhebung verbessert wird (Parshall, Spray, Kalohn & Davey, 2002; van der Linden & Glas, 2000, 2010; Wainer & Dorans, 2000). In der aktuellen PISA-Studie 2018 geschah dies konkret durch die Einführung eines mehrstufig adaptiven Testdesigns. Außerdem können bei einer computerbasierten Testung zusätzliche „Meta- oder Prozessdaten“ erhoben werden (beispielsweise Bearbeitungszeiten für Teilaufgaben), welche

1 Den Schülerinnen und Schülern wurden insgesamt 173 neue entwickelte Aufgaben im Rahmen der computerbasierten Erfassung der Lesekompetenz dargeboten. Allerdings konnte eine der Aufgaben (CR563Q12) nicht für den internationalen Vergleich aller Staaten verwendet werden. Aus der psychometrischen Perspektive der Skalierung beträgt die Anzahl der neuen Items (Aufgaben) daher nur 172.

zu einer erweiterten Diagnostik der Antworten der Schülerinnen und Schüler eingesetzt werden können (z. B. Bradshaw & Templin, 2014; Goldhammer, Naumann, Rölke, Stelter & Tóth, 2017). Aus inhaltlicher Perspektive kann angeführt werden, dass die Messung von Kompetenzen nach zeitgemäßen Definitionen und Rahmenkonzeptionen durch rein papierbasierte Tests nicht (mehr) angemessen erfasst werden können (Bennett, 2002; Parshall, Harmes, Davey & Pashley, 2010). So werden die Rahmenkonzeptionen zu den PISA-Domänen zunehmend durch interaktive Elemente erweitert, welche sich sinnvoll nur über eine computerbasierte Testung realisieren lassen, wie zum Beispiel Simulationsszenarien im Bereich Naturwissenschaften oder innovative Frageformate. Schließlich kann die Datenerhebung für die Testadministratoren und für die Schülerinnen und Schüler strukturierter, zeitsparender und damit effizienter gestaltet werden (z. B. Frey & Ehmke, 2008; Frey & Seitz, 2009). Es ist offensichtlich, dass bei einem weltumspannenden Projekt mit mehr als 600 000 teilnehmenden Schülerinnen und Schülern an etwa 22 000 Schulen, die insgesamt 126 unterschiedliche Sprachversionen der Testinstrumente bearbeiten, die Effizienz der computerbasierten Datenerhebung und Verarbeitung einen zentralen Stellenwert hat.

10.2.2 Mehrstufig adaptives Testdesign

Im Rahmen der *Operationalisierung* werden die Testaufgaben in einer geeigneten Art und Weise – dem *Testdesign* – zu einem Test oder einer Fragebogenskala zusammengestellt. Mit dem Begriff *Testdesign* wird die systematische und psychometrisch begründete Anordnung der Testaufgaben in einem Test bezeichnet. Auf einer obersten Ebene unterschiedlicher Testdesigns kann zunächst zwischen einem *universellen Testdesign* und einem *rotierten Booklet-Testdesign* unterschieden werden. Während bei einem universellen Testdesign die Anzahl und auch die Abfolge der Testaufgaben oder Fragen für alle Personen gleich ist (z. B. Klassenarbeiten), werden die zur Verfügung stehenden Aufgaben oder Fragen bei einem rotierten Booklet-Testdesign systematisch über eine unterschiedliche Anzahl von Testheften verteilt (vgl. z. B. Frey, Hartig & Rupp, 2009; Gonzalez & Rutkowski, 2010; Kubinger et al., 2011, für die technisch, praktischen Aspekte solcher Booklet-Designs). Bei rotierten Booklet-Testdesigns bearbeitet also eine einzelne Schülerin beziehungsweise ein einzelner Schüler jeweils nur einen Teil der existierenden Testaufgaben. Bei PISA wird seit 2000 jeweils ein rotiertes Booklet-Design zur Erfassung der einzelnen Kompetenzdomänen eingesetzt. Dies ergibt sich aus dem Kerngedanken der Studie, Aussagen auf Populationsebene zu machen. Dabei soll zur Messung eine große Anzahl von Aufgaben eingesetzt werden und gleichzeitig die Testdauer für die einzelnen Schülerinnen und Schüler möglichst kurz ausfallen (Heine, Sälzer, Borchert, Siberns & Mang, 2013; Heine et al., 2016). Ausgehend von diesem Prinzip des Testdesigns wurde in PISA 2018 für die Hauptdomäne Lesen ein computerbasiertes *mehrstufig adaptives Testdesign* (*Multistage Adaptive Testdesign* – MAT) eingeführt. Solch ein Testdesign verwendet während der Messung adaptive Algorithmen, um die Testinhalte systematisch

und dynamisch auszuwählen. Der Begriff „adaptiv“ beschreibt dabei, dass die Auswahl und Reihenfolge der präsentierten Testaufgaben individuell an die Kompetenz der Schülerinnen und Schüler angepasst werden. Vollständig adaptive Testdesigns nutzen dabei Modelle aus der *Item Response Theory* (IRT) und stützen die individuelle Zuweisung von Aufgaben zu Testpersonen auf a priori bekannte Koeffizienten der Aufgabenschwierigkeiten und vorläufige, im Verlaufe der Testung immer präziser ausfallende Schätzer der Personenfähigkeit (Kompetenz). Typischerweise wird die Anwendung adaptiver Testdesigns damit begründet, einen präziseren Messwert des zu erfassenden Konstrukts für eine Testperson in kürzerer Zeit zu erzielen (z. B. Diedenhofen & Musch, 2018; Kubinger, 2017; Weiss, 2011; Yamamoto, Khorramdel & Shin, 2018). Die adaptive Auswahl von Testaufgaben maximiert so die gewonnene Testinformation. Ein Spezialfall des adaptiven Testens stellt die *Testlet*-basierte, adaptive Zusammenstellung der einzelnen Testaufgaben dar (Frey, Seitz & Brandt, 2016). Als *Testlets* werden Gruppen von Aufgaben bezeichnet, die nach theoretischen und inhaltlichen Aspekten zusammengestellt wurden. Demgegenüber werden mit dem Begriff *Unit* (Aufgabeneinheit), der im Kontext von PISA verwendet wird, Testaufgaben bezeichnet, welche denselben *Stimulus* (Aufgabenstamm) teilen, also einen gemeinsamen Kontext, der beispielsweise durch eine Grafik, ein Bild oder Texte vermittelt wird. Insofern können die *Testlets* bei PISA nur eine einzige, oder aber auch mehrere Units umfassen. Diese *Testlets* werden bei PISA für die Hauptdomäne Lesekompetenz in adaptiver Form einzelnen Schülerinnen und Schülern zugewiesen. Das bei PISA 2018 für die Kompetenzdomäne Lesen eingesetzte mehrstufige adaptive Testdesign besteht auf einer übergeordneten Ebene aus drei Stufen (siehe Abbildung 10.1).

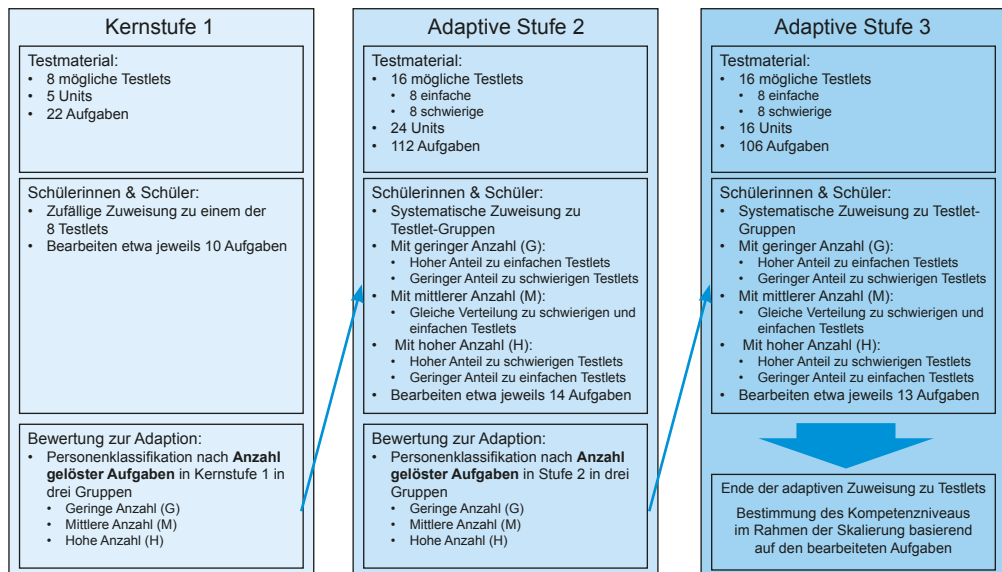


Abbildung 10.1: Schematische Darstellung des mehrstufigen adaptiven Testdesigns (MAT) für den Kompetenzbereich Lesen

Innerhalb dieser Stufen sind die einzelnen Aufgaben in *Testlets* zusammengefasst. Für die beiden adaptiven Stufen lassen sie sich in einfache und schwierige *Testlets* gruppieren (siehe Abbildung 10.1). Nach der Bearbeitung der Aufgaben in einer (vorangegangenen) Stufe werden die Schülerinnen und Schüler anhand des Kriteriums der Anzahl richtig gelöster Aufgaben (temporär) in drei Gruppen eingeteilt. Dies sind die Gruppen mit geringer, mittlerer und hoher Anzahl gelöster Aufgaben (siehe Abbildung 10.1, linke und mittlere Säule unten). Diese drei Gruppen von Schülerinnen und Schülern werden nun in der jeweils folgenden Stufe in bestimmten prozentualen Anteilen den einfacheren und schwierigeren *Testlets* zugewiesen (siehe Pfeile in Abbildung 10.1). Aus der Gruppe der Schülerinnen und Schüler mit hoher Anzahl gelöster Aufgaben werden etwa zehn Prozent den leichten *Testlets* und etwa 90 Prozent den schwierigen *Testlets* zugeordnet, aus der Gruppe mit niedriger Anzahl gelöster Aufgaben werden etwa zehn Prozent den schwierigen *Testlets* und etwa 90 Prozent den leichten *Testlets* zugeordnet. Für die Gruppe mit mittlerer Anzahl gelöster Aufgaben ist die Zuordnung zu den leichten und schwierigen *Testlets* jeweils 50 Prozent. Erst bei der Datenauswertung wird die Lesekompetenz im Rahmen der *Skalierung* mit Modellen aus der IRT bestimmt. Diese stützt sich so für alle Schülerinnen und Schüler auf unterschiedliche Aufgaben, die jeweils an das individuelle Kompetenzniveau angepasst sind. Das in Abbildung 10.1 skizzierte adaptive Testdesign wurde in PISA 2018 in zwei unterschiedlichen Varianten (A und B) implementiert. Beide Varianten unterscheiden sich im Wesentlichen durch die unterschiedliche Reihenfolge der adaptiven Stufen 2 und 3 (siehe mittlere und rechte Säule in Abbildung 10.1). Die Variante A stellt dabei eine im Hinblick auf die Testeffizienz optimierte Version dar. Die zusätzliche Ergänzung der Variante B wird damit begründet, dass damit Effekte der Aufgabenreihenfolge (z. B. Debeer, Buchholz, Hartig & Janssen, 2014; Frey, Bernhardt & Born, 2017) kontrolliert und analysiert werden können.

Für eine detaillierte Darstellung des in PISA 2018 eingesetzten adaptiven Testdesigns wird auf die online verfügbaren Publikationen von *Educational Testing Service* (ETS) für PISA 2018 (z. B. ETS, 2016a) sowie auf den erwarteten internationalen *Technical Report* zu PISA 2018 der OECD verwiesen. Yamamoto et al. (2018) behandeln die Implementation und berichten die Effekte einer gesteigerten Messgenauigkeit des MAT im Rahmen der internationalen Studie zur Untersuchung von Alltagsfertigkeiten Erwachsener (PIAAC), deren grundlegendes Testdesign dem bei PISA 2018 eingesetzten adaptiven Testdesign gleicht.

10.3 Aspekte der technischen Durchführung 2018 in Deutschland

10.3.1 Datengewinnung – Ziehung der Stichprobe

Wie bei vergleichbaren Bildungsstudien wird auch in PISA eine systematische Teilerhebung der zugrundeliegenden *Population* oder *Grundgesamtheit* realisiert. Dies geschieht durch Ziehung einer geeigneten *Stichprobe* von fünfzehnjährigen Schülerinnen und Schülern. Die systematische, geplante Ziehung dieser Stichprobe erlaubt eine Verallgemeinerung der Befunde aus der Stichprobe auf die Population aller fünfzehnjährigen Schülerinnen und Schüler in Deutschland (vgl. Bortz & Schuster, 2010; Brown, 2010; Kish, 1965; Levy & Lemeshaw, 2008; Thompson, 2012). Wie in den meisten Teilnehmerstaaten², so erfolgt auch in Deutschland die Stichprobenziehung nach einem zweistufigen Prinzip. Die Schulen sind dabei auf der ersten Ziehungsebene (Stufe). Eine zweite Ziehungsebene stellen die eigentlichen Untersuchungseinheiten dar. Das sind die Schülerinnen und Schüler (für den Fragebogen und die PISA-Kompetenzdomänen) sowie die davon abhängige Stichprobe der Eltern (für den Elternfragebogen) und die unabhängige Stichprobe der Lehrkräfte (für den Lehrerfragebogen). Für Deutschland wurde dieses generelle Prinzip der Stichprobenziehung durch ein erweitertes Studiendesign ergänzt, um auch zusätzlich neunte Klassen ziehen zu können. Details zu dem erweiterten Stichprobendesign finden sich bei Mang et al. (2019). Die Stichprobenziehung stützt sich international auf eine *formale Definition der Zielpopulation* der Schülerinnen und Schüler (siehe Kasten).

Formale Definition der Zielpopulation (Schülerinnen und Schüler):

Alle Schülerinnen und Schüler („Fünfzehnjährige“), die im Zeitraum 1. Januar 2002 bis 31. Dezember 2002 (einschließlich) geboren sind und mindestens die 7. Klassenstufe besuchen.

Neben dieser formalen Definition stützt sich die Stichprobenziehung auf weitere Informationen und Variablen, welche dazu geeignet sind, die Population angemessen zu beschreiben. International werden diese Informationen im so genannten *Sampling Frame* festgehalten (vgl. Westat, 2017). Im *Sampling Frame* wird durch *explizite* und *implizite Stratifizierungsvariablen* sichergestellt, dass die zu ziehende Stichprobe der zugrundeliegenden Grundgesamtheit in zentralen Merkmalen möglichst ähnlich ist. Die Anzahl und Eigenschaft der Stratifizierungsvariablen muss daher (naturgemäß) zwischen den einzelnen an PISA teilnehmenden Staaten variieren, um die jeweilige Population mit beispie-

2 Die Ziehungsverfahren können im Detail zwischen den teilnehmenden Staaten aufgrund der unterschiedlichen empirischen Gegebenheiten im jeweiligen Bildungssystem variieren – auch werden in kleineren Staaten Vollerhebungen anstelle der Erhebung von Stichproben durchgeführt (z. B. OECD, 2018).

weise unterschiedlichen Schulsystemen oder unterschiedlichen föderalen Strukturen angemessen beschreiben zu können. Zentrale implizite und explizite Stratifizierungsvariablen stellen für die deutsche Teilstichprobe bei PISA die Variablen Bundesland und Schulart dar (vgl. Mang et al., 2019 für eine detaillierte Darstellung der Stratifizierung und Stichprobenziehung in Deutschland). Damit erfolgt die PISA-Stichprobenziehung so, dass sich die Anteile der in der Stichprobe befindlichen Schulen, Schularten sowie der Schülerinnen und Schüler über die einzelnen Bundesländer hinweg möglichst proportional zu den Verhältnissen in der Grundgesamtheit in ganz Deutschland verteilen. Für die Ziehung der Stichprobe werden vom internationalen Konsortium detaillierte Vorgaben gemacht (vgl. Westat, 2017) damit die Stichproben aller teilnehmenden Staaten in gleicher Weise deren Population fünfzehnjähriger Schülerinnen und Schüler repräsentieren.

10.3.2 Das Testsystem – Student Delivery System (SDS)

Die PISA-Datenerhebung erfolgt für die Schülerinnen und Schüler, Eltern und Lehrkräfte in unterschiedlicher Form. Während die Befragung der Eltern wegen der einfacheren Erreichbarkeit über papierbasierte Fragebögen erfolgt, werden die Daten von Schülerinnen und Schülern, Lehrkräften sowie Schulleiterinnen und Schulleitern mit computerbasierten Systemen in der jeweiligen Schule erfasst. Die Erhebung der PISA-Kompetenzdomänen und auch der Hintergrundinformationen erfolgt bei den Schülerinnen und Schülern über ein (offline) computerbasiertes System (vgl. ETS, 2016b). Dieses so genannte *Student Delivery System* (SDS) wird auf einem USB-Stick jeweils individuell für die Schülerinnen und Schüler zur Erhebung in den Schulen von den *Testleiterinnen* oder *Testleitern* mitgebracht. Diese sind schulexterne Personen, die im Vorfeld ausgewählt werden. Das SDS läuft während der Testung entweder auf den vorhandenen Schulcomputern oder auf Laptops, die am Testtag bei Bedarf gestellt werden. Zur Vorbereitung des PISA-Testtags wurde in einem ersten Schritt an den betreffenden Schulen erhoben, wie viele schuleigene Computer mit entsprechend notwendigen technischen Mindestanforderungen für die Durchführung von PISA zur Verfügung stehen. Für diese Überprüfung und die sonstige schulorganisatorische Planung des Testtages wurde zuvor von jeder Schule eine *Schulkoordinatorin* oder ein *Schulkoordinator* bestimmt. Es handelt sich dabei um schulinterne Lehrkräfte, die zwar die Jugendlichen während der PISA-Testung beaufsichtigen, aber an der eigentlichen Durchführung nicht beteiligt sind. Die computerbasierte Testsitzung verläuft für die Schülerinnen und Schüler in drei Schritten. Nach einer Begrüßung durch die Testleiterin oder den Testleiter erfolgt eine standardisierte mündliche Einweisung. Die Schülerinnen und Schüler erhalten unmittelbar vor der Erhebung individuelle Login-Daten, um sich nach Auswahl des Erhebungsabschnittes (siehe Abbildung 10.2, oben) an dem Computersystem anzumelden (siehe Abbildung 10.2, unten). In der folgenden computerbasierten Einweisung in das SDS (siehe exemplarisch Abbildung 10.3) erhalten die Schülerinnen und Schü-

ler eine Einführung in die Struktur und die einzelnen Elemente des Testsystems (siehe exemplarisch Abbildung 10.3).

Wähle deinen Abschnitt aus
Bitte warte, bis du erfährst, mit welchem Abschnitt du beginnen sollst.

Abschnitt 1	Die PISA-Tests
Abschnitt 2	Die PISA-Fragebögen

Klicke deine Auswahl an.

Login zu Abschnitt 1
Gib deine Schüler-ID und dein Passwort ein.

Schüler-ID: - - - -

Passwort: - -

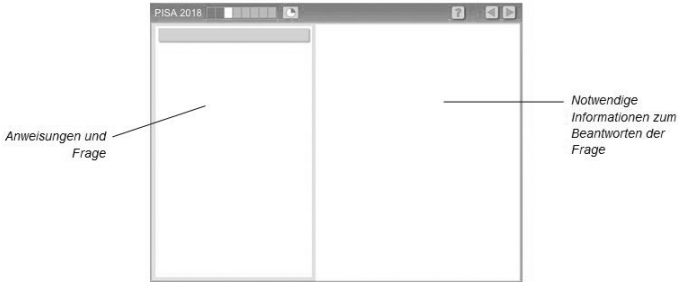
Abbildung 10.2: Auswahl- und Login-Bildschirm des PISA Student Delivery Systems (SDS)

PISA 2018 ? ◀ ▶

Allgemeine Einleitung
Das Test-Fenster

Bei vielen Fragen des Tests ist das Fenster in zwei Bereiche aufgeteilt, wie unten abgebildet.

Die Anweisungen und die Frage werden auf der linken Seite des Fensters angezeigt und die Informationen, die du zum Beantworten der Frage benötigst, werden auf der rechten Seite angezeigt.



Bitte klicke auf den WEITER-Pfeil, um fortzufahren.

Abbildung 10.3: Bildschirm mit allgemeiner Einführung zur Erklärung des PISA-Testsystems (SDS)

Der erste Abschnitt der Testung beginnt mit zwei 60-minütigen Blöcken zu den PISA-Kompetenzbereichen, welche durch eine Pause von fünf Minuten unterbrochen sind. Es folgen eine 15-minütige Pause und der zweite Testabschnitt mit 50 Minuten Dauer, in dem die PISA-Fragebögen bearbeitet werden (siehe Abbildung 10.2 oben). Im Anschluss lösen die Schülerinnen und Schüler einige „Logikaufgaben“, wofür 15 Minuten vorgesehen sind. Es handelt sich dabei um die Teilskala „Schlussfolgerndes Denken figuraler Teil“ des Berliner Tests zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Klassenstufe (BEFKI 8–10 – Wilhelm, Schroeders & Schipolowski, 2014).

10.3.3 Prozeduren der Datenaufbereitung

Am Ende der Testung an einer Schule werden die USB-Sticks mit den darauf befindlichen verschlüsselten Antwortdaten der Schülerinnen und Schüler von der Testleiterin oder dem Testleiter eingesammelt und an die IEA Hamburg gesendet. Die einzelnen USB-Sticks werden dort nach einer vom internationalen Konsortium definierten Prozedur ausgelesen und lokal in einer Software zur Datenverwaltung gespeichert. Eingesetzt wird hierbei die Software *IEA Data Management Expert* (DME), welche in einer für die PISA-Erhebung angepassten Version vom internationalen Konsortium bereitgestellt wird. Die DME-Software ist eine eigenständige Anwendung, die für den Betrieb keine Internetverbindung benutzt. Die DME-Anwendung arbeitet mit einer lokal gespeicherten Datenbankdatei, welche international anhand struktureller und relationaler Vorgaben erstellt wird und das *Codebuch* definiert. Das Codebuch ist ein vollständiger Katalog aller zu erhebenden und zu verwaltenden Variablen, deren (Antwort-)Kategorien und deren Position in den einzelnen Datensätzen (z. B. Datensätze der Schülerinnen und Schüler, Eltern und Lehrkräfte). Die Software erleichtert so die Adaption und Plausibilitätsüberprüfung der Daten im Rahmen verschiedener Prozeduren zur Analyse der Datenkonsistenz. Ein Hauptziel dieser Prozeduren ist zu prüfen, ob alle Informationen in der Datenbank zu der international vorgegebenen und definierten Datenstruktur passen. Eventuell notwendige nationale Anpassungen der Fragebögen im Rahmen der vorgegebenen Leitlinien zur Adaption und Übersetzung (vgl. cApStAn & Halleux, 2016, 2018) sollen angemessen berücksichtigt werden. Dies sind in Deutschland beispielsweise Fragen zur besuchten Schulart oder zum erreichten Bildungsabschluss der Eltern. Es soll sichergestellt werden, dass die für internationale Vergleiche herangezogenen Variablen für alle teilnehmenden Staaten – trotz unterschiedlicher Bildungssysteme – vergleichbar aufbereitet sind. Ein weiterer wesentlicher Aspekt beim Einsatz des DME besteht darin, dass keinerlei personenbezogene Informationen in den Datensätzen enthalten sind und die Daten somit in *anonymisierter Form*³ zur Skalierung

3 Die Daten liegen zum Zeitpunkt der Übermittlung an das internationale Konsortium in *anonymisierter Form* vor, da zeitgleich mit der Datenübermittlung die Schlüssel, welche die Codes der Teilnehmerinnen und Teilnehmer mit deren Klarnamen verknüpfbar machen, in den Schulen gelöscht sind.

rung und Auswertung an das internationale Konsortium übermittelt werden können. Detaillierte Informationen zum gesamten Dateneingabe- und Aufbereitungsprozess, welcher im Hinblick auf grundsätzliche Abläufe und Strukturen identisch mit den bei PISA 2015 angewendeten Prozessen ist, sind bei Mang et al. (2019) sowie im internationalen Technischen Bericht zu PISA 2015 beschrieben (vgl. OECD, 2018, S. 187 ff.).

10.4 PISA-Daten und Methoden für Sekundäranalysen

Die Erhebung von PISA-Daten ist mit einer nicht unerheblichen Belastung der Teilnehmerinnen und Teilnehmer sowie mit finanziellem Aufwand aus öffentlichen Geldern verbunden. Nicht zuletzt daraus ergibt sich die legitime Forderung nach einer öffentlichen Verfügbarkeit der Daten und Transparenz der eingesetzten Methoden. Dies bildet die Basis für die Kontrolle und Akzeptanz zentraler Befunde im Rahmen der Berichterstattung rund um PISA. Allerdings erfordern die komplexen Stichproben- und Testdesigns bei den Analysen zur Auswertung von PISA-Daten, neben deren grundsätzlicher Verfügbarkeit, die Anwendung fortgeschrittener statistischer Auswertungsmethoden. So müssen beispielsweise aufgrund der in den einzelnen Teilnehmerstaaten unterschiedlichen Stratifizierung der Stichproben so genannte *Replikationsverfahren* zur Schätzung der *Varianz* von statistischen Koeffizienten eingesetzt werden (vgl. z. B. Arrenberg, 1998; Fay, 1989; Judkins, 1990). Diese bilden die Grundlage für die Beurteilung der *Signifikanz* der statistischen Koeffizienten. Die folgenden beiden Abschnitte sollen daher praktische Hinweise zur Verfügbarkeit und zur Auswertung der PISA-Daten an interessierte Leserinnen und Leser geben.

10.4.1 Verfügbarkeit der PISA-Daten

Die internationale Datenbasis aus den PISA-Erhebungen wird von den jeweils durch die OECD beauftragten Konsortien (ETS für 2015 und 2018) aufbereitet und im Anschluss über die Webseiten der OECD der interessierten Öffentlichkeit und Wissenschaft zur Verfügung gestellt (vgl. <http://www.oecd.org/pisa/data>). Daneben gibt es den PISA-Data-Explorer (PDE), eine webbasierte Benutzeroberfläche, die es der Anwenderin oder dem Anwender ermöglicht, die internationale PISA-Datenbasis für Auswertungen zu nutzen. Schließlich werden die Daten aus den *nationalen Ergänzungen* für Deutschland (zusätzliche Variablen und Teilstichproben) über das Forschungsdatenzentrum (FDZ), ein Teil des *Zentrums für internationale Bildungsvergleichsstudien* (ZIB), zur Verfügung gestellt (z. B. für PISA 2012: Prenzel et al., 2015; für PISA 2015: Reiss et al., 2019). Alle Datenbasen lassen sich zur Replikation der publizierten Befunde oder auch zu weitergehenden Sekundäranalysen nutzen.

10.4.2 Praktische Hinweise zur Analyse von PISA-Daten

Für die Auswertung von PISA-Daten stehen unterschiedliche „Werkzeuge“ zur Verfügung. Im Folgenden werden drei Möglichkeiten zur Auswertung von PISA-Daten skizziert. Diese erfordern zunehmend größere methodische Kenntnisse und Fertigkeiten bei den Anwenderinnen und Anwendern, wobei auch zunehmend komplexere Analysen möglich werden.

PISA-Data-Explorer – webbasierte Anwendung

Beim *PISA-Data-Explorer* (PDE) handelt es sich um eine webbasierte Benutzeroberfläche. Sie ermöglicht es, die von der OECD bereitgestellte internationale PISA-Datenbasis über einen Webbrowser abzufragen. Die Datenbasis umfasst die Daten aller PISA-Runden und jeweils teilnehmenden Staaten (seit 2000). Der PDE ist verfügbar unter <http://www.oecd.org/pisa/data/> wobei die jeweils aktuellste Version (derzeit die Versionen 2012 und 2015⁴) ausgewählt werden muss. Die dem PDE zugrundeliegende Datenbank speist sich aus den *Public Use Files* und umfasst damit mehr als 2,4 Millionen Fälle. Mit dem PDE können eine Vielzahl von Statistiken berechnet werden, wie etwa Mittelwerte von Teilnehmerstaaten, Standardabweichungen, Standardfehler, Prozentsätze für Untergruppen (z. B. Geschlecht oder Schulart), Prozentsätze für PISA-Kompetenzstufen sowie Perzentile. Alle Statistiken werden unter Berücksichtigung des Stichproben- und Testdesigns mit entsprechenden *Replikationsverfahren* berechnet. Es besteht die Möglichkeit, die Ergebnisse in grafischer Form darzustellen und diese in den Formaten Microsoft® Word, Microsoft® Excel® oder als HTML-Datei zu exportieren. Durch die browserbasierte Benutzeroberfläche sind die technischen Anforderungen (plattformübergreifend) sowie die Anforderungen hinsichtlich des technischen Wissens der Anwenderin oder des Anwenders vergleichsweise gering. Sie beschränken sich letztlich auf die Verfügbarkeit eines einfachen handelsüblichen Computers mit Internetzugang und aktuellen Browsers sowie basale Fertigkeiten im Umgang mit interaktiven Webseiten.

Der IEA-International Database-Analyzer

Der *International Database Analyzer* (IEA IDB Analyzer) wird von der *International Association for the Evaluation of Educational Achievement* (IEA) entwickelt und über die Webseiten der IEA (<http://www.iea.nl/data.html>) bereitgestellt. Mit dem IEA IDB Analyzer können Datensätze aus verschiedenen *International Large Scale Assessments* (ILSA) zusammengestellt und unter korrekter Berücksichtigung der Stichprobeneigenschaften und Gewichtung- und Replikationsvariablen analysiert werden. Bei dem IEA IDB Analyzer handelt es sich um eine eigenständige Softwareanwendung für Win-

4 Die öffentliche Verfügbarkeit der Version des PDE für die PISA-Datenbasis 2018 wird für Ende 2019 beziehungsweise Anfang 2020 erwartet.

dows Betriebssysteme, welche über ein interaktives Menüsystem eine Analysesyntax zur Datenauswertung für die Statistikanwendungen SAS oder SPSS erzeugt. Die technischen Voraussetzungen für den erfolgreichen Einsatz des IEA IDB Analyzers bestehen in einem leistungsfähigen aktuellen Computersystem mit einer Installation der (kommerziellen) Statistikanwendungen SAS oder SPSS. Auf Seiten der Anwenderinnen und Anwender werden wenigstens grundlegende Kenntnisse in statistischen Analysen sowie vertiefte Kenntnisse bezüglich der Besonderheiten der Stichprobenziehung und der eingesetzten *Replikationsverfahren* zu Varianzschätzung in ILSAs vorausgesetzt. Ferner sollten wenigstens grundlegende Kenntnisse in den jeweils eingesetzten Statistikanwendungen (SAS oder SPSS) vorhanden sein.

Die freie Statistikumgebung R und verschiedene R-Pakete

Insbesondere in den vergangenen Jahren hat sich in der quantitativ, methodisch orientierten Praxis der Datenauswertung in den Sozialwissenschaften die freie Statistikumgebung *R* (R Core Team, 2019) zunehmend durchgesetzt. Neben *R*-Paketen, welche unterschiedliche Algorithmen zur Skalierung von Test- und Fragebogenskalen im Rahmen der Item Response Theory (IRT) bereitstellen, wie die Pakete *TAM* (Robitzsch, Kiefer & Wu, 2019), *eRm* (Mair, Hatzinger, Maier & Rusch, 2019), *pairwise* (Heine, 2019) und *ltm* (Rizopoulos, 2018), stehen zunehmend auch *R*-Pakete zur Verfügung, welche sich mit speziellen Funktionen der Auswertungen von Daten aus komplexen Stichproben widmen, wie sie bei ILSAs wie PISA anfallen. Zum derzeitigen Stand sind hier beispielsweise die *R*-Pakete *survey* (Lumley, 2019b), *mitools* (Lumley, 2019a), *BIFIEsurvey* (BIFIE, Robitzsch & Oberwimmer, 2019), *intsvy* (Caro & Biecek, 2019) sowie die datenorientierten, nur über GitHub (<https://github.com/>) verfügbaren *R*-Pakete *PISA2000lite*, *PISA2003lite*, *PISA2006lite*, *PISA2009lite* und *PISA2012lite* (z. B. Biecek, 2013) zu nennen. Bei den letztgenannten handelt es sich um *R*-Pakete, die „lediglich“ die im *R*-Format aufbereiteten internationalen Daten der jeweiligen PISA-Runde enthalten. Der Vorteil der Verwendung dieser Pakete besteht darin, dass sich die *R*-Anwenderin oder der *R*-Anwender den oft als kompliziert und fehleranfällig empfundenen Prozess des „Dateneinlesens“ in *R* erspart. Dieser Prozess des Dateneinlesens wird in dem Paket *intsvy* vereinfacht, indem speziell auf die PISA-Runden 2000 bis 2012 sowie auf die PISA-Runden 2015 bis einschließlich 2018 zugeschnittene Funktionen bereitgestellt werden, um die über die OECD-Seiten öffentlichen oder das FDZ verfügbaren Daten im SPSS-Format einfach einlesen zu können. Daneben werden in dem Paket *intsvy* Funktionen bereitgestellt, welche deskriptive Auswertungen ermöglichen, die sich im Wesentlichen mit dem Funktionsumfang des IEA Analyzers decken. Das *R*-Paket *BIFIEsurvey* geht im Hinblick auf die Komplexität der angebotenen Analysemöglichkeiten über diese in *intsvy* angebotene Basisfunktionalität hinaus. Hier besteht beispielsweise die Möglichkeit, Datenstrukturen mit *genesteten multiplen Imputationen* (z. B. Weirich et al., 2014) zu analysieren sowie die *Mehrebenenstruktur* der in PISA erhobenen Daten (Schülerinnen und Schüler in Klassen in Schulen) zu berücksichtigen. Die beiden Pakete *BIFIEsurvey* und *intsvy* basieren

im Hinblick auf ihre Analysefunktionen auf der grundlegenden Funktionalität, die das R-Paket *survey* (Lumley, 2019b) bereitstellt. Dieses Paket stellt mit maximaler Flexibilität allgemeine und grundlegende Funktionen zur angemessenen Behandlung von stratifizierten Datensätzen zur Verfügung. Zur spezifischen Anwendung bei PISA-Datensätzen sind hier aber oft als kompliziert empfundenen Anpassungen im Funktionsaufruf erforderlich, was letztlich, für den Kontext von PISA, die Entwicklung und Anwendung der Pakete *intsvy* und *BIFIEsurvey* rechtfertigt.

10.5 Zusammenfassung

Im vorliegenden Kapitel wurden die zentralen Veränderungen bei der Durchführung und Auswertung der PISA-Studie 2018 skizziert. Einen Schwerpunkt bilden dabei die methodischen Aspekte. Insgesamt umfasst die mit PISA verbundene Methodologie komplexe Testdesigns, Stichprobenziehungsverfahren und Auswertungsverfahren, welche im Rahmen dieses Kapitels jeweils im Überblick angesprochen wurden. Die psychometrische Modellierung im Rahmen der Skalierung der PISA-Daten hat sich für die PISA-Runde 2018 im Vergleich zu 2015, nach den grundlegenden Veränderungen zwischen den Runden 2012 und 2015 (vgl. Heine et al., 2016), nicht verändert. Für eine vergleichende Darstellung unterschiedlicher Skalierungsmodelle und Aspekte der psychometrischen Modellierung sei daher auf die technischen Kapitel früherer Berichtsbände (Heine et al., 2016; Heine & et al., 2013) sowie auf die detaillierte Dokumentation der Skalierung von PISA-Daten im internationalen technischen Bericht zu PISA 2015 verwiesen (OECD, 2018).

Ein für PISA 2018 zentraler und neuer technischer Aspekt ist der Einsatz eines adaptiven, computerbasierten Testdesigns (Multistage Adaptive Testdesign – MAT). Neben der Darstellung der methodologischen Veränderungen wurden in diesem Kapitel Aspekte der praktischen Durchführung der PISA-Erhebung an den Schulen in Deutschland skizziert. Für replizierende und weiterführende Analysen zu PISA wurde die Verfügbarkeit der Daten sowie Methoden zu deren Auswertung behandelt. Für eine vertiefende Betrachtung der mit PISA verbundenen Methodologie wird auf die zitierten grundlegenden wissenschaftlichen Quellen verwiesen. Zusätzlich sind die spezifisch zur PISA-Runde 2018 verfügbaren Dokumente der OECD und des internationalen PISA-Konsortiums zu nennen, welche in der OECD-Library (<http://www.oecd.org/pisa/publications/>) über die darin verfügbare Suchmaske durch Eingabe der Fachbegriffe erreichbar sind.

Literatur

- Arrenberg, J. (1998). *Schätzung der Varianz von Mittelwertschätzern in endlichen Populationen*. Göttingen: Vandenhoeck & Ruprecht.
- Bennett, R. E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *The Journal of Technology, Learning and Assessment*, 1(1). Verfügbar unter <https://ejournals.bc.edu/index.php/jtla/article/view/1667>
- Biecek, P. (2013). Pbiecek/PISA2000lite (R package Version 1.0). Verfügbar unter <https://github.com/pbiecek/PISA2000lite>
- BIFIE, Robitzsch, A. & Oberwimmer, K. (2019). *BIFIEsurvey: Tools for Survey Statistics in Educational Assessment* (R package version 3.3-12). Verfügbar unter <https://CRAN.R-project.org/package=BIFIEsurvey>
- BMBF = Bundesministerium für Bildung und Forschung, Jude, N., Hartig, J. & Klieme, E. (Hrsg.). (2008). *Kompetenzerfassung in pädagogischen Handlungsfeldern Theorien, Konzepte und Methoden* (Bd. 26). Bonn, Berlin.
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer. doi: <https://doi.org/10.1007/978-3-642-12770-0>
- Bradshaw, L. & Templin, J. (2014). Combining item response theory and diagnostic classification models: A psychometric model for scaling ability and diagnosing misconceptions. *Psychometrika*, 79(3), 403–425. <https://doi.org/10.1007/s11336-013-9350-4>
- Brown, R. S. (2010). Sampling. In P. Peterson, E. Baker & B. McGaw (Hrsg.), *International Encyclopedia of Education* (S. 142–146). Oxford: Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.00294-3>
- Browne, M. W. (2000). Psychometrics. *Journal of the American Statistical Association*, 95 (450), 661–665. <https://doi.org/10.2307/2669413>
- cApStAn & Halleux, B. (2016). PISA 2018 translation and adaptation guidelines. (OECD, Hrsg.). Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/PISA-2018-TRANSLATION-AND-ADAPTATION-GUIDELINES.pdf>
- cApStAn & Halleux, B. (2018). PISA 2018 translation kit “read-me” file. (OECD, Hrsg.). Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/PISA-2018-TRANSLATION-KIT.pdf>
- Caro, D. & Biecek, P. (2019). *intsvy: International assessment data manager* (R package version 2.4). Verfügbar unter <https://CRAN.R-project.org/package=intsvy>
- Debeer, D., Buchholz, J., Hartig, J. & Janssen, R. (2014). Student, school, and country differences in sustained test-taking effort in the 2009 PISA reading assessment. *Journal of Educational and Behavioral Statistics*, 39(6), 502–523. <https://doi.org/10.3102/1076998614558485>
- Diedenhofen, B. & Musch, J. (2018). An investigation into the usefulness of time-efficient item selection in computerized adaptive testing. *Psychological Test and Assessment Modeling*, 60(3), 289–308.
- ETS. (2016a, März). PISA 2018 integrated design. (OECD, Hrsg.). Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/PISA-2018-INTEGRATED-DESIGN.pdf>
- ETS. (2016b, November). Student delivery system manual field trial. (OECD, Hrsg.). Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/STUDENT-DELIVERY-SYSTEM-MANUAL-Field-Trial.pdf>
- ETS. (2018a). PISA 2018 Released field trial new reading items. (OECD, Hrsg.). Verfügbar unter http://www.oecd.org/pisa/test/PISA_2018_FT_Released_New_Reading_Items.pdf

- ETS. (2018b). Beyond PISA 2015 a longer term strategy of PISA. (OECD, Hrsg.). Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/BEYOND-PISA-2015-A-LONGER-TERM-STRATEGY-OF-PISA.pdf>
- Fay, R. E. (1989). Theory and Application of Replicate Weighting for Variance Calculations. In *Proceedings of the Section on Survey Research Methods of the American Statistical Association* (S. 212–217). Alexandria: American Statistical Association.
- Frey A., Bernhardt, R. & Born, S. (2017). Handling of Item Positions Effects in the Development of Computerized Adaptive Tests. *Diagnostica*, 63(3), 167–178. <https://doi.org/10.1026/0012-1924/a000173>
- Frey, D. A. & Ehmke, D. T. (2008). Hypothetischer Einsatz adaptiven Testens bei der Überprüfung von Bildungsstandards. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 169–184). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90865-6_10
- Frey, A., Hartig, J. & Rupp, A. A. (2009). Booklet designs in large-scale assessments of student achievement: theory and practice. *Educational Measurement: Issues and Practice*, 28(3), 39–53. <https://doi.org/10.1111/j.1745-3992.2009.00154.x>
- Frey, A. & Seitz, N.-N. (2009). Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation*, 35(2–3), 89–94. <https://doi.org/10.1016/j.stueduc.2009.10.007>
- Frey, A., Seitz, N.-N. & Brandt, S. (2016). Testlet-Based Multidimensional Adaptive Testing. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01758>
- Goldhammer, F., Naumann, J., Rölke, H., Stelter, A. & Tóth, K. (2017). Relating product data to process data from computer-based competence assessment. In D. Leutner, J. Fleischer, J. Grünkorn & E. Klieme (Hrsg.), *Competence Assessment in Education: Research, Models and Instruments* (S. 407–425). Cham: Springer. https://doi.org/10.1007/978-3-319-50030-0_24
- Gonzalez, E. & Rutkowski, L. (2010). Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. In M. Von Davier & D. Hastedt (Hrsg.), *Issues and Methodologies in Large-Scale Assessments* (S. 125–156). Hamburg/Princeton: IER Institut.
- Hartig, J. (2008). Kompetenzen als Ergebnisse von Bildungsprozessen. In BMBF = Bundesministerium für Bildung und Forschung, N. Jude, J. Hartig & E. Klieme (Hrsg.), *Kompetenzerfassung in pädagogischen Handlungsfeldern Theorien, Konzepte und Methoden* (Bd. 26). Bonn/Berlin: Bundesministerium für Bildung und Forschung.
- Heine, J.-H. (2019). *pairwise: Rasch Model Parameters by Pairwise Algorithm* (R package version 0.4.4-5.2). Verfügbar unter <https://CRAN.R-project.org/package=pairwise>
- Heine, J.-H., Mang, J., Borchert, L., Gomolka, J., Kröhne, U., Goldhammer, F. & Sälzer, C. (2016). Kompetenzmessung in PISA 2015. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015: Eine Studie zwischen Kontinuität und Innovation* (S. 383–430). Münster: Waxmann.
- Heine, J.-H., Sälzer, C., Borchert, L., Siberns, H. & Mang, J. (2013). Technische Grundlagen des fünften internationalen Vergleichs. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 309–346). Münster: Waxmann.
- Judkins, D. R. (1990). Fay's method for variance estimation. *Journal of Official Statistics*, 6(3), 223–239.
- Kish, L. (1965). *Survey sampling*. New York: Wiley.

- Klieme, P. D. E. & Hartig, J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (S. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90865-6_2
- Kubinger, K. D. (2017). Adaptive testing. In K. Schweizer & C. DiStefano (Hrsg.), *Principles and methods of test construction: standards and recent advances* (Bd. 3, S. 104–119). Göttingen: Hogrefe.
- Kubinger, K. D., Hohensinn, C., Hofer, S., Khorramdel, L., Frebort, M., Holocher-Ertl, S., Reif, M. & Sonnleitner, P. (2011). Designing the test booklets for Rasch model calibration in a large-scale assessment with reference to numerous moderator variables and several ability dimensions. *Educational Research and Evaluation*, 17(6), 483–495. <https://doi.org/10.1080/13803611.2011.632666>
- Levy, P. S. & Lemeshow, S. (2008). *Sampling of populations: Methods and applications* (4. Aufl.). Hoboken: Wiley. <https://doi.org/10.1002/9780470374597>
- Lumley, T. (2019a). *mitools: Tools for multiple imputation of missing data* (R package Version 2.4). Verfügbar unter <http://cran.r-project.org/package=mitools>
- Lumley, T. (2019b, August). *survey: Analysis of complex survey samples* (R package Version 3.36). Verfügbar unter <http://cran.r-project.org/package=survey>
- Mair, P., Hatzinger, R., Maier, M. J. & Rusch, T. (2019). *eRm: Extended Rasch Modeling* (R package Version 1.0-0). Verfügbar unter <https://CRAN.R-project.org/package=eRm>
- Mang, J., Wagner, S., Gomolka, J., Schäfer, A., Meinck, S. & Reiss, K. (2019). *Technische Hintergrundinformationen PISA 2018*. München: Technische Universität München. <http://doi.org/10.14459/2019md1518258>
- Mummendey, H. D. & Grau, I. (2014). *Die Fragebogen-Methode*. Göttingen: Hogrefe.
- Narens, L. (1981). On the scales of measurement. *Journal of Mathematical Psychology*, 24(3), 249–275. [https://doi.org/10.1016/0022-2496\(81\)90045-6](https://doi.org/10.1016/0022-2496(81)90045-6)
- OECD. (2018). *PISA 2015 technical report*. Paris: OECD Publishing. Verfügbar unter <https://www.oecd.org/pisa/data/2015-technical-report/>
- OECD. (2019). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- Parshall, C. G., Harmes, J. C., Davey, T. & Pashley, P. J. (2010). Innovative items for computerized testing. In W. J. van der Linden & C. A. W. Glas (Hrsg.), *Elements of adaptive testing* (S. 215–230). New York: Springer. https://doi.org/10.1007/978-0-387-85461-8_11
- Parshall, C. G., Spray, J. A., Kalohn, J. C. & Davey, T. (Hrsg.). (2002). *Practical considerations in computer-based testing*. New York: Springer. <https://doi.org/10.1007/978-1-4613-0083-0>
- Prenzel, M., Gogolin, I. & Krüger, H.-H. (Hrsg.). (2007). *Kompetenzdiagnostik*. Wiesbaden: VS Verlag für Sozialwissenschaft. <https://doi.org/10.1007/978-3-531-90865-6>
- Prenzel, M., Sälzer, C., Klieme, E., Köller, O., Mang, J., Heine, J.-H., Schiepe-Tiska, A. & Müller, K. (2015). *Programme for International Student Assessment 2012 (PISA 2012)*. Version: 3. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. http://doi.org/10.5159/IQB_PISA_2012_v3
- R Core Team. (2019). *R: A language and environment for statistical computing* (Version 3.6.1). Verfügbar unter <https://www.R-project.org/>
- Reiss, K., Sälzer, C., Schiepe-Tiska, A., Mang, J., Heine, J.-H., Weis, M., Klieme, E. & Köller, O. (2019). *Programme for International Student Assessment 2015 (PISA 2015)*. Version 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. http://doi.org/10.5159/IQB_PISA_2015_v1

- Rizopoulos, D. (2018). *ltm: Latent trait models under IRT* (R Version 1.1-1). Verfügbar unter <https://CRAN.R-project.org/package=ltm>
- Robitzsch, A., Kiefer, T. & Wu, M. (2019). *TAM: Test analysis modules* (R package version 3.3-10). Verfügbar unter <https://CRAN.R-project.org/package=TAM>
- Steyer, R. & Eid, M. (2000). *Messen und Testen*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-56924-1>
- Thompson, S. K. (2012). *Sampling* (3. Aufl.). Hoboken: Wiley. <https://doi.org/10.1002/9781118162934>
- Torgerson, W. S. (1961). Scaling and test theory. *Annual Review of Psychology*, 12(1), 51–70. <https://doi.org/10.1146/annurev.ps.12.020161.000411>
- van der Linden, W. J. & Glas, C. A. W. (2000). *Computerized adaptive testing: Theory and practice*. New York: Kluwer Academic Publishers. <https://doi.org/10.1007/0-306-47531-6>
- van der Linden, W. J. & Glas, C. A. W. (Hrsg.). (2010). *Elements of adaptive testing*. New York: Springer. <https://doi.org/10.1007/978-0-387-85461-8>
- Wainer, H. & Dorans, N. J. (2000). *Computerized adaptive testing: A primer*. Hillsdale: Lawrence Erlbaum. <https://doi.org/10.4324/9781410605931>
- Weirich, S., Haag, N., Hecht, M., Böhme, K., Siegle, T. & Lüdtke, O. (2014). Nested multiple imputation in large-scale assessments. *Large-Scale Assessments in Education*, 2(1), 1–18. <https://doi.org/10.1186/s40536-014-0009-0>
- Weiss, D. J. (2011). Better data from better measurements using computerized adaptive testing. *Journal of Methods and Measurement in the Social Sciences*, 2(1), 1–27. <https://doi.org/10.2458/jmm.v2i1.12351>
- Westat, (2017). *Main survey school sampling preparation manual*. Paris: OECD Publishing. Verfügbar unter <http://www.oecd.org/pisa/pisaproducts/MAIN-SURVEY-SCHOOL-SAMPLING-PREPARATION-MANUAL.pdf>
- Wilhelm, O., Schroeders, U. & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluider und kristalliner Intelligenz für die 8. bis 10. Jahrgangsstufe (BEFKI 8–10)*. Göttingen: Hogrefe.
- Yamamoto, K., Khorramdel, L. & Shin, H. J. (2018). Introducing multistage adaptive testing into international large-scale assessments designs using the example of PIAAC. *Psychological Test and Assessment Modeling*, 60(3), 347–368.

Abbildungsverzeichnis

Abbildung 2.1:	Modell der Leseprozesse in PISA 2018. Abbildung adaptiert nach OECD, 2019, S. 33	29
Abbildung 2.2:	Aufgabeneinheit „Hühnerforum“ mit Beispielaufgaben 1, 3 und 6	37
Abbildung 2.3:	Aufgabeneinheit „Die Osterinsel“ mit den Beispielaufgaben 3, 6 und 7	42
Abbildung 3.1:	Mittelwerte, Streuungen und Perzentilbänder der Gesamtskala Lesekompetenz der OECD-Staaten	59
Abbildung 3.2:	Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe Ia, Ib, Ic oder darunter sowie auf Kompetenzstufe V oder VI für die Gesamtskala Lesekompetenz in den OECD-Staaten.....	61
Abbildung 3.3:	Mittelwerte der Lesekompetenz nach Geschlecht in den OECD-Staaten.....	63
Abbildung 3.4:	Mittelwerte und Streuungen der Teilskala <i>Lokalisieren von Informationen</i> in den OECD-Staaten	64
Abbildung 3.5:	Mittelwerte und Streuungen der Teilskala <i>Textverstehen</i> in den OECD-Staaten	65
Abbildung 3.6:	Mittelwerte und Streuungen der Teilskala <i>Bewerten und Reflektieren</i> in den OECD-Staaten	66
Abbildung 3.7:	Perzentilbänder für die Lesekompetenz für die Gesamtstichprobe und nach Schulart in Deutschland.....	69
Abbildung 3.8:	Prozentuale Anteile der Fünfzehnjährigen auf den Stufen der Lesekompetenz nach Schulart in Deutschland	70
Abbildung 3.9:	Prozentuale Anteile auf den Stufen der Lesekompetenz nach Geschlecht in Deutschland	71
Abbildung 3.10:	Mittelwerte der Lesekompetenz in PISA 2009, 2015 und 2018 in den OECD-Staaten	73
Abbildung 3.11:	Veränderung der mittleren Lesekompetenz von 2000 bis 2018 in Deutschland.....	74
Abbildung 3.12:	Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe Ia, Ib, Ic oder darunter sowie auf Kompetenzstufe V oder VI für die Gesamtskala Lesekompetenz bei PISA 2018, 2015 und 2009 in Deutschland.....	75
Abbildung 4.1:	Lesefreude im internationalen Vergleich.....	86
Abbildung 4.2:	Lesemenge im internationalen Vergleich	88
Abbildung 4.3:	Online-Lesen im internationalen Vergleich.....	89
Abbildung 4.4:	Online-Lesen differenziert nach Geschlecht.....	91
Abbildung 4.5:	Lesen zum Vergnügen in Abhängigkeit von der Lesemenge in Deutschland.....	91
Abbildung 4.6:	Wahrgenommene Kompetenz im internationalen Vergleich.....	92
Abbildung 4.7:	Wahrgenommene Schwierigkeit im internationalen Vergleich	93
Abbildung 4.8:	Lesemotivation und lesebezogenes Selbstkonzept differenziert nach Schulart.....	95
Abbildung 4.9:	Online-Lesen differenziert nach Schulart	95
Abbildung 4.10:	Veränderung der Lesefreude in Deutschland differenziert nach Schulart und Geschlecht	96
Abbildung 4.11:	Lernszenarien Lesestrategiewissen – <i>Zusammenfassen von Informationen</i>	99
Abbildung 4.12:	Lernszenarien Lesestrategiewissen – <i>Verstehen und Erinnern eines Textes</i>	100
Abbildung 4.13:	Lernszenarien Lesestrategiewissen – <i>Beurteilung der Qualität und Vertrauenswürdigkeit einer Online-Quelle</i>	100
Abbildung 4.14:	Lesestrategiewissen im internationalen Vergleich.....	102
Abbildung 4.15:	Lesestrategiewissen nach Schularten	103
Abbildung 4.16:	Veränderungen des Lesestrategiewissens in den Szenarien <i>Erinnern und Zusammenfassen</i> in Deutschland zwischen PISA 2009 und PISA 2018	105

Abbildung 5.1:	Entwicklung der Anzahl tragbarer Computer pro Schülerin beziehungsweise Schüler in der neunten Klassenstufe zwischen PISA 2015 und PISA 2018 nach Schulart.....	115
Abbildung 5.2:	Zusatzangebote für den Unterricht der Landessprache im OECD-Durchschnitt, in Deutschland insgesamt und nach Schularten.....	122
Abbildung 6.1:	Mittelwerte, Streuungen und Perzentilbänder des sozioökonomischen beruflichen Status (HISEI) der OECD-Staaten	137
Abbildung 6.2:	Lesekompetenz und Varianzaufklärung (R^2) durch den sozioökonomischen und -kulturellen Status (ESCS) in den OECD-Staaten	142
Abbildung 6.3:	Mittelwerte und Streuungen der Lesekompetenz differenziert nach EGP-Klassen (Bezugsperson) bei PISA 2000, PISA 2009 und 2018 in Deutschland.....	146
Abbildung 6.4:	Prozentuale Anteile auf den Stufen der Lesekompetenz nach Zuwanderungsstatus in Deutschland.....	157
Abbildung 7.1:	Untersuchungsdesign der nationalen Ergänzung mit Typ-A- und Typ-B-Schulen (2. Testtag) in Anknüpfung an das Design von PISA 2018 (1. Testtag).....	171
Abbildung 7.2:	Erwarteter Lösungserfolg bei Leseitems in Abhängigkeit der Ausprägung einer Itemeigenschaft und des Modus; die jeweilige Anzahl der Items ist in Klammern angegeben	175
Abbildung 7.3:	Zusammenhang der Itemdiskriminationen der CBA-Messung und der PBA-Messung je Domäne. Die gefüllten Punkte zeigen Items mit signifikanten Unterschieden zwischen CBA und PBA an.....	176
Abbildung 7.4:	Zusammenhang der Itemschwierigkeiten der CBA-Messung und der PBA-Messung je Domäne. Die gefüllten Punkte zeigen Items mit signifikanten Unterschieden zwischen CBA und PBA an. Die gestrichelte Linie stellt die Trendlinie der invariant klassifizierten Items dar.....	177
Abbildung 7.5:	Schematische Darstellung des Vorgehens bei der marginalen Trendschätzung. Die Stichproben der Typ-B-Schulen der nationalen Ergänzungsstudie 2018 („2018 PBA“) und der internationalen PISA-Studie („2018 CBA“) werden als zufallsäquivalent angesehen	179
Abbildung 7.6:	Originaler und marginaler Trend (C2) der Leistungen in den drei Domänen zwischen 2009 und 2018.....	182
Abbildung 8.1:	Beispielaufgaben der Aufgabeneinheit „Bergsteigen am Mount Fuji“ (mit Genehmigung entnommen aus Hammer et al., 2016, S. 228)	194
Abbildung 8.2:	Perzentilbänder mathematischer Kompetenz in den OECD-Staaten.....	196
Abbildung 8.3:	Prozentuale Anteile der Schülerinnen und Schüler auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI.....	198
Abbildung 8.4:	Mittelwerte mathematischer Kompetenz von Mädchen und Jungen in den OECD-Staaten.....	200
Abbildung 8.5:	Perzentilbänder mathematischer Kompetenz in Deutschland nach Schulart und für die Gesamtstichprobe.....	202
Abbildung 8.6:	Prozentuale Anteile von Schülerinnen und Schülern auf den Stufen mathematischer Kompetenz in Deutschland nach Schulart.....	203
Abbildung 8.7:	Prozentualer Anteil der Mädchen und Jungen auf den Stufen der mathematischen Kompetenz in Deutschland.....	204
Abbildung 8.8:	Mathematische Kompetenz in Deutschland und an Gymnasien von PISA 2003 bis PISA 2018.....	205
Abbildung 9.1:	Die PISA-2018-Rahmenkonzeption naturwissenschaftlicher Grundbildung.....	213
Abbildung 9.2:	Aufgabeneinheit <i>Energieeffiziente Häuser</i> , Einleitung.....	217
Abbildung 9.3:	Aufgabeneinheit <i>Energieeffiziente Häuser</i> , Erläuterung der Simulation.....	218
Abbildung 9.4:	Aufgabeneinheit <i>Energieeffiziente Häuser</i> , Aufgabe 1	218
Abbildung 9.5:	Aufgabeneinheit <i>Energieeffiziente Häuser</i> , Aufgabe 2	219

Abbildung 9.6:	Aufgabeneinheit <i>Energieeffiziente Häuser</i> , Aufgabe 3	220
Abbildung 9.7:	Mittelwerte, Streuungen und Perzentilbänder der naturwissenschaftlichen Kompetenz der OECD-Staaten.....	224
Abbildung 9.8:	Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI	227
Abbildung 9.9:	Mittelwerte naturwissenschaftlicher Kompetenz nach Geschlecht in den OECD-Staaten	229
Abbildung 9.10:	Perzentilbänder naturwissenschaftlicher Kompetenz in Deutschland für die Gesamtstichprobe und nach Schulart.....	231
Abbildung 9.11:	Prozentuale Anteile Fünfzehnjähriger auf den Stufen der naturwissenschaftlichen Kompetenz in Deutschland nach Schulart.....	232
Abbildung 9.12:	Mittelwerte naturwissenschaftlicher Kompetenz in den OECD-Staaten in PISA 2018, PISA 2015 und PISA 2006	235
Abbildung 10.1:	Schematische Darstellung des mehrstufig adaptiven Testdesigns (MAT) für den Kompetenzbereich Lesen.....	245
Abbildung 10.2:	Auswahl- und Login-Bildschirm des PISA Student Delivery Systems (SDS)	249
Abbildung 10.3:	Bildschirm mit allgemeiner Einführung zur Erklärung des PISA-Testsystems (SDS).....	249

Tabellenverzeichnis

Tabelle 2.1:	Beispielaufgaben zur Leseflüssigkeit	27
Tabelle 2.2:	Kodieranweisungen zu der in Abbildung 2.2 dargestellten Aufgabe 6 aus der Aufgabeneinheit „Hühnerforum“	38
Tabelle 3.1:	Überblick über die typischen Anforderungen der acht Kompetenzstufen (Ic - VI) der Lesekompetenz (Tabelle adaptiert nach OECD, 2019b)	53
Tabelle 3.2:	Mittelwerte und Streuungen der Teilskalen der Lesekompetenz nach Schulart in Deutschland	70
Tabelle 3.3:	Mittelwerte und Streuungen der Gesamtskala Lesekompetenz nach Geschlecht und Schulart bei PISA 2018, 2015 und 2009 in Deutschland	75
Tabelle 4.1:	Übersicht der lesebezogenen Schülermerkmale	83
Tabelle 4.2:	Prozent der Zustimmung zu Einzelitems der Lesefreude im Vergleich zum OECD-Durchschnitt	87
Tabelle 4.3:	Veränderung der Anteile der „Lesemuffel“, „Normal-Lesenden“ und „Viel-Lesenden“ zwischen 2009 und 2018 nach Schulart und Geschlecht	97
Tabelle 4.4:	Veränderung des Online-Lesens zwischen 2009 und 2018	98
Tabelle 5.1:	Ausstattung mit materiellen ICT-Ressourcen pro Schülerin beziehungsweise Schüler in der neunten Klassenstufe	114
Tabelle 5.2:	Qualität materieller sowie personelle ICT-Ressourcen	116
Tabelle 5.3:	Integration digitaler Medien in den Schulalltag	118
Tabelle 5.4:	Zeit, die in einer normalen Schulwoche während des Unterrichts der Landessprache mit der Nutzung digitaler Geräte verbracht wird	120
Tabelle 5.5:	Maßnahmen zur Förderung von Schülerinnen und Schülern mit anderer Herkunftssprache in PISA 2009 und PISA 2018	125
Tabelle 6.1:	Zusammenhang zwischen der Lesekompetenz und dem sozioökonomischen beruflichen Status in den OECD-Staaten (Prädiktorvariable: HISEI)	139
Tabelle 6.2:	Zusammenhang zwischen der Lesekompetenz und dem sozioökonomischen und -kulturellen Status in den OECD-Staaten (Prädiktorvariable: ESCS)	141
Tabelle 6.3:	Merkmale der sozialen Herkunft und der Bildungswege differenziert nach EGP-Klassen	144
Tabelle 6.4:	Merkmale der sozialen Herkunft differenziert nach EGP-Klassen in den Gruppen der lesestarken (Kompetenzstufe V und VI) und leseschwachen (unter Kompetenzstufe II) Schülerinnen und Schüler	145
Tabelle 6.5:	Unterschiede im sozioökonomischen beruflichen Status der Eltern (HISEI) zwischen Jugendlichen mit und ohne Zuwanderungshintergrund in ausgewählten europäischen Staaten	148
Tabelle 6.6:	Lesekompetenz und Zuwanderungshintergrund in ausgewählten europäischen Staaten	150
Tabelle 6.7:	Prozentuale Anteile der Fünfzehnjährigen mit Zuwanderungs- hintergrund in Deutschland	151
Tabelle 6.8:	Unterschiede in der zu Hause gesprochenen Sprache sowie in der sozialen Herkunft zwischen Jugendlichen mit und Jugendlichen ohne Zuwanderungshintergrund	153
Tabelle 6.9:	Prozentuale Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund an Gymnasien und nicht gymnasialen Schularten	154
Tabelle 6.10:	Regressionsmodelle zu Zuwanderungshintergrund, sozialer Herkunft und Lesekompetenz für PISA 2018 und 2009	156
Tabelle 7.1:	Haupt-/Nebendomäne sowie wesentliche Änderungen nach Erhebungsrunde	168

Tabelle 7.2:	Originale und marginale Trendschätzungen von 2009 nach 2018 in Lesen, Mathematik und Naturwissenschaften	181
Tabelle 8.1:	Stufen mathematischer Kompetenz in PISA	192
Tabelle 8.2:	Mittelwerte und Standardabweichungen für die Gesamtstichprobe und nach Schulart.....	202
Tabelle 9.1:	Verteilung der Aufgaben des Naturwissenschaftstests auf die drei Teilkompetenzen, Wissensbereiche und Wissenssysteme	216
Tabelle 9.2:	Einordnung der Beispielaufgaben der Aufgabeneinheit <i>Energieeffiziente Häuser</i>	217
Tabelle 9.3:	Stufen naturwissenschaftlicher Kompetenz in PISA 2018	221
Tabelle 9.4:	Naturwissenschaftliche Kompetenz in Deutschland bei PISA 2018, 2015 und 2006 für die Gesamtstichprobe und getrennt nach Geschlecht und Schulart.....	233

Die Autorinnen und Autoren dieses Berichtsbandes

Prof. Dr. Kristina Reiss (ZIB/TU München)
Prof. Dr. Eckhard Klieme (ZIB/DIPF Frankfurt)
Prof. Dr. Olaf Köller (ZIB/IPN Kiel)
Prof. Dr. Cordula Artelt (Universität Bamberg)
Prof. Dr. Michael Becker-Mrotzek (Universität zu Köln)
Prof. Dr. Frank Goldhammer (ZIB/DIPF Frankfurt)
Prof. Dr. Aiso Heinze (IPN Kiel)
Prof. Dr. Doris Holzberger (ZIB/TU München)
Prof. Dr. Thomas Lindauer (Fachhochschule Nordwestschweiz)
Prof. Dr. Oliver Lüdtke (ZIB/IPN Kiel)
Prof. Dr. Nele McElvany (TU Dortmund)
Prof. Dr. Katharina Müller (Leibniz Universität Hannover)
Prof. Dr. Knut Neumann (IPN Kiel)
Prof. Dr. Maximilian Pfof (Universität Bamberg)
Dr. Sarah Bürger (Universität Wien)
Dr. Jennifer Diedrich (ZIB/TU München)
Dr. Carolin Hahnel (ZIB/DIPF Frankfurt)
Dr. Scott Harrison (ZIB/DIPF Frankfurt)
Dr. Sarah Hofer (ZIB/TU München)
Dr. Ulf Kroehne (DIPF Frankfurt)
Dr. Nicole Mahler (IQB Berlin)
Dr. Frank Reinhold (TU München)
Dr. Alexander Robitzsch (ZIB/IPN Kiel)
Dr. Silke Rönnebeck (IPN Kiel)
Dr. Anja Schiepe-Tiska (ZIB/TU München)
Dr. Mirjam Weis (ZIB/TU München)
Anastasia Doroganova (ZIB/TU München)
Jörg-Henrik Heine (ZIB/TU München)
Julia Mang (ZIB/TU München)
Anselm Strohmaier (TU München)
Ana Tupac-Yupanqui (ZIB/TU München)
Lisa Ziernwald (ZIB/TU München)

Die *TUM School of Education* an der Technischen Universität München ist Deutschlands erste Fakultät für Lehrerbildung und Bildungsforschung. Sie widmet sich der Aus- und Weiterbildung von Lehrkräften an Gymnasien in den MINT-Fächern (Mathematik, Informatik, Naturwissenschaften, Technik) sowie der Aus- und Weiterbildung von Lehrkräften an beruflichen Schulen im gewerblich-technischen Bereich. Das nationale Pro-

jektmanagement für PISA ist seit der Erhebung 2012 an der *TUM School of Education* angesiedelt.

Das *DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation* in Frankfurt am Main unterstützt Forschung, Politik und Praxis im Bildungsbereich durch wissenschaftliche Infrastruktur und vielfältige Forschungstätigkeit. Als Mitglied der Leibniz-Gemeinschaft verbindet das Institut erkenntnisorientierte Grundlagenforschung mit innovativen Entwicklungsarbeiten und Anwendungen zum Nutzen der Gesellschaft und ihrer Mitglieder. In PISA 2009 war das DIPF verantwortlich für das nationale Projektmanagement und in PISA 2015/2018 für die internationale Fragebogenentwicklung.

Das *IPN (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik)* an der Universität Kiel hat zum Auftrag, durch seine Forschungen die Pädagogik der Naturwissenschaften und der Mathematik weiterzuentwickeln und zu fördern. Die interdisziplinären Arbeiten des IPN umfassen Grundlagenforschung in Fragen des Lehrens und Lernens. Das IPN hatte das nationale Projektmanagement für PISA 2003 und 2006 inne.

Das *ZIB (Zentrum für internationale Vergleichsstudien, www.zib.education)* ist ein An-Institut der Technischen Universität München und ein gemeinnütziger Verein. Beteiligt sind am ZIB die *TUM School of Education*, das *IPN Kiel* und das *DIPF Frankfurt*. Der Zusammenschluss der drei Institutionen ermöglicht eine effiziente Bündelung von Kompetenz und Expertise im Bereich der Large-Scale-Assessments.

Das *IQB (Institut zur Qualitätsentwicklung im Bildungswesen)* in Berlin unterstützt die Arbeiten der Länder in der Bundesrepublik Deutschland bei der kontinuierlichen Weiterentwicklung und Sicherung von Bildungserträgen im Schulsystem. Eine zentrale Grundlage dieser Arbeiten bilden die länderübergreifenden Bildungsstandards der Kultusministerkonferenz, die definieren, welche Kompetenzen Schülerinnen und Schüler bis zu bestimmten Zeitpunkten in ihrer schulischen Laufbahn erwerben sollen. Das IQB überprüft regelmäßig, inwieweit diese Kompetenzziele in deutschen Schulen erreicht werden und unterstützt die Länder bei der Umsetzung der Bildungsstandards.